PROBABILIDAD Y ESTADISTICA (Fundamentos y Aplicaciones)

1 9 8 5 .

Fecha	Tema	Horario	Profesor
Del 1° al 27 de Febrero	INTRODUCCION ESTADISTICA DESCRIPTIVA PROBABILIDAD	18 a 21 h c/día Lunes, miércoles y viernes.	DR. OCTAVIO A. RASCON M. EN I. RUBEN TELLEZ SANCHEZ
Del 1° de Marzo al 17	R E C E S O	VI FERIA INERNACIONAL DEL LIBRO	
18 al 29 de Marzo	INFERENCIA ESTADISTICA	18 a 21 h c/día Lunes, miércoles y viernes.	M. EN I. AUGUSTO VILLARREAL ARANDA M. EN I. RUBEN TELLEZ SANCHEZ

'edcs.

PROBABILIDAD Y ESTADISTICA: FUNDAMENTOS Y APLICACIONES

BIBLIOGRAFIA

ENERO DE 1985

PROBABILIDAD Y ESTADISTICA : FUNDAMENTOS Y APLICACIONES

Coordinador: Dr. Octavio A. Rascón Chávez

El siguiente material se encuentra disponible en el Centro de Información y Documentación "Ing. Bruno Mascanzoni".

CANAVOS, George C. Applied Probability and Statistical Methods. -- Boston : Little, Brown and Company, 1984.

ROHATGI, Vijay K. Statistical Inference. -- New York : John Wiley & Sons, 1984.

BARR, Donald R. and Peter W. Zehna. Probability : modeling Uncertainty. -- Massachusetts : Addison-Wesley Publishing Company, 1983.

SIDDALL, James N. Probabilistic Engineering Design. -- New York : Marcel Dekker, Inc., 1983.

HUNTER, Jeffrey J. Mathematical Techniques of Applied Probability. Volume 1. -- New York : Academic Press, Inc., 1983.

BIBLIOGRAFIA ADICIONAL

CAPITULOS DE LIBROS DISPONIBLES EN EL CENTRO DE INFORMACION Y DOC.

"Introduction". -- p 1-7. -- En : Montgomery, Duglas C. Design and Analysis of Experiments. -- New York : John Wiley & Sons, 1984.

"Simple Comparative Experiments". -- p. 9-42. -- En : Montgomery, Duglas C. Design and Analysis of Experiments. -- New York : John Wiley & Sons, 1984.

"Statistical Technology of Quality". -- p. 343-512 -- En : Feigenbaum A. V. Total Quality Control. -- New York : Mc Graw-Hill, 1983

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 CH.

FEBRERO, 1985

Palacio de Minería Calle de Tacuba 5 primer piso Deleg. Cuauhtémoc 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:

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

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

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

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: $b \pm t_c\,\sigma_b$

$$b \pm t_c\,S_{y|x}\sqrt{\frac{\bar{x}^2}{nS_x^2}}$$

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_x^2 ES LA VARIANCIA (SESGADA) DE LA MUESTRA DE x.

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

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

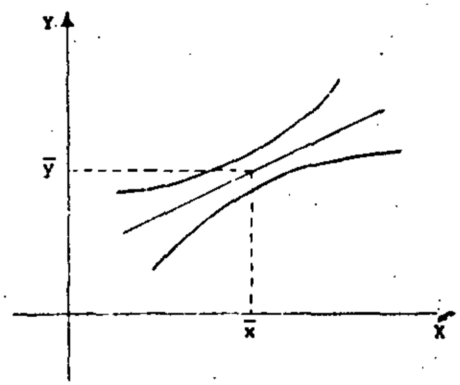
c. PARA LA PREDICCION, Y_i: $\tilde{Y}_i \pm t_c \sigma_{\tilde{Y}}$

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

SI x_i ESTA DENTRO DEL RANGO DE LA MUESTRA, O

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

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, lt	20.2	23.1	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 :

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

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

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

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

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.

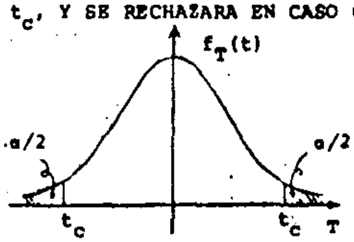
SI SE DESA PROBAR LA HIPOTESIS

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

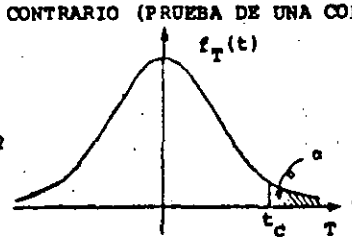
BASTA SUSTITUIR A $B = b_0$ EN LA ECUACION ANTERIOR Y EVALUAR $T = t$, ES DECIR,

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

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	0.08	0.23	0.60	0.39	0.55	0.75	0.81	0.85

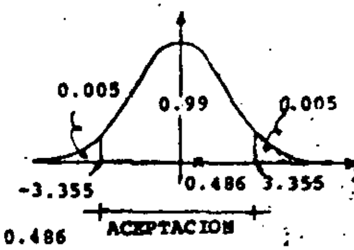
$m = 0.093$, $b = 0.032$, $S^2_{y|x} = 0.01258$

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

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

$$t = \frac{m - m_0}{\frac{S_{y|x}}{s_x\sqrt{n}}} = \frac{0.093 - 0.1}{\frac{\sqrt{0.01258}}{\sqrt{8.25 \times 10}}} = 0.567 < 3.355$$

SE ACEPTA H_0 CON 99% DE NIVEL DE CONFIANZA.

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

PRUEBA

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

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 HIPOTESIS DE QUE $\rho_{xy} = 0$. USAR $\alpha = 0.05$.

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

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

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

$\therefore$ 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

A N E X O S

FEBRERO, 1985

PR[illegible]ABILIDAD Y ESTADISTICA PARA IN[illegible]NIEROS - I. MILLER, J. FREUND

CAP. 11 . EDITORIAL REVERTE

4. Utilizando las 20 lecturas de temperatura dadas en el problema 4 de la página 114, estimar la media de la población a partir de la cuartila media.
5. Empleando la distribución de pesos de recubrimientos de estaño obtenidos en el ejercicio 3 de la página 106, determinar la mediana y la cuartila media y comparar con la media obtenida en el problema 10 de la página 114.
6. Usar la distribución de velocidades obtenida en el problema 10 de la página 115 para calcular la cuartila media y estimar μ. Comparar este resultado con $\bar{x}$ y con una estimación de μ obtenida dibujando la gráfica en papel de probabilidades.
7. Utilizar los dos estimadores de cuantilas de la página 197 para dar una estimación de σ para los ingresos no agrupados del problema 6 de la página 105. Calcular la desviación típica muestral y comprarar las estimaciones. Estimar, también, σ tomando un cuarto de la diferencia entre la media del 5% superior y el 5% inferior de los datos.
8. Con respecto a las lecturas de temperatura del ejercicio 4 de la página 114, estimar la desviación típica de la población utilizando (a) la desviación típica muestral, (b) las dos fórmulas dadas en la página 197, y (c) una gráfica de probabilidad. Comparar los resultados obtenidos.
9. Emplear la distribución de pesos de recubrimiento de estaño obtenida en el problema 3 de la página 106 para calcular los dos estimadores de cuantilas de σ dados en la página 197. También, comparar con el valor de la desviación típica de la muestra obtenida en el problema 11 de la página 114.
10. Comparar el valor estimado de σ obtenido de la gráfica de probabilidad del ejercicio 6 con los valores estimados basados en las dos fórmulas de la página 197. Calcular las cuantilas necesarias a partir de los datos agrupados.

11.3 Tests de los signos

En esta sección describiremos tests no paramétricos basados en clasificar los datos de acuerdo con dos tributos, representados convenientemente por *signos más* y *signos menos*. Por ejemplo, si queremos contrastar la hipótesis nula $H_0: \mu = \mu_0$ sobre la base de una muestra de azar de tamaño n, podemos cambiar cada observación que exceda de μ_0 por un signo más y, cada observación menor que μ_0, por un signo menos. Si la población de la que obtenemos las muestras es *continua y simétrica* la probabilidad de que una observación sea cambiada por un signo más es igual a 1/2 cuando H_0 es cierta. En consecuencia, la prueba de la hipótesis nula $\mu = \mu_0$ es equivalente a una prueba de la hipótesis nula $p = 1/2$, donde p es el parámetro de una distribución binómica. La alternativa bilatera $\mu \neq \mu_0$ es equivalente, ahora, a $p \neq 1/2$, y las alternativas unilaterales $\mu < \mu_0$ y $\mu > \mu_0$ son equivalentes a $p < 1/2$ y $p > 1/2$, respectivamente, siendo p la probabilidad de encontrar un signo más, es decir, una observación mayor que μ_0.

Para ilustrar el *test de los signos de una sola muestra* que acabamos de describir, vamos a contrastar la hipótesis de que la temperatura media a la que opera un termostato es 28° C., utilizando los resultados siguientes obtenidos de 20 termostatos:

29.9	28.2	32.0	30.5	29.3	30.1	27.7	31.4	28.6	27.9
+	+	+	+	+	+	−	+	+	−
26.8	30.3	29.0	28.8	28.0	31.4	32.1	27.8	31.7	29.2
−	+	+	+		+	+	−	+	+

Notemos que hay 15 observaciones mayores que 28.0, 4 observaciones menores que 28.0 y una observación igual a 28.0. Aunque la probabilidad de que una observación, en una población continua, sea exactamente igual a 28.0, es nula, los números anteriores están redondeados y, como no sabemos si el número 28.0 representa un valor mayor o menor que 28, descartamos esta observación. Así pues, debemos determinar si los 15 signos más y los 4 signos menos, o sea, 15 "casos favorables" en 19 pruebas, comprueban la hipótesis de que $p = 1/2$. Aplicando el criterio del test exacto dado en la página 180 con $\alpha = 0.05$, encontramos en la tabla de probabilidades binómicas que $k_{.025} = 4$ y $k'_{.025} = 15$; como hubo 15 signos más y 4 signos menos, se deduce que la hipótesis nula debe ser rechaza. Nótese que, si el tamaño de la muestra es suficientemente grande, podemos usar la curva normal como aproximación de la distribución binómica y los tests dados en la tabla de la página 164.

Utilizando muestras apareadas, podemos extender inmediatamente el test de los signos a test de diferencias entre dos medias de población. En este caso, el test de los signos se puede emplear como una alternativa no paramétrica del test t para muestras apareadas introducida en la página 157. Dadas n observaciones apareadas, con la primera observación proviniente de la población uno y la segunda de la población dos, usamos un *signo más* para reemplazar cada par para el cual la observación de la primera población excede a la de la segunda, y un *signo menos* para substituir cada par en el que la observación de la segunda población excede al de la primera. En el caso en que dos observaciones pareadas sean iguales, se omite esta pareja y la prueba se desarrolla como en el caso de una sola muestra antes descrita.

Para ilustrar el *test de los signos de muestras apareadas*, vamos a comparar dos métodos para anodizar aluminio atendiendo a la apariencia de las piezas anodizadas (brillo, color, etc.). Aunque resulta difícil asignar valores numéricos a estas cualidades, no es difícil comparar piezas apareadas y decidir cuál tiene el aspecto más agradable. Supongamos que 40 unidades apareadas se juzgan, dando un signo más o un signo menos a cada par, de acuerdo con que el anodizado del primer método o el del segundo se considere superior. Dado que hubo 24 signos más, 11 signos menos y 5 empates, queremos probar si el primer método es realmente superior. Empleando la curva normal (ver página 180) y $\alpha = 0.05$ calculamos primero

$$z = \frac{24 - 17.5}{\sqrt{35 \cdot \frac{1}{2} \cdot \frac{1}{2}}} = 2.20$$

y como este valor excede a 1.645, valor crítico para un test unilatero, con un nivel de significación de 0.05, concluimos que la hipótesis nula debe ser rechazada. En otras palabras, concluimos que el primer método de anodización es mejor.

La eficacia del test de los signos es bastante alta para muestras pequeñas, 95% para $n = 6$, pero disminuye a medida que el tamaño de la muestra aumenta hasta llegar a una eficacia límite de 63%. Las hipótesis necesarias para aplicar el test de los signos, lo mismo en el caso de una sola muestra que en el de muestras apareadas, son que las poblaciones consideradas sean continuas y simétricas. Si la población no fuese continua, podría haber una probabilidad positiva de que una observación fuera realmente igual a μ_0 en el caso de una sola muestra, o que las observaciones apareadas fueran exactamente iguales en el caso de muestras apareadas. Entonces, dejará de ser válida la hipótesis de ser $p = 1/2$, a menos que impongamos mayores restricciones. Si la población (o poblaciones) no fueran simétricas, la probabilidad de que una observación fuera mayor que la media, o de que la diferencia entre dos observaciones apareadas fuera mayor que cero, no iguala necesariamente un medio, bajo la hipótesis nula $\mu = \mu_0$ en el caso de una sola muestra, o $\mu_1 = \mu_2$ en el caso de muestras apareadas. Sin embargo, es posible modificar el test de los signos para eliminar la hipótesis de simetría. Para concluir esto, sólo tenemos que considerar las hipótesis concernientes a las medianas de la población en lugar de las concernientes a las medias.

11.4 Tests por suma de números de orden

El test de los signos de muestras apareadas es uno de los métodos no paramétricos para contrastar la hipótesis nula de que dos muestras provienen de poblaciones continuas idénticas, frente a la alternativa de que las poblaciones tienen medias diferentes. Una clase altamente eficaz de tests no paramétricos de esta hipótesis, y otras similares, se basa en la *suma de los números de orden;* esto es, se dan números de orden a las observaciones de acuerdo con su magnitud y los tests se realizan sobre la base de ciertas sumas de estos número de orden. En esta sección, introduciremos tres tests basados en sumas de números de orden. *El test U de Mann-Whitney* se presenta como un substituto del test t de dos muestras, y tiene una eficacia límite de 95.5% cuando las hipótesis necesarias para el test t correspondiente quedan satisfechas. Un test similar al test U, que se puede emplear cuando la hipótesis alternativa especifica que las dos poblaciones tienen *dispersiones* diferentes, se considerará a continuación. Finalmente, introduciremos el *test H de Kruskal-Wallis H* para contrastar si k muestras provienen de poblaciones idénticas, frente a la alternativa de que las poblaciones tengan medias diferentes. Como el test U, el test H tiene, también, una eficacia de 95.5% cuando se compara con el procedimiento "normal" correspondiente, que se analizará en el capítulo 13.

Vamos a describir, en primer lugar, el test U de Mann-Whitney por medio del ejemplo siguiente. Supongamos que queremos comparar dos instrumentos de control o "registradores" diferentes utilizados para la determinación del contenido de humedad dentro de un semiconductor, a partir de las siguientes corrientes medidas en microamperes:

Registrador A:	1.3	0.9	0.8	0.2	0.4	0.6	0.1	5.1	0.2	
Registrador B:	1.7	3.5	7.8	0.9	0.7	2.6	0.2	1.5	15.3	0.7

Primero, ordenamos *conjuntamente* las 19 observaciones de acuerdo con su tamaño, reteniendo la identidad de la muestra en cada observación. Después, asignamos a estas observaciones los puestos 1, 2, 3, . . . y 19, como se indica en la tabla siguiente:

Registrador:	*A*	*A*	*A*	*B*	*A*	*A*	*B*	*B*	*A*	*A*	*B*	*A*
Observación:	0.1	0.2	0.2	0.2	0.4	0.6	0.7	0.7	0.8	0.9	0.9	1.3
Rango:	1	3	3	3	5	6	7.5	7.5	9	10.5	10.5	12

Registrador:	*B*	*B*	*B*	*B*	*A*	*B*	*B*
Observación:	1.5	1.7	2.6	3.5	5.1	7.8	15.3
Rango:	13	14	15	16	17	18	19

Nótese que, si dos o más observaciones están empatadas en el mismo lugar, damos a cada una de ellas el medio de los lugares que ocupan en conjunto.

Si denotamos los tamaños respectivos de las muestras por n_1 y n_2 y la suma de los números de orden ocupados por la primera muestra por R_1, se puede demostrar que la media y la varianza de la distribución muestral del estadístico

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

están dadas por

$$\mu_U = \frac{n_1 n_2}{2} \qquad \text{y} \qquad \sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

(Si hay empates en la ordenación, estas fórmulas son correctas sólo aproximadamente). Si n_1 y n_2 son, ambas, mayores que 8, la distribución del estadístico U se puede aproximar con bastante exactitud por una distribución normal y, por lo tanto, el test se puede basar en el estadístico

$$z = \frac{U - \mu_U}{\sigma_U}$$

y en la tabla III. Existen, también, tablas en las que se pueden basar tests exactos cuando n_1 y n_2 son pequeños (ver la tabla de D. B. Owen mencionada en la bibliografía). Observemos que no tiene consecuencias saber qué muestra se considera "primera", por lo que podemos trabajar con cualquier suma de números de orden, escogiendo la que sea más fácil de obtener.

