DIRECTORIO DE PROFESORES DEL CURSO: PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES. FEBRERO-MARZO 1983.

DR. OCTAVIO A. RASCON CHAVEZ (COORDINADOR)
DIRECTOR
FACULTAD DE INGENIERIA
UNAM
548 33 54

M. EN I. AUGUSTO VILLARREAL ARANDA
GERENTE DE OPERACIONES
GRUPO VEA
ASIA NO. 31
MEXICO 21, D.F.
554 45 31 Y 554 41 31

U.N.A.M. FACULTAD DE INGENIERIA
DIVISION DE EDUCACION CONTINUA

PROGRAMA DEL CURSO: PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES
QUE SE IMPARTIRA DEL 7 de Febrero al 18 de Marzo de DE 1983.

FECHA	HORARIO	TEMA	PROFESOR
7 de Feb. a 2 de Marzo.	18 a 21 h c/d.	INTRODUCCION Probabilidad. Estadística descriptiva e inferencia estadística. ESTADISTICA DESCRIPTIVA Obtención de datos: muestreo aleatorio simple. Procesamiento de información. Tablas de frecuencia. Histogramas. Polígonos de frecuencias. Medidas de tendencia central y de dispersión. Fractiles. Distribución conjunta de frecuencias. Regresión y correlación lineal. Análisis de series en el tiempo; predicción. Ejemplos y aplicaciones. PROBABILIDAD Eventos. Teoría de conjuntos. Espacio de eventos Probabilidad condicional. Independencia. Teorema de Bayes. Variables aleatorias continuas y discretas. Funciones de distribución. Momentos y esperanzas. Distribuciones de Bernoulli, hipergeométrica, binomial y de Poisson. Proceso de Poisson simple. Distribuciones uniforme, exponencial, normal y extremos. Ejemplos y Aplicaciones.	Dr. Octavio A. Rascón Chávez
al 18 de Mzo.	18 a 21 h c/día.	INFERENCIA ESTADISTICA Estimación puntual de los parámetros de una distribución de probabilidades. Estimación por intervalos Distribuciones muestrales Pruebas de hipótesis que involucren medias varianzas o proporciones. Prueba de bondad de ajuste en regresión lineal y en distribuciones de probabilidades. Ejemplos y Aplicaciones.	M. en I. Augusto Villarreal Aranda

EVALUACION DEL CURSO

	CONCEPTO	EVALUACION
1.	APLICACION INMEDIATA DE LOS CONCEPTOS EXPUESTOS	
2.	CLARIDAD CON QUE SE EXPUSIERON LOS TEMAS	
3.	GRADO DE ACTUALIZACION LOGRADO CON EL CURSO	
4.	CUMPLIMIENTO DE LOS OBJETIVOS DEL CURSO	
5.	CONTINUIDAD EN LOS TEMAS DEL CURSO	
6.	CALIDAD DE LAS NOTAS DEL CURSO	
7.	GRADO DE MOTIVACION LOGRADO CON EL CURSO	

ESCALA DE EVALUACION DE 1 A 10

1. ¿Qué le pareció el ambiente en la División de Educación Continua?

MUY AGRADABLE	AGRADABLE	DESAGRADABLE

2. Medio de comunicación por el que se enteró del curso:

PERIODICO EXCELSIOR ANUNCIO TITULADO DIVISION DE EDUCACION CONTINUA	PERIODICO NOVEDADES ANUNCIO TITULADO DIVISION DE EDUCACION CONTINUA	FOLLETO DEL CURSO

CARTEL MENSUAL	RADIO UNIVERSIDAD	COMUNICACION CARTA, TELEFONO, VERBAL, ETC.

REVISTAS TECNICAS	FOLLETO ANUAL	CARTELERA UNAM "LOS UNIVERSITARIOS HOY"	GACETA UNAM

3. Medio de transporte utilizado para venir al Palacio de Minería:

AUTOMOVIL PARTICULAR	METRO	OTRO MEDIO

4. ¿Qué cambios haría usted en el programa para tratar de perfeccionar el curso?

5. ¿Recomendaría el curso a otras personas?

SI	NO

6. ¿Qué cursos le gustaría que ofreciera la División de Educación Continua?

____

____

7. La coordinación académica fue:

EXCELENTE	BUENA	REGULAR	MALA

8. Si está interesado en tomar algún curso <u>intensivo</u> ¿Cuál es el horario más conveniente para usted?

LUNES A VIERNES DE 9 A 13 H. Y DE 14 A 18 H. (CON COMIDAS)	LUNES A VIERNES DE 17 A 21 H.	LUNES, MIERCOLES Y VIERNES DE 18 A 21 H.	MARTES Y JUEVES DE 18 A 21 H.

VIERNES DE 17 A 21 H. SABADOS DE 9 A 14 H.	VIERNES DE 17 A 21 H. SABADOS DE 9 A 13 Y DE 14 a 18 H.	O T R O

9. ¿Qué servicios adicionales desearía que tuviese la División de Educación Continua, para los asistentes? ______________________

____

10. Otras sugerencias:

____

____

____

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

PROBABILIDAD Y ESTADISTICA

FUNDAMENTOS Y APLICACIONES

E J E R C I C I O S

DR. OCTAVIO A. RASCON CHAVEZ

FEBRERO , 1983

1

EJEMPLO

EN UNA PRUEBA DE APTITUD QUE SE APLICO A 17 ALUMNOS DE PRIMERO DE SECUNDARIA, SELECCIONADOS AL AZAR DE LAS SECUNDARIAS DE UNA CIUDAD, SE OBTUVO UN PROMEDIO DE LAS CALIFICACIONES IGUAL A 34.35 PUNTOS, Y UNA DESVIACION ESTANDAR DE 3.9 PUNTOS. PROBAR LA HIPOTESIS DE QUE LA MEDIA DE LA VARIABLE ALEATORIA "CALIFICACION EN LA PRUEBA DE APTITUD DE LOS ALUMNOS DE 1° DE SECUNDARIA DE ESA CIUDAD" ES DE 38 PUNTOS, CONTRA LA DE QUE ES MENOR QUE 38. TOMAR $\alpha = 0.05$.

$$H_0: \mu = 38$$

$$H_1: \mu < 38$$

$$T = \frac{\bar{X} - \mu}{S_x}\sqrt{n-1} = \frac{34.35 - 38}{3.9}\sqrt{17-1} = -3.74$$

$$t_{c,0.05,16} = -1.746 > -3.74$$

PUESTO QUE $T < t_{c,0.05,16}$ SE RECHAZA LA HIPOTESIS NULA, CON 95% DE NIVEL DE CONFIANZA.

2

EJEMPLO

EN UN ESTUDIO DE MERCADOTECNIA SE TOMO UNA MUESTRA DE 20 PRECIOS DE CARNE EN 20 TIENDAS DISTINTAS PARA ESTIMAR SU VARIABILIDAD. LOS DATOS ARROJARON UN PROMEDIO $\bar{X}$ = \$32.00 Y UNA DESVIACION ESTANDAR S = \$8.00. CALCULAR EL INTERVALO DEL 95% DE NIVEL DE CONFIANZA DE LA VARIANCIA

$$I.C. = \left(\frac{20(8)^2}{34.9}, \frac{20(8)^2}{8.91}\right) = (38.91,\ 143.66)\ \$^2$$

EL INTERVALO DE CONFIANZA PARA LA DESVIACION ESTANDAR ES

$$(\sqrt{38.91}, \sqrt{143.66}) = (6.24\ ,\ 11.99)\ \$$$

EJERCICIO

LA DURACION DE LOS TRANSFORMADORES PRODUCIDOS EN UNA FABRICA FUE MEDIDA EN UNA MUESTRA DE 50 ELEMENTOS TOMADOS AL AZAR, OBTENIENDOSE LA SIGUIENTE DISTRIBUCION DE FRECUENCIAS:

INTERVALO N°	1	2	3	4
INTERVALO DE TIEMPO, AÑOS	$0 \le t < 1$	$1 \le t < 2$	$2 \le t < 3$	$t \ge 3$
FRECUENCIA	21	16	9	4

PROBAR LA HIPOTESIS DE QUE LA DISTRIBUCION DE PROBABILIDADES DE LA VARIABLE ALEATORIA "DURACION DE LOS TRANSFORMADORES" ES EXPONENCIAL CON PARAMETRO $\lambda = 0.45$ AÑOS^{-1}. USAR $\alpha = 0.05$.

SOLUCION

LAS FRECUENCIAS ESPERADAS SON: $nP(x_1 \le X < x_2)$

DONDE n = TAMAÑO DE LA MUESTRA

$p_1 = P(0 \le X < 1) = \int_0^1 (0.45)e^{-0.45t}\,dt = 0.362;\ 50p_1 = 18.10$

$p_2 = P(1 \le X < 2) = \int_1^2 (0.45)e^{-0.45t}\,dt = 0.232;\ 50p_2 = 11.60$

$p_3 = P(2 \le X < 3) = \int_2^3 (0.45)e^{-0.45t}\,dt = 0.145;\ 50p_3 = 7.25$

$p_4 = P(X > 4) = \int_3^\infty (0.45)e^{-0.45t}\,dt = \frac{0.259}{0.938}\ 41;\ 50p_4 = \frac{12.95}{\Sigma = 49.90} \ne 50$

$$\chi^2 = \frac{(21-18.10)^2}{18.10} + \frac{(16-11.6)^2}{11.6} + \frac{(9-7.25)^2}{7.25} + \frac{(4-12.95)^2}{12.95} = \underline{8.71}$$

$\chi^2_{0.95,3} = 7.81 < 8.74$

POR LO QUE SE RECHAZA LA HIPOTESIS NULA CON UN 5% DE NIVEL DE SIGNIFICANCIA.

EJERCICIO

SE PIENSA QUE LA EMISION DE PARTICULAS RADIOACTIVAS DE CIERTA FUENTE OCURRE SEGUN UNA DISTRIBUCION DE PROBABILIDADES DE POISSON. EL NUMERO DE PARTICULAS EMITIDAS EN 100 INTERVALOS CONSECUTIVOS DE 10 SEG QUEDO DISTRIBUIDO DE LA SIGUIENTE MANERA

N° DE PARTICULAS	0	1	2	3	4	>4
N° DE INTERVALOS (FRECUENCIA)	11	30	25	20	10	4

PROBAR LA HIPOTESIS DE QUE EFECTIVAMENTE SE TRATA DE UNA DISTRIBUCION DE POISSON. USAR $\alpha = 0.01$.

SOLUCION

PUESTO QUE NO NOS INDICAN UN VALOR DEL PARAMETRO DE LA DISTRIBUCION NECESITAMOS ESTIMARLO A PARTIR DE LA INFORMACION DADA ARRIBA:

$\lambda = \{(0 \times 11) + (1 \times 30) + (2 \times 25) + (3 \times 20) + (4 \times 10) + (5 \times 4)\}/100$

$= 2.00$ PARTICULAS/INTERVALO

LA DISTRIBUCION DE POISSON ES ENTONCES:

$$f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{2^x e^{-2}}{x!} = P(X = x)$$

$P_1 = f_X(0) = 2^0 e^{-2}/0! = 0.135;\ np_1 = 100 \times 0.135 = 13.5$

$P_2 = f_X(1) = 2^1 e^{-2}/1! = 0.270;\ np_2 = 100 \times 0.270 = 27.0$

$P_3 = f_X(2) = 2^2 e^{-2}/2! = 0.270;\ np_3 = 27.0$

$P_4 = f_X(3) = 2^3 e^{-2}/3! = 0.180;\ np_4 = 18.0$

$P_5 = f_X(4) = 2^4 e^{-2}/4! = 0.090;\ np_5 = 9.0$

$P_6 = P(X \geq 5) = 1 - F_X(4) = 0.055;\ np_6 = 5.5$

$$\chi^2 = \frac{(11-13.5)^2}{13.5} + \frac{(30-27.0)^2}{27.0} + \frac{(25-27.0)^2}{27.0} + \frac{(20-18)^2}{18} + \frac{(10-9.0)^2}{9.0} + \frac{(4-5.5)^2}{5.5}$$

$= \underline{1.687}$ $\quad \nu = 6 - 1 - 1 = 4$ $\quad$ ($\nu = N - r - 1$; r = N° DE ESTIMACIONES HECHA CON LOS DATOS)

$\chi^2_{0.99,\ 4} = 13.277 > 1.687$ ∴ <u>SE ACEPTA</u> LA HIPOTESIS NULA

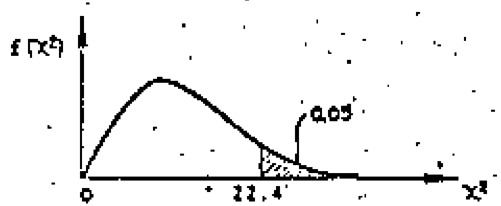
EJERCICIO

EN UN ESTUDIO CON FINES ANTROPOLOGICOS SE OBTUVO UNA MUESTRA ALEATORIA DEL TAMAÑO DE LA CABEZA DE LOS INDIGENAS ORIGINARIOS DE CIERTA REGION TROPICAL. LOS DATOS AGRUPADOS SE PRESENTAN EN LA SIGUIENTE TABLA. PROBAR LA HIPOTESIS DE QUE ESTOS DATOS CORRESPONDEN A UNA VARIABLE CON DISTRIBUCION NORMAL.

INTERVALO DE VALORES, mm	FRECUENCIA OBSERVADA, f_i	FRECUENCIA ESPERADA, e_i	$f_i - e_i$	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
<171.5	0	0.4	0.4	0.16	0.40
171.5-175.5	3	2.4	0.6	0.36	0.15
175.5-179.5	9	10.5	-1.5	1.25	0.12
179.5-183.5	29	33.1	-4.1	16.81	0.51
183.5-187.5	76	71.3	4.7	22.09	0.31
187.5-191.5	104	104.2	-0.2	0.04	0.00
191.5-195.5	110	108.2	1.8	3.24	0.03
195.5-199.5	88	77.3	10.7	114.49	1.48
199.5-203.5	30	37.5	-7.5	56.25	1.50
203.5-207.5	6	13.0	-7.0	49.00	3.77
207.5-211.5	4	3.0	1.0	1.00	0.33
211.5-215.5	2	0.5082	1.4918	2.23	4.53
215.5-219.5	1	0.0462	0.9538	0.910	19.69
> 219.5	0	≈ 0		TOTAL:	32.67

$$\chi^2 = \sum_{i=1}^{13} \frac{(f_i - e_i)^2}{e_i}, \qquad \chi^2 = 32.67 > 22.4 = \chi^2_{0.95,13} = \chi^2_c$$

POR LO QUE LA HIPOTESIS NULA NO PUEDE RECHAZARSE CON UN 95% DE NIVEL DE CONFIANZA.



EJERCICIO

SACAR UNA MUESTRA DE 50 NUMEROS DE UNA TABLA DE NUMEROS ALEATORIOS Y PROBAR LA HIPOTESIS DE QUE PROVIENEN DE UNA DISTRIBUCION UNIFORME DE 0 A 1, PREVIA REDUCCION A DECIMALES. USAR α = 0.05.

SOLUCION

UTILIZANDO LOS RENGLONES 1, 3, 5, 7, 9 DE LA TABLA DE NUMEROS ALEATORIOS PRESENTADA EN EL VOL. 1 DE ESTADISTICA DESCRITIVA*, MULTIPLICANDO x 10^{-5} CADA NUMERO Y ELIMINANDO LOS 3 ULTIMOS DIGITOS SE OBTIENE LA SIGUIENTE MUESTRA:

0.16 - 0.81 - 0.04 - 0.53 - 0.79 - 0.21 - 0.83 - 0.92 - 0.36 - 0.31
0.59 - 0.73 - 0.47 - 0.47 - 0.87 - 0.99 - 0.00 - 0.88 - 0.71 - 0.18
0.20 - 0.23 - 0.33 - 0.03 - 0.23 - 0.14 - 0.15 - 0.45 - 0.22 - 0.39
0.09 - 0.74 - 0.68 - 0.96 - 0.20 - 0.42 - 0.78 - 0.05 - 0.22 - 0.24
0.54 - 0.35 - 0.19 - 0.11 - 0.31 - 0.76 - 0.17 - 0.03 - 0.44 - 0.64

AGRUPANDO DATOS EN 10 INTERVALOS TENEMOS:

INTERVALO	f_i	e_i	$f_i - e_i$	$(f_i - e_i)^2$	$(f_i - e_i)^2/e_i$
0.000-0.105	6	5	1	1	0.20
0.105-0.205	10	5	5	25	5
0.205-0.305	7	5	2	4	0.80
0.305-0.405	4	5	-1	1	0.20
0.405-0.505	5	5	0	0	0
0.505-0.605	3	5	-2	4	0.80
0.605-0.705	2	5	-3	9	1.80
0.705-0.805	6	5	1	1	0.20
0.805-0.905	4	5	-1	1	0.20
0.905-1.005	3	5	2	4	0.80
					Σ = 10.0

$$\chi^2 = \sum_{i=1}^{n} \frac{(f_i - e_i)^2}{e_i} = 10.0$$

$$\chi^2_{0.95,9} = 16.9 > 10$$

POR LO TANTO SE ACEPTA LA HIPOTESIS DE QUE LOS NUMEROS CORRESPONDEN A UNA DISTRIBUCION UNIFORME, CON UN NIVEL DE SIGNIFICANCIA DE 0.05.

*PASCON, D. A., "INTRODUCCION A LA ESTADISTICA DESCRIPTIVA", VOL. I, 3a. ED. (1977), PAG. 21.

It is sometimes difficult to decide whether or not a frequency distribution is sufficiently near to the normal type to be fitted by a normal curve. A preliminary decision in a given case is largely the result of experience—of good guessing. Such a decision, however, can be reinforced by a fairly simple test involving the use of arithmetic probability paper.

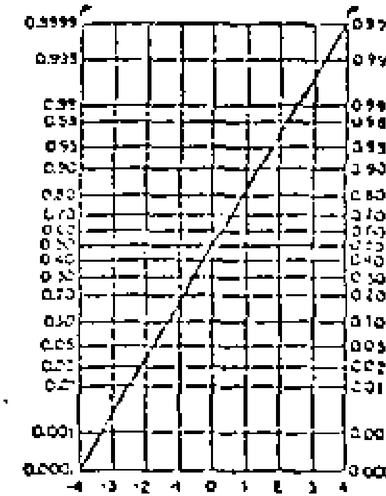


Figure 7-29

Since the area under the normal curve is unity, the partial areas, P_n, represent the *percentage cumulative frequencies* of a normal curve. For example, if we refer to Figure 7-18, we find that about 2 per cent of the normally distributed z's have values less than -2, about 16 per cent have values less than -1, 50 per cent less than 0, and so on.

We illustrate the use of the paper with the aid of Table 7-2 and Figure 7-30. Table 7-2 contains the familiar data of head lengths. Inasmuch as cumulative frequencies are of prime importance here, we are interested only in boundary values and not mid values. The last column of values is found from the formula $100 \times \text{cum } f/N$. For example, the tenth number in the last column, 23.3, equals $100 \times 113/482$.

EJERCICIO

CON OBJETO DE VERIFICAR LA CONSISTENCIA INTERNA DE UNA PRUEBA PSICOLOGICA, ESTA SE APLICO DOS VECES A CADA UNA DE DOS MUESTRAS ALEATORIAS. ESTAS MUESTRAS SE EXTRAJERON DE NIÑOS DEL CUARTO GRADO DE DOS ESCUELAS DISTINTAS, "A" y "B". LAS CALIFICACIONES DE LA PRIMERA APLICACION CORRESPONDEN A LA VARIABLE X; LAS DE LA SEGUNDA APLICACION (15 DIAS DESPUES DE LA PRIMERA), CORRESPONDEN A LA VARIABLE Y.

a. CALCULAR EL COEFICIENTE DE CORRELACION DE X y Y PARA CADA ESCUELA, Y PARA LAS DOS ESCUELAS JUNTAS, Y PROBAR LA HIPOTESIS DE QUE $\rho_{XY}>0$ EN CADA CASO.

b. PROBAR LA HIPOTESIS DE QUE $\mu_X = \mu_Y$ PARA AMBAS ESCUELAS JUNTAS, Y PARA CADA ESCUELA POR SEPARADO.

c. PROBAR LA HIPOTESIS DE QUE

1. $\mu_{X_A} = \mu_{X_B}$

2. $\mu_{Y_A} = \mu_{Y_B}$

3. $\sigma^2(X_A) = \sigma^2(X_B)$

4. $\sigma^2(Y_A) = \sigma^2(Y_B)$

FORMULAS

$\bar{X} = \Sigma X_i/n$, $\bar{Y} = \Sigma Y_i/n$, $S^2(X) = \Sigma X^2/n - \bar{X}^2$, $S^2(Y) = \Sigma Y_i^2/n - \bar{Y}^2$,

$S^2(d) = \Sigma d_i^2/n - \bar{d}^2$, $t_d = \frac{(\bar{X}-\bar{Y})\sqrt{n-1}}{S_d}$, $t_r = r_{XY}\sqrt{\frac{n-2}{1-r^2}}$

$$t_{\bar{X}-\bar{Y}} = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{n_X S^2(X) + n_Y S^2(Y)}{n_X+n_Y-2}\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}}\,; \quad F = \frac{S_M^2}{S_m^2}$$

DONDE S_M^2 y S_m^2 SON ESTIMACIONES INSESGADAS DE LAS VARIANCIAS MAYOR Y MENOR, RESPECTIVAMENTE, DE LAS DOS QUE SE ESTAN COMPARANDO.

RESPUESTAS A LOS INCISOS a y b

ESCUELA A

X	Y	X^2	Y^2	XY	d=x-y	d^2
34	35	1156	1225	1190	-1	1
39	36	1521	1295	1404	3	9
40	40	1600	1600	1600	0	0
35	38	1225	1444	1330	-3	9
30	29	900	841	870	1	1
28	26	784	676	728	2	4
33	34	1089	1156	1122	-1	1
38	40	1444	1600	1520	-2	4
32	39	1024	1521	1248	-7	49
37	35	1369	1225	1295	2	4
26	26	676	676	676	0	0
40	39	1600	1521	1560	1	1
32	30	1024	900	960	2	4
33	34	1089	1156	1122	-1	1
38	33	1444	1089	1254	5	25
34	39	1156	1521	1326	-5	25
35	37	1225	1369	1295	-2	4
584	590	20326	20516	20530	-6	142

$\bar{X} = \frac{584}{17} = 34.352941$; $\bar{X}^2 = 1180.1245$

$\bar{Y} = \frac{590}{17} = 34.705882$; $\bar{Y}^2 = 1204.4982$

$\bar{d} = -6/17 = -0.3529411$; $\bar{d}^2 = 0.1245674$

$S^2(X) = \frac{20326}{17} - 1180.1245 = 15.5225$; $S(X) = 3.9398604$

$S^2(Y) = \frac{20516}{17} - 1204.4982 = 19.9723$; $S(Y) = 4.4690379$

$S_d^2 = \frac{142}{17} - 0.1245674 = 8.2283737$; $S_d = 2.8685141$

$r_{XY} = \frac{(20530/17)-(34.352941)(34.705882)}{(3.9398604)(4.4690379)} = 0.7242882$

$H_0: \rho = 0$, $H_1: \rho \neq 0$, $t = t_{0.975,15} = 2.13$

$$t_\rho = 0.774 \sqrt{\frac{17-2}{1-0.774^2}} = 0.774 \times 6.116 = 4.73 > 2.13$$

POR LO QUE SE RECHAZA LA HIPOTESIS DE QUE $\rho_{xy} = 0$ CON UN NIVEL DE CONFIANZA DEL 95%. $t_c = t_{0.975,16} = 2.12$

$$t_d = \left|\frac{(34.353-34.706)\sqrt{16}}{2.869}\right| = 0.492 < 2.12$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\mu_x = \mu_y$ CON UN 95% DE NIVEL DE CONFIANZA.

ESCUELA B

X	Y	X^2	Y^2	XY	d=X-Y	d^2
39	41	1521	1681	1599	-2	4
27	36	729	1296	972	-9	81
33	31	1089	961	1023	2	4
37	36	1369	1296	1332	1	4
35	36	1225	1296	1260	-1	1
31	33	961	1089	1023	-2	4
33	32	1089	1024	1056	1	1
39	40	1521	1600	1560	-1	1
39	35	1521	1225	1365	4	16
27	29	729	841	783	-2	4
32	36	1024	1296	1152	-4	16
34	35	1156	1225	1190	-1	1
35	34	1225	1156	1190	1	1
36	42	1296	1764	1512	-6	36
34	34	1156	1156	1156	0	0
29	31	841	961	899	-2	4
545	561	18614	19867	19072	-21	175

$\bar{x} = \frac{545}{16} = 33.75$; $\bar{x}^2 = 1139.0625$; $\bar{d} = -1.3125$; $\bar{d}^2 = 1.7226562$

$\bar{y} = \frac{561}{16} = 35.0625$; $\bar{y}^2 = 1229.3789$

$S_x^2(x) = \frac{18614}{16} - 1139.0625 = 24.3125$; $S(x) = 4.9307707$

$S^2(Y) = \frac{19867}{16} - 1229.3789 = 12.3086$; $S(Y) = 3.5083614$

$S_d^2 = \frac{175}{16} - 1.7226562 = 9.214844$ $S_d = 3.0355961$

$$r_{xy} = \frac{(19072/16)-(33.75)(35.0625)}{(4.9307707)(3.5083614)} = \underline{0.4994934}$$

$$t_\rho = 0.499\sqrt{\frac{14}{0.751}} = 2.154 < 2.15$$

POR LO QUE SE RECHAZA LA HIPOTESIS DE QUE $\rho_{xy} = 0$, CON UN 95% DE NIVEL DE CONFIANZA.

$$t_d = \left|\frac{(33.75-35.063)\sqrt{15}}{3.036}\right| = 1.67 < 2.13$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\mu_x = \mu_y$, CON UN 95% DE NIVEL DE CONFIANZA.

AMBAS ESCUELAS JUNTAS

$\Sigma x_i = 1124$, $\Sigma y_i = 1151$, $\Sigma x_i^2 = 38940$, $\Sigma y_i^2 = 39572$, $\Sigma d_i = -27$, $\Sigma d_i^2 = 317$

$\bar{x} = \frac{1124}{33} = 34.060606$; $\bar{x}^2 = 1160.1248$; $\bar{d} = \frac{-27}{33} = -0.8181818$; $\bar{d}^2 = 0.6694214$

$\bar{y} = \frac{1151}{33} = 34.878787$; $\bar{y}^2 = 1216.5297$

$S^2(X) = \frac{38940}{33} - 1160.1248 = 19.8752$; $S(X) = 4.458161$

$S^2(Y) = \frac{40683}{33} - 1216.5297 = 16.2884$; $S(Y) = 4.0358889$

$S_d^2 = \frac{317}{33} - 0.6694214 = 8.9366392$; $S_d = 2.9894212$

$$r_{xy} = \frac{(39572/33)-(34.060606)(34.878787)}{(4.4581613)(4.0358889)} = \underline{0.6201924}$$

$$t_d = \left| \frac{(33.061 - 34.879)\sqrt{32}}{2.989} \right| = 1.548 < 2.04$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\mu_X = \mu_Y$, CON UN 95% DE NIVEL DE CONFIANZA.

$$t_\rho = 0.620\sqrt{\frac{31}{0.616}} = 4.398 > 2.04$$

POR LO QUE SE RECHAZA LA HIPOTESIS DE QUE $\rho_{xy} = 0$, CON UN NIVEL DE CONFIANZA DEL 95%.

<u>RESPUESTAS AL INCISO c</u>

$$t_{\bar{x}_A - \bar{x}_B} = \frac{34.35 - 33.75}{\sqrt{\frac{17 \times 15.52 + 16 \times 24.31}{31}\left[\frac{1}{17} + \frac{1}{16}\right]}} = \frac{0.6}{\sqrt{\frac{263.84 + 388.96}{31}(0.121)}} =$$

$$= \underline{0.368 < 2.04}$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\mu_{X_A} = \mu_{X_B}$, CON UN 95% DE NIVEL DE CONFIANZA.

$$t_{\bar{Y}_A - \bar{Y}_B} = \frac{|34.71 - 35.06|}{\sqrt{\frac{17 \times 19.97 + 16 \times 12.31}{31}(0.121)}} = \frac{|-0.35|}{\sqrt{\frac{339.49 + 196.96}{31}(0.121)}}$$

$$= 0.24 < 2.04$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\mu_{Y_A} = \mu_{Y_B}$, CON UN 95% DE NIVEL DE CONFIANZA.

PARA LA PRUEBA DE IGUALDAD DE VARIANCIAS USAREMOS

$$F = \frac{24.31\sqrt{\frac{16}{15}}}{15.52\sqrt{\frac{17}{16}}} = 1.57 < 3.41 = F_{0.01}(15,16)$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\sigma^2_{X_A} = \sigma^2_{X_B}$ CON UN 98% DE NIVEL DE CONFIANZA.

$$F = \frac{19.97\sqrt{\frac{17}{16}}}{12.31\sqrt{\frac{16}{15}}} = \frac{20.58}{12.71} = 1.62 < 3.41$$

POR LO QUE SE ACEPTA LA HIPOTESIS DE QUE $\sigma^2_{Y_A} = \sigma^2_{Y_B}$ CON UN 98% DE NIVEL DE CONFIANZA.

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

PROBABILIDAD Y ESTADISTICA. FUNDAMENTOS Y APLICACIONES

PRUEBAS DE HIPOTESIS E INTERVALOS DE CONFIANZA EN REGRESION LINEAL

DR. OCTAVIO A. RASCON CHAVEZ

FEBRERO, 1983

o de Minería Calle de Tacuba 5 primer piso Deleg. Cuauhtemoc 06000 México, D.F. Tel.: 521-40-20 Apdo. Postal M-2285

PRUEBAS DE HIPOTESIS E INTERVALOS DE CONFIANZA EN REGRESION LINEAL

SI EL MODELO QUE RELACIONA A Y CON X ES LINEAL, ENTONCES

$$Y = MX + B$$

SI NO SE CONOCEN M Y B, ES NECESARIO ESTIMARLOS CON BASE EN UNA MUESTRA, CON LO CUAL SE OBTIENE

$$\tilde{Y} = mX + b$$

EN DONDE m ES EL ESTIMADOR DE M, Y b, EL DE B. SEA $\sigma^2_{y|x}$ LA VARIANCIA DE LA ESTIMACION DE Y CON BASE EN X.

SE PUEDE DEMOSTRAR QUE, SI SE CONOCE $\sigma^2_{y|x}$, ENTONCES:

$$Var(m) = \sigma^2_m = \sigma^2_{y|x}/\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sigma^2_{y|x}/nS^2_x$$

$$Var(b) = \sigma^2_b = \sigma^2_{y|x}/n + \frac{\bar{x}^2\sigma^2_{y|x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sigma^2_{y|x}\left(\frac{1}{n} + \frac{\bar{x}^2}{nS^2_x}\right)$$

$$Var(mX + b) = \sigma^2_{y|x}/n + \frac{\sigma^2_{y|x}(x - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sigma^2_{\tilde{y}} = \sigma^2_{y|x}\left(\frac{1}{n} + \frac{(x - \bar{x})^2}{nS^2_x}\right)$$

SI $\sigma^2_{y|x}$ NO SE CONOCE, SE PUEDE OBTENER UNA <u>ESTIMACION INSESGADA</u> DE ELLA MEDIANTE LA ECUACION

$$S^2_{y|x} = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

INTERVALOS DE CONFIANZA; $\sigma_{y|x}$ CONOCIDA

a. PARA LA ORDENADA EN EL ORIGEN, B,

$$b \pm z_c\sigma_b$$

DONDE $z_c = P(z < z_c) = 1 - \alpha/2$; α = NIVEL DE SIGNIFICANCIA

b. PARA LA PENDIENTE, M:

$$m \pm z_c\sigma_m$$

c. PARA LA PREDICCION, Y_i:

$$\tilde{Y}_i \pm z_c\sigma_{\tilde{y}}$$

EN CASO DE QUE $\sigma_{y|x}$ SEA DESCONOCIDA (ES LO USUAL), DEBE ESTIMARSE A PARTIR DE LA MUESTRA MEDIANTE $S_{y|x}$. EN TAL CASO LOS INTERVALOS DE CONFIANZA CAMBIAN A:

a. PARA LA ORDENADA EN EL ORIGEN, B: $b \pm t_c\sigma_b$

$$b \pm t_c S_{y|x}\sqrt{\frac{\overline{x^2}}{nS^2_x}}$$

DONDE t_c ES EL VALOR CRITICO DE UN NIVEL DE SIGNIFICANCIA α, CORRESPONDIENTE A UNA DISTRIBUCION t DE STUDENT CON $\nu = n - 2$ GRADOS DE LIBERTAD, Y S^2_x ES LA VARIANCIA (SESGADA) DE LA MUESTRA DE x.

b. PARA LA PENDIENTE, M: $m \pm t_c\sigma_m$

$$m \pm t_c S_{y|x}/\sqrt{nS^2_x} \quad \text{O} \quad m \pm t_c\frac{S_{y|x}}{S_x\sqrt{n}}$$

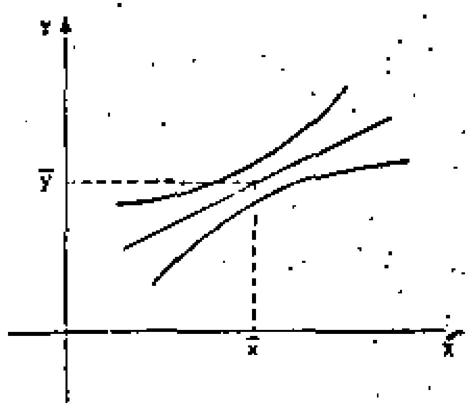
c. PARA LA PREDICCION. Y_i: $\hat{Y}_i \pm t_c \sigma_{\hat{y}}$

$$\hat{Y}_i \pm t_c\, S_{y|x} \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{nS_x^2}}$$

SI x_i ESTA DENTRO DEL RANGO DE LA MUESTRA, O

$$\hat{Y}_i \pm t_c\, S_{y|x} \sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{nS_x^2}}$$

SI x_i ESTA FUERA DEL RANGO.



EJEMPLO

LA FORMACION DEL ALCOHOL EN UN PROCESO DE FERMENTACION SE RELACIONA CON LA TEMPERATURA. EN UNA SERIE DE SEIS MEDICIONES A DISTINTAS TEMPERATURAS SE OBTUVO LO SIGUIENTE:

TEMPERATURA, x, °C	35	40	45	50	55	60
ALCOHOL, y, 1%	20.2	21.1	23.2	23.6	25.8	26.3

SI SE AJUSTA UNA RECTA POR MINIMOS CUADRADOS SE OBTIENE

$$\tilde{y} = 0.225\, x + 13.01$$

$(\bar{x} = 47.5,\ \bar{y} = 23.7)$

INTERVALOS DE CONFIANZA CON $\sigma^2_{y|x} = 0.8$ (CONOCIDA), $\alpha = 0.05$.

$$\sigma_b = \sqrt{\frac{0.8}{6} + \frac{0.8 \times 47.5^2}{437.5}} = 1.845$$

DONDE $\sum_{i=1}^{6} (x_i - \bar{x})^2 = 437.5$

$$b \pm z_c \sigma_b = 13.01 \pm 1.96 \times 1.845 = (9.39,\ 16.63)$$

$$\sigma_m = \sqrt{\frac{0.8}{437.5}} = 0.0428$$

$$m \pm z_c \sigma_m = 0.225 \pm 1.96 \times 0.0428 = 0.225 \pm 0.084 =$$

$$= (0.141,\ 0.309)$$

EJERCICIO

PARA LOS DATOS DE X y Y PRESENTADOS EN LA SIGUIENTE TABLA, CALCULAR $S_{y|x}$ Y LOS INTERVALOS DE CONFIANZA DE B y M PARA $\alpha = 0.05$, Y PARA Y CORRESPONDIENTE A X=50.

Temp. x, °C	Alcohol lts.	$\hat{y}_i$	$y_i-\hat{y}_i$	$(y_i-\hat{y}_i)^2$	$x_i-\bar{x}$,	$(x_i-\bar{x})^2$
35	20.2	20.9	-0.7	0.49	-12.5	156.2
40	23.1	22.0	1.1	1.21	- 7.5	56.2
45	23.2	23.1	0.1	0.01	- 2.5	6.2
50	23.6	24.3	-0.7	0.49	2.5	6.2
55	25.8	25.4	0.4	0.16	7.5	56.2
60	26.3	26.5	-0.2	0.04	12.5	156.2
$\Sigma=285$				$\Sigma=2.40$		$\Sigma=437.2$

SABEMOS QUE $\hat{y} = 0.225\,x + 13.01$

$\hat{y}(35)=0.225(35)+13.01=20.9,\qquad \hat{y}(40)=0.225\,(40)+13.01=22.0,$ etc.

$$\bar{x} = \frac{285}{6} = 47.5, \qquad S_x^2 = \frac{437.2}{6} = 72.9$$

INTERVALOS DE CONFIANZA:

a) PARA B:

$$13.01 \pm t_c\, S_{y|x}\sqrt{\frac{\bar{x}^2}{nS_x^2} + \frac{1}{n}}$$

$$t_c = t_{0.975,4} = 2.776, \qquad S^2_{y|x} = \frac{1}{n-2}\,\Sigma(y_i-\hat{y}_i)^2 = \frac{1}{4}\,2.4 = 0.6,$$

$$S_{y|x} = \sqrt{0.6} = 0.77\ .$$

$$13.01 \pm 2.776 \times 0.77\sqrt{\frac{47.5^2}{6(72.9)} + \frac{1}{6}} = 13.01 \pm 4.93 = (8.08, 17.94)$$

b) PARA M:

$$0.225 \pm t_c \frac{S_{y|x}}{\sqrt{nS_x^2}} = 0.225 \pm 2.776\,\frac{0.77}{\sqrt{6\,(72.9)}} =$$

$$= 0.225 \pm 0.102 = (0.123, 0.327)$$

c) PARA $y_i(x=50)$: $y_i(50) = 24.3$

$$24.3 \pm t_c S_{y|x}\sqrt{\frac{1}{n} + \frac{(x_i-\bar{x})^2}{n\,S_x^2}} = 24.3 \pm 2.776 \times 0.77\sqrt{\frac{1}{6} + \frac{(50-47.5)^2}{6(72.9)}} =$$

$$= 24.3 \pm 0.9 = (23.4, 25.2)$$

TAREA: HACER ESTIMACIONES DE INTERVALOS DE CONFIANZA PARA $\alpha = 0.05$ Y $\alpha = 0.01$, DE b, m Y Y_i, ESTE ULTIMO PARA UN $X = x_i$ QUE SELECCIONE CADA QUIEN. UTILIZAR UNO DE LOS PROBLEMAS DE REGRESION DEJADOS COMO TAREA ANTERIORMENTE.

PRUEBAS DE HIPOTESIS

a. PARA LA ORDENADA EN EL ORIGEN

SE DEMUESTRA QUE

$$\frac{B - b_0}{S_{y|x}\sqrt{\dfrac{\overline{x^2}}{nS_x^2}}} = \frac{B - b_0}{\dfrac{S_{y|x}}{S_x}\sqrt{\dfrac{\overline{x^2}}{n}}} = T$$

TIENE DISTRIBUCION t DE STUDENT CON $\nu = n - 2$ GRADOS DE LIBERTAD.

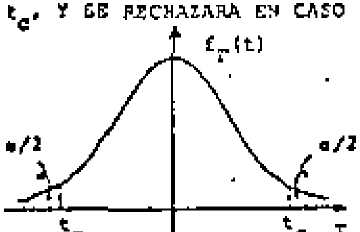
SI SE DESEA PROBAR LA HIPOTESIS

$$H_0 : B = b_0$$
$$H_1 : B \neq b_0$$

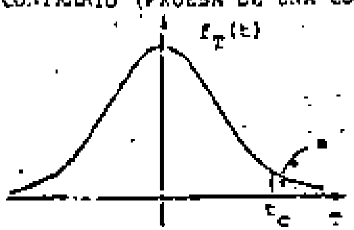
BASTA SUSTITUIR A $B = b_0$ EN LA ECUACION ANTERIOR Y EVALUAR $T = t$, ES DECIR,

$$t = \frac{b - b_0}{\dfrac{S_{y|x}}{S_x}\sqrt{\dfrac{\overline{x^2}}{n}}}$$

SE ACEPTARA H_0 SI $|t| < |t_c|$; EN CASO CONTRARIO SE RECHAZARA (PRUEBA DE DOS COLAS). SI H_1 FUERA $B > b_0$, SE ACEPTARA SI $t < t_c$, Y SE RECHAZARA EN CASO CONTRARIO (PRUEBA DE UNA COLA)



PRUEBA DE DOS COLAS



PRUEBA DE UNA COLA

b. PARA LA PENDIENTE, M

ANALOGAMENTE, PARA M, LA ESTADISTICA

$$\frac{M - m_0}{S_{y|x}/\sqrt{nS_x^2}} = \frac{M - m_0}{\dfrac{S_{y|x}}{S_x\sqrt{n}}} = T,$$

DONDE m_0 = VALOR DE M BAJO LA HIPOTESIS NULA $H_0 : M = m_0$.

TAMBIEN TIENE DISTRIBUCION t DE STUDENT CON $\nu = n - 2$ GRADOS DE LIBERTAD:

$$t = \frac{m - m_0}{\dfrac{S_{y|x}}{S_x\sqrt{n}}}$$

EJEMPLO

CONSIDERE LOS DATOS SIGUIENTES:

x	0	1	2	3	4	5	6	7	8	9
y	0.16	0.09	0.09	0.23	0.60	0.39	0.55	0.75	0.81	0.85

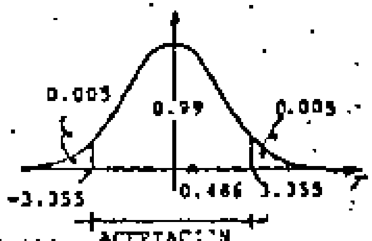
$m = 0.093$, $b = 0.032$, $S^2_{y|x} = 0.01258$

$$S_x^2 = \frac{1}{10}\sum_{i=1}^{10}(x_i - \bar{x})^2 = 82.50/10 = 8.25;\ \Sigma x_i^2 = 285,\ \overline{x^2} = \frac{285}{10} = 28.5$$

a. PROBAR LA HIPOTESIS DE QUE $B = 0$

b. PROBAR LA HIPOTESIS DE QUE $M = 0.1$

CON $\alpha = 0.01$ Y $S_{y|x}$ DESCONOCIDA.



a. $H_0 : B = 0$; $H_1 : B \neq 0$

$$t = \frac{b - b_0}{S_{y|x}\sqrt{\dfrac{\overline{x^2}}{nS_x^2}}} = \frac{0.032 - 0}{\sqrt{0.01258}\sqrt{\dfrac{28.5}{10 \times 8.25}}} = 0.486$$

$t_c = t_{0.995,\ 8} = 3.355 > 0.486 \therefore$ SE ACEPTA H_0.

b. $H_0 : M = 0.1;$ $\quad H_1 : M \neq 0.1$

$$t = \frac{m - m_0}{\dfrac{S_{y|x}}{S_x\sqrt{n}}} = \frac{0.053 - 0.1}{\dfrac{\sqrt{0.01252}}{\sqrt{8.25 \times 10}}} = 3.25^{*} < 3.355$$

SE ACEPTA H_0 CON 99% DE NIVEL DE CONFIANZA.

PRUEBA DE HIPOTESIS PARA EL COEFICIENTE DE CORRELACION, ρ_{xy}

PRUEBA

$H_0 : \rho_{xy} = 0 ;$ $\quad H_1 : \rho_{xy} \gtrless 0$

SE DEMUESTRA QUE EN CASO DE QUE X Y Y SON INDEPENDIENTES ($\rho = 0$), LA ESTADISTICA

$$T = r_{xy}\sqrt{\frac{n - 2}{1 - r_{xy}^2}}$$

TIENE DISTRIBUCION t DE STUDENT CON $n - 2$ GRADOS DE LIBERTAD.

EJEMPLO

EN BASE A UNA MUESTRA ALEATORIA DE 30 DATOS SOBRE LA TEMPERATURA MEDIA DURANTE UN MES, X, Y EL PESO MEDIO DE LOS TOMATES PISCADOS, Y, SE OBTUVO UN COEFICIENTE DE CORRELACION $r_{xy} = 0.931$.

PROBAR LA HIPÓTESIS DE QUE $\rho_{xy} = 0$. USAR $\alpha = 0.05$.

$H_0 : \rho_{xy} = 0 ;$ $\quad H_1 : \rho_{xy} \neq 0$

$$t = r_{xy}\sqrt{\frac{n - 2}{1 - r_{xy}^2}} = 0.931\sqrt{\frac{30 - 2}{1 - 0.931^2}} = 13.448$$

$$t_c = t_{0.975,\ 28} = 2.048 < 13.448$$

$\therefore$ SE RECHAZA H_0 A UN NIVEL DE CONFIANZA DEL 95%.

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

PROBABILIDAD Y ESTADISTICA: FUNDAMENTOS Y APLICACIONES

INTRODUCCION A LA ESTADISTICA DESCRIPTIVA

VOLUMEN I

Dr. Octavio A. Rascón Chávez

160

FEBRERO, 1983

INTRODUCCION A LA ESTADISTICA DESCRIPTIVA

VOLUMEN I

OCTAVIO A. RASCON CH.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
DIRECCION GENERAL DE PUBLICACIONES
MEXICO 1977

COMISION DE NUEVOS METODOS DE ENSEÑANZA

TEXTOS PROGRAMADOS

CONTENIDO

PRESENTACION

Una de las técnicas modernas de la enseñanza es la instrucción programada. Dentro de esta técnica el contenido de los textos se ordena de modo sistemático y se adapta al ritmo de asimilación de cada estudiante; existe la participación activa a lo largo del curso y se controla constantemente el aprendizaje mediante el planteamiento de preguntas y la inmediata verificación o corrección de las respuestas.

El material de un texto programado se organiza empleando ideas simples, las cuales se presentan en un orden que facilite la comprensión del estudiante. Las ideas o segmentos de información se ofrecen en cuadros lógicamente coordinados, con lo cual se evita que, en alguna de sus fases, la enseñanza llegue a ser demasiado fácil o muy difícil, a la vez que se logra mantener abierto el camino para que todos puedan adelantar según su propio ritmo de trabajo y comprensión.

Conforme van apareciendo los nuevos segmentos de información, se plantean al estudiante cuestiones específicas que, por obligarlo a intervenir activamente, le reafirman los conocimientos aprendidos, le estimulan y le mantienen atento. Para ello, cada pregunta va seguida inmediatamente de la respuesta correcta. Así el estudiante se siente alentado por sus aciertos y corrige de inmediato sus errores.

Debido a tales características el texto programado se adapta a la velocidad particular de cada estudiante, lo que no puede hacer el maestro encargado de un grupo numeroso.

Con la enseñanza programada no se pretende sustituir al profesor. Se utiliza sólo como un instrumento auxiliar, útil para iniciar estudios, subsanar lagunas, desarrollar habilidades o complementar el aprendizaje de un gran número de materias.

Los textos programados son productos fundamentales de la nueva etapa experimental de la enseñanza. Su elaboración requiere que el contenido se apruebe, reiteradamente, con grupos de alumnos, a fin de simplificar al máximo el proceso de aprendizaje y de llevar a un punto óptimo su eficiencia y aprovechamiento.

Este libro pertenece a la primera serie de textos programados que prepara la Universidad Nacional Autónoma de México para ofrecerlos, a sus estudiantes y profesores, como parte de un programa general de modernización de los métodos docentes y de superación académica.

COMISIÓN DE NUEVOS MÉTODOS DE ENSEÑANZA

PROLOGO

La elaboración del texto *Introducción a la estadística descriptiva*, volumen I, fue planeada, conjuntamente, por la Comisión de Nuevos Métodos de Enseñanza y el Instituto de Ingeniería de la Universidad Nacional Autónoma de México, como parte de su programa general de actualización de los métodos docentes.

Los objetivos generales que persigue este texto son los de enseñar a:

1. Obtener una muestra representativa de la población que se estudie estadísticamente (unidades 1 y 2).
2. Determinar las frecuencias de ocurrencia de los valores o intervalos de valores que constituyen la muestra (unidades 2 y 3).
3. Representar la muestra en forma tabular y gráfica (unidades 3 y 4).
4. Calcular medidas numéricas que representen las tendencias de concentración y de dispersión de los valores que constituyen la muestra (unidades 5, 6 y 7).

Los conocimientos previos para entender el contenido del texto se limitan a la aritmética y al trazo de gráficas en el plano; por lo tanto, todo estudiante de bachillerato o de nivel profesional debe estar capacitado para estudiarlo.

Agradezco a todos los que participaron en la elaboración de este texto, en particular al doctor Roger Díaz de Cossío, director del Instituto de Ingeniería, UNAM, por haberme brindado todas las facilidades para realizarlo.

Octavio A. Rascón Ch.

INSTRUCCIONES

Este libro puede utilizarse en diversas formas; el profesor deberá tomar la decisión definitiva sobre cómo usarlo. Algunos profesores preferirán asignar temas para estudiar durante toda la hora de clase y además material para tareas; este sistema tiene la desventaja de impedir la relación directa entre profesor y alumno. Otros profesores optarán por dejar tareas que se discutirán al principiar la siguiente clase, y, en caso necesario, se ampliará algún tema específico.

Se recomienda que las sesiones de trabajo no excedan de una hora para evitar fatiga mental y reducción de la capacidad de asimilación; si se desea seguir estudiando, habrá que tomar un descanso de algunos minutos. No es conveniente interrumpir el estudio en cualquier punto, sino al finalizar un tema.

A continuación se muestra un ejemplo del formato en que están dispuestos los cuadros: con fondo blanco se presenta la información o la pregunta, y con fondo gris la respuesta correspondiente.

Este es un libro programado de Estadística ____________.

Descriptiva

Este texto de Estadística Descriptiva es ____________.
(programado/convencional)

programado

El presente libro ha sido diseñado para que el estudiante lo aprenda por sí mismo en un tiempo promedio de 15 horas. Para lograr el máximo aprovechamiento de él, se requiere del lector una participación activa y que atienda las siguientes indicaciones:

1. Al estudiar el material no se apresure, hágalo cuidadosamente.
2. Mientras lee un cuadro, mantenga oculta la respuesta correcta (zona sombreada)

con el cobertor de respuestas; luego escriba su contestación. *Nunca responda antes de leer totalmente un cuadro.*

3. Moviendo cuidadosamente hacia abajo el cobertor hasta dejar descubierta totalmente la parte sombreada, destape la respuesta correspondiente al cuadro que leyó y cotéjela con la suya; si está correcta pase al siguiente cuadro; en caso contrario *corríjala y lea nuevamente el cuadro.*
4. Al terminar cada unidad, resuelva totalmente el examen correspondiente y después verifique sus respuestas; cada una vale un punto. Si su calificación es de 85 por ciento, o mayor, repase los temas en los cuales se equivocó; si la calificación es menor, estudie nuevamente *toda* la unidad.

Al final de cada unidad se encuentran las hojas de trabajo que se emplearán en ella. (se le indique usar una, debe separarla del libro para facilitar su manejo.

UNIDAD I POBLACION Y MUESTRA

PREFACIO

En esta unidad se introducen algunos conceptos fundamentales que se utilizarán a menudo en las unidades subsecuentes. Se definen *dato u observación*, *población* y *muestra*. Se indican además algunas formas sencillas para obtener una *muestra aleatoria*, se da la definición de Estadística y se clasifica a ésta; de acuerdo con el tipo de reglas que proporciona, en *Estadística Descriptiva e Inferencia Estadística*: la primera da las reglas para describir los datos, y la última para inferir características de la población a partir de la información limitada obtenida de una muestra.

PARTE A. POBLACION

1 Al realizar varias veces un experimento para conseguir información acerca de un problema se obtiene un grupo de resultados; a cada resultado se le denomina *dato* u *observación*. Si se repite tres veces un experimento consistente en lanzar una moneda, y anotar el lado que queda hacia arriba, se obtendrá un grupo de tres ________
(datos/monedas)

datos (observaciones)

2 Supongamos que un fabricante de focos realiza una prueba con una serie de focos nuevos. El experimento consiste en seleccionar un foco, conectarlo y encenderlo. Si el foco prende se anota una S; en caso contrario se anota una N. Si el experimento se repite cinco veces, se obtendrá un grupo de cinco datos u observaciones. Si los tres primeros focos sí prendieron y los otros no, se tendrán las siguientes anotaciones: S, S, ___, N, ___.

S, N

3 Si un experimento consiste en lanzar una moneda, y anotar el lado que queda hacia arriba, se pueden obtener sólo ______ tipos de observaciones diferentes, cara y cruz.
(cuántos)

dos

4 Si una moneda se lanza diez veces y cada vez se anota el lado que queda hacia arriba, ¿cuántos datos se obtendrán? ______.

diez

5 Horacio Molina es un atleta que corre los cien metros planos; en las últimas carreras ha registrado los siguientes tiempos:

10.3, 11.1, 10.8, 10.6, 10.4 y 10.4 segundos

¿De cuántos datos se dispone, acerca de los tiempos empleados en correr los cien metros? ______.

Seis.

6 ¿Cuáles son las diferentes observaciones que se pueden obtener al efectuar un experimento consistente en lanzar un dado y anotar el número de la cara que queda hacia arriba? ______.

Los números del 1 al 6.

7 A cada resultado que se obtiene al realizar un experimento se le llama ______.

dato (observación)

8 Si queremos obtener cien datos de estaturas de estudiantes de Bachillerato necesitamos seleccionar ______ estudiantes de Bachillerato.
(cuántos)

cien

9 Si se quieren obtener datos de las edades de los 353 empleados de una oficina, el número mínimo de datos que se puede obtener es cero; ¿cuál es el máximo? ______.

353

10 Puesto que el resultado del experimento consistente en lanzar una moneda y anotar el lado que queda hacia arriba está sujeto al azar, *no podemos predecir con certeza* el dato que se obtendrá en cada realización del mismo.
¿Podría usted predecir con certeza la estatura de un estudiante de Ingeniería seleccionado al azar? ______.

No.

11 Cuando el resultado de un experimento está sujeto al azar, no se puede predecir con ______ el dato que se obtendrá cada vez que éste se realice.

certeza

12 Al realizar un experimento cuyo resultado depende del ______, no se puede predecir con certeza cuál será la observación.
¿Podría usted decir con certeza cuál será el resultado de un experimento consistente en extraer al azar una carta de un paquete de naipes bien barajados y anotar el número y palo observado? ______

azar, no

13 Cuando se realiza un experimento cuyo resultado depende del azar, ______ se puede predecir con certeza cuál va a ser la observación.
(no/sí)

no

14 Al efectuar un experimento es necesario definir claramente qué datos se pretende obtener. El grupo formado por el *número total* de datos que se pueden obtener, al efectuar una *secuencia exhaustiva* de experimentos, se denomina *población*.
La secuencia hipotética de caras y cruces, obtenida al efectuar un experimento consistente en lanzar una moneda un número infinito de veces, constituye una población de caras y ______.

cruces

15 Un juego de naipes americanos consta de 52 cartas. Si los naipes están barajados, cada vez que se extraiga una carta se obtendrá una _________________. El
(observación/población)
número total de cartas que se pueden extraer es 52, es decir, hay sólo 52 observaciones posibles.
Por lo tanto la población consta de _____ observaciones.
(52/9)

observación, 52

---

1 ¿De cuántos elementos se compondrá la población de estaturas de los 353 empleados de una oficina?

353

---

17 El grupo formado por el número total de datos que se pueden obtener al efectuar una secuencia exhaustiva de experimentos, se denomina _______________.
(observación/población)

población

---

18 Una población se compone de ___________ las observaciones que se pueden
(todas/parte de)
obtener al efectuar una secuencia exhaustiva de experimentos.

todas

---

1 Si una urna contiene 10,000 bolas numeradas progresivamente del 1 al 10,000, se tiene una población consistente en los números enteros del 1 al 10,000. ¿Cuántos elementos componen la población? ________.

10,000

---

20 Si se repite diez veces un experimento que consiste en extraer una bola de una urna que contiene 10,000 bolas numeradas progresivamente del 1 al 10,000, se obtendrán ________ observaciones de elementos de la *población*.
(cuántas)

diez

---

21 Si se seleccionan cinco profesores de los 300 que laboran en una Universidad para someterlos a un examen sicológico, se obtendrán cinco datos de una población que tiene ________ elementos.
(cuántos)

300

---

22 Al conjunto total de datos que se pueden obtener, al efectuar una secuencia exhaustiva de experimentos, se le denomina _______________.

población

---

23 Si se realiza hipotéticamente un experimento, consistente en lanzar un dado un número infinito de veces y anotar cada vez la cara que queda hacia arriba, tendrá una población cuyos elementos son _____.
(a/b/c)

a. números enteros del 1 al 6
b. caras y cruces
c. números del uno al infinito

a. números enteros del 1 al 6

---

24 Si en un laboratorio de botánica se realiza un experimento consistente en sembrar 50 semillas extraídas de un granero, y anotar una E de éxito, por cada semilla que germina, y una F, de falla, por cada una que no germina, se tendrán ________ datos representados por los correspondientes símbolos ___ y ___.
(cuántos)

50, E, F

---

25 Cuando una población consta de un número *limitado* (finito) de elementos se denomina *población finita*. Si se desea obtener información acerca de la capacidad didáctica de los 300 profesores de una Universidad, ¿se obtendrán acerca de una población finita?

Sí.

---

26 Si se tienen 100,000 bolas de diferentes colores en una urna y se extraen 50

bolas, anotando el color de cada una, se obtendrán datos de una población ______ (finita/incompleta)

finita

27 Si se extraen cartas de un juego de naipes bien barajados, sólo se podrán obtener 52 resultados, o sea que la población consta de 52 elementos, por lo cual se trata de una población ________.

finita

28 Si una población tiene un número infinito de elementos, ¿es una población finita?

No.

29 Los datos recolectados al lanzar un dado 10 veces y anotar la cara que queda hacia arriba no pertenecen a una población finita puesto que el dado puede lanzarse un número infinito de veces.
¿Pertenecen a una población finita los datos obtenidos al lanzar 100 veces un dado y anotar cada vez la cara que queda hacia arriba?

No.

30 Cuando una población consta de un número infinito de elementos se denomina *población infinita*.
¿Pertenecen a una población infinita los datos obtenidos al lanzar 200 veces un dado y anotar cada vez el número que queda hacia arriba?

Sí.

31 Los datos obtenidos al lanzar 200 veces un dado pertenecen a una población con un número Infinito de elementos, puesto que hipotéticamente el dado puede lanzarse un número infinito de veces, por lo tanto, es una población ________.

infinita

32 La población asociada al experimento de medir la velocidad de la luz tiene un número infinito de elementos, puesto que, hipotéticamente, podrían realizarse una infinidad de mediciones. Esta es, por lo tanto, una población ________.

Infinita

33 En vista de lo anterior una población es *finita* o *infinita*. Supongamos que se van a efectuar elecciones para presidente de la sociedad de alumnos de una Universidad con 300 estudiantes, y se quiere tener una idea de las posibilidades que cada uno de los tres candidatos postulados tiene de ganar. Para esto se pregunta a 50 estudiantes, seleccionados al azar, el nombre del candidato por el cual van a votar, obteniéndose así 50 observaciones de elementos de una población ________.

finita

34 Si se tiene una población de 1,300 calificaciones en Geografía, de una escuela de Bachillerato, ¿qué tipo de población es? ________.

Finita.

35 La población asociada al experimento consistente en medir la velocidad del sonido contiene un número ____________ de elementos, por lo cual se trata de una población infinita.

infinito

36 La población asociada al experimento consistente en lanzar indefinidamente dos dados y anotar la suma de los números que quedan hacia arriba, contiene un número _____ de elementos, por lo cual se trata de una población ________.

infinito, infinita

37 Las poblaciones se clasifican en finitas e ________.

infinitas

38 Las poblaciones finitas tienen un número __________ de elementos.

finito (limitado)

39 Las poblaciones infinitas tienen un número __________ de elementos.

infinito

40 Las poblaciones se clasifican de acuerdo con el número de elementos que las componen en __________ e __________.

finitas, infinitas

PARTE B. MUESTRA

41 La mayoría de las veces no es posible o práctico observar todos los elementos de la población, en tal caso se toma sólo una parte de ella, llamada *muestra*. Por ejemplo, si se tienen 50,000 bolas de colores en una urna, se extraen 50, y se anota el color de cada una, se obtendrá una muestra consistente en 50 observaciones. Una muestra comprende __________ de la población.
(una parte/la totalidad)

una parte

42 ¿Tiene una muestra más elementos que la población?

No.

43 Un grupo de datos extraídos de una población constituye una __________.

muestra

44 La muestra es una parte de la __________.

población

45 A veces, las observaciones son características referentes a personas tales como peso, edad, sexo o estatura. En ese caso, la población considerada no está formada por todas las personas, sino por todas las observaciones acerca de la característica particular bajo estudio.

Si una escuela tiene 1,000 estudiantes, y se requiere hacer un estudio de sus *estaturas*, la población de interés no está constituida por los 1,000 estudiantes, sino por sus 1,000__________.

estaturas

46 Al ingresar los estudiantes a una universidad, se les somete a un examen de admisión. Todas las calificaciones de dicho examen constituyen una población de observaciones numéricas. Una parte de estas calificaciones constituirá una __________

muestra

47 Para ingresar los estudiantes a una universidad se les somete a un examen de admisión; si se pretende analizar el resultado de dichos exámenes, la población bajo estudio no la componen los estudiantes, sino sus __________

calificaciones

48 Si se desea conocer la duración de los focos de una cierta marca, sería antieconómico probar todos los focos. En tal caso conviene seleccionar una parte de los focos, probarlos y anotar la duración de cada uno. El grupo de datos obtenidos constituirá una __________ de la __________, cuyos elementos son la duración de los focos.

muestra, población

49 Si se lanza un dado diez veces y se anota cada vez el número de la cara que queda hacia arriba, se tendrá una muestra consistente en __________ observaciones de una
(cuántas)
población __________.
(finita/infinita)

diez, infinita

50 Para obtener una muestra que realmente represente a una población, es necesario que *no* intervenga preferencia del muestreador por algún elemento de la población. En otras palabras, cada elemento de la población deberá tener *igual oportunidad* de ser seleccionado.

Al excluir algunos elementos de la población que se va a muestrear, éstos tendrán ____________ oportunidad que los demás elementos de la población de ser seleccionados, ya que tal oportunidad será nula.
(igual/diferente)

diferente

---

51 Alguien que desea predecir el ganador de las elecciones para Presidente de una nación, obtiene una muestra de los nombres de los candidatos por los que votarán los residentes de la capital del país. La *muestra* obtenida no será representativa, debido a que se ____________ las opiniones de los residentes de las demás ciudades y pueblos del país.
(incluyeron/excluyeron)

excluyeron

---

52 Al excluir algunos elementos de una población que se va a muestrear, éstos tendrán diferente ____________ de ser seleccionados que los demás elementos de la población.

oportunidad

---

53 Si la obtención de un dato para la muestra no afecta la oportunidad que los demás elementos de la población tienen de ser seleccionados, los elementos son *independientes*. Para que los elementos de una población sean independientes, se requiere que la selección de cualquiera de ellos ______ afecte la oportunidad que cada uno de los demás elementos tienen de ser incluidos en la muestra.
(no/sí)

no

---

54 Si se extrae al azar un as de un juego de naipes bien barajados y no se repone esa carta antes de sacar la siguiente, se alterará la oportunidad de que la segunda carta, extraída también al azar, sea otra vez un as. Entonces, las dos observaciones obtenidas ______ son independientes.
(no/sí)

---

55 Para que los elementos de una población sean ____________, se necesita que la selección de cualquiera de ellos no afecte la oportunidad que cada uno de los demás elementos tienen de ser seleccionados.

independientes

---

56 Si se lanza una vez una moneda y se observa una cruz, este resultado *no* afectará la posibilidad que hay en el siguiente lanzamiento de que salga cualquiera de los dos elementos de la población (cara o cruz); por lo tanto, las observaciones obtenidas son ____________.
(independientes/dependientes)

independientes

---

57 Si se extrae un as de un juego de naipes bien barajado, luego se repone y se baraja, el hecho de que esa carta haya sido un as, no afecta la posibilidad de que la siguiente carta sea también as. Entonces las dos observaciones son ________.

independientes

---

58 Siempre que se trabaje con una población *finita*, es necesario *reponer* cada elemento de la muestra para lograr que sus elementos sean independientes.

Si una urna contiene 1,000 bolas de colores, se tiene una población finita; ¿sería necesario reemplazar cada bola extraída antes de sacar la siguiente, para conseguir que las observaciones sean independientes?

Sí.