Volviendo ahora a nuestro ejemplo, tenemos $n_1 = 9$, $n_2 = 10$, $R_1 = 66.5$, y, por consiguiente,

$$U = 9 \cdot 10 + \frac{9 \cdot 10}{2} - 66.5 = 68.5$$

$$\mu_U = \frac{9 \cdot 10}{2} = 45.0$$

$$\sigma_U^2 = \frac{9 \cdot 10 \cdot 20}{12} = 150.0$$

Así,

$$z = \frac{68.5 - 45.0}{\sqrt{150.0}} = 1.93$$

y como este valor se encuentra entre -1.96 y 1.96, valores críticos de una alternativa bilatera con $\alpha = 0.05$, llegamos a la conclusión de que la hipótesis nula de poblaciones idénticas no se puede rechazar.

Si se asignan los números de orden de alguna otra forma diferente, se puede emplear el mismo estadístico U para contrastar la hipótesis nula de poblaciones idénticas frente a la alternativa de que las poblaciones tienen *dispersiones distintas.* Los números de orden se asignan "desde ambos extremos hacia el medio", dando el número 1 a la menor observación; los números 2 y 3, a la mayor y segunda mayor observaciones; los números 4 y 5, a la segunda y tercera menores; los 6 y 7, a la tercera y cuarta mayores, y así sucesivamente. Todos los demás aspectos de este test con dispersiones diferentes son idénticos a los del test U de Mann-Whitney.

El test Kruskal-Wallis para decidir si k muestras independientes provienen de poblaciones idénticas, se desarrolla en una forma similar al test U. Como antes, las observaciones se tratan *en conjunto* para darles el lugar de orden, y si R_i es la suma de los números de orden ocupados por las n_i observaciones de la i-ésima muestra, el test se basa en el estadístico

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

donde $n = n_1 + n_2 + \ldots + n_k$. Cuando $n_i > 5$ para todas las i y la hipótesis nula es válida, la distribución del estadístico H se puede aproximar bastante por la distribución X-cuadrado, con $k-1$ grados de libertad. En la tabla de D. B. Owen, mencionada en la bibliografía, se encuentran tablas especiales para aplicar con valores pequeños seleccionados de las n_i y k.

Para ilustrar la prueba de Kruskal-Wallis H, supondremos que el experimento descrito en la página 200, se amplió para incluir cuatro registradores diferentes, con los resultados mostrados en la tabla siguiente. (Nótese que a las observaciones empatadas se les asignan nuevamente, el medio de los puestos que ocupan en conjunto.)

Registrador A:	0.2	0.3	0.4	0.5	1.7	1.9	2.0
Registrador B:	0.8	1.1	1.3	1.9	2.5	7.8	
Registrador C:	0.7	0.9	8.2	12.0	12.1	15.3	
Registrador D:	0.1	0.1	0.3	0.5	2.9	13.8	

Las observaciones son, otra vez, corrientes de retorno en microamperes. Como se puede verificar fácilmente, las observaciones de la primera muestra ocupan los puestos 3, 4.5, 6, 7.5, 14, 15.5 y 17, por lo que $R_1 = 67.5$. Similarmente, las observaciones de la segunda muestra ocupan los puestos 10, 12, 13, 15.5, 18 y 20, por lo que $R_2 = 88.5$; las observaciones de la tercera muestra ocupan los puestos 9, 11, 21, 22, 23 y 25, por lo que $R_3 = 111.0$; y las observaciones de la cuarta muestra ocupan los puestos 1.5, 1.5, 4.5, 7.5, 19 y 24, por lo que $R_4 = 58.0$. Substituyendo en la fórmula de H, encontramos

$$H = \frac{12}{25 \cdot 26}\left(\frac{67.5^2}{7} + \frac{88.5^2}{6} + \frac{111.0^2}{6} + \frac{58.0^2}{6}\right) - 3(26) = 6.4$$

y si comparamos este valor con 7.815, valor de $\chi^2_{.05}$ con 3 grados de libertad, vemos que no se puede rechazar la hipótesis nula. En otras palabras, no podemos rechazar la hipótesis nula de que las muestras provienen de poblaciones idénticas frente a la alternativa de que las medias de las poblaciones no son iguales.

EJERCICIOS

1. Con respecto a las 100 medidas de los pesos de planchas de estaño de galvanizado electrolítico del problema 3 de la página 106, utilizar el test de los signos con $\alpha = 0.05$ para contrastar la hipótesis nula $\mu = 0.33$ frente a la alternativa $\mu < 0.33$, donde μ es la media de la población de pesos de la que se obtuvo la muestra.
2. Empleando el test de los signos de una sola muestra, contrastar la hipótesis nula de que el octanage medio de la gasolina de la que se tomaron las 16 muestras siguientes es 100, frente a la hipótesis alternativa de que es mayor de cien.

101.6	98.2	104.5	99.0	102.8	105.4	107.7	99.4
103.3	100.0	102.5	97.1	103.6	101.0	98.7	101.0

Emplear un nivel de significación de 0.05.

3. Utilizar el test de los signos y un nivel de significación de 0.10 para decidir si hay una diferencia sistemática entre las lecturas obtenidas de los dos instrumentos del ejercicio 19 de la página 161.
4. Se sierran vigas de acero por dos métodos, consistente el primero en aserrarlas cuando aún están calientes, y el segundo cuando se han enfriado. Las longitudes finales resultantes (en pies) una vez que todas las vigas se han enfriado hasta la temperatura ambiente, son las siguientes:

Vigas serradas en caliente:	31.6	30.5	31.1	29.7	27.9	[illegible]
	30.5	31.8	32.6	28.8	29.6	[illegible]
	28.9	29.9	31.6	30.7	30.3	31.5
Vigas serradas en frío:	30.1	31.0	29.9	29.8	30.0	30.5
	30.6	30.2	31.1	29.8	29.7	[illegible]
	31.3	30.5	30.1	30.0	30.8	30.3

Apareando al azar las 18 observaciones de las dos muestras, usar el test de los signos de dos muestras, con un nivel de significación de 0.05, para determinar si hay alguna diferencia significante en las longitudes finales medias.

5. En pruebas repetidas, un motor experimental operó, respectivamente, durante 20, 19, 22, 17, 18, 20, 23, 19, 20, 15, 24, 21, 18, 20, 24, 23, 20, 17, 25 y 28 minutos, con un galón de cierta clase de combustible. Empleando el test de los signos y un nivel de significación de 0.01 contrastar la hipótesis nula $\mu = 20$ frente a la alternativa $\mu \neq 20$.
6. Para comparar una bebida con la de una marca de la competencia, 50 personas probaron una bebida y luego la otra y después se les pidió que indicaran su preferencia. (El orden de las marcas se escogió al azar para cada persona.) Si 27 prefirieron la marca dada, 18 prefirieron la de la competencia, y 5 no encontraron diferencia en el sabor, contrastar, con un nivel de significación de 0.01, si la marca dada es superior en sabor a la de la competencia.
7. Un experimento para comparar la resistencia a la tensión de dos clases de hilos dio los resultados siguientes (en libras):

Hilo A:	143.6	114.8	145.2	144.8	145.6	146.0
	143.0	147.4	144.0	145.6	145.5.	144.8
Hilo B:	146.6	147.8	144.4	140.8	143.0	148.8
	153.0	142.4	146.8	143.2	140.9	150.6

Usar el test U y un nivel de significación de 0.05 para contrastar la hipótesis nula de que las dos muestras provienen de poblaciones idénticas, frente a la alternativa de que las dos poblaciones tienen medias diferentes.

8. Repetir el ejercicio 4, usando el test U de Mann-Whitney. También contrastar si las dispersiones de las dos poblaciones son iguales, utilizando el test de suma de números de orden mencionada en la página 211.
9. En el ejercicio 7 utilizar $\alpha = 0.05$ para contrastar la hipótesis nula de que las muestras provienen de poblaciones idénticas, frente a la alternativa de que las dos poblaciones tienen dispersiones diferentes.
10. Los tests denominados de Franklin se establecieron para determinar las propiedades de aislamiento de aceros al silicio de granos orientados que fueron recocidos en cinco atmósferas diferentes, con los resultados siguientes :

Atmosfera	*Resultados del ensayo (amperios)*							
1	0.58	0.61	0.69	0.79	0.61	0.59		
2	0.37	0.37	0.58	0.40	0.28	0.44	0.35	
3	0.29	0.19	0.34	0.17	0.29	0.16		
4	0.81	0.69	0.75	0.72	0.68	0.85	0.57	0.77
5	0.26	0.34	0.29	0.47	0.30	0.42		

Emplear el test H de Kruskal-Wallis H y un nivel de significación de 0.05 para decidir si se puede aceptar que estas cinco muestras proceden de poblaciones idénticas.

11. Para investigar tres medidas preventivas contra la corrosión, se probaron muestras al azar de 10 piezas de alambre para cada una de las tres medidas preventivas, dando los siguientes resultados para las profundidades máximas de las partes erosionadas (en milésimas de pulgada):

A:	45	53	60	48	57	62	49	55	53	52
B:	62	58	47	59	63	48	58	52	50	49
C:	57	45	60	54	57	55	48	59	62	60

Contrastar, con un nivel de significación de 0.05, si hay alguna diferencia en la eficacia de las tres medidas preventivas contra la corrosión.

11.5 Tests de las series de términos iguales

Al discutir las muestras aleatorias en el capítulo 7, presentamos varios métodos que daban por anticipado cierta seguridad de que una muestra fuera de azar. Sin embargo, es útil tener una técnica para contrastar si una muestra se puede considerar como aleatoria *después de haber sido obtenida.* Una de estas técnicas está basada en el orden en que se obtuvieron los valores de las muestras; más concisamente, se basa en el número de series de términos iguales mostradas en los resultados de las muestras.

Dada una sucesión de dos símbolos, tales como H y T (que pueden representar, por ejemplo, las caras y las cruces en tiradas sucesivas), una "serie de iguales" es una sucesión de símbolos idénticos comprendidos entre dos símbolos diferentes o sin estos últimos. Es decir, la serie

$$\underline{T\ T}\ \ \underline{H\ H}\ \ \underline{T\ T}\ \ \underline{H\ H\ H}\ \ \underline{T}\ \ \underline{H\ H\ H}\ \ \underline{T\ T\ T\ T}\ \ \underline{H\ H\ H}$$

contiene 8 series de iguales, como indican los subrayados. El número total de series de iguales en una sucesión de n ensayos da una indicación de si la sucesión se puede considerar como de azar. Entonces, si sólo ha habido dos series de iguales, consistentes en diez caras, seguidas por diez cruces, se puede suponer que la probabilidad de un caso favorable no ha permanecido constante de una prueba a la siguiente. Por otra parte, si la sucesión consta de veinte tiradas formadas alternativamente por caras y cruces, se puede suponer que los ensayos no han sido independientes. En cualquier caso, hay razones para suponer que no se trata de un azar. Notemos que nuestra suposición no procede del *número* de caras y cruces, si no del *orden* en que aparecen.

Si una sucesión contiene n_1 símbolos de una clase y n_2 de otra (y ni n_1 ni n_2 son muy pequeños), la distribución muestral del número total de series de iguales se puede representar muy aproximadamente por una distribución normal de media

$$\mu_u = \frac{2n_1n_2}{n_1 + n_2} + 1$$

y desviación típica

$$\sigma_u = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

donde u denota el número total de series de iguales. Entonces, el test de la hipótesis nula de que la ordenación de los símbolos (y, por consiguiente, de la muestra) sea aleatoria, se puede basar en el estadístico

$$z = \frac{u - \mu_u}{\sigma_u}$$

y en la tabla III. Este test da una excelente aproximación cuando ni n_1 ni n_2 son menores de 10. Se pueden encontrar tablas especiales para hacer tests exactos cuando n_1 o n_2 son pequeñas, en las tablas de D. B. Owen, citadas en la bibliografía.

Para ilustrar este test, examinaremos la siguiente sucesión de 32 vuelos de prueba de un cohete, donde S y F marcan, sucesivamente, los éxitos y los fallos:

$$\underline{F\ F\ F}\ \underline{S\ S}\ \underline{F\ F}\ \underline{S\ S\ S}\ \underline{F}\ \underline{S}\ \underline{F}\ \underline{S\ S\ S\ S}\ \underline{F}\ \underline{S\ S\ S}\ \underline{F}\ \underline{S\ S\ S\ S\ S}\ \underline{F}\ \underline{S\ S\ S\ S}$$

Como hay 22 éxitos, 10 fallos y 14 series de iguales, substituimos $n_1 = 22$, $n_2 = 10$, $u = 14$, y obtenemos

$$\mu_u = \frac{2 \cdot 22 \cdot 10}{32} + 1 = 14.75$$

$$\sigma_u = \sqrt{\frac{2 \cdot 22 \cdot 10(2 \cdot 22 \cdot 10 - 22 - 10)}{(22 + 10)^2(22 + 10 - 1)}} = 2.38$$

y
$$z = \frac{14 - 14.75}{2.38} = -0.31$$

Como este valor queda entre —1.96 y 1.96, no podemos rechazar (con un nivel de significación de 0.05) la hipótesis nula de que la ordenación sea aleatoria. Evidentemente, no hay razones suficientes para concluir que hay una fiabilidad real.

Se puede emplear también el test de series de iguales para contrastar la casualidad que hay en las muestras formadas por datos numéricos, contando las series de iguales a partir de la mediana *por encima y por debajo de ésta.* Si denotamos una observación mayor que la mediana de la muestra por la letra a y una observación menor que la mediana por la letra b, podemos utilizar la sucesión resultante de a y b para constrastar la casualidad, siguiendo el método indicado antes. Una aplicación frecuente de este método es en el control de calidad, donde las medias de muestras pequeñas sucesivas se representan en una gráfica en orden cronológico. El test de series de iguales se puede usar entonces para comprobar si hay alguna tendencia en los datos; que nos indique que es necesario ajustar una máquina o hacer algún otro proceso antes de que ocurra algún daño grave.

Para ilustrar un test de series de iguales por encima y por debajo de la mediana, supongamos que un ingeniero se interesa en la posibilidad de que se hayan hecho demasiados cambios en el ajuste de un torno automático. Para contrastar esta hipótesis, se obtuvieron los siguientes diámetros medios (en pulgadas) de 40 ejes torneados sucesivamente en el torno:

0.261	0.258	0.249	0.251	0.247	0.256	0.250	0.247	0.255	0.243
0.252	0.250	0.253	0.247	0.251	0.243	0.258	0.251	0.245	0.250
0.248	0.252	0.254	0.250	0.247	0.253	0.251	0.246	0.249	0.252
0.247	0.250	0.253	0.247	0.249	0.253	0.246	0.251	0.249	0.253

La mediana de estas medidas es 0.250 y, cambiando cada una de ellas por una letra a si excede de 0.250, por una b si es menor que 0.250, y omitiendo las cinco que son iguales a 0.250, obtenemos la sucesión

$$aababababaababaabbaabaabbababbababa$$

que tiene 27 series de iguales. Entonces, $n_1 = 19$, $n_2 = 16$, $u = 27$, tendremos:

$$\mu_u = \frac{2 \cdot 19 \cdot 16}{35} + 1 = 18.37$$

$$\sigma_u = \sqrt{\frac{2 \cdot 19 \cdot 16(2 \cdot 19 \cdot 16 - 19 - 16)}{(19+16)^2(19+16-1)}} = 2.89$$

y
$$z = \frac{27 - 18.37}{2.89} = 2.98$$

Como este valor excede a 1.96, podemos rechazar la hipótesis nula de que la sucesión de medidas sea aleatoria. Como el número de series de iguales es mayor que el que se podría esperar, debido al azar, es razonable suponer que el torno se ha ajustado demasiado: es probable que se haya hecho un ajuste después de tornear cada pieza, tratando, con ello, de compensar cualquier discrepancia que se haya observado con respecto al diámetro nominal de 0.250 pulgadas.

11.6 Tests de Kolmogorov-Smirnov

Los tests de Kolmogorov-Smirnov son test no paramétricos para diferencias entre dos distribuciones totales o acumulativas. El *test uni-muestral* se refiere a la concordancia entre una distribución acumulativa observada de valores de una muestra y una función de distribución continua especificada; es decir, se trata de una prueba de bondad de ajuste. El *test bi-muestral* se refiere a la concordancia entre dos distribuciones acumulativas observadas; se contrasta la hipótesis de si dos muestras independientes provienen de distribuciones continuas idénticas, y es sensible a las diferencias de población en lo que se refiere a la localización, dispersión, o disimetría.

El test uni-muestral de Kolmogorov-Smirnov es, en general, más eficaz que el test X-cuadrado para la bondad de ajuste de muestras pequeñas y puede usarse con muestras muy pequeñas en las que el test X-cuadrado no es aplicable. Debemos recordar, sin embargo, que el test X-cuadrado de la sección 10.5 se puede usar con distribuciones discretas, mientras que el test de Kolmogorov-Smirnov no puede usarse.

El test uni-muestral se basa en la diferencia absoluta máxima D entre los valores de la distribución acumulativa de una muestra aleatoria de tamaño n y una distribución teórica especificada. Como ilustración de este test, se quiere comprobar si los agujeros para clavijas en una placa de hojalata electrolítica están distribuidos uniformemente para lo cual se han tomado medidas de las siguientes distancias (en pulgadas) de 10 agujeros a partir de un extremo de una tira grande de placa de hojalata de 30 pulgadas de ancho:

4.8 14.8 28.2 23.1 4.4 28.7 19.5 2.4 25.0 6.2

Bajo la hipótesis nula de que los agujeros están uniformemente repartidos, la distribución acumulativa teórica con la que queremos comparar la distribución acumulativa observada está dada por

$$F(x) = \begin{cases} 0 & \text{para } x \leq 0 \\ x/30 & \text{para } 0 < x < 30 \\ 1 & \text{para } x \geq 30 \end{cases}$$

La gráfica de esta distribución acumulativa teórica se muestra, junto con la de la distribución acumulativa observada, en la figura 11.1. Como se indica en esta figura, la diferencia máxima entre las dos distribuciones acumulativas es 0.193.

Para determinar si esta diferencia es mayor que lo que se puede esperar razonablemente, encontramos el valor crítico de D en la tabla IX. Para $n = 10$ $\alpha = 0.05$, el valor crítico es $D_{.05} = 0.410$, y de aquí que la hipótesis nula (que los agujeros están uniformemente distribuidos) no se puede rechazar.

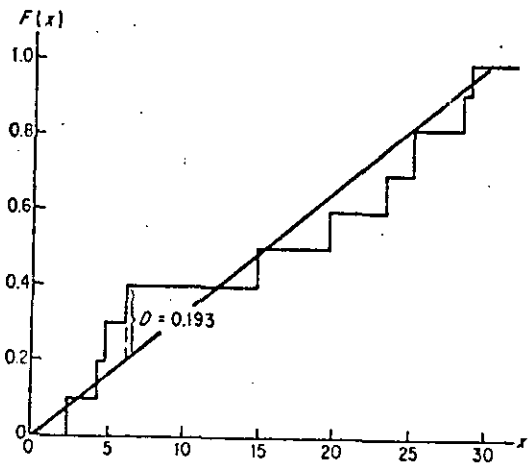


Fig. 11.1 Prueba Kolmogorov-Smirnov

El test bimuestral de Kolmogorov-Smirnov se basa en la diferencia absoluta máxima entre los valores de las dos distribuciones acumulativas observadas. En principio, es muy similar al test uni-muestral, y los valores críticos necesarios se pueden obtener de tablas especiales (por ejemplo, las de D. B. Owen, citada en la bibliografía).

EJERCICIOS

1. Para comprobar si cierta señal de radio contiene un mensaje, se puede subdividir un intervalo de tiempo en cierto número de intervalos muy cortos y determinar después si la fuerza de la señal excede cierto nivel (ruido de fondo) en cada corto intervalo. Supongamos que la siguiente es parte de una observación de este tipo, donde *H* indica una señal fuerte y *L* que la señal no excede cierto nivel de ruido.

   L L H L H L H L H H H L H H H L H H H L H L H L H L L
   L L H L H L H L H H H L H H H L H H H L H L H L H L L

   Verificar si esta sucesión se debe al azar (usando un nivel de significación de 0.05) y comprobar si es razonable suponer que la señal contiene un mensaje.
2. La siguiente es una lista que nos da, leyendo las filas sucesivas de izquierda a derecha, el orden en que una máquina produjo piezas defectuosas *(D)* y no defectuosas *(N)* durante cierto periodo:

   N N N N N N N N N D D D N N N N N N N N
   N N N D N N N N N N N N D D N N D D N N
   N N N N N N N N N N N N N D D D N N D N
   D N N D N N N N N N N N D D N N N D N N
   N N N N N N N N D D D N N N N N N N N N

   Comprobar, con un nivel de significación de 0.05, si este arreglo es debido al azar.
3. En la página 92 se describió un método para generar dígitos al seudoazar, en el cual un número de 4 dígitos se elevaba al cuadrado, los 4 números de la parte media del número resultante se elevaban al cuadrado, y así sucesivamente. Comenzando con el número 3571, usar una tabla de cuadrados o una calculadora para continuar este proceso hasta obtener una sucesión de 48 dígitos. Contrastar si la sucesión resultante, para ver si es de azar, utilizando series de términos iguales por encima y por debajo de la mediana y un nivel de significación de 0.05.
4. En una fábrica, el tiempo que no trabaja una máquina durante las horas de trabajo, debido a dificultades tales como roturas o fallos, se llama "tiempo muerto". La tabla que sigue corresponde a 50 tiempos muertos (en minutos) consecutivos observados por un ingeniero de control durante cierto periodo (léanse las filas sucesivas de izquierda a derecha):

18	25	28	21	29	30	34	30	25	21
22	29	30	24	35	37	20	27	30	25
30	21	26	33	36	35	31	20	28	39
40	30	36	34	35	39	42	30	35	41
38	35	50	46	34	37	39	42	48	51

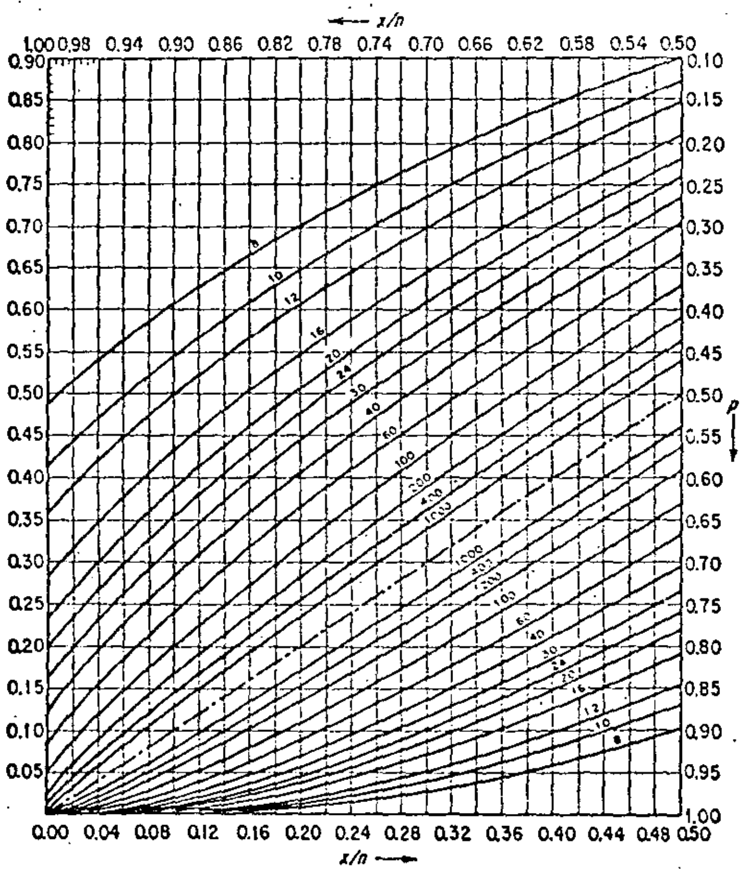
   Usar el test de las series de iguales por encima y por debajo de la mediana, a un nivel de significación de 0.05, para contrastar la hipótesis de que los datos marcan una tendencia.
5. Las temperaturas horarias de un horno (en grados centígrados) tomadas durante un periodo de 24 horas son las siguientes:

   269, 265, 271, 268, 270, 266, 273, 271, 275, 269, 271, 273
   275, 268, 276, 270, 273, 266, 270, 268, 272, 271, 278, 267

   Contrastar si esta disposición es al azar, con un nivel de significación de 0.01, para investigar si el horno está trabajando cíclicamente en intervalos de dos horas.
6. El problema 13 de la página 107 contiene el número de imperfecciones en muestras tomadas de 50 piezas de tela. Suponiendo que el orden de las muestras es el mismo que el orden en que se han producido las piezas de tela, contrastar la hipótesis de que la presencia de muestras sin imperfecciones se debe al azar, frente a la alternativa de que existe un agrupamiento. (Se toma $\alpha = 0.05$).
7. Utilizar el test de Kolmogorov-Smirnov, con $\alpha = 0.01$, para decidir si las resistencias a la compresión del ejercicio 9 de la página 107 se pueden suponer provinientes de una distribución normal con media de 50.000 libras por pulgada cuadrada y desviación típica de 10.000 libras por pulgada cuadrada. [*Sugerencia:* utilizar papel gráfico de probabilidades.]
8. En un estudio de vibraciones, se sometieron las componentes de un avión a fuertes vibraciones, hasta que se originaron fallos estructurales. Los siguientes son los tiempos obtenidos (en minutos) 4.1, 0.8, 5.3, 5.0, 8.3, 1.7, 2.5, 6.2, 7.3, 9.0, 1.2, 3.7, 9.5, 10.5. Contrastar si se pueden considerar estos datos como una muestra proviniente de una población exponencial con una media de 5 minutos. (Se toma $\alpha = 0.05$.)

Tabla VIII(b)

0.99 INTERVALOS DE CONFIANZA PARA PROPORCIONES*



* Reproducción de la tabla 41 de *Biometrika Tables for Statisticians*, Vol. I (New York: Cambridge University Press, 1954) con autorización de *Biometrika* fideicomiso.

Tabla IX

VALORES CRITICOS DE D*

Número de muestra n	$D_{.10}$	$D_{.05}$	$D_{.01}$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.828
4	0.564	0.624	0.733
5	0.510	0.565	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.410	0.490
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.356
25	0.24	0.27	0.32
30	0.22	0.24	0.29

* Adaptado de F. J. Massey, Jr., "The Kolgomorov-Smirnov test for goodness of flit" *J. Amer. Statist. Ass.*, Vol. 46 (1951), p. 70 con autorización de autor y editores.

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

PROBABILIDAD Y ESTADISTICA. FUNDAMENTOS Y APLICACIONES

EJERCICIOS

DR. OCTAVIO A. RASCON CHAVEZ

FEBRERO, 1985

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 $\bar{X}$ = \$92.00 Y UNA DESVIACION ESTANDAR S = \$8.00. CALCULAR EL INTERVALO DEL 95% DE NIVEL DE CONFIANZA DE LA VARIANCIA

$$\text{I.C.} = \left(\frac{20(8)^2}{32.9}, \frac{20(8)^2}{8.91}\right) = (38.91, 143.66)\ \$^2$$

EL INTERVALO DE CONFIANZA PARA LA DESVIACION ESTANDAR ES

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

3

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	$0 \leq t < 1$	$1 \leq t < 2$	$2 \leq t < 3$	$t \geq 3$
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: $nP(x_1 \leq X < x_2)$

DONDE n = TAMAÑO DE LA MUESTRA

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

$$p_2 = P(1 \leq X < 2) = \int_1^2 (0.45)e^{-0.455}\,dt = 0.232; \quad 50p_2 = 11.60$$

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

$$p_4 = P(X > 4) = \int_3^\infty (0.45)e^{-0.45t}\,dt = \frac{0.259}{\Sigma = 0.998 \doteq 1}; \quad 50p_4 = \frac{12.95}{\Sigma = 49.90 \doteq 50}$$

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

4

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

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:

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

$p_1 = f_X(0) = 2^0 e^{-2}/0! = 0.135;\ np_1 = 100 \times 0.135 = 13.5$

$p_2 = f_X(1) = 2^1 e^{-2}/1! = 0.270;\ np_2 = 100 \times 0.270 = 27.0$

$p_3 = f_X(2) = 2^2 e^{-2}/2! = 0.270;\ np_3 = 27.0$

$p_4 = f_X(3) = 2^3 e^{-2}/3! = 0.180;\ np_4 = 18.0$

$p_5 = f_X(4) = 2^4 e^{-2}/4! = 0.090;\ np_5 = 9.0$

$p_6 = P(X \geq 5) = 1 - F_X(4) = 0.055;\ np_6 = 5.5$

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

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

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

7

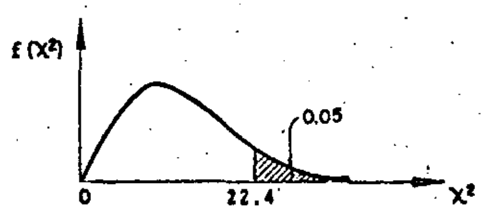
EJERCICIO

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	108.8	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	4.58
215.5-219.5	1	0.0462	0.9538	0.910	19.69
> 219.5	0	≈ 0	TOTAL:		32.67

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

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



6

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 $\alpha = 0.05$.

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 - 0.30 - 0.03 - 0.23 - 0.14 - 0.15 - 0.45 - 0.22 - 0.19
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
					$\Sigma = 10.0$

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

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

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

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

10

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.

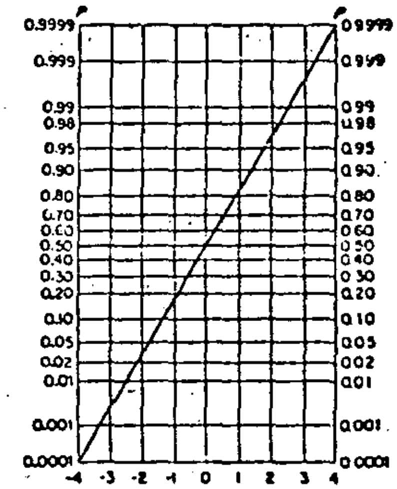


Figure 7-29

Since the area under the normal curve is unity, the partial areas, P_a, 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 fifth number in the last column, 25.3, equals $100 \times 117/462$.

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

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

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

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

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	1296	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	20816	20500	-6	142

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

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

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

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

$S^2(Y) = \frac{20816}{17} - 1204.4982 = 19.9723$; $S(Y) = 4.4690379$

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

$r_{xy} = \frac{(20500/17)-(34.352941)(34.705882)}{(3.9398604)(4.4690379)} = 0.77$[illegible]

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

$$[illegible]\; 0.774 \sqrt{\frac{17-2}{1-0.744^2}} = 0.774 \times 6.116 = 4.73 > 2.13$$

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	891	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
540	561	18614	19867	19072	-21	175

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

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

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

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

$S_d^2 = \frac{175}{16} - 1.7226562 = 9.214844$ $\qquad S_d = 3.03$[illegible]

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

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

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

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

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

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

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

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

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

$$r_{xy} = \frac{(39572/33)-(34.060606)(34.878787)}{(4.458161)(4.0358889)} = \underline{\underline{0.6201924}}$$

$$t_d = \left| \frac{(34.061 - 34.879)\sqrt{32}}{2.989} \right| = 1.548 < 2.04$$

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%.

RESPUESTAS AL INCISO c

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

$$= 0.368 < 2.04$$

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

ESTADISTICAS NO PARAMETRICAS

FEBRERO, 1985

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

Estadísticas no paramétricas 13

13.1 INTRODUCCION

La mayoría de los procedimientos para prueba de hipótesis descritos hasta ahora se basan en la suposición de que las muestras aleatorias se seleccionan de poblaciones normales. Por fortuna casi todas siguen siendo bastante confiables para ligeras desviaciones de la normalidad, particularmente cuando el tamaño de la muestra es grande. En este capítulo se examinarán ciertos procedimientos de prueba que no suponen algún conocimiento previo acerca de la distribución de la población fundamental. A la prueba efectuada sin esta información se le llama *no paramétrica* o *prueba libre de distribución*

Las pruebas no paramétricas han ganado cierto terreno en los últimos años por varias razones. Primero, los cálculos que requieren son más rápidos y fáciles de efectuar. Segundo, los datos no necesitan ser mediciones cuantitativas ya que pueden ser respuestas cualitativas; por ejemplo, "defectuosa" o "no defectuosa", "sí" o "no", etc; a menudo son valores de una escala ordinal a la cual se le asignan rangos. En una escala ordinal los sujetos se colocan en rangos según un orden específico y luego se analizan los diferentes rangos mediante una prueba no paramétrica. Por ejemplo, dos jueces pueden catalogar cinco marcas de cerveza asignando el rango 1 a la cerveza que creen tiene las mejores cualidades, el rango 2 a la segunda mejor y así sucesivamente. La prueba no paramétrica debe usarse entonces para determinar si los dos concuerdan. La tercera y, quizá la más importante ventaja de las pruebas no paramétricas es que contienen menos suposiciones restrictivas que las pruebas paramétricas correspondientes, pues suelen suponer que las distribuciones fundamentales son continuas y simétricas.