---

59 Supongamos que se tiene una lista de hombres casados que son votantes para unas elecciones. Si de la lista se extrae un nombre de cada mil y se pregunta a éstos y a sus respectivas esposas por el nombre del candidato por el cual van a votar, los datos obtenidos no serán ____________ puesto que las mujeres incluidas en la muestra no han sido seleccionadas independientemente de los demás elementos de la muestra.

independientes

---

60 En el ejemplo presentado en el cuadro anterior, la inclusión simultánea en una muestra de vo de un señor y de su esposa conduce a datos ________.

▼

porque en tal caso la inclusión de la opinión de la mujer no es independiente de la de su marido.
Además ______ los votantes tendrían igual oportunidad de ser selecciona-
(todos/no todos)
dos, puesto que no habría posibilidad de incluir las opiniones de las personas solteras.

independientes, no todos

---

61 Si se lanza un dado dos veces, el resultado del primer lanzamiento ______ afecta
(no/sí)
la oportunidad que tienen los demás elementos de la población de quedar hacia arriba en el siguiente lanzamiento. En tal caso, los dos datos obtenidos son ______.

no, independientes

---

62 Se pretende obtener una muestra de las calificaciones en Historia de los estudiantes de una escuela. El muestreo se realizará extrayendo al azar de una urna varias tiras de papel, cada una de las cuales tiene una calificación anotada, pero en una tira aparecen impresas tres calificaciones. Si se selecciona ésta última, necesariamente quedarán incluidas en la muestra las tres calificaciones, por lo cual, la oportunidad de seleccionar cualquiera de estas calificaciones no es ______ ______ de las otras escritas en el papel.
(dependiente/independiente)

Independiente

---

63 Cuando se selecciona una muestra de una población en la que todos los elementos son *independientes* y tienen *igual oportunidad* de ser seleccionados, se tiene una *muestra aleatoria o representativa.*
En una muestra aleatoria, todos los elementos de la población son ______ y tienen ______ oportunidad de ser seleccionados.

independientes, igual (la misma)

---

64 Si se quisiera muestrear el grado de inteligencia de los estudiantes de una escuela, y si dicha muestra se seleccionara extrayendo de una urna los nombres de cada alumno, escritos en una tira de papel, reponiendo cada tira extraída, todos los alumnos tendrían igual oportunidad de ser seleccionados y los elementos serían independientes, por lo cual, se obtendría una muestra ______.

aleatoria (representativa)

---

65 Si en el muestreo del cuadro anterior se escribiera el nombre del alumno Horacio Molina en diez papeles distintos, ¿se obtendría una muestra aleatoria?

No.

---

66 Si el nombre de un alumno, tal como el de Horacio Molina, se escribiera en varias tiras de papel, la muestra extraída no sería aleatoria porque él tendría ______ ______ oportunidad de ser seleccionado que los demás alumnos.
(igual/menor/mayor)

mayor

---

67 Cuando algunos elementos de la población no tienen la misma oportunidad que los demás de ser incluidos en una muestra, dicha muestra no es ______.

aleatoria (representativa)

---

68 Una muestra es *sesgada* cuando no todos los elementos de la población tienen igual ______ de ser seleccionados, o cuando no son ______.

oportunidad, independientes

---

69 Cuando se extrae una muestra con datos independientes, pero con diferente oportunidad de ser seleccionados, la muestra es ______.
(aleatoria/sesgada)

sesgada

---

70 Cuando se extrae una muestra en la que todos los elementos de la población tienen la misma oportunidad de ser seleccionados, pero contiene observaciones que no son independientes, se obtiene una muestra ______.

sesgada

---

71 Para que una muestra sea aleatoria, es necesario que cada elemento de la población correspondiente tenga ______ oportunidad que los demás de ser seleccionado, y que los datos sean independientes.

igual (la misma)

---

72 Si la muestra no es aleatoria, se dice que es ____________.

sesgada

---

73 Para que una muestra sea aleatoria, se requiere que todos los elementos de la población tengan ____________ oportunidad de ser seleccionados y que, además, las observaciones sean ____________.

Igual (la misma), independientes

---

74 Una muestra es aleatoria o es ____________

sesgada

---

75. Una muestra obtenida de una población en la que cada elemento tiene la misma oportunidad de ser seleccionado, y además es independiente de los demás, se llama ____________ ____________.

muestra aleatoria (representativa)

---

76 Si se extrae una muestra de una población, y el muestreo se realiza en tal forma que *no todos* los elementos de la población tengan la misma oportunidad de ser incluidos, se obtiene una muestra ____________.

sesgada

---

77 Un procedimiento para obtener una muestra aleatoria consiste en asignarle un *número diferente* a los diversos elementos de la población y extraer al azar de una urna, por ejemplo, algunas bolas que tienen un solo número anotado, regresando cada una antes de extraer la siguiente; cada número seleccionado identificará a un elemento de la población que debe incluirse en la muestra. Para que la muestra sea aleatoria es necesario que cada bola tenga ____________ (la misma/diferente) oportunidad de ser seleccionada.

la misma

---

78 Si se utiliza el procedimiento anterior para obtener una muestra aleatoria, se requiere que haya una bola, con un número anotado, correspondiente a cada elemento de la población. Si la población tiene 500 elementos se requiere que la urna contenga ________ (cuántas) bolas.

500

---

79 Es conveniente, para ahorrar tiempo durante la selección de la muestra, que la numeración sea progresiva, comenzando con el número uno.
Si se procede en esta forma para una población de 500 elementos, los números asignados a cada uno irán del 1 al ____.

500

---

80 Si se utiliza el procedimiento de la urna para seleccionar una muestra aleatoria, cada bola de la urna representa a un elemento de la ____________ y cada bola extraída identifica a un elemento de la muestra.

población

---

81 Existen *tablas de números aleatorios* para facilitar la selección de una muestra aleatoria. Estos *números* se calculan en tal forma que todos ellos tengan la misma oportunidad de ser seleccionados y sean independientes. Entonces, si se utiliza una tabla de números aleatorios para obtener una muestra, ésta resulta ____________ (sesgada/aleatoria).

aleatoria

---

82 Si a cada elemento de una población se le asigna un número diferente, y se seleccionan diez números de una tabla de números aleatorios, se obtiene una muestra ____________, la cual contiene ________ (cuántos) datos.

aleatoria, diez

---

83 Mediante el uso de una tabla de números aleatorios, se obtiene una muestra representativa (o aleatoria) porque todos los miembros de la población tienen ____________ (igual/diferente) oportunidad de ser seleccionados y además son independientes entre sí.

igual

---

84 Para facilitar el control, es conveniente utilizar numeración corrida para identificar a los elementos de una población. Entonces, si se tiene una población finita de 360 elementos, los números que les corresponden son 1, 2, . . . hasta ________.

360

Si la población que se va a muestrear consta de 6,600 elementos y se utiliza numeración corrida para identificarlos, se requieren números aleatorios con cuatro *dígitos* (cuatro cifras) para seleccionar la muestra aleatoria. Si la población constara de 50 elementos, se necesitarían números aleatorios con ________ dígitos para seleccionar los elementos de una muestra aleatoria. (1/2/50)

dos

---

86 El número 25,436 tiene cinco cifras o dígitos. El 316,210 tiene seis dígitos. ¿Cuántos dígitos tienen los números 4,625 y 27?

cuatro, dos

---

87 El número 365 tiene ________ dígitos. El número 31,100 tiene cinco ________.

tres, dígitos

Si los números de las tablas tienen cinco dígitos y sólo necesitamos cuatro, hay que excluir uno de ellos. Por ejemplo, si los números aleatorios son:

16408
18629
73115

entonces, para seleccionar la muestra *representativa*, se excluye el dígito del extremo *izquierdo* y sólo se consideran los números 6408, 8629, 3115; si sólo se necesitaran dos dígitos (dos cifras), usaríamos los números ________ ________. (408, 629, 115/08, 29, 15)

08, 29, 15

---

89 *i* nún *s* con dígi y se e una tabla que contiene números con cinco dígitos, hay que excluir las tres cifras de la izquierda y leer sólo las últimas dos cifras, así:

164 08
186 29
731 15

Si se necesitaran números con un solo dígito se excluirían las ________ (cuántas) cifras de la izquierda y se leerían los números 8, 9, 5.

cuatro

---

90 Supongamos que se tienen los siguientes números aleatorios:

345179
964736
146328

Si para obtener una muestra aleatoria se necesitaran números de tres dígitos, se usarían los números 179, 736, 328. Si se necesitaran de cinco dígitos, se emplearían los números ______, ______, ______.

45179, 64736, 46328

---

91 Si se tiene una población con 317 elementos, se necesitan números de tres dígitos. Por ejemplo, el número 17301, de la tabla de números aleatorios, identifica el elemento número 301, ya que se excluyen las dos cifras de la izquierda. Por la misma razón, los números 25213 y 73115 identifican a los elementos 213 y 115. ¿A qué elemento corresponden los números 14101, 21003 y 37085?

101, 3 y 85 respectivamente

---

92 Si la población que se va a muestrear consta de 6,600 elementos, se necesitan números con _____ dígitos para seleccionar la muestra. El número 16408 corresponde al elemento ________. (08/6408/16408)

cuatro, 6408

---

93 Si se tiene una población de 6,600 elementos, a a u se le asigna un número

progresivo del 1 al ____, el mayor número de la población es, entonces, el ____.

6,600, 6,600

Sea una población de 6,600 elementos. Si se tiene el número 18629 de una tabla de números aleatorios, simplemente se considerará el 8,629 que resulta de excluir la cifra de la izquierda. Este número es mayor que el más grande de la población (6,600), por lo cual no corresponde a ningún elemento de la misma. ¿Identifica el número 24160, leído en la tabla, algún elemento de esa población?

Sí

95 El número 18629 *no* representa a ningún elemento de una población de 6,600 elementos, puesto que *no* existe ninguno identificado con el número 8,629. ¿Representa el número 79639, leído en una tabla de números aleatorios, algún elemento de esa población?

No.

9 El número 79639, leído en la tabla de números aleatorios, no representa a ningún elemento de una población de 6,600 elementos, puesto que 9639 es mayor que 6,600.
¿Corresponde el número aleatorio 25436 a algún elemento de esa población?

Sí.

97 Cuando se obtiene un número aleatorio que no representa a ningún elemento de la población, se desecha dicho número, y se continúa con el muestreo. Si se tiene una población de 6,600 elementos, de la cual se quiere obtener una muestra representativa mediante el uso de una tabla de números aleatorios, y si se encuentra con la siguiente porción de la tabla:

16408
18629
73115

a. ¿Qué elemento de esa población representa el número 16408?
b. ¿Representa número 18629 algún elemento?
c. ¿Representa número 73115 algún elemento?
d. ¿Qué elemento representa el número 73115?

a. (Al) 6,408.
b. No.
c. Sí.
d. (Al) 3,115

98 Si el número extraído fuese el 57591, ¿representaría éste a algún elemento de una población de 500 elementos?

No.

99 ¿Qué haría en caso de que un número no representara a un elemento de la población? ____ (a/b)

a. Detener el muestreo
b. Desechar el número y proseguir el muestreo

b. Desechar el número y proseguir el muestreo

100 La manera de escoger los números aleatorios es arbitraria, puesto que todos tienen la misma oportunidad de ser seleccionados. Se puede convenir, por ejemplo, en obtener la muestra usando sólo los números de los renglones pares, o los de las columnas pares, etcétera.
Si se siguiera este último procedimiento, los primeros tres números que se leerían en la columna 2 de la tabla 1.1 de números aleatorios, serían: 81899, 81953 y 35101. (Vea la tabla 1.1 verifique estos números y diga cuáles serían los tres siguientes.)

10703, 83946, 35006

101 Si se usan los números anotados en los renglones pares para obtener la muestra aleatoria, los primeros tres números del segundo renglón son: 18629, 81953, y 05520 (vea la tabla 1.1 y verifíquelos).
¿Cuáles son los tres siguientes?

91962, 04739, 13092

102 Si se emplean los renglones pares de la tabla 1.1, empezando por el renglón 2, para extraer una muestra aleatoria de una población de 6,600 elementos, el

primer número aleatorio es el _______ (vea la tabla 1.1). ¿Representa este número a algún elemento de la población?

18620, No.

103. El segundo número aleatorio del segundo renglón es el _______ (vea la tabla 1.1). ¿Representa este número a algún elemento de una población de 6,600 elementos?
¿A cuál?

81953, Sí, (Al) 1053.

104 Entonces, el primer elemento de la muestra es el identificado con el número 1953. Prosiguiendo el muestreo en la misma forma (leyendo en el renglón 2 de la tabla 1.1), el segundo elemento de la muestra será el identificado con el número 5520.
¿Cuál será el tercero?
¿Cuál será el cuarto?

1962, 4730

105 Si la tabla de números aleatorios se usa en otra forma, se obtendrá una muestra distinta. Si se utilizan las columnas pares de la tabla 1.1 (empezando con la columna 2), ¿cuáles serán los primeros tres miembros de una muestra extraída de una población de 6,000 elementos?

1800, 1953, 5101

1 Para usar una tabla de números aleatorios es necesario asignar un _______ diferente a cada miembro de la población.

número

107 Mediante el uso de una tabla de números aleatorios es posible obtener una muestra _______.

aleatoria (representativa)

108 Los datos originales que constituyen una muestra se conocen como *datos bási-*

Tabla 1.1
Números aleatorios

renglón \ columna	1	2	3	4	5	6	7	8	9	10	11
1	10480	01059	04153	53381	79401	21438	03035	97350	30743	31223	57647
2	10629	01953	06520	91352	34739	13292	37662	24922	94730	06496	35090
3	73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153
4	57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812
5	30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668
6	16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817
	96773	20206	42559	78985	05300	22164	24369	54224	35083	19687	11052
	38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318
	31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614
	78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806
11	03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833
	74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970
	09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375
	42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689
	16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510
16	21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240
	21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995
	55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525
	44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911
	91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203
21	91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867
	50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405
	65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202
	27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142
	37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330
26	11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081
	37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651
	46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345
	30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339
	63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377
31	82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870
	21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194
	60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535
	43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434
	97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534
36	03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303
	79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076
	85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327
	18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398
	08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639
41	79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732
	92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022
	23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543
	09915	96306	05908	97901	28395	14186	80703	70426	70426	75647	76310
	42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799
46	46764	86273	93017	31204	36692	40202	35275	57306	63003	55543	53203
	03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216
	86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548
	38534	01715	94964	87288	65680	43772	39560	12918	86537	62738	19636
	06652	11645	40876	20507	62598	95715	03585	64303	36732	70185	56008

cos, ya que sirven de base para obtener otro tipo de información (de datos) acerca de la población bajo estudio.
Los datos __________ son los datos originales que constituyen una muestra.

básicos

---

109 La secuencia original de datos: 1,3,6,2,1,4,5,2, obtenidos al lanzar un dado ocho veces, constituye un conjunto de __________ básicos.

datos

---

110 Los datos originales de una muestra se conocen como datos ________.

básicos

---

111 Las observaciones originales de las estaturas de 60 estudiantes de Bachillerato, constituyen un grupo de ________ ________.

datos básicos

---

112 La *Estadística* es la rama de las Matemáticas que estudia las reglas para recolectar, organizar y procesar datos, y para utilizarlos con objeto de extraer conclusiones acerca de una población.
Al indicar la forma de obtener una muestra aleatoria, mediante el uso de números aleatorios, se dio una regla para recolectar datos. Cuando se seleccionan los elementos de la muestra se están __________________ los datos.
(organizando/recolectando)

recolectando

---

113 La Estadística proporciona, por lo tanto, las reglas para trabajar con muestras ______________.
(sesgadas/aleatorias)

aleatorias

---

114 La rama de las Matemáticas que estudia las reglas para trabajar con muestras aleatorias se llama ____________.

Estadística

---

115 Para hacer un estudio estadístico de una población infinita, es forzoso analizar ______________________.
(una muestra/la población completa)

una muestra

---

116 No sólo la imposibilidad de obtener información completa acerca de una población, como sucede en una población infinita, conduce a analizar una muestra de la misma: existen también razones de *tiempo* y *dinero*. ¿Sería razonable obtener la opinión de cada uno de los votantes de la República Mexicana, antes de una elección presidencial, para predecir el ganador?

No.

---

117 Cuando por razones de tiempo o dinero no se puede estudiar una población completa, podemos recurrir a la __________________, ya que es la rama de las Matemáticas que nos proporciona las reglas para extraer (inferir) conclusio acerca de la población mediante el uso de una muestra aleatoria.

Estadística

---

118 Al realizar un estudio estadístico, la muestra que se utilice debe ser ____________.
(aleatoria/sesgada)

aleatoria

---

119 Hasta ahora, sólo se han dado las reglas para recolectar datos.
Más adelante se darán las reglas para organizarlos y procesarlos.
Todas estas reglas sirven para *describir* los datos, y son tratadas por una rama la Estadística llamada *Estadística Descriptiva*.
La Estadística Descriptiva proporciona las reglas para ______________ los datos.

describir

---

120 La rama de la Estadística que proporciona las reglas para describir los datos llama ______________________.
(Estadística Descriptiva/Algebra)

Estadística Descriptiva

---

121 La rama de la Estadística que proporciona las reglas para recolectar, organizar y procesar los datos, es decir, para describirlos, se llama ______ ______.

Estadística Descriptiva

122 En la Estadística Descriptiva se estudian las reglas para recolectar, organizar y ______ los datos.

procesar

123 En la Estadística Descriptiva se estudian las reglas para recolectar, ______ y procesar los datos.

organizar

124 Para recolectar, organizar y procesar los datos se utilizan reglas que proporciona la rama de la Estadística llamada ______ ______.

Estadística Descriptiva

125 En Estadística no sólo se estudian las reglas para describir los datos, sino también las reglas para analizarlos, extraer conclusiones y tomar decisiones. La rama de la Estadística que proporciona las reglas para analizar los datos, extraer conclusiones y tomar decisiones, se llama *Inferencia Estadística*. La Estadística se divide, por lo tanto, en Estadística Descriptiva e ______ ______.

Inferencia Estadística

126 La ______ Estadística da las reglas para *inferir* características de la población mediante el uso de una muestra.

Inferencia

127 La rama de la Estadística que proporciona las reglas para extraer conclusiones acerca de una población, con base en la información contenida en una muestra, se llama ______ ______.

Inferencia Estadística

128 Del estudio de la Estadística se aprende que las inferencias acerca de una población serán de confianza cuando los datos se recolecten, organicen y procesen de acuerdo con ciertas reglas. La rama de la Estadística que proporciona dichas reglas se llama ______ ______.

Estadística Descriptiva

129 Al emplear la información extraída de una muestra que ya ha sido organizada y procesada, para investigar cuánto se espera que duren las llantas de un cierto tipo y marca, se tendrá un problema de ______.
(Estadística Descriptiva/Inferencia Estadística)

Inferencia Estadística

130 La descripción de los datos usados en un problema para inferir cierta característica de una población, corresponde a la rama de la Estadística llamada ______ ______.

Estadística Descriptiva

131 La Estadística se divide en dos ramas: una que da las reglas para describir los datos, llamada ______ ______, y otra que da las reglas para inferir ciertas características de la población, llamada ______ ______.

Estadística Descriptiva, Inferencia Estadística

132 ¿Enseña la Estadística las reglas para describir los datos e inferir, a partir de la muestra, característica de la población?

Sí.

133 En este libro sólo se dan las reglas para describir datos; por lo tanto, es un libro de ______ ______.

Estadística Descriptiva

Revisión

134 Cuando una población es infinita o, cuando por razones prácticas, no es conve-

niente estudiarla totalmente, se obtiene un grupo de observaciones de la misma, el cual constituye una ________.

muestra

135 Para que una muestra sea aleatoria se requiere que todos los miembros de la población tengan ________ oportunidad de ser seleccionados y que sean ________________ entre sí.

igual, independientes

136 El conjunto de todas las observaciones posibles, cuando se realiza una secuencia exhaustiva de experimentos, se denomina ________.

población

137 La Inferencia Estadística utiliza las muestras para inferir características de la ________.

población

138 La rama de las Matemáticas que da las reglas para describir los datos de una muestra, e inferir características de la población a partir de la muestra, se llama ____________.

Estadística

139 La Estadística se divide en dos partes, una que da las reglas para describir los datos, llamada ________ ________, y otra que da las reglas para inferir características de la población, llamada ________ ________

Estadística Descriptiva, Inferencia Estadística

140 De acuerdo con la cantidad de elementos que contienen, las poblaciones se clasifican en ________ e ________.

finitas, infinitas

141 Una muestra obtenida en tal forma que no todos los elementos de la población correspondiente tienen igual oportunidad de ser seleccionados, o no son independientes entre sí, se llama ________ ________.

muestra sesgada

142 Una tabla de ________ ________ es útil, entre otras cosas, porque facilita la obtención de una muestra aleatoria.

números aleatorios

143 Empezando con el renglón 46 de la tabla 1.1, obtenga una muestra aleatoria cinco datos, de una población de 4,000 elementos.

3017, 1204, 202, 3003, 3203

EXAMEN

1. *Al resultado de un experimento se le llama ______________.*

2. *El grupo de todos los resultados posibles al realizar una secuencia exhaustiva de experimentos, se denomina ____________.*

3. *Un grupo de datos constituyen una ____________.*

4. *Relacione los nombres de la columna de la izquierda con las descripciones que aparecen a la derecha, y que están asociadas al experimento que consiste en extraer una carta de un juego de 52 naipes, sin regresar cada carta extraída.*

1. Población	a. Cada uno de los resultados
2. Muestra	b. Menos de 52 observaciones
3. Dato	c. 52 observaciones .

5. *Dependiendo del número de elementos que contienen, las poblaciones se clasifican en __________ e __________.*

6. *Diga a qué tipo de población se asocian los siguientes experimentos:*

   a. *Lanzamiento de un dado.*
   b. *Extracción de una bola de una urna que contiene 150 blancas y 50 negras.*
   c. *Medición de la longitud de espigas de trigo.*
   d. *Medición del peso de 500 cabezas de ganado vacuno.*
   e. *Registro del número de defectos en todas las piezas de casimir que se producen en una fábrica.*

7. *Una muestra puede ser ____________ o ____________.*

8. *¿Cuáles son los dos requisitos que necesita satisfacer una muestra para ser aleatoria?*

9. *Una tabla de __________ ____________ sirve para obtener una muestra aleatoria.*

10. *Una hacienda exportará 3,200 cabezas de ganado vacuno; para tener una idea de la ganancia que se obtendrá, se seleccionan al azar 10 animales y se pesan; para lo cual, se numeran en orden progresivo empezando por el 1.*

    a. *¿Cuántos elementos tiene esta población?*
    b. *¿La población es finita o infinita?*
    c. *La muestra consta de ____ observaciones.*
    d. *¿Cuáles son las observaciones que interesan; los animales o sus pesos?*
    e. *Use la tabla 1.1 de números aleatorios para obtener la muestra aleatoria se desea. Use los renglones rom empezando con el renglón 11. (10 puntos).*

11. *¿Con qué nombre se conoce una muestra que no es aleatoria?*

12. *Las observaciones originales que constituyen una muestra se conocen como datos ________.*

13. *Para que los estudios estadísticos sean válidos, se requiere que las muestras ________.*

14. *La rama de las Matemáticas que estudia las reglas para trabajar con muestras aleatorias, se llama __________.*

15. *La Estadística se clasifica en Estadística __________ e __________ ________.*

16. *La rama de la Estadística que proporciona las reglas para recolectar, organizar y procesar los datos, se llama __________ __________.*

17. *La rama de la Estadística que estudia las reglas para inferir características de una población, a partir de una muestra, se denomina __________ __________.*

TOTAL: 40 puntos.

RESPUESTAS

1. *dato (observación)*

2. *población*

3. *muestra*

4. *1-c*
   2-b
   3-a

5. *finitas*
   infinitas

6. a. *Infinita*
   b. *Finita*
   c. *Infinita*
   d. *Finita*
   e. *Infinita*

7. *aleatoria*
   sesgada

8. 1. *Que todos los elementos de la población tengan la misma oportunidad de ser seleccionados.*
   2. *Que las observaciones sean independientes.*

9. *números aleatorios*

10. a. *3,200*
    b. *Finita*
    c. *10*
    d. *Sus pesos*
    e. *2825, 1284, 903, 795, 2648, 1125, 1744, 1959, 33, 1681.*

11. *Muestra sesgada.*

12. *básicos*

13. *aleatorias*

14. *Estadística*

15. *Estadística Descriptiva*
    Inferencia Estadística

16. *Estadística Descriptiva*

17. *Inferencia Estadística*

UNIDAD II VARIABLES

PREFACIO

En esta unidad se estudian las *variables*, y se les clasifica en determinísticas y aleatorias; a su vez, cada uno de estos grupos se subdividen en variables escalares, nominales, discretas y continuas. Se hace, además, un breve estudio de las técnicas de medición.

1 Si se anotan las estaturas de los estudiantes de un curso de Estadística, es casi seguro que las medidas no coincidirán, es decir, que variarán de unos alumnos a otros. Si además se pesan esos mismos alumnos, es muy posible que los pesos sean ________
(iguales/diferentes)

diferentes

---

2 Si se examina de Estadística Descriptiva a los alumnos que ya tomaron este curso, es casi seguro que ________ las calificaciones serán iguales; es decir, habrá variación de unas a otras.
(todas/no todas)

no todas

---

3 Si se anota el sexo de los niños nacidos en un hospital durante un mes, se verá que unos son hombres y otros son mujeres; es decir ________ son del mismo sexo.
(todos/no todos)

no todos

---

4 Cuando una característica de personas o cosas puede tomar diferentes valores se llama *variable*.

Los niños nacidos en un hospital durante un año serán de ________ sexo.
(igual/diferente)

diferente

---

5 Puesto que el sexo varía de unas personas a otras, el sexo es una ________

variable

---

6 Es casi seguro que las calificaciones obtenidas en el examen de admisión de una Universidad, por los alumnos que ingresen en 1973, serán ________ entre sí.
(diferentes/iguales)

diferentes

---

7 Las calificaciones del examen de admisión a una Universidad presentado por los aspirantes a ingresar en ella varían de unas a otras; por lo tanto, dicha calificación es una ________.

variable

---

8 Supongamos que se efectúa el siguiente experimento:
Se vacunan 10 ratones y luego se exponen, durante cierto tiempo, a la enfermedad contra la cual fueron vacunados. Por cada ratón que adquiera la enfermedad se anota una F, de falla, y por cada uno que no se contagie se anota una E, de éxito. Supongamos que los resultados son los siguientes: F,E,E,F,E,E,E,F,F,E. Se observa que no todas las anotaciones son iguales; es decir, varían de unas a otras. Por lo tanto, la efectividad del suero es una ________.

variable

---

9 Si se lanza varias veces una moneda que tiene un águila de cada lado y se anota cada vez la cara que queda hacia arriba, se obtendrá una secuencia de observaciones de la forma: águila, águila, . . . , etcétera; es imposible que las observaciones varíen de un lanzamiento a otro y, por lo tanto, ________ se trata de una variable.
(no/sí)

no

---

10 Algunas variables asumen *valores numéricos*; tal es el caso de la estatura y peso de los estudiantes de este curso de Estadística Descriptiva.
Las calificaciones obtenidas por cinco alumnos en un examen de Matemáticas fueron 5,7,5,10,8. Esta variable asume, por lo tanto, valores ______

numéricos

---

11 Si con el fin de ejercitarse para una competencia, un atleta corre varias veces los 100 metros planos, la duración de la carrera será distinta en cada ocasión. Por lo tanto, el tiempo que necesita un atleta para correr los 100 metros planos es una ______ que asume ______ numéricos.

variable, valores

---

12 Si se observa la lista de los ingresos anuales de las familias de la ciudad de México, se concluye que éstos varían de unas familias a otras; por lo tanto, el nivel de ingresos por familia es una ______ que asume valores ______

variable, numéricos

---

13 Una variable que sólo toma valores numéricos se llama *variable escalar*.
La estatura de los estudiantes de este curso de Estadística ______ (es/no es) una variable escalar.

es

---

14 Las variables escalares son aquellas que asumen valores ______

numéricos

---

15 ¿Es escalar la variable "año de nacimiento de los profesores de la Universidad de México"?

Sí.

---

16 La variable "año de nacimiento de los profesores de la Universidad de México" sí es una variable ______ porque sólo puede tomar valores ______

escalar, numéricos

---

17 La variable "resultado del lanzamiento de un dado", cuando el dado tiene número anotado en cada cara, es una variable ______ porque sólo p asumir valores numéricos.

escalar

---

18 Una variable que sólo puede tomar valores numéricos se denomina ______

variable escalar

---

19 Existen variables que *no* pueden tomar valores numéricos. Tal es el caso, de la variable "sexo de los estudiantes de la Universidad de Guanajuato"; ésta puede tomar solamente los valores masculino y ______.

femenino

---

20 Existen variables que no pueden tomar valores numéricos sino sólo valores que pueden designarse mediante nombres o atributos. Por ejemplo los ______ (nombres/numeros) que puede asumir la variable "resultado de un tratamiento para reducir de peso", son éxito o fracaso.

nombres

---

21 Cuando una variable sólo puede tomar valores designables por *nombres* o *atributos* se dice que toma *valores nominales*. Los valores ______ (numéricos/nominales) de la variable "resultado de lanzar una moneda", son cara y cruz.

nominales

---

22 La variable "color de las bolas contenidas en una urna" toma sólo valores ______

nominales

---

23 Una variable que no puede tomar valores numéricos, forzosamente toma valores ______

nominales

24 Una variable que sólo puede asumir valores nominales se llama variable *nominal* o *no escalar*.
¿Es la variable "deporte predilecto de los profesores de la Universidad de México" una variable nominal?

Sí.

25 La variable "deporte predilecto de los profesores de la Universidad de México" no es una variable escalar, porque sólo puede tomar valores nominales. Por lo tanto, esta variable es ______________.

nominal (no escalar)

26 Diez personas demasiado gordas fueron sometidas, durante un mes, a un proceso para reducir de peso. Por cada persona que redujo 10 kilos se anotó una E, de éxito, y por cada una que redujo menos de 10 kilos se anotó una F, de falla. Se obtuvieron los siguientes resultados: E,E,F,E,F,F,F,E,E,E. La variable "resultado del tratamiento" sólo puede tomar los valores nominales E o F; por lo tanto, es una variable ________________

nominal

27 La variable "resultado del lanzamiento de una moneda" es una variable nominal (o no escalar), porque ______ puede tomar valores numéricos.
(sí/no)

no

28 Una variable que sólo puede tomar valores nominales se conoce como variable ____________.
(escalar/nominal)

nominal

29 La variable "resultado del lanzamiento de una moneda" es una variable ____________.
(escalar/nominal)

nominal

30 ¿Es la variable "estado civil de los habitantes de Parral, Chihuahua" una variable escalar? ______________.

No.

31 La variable "estado civil de los habitantes de Parral, Chihuahua", no es una variable escalar porque sólo puede tomar valores ___________.

nominales

32 Una variable es escalar o ________________.

nominal (no escalar)

33 Las variables nominales sólo pueden tomar un número *limitado*, o *finito* (no infinito) de valores claramente diferenciados. La variable nominal "sexo" solamente puede tomar ________ valores: masculino o femenino.
(cuántos)

dos

34 La variable nominal "resultado de un tratamiento para reducir de peso 10 kilos o más" medida en términos de éxito o fracaso, sólo puede tomar dos valores distintos: éxito o __________.

fracaso

35 La variable nominal "color de las bolas contenidas en una urna", si se sabe que hay bolas verdes, rojas, blancas y negras, sólo puede tomar ________ valores distintos.
(cuántos)

cuatro

36 Algunas variables escalares pueden tomar también un número limitado (finito) de valores ______ claramente diferenciados. Por ejemplo, la variable "resultado de lanzar un dado", sólo puede tomar los valores 1,2,___, 4, 5 y ___.

numéricos, 3, 6

37 La variable escalar "número de hijos de un matrimonio", sólo se puede medir mediante los números enteros 1, 2, etcétera, siendo en todos los casos un número ________.
(finito/infinito)

finito

38 Si a un grupo de estudiantes se le hace un examen de Filosofía que consta de diez preguntas concretas, las calificaciones serán números enteros del uno al diez. Por lo tanto, la variable escalar "calificación en el examen de Filosofía" sólo puede tomar ________ valores distintos, y tiene un número finito de valores.
(cuántos)

diez

39 Para que una variable escalar tome un número finito de valores, no es forzoso que éstos sean números enteros. La variable puede también tomar valores fraccionarios, siempre que estén perfectamente definidos. Si se hace un experimento en el cual la variable puede tomar los valores $0, \frac{1}{2}, 2, 2\frac{1}{2}, \ldots, 10$, ésta tomará un número ________ de valores.
(finito/infinito)

finito

40 En algunas ocasiones una variable escalar puede tomar un número infinito de valores, que se pueden numerar del uno al infinito. En este caso los valores están claramente diferenciados, ya que son 1, 2, 3, ___, ___, ___, 7, 8, ..., n, etcétera.

4, 5, 6

41 Sea la variable escalar "número necesario de lanzamientos de una moneda para que aparezca por primera vez una cruz", los valores que esta variable puede tomar son 1, 2, 3, ..., n, ..., ∞, donde el símbolo ∞ significa infinito; por lo tanto, es una variable escalar que puede tomar un número ________ de valores que se pueden numerar.
(finito/infinito)

infinito

42 Cuando una variable puede tomar un número *finito, o infinito pero numerable* (que se puede numerar) de valores, se llama *variable discreta*.

Una variable nominal toma forzosamente un número finito de valores, por lo cual, es una variable ________

discreta

43 Sea la variable escalar "número de personas que usan zapatos"; ésta sólo puede tomar un número finito de valores claramente diferenciados; es decir, los números 1, 2, 3, ..., n, y, por lo tanto, se trata de una variable ________

discreta

44 La variable escalar "número necesario de lanzamientos de un par de dados para que la suma de los resultados sea igual a siete", puede asumir los valores 1, 2, 3, ..., n, ..., ∞, por lo cual puede tomar un número ________ numerable de valores, siendo entonces una variable ________
(finito/infinito)

infinito, discreta

45 Una variable discreta puede tomar un número ________ de valores, o un número infinito pero ________.
(finito/infinito)
(numerable/no numerable)

finito, numerable

46 Una variable que puede tomar un número finito o un número infinito numerable de valores, se denomina variable ________.

discreta

47 La variable "sexo de los niños nacidos en un hospital" es una variable nominal, por lo cual sólo puede tomar un número finito de valores. ¿Es ésta una variable discreta?

Sí.

48 Las variables nominales sólo pueden tomar un número finito de valores, por lo tanto, son variables ________

discretas

49 Una variable que puede tomar un número infinito de valores *no numerables* (que no se pueden numerar) se llama *variable continua*. Entonces, una variable es discreta o es ______________

continua

50 Al observar el velocímetro de un automóvil, se nota que, al acelerar, la manecilla se mueve ______________ (a saltos/en forma continua), por lo cual, la variable "velocidad del automóvil", puede tomar valores aproximados a cualquier fracción de kilómetro por hora. ¿Es ésta una variable continua?

en forma continua, Sí.

51 Supongamos que se mide la estatura de un individuo y el resultado es 171 cm; en realidad ______ (no/sí) mide exactamente 171 cm, sino que, por razones de aproximación debidas al instrumento de medición, se le asigna ese valor.

no

52 La variable escalar "estatura de los estudiantes de este curso de Estadística Descriptiva" ________ (es/no es) una variable discreta porque toma valores aproximados a cualquier fracción de centímetro, los cuales no pueden numerarse, ya que sus valores reales pueden situarse a lo largo de una línea recta continua. ¿Es ésta una variable continua?

no es. Sí.

53 La variable escalar "peso de los estudiantes de la Universidad de Chihuahua", puede asumir cualquier número positivo, es decir, puede tomar un número infinito no numerable de valores, por lo que se trata de una variable ______________ (discreta/continua)

continua

54 Una variable continua puede tomar valores que no están restringidos a números aislados, sino que puede tomar *cualquier* valor numérico comprendido en un cierto intervalo de valores. La cantidad de agua que fluye por un río durante un día es una variable que puede tomar cualquier valor mayor o igual que cero, por lo cual es una variable ______________ (discreta/continua).

continua

55 Una variable es discreta o es ______________

continua

56 Una variable continua sólo puede tomar valores numéricos, por lo cual es forzosamente una variable ______________ (escalar/nominal).

escalar

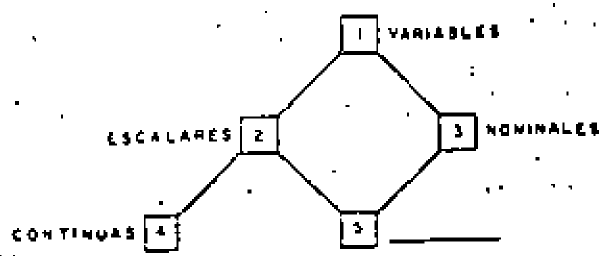
57 Por lo visto anteriormente, las variables se clasifican, de acuerdo con el tipo de valores que pueden tomar, ya sean numéricos o nominales, en escalares y ______________ respectivamente.

nominales

58 Las variables nominales sólo toman un número finito de valores, por lo que forzosamente son discretas. Algunas variables escalares pueden tomar un número finito o infinito numerable de valores; otras, un número infinito no numerable de valores. De acuerdo con esto, las variables escalares se clasifican en discretas y ______________

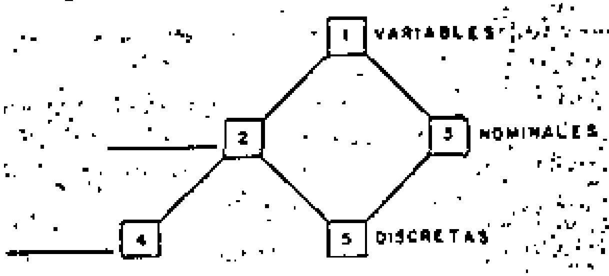
continuas

59 Complete el siguiente esquema de clasificación



[5] DISCRETAS

60 Complete el siguiente esquema de clasificación

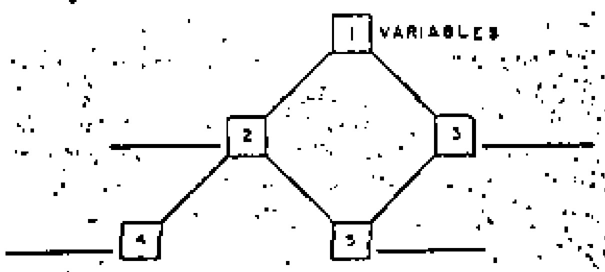


[2] ESCALARES, [4] CONTINUAS

61 Una variable nominal es forzosamente discreta, pero una variable escalar puede ser ______________ o ______________.

discreta, continua (en cualquier orden).

62 Complete el siguiente esquema de clasificación:



[2] ESCALARES, [3] NOMINALES, [4] CONTINUAS, [5] DISCRETAS

63 Al realizar un experimento se obtiene un dato u observación que forma parte de una muestra de la población cuyas características se tratan de inferir. Las observaciones, en general, varían de un experimento a otro, por lo cual ______ (son/no son) valores de una variable.

son

64 Por lo visto en el cuadro anterior, siempre existe una variable asociada a cada tipo de experimento, y los elementos de la muestra ______ (son/no son) valores que dicha variable ha tomado. Entonces, la variable es precisamente la característica bajo estudio.

son

65 Sea el siguiente experimento: se extraen, al azar, cinco elementos de una urna que contiene bolas de colores, con el objeto de inferir cuántas hay en cada color. Las observaciones son: dos rojas, una negra, una blanca y una azul; estas observaciones ______ (son/no son) los elementos de la muestra y ______ (son/no son) valores que ha tomado la variable nominal "color de las bolas contenidas en la urna".

son, son

66 Todos los elementos diferentes de la población asociada a un experimento constituyen los diversos valores que puede tomar la variable asociada al mismo experimento. Por ejemplo, la población asociada al lanzamiento de una moneda es una secuencia infinita de caras y cruces, la cual sólo tiene dos elementos diferentes: cara y cruz. ¿Son éstos los valores que pueden tomar la variable "cara de la moneda que queda hacia arriba"?

Sí.

67 Al obtener una muestra de estudiantes de la Universidad de Guadalajara para inferir qué proporción de ellos son varones, se obtiene una tabla con anotaciones "masculino" y "femenino". Estos son los datos básicos de la muestra y, a la vez, son valores de la ______________ "sexo de los estudiantes de la Universidad de Guadalajara".

variable

68 En algunos experimentos sucede que, antes de realizarlos, no podemos decir con certeza cuál será el resultado; es decir, ______ (sí/no) podemos asegurar cuál será el valor que tome la variable asociada al experimento.

no

69 Se dice que una variable es *aleatoria* cuando *no* se puede predecir con *certeza* el

valor que tomará al realizarse un experimento.
¿Es aleatoria la variable "resultado de lanzar un dado"?

Sí.

---

70 El nombre de "aleatoria" proviene del hecho de que la observación está sujeta al azar. ¿Podría usted predecir con certeza cuál será el número de la cara que quedará hacia arriba al lanzar un dado?

No.

---

71 Puesto que no se puede predecir cuál será el número que quedará hacia arriba al lanzar un dado, la variable "número que queda hacia arriba" ________ (es/no es) una variable aleatoria.

es

---

72 Puesto que tampoco puede predecirse con certeza el valor de la variable "tiempo que tarda Horacio Molina en correr los 200 metros planos", ésta es una variable ________

aleatoria

---

73 ¿Es aleatoria la variable "precipitación pluvial diaria en Parral, Chihuahua"?

Sí.

---

74 ¿Es aleatoria la variable "resultado del trasplante de un corazón"?

Sí.

---

75 No podemos asegurar cual será la intensidad del próximo temblor que ocurra en la Ciudad de México; la intensidad del mismo es, entonces, una variable ________

aleatoria

---

76 Si se puede predecir con certeza cuál será el valor de una variable, entonces esa variable ________ (es/no es) aleatoria. Cuando una variable no es aleatoria se llama *determinística*

no es

---

77 Las variables determinísticas son aquellas cuyos valores ______ (sí/no) se pueden predecir con certeza.

sí

---

78 Sea el experimento que consiste en soltar suavemente un cuerpo sobre la superficie de un recipiente con agua destilada y anotar una E, de éxito, si el cuerpo se hunde, y una F, de fracaso, si el cuerpo flota. Si se conoce la densidad del cuerpo que se usará, se puede predecir con certeza el resultado del experimento; por lo tanto, se trata de una variable ________ (determinística/aleatoria).

determinística

---

79 De acuerdo con esto, una variable puede ser determinística o ________

aleatoria

---

80 Si estamos seguros acerca de cuál será el valor que tome una variable al realizar un experimento, entonces ésta es ________ (determinística/aleatoria). Entonces no tiene caso efectuar estudios estadísticos sobre variables determinísticas.

determinística

---

81 A la Estadística interesan únicamente variables ________ (aleatorias/determinísticas).

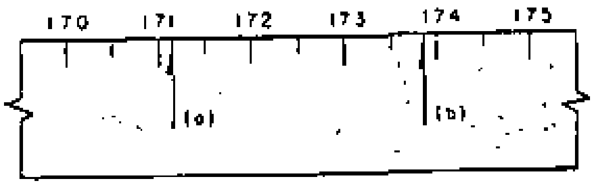
aleatorias

---

82 Como las variables continuas son forzosamente ________ (escalares/nominales), tenemos que utilizar una *escala* para poder obtener el valor de una observación.

escalares

83 Al tratar de medir un valor específico de una variable continua, tenemos que conformarnos con la aproximación que nos proporciona la escala de medición. En otras palabras, tenemos que aproximar la medida al centímetro, al gramo, etcétera. Supongamos que las estaturas de dos estudiantes (a) y (b) son las indicadas por las flechas en la siguiente escala.



Si aproximáramos la estatura al centímetro (cm) más cercano a la flecha, se asignaría una estatura de 171 cm para (a) y de ________ cm para (b).
(cuántos)

174

---

84 Cuando decimos que la estatura de una persona es 174 cm, no necesariamente aseguramos que ésta mide exactamente 174 cm. Al anotar aproximando al centímetro *más* cercano, la lectura puede ser realmente 174.2 ó 173.7 cm; es decir, puede tener una estatura comprendida entre 173.5 y 174.__ cm.

5

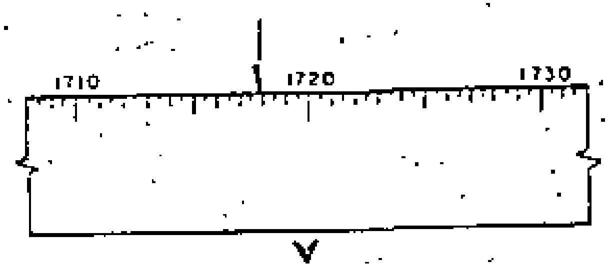
---

85 Observe que al redondear las mediciones al centímetro más cercano, se están dando medidas consistentes en números enteros; en ese caso la variable ______ (continua/discreta) se está tratando como si fuese variable discreta.

continua

---

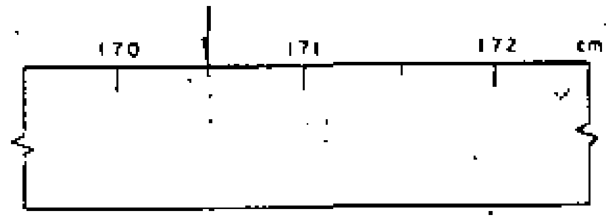
86 Si se quiere aproximar la medición indicada con la flecha en la siguiente escala al milímetro (mm) más cercano, ésta se anotará de ________ mm.
(cuántos)



1718

---

87 Sea la medida indicada por la flecha en la siguiente escala.

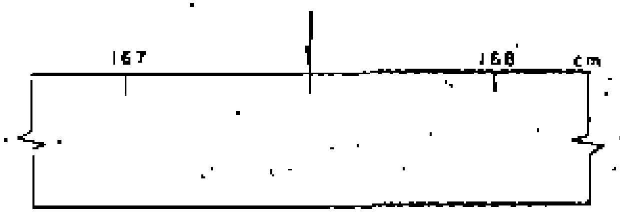


Si se requiere la medida aproximada al centímetro más cercano, será igualmente legítimo considerarla de 170 o de ________ cm, puesto que aparece a la mi del intervalo de estos valores.

171

---

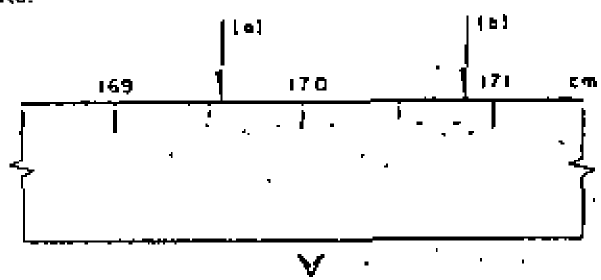
88 Cuando una medida se localiza a la mitad del intervalo de dos valores, es necesario convenir si ésta se considerará igual al número inmediato superior o al inmediato inferior. Conviniendo en esto último, la medida indicada por la flecha en la siguiente escala se anotará de ________ cm.



167

---

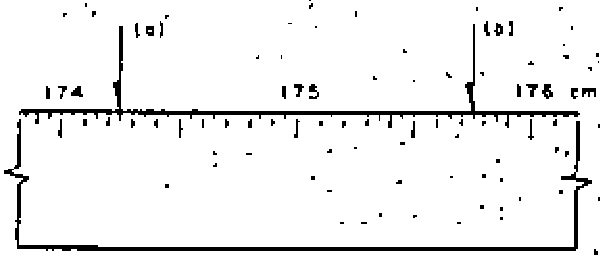
89 Si se quiere aproximar una medida al medio centímetro más cercano las lecturas (a) y (b) de la siguiente escala serán ________ cm, y ________ cm respectivamente.



(a) 169.5, (b) 171.0

---

90 Si se conviene en aproximar una medida al medio centímetro más cercano, y, en caso de que una sea equidistante de dos valores, ésta se considera igual al número inferior, las medidas (a) y (b) de la siguiente escala se anotarán de ______ y ______ cm, respectivamente.



(a) 174.0, (b) 175.5

---

91 Al aproximar las medidas al centímetro más cercano, todas las medidas comprendidas entre los *límites* 167.5 cm y 168.5 cm se anotarán de ______ cm.

168.0

---

92 Al aproximar las medidas al medio centímetro más cercano, todas las medidas comprendidas entre los *límites* :177.25 cm y 177.75 cm se anotarán de ______ cm.

177.5

---

93 ¿Cuáles son los límites para una medida considerada como de 183.0 cm, cuando se sabe que fue aproximada al centímetro más cercano?

182.5 cm, 183.5 cm

---

94 ¿Cuáles son los límites para una medida considerada como de 180.5 cm, cuando se sabe que fue aproximada al medio centímetro más cercano?

180.25 cm, 180.75 cm

---

95 Los límites de las medidas del tipo dado en los cuadros anteriores, se conocen como *límites reales*; el menor se denomina *límite real inferior* y el mayor *límite real superior*. Si los límites reales de una medida son 180.25 cm y 180.75 cm, el límite real inferior es __________ y el límite real superior es __________ cm.

180.25 cm, 180.75

---

96 Los límites __________ de una medida de 170.0 cm que fue aproximada al centímetro más cercano son 169.5 cm y __________ cm.

reales, 170.5

---

97 Si los límites reales de una medida son 169.5 cm y 170.5 cm, el menor de ellos se llama ________ ________ ________, y el mayor ________ ________ ________.

límite real inferior, límite real superior

---

98 Si 71.5 kg es el límite real inferior de una medida aproximada al kilogramo (kg) más cercano, el límite real superior es ______ kg.

72.5

---

99 ¿Cuáles son los límites reales de una medida que se anota como de 18.3 toneladas (ton) cuando se sabe que fue aproximada al décimo de tonelada más cercano? __________, __________.

18.25, 18.35 ton

REVISION

100 Cuando una característica de personas o cosas puede adquirir diferentes valores se denomina __________.

variable

---

101 Dependiendo del tipo de valores (numéricos o nominales), que pueden adquirir las variables, éstas se clasifican en variables __________ y __________.

escalares, nominales (no escalares) (en cualquier orden).

---

102. Una variable, ya sea escalar o nominal, que puede adquirir

a. sólo un número finito o
b. un número infinito, pero numerable, de valores se denomina variable __________.

discreta

---

103 Una variable que no es discreta se llama variable __________

continua

---

104 Las variables cuyos valores no pueden predecirse con certeza se denominan variables __________.

aleatorias

---

105 Una variable es aleatoria o es __________

determinística

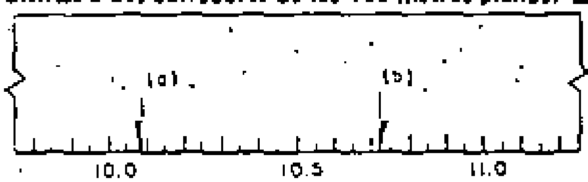
---

106 La Estadística se interesa únicamente en las variables __________ (aleatorias/determinísticas)

aleatorias

---

107 Si las medidas se aproximan al décimo de segundo (seg) más cercano, ¿qué anotaciones haría para los tiempos (a) y (b) indicados en la siguiente escala, correspondientes a dos corredores de los 100 metros planos? ________, ________.



a. 10.1 seg b. 10.7 seg

---

108 Los límites de aproximación de una medida se llaman __________

límites reales

---

109 El menor de los límites reales correspondientes a un intervalo de medidas se llama __________ __________ __________, y el mayor __________ __________ __________.

límite real inferior, límite real superior.

---

110 El límite real inferior de una medida anotada como de 137.5 cm/seg cuando fue aproximada al décimo de cm/seg más cercano vale __________.
El límite real superior vale __________.

137.45 cm/seg, 137.55 cm/seg

EXAMEN

1. *Cuando una característica de personas o cosas puede asumir diferentes valores, se llama ____________.*

2. *Las variables pueden asumir valores ____________ o ____________.*

3. *Las variables que toman valores numéricos, se llaman __________; las que asumen valores nominales, se llaman __________.*

4. *Relacione los nombres de la columna de la izquierda con las descripciones que aparecen a la derecha (4 puntos):*

1. Variable escalar	*a. Calificación obtenida en el examen de Estadística*
	b. Color de los ojos
2. Variable nominal	*c. Estatura de los estudiantes de Bachillerato en cm*
	d. Clasificación de un producto en aceptable o defectuoso

5. *Las variables que pueden asumir un número finito o infinito pero numerable de valores se llaman ____________.*

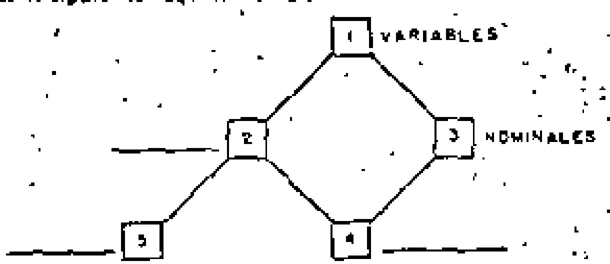
6. *Las variables que toman un número infinito no numerable de valores, se llaman __________.*

7. *Relacione los nombres de la columna de la izquierda con las variables que aparecen a la derecha (4 puntos):*

1. Variable continua	*a. Cualquier variable nominal*
	b. Resistencia, en kg, de los cables de acero producidos en una fábrica
2. Variable discreta	*c. Resultados al extraer las 52 cartas, de una en una, de un juego de naipes*
	d. Tiempo en que un atleta recorre 100 metros

8. *Una variable para la cual se puede predecir con certeza el valor que va a tomar se llama determinística; en caso contrario se llama __________.*

9. *Complete el siguiente esquema de clasificación:*



10. *Las variables continuas son, forzosamente, ________________.*
(nominales/escalares)

11. *Al medir el valor específico que asume una variable continua necesitamos cierta aproximación en la medición. Si se desea aproximar al décimo de kilo más cercano, los pesos de tres niños recién nacidos, ¿qué cantidades anotaría usted, si la báscula registró lo siguiente?*

12. *¿Cuáles son los límites reales de una medida considerada de 3.4 kg, cuando se sabe que se aproximó al décimo de kg más cercano?*

13. *A los límites de una medida se les llama límites ________. El menor de ellos, es el límite ________ __________ y el mayor, el límite ________ __________.*

14. *El límite real inferior de una anotación de 89 km/h, cuando se sabe que se aproximó al km/h más cercano, es ________; el superior es ______.*

TOTAL: 30 puntos.

RESPUESTAS

1. *Variable*

2. *numéricos*
   nominales

3. *escalares*
   nominales

4. *1-a, c*
   2-b, d

5. *discretas*

6. *continuas*

7. *1-b, d*
   2-a, c

8. *aleatoria*

9. $\boxed{2}$ *Escalares*

   $\boxed{4}$ *Discretas*

   $\boxed{5}$ *Continuas*

10. *escalares*

11. *a. 3.1 kg*
    b. 3.4 o 3.5 kg
    c. 4.0 kg

12. *3.35 y 3.45 kg*

13. *reales*
    real inferior
    real superior

14. *88.5*
    89.5

UNIDAD III AGRUPAMIENTO DE DATOS

PREFACIO

En esta tercera unidad se define lo que es *evento, frecuencia* y *frecuencia relativa*. Se dan, además, algunas reglas para organizar y procesar los datos básicos, y para presentar los resultados en forma de tablas.

PARTE A. FRECUENCIA

1 Al lanzar una vez un dado y anotar la cara que queda hacia arriba, se puede obtener cualquiera de los números 1,2,3,4,5, ó 6; si es el 3, se dice que ha ocurrido el *evento* "número que queda hacia arriba = 3". Si el dado se lanza y cae el 5, se dice que ha ocurrido el evento "número que queda hacia arriba = ____".
(3/5)

5

---

2 Dividamos en tres grupos (eventos) los diferentes elementos de la población asociada al experimento del lanzamiento de un dado:

grupo A: 1,2
grupo B: 3,4
grupo C: 5,6

Si lanzamos una vez el dado y cae el número 2, es decir, un número del grupo A, se dice que *ha ocurrido* el evento A. Si en el siguiente lanzamiento cae el número 6, se dice que ha ocurrido el evento ____.
(A/B/C)

C

---

3 Si ahora tenemos a los diferentes elementos de una población agrupados con sigue:

grupo A: 1,2,3,4
grupo B: 5,6

y en un lanzamiento se obtiene el número 6, se dice que ________ el evento B, por pertenecer este número al grupo B. Para que ocurra el ____________ A se necesita obtener cualesquiera de los números 1, ____, 3 ó 4.

ocurrió (ha ocurrido), evento, 2

---

4 Al reunir los diferentes elementos de una población en grupos de *uno o más*, cada uno de esos grupos constituye lo que se conoce como un *evento*.
Al lanzar simultáneamente dos dados y observar la suma de los números que quedan hacia arriba, se tiene que los diferentes elementos de la población son los números enteros del 2 al 12. Si reunimos estos elementos en los grupos:

A: 7,11
B: 2,3,4,5,6,8,9,10,12

y en un lanzamiento obtenemos que la suma es 7, se dice que ha ocurrido el evento _____ (A/B). Si la suma es 12 se dice que ha ocurrido el ___________ ____.

A, evento B

---

5 Para que ocurra un evento, al realizar un experimento, basta con que la observación sea uno cualesquiera de los elementos que lo forman. Si los diversos elementos de una población son los números enteros del 2 al 12 y éstos se dividen en los grupos:

A: 2,3,4
B: 5,6,7,8
C: 9,10,11,12

para que ocurra el evento B es necesario que la observación sea *cualesquiera* de los números ___, ___, ___ u ___.

5, 6, 7, 8

---

6 Si se tienen bolas rojas, verdes, azules, blancas y negras en una urna, y dividimos esta población de colores en los grupos (eventos):

A: negra, blanca
B: roja, verde
C: azul

al extraer al azar una bola azul se dice que ocurrió el evento______. Para que ocurra el evento A se necesita extraer una bola __________ o una __________.

C, negra, blanca

---

7 Si una unidad de un producto elaborado en una fábrica no satisface las normas de calidad especificadas, se anota una D de defectuoso; en caso contrario anota una A de aceptado. La población asociada a este experimento es una secuencia de las letras "D" y letras "A". Consideremos ahora el evento definido así: "el producto es defectuoso". Este evento ocurrirá cada vez que al probar una unidad producida, ésta resulte _______________ (defectuosa/aceptable), es decir, cada vez que anota una D en la lista de datos básicos.

defectuosa

---

8 Si la población de estaturas de los estudiantes de la Universidad de Chihuahua se divide en los eventos:

A: 160 cm o menos
B: de 161 a 180 cm
C: 181 cm o más

y se anota la estatura de un estudiante seleccionado al azar, aproximando la medida al centímetro más cercano, para que ocurra el ______ C se necesita que el estudiante mida 181 cm o más; para que ocurra el evento A se necesita que mida 160 cm o __________ (más/menos); para que ocurra el evento B se necesita que mida de ___ a ___ cm.

evento, menos, 161, 180

---

9 Sea el evento "Jesús Rivas corre los 100 metros planos en menos de 11 segundos". Para que tal evento ocurra se necesita que este atleta recorra esa distancia en menos de ________ (cuántos) seg.

11

---

10 Los eventos suelen denotarse con *letras mayúsculas*. Por ejemplo, en el experimento del lanzamiento de una moneda, sea A el evento "cae cara" y sea B el evento "cae cruz"; si cae cara se dice que ocurrió el evento A; si cae cruz se dice que ocurrió el ________ ___.

evento B

---

11 En el experimento del lanzamiento de un dado, sean los eventos:

C: número par
D: número impar

si se observa el 2 se dice que ocurre el evento C; si cae 4, ________ el evento ____; si cae 5 ocurre el ________ ___.
(C/D)

ocurre, C, evento D

---

12 Sean los eventos asociados el número de estudiantes reprobados anualmente en una universidad:

A: 50 o menos
B: de 51 a 70
C: 71 o más

Si se selecciona al azar el expediente escolar correspondiente a un año, para que ocurra el evento B se necesita que haya habido de ______ a ______ reprobados. Para que ocurra el evento A se requiere que haya habido 50 o ________. Para
(más/menos)
que ocurra el evento C se necesita que haya habido ______ o más reprobados.
(cuántos)

51, 70, menos, 71

---

13 Denotemos con n(A) (léase n de A) el número de veces que ocurre el evento A. Análogamente, el número de veces que ocurre el evento C se denota con ___ (C).

n

---

14 El número de veces que ocurre el evento D se denota con n(__).

D

---

15 El número de veces que ocurre el evento B se denota con ___ (___).

n(B)

---

16 Sea A el evento "cae cruz"; si al repetir 100 veces el experimento de lanzar una moneda, se observan 41 cruces, entonces el número de veces que ocurrió el evento A es n(A) = ______.
(100/41)

41

---

17 Sea B el evento "cae el número 6". Al lanzar un dado 60 veces y observar 12 veces el 6, el número de veces que ocurre el evento B es n(B) = ______.

12

---

18 Denotemos con N el número total de veces que se repite un experimento. Si lanzamos 100 veces un dado, entonces N = 100. Si lanzamos 15 veces un dado, entonces N = ______.

15

---

19 Si anotamos las estaturas de 118 estudiantes, entonces N = ___.

118

---

20 Si el experimento de lanzar un dado se repite 366 veces, y si el evento A "cae el número seis" ocurre 58 veces, entonces n(A) = ______ y N = ___.

58, 366

---

21 El número de veces que ocurre un evento se llama *frecuencia* del evento. Si A es un evento y n(A) = 27, entonces la frecuencia de A es 27.
Si B es otro y n(B) = 36, entonces la frecuencia de B es ____.

36

---

22 En vista de la definición dada en el cuadro anterior, n(A) denota la frecuencia del evento A y n(B) denota la ________ del evento B.

frecuencia

---

23 La frecuencia del evento C es el número de veces que ocurre el evento ____.
(cuál)

C

---

24 La frecuencia del evento B es el ____________ de veces que ocurre el evento B.

número

---

25 El número de veces que ocurre el evento E se denomina ________ del evento E.

frecuencia

---

26 En un laboratorio se aplicó una medicina para curar la hepatitis a 100 personas enfermas. Se anotó una E por cada persona que sanó en un término de 30 días y una F por cada una que no sanó en ese lapso. El número de veces que ocurrieron los eventos A: "éxito" y B: "fracaso", fueron n(A) = 81 y n(B) = 19. La frecuencia de A es, entonces, 81 y la frecuencia de B es ________.

19

---

27 En un día nacieron en un hospital local 18 niños y 10 niñas. Si A es el evento "sexo masculino" y B el evento "sexo femenino", entonces n(A) = __________, n(B) = ________ y, por lo tanto, la frecuencia de A es ________ y la de B es ________.

18, 10, 18, 10

---

28 Use la tabla 3,1

La tabla 3.1 presenta los datos básicos correspondientes a las calificaciones en un examen de Pedagogía, y a las estaturas de 30 profesores de la Facultad de Derecho. Si A es el evento "calificación de 51 a 60", y contamos el número de calificaciones de la tabla 3.1 que quedan en ese intervalo de valores, encontramos que

n(A) = 2 (vea en la tabla que las dos calificaciones señaladas con una flecha son las únicas que caen dentro de este intervalo). Esto indica que la ______________ del evento A es 2.
(frecuencia/población)

TABLA 3.1

[illegible]	[illegible]	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible] ←	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible] ←	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

frecuencia

---

29 Use la tabla 3.1.

Si D es el evento "calificación de 81 a 90" y contamos el número de calificaciones de 81 a 90, encontramos que n(D) = 11 (verifique que los once elementos indicados con una cruz en la tabla 3.1 corresponden realmente a este evento). Entonces, la frecuencia del evento D es ______.

11

---

30 Use la tabla 3,1.

¿Cuál es la frecuencia del evento B "calificación de 61 a 70"?

6.

---

31 Use la tabla 3.1.

¿Cuál es el valor de n(C) si C es el evento "calificación de 71 a 80"?

32 Use la tabla 3.2.

Para encontrar fácilmente la frecuencia de un _______, se pueden *organizar* los datos como se ha hecho en la tabla 3.2. Vea los encabezados de las columnas: en la primera se indican los eventos y en la segunda se anotan los elementos que ocurrieron de cada evento.

Observe la tabla 3.1. ¿Qué calificaciones hacen que ocurra el evento A: 51-60? _______.

TABLA 3.2

Evento (intervalo de calificaciones)	Elementos correspondientes a los intervalos
A: 51-60	59, 57
B: 61-70	67, 65, 69, 67, 67
C: 71-80	72, 73, 73, 77, 78, 78
D: 81-90	83, 88, 84, 89, 83, 84, 83, 87, 81, 83, 81
E: 91-100	99, 91, 97, 95, 91, 93

evento, 59, 57

33 Observe que los números de la respuesta correcta al cuadro anterior se han anotado en el renglón del evento A de la tabla 3.2.