También cabe advertir que tienen varias desventajas. En primer lugar, no utilizan toda la información proporcionada por la muestra y de ahí que sean menos eficaces que el procedimiento paramétrico correspondiente, cuando ambos son aplicables. Consecuentemente, una prueba no paramétrica requerirá una muestra mayor que la de la prueba paramétrica correspondiente, si es que se quiere lograr la misma probabilidad de cometer un error tipo II.

En resumen, si las pruebas paramétrica y no paramétrica son aplicables al mismo conjunto de datos, se aplicará siempre la técnica más eficaz: la prueba paramétrica. Sin embargo, dado que muchas veces la suposición de normalidad no se puede justificar y que no siempre se tienen mediciones cuantitativas, es una fortuna que los estadigrafos hayan ideado varios procedimientos no paramétricos útiles.

13.2 PRUEBA DE WILCOXON PARA DOS MUESTRAS

Los procedimientos explicados en la sección 7.4 para la prueba de hipótesis acerca de la diferencia entre dos medias sólo son válidos si las poblaciones son aproximadamente normales o si la muestras son grandes. En 1945, Frank Wilcoxon propuso un procedimiento no paramétrico muy sencillo para la comparación de dos poblaciones continuas cuando se disponia solamente de pequeñas muestras independientes y que las poblaciones de donde se habian seleccionado no eran normales. Este procedimiento se llama ahora *prueba de Wilcoxon para dos muestras* o *prueba Wilcoxon para la suma de rangos.*

Se probará la hipótesis nula H_0 de que $\mu_1 = \mu_2$ contra alguna alternativa apropiada. Primero se selecciona una muestra aleatoria de cada una de las poblaciones. Sea n_1 el número de observaciones en la muestra más pequeña y n_2 el número de observaciones en la muestra más grande. Cuando las muestras son de igual tamaño, n_1 y n_2 se pueden asignar al azar. Las observaciones de las muestras combinadas, $n_1 + n_2$, se arreglan en orden ascendente y se les asigna un rango de $1, 2, \ldots, n_1 + n_2$ en cada observación. En el caso de observaciones idénticas, se reemplazan por la media de los rangos que tendrian las observaciones si no lo fueran. Por ejemplo, si la séptima y octava observaciones son idénticas, se les puede asignar un rango de 7.5.

La suma de los rangos correspondiente a las observaciones n_1 de la muestra más pequeña se indica por w_1. De la misma manera, w_2 representa la suma de los n_2 rangos correspondientes a la muestra más grande. El total $w_1 + w_2$ depende solamente del número de observaciones en las dos muestras y en él no influyen en absoluto los resultados del experimento. Por ejemplo, si $n_1 = 3$ y $n_2 = 4$, entonces $w_1 + w_2 = 1 + 2 + \cdots + 7 = 28$, cualesquiera que sean los valores numéricos de las observaciones. En general,

$$w_1 + w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2},$$

es la suma aritmética de los enteros $1, 2, \ldots, n_1 + n_2$. Una vez determinada w_1 puede ser más fácil encontrar w_2 mediante la fórmula

$$w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - w_1.$$

Al seleccionar varias muestras de tamaños n_1 y n_2, cabe esperar que w_1 y, por lo tanto w_2, varien. Así, se puede pensar que w_1 y w_2 son repectivamente valores de las variables aleatorias W_1 y W_2. La hipótesis nula $\mu_1 = \mu_2$ será rechazada en favor de la alternativa $\mu_1 < \mu_2$ sólo si w_1 es pequeña y w_2 grande. Igualmente, la alternativa $\mu_1 > \mu_2$ sólo puede aceptarse si w_1 es grande y w_2 pequeña. En una prueba de dos colas, se puede rechazar H_0 en favor de H_1 si w_1 es pequeña y w_2 grande o si w_1 es grande y w_2 pequeña. A consecuencia de la simetria en las distribuciones de W_1 y W_2, las probabilidades de la cola superior se pueden obtener partiendo de las probabilidades de la cola inferior. Por lo tanto, sin importar cuál pueda ser la hipótesis alternativa, la hipótesis nula se rechaza cuando la menor de w_1 y w_2 es suficientemente pequeña. Supóngase que, en un experimento, $w_1 < w_2$. Si conocemos la distribución de W_1, se puede determinar $P(W_1 \leq w_1 | H_0$ es verdadera). Si esta probabilidad es menor que o igual a 0.05, la prueba es significativa y se puede rechazar H_0 en favor de la alternativa unilateral apropiada. Cuando la probabilidad no excede de 0.01, la prueba es altamente significativa. En el caso de una prueba de dos colas, la simetria permite basar la decisión en el valor de $2P(W_1 \leq w_1 | H_0$ es verdadera). Por lo tanto, cuando $2P(W_1 \leq w_1 | H_0$ es verdadera) < 0.05, la prueba es significativa y se concluye que $\mu_1 \neq \mu_2$.

La distribución de W_1, cuando H_0 es verdadera, se basa en el hecho de que todas las observaciones en la muestra menor podrian tener rangos asignados al azar a medida que sus sumas fueran menores o iguales a w_1. El número total de formas para asignar $n_1 + n_2$ rangos a n_1 observaciones, de tal manera que la suma de los rangos no exceda a w_1, se indica por $n(W_1 \leq w_1)$. Hay $\binom{n_1 + n_2}{n_1}$ formas iguales probables de asignar los $n_1 + n_2$ rangos a las n_1 observaciones, que dan todos los valores posibles de W_1. De donde

$$P(W_1 \leq w_1 | H_0 \text{ es verdadera}) = \frac{n(W_1 \leq w_1)}{\binom{n_1 + n_2}{n_1}}, \quad \text{para } n_1 \leq n_2.$$

Es posible encontrar $n(W_1 \leq w_1)$ para cualquier prueba listando todos los casos y contándolos. Así, cuando $n_1 = 3$ y $n_2 = 5$, el número de casos donde la suma de los rangos en la muestra menor es menor que o igual a 8 puede listarse del modo siguiente:

$$1 + 2 + 3 = 6$$

$$1 + 2 + 4 = 7$$

$$1 + 3 + 4 = 8$$

$$1 + 2 + 5 = 8.$$

Por lo tanto, en $\binom{8}{3} = 56$ casos iguales posibles hay 4 favorables. De donde

$$P(W_1 \leq 8 \mid H_0 \text{ es verdadera}) = \tfrac{4}{56} = 0.0714.$$

Para el caso en que $w_2 < w_1$, se puede proceder como antes y determinar. $P(W_2 \leq w_2 \mid H_0$ es verdadera). Sin embargo, en ambos casos generalmente es más fácil encontrar la probabilidad deseada usando la tabla XVI cuando n_2 no excede de ocho. La tabla XVI se basa en la estadística U, el mínimo de U_1 y U_2, donde

$$U_1 = W_1 - \frac{n_1(n_1 + 1)}{2}$$

y

$$U_2 = W_2 - \frac{n_2(n_2 + 1)}{2}.$$

Si $P(U \leq u \mid H_0$ es verdadera$) \leq \alpha$, la prueba es significativa y se rechaza H_0 en favor de la alternativa unilateral apropiada. En una prueba de dos colas, la prueba es significativa cuando $2P(U \leq u \mid H_0$ es verdadera$) \leq \alpha$, y entonces se acepta la hipótesis alternativa de que $\mu_1 \neq \mu_2$.

En la ilustración anterior, donde se tenía $n_1 = 3$, $n_2 = 5$ y $w_1 = 8$, se encuentra que $w_2 = [(8)(9)/2] - 8 = 28$ de donde obtenemos

$$u_1 = 8 - [(3)(4)/2] = 2$$
$$u_2 = 28 - [(5)(6)/2] = 13.$$

Usando la tabla XVI, con $u = 2$, se tiene

$$P(U \leq 2 \mid H_0 \text{ es verdadera}) = 0.071,$$

que coincide con la respuesta anterior. Si en la misma ilustración $w_1 = 7$ de manera que $u = 1$, se obtiene

$$P(U \leq 1 \mid H_0 \text{ es verdadera}) = 0.036,$$

que es significativa para un prueba de una cola al nivel de 0.05 pero no al nivel de 0.01. En una prueba de dos colas, la probabilidad de que las medias muestrales difieran en una cantidad mayor o igual que la observada es

$$2P(U \leq 1 \mid H_0 \text{ es verdadera}) = (2)(0.036) = 0.072,$$

de lo cual se concluye que H_0 es verdadera.

Cuando n_2 se encuentra entre 9 y 20 se puede usar la tabla XVII. Si el valor observado de U es menor o igual al valor tabulado, la hipótesis nula puede ser rechazada al nivel de significancia indicado en la tabla. La tabla proporciona valores críticos de U para niveles de significancia iguales a 0.001, 0.01, 0.025 y 0.05 para una prueba de una cola. En el caso de una prueba de dos colas, los valores críticos de U corresponden a los niveles de significancia 0.002, 0.02, 0.05 y 0.1. Cuando n_1 y n_2 aumentan de tamaño, la distribución muestral de U se aproxima a la distribución normal con media

$$\mu_U = \frac{n_1 n_2}{2}$$

y variancia

$$\sigma_U^2 = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}.$$

En consecuencia, cuando n_2 es mayor que 20 se puede usar como prueba la estadística $Z = (U - \mu_U)/\sigma_U$ cuya región crítica se encuentra en una o ambas colas de la distribución normal estándar, según la forma de H_1.

Para probar la hipótesis nula de que las medias de dos poblaciones no normales son iguales cuando se dispone solamente de muestras independientes pequeñas, se procede a través de los pasos siguientes:

1. H_0: $\mu_1 = \mu_2$.
2. H_1: Las alternativas son $\mu_1 < \mu_2$, $\mu_1 > \mu_2$, o $\mu_1 \neq \mu_2$.
3. Seleccione un nivel de significancia igual a α.
4. Región crítica:
   a) Todos los valores u para los cuales $P(U \leq u \mid H_0$ es verdadera$) < \alpha$ cuando $n_2 \leq 8$ y la prueba es de una cola.
   b) Todos los valores u para los cuales $2P(U \leq u \mid H_0$ es verdadera$) < \alpha$ cuando $n_2 \leq 8$ y la prueba es de dos colas.
   c) Todos los valores u menores que o iguales al valor crítico apropiado, en la tabla XVII, cuando $9 \leq n_2 \leq 20$.
5. Calcule w_1, w_2, u_1 y u_2 partiendo de las muestras independientes de tamaños n_1 y n_2, donde $n_1 \leq n_2$. Usando como u, el menor de los valores u_1 y u_2 determine si u se encuentra en la región de aceptación o en la región crítica.
6. Conclusión: Rechace H_0 si u cae en la región crítica; de otra forma, acepte H_0.

Ejemplo 13.1 Para averiguar si una nueva vacuna detendrá la leucemia, se seleccionan nueve ratones que están en una etapa avanzada de la enfermedad y el tratamiento se aplica únicamente a cinco ratones. A partir del inicio del experimento los años de supervivencia son:

Vacunados	2.1	5.3	1.4	4.6	0.9
No vacunados	1.9	0.5	2.8	3.1	

Al nivel de significancia de 0.05, ¿se puede afirmar que la vacuna es eficaz?

Solución Se seguirá el procedimiento de seis pasos con $n_1 = 4$ y $n_2 = 5$.

1. H_0: $\mu_1 = \mu_2$.
2. H_1: $\mu_1 < \mu_2$.
3. $\alpha = 0.05$.
4. Región crítica: Todos los valores de u para los cuales $P(U \leq u | H_0$ es verdadera) < 0.05.
5. Las observaciones se arreglan en orden ascendente y se asignan rangos de 1 a 9.

Datos iniciales	0.5	0.9	1.4	1.9	2.1	2.8	3.1	4.6	5.3
Rangos	1	2	3	4	5	6	7	8	9

Para hacer la identificación se subrayan las observaciones de los ratones vacunados. Ahora,

$$w_1 = 1 + 4 + 6 + 7 = 18$$

y

$$w_2 = [(9)(10)/2] - 18 = 27.$$

Por lo tanto,

$$u_1 = 18 - [(4)(5)/2] = 8$$

$$u_2 = 27 - [(5)(6)/2] = 12.$$

de manera que $u = 8$. Como $P(U \leq 8 | H_0$ es verdadera) $= 0.365 > 0.05$, el valor $u = 8$ cae dentro de la región de aceptación.

6. Conclusión: Se acepta H_0 y se concluye que la vacuna no prolonga la vida por detención de la leucemia.

Ejemplo 13.2 El contenido de nicotina encontrado en dos marcas de cigarrillos, medido en miligramos, fue el siguiente:

Marca A	2.1	4.0	6.3	5.4	4.8	3.7	6.1	3.3		
Marca B	4.1	0.6	3.1	2.5	4.0	6.2	1.6	2.2	1.9	5.4

Al nivel de significancia de 0.05, pruebe la hipótesis de que el contenido promedio de nicotina de las dos marcas es el mismo, frente a la alternativa de que no lo es.

Solución Se procede con la regla de seis pasos con $n_1 = 8$ y $n_2 = 10$.

1. H_0: $\mu_1 = \mu_2$.
2. H_1: $\mu_1 \neq \mu_2$.
3. $\alpha = 0.05$.
4. Región crítica: $U \leq 17$ (de la tabla XVII).
5. Cálculos: Las observaciones se arreglan en orden ascendente y se asignan rangos de 1 a 18.

Datos iniciales	Rangos
0.6	1
1.6	2
1.9	3
2.1	4
2.2	5
2.5	6
3.1	7
3.3	8
3.7	9
4.0	10.5
4.0	10.5
4.1	12
4.8	13
5.4	14.5
5.4	14.5
6.1	16
6.2	17
6.3	18

En este caso se subrayan las observaciones pertenecientes a la muestra menor. Ahora

$$w_1 = 4 + 8 + 9 + 10.5 + 13 + 14.5 + 18 = 93$$

y

$$w_2 = \left[\frac{(18)(19)}{2}\right] - 93 = 78.$$

Por lo tanto,

$$u_1 = 93 - \left[\frac{(8)(9)}{2}\right] = 57$$

$$u_2 = 78 - \left[\frac{(9)(10)}{2}\right] = 33.$$

de tal manera que $u = 33$.

6. Conclusión: Se acepta H_0 y se concluye que no hay diferencia en el contenido promedio de nicotina de las dos marcas de cigarrillos.

La prueba de dos muestras de Wilcoxon no está restringida a poblaciones no normales. Se puede usar el lugar de la prueba t cuando las poblaciones son normales, aunque la probabilidad de cometer un error tipo II será mayor. Será siempre superior a la prueba t para poblaciones no normales.

13.3 PRUEBA DEL SIGNO

Supóngase que se seleccionan n pares de observaciones tomadas de dos poblaciones no normales definidas sobre un espacio muestral *continuo*. Para n grandes, en muestreos repetidos la distribución de la media de las diferencias de los pares acoplados de observaciones es aproximadamente normal y las pruebas de las hipótesis relativas a las dos medias poblacionales se pueden efectuar mediante la estadística

$$T = \frac{\bar{D} - d_0}{S_d/\sqrt{n}},$$

dada en la tabla 7-1. Sin embargo, si $n < 30$ y la población de diferencias no es claramente normal, se debe recurrir a una prueba no paramétrica. La más fácil y rápida de efectuar es quizá la llamada *prueba del signo*. Al probar la hipótesis nula H_0 de que $\mu_1 = \mu_2$ o que $\mu_D = 0$, se asigna un signo de *mas* o *menos* a cada diferencia d_i de las observaciones apareadas, según d_i sea positiva o negativa. Si la hipótesis nula es verdadera y las poblaciones son simétricas, la suma de los signos más deberá ser aproximadamente igual a la suma de los signos menos. Cuando un signo aparece con más frecuencia de lo que debiera, basado en el solo resultado, se rechaza la hipótesis de que las medias poblacionales son iguales.

Para proporcionar una estadística apropiada de la prueba del signo, se representará mediante r_- y r_+ la cantidad de signos menos y más contenidos en la muestra aleatoria de observaciones apareadas. La estadística de prueba es definida como la variable aleatoria R que toma el valor r en un experimento particular, donde

$$r = \text{la menor de } r_- \text{ y } r_+.$$

La prueba del signo solamente es aplicable en situaciones donde no ocurran diferencias iguales a cero entre las observaciones apareadas. Aunque una diferencia cero es teóricamente imposible, pues en la práctica las poblaciones continuas se presentan por falta de precisión al registrar los datos. Cuando se observan tales diferencias deben excluirse del análisis, reduciendo el número correspondiente de observaciones apareadas. Así, en un experimento particular se tiene que $n = r_- + r_+$.