Observe la tabla 3.1. ¿Qué calificaciones corresponden al evento E: 91-100? _______

_______.

TABLA 3.1 REPETIDA

Número del alumno	Calificación en el examen de Estadística	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible]	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible]	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

99, 91, 97, 95, 91, 93.

34 Verifique que los valores de la respuesta correcta al cuadro anterior estén anotados en el renglón del evento E de la tabla 3.2.

Observe la tabla 3.1. ¿Qué calificaciones hacen que ocurra el evento B: 61-70?

_______.

67, 65, 69, 67, 67.

35 Verifique que los valores de la respuesta correcta al cuadro anterior se encuentren anotados en el renglón del evento B de la tabla 3.2. La secuencia indicada en los cuadros anteriores se prosigue hasta agotar todos los eventos.

Observe la tabla 3.2. ¿Cuántos elementos del evento A ocurrieron? _______.

2.

36 La tabla 3.2 se puede elaborar con facilidad si previamente se *organizan* los datos en la forma presentada en la tabla 3.3 en la cual aparecen los datos *ordenados* en forma creciente (obsérvela). ¿Qué calificaciones de la tabla 3.3 hacen que ocurra el evento A: 51-60? _______.

TABLA 3.3

Calificación en Estadística	Número del alumno
57	12
59	1
55	11
67	8
67	22
67	24
69	17
72	6
73	16
73	18
77	3
78	21
78	28
81	25
81	30
83	1
83	15
83	27
84	9
84	17
87	23
88	4
89	12
89	21
91	5
91	10
93	24
95	14
97	13
99	2

57, 59.

37 Observe en la tabla 3.2 que los valores de la respuesta al cuadro anterior están anotados en el lugar correspondiente al evento A.
Observe la tabla 3.3. ¿Qué calificaciones hacen que ocurra el evento E: 91-100?

TABLA 3.2 REPETIDA

Evento (intervalo de calificaciones)	Elementos correspondientes a los intervalos
A: 51-60	59, 57
B: 61-70	67, 65, 69, 67, 69
C: 71-80	72, 73, 73, 77, 78, 78
D: 81-90	83, 82, 84, 89, 83, 84, 89, 87, 84, 83, 81
E: 91-100	93, 91, 97, 99, 91, 95

91, 91, 93, 95, 97, 99.

---

38 Observe en la tabla 3.2 que los valores de la respuesta al cuadro anterior están anotados en su lugar correspondiente, aunque en diferente orden.
Una tabla en la que se presentan los datos en forma ordenada se denomina *tabla de datos ordenados*. Por ejemplo, la tabla 3.3 es una tabla de ______ ______.

datos ordenados

---

39 La tabla de datos ordenados es útil porque en ella se pueden localizar fácilmente los datos que corresponden a un evento.
Observe la tabla 3.3. ¿Cuáles datos corresponden al evento "calificación de 55 a 65 inclusive"?

TABLA 3.3 REPETIDA

Calificación en Pedagogía	Número del profesor
57	13
59	7
65	11
67	9
67	22
69	29
69	17
72	5
73	16
73	18
77	3
78	21
78	23
81	26
81	30
83	1
83	15
83	27
84	8
84	19
87	25
89	4
89	12
89	14
91	9
91	17
93	24
95	12
97	13
99	2

57, 59, 65.

---

40 Una tabla que presenta los datos en forma ordenada se denomina tabla de ________ ordenados. Al ordenar éstos a partir de la tabla de datos básicos, es conveniente poner una señal a un lado de cada número que se va ordenando para estar seguro de que no faltó ninguno de ser incluido.
Observe que en la tabla 3.1 se usó una marca a la izquierda de cada número al ser incluido en la tabla 3.3.

TABLA 3.1 REPETIDA

Número del Profesor	Calificación en el examen de Pedagogía	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible]	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible]	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

datos

---

41 Recuerde que el número de ocurrencias de un evento se denomina **frecuencia**. En la tabla 3.2 es fácil encontrar las frecuencias simplemente *contando* los elementos correspondientes a cada evento anotado en la segunda columna.
Observe la tabla 3.2. ¿Cuál es la frecuencia del evento A?

2.

---

42 Observe la tabla 3.2.
¿Cuál es la frecuencia del evento B?

5.

43 Observe la tabla 3.2.
¿Cuál es la frecuencia del evento E?

TABLA 3.2 [illegible]

Evento (intervalo de calificaciones)	Elementos correspondientes a los intervalos
A: 51-60	[illegible]
B: 61-70	[illegible]
C: 71-80	[illegible]
D: 81-90	[illegible]
E: 91-100	[illegible]

6.

44 Una tabla en la que se anotan los datos básicos en grupos se conoce como *tabla de datos agrupados*.
Observe la tabla 3.2. ¿Es ésta una tabla de datos agrupados?

SI.

45 En una tabla de datos agrupados se ______________ los datos básicos.
(separan/agrupan)

agrupan

46 Una tabla que presenta los datos básicos agrupados, se conoce como tabla de ______ ______.

datos agrupados

47 Los datos básicos pueden organizarse en dos formas distintas; en una tabla de ______ ordenados y en una de datos ____________.

datos, agrupados

48 Observe que en la tabla 3.4 se encuentran anotadas las frecuencias de cada evento.
¿Cuál es la frecuencia del evento D?

TABLA 3.4

Evento (intervalo de calificaciones)	Frecuencia
A: 51-60	2
B: 61-70	5
C: 71-80	6
D: 81-90	11
E: 91-100	6

11.

49 Observe la tabla 3.4.
¿Cuál es la frecuencia de las calificaciones de 91 a 100? ___________.

6.

50 Observe la tabla 3.4.
¿Cuál es la frecuencia de las calificaciones de 61 a 70? ___________.

5.

51 Una tabla en la que se presentan los eventos con su correspondiente frecuencia se llama *tabla de distribución de frecuencias*.
Observe la tabla 3.4; ¿es ésta una tabla de distribución de frecuencias?________.

SI.

52 Una tabla de distribución de frecuencias es una tabla en la que se presentan los eventos con su correspondiente ____________.
(nombre/frecuencia)

frecuencia

53 Una tabla en la que se presentan los eventos con su correspondiente frecuencia, se llama tabla de ___________ de frecuencias.

distribución

54 En la tabla 3.4 se presentan algunos eventos con su correspondiente frecuencia, por lo cual es una tabla de ___________ __ ___________.

distribución de frecuencias

PARTE B. FRECUENCIA RELATIVA

55 Si se nos dice que la frecuencia de un evento es 4, no podremos interpretar esta

información apropiadamente, a menos que se nos diga cuántas veces se realizó el experimento.
Si un experimento se realiza N veces, y n(A) es la frecuencia del evento A, la *frecuencia relativa* del evento A es igual a $n(A)/N$.
Si $n(A) = 4$ y $N = 10$, la frecuencia relativa de A es $n(A)/N$ = ________ /10 = ________.

4, 0.4

---

56 Si el equipo de basket-ball de la Facultad de Ingeniería ha ganado 6 juegos durante la presente temporada, y el de la Facultad de Medicina 8, no podremos opinar que este último equipo sea mejor que el de Ingeniería, a menos que se nos diga cuántos partidos ha jugado cada uno. Si el de Ingeniería ha jugado 8 y el de Medicina 16, la frecuencia relativa de partidos ganados por Ingeniería es 6/8 = 0.75 y la de Medicina es ____/____ = 0.50. Estas cifras harían favorito al equipo de Ingeniería si jugara contra el de Medicina en este momento.

8/16

---

57 La fórmula para calcular la frecuencia relativa h(B), del evento B es

$$(?) = \frac{n(B)}{(?)}$$

Si la frecuencia del evento B es 5 y el número total de experimentos es 10, su frecuencia relativa es:

$$h(B) = \frac{\quad}{\quad} = \_\_\_\_$$

$h(B) = \frac{n(B)}{N}$ $\frac{5}{10} = 0.5$

---

58 Un biólogo aplica una vacuna a 100 ratones. Después de exponerlos a la enfermedad contra la cual fueron vacunados, se anota una E por cada ratón que no contrajo la enfermedad y una F por cada uno que sí la contrajo. Si $n(E) = 80$ y $n(F) = 20$, la frecuencia relativa del evento F es $h(F) = 20/100 = 0.2$ y la del evento E es ________.

$h(E) = 0.8$; $(80/100 = 0.8)$

---

59 Si la frecuencia relativa del evento E es 0.8 y el número de veces que se realizó el experimento es 100, ¿cuántas veces ocurrió el evento E? ________.

80; $(0.8 \times 100 = 80)$

---

60 Un fabricante de radios prueba los radios recién producidos; si funcionan se anota una E; si no una F. En una hora de pruebas se obtuvieron los siguientes datos: E,E,E,E,F,E,E,E,E,E. ¿Cuánto vale

a. N ________ d. h(E) ________
b. n(E) ________ e. h(F)? ________
c. n(F) ________

a. 10, b. 9, c. 1, d. 0.9; (9/10), e. 0.1; (1/10).

---

61 Note que el número de ocurrencias de un evento no puede ser mayor que el número de veces que se realiza el experimento correspondiente. Esto implica que la frecuencia relativa es *siempre menor o igual que 1*.
Si en un experimento, E denota éxito y F fracaso, y si se obtienen las observaciones: E,E,E,E,E,E,E,E,E,E, entonces $n(E)$ = _____ y N = _____, por lo tanto

$$h(E) = \_\_\_\_ = \_\_\_\_$$

10, 10, $\frac{10}{10} = 1$

---

62 Note también que la frecuencia relativa *no* puede ser *menor* que cero, puesto que el mínimo valor posible de ocurrencias de un evento es cero. Si E indica éxito y F fracaso, y si las observaciones de un experimento son: E,E,E,E,E,E,E,E,E,E, entonces N = _____, y $n(F)$ = _____, por lo tanto;

$$h(F) = \_\_\_\_ = \_\_\_\_$$

10, 0, $\frac{0}{10} = 0$

---

63 De los cuadros anteriores se concluye que la frecuencia relativa puede adquirir valores de cero a uno. Entonces, el mínimo valor de la frecuencia relativa es cero y el máximo valor es ________ (cuánto)

uno

64 Use la tabla 3.4.

Las calificaciones de 51 a 60 (evento A) aparecen con una frecuencia de ________. Puesto que el número total de observaciones es 30, la frecuencia relativa correspondiente, la cual se indica con el símbolo ________, es

$$h(A) = \frac{\_\_\_}{\_\_\_} =$$

TABLA 3.4 EVENTOS

Evento (intervalo de calificaciones)	Frecuencia
A: 51–60	2
B: 61–70	5
C: 71–80	6
D: 81–90	11
E: 91–100	6

2. h(A), $\frac{2}{30} = 0.067$

---

65 Observe la tabla 3.4.

Las calificaciones de 81 a 90 (evento D) aparecen con una frecuencia de 11. Puesto que el número total de observaciones es 30, la frecuencia relativa correspondiente es

$$\underset{\text{(símbolo)}}{\_\_\_\_\_\_} = \_\_\_\_\_\_ = 0.367$$

$h(D) = \frac{11}{30}$

---

66 Use la tabla 3.4.

¿Cuánto vale la frecuencia relativa de los eventos B, C y E?_____, _________

h(B) = 0.166; (5/30 = 0.166).
h(C) = 0.200; (6/30 = 0.200).
h(E) = 0.200; (6/30 = 0.200).

---

67 Una tabla que presenta los eventos con su correspondiente frecuencia relativa se denomina *tabla de distribución de frecuencias relativas.*

Observe la tabla 3.5. ¿Es ésta una tabla de distribución de frecuencias relativas?_____.

TABLA 3.5

Evento (intervalo de calificaciones)	Frecuencia relativa
A: 51–60	0.067
B: 61–70	0.166
C: 71–80	0.200
D: 81–90	0.367
E: 91–100	0.200

Sí.

---

68 En una tabla de distribución de frecuencias relativas se presentan los eventos y su correspondiente ______________________.
(frecuencia/frecuencia relativa)

frecuencia relativa.

---

69 Una tabla que presenta los eventos con su correspondiente frecuencia relativa se conoce como tabla de ________ de frecuencias ________.

distribución, relativas

---

70 Es común expresar las frecuencias relativas en forma decimal o en porcentaje. Si h(J) = 0.3, es equivalente decir que h(J) = 30%. Si h(T) = 0.97, entonces h(T) = ________%.

97

---

71 Observe la tabla 3.5.

¿Cuáles son las frecuencias relativas de los eventos B y D?

h(B) = 0.166, h(D) = 0.367.

---

72 De una urna que contiene 10,000 bolas de colores, rojas (R), blancas (B), negras (N) y azules (A), se obtuvieron las 30 observaciones dadas a continuación

B B R A A N B N A R
R N B N B B B A R N
A R B B N R A B B N

Complete la siguiente tabla:

Evento	Elementos que ocurrieron	Frecuencia	Frecuencia relativa
Sale R	R R R R R R		
Sale B			
Sale N	N N N N N N N		
Sale A			

Evento	Elementos que ocurrieron	Frecuencia	Frecuencia relativa
Sale R	R R R R R R	6	0.200
Sale B	B B B B B B B B B B B	11	0.367
Sale N	N N N N N N N	7	0.233
Sale A	A A A A A A	6	0.200

73 Si ahora se divide la población del cuadro anterior en los eventos:

S: sale blanca o negra
T: sale roja o azul

y en una extracción se observa que la bola es blanca o negra, entonces se dice que ocurrió el evento S. Para que ocurra el evento T se necesita sacar una bola ________ o ________.

roja, azul (en cualquier orden.)

74 Si se emplean los 30 datos básicos del cuadro 72, agrupados en los eventos S y T se obtiene la siguiente tabla. LLene las columnas tercera y cuarta.

Evento	Elementos que ocurrieron	Frecuencia	Frecuencia relativa
S: Sale blanca o negra	B B N B N N B N B N N N N B N N B N		
T: Sale roja o azul	R A A A R R A R A R R A		

Frecuencia	Frecuencia relativa
18	0.600
12	0.400

75 La frecuencia del evento S "sale blanca o negra" se obtiene contando las bolas blancas y las negras. Esto equivale a sumar la frecuencia de las bolas blancas y la ________ de las bolas ________.

frecuencia, negras

76 Si la frecuencia de bolas blancas es 11 y la de negras es 7, la frecuencia del evento S "sale blanca o negra" será la suma de las frecuencias de los eventos "sale blanca" y "sale negra", es decir, $n(S) = 11 + 7 = 18$.
¿Cuánto vale la frecuencia del evento T "sale roja o azul" si se sabe que la frecuencia del evento "sale roja" es 6 y la del evento "sale azul" es 6?

$n(T) = 12$; $(6 + 6 = 12)$

77 Al ordenar los datos básicos, se puede describir más claramente la distribución de sus valores. Por este motivo es conveniente *organizar* los ________ y presentarlos en forma tabular mediante una *tabla de datos ordenados.*

datos

78 Use la tabla 3.2.
Una manera de obtener una idea aún más clara de la distribución de los valores es *agrupando* los datos. En la tabla 3.2, por ejemplo, se agruparon los datos en los *intervalos* de calificaciones: 51 – 60, 61 – 70, ________, ________ y ________.

TABLA 3.2 [illegible]

Evento (Intervalo de calificaciones)	Elementos correspondientes a los intervalos
A: 51–60	[illegible]
B: 61–70	[illegible]
C: 71–80	[illegible]
D: 81–90	[illegible]
E: 91–100	[illegible]

71-80, 81-90, 91-100

79 Los intervalos de agrupamiento de datos se conocen como *intervalos de clase.* Los intervalos de clase son intervalos de ________ de datos.
(separación/agrupamiento)

agrupamiento

80 Los ________ de clase son intervalos de agrupamiento de datos.

intervalos

81 Los intervalos de agrupamiento de datos se denominan ________ ________.

intervalos de clase

82 Use la tabla 3.3. La *primera* etapa en el proceso de agrupamiento de datos consiste en la determinación de los intervalos de clase. Para hacer esto es necesario encontrar primero los valores mínimo y

máximo de los datos, lo cual se puede hacer fácilmente a partir de una tabla de datos ordenados. ¿Cuál es el valor mínimo de los datos presentados en la tabla 3.3?

TABLA 3.3 REPETIDA

Calificación en Probabilidad	Número del estudiante
57	13
59	7
65	11
57	8
67	22
57	29
69	17
72	6
71	16
71	15
77	3
71	21
78	23
81	25
81	30
83	[illegible]
83	15
83	27
84	9
85	19
87	23
88	4
89	12
89	21
91	5
91	25
93	24
95	14
97	10
99	2

57.

83 Observe la tabla 3.3, de datos ordenados; ¿cuál es el valor máximo de los datos?

99.

84 La diferencia (resultado de la resta) de los valores máximo y mínimo se conoce con el nombre de *rango* de los datos. Los valores mínimo y máximo de los datos de la tabla 3.3 son 57 y 99 respectivamente. El rango de los datos es, por lo tanto, 99 − _____ = _____.

57, 42

85 Una tabla de datos ordenados sirve, entre otras cosas, para localizar fácilmente los valores mínimo y ________ de los datos, los cuales son necesarios para calcular el rango.

máximo

86 Puesto que el primer paso en el proceso de agrupamiento consiste en encontrar los valores mínimo y máximo de los datos, y el rango es la diferencia entre ellos, este primer paso consiste básicamente en encontrar ________________ de los datos. (el rango/la frecuencia)

el rango

87 El rango es la diferencia entre los valores máximo y ________ de los datos.

mínimo

88 Si al valor máximo de los datos se le resta el valor mínimo de los mismos, se obtiene el número llamado ________.

rango

89 Si los valores máximos y mínimos de una muestra son 1947 y 1347, el rango de los datos es ________.
(cuánto)

600;(1947-1347 = 600)

90 El primer paso en el proceso de agrupamiento de datos consiste en calcular el rango; el *segundo* consiste en decidir *cuántos* intervalos de clase se usarán. Es usual, dependiendo del número de observaciones, que el número de intervalos varíe entre 5 y 20. Tome la tabla 3.4 y diga cuántos intervalos se usaron para construirla.

TABLA 3.4 REPETIDA

Clase (intervalo de calificaciones)	Frecuencia
A) 51–60	2
B) 61–70	5
C) 71–80	6
D) 81–90	11
E) 91–100	6

5

91 A mayor número de datos se usará mayor número de intervalos de clase. Para aclarar esta idea, fijemos por ejemplo el número de intervalos suponiendo que caigan unos 5 elementos de la muestra en cada uno. De esta forma, si se tuvieran 30 elementos, el número de intervalos sería 30/5 = 6.
Si se tuvieran 75 elementos, el número de intervalos sería ____/5 = 15.

75

92 Si al dividir entre 5 para obtener el número de intervalos de clase que se van a usar, se obtiene un número fraccionario, éste se redondea aproximándolo al entero más cercano. Si el número de datos es 32, la cantidad de intervalos sería aproximadamente de 32/____=6.4. Al redondear este número al entero más cercano, se obtiene el número _____.

5, 6

---

93 Si se tienen 100 datos, un número conveniente de intervalos de clase es ___/___ = 20. Este es el máximo número de intervalos que se recomienda, por lo cual, si el número de datos es *mayor* de 100, es conveniente (no necesario) usar sólo 20 intervalos.

100, 5

---

94 Si aplicamos esta regla a los 30 datos de la segunda columna de la tabla 3.1, el número de intervalos de clase que conviene usar es aproximadamente ___/___ = ___. Observe la tabla 3.2. ¿Cuántos intervalos de clase se usaron?

TABLA 3.2 REPETIDA

Suerte (intervalo de calificaciones)	Elementos correspondientes a los intervalos
A: 51–60	55, 57
B: 61–70	61, 65, 66, 67, 69
C: 71–80	72, 73, 73, 77, 78, 78
D: 81–90	83, 84, 84, 85, 85, 86, 88, 87, 87, 88, 89
E: 91–100	92, 93, 97, 95, 91, 93

TABLA 3.1 REPETIDA

[illegible]	[illegible]	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible]	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible]	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

30/5 = 6, 5.

---

95 En la tabla 3.2 se usaron 5 intervalos de clase en vez de los 6 que indica la regla. Lo que sucede es que la regla *sólo* proporciona un valor *alrededor* del cual conviene que se tome el número de __________ de clase.

intervalos

---

96 Para decidir el número de intervalos de clase que conviene usar en cada caso, cada individuo deberá basarse en su experiencia en el manejo de datos. Mientras se adquiere esta experiencia se recomienda usar la regla:
Número de intervalos = número de datos / 5 (aproximando el resultado al entero más cercano).
¿Cuántos intervalos usaría según esta regla, si tuviera 83 datos?

17; (83/5 = 16.6, valor que aproximado al entero más cercano da 17).

---

97 El primer paso en el proceso de agrupamiento de datos consiste en calcular ______________
(el rango/la frecuencia)

el rango

---

98 El segundo paso en el proceso de agrupamiento de datos consiste en fijar ______________ de intervalos de clase.
(el número/la frecuencia)

el número

---

99 El *tercer* paso en el proceso de agrupamiento de datos consiste en fijar los números que limitan cada intervalo de clase; estos números se denominan *límites de clase*.
Observe la tabla 3.2, primera columna. Los límites de clase del primer intervalo son 51 y 60. ¿Cuáles son los límites de clase del segundo intervalo?

61 y 70

---

100 Observe la tabla 3.2, ¿cuáles son los límites de clase del último intervalo?

91 y 100

101 Los números que limitan los intervalos de clase se denominan ________ ________ de clase. (límites/frecuencias)

límites

---

102 Los ________ ___ ______ son los números que limitan los intervalos de clase.

límites de clase

---

103 El menor de los límites de clase de un intervalo se llama *límite inferior de clase*. El mayor se llama *límite superior de clase*.
Observe la tabla 3.2, primera columna. El límite inferior de clase del primer intervalo es el número 51; ¿cuál es el límite superior de clase de ese mismo intervalo?

TABLA 3.2 [illegible]

[illegible]	[illegible]
A: 51–60	[illegible]
B: 61–70	[illegible]
C: 71–80	[illegible]
D: 81–90	[illegible]
E: 91–100	[illegible]

60.

---

104 Observe la tabla 3.2.
¿Cuál es el límite inferior de clase del segundo intervalo?
¿Cuál es el límite superior?

61. 70.

---

105 El menor de los límites de clase de un intervalo se llama límite ______ de clase. El mayor se llama ______ ______ ______ ______.

inferior, límite superior de clase

---

106 El límite superior de clase de un intervalo es el mayor de los ________ de clase de ese intervalo.

límites

---

107 Si las observaciones se aproximan al entero más cercano, los *límites reales* de una anotación de 51 son 50.5 y 51.5. Los de una anotación de 60 son _____ y _____.

59.5, 60.5

---

108 Puesto que los límites reales de 51 son 50.5 y 51.5, y los de 60 son 59.5 y 60.5, cuando las observaciones se aproximan al entero más cercano, los límites reales del intervalo 51-60 son 50.5 y 60.5. Análogamente, los límites reales del intervalo 61–70 son ______ y ______.

60.5, 70.5

---

109 Si los datos se aproximan al entero más cercano, los ______ reales del intervalo 71-80 son ________ y ________.

límites, 70.5, 80.5

---

110 Si las observaciones se aproximan al medio centímetro más cercano, los límites reales del intervalo 51–60 son 50.75 y ______.

60.25

---

111 El menor de los límites reales de un intervalo se llama *límite real inferior*.
Si se aproximan los datos al número entero más cercano, el límite real inferior del intervalo 51-60, es 50.5. El del intervalo 61-70 es ________.

60.5

---

112 El mayor de los límites reales de un intervalo se llama *límite real superior*.
Si se aproximan los datos al número entero más cercano, el límite real superior del intervalo 51-60 es 60.5. El del intervalo 61-70 es ______.

70.5

---

113 Complete la siguiente tabla (considere que los datos se aproximaron al número entero más cercano).

Intervalo	Límite real inferior	Límite real superior
51-60	50.5	60.5
61-70	60.5	70.5
71-80		
81-90		
91-100		

70.5	80.5
80.5	90.5
90.5	100.5

114 La diferencia entre los límites reales superior e inferior de un intervalo se denomina *amplitud o ancho del intervalo*. Así, la amplitud del intervalo cuyos límites reales son 50.5 y 60.5 es 60.5−50.5 = 10.0
Si los límites reales fuesen 50.75 y 60.25, la amplitud sería ______ − ______ = ______.

60.25 − 50.75 = 9.50

115 Si los límites __________ de un intervalo son 120.75 y 135.25, la amplitud (ancho) del intervalo es ______ − ______ = ______.

reales, 135.25 − 120.75 = 14.50

116 La diferencia entre los límites reales superior e inferior de un intervalo se llama ______________ del intervalo.

amplitud (ancho)

117 La amplitud del intervalo cuyos límites reales son 23.5 y 25.5 es ______.

2.0;(25.5 - 23.5 = 2.0)

118 Observe la tabla 3.2.
¿Tienen todos los intervalos de clase igual amplitud?

TABLA 3.2 [illegible]

Clase	(Intervalo de calificaciones)	Elementos correspondientes a los intervalos
A:	51-60	58, 57
B:	61-70	[illegible]
C:	71-80	[illegible]
D:	81-90	[illegible]
E:	91-100	[illegible]

Sí.

119 Habíamos dicho que el *tercer* paso en el proceso de agrupamiento de datos consiste en fijar todos los límites de clase. Para esto es necesario determinar primero la ______________ o ancho de los ______________ de clase.

amplitud, intervalos

120 Si se desea que todos los intervalos de clase tengan *igual* amplitud, ésta se calcula dividiendo el rango de los datos entre el número deseado de intervalos, aproximando el resultado al número entero inmediato superior. Por ejemplo, si el rango es 42, y se quieren 5 intervalos de clase, la amplitud de los mismos será ___/___ = 8.4, valor que aproximado al entero superior más cercano da 9.

42/5 = 8.4

121 La amplitud (ancho) de los intervalos de clase, se calcula dividiendo el rango de los datos entre el número deseado de intervalos. Si el rango es 49 y el número de intervalos es 7, la amplitud es ______.

7;(49/7 = 7)

122 El ancho de los intervalos de clase se calcula dividiendo el ________ de los datos entre el número deseado de intervalos.

rango

123 El cociente del __________ de los datos entre el número deseado de intervalos, nos proporciona la _________ de los intervalos de clase.

rango, amplitud

124 En ocasiones es conveniente utilizar una amplitud de los intervalos de clase diferente a la calculada. En la tabla 3.2, por ejemplo, se usó una amplitud de 10 unidades en vez de 9 (valor calculado), ya que por tratarse de calificaciones, es útil tener los datos agrupados por decenas. Observe la tabla 3.2.
¿Está el valor mínimo de los datos (57), incluido en el primer intervalo de clase?

Sí.

125 Observe la tabla 3.2.

¿Está el valor máximo de los datos (99), incluido en el último intervalo de clase?

TABLA 3.2 [illegible]

Clase (intervalo de calificaciones)	[illegible]
A [illegible]	[illegible]
B [illegible]	[illegible]
C [illegible]	[illegible]
D [illegible]	[illegible]
E [illegible]	[illegible]

Sí.

---

126 Al fijar los límites de clase, es necesario tomar en cuenta que el valor mínimo de los datos debe quedar incluido en el primer intervalo de clase y el valor máximo en el último. Observe la tabla 3.2.

¿Cumplen los límites de clase con esta condición? (Recuerde que el valor mínimo es 57 y el máximo 99.)

Sí.

---

127 Para que el valor __________ (mínimo/máximo) de los datos quede incluido en el primer intervalo de clase, el primer límite inferior de clase deberá escogerse en tal forma que sea *igual o menor* que él.

mínimo

---

128 De la misma manera, para que el valor __________ (mínimo/máximo) de los datos quede incluido en el último intervalo de clase, el último límite superior de clase deberá ser *igual o mayor* que él.

máximo

---

129 ¿Debe ser el primer límite inferior de clase *menor o igual* que el dato de mínimo valor?

Sí.

---

130 El último límite superior de clase debe ser __________ (menor/mayor) o igual que el dato de máximo valor.

mayor

---

131 En ocasiones es necesario considerar algunas condiciones particulares de los datos antes de fijar los límites de clase. Si en el caso de las calificaciones de la tabla 3.1 se sabe que la máxima calificación posible es 100, entonces el último límite superior de clase, convendría que no fuera mayor que 100. Por ejemplo, si en este caso los intervalos se escogieran en tal forma que el *último* intervalo fuera 94-103, el mayor límite de clase sería _____, el cual es mayor que la calificación máxima posible (100), por lo cual convendría cambiar los intervalos.

TABLA 3.1 [illegible]

[illegible]	[illegible]	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible]	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible]	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

103

---

132 Supongamos que ahora queremos agrupar los datos de calificaciones de la tabla 3.1 en tal forma que el ancho de los intervalos de clase sea 9 (valor calculado anteriormente para cinco intervalos).

Si tomamos como primer límite inferior de clase el mínimo valor de los datos (57), el primer intervalo será 57-65 y el segundo será 66-74; ¿cuáles serán el tercero y el cuarto?

75-83, 84-92.

---

133 El quinto y último intervalo para el ejemplo del cuadro anterior, es 93-101; ¿cuál es el límite superior de clase en este intervalo?

101

134 Este número (101) es mayor que el máximo valor que es posible obtener (100), por lo tanto, conviene cambiar los límites de clase.
¿Cuáles serán los 5 intervalos de clase, si el primer límite inferior de clase es 53 y la amplitud del intervalo es 9? ________, ________, ________, ________, ________.

53-61, 62-70, 71-79, 80-88, 89-97.

---

135 ¿Está el dato de máximo valor (99) comprendido en el último intervalo de clase?
________.

No.

---

136 ¿Es válido que el dato de máximo valor no quede comprendido en el último intervalo de clase? ____________.

No.

---

137 Puesto que no es válido que el dato de máximo valor no quede comprendido en el último intervalo de clase, ________ necesario cambiar los límites de clase.
(es/no es)

es

---

138 Si ahora comenzamos con el número 55 como primer límite de clase, los intervalos serán: 55-63, 64-72, ________, ________, ________.

73-81, 82-90, 91-99

---

139 ¿Incluye el primer intervalo de clase (55-63) al dato de mínimo valor (57)? _____.

Sí.

---

140 ¿Incluye el último intervalo de clase (91-99) el dato de máximo valor (99)? _____.

Sí.

---

141 ¿Es el último límite de clase, mayor que el máximo valor posible de las calificaciones (100)? ____________.

No.

---

142 Puesto que ahora sí se cumplen las condiciones requeridas, estos límites de clase son aceptables.
Recuerde que la frecuencia de un evento es el número de veces que éste ocurre. Observe la tabla 3.3. ¿Cuál es la frecuencia del evento "calificaciones de 55 a 63"?

TABLA 3.3 [illegible]

Calificación en Porcentaje [illegible]	Número del [illegible]
57	11
[illegible]	7
61	11
67	8
67	32
67	24
69	17
72	5
73	15
73	15
77	3
78	21
79	24
81	25
81	30
83	1
83	13
83	27
84	9
85	19
87	25
81	4
89	15
89	23
91	5
91	10
93	24
95	14
97	10
59	2

2.

---

143 Resumiendo, la primera etapa en el agrupamiento de datos consiste en determinar los ________ de clase. Consideraremos que son cuatro los pasos de esta etapa, a saber:

1. Calcular el rango de los datos.
2. Determinar el número de intervalos.
3. Calcular la amplitud de los intervalos.
4. Fijar los límites de clase.

intervalos

---

144 Indique los cuatro pasos de la primera etapa en el agrupamiento de datos, cuyo objeto es determinar los intervalos de clase.

1.
2.
3.
4.

1. Calcular el rango de los datos.
2. Determinar el número de intervalos.
3. Calcular la amplitud de los intervalos.
4. Fijar los límites de clase.

145 La *segunda* etapa en el agrupamiento de datos consiste en encontrar las frecuencias de cada intervalo. Si un evento ocurre dos veces, su frecuencia es igual a _____

Use la tabla 3.3 para completar la siguiente.

Intervalo de clase	Elementos que pertenecen	Frecuencia
55-63	57,59	2
64-72	65,67,67,67,69,72	(a)
73-81	73,73,77,78,78,81,81	(b)
82-90	(c)	9
91-99	(d)	6

TABLA 3.3

Calificación en porcentaje	Número del profesor
57	13
59	7
65	11
67	8
67	22
67	29
69	17
72	5
73	16
73	18
77	3
78	21
78	28
81	26
81	30
83	1
83	15
83	27
84	9
84	19
87	25
88	4
89	12
89	23
91	5
91	20
93	24
95	14
97	13
99	2

2, (a) 6, (b) 7, (c) 83,83,83,84,84,87,88,89,89
(d) 91,91,93,95,97,99

146 A la frecuencia de cada intervalo se le llama *frecuencia de clase* (o simplemente frecuencia). Entonces, una parte de la segunda etapa en el proceso de agrupamiento de datos consiste en encontrar las __________ de clase.

frecuencias

147 Una parte del proceso de agrupamiento de datos, consiste en obtener las __________ __________

frecuencias de clase

148 Existe otro tipo de frecuencia llamado *frecuencia relativa de clase* (o simplemente frecuencia relativa). Ésta es el resultado de dividir la frecuencia de un intervalo entre el número total de observaciones. Por ejemplo, si la frecuencia de un intervalo es 2 y se dispone de 30 observaciones, la frecuencia __________ de ese intervalo es 2/30 = 0.067.

relativa

149 Complete la siguiente tabla, que es continuación de la usada en el cuadro 145, tomando en cuenta que el número total de observaciones es 30.

Intervalo de clase	Frecuencia	Frecuencia relativa
55-63	2	2/30 = 0.067 (6.7%)
64-72	6	(a)
73-81	7	7/30 = 0.233 (23.3%)
82-90	9	(b)
91-99	6	(c)

(a) 6/30 = 0.200; (20%), (b) 9/30 = 0.300; (30%),
(c) 6/30 = 0.200; (20%)

150 Observe la tabla 3.4.
¿Cuáles son las frecuencias de clase de los intervalos 81-90 y 91-100? __________, __________

TABLA 3.4

Evento (intervalo de calificaciones)	Frecuencia
A: 51-60	2
B: 61-70	5
C: 71-80	6
D: 81-90	11
E: 91-100	6

11, 6.

151 Observe la tabla 3.4. ¿A qué intervalo corresponde una frecuencia de clase de 5?

61-70.

152 Observe la tabla 3.4. La frecuencia de ________ del primer intervalo es ___.

clase, 2

153 Observe la tabla 3.5. ¿Cuál es la frecuencia relativa de clase del intervalo 91-100?

TABLA 3.5

Evento (intervalo de calificaciones)	Frecuencia relativa
A: 51-60	[illegible]
B: 61-70	[illegible]
C: 71-80	0.200
D: 81-90	0.367
E: 91-100	0.200

0.200

---

154 Observa la tabla 3.5. La frecuencia ________________ de clase del primer intervalo es ________.

relativa, 0.067

PARTE C. FRECUENCIA RELATIVA ACUMULADA

155 Existe otro tipo de frecuencia llamado *frecuencia acumulada de clase* (o simplemente frecuencia acumulada), la cual, como su nombre lo indica, es la suma de las frecuencias de clase del intervalo en consideración y de los intervalos anteriores.

Observe la tabla 3.4. La frecuencia acumulada de clase del segundo intervalo es 5 + 2 = 7; el 5 corresponde al segundo intervalo y el 2 al primero.

Análogamente la frecuencia acumulada del tercer intervalo es 6 + ____ + 2 = 13.

TABLA 3.4

Evento (intervalo de calificaciones)	Frecuencia
A: 51-60	2
B: 61-70	5
C: 71-80	6
D: 81-90	11
E: 91-100	6

5

---

156 Use la tabla 3.4. La frecuencia acumulada de clase del cuarto intervalo es 11 + (+ ___ + ___ = 24.

6, 2

---

157 La frecuencia acumulada de un intervalo es la suma de la ________ del intervalo en consideración y de las frecuencias de los intervalos anteriores.

frecuencia

---

158 Use la tabla 3.4. ¿Cuál es la frecuencia acumulada de clase del último intervalo?

30;(6 + 11 + 6 + 5 + 2 = 30)

---

159 Recuerde que el número de datos que integran la tabla 3.4 es 30.
La frecuencia acumulada del último intervalo, calculada en el cuadro anterior, es también 30. Esto indica que la frecuencia acumulada de clase del último intervalo es ______________ al número total de datos. De este hecho se obtiene una
(igual/diferente)
regla para verificar que las frecuencias estén bien calculadas: la suma de todas las frecuencias debe ser igual al número ________ de datos.

igual, total

---

160 La frecuencia __________ de clase de un intervalo es la suma de la frecuencia del intervalo y de las ___________ de los intervalos anteriores. Por esto la frecuencia acumulada de clase representa la frecuencia de los valores *menores o iguales* que el límite real *superior* del intervalo en consideración.

acumulada, frecuencias

---

161 Observe la tabla 3.6. La frecuencia acumulada de clase del segundo intervalo es ________. ¿Cuál es el límite real superior de este intervalo, cuando las observaciones se aproximan al entero más cercano?

TABLA 3.6

Evento (intervalo de calificaciones)	Frecuencia	Frecuencia acumulada
A: 51-60	2	2
B: 61-70	5	7
C: 71-80	6	13
D: 81-90	11	24
E: 91-100	6	30
Total	30	

7, 70.5

---

162 Puesto que 7 es la frecuencia acumulada del segundo intervalo, cuyo límite real superior es 70.5, dicha frecuencia representa la frecuencia de los valores ______ (menores/mayores) ___________ o iguales que 70.5.

menores

163 La frecuencia ________ de clase de un intervalo representa la frecuencia de todos los valores menores o iguales que el límite real ________ de ese intervalo.
(inferior/superior)

acumulada, superior

---

164 La frecuencia acumulada de un intervalo representa la frecuencia de todos los valores menores o ____________ que el límite ________ superior de ese intervalo.
(iguales/diferentes)

iguales, real

---

165 Observe la tabla 3.6. La frecuencia de los valores menores o iguales que 80.5 es el valor de la frecuencia ____________ del intervalo cuyo límite real superior es ________; ésta es igual a ________.

TABLA 3.6 [illegible]

Clase (intervalo de calificaciones)	Frecuencia	Frecuencia acumulada
A: 51–60	7	7
B: 61–70	[illegible]	7
C: 71–80	6	13
D: 81–90	11	24
E: 91–100	6	30
TOTAL:	30	

acumulada, 80.5, 13

---

166 Observe la tabla 3.6. ¿Cuál es la frecuencia de los valores menores o iguales que 90.5?

24

---

167 Observe en la tabla 3.6, que la frecuencia acumulada de los valores menores o iguales que 100.5 es ________. Es lógico que esta frecuencia sea igual al número total de datos, puesto que, siendo 30 estudiantes, y 100 la máxima calificación posible, todos debieron obtener una calificación menor o igual que 100.

30

---

168 El valor máximo de las frecuencias acumuladas es, por lo tanto, igual al número total de ______, o lo que es lo mismo, la suma de todas las ________ debe ser igual al ________ total de datos.

datos (observaciones), frecuencias, número

---

169 Use la tabla 3.6. Una tabla que presenta las frecuencias acumuladas de clase se llama *tabla de distribución de frecuencias acumuladas*.
La tabla 3.6 es una tabla de ____________ ____________.

distribución de frecuencias acumuladas

---

170 Una tabla que presenta las frecuencias acumuladas de clase se llama tabla de distribución de ____________ ____________.

frecuencias acumuladas

---

171 Observe que para calcular la frecuencia acumulada del tercer intervalo no necesita hacer la suma 7+6+2, sino que es suficiente con sumar 7+8, donde 8 es la frecuencia acumulada del intervalo anterior (complete la tabla).

Intervalo de clase	Frecuencia	Frecuencia acumulada
55–63	2	2
64–72	6	8;(2+6=8)
73–81	7	15;(7+8=15)
82–90	9	
91–99	6	
TOTAL:	30	

24;(15 + 9 = 24), 30;(24 + 6 = 30)

---

172 Se entiende por *frecuencia relativa acumulada de clase* (o simplemente frecuencia relativa acumulada) de un intervalo, la frecuencia acumulada de ese intervalo dividida entre el número total de datos de la muestra.
Observe la tabla 3.6. La frecuencia acumulada del primer intervalo es ________ y el total de observaciones es 30. Entonces, la frecuencia relativa acumulada de ese intervalo es 2/ _____ = 0.067.

2, 30

---

173 La frecuencia ____________ acumulada de un intervalo es igual a la frecuencia acumulada de ese intervalo dividida entre el número total de observaciones.

relativa

174 Use la tabla 3.6.

La ______________ ______________

acumulada del segundo intervalo es

______/30 = 0.233.

TABLA 3.6 REPETIDA

Clase (intervalo de calificaciones)	Frecuencia	Frecuencia acumulada
A: 51–60	2	2
B: 61–70	5	7
C: 71–80	6	13
D: 81–90	11	24
E: 91–100	6	30
TOTAL	30	

frecuencia relativa, 7

175 Use la tabla 3.6. ¿Cuáles son las frecuencias relativas acumuladas de los 3 últimos intervalos?

13/30;(0.433), 24/30;(0.800), 30/30;(1.000).

176 Una tabla que presenta las frecuencias relativas acumuladas de clase se conoce como *tabla de distribución de frecuencias relativas acumuladas*.

Observe la tabla 3.7. ¿Es ésta una tabla de distribución de frecuencias relativas acumuladas?

TABLA 3.7

Intervalo de calificaciones	Frecuencia relativa acumulada
51–60	0.067
61–70	0.233
71–80	0.433
81–90	0.800
91–100	1.000

Sí.

177 Continuando con el ejemplo que se ha venido desarrollando, complete la siguiente tabla de distribución de frecuencias relativas acumuladas de clase. Aproxime al milésimo más cercano.

(Recuerde que el total de datos es 30).

Intervalo	Frecuencia acumulada	Frecuencia relativa acumulada
55–63	2	0.067=6.7%
64–72	8	
73–81	15	
82–90	24	
91–99	30	

Frecuencia relativa acumulada
0.067=6.7%
0.267=26.7%;(8/30=0.266)
0.500=50.0%;(15/30=0.500)
0.800=80.0%;(24/30=0.800)
1.000=100.0%;(30/30=1.000)

178 Observe en la respuesta al cuadro anterior que la *mayor* frecuencia relativa acumulada es 1.000, es decir, todos los datos tienen valores ______________ o
(menores/mayores)
iguales que el mayor de los límites reales superiores. De este hecho se deriva la regla para verificar que las frecuencias ______________ estén bien calculadas; la suma de todas ellas debe ser igual a 1.0.

menores, relativas

179 Además, puesto que *no hay* ningún dato menor que el primer límite real inferior, la frecuencia relativa de los valores menores que ese límite es ________.
(cero/uno)

cero

180 Por lo visto anteriormente, las frecuencias relativas acumuladas de clase están comprendidas entre cero y uno.

¿Cuál es el mínimo valor posible de las frecuencias relativas acumuladas de clase?

¿Cuál es el máximo?

Cero. Uno.

181 Es conveniente presentar los diversos tipos de frecuencias en una sola tabla, la cual se denomina *tabla de distribuciones de frecuencias*.

Observe la tabla 3.8. ¿Se presentan en ella todos los tipos de frecuencias?

TABLA 3.8

Intervalos de clase	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
51–60	2	0.067	2	0.067
61–70	5	0.166	7	0.233
71–80	6	0.200	13	0.433
81–90	11	0.367	24	0.800
91–100	6	0.200	30	1.000
TOTAL	30	1.000		

Sí.

182 Una tabla que presenta *dos o más* tipos de frecuencias se llama tabla de distribuciones de ___________.

frecuencias

183 Otra manera de calcular las frecuencias relativas acumuladas, consiste en sumarle a la frecuencia relativa de cada intervalo las correspondientes a todos los intervalos anteriores.
Observe la tabla 3.8. ¿Cuál es la frecuencia relativa acumulada del segundo intervalo?

TABLA 3.8 PUNTOS

Intervalos de clase	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
51-60	2	0.067	2	0.067
61-70	5	0.166	7	0.233
71-80	6	0.200	13	0.433
81-90	11	0.367	24	0.800
91-100	6	0.200	30	1.000
TOTAL	30	1.000		

0.233

184 Use la tabla 3.8. ¿Cuánto vale la suma de las frecuencias relativas del primero y segundo intervalos? ¿Es el resultado igual a la frecuencia relativa acumulada del segundo intervalo? ___________, ___________.

0.233;(0.067 + 0.166 = 0.233), Sí.

185 La equivalencia de ambos procedimientos radica en que uno primero suma las frecuencias y luego divide entre el total de observaciones, y el otro hace la operación inversa; es decir, primero divide cada frecuencia entre el total de observaciones y luego suma los resultados.
Use la tabla 3.8. Siguiendo este último procedimiento, la frecuencia relativa acumulada del tercer intervalo vale 0.200 + ______ + ______ = 0.433.

0.166, 0.067

186 *Resumiendo*: la *primera* etapa en el proceso de agrupamiento de datos consiste en fijar los ___________________________.

Intervalos de clase

187 La *segunda* etapa en el proceso de agrupamiento de datos consiste en calcular los cuatro diferentes tipos de frecuencia de los intervalos. ¿Cuáles son?

Frecuencia, Frecuencia relativa, Frecuencia acumulada, Frecuencia relativa acumulada.

188 La primera etapa en el proceso de agrupamiento de datos se divide en cuatro partes:

1. Calcular el rango de los datos.
2.
3.
4.

2.Determinar el número de intervalos. 3.Calcular la amplitud de los intervalos. 4.Determinar los límites de clase.

189 Para fijar los límites de clase es necesario calcular primero el *ancho* (amplitud de los intervalos de clase) éste se calcula dividiendo el _______________ entre el número de intervalos deseado.
(rango/valor mínimo)

rango

190 La *segunda* etapa en el proceso de agrupamiento de ________ consiste en calcular las frecuencias correspondientes a cada intervalo de clase. Para comprobar los cálculos, la suma de todas las ___________ debe ser igual al número total de datos, y la suma de todas las ___________ ___________ debe ser 1.0.
Observe la tabla 3.8. ¿Se satisfacen estas condiciones?

datos, frecuencias, frecuencias relativas, Sí.

191 Como resultado de la aplicación de estas dos etapas del agrupamiento de datos, se obtiene una serie de tablas. Por esto se dice que se ha llegado a una *presentación tabular de los datos*. Análogamente, en las tablas 3.1, 3.2 y 3.3 se ha hecho la presentación ___________ de los datos básicos.

tabular

192 Cuando se presentan los datos en forma de tablas, se dice que se tiene una presentación ______________ de los datos.
Las tablas de la 3.4 a la 3.8 inclusive constituyen la ______________ ______________ de los datos *no básicos*, ya que la información contenida en ellas no es la inicial, sino una información obtenida a partir de los datos básicos.

tabular, presentación tabular

---

193 Practiquemos de nuevo todo el proceso de agrupamiento de datos, usando las observaciones de estaturas presentadas en la tercera columna de la tabla 3.1. Ordene los datos en forma creciente; hágalo en la hoja de trabajo 3.1.

[illegible]

[illegible]	[illegible]	[illegible]
1	[illegible]	[illegible]
2	[illegible]	[illegible]
3	[illegible]	[illegible]
4	[illegible]	[illegible]
5	[illegible]	[illegible]
6	[illegible]	[illegible]
7	[illegible]	[illegible]
8	[illegible]	[illegible]
9	[illegible]	[illegible]
10	[illegible]	[illegible]
11	[illegible]	[illegible]
12	[illegible]	[illegible]
13	[illegible]	[illegible]
14	[illegible]	[illegible]
15	[illegible]	[illegible]
16	[illegible]	[illegible]
17	[illegible]	[illegible]
18	[illegible]	[illegible]
19	[illegible]	[illegible]
20	[illegible]	[illegible]
21	[illegible]	[illegible]
22	[illegible]	[illegible]
23	[illegible]	[illegible]
24	[illegible]	[illegible]
25	[illegible]	[illegible]
26	[illegible]	[illegible]
27	[illegible]	[illegible]
28	[illegible]	[illegible]
29	[illegible]	[illegible]
30	[illegible]	[illegible]

Datos de estaturas ordenadas en forma creciente	
Estatura, en cm	Número del profesor
160	5
161	17
163	6
163	14
163	22
167	1
167	13
167	18
167	26
168	8
168	24
168	30
168	2
169	29
170	25
171	7
171	28
173	23
174	12
175	3
175	9
176	15
178	16
179	10
181	4
181	11
183	19
184	27
187	21
191	20

194 Observe la tabla 3.9 que presenta los datos ordenados del problema del cuadro anterior.

a) ¿Cuál es el dato de valor mínimo?
b) ¿Cuál es el dato de valor máximo?
c) ¿Cuál es el rango?

TABLA 3.9

Datos ordenados en forma creciente	
Estatura, en cm	Número del profesor
160	5
161	17
163	6
163	14
163	22
167	1
167	13
167	18
167	25
168	8
168	24
168	30
169	7
169	29
170	26
171	7
171	28
173	23
174	12
175	3
175	9
175	15
178	16
179	10
181	4
181	11
183	19
184	27
187	21
191	20

160, 191, 31;(191 - 160 = 31).

195 Determine el número de intervalos de clase, tratando de incluir alrededor de 6 datos en cada intervalo (recuerde que hay 30 datos).

6;(30/5 = 6).

196 Calcule el ancho de los intervalos de clase (recuerde que el rango es 31).

6;(31/6 = 5.1, el cual redondeado al número entero superior más próximo da 6).

197 Determine los límites de clase (recuerde que el valor mínimo (160) debe quedar incluido en el primer intervalo, y el valor máximo (191) en el último y que, además, debe tener 6 intervalos de 6 unidades de ancho cada uno).

Existen varias soluciones pero el primer límite de clase debe ser 156, 157, 158, 159 ó 160. A continuación se da una de ellas: 160-165; 166-171; 172-177; 178-183; 184-189; 190-195.

198 Use la hoja de trabajo 3.2 y la tabla 3.9. Complete la *tabla de distribuciones de frecuencias* que se presenta en la hoja de trabajo, aprovechando los datos ordenados de la tabla. Verifique que la suma de todas las frecuencias sea 30 (número total de datos), y que la suma de todas las frecuencias relativas sea 1.0.

Intervalo de estaturas	Elementos que contienen	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
157-162	160, 161	2	0.067 (6.7%)	2	0.067
163-168	163, 163, 163, 167, 167, 167, 167, 168, 168, 168	10	0.333 (33.3%)	12	0.400
169-174	169, 169, 170, 171, 171, 173, 174	7	0.233 (23.3%)	19	0.633
175-180	175, 175, 175, 178, 179	5	0.167 (16.7%)	24	0.800
181-186	181, 181, 183, 184	4	0.133 (13.3%)	28	0.933
187-192	187, 191	2	0.067 (6.7%)	30	1.000
	TOTAL	30	1.000 (100%)		

199 Observe la hoja de trabajo 3.2.
Una tabla, como la presentada en la hoja de trabajo 3.2, en la que se presentan diversos tipos de frecuencias, se llama ______________________

tabla de distribuciones de frecuencias

REVISION

200 Al reunir los diferentes elementos de una población en grupos de uno o más, cada uno de estos grupos constituye un ______________

evento

201 Para que un evento ocurra, se necesita que al realizar un experimento, la observación sea ______________ (igual a/diferente que) alguno de sus elementos.

igual a

202 El número de veces que ocurre un evento se llama ______________.

frecuencia

203 Si el evento A es "número comprendido en el intervalo del 3 al 5 inclusive", y las observaciones en varios experimentos son 1.5, 3.0, 2.5, 4.0, 5.5, 4.5, 9.0, 7.5, 4.0 y 8.5, ¿cuál es la frecuencia del evento A?

Cuatro.

---

204 Al dividir la frecuencia de un evento entre el número total de observaciones, se obtiene la ______________ ______________.

frecuencia relativa

---

205 Relacione los símbolos de la izquierda con las descripciones de la derecha.

1.	$n(A)$	I.	Frecuencia relativa del evento A
2.	$h(A)$	II.	Número total de datos
3.	N	III.	Frecuencia del evento A

1-III, 2-I, 3-II

---

206 Complete la fórmula para calcular la frecuencia relativa del evento A:

$$(?) = \frac{(?)}{(?)}$$

$$h(A) = \frac{n(A)}{N}$$

---

207 Relacione los nombres de la izquierda con las descripciones de la derecha.

1.	Tabla de datos agrupados	I.	Contiene los eventos y sus respectivas frecuencias.
2.	Tabla de datos ordenados	II.	Contiene los datos básicos agrupados.
3.	Tabla de distribución de frecuencias relativas	III.	Contiene los eventos y sus respectivas frecuencias relativas.
4.	Tabla de distribución de frecuencias	IV.	Contiene los datos ordenados.

1-II, 2-IV, 3-III, 4-I

---

208 Para tener una descripción más concreta de la distribución de los valores de los datos, se agrupan éstos en ______________ ___ ______________, y se obtienen las frecuencias, frecuencias relativas, frecuencias acumuladas y frecuencias relativas acumuladas.

intervalos de clase

---

209 Relacione los nombres de la izquierda con las descripciones de la derecha.

1.	Frecuencia relativa de clase	I.	Es el número de elementos que pertenecen al intervalo de clase.
2.	Frecuencia acumulada de clase	II.	Es la suma de la frecuencia del intervalo y de las frecuencias de los intervalos anteriores.
3.	Frecuencia de clase	III.	Es la frecuencia acumulada dividida entre el número total de datos.
4.	Frecuencia relativa acumulada de clase	IV.	Es la frecuencia dividida entre el número total de datos.

1-IV, 2-II, 3-I, 4-III

---

210 Al agrupar los datos en intervalos de clase se siguen dos etapas.
La primera consiste en ____ (a/b), y la segunda en ____ (a/b).

a. calcular las frecuencias de clase
b. determinar los intervalos de clase

b. determinar los intervalos de clase
a. calcular las frecuencias de clase

---

211 La primera etapa en el agrupamiento de datos consiste en los cuatro pasos indicados a continuación. Diga cuál es el orden correcto.

a. Fijar el número de intervalos.
b. Calcular el rango de los datos.
c. Fijar los límites de clase.
d. Calcular la amplitud de los intervalos.

b-a-d-c

212 Relacione los nombres de la izquierda con las descripciones de la derecha.

1. Rango	I. Es la diferencia entre los límites reales superior e inferior correspondientes.
2. Límites de clase	II. Es la diferencia entre los valores máximo y mínimo de los datos.
3. Amplitud del intervalo de clase	III. Son los límites de los intervalos de clase.

1-II, 2-III, 3-I

213 Supóngase que los datos fueron aproximados al *medio* centímetro más cercano y se agruparon en la siguiente forma:

(a)	(b)
20.0–22.5	19.75–22.75
23.0–25.5	22.75–25.75
26.0–28.5	25.75–28.75
etc.	etc.

Los límites de clase son los de la columna ____ (a/b), y los límites reales de clase son los de la columna ____ (a/b).

a, b

214 De acuerdo con la regla presentada en esta unidad, si se tienen 25 datos, ¿cuántos intervalos de clase usaría?

5 (25/5 = 5).

215 Use la tabla 3.10, de datos ordenados, correspondiente a la resistencia, en toneladas, de los cables de acero producidos en una fábrica.
¿Cuánto vale el rango?

TABLA 3.10

47.5	51.0	53.5	56.0	60.0
48.0	52.0	53.5	56.5	60.5
50.0	53.0	54.0	57.0	61.0
50.5	53.0	54.0	[illegible]	62.5
51.0	53.5	55.0	59.0	63.0

15.5 (63.0 – 47.5 = 15.5 ton).

216 Si el rango de los datos del cuadro anterior es 15.5 ton y se usarán 5 intervalos de clase, ¿cuál será la amplitud de los intervalos?

4; (15.5/5 = 3.1, valor que redondeado al entero superior más cercano da 4 ton).

217 Complete la siguiente tabla tomando en cuenta que hay 5 intervalos de 4 ton de amplitud, y que los datos fueron aproximados a la media tonelada más cercana. Tome como primer límite de clase a 47.0 ton.

Intervalo de clase	
Límites de clase	Límites reales de clase

Límites de clase	Límites reales de clase
47.0–50.5	46.75–50.75
51.0–54.5	50.75–54.75
55.0–58.5	54.75–58.75
59.0–62.5	58.75–62.75
63.0–66.5	62.75–66.75

218 Complete la siguiente tabla de distribuciones de frecuencias, correspondiente a los datos de la tabla 3.10.

Intervalo de clase	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
47.0–50.5				
51.0–54.5				
55.0–58.5				
59.0–62.5				
63.0–66.5				
TOTAL				

Intervalo de clase	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
47.0–50.5	4	0.16	4	0.16
51.0–54.5	10	0.40	14	0.56
55.0–58.5	5	0.20	19	0.76
59.0–62.5	5	0.20	24	0.96
63.0–66.5	1	0.04	25	1.00
TOTAL:	25	1.00		

219 Al calcular las frecuencias de clase, se pueden comprobar los resultados, ya que la suma de todas ellas debe ser igual al número ________ de ________.

total, datos

220 De la misma manera, la suma de todas las frecuencias relativas debe ser ________ (cuánto)

1

221 En una colonia de la Ciudad de México se obtuvo una muestra aleatoria del monto del consumo mensual de corriente eléctrica. Los datos ordenados son los siguientes

Consumo mensual de corriente eléctrica, en pesos				
20	30	39	45	58
21	31	40	46	65
22	33	40	46	70
23	34	40	50	80
24	35	41	50	105
24	36	41	51	110
25	36	43	51	145
25	37	44	53	145
27	37	45	54	150
29	38	45	55	203

a. ¿Cuál es el rango?
b. De acuerdo con la regla presentada en esta unidad, ¿cuántos intervalos de clase hay que usar?
c. Calcule la amplitud de los intervalos de clase.

a. 203-20 = 183. b. 50/5 = 10. c. 19; (183/10 = 18.3).

222 Complete la siguiente tabla para los datos del cuadro anterior. Use una amplitud de 20 para los intervalos de clase.

Límites de clase		Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Inferior	Superior				
20					

Límites de clase		Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Inferior	Superior				
20	39	21	21	0.42	0.42
40	59	20	41	0.40	0.82
60	79	2	43	0.04	0.86
80	99	1	44	0.02	0.88
100	119	2	46	0.04	0.92
120	139	0	46	0.00	0.92
140	159	3	49	0.06	0.98
160	179	0	49	0.00	0.98
180	199	0	49	0.00	0.98
200	219	1	50	0.02	1.00
		TOTAL 50		1.00	

EXAMEN

1. *Al dividir los diferentes elementos de una población en grupos de uno o más elementos, cada grupo constituye un ________________.*

2. *Si la población "pesos de los niños recién nacidos" se divide en los eventos*

   A: 3 kg o menos
   B: de 3.1 kg a 5.0 kg
   C: 5.1 kg o más

   y si un niño pesa 4.8 kg, ocurre el evento____; si otro pesa 5.6 kg, ocurre el evento ___.

3. *Sea el evento "Evaristo Hernández lanza la jabalina a 50 m". Para que este evento ocurra se necesita que en un lanzamiento la jabalina caiga a ___ m o más.*

4. *El número de veces, n(A), que ocurre un evento al repetir un experimento N veces, se llama ________________ del evento.*

5. *Al cociente n(A)/N se le llama ____________ ____________.*

6. *Evaristo Hernández lanzó la jabalina 10 veces, 8 de ellas a una distancia de 30 m o más (evento A). En tal caso:*

   $n(A) =$ ____
   $N =$ ____

   frecuencia del evento A = ____
   frecuencia relativa de A = ____

7. *Una tabla que muestra los eventos con sus respectivas frecuencias se llama ____.*

8. *Los pesos de 10 niños recién nacidos son: 2.7, 3.6, 3.4, 4.1, 5.1, 3.2, 2.6, 6.3, 2.9 y 5.0 kg.*

   a. *Organice los datos en una tabla de datos ordenados.* (10 puntos)
   b. *¿Cuál es el rango de las observaciones?* ________________.
   c. *Organice los datos en una tabla de datos agrupados, de acuerdo con los eventos de la pregunta 2.* (13 puntos)
   d. *De acuerdo con los eventos de la pregunta 2, complete la siguiente tabla* (11 puntos)

Evento	Datos

Evento	Frecuencia	Frecuencia relativa

9. *La suma de las ________________ de todos los eventos debe ser igual al número total de datos.*

10. *El mínimo valor que puede asumir una frecuencia relativa es ___; el máximo es ___.*

11. *Relacione los nombres de la columna de la izquierda con las descripciones que aparecen a la derecha.*

    1. *Frecuencia de clase* — a. *Son los límites de los intervalos de clase y tienen el mismo número de cifras decimales que los datos básicos.*
    2. *Límites de clase* — b. *Es el número de datos que corresponden al intervalo.*
    3. *Límites reales* — c. *Son límites de los intervalos de clase y tienen una cifra decimal más que los datos básicos.*

12. *El menor de los límites de clase de un intervalo se llama ________________ ________________; el mayor se llama ________________ ________________.*

13. *En un proceso de control de calidad de los productos de una fábrica de explosivos se obtuvo la siguiente muestra aleatoria del peso de los cartuchos de dinamita:*

Peso, en gramos

38.4	38.3	36.1	39.8
37.1	37.7	37.6	40.8
38.6	37.4	38.3	38.1
38.5	37.1	39.2	37.8
37.4	36.5	38.7	36.7
37.3	36.3	38.2	38.3
39.0	38.0	36.2	39.0
37.7	39.2	38.8	38.3
39.5	37.0	39.5	36.9
37.4	38.2	39.2	38.8

a. *¿Cuál es el rango?*
b. *¿Cuántos intervalos de clase hay que usar, si incluimos 5 datos en cada intervalo?*
c. *Calcule la amplitud de los intervalos.*
d. *Tomando como primer límite inferior de clase a 36.1, complete la siguiente tabla (34 puntos):*

Límites de clase		Límites reales	
Inferior	Superior	Inferior	Superior
36.1			

e. *Complete la siguiente tabla (34 puntos):*

Límites de clase		Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Inferior	Superior				
36.1	36.6				
36.7	37.2				
37.3	37.8				
37.9	38.4				
38.5	39.0				
39.1	39.6				
39.7	40.2				
40.3	40.8				
TOTAL		———		———	

TOTAL: 122 puntos.

RESPUESTAS

1. *evento*

2. *B*
C

3. *30*

4. *frecuencia*

5. *frecuencia relativa*

6. *8*
10
8
0.8

7. *tabla de distribución de frecuencias*

8. a. *2.6, 2.7, 2.9, 3.2, 3.4, 3.6, 4.1, 5.0, 5.1, 6.3 kg*
b. *6.3 − 2.6 = 3.7 kg*
c.

Evento	Datos
A	2.6, 2.7, 2.9
B	3.2, 3.4, 3.6, 4.1, 5.0
C	5.1, 6.3

d.

Evento	Frecuencia	Frecuencia relativa
A	3	0.3
B	5	0.5
C	2	0.2
TOTAL:	10	1.0

9. *frecuencias*

10. *0*
1

11. *1-b*
2-a
3-c

12. *límite inferior de clase*
límite superior de clase

13. a. $40.8 - 36.1 = 4.7$
b. $40/5 = 8$
c. $4.7/8 = 0.6$
d.

Límites de clase		Límites reales	
Inferior	Superior	Inferior	Superior
36.1	36.6	36.05	36.65
36.7	37.2	36.65	37.25
37.3	37.8	37.25	37.85
37.9	38.4	37.85	38.45
38.5	39.0	38.45	39.05
39.1	39.6	39.05	39.65
39.7	40.2	39.65	40.25
40.3	40.8	40.25	40.85

e.

	Frecuencia	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
	4	4	0.100	0.100
	5	9	0.125	0.225
	8	17	0.200	0.425
	9	26	0.225	0.650
	7	33	0.175	0.825
	5	38	0.125	0.950
	1	39	0.025	0.975
	1	40	0.025	1.000
TOTAL:	40		1.000	

HOJA DE TRABAJO 3.1

Datos ordenados en forma creciente	
Estatura, en cm	Número del profesor

HOJA DE TRABAJO 3.2

Intervalo de estaturas	Elementos que ocurrieron	Frecuencia	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativa acumulada
157-162					
163-168					
169-174					
175-180					
181-186					
187-192					
	TOTAL:				

UNIDAD IV PRESENTACION GRAFICA DE LAS DISTRIBUCIONES DE FRECUENCIAS

PREFACIO

En la unidad anterior estudiamos la forma de presentar tabularmente los datos; en ésta se estudiarán varias formas gráficas de presentarlos, las cuales proporcionan una manera objetiva de observar los resultados.