Si la hipótesis nula de que $\mu_1 = \mu_2$ es verdadera, la probabilidad de que la diferencia de un par acoplado resulte con signo más o con signo menos es igual a $1/2$. Por lo tanto la estadística de prueba R tiene una distribución de probabilidad binomial con parámetro $p = 1/2$ cuando H_0 es verdadera y se pueden calcular los niveles de significancia tanto para la alternativa unilateral como para la bilateral. Por ejemplo, al probar

$$H_0: \quad \mu_1 - \mu_2 = 0,$$

$$H_1: \quad \mu_1 - \mu_2 < 0,$$

se rechazará H_0 en favor de H_1 solamente si la proporción de signos más es suficientemente menor que $1/2$, esto es, cuando $r = r_+$ es pequeña. Por lo tanto, la región crítica se establece al formar la desigualdad $R < r^*$, donde r^* es un entero positivo apropiado menor que $n/2$ de manera que proporcione un valor razonable para el nivel de significancia,

$$\alpha = P(R < r^* \mid H_0 \text{ es verdadera}).$$

Como la distribución binomial toma valores discretos no se puede esperar encontrar valores de r^* que establezcan las regiones críticas normales de tamaño exactamente igual a 0.01 ó 0.05. Sin embargo, este problema sólo se presenta en muestras de tamaño pequeño. Por ejemplo, cuando $n = 15$ y $r^* = 5$, de la tabla II (cf. Tablas estadísticas) se encuentra que

$$\alpha = P(R < 5) = \sum_{x=0}^{4} b(x; 15, 0.5) = 0.0592.$$

de manera que la región crítica $R < 5$ resulta una estadística de prueba al nivel de significancia 0.0592. Sin embargo, si $n = 6$ y $r^* = 1$, se encuentra

$$\alpha = P(R < 1) = b(0; 15, 0.5) = 0.0156$$

en tanto que con $r^* = 2$ se obtiene

$$\alpha = P(R < 2) = b(0; 15, 0.5) + b(1; 15, 0.5)$$
$$= 0.1094,$$

por lo que ningún valor de r^* produce un nivel de significancia de tamaño α cercano a 0.05.

Para probar la hipótesis

$$H_0: \ \mu_1 - \mu_2 = 0,$$
$$H_1: \ \mu_1 - \mu_2 > 0,$$

se rechaza H_0 en favor de H_1 sólo si la proporción de signos *menos* es suficientemente menor que 1/2, esto es, cuando $r = r_-$ es pequeña. Por lo tanto, la región crítica de tamaño α se establece una vez más formando el intervalo $R < r^*$ donde α se calcula, como antes, mediante la distribución de probabilidad binomial. Por último, para probar la hipótesis

$$H_0: \ \mu_1 - \mu_2 = 0,$$
$$H_1: \ \mu_1 - \mu_2 \neq 0,$$

se rechaza H_0 en favor de H_1 cuando la proporción de signos de *más* (o signos de *menos*) es significativamente menor que o mayor que 1/2. Esto equivale a decir que tanto r_+ o r_- son suficientemente pequeñas o grandes. Sin embargo, la simetría permite usar nuevamente una región crítica unilateral $R < r^*$ en esta prueba de dos colas con el nivel de significancia dado ahora por

$$\alpha = 2P(R < r^* \mid H_0 \text{ es verdadera}).$$

Como los valores de α están calculados a partir de una distribución binomial con parámetro $p = 1/2$, se podría usar la curva normal para aproximar α, como se dijo en la sección 4.3, siempre que $n > 10$. Por lo tanto, para $n = 15$ y $r^* = 5$,

$$\alpha = P(R < 5) \simeq P(X < 4.5),$$

donde X es una variable aleatoria normal con media

$$\mu = np = (15)(0.5) = 7.5$$

y desviación estándar

$$\sigma = \sqrt{npq} = \sqrt{(15)(0.5)(0.5)} = 1.936.$$

Por lo tanto,

$$z = \frac{4.5 - 7.5}{1.936} = -1.55$$

y

$$\alpha = P(R < 5) \simeq P(Z < -1.55)$$
$$= 0.0606,$$

que coinciden aproximadamente con el valor exacto de 0.0592 calculado antes.

Los niveles de significancia correspondientes a las diferentes regiones críticas se pueden obtener rápidamente de la tabla XVIII (cf. Tablas estadísticas), donde $P(R < r^* \mid H_0$ es verdadera) está dada para valores significativos de r^* y valores de $n = 5, 6, \ldots, 25$. Ya que H_0 se rechaza siempre en favor de cada H_1 cuando $R < r^*$, la tabla sólo da niveles de significancia para alternativas unilaterales. Los valores de la tabla deben duplicarse para proporcionar los niveles de significancia correctos para alternativas bilaterales. Por ejemplo, si $n = 21$, una prueba de una cola rechazará H_0 cuando $R < 7$ al nivel de significancia de 0.039, en tanto que una prueba de dos colas rechazará H_0 en favor de H_1 cuando $R < 6$ al nivel de significancia (2) (0.013) = 0.026.

Ejemplo 13.3 Una compañía de taxis está estudiando la conveniencia de substituir los neumáticos normales por neumáticos radiales y economizar más combustible. Doce automóviles se equiparon con neumáticos radiales y se asignaron a una ruta determinada de antemano. Sin cambiar de conductores, se equiparon con neumáticos normales y se asignaron a la misma ruta. Los consumos de gasolina, en km por litro, se registraron así:

Kilómetros por litro

Automóvil	*Neumáticos radiales*	*Neumáticos normales*
1	4.2	4.1
2	4.7	4.9
3	6.6	6.2
4	7.0	6.9
5	6.7	6.8
6	4.5	4.4
7	5.7	5.7
8	6.0	5.8
9	7.4	6.9
10	4.9	4.7
11	6.1	6.0
12	5.2	4.9

¿Se puede concluir que los automóviles equipados con neumáticos radiales economizan más combustible que los equipados con neumáticos normales?

Solución Se representarán con μ_1 y μ_2 los kilometrajes por litro de gasolina de los automóviles equipados con neumáticos radiales y normales respectivamente. El examen de los datos indica nueve signos más, dos signos menos y un cero. Por lo tanto, con $n = 11$, después de excluir el cero se procede así:

1. H_0: $\mu_1 - \mu_2 = 0$.
2. H_1: $\mu_1 - \mu_2 > 0$.
3. $\alpha = 0.033$ (de la tabla XVIII).
4. Región crítica: $R < 3$.
5. Cálculos: $r_+ = 9$, $r_- = 2$, así que $r = 2$, la menor de r_+ y r_-.
6. Conclusión: Se rechaza H_0 y se concluye que, en promedio, los neumáticos radiales mejoran la economía de combustible.

La prueba del signo para observaciones apareadas puede utilizarce también para probar la hipótesis nula de que $\mu_1 - \mu_2 = d_0$. Simplemente se asignan signos de más o menos después que cada d_i ha sido ajustada al restarle d_0 y se procede como antes. También se puede aplicar para probar la hipótesis $\mu = \mu_0$ en una muestra aleatoria tomada de una sola población. En este caso se asignan signos de más o menos a las diferencias $(x_i - \mu_0)$ y se aplican el mismo procedimiento.

La prueba del signo no sólo es uno de los procedimientos paramétricos más fáciles de aplicar sino que además tiene la ventaja de poder aplicarse a datos dicotómicos que no es posible registrar sobre una escala numérica pero que pueden representarse mediante respuestas positivas y negativas. Por ejemplo, se usa en experimentos en que se registran respuestas cualitativas tales como "acierto" o "falla" y en experimentos de tipo sensorial donde se registra un signo de más o menos según el catador identifique correcta o incorrectamente el ingrediente deseado.

13.4 PRUEBA DE WILCOXON PARA OBSERVACIONES APAREADAS

La prueba del signo muestra, mediante la asignación de un signo de más o menos, cuál es el miembro más grande en un par de observaciones pero no indica la magnitud de la diferencia. En 1945, Wilcoxon propuso una prueba que utilizara tanto la dirección como la magnitud; se le conoce con el nombre de *prueba de Wilcoxon para observaciones apareadas*. Detecta con mayor sensibilidad que la prueba del signo una diferencia en la medias poblacionales y por lo tanto se analizará detalladamente.

Para probar la hipótesis de que $\mu_1 = \mu_2$ mediante la prueba de Wilcoxon, primero se descartan todas las diferencias iguales a cero y después se asignan rangos a las diferencias d_i restantes sin atender a su signo. El rango 1 se asigna a la d_i que tenga el menor valor absoluto y el rango 2 a la que ocupe el segundo lugar y así sucesivamente. Cuando el valor absoluto de dos o más diferencias

es el mismo, se asignan a cada una el promedio de los rangos que les corresponderían si no fueran iguales. Si no hay diferencia entre las medias poblacionales, el total de rangos correspondientes a las diferencias negativas. Los totales se representarán respectivamente por w_+ y w_- Se designa con w a la menor de w_+ y w_- y se encuentra la probabilidad de obtener, sólo por azar, un valor menor o igual a w cuando H_0 es verdadera.

Al seleccionar muestras repetidas de observaciones apareadas, se puede esperar que w varíe. Esto hace pensar que w es un valor de alguna variable aleatoria W. Una vez que se conoce la distribución de W, se puede determinar $P(W \le w \mid H_0$ es verdadera). En un nivel de significancia igual a α, H_0 se rechaza cuando

$$P(W \le w \mid H_0 \text{ es verdadera}) < \alpha$$

y se acepta la alternativa unilateral apropiada. En el caso de una prueba de dos colas, se rechaza H_0 al nivel de significancia α en favor de la hipótesis alternativa bilateral de que $\mu_1 \ne \mu_2$ cuando

$$2P(W \le w \mid H_0 \text{ es verdadera}) < \alpha.$$

Si se supone que no hay diferencia en las medias poblacionales, cada d_i tiene la misma posibilidad de ser positiva o negativa. Por lo tanto, un rango puede recibir su signo de dos maneras. Para n diferencias, hay 2^n maneras iguales de que los n rangos reciban su signo. Sea $n(W \le w)$ el número de las 2^n maneras de asignar signo a los n rangos, de modo que el valor de W no exceda a w. Entonces

$$P(W \le w \mid H_0 \text{es verdadera}) = \frac{n(W \le w)}{2^n}.$$

Por ejemplo, pongamos el caso de $n = 6$ pares acoplados que proporcionan un valor $w = 5$. ¿Cuál es la probabilidad de que $W \le 5$ cuando las dos medias poblacionales son iguales? Los conjuntos de rangos cuyo total no excede de 5 se pueden listar así:

Valor de W	*Conjuntos de rangos que suman W*
0	$\varnothing$
1	{1}
2	{2}
3	{3}, {1, 2}
4	{4}, {1, 3}
5	{5}, {1, 4}, {2, 3}

Por lo tanto, en $2^6 = 64$ hay $n(W \leq 5) = 10$ casos igualmente posibles. De donde

$$P(W \leq 5 \mid H_0 \text{ es verdadera}) = \tfrac{10}{64} = 0.1563,$$

un resultado que tiene bastante probabilidad de ocurrir cuando $\mu_1 = \mu_2$.

Cuando $5 \leq n \leq 30$ la tabla XIX (cf. Tablas estadísticas) da en forma aproximada los valores críticos de W para niveles de significancia iguales a 0.01, 0.025 y 0.05 para la prueba de una cola e iguales a 0.02, 0.05 y 0.10 para la prueba de dos colas. En el ejemplo anterior, con $n = 6$ la tabla muestra que se requiere un valor de $W \leq 2$ para que la prueba de una cola sea significativa al nivel de 0.05. Cuando $n > 30$, la distribución muestral de W se aproxima a la distribución normal con media

$$\mu_W = \frac{n(n+1)}{4}$$

y variancia

$$\sigma_W^2 = \frac{n(n+1)(2n+1)}{24}.$$

En este caso con la estadística $Z = (W - \mu_W)/\sigma_W$ se puede determinar la región crítica de la prueba.

La prueba de Wilcoxon para observaciones apareadas sirve también para probar la hipótesis nula de que $\mu_1 - \mu_2 = \mu_D = d_0$. Simplemente se aplica el mismo procedimiento anterior después que cada d_i se ajusta al restarle d_0. Por lo tanto, para probar una hipótesis acerca de la diferencia entre las medias de dos poblaciones cuyas distribuciones son desconocidas y en las que las observaciones ocurren en pares y el tamaño de la muestra es pequeño, se procede con los seis pasos siguientes:

1. H_0: $\mu_1 - \mu_2 = \mu_D = d_0$.
2. H_1: Las alternativas son $\mu_1 - \mu_2 < d_0$, $\mu_1 - \mu_2 > d_0$, o $\mu_1 - \mu_2 \neq d_0$.
3. Escoja un nivel de significancia igual a α.
4. Región crítica:
   a) Todos los valores w para los cuales $P(W \leq w \mid H_0$ es verdadera$) < \alpha$ cuando $n<5$ y la prueba es de una cola.
   b) Todos los valores w para los cuales $2P(W \leq w \mid H_0$ es verdadera$) < \alpha$ cuando $n<5$ y la prueba es de dos colas.
   c) Todos los valores w menores que o iguales al valor crítico apropiado, en la tabla XIX, cuando $5 \leq n \leq 30$.
5. Asigne rangos a las n diferencias $d_i - d_0$, sin tomar en cuenta su signo y después calcular w.
6. Conclusión: Rechace H_0 si w cae en la región crítica; en caso contrario, acepte H_0.

Ejemplo 13.4 Se dice que un estudiante de bachillerato puede mejorar su puntuación en los exámenes de posgrado al menos en 50 puntos, si se le proporciona con anticipación una muestra de los problemas. Para probar lo anterior se seleccionan 20 estudiantes y se forman 10 parejas, de manera que una haya tenido casi la misma puntuación promedio durante los años de estudio. La muestra de los problemas y sus respuestas se entregaron al azar a un miembro de cada pareja una semana antes del examen final. Se obtuvieron estas puntuaciones:

	Pareja									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Con muestra de problemas	531	621	663	579	451	660	591	719	543	575
Sin muestra de problemas	509	540	688	502	424	683	568	748	530	524

Al nivel de significancia de 0.05, pruebe la hipótesis nula de que la muestra de problemas incrementa las puntuaciones en 50 puntos contra la hipótesis alternativa de que las incrementa en menos de 50 puntos.

Solución Sean μ_1 y μ_2 las medias de las puntuaciones de los estudiantes que recibieron la muestra de problemas y de los que no la recibieron. Con el procedimiento de seis pasos, se tiene:

1. H_0: $\mu_1 - \mu_2 = 50$.
2. H_1: $\mu_1 - \mu_2 < 50$.
3. $\alpha = 0.05$.
4. Región crítica: Como $n = 10$, la tabla XIX muestra que la región crítica es $W \leq 11$.
5. Cálculos:

	Pareja									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
d_i	22	81	−25	77	27	−23	23	−29	13	51
$d_i - d_0$	−28	31	−75	27	−23	−73	−27	−79	−37	1
Rangos	5	6	9	3.5	2	8	3.5	10	7	1

Ahora, $w_+ = 10.5$ y $w_- = 44.5$, por lo que $w = 10.5$, la menor de w_+ y w_-.

6. Conclusión: Se rechaza H_0 y se concluye que la muestra de problemas, en promedio, no incrementa la puntuación de los exámenes finales en 50 puntos.

13.5 PRUEBA DE CORRIDAS

Al aplicar todos los conceptos estadísticos descritos a lo largo de este libro, se ha supuesto que los datos muestrales se reúnen mediante un proceso aleatorio. La *prueba de ensayos* se basa en el orden en que se obtienen las observaciones muestrales y es una técnica útil para probar la hipótesis nula H_0 de que las observaciones se han tomado efectivamente al azar.

Para ilustrar este tipo de prueba supóngase que se hace una encuesta entre 12 personas para averiguar si usan cierto producto. Se podría poner en duda la aleatoriedad de la muestra si las 12 personas fueran del mismo sexo. Si se indica a un hombre por H y a una mujer por M y se registran las respuestas a partir del sexo y en el orden en que ocurren, una secuencia típica para este experimento podría

$$\underline{H\ H}\quad \underline{M\ M\ M}\quad H\quad \underline{M\ M}\quad \underline{M\ M\ M\ M},$$

donde se han agrupado las subsecuencias de símbolos similares. A tales agrupamientos se les llama corridas.