PARTE A. HISTOGRAMA

1 Existen varias formas de presentar gráficamente una distribución de frecuencias; la más usual se llama *histograma*.
El __________ es una gráfica de una distribución de frecuencias.

histograma

2 Un histograma constituye una representación __________ (gráfica/tabular) de una distribución de frecuencias. Primero estudiaremos la forma de trazar un histograma correspondiente a los datos de variables escalares; es decir, de aquéllas que sólo asumen valores __________ (nominales/numéricos).

gráfica, numéricos

3 Una forma común de representar gráficamente una distribución de frecuencias es mediante el uso de un __________.

histograma

4 Para dibujar un histograma necesitamos emplear un sistema de coordenadas *rectangulares o cartesianas.*
¿Qué sistema de coordenadas utilizaremos para dibujar un histograma?

Rectangulares (o cartesianas).

5 El sistema de coordenadas rectangulares que usaremos en esta parte consiste en dos ejes: uno horizontal, llamado eje de las abscisas, y otro vertical, llamado eje de las ordenadas.
¿Cuántos ejes vamos a utilizar al dibujar un histograma?

Dos.

6 El eje horizontal, o de las abscisas, representará la *variable aleatoria* bajo estudio. Si se está realizando un estudio estadístico del coeficiente de inteligencia de los estudiantes de bachillerato de la República Mexicana, en el eje de las abscisas se indicarán ____ (a/b).

a. los coeficientes de inteligencia
b. las edades

a. los coeficientes de inteligencia

7 El eje vertical, o de las ordenadas, representará las *frecuencias de clase* divididas entre la amplitud del intervalo de clase correspondiente.
El eje vertical de un sistema de coordenadas rectangulares, empleado para dibujar un histograma, representa las ____ (a/b).

a. frecuencias acumuladas
b. frecuencias de clase divididas entre la amplitud (o ancho) del intervalo de clase correspondiente

b. frecuencias de clase divididas entre la amplitud (o ancho) del intervalo de clase correspondiente

8 Al dibujar un histograma, el eje de las abscisas (horizontal) representa ____ (a/b/c).

a. la frecuencia acumulada

b. la variable aleatoria bajo estudio
c. los límites de clase

b. la variable aleatoria bajo estudio

9 Al dibujar un histograma, el eje de las ordenadas (vertical) representa ______ (a/b)

a. la población
b. las frecuencias de clase divididas entre la amplitud del intervalo de clase correspondiente

b. las frecuencias de clase divididas entre la amplitud del intervalo de clase correspondiente

10 Las frecuencias de clase divididas entre el ancho del intervalo de clase correspondiente, las llamaremos *frecuencias de clase normalizadas*, o simplemente *frecuencias normalizadas*.
Si una frecuencia de clase es igual a 6 y el ancho de su intervalo de clase es 10, la frecuencia de clase normalizada es 6/10 = 0.6. ¿Cuánto valdría la frecuencia de clase normalizada si la frecuencia fuese 8 en vez de 6?

0.8;(8/10=0.8).

11 Una frecuencia de clase normalizada es igual a la frecuencia de clase dividida entre el ______________ del ______________ de clase correspondiente.

ancho, intervalo

12 Una frecuencia de clase dividida entre el ancho (la amplitud) del intervalo de clase correspondiente se denomina ______________ de clase normalizada.

frecuencia

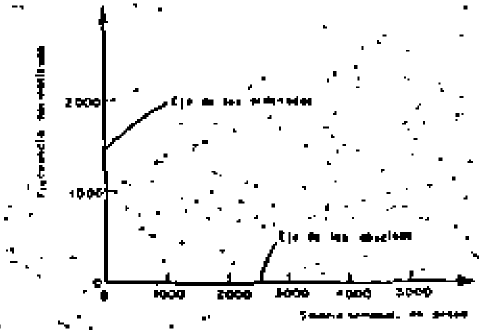
13 Si una frecuencia de clase es 15 y el ancho del intervalo correspondiente es 5, la frecuencia de clase ______________ es ______________.

normalizada, 3;(15/5=3)

14 Para dibujar un histograma, empleando un sistema de coordenadas rectangulares cuyo eje horizontal tenga anotados valores numéricos, se requiere que la variable aleatoria sea ______________ (nominal/escalar)

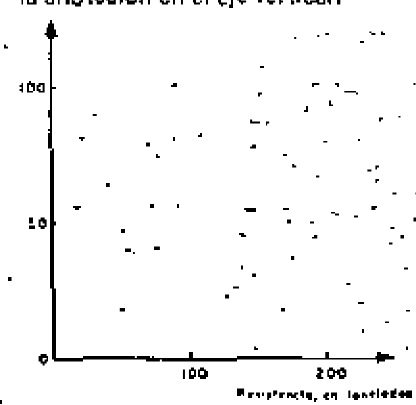
escalar

15 Si se está realizando un estudio estadístico del salario semanal que perciben los trabajadores mexicanos, se empleará el siguiente sistema de coordenadas; ¿están correctos los nombres de los ejes y los títulos de cada uno?



Sí.

16 Si se realiza un estudio estadístico de la carga que resisten los cables de acero producidos en una cierta fábrica, se usará el sistema de coordenadas mostrado abajo. En el eje horizontal se ha anotado "Resistencia, en toneladas"; ¿cuál será la anotación en el eje vertical?



Frecuencia normalizada.

17 Un histograma es una gráfica de *barras o rectángulos*; el ancho de cada rectángulo es igual al *ancho* del intervalo de clase; la altura de cada rectángulo es igual a la ***frecuencia de clase normalizada***.
Un histograma es una gráfica ____.
(a/b)

a. circular
b. de barras o rectángulos

b. de barras o rectángulos

---

18 En un histograma, el ancho de cada rectángulo es igual al _____ del intervalo de clase que representa.

ancho

---

19 **El ancho de cada rectángulo es igual al ancho del intervalo de clase que representa, y su altura es igual a** ____________________
(la frecuencia de clase normalizada/el rango)

la frecuencia de clase normalizada

---

20 Si todos los intervalos de clase tienen el mismo ancho, todos los rectángulos del histograma tendrán ____________ ancho.
(igual/diferente)

igual

---

21 Cuando el ancho (amplitud) de *todos* los intervalos de clase es el *mismo*, es común representar en el eje de las ordenadas la frecuencia de clase sin haberla dividido entre el ancho de los intervalos. Entonces, dependiendo de que el ancho de todos los intervalos de clase sea el mismo o no, el eje vertical representará las frecuencias de clase ________________ o simplemente las frecuencias de clase.

normalizadas

---

22 Cuando el eje vertical representa las frecuencias de clase normalizadas, las *áreas* de cada rectángulo son las frecuencias de clase, f. Esto se demuestra fácilmente, puesto que, si b es el ancho de un intervalo, la frecuencia normalizada es f/b. Al multiplicar la frecuencia de clase normalizada (la altura), por el ancho del intervalo para obtener el área del rectángulo se tendrá

$$\frac{f}{b} \times b = f$$

¿Representa el *área* de cada rectángulo la frecuencia del intervalo correspondiente, cuando en el eje de las ordenadas se anotan las frecuencias de clase normalizadas?

Sí.

---

23 **Al anotar en el eje vertical las frecuencias de clase normalizadas, el área de cada rectángulo representa la ________________ de clase correspondiente.**

frecuencia

---

24 **La frecuencia de clase es igual a ____________ del rectángulo correspondiente**
(la altura/el área)
cuando en el eje vertical se anotan las frecuencias de clase normalizadas.

el área

---

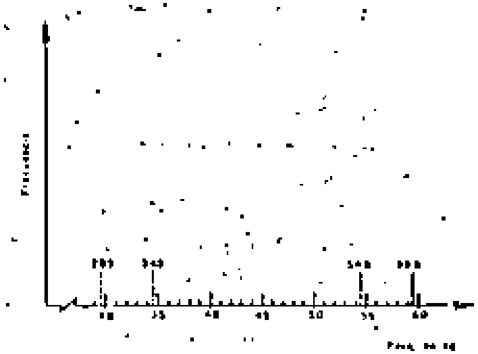
25 El extremo *izquierdo* de un rectángulo de un histograma, corresponde al *límite real inferior* del intervalo; el extremo *derecho* coincide con el *límite real superior*. La siguiente tabla corresponde a una muestra de los pesos de estudiantes de secundaria aproximados al kilogramo más cercano; complete la tabla.

Intervalo de clase	Límite real inferior	Límite real superior
30–34	29.5	34.5
35–39	34.5	39.5
40–44	(a) ____	(b) ____
45–49	(c) ____	(d) ____
50–54	49.5	54.5
55–59	54.5	59.5

(a) 39.5, (b) 44.5, (c) 44.5, (d) 49.5

---

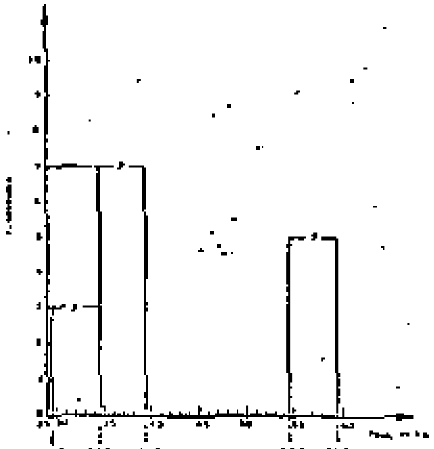
26 Sobre la gráfica que se presenta a continuación, marque en el eje de las abscisas los límites reales que se dan como respuesta al cuadro anterior.

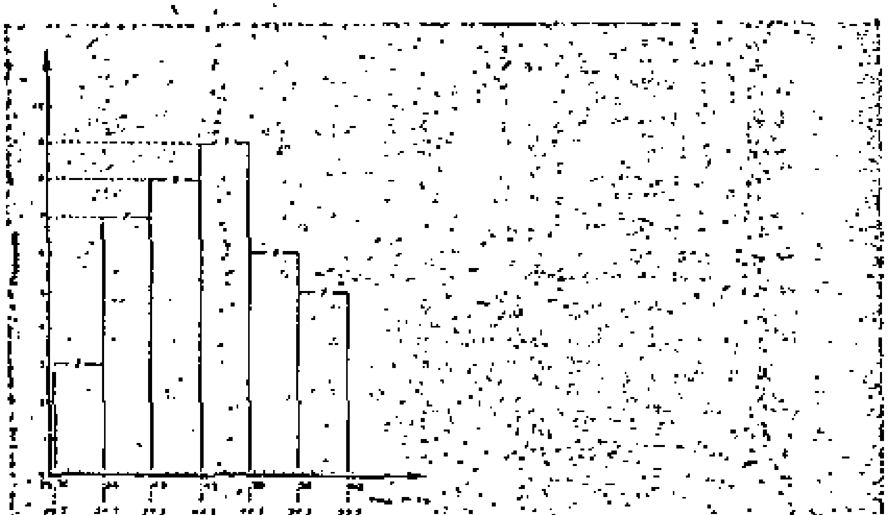


27 Complete el siguiente histograma, correspondiente a los datos dados en el cuadro 25, teniendo en cuenta la distribución de frecuencias dada en la tabla 4.1. Recuerde que los extremos de los rectángulos son los límites reales y que sus alturas son las frecuencias de clase cuando todos los intervalos de clase tienen el mismo ancho.

Tabla 4.1

Intervalo de clase	Límite real inferior	Límite real superior	Frecuencia
30–34	29.5	34.5	3
35–39	34.5	39.5	7
40–44	39.5	44.5	8
45–49	44.5	49.5	9
50–54	49.5	54.5	6
55–59	54.5	59.5	5



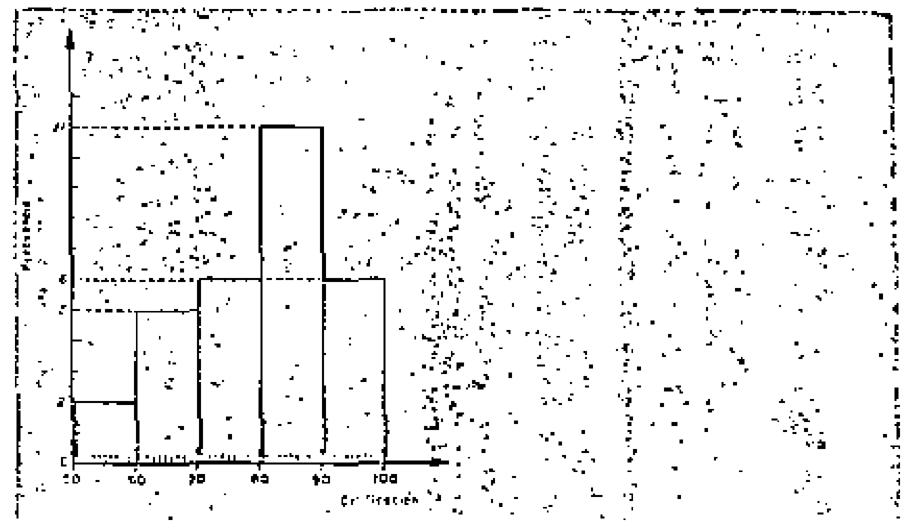


En el cuadro anterior se anotaron las frecuencias en la parte superior de cada rectángulo. Esto normalmente no es necesario, ya que en el eje vertical se encuentra la escala de frecuencias. Por la misma razón no es necesario anotar los límites reales.
La representación gráfica de la distribución de frecuencias que estamos estudiando se llama ________________

histograma

29 Tome la hoja de trabajo 4.1, y dibuje el histograma correspondiente a la distribución de frecuencias que se presenta a continuación.

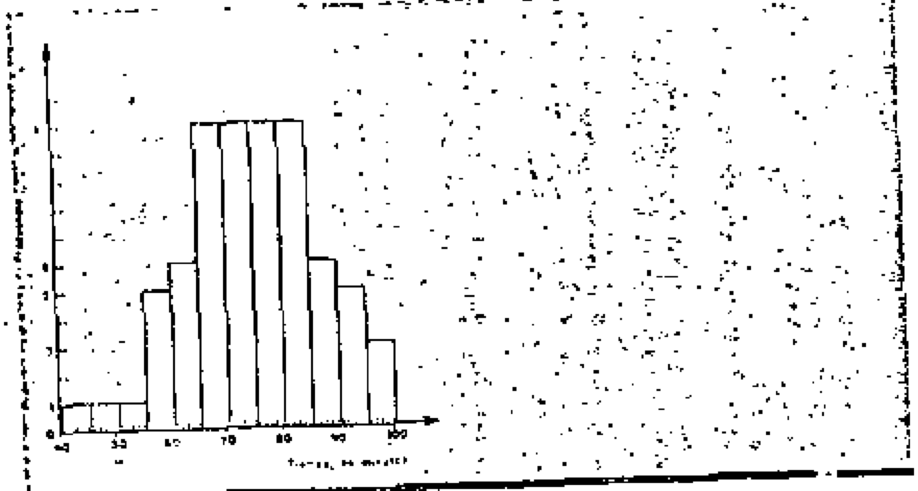
Intervalo de clase	Frecuencia
51–60	2
61–70	5
71–80	6
81–90	11
91–100	6



30 Use la hoja de trabajo 4.2 y la tabla 4.2. Dibuje, en la hoja de trabajo, el histograma correspondiente a la distribución de frecuencias de la tabla, la cual presenta los tiempos que necesitaron 30 alumnos de primer año de Ingeniería para efectuar una prueba de velocidad de aprendizaje. Los tiempos se aproximaron al minuto más cercano.

TABLA 4-2

Intervalo de clase, en minutos	Frecuencia
41–45	1
46–50	1
51–55	1
56–60	3
61–65	6
66–70	11
71–75	11
76–80	11
81–85	11
86–90	6
91–95	5
96–100	3



31 Observe la tabla 4.3, en ella se presenta otra distribución de frecuencias de los mismos datos del cuadro anterior. ¿Son todos los intervalos de clase del mismo ancho?

TABLA 4.3

Intervalo de clase, en minutos	Frecuencia
41–55	3
56–60	5
61–65	6
66–70	11
71–75	11
76–80	11
81–85	11
86–90	6
91–100	6

No.

32 Observe la tabla 4.3. ¿Cuál es el ancho (la amplitud) del primer intervalo de clase?

15.0.

33 Observe la tabla 4.3. ¿Cuál es el ancho de los intervalos de clase, del segundo al octavo?

TABLA 4.3 (repetida)

Intervalo de clase, en minutos	Frecuencia
41–55	3
56–60	5
61–65	6
66–70	11
71–75	11
76–80	11
81–85	11
86–90	6
91–100	6

5.0.

34 Observe la tabla 4.3. ¿Cuál es el ancho del último intervalo de clase?

10.0.

35 Por los resultados de los tres cuadros anteriores, se concluye que los intervalos de clase de la tabla 4.3 tienen _______________ ancho.
(el mismo/diferente)

diferente.

36 Puesto que los intervalos de clase de la tabla 4.3 tienen diferente ancho, es necesario, al dibujar el histograma correspondiente, anotar en el eje vertical las _______________
(frecuencias/frecuencias normalizadas)

frecuencias normalizadas

37 Recuerde que una frecuencia normalizada es igual a la __________ dividida entre el ancho del intervalo de clase correspondiente.

frecuencia

38 Use la tabla 4.3. ¿Cuánto vale la frecuencia normalizada del primer intervalo?

0.2; (3/15 = 0.2).

39 Observe la respuesta correcta del cuadro anterior se ha anotado en el lugar apropiado de la hoja de trabajo 4.3.

Use la tabla 4.3. ¿Cuánto vale la frecuencia normalizada de los intervalos segundo, tercero y cuarto? Después de verificar sus respuestas anótelas correctamente en la hoja de trabajo 4.3.

1; (5/5 = 1). 1.2; (6/5 = 1.2). 2.2; (11/5 = 2.2).

---

40 Use la tabla 4.3. ¿Cuánto vale la frecuencia normalizada del último intervalo? Anote su respuesta en la hoja de trabajo 4.3.

0.8;(8/10=0.8).

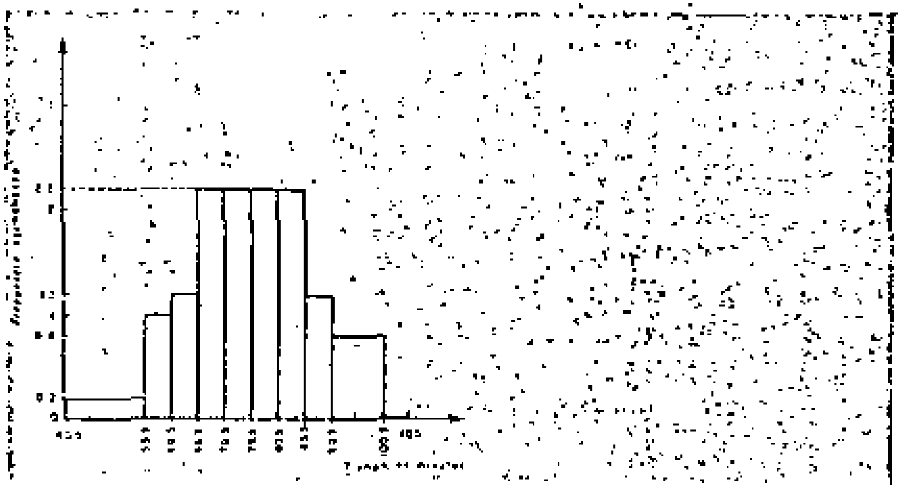
---

41 Use la hoja de trabajo 4.3. ¿Cuánto valen los límites reales superior e inferior, correspondientes a los tres primeros intervalos de clase? Recuerde que los tiempos se aproximaron al minuto más cercano. Después de verificar su respuesta anótela correctamente en la misma hoja de trabajo.

Inferior	Superior
41.5	55.5
55.5	60.5
60.5	65.5

---

42 Use la hoja de trabajo 4.4, para dibujar el histograma de los datos presentados en la hoja de trabajo 4.3 (recuerde que en el eje vertical tiene que anotar las frecuencias normalizadas). Indique en la figura los límites reales.



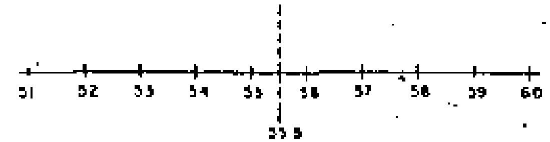
43 En ocasiones es conveniente indicar en el eje horizontal los [illegible] de los puntos *medios* de cada intervalo, los cuales se denominan *marcas de clase*.

El punto *medio* de cada intervalo de clase se llama ______________

marca de clase

---

44 El punto __________ de cada intervalo de clase se denomina *marca de clase*. Vea la siguiente figura y diga cuál es el punto medio del intervalo 51-60.



medio, 55.5.

---

45 Puesto que 55.5 es el punto medio del intervalo 51-60, éste es precisamente la marca de clase de ese intervalo. La marca de clase de un intervalo es el punto __________ de ese intervalo.

medio

---

46 El punto medio de un intervalo de clase se llama __________ de clase.

marca

---

47 Puesto que la marca de clase es el punto medio de un intervalo, éste se puede calcular sumándole al límite inferior de clase del intervalo, la mitad de la distancia entre los límites de clase. Siguiendo este procedimiento, la marca de clase del intervalo 51-60 es:

$$51 + \frac{60 - 51}{2} = 51 + 4.5 = 55.5$$

y del intervalo 61-70 es:

$$61 + \frac{\qquad}{\qquad} = \underline{\qquad} = \underline{\qquad}.$$

$$61 + \frac{70 - 61}{2} = 61 + 4.5 = 65.5$$

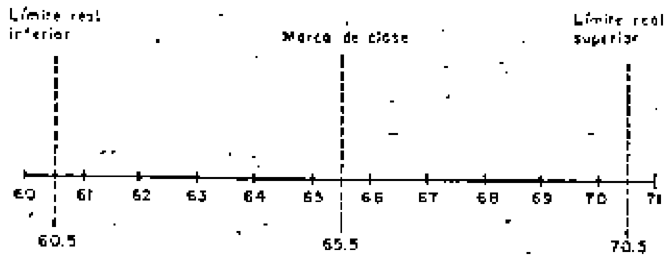
48 ¿Cuál es la marca de clase del intervalo 91-100?

95.5; (91+4.5=95.5).

49 Si las medidas están redondeadas al entero más cercano, los límites reales del intervalo 51 – 60 son 50.5 y 60.5. ¿Cuáles son los límites reales del intervalo 61 – 70?

60.5 y 70.5.

50 Observe la siguiente figura:



¿Corresponde la marca de clase al punto medio entre los límites reales?

Sí.

51 Puesto que la marca de clase es el punto medio entre los límites reales, ésta se puede calcular también sumándole al límite real inferior la mitad de la amplitud del intervalo comprendido entre los límites reales:

$$60.5 + \frac{70.5 - 60.5}{2} = 60.5 + 5 = 65.5$$

Siguiendo este procedimiento, calcule la marca de clase del intervalo 81 – 90, cuando las medidas están redondeadas al entero más cercano.

80.5 + 5 = 85.5

52 Utilizando los límites reales calcule la marca de clase del intervalo 157 – 162.

$159.5;\ (156.5 + \dfrac{162.5 - 156.5}{2} = 159.5)$

53 Si en el intervalo 41 – 55 de la tabla 4.3 hubo tres observaciones, éstos se le asignan a la marca de clase de ese intervalo, es decir a 48.0. ¿Cuántas observaciones se le asignan a la marca de clase del segundo intervalo (58)?

Intervalo de clase, en minutos	Frecuencia
41–55	3
56–60	5
61–65	6
66–70	11
71–75	11
76–80	11
81–85	11
86–90	6
91–100	8

5

54 Use la tabla 4.3.

a. ¿Cuál es la marca de clase del octavo intervalo?

b. ¿Qué frecuencia se le asigna a 88.0?

88. 6.

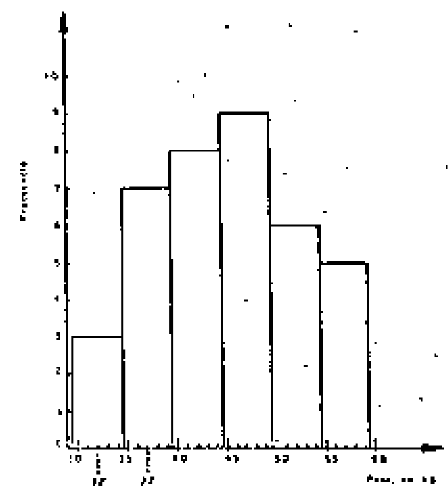
55 ¿Cuál es la marca de clase del intervalo 30 – 34, si se sabe que las observaciones se aproximaron al entero más cercano?

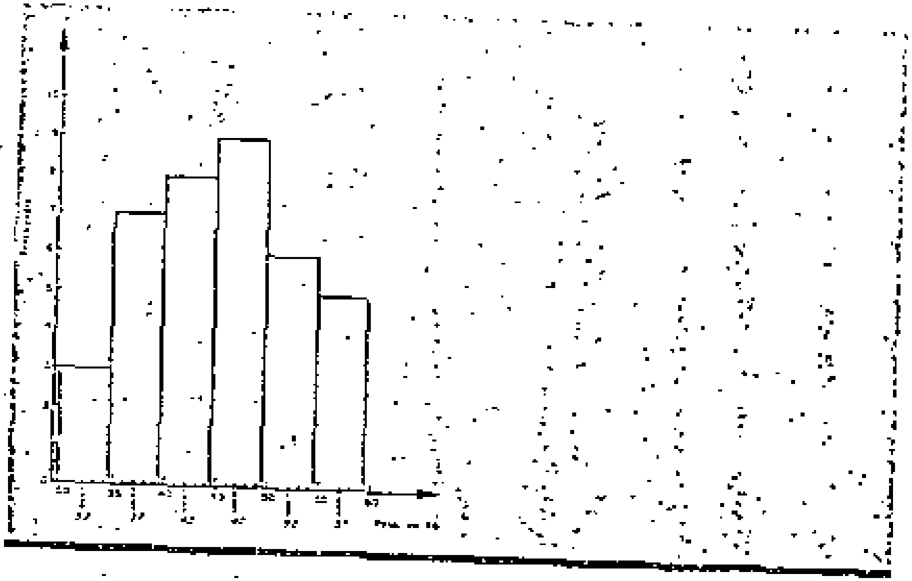
¿Cuál es la marca de clase del intervalo 35 – 39?

$32.0;\ (29.5 + \dfrac{34.5 - 29.5}{2} = 32.0).\quad 37.0;\ (34.5 + \dfrac{39.5 - 34.5}{2} = 37.0).$

PARTE B. POLIGONO DE FRECUENCIAS

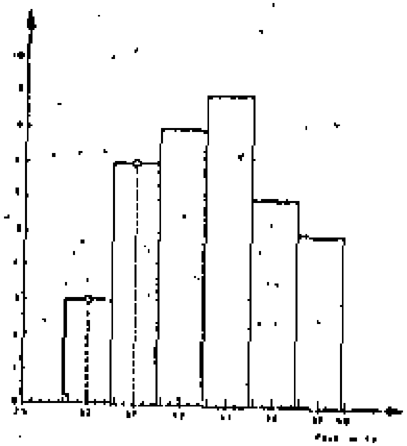
56 En el siguiente histograma se han indicado las dos marcas de clase (32 y 37) calculadas en el cuadro anterior. Indique en forma semejante las marcas de clase que faltan.

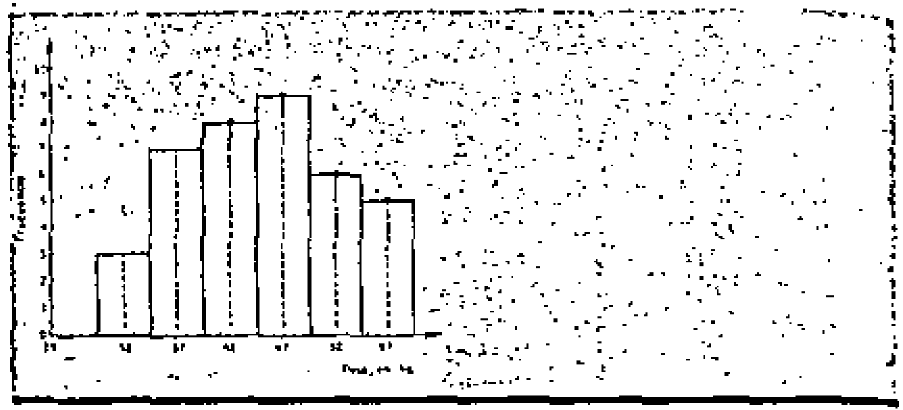




57 En la hoja de trabajo 4.4, indique en el eje horizontal todas las marcas de clase.

58 Observe que en el siguiente histograma se han indicado con un punto las marcas de clase (32 y 37) en la parte superior de los rectángulos correspondientes. Haga lo mismo con los demás intervalos de clase (márquelos directamente sobre el histograma).

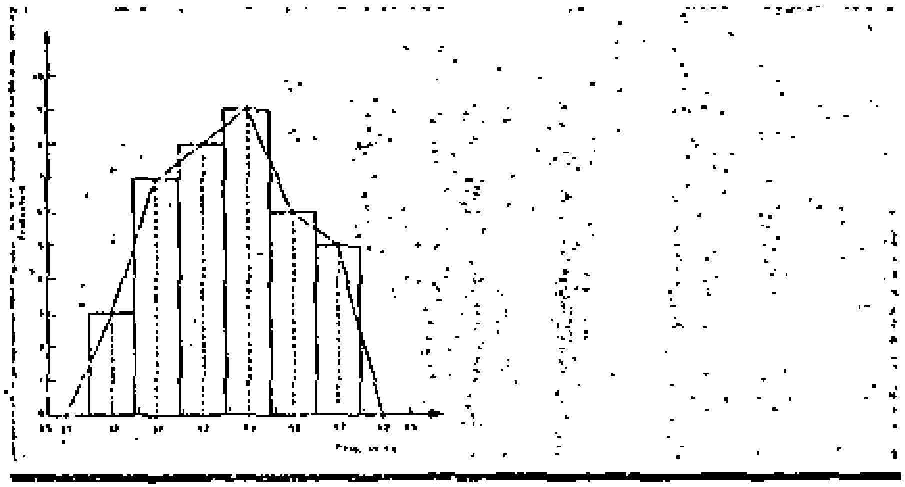




59 Vea la hoja de trabajo 4.5, en la que se muestra el histograma del cuadro anterior, con todas las marcas de clase señaladas en la parte superior de los rectángulos.

Observe que se han marcado dos puntos sobre el eje horizontal, en 27 y 62, los cuales corresponden a las marcas de clase de dos intervalos con *frecuencia nula*, cada uno de éstos con *igual amplitud* que el intervalo de clase adyacente. Observe, también, que los primeros tres puntos, correspondientes a las marcas de clase 27, 32, y 37, se unieron mediante líneas rectas.

Una los demás puntos directamente sobre el histograma.



60 Un polígono, como el dibujado en el cuadro anterior, que une todas las marcas de clase señaladas en la parte superior de los rectángulos, se conoce como *polígono de frecuencias*. Para dibujar los extremos del polígono de frecuencias, es necesario indicar dos marcas de clase adicionales, adyacentes una a la ______ (derecha/izquierda) ______ del primer intervalo y la otra a la ______ del último.

izquierda, derecha

61 Para dibujar los extremos de un polígono de frecuencias es necesario suponer que existe un intervalo de clase adyacente, a la izquierda del primer intervalo, con amplitud igual a la de éste y frecuencia de clase igual a cero. Además hay que señalar sobre el eje horizontal la ________ ___ ________ correspondiente.

marca de clase

62 Además del intervalo adicional, a la izquierda del histograma, es necesario suponer que existe otro intervalo en el extremo derecho, con frecuencia nula y amplitud igual a la del último intervalo de clase; también es necesario señalar su ________ ___ ________.

marca de clase

63 El polígono que une las marcas de clase señaladas en los extremos superiores de los rectángulos del histograma se conoce como ____________________ (polígono de frecuencias/histograma)

polígono de frecuencias

64 El polígono de __________ es el polígono formado al unir todas las marcas de clase, señaladas en las partes superiores de los rectángulos del __________.

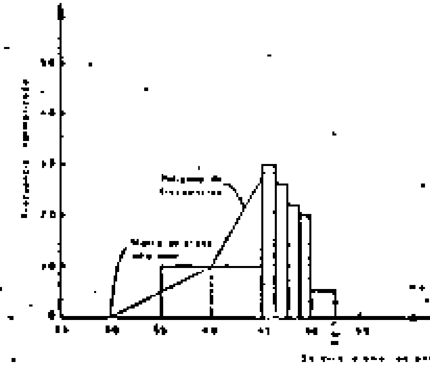
frecuencias, histograma

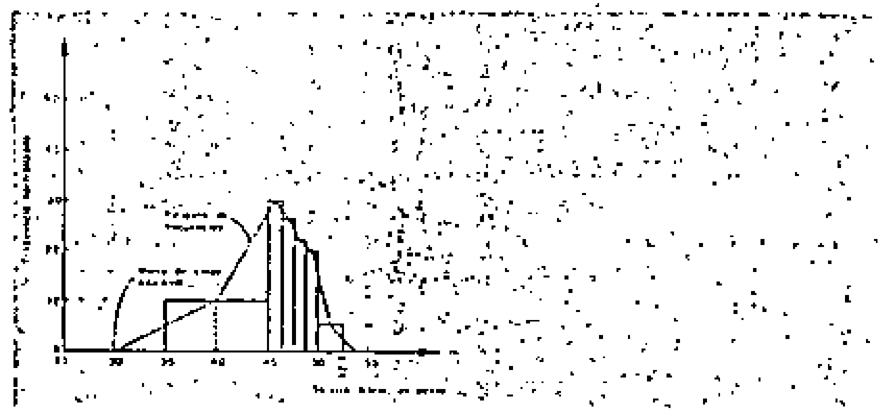
65 Al dibujar el ______________ de ______________ es necesario señalar dos marcas de clase adicionales, una en cada extremo del histograma.

polígono, frecuencias

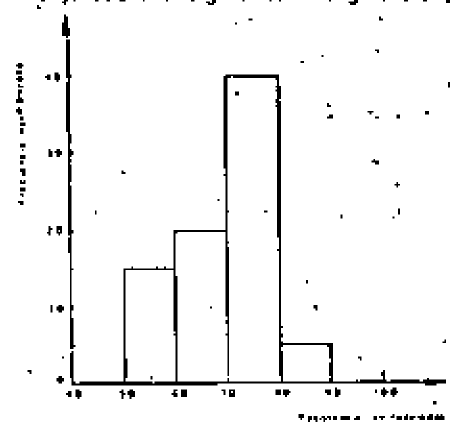
66 Observe en el siguiente histograma que la marca de clase adicional del extremo izquierdo se indicó suponiendo que corresponde a un intervalo de clase de igual ancho que el primero.

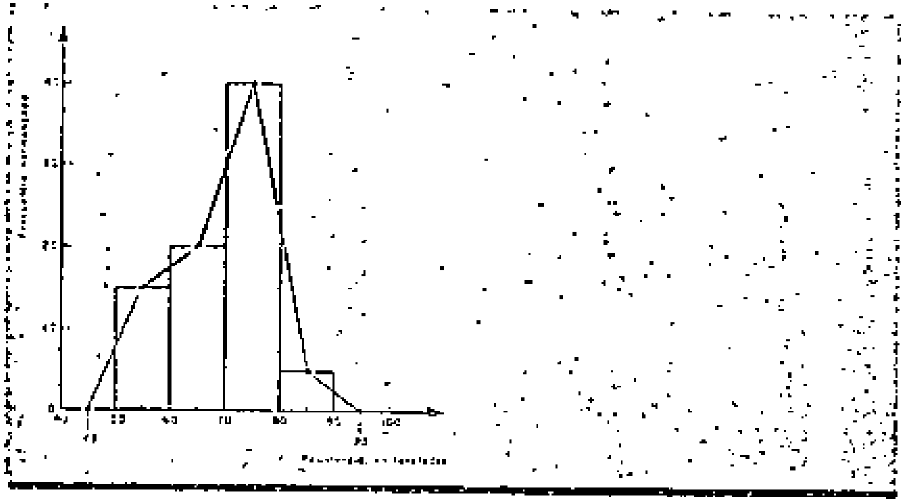
Señale, directamente sobre el histograma, la marca de clase a[illegible]al del extremo derecho, y complete el polígono de frecuencias.





67 Dibuje sobre el siguiente histograma el polígono de frecuencias correspondiente.





68 En el histograma del cuadro anterior, *todos* los intervalos de clase tienen el *mismo* ancho.

Demostraremos a continuación que, cuando esto sucede, la suma de las *áreas* de todos los rectángulos, es *igual al área* comprendida entre el polígono de frecuencias y el eje horizontal. Para esto debemos recordar que las áreas de dos triángulos semejantes ________ proporcionales.
(son/no son)

son

69 En la figura 4.1 se reproduce el histograma y el polígono de frecuencias del cuadro 67.

Observe que los triángulos *a* y *b* tienen igual área por ser triángulos semejantes y porque sus lados son iguales; por lo mismo, los triángulos *c* y *d* tienen igual área.

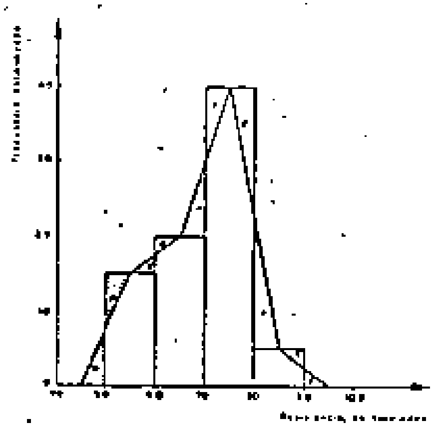


Fig. 4.1

1. El triángulo _____ tiene igual área que el *f*.
(cuál)
2. El triángulo *h* tiene igual área que el ____.
3. El triángulo ____ tiene igual área que el *i*.

1. *e*
2. *g*
3. *j*

70 Observe que en la figura 4.1, al trazar el primer lado del polígono de frecuencias se eliminó el área del triángulo *b*, pero se sustituyó con un *área* de *igual* magnitud correspondiente al triángulo *a*. Una cosa semejante sucede con los tra:os restantes del polígono de frecuencias; por ejemplo, al eliminar el triángulo *d*, éste es sustituido por el triángulo *c*, de igual área. Los triángulos *f*, *g*, e *i*, fueron sustituidos respectivamente por los triángulos ___, ___ y ___.

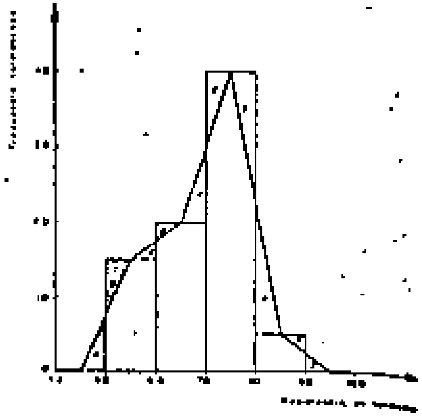


Fig. 4.1

e, *h*, *j*

71 Puesto que al trazar el polígono de frecuencias se eliminan algunas partes de los rectángulos del histograma, pero son sustituidas por otras de igual área, ¿se puede concluir que la suma de las áreas de todos los rectángulos es igual al área encerrada entre el polígono de frecuencias y el eje horizontal? ________.

Sí.

72 Puesto que la suma de todas las áreas de los rectángulos del histograma es igual al número total de elementos de la muestra, por ser el área de cada rectángulo igual a la frecuencia de clase correspondiente, el área encerrada entre el polígono de frecuencias y el eje horizontal ________ igual al número de elementos de la muestra.
(es/no es)

es

73 El polígono de frecuencias es útil porque presenta una forma más *suave* de la distribución de frecuencias que la que proporciona el histograma correspondiente, es decir, ______ presenta los saltos bruscos, en forma de escalón, que presenta el histograma.
(sí/no)

no

74 Por lo vi... en los cuadros anteriores, se concluye que el polígono de frecuencias constituye otra forma de *presentación gráfica* de los datos. ¿Cuáles son éstas?

El histograma y el polígono de frecuencias

---

75 ¿Se obtiene, al observar un histograma o un polígono de frecuencias, una idea inmediata de los intervalos en que los datos aparecieron con mayor o menor frecuencia?

Sí.

---

76 ¿Al observar un histograma o un polígono de frecuencias se obtiene una idea inmediata de la tendencia de los datos a aglomerarse en la zona central, en la de valores pequeños, o en la de valores grandes de la variable?

Sí.

---

77 Entonces, al observar un ________ o un ____________________ se obtiene una idea inmediata de los intervalos en que los datos aparecieron con mayor frecuencia, y de la tendencia a aglomerarse en la zona central, en la de valores pequeños, o en la de valores grandes de la variable.

histograma, polígono de frecuencias

---

78 Una representación gráfica del tipo de barras de una distribución de frecuencias se llama ____________________.
(histograma/polígono de frecuencias)

histograma

---

79 Una representación gráfica de una distribución de frecuencias, que se forma uniendo todos los puntos definidos por las marcas de clase y las frecuencias de los intervalos correspondientes, se llama ____________________
(histograma/polígono de frecuencias)

polígono de frecuencias

---

80 Una distribución de frecuencias se presenta gráficamente mediante un ____________ o mediante un ________ ______ ____________

histograma, polígono de frecuencias

---

81 Las barras de un histograma quedan delimitadas por los límites ________

reales

---

82 Un ____________ ____ ____________ se traza uniendo los puntos definidos por las marcas de clase y las frecuencias de los intervalos correspondientes.

polígono de frecuencias

---

83 Al trazar un histograma, en el eje horizontal se anotan los valores de la ____________.

variable

---

84 Al trazar un histograma para datos agrupados en intervalos de igual amplitud, es común anotar en el eje vertical las ________________ ____ ____________

frecuencias de clase (o frecuencias)

---

85 Si los intervalos de clase no son de igual amplitud, es necesario anotar en el eje vertical las ____________ ____ ____________ ____________

frecuencias de clase normalizadas (o frecuencias normalizadas)

---

86 Cuando en el eje vertical se anotan las frecuencias normalizadas, la frecuencia de clase está dada por el ________ encerrada en el rectángulo correspondiente.

área

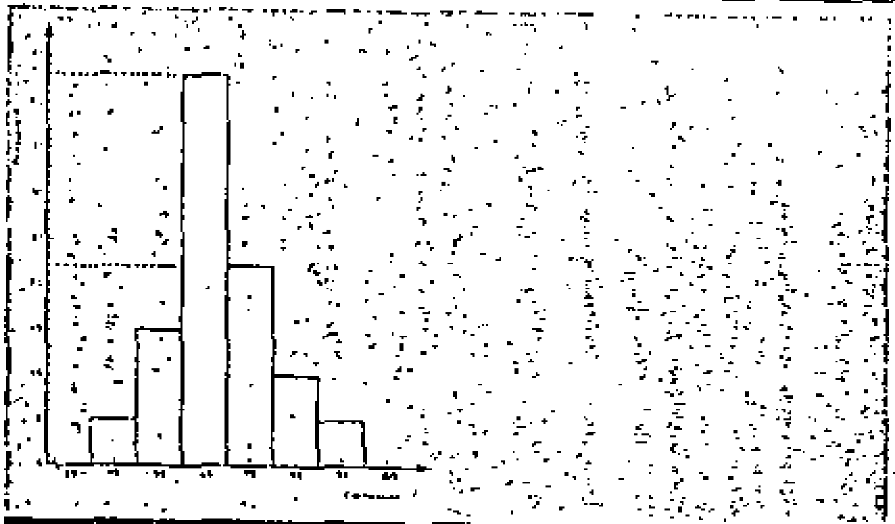
---

87 Al trazar un polígono de frecuencias, es necesario considerar un intervalo de clase con frecuencia cero en cada uno de los extremos; el ancho de cada uno de estos intervalos es igual al ancho del ________ adyacente.

intervalo

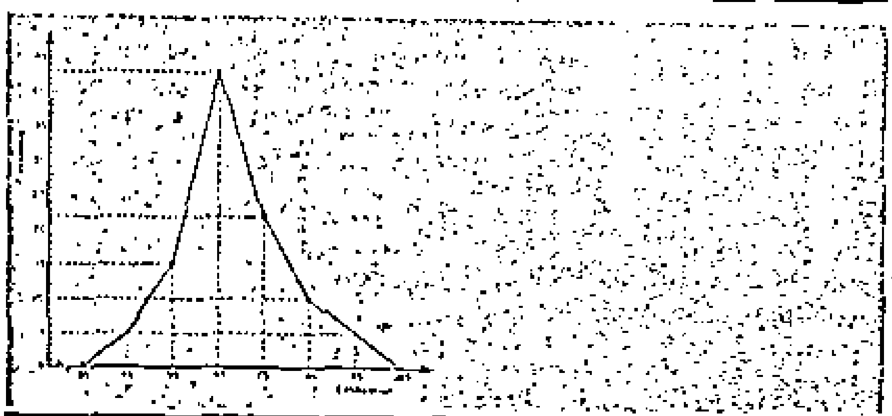
88 Use la [illegible] de trabajo 4.6 para trazar el histograma correspondiente a los datos presentados en la siguiente tabla.

Marca de clase	Frecuencia
45	5
55	15
65	43
75	22
85	10
95	5



89 Use la hoja de trabajo 4.7 para trazar el polígono de frecuencias correspondiente a la siguiente distribución de frecuencias.

Calificación en Electricidad (marca de clase)	Frecuencia
45	5
55	15
65	43
75	22
85	10
95	5



90 Existe otra forma gráfica de presentar una distribución de frecuencias. En el eje vertical de ésta se lee la *frecuencia acumulada*, por lo cual la gráfica recibe el nombre de *curva de distribución de frecuencias acumuladas*. En una curva de distribución de frecuencias acumuladas, en cada punto del eje vertical se indica la ______________, y en el eje horizontal el valor de la variable.
(frecuencia/frecuencia acumulada)

frecuencia acumulada

91 Cada punto del eje vertical de una curva de distribución de frecuencias acumuladas indica la frecuencia acumulada correspondiente al valor de la variable, el cual se lee en el eje horizontal; ¿es esto equivalente a decir que indica la frecuencia de observaciones de la variable *menores* o *iguales* que el valor considerado? ____

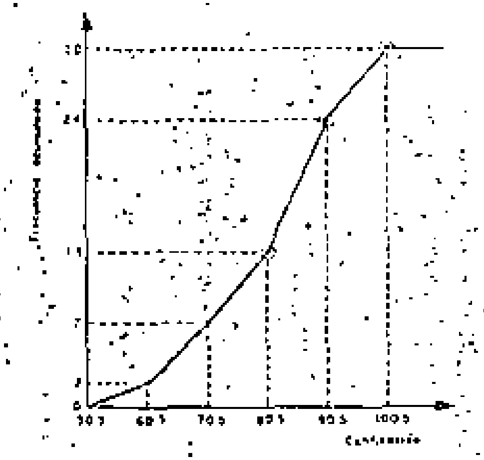
Sí.

92 La tabla 4.4 presenta una distribución de frecuencias acumuladas. La gráfica se dibuja anotando, para cada límite real superior, la frecuencia acumulada del intervalo correspondiente. Tome la hoja de trabajo 4.8 y observe que la frecuencia acumulada hasta el primer límite real inferior (50.5) es cero; esto queda indicado en la gráfica con el punto *A*. La frecuencia acumulada hasta el primer límite real superior (60.5) es 2 (punto *B*). La frecuencia acumulada hasta 70.5 es ________ (punto *C*). Indique en la misma gráfica los puntos de la curva de frecuencias acumuladas, correspondientes a los límites 80.5, 90.5 y 100.5; complete la curva.

TABLA 4.4

Intervalo de calificaciones	Frecuencia acumulada	Límites reales Inferior	Límites reales Superior
51–60	2	50.5	60.5
61–70	7	60.5	70.5
71–80	13	70.5	80.5
81–90	24	80.5	90.5
91–100	30	90.5	100.5

7.



93 Observe ... curva de distribución de frecuencias acumuladas que se presenta en la figura 4.2.

La frecuencia de valores menores que: 70.5 es 7.

La de 85 es ______.
(30/18/15)

La de 100.5 es ______.
(30/18/15)

La de 50.5 es ______

La de ______ es 18.

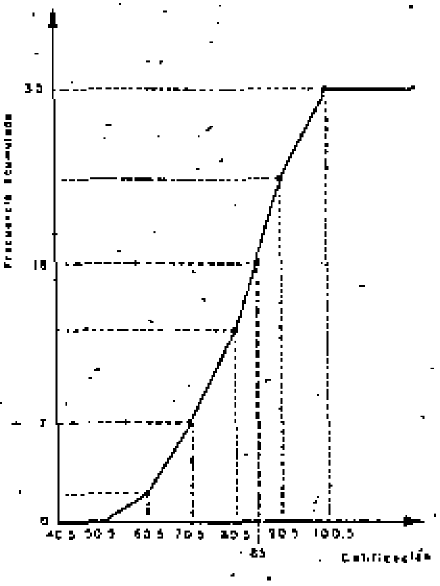


Fig 4.2

18. 30. 0. 85

94 Recuerde que la máxima frecuencia acumulada es igual al número total de datos y corresponde al último límite *real* superior. Esto implica que cualquier valor de la variable *mayor* que el último límite real superior también tendrá la máxima frecuencia acumulada.

En la figura 4.2 la frecuencia acumulada para 100.5 es 30 (número total de datos). Si los valores de los 30 datos son menores que 100.5, también serán menores que cualquier número ______ que 100.5.
(menor/mayor)

mayor

95 Recuerde que el mínimo valor de la frecuencia acumulada es cero, y corresponde al primer límite real inferior. Esto implica que para cualquier valor *menor* que el primer límite real inferior la frecuencia acumulada es nula.

Por ejemplo, en la figura 4.2 la frecuencia de los valores menores que 50.5 es ______

cero

96 Observe la figura 4.2.

La frecuencia acumulada de los valores mayores que el último límite real superior es igual al número total de datos; esto se representa en ... áfica de frecuencias acumuladas mediante una línea horizontal, trazada hacia la ______ del último límite real superior.
(derecha/izquierda)

derecha.

97 Observe la figura 4.2.

La frecuencia acumulada de los valores *menores* que el primer límite real inferior es igual a cero; esto se representa en la gráfica de frecuencias acumuladas mediante una línea horizontal trazada hacia la ______ del primer límite real inferior.
(derecha/izquierda)

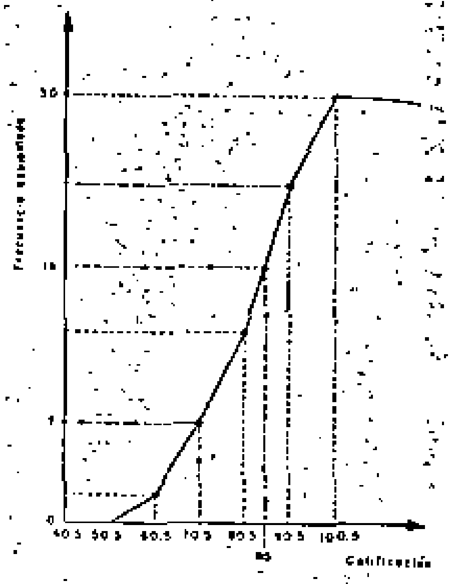
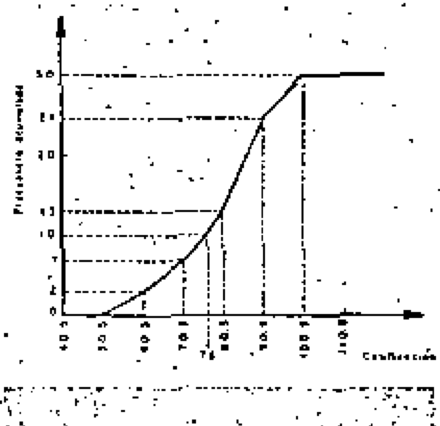


Fig 4.2 REPETIDA

izquierda

98 Observe que la siguiente curva de distribución de frecuencias acumuladas presenta las líneas horizontales mencionadas en los dos cuadros anteriores. Cuál es la frecuencia de los valores:

a. menores o iguales que 106 ______
b. menores o iguales que 43 ______
c. menores o iguales que 76 ______

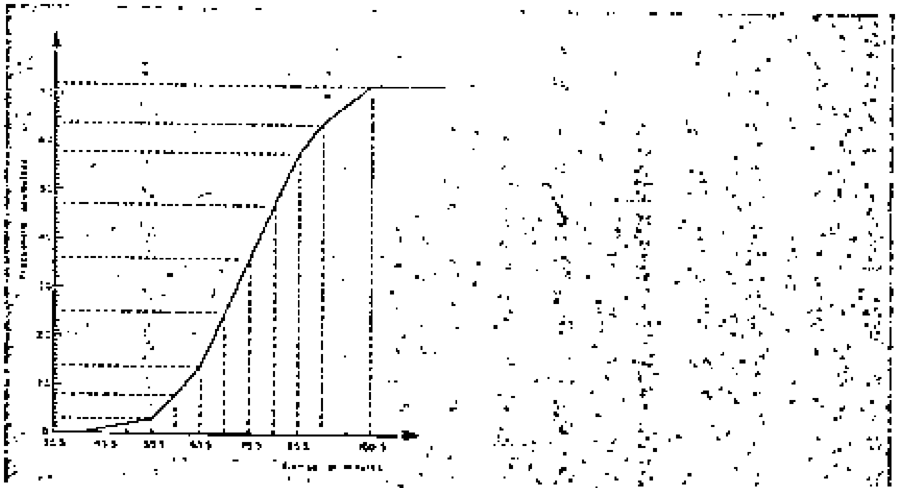


30. 0. 10.

99 Tome [illegible]oja de trabajo 4.9 y dibuje en ella la curva de distribución de frecuencias acumuladas correspondiente a la tabla 4.5.
No olvide dibujar las líneas horizontales a la izquierda y a la derecha del primero y del último límites reales.

TABLA 4.5

Intervalo de clase, en minutos	Límites reales		Frecuencia acumulada	Frecuencia acumulada complementaria
	Inferior	Superior		
51–55	50.5	55.5	3	69
56–60	55.5	60.5	8	64
61–65	60.5	65.5	14	58
66–70	65.5	70.5	25	47
71–75	70.5	75.5	36	36
76–80	75.5	80.5	47	25
81–85	80.5	85.5	58	14
86–90	85.5	90.5	64	8
91–100	90.5	100.5	72	0



100 La curva mediante la cual se obtienen las frecuencias acumuladas, se llama curva de ______________ ______________ acumuladas.

distribución de frecuencias

101 La gráfica del tipo de barras mediante la cual se leen las frecuencias de clase se llama ______________.

histograma

102 La gráfica formada mediante la unión de los puntos definidos por las marcas de clase y las frecuencias de clase correspondientes se llam. ______________ ______________

polígono de frecuencias

103 Si se tiene un total de 30 datos y 13 de ellos tienen valores *menores* que 80.5, el resto (30−13=17) tiene valores *mayores* que 80.5. Si 24 de los datos tienen valores menores que 90.5, ________ datos tendrán valores mayores que 90.5.
(cuántos)

6; (30 − 24 = 6)

104 Entonces, la diferencia entre el número de datos y una frecuencia acumulada cualquiera, nos da la frecuencia de valores ______________ que el valor correspondiente de la variable.
(menores/mayores)

mayores

105 A la frecuencia de valores *mayores* que un cierto número la denominaremos *frecuencia acumulada complementaria.*
La frecuencia acumulada complementaria es la ____________ de valores mayores que un cierto número.

frecuencia

106 La frecuencia acumulada complementaria se calcula restándole al número de datos la ______________.
(frecuencia/frecuencia acumulada)

frecuencia acumulada

107 Si se tiene un total de 50 datos y la frecuencia acumulada correspondiente a un cierto valor es 33, la frecuencia acumulada ____________ es igual a 50 − 33 = 17.

complementaria

108 Complete la tercera columna de la siguiente tabla. (Recuerde que la última frecuencia acumulada es igual al total de observaciones.)

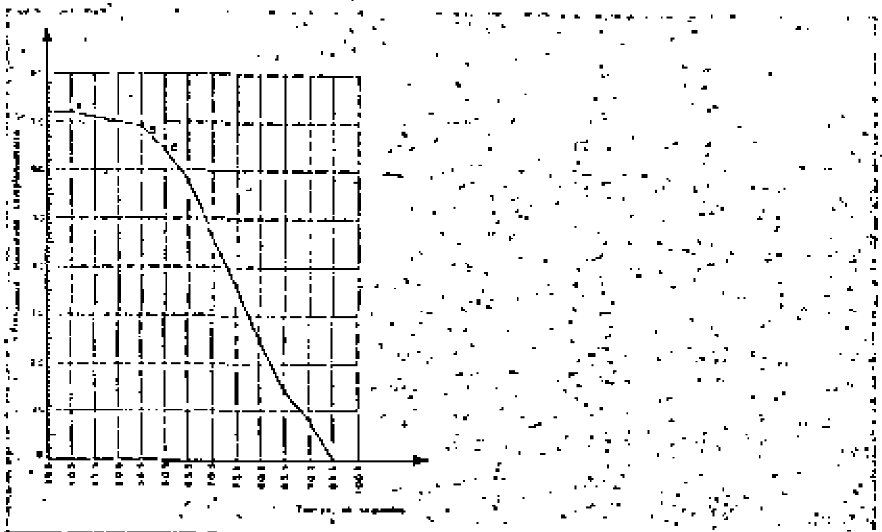
Intervalo de clase	Frecuencia acumulada	Frecuencia acumulada complementaria
41-55	3	69 (72-3)
56-60	8	64 (72-8)
61-65	14	
66-70	25	
71-75	36	
76-80	47	
81-85	58	
86-90	64	
91-100	72	

58, 47, 36, 25, 14, 8, 0

109 Tome la hoja de trabajo 4.10 y complete la curva de frecuencias acumuladas complementarias correspondiente a los datos presentados en la tabla 4.5. Observe que se anotó una frecuencia acumulada complementaria de 72 para un tiempo de 40.5, puesto que todos los datos fueron mayores que 40.5 (punto *A*). El punto *B* corresponde al límite real superior de 55.5 el cual tiene una frecuencia acumulada complementaria de 69. Al siguiente límite real superior (60.5), corresponde una frecuencia acumulada complementaria de 61 (punto *C*).

TABLA 4.5 [illegible]

Intervalo de clase, en minutos	Límites reales		Frecuencia acumulada	Frecuencia acumulada complementaria
	Inferior	Superior		
41-55	40.5	55.5	3	69
56-60	55.5	60.5	8	64
61-65	60.5	65.5	14	58
66-70	65.5	70.5	25	47
71-75	70.5	75.5	36	36
76-80	75.5	80.5	47	25
81-85	80.5	85.5	58	14
86-90	85.5	90.5	64	8
91-100	90.5	100.5	72	0



10 Puesto que todos los datos son ____________ (menores/mayores) que el [illegible]r límite real inferior, al dibujar la curva de frecuencias acumuladas complementarias, se traza una línea horizontal hacia la izquierda de este punto.

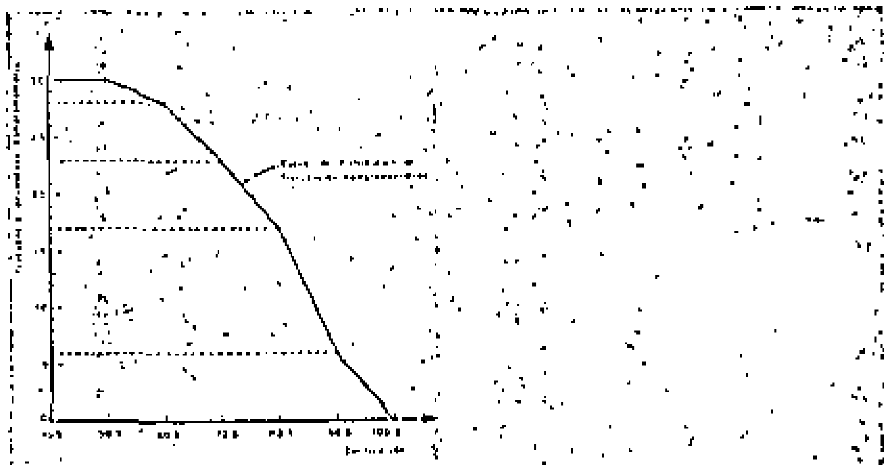
mayores

111 Puesto que todos los datos son ____________ (menores/mayores) que el último límite real superior, al dibujar la curva de frecuencias acumuladas complementarias, se traza una línea horizontal hacia la derecha de este punto.

menores

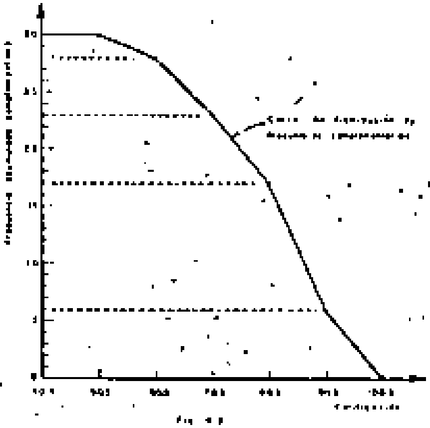
112 Dibuje, en la hoja de trabajo 4.11, la curva de frecuencias acumuladas complementarias correspondiente a los datos de la siguiente tabla (no olvide las líneas horizontales en los extremos). El total de datos es 30.

Intervalo de calificaciones	Límites reales		Frecuencia acumulada complementaria
	Inferior	Superior	
51-60	50.5	60.5	28
61-70	60.5	70.5	23
71-80	70.5	80.5	17
81-90	80.5	90.5	6
91-100	90.5	100.5	0



113 Observe la figura 4.3 en la cual se muestra la respuesta al cuadro anterior.

¿Cuál es la frecuencia de valores mayores que: 103, 43, 55.5 y 90.5?



0, 30, 29, 5.

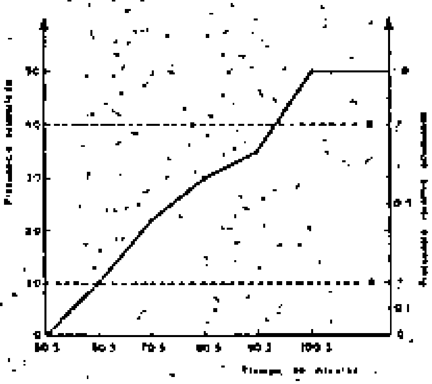
114 En ocasiones es conveniente indicar en la gráfica de frecuencias acumuladas una escala en la cual se lean las *frecuencias relativas acumuladas*. Recuerde que la frecuencia relativa acumulada varía de cero a ______.

uno

115 La frecuencia relativa acumulada igual a cero corresponde al cero de la frecuencia acumulada, y la frecuencia relativa igual a _____ al número total de observaciones.

uno

116 Anote los valores correspondientes a los puntos A y B, que se indican en la escala (de la derecha) de frecuencias relativas acumuladas de la siguiente figura. (Recuerde que la frecuencia relativa acumulada se calcula dividiendo la frecuencia acumulada entre el número total de datos.)



A: 0.2, B: 0.8

117 ¿Cuáles son los cuatro tipos de representaciones gráficas de datos, que hemos estudiado?

1. Histograma, 2. Polígono de frecuencias, 3. Curva de distribución de frecuencias acumuladas, 4. Curva de distribución de frecuencias acumuladas complementarias.

118 La curva que une los puntos definidos por las marcas de clase y las frecuencias de los intervalos correspondientes se llama ______ ______ ______.

polígono de frecuencias

119 Una curva que define las frecuencias de valores menores que un cierto valor, se denomina ______ (a/b/c)

a. curva de frecuencias acumuladas
b. curva de frecuencias acumuladas complementarias
c. histograma

a. curva de frecuencias acumuladas

120 Una gráfica del tipo de barras en la cual se leen las frecuencias correspondientes a cada intervalo de clase se denomina ______________.

histograma

121 Hasta ahora sólo hemos presentado gráficamente los datos de variables aleatorias continuas, las cuales son siempre escalares. Estas presentaciones gráficas se pueden realizar para algunas variables aleatorias *discretas*; la única condición requerida, es que dichas variables sean ____________ (nominales/escalares). Si la variable es nominal (no toma valores numéricos), ______ (sí/no) es posible trazar una curva de distribución de frecuencias acumuladas.

escalares, no

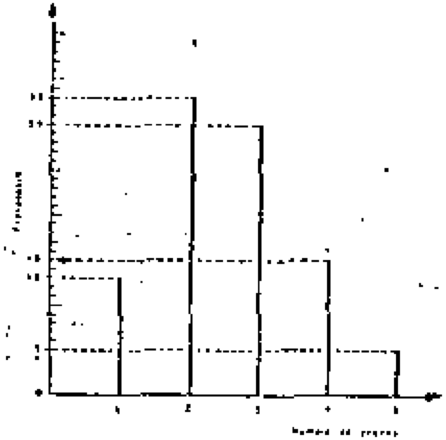
122 El histograma correspondiente a los datos de una variable aleatoria discreta se forma a base de barras angostas asociadas a cada __________ (valor/altura) de la variable, y no a base de rectángulos como acontece con las variables ____________.

valor, continuas

123 Se lanzaron 100 veces cinco monedas simultáneamente; en cada lanzamiento se anotó el número de cruces que quedaron hacia arriba, obteniéndose la distribución de frecuencias presentada a continuación.

Complete el histograma correspondiente que se presenta en la hoja de trabajo 4.12.

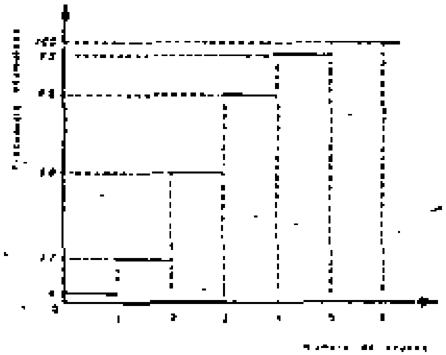
Número de cruces	Frecuencia
0	4
1	15
2	33
3	30
4	15
5	3
TOTAL	100



124 A continuación se presenta la tabla de distribución de frecuencias acumuladas correspondiente a los datos del cuadro anterior.

Complete la gráfica de distribución de frecuencias acumuladas que se presenta en la hoja de trabajo 4.13.

Número de cruces	Frecuencia acumulada
menos de 0	0
menos de 1	4
menos de 2	19
menos de 3	52
menos de 4	[illegible]
menos de 5	95
menos de 6	100



125 El histograma, el polígono de frecuencias y los dos tipos de curvas de distribución de frecuencias acumuladas constituyen diferentes formas de presentación ______________ de las distribuciones de frecuencias.
(gráfica/tabular)

gráfica

PARTE C. PERCENTILES, DECILES, CUARTILES Y MEDIANA

126 Observe, en la figura 4.4, la curva de distribución de frecuencias relativas acumuladas correspondiente a la resistencia de un cierto tipo de cables de acero.

Note que en la escala vertical, dividida en 100 partes iguales, se han marcado algunos puntos divisorios que se identifican con una letra P. Así, P_1 representa el primer punto divisorio; P_5 representa el ______________, etcétera.
(quinto/décimo)

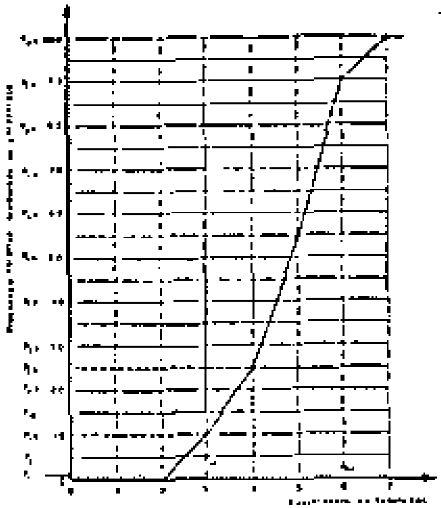


Fig. 4.4

quinto

127 Observe la figura 4.4.

P_{20} representa el 20º (vigésimo) punto de los 100 en que se ha dividido el eje vertical; P_{25} representa el ______ punto; P_{50} representa el ________ punto.

25º, 50º

128 Observe la figura 4.4.

P_{60} representa el 60º punto de los 100 en que se ha dividido el eje vertical; ________ representa el 75º punto. El 90º punto se representa con ________.

P_{75}, P_{90}

129 Hemos usado la letra P para identificar puntos que marcan la división de una distribución de frecuencias acumuladas (o de frecuencias relativas acumuladas) en __________ partes iguales.
(10/50/100)

100

---

130 Cada uno de los puntos que dividen una distribución de frecuencias relativas acumuladas en 100 partes iguales se relaciona, mediante la curva de distribución de frecuencias acumuladas, con un número del eje horizontal.
Por ejemplo, en la figura 4.4, a P_{10} le corresponde el valor 3.0 de la variable.
Observe que a P_{33} le corresponde el valor ______ de la variable.

4.0

---

131 Observe la figura 4.4.
A P_{50} le corresponde el valor de la variable __________.
(4.35/6.28/4.86)
¿Qué valor le corresponde a P_{94}?

4.86. 6.40.

---

132 P_{11} corresponde a una frecuencia relativa acumulada de 11%. P_{19} corresponde a una frecuencia relativa acumulada de _____ %.

19

---

133 El valor de la variable que corresponde a P_{94} lo representaremos con $\overline{P}_{94}$. El que corresponde a P_{41} lo representaremos con __________.
($\overline{P}_{51}$/$\overline{P}_{49}$/$\overline{P}_{41}$)
¿Con qué símbolo denotamos el valor de la variable que corresponde a P_{10}?

$\overline{P}_{41}$. $\overline{P}_{10}$.

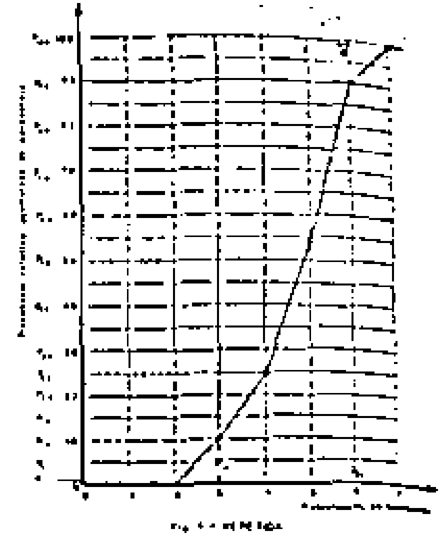
---

134 Observe la figura 4.4.
¿Cuánto vale $\overline{P}_{10}$?

3.0.

---

135 Observe la figura 4.4.
$\overline{P}_{37}$ es el valor de la variable correspondiente a una frecuencia relativa acumulada de _________ %.
¿Cuánto vale $\overline{P}_{37}$?



37. 4.4 (aproximadamente).

---

136 ¿Qué representa P_{94}? _____.
(a/b/c)

a. 94 datos
b. Una frecuencia relativa acumulada de 94%
c. Una frecuencia de 94

b. Una frecuencia relativa acumulada de 94%

---

137 Puesto que P_{11} representa una frecuencia relativa acumulada de 11%, el valor correspondiente de la variable, $\overline{P}_{11}$, es tal que el 11% de los valores son menores que $\overline{P}_{11}$.
De la misma manera $\overline{P}_{94}$ es tal que el ____ % de los valores son menores que él.

94

---

138 El 75% de los datos son menores que __________.
(P_{75}/$\overline{P}_{75}$/75)

$\overline{P}_{75}$

---

139 El valor de la variable cuya frecuencia relativa acumulada es de 13% se representa con ________.

$\overline{P}_{13}$

140 De lo visto anteriormente, se observa que los valores de $\overline{P}$ son los *valores de la variable* correspondiente a las frecuencias P, que dividen la distribución de frecuencias relativas acumuladas en ________ partes iguales.
(cuántas)

100

141 Los valores de $\overline{P}$ son aquellos de la variable que corresponden a una división de la distribución de frecuencias relativas ________________ en cien partes iguales.

acumuladas

142 Debido a que $\overline{P}$ representa los *valores de la variable* que dividen la distribución de frecuencias relativas acumuladas en 100 *partes iguales*, se le denomina *percentil*. Así $\overline{P}_1$ es el primer percentil; $\overline{P}_3$ es el ____________ percentil; $\overline{P}_{20}$ es el
(quinto/tercer)
________________________.

tercer, vigésimo (20º) percentil

143 $\overline{P}_{39}$ es el 39º ________________.
¿Con qué símbolo se denota el 56º percentil?

percentil, $\overline{P}_{56}$

144 El 49º percentil se representa con ______.

$\overline{P}_{49}$

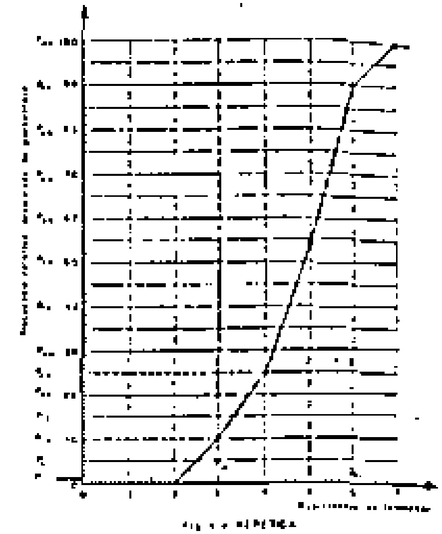
145 Los percentiles son los valores de la variable que dividen la distribución de ________________ relativas acumuladas en ________ partes iguales.
(cuántas)

frecuencias, 100

146 Los *valores de la variable* que corresponden a una división de la distribución de frecuencias relativas acumuladas en 100 partes iguales se conocen como ________________.

percentiles

147 Observe la figura 4.4.
¿Cuánto vale el 49º percentil?



4.8 (aproximadamente).

148 Observe la figura 4.4.

a. ¿Cuánto vale el sexto percentil?
b. El valor 6.5 de la variable corresponde al __________ percentil.

2.6, 95º ($\overline{P}_{95}$)

149 Para que *todos* los percentiles queden definidos en forma precisa es necesario que la variable aleatoria bajo estudio sea continua. Cuando la variable es discreta, es necesario definir los percentiles en forma aproximada, como se verá más adelante.

No hay respuesta.

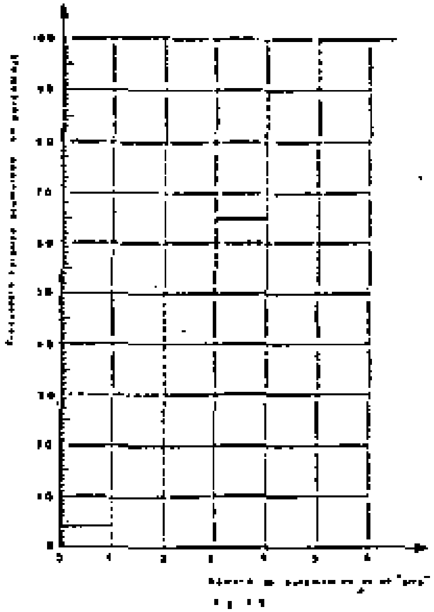
150 Para hablar en forma precisa de percentiles, se requiere que la variable en cues-

tión sea continua porque, en general, los percentiles asumen valores ______________ (enteros/fraccionarios)

fraccionarios

---

151 Es necesario que la variable sea continua para poder calcular percentiles en forma precisa. Vea la figura 4.5. En ella se observa que el 4% de los datos fueron menores que uno; que el 10% fueron menores que __________; que el __________% fueron menores que tres, etcétera.



2, 50

---

152 Observe la figura 4.5.
Al tratar de encontrar el valor de $\overline{P}_{30}$, por ejemplo, lo único que podemos decir es que está comprendido entre 1 y 2. Análogamente el valor de $\overline{P}_{80}$ está comprendido entre __________ y __________.

3, 4

---

153 Observe en la gráfica de frecuencias acumuladas de la figura 4.5 que el quincuagésimo percentil sí queda definido; note que vale 3.
¿Cuánto vale:

a. $\overline{P}_{65}$
b. $\overline{P}_{40}$?

4, entre 1 y 2

---

154 De los cuadros anteriores se concluye que *no todos* los perc es pueden definirse exactamente cuando se trata de variables ______________ (discretas/continuas), es decir, no a todas las frecuencias relativas acumuladas les corresponde uno de los valores que puede asumir la variable aleatoria discreta bajo estudio.

discretas

---

155 En forma semejante a la definición de un percentil, podemos definir lo que se entiende por un *decil*. Los deciles son los *valores de la variable* que dividen una distribución de frecuencias acumuladas, en *diez partes* ______________ (iguales/diferentes).

iguales

---

156 Los deciles son valores de la variable que dividen una distribución de frecuencias ______________ en ________ (cuántas) partes iguales.

acumuladas, diez

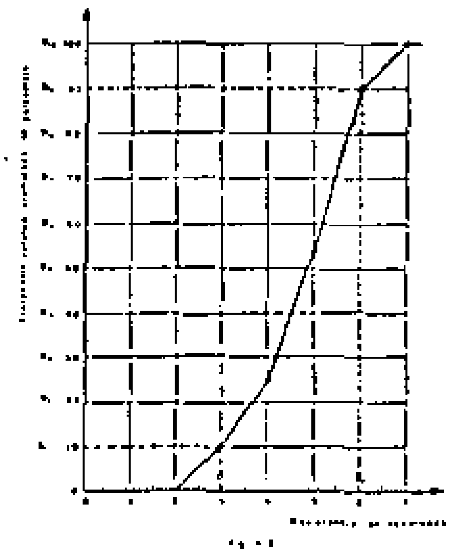
---

157 Los valores de la variable que dividen una distribución de frecuencias acumuladas en diez partes iguales, se llaman __________.

deciles

---

158 Observe la figura 4.6.
La distribución de frecuencias relativas acumuladas se ha dividido en diez partes iguales; cada punto divisorio se ha marcado con una letra D. Así D_1 representa una frecuencia relativa acumulada de 10%. D_3 representa una de %, etcétera.



159 Una frecuencia relativa acumulada de 100%, en deciles, se representa con ______
$(D_1/D_{10}/D_{100})$

D_{10}

160 Mediante la curva de distribución de frecuencias relativas acumuladas, a cada D se le asocia un valor de la variable. Así, a D_1 se le asocia $\overline{D}_1$, a D_5 corresponde ______ y a D_8 ______.
$(\overline{P}_5/\overline{D}_5)$

$\overline{D}_5$, $\overline{D}_8$

161 Puesto que $\overline{D}$ denota los valores de la variable que dividen una distribución de ______ ______ en diez partes iguales, ésta representa los ______.
(deciles/percentiles)

frecuencias acumuladas (o frecuencias relativas acumuladas), deciles

162 $\overline{D}_1$ se conoce como primer decil $\overline{D}_2$ como segundo decil, etcétera. $\overline{D}_{10}$ se conoce, entonces, como ______ ______.

décimo decil

163 Observe la figura 4.6.
El primer decil, $\overline{D}_1$, es el valor de la variable asociado a D_1. En la figura 4.6 se observa que $\overline{D}_1$ vale 3.0. ¿Qué valor tiene $\overline{D}_9$?

6.0.

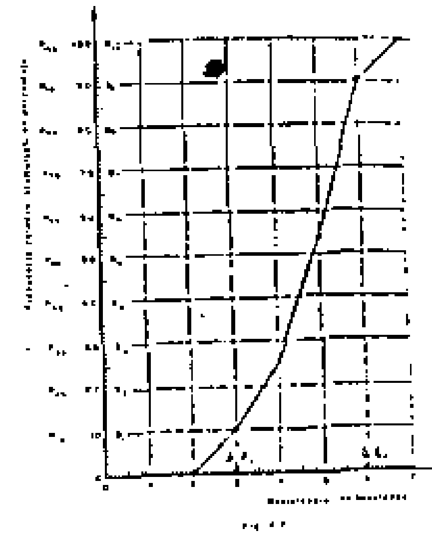
164 Observe la figura 4.6.
a. El quinto decil, $\overline{D}_5$, vale ______
(4.86/4.30/9.10)
b. ¿Cuánto vale $\overline{D}_8$?

4.86. 5.7 (aproximadamente).

165 Observe la figura 4.6.
El valor 5.2, de la variable, corresponde al ______ decil.
(cuál)

sexto $(\overline{D}_6)$

166 Observe la figura 4.7.
Existen diez percentiles que coinciden con los diez deciles: $\overline{P}_{10}$ coincide con $\overline{D}_1$; $\overline{P}_{20}$ coincide con $\overline{D}_2$ y así sucesivamente. ¿Con qué decil coincide $\overline{P}_{100}$?



Con $\overline{D}_{10}$.

167 ¿Con qué percentil coincide $\overline{D}_5$?

Con el 50º $(\overline{P}_{50})$.

168 Los valores de la variable que dividen una distribución de frecuencias acumuladas en cuatro partes iguales se conocen como *cuartiles*. En una distribución de frecuencias acumuladas sólo puede haber ______ cuartiles.
(4/10/100)

cuatro

169 Los cuartiles son los valores de la ______________ que dividen una distribución de frecuencias acumuladas en ______________ partes iguales.

variable, cuatro

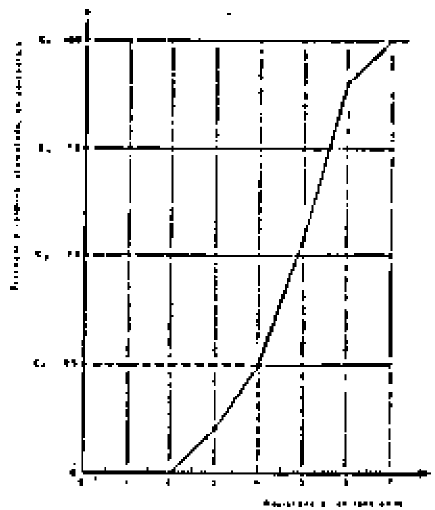
---

170 Los valores de la variable que dividen una distribución de frecuencias acumuladas en cuatro partes iguales se llaman ______________ (deciles/cuartiles/percentiles)

cuartiles

---

171 Observe la figura 4.8.
En el eje vertical se han señalado con la letra C cuatro valores de la frecuencia relativa acumulada. *Así*, C_1 representa una frecuencia acumulada de 25%. ¿Qué frecuencia relativa acumulada corresponde a C_3?



75%.

---

172 ¿Qué frecuencia relativa acumulada corresponde a C_4?

100%.

---

173 Mediante la curva de distribución de frecuencias relativas acumuladas a cada C se asocia un valor de la variable. A C_1 se asocia $\bar{C}_1$; a C_2 le corresponde $\bar{C}_2$; a C_4, ______

$\bar{C}_4$

---

174 Puesto que la letra $\bar{C}$ representa los *valores de la variable* dividen una distribución de frecuencias en cuatro partes ______________ (iguales/diferentes), ésta representa los ______________ (cuartiles/deciles/percentiles).

iguales, cuartiles

---

175 A $\bar{C}_1$ se le conoce como primer cuartil, a $\bar{C}_2$ como segundo ________, etcétera.

cuartil

---

176 Los cuatro cuartiles se representan con los símbolos _____, _____, _____ y _____

$\bar{C}_1$, $\bar{C}_2$, $\bar{C}_3$ y $\bar{C}_4$

---

177 ¿Qué representa $\bar{C}_3$?

El tercer cuartil.

---

178 Observe la figura 4.8.
El primer cuartil, $\bar{C}_1$, correspondiente a una frecuencia relativa acumulada de 25%, tiene un valor de 4.0. ¿Cuánto vale $\bar{C}_3$?

5.6 (aproximadamente).

---

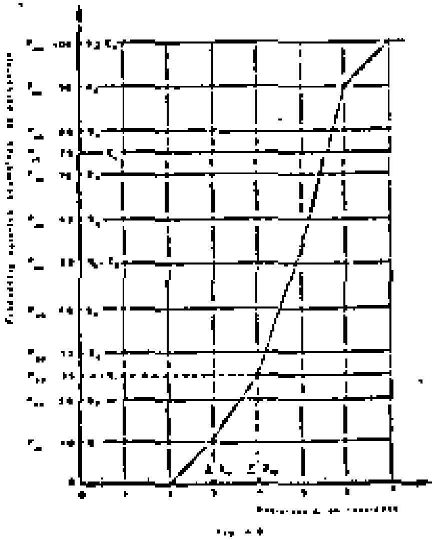
179 Relacione los nombres de la columna de la izquierda con las descripciones de la derecha.

1.	Deciles	a.	Dividen la distribución de frecuencias acumuladas en cuatro partes iguales
2.	Cuartiles	b.	La dividen en 30 partes iguales
3.	Percentiles	c.	La dividen en diez partes iguales
		d.	La dividen en 100 partes iguales

1—c, 2—a, 3—d

---

180 Observe la figura 4.9.
$\overline{C}_1$ coincide con el percentil $\overline{P}_{25}$;
_____ coincide con $\overline{P}_{50}$ y con $\overline{D}_5$
(cuál)



$\overline{C}_2$

181 Observe la figura 4.9.
$\overline{C}_4$ coincide con ______ y _____.

$\overline{D}_{10}$, $\overline{P}_{100}$

182 El valor de la variable que divide una distribución de frecuencias acumuladas en *dos partes iguales* se denomina *mediana*.
La mediana es el valor de la variable que divide una distribución de frecuencias acumuladas en _____ partes ________________.

2, iguales

183 La ______________ es el *valor de la variable* que divide una distribución de
(mediana/medial)
frecuencias acumuladas, en dos partes iguales.

mediana

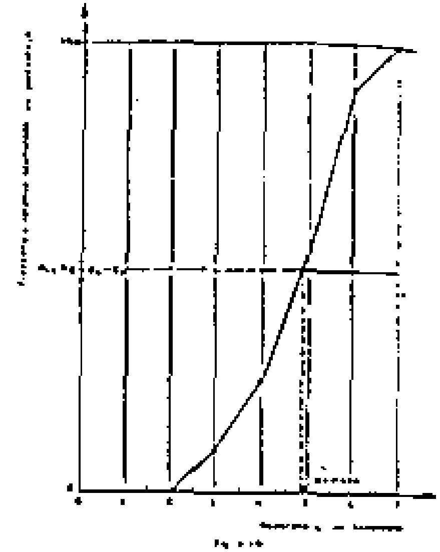
184 El valor de la ______________ que divide una distribución de frecuencias acumuladas en dos partes iguales, se denomina ______________.

variable, mediana

185 El ____% de los valores de la variable son menores que la me[illegible].

50

186 Observe la figura 4.10.
El valor de la variable correspondiente a una frecuencia acumulada de 50% es
____
(4.50/4.86/5.91)



4.86 (aproximadamente)

187 En los datos de la figura 4.10, el 50% de los valores son menores que 4.86, entonces, la mediana vale _____.

4.86

188 ¿Qué porcentaje de los valores de la variable son menores que la mediana?

50%

189 ¿Qué porcentaje de los valores de la variable son mayores que la mediana?

50%

190 Relacione los nombres de la columna de la izquierda con las descripciones de la derecha.

1. Mediana
2. Deciles
3. Percentiles
4. Cuartiles

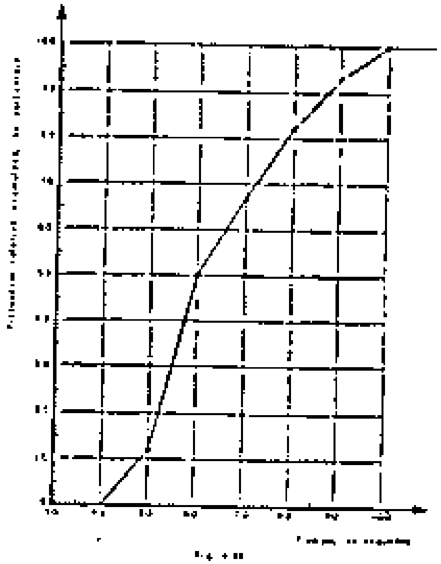
a. Dividen la distribución de frecuencias acumuladas en cuatro partes iguales
b. La dividen en dos partes iguales
c. La dividen en 100 partes iguales

1—b, 3—c, 4—a

---

191 Observe la figura 4.11: en ella se presenta la curva de distribución de frecuencias relativas acumuladas correspondiente a los tiempos que un ciclista tarda en dar una vuelta al circuito de un velódromo.
¿Cuánto vale:

a. $\bar{D}_4$
b. $\bar{P}_{40}$?



57.5, 59.3 (aproximadamente).

---

192 Observe la figura 4.11.
¿Cuánto vale:
a. la mediana
b. $\bar{D}_{10}$?

60.0, 100.0

---

193 Observe la figura 4.11.
¿Cuánto vale:
a. $\bar{C}_1$
b. $\bar{P}_3$?

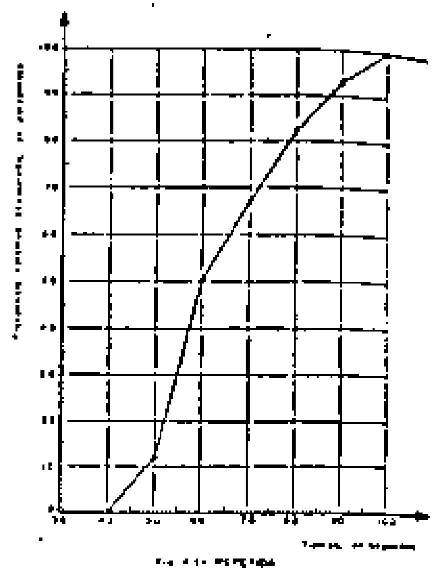
53.5, 41.5 (aproximadamente).

---

194 Observe la figura 4.11.
¿Cuánto vale:

a. $\bar{C}_3$

b. $\bar{D}_t$?



75.3, 48.0.

---

195 En la notación que hemos usado en esta unidad, la letra $\bar{D}$ representa los ________; la $\bar{P}$ los ________; y la $\bar{C}$ los ________.

deciles, percentiles, cuartiles

---

196 Al igual que para los percentiles, para definir en forma *precisa* la totalidad de los deciles y cuartiles, y la mediana, es necesario que la variable aleatoria bajo estudio sea ________ (discreta/continua).

continua

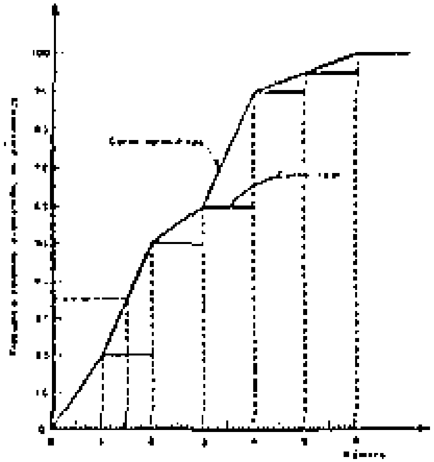
---

197 Un método *aproximado* para calcular percentiles, deciles, cuartiles y medianas, cuando las variables son discretas, consiste en *suponer* que son continuas. En tal caso, la curva de frecuencias acumuladas no queda en forma de escalones y, por lo tanto, queda definido, para cada valor de la frecuencia relativa ________, un valor de la variable.

acumulada

198 Al suponer que las variables discretas son continuas para efectos de calcular percentiles, deciles, cuartiles y medianas, se obtiene una curva de frecuencias acumuladas que no queda en forma de escalones, pudiendo así obtenerse valores fraccionarios de la variable.

De la siguiente figura, $\overline{P}_{35}$ vale 1.5; ¿cuánto valen $\overline{P}_{75}$ y $\overline{D}_4$? ______ , ______



3.5, 1.6 (aproximadamente).

REVISION

199 Una distribución de frecuencias acumuladas se presenta gráficamente mediante una ______ ______ ______ ______.

curva de frecuencias acumuladas

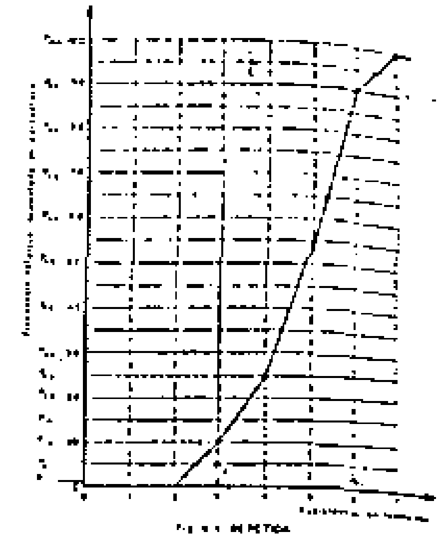
200 La frecuencia acumulada complementaria se define como la diferencia entre el número de datos y la ______ ______ correspondiente.

frecuencia acumulada

201 Si la frecuencia acumulada correspondiente a 165 ton es 37, ¿cuánto vale la frecuencia complementaria si se sabe que se utilizaron 100 datos? ______.

63; (100 − 37 = 63).

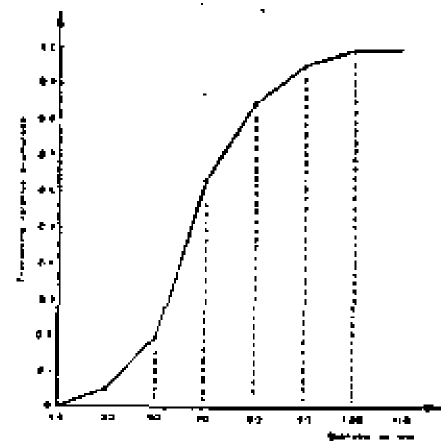
202 Observe la figura 4.4. En ella se ha dibujado una ______ ______ ______ ______ ______.



curva de frecuencias relativas acumuladas

203 Dibuje en la hoja de trabajo 4.14 la curva de distribución de frecuencias relativas acumuladas correspondiente a los datos presentados en la siguiente tabla.

Límite real superior	Frecuencia	Frecuencia relativa	Frecuencia relativa acumulada
50	5	0.05	0.05
60	15	0.15	0.20
70	43	0.43	0.63
80	22	0.22	0.85
90	10	0.10	0.95
100	5	0.05	1.00



204 ¿Qué son los percentiles?

Los percentiles son los valores de la variable que dividen una distribución de frecuencias (relativas) acumuladas, en cien partes iguales.

205 Observe la figura 4.4.

a. ¿A qué percentil corresponde un valor de 5.45 ton?

b. ¿Cuánto vale el 30º percentil?

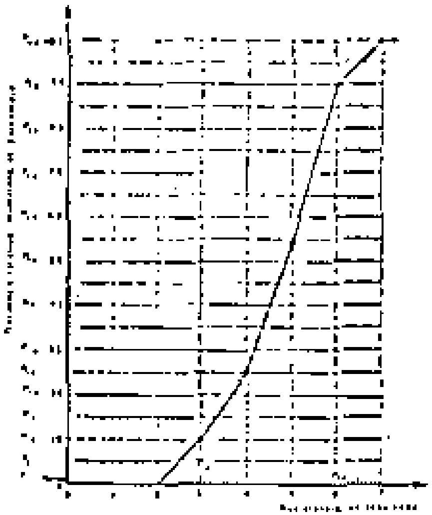


Fig. 4.4 REPETIDA

al 70º. 4.2 (aproximadamente).

206 Observe la figura 4.4.
¿Cuánto vale

a. el tercer cuartil
b. la mediana?

5.6 (aproximadamente). 4.86 (aproximadamente).

207 Observe la figura 4.11.

a. ¿Cuánto vale la mediana?
b. ¿A qué decil corresponde un valor de 52 seg?

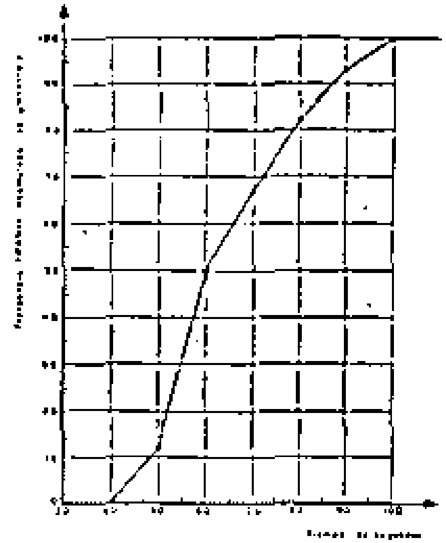


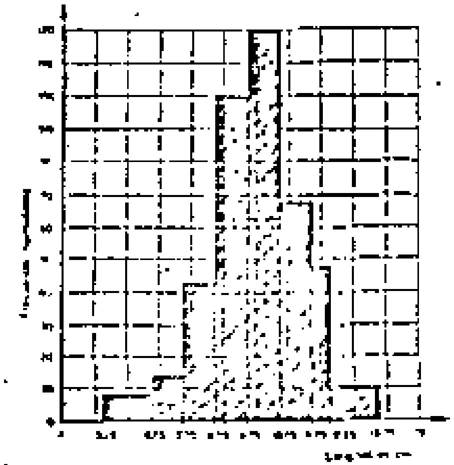
Fig. 4.11 REPETIDA

60 seg. Al segundo decil.

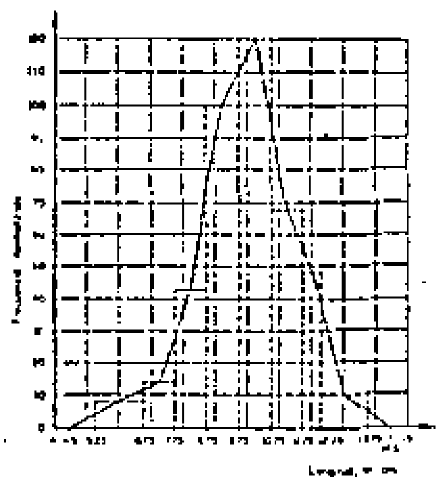
208 **En la tabla 4.6 se presentan las distribuciones de frecuencias de las longitudes de las espigas de trigo, medidas en una investigación de agricultura. Los valores se aproximaron al décimo de centímetro. En la hoja de trabajo 4.15, trace el histograma correspondiente.**

TABLA 4.6

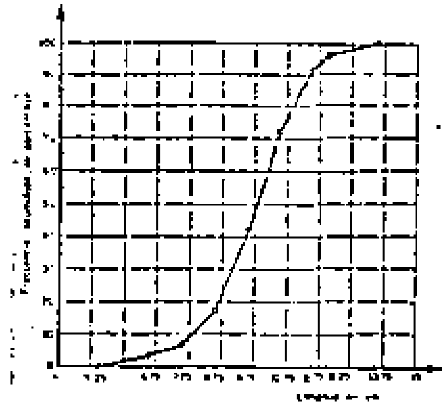
Intervalo de clase, en cm	Marca de clase, en cm	Frecuencia	Frecuencia normalizada	Frecuencia acumulada	Frecuencia relativa acumulada, en %
5.3-6.7	6.0	12	[illegible]	12	3.0
6.8-7.7	7.25	14	14	26	6.5
7.8-8.7	8.25	43	43	69	17.3
8.8-9.7	9.25	100	100	169	42.3
9.8-10.7	10.25	120	120	289	72.3
10.8-11.7	11.25	[illegible]	60	361	90.3
11.8-12.7	12.0	24	48	385	96.3
12.8-13.7	13.0	10	[illegible]	400	100.0
TOTAL		400			



209 Si cometió algún error, corrija su respuesta en la hoja de trabajo 4.15; luego trace, en la misma, el polígono de frecuencias correspondiente (si desea, vea la tabla 4.6).



210 Trace en la hoja de trabajo 4.16 la curva de distribución de frecuencias relativas acumuladas correspondiente a los datos de la tabla 4.6.



211 Si cometió algún error, corrija su respuesta en la hoja de trabajo 4.16 y luego diga cuánto vale:

a. el tercer percentil

b. el décimo decil

a. 6.75 b. 13.75

212 De la hoja de trabajo 4.16, ¿cuánto vale:

a. la mediana

b. el primer cuartil

c. el octavo decil?

a. 10.1, b. 9.1, c. 11.2 (aproximadamente).

EXAMEN

1. *Una distribución de frecuencias se representa gráficamente mediante un ____________________.*

2. *El eje horizontal de un sistema de coordenadas rectangulares se llama eje de las ________ y el vertical, eje de las ________.*

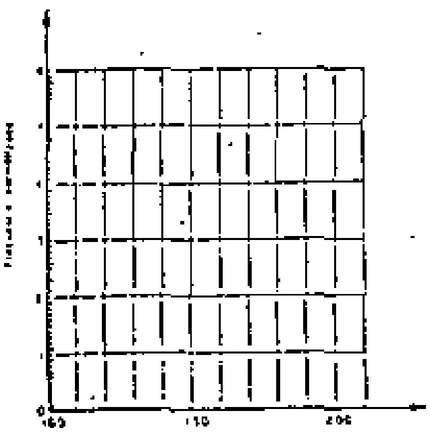
3. *Un histograma es una gráfica de barras. El ancho de cada barra es igual a la ________ del intervalo de clase correspondiente; su altura es la ________ de clase normalizada.*

4. *Una frecuencia normalizada se calcula dividiendo la ________ del intervalo entre su ________.*

5. *Los puntos medios de cada intervalo se denominan ________ ________ ________.*

6. *En la siguiente tabla se presenta la distribución de frecuencias correspondiente a la cantidad de acero producida en un alto horno, en ton/mes. En el sistema de coordenadas dibuje el histograma y el polígono de frecuencias. (25 puntos.)*

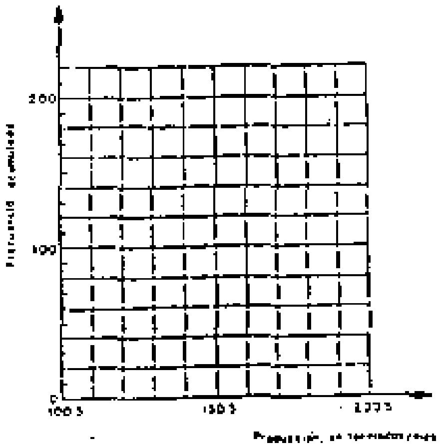
Límites reales		Marcas de clase	Frecuencia	Frecuencia normalizada
Inferior	Superior			
110.5	130.5		9	
130.5	140.5		29	
140.5	150.5		54	
150.5	160.5		58	
160.5	170.5		44	
170.5	180.5		22	



7. *El área de una barra de un histograma es igual a la ________ tervalo de clase correspondiente. El área de todas las barras es igual al ________ de datos.*

8. *Al trazar una curva de frecuencias acumuladas, cada punto queda definido por el límite ________ ________ del intervalo correspondiente, en el eje de las abcisas, y por la ________ ________, en el de las ordenadas.*

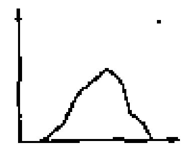
9. *Trace en el siguiente sistema de coordenadas rectangulares la curva de frecuencias acumuladas de los siguientes datos. (7 puntos.)*



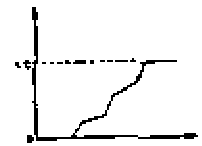
Límites reales		Frecuencia acumulada
Inferior	Superior	
110.5	130.5	9
130.5	140.5	38
140.5	150.5	92
150.5	160.5	150
160.5	170.5	194
170.5	180.5	214

10. *Relacione los nombres de la izquierda con las figuras que aparecen a la derecha.*

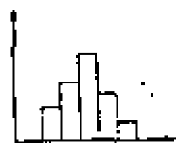
1. *Histograma* — a.



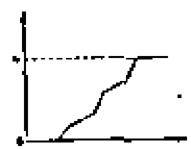
2. *Polígono de frecuencias* — b.



3. *Curva de frecuencias acumuladas* — c.



4. *Curva de frecuencias relativas acumuladas* — d.



11. *La curva de frecuencias acumuladas proporciona las frecuencias de los valores __________ o iguales que un valor dado de la variable. En cambio, la curva de frecuencias acumuladas ____________________ proporciona la frecuencia de los valores mayores.*

12. *Los valores de la variable que dividen una distribución de frecuencias en 100 partes iguales se llaman ____________________; los que la dividen en cuatro partes iguales se llaman ____________________.*

13. *El 25º percentil coincide con el __________ cuartil. El quinto decil coincide con el __________ percentil, con el __________ cuartil y con la __________.*

14. *La siguiente figura corresponde a la venta diaria de un almacén, en miles de pesos. ¿cuánto vale*

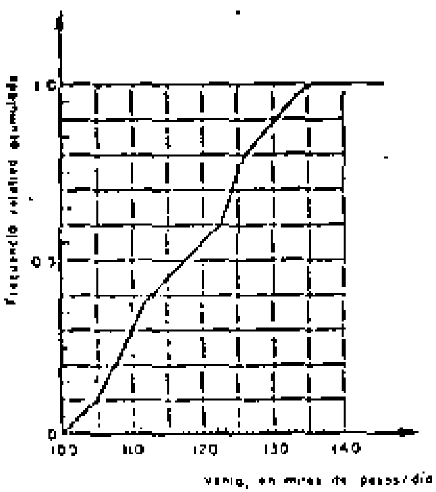
   a. $\bar{P}_{10}$

   b. $\bar{D}_3$

   c. $\bar{C}_3$

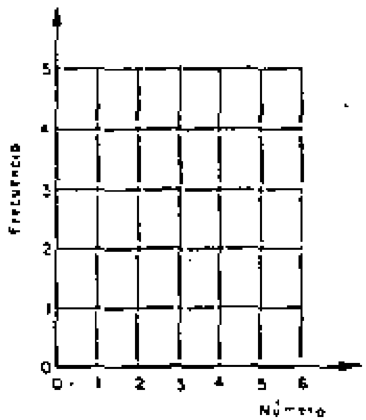
   d. $\bar{P}_{100}$

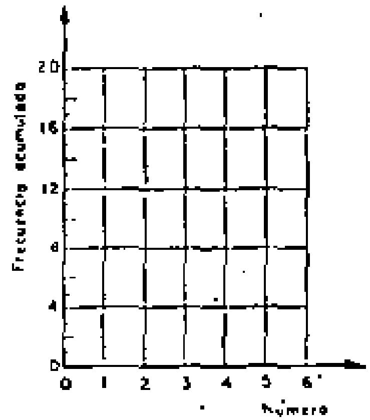
   e. *la mediana?*



15. *En una serie de 20 experimentos para estudiar si un dado está cargado, se obtubieron los siguientes resultados:*

3, 5, 6, 6, 2, 5, 1, 6, 2, 4, 1, 4, 3, 2, 5, 6, 3, 2, 5, 1. Trace el histograma y la curva de frecuencias acumuladas. (12 puntos.)





TOTAL: 73 puntos

RESPUESTAS

1. *Histograma*

2. *abscisas*
   ordenadas

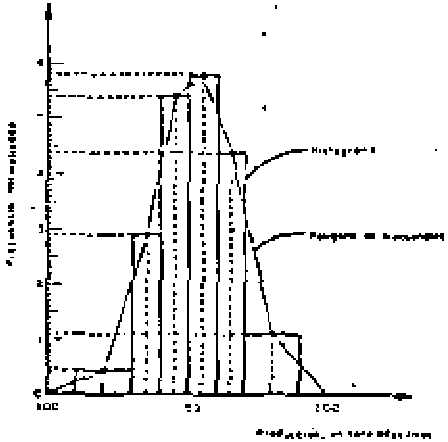
3. *amplitud*
   frecuencia

4. *frecuencia*
   amplitud

5. *marcas de clase*

6.

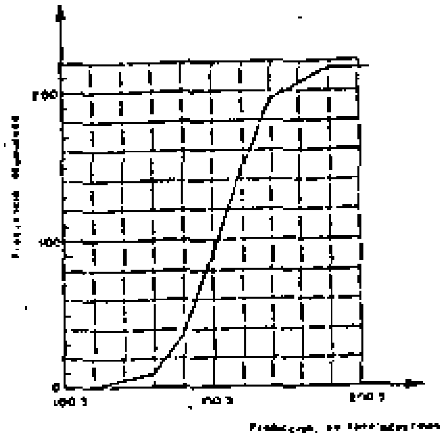
Marcas de clase	Frecuencia	Frecuencia [illegible]
[illegible]	9	0.45
135.5	[illegible]	2.90
145.5	[illegible]	5.40
155.5	[illegible]	5.50
[illegible]	[illegible]	4.40
[illegible]	22	1.10



7. *frecuencia*
   número (total)

8. *real superior*
   frecuencia acumulada

9.



10. 1-c 2-a 3-d 4-b

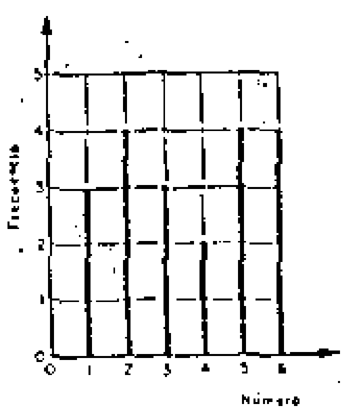
11. *menores*
    complementarias

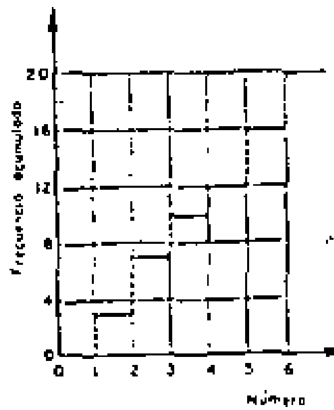
12. *percentiles*
    cuartiles

13. *primer*
    10º
    segundo
    mediana

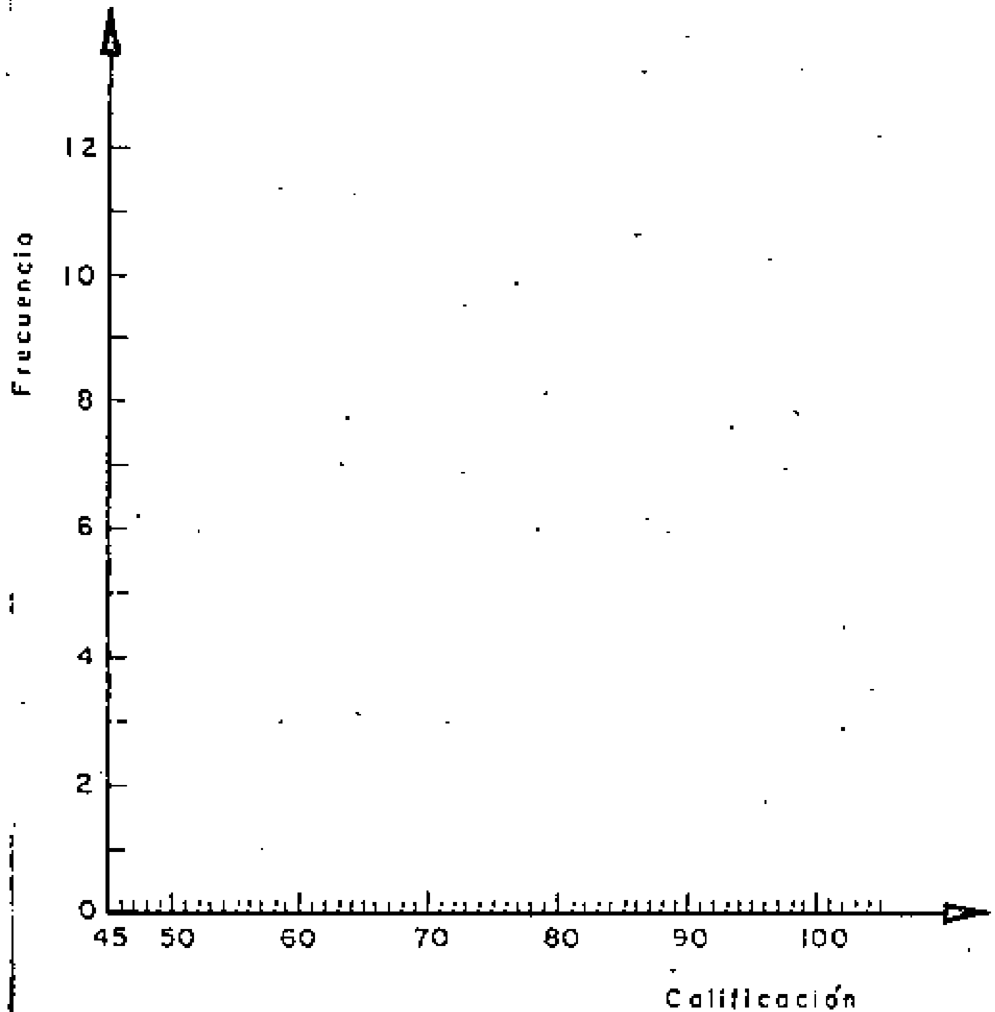
14. a. 105.0 b. 110.0 c. 125.0 d. 135.0 e. 117.5.

15.

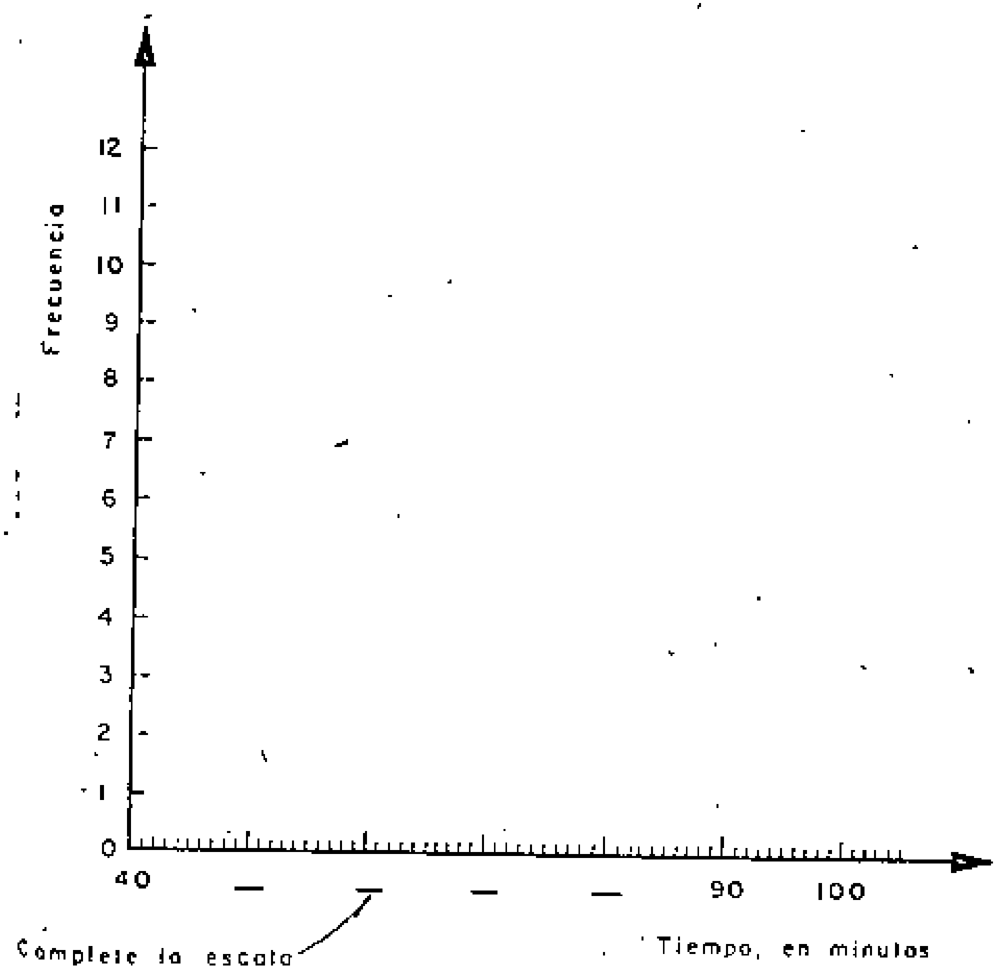




Hoja de trabajo 4.1



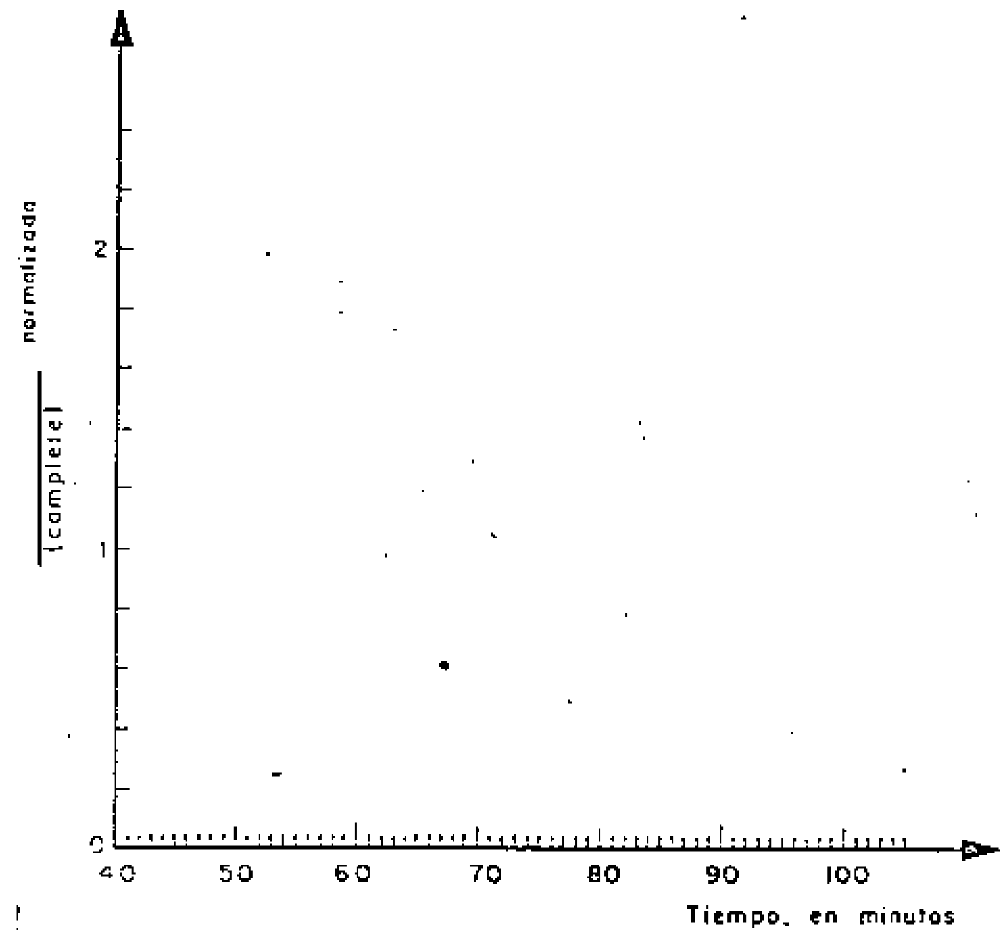
Hoja de trabajo 4.2



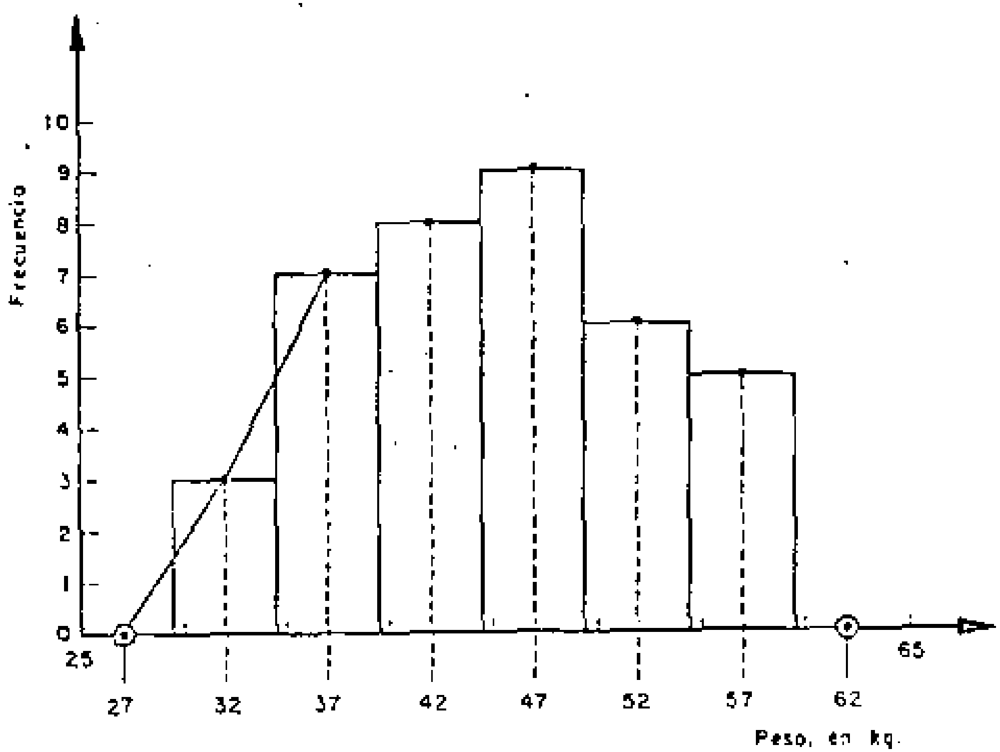
HOJA DE TRABAJO 4.3

Intervalo de clase	Frecuencia	Frecuencia normalizada	Límites reales	
			Inferior	Superior
41-55	3	0.2		
56-60	5			
61-65	6			
66-70	11		65.5	70.5
71-75	11	2.2	70.5	75.5
76-80	11	2.2	75.5	80.5
81-85	11	2.2	80.5	85.5
86-90	6	1.2	85.5	90.5
91-100	8		90.5	100.5