> DEFINICION 13.1 *Una* corrida *es una subsecuencia de uno o más símbolos idénticos que representan una propiedad común de los datos.*

Sin importar si las mediciones muestrales representen datos cualitativos o cuantitativos, la prueba los divide en dos categorías mutuamente excluyentes: hombre o mujer, defectuoso o no defectuoso, caras o sellos, arriba o abajo de la mediana, etc. Por eso una secuencia siempre estará limitada a dos símbolos distintos. Sea n_1 el número de símbolos asociados con la categoría que ocurre menos veces y n_2 el número de símbolos que pertenece a la otra categoría. El tamaño de la muestra será entonces $n = n_1 + n_2$.

Para los $n = 12$ símbolos de la encuesta se tienen cinco corridas, en las que la primera contiene dos H, la segunda tres M, etc. Si el número de corridas es mayor o menor de lo que cabría esperar, debe rechazarse la hipótesis de que la muestra fue tomada al azar. Así, una muestra que produzca sólo dos corridas,

$$M\quad M\ M\ M\ M\ M\ M\ H\quad H\ H\ H\ H,$$

o a la inversa, tiene pocas probabilidades de ocurrir en un proceso de selección al azar. Tal resultado indica que las primeras siete personas entrevistadas fueron hombres y las cinco restantes mujeres. Asimismo si el resultado es el

número máximo de 12 corridas, como sucede en la secuencia alternante se puede sospechar del orden en que se seleccionaron los individuos.

$$H\ M\ H\ M\ H\ M\ H\ M\ H\ M\ H\ M,$$

La prueba de corridas por aleatoriedad se basa en la variable aleatoria V, el número total de corridas que ocurren en la secuencia completa de un experimento. Una región de rechazo para una prueba de dos colas puede expresarse así: $V \le a$ y $V \ge b$. Conocer la distribución muestral de V permite establecer los niveles de significancia para tal prueba. En la tabla XX (cf. Tablas estadísticas) se dan los valores de $P(V \le a | H_0$ es verdadera) para valores de n_1 y n_2 menores que o iguales a 10. Usando los valores tabulados se pueden obtener las regiones de rechazo para las pruebas de una y dos colas. En la encuesta efectuada anteriormente se presentan un total de cinco M y siete H. Así, con $n_1 = 5$, $n_2 = 7$ y $v = 5$ la tabla XX indica que $P(V \le 5 | H_0$ es verdadera) = 0. 197 > 0.025. Es decir, el valor $v = 5$ es razonable cuando H_0 es verdadera y, por tanto, no se tiene suficiente evidencia para rechazar la hipótesis de aleatoriedad en la muestra. Los valores críticos de una prueba de dos colas, para un nivel de significancia exactamente de 0.05 ó 0.01, no se puede encontrar a partir de la tabla XX. Sin embargo, para $n_1 = 5$ y $n_2 = 7$ se nota que $P(V \le 3 | H_0$ es verdadera) = 0.015 y $P(V \ge 11 | H_0$ es verdadera) = 1 − $P(V \le 10 | H_0$ es verdadera) = 1 − 0.992 = 0.008. Por lo tanto, la hipótesis de aleatoriedad se podría rechazar cuando $V \le 3$ ó $V \ge 11$ al nivel de significancia de $\alpha = 0.015 + 0.008 = 0.023$. Como ya se ha visto, el valor $v = 5$ de esta muestra cae dentro de la región de aceptación.

La prueba de corridas se puede usar también para detectar las desviaciones de la aleatoriedad que con el tiempo aparecen en una secuencia de mediciones cuantitativas y que provienen de tendencias o periodicidades. Al reemplazar cada medición, en el orden en que se recoge (por un símbolo *más* si cae arriba de la mediana, por un símbolo *menos* si cae abajo de la mediana) y al omitir todas las mediciones que son exactamente iguales a la mediana, se genera una secuencia de símbolos más y menos que puede probarse por la aleatoriedad como se ilustra en el ejemplo siguiente.

Ejemplo 13.5 Una máquina se ajusta para surtir un líquido en un recipiente. Si se encuentra que las mediciones efectuadas en 15 recipientes consecutivos son 3.6, 3.9, 4.1, 3.6, 3.8, 3.7, 3.4, 4.0, 3.8, 4.1, 3.9, 4.0, 3.8, 4.2 y 4.1 litros, ¿se puede decir que la cantidad de líquido surtida por esta máquina varía aleatoriamente?

Solución Para la muestra dada se encuentra que $\tilde{x} = 3.9$. Reemplazando cada medición por el símbolo "+" si cae arriba de 3.9, por el símbolo "−" si cae abajo de 3.9 y omitiendo las dos mediciones iguales a 3.9, se obtiene la secuencia

$$-\ +\ -\ -\ -\ -\ +\ -\ +\ +\ -\ +\ +$$

para la cual $n_1 = 6$, $n_2 = 7$ y $v = 6$. Si se consulta la tabla XX se encuentra que $P(V \le 6 \mid H_0$ es verdadera$) = 0.298 > 0.025$. Por lo tanto, se acepta la hipótesis de que la secuencia de las mediciones varía aleatoriamente.

Aunque la prueba de corridas tiene poca potencia, sirve también de alternativa a la prueba de dos muestras de Wilcoxon para probar el enunciado de que dos muestras aleatorias provienen de poblaciones con las mismas distribuciones y, por lo tanto, con medias iguales. Si las poblaciones son simétricas, rechazar el enunciado de que las distribuciones son iguales equivale a aceptar la hipótesis alterna de que las medias no son iguales. Al efectuar la prueba, se combinan primero las observaciones de ambas muestras y se arreglan en orden ascendente. Después se asigna la letra A a cada observación tomada de una de las poblaciones y la letra B a cada observación tomada de la segunda población, generando así una secuencia consistente de los símbolos A y B.

Para los años de supervivencia de los pacientes del ejemplo 13.1, se tiene

0.5	0.9	1.4	1.9	2.1	2.8	3.1	4.6	5.3
B	A	A	B	A	B	B	A	A.

cuyo resultado es $v = 6$ corridas. Si las dos poblaciones simétricas tienen medias iguales, las observaciones de las dos muestras se entremezclarán y originarán muchas corridas. Sin embargo, si las medias poblacionales son significativamente diferentes, cabe esperar que la mayoría de las observaciones para una de las dos muestras serán menores que las de la otra muestra. En el caso extremo en que las poblaciones no se traslapan, se puede obtener una secuencia como la siguiente:

$$A\ A\ A\ A\ A\ B\ B\ B\ B \quad \text{o} \quad B\ B\ B\ B\ A\ A\ A\ A\ A.$$

y en cualquier caso, sólo habrá dos corridas. Así pues, se rechazará la hipótesis de igualdad de las medias poblacionales cuando V es pequeña y cae dentro de la región crítica $V \le v_\alpha$, lo cual indica una prueba de una cola.

Volvamos ahora a los datos del ejemplo 13.1 para el que $n_1 = 4$, $n_2 = 5$, y $v = 6$, de la tabla XX se encuentra que $P(V \le 6 \mid H_0$ es verdadera$) = 0.786$ y por lo tanto se acepta la hipótesis nula de que las medias son iguales. Con objeto de rechazar H_0 incluso al nivel de significancia de 0.071, se requerirá un valor de $V \le 3$. De aquí se concluye que la nueva vacuna no prolonga la vida deteniendo la leucemia.

Cuando n_1 y n_2 aumentan de tamaño, la distribución muestral de V se aproxima a la distribución normal con media

$$\mu_V = \frac{2n_1 n_2}{n_1 + n_2} + 1$$

y variancia

$$\sigma_V^2 = \frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

Así pues, cuando n_1 y n_2 son mayores que 10 se puede usar la estadística para establecer la región crítica en la prueba de corridas.

$$Z = \frac{V - \mu_V}{\sigma_V}$$

13.6 LIMITES DE TOLERANCIA

Los límites de tolerancia en una distribución normal de mediciones se explicaron en el capítulo 6. En esta sección se describirá un método para construir los intervalos de tolerancia que no dependen de la forma de la distribución fundamental. Como cabe suponer para un grado de confianza razonable, serán substancialmente más amplios que los que se construyen suponiendo normalidad y el tamaño de muestra requerido suele ser muy grande. Los límites de tolerancia no paramétrica se establecen en términos de las observaciones más pequeña y más grande de la muestra.

> LIMITES DE TOLERANCIA BILATERALES *En cualquier distribución de mediciones, los límites de tolerancia bilaterales están dados por las observaciones más pequeña y más grande en una muestra de tamaño n, donde n está determinada de manera que se puede afirmar con una confianza del* $100\gamma\%$ *que* al menos *la proporción* $1 - \alpha$ *de la distribución está incluida entre los extremos de la muestra.*

La tabla XXI (ver Tablas estadísticas) proporciona los tamaños de muestra requeridos para algunos valores de γ y $1 - \alpha$. Por ejemplo: cuando $\gamma = 0.99$ y $1 - \alpha = 0.95$, se debe escoger una muestra aleatoria de tamaño $n = 130$ con objeto de tener una confianza del 99% de que al menos el 95% de los valores de la distribución estén incluidos entre los extremos de la muestra.

En lugar de determinar un tamaño n de la muestra que contenga una proporción específica de las mediciones entre los extremos, en algunos procesos industriales conviene determinar un tamaño de manera que una proporción fija de la poblaciones caiga por debajo de la observación mayor (o por arriba de la más pequeña) de la muestra. Tales límites se llaman límites de tolerancia unilaterales.

LIMITES DE TOLERANCIA UNILATERALES *Para cualquier distribución de mediciones, los límites de tolerancia unilaterales están dados por la observación más pequeña (o más grande) en una muestra de tamaño n donde n está determinada de manera que se puede afirmar, con una confianza del* $100\gamma\%$ *que* al menos *la proporción* $1-\alpha$ *de la distribución excederá (será menor que) la observación más pequeña (más grande) de la muestra.*

La tabla XXII (cf. Tablas estadísticas) proporciona los tamaños de muestra requeridos correspondientes a algunos valores de γ y $1-\alpha$. Así, cuando $\gamma = 0.95$ y $1-\alpha = 0.70$, se debe escoger una muestra de tamaño $n = 9$ con objeto de tener una confianza del 95% de que el 70% de valores de la distribución exceda la observación más pequeña de la muestra.

13.7 COEFICIENTE DE CORRELACION DEL RANGO

En el capítulo 8 se usó el coeficiente de correlación muestral r para medir la relación lineal entre dos variables continuas X y Y. Si se asignan los rangos 1, 2,. . . , n a las observaciones x por orden de magnitud y a las observaciones y y si luego se substituyen estos rangos por sus valores numéricos reales en la fórmula de r, se obtiene la contraparte no paramétrica del coeficiente de correlación ordinario. Un coeficiente de correlación calculado en esta forma se conoce como *coeficiente de correlación del rango de Spearman* y se denota por r_s. Cuando no hay valores iguales dentro de cualquier conjunto de mediciones, la fórmula para r_s se reduce a una expresión más simple, que se define a continuación.

COEFICIENTE DE CORRELACION DEL RANGO DE SPEARMAN *Una medida no paramétrica de la asociasión entre dos variables X y Y está dada por el coeficiente de correlación del rango*

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)},$$

donde d_i *es la diferencia entre los rangos asignados a* x_i *y* y_i*; y n el número de parejas de datos.*

En la práctica, la fórmula anterior se usa también cuando hay valores iguales dentro de cualquiera de las observaciones x o y. En el caso de observaciones de valores iguales, los rangos se asignan como en la prueba de Wilcoxon para observaciones apareadas, promediando los rangos que podrían haberse asignado si las observaciones fueran distintas.

Casi siempre el valor de r_s está cercano al valor obtenido al encontrar r, basado en las mediciones numéricas, y se interpreta en la misma forma. También en este caso los valores de r_s variarán desde -1 hasta $+1$. Un valor $+1$ ó -1 indica una perfecta asociación entre X y Y, el signo más ocurre para rangos idénticos y el signo menos para rangos opuestos. Cuando r_s se aproxima a cero, se puede concluir que las variables están correlacionadas.

Ejemplo 13.6 Las cifras siguientes muestran los miligramos de alquitrán y nicotina encontrados en 10 marcas de cigarrillos.

Marca de cigarrillo	*Contenido de alquitrán*	*Contenido de nicotina*
Viceroy	14	0.9
Marlboro	17	1.1
Chesterfield	28	1.6
Kool	17	1.3
Kent	16	1.0
Raleigh	13	0.8
Old Gold	24	1.5
Philip Morris	25	1.4
Oasis	18	1.2
Players	31	2.0

Calcule el coeficiente de correlación del rango para medir el grado de relación entre los contenidos de alquitrán y nicotina de los cigarrillos.

Solución Sean X y Y los contenidos de alquitrán y nicotina, respectivamente. Primero se asignan rangos a cada conjunto de mediciones, asignando el rango 1 al número menor de cada conjunto, el rango 2 al siguiente número menor, etc., hasta asignar el rango 10 al número más grande. La tabla 13-1 muestra los rangos individuales de las mediciones y sus diferencias para los 10 pares de observaciones.

Substituyendo en la fórmula para r_s, se encuentra que

$$r_s = 1 - \frac{(6)(5.5)}{(10)(100-1)} = 0.9667,$$

indicando una alta correlación positiva entre las cantidades de alquitrán y nicotina encontradas en los cigarrillos.

El uso de r_s en lugar de r ofrece varias ventajas. Por ejemplo, no se supone que la relación fundamental entre X y Y sea lineal y, por lo tanto, cuando los datos poseen una relación curvilínea distinta, el coeficiente de correlación del

Tabla 13-1 Rangos para los contenidos de alquitrán y nicotina

Marca de cigarrillos	x	y	d
Viceroy	2	2	0
Marlboro	4.5	4	0.5
Chesterfield	9	9	0
Kool	4.5	6	−1.5
Kent	3	3	0
Raleigh	1	1	0
Old Gold	7	8	−1
Philip Morris	8	7	1
Oasis	6	5	1
Players	10	10	0

rango será más confiable que la medida normal. Una segunda ventaja es que no se hacen suposiciones de normalidad con respecto a las distribuciones de X y Y. Quizá la ventaja más grande se logra cuando es imposible realizar mediciones numéricas significativas y a pesar de ello se pueden establecer los rangos. Eso ocurre, por ejemplo, cuando jueces diferentes asignan rangos a un grupo de individuos a partir de alguna cualidad, el coeficiente de correlación del rango sirve entonces de medida a la concordancia de los dos jueces.

Para probar la significancia del coeficiente de correlación del rango es necesario considerar la distribución de los valores r_s bajo la suposición de que X y Y son independientes. En la tabla XXIII (cf. Tablas estadísticas) se dan los valores críticos calculados para $\alpha = 0.05$, 0.025, 0.01 y 0.005. La tabla tiene forma similar a la de los valores críticos para la distribución t, excepto que ahora la columna izquierda contiene el número de parejas de observaciones y no el de los grados de libertad. Como la distribución de los valores r_s es simétrica con respecto a $r_s = 0$, el valor r_s que deja un área de α a la izquierda, es igual al negativo del valor r_α que deja un área de α a la derecha. En una hipótesis alterna bilateral, la región crítica de tamaño α cae igualmente dentro de las dos colas de la distribución. En una prueba en la que la hipótesis alternativa es negativa, la región crítica se encuentra totalmente en la cola izquierda de la distribución, y cuando la alternativa es positiva está colocada íntegramente en la cola derecha. En el ejemplo 13.6 el valor crítico para probar la hipótesis nula H_0 de que el coeficiente de correlación del rango es cero contra la hipótesis alterna H_1 de que es mayor que cero, con $\alpha = 0.01$ y $n = 10$, es 0.745. Es decir, se rechaza H_0 si $r_s > 0.745$ y como el valor calculado fue $r_s = 0.9667$ concluye, al nivel de significancia de 0.01, que existe una gran correlación positiva entre las cantidades de alquitrán y nicotina halladas en los cigarrillos.

Cuando X y Y son independientes, se puede demostrar que la distribución de los valores r_s se aproxima a la distrubución normal con una media de cero y una desviación estándar de $1/\sqrt{n-1}$ a medida que n aumenta. Así pues, cuando n excede los valores dados en la tabla XXIII, uno puede probar si la correlación es significativa al calcular

$$z = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s\sqrt{n-1}$$

y compararlo después con los valores críticos obtenidos de la curva normal estándar.