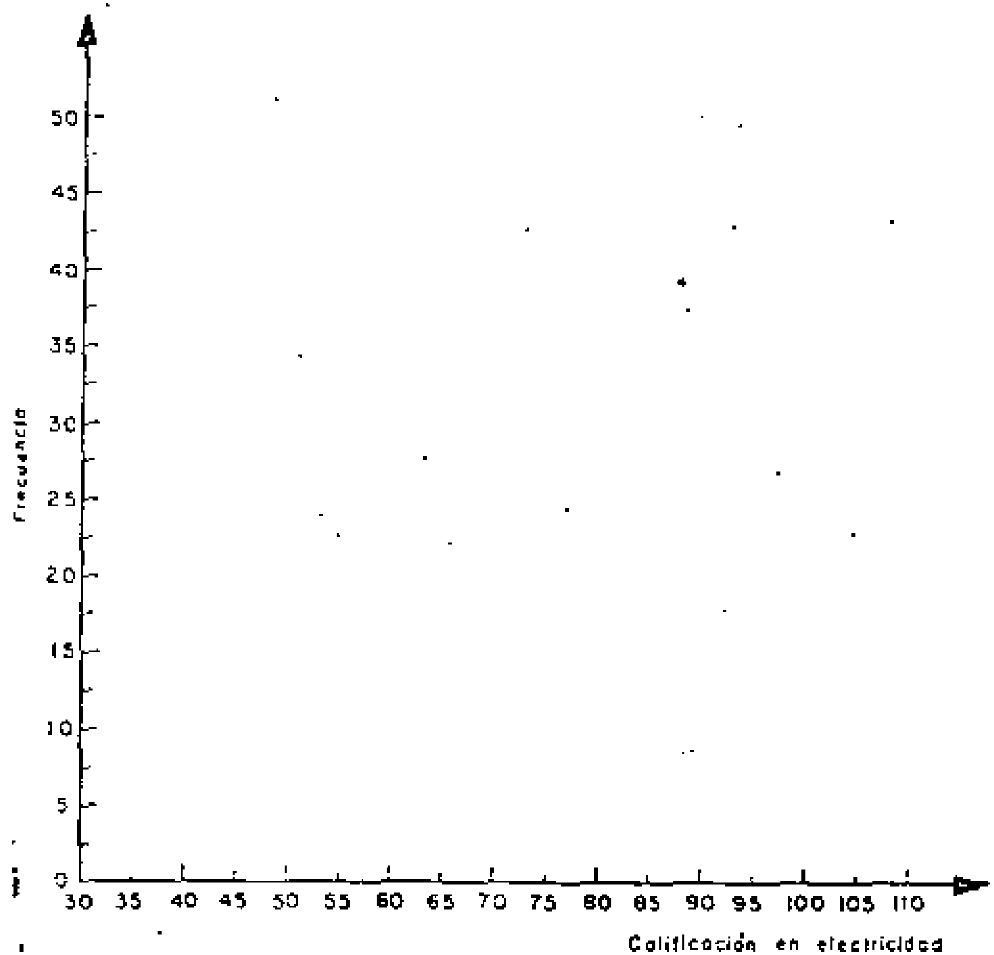
Hoja de trabajo 4.4



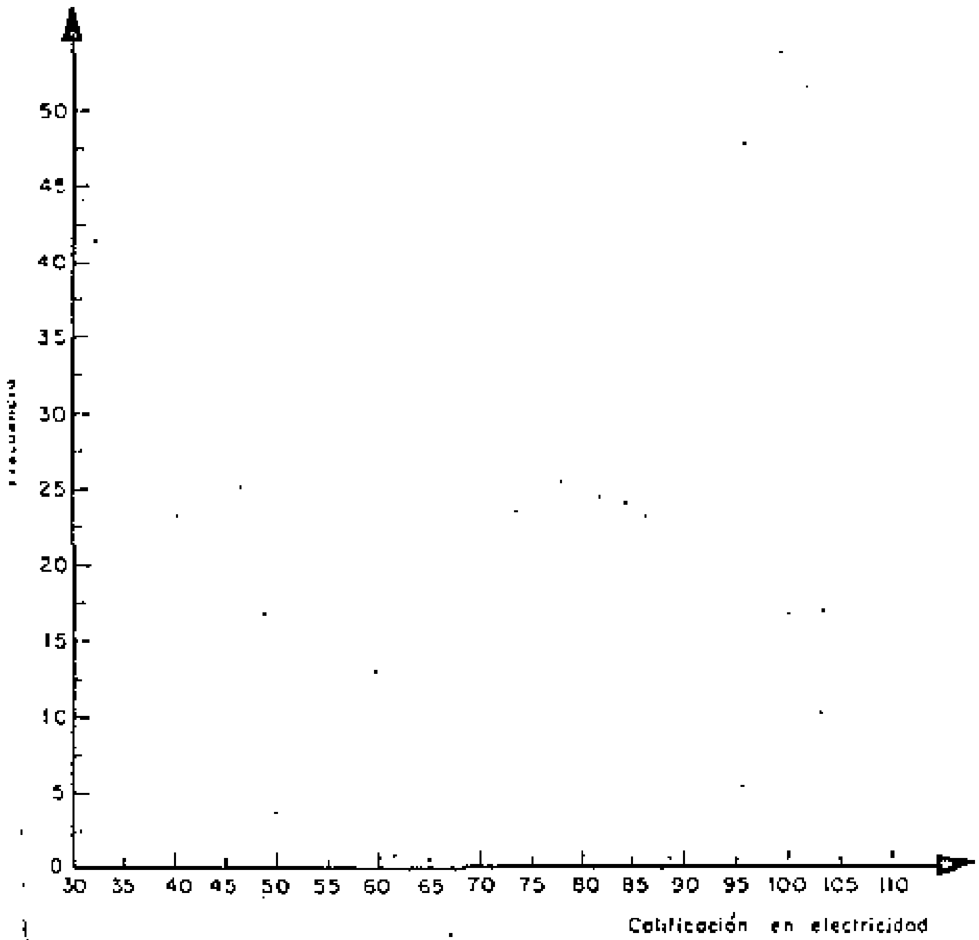
Hoja de trabajo 4.5



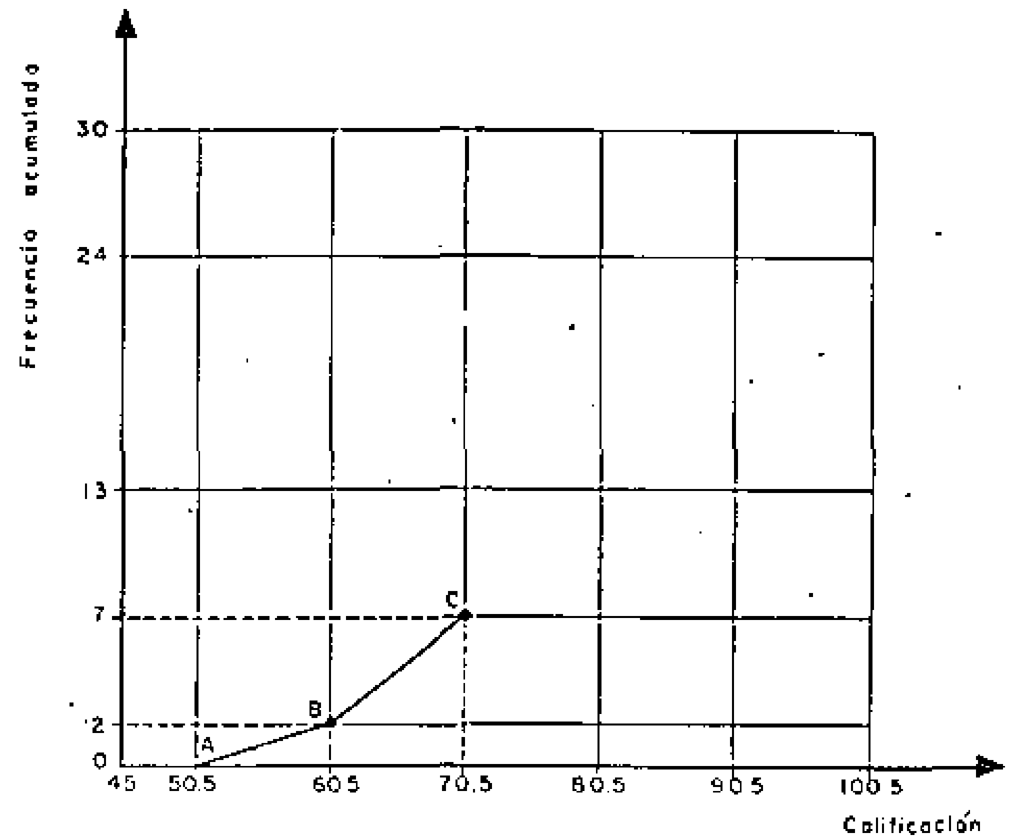
Hoja de trabajo 4.6



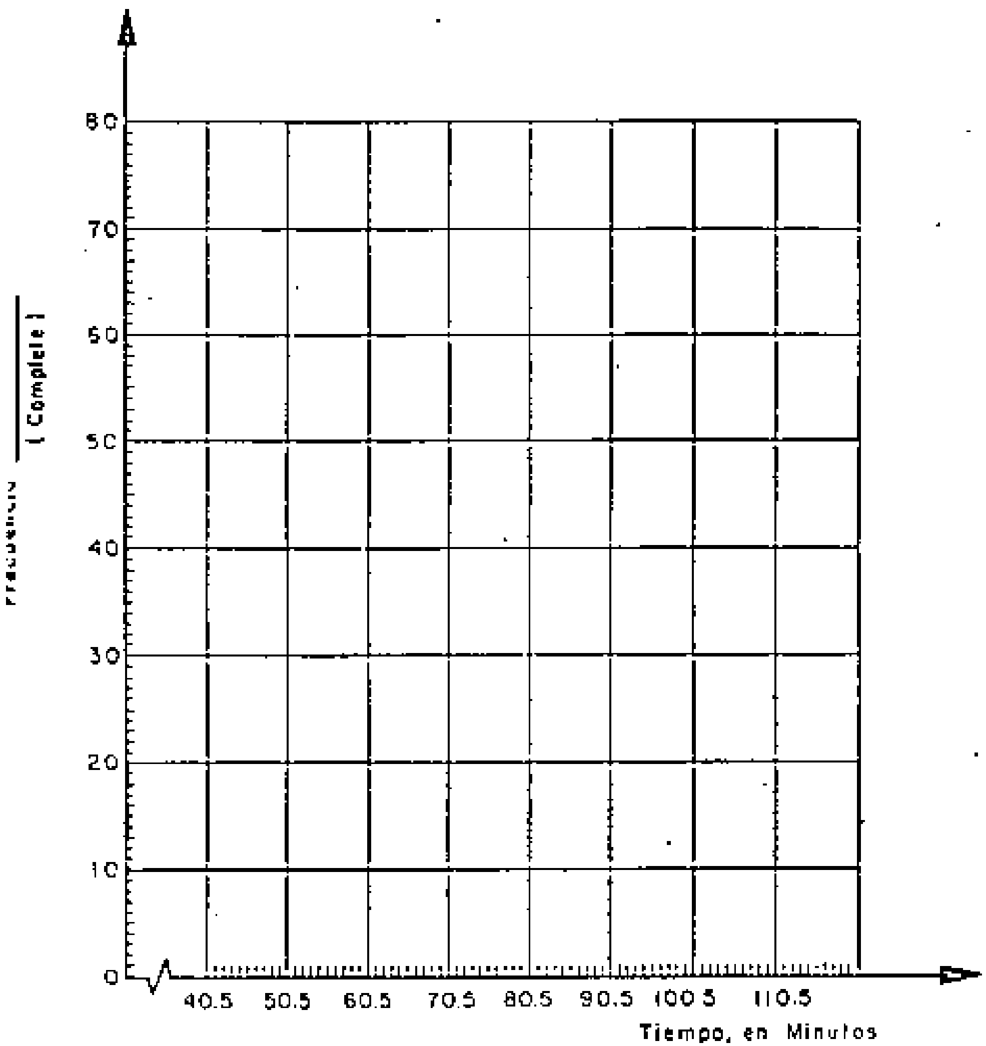
Hoja de Trabajo 4.7



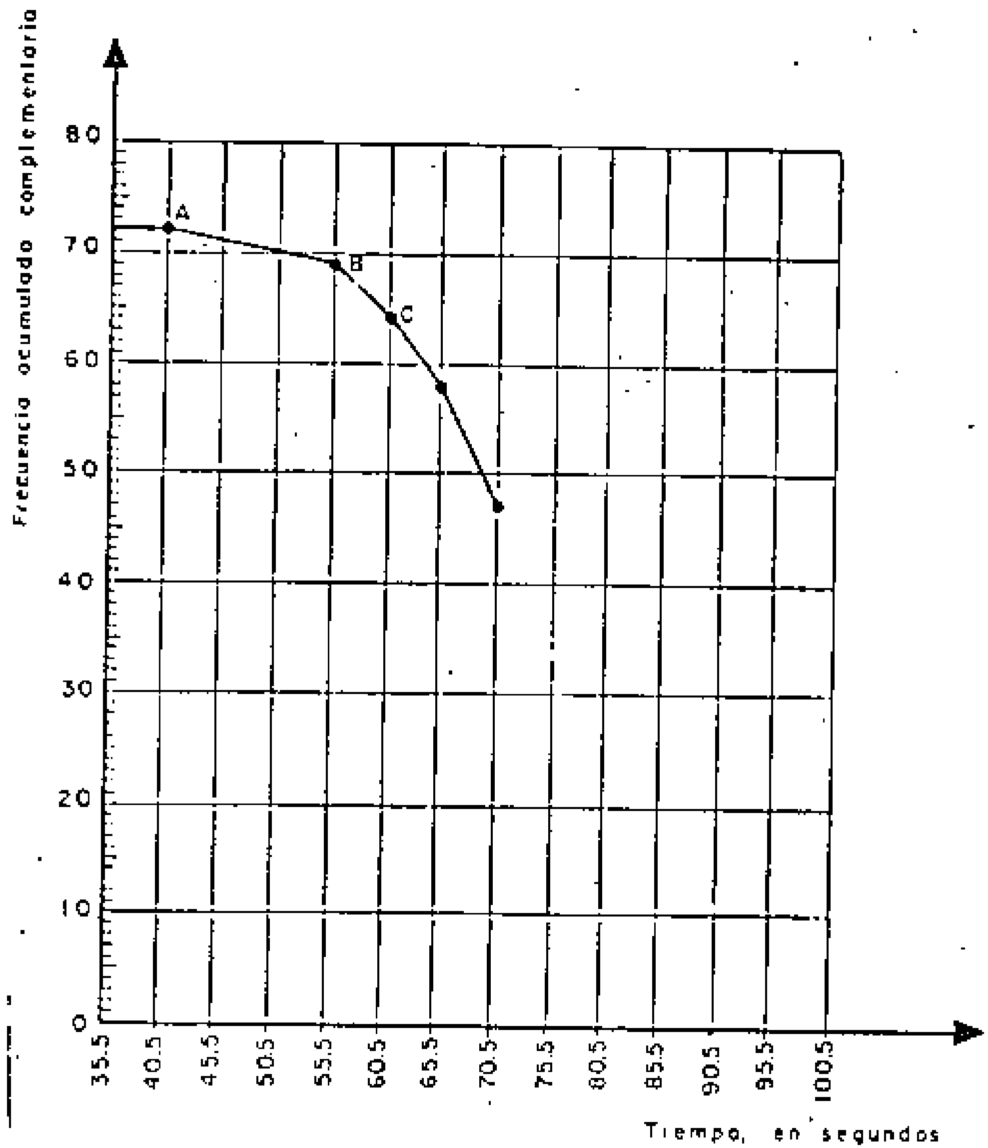
Hoja de Trabajo 4.8



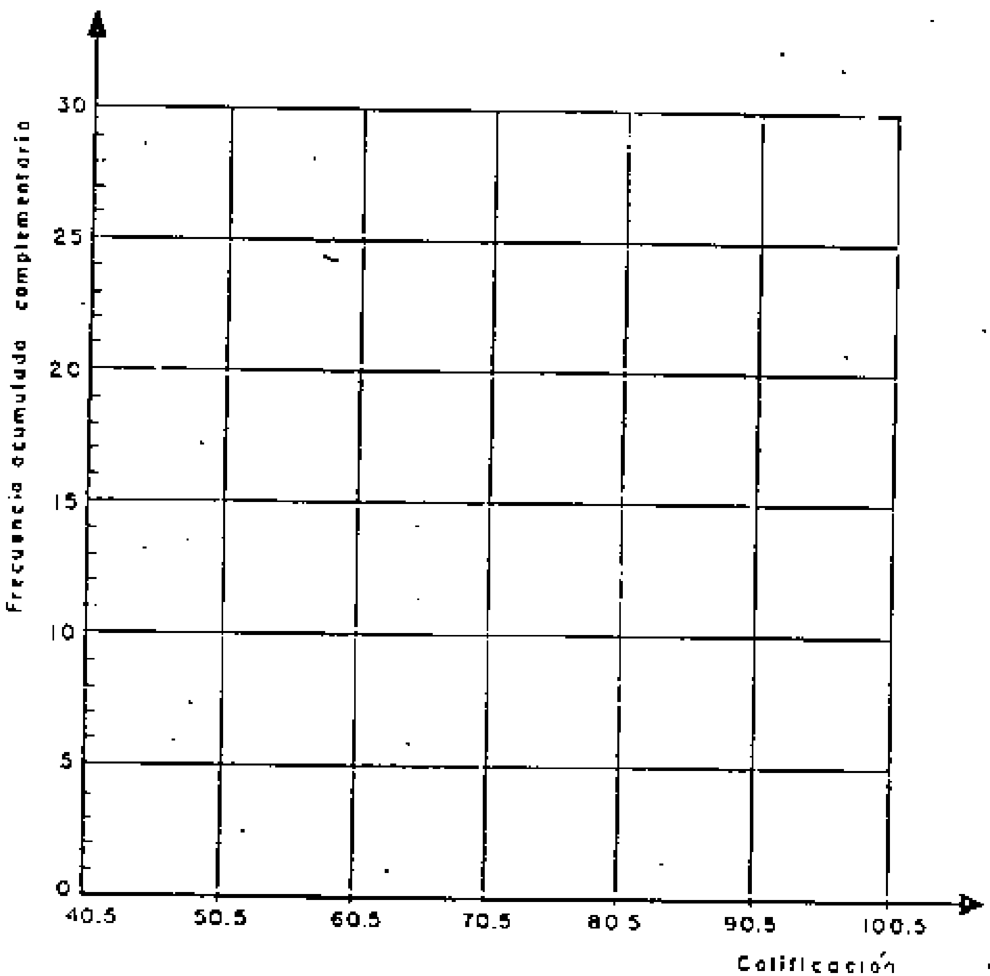
Hoja de trabajo 4.9



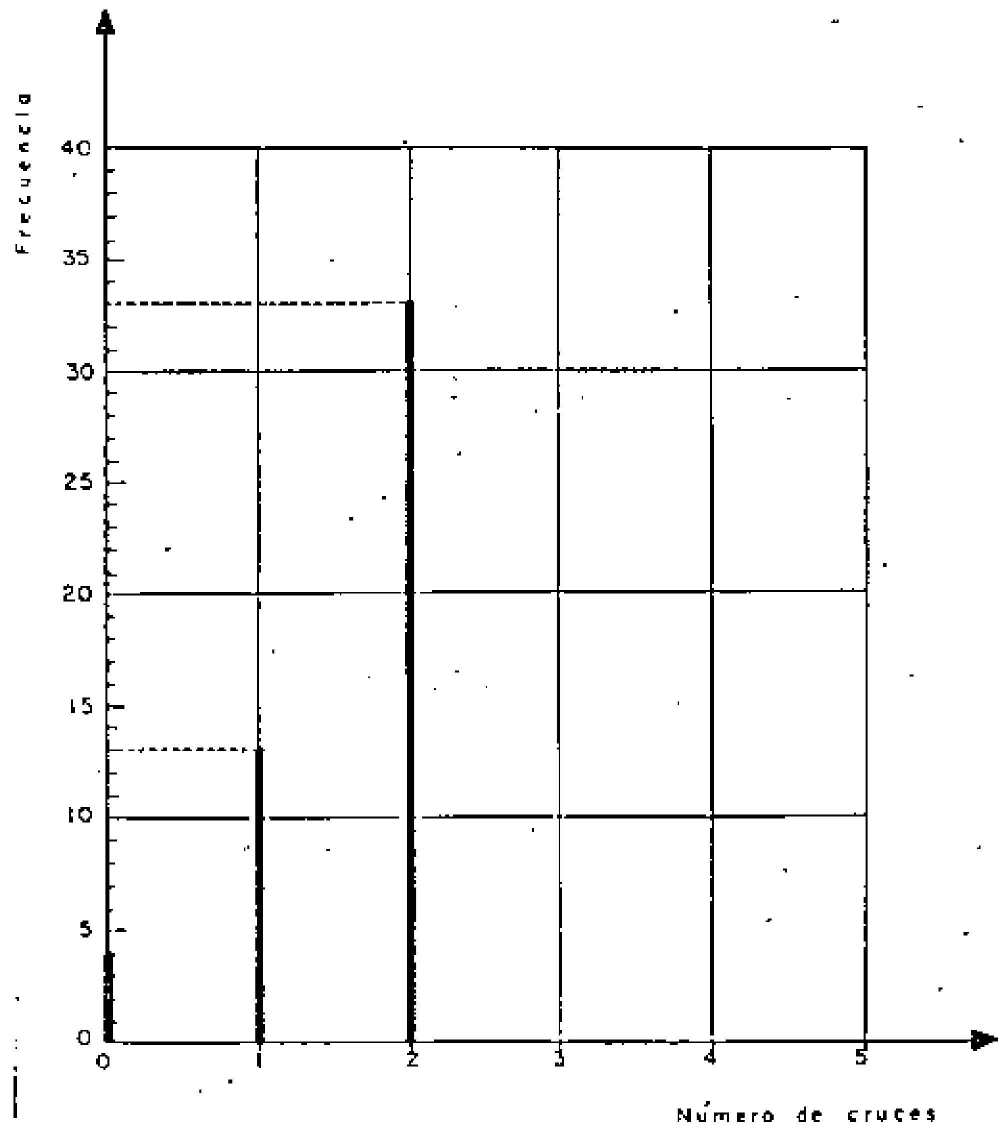
Hoja de trabajo 4.10



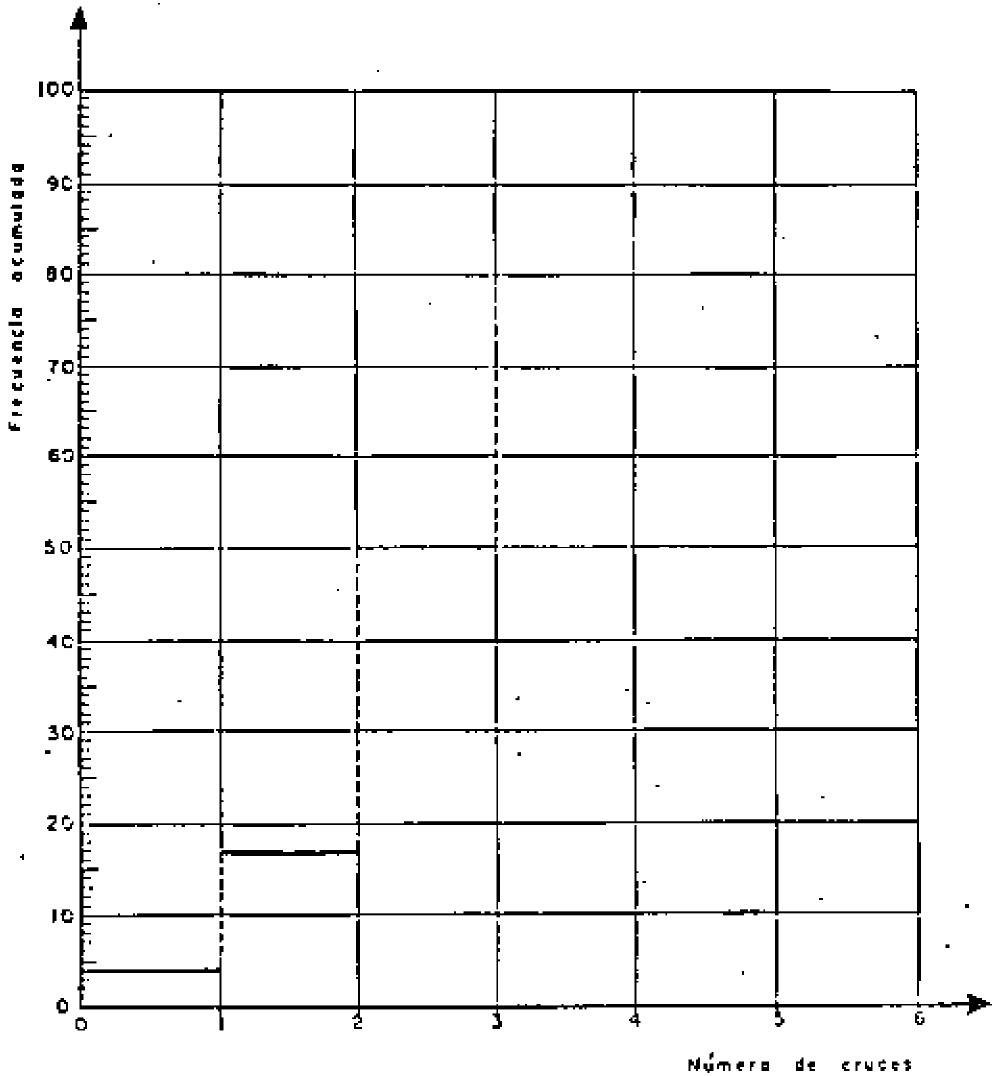
Hoja de trabajo 4.11



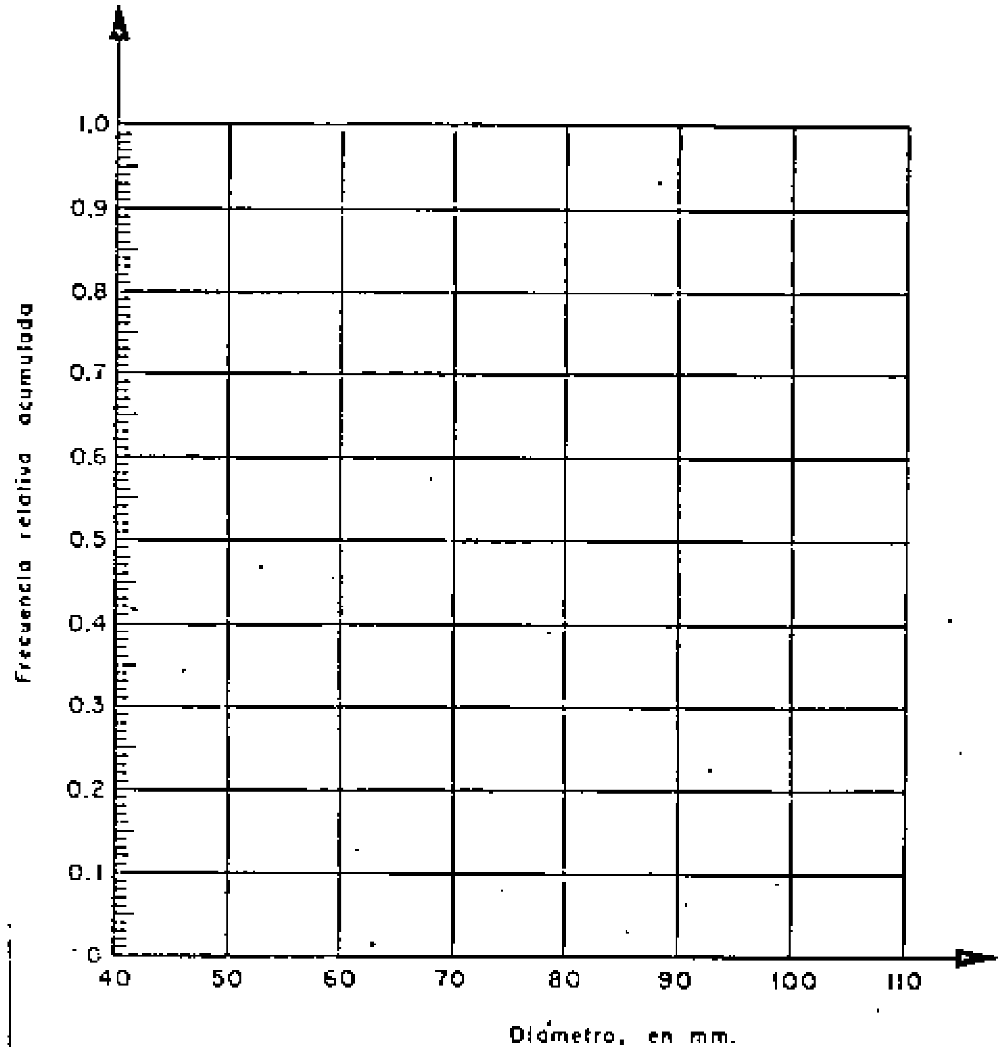
Hoja de trabajo 4.12



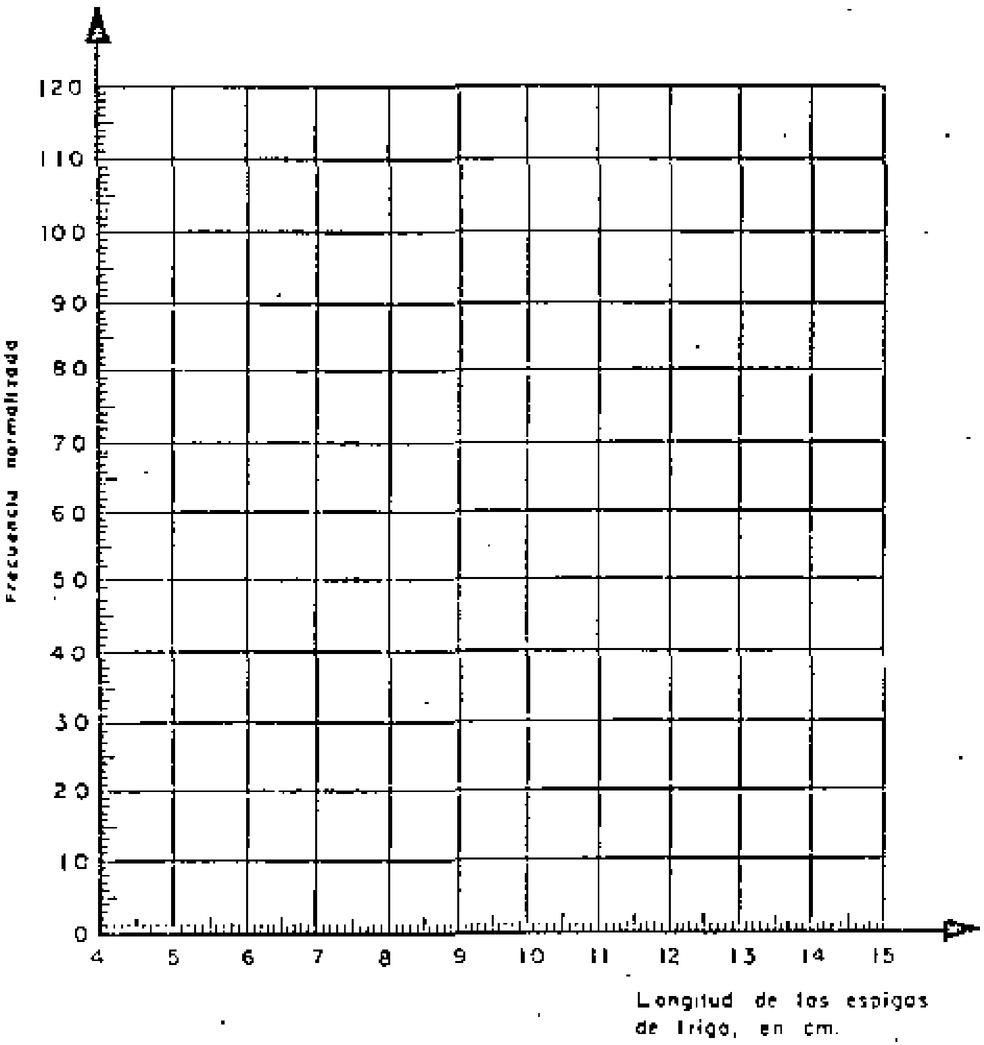
Hoja de trabajo 4.13



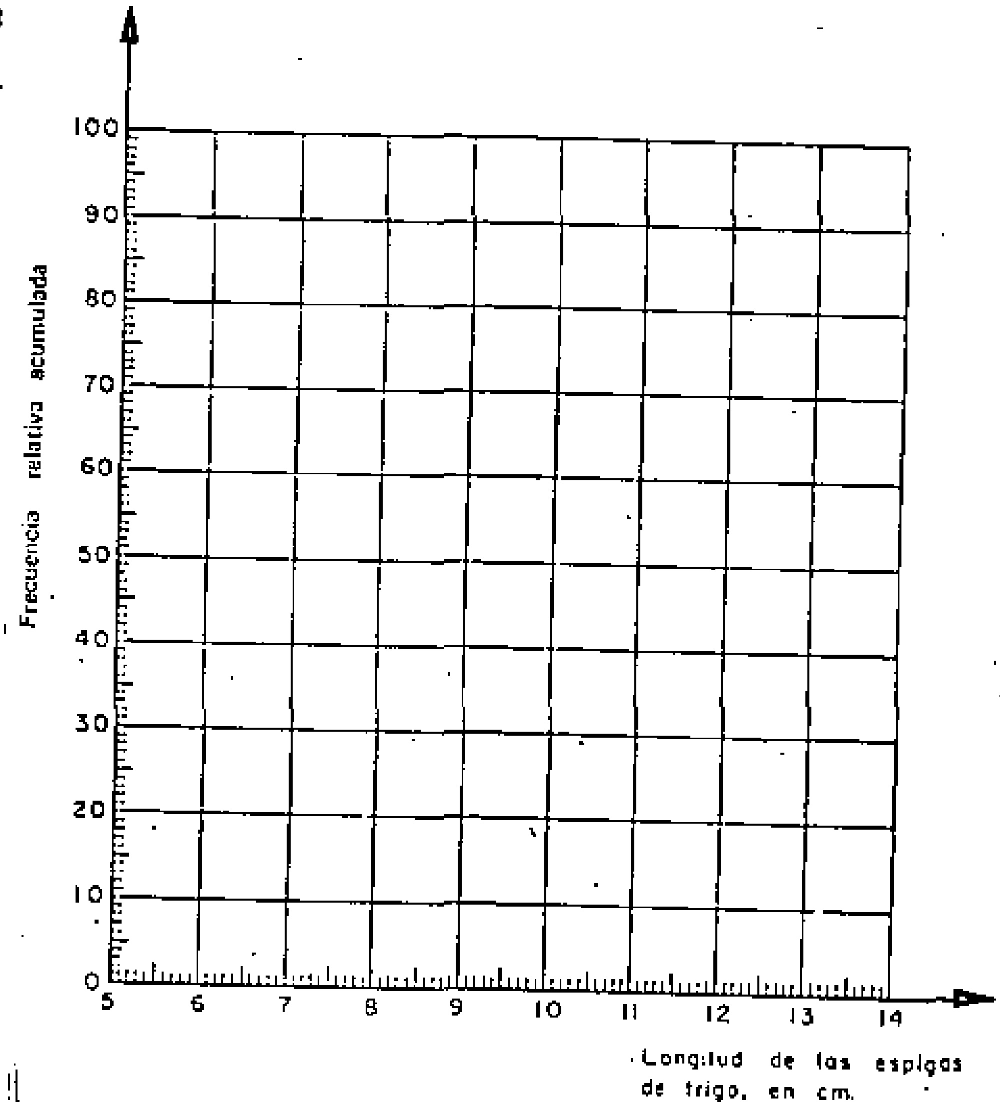
Hoja de trabajo 4.14



Hoja de trabajo 4.15



Hoja de trabajo 4.16



UNIDAD V MEDIDAS DE TENDENCIA CENTRAL

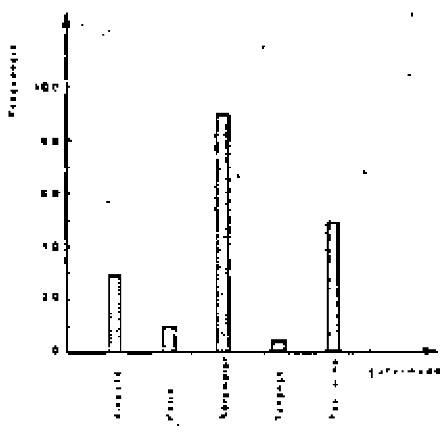
PREFACIO

La razón principal para agrupar los datos, calcular las distribuciones de frecuencias y presentar gráficamente los resultados, es determinar el comportamiento del fenómeno que interesa analizar. Aunque un histograma, por ejemplo, proporciona bastante información, en ocasiones es suficiente contar con algunas descripciones numéricas de la distribución; tales números proporcionan una idea de los valores de la variable alrededor de los cuales tienden a aglomerarse las observaciones *(medidas de posición o tendencia central)*, o dan una idea de la dispersión o variabilidad de las observaciones *(medidas de dispersión o variabilidad)*.

En esta unidad nos concretaremos a estudiar algunas de las medidas de tendencia central más útiles.

PARTE A. EL MODO

1 Observe que en la figura 5.1 se presenta un histograma correspondiente al número de pacientes que durante un mes ingresaron a un hospital infantil con diferentes enfermedades. Estos datos corresponden, por lo tanto, a una variable aleatoria que asume valores ______
(numéricos/nominales)



nominales

---

2 Puesto que las variables nominales no tienen una forma natural de ser ordenadas (del valor mínimo al máximo, o viceversa), los diferentes atributos pueden ordenarse de ______ manera (s).
(una sola/varias)

varias

---

3 Puesto que las variables nominales pueden ordenarse en forma ______ no es posible pensar en un "centro" de la distribución de frecuencias. Sin embargo, la distribución tiende a aglomerarse en ciertos puntos donde la frecuencia es máxima.
(única/arbitraria)

arbitraria

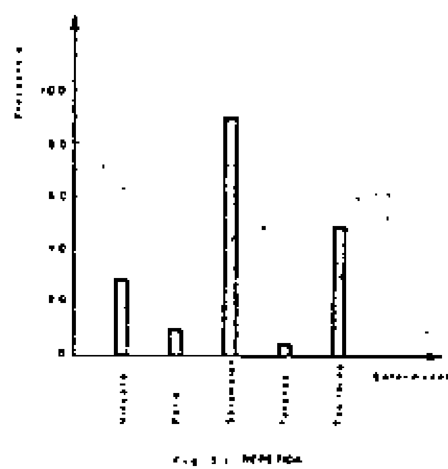
---

4 Observe la figura 5.1. ¿Cuál es la clase (el valor nominal) de máxima frecuencia?

Sarampión.

---

5 A la clase de mayor frecuencia en una distribución se le conoce con el nombre de *modo, moda o valor modal.* Observe la figura 5.1. El modo corresponde a la clase ______
(sarampión/tos ferina)



sarampión

---

6 El "sarampión" constituye el modo de la distribución de frecuencia de la figura 5.1 porque es la clase de ______ frecuencia.
(mayor/menor)

mayor

7 Si en una distribución de frecuencias se tiene que la frecuencia de una clase no es superada por ninguna otra, se dice que esta clase es ______________ de la distribución.
(la muestra/el modo)

el modo

8 El modo de una distribución es la clase de______________ frecuencia.
(mínima/máxima)

máxima

9 Para determinar el modo (moda o valor modal) de una distribución es necesario agrupar previamente los datos en una tabla de ______________
(frecuencias/frecuencias acumuladas)

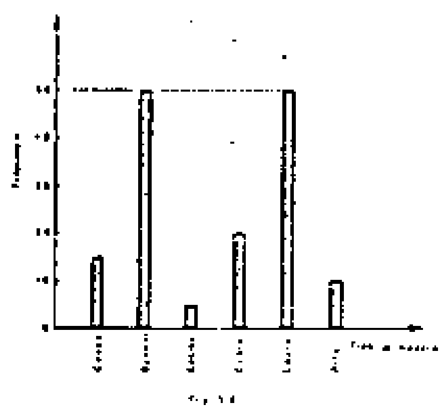
frecuencias

10 En una distribución de frecuencias, la clase de máxima frecuencia se llama ______________.

modo

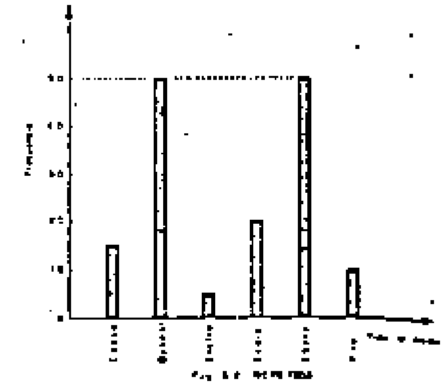
11 Observe la figura 5.2 en la que se presenta el histograma correspondiente a los datos de la variable ______________ "tipo de madera".
(escalar/nominal)

La frecuencia indica metros cúbicos por kilómetro cuadrado (m^3/km^2). ¿Existen dos clases que tienen la mayor frecuencia?



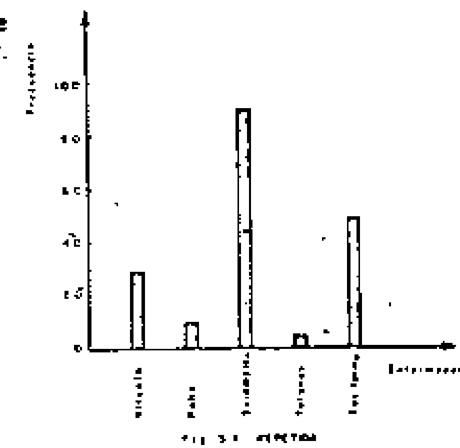
nominal, Sí

12 Cuando existen dos clases que tienen la mayor frecuencia, se dice que la distribución es *bimodal* (tiene dos modos). Observe la figura 5.2.
¿Cuáles son los dos modos de esta distribución?



Oyamel y Ebano.

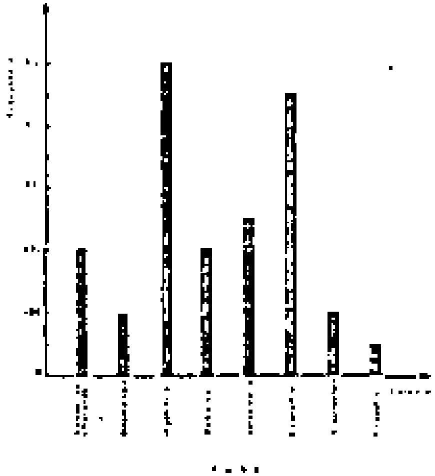
13 Una distribución de frecuencias que tiene dos modos se conoce como *bimodal*.
Observe la figura 5.1.
¿Es bimodal esta distribución?



No.

14 Aun cuando las dos frecuencias más grandes sean sólo aproximadamente iguales, es costumbre decir que la distribución es bimodal. En la figura 5.3 se presenta el histograma correspondiente a los estudiantes de diversas carreras que tomaron un curso de Estadística durante 1968. La frecuencia de la clase "Economía" es ______________ igual a la de la clase "Ingeniería", por lo cual
(aproximadamente/exactamente)

la distribución se puede considerar
______________.
(unimodal/bimodal)



aproximadamente, bimodal

15 El modo es una medida de tendencia central particularmente apropiada para indicar la tendencia de concentración (aglomeración) de los datos de variables nominales, ya que éstas no pueden tomar valores ______________.
(numéricos/nominales)

numéricos

16 Aunque el modo es una medida natural de tendencia o posición central de los datos de variables nominales, también se puede utilizar para variables ______________, que son las que asumen valores numéricos. (escalares/nominales)

escalares

17 Si los datos de una variable escalar *no* están agrupados en intervalos de clase, el ______________ de la distribución es el valor del *dato de mayor frecuencia.*

modo

18 Si los datos *están* agrupados en intervalos de clase, el modo es la *marca de clase* del intervalo de ______________ frecuencia.

mayor

19 En el caso de una variable escalar cuyos datos no están agrupados en intervalos de clase, el modo es el valor de la observación que ocurre ______________ frecuentemente.
(menos/más)

más

20 En el caso de una variable escalar cuyos datos están agrupados en intervalos de clase, el modo es la ______________ de clase del intervalo de ______________ frecuencia.

marca, máxima

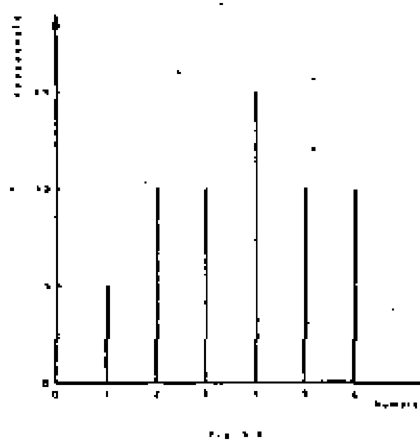
21 Sea la siguiente distribución de frecuencias.

Valor de la variable	Frecuencia
9	2
4	6
3	3
1	4
2	1
8	1
6	5
0	0
5	3
7	4

El dato de mayor frecuencia es ___.
Por lo tanto, el modo de esta distribución es ___.

4, 4

22 En la figura 5.4 se presenta el histograma correspondiente a los números que quedaron hacia arriba al lanzar un dado 60 veces. ¿Esta distribución es unimodal o bimodal? ______________



unimodal

23 ¿Cuál es el modo de la distribución presentada en la figura 5.4 ? ________ .

4

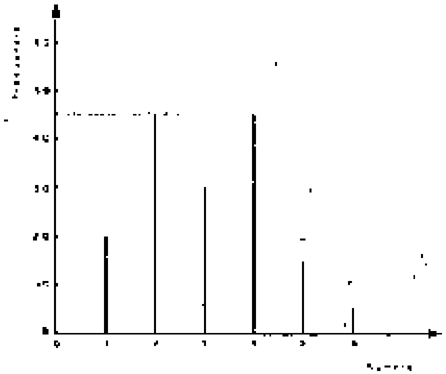
---

24 Cuando se trata de una distribución de frecuencias de los datos de una variable escalar, el modo se da mediante un valor ____________ .
(nominal/numérico)

numérico

---

25 Observe la figura 5.5 en la que se presenta el histograma correspondiente al número de hijos de un matrimonio.
¿Es bimodal esa distribución de frecuencias? ________________ .



Sí.

---

26 La distribución de frecuencias de la figura 5.5 sí es bimodal porque presenta dos clases cuya frecuencia es la ____________ .
(mínima/máxima)

máxima

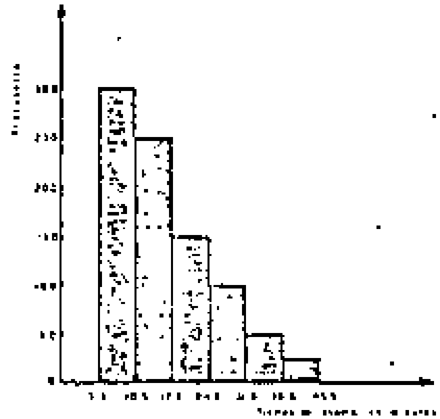
---

27 Observe la figura 5.5.
¿Cuáles son los dos modos de esa distribución? ________________

2 y 4

---

28 Observe la figura 5.6 que presenta el histograma de los datos de la variable continua "tiempo de espera para obtener líneas para una llamada telefónica de larga distancia". ¿Cuál es el intervalo de clase de máxima frecuencia?



3.5-10.5

---

29 Recuerde que cuando se trata de datos agrupados en intervalos de clase, el modo corresponde a la ________ de clase del intervalo de mayor frecuencia.

marca

---

30 Observe la figura 5.6.
¿Cuál es la marca de clase del primer intervalo?

7.0 min $\left(3.5 + \frac{10.5 - 3.5}{2} = 7 \text{ min}\right)$

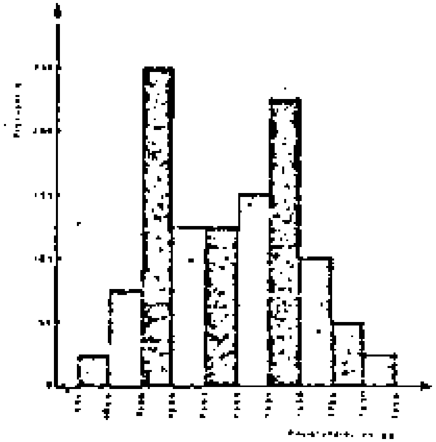
---

31 Si la marca de clase del intervalo de mayor frecuencia es 7.0, el modo de la distribución vale ________.

7.

---

32 En la figura 5.7 se presenta el histograma de los datos de la variable "resistencia de las varillas de acero marca XX". ¿Cuál es el modo de esa distribución?



1200 kg $\left(1150 + \frac{100}{2} = 1200 \text{ kg}\right)$

33 Aunque en forma rigurosa, la distribución de frecuencias de la figura 5.7 sólo tiene un intervalo de clase de mayor frecuencia, existe otro intervalo cuya frecuencia se le aproxima bastante. ¿Cuál es?

1550-1650 kg.

34 ¿Cuál es la marca de clase del intervalo 1550-1650?

1600 kg.

35 Observe la figura 5.7.
Puesto que hay dos intervalos de clase cuyas frecuencias sobresalen marcadamente de las demás, esa distribución de frecuencias puede considerarse como ______

(bimodal/trimodal)

bimodal

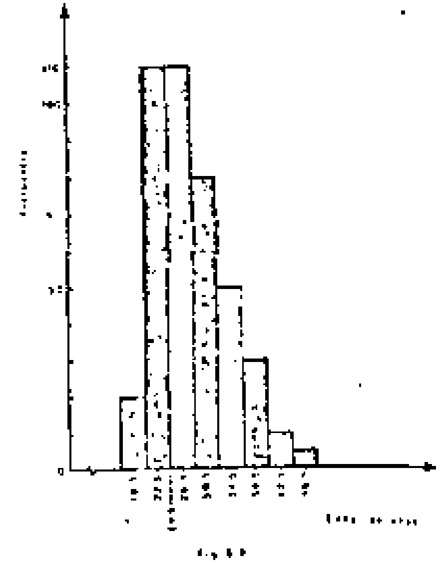
36 Observe la figura 5.7.
Puesto que esa distribución de frecuencias puede considerarse como bimodal, ¿cuáles son los valores de los dos modos?

1200 y 1600 kg.

37 Los dos modos de la distribución presentada en la figura 5.7 (1200 y 1600 kg; es decir, las ______ de clase de los intervalos de ..., or frecuencia.

marcas

38 Observe la figura 5.8 en la que se presenta el histograma correspondiente a la edad en la cual contrajeron matrimonio los varones de un pueblo. Esta distribución tiene un solo modo a pesar de existir dos intervalos de mayor frecuencia. Se considera que es unimodal porque sus ______ de máxima frecuencia están juntos.



intervalos

39 Puesto que la distribución de frecuencias presentada en la figura 5.8 (obsérvela) tiene dos intervalos adyacentes, cuyas frecuencias son las máximas, lo más lógico es tomar como modo el valor ______ años.
(22.5/24.5/26.5)

24.5

40 ¿Cuál es el modo de la distribución de frecuencias presentadas en la siguiente tabla, correspondiente a la resistencia de un pegamento para ensamblar piezas de avión?

Intervalo, en kg/cm²	Frecuencia
17-20	20
21-24	110
25-28	90
29-32	70
33-36	40
37-40	30
41-44	17
45-48	6
49-52	1

22.5

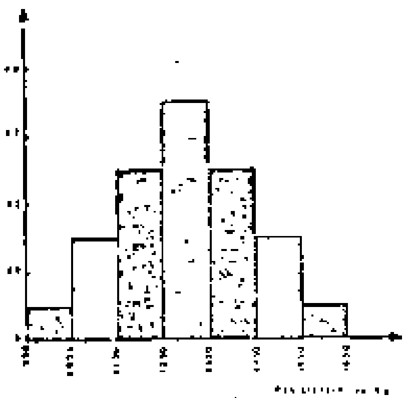
41 El modo de la distribución de frecuencias del cuadro anterior vale 22.5, porque éste es el valor de la __________ __ ____ del intervalo 21-24, que es el de máxima frecuencia.

marca de clase

42 Conviene mencionar que no todas las distribuciones de frecuencias están cargadas hacia el mismo lado. Así, en la figura 5.8 la distribución de frecuencias está cargada hacia ________________ .
(la izquierda/la derecha/el centro)

la izquierda

43 Observe la figura 5.9, correspondiente a la resistencia de algunos especímenes de madera sujetos a fuerza de compresión. La distribución de frecuencias está cargada hacia ______________ ________.



el centro

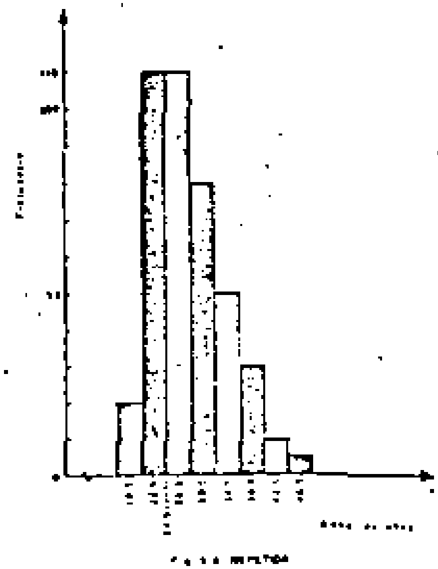
44 Observe la figura 5.9.
¿Los intervalos de clase adyacentes al intervalo que contiene el modo tienen la misma frecuencia? ¿Cuánto vale? __________ .

Sí, 50

45 Observe la figura 5.9.
Cada intervalo de clase situado a la derecha del intervalo que contiene al modo, ¿tiene la misma frecuencia que el respectivo de la izquierda que está a la misma distancia del modo? ________ .

Sí.

46 Cuando esto sucede, es decir, cuando hay *simetría* con respecto al *modo* en una distribución de frecuencias unimodal, se dice que la distribución es *simétrica*.
Vea la figura 5.8; ¿es simétrica la distribución de frecuencias presentada en ella? __________.

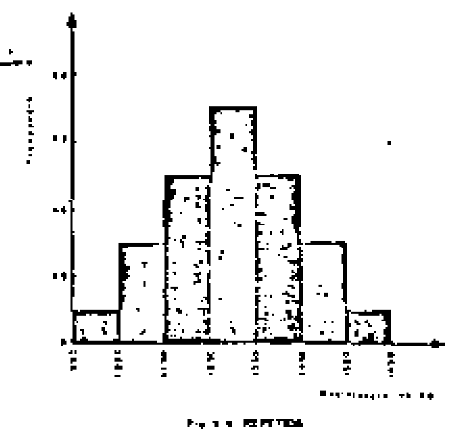


No.

47 Se dice que una distribución unimodal es simétrica cuando hay simetría con respecto al ______ de la distribución.

modo

48 Observe la figura 5.9.
¿Es simétrica esta distribución? ________.



Sí.

49 Se dice que una distribución de frecuencias unimodal es __________ cuando hay simetría con respecto al modo.

simétrica

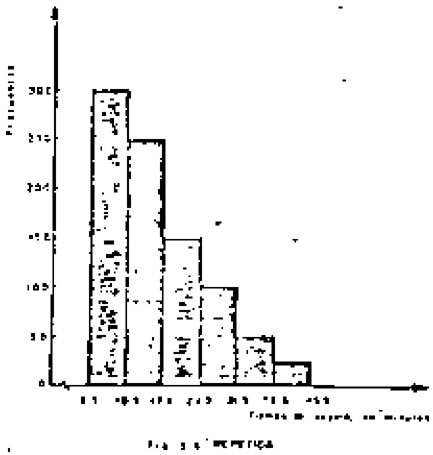
50 Cuando una distribución no es simétrica se dice que es *asimétrica*.
Observe las figuras 5.8 y 5.9.
¿Cuál de las dos distribuciones es asimétrica? __________.

La de la figura 5.8.

51 Se dice que una distribución es asimétrica cuando *no es* simétrica. Así, una distribución unimodal es asimétrica cuando no hay _____ con respecto al modo.

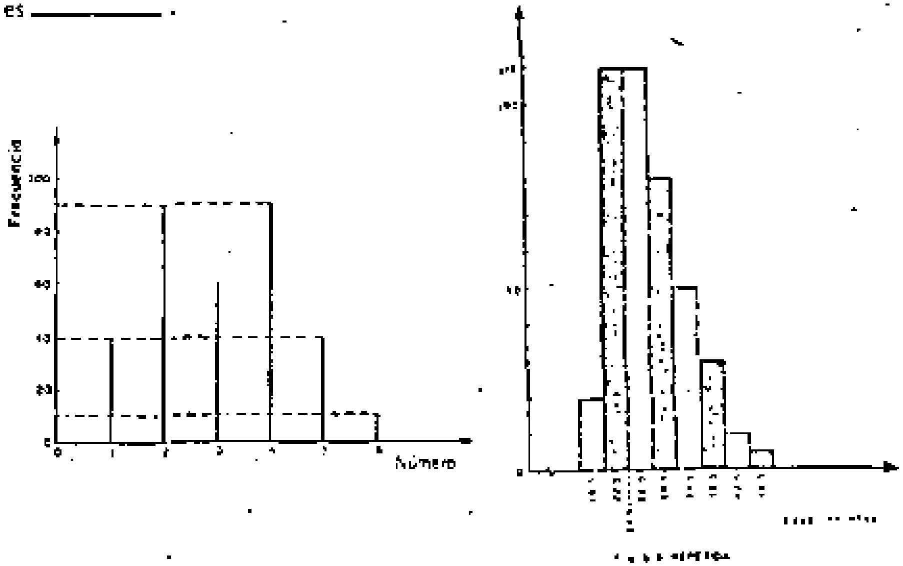
simetría

52 Observe la figura 5.6.
Esa distribución de frecuencias es __________
(asimétrica/simétrica)



asimétrica

53 La definición de simetría que hemos visto ahora se ha limitado a distribuciones unimodales. En general, se dice que una distribución es simétrica si hay simetría de las frecuencias respecto a *algún valor* de la variable.
Así, en la siguiente figura, en la que se observa una distribución bimodal, se observa que hay simetría respecto al número______ (1/3/4), por lo que la distribución es ______.



3, simétrica

54 Cuando una distribución de frecuencias tiene un solo modo y es simétrica, éste cae exactamente __________ (en el centro/en un extremo) de la distribución. Si tiene dos modos, éstos pueden quedar juntos al centro, o uno a cada lado del mismo.

en el centro

55 Puesto que el modo de la distribución unimodal de la figura 5.8 no cae en el centro, esa distribución no es __________, sino __________

simétrica, asimétrica

56 Las distribuciones asimétricas se conocen también con el nombre de *sesgadas*. Observe la figura 5.8. Esta distribución ________ (es/no es) sesgada porque está cargada hacia __________ (un lado/el centro)

es, un lado

57 Una distribución que es asimétrica se conoce también con el nombre de ______.

sesgada

58 Una distribución asimétrica o sesgada puede tener el modo hacia la derecha o hacia la ______ del centro de la distribución.

izquierda

PARTE B. LA MEDIANA

59 Existen otras medidas de posición o tendencia central. Una de ellas es la *mediana.*
Recuerde que la mediana es el valor de una variable para el cual el 50% de la distribución de datos agrupados queda a la izquierda y el ______ % a la ______

50, derecha

60 Si se tienen los siguientes datos *ordenados*:

14, 21, 23, *24*, 27, 30, 31

su mediana es 24, porque hay ______ (cuántos) datos a cada lado de ella.

3

61 Cuando los datos no están agrupados, pero sí ordenados, la mediana se define como el *valor central* de los datos.
Sea la secuencia de datos ordenados del cuadro anterior:

14, 21, 23, 24, 27, 30, 31

¿cuál es el valor central de estos datos? ______.

24

62 El dato cuyo valor es 24, es la mediana porque hay exactamente tres valores ______ (a un lado/a cada lado) de ella, es decir, tres valores son menores y tres son mayores que 24.

a cada lado

63 La regla para seleccionar el valor central de una serie de datos ordenados (no agrupados), varía ligeramente según se trate de un número *par* o *impar* de datos. Si hay un número impar, la ______ es precisamente el valor central.

mediana

64 ¿Cuál es la mediana de los siguientes datos?

3, 1, 8, 6, 2

3 (los datos ordenados son 1,2,3,6,8)

65 La mediana de los datos del cuadro anterior vale 3, porque al ordenar los datos nos queda

1, 2, 3, 6, 8

es decir, hay igual número de datos a ______ (un lado/cada lado) del valor 3.

cada lado

66 Entonces, para estimar la mediana de los datos básicos, lo primero que hay que hacer es ______ (ordenar/agrupar) los datos.

ordenar

67 Una vez que se han ordenado los datos, la mediana será el valor ______ (central/lateral) de los mismos.

central

68 Si el número de datos es par, la mediana es el valor que queda en medio de los dos datos centrales. En la serie ordenada de datos

4, 5, 7, 10, 14, 20, 21, 29

los dos datos centrales son ______ y ______.

10, 14

69 Puesto que los dos datos centrales del cuadro anterior son 10 y 14, el valor que queda en medio de éstos es ______.
(13/11/12)

12

---

70 El hecho de que sea 12 la mediana de la distribución del cuadro 68, la cual contiene 8 datos (un número par), significa que 4 datos son menores que 12 y 4 son ______.

mayores

---

71 Puesto que para calcular la mediana de una distribución es necesario, como primer paso, arreglar los datos en orden creciente o decreciente, la mediana no puede ser medida de tendencia central de datos de variables ______
(escalares/nominales)

nominales

---

72 La mediana es, entonces, una medida de posición o tendencia ______ para variables ______
(nominales/escalares)

central, escalares

---

73 La medida de tendencia o posición central apropiada para una variable aleatoria nominal es ______.
(la mediana/el modo/el rango)

el modo

---

74 Relacione los conceptos de la izquierda con las descripciones de la derecha.

a. Modo
b. Mediana

1. La clase de máxima frecuencia
2. La marca de clase del intervalo de máxima frecuencia
3. El valor de la variable para el cual la mitad de los datos son menores que él y la otra mitad son mayores

a-1,2, b-3

---

75 Una de las respuestas del cuadro anterior es a-1,2 porque el modo (a) es la clase de máxima frecuencia (1) cuando se trata de datos ______ (agrupados/no agrupados), y la marca de clase del intervalo de máxima frecuencia (2) cuando se trata de datos ______ en intervalos.
(agrupados/no agrupados)

no agrupados, agrupados

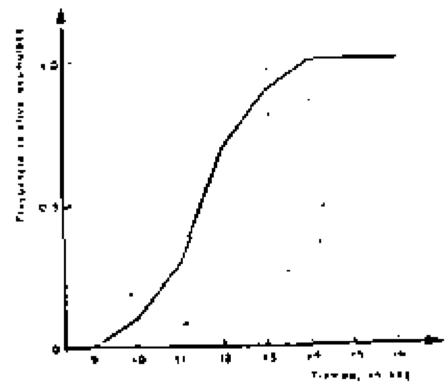
---

76 Si los datos están agrupados en intervalos de clase, la mediana es el valor de la variable para el cual el ______% de los valores de los datos son menores y el
(10/25/50)
otro ______% son mayores que ella.

50, 50

---

77 ¿Recuerda usted cómo se calcula la mediana a partir de una curva de distribución de frecuencias relativas acumuladas? Demuéstrelo calculando la mediana correspondiente a la siguiente distribución: ______.



11.6 seg.

---

78 Veremos ahora una forma de calcular la mediana a partir de una tabla de distribución de frecuencias acumuladas.
La mediana queda en el intervalo para el cual menos del 50% de los valores son menores que su límite real inferior de clase y menos del ______% de los valores son mayores que su límite real superior de clase.

50

79 Observe la distribución de frecuencias presentada en la tabla 5.1, correspondiente al número de bacterias por centímetro cúbico de agua contaminada. ¿Cuántos datos se utilizaron? _______

TABLA 5.1

Límites reales		Frecuencia	Frecuencia acumulada
Inferior	Superior		
[illegible]	55.5	3	3
55.5	60.5	5	8
60.5	65.5	6	14
65.5	70.5	11	25
70.5	75.5	13	38
75.5	80.5	9	47
80.5	85.5	[illegible]	56
85.5	90.5	[illegible]	[illegible]
90.5	100.5	4	72

72.

---

80 Si 72 es el número de datos, deberá haber 36 datos a cada lado de la _______

mediana

---

81 Observe la tabla 5.1.
La frecuencia acumulada hasta 70.5 es de _____ y la frecuencia acumulada hasta 75.5 es de _____.

25, 38

---

82 Puesto que estamos buscando el valor de la variable (la mediana) para el cual 36 datos son menores y 36 datos son mayores que él, dicho valor tendrá una frecuencia acumulada de _______.

36

---

83 Puesto que la frecuencia acumulada hasta 70.5 es 25 y hasta 75.5 es 38, la frecuencia acumulada de 36 (a la cual corresponde la mediana) queda en el intervalo de 70.5 a _______

75.5

---

84 Observe la tabla 5.1.
¿Cuál es la frecuencia del intervalo 70.5 - 75.5? _______.

13.

---

85 Observe la tabla 5.1.
¿Cuál es el ancho del intervalo de clase que contiene a la mediana, esto es, del intervalo 70.5-75.5? _______.

TABLA 5.1 [illegible]

Límites reales		Frecuencia	Frecuencia acumulada
Inferior	Superior		
[illegible]	55.5	3	3
55.5	60.5	5	8
60.5	65.5	6	14
65.5	70.5	11	25
70.5	75.5	13	38
75.5	80.5	9	47
80.5	85.5	11	56
85.5	90.5	6	64
90.5	100.5	8	72

5.

---

86 Para calcular el valor de la mediana necesitamos suponer que las 13 observaciones correspondientes al intervalo 70.5-75.5 están *uniformemente espaciadas* en ese intervalo de 5 unidades de amplitud.
Si 25 es la frecuencia acumulada hasta 70.5, ¿cuántos datos nos faltan para completar 36? _______.

11; (36−25=11).

---

87 Puesto que se supuso que los datos en el intervalo que contienen a la mediana están espaciados _______ (al azar/uniformemente), el valor de la mediana estará a 11/13 del camino de 70.5 a 75.5 (esto resulta de la regla de tres 5:13::d:11, de la cual se obtiene $d = \frac{11}{13} \times 5$, donde d denota la distancia de 70.5 a la mediana).

uniformemente

---

88 Puesto que la amplitud del intervalo 70.5-75.5 es de 5, 11/13 de esta amplitud es _______.

$\frac{55}{13}$; $\left(\frac{11}{13} \times 5 = \frac{55}{13} = 4.23\right)$

---

89 Redondeando a dos cifras decimales, 55/13 = 4.23. Entonces, la mediana es igual a 70.5 + _______ = 74.73.

4.23

90 Revisemos el procedimiento para calcular la mediana de una distribución de frecuencias acumuladas de datos agrupados.

1. Determine el número de puntuaciones que debe haber a cada lado de la mediana. Si N es el número total de datos, debe haber N/2 datos a cada lado de la ________.

mediana

---

91 Observe la tabla 5.1.
¿Cuántos datos debe haber a cada lado de la mediana de esa distribución?

36; (72/2=36).

---

92 2. Localice el intervalo de clase que contiene a la mediana.
Observe la tabla 5.1.
¿Cuál es el intervalo que contiene la mediana de esa distribución?

70.5-75.5.

---

93 3. Réstele al valor de N/2 la frecuencia acumulada hasta el límite real inferior del intervalo que contiene a la mediana. Si F_M denota dicha frecuencia, esta etapa consiste en calcular el valor de N/2—____.

F_M

---

94 Observe la tabla 5.1. ¿Cuánto vale la frecuencia acumulada hasta el límite real inferior del intervalo que contiene a la mediana, es decir, cuánto vale F_M?
Por lo tanto, $N/2-F_M$=________.

25, 11

---

95 4. Calcule el valor del cociente de $N/2 \cdot F_M$, entre la frecuencia del intervalo que contiene a la mediana.
Si f_M denota dicha frecuencia, esta etapa consiste en calcular el valor de $(N/2 - F_M)/(?)$.

$(N/2 - F_M)/f_M$

---

96 Observe la tabla 5.1.
¿Cuánto vale f_M, es decir, la frecuencia del intervalo que contiene a la mediana?
Recuerde que para esta distribución $N/2-F_M = 11$; ¿cuánto vale, entonces, $(N/2-F_M)/f_M$?

Tabla 5.1

Límites reales		Frecuencia	Frecuencia
Inferior	Superior		acumulada
50.5	55.5	3	3
55.5	60.5	5	8
60.5	65.5	6	14
65.5	70.5	11	25
70.5	75.5	13	38
75.5	80.5	9	47
80.5	85.5	11	58
85.5	90.5	6	64
90.5	100.5	8	72

13, 11/13

---

97 5. Multiplique $(N/2-F_M)/f_M$ por el ancho del intervalo de clase que contiene a la mediana. Si d_M denota este ancho, esta etapa consiste en calcular el valor de

$$\frac{N/2 - F_M}{f_M} \times (?)$$

$$\frac{N/2 - F_M}{f_M} d_M$$

---

98 Use la tabla 5.1.
¿Cuánto vale

$$\frac{N/2 - F_M}{f_M} \times (?)$$

para esa distribución? (Recuerde que $(N/2 - F_M)/f_M$ vale 11/13 y que d_M denota el ancho del intervalo de clase que contiene a la mediana.)

d_M, 55/13; $(\frac{11}{13} \times 5 = 55/13 = 4.23)$.

---

99 6. Súmele al límite real inferior del intervalo que contiene a la mediana, el valor de $(N/2 - F_M)/f_M \times d_M$. Si L_M denota dicho límite, esta etapa consiste en calcular el valor de (complete la fórmula).

$$(?) + \frac{N/2 - F_M}{(?)} d_M$$

$$L_M + \frac{N/2 - F_M}{f_M} d_M$$

---

100 Use la tabla 5.1.

¿Cuánto vale $L_M + \frac{N/2 - F_M}{f_M} \times d_M$

para esa distribución? (Recuerde que $\frac{N/2 - F_M}{f_M} d_M = 4.23$)

74.73;(70.5 + 4.23 = 74.73).

---

101 Relacione los símbolos de la columna de la izquierda con las descripciones de la derecha.

a. N — 1. Frecuencia acumulada hasta el límite real inferior del intervalo que contiene a la mediana

b. F_M — 2. Número total de datos en la muestra

a-2, b-1

---

102 Relacione los símbolos de la columna de la izquierda con las descripciones de la derecha.

a. f_M — 1. Frecuencia acumulada hasta el límite real inferior del intervalo que contiene a la mediana.

b. F_M — 2. Número de datos en la muestra

c. N — 3. Frecuencia del intervalo de clase que contiene a la mediana

a-3, b-1, c-2

---

103 Relacione los símbolos de la columna de la izquierda con las descripciones de la derecha.

a. f_M — 1. Ancho del intervalo ... contiene a la mediana

b. d_M — 2. Límite real inferior del intervalo que contiene a la mediana

c. L_M — 3. Frecuencia del intervalo que contiene a la mediana

a-3, b-1, c-2

---

104 El proceso descrito anteriormente, para calcular la mediana de una distribución de frecuencias, puede resumirse en una simple fórmula (obsérvela detenidamente).

$$\text{Mediana} = L_M + \frac{N/2 - F_M}{f_M} d_M$$

En esta expresión L_M denota el límite real ______________ (inferior/superior) del intervalo que contiene a la __________.

inferior, mediana

---

105 Diga cuál de los siguientes símbolos, d_M, L_M, F_M sirve para completar la fórmula para calcular la mediana.

$$\text{Mediana} = (?) + \frac{N/2 - F_M}{f_M} d_M$$

L_M

---

106 Complete la siguiente fórmula para calcular la mediana de una distribución de frecuencias.

$$\text{Mediana} = (?) + \frac{(?) - F_M}{f_M} d_M$$

$$\text{Mediana} = L_M + \frac{N/2 - F_M}{f_M} d_M$$

---

107 F_M es el símbolo que denota la frecuencia acumulada hasta el límite real ______ del intervalo que contiene a la ________.

inferior, mediana

---

108 Complete la fórmula para calcular la mediana.

$$\text{Mediana} = L_M + \frac{N/2 - (?)}{f_M}\ (?)$$

$$\text{Mediana} = L_M + \frac{N/2 - F_M}{f_M}\ d_M$$

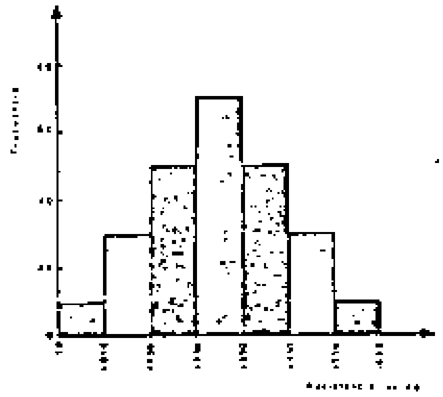
---

109 Use la figura 5.9.

a. ¿Cuál intervalo contiene a la mediana?

b. ¿Cuánto vale N?

c. ¿Cuánto vale F_M?



a. 1250-1350 kg, b. 250, c. 90 kg.

---

110 f_M denota la ____________ del intervalo de clase que contiene a la mediana.

frecuencia

---

111 Use la figura 5.9.

¿Cuánto vale f_M?

70 kg.

---

112 Complete la siguiente fórmula para calcular la mediana de __s agrupados en intervalos de clase.

$$\text{Mediana} = L_M + \frac{(?) - F_M}{(?)}\ d_M$$

$$\text{Mediana} = L_M + \frac{\frac{N}{2} - F_M}{f_M}\ d_M$$

---

113 El símbolo d_M denota el ________ del intervalo de clase que contiene a la mediana.

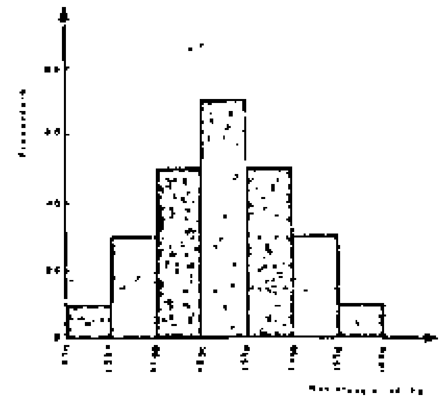
ancho

---

114 Use la figura 5.9.

a. ¿Cuánto vale d_M?

b. ¿Cuánto vale L_M?



a. 100 kg. b. 1250 kg.

---

115 Escriba la fórmula para calcular la mediana de datos agrupados en intervalos de clase.

$$\text{Mediana} = L_M + \frac{\frac{N}{2} - F_M}{f_M} \times d_M$$

---

116 Si $N = 250$, $F_M = 90$, $f_M = 70$, $d_M = 100$ y $L_M = 1250$. ¿Cuánto vale la mediana?

1300 kg ($1250 + \frac{125-90}{70} \times 100 = 1250 + 50 = 1300$).

---

117 Observe la distribución ______________ de la figura 5.9.
(simétrica/asimétrica)
¿Cuánto vale el modo de esa distribución? ______________.

simétrica, 1300 kg.

---

118 Para la distribución simétrica de la figura 5.9, el modo vale 1300 y la mediana *también* vale 1300. En general, para toda distribución ______________ de
(simétrica/asimétrica)
frecuencias unimodal, el modo y la mediana tienen el *mismo* valor.

simétrica

---

119 En una distribución de frecuencias simétrica y unimodal, el modo y la mediana tienen ______________ valor. Si no es unimodal, pero es simétrica, la me-
(el mismo/diferente)
diana coincidirá con el punto de simetría.

el mismo

---

120 Hasta ahora hemos descrito dos medidas de tendencia central; ¿cuáles son? ______.

la mediana y el modo.

PARTE C. LA MEDIA

121 Describiremos ahora otra medida de tendencia central, pero antes se introducirá la notación que necesitamos.
Sea X el símbolo para denotar el valor que toma una variable aleatoria al realizar un experimento.
Si al realizar un experimento el resultado es 37, decimos que X = ______.

37

---

122 Si al realizar un experimento se obtiene una observación 8, decimos que X = ______.

328

---

123 El símbolo que utilizamos para denotar el valor que toma la variable es ___.

X

---

124 Si al repetir dos veces el mismo experimento las observaciones son 10.1 y 10.9, decimos que para el primer resultado X = ______ y para el segundo X = ______.

10.1, 10.9

---

125 Para aclarar que el *primer* resultado fue X = 10.1, ponemos un subíndice 1 a X, así $X_1 = 10.1$. Para el *segundo* resultado usamos un subíndice 2, por lo cual escribiremos ______ = 10.9.

X_2

---

126 Este proceso se puede generalizar si *numeramos* las observaciones; para el primer resultado escribimos X_1, para el segundo X_2, etc. ¿Cómo indicaremos la novena observación? ______________.

X_9

---

127 Si necesitamos un término genérico para denotar *cualquier* elemento de la muestra, usaremos un subíndice i para las X; este subíndice podrá tomar valores desde 1 hasta N, donde *N* es el ________ total de elementos de la muestra.

número

---

128 Si tenemos $X_7 = 9.9$, entonces i = ______.

7

---

129 Si tenemos $X_{16} = 22.3$, entonces i = ____.

130 Si $i = 31$, entonces $X_{(\ )} = 22.3$.

31; (X_{31})

---

131 Si queremos sumar *todos* los elementos de la muestra escribiremos

$$\text{suma} = X_1 + X_2 + \ldots + X_N$$

donde X_N es el ______________ elemento de la muestra.
(primer/último)

último

---

132 Si se tiene la siguiente muestra: 21,18,17,17,20,24, entonces $X_1 = 21$, $X_2 = 18$, $X_3 = 17$, $X_4 =$ ______, $X_5 =$ ______ y $X_6 =$ ______.

17, 20, 24

---

133 Para no escribir explícitamente $X_1 + X_2 + \ldots + X_N$, se utiliza el *símbolo de suma* que es la S del alfabeto griego, la cual se escribe así: Σ. Entonces

$$\sum_{i=1}^{N} X_i = X_1 + X_2 + \ldots + X_N$$

El símbolo Σ sirve para indicar una ______________.
(suma/resta/división)

suma

---

134 Al simbolizar una suma en la forma

$$\text{suma} = \sum_{i=1}^{N} X_i$$

ponemos debajo del símbolo Σ una anotación que indica el valor de i, por el cual se comienza la suma; arriba del símbolo ponemos el máximo valor que tomará i en esa suma. Así, en la fórmula anterior el primer valor de i es ________.
(1/N/10)

135 El símbolo Σ se llama *sigma*, la *sigma* sirve, entonces, para indicar una ______.

suma

---

136 La notación $\sum_{i=1}^{N}$ significa que en la suma se incluyen las X_i con índices del 1 al ______.

N

---

137 Si se tiene la muestra: 21, 18, 17, 17, 20 y 24 ($N = 6$), entonces

$$\sum_{i=1}^{6} X_i = \_\_\_\_ + 18 + \_\_\_\_ + 17 + \_\_\_\_ + 24$$

21, 17, 20

---

138 ¿Cómo denotaría la suma de las X_i, con i desde 1 hasta 27?

$\sum_{i=1}^{27} X_i$

---

139 La otra medida de tendencia o posición central que estudiaremos en esta unidad se llama media. Entonces, las tres medidas de tendencia central que se estudian en este libro son ______________, ______________ y ______________.

el modo, la mediana, la media (En cualquier orden.)

---

140 La media es una medida de posición o tendencia __________.

central

---

141 La media se define como

$$\text{media} = \frac{\sum_{i=1}^{N} X_i}{N}$$

donde $\sum_{i=1} X_i$ significa que se suman ______ (sólo algunos/todos) los valores de los datos de la muestra.

todos

---

142 Si se tiene una muestra con seis elementos entonces

$$\text{media} = \frac{\sum_{i=1}^{(?)} X_i}{(?)}$$

6, 6

---

143 Si la muestra de 6 elementos es: 21,18,17,17,20 y 24, entonces

$$\text{media} = \frac{\sum_{i=1}^{6} X_i}{6} = \frac{21 + 18 + 17 + \_\_ + \_\_ + \_\_}{6}$$

17, 20, 24

---

144 La media se define como

$$\text{media} = \frac{\sum_{i=1}^{N} (?)}{(?)}$$

$$\text{media} = \frac{\sum_{i=1}^{N} X_i}{N}$$

---

145 Si usamos una equis con barra (tilde), $\overline{X}$, para denotar la media, entonces

$$\overline{X} = \frac{(?)}{(?)}$$

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

---

146 La media se conoce también con los nombres de *valor medio* y *medio aritmético* (al decir "promedio" nos referiremos al promedio aritmético).
¿Cuál es la fórmula para calcular el promedio? ______

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

---

147 Valor medio y promedio aritmético (o promedio) son sinónimos de la medida de tendencia central llamada ______ (modo/mediana/media).

media

---

148 De los siguientes conceptos, ¿cuáles son sinónimos de la media?

a. Rango
b. Promedio
c. Promedio aritmético
d. Población
e. Modo
f. Valor Medio ______

b, c, f

---

149 ¿Cuál es la fórmula para calcular el valor medio de los datos de una muestra?
______

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

---

150 Se obtuvo una muestra de 50 tiempos, requeridos por los estudiantes de nuevo ingreso en la Universidad, para resolver una prueba de "retención de conceptos". La suma de los 50 tiempos vale 3 150 minutos.
¿Cuánto vale el promedio de estos tiempos? ______

$\overline{X} = 63$ min ($\frac{3150}{50} = 63$).

---

151 Entonces, para calcular el promedio de un grupo de datos, lo que hay que hacer es ________________ todos los datos y el resultado dividirlo entre _____.
(sumar/restar/dividir)

sumar, N (número total de datos)

---

152 Si los datos están agrupados, siendo f_i la frecuencia del valor X_i, entonces, un número f_i de elementos tiene un valor igual a ________.
$(f_i/X_i/\overline{X})$

X_i

---

153 Use la tabla 5.2. Si aplicamos a los datos de esta tabla la fórmula para calcular la media tendremos:

$$\overline{X} = \frac{3+3+4+6+6+6+\ldots}{15}$$

TABLA 5.2

X	Frecuencia
3	2
4	1
6	3
8	5
10	4

es decir, sumamos dos veces 3, una vez 4, ________ (cuántas) veces 6, ________ veces 8 y cuatro veces ________.

tres, cinco, 10

---

154 Al calcular la media de los datos de la tabla 5.2, sumamos tres veces 6 porque la frecuencia de este valor es ________ (observe la tabla).

3

---

155 El sumar tres veces 6 es equivalente a hacer la multiplicación 6 x 3, es decir, a multiplicar el valor de la variable por su ____________ correspondiente.

frecuencia

---

156 Entonces, el cálculo del promedio de datos agrupados puede simplificarse si, en vez de sumar todos los datos, sumamos los productos de cada dato por su ____________ correspondiente.

frecuencia

---

157 Use la tabla 5.2.
El promedio de datos agrupados en la tabla, puede calcularse en la forma

$$\overline{X} = \frac{3X2+4X1+6X\_\_\_ + \_\_\_\_\_\_ + \_\_\_\_\_\_}{15}$$

TABLA 5.2

X	Frecuencia
3	2
4	1
6	3
8	5
10	4

3, 8 X 5, 10 X 4

---

158 Si f_i denota la frecuencia de X_i, la fórmula para calcular la media de datos agrupados se simplifica a

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

donde K es el número total de clases.
La frecuencia de X_4 se denota con f_4. La frecuencia de X_{10} se denota con ____.

f_{10}

---

159 Observe la tabla 5.2.
¿Cuánto vale K (número de clases)?

6.

---

160 Complete la siguiente fórmula para el cálculo del promedio de los datos agrupados en K intervalos.

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{(?)}$$

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i \ f_i}{N}$$

161 El número total de intervalos lo indicamos con la letra _____ y el número total de datos con la _____.

K, N

162 Complete la siguiente fórmula

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i (?)}{(?)}$$

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

163 f_i es la _____________ de X_i.

frecuencia

164 Use la tabla 5.2.
La media de los datos en la tabla 5.2 vale

$$\overline{X} = \frac{\sum_{i=1}^{5} X_i \ f_i}{N} = \frac{3X2+4X1+\_\_\_\_+\_\_\_\_+\_\_\_\_}{15} = 7.2$$

6 X 3, 8 X 5, 10 X 4

165 El cálculo del promedio de datos agrupados se simplifica si a la tabla de distribución de frecuencias le agregamos una columna con encabezado "Xf" (significa X_i por f_i).

Complete la siguiente tabla.

X	f	Xf
3	2	6
4	1	4
6	3	
8	5	
10	4	
TOTAL:	15	108

18, 40, 40

166 Puesto que la suma de todas las $X_i \ f_i$ es 108 y se tienen 15 datos, la media vale

$$\overline{X} = \frac{(?)}{(?)} = \_\_\_\_$$

$$\frac{108}{15} = 7.2$$

167 Puesto que para calcular la media es necesario efectuar operaciones numéricas, ésta se puede obtener sólo para variables _____________ (escalares/nominales).

escalares

168 La mediana, la media y el modo son medidas de posición o tendencia _________ de variables escalares. De éstas, *solamente* el modo puede utilizarse para variables _____________ (escalares/nominales).

central, nominales

169 ¿Cuáles son las tres medidas de tendencia central que hemos estudiado?

modo, mediana, media (en cualquier orden).

170 Relacione los nombres de la columna de la izquierda con las descripciones dadas en la de la derecha.

a. Mediana
b. Media
c. Modo

1. $\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$
2. La marca de clase del intervalo de mayor frecuencia
3. $\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$
4. El valor de la variable para el cual 50% de los datos tienen un valor menor y el 50% tienen un valor mayor que él.

a-4, b-1,3, c-2

171 Si los datos están agrupados en intervalos de clase, al calcular la media se les asigna a las *marcas de clase* las *frecuencias* de clase de los intervalos correspondientes.
Observe la figura 5.8.
Según este criterio, la frecuencia de la marca de ________ 34.5 vale ____

clase, 50

172 Al calcular la media de datos agrupados en intervalos, se supone que la ____________________ de un intervalo corresponde a su marca de clase.

frecuencia

173 Use la figura 5.8.
Si queremos calcular la media de esa distribución, las frecuencias de los primeros tres intervalos corresponderán a los valores de la variable: ________, ____________ y ____________ respectivamente.

18.5, 22.5, 26.5

174 Puesto que a las marcas de clase se les asocian las frecuencias de los intervalos correspondientes, la media de los datos agrupados en intervalos la podemos calcular mediante la segunda de las dos fórmulas estudiadas anteriormente:

$$\overline{X} = \frac{\sum_{i=1}^{k} f_i X_i}{N}$$

donde las X_i denotan ahora las __________ de clase.

marcas

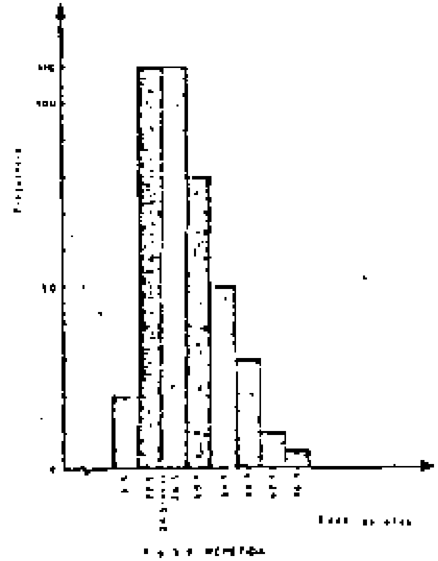
175 Si en la fórmula

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

las X_i denotan las marcas de clase, entonces las f_i denotan las ____________ de los intervalos de clase correspondiente.

frecuencias

176 Use la figura 5.8.
Complete el cálculo de la media para esa distribución de frecuencias



$$\overline{X} = \frac{18.5\times20+22.5\times110+26.5\times110+30.5\times80+\_\_\_\_+\_\_\_\_+\_\_\_\_+46.5\times5}{415}$$

34.5X50, 38.5X30, 42.5X10

177 La media de los datos agrupados en intervalos de clase se calcula en forma sistemática así (completa la tabla):

Marca de clase, X_i en años	Frecuencia, f	Xf en años
18.5	20	
22.5	110	2475
26.5	110	2915
30.5	80	2440
34.5	50	1725
38.5	30	1155
42.5	10	
46.5	5	
TOTAL:	$N = \sum f_i = 415$	$\sum X_i f_i = 11737.5$

370, 425, 232.5

---

178 Si $\sum_{i=1}^{8} X_i f_i = 11737.5$ años y $N = 415$, entonces $\overline{X} = \frac{(?)}{(?)} = 28.28$ años.

$$\frac{11737.5}{415}$$

---

179 Se dispone de la distribución de frecuencias correspondiente al tiempo que un individuo tarda para reaccionar a ciertos estímulos sicológicos. Anote en la tabla el encabezado que falta, complétela y calcule la media.

Marca de clase, X_i en seg.	f	(?), en seg
0.10	2	
0.15	7	
0.20	14	
0.25	4	
0.30	3	
TOTAL:	N= ____	$\sum_{i=1}^{5} X_i f_i =$ ____ seg

$$\overline{X} = \frac{(?)}{(?)} = 0.198 \text{ seg}$$

Marca de clase, X_i en seg	f	Xf, en seg
0.10	2	0.20
0.15	7	1.05
0.20	14	2.80
0.25	4	1.00
0.30	3	0.90
TOTAL:	N = 30	$\sum_{i=1}^{5} X_i f_i = 5.95$

$$\overline{X} = \frac{5.05}{30} = 0.198 \text{ seg}$$

REVISION

180 En la siguiente tabla calcule el promedio aritmético de los datos agrupados correspondientes a la resistencia de los alambres de acero producidos por una fábrica.

Marca de clase, X_i en kg	f	(?) en kg
40	2	
45	4	
50	5	
55	10	
60	4	
TOTAL:		

Marca de clase, X_i en kg	f	Xf, en kg
40	2	80
45	4	180
50	5	250
55	10	550
60	4	240
TOTAL:	25	1300

$$\overline{X} = \frac{\sum_{i=1}^{5} X_i f_i}{25} = \frac{1300}{25} = 52 \text{ kg}$$

---

181 La distribución de frecuencias del cuadro anterior fue

Marca de clase, X, en kg	f
40	2
45	4
50	5
55	10
60	4

¿Cuánto vale el *modo*?

55.

---

182 La fórmula para calcular la mediana de datos agrupados en intervalos de clase es

$$\text{Mediana} = (?) + \frac{(?) - (?)}{(?)} \times (?)$$

$$\text{Mediana} = L_M + \frac{N/2 - F_M}{f_M} \times d_M$$

---

183 La distribución de frecuencias del cuadro 180 fue

Marca de clase, X, en kg	f
40	2
45	4
50	5
55	10
60	4
TOTAL :	25

Para éste se tiene:

Límites reales del intervalo que contiene a la mediana (recuerde que los límites reales se obtienen sumándole y restándole a la marca de clase la mitad del ancho del intervalo correspondiente): ________; N = ________; L_M = ________; F_M = ________; d_M = ________.

52.5 y 57.5, N = 25, L_M = 52.5, f_M = 10, F_M = 11, d_M = 5

---

184 Para los datos del cuadro anterior se obtuvo N= 25, L_M=52.5, f_M=10, F_M = 11 y d_M= 5.

¿Cuánto vale la mediana correspondiente?

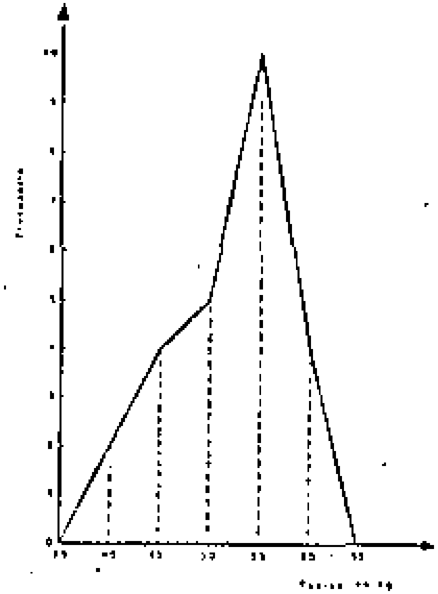
$$53.25; \left(52.5 + \frac{\frac{25}{2} - 11}{10}\, 5 = 53.25\right)$$

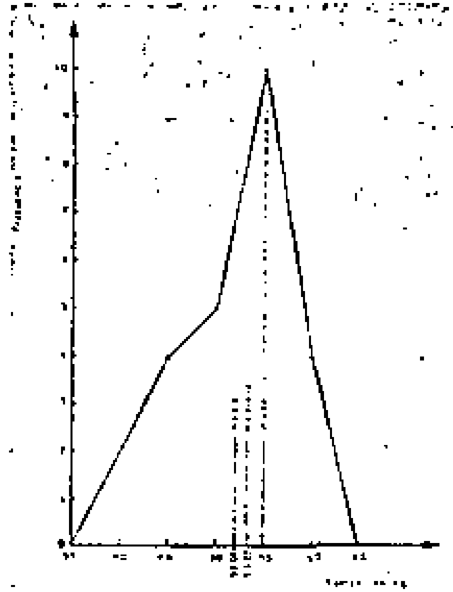
---

185 En los cuadros anteriores se obtuvo:

modo = 55 kg
mediana = 53.25 kg
media = 52.00 kg

a. Marque estos valores en el polígono de frecuencias presentado en la figura 5.10.
b. ¿Es simétrica esta distribución de frecuencias?





b No.

186 Cuando una distribución de frecuencias no es simétrica, es decir, cuando es sesgada, no coinciden los valores de las tres medidas de tendencia central aquí estudiadas, a saber, __________, __________ y __________.

modo, mediana, media (en cualquier orden).

---

187 Mientras más sesgada sea una distribución de frecuencias, mayor diferencia existe entre la media, la mediana y el modo.
Si la distribución tiene un solo modo (unimodal) y es simétrica, entonces estas tres medidas de tendencia central adquieren __________ valor.
(el mismo/diferente)

el mismo

---

188 El modo es la única medida de tendencia central que se puede utilizar cuando se tienen datos de variables __________.
(escalares/nominales)

nominales

---

189 La media es la medida de posición o __________ __________ que más se utiliza cuando se tienen distribuciones de frecuencias que no sean demasiado asimétricas.

tendencia central

---

190 La mediana es la medida de tendencia central que se utiliza con más frecuencia cuando se tienen distribuciones francamente asimétricas (o sesgadas).
Mientras más sesgada es una distribución, la diferencia entre la media, la mediana y el modo es __________.
(menor/mayor)

mayor

---

191 Calcule el modo, la mediana y la media de los siguientes números:

2, 2, 4, 5, 5, 5, 6, 7

__________, __________, __________.

modo = 5 (la frecuencia de 5 es la máxima)
mediana = 5 (los dos datos centrales valen 5)
media = 4.5; $\left(\frac{2\times 2+4+5\times 3+6+7}{8} = 4.5\right)$

J

EXAMEN

1. *¿Cuáles son las medidas de tendencia o posición central que se estudiaron?* __________ __________.

2. *El modo de una distribución de frecuencias es el valor de la variable cuya frecuencia es la* __________.

3. *Si los datos están agrupados en intervalos de clase, el modo es la* __________ __________ *del intervalo de máxima frecuencia.*

4. *Cuando los intervalos a la izquierda del modo tienen igual frecuencia que los situados a igual distancia a la derecha, se dice que la distribución de frecuencias es* __________, *en caso contrario es* __________.

5. *¿Cuál es el modo de la siguiente distribución de frecuencias? ¿Es simétrica?* __________, __________.

[illegible]	[illegible]
[illegible]	3
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
90	5

6. *¿Cuál es la mediana de los siguientes datos: 2, 8, 3, 5, 9, 11, 1?* __________.

7. *¿Cuál es la mediana de los siguientes datos: 2, 8, 3, 5, 9, 11, 1, 7?* __________.

8. *¿Cuándo coinciden la mediana, el modo y la media; cuando la distribución es simétrica o cuando es asimétrica?* __________.

9. *¿Cuál es la mediana de la distribución de frecuencias presentada en la pregunta 5?* __________.

10. *¿En qué intervalo de la siguiente distribución de frecuencias queda la mediana?* __________.

Límites reales	Frecuencia
20 - 30	5
30 - 40	7
40 - 50	10
50 - 60	4
60 - 70	2
TOTAL:	28

11. *La fórmula para calcular la mediana de datos agrupados en intervalos de clase es:*

$$mediana = L_M + \frac{N/2 - F_M}{f_M}\, d_M$$

Para la distribución de frecuencias de la pregunta anterior, cuánto vale:

a. f_M ________________.
b. $N/2$ ________________.
c. F_M ________________.
d. f_M ________________.
e. d_M ________________.
f. la mediana

12. *Escriba la fórmula para calcular la media de datos no agrupados.* ____________.

13. *Calcule el promedio aritmético de los siguientes datos: 2, 8, 3, 5, 9, 10, 11, 1, 7, 6.* ________________.

14. *Escriba la fórmula para calcular la media de datos agrupados.* ____________.

15. *Cuando los datos están agrupados en intervalos de clase, las X_i que aparecen en la fórmula para calcular $\overline{X}$ son las* ________ ________ ________ *de los intervalos.*

16. *Calcule el valor medio de la siguiente distribución de frecuencias:* ____________.

Límites reales	Frecuencia
20 – 30	5
30 – 40	7
40 – 50	18
50 – 60	8
60 – 70	2

________________.

17. *¿Cuál es la única medida de tendencia central que se puede utilizar para datos de variables nominales?* ________________.

18. *¿Cuál es el modo de la siguiente distribución de frecuencias?* ________________.

Nacionalidad	Frecuencia
[illegible]	30
[illegible]	[illegible]
[illegible]	375
[illegible]	45
China	131

19. *La diferencia entre la mediana, la media y el modo de una distribución de frecuencias es mayor mientras más* ____________ *sea.*

TOTAL: 28 puntos

RESPUESTAS

Modo, mediana, media.

máxima (mayor)

marca de clase

simétrica
asimétrica (sesgada)

60
Sí.

5

6

Cuando es simétrica

60

En el intervalo 40-50.

a. 40
b. 18
c. 12
d. 18
e. 10
f. $\text{mediana} = 40 + \dfrac{18-12}{18} \times 10 = 43.3$

12. $$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

13. $$\overline{X} = \frac{62}{10} = 6.2$$

14. $$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

15. *marcas de clase*

16. $\bar{X} = \dfrac{5X25 + 7X35 + 18X45 + 4X55 + 2X65}{36} = 42.5$

17. *El modo*

18. *Hebrea*

19. *asimétrica*

UNIDAD VI MEDIDAS DE DISPERSION

PREFACIO

En esta unidad estudiaremos cómo determinar el grado de dispersión o variabilidad de los datos respecto a alguna de las medidas de tendencia central. Es decir, trataremos de conocer si los datos están muy esparcidos, mediante los criterios que llamaremos *medidas de dispersión*.

PARTE A. EL RANGO

1 **¡Recuerde!**
El *rango* de una muestra es igual a la diferencia entre el dato de máximo valor y el de __________ valor.

mínimo (o menor)

---

2 Use la tabla 6.1.
¿Cuánto vale el rango de los datos ordenados de las calificaciones en Pedagogía? __________.

TABLA 6.1

Calificaciones en Pedagogía, X			Calificaciones [illegible], Y		
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
$\bar{X}$ = [illegible]			$\bar{Y}$ = [illegible]		

42; (99 − 57 = 42).

3 Use la tabla 6.1.
¿Cuánto vale el rango de los datos ordenados de las calificaciones en Historia?

28; (100 − 72 = 28).

---

4 Recuerde que $\overline{X}$ denota la media de los datos de la variable X. De la misma manera, $\overline{Y}$ denota la __________ de los datos de la variable ______.
Observe la tabla 6.1.
¿Cuánto valen los promedios (las medias) de las calificaciones en Pedagogía y en Historia?

TABLA 6.1 REPETIDA

Calificaciones en Pedagogía X			Calificaciones en Historia Y		
[illegible]	77	97	72	[illegible]	[illegible]
59	[illegible]	[illegible]	72	77	[illegible]
65	79	[illegible]	72	77	[illegible]
67	81	89	74	77	[illegible]
67	81	91	[illegible]	77	[illegible]
57	83	91	75	78	[illegible]
[illegible]	83	93	75	78	[illegible]
[illegible]	83	[illegible]	[illegible]	79	[illegible]
73	84	97	77	79	[illegible]
73	85	99	[illegible]	79	[illegible]
$\overline{X}$ = 80.00			$\overline{Y}$ = 80.0		

media, Y, 80.0, 80.0.

---

5 Los dos grupos de datos, presentados en la tabla 6.1, tienen __________ (diferente/la misma) *media* pero __________ (el mismo/diferente) rango.

la misma, diferente

---

6 La media nos da una idea de la tendencia __________ de los datos; el rango, en cambio, nos la da de la variabilidad o dispersión de los datos. A mayor rango mayor variabilidad; a menor __________ menor variabilidad.

central, rango

---

7 Los dos grupos de datos de la tabla 6.1 tienen el mismo promedio pero diferente rango, siendo menor el de las calificaciones en Historia; éstas tienen, por lo tanto, __________ (menor/mayor) *variabilidad* que las calificaciones en Pedagogía, es decir, los datos están __________ (más/menos) dispersos.

menor, menos

---

8 Puesto que el rango nos da una idea de la variabilidad de los datos, lo consideraremos como una medida de __________ (tendencia central/dispersión).

dispersión

---

9 Observe la tabla 6.1.
Si la penúltima calificación en Historia hubiese sido 99 en vez de 91, la media hubiera __________ (disminuido/aumentado/permanecido igual), y el rango hubiera __________ (disminuido/aumentado/permanecido igual).

aumentado, permanecido igual

---

10 El rango no se altera cuando se cambia alguno de los datos __________ (extremos/intermedios), pero varía cuando se cambia cualquiera de los valores __________ (intermedios/extremos) (el mínimo o el máximo).

intermedios, extremos

---

11 En vista de lo anterior, el *rango*, que es una medida de __________ (dispersión/tendencia central), es insensible a cambios de los valores intermedios de los datos y muy sensible a los cambios de los valores __________.

dispersión, extremos

---

12 El *rango* es una medida de __________.

dispersión

---

13 ¿Alguno de los siguientes conceptos es una medida de dispersión?

a. Media
b. Mediana
c. Modo

No.

14 ¿Cuál es la medida de dispersión que hemos estudiado hasta ahora?

El rango.

---

15 El rango es sensible a cambios en los valores ________________ de los datos,
(intermedios/extremos)
es decir, a cambios de los valores mínimo o máximo.

extremos

---

16 Puesto que el rango es sensible a cambios en los valores extremos, posiblemente si aumentamos el tamaño de la muestra (el número de datos), el rango ________
(disminuya/aumente)

aumente

---

17 Observe la tabla 6.1.
A simple vista, ¿en cuál grupo de datos hay mayor concentración de valores alrededor del promedio?

TABLA 6.1 [illegible]

Calificaciones en Pedagogía			Calificaciones en Historia		
57	77	87	72	77	81
58	79	88	72	77	81
65	78	89	72	77	81
57	[illegible]	99	74	77	[illegible]
67	81	91	75	77	83
63	82	91	75	78	86
[illegible]	83	93	75	78	90
72	83	95	75	79	91
73	84	97	77	79	91
74	85	99	77	79	100
$\bar{x} = 80.00$			$\bar{x} = 80.0$		

En el de calificaciones en Historia.

---

18 Si en la lista de datos de calificaciones en Pedagogía conservamos el dato de mínimo valor y aumentamos el valor de los cinco o seis datos siguientes, el rango tendrá ________________ valor pero habrá mayor concentración de datos
(el mismo/diferente)
alrededor de la media.

el mismo

---

19 Puesto que el rango no se altera cuando se cambian algunos valores de los datos intermedios, en tal forma que haya mayor o menor concentración de valores alrededor de la media, el rango se considera como una medida ____________ de
(exacta/burda)
____
(dispersión/tendencia central)

burda, dispersión

---

20 El rango es sólo una medida burda de ____________ porque es sensible a cambios de los valores ____________ e insensible a cambios de los valores ____________.

dispersión, extremos, intermedios

---

21 Aunque el ____________ es sólo una medida burda de dispersión, proporciona rápidamente una idea de la variabilidad de los ____________.

rango, datos

---

22 Puesto que el rango es insensible a cambios de los valores intermedios, es una medida ________________ de dispersión.
(exacta/inexacta)

inexacta

---

23 El rango es una medida burda de dispersión ya que ________ sensible a cambios
(no es/es)
de los valores extremos y, por lo tanto, al tamaño de la muestra (número de datos).

es

PARTE B. LA VARIANCIA Y LA DESVIACION ESTANDAR

24 En virtud de que el rango es una medida burda de dispersión, se ve la necesidad de tener una medida más precisa que dependa de ________________ los datos
(algunos de/todos)
de la muestra y no solamente de los valores máximo y mínimo, como sucede con el rango.

todos

---

25. Una medida de dispersión que toma en cuenta todos los datos, es la llamada variancia.

La variancia es una medida de ______________.

dispersión

---

26 El ________ y la variancia son medidas de ____________.

rango, dispersión

---

27 El cálculo de la variancia sigue la secuencia:

1. *Calcular las desviaciones o diferencias de cada dato con respecto a su media.*

Si $\overline{X} = 7$ y el primer dato es $X_1 = 3$, su desviación con respecto a la media vale $X_1 - \overline{X} = 3 - 7 = -4$. Si el segundo dato vale 4, su ______________ con respecto a la media es $X_2 - \overline{X} =$ ______ $-$ ______ $=$ ______.

desviación, $4 - 7 = -3$

---

28 Complete la siguiente tabla ($\overline{X}$=7).

x	Desviaciones, $X - \overline{X}$
3	-4
4	-3
5	
8	
9	
12	

-1, 1, 2, 5

---

29 La tabla completa del cuadro anterior es:

x	Desviaciones, $X - \overline{X}$
3	-4
4	-3
6	-1
8	1
9	2
12	5

¿Cuánto vale la suma de todas las desviaciones $X_i - \overline{X}$?

Cero.

---

30 Puesto que las desviaciones se miden con respecto a un valor central (la media), siempre se tendrá que algunas de éstas resultan positivas y las demás ________.

negativas

---

31 Al sumar valores positivos y negativos, el resultado no necesariamente es cero. Sin embargo, se puede demostrar que la suma de todas las desviaciones con respecto a la __________ es igual a cero.

media

---

32 A $X_i - \overline{X}$ se le denomina ______________ de X_i con respecto a la __________.

desviación, media

---

33 Si los datos están agrupados, para calcular las desviaciones será necesario multiplicar la desviación de cada dato por su ______________.
(rango/frecuencia)

frecuencia

---

34 Si la media de unos datos agrupados es 7.2 y la frecuencia de un dato, cuyo valor es 3, es igual a 2, la desviación total con respecto a la media es $(3 - 7.2) \times 2 = -4.2 \times 2 = -8.4$.

Si 3 es la frecuencia del valor 6, su desviación vale $(6 - 7.2) \times$ ______ $=$ __________.

3, -3.6

---

35 Complete la siguiente tabla ($\overline{X}$=7.2). Recuerde que la letra f denota frecuencia.

x	f	$x - \overline{x}$	$(x - \overline{x})\ f$
3	2	-4.2	-8.4
4	1	-3.2	-3.2
5	3	-1.2	
8	5		
10	4		

$x - \overline{x}$	$(x - \overline{x})\ f$
-4.2	-8.4
-3.2	-3.2
-1.2	-3.6
0.8	4.0
2.8	11.2

36 Observe la respuesta correcta del cuadro anterior. ¿Cuánto vale la suma:

$\sum_{i=1}^{5}(X_i - \overline{X}) f_i$? ______________.

Cero.

---

37 Si los datos están agrupados en intervalos de clase, al calcular la desviación con respecto a la media se toma como X_i a la *marca de clase* del iésimo ________.

intervalo

---

38 Si los datos están agrupados en intervalos, las X_i serán las marcas de clase y las f_i las ____________ de los intervalos correspondientes.

frecuencias

---

39 Complete la siguiente tabla correspondiente a la edad en que murieron 25 personas de fiebre tifoidea. Sólo se consideraron edades entre 40 y 60 años ($\overline{X}$=52).

Edad, X, en años (marca de clase)	f	$x - \overline{x}$	$(x - \overline{x}) f$
40	2		
45	4		
50	5		
55	10		
60	4		

$x - \overline{x}$	$(x - \overline{x}) f$
−12	−24
−7	−28
−2	−10
3	30
8	32

---

40 Observe la respuesta correcta del cuadro anterior. ¿Cuánto vale $\sum_{i=1}^{5}(X_i - \overline{X}) f$? ______________.

Cero.

---

41 En todos los ejemplos que hemos presentado se ha verificado que la suma de *todas* las desviaciones con respecto a la *media* es ________.

cero

---

42 Puesto que la suma de todas las desviaciones con respecto a la media es cero, su promedio también vale ________, ya que éste se calcula como el cociente de la suma de las desviaciones, entre el número total de datos.

cero

---

43 Puesto que el promedio aritmético de todas las desviaciones con respecto a la media es siempre ________, no puede usarse este promedio como una medida de dispersión.

cero

---

44 Para evitar una dispersión promedio, respecto a la media, igual a cero, se elevan al cuadrado todas las desviaciones y luego se suman. Esto conduce a la *segunda* etapa en el cálculo de la *variancia*, la cual es una medida de ________.

dispersión

---

45 La *segunda* etapa para calcular la variancia consiste en:

2. ***elevar al cuadrado cada una de las desviaciones con respecto a la media y sumar todos los resultados.***

Si $X_i = 3$ y $\overline{X} = 5$, entonces $(X_i - \overline{X})$ = ________ y $(X_i - \overline{X})^2$ = ________.

−2, 4

---

46 Complete la siguiente tabla ($\overline{X} = 7$).

x	$x-\overline{x}$	$(x-\overline{x})^2$
3		
4		
6	−1	1
8	1	1
9	2	4
12		
TOTAL:		

x	$x-\bar{x}$	$(x-\bar{x})^2$
3	−4	16
4	−3	9
6	−1	1
8	1	1
9	2	4
12	5	25
TOTAL:	0	56

47 Complete la siguiente tabla de datos agrupados ($\bar{X} = 7.2$).

x	f	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
3	2			
4	1	−3.2	10.24	10.24
6	3	−1.2	1.44	4.32
8	5			
10	4			
TOTAL:				______

x	f	$(x-\bar{x})$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
3	2	−4.2	17.64	35.28; (17.64×2)
4	1	−3.2	10.24	10.24; (10.24×1)
6	3	−1.2	1.44	4.32; (1.44×3)
8	5	0.8	0.64	3.20; (0.64×5)
10	4	2.8	7.84	31.36; (7.84×4)
TOTAL:				84.40

48 La segunda etapa en el cálculo de la variancia consiste en calcular los cuadrados de las ____________ con respecto a la ________ y luego ___________ los resultados.

desviaciones, media, sumar

49 Anote los encabezados de las columnas que se utilizan hasta terminar con la segunda etapa del cálculo de la variancia:

X	f	(?)	(?)	(?)

X	f	$X - \bar{X}$	$(X - \bar{X})^2$	$(X - \bar{X})^2 f$

50 La *tercera* y última etapa para calcular la variancia consiste en:

3. dividir entre N *la suma de los cuadrados de todas las desviaciones con respecto a la media.*

Recuerde que N representa al número total de datos.

Si $\sum_{i=1}^{N} (X_i - \bar{X})^2 = 56$ y $N = 6$, el resultado de esta etapa es

$$\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N} = \frac{(?)}{(?)} = 9.33$$

$\frac{56}{6}$

51 Puesto que la última etapa para calcular la variancia consiste en dividir entre *N* la suma de todos los cuadrados de las desviaciones, la fórmula para calcular la variancia de datos no agrupados será:

$$\text{variancia} = \frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{(?)}$$

N

52 La fórmula para calcular la variancia cuando se usan datos no agrupados es

$$\text{variancia} = \frac{(?)}{(?)}$$

$$\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}$$

53 Si la suma de los cuadrados de las desviaciones con respecto a la media es de 235.2 y el total de datos es 15, la variancia vale:

$$\text{varianza} = \frac{(\,?\,)}{(\,?\,)} = 15.68$$

$\frac{235.2}{15}$

---

54 Si usamos el símbolo S^2 para denotar la variancia, entonces la fórmula de ésta para datos no agrupados (complete la fórmula) es

$$S^2 = \frac{(\,?\,)}{(\,?\,)}$$

$$S^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

---

55 ¿Qué símbolo usaremos para denotar la variancia? ______________

S^2.

---

56 Escriba la fórmula para calcular la variancia de datos no agrupados. ______________

$$S^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

---

57 Si los datos están agrupados, la suma de todas las desviaciones elevadas al cuadrado está dada por $\sum_{i=1}^{K} (X_i - \overline{X})^2 f_i$; donde f_i es la ____________ del dato X_i y K es el número de grupos.

frecuencia

---

58 Entonces, si los datos están agrupados, la variancia se calcula con la fórmula

$$S^2 = \frac{\sum_{i=1}^{K} (X_i - \overline{X})^2 f_i}{(\,?\,)}$$

N

---

59 La fórmula para calcular la variancia de datos agrupados en K clases es

$$(\,?\,) = \frac{(\,?\,)}{(\,?\,)}$$

$$S^2 = \frac{\sum_{i=1}^{K} (X_i - \overline{X})^2 f_i}{N}$$

---

60 Si $\sum_{i=1}^{K} (X_i - \overline{X})^2 f_i = 235.2$ y $N = 15$, entonces

$$S^2 = \frac{(\,?\,)}{(\,?\,)}$$

$\frac{235.2}{15}$

---

61 El símbolo S^2 sirve para denotar la ____________.

variancia

---

62 Si los datos están agrupados en K intervalos de clase, se utiliza la fórmula

$$S^2 = \frac{\sum_{i=1}^{K} f_i (X_i - \overline{X})^2}{N}$$

en este caso, X_i es la *marca de clase* y f_i la ______________.

frecuencia

---

63 Cuando los datos están agrupados en K intervalos, las X_i representan las ______________.

marcas de clase

64 Use la hoja de trabajo 6.1 para calcular la variancia de los datos en ella presentados, los cuales corresponden a la edad a la cual murieron 25 personas de fiebre tifoidea. Sólo se consideraron edades entre 40 y 60 años. Use $\overline{X} = 52$.

Edad, X, en años (marca de clase)	Frecuencia	$x - \overline{x}$	$(x - \overline{x})^2$	$(x - \overline{x})^2 f$
40	2	−12	144	288
45	4	−7	49	196
50	5	−2	4	20
55	10	3	9	50
60	4	8	64	256
TOTAL:	25			650

$$S^2 = \frac{650}{25} = 34.0$$

65 El rango y la ______________ son medidas de dispersión o variación.

variancia

66 Si las unidades de X son *metros*, ¿cuáles son las unidades de $(X_i - \overline{X})^2$?

Metros al cuadrado (m^2).

67 Si las unidades de $(X_i - \overline{X})^2$ son metros al cuadrado, ¿cuáles son las unidades de

$\sum_{i=1}^{N} (X_i - \overline{X})^2/N$?

Metros al cuadrado (m^2).

68 Si las unidades de $\sum_{i=1}^{N} (X_i - \overline{X})^2/N$ son metros al cuadrado, las unidades de la variancia también son metros al cuadrado.

Las unidades de la variancia son ______.
(a/b/c)

a. las mismas que las de la variable

b. las de la variable al cuadrado

c. metros al cuadrado

b.

69 Use la hoja de trabajo 6.2 para calcular la variancia de los siguientes datos. La media de estas calificaciones es 80.3.

Intervalo de calificaciones	Marca de clase	Frecuencia
55–63	59	2
64–72	68	6
73–81	77	7
82–90	86	9
91–99	95	6

x	f	$(x - \overline{x})$	$(x - \overline{x})^2$	$(x - \overline{x})^2 f$
59	2	−21.3	453.7	907.4
68	6	−12.3	151.3	907.8
77	7	−3.3	10.9	76.3
86	9	5.7	32.5	292.5
95	6	14.7	216.1	1,296.6
TOTAL:	30			3,480.6

$$S^2 = \frac{3{,}480.6}{30} = 116.0$$

70 ¿Cuáles son las medidas de dispersión que hemos visto hasta ahora?

El rango y la variancia (en cualquier orden).

71 Otra medida de dispersión muy útil es la llamada *desviación estándar*, la cual es igual a la *raíz cuadrada* de la variancia, o sea que la fórmula de la desviación estándar es ______.
(a/b/c)

a. desviación estándar = $\overline{X}$

b. desviación estándar = S^2

c. desviación estándar = $\sqrt{S^2} = S$

c. desviación estándar = $\sqrt{S^2} = S$

72 La desviación estándar es igual a la ______________ de la variancia.

raíz cuadrada

73 Puesto que la raíz cuadrada de la variancia es $\sqrt{S^2} = S$, entonces S es la ______ ______.

desviación estándar

74 Puesto que las unidades de la variancia son las de la variable al cuadrado, las unidades de la desviación estándar son ______ (a/b/c).

a. las mismas que las de la variable
b. las de la variable al cuadrado
c. metros cuadrados

a. las mismas que las de la variable.

75 ¿Cuál es el símbolo para denotar la desviación estándar?

S.

76 Relacione los símbolos de la izquierda con los conceptos de la derecha.

1. $\bar{X}$	a. Una variable
2. S^2	b. La media
3. X	c. La desviación estándar
4. S	d. La mediana
	e. La variancia

1–b, 2–e, 3–a, 4–c

77 La desviación estándar es una medida de ______.

dispersión

78 La desviación estándar es igual a la ______ ______ de la ______.

raíz cuadrada, variancia

79 Si la variancia de unos datos vale 400 cm², ¿cuánto vale la desviación estándar?

20 cm ($\sqrt{400 \text{ cm}^2} = 20$ cm)

80 La desviación estándar tiene ______ (diferentes/las mismas) unidades que la variable.

las mismas

81 Identifique la fórmula para calcular la desviación estándar.

a. $\dfrac{\sum_{i=1}^{N} X_i}{N}$ b. $\dfrac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}$ c. $L_M + \dfrac{\frac{N}{2} - F_M}{f_M} d_M$

d. $\sqrt{\dfrac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}}$.

d. $S = \sqrt{\dfrac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N}}$

82 Si $\sum_{i=1}^{N} (X_i - \bar{X})^2 = 90$ cm² y N = 10, entonces S = ______.

3; ($\sqrt{\dfrac{90}{10}} = 3$)

83 Se puede demostrar la siguiente ecuación: $S^2 = \dfrac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N} = \dfrac{\sum_{i=1}^{N} X_i^2}{N} - \bar{X}^2$
El miembro de la derecha de esta ecuación nos proporciona una manera más fácil para calcular la variancia y, por tanto, la ______ ______.

desviación estándar

84 Note que $\sum_{i=1}^{N} X_i^2/N$ es el promedio aritmético de las X_i^2. La fórmula simplificada para calcular S^2 es

$$S^2 = \frac{\sum_{i=1}^{N} X_i^2}{N} - \overline{X}^2$$

donde $\sum_{i=1}^{N} X_i^2/N$ es el ______ ______

promedio aritmético

---

85 La *fórmula simplificada* para calcular la variancia es:

$$S^2 = \frac{\sum_{i=1}^{N} X_i^2}{(\ ?\ )} - \overline{X}^2$$

$$S^2 = \frac{\sum_{i=1}^{N} X_i^2}{N} - \overline{X}^2$$

---

86 Puesto que $\sum_{i=1}^{N} X_i^2/N$ es el promedio aritmético de las X_i^2, introduciremos la siguiente notación:

$$\overline{X^2} = \frac{\sum_{i=1}^{N} X_i^2}{N}$$

(note que la tilde va arriba del exponente, esto es $\overline{X^2}$).
Entonces, usando esta notación:

$$S^2 = \frac{\sum_{i=1}^{N} X_i^2}{(\ ?\ )} - \overline{X}^2$$

$$S^2 = \overline{X^2} - \overline{X}^2$$

---

87 Complete la siguiente fórmula

$$\overline{X^2} = \underline{\qquad\qquad}$$

$$\overline{X^2} = \frac{\sum_{i=1}^{N} X_i^2}{N}$$

---

88 Complete la siguiente fórmula:

$$S^2 = \overline{X^2} - (\ ?\ )$$

$$S^2 = \overline{X^2} - \overline{X}^2$$

---

89 A $\overline{X^2}$ se le llama *promedio cuadrático*. Entonces, ya que $S^2 = \overline{X^2} - \overline{X}^2$, la variancia es igual al promedio cuadrático menos el cuadrado de ______.

el promedio (la media, $\overline{X}$)

---

90 El promedio cuadrático se calcula con la fórmula

$$\overline{X^2} = \frac{(\ ?\ )}{(\ ?\ )}$$

$$\overline{X^2} = \frac{\sum_{i=1}^{N} X_i^2}{N}$$

---

91 Escriba la fórmula simplificada para calcular la variancia. ______.

$$S^2 = \overline{X^2} - \overline{X}^2$$

---

92 Puesto que $S = \sqrt{S^2}$, ¿cuál es la fórmula simplificada para calcular la desviación estándar? ______.

$$S = \sqrt{\overline{X^2} - \overline{X}^2}$$

---

93 La siguiente tabulación se utiliza para calcular la variancia. Complete la tabla, y calcule la media y la variancia utilizando la fórmula simplificada.

x	x^2
2	4
4	16
5	25
7	49
9	
10	
TOTAL: (?)	275

$\overline{X} = \_\_\_\_\_\_ = \_\_\_\_\_\_$

$\overline{X^2} = \_\_\_\_\_\_ = \_\_\_\_\_\_$

$S^2 = \_\_\_\_\_\_$

x	x^2
2	4
4	16
5	25
7	49
9	81
10	100
TOTAL: 37	275

$\overline{X} = \dfrac{37}{6} = 6.17$

$\overline{X^2} = \dfrac{275}{6} = 45.83$

$S^2 = 45.83 - 6.17^2 = 7.76$

---

94 Puesto que la variancia de los datos del cuadro anterior vale 7.76, ¿cuánto vale la desviación estándar?

2.78; ($\sqrt{7.76} = 2.78$).

---

95 ¿Cuáles son las tres medidas de dispersión que hemos estudiado hasta ahora?

El rango, la variancia y la desviación estándar.

---

96 Relacione los nombres de la columna de la izquierda con las fórmulas de la derecha.

1. Desviación estándar — a. $\sum_{i=1}^{N} X_i^2/N$
2. Media — b. $\sum_{i=1}^{N} X_i/N$
3. Variancia — c. $\sqrt{\overline{X^2} - \overline{X}^2}$
4. Promedio cuadrático — d. $\dfrac{\sum_{i=1}^{N} X_i^2}{N} - \overline{X}^2$

1-c, 2-b, 3-d, 4-a

---

97 Si los datos están agrupados, la fórmula para calcular el promedio cuadrático es:

$$\overline{X^2} = \frac{\sum_{i=1}^{K} f_i X_i^2}{(?)}$$

donde f_i es la ______ de la iésima clase y K es el ______ de grupos.

N, frecuencia, número (total)

---

98 ¡Recuerde!

La fórmula para calcular la media de datos agrupados es

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{(?)}$$

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

---

99 Relacione los símbolos de la columna de la izquierda con las fórmulas de la derecha.

1. $\overline{X}$ — a. $\sum_{i=1}^{K} f_i X_i^2/N$
2. $\overline{X^2}$ — b. $\overline{X^2} - \overline{X}^2$
3. S^2 — c. $\sum_{i=1}^{K} X_i/N$
 — d. $\sum_{i=1}^{N} X_i/N$

1-d, 2-a, 3-b

Complete la siguiente fórmula para el cálculo del promedio cuadrático de datos agrupados:

$$\overline{X^2} = \sum_{i=1}^{(?)} X_i^2 f_i/(?)$$

$$\overline{X^2} = \sum_{i=1}^{K} X_i^2 f_i/N$$

101 La fórmula simplificada para calcular la variancia es

$$S^2 = \overline{X^2} - \overline{X}^2$$

Si los datos están agrupados, $\overline{X^2}$ y $\overline{X}$ se calculan con las fórmulas (escríbalas):

$\overline{X^2} =$

$\overline{X} =$

$$\overline{X^2} = \frac{\sum_{i=1}^{K} f_i X_i^2}{N}$$

$$\overline{X} = \frac{\sum_{i=1}^{K} X_i f_i}{N}$$

---

102 Si $\sum_{i=1}^{10} X_i^2 = 500 \text{ cm}^2$, $\sum_{i=1}^{10} X_i = 50$ y $N = 10$, entonces

$\overline{X^2} =$ ____________

$\overline{X} =$ ____________

$S^2 =$ ____________

$S =$ ____________

50; (500/10), 5; (50/10), 25; (50 — 25), 5; $(\sqrt{25})$

---

103 Si los datos están agrupados, se puede utilizar una tabla como la siguiente para calcular la variancia mediante la fórmula simplificada. Complete y calcule la variancia (haga todo el trabajo en esta hoja). Recuerde que $N = \Sigma f_i$.

x	f	xf	x^2	x^2f
3	2	6	9	18
4	1	4	16	16
6	3	18	36	
8	5	40	64	
10	4	40	100	
TOTALES				

$\overline{X} =$ ____________

$\overline{X}^2 =$ ____________

$\overline{X^2} =$ ____________

$S^2 =$ ____________

x	f	xf	x^2	x^2f
3	2	6	9	18
4	1	4	16	16
6	3	18	36	108
8	5	40	64	320
10	4	40	100	400
TOTALES 15		108		862

$$\overline{X} = \frac{108}{15} = 7.20$$

$$\overline{X}^2 = 51.84$$

$$\overline{X^2} = \frac{862}{15} = 57.45$$

$$S^2 = 57.46 - 51.84 = 5.62$$

---

104 Si los datos están agrupados en intervalos de clase, al aplicar la fórmula para calcular la variancia, las X_i representan las ____________ de clase, y las f_i las ____________ de clase.

marcas, frecuencias

---

105 Anote los encabezados de las columnas de la tabla usada para calcular la variancia mediante la fórmula simplificada.

x	f	(?)	(?)	(?)

x	f	xf	x^2	x^2f

---

106 Use la hoja de trabajo 6.3 para calcular, mediante el método simplificado, la desviación estándar de la edad a la que fallecieron 25 personas de fiebre tifoidea. Sólo se consideraron edades entre 40 y 60 años. La media de esos datos es 52.

Edad X_i en años (marca de clase)	Frecuencia	x^2	x^2f
40	2	1,600	3,200
45	4	2,025	8,100
50	5	2,500	12,500
55	10	3,025	30,250
60	4	3,600	14,400
TOTALES	25		68,450

$$\overline{X^2} = \frac{68{,}450}{25} = 2{,}738$$

$$S^2 = 2{,}738 - 2{,}704 = 34$$

$$S = \sqrt{34} = 5.83$$

PARTE C. EL COEFICIENTE DE VARIACION

107 Cuando se comparan las *dispersiones* de varios grupos de datos, es conveniente usar para ello un parámetro *adimensional* llamado *coeficiente de variación*. Este se define como el cociente $S/\overline{X}$, donde S es la ____________ ____________ y $\overline{X}$ es ____________ ____________.

desviación estándar, la media (o el promedio)

---

108 ______________ es igual a $S/\overline{X}$.
(El coeficiente de variación/La varianza)

El coeficiente de variación

---

109 El cociente de la desviación estándar, S, entre la media, $\overline{X}$, define otra medida de dispersión llamada coeficiente de ______________.

variación

---

110 La fórmula para calcular el coeficiente de variación es:

$$\text{Coeficiente de variación} = \frac{(\ ?\ )}{(\ ?\ )}.$$

$$\text{Coeficiente de variación} = \frac{S}{\overline{X}}$$

---

111 Usaremos la letra v para denotar el coeficiente de variación, por lo cual:

$$v = \frac{(\ ?\ )}{(\ ?\ )}.$$

$$v = \frac{S}{\overline{X}}$$

---

112 El coeficiente de variación lo denotaremos con la letra ___.

v

---

113 Escriba la fórmula para calcular el coeficiente de variación. ______________

$$v = \frac{S}{\overline{X}}$$

---

114 Las unidades de S son las mismas que las de ______.
(a/b/)

a. $\overline{X}$
b. S^2

a. $\overline{X}$

---

115 Se dice que un valor es *adimensional* cuando se expresa mediante una cifra no acompañada por una unidad de medida. Por ejemplo: 3.16, 4,326, etcétera. En caso contrario, el valor tiene la dimensión de la unidad de medida que lo acompaña. Por ejemplo: 4.1 km, 3 lt, 17 pesos, 80 km/h, etcétera.
Diga cuáles de los siguientes valores son adimensionales. ______________.
a. 217 km/h
b. 30.1
c. 17 m^3
d. 3,129

b. 30.1, d. 3,129

---

116 Puesto que las unidades de S y $\overline{X}$ son las mismas, el coeficiente de variación, v, ______________.
(tiene las unidades de S/es adimensional)

es adimensional

---

117 Decíamos que el coeficiente de variación es útil cuando se comparan las dispersiones de varios grupos de datos. Por ejemplo, si de dos fábricas, A y B, de remaches, se obtienen datos de los diámetros del producto, y en una se obtiene un promedio de 5 mm y una desviación estándar de 0.05 mm, entonces

$$v = \frac{(\ ?\ )}{(\ ?\ )} = \underline{\qquad\qquad}.$$

$$v = \frac{0.05}{5} = 0.01$$

---

118 Si de la otra fábrica se obtienen los resultados: $\overline{X} = 5.5$ mm y $S = 0.051$, entonces:

$$v = \frac{(\ ?\ )}{(\ ?\ )} = 0.0093$$

$$v = \frac{0.051}{5.5} = 0.0093$$

---

119 Los resultados de las dos fábricas fueron

Parámetro	Fábrica A	Fábrica B
Media (mm)	5	[illegible]
Desviación estándar	[illegible]	[illegible]
Coeficiente de variación	0.01	[illegible]

a. ¿Cuál fábrica produce con mayor desviación estándar?
b. ¿Cuál fábrica produce con menor coeficiente de variación?

Fábrica B. Fábrica B.

---

120 La fábrica B produce con mayor desviación estándar que la A, pero con menor coeficiente de __________.

variación

---

121 En algunas normas de fabricación (tales como las de concreto para construcción), se especifica el grado de calidad del producto de acuerdo con el coeficiente de variación; a mayor coeficiente de variación menor control en la calidad del producto; a menor __________ ______ ____________ mayor control en la calidad del producto.

coeficiente de variación

---

122 Relacione los conceptos de la columna de la izquierda con los nombres de la derecha.

1. Medidas de tendencia central
2. Medidas de dispersión

a. Rango
b. Mediana
c. Modo
d. Variancia
e. Media
f. Coeficiente de variación
g. Desviación estándar

1-b,c,e, 2-a,d,f,g

---

123 Relacione los nombres de la columna de la izquierda con las fórmulas de la derecha.

1. Media
2. Variancia
3. Coeficiente de variación
4. Mediana

a. $S/\overline{X}$

b. $\sum_{i=1}^{N} \frac{X_i^2}{N} - \overline{X}$

c. $\sqrt{\overline{X^2} - \overline{X}^2}$

d. $L_M + \frac{N/2 - F_M}{f_M} d_M$

e. $\sum_{i=1}^{N} X_i/N$

1-e, 2-b, 3-a, 4-d

---

124 Si la desviación estándar de unos datos vale 10.77 y la media es 80.30, ¿cuánto vale el coeficiente de variación?

0.13; $\left(\frac{10.77}{80.30} = 0.13\right)$.

REVISION

125 El rango es una medida de __________

dispersión

---

126 ¿Cómo se calcula el rango? ____________________

Se resta al dato de máximo valor el de mínimo valor.

---

127 Escriba la fórmula simplificada para calcular la variancia (en términos del promedio cuadrático). __________________.

$S^2 = \overline{X^2} - \overline{X}^2$

---

128 Escriba la fórmula para calcular el promedio cuadrático de datos no agrupados.

____

$$\overline{X^2} = \frac{\sum_{i=1}^{N} X_i^2}{N}$$

129 Escriba la fórmula para calcular el promedio cuadrático de datos agrupados____.

$$\overline{X^2} = \frac{\sum_{i=1}^{K} X_i^2 \, f_i}{N}$$

130 ¿Cuáles son las cuatro medidas de dispersión que hemos estudiado en esta unidad? ____________

El rango, la variancia, la desviación estándar y el coeficiente de variación.

131 Escriba la fórmula para calcular la desviación estándar de datos agrupados en intervalos de clase, en términos de X_i y de f_i.____________

$$S = \sqrt{\frac{\sum_{i=1}^{K} (X_i - \overline{X})^2 \, f_i}{N}}$$

132 Cuando los datos están agrupados en intervalos, las X_i que aparecen en la fórmula para la variancia identifican a las ____________ y las f_i a las ____________ de clase.

marcas de clase, frecuencias

133 ¿Con qué letra se denota el coeficiente de variación?____________.

Con la v.

134 ¿Cuál es la fórmula que sirve para calcular el coeficiente de variación?________.

$$v = \frac{S}{\overline{X}}$$

135 Las unidades de la desviación estándar son ____________ (diferentes/las mismas) que las de la variable.

las mismas

136 ¿Cuáles son las unidades del coeficiente de variación?

No tiene unidades (es adimensional).

137 Escriba la fórmula para calcular la variancia de datos no agrupados, en términos de X_i, $\overline{X}$ y N.

$$S^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

138 La distribución de frecuencias de los ingresos diarios de las familias de mil alumnos de la Universidad Nacional Autónoma de México, es aproximadamente la siguiente:

Marca de clase, en pesos	Frecuencia
30	149
100	234
200	4[illegible]5
300	62
400	4[illegible]
500	5
600	2
Fuente: Gaceta Universitaria, 19[illegible].	

Calcule, en la hoja de trabajo 6.4, la media, el promedio cuadrático, la variancia, la desviación estándar y el coeficiente de variación de dichos ingresos.

x_i en pesos (marca de clase)	f (frecuencia)	xf	x^2	x^2f
30	149	4,470	900	134,100
100	234	23,400	10,000	2,340,000
200	425	85,000	40,000	17,000,000
300	62	[illegible]	90,000	[illegible]
400	[illegible]	[illegible]	160,000	[illegible]
500	5	[illegible]	250,000	[illegible]
600	2	[illegible]	360,000	[illegible]
TOTALES	1,000	167,470		38,794,100

$$\overline{X} = \frac{167{,}470}{1{,}000} = 167.47 \text{ pesos}$$

$$\overline{X^2} = \frac{38{,}794{,}100}{1{,}000} = 38{,}794.1 \text{ pesos}^2$$

$$S^2 = 38{,}794.1 - (167.47)^2 = 38{,}794.1 - 28{,}046.2 = 10{,}747.9 \text{ pesos}^2$$

$$S = \sqrt{10{,}747.9} = 327.79 \text{ pesos}$$

139 En una investigación realizada para determinar la velocidad de propagación del sonido a través de un cierto material, se obtuvieron las siguientes mediciones en km/seg : 6.3, 5.9, 6.8, 6.0, 6.5, 6.1, 7.0, 6.2, 6.6 y 6.4. En la hoja de trabajo 6.5 calcule la media, el rango, la variancia, la desviación estándar y el coeficiente de variación. Al calcular la media aproxime al décimo más cercano; en lo demás aproxime al centésimo más cercano.

x	$x-\bar{x}$	$(x-\bar{x})^2$
6.3	−0.1	0.01
5.9	−0.5	0.25
6.8	0.4	0.16
6.0	−0.4	0.16
6.5	0.1	0.01
6.1	−0.3	0.09
7.0	0.6	0.36
6.2	−0.2	0.04
6.6	0.2	0.04
6.4	0	0
63.8		1.12

$$\bar{X} = \frac{63.8}{10} = 6.4$$

Rango = 7.0 − 5.9 = 1.1

$$S^2 = \frac{1.12}{10} = 0.11$$

$$S = \sqrt{0.11} = 0.33$$

$$v = 0.33/6.4 = 0.05$$

EXAMEN

1. *Las medidas que dan idea de qué tan dispersos están los datos, o sea, de la variabilidad de los mismos, se llaman ____________ ________ __________.*

2. *¿Cuáles son las medidas de dispersión que se estudiaron? __________________.*

3. *¿Cómo se calcula el rango de un grupo de datos? ____ ____________________.*

4. *El rango de un grupo de datos es ________________ a los cambios de los valores intermedios.*

5. *Una medida de dispersión que toma en cuenta todos los datos es la variancia. Escriba la fórmula para calcularla, cuando los datos no están agrupados. ________.*

6. *Calcule la variancia de los siguientes datos: 2, 8, 3, 5, 9, 10, 11, 1, 7, 6. Tome en cuenta que la media es 6.2. ____.*

7. *Escriba la fórmula para calcular la variancia de datos agrupados. _____ __________.*

8. *Calcule la variancia de la siguiente distribución de frecuencias; la media es 5.0.*

Límites reales	Frecuencia
0 − 2	2
2 − 4	4
4 − 6	8
6 − 8	4
8 − 10	2

9. *Otra medida de dispersión es la ___________ _______________, la cual se define como la raíz cuadrada de la ____________________.*

10. *Si la variancia de ciertos datos es 4.8, ¿cuál es la desviación estándar?*

11. *Al dividir la desviación estándar de ciertos datos entre la media, se obtiene el ___________ ________ __________.*

12. *Relacione los nombres de la columna de la izquierda con las descripciones que aparecen a la derecha.*

1.	*Coeficiente de variación*	*a.*	*Unidades iguales que los datos*
2.	*Variancia*	*b.*	*Es adimensional*
3.	*Desviación estándar*	*c.*	*Unidades de los datos al cuadrado*

13. *A* $\overline{X^2} = \sum_{i=1}^{N} X_i^2 / N$ *se le llama __________ __________ __________.*

14. *La fórmula para calcular la variancia en términos de $\overline{X^2}$ y de $\overline{X}$ es* _____.