La prueba de independencia de dos *variables continuas* presentada en esta sección es una alternativa simplificada del elaborado procedimiento de ji cuadrada con las tablas de contingencia descritas en la sección 7.9. Desafortunadamente, muchas veces todos los valores de las dos variables aleatorias deben caer dentro de ciertas categorías establecidas y por lo tanto no puede medirse sobre una escala continua; se necesitan cálculos más complejos asociados con las tablas de contingencia.

13.8 PRUEBA DE KRUSKAL-WALLIS

La prueba de Kruskal-Wallis, llamada también prueba H de Kruskal-Wallis, es una generalización de la prueba de dos muestras de Wilcoxon para el caso de $k > 2$ muestras. Sirve para probar la hipótesis nula H_0 de que k muestras independientes provienen de poblaciones idénticas. Dada a conocer en 1952 por W. H. Kruskal y W. A. Wallis, es un procedimiento no paramétrico alterno de la prueba F, utilizada para probar la igualdad de medias en el análisis de la varianza de un factor, cuando el investigador desea eliminar la suposición de que las muestras fueron seleccionadas de poblaciones normales.

Sea n_i $(i = 1, 2, \ldots, k)$ el número de observaciones en la i-ésima muestra. Primero se combinan todas las k muestras y se arreglan las $n = n_1 + n_2 + \cdots + n_k$ observaciones en orden ascendente, substituyendo el rango apropiado, $1, 2, \ldots, n$, en cada observación. En el caso de observaciones idénticas, se sigue el procedimiento usual de remplazar las observaciones por las medias de los rangos que tendrían si fueran distintas. La suma de los rangos correspondientes a las n_i observaciones de la i-ésima muestra se indica por la variable aleatoria R_i. Examinemos ahora la estadística

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1),$$

que puede aproximarse bastante bien mediante una distribución ji cuadrada con $k - 1$ grados de libertad cuando H_0 es verdadera y si cada muestra consta al menos de cinco observaciones. Nótese que la estadística H toma el valor h, dónde

$$h = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{r_i^2}{n_i} - 3(n+1),$$

cuando . . . ma el valor r_1, R_2 toma el valor r_2, y así sucesivamente. El hecho de que h sea grande cuando las muestras independientes provienen de poblaciones que no son idénticas permite establecer el siguiente criterio de decisión para probar

> PRUEBA DE KRUSKAL—WALLIS *Para probar la hipótesis nula xx de que k muestras independientes provienen de poblaciones idénticas, se calcula*
>
> $$h = \frac{12}{n(n+1)} \cdot \sum_{i=1}^{k} \frac{r_i^2}{n_i} - 3(n+1).$$
>
> *Si h cae dentro de la región crítica $H > \chi^2_\alpha$ con $v = k - 1$ grados de libertad, se rechaza H_0 al nivel de significancia α; de lo contrario, se acepta H_0.*

Ejemplo 13.7 En un experimento cuya finalidad es determinar cuál de tres sistemas de proyectiles es mejor, se midió la velocidad de combustión del propulsante. Los datos, después de ser codificados, se dan en la tabla 13-2.

Tabla 13-2 Velocidad de combustión de los propulsantes

Sistema de proyectiles		
1	*2*	*3*
24.0	23.2	18.4
16.7	19.8	19.1
22.8	18.1	17.3
19.8	17.6	17.3
18.9	20.2	19.7
	17.8	18.9
		18.8
		19.3

Utilice la prueba de Kruskal-Wallis a un nivel de significancia de $\alpha = 0.05$ para probar la hipótesis de que las velocidades de combustión del propulsante son iguales en los tres sistemas de proyectiles.

Solución En la tabla 13-3 se han convertido a rangos las 19 observaciones y dichos rangos se han sumado para cada sistema de proyectiles.

Tabla 13-3 Rangos para la velocidad de combustión de los propulsantes

Sistema de proyectiles		
1	*2*	*3*
19	18	7
1	14.5	11
17	6	2.5
14.5	4	2.5
9.5	16	13
$r_1 = 61.0$	5	9.5
	$r_2 = 63.5$	8
		12
		$r_3 = 65.5$

Ahora bien, al substituir $n_1 = 5$, $n_2 = 6$, $n_3 = 8$, y $r_1 = 61.0$, $r_2 = 63.5$, $r_3 = 65.5$, la estadística de prueba H toma el valor

$$\begin{aligned} h &= \frac{12}{(19)(20)}\left(\frac{61.0^2}{5} + \frac{63.5^2}{6} + \frac{65.5^2}{8}\right) - (3)(20) \\ &= 1.6586. \end{aligned}$$

Usando un nivel de significancia de 0.05 con $k - 1 = 2$ grados de libertad, se encuentra $\chi^2_{0.05} = 5.991$. Como $h = 1.6586$ no cae en la región crítica $H > 5.991$ no se tienen suficientes datos para rechazar la hipótesis de que las velocidades de combustión del propulsante son iguales en los tres sistemas de proyectiles.

EJERCICIOS

1. Un fabricante de cigarrillos anuncia que el contenido de alquitrán de los cigarrillos marca B es menor que la de marca A. Para probarlo, se anotaron en mg los contenidos de alquitrán:

Marca A	12	9	13	11	14
Marca B	8	10	7		

Utilice la prueba de dos muestras de Wilcoxon con un nivel $\alpha = 0.05$ para determinar si el anuncio es válido.

2. Los datos siguientes representan el número de horas que dos calculadoras de bolsillo pueden operar antes que haya que recargarlas.

Calculadora A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1
Calculadora B	3.8	4.8	4.3	4.2	4.0	4.9	4.5	5.2	4.5

Use la pruebas de dos muestras de Wilcoxon, con $\alpha = 0.01$, para determinar si la calculadora A opera más tiempo que la calculadora B con la carga total de las baterias.

3. Los siguientes datos representan el peso del equipaje de los integrantes de dos equipos de béisbol que viajan en un avión:

Equipo A	34	39	41	28	33	
Equipo B	36	40	35	31	39	36

Use la prueba de dos muestras de Wilcoxon, con $\alpha = 0.05$, para probar la hipótesis de que los dos equipos cargan en promedio la misma cantidad de equipaje contra la hipótesis alternativa de que el peso promedio de equipaje del equipo B es mayor que el del equipo A.

4. Una cuerda para pescar se fabrica mediante dos procesos. Para determinar si existe diferencia en la resistencia media a la rotura de las cuerdas, se seleccionan 10 piezas de cada proceso y se prueba su resistencia. He aquí los resultados:

Proceso 1	10.4	9.8	11.5	10.0	9.9	9.6	10.9	11.8	9.3	10.7
Proceso 2	8.7	11.2	9.8	10.1	10.8	9.5	11.0	9.8	10.5	9.9

Utilice la prueba de dos muestras de Wilcoxon, con $\alpha = 0.1$, para determinar si existe diferencia entre la resistencia media a la rotura de las cuerdas fabricadas con los dos procesos.

5. En un curso de matemáticas, 12 estudiantes de igual capacidad usan materiales programados; a 5 de ellos, que fueron seleccionados al azar, el maestro les imparte instrucciones complementarias para el examen final. Los resultados fueron los siguientes:

	Calificación						
Instrucciones complementarias	87	69	78	91	80		
Sin instrucciones complementarias	75	88	64	82	93	79	67

Utilice la prueba de dos muestras de Wilcoxon, con $\alpha = 0.05$, para determinar si las instrucciones complementarias afectan la calificación promedio.

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

PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES

PRUEBA DE HIPOTESIS

FEBRERO, 1985

Prueba de hipótesis 7

7.1 HIPOTESIS ESTADISTICAS

La prueba de hipótesis estadísticas es quizá el área más importante de la teoría de decisión. En primer término, se definirá lo que se entiende por hipótesis estadística.

> DEFINICION 7.1 Hipótesis estadística *es una suposición o enunciado, que pueden o no ser verdaderos, relativos a una o más poblaciones.*

La verdad o falsedad de una hipótesis nunca se conoce con certeza si no se examina toda la población. Como las más de las veces eso resulta poco práctico, se toma una muestra aleatoria de la población y se usa la información contenida en ella para decidir si la hipótesis es verdadera o falsa. Si los datos de la muestra se oponen a la hipótesis, ésta se la rechaza, mientras que si la apoyan se la acepta. Se debe aclarar aquí que la aceptación de una hipótesis estadística es un resultado de evidencia insuficiente para rechazarla y no implica necesariamente que sea verdadera. Por ejemplo, al lanzar una moneda 100 veces se puede probar la hipótesis de que la moneda está balanceada. En términos de parámetros de la población, se está probando la hipótesis de que la proporción de caras sería $p = 0.5$ si la moneda fuese lanzada indefinidamente. Un resultado de 48 caras no sería sorprendente si la moneda está balanceada y apoyará con seguridad la hipótesis $p = 0.5$. Se puede afirmar que tal ocurrencia también concuerda con la hipótesis de que $p = 0.45$. Así, al aceptar la hipótesis, la única cosa de que podemos estar bastante seguros es que la verdadera proporción de caras se acerca a la mitad. Si de los 100 intentos se obtienen solamente 35 caras, se contará entonces con datos para apoyar el rechazo de la hipótesis. La probabilidad de obtener 35 caras o menos en 100 lanzamientos de una moneda balanceada es aproximadamente 0.002; Por tanto, es posible que haya ocurrido un evento muy raro o que tengamos razón al concluir que $p \neq 0.5$.

En este capítulo se utilizan frecuentemente los términos *aceptar o rechazar*, pero cabe advertir que el rechazo de una hipótesis implica meramente que no se tiene evidencia para creer lo contrario. Debido a esta terminología, el estadígrafo y el investigador han de establecer siempre como sus hipótesis a aquellas que esperan rechazar. Por ejemplo, si les interesa una nueva vacuna contra el resfriado común, deben suponer que ésta no es mejor que la vacuna que se vende actualmente y después prepararse para rechazar esta afirmación. He aquí otro ejemplo, para probar que una técnica de enseñanza es superior a otra, se prueba la hipótesis de que no hay diferencia entre ambas.

La hipótesis que se formula con la esperanza de rechazarla se llama *hipótesis nula* y se indica por H_0 su rechazo lleva a la aceptación de una *hipótesis alterna*, (alternativa) indicada por H_0. Por lo tanto, si H_0 es la hipótesis nula $p = 0.5$ si para una población binomial, la hipótesis alterna H_1 puede ser $p = 0.75$, $p > 0.5$, $p < 0.5$, o $p \neq 0.5$. H_1, etc.

7.2 ERRORES TIPO I Y TIPO II

Para ilustrar los conceptos empleados al probar una hipótesis estadística acerca de una población, considérese el ejemplo siguiente. Se sabe que cierto tipo de vacuna contra el resfriado común sólo es eficaz en un 25% al cabo de 2 años. Para precisar si una nueva vacuna un poco más costosa ofrece mejor protección durante un periodo mayor, se inoculan 20 personas escogidas al azar. Si 9 o más de ellas rebasan ese lapso sin contraer el virus, se considerará que la nueva vacuna es superior a la actual. Escoger el número es arbitrario pero parece razonable puesto que representa una modesta ganancia sobre las 5 personas que podría esperarse que obtuvieran protección, si las 20 hubieran sido inoculadas con la vacuna actual. Esencialmente se está probando la hipótesis nula de que la nueva vacuna es igualmente eficaz que la actual después de 2 años, contra la hipótesis alterna de que la nueva vacuna es realmente superior. Esto equivale a probar la hipótesis de que el parámetro binomial para la probabilidad de un éxito en un ensayo dado es $p = 1/4$, contra la alternativa de que $p > 1/4$. Esto se escribe generalmente como sigue:

$$H_0: \quad p = 1/4,$$
$$H_1: \quad p > 1/4.$$

La estadística en que se basa esta decisión es X, o sea el número de personas en la muestra que reciben protección con la nueva vacuna por un periodo mínimo de 2 años. Los valores posibles, desde cero hasta 20, se dividen en dos grupos: los menores que 9 y los mayores o iguales a él. Todas las puntuaciones posibles situadas arriba de 8.5 constituyen la *región crítica* y todas las puntuaciones posibles que caigan por debajo pertenecen a la *región de aceptación*. Al número 8.5, que separa estas dos regiones, se le llama valor crítico. Si la estadística X cae en la región crítica, se rechaza H_0 en favor de la hipótesis alterna H_1. Si cae en la región de aceptación, H_0 se acepta.

El procedimiento de decisión que acabamos de describir puede conducir a uno u otro tipo de error. Por ejemplo, la nueva vacuna quizá no sea mejor que la actual y puede suceder que para este grupo particular de individuos seleccionados al azar, 9 o más personas rebasen el periodo de 2 años sin contraer el virus. Así, tal vez se esté cometiendo un error al rechazar H_0 en favor de H_1 cuando, de hecho, H_0 es verdadera. A tal error se le conoce con el nombre de *error tipo I*.

> DEFINICION 7.2 *Se ha cometido un* error tipo I *si se rechaza la hipótesis nula cuando ésta es verdadera.*

Se comete una segunda clase de error, si menos de 9 personas del grupo rebasan el periodo de 2 años sin enfermarse y se concluye que la nueva vacuna no es mejor que la actual. En este caso se podría aceptar H_0 cuando es falsa. Esto se llama *error tipo II*.

> DEFINICION 7.3 *Se ha cometido un* error tipo II *si se acepta la hipótesis nula cuando ésta es falsa.*

A la probabilidad de cometer un error tipo I se le llama *nivel de significancia* de la prueba y se denota por α. En el ejemplo en cuestión, ocurrirá un error tipo I cuando 9 o más individuos rebasan el periodo de 2 años sin contraer el virus y usan una nueva vacuna que en realidad es equivalente a la actual. Por lo tanto, si X es el número de personas que permanecen sanas al menos durante 2 años,

Aquí se dice que la hipótesis nula, $p = 1/4$, se está probando al nivel de significancia $\alpha = 0.0409$. Algunas veces, al nivel de significancia se denomina el *tamaño* de la región crítica. Una región crítica 0.0409 de tamaño es muy pequeña y por lo tanto es poco probable que se haya cometido un error tipo I.

$$\begin{aligned}\alpha &= P(\text{error tipo I})\\ &= P(X \geq 9 \mid p = \tfrac{1}{4})\\ &= \sum_{x=9}^{20} b(x; 20, \tfrac{1}{4})\\ &= 1 - \sum_{x=0}^{8} b(x; 20, \tfrac{1}{4})\\ &= 1 - 0.9591\\ &= 0.0409\end{aligned}$$

Así pues, será factible que 9 o más individuos permanezcan inmunes al virus durante 2 años usando una nueva vacuna que equivale a la actual.

La probabilidad de cometer un error tipo II, indicada por β, es imposible de calcular, a menos que se tenga una hipótesis alterna específica. Si se prueba la hipótesis nula de que $p = 1/4$ contra la hipótesis alterna de que $p = 1/2$, entonces se puede calcular la probabilidad de aceptar H_0 cuando ésta es falsa. Basta encontrar la probabilidad de obtener menos de 9 personas en el grupo que rebasen el periodo de 2 años cuando $p = 1/2$. En este caso

$$\begin{aligned}\beta &= P(\text{error tipo II})\\ &= P(X < 9 \mid p = \tfrac{1}{2})\\ &= \sum_{x=0}^{8} b(x; 20, \tfrac{1}{2})\\ &= 0.2517.\end{aligned}$$

Esta es una probabilidad muy alta que indica un pobre procedimiento de prueba. Hay muchas probabilidades de que se rechace la nueva vacuna cuando de hecho es superior a la actual. En teoría, es preferible usar un procedimiento de prueba en el que tanto el error tipo I como el tipo II sean pequeños.

Es posible que el director del programa de pruebas quiera cometer un error tipo II si la vacuna más costosa no es significativamente superior. La única vez que desea protegerse contra el error tipo II, es cuando el verdadero valor de p es la menos 0.7. Si $p = 0.7$, este procedimiento de prueba da

$$\begin{aligned}\beta &= P(\text{error tipo II})\\ &= P(X < 9 \mid p = 0.7)\\ &= \sum_{x=0}^{8} b(x; 20, 0.7)\\ &\doteq 0.0051.\end{aligned}$$

Con una probabilidad tan pequeña de cometer un error tipo II, hay poquísimas probabilidades de que la nueva vacuna sea rechazada cuando es eficaz en un 70% al cabo de 2 años. A medida que la hipótesis alterna se aproxima a la unidad, el valor de β disminuye hacia cero.

Supóngase que el director del programa de pruebas no desea cometer un error tipo II cuando la hipótesis alterna $p = 1/2$ es verdadera, aunque se haya encontrado que la probabilidad de tal error es $\beta = 0.2517$. Siempre es posible una reducción de β al aumentar el tamaño de la región crítica. Por ejemplo, veamos qué sucede a los valores de α y β cuando se cambia el valor crítico a 7.5 para que todas las puntuaciones de 8 o mayores caigan en la región crítica y las menores de 8 caigan en la región de aceptación. Ahora bien, al probar $p = 1/4$ contra la hipótesis alterna de que $p = 1/2$, se encuentra que

$$\alpha = \sum_{x=8}^{20} b(x; 20, \tfrac{1}{4})$$

$$= 1 - \sum_{x=0}^{7} b(x; 20, \tfrac{1}{4})$$

$$= 1 - 0.8982$$

$$= 0.1018$$

y

$$\beta = \sum_{x=0}^{7} b(x; 20, \tfrac{1}{2})$$

$$= 0.1316.$$

Mediante al adopción de un nuevo procedimiento de decisión se ha reducido la probabilidad de cometer un error tipo II a riesgo de aumentar la probabilidad de incurrir en un error tipo I. Para una muestra de tamaño fijo, una disminución de la probabilidad de un error resulta generalmente en un aumento de la probabilidad del otro error. Afortunadamente, la probabilidad de cometer ambos se aminora si se aumenta el tamaño de la muestra. Considerando el mismo problema con una muestra aleatoria de 100 individuos, si 37 o más rebasan el periodo de 2 años, se rechaza la hipótesis nula de que $p = 1/4$ y se acepta la hipótesis alterna $p > 1/4$ y ahora el valor crítico es 36.5. Todas las puntuaciones posibles situadas arriba de 36.5 constituyen la región crítica y todas las que se hallan por debajo de 36.5 caen en la región de aceptación.

Para determinar la probabilidad de cometer un error tipo I, se empleará la aproximación a la curva normal con

$$\mu = np = (100)(\tfrac{1}{4}) = 25$$

y

$$\sigma = \sqrt{npq} = \sqrt{(100)(\tfrac{1}{4})(\tfrac{3}{4})} = 4.33.$$

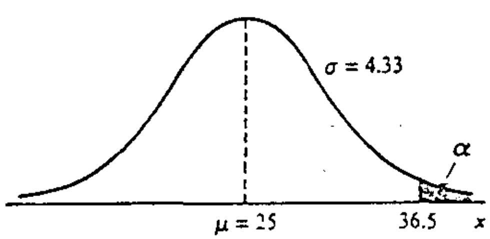


Fig. 7-1 Probabilidad de un error de tipo I.

En la figura 7-1

$$\alpha = P(\text{error tipo I})$$
$$= P(X > 36.5 \mid H_0 \text{ es verdadera}).$$

El valor z correspondiente a $x = 36.5$ es

$$z = \frac{36.5 - 25}{4.33} = 2.656.$$

Por lo tanto,

$$\alpha = P(Z > 2.656)$$
$$= 1 - P(Z < 2.656)$$
$$= 1 - 0.9961$$
$$= 0.0039.$$

Si H_0 es falsa y el verdadero valor de H_1 es $p = 1/2$, se puede determinar la probabilidad de un error tipo II empleando la aproximación a la curva normal con

$$\mu = np = (100)(\tfrac{1}{2}) = 50$$

y

$$\sigma = \sqrt{npq} = \sqrt{(100)(\tfrac{1}{2})(\tfrac{1}{2})} = 5.$$

La probabilidad de caer en la región de aceptación cuando H_1 es verdadera está dada por el área de la región sombreada de la figura 7-2. De donde

$$\beta = P(\text{error tipo II})$$
$$= P(X < 36.5 \mid H_1 \text{ es verdadera}$$

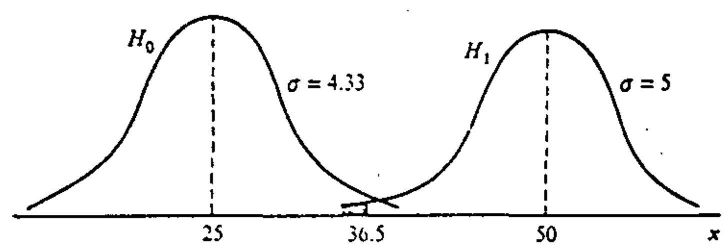


Fig. 7-2 Probabilidad de un error de tipo II.

El valor z correspondiente a $x = 36.5$ es

$$z = \frac{36.5 - 50}{5} = -2.7.$$

Por lo tanto,

$$\begin{aligned}\beta &= P(Z < -2.7)\\ &= 0.0035.\end{aligned}$$

Obviamente, los errores tipo I y tipo II ocurrirán rara vez si el experimento se hace con 100 individuos.

Los conceptos expuestos pueden verse en forma gráfica cuando la población es continua. Considérese la hipótesis nula de que el peso promedio de los estudiantes varones de cierta escuela es de 68 kilogramos contra la hipótesis alterna de que no es igual a 68. Esto es, se desea probar

$$\begin{aligned}H_0&:\quad \mu = 68.\\ H_1&:\quad \mu \neq 68.\end{aligned}$$

La hipótesis alterna permite la posibilidad de que $\mu < 68$ ó $\mu > 68$.

Supóngase que la desviación estándar de la población de pesos es $\sigma = 3.6$. La estadística de decisión, basada en una muestra de tamaño $n = 36$, será $\bar{X}$, el estimador más eficiente de μ. En el capítulo 5 se dijo que la distribución muestral de $\bar{X}$ es una distribución aproximadamente normal con desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 3.6/6 = 0.6$.

Una media de la muestra que caiga cerca del valor hipotético de 68 será considerada como evidencia en favor de H_0. Por otra parte, una media de la muestra que sea considerablemente menor o mayor que 68; será una evidencia contraria a H_0 y, por lo tanto, acorde con H_1. Se escoge arbitrariamente una región crítica, indicada por el área sombreada de la figura 7.3, que es $\bar{X} < 67$ y $\bar{X} > 69$. Por lo tanto, la región de aceptación será $67 < \bar{X} < 69$. De ahí que, si la media de la muestra $\bar{x}$ cae dentro de la región crítica, H_0 se rechaza; en caso contrario H_0 se acepta.

La probabilidad de cometer un error tipo I, o sea el nivel de significancia de la prueba, es igual a la suma de las áreas que se han sombreado en cada extremo de la distribución de la figura 7-3.

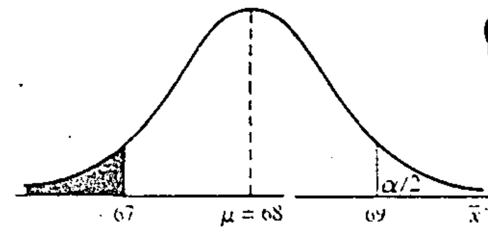


Fig. 7-3 Región crítica para probar $\mu = 68$ frente a $\mu \neq 68$.

Por lo tanto,

$$\alpha = P(\bar{X} < 67 \mid H_0 \text{ es verdadera}) + P(\bar{X} > 69 \mid \text{si } H_0 \text{ es verdadera}$$

Los valores de z correspondientes a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando H_0 es verdadera, son

$$z_1 = \frac{67 - 68}{0.6} = -1.67$$

$$z_2 = \frac{69 - 68}{0.6} = 1.67.$$

Por lo tanto,

$$\begin{aligned}\alpha &= P(Z < -1.67) + P(Z > 1.67)\\ &= 2P(Z < -1.67)\\ &= 0.0950.\end{aligned}$$

Así, el 9.5% de todas las muestras de tamaño 36 conducirá al rechazo de $\mu = 68$ kilogramos cuando es verdadera. Para reducir α, se puede aumentar el tamaño de la muestra o ampliar la región de aceptación. Supóngase que el tamaño de la muestra se aumenta a $n = 64$. Entonces $\sigma_{\bar{X}} = 3.6/8 = 0.45$. Ahora

$$z_1 = \frac{67 - 68}{0.45} = -2.22$$

$$z_2 = \frac{69 - 68}{0.45} = 2.22.$$

De donde

$$\begin{aligned}\alpha &= P(Z < -2.22) + P(Z > 2.22)\\ &= 2P(Z < -2.22)\\ &= 0.0264.\end{aligned}$$

La reducción en α no basta para garantizar un buen procedimiento de prueba. Se debe evaluar β para diferentes hipótesis alternas que pensamos que debieran aceptarse si son verdaderas. Por lo tanto, si es importante rechazar a H_0 cuando la media verdadera es algún valor $\mu \geq 70$ ó $\mu \leq 66$, la probabilidad de cometer un error tipo II debe calcularse y examinarse para las alternativas $\mu = 66$ y $\mu = 70$. Por razones de simetría, sólo es necesario considerar la probabilidad de aceptar la hipótesis nula de que $\mu = 68$ cuando la alternativa $\mu = 70$ es verdadera. Un error tipo II resultará cuando la media de la muestra $\bar{x}$ caiga entre 67 y 69 cuando H_1 es verdadera. Por lo tanto, en la figura 7-4,

$$\beta = P(67 < \bar{X} < 69 \mid \text{si } H_1 \text{ es verdadera}).$$

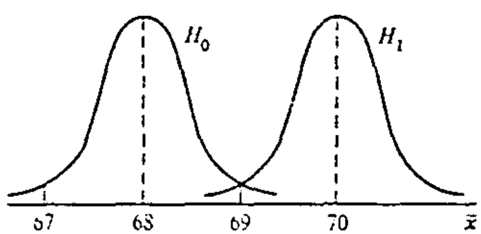


Fig. 7-4 Error de tipo II para probar $\mu = 68$ frente a $\mu = 70$.

Los valores z correspondientes a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando H_1 es verdadera, son

$$z_1 = \frac{67 - 70}{0.45} = -6.67$$

$$z_2 = \frac{69 - 70}{0.45} = -2.22.$$

Por lo tanto,

$$\begin{aligned}\beta &= P(-6.67 < Z < -2.22)\\ &= P(Z < -2.22) - P(Z < -6.67)\\ &= 0.0132 - 0.0000\\ &= 0.0132.\end{aligned}$$

Si el verdadero valor de μ es la alternativa $\mu = 66$, el valor β será nuevamente 0.0132. Para todos los valores posibles de $\mu < 66$ ó $\mu > 70$, el valor de β será menor cuando $n = 64$, y en consecuencia habrá poca probabilidad de aceptar H_0 cuando sea falsa.

La probabilidad de cometer un error tipo II aumenta rápidamente cuando el verdadero valor de μ se aproxima, pero sin ser igual, al valor hipotético. Por supuesto, esto sucede generalmente cuando no importa incurrir en él. Por ejemplo, si la hipótesis alterna $\mu = 68.5$ es verdadera, no importa cometer un error tipo II al concluir que la verdadera respuesta es $\mu = 68$. La probabilidad de hacer ese error será alta cuando $n = 64$. En la figura 7-5, se tiene

$$\beta = P(67 < \bar{X} < 69 \mid H_1 \text{ es verdadera}).$$

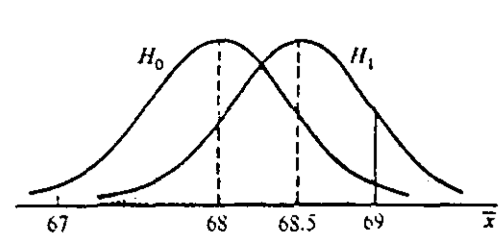


Fig. 7-5 Error de tipo II para probar $\mu = 68$ frente a $\mu = 68.5$.

Los valores z correspondientes a $\bar{x}_1 = 67$ y $\bar{x}_2 = 69$ cuando $\mu = 68.5$, son

$$z_1 = \frac{67 - 68.5}{0.45} = -3.33$$

$$z_2 = \frac{69 - 68.5}{0.45} = 1.11.$$

Por lo tanto,

$$\begin{aligned}\beta &= P(-3.33 < Z < 1.11)\\ &= P(Z < 1.11) - P(Z < -3.33)\\ &= 0.8665 - 0.0004\\ &= 0.8661.\end{aligned}$$

Los ejemplos anteriores ilustran las propiedades siguientes:

1. El error tipo I y el error tipo II están relacionados. Una disminución en la probabilidad de uno resulta generalmente en un aumento de la probabilidad del otro.
2. El tamaño de la región crítica, y por lo tanto la probabilidad de cometer un error tipo I, puede reducirse siempre ajustando el valor o los valores críticos.
3. Un aumento del tamaño de la muestra n reducirá α y β simultáneamente.
4. Si la hipótesis nula es falsa, β es un máximo cuando el verdadero valor de un parámetro se acerca al valor supuesto. Cuando mayor sea la distancia entre el valor verdadero y el valor supuesto, menor será β

7.3 PRUEBAS DE UNA COLA Y DE DOS COLAS

Una prueba de cualquier hipótesis estadística en que la alternativa sea *unilateral* o de *una cola* tal como

$$H_0:\quad \theta = \theta_0,$$
$$H_1:\quad \theta > \theta_0,$$

o quizás

$$H_0:\quad \theta = \theta_0,$$
$$H_1:\quad \theta < \theta_0,$$

se llama *prueba de una cola o unilateral*. La región crítica para la hipótesis alterna $\theta > \theta_0$ se encuentra totalmente en el extremo derecho de la distribución, mientras que la región crítica para la hipótesis alterna $\theta < \theta_0$ se encuentra totalmente en el extremo izquierdo. Por ejemplo, el director del programa de pruebas utilizó una prueba unilateral para probar la hipótesis $p = 1/4$ contra la alternativa de un lado $p > 1/4$ para la distribución binomial.

Una prueba de cualquier hipótesis estadística en que la alternativa es *bilateral* o de *dos colas* tal como

$$H_0:\quad \theta = \theta_0,$$
$$H_1:\quad \theta \neq \theta_0,$$

recibe el nombre de *prueba de dos colas o bilateral*. La hipótesis alterna establece que $\theta < \theta_0$ y que $\theta > \theta_0$. Los valores en ambos extremos de la distribución constituyen la región crítica. Se utilizó una prueba de dos colas para probar la hipótesis nula de que $\mu = 68$ kilogramos contra la alternativa de dos colas $\mu \neq 68$ para la población continua de los pesos de los estudiantes varones.

El formular una hipótesis alterna unilateral o bilateral dependerá de la conclusión que se saque si H_0 es rechazada. La posición de la región crítica puede determinarse sólo después de establecer H_1. Por ejemplo, al probar un

nuevo medicamento, se establece la hipótesis alterna de que el medicamento no es mejor que los medicamentos similares con que se cuenta y se prueba contra la hipótesis alterna de que es superior. Tal hipótesis alterna resultará en una prueba unilateral con la región crítica en el extremo derecho. Sin embargo, si se desea determinar que los dos procedimientos de enseñanza tienen igual eficacia, la hipótesis alterna permitirá que cualquier procedimiento sea superior. Por lo tanto, la prueba es dos extremos (colas) con la región crítica dividida de modo que se encuentre en los extremos izquierdo y derecho de la distribución.

Al probar las hipótesis relativas a poblaciones discretas, la región crítica se escoge arbitrariamente y se determina su tamaño. Si el tamaño α es demasiado grande, puede reducirse haciendo un ajuste en el valor crítico. Quizá haya que aumentarlo para compensar el incremento automático que ocurre en β. Al probar las hipótesis relativas a poblaciones continuas, se acostumbra escoger valores de α de 0.05 ó 0.01 y después encontrar el valor o los valores críticos. Por ejemplo, en una prueba de dos colas al nivel de significancia 0.05, los valores críticos de una estadística que tiene una distribución normal estándar serán $-z_{0.025} = -1.96$ y $z_{0.025} = 1.96$. En términos de los valores z, la región crítica de tamaño 0.05 será $Z < -1.96$ y $Z > 1.96$ y la región de aceptación será $-1.96 < Z < 1.96$. Se dice que una prueba es *significativa* si la hipótesis nula se rechaza al nivel de significancia 0.05, y se considera *altamente significativa* si la hipótesis nula se rechaza al nivel de significancia 0.01.

En las secciones restantes de este capítulo se expondrán varias pruebas de hipótesis especiales usadas frecuentemente por los estadígrafos.

7.4 PRUEBAS DE MEDIAS Y VARIANCIAS

Tomemos el problema de probar la hipótesis de que la media μ de una población, con variancia σ^2, conocida, sea igual a un valor especificado μ_0 contra la alternativa de dos colas de que la media no es igual a μ_0: Es decir, deberá probarse

$$H_0:\quad \mu = \mu_0,$$
$$H_1:\quad \mu \neq \mu_0.$$

Una estadística apropiada en la cual se base el criterio de decisión es la variable aleatoria $\bar{X}$. En el capítulo 5 se dijo que la distribución muestral de $\bar{X}$ es una distribución aproximadamente normal con media $\mu_{\bar{X}} = \mu$ y variancia $\sigma^2_{\bar{X}} = \sigma^2/n$, donde