15. *Para cierto grupo de datos, se tiene $\overline{X} = 5$ y $\overline{X^2} = 41$; cuánto vale:*

   a. *la variancia* ____________________,
   b. *la desviación estándar* ________________,
   c. *el coeficiente de variación?* ____________.

TOTAL: 23 puntos

RESPUESTAS

1. *medidas de dispersión*

2. *Rango, variancia, desviación estándar y coeficiente de variación.*

3. *Obteniendo la diferencia entre los valores máximo y mínimo de los datos.*

4. *insensible*

5. $$S^2 = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

6. $$S^2 = \frac{(-4.2)^2 + 1.8^2 + (-3.2)^2 + (-1.2)^2 + 2.8^2 + 3.8^2 + 4.8^2 + (-5.2)^2 + 0.8^2 + (-0.2)^2}{10}$$
$$= \frac{105.60}{10} = 10.56$$

7. $$S^2 = \frac{\sum_{i=1}^{K} f_i (X_i - \overline{X})^2}{N}$$

8. $$S^2 = \frac{2X(-4)^2 + 4X(-2)^2 + 8X(0)^2 + 4X(2)^2 + 2X(4)^2}{20} = \frac{96}{20} = 4.8$$

9. *desviación estándar*
   variancia

10. *2.2*

11. *coeficiente de variación*

12. *1—b*
   2—c
   3—a

13. *valor medio cuadrático*

14. $S^2 = \overline{X^2} - \overline{X}^2$

15. a. $S^2 = 41 - 25 = 16$
   b. $S = 4$
   c. $v = \frac{4}{5} = 0.8$

HOJA DE TRABAJO 6.1

Edad, X, en años (marca de clase)	Frecuencia	$X - \bar{X}$	$(X - \bar{X})^2$	$(X - \bar{X})^2 f$
40	2			
45	4			
50	5			
55	10			
60	4			
TOTAL:				

$s^2 =$ ________ = ________

HOJA DE TRABAJO 6.2

X	f	$(X - \bar{X})$	$(X - \bar{X})^2$	$(X - \bar{X})^2 f$
			453.7	
			151.3	
			10.9	
			32.5	
			216.1	
TOTAL:				

$s^2 =$ ________ = ________

HOJA DE TRABAJO 6.3

Edad, X, en años (marca de clase)	Frecuencia	X^2	$X^2 f$
40	2		
45	4		
50	5		
55	10		
60	4		
TOTAL:			

$\overline{X^2} =$ ________ = ________

$s^2 =$ ________________

$s =$ ________________

HOJA DE TRABAJO 6.4

X, en pesos marca de clase	f (frecuencia)	Xf	X^2	$X^2 f$
30	143			
100	298			
200	425			
300	62			
400	49			
600	5			
900	3			
TOTAL:				

HOJA DE TRABAJO 6.5

$\overline{X}$ = ______ = ______

Rango = ______

S^2 = ______ = ______

S = ______

v = ______ = ______

UNIDAD VII TRANSFORMACION DE VARIABLES

PREFACIO

En esta unidad estudiaremos el comportamiento de algunas medidas de tendencia central y de dispersión, cuando se efectúan transformaciones de variables aleatorias.
En particular, se hará hincapié sobre la transformación en *puntuaciones estándar*, las cuales son muy útiles en estudios de Inferencia Estadística y de Probabilidades.

PARTE A. TRANSFORMACIONES Y = X + C y Y = X − C

1 Escriba la fórmula para calcular la media de un grupo de datos no agrupados. ______

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

2 Si efectuamos alguna operación algebraica con la variable aleatoria X, decimos que la estamos *transformando*. Por ejemplo, si multiplicamos X por una constante decimos que estamos ______ la variable X.

transformando

3 Al efectuar alguna operación algebraica con la variable aleatoria X, decimos que se está haciendo una ______ de X.

transformación

4 En esta unidad nos concretamos a estudiar transformaciones que involucran una constante C. Analizaremos los casos X + C, X − C, XC y X/C. En estas transformaciones C denota una ______ (constante/variable)

constante

5 Empezaremos por estudiar la relación que existe entre las medias (promedios aritméticos) de los datos de X y de X + C. Usaremos el símbolo Y para denotar una variable aleatoria transformada. En este caso, Y = X + (?).

Y = X + C

6 La letra C nos sirve para denotar una ______.

constante

7 La letra Y nos sirve para denotar ______ (a/b/c)

a. la variable aleatoria original
b. una constante
c. la variable aleatoria transformada

c. la variable aleatoria transformada

8 Para la transformación X + C se tiene que Y= ______.

Y = X + C

9 Si transformamos la variable aleatoria X en otra variable aleatoria, Y, mediante la adición a X de una *constante* C, los datos de Y, en términos de los datos de X, serán:

$$Y_i = X_i + C$$

Por lo tanto, la media de los datos de Y es

$$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{\sum_{i=1}^{N} [X_i + (?)]}{(?)}$$

$$\overline{Y} = \frac{\sum_{i=1}^{N} (X_i + C)}{N}$$

10 Desarrollando la suma que aparece en el miembro derecho de la ecuación para calcular $\overline{Y}$ se obtiene:

$$\frac{\sum_{i=1}^{N} (X_i + C)}{N} = \frac{(X_1 + C) + (X_2 + C) + \ldots + (X_N + C)}{N} = \frac{\sum_{i=1}^{N} X_i + NC}{N}$$

$$= \frac{\sum_{i=1}^{N} X_i}{N} + (?).$$

C

11 En el cuadro anterior se llegó a la ecuación

$$\overline{Y} = \frac{\sum_{i=1}^{N} X_i}{N} + C$$

pero, tomando en cuenta que $\sum_{i=1}^{N} X_i / N = (?)$, se concluye que $\overline{Y} = \overline{X} + C$.

$\overline{X}$

12 De esta manera la media de los datos de Y = X + C es igual a la media de ______ más la constante C.

X

13 Cuando cada observación de una variable aleatoria X se incrementa en una constante, la media que resulta es igual a la media de las X_i más la ______.

constante

14 Si Y = X + C, entonces $\overline{Y}$ = (?) + (?)

$\overline{Y} = \overline{X} + C$

15 Si $Y = X + C$, $X = 17.35$ y $C = 16$, ¿cuánto vale $\overline{Y}$? ________

33.35; ($\overline{Y} = \overline{X} + C = 17.35 + 16 = 33.35$).

---

16 Por analogía, si la transformación es $Y = X - C$, entonces $\overline{Y} = (?) - (?)$.

$\overline{Y} = \overline{X} - C$

---

17 Cuando hacemos la transformación $Y = X - C$, la media de los datos de Y es igual a la media de los datos de ____ ________ (más/menos) la constante C.

X, menos

---

18 Si $Y = X + 3$, calcule la media de los datos de Y correspondientes a los datos de X presentados en la siguiente tabla.

$\overline{Y} = \frac{\quad}{\quad} = $ ________ cm.

X_i en cm	Y_i en cm
3	6
4	7
6	
8	
9	
12	
TOTAL:	

X_i en cm	Y_i en cm
3	6
4	7
6	9
8	11
9	12
12	15
TOTAL:	60

$\overline{Y} = \frac{60}{6} = 10$ cm.

---

19 Si se dispone del valor de Y para la transformación $Y = X + C$, se puede calcular $\overline{X}$, ya que de la ecuación $\overline{Y} = \overline{X} + C$ se obtiene $\overline{X} = \overline{Y} - C$. Si la transformación es $Y = X - C$, ¿con qué ecuación calcularía $\overline{X}$ en términos de $\overline{Y}$ y C?

$\overline{X} = \overline{Y} + C$.

---

20 Si $Y = X + 3$ y $\overline{Y} = 10$, ¿cuánto vale $\overline{X}$?

$\overline{X} = \overline{Y} - 3 = 7$.

---

21 Si $Y = X - C$ y $C = \overline{X}$, entonces $\overline{Y} =$ ________.

$\overline{Y} = 0$

---

22 Si a cada dato X_i le sustraemos la media $\overline{X}$, el promedio de los datos transformados es igual a ______.

cero

---

23 Recuerde que en la Unidad VI a $X_i - \overline{X}$ le llamamos ______ (a/b/c).

a. desviación estándar
b. desviación respecto a la media
c. variancia

b. desviación respecto a la media

---

24 Recuerde que la suma de *todas* las desviaciones de los datos con respecto a la media es igual a ______ (a/b/c).

a. la media
b. cero
c. la mediana

b. cero

---

25 El resultado de que la suma de todas las desviaciones de los datos con respecto a la media es igual a cero, es congruente con el hecho de que si $Y = X - \overline{X}$, entonces $\overline{Y} =$ ______.

cero

---

26 Cuando se tiene la transformación $Y = X - C$, $\overline{X}$ se calcula mediante la ecuación (escríbala):

$\overline{X} = \overline{Y} + C$

---

27 Para ahorrar trabajo en el cálculo de las medias de datos cuyos valores sean grandes, podemos trabajar primero con valores pequeños obtenidos mediante la

transformación $Y = X - C$. Por ejemplo, vea la siguiente tabla, en la cual se presentan los datos, X_i, de la capacidad de carga de seis vigas de concreto nominalmente iguales. Los valores de X_i son grandes, por lo que podemos ahorrarnos trabajo si en vez de calcular directamente $\overline{X}$, calculamos primero la media de las $Y_i = X_i - 3{,}000$. Hágalo.

X_i en ton	Y_i en ton
2,150	−850
2,200	
2,900	
3,200	
3,700	700
3,750	
TOTALES	

$\overline{Y} =$ ________ $=$ ________ ton.

X_i en ton	Y_i en ton
2,150	−850
2,200	−800
2,900	−100
3,200	200
3,700	700
3,750	750
TOTALES	−100

$\overline{Y} = \dfrac{-100}{6} = -16.67$ ton.

28 En el cuadro anterior teníamos $Y = X - 3{,}000$ y $\overline{Y} = -16.67$, ¿cuánto vale $\overline{X}$, es decir, la media de los datos originales? ________

$\overline{X} = \overline{Y} + 3{,}000 = -16.67 + 3{,}000 = 2{,}983.33$ ton.

29 Al calcular la media de los datos de X, mediante el cálculo previo de la media de los datos de $Y = X - C$, es conveniente que C tome un valor cercano a la media de los datos de X. Esto se logra de la simple observación de los datos básicos. Mientras más cercano esté el valor de C al de X ________ (más cercano/más lejano) a cero estará el valor de $\overline{Y}$.

más cercano

30 Por lo tanto, mientras más próximo esté el valor de C al de $\overline{X}$, al calcular el valor de $\overline{Y}$ manejaremos números ________ (más pequeños/más grandes).

más pequeños

31 Identifique la fórmula para calcular la media de los datos de X, agrupados en intervalos de clase. ________

I. $\sum_{i=1}^{N} X_i/N$

II. $\sqrt{\overline{X^2} - \overline{X}^2}$

III. $\sum_{i=1}^{K} f_i X_i/N$

IV. $\sum_{i=1}^{N} f_i X_i^2/N$

III. $\sum_{i=1}^{K} f_i X_i/N$

32 En la expresión $\overline{X} = \sum_{i=1}^{K} f_i X_i/N$, ¿qué representa f_i?

La frecuencia de clase del intervalo i.

33 En la expresión $\overline{X} = \sum_{i=1}^{K} f_i X_i/N$:

a. ¿qué representa X_i?

b. ¿qué representa K?

X_i es la marca de clase del intervalo i., K es el número de intervalos

34 Si hacemos la transformación $Y = X - C$, entonces $X = Y + C$, por lo cual,

$$\overline{X} = \sum_{i=1}^{K} f_i (Y_i + C)/N = \frac{\sum_{i=1}^{K} (?)(?)}{N} + \frac{C \sum_{i=1}^{K} f_i}{N}$$

$$\frac{\sum_{i=1}^{K} f_i Y_i}{N}$$

35 Recuerde que la suma de todas las frecuencias de clase de una distribución es igual al número total de datos, N, o sea: $\sum_{i=1}^{N} f_i = (?)$

N

36 En un cuadro anterior se obtuvo:

$$\overline{X} = \sum_{i=1}^{K} f_i Y_i/N + C \sum_{i=1}^{K} f_i/N$$

Tomando en cuenta que $\sum_{i=1}^{K} f_i = N$, esta ecuación nos queda en la forma:

$$\overline{X} = \frac{\sum_{i=1}^{K} f_i Y_i}{N} + \frac{CN}{N} = \frac{\sum_{i=1}^{K} f_i Y_i}{N} + C$$

Además, tomando en cuenta que $\sum_{i=1}^{K} f_i Y_i / N = \overline{Y}$, se tiene que $\overline{X} = (?) + C$.

$\overline{X} = \overline{Y} + C$

37 Entonces, si le sustraemos un valor constante a *todas* las marcas de clase X_i, obtenemos los correspondientes valores de Y_i. Si calculamos $\overline{Y}$ podremos calcular $\overline{X}$ mediante la fórmula (escríbala):

$\overline{X} = \overline{Y} + C$

38 **En la siguiente tabla se muestra la distribución de frecuencias de los salarios, X_i, del personal de confianza de una empresa. Calcule la media de $Y = X - 5{,}500$.**

Límites reales en pesos	Marca de clase, X_i en pesos	f	Y_i en pesos	fY_i en pesos
2,000 - 3,000	2,500	5	-3,000	-15,000
3,000 - 4,000	3,500	10		
4,000 - 5,000	4,500	35		
5,000 - 6,000	5,500	35		
6,000 - 7,000	6,500	10	1,000	10,000
7,000 - 8,000	7,500	5	2,500	10,000
TOTAL:				

$\overline{Y} = \text{———} = \text{———}$

Límites reales, en pesos	Marca de clase, X_i en pesos	f	Y_i en pesos	fY_i en pesos
2,000 - 3,000	2,500	5	-3,000	-15,000
3,000 - 4,000	3,500	10	-2,000	-20,000
4,000 - 5,000	4,500	35	-1,000	-35,000
5,000 - 6,000	5,500	35	0	0
6,000 - 7,000	6,500	10	1,000	10,000
7,000 - 8,000	7,500	5	2,000	10,000
TOTAL:		100		-50,000

$$\overline{Y} = \frac{-50{,}000}{100} = -500 \text{ pesos}$$

39 Para el ejemplo del cuadro anterior se tiene $Y = X - 5{,}500$ y $\overline{Y} = -500$ pesos; ¿cuánto vale $\overline{X}$?

$\overline{X} = -500 + 5{,}500 = 5{,}000$ pesos

40 Al obtener la media de X mediante el cálculo previo de la media de $Y = X - C$, debemos procurar que C tenga un valor ________ (cercano al/alejado del) valor de $\overline{X}$.
Si hacemos esto trabajaremos con cifras ________ (grandes/pequeñas)

cercano al, pequeñas

41 En la siguiente tabla se presenta la distribución de frecuencias de las calificaciones, X_i, en el examen de admisión a una universidad.
Calcule la media de los datos de X utilizando la transformación $Y = X - 65$.

Marca de clase, X	f	Y	fY
5	1		
15	0		
25	3		
35	5		
45	7		
55	10		
65	13		
75	20		
85	5		
95	2		
TOTAL:			

$\overline{Y} = \text{———} = \text{———}$

$\overline{X} = \text{———} = \text{———}$

Marca de clase, X	f	Y	fY
5	1	-60	-60
15	0	-50	0
25	3	-40	-120
35	5	-30	-150
45	7	-20	-140
55	10	-10	-100
65	13	0	0
75	20	10	200
85	5	20	100
95	2	30	60
TOTAL:	73		-110

$$\overline{Y} = \frac{-110}{73} = -1.51$$

$$\overline{X} = \overline{Y} + 65 = -1.51 + 65 = 63.49$$

PARTE 8. TRANSFORMACIONES Y = XC y Y = X/C

42 **Sea ahora la transformación $Y = XC$, donde C es una constante. El promedio de los datos de Y será**

$$\overline{Y} = \frac{\sum_{i=1}^{N} Y_i}{N} = \frac{\sum_{i=1}^{N} X_i C}{N} = \frac{C \sum_{i=1}^{N} X_i}{N}$$

pero $\sum_{i=1}^{N} X_i/N = (?)$

por lo cual $\overline{Y} = (?)\ (?)$.

$\sum_{i=1}^{N} X_i/N = \overline{X}$ $\qquad \overline{Y} = C\overline{X}$

---

43 Es decir, si multiplicamos cada uno de los datos de X por una constante, el promedio que se obtiene es igual al promedio de X multiplicado por la ________.

constante

---

44 Si Y = CX entonces $\overline{Y} = (?)\ (?)$.

$\overline{Y} = C\overline{X}$

---

45 Si Y = CX, entonces X = Y/C; por lo tanto, $\overline{X} = (?)\ /\ (?)$.

$\overline{X} = \overline{Y}/C$

---

46 Si $\overline{X} = 5$ y C = 1,000, ¿cuánto vale $\overline{Y}$ cuando Y = CX?

$\overline{Y} = 1{,}000 \times 5 = 5{,}000$.

---

47 Calcule el promedio de los datos de X presentados en la siguiente tabla; calcule primero la media de Y = X/1,000.

X	Y
−15,000	−15
−20,000	
−35,000	
0	
10,000	
10,000	
TOTAL:	

$\overline{Y} =$ ____________ = ____________

$\overline{X} =$ ________________________

X	Y
−15,000	−15
−20,000	−20
−35,000	−35
0	0
10,000	10
10,000	10
TOTAL:	−50

$\overline{Y} = \frac{-50}{6} = -8.33333$

$\overline{X} = -8.33333 \times 1{,}000 = -8333.33$

---

48 Relacione las transformaciones de la izquierda con las fórmulas de la derecha.

I.	Y = X/C	1.	$\overline{Y} = \overline{X} - C$
II.	Y = X − C	2.	$\overline{Y} = \overline{X}\,C$
III.	Y = XC	3.	$\overline{Y} = \overline{X} + C$
IV.	Y = X + C	4.	$\overline{Y} = \overline{X}/C$

I-4, II-1, III-2, IV-3

PARTE C. METODO CORTO PARA CALCULAR LA MEDIA

49 Si hacemos una *doble* transformación de la forma

$$Y = (X - C_1)/C_2$$

donde C_1 y C_2 son constantes, entonces

$$\overline{Y} = \frac{(?) - C_1}{C_2}$$

$$\overline{Y} = \frac{\overline{X} - C_1}{C_2}$$

---

50 Si $\overline{Y} = (\overline{X} - C_1)/C_2$, ¿cuánto vale $\overline{X}$ en términos de $\overline{Y}$? ____________

$\overline{X} = C_2\overline{Y} + C_1$

---

51 Mediante la transformación $Y = (X - C_1)/C_2$ podemos simplificar el cálculo de la media de datos cuyos valores son grandes, obteniendo primero el promedio aritmético de los datos de Y y luego, en términos de $\overline{Y}$, la media de los datos de X mediante la fórmula (complétela):

$$\overline{X} =$$

$\overline{X} = C_2\overline{Y} + C_1$

52 Al hacer la transformación $Y = (X - C_1)/C_2$ para calcular $\overline{X}$, debemos procurar que C_1 tenga un valor ____________ (cercano al/alejado del) valor de $\overline{X}$.

cercano al

53 La doble transformación que hemos utilizado en los cuadros inmediatos anteriores es

$$Y = (? - ?)/?$$

$Y = (X - C_1)/C_2$

54 ¿A qué transformación corresponde la ecuación $\overline{X} = C_2\overline{Y} + C_1$?

$Y = (X - C_1)/C_2$

55 Use la tabla siguiente, en la cual se muestra la distribución de frecuencias de los salarios del personal de confianza de una empresa.
Calcule Y, empleando la transformación $Y = (X - 5{,}500)/1{,}000$

Marca de clase, X, en pesos	f	Y	fY
2,500	5	-3	-15
3,500	10	-2	-20
4,500	35		
5,500	35		
6,500	10		
7,500	5	2	10
TOTAL			

$\overline{Y} = \_\_\_\_ = \_\_\_\_$

Marca de clase, X, en pesos	f	Y	fY
2,500	5	-3	-15
3,500	10	-2	-20
4,500	35	-1	-35
5,500	35	0	0
6,500	10	1	10
7,500	5	2	10
TOTAL	100		-50

$$\overline{Y} = \frac{-50}{100} = -0.5$$

56 En el cuadro anterior la transformación fue $Y = (X - 5{,}500)/1{,}000$, y el valor de $\overline{Y}$ fue de -0.5. ¿Cuánto vale $\overline{X}$? ____________.

5,000 pesos; ($\overline{X} = 1{,}000\ \overline{Y} + 5{,}500 = -500 + 5{,}500 = 5{,}000$)

57 En el cuadro 55 se usó el siguiente agrupamiento

a. ¿Tienen todos los intervalos igual amplitud? ______.
b. ¿Cuánto vale la amplitud? ______.
c. ¿Corresponde el valor de 5,500 a una marca de clase? ______.

Marca de clase, X, en pesos
2,500
3,500
4,500
5,500
6,500
7,500

a. Sí. b. 1,000 pesos. c. Sí.

58 En la transformación $Y = (X - 5{,}500)/1{,}000$ usamos un valor de C_1 igual a una marca de clase, y el valor de C_2 igual a la ________ de los intervalos.

amplitud (el ancho)

59 Cuando todos los intervalos tienen *la misma* amplitud y tomamos C_1 igual a una marca de clase y C_2 igual a la amplitud de los intervalos, en la columna para Y nos quedan sólo números enteros. Por ejemplo, para $Y = (X - 5{,}500)/1{,}000$ (complete la tabla):

X, en pesos	Y
2,500	-3
3,500	-2
4,500	(?)
5,500	(?)
6,500	(?)
7,500	2

X, en pesos	Y
2,500	−3
3,500	−2
4,500	−1
5,500	0
6,500	1
7,500	2

60 Cuando C_1 es una *marca de clase* y C_2 igual a la amplitud de los intervalos, los valores de Y son números enteros sucesivos (vea la respuesta del cuadro anterior); el cero corresponde al intervalo cuya marca de clase se usó como C_1. Los enteros negativos corresponden a los intervalos con marcas de clase ______________ (menores/mayores) que C_1.

menores

61 Los valores enteros positivos de Y corresponden a los intervalos cuyas marcas de clase son ______________ (menores/mayores) que C_1.

mayores

62 Complete la siguiente tabla para $Y = \frac{(X - C_1)}{C_2}$ en donde $C_1 = 6{,}500$ pesos y $C_2 = 1{,}000$ pesos.

X, en pesos	Y
2,500	
3,500	
4,500	
5,500	
6,500	
7,500	

X, en pesos	Y
2,500	−4
3,500	−3
4,500	−2
5,500	−1
6,500	0
7,500	1

63 Al método para calcular $\overline{X}$ tomando C_1 igual a una marca de clase y C_2 a la amplitud de los intervalos se le llama *método corto, rápido o abreviado*.

Use el método corto para calcular la media de los datos presentados en la tabla siguiente, correspondientes a las calificaciones en el examen de admisión a una universidad. Use $C_1 = 65$.

Marca de clase, X	f	Y	fY
5	1	−6	−6
15	0	−5	0
25	3	−4	−12
35	5	−3	−15
45	7	−2	
55	10	−1	
65	10	0	
75	30	1	
85	5	2	
95	2		
TOTAL:	73		

$\overline{Y} =$ ________ = ________

$C_2 =$ ________

$\overline{X} =$ ________

Marca de clase, X	f	Y	fY
5	1	−6	−6
15	0	−5	0
25	3	−4	−12
35	5	−3	−15
45	7	−2	−14
55	10	−1	−10
65	10	0	0
75	30	1	30
85	5	2	10
95	2	3	6
TOTAL:	73		−11

$$\overline{Y} = \frac{-11}{73} = -0.151$$

$C_2 = 10$ (ancho de los intervalos)

$$\overline{X} = C_2\,\overline{Y} + C_1 = 10(-0.151) + 65 = -1.51 + 65 = 63.49$$

64 En el método corto (rápido o abreviado) para calcular $\overline{X}$ se usa C_1 igual a una ________ ________ ________, y C_2 igual a la ________ ________ ________.

marca de clase, amplitud de los intervalos

65 En el ________ ________ para calcular $\overline{X}$, ______ asume el valor de la marca de clase cerca de la cual se supone caerá el valor de la media. Además, cuando *todos* los intervalos tienen la misma amplitud, se toma C_2 igual a ________

método corto (abreviado o rápido), C_1, la amplitud

66 Si no todos los intervalos tienen igual amplitud, C_2 puede tomar el valor de la amplitud de la mayoría de los intervalos; en este caso los valores de Y ______ serán números enteros sucesivos. (sí/no)

no

---

67 En la siguiente tabla se presenta la distribución de frecuencias de las calificaciones, X_i, en el examen de admisión a una universidad.
Calcule $\overline{X}$ mediante el método corto. Use $C_1 = 65$ y $C_2 = 10$. Observe que no todos los intervalos tienen la misma amplitud.

Marca de clase, X	f	Y	fY
5	2		
35	7		
45	7		
55	10		
65	10		
75	30		
90	7		
TOTAL			

$Y = (X - C_1)/C_2$

$C_1 = 65; \; C_2 = 10$

$\overline{Y} =$ ______ = ______

$\overline{X} =$ ______

Marca de clase, X	f	Y	fY
5	2	−6	−12
35	7	−3	−21
45	7	−2	−14
55	10	−1	−10
65	10	0	0
75	30	1	30
90	7	2.5	17.5
TOTAL	73		−9.5

$Y = (X - C_1)/C_2$
$C_1 = 65, \; C_2 = 10$

$\overline{Y} = \dfrac{-9.5}{73} = -0.13$

$\overline{X} = -1.3 + 65 = 63.7; \; (\overline{X} = C_2\overline{Y} + C_1)$

---

68 ¿Qué transformación se utiliza en el método corto para calcular $\overline{X}$? ______

$Y = (X - C_1)/C_2$

PARTE D. METODO CORTO PARA CALCULAR LA VARIANCIA

69 Calculemos ahora la variancia de los datos de Y, donde $Y = X + C$.
Para esto, introduzcamos la notación S_Y^2 para la variancia de los datos de la variable aleatoria Y.
La ecuación para calcular S_Y^2, en términos de las Y_i es

$$S_Y^2 = \frac{\sum_{i=1}^{N} ((?) - \overline{Y})^2}{(?)}$$

$$S_Y^2 = \frac{\sum_{i=1}^{N} (Y_i - \overline{Y})^2}{N}$$

---

70 S_Y^2 es la notación para ______

la variancia de Y

---

71 Se tiene $S_Y^2 = \dfrac{\sum_{i=1}^{N} (Y_i - \overline{Y})^2}{N}$
donde $Y_i = X_i + C$ y, en consecuencia, $\overline{Y} = (?) + C$.
Por lo tanto,

$$S_Y^2 = \frac{\sum_{i=1}^{N} [(X_i + C) - (\overline{X} + C)]^2}{N} = \frac{\sum_{i=1}^{N} (X_i + C - \overline{X} - C)^2}{N} = \frac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

$\overline{Y} = \overline{X} + C$

---

72 En el cuadro anterior se obtuvo $S_Y^2 = \dfrac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$

pero $\dfrac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N} = S_x^2$ (variancia de los datos de X) por lo cual, $S_Y^2 = (?)$.

$S_Y^2 = S_x^2$

---

73 Es decir, la variancia de los datos de Y es igual a la ______ de los datos de X cuando $Y = X + C$.

variancia

---

74 La variancia de $Y = X + C$ es igual a la ______

variancia de X

75 La desviación estándar es igual a la ________________________ de la varianza.

raíz cuadrada

---

76 Puesto que $S_Y^2 = S_X^2$ entonces $S_Y = (?)$.

$S_Y = S_X$

---

77 Es decir, la desviación estándar de Y es igual a la ________________________ de X, cuando $Y = X + C$.

desviación estándar

---

78 Para la transformación $Y = X + C$ se tiene que la varianza de Y es igual a la ____________ de ________ y la ________________________ de Y es igual a la desviación estándar de X.

varianza, X, desviación estándar

---

79 En forma análoga se concluye que $S_Y^2 = S_X^2$ y $S_Y = (?)$ para $Y = X - C$.

$S_Y = S_X$

---

80 Si $S_X = 5$, ¿cuánto valen S_Y^2 y S_Y? ______________

$S_Y^2 = 25$; $(S_Y^2 = 5^2 = 25)$, $S_Y = 5$; $(S_Y = S_X = 5)$

---

81 Si Y=XC, entonces

$$S_Y^2 = \frac{\sum_{i=1}^{N} (Y_i - \overline{Y})^2}{N}$$

y, en términos de $\overline{X}$, $\overline{Y} = (?)$.

$\overline{Y} = C\overline{X}$

---

82 Para $Y = CX$ se tiene $\overline{Y} = C\overline{X}$ y

$$S_Y^2 = \frac{\sum_{i=1}^{N} (Y_i - \overline{Y})^2}{N} = \frac{\sum_{i=1}^{N} [CX_i - (?)]^2}{N}$$

$$S_Y^2 = \frac{\sum_{i=1}^{N} [CX_i - C\overline{X}]^2}{N}$$

---

83 Factorizando a C en la ecuación para S_Y^2 se obtiene:

$$S_Y^2 = \frac{\sum_{i=1}^{N} [C(X_i - \overline{X})]^2}{N} = \frac{\sum_{i=1}^{N} C^2 (X_i - \overline{X})^2}{N} = \frac{C^2 \sum_{i=1}^{N} (X_i - \overline{X})^2}{N}$$

donde $\dfrac{\sum_{i=1}^{N} (X_i - \overline{X})^2}{N} = \underline{\qquad\qquad}$ $(\overline{X}/S_X/S_X^2)$.

S_X^2

---

84 Por lo tanto,

$$S_Y^2 = C^2 S_X^2 .$$

es decir, la varianza de $Y = CX$ es igual a la constante elevada al cuadrado y multiplicada por la ________________ de X.

varianza

---

85 Si $Y = CX$, entonces

$$S_Y^2 = (?)\ (?)$$

$S_Y^2 = C^2 S_X^2$

---

86 En forma análoga se obtiene, para la transformación $Y = X/C$,

$$S_Y^2 = S_X^2/C^2$$

y, por lo tanto, $S_Y = (?)/(?)$.

$S_y = S_x/C$

87 Si $Y = CX$, $S_y^2 = (?)$ y $S_y = (?)$.

$S_Y^2 = C^2 S_x^2$; $S_y = C\,S_x$

88 Si $Y = X/C$, $S_Y^2 = (?)$ y $S_y = (?)$.

$S_Y^2 = S_x^2/C^2$, $S_y = S_x/C$

89 Relacione las transformaciones de la izquierda con las fórmulas de la derecha.

1. $Y = X - C$ I. $S_y = S_x/C$
2. $Y = X/C$ II. $\overline{Y} = \overline{X} + C$
3. $Y = X + C$ III. $S_y = S_x$
   IV. $\overline{Y} = \overline{X}/C$

1-III,   2-I, IV,   3-II, III

90 Relacione las transformaciones de la izquierda con las fórmulas de la derecha.

1. $Y = CX$ I. $\overline{Y} = \overline{X} - C$
2. $Y = X - C$ II. $\overline{Y} = \dfrac{\overline{X} - C_1}{C_2}$
3. $Y = \dfrac{X - C_1}{C_2}$ III. $S_Y^2 = C^2\,S_x^2$
   IV. $Y = C\overline{X}$

1–III, IV,   2–I,   3–II

91 Para la doble transformación

$$Y = \frac{X - C_1}{C_2}$$

se obtiene $S_Y^2 = S_X^2/C_2^2$ y $S_Y = (?)/(?)$.

$S_y = S_x/C_2$

92 Entonces, para $Y = (X - C_1)/C_2$ se tiene que (escriba las fórmulas):

$\overline{Y} =$

$S_Y^2 =$

$\overline{Y} = (\overline{X} - C_1)/C_2$, $S_Y^2 = S_X^2/C_2^2$

93 Si en la transformación $Y = \dfrac{X - C_1}{C_2}$ tomamos $C_1 = \overline{X}$, entonces

$$Y = \frac{(?) - (?)}{(?)}$$

$Y = \dfrac{X - \overline{X}}{C_2}$

94 Si en la transformación $Y = (X - C_1)/C_2$ tomamos $C_1 = \overline{X}$, entonces $\overline{Y} =$ ——

$\overline{Y} = 0$

95 Si tomamos $C_1 = \overline{X}$ y $C_2 = S_x$, tendremos

$$Y = \frac{(?) - (?)}{(?)}$$

$Y = \dfrac{X - \overline{X}}{S_x}$

96 Si $Y = \dfrac{X - \overline{X}}{S_x}$, $\overline{Y} =$ —— y $S_Y^2 =$ ——.

$\overline{Y} = 0$, $S_Y^2 = 1$

97 Entonces, para la transformación $Y = (X - \overline{X})/S_x$ se tiene que la media de los datos de Y es cero y la variancia es ________

uno

98 Las puntuaciones Y_i, obtenidas mediante la transformación $Y_i = (X_i - \overline{X})/S_x$ se llaman *puntuaciones estándar*.
Las puntuaciones estándar tienen un promedio igual a ________ (cuánto), y una desviación estándar igual a ________.

cero, uno

99 Las puntuaciones $Y_i = (X_i - \overline{X})/S_x$ se denominan puntuaciones ________.

estándar

100 Las puntuaciones estándar se obtienen mediante la transformación

$$Y_i = \frac{(?)}{(?)}$$

$$Y_i = \frac{X_i - \overline{X}}{S_x}$$

101 Para $Y = (X - \overline{X})/S_x$ se obtiene $\overline{Y}$ = ______ y S_y = ______.

$\overline{Y} = 0$, $S_y = 1$

102 Es común utilizar la letra Z para la transformación $(X - \overline{X})/S_x$.
En adelante, la letra Z nos servirá para identificar la transformación

$$Z = \frac{X - (?)}{(?)}$$

$$Z = \frac{X - \overline{X}}{S_x}$$

103 Entonces, las puntuaciones $Z_i = \frac{(X_i - \overline{X})}{S_x}$ son las puntuaciones ________.

estándar

104 ¿Qué símbolo utilizaremos para denotar las puntuaciones estándar?

Z_i

105 Así como formulamos un método corto (rápido o abreviado) para calcular $\overline{X}$, podemos formular otro para calcular S_x^2. También aquí utilizaremos la transformación $Y = (X - C_1)/C_2$, tomando C_1 el valor ________ (de una marca de clase/del ancho de los intervalos) y C_2 el valor ________ (de una marca de clase/del ancho de los intervalos).

C_1 = una marca de clase, C_2 = ancho de los intervalos

106 Al asumir C_1 el valor de una marca de clase, ésta debe estar cercana al valor que *esperamos* tenga $\overline{X}$.

Si $Y = (X - C_1)/C_2$, $S_y^2 = (?) / (?)$.

$S_y^2 = S_x^2/C_2^2$

107 Puesto que para la transformación $Y = (X - C_1)/C_2$, $S_y^2 = S_x^2/C_2^2$, entonces $S_x^2 = (?)\ (?)$.

$S_x^2 = C_2^2\ S_y^2$

108 Si calculamos S_y^2, podemos calcular S_x^2 mediante la fórmula

____

$S_x^2 = C_2^2\ S_y^2$

109 La fórmula simplificada para calcular S_X^2 es $S_X^2 = \overline{X^2} - \overline{X}^2$, donde $\overline{X}$ y $\overline{X^2}$ son respectivamente, la media y el valor medio cuadrático de los datos de X. Por analogía se tiene que $S_Y^2 = \overline{Y^2} - (?)$.

$$S_Y^2 = \overline{Y^2} - \overline{Y}^2$$

110 Recuerde que si los datos están agrupados en intervalos de clase, la fórmula para calcular el promedio cuadrático es

$$\overline{Y^2} = \sum_{i=1}^{K} f_i\, Y_i^2/N$$

y la fórmula para obtener la media es

$$\overline{Y} = (?)$$

$$\overline{Y} = \sum_{i=1}^{K} f_i\, Y_i/N$$

111 Si los datos están agrupados en intervalos de clase, la fórmula para obtener el valor medio cuadrático de los datos de Y es

$$\overline{Y^2} = (?)$$

$$\overline{Y^2} = \sum_{i=1}^{K} f_i\, Y_i^2/N$$

112 En la siguiente tabla se presenta la distribución de frecuencias de los datos de la variable aleatoria "salario del personal de confianza de una empresa". Calcule la variancia de Y. Complete la tabla usando C_1 = 5,500 pesos y C_2 = 1,000 pesos.

Marca de clase, X_i en pesos	f	Y	fY	Y^2	fY^2
2,500	5				
3,500	10				
4,500	35	−1	−35		
5,500	35	0	0		
6,500	10	1	10		
7,500	5	2	10		
TOTAL:					

$\overline{Y}$ = ______ = ______

$\overline{Y^2}$ = ______ = ______

S_Y^2 = ______

Marca de clase, X_i en pesos	f	Y	fY	Y^2	fY^2
2,500	5	−3	−15	9	45
3,500	10	−2	−20	4	40
4,500	35	−1	−35	1	35
5,500	35	0	0	0	0
6,500	10	1	10	1	10
7,500	5	2	10	4	20
TOTAL:	100		−50	-	150

$$\overline{Y} = \frac{-50}{100} = -0.5$$

$$\overline{Y^2} = \frac{150}{100} = 1.50$$

$$S_Y^2 = \overline{Y^2} - \overline{Y}^2 = 1.50 - (0.5)^2 = 1.25$$

113 En vista de que $S_Y^2 = 1.25$, S_X^2 = ______ y S_X = ______

(C_2 = 1,000 pesos.)

$S_X^2 = 1.25\ C_2^2 = 1{,}250{,}000$ pesos², $S_X = \sqrt{1{,}250{,}000} = 1{,}118$ pesos

114 En la siguiente tabla se presenta la distribución de frecuencias de la capacidad de carga, X, de unas vigas de concreto. Calcule la variancia de los datos de X.

Marca de clase, X_i en ton	f	Y	fY	Y^2	fY^2
2,150	1				
2,250	5				
2,350	10				
2,450	9				
2,550	20				
2,650	50				
2,750	5				
TOTAL:					

C_1 = 2,550 ton.

C_2 = ______

$\overline{Y}$ = ______ = ______

$\overline{Y^2}$ = ______ = ______

S_Y^2 = ______

S_X^2 = ______

Marca de clase, X_i en ton	f_i	Y	fY	Y^2	fY^2
2,150	1	−4	−4	16	16
2,250	5	−3	−15	9	45
2,350	10	−2	−20	4	40
2,450	9	−1	−9	1	9
2,550	20	0	0	0	0
2,650	10	1	10	1	10
2,750	5	2	10	4	20
TOTALES	100		12		180

$C_1 = 2{,}550$ ton

$C_2 = 100$ ton

$$\overline{Y} = \frac{\sum_{i=1}^{7} f_i Y_i}{N} = \frac{12}{100} = 0.12$$

$$\overline{Y^2} = \frac{\sum_{i=1}^{7} f_i Y_i^2}{N} = \frac{180}{100} = 1.8$$

$$S_y^2 = \overline{Y^2} - \overline{Y}^2 = 1.8 - (0.12)^2 = 1.7856$$

$$S_x^2 = S_y^2\ C_2^2 = 1.786 \times 100^2 = 17856 \text{ ton}^2$$

115 Si $C_1 = 2{,}550$ ton, $C_2 = 100$ ton, $\overline{Y} = 0.12$ y $S_x^2 = 17{,}856$ ton

a. ¿cuánto vale $\overline{X}$? ______

b. ¿cuánto vale la desviación estándar de los datos de X? ______

2,562 ton ($\overline{X} = C_2\ \overline{Y} + C_1 = 12 + 2{,}550 = 2{,}562$ ton).

$S_x = \sqrt{17{,}856} = 133.63$ ton.

PARTE E. PUNTUACIONES ESTANDAR

116 Recuerde que las puntuaciones $Z_i = (X_i - \overline{X})/S_x$ se denominan puntuaciones ______

estándar

117 Las puntuaciones estándar se denotan con la letra ______.

Z_i

118 Al hacer la transformación de los datos en puntuaciones estándar, la distribución de frecuencias no se altera, sino sólo cambia la escala horizontal.

Observe la figura 7.1. En ella se ha trazado un ______ (histograma/polígono de frecuencias)

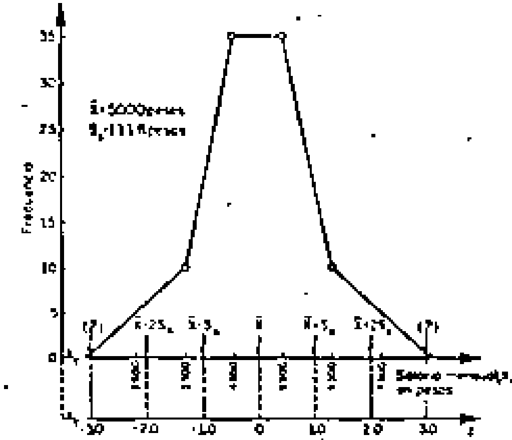


Fig 7.1

polígono de frecuencias

119 Observe que en la figura 7.1 se han trazado dos ejes horizontales con diferentes escalas, una correspondiente a los valores de los datos básicos, y la otra a las ______.

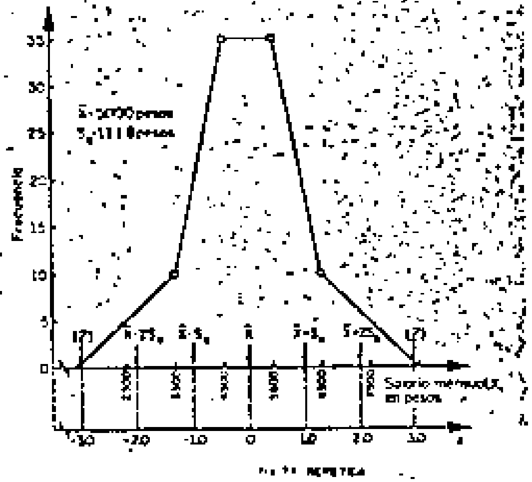
puntuaciones estándar

120 Observe que en la figura 7.1, en la escala horizontal superior se ha indicado, en términos de S_x, las correspondencias entre las dos escalas. Así, a Z = 1 corresponde un valor de X igual a $\overline{X} + S_x$, a Z = 2 corresponde $\overline{X} + 2S_x$, a Z = 3 corresponde ______, etcétera.

$\overline{X} + 3S_x$

121 Observe la figura 7.1.
A Z = − 1 corresponde un valor de X igual a $\overline{X} - S_x$, a Z = − 2 corresponde

$\overline{X} - 2S_x$, a $Z = -3$ corresponde ______



$\overline{X} - 3S_x$

---

122. Esta correspondencia entre las puntuaciones de X y Z se generaliza en la siguiente forma

Se tiene $Z_i = \dfrac{X_i - \overline{X}}{S_x}$

por lo cual $X_i - \overline{X} = Z_i S_x$

de donde $X_i = (?) + (?)$.

$X_i = \overline{X} + Z_i S_x$

---

123 Puesto que $X_i = \overline{X} + Z_i S_x$, para $Z_i = 1$ se tiene $X_i = X + S_x$, para $Z_i = -5$ se tiene $X_i =$ ________.

$\overline{X} - 5S_x$

---

124 Ya que $X_i = \overline{X} + Z_i S_x$, las Z_i nos indican qué tan alejados de X, en términos de S_x, se localizan los diversos valores de los datos básicos. Así, si $Z_i = 3$, se dice que X_i está a tres desviaciones estándar de $\overline{X}$. Si $Z_i = -2$ se dice que X_i está a ______ desviaciones estándar de ________.
(cuántas)

menos dos (-2), $\overline{X}$

---

125 A una puntuación situada a dos desviaciones estándar de X le corresponde $Z_i =$ ____.

2

---

126 La forma de la distribución de frecuencias ______ se altera al transformar de puntuaciones X_i a Z_i.
(no/sí)

no

---

127 Si Z_i asume un valor negativo, entonces la X_i correspondiente es ________ que X.
(menor/mayor)

menor

---

128 Los valores positivos de la Z_i corresponden a valores de la X_i ________ que X.
(menores/mayores)

mayores

---

129 Las puntuaciones estándar, $Z_i = (?) / (?)$, al igual que los percentiles, permiten comparar los resultados que se obtienen en dos diferentes experimentos realizados por una misma persona. Por ejemplo, sirven para contestar preguntas como ésta:

¿En qué examen tuvo Jesús Rivas mejor actuación con respecto a un grupo: en el de "inteligencia" o en el de la "aptitud hacia la Estadística", si en ambos obtuvo una calificación de 69?

$Z_i = (X_i - \overline{X})/S_x$

130 Observe la figura 7.2. En ella se muestran ______ (a/b/c)

a. dos histogramas
b. dos curvas de frecuencias relativas acumuladas
c. dos polígonos de frecuencias

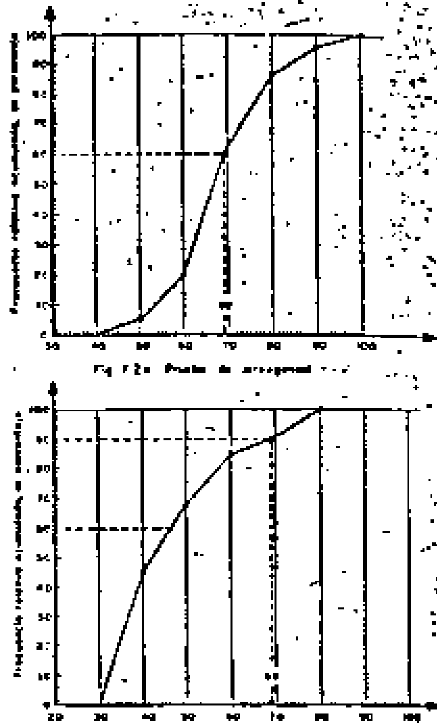


Fig. 7.2a Prueba de inteligencia

Fig. 7.2b Prueba de aptitud hacia la Estadística

Fig. 7.2

b. dos curvas de frecuencias relativas acumuladas

131 Recuerde que los percentiles son los valores de la variable que dividen una distribución de frecuencias acumuladas, en ______ (10/2/100) partes iguales.

100

132 Observe la figura 7.2a.
¿A qué percentil corresponde un valor de 69 de la variable?

Al 60º percentil

133 Observe la figura 7.2b.
¿A qué percentil corresponde un valor de 69 de la variable?

Al 90º percentil.

134 El que una calificación igual a 69 corresponda al 60º percentil en la prueba de inteligencia (el 60% de los alumnos tuvieron calificación menor que 69), y al 90º percentil en la prueba de aptitud, indica que, con referencia al grupo, se tuvo ______ (mejor/peor) desempeño en la prueba de aptitud.

mejor

135 Use la figura 7.2.
¿Qué calificación se necesita obtener para que el desempeño en la prueba de aptitud sea igual al de la prueba de inteligencia; es decir, qué calificación se necesita para que en ambas pruebas se obtenga el 60º percentil?

47 (aproximadamente)

136 Use las figuras 7.2a y 7.2b.
Si Horacio Molina obtuvo 80 de calificación en la prueba de inteligencia y 55 en la de aptitud:

a. ¿a cuál percentil corresponde la calificación de 80 en la prueba de inteligencia?
b. ¿a cuál corresponde la calificación de 55 en la prueba de aptitud?
c. ¿en cuál prueba tuvo mejor desempeño con respecto al grupo que realizó los exámenes?

a. Al 86º, b. Al 76º, c. En la prueba de inteligencia.

137 Si no contamos con la distribución de frecuencias, y sólo sabemos que el promedio de la prueba de inteligencia es 68.2 y la desviación estándar es 17.4, podemos utilizar las puntuaciones estándar para hacer un estudio semejante al de los cuadros anteriores. ¿Qué valor de Z le corresponde a la calificación 69?

0.046; $(Z = \frac{69-68.2}{17.4} = \frac{0.8}{17.4} = 0.046)$

138 En forma análoga, si sólo sabemos que la media en la prueba de aptitud es 46.4 y la desviación estándar es 13, el valor que se obtiene para la puntuación estándar correspondiente a la calificación 69 es ______ (cuánto)

1.74; $(\frac{69-46.4}{13} = \frac{22.6}{13} = 1.74)$

139 Una vez obtenidos los valores de las puntuaciones estándar correspondientes a la calificación 69 en las dos pruebas, se puede decir en cuál tuvo mejor desempeño con respecto al grupo. Esto corresponde a la prueba en la cual se obtuvo la ____________ puntuación estándar.
(menor/mayor)

mayor

140 En la prueba de inteligencia Jesús Rivas obtuvo $Z = 0.046$ y en la prueba de aptitud obtuvo $Z = 1.74$. Estos resultados nos indican que el mejor desempeño lo tuvo en la prueba de ________, puesto que la calificación estándar, es más alta.

aptitud

141 Horacio Molina logró 75 de calificación en la prueba de inteligencia para la cual el promedio fue de 68.2 y la desviación estándar fue de 17.4. En la prueba de aptitud obtuvo 65, para la cual la media fue de 46.4 y la desviación estándar fue de 13. ¿En cuál prueba tuvo mejor desempeño, respecto al grupo que presentó los exámenes? (diga por qué). ____________________

Tuvo mejor desempeño en la prueba de aptitud porque para ella $Z = 1.43$, la cual es mayor que $Z = 0.39$ que corresponde a la prueba de inteligencia.

142 Cuando no disponemos de las distribuciones de frecuencias, podemos hacer uso de las puntuaciones estándar para comparar el desempeño relativo de un mismo individuo en dos pruebas (o experimentos) distintos. Para calcular las puntuaciones estándar, la única información que necesitamos son los valores de la ____________ y de la ____________

media, desviación estándar

REVISIÓN

143 Cuando se efectúan una o varias operaciones algebraicas con una variable aleatoria, se dice que se hace ________.
(a/b/c)

a. un agrupamiento de datos
b. un cálculo de frecuencias
c. una transformación

c. una transformación

144 ¿Cuánto vale la media, $\overline{Y}$, de $Y = X + C$, en términos de $\overline{X}$? ________

$\overline{Y} = \overline{X} + C.$

145 ¿Cuánto vale la variancia S_Y^2 de $Y = X - C$ en términos de S_X^2? ________

$S_Y^2 = S_X^2.$

146 Si $Y = X/C$, ¿cuánto vale $\overline{Y}$? ________

$\overline{Y} = \overline{X}/C.$

147 Si $Y = XC$, ¿cuánto vale S_Y^2? ________

$S_Y^2 = S_X^2\, C^2.$

148 Si $Y = (X - C_1)/C_2$, ¿cuánto valen $\overline{Y}$ y S_Y? ________

$\overline{Y} = (\overline{X} - C_1)/C_2.$ $S_y = S_x/C_2.$

149 Las puntuaciones estándar las denotamos con el símbolo ________.

Z_i

150 ¿Qué transformación se hace para obtener puntuaciones estándar? ________

$$Z_i = \frac{X_i - \overline{X}}{S_x}$$

151 El promedio de las puntuaciones estándar vale ________.

cero

---

152 La varianza de las puntuaciones estándar vale ________________.

1; $(S_y^2 = S_x^2/S_x^2 = 1)$

---

153 ¿Cambia la distribución de frecuencias al transformar datos básicos en puntuaciones estándar?

No.

---

154 Si en una prueba A se obtiene $Z = -0.90$ y en otra B se obtiene $Z = -0.68$, ¿en cuál se tuvo mejor desempeño respecto al grupo de individuos que presentó las pruebas?

B (en la que se obtuvo $Z = -0.68$).

---

155 Utilice la siguiente tabla para calcular la media, la varianza y la desviación estándar de los datos de la variable aleatoria, X, "calificación en Botánica", que en ella se presentan. Use los métodos cortos (rápidos o abreviados) con $C_1 = 65$.

Marca de clase, X	Frecuencia	(?)	(?)	(?)	(?)
45	5				
55	15				
65	43				
75	22				
85	10				
90	5				
TOTAL					

$N =$ ____________ ; $C_2 =$ ____________

$\overline{Y} =$ ____________ $=$ ____________ ; $\overline{X} =$ ____________

$\overline{Y^2} =$ ________ $=$ ________ ; $S_Y^2 =$ ____________

$S_x^2 =$ ____________ ; $S_x =$ ____________

Marca de clase, X	Frecuencia	Y	fY	Y^2	fY^2
45	5	−2	−10	4	20
55	15	−1	−15	1	15
65	43	0	0	0	0
75	22	1	22	1	22
85	10	2	20	4	40
95	5	3	15	9	45
TOTALES	100		32		142

$N = 100$; $C_2 = 10$

$\overline{Y} = \dfrac{32}{100} = 0.32$; $\overline{X} = 65 + 0.32\,(10) = 68.2$

$\overline{Y^2} = \dfrac{142}{100} = 1.42$; $S_Y^2 = 1.42 - 0.32^2 = 1.32$

$S_x^2 = C_2^2\, S_Y^2 = (100)\,(1.32) = 132$; $S_x = \sqrt{132} = 11.4$

EXAMEN

1. *Si efectuamos alguna operación algebraica sobre la variable aleatoria X, se dice que se hace una ____________.*

2. *Escriba las fórmulas para calcular $\overline{X}$ en términos de $\overline{Y}$ para las siguientes transformaciones:*

   a. $Y = XC$ ______________________.
   b. $Y = X-C$ ______________________.
   c. $Y = X/C$ ______________________.
   d. $Y = X+C$ ______________________.

3. *Escriba las fórmulas para calcular S^2_x en términos de S^2_y para las transformaciones de la pregunta anterior.* ______________________________

4. *Si la transformación es $Y=(X-C_1)/C_2$, ¿cuánto valen $\overline{X}$ y S^2_x, en términos de $\overline{Y}$ y S^2_y?* ______________________________.

5. *¿Qué transformación se utiliza en los métodos cortos para calcular $\overline{X}$ y S^2_x?* ________.

6. *En los métodos cortos para calcular $\overline{X}$ y S^2_x, C_1 debe ser una ______________ ______________________ cuyo valor esté cercano al valor que se espera tome $\overline{X}$, y C_2 es igual a la ____________________ de los intervalos de clase.*

7. *Los resultados de una prueba de inteligencia (IQ) aplicada a 100 estudiantes universitarios, de 20 a 25 años de edad, se muestran en la siguiente tabla. Calcule $\overline{X}$, S^2_x y S_x mediante los métodos cortos. Tome $C_1 = 104.5$.*

Marca de clase	Frecuencia				
144.5	1				
134.5	3				
124.5	9				
114.5	17				
104.5	23				
94.5	22				
84.5	15				
74.5	6				
64.5	2				
54.5	2				
TOTAL	100				

(48 puntos)

8. *Escriba la fórmula para transformar datos básicos en puntuaciones estándar.* ________

9. *¿Cuál es la puntuación estándar correspondiente a $X = 135.1$ si se sabe que $\overline{X} = 100.5$ y $S_x = 17.3$?*

10. *En la pregunta anterior se obtuvo $Z = 2$; esto significa que $X = 135.1$ está ________ (cuántas) desviaciones estándar a la ______________ (izquierda/derecha) de $\overline{X}$.*

11. *Relacione los elementos de la columna de la izquierda con los de la derecha.*

1. Z es mayor de cero	a. X igual a $\overline{X}$
2. Z igual a cero	b. X es mayor de $\overline{X}$
3. Z es menor de cero	c. X es menor de $\overline{X}$

12. *¿Cambia la distribución de frecuencias al transformar los datos básicos a puntuaciones estándar?*

TOTAL: 70 puntos.

1. *transformación*

2. a. $\overline{X} = \overline{Y}/C$
   b. $\overline{X} = \overline{Y} + C$
   c. $\overline{X} = \overline{Y}C$
   d. $\overline{X} = \overline{Y} - C$

3. a. $S_x^2 = S_y^2/C^2$

   b. $S_x^2 = S_y^2$

   c. $S_x^2 = S_y^2\, C^2$

   d. $S_x^2 = S_y^2$

4. $\overline{X} = C_2\, \overline{Y} + C_1$

   $S_x^2 = C_2^2\, S_y^2$

5. $Y = (X - C_1)/C_2$

6. *marca de clase*
   amplitud

7.

Marca de clase	Frecuencia	y	yf	y^2	y^2f
144.5	1	4	4	16	16
134.5	3	3	9	9	27
124.5	9	2	18	4	36
114.5	17	1	17	1	17
104.5	25	0	0	0	0
94.5	22	−1	−22	1	22
84.5	15	−2	−30	4	60
74.5	6	−3	−18	9	54
64.5	2	−4	−8	16	32
54.5	2	−5	−10	25	50
TOTAL:	100		−40		314

$$\overline{Y} = \frac{-40}{100} = -0.4$$

$$\overline{X} = -0.4 \times 10 + 104.5 = 100.5$$

$$\overline{Y^2} = \frac{314}{100} = 3.14$$

$$S_y^2 = 3.14 - (-0.4)^2 = 2.98$$

$$S_x^2 = 2.98 \times 10^2 = 298$$

$$S_x = \sqrt{298} = 17.3$$

8. $Z = \dfrac{X - \overline{X}}{S_x}$

9. $Z = \dfrac{135.1 - 100.5}{17.3} = 2$

10. 2
    derecha

11. 1–b
    2–a
    3–c

12. No.

DIRECTORIO DE ALUMNOS DEL CURSO: PROBABILIDAD Y ESTADISTICA: FUNDAMENTOS Y APLICACIONES 1983.

1. Rosa M. Alatorre Salgado
   Comisión Nal. Bancaria y de Seguros
   Rép. del Salvador 47
   Centro
   Cuauhtémoc
   México,D.F.
   521 99 63

   Nte. 92 # 4537
   Nva. Tenochtitlán
   G.A.Madero
   México,D.F.

2. Javier Anaya Ruiz
   Cía. de Luz y Fza. del Centro
   Av. de los Angeles 89
   San Martín Xochinahuac,Azcapotzalco
   382 15 83

   Daniel Delgadillo 3
   Magisterial
   Vistabella
   Tlalnepantla, Edo. de Méx.
   398 41 85

3. Carlos Angeles Medina
   U N A M
   Coord. Ext. Universitaria
   Torre de Rectoría 12° Piso
   Coyoacán
   México,D.F.
   550 52 15 Ext.3235

   Unidad Esperanza 603
   Narvarte
   B.Juárez
   03020 México,D.F.
   519 44 74

4. Rubén Avila Espinosa
   IPESA Consultores, S.C.

5. Ricardo Aviles Escamilla
   C. F.E.
   Río Ródano 14-109 B
   Cuauhtémoc
   México,D.F.
   553 65 20

   Av.Fco. del Paso y T. 481
   J.Balbuena
   V.Carranza
   México,D.F.
   552 22 11

6. Celia Benet Jiménez
   Coordinación Gral. de Planif. Fam.
   S. S. A.
   Insurgentes Sur 1397-7°
   México,D.F.

   La Corona 321
   Col. Industrial
   G.A.Madero
   07800
   781 68 59

7. Sergio Antonio Bonilla Alonso
   Construcción Lozada
   16 de Septiembre 706
   72000 Puebla, Pue.
   41 44 37

8. Joaquín Bravo Pérez
   Esc. Sup. de Turismo
   I.P.N.
   Av. La Salle 39
   La Escalera
   G.A.Madero
   México,D.F.
   586 49 31

   Ahuehuetes 27
   Villa de las Flores
   Coacalco de B., México
   915 91 51207

9. Víctor E. Cocholia Flores
   ICA Nuclear , S.A.
   V. M. Alemán 81-2°
   Escandón
   México, D.F.
   515 42 89

   Serranía 267
   Pedregal Ote.
   Coyoacán
   04530 México,D.F.
   566 80 90

10. Ma. de la Luz Castro Sánchez
   ALCAN Aluminio, S.A. de C.V.
   Vía Morelos 347
   Tulpetlac Ecatepec, Edo. de Méx.
   569 93 00

   La Llanura 109
   Los Pastores
   Estado de México
   560 71 70

11. Jesús R. Cedeño Cedeño
   Aguascalientes 56-7
   Roma
   06700 México,D.F.
   584 51 86

12. Rubén A. Estrada Villegas
   Escuela Nacional de Trabajo Social

   Prolongación Uxmal 975-10 A
   Col. Gral. Anaya
   México,D.F.
   575 09 74

13. Margarita Ferat Toscano
   AHOP
   BLVD. Pilila No. 1
   Presa Sn. Joaquín
   Tecamachalco,Edo. de México
   Tlalnepantla
   589 07 75

   Adolfo Prieto 1425
   Casa C
   Del Valle
   B.Juárez
   03100
   575 10 17

14. Rosa M. Fuentes Solórzano
   Dir. Gral. de Planificación
   Familiar
   Insurgentes Sur 1397-7°
   México,D.F.
   563 54 78

   Av. 577 No.54
   Unidad Aragón
   796 51 80

15. Francisco J. Galindo Herrera
   IEM de México
   Legaria No. 653-4°
   Irrigación
   M. Hidalgo
   11500 México, D.F.
   557 85 88

   Michelet 51
   Anzures
   M. Hidalgo
   11590 México, D.F.
   545 11 12

16. Graciela Galindo Orozco
   Dirección General Adjunta de
   Contenidos y Métodos Educativos
   S. E. P.
   Cozumel 47-6° Piso
   Roma
   B.Juárez
   México, D.F.
   533 42 10

   Amores 1037-301
   Del Valle
   B.Juárez
   03100 México,.D.F.
   559 88 83

17. Genaro Guerrero Chávez
Preparatoria, Instituto B.Juárez
Cuauhtémoc 99
Coyoacán
México,D.F.

Calle 61 -125
Sta. Cruz Meyehualco
Iztapalapa
09290 México,D.F.
691 14 74

18. Raúl Gutiérrez Gutiérrez
ISSSTE
Viena 335
Del Carmen Coyoacán
México,D.F.
534 66 55

Av. México 1256-109
Sta. Teresa
Contreras
México,D.F.
568 29 01

19. Margarita Juárez Nájera
Centro de Ciencias de la Atmósfera
Circ. Ext. de C.U.
Coyoacán
México,D.F.
548 81 90

Norte 72 A # 5414
Bondojito
G.A.Madero
07850 México,D.F.
551 67 30

20. Gustavo Manzo García
S. C. T.
Jalapa 147-2°
Roma
Cuauhtémoc
06700 México,D.F.
574 82 17

Chihuahua 110-503
Roma
Cuauhtémoc
06700 México,D.F.
574 01 18

21. Sergio Martín Moreno
UNAM

Xochicalco 52-402
Narvarte
B.Juárez
03020 México,D.F.
530 96 07

22. Jorge Mendoza Larraquivel
Ing. de Sistemas
Legaría 853
Irrigación
M.Hidalgo
11500 México,D.F.
557 85 89

Sur 73 No.4311
Viaducto Piedad
Iztacalco
519 90 65

23. Eliut Rafael Quintana Bache
Romero de Terreros 1009-8
Del Valle
México,D.F.
584 51 86

24. Fausto A.Ramos Danache
Hacienda de Torrecillas 39
Villa Quietud
Coyoacán
04960 México,D.F.
594 45 54

25. Víctor M. Ramírez Gallegos
Construestudios, S.A. de C.V.
Gabriel Mancera 1258-B-5
Del Valle
03100 México,D.F.
575 82 61

Fco. I. Madero 32
Col. Nva. Aragón
Ecatepec, Edo. de México

26. Bernardina Rodríguez Arroyo
Av. Amsterdam 213-602
Hipódromo Condesa
Cuauhtémoc
06100 México, D.F.
584 03 22

27. Guillermo F. Salazar Valdez
Litografía 212
20 de Nov.
V.Carranza
15300 México,D.F.
789 56 13

28. Carlos M. Tovar Massanille
S. E. P.
Dirección Gral. de Organización y Métodos
Centeno 670 -5°
Granjas México
Iztacalco
650 19 62

Antillas 614-304
Portales
B.Juárez
03300 México,D.F.

29. Alberto Velázquez Jiménez
Edgar Alan Poe 48-1
Polanco
M.Hidalgo
11560 México,D.F.
250 23 62

Insurgentes Sur 2387
San Angel
A.Obregón
050 90 00 Ext.197

30. Rubén Yesin Toledo
Banco de México
Condesa 6-2° Piso
Cuauhtémoc
México,D.F.
518 05 00 Ext.669

Sur 59 No. 148
Prado Ermita
B.Juárez
03590 México,D.F.
539 95 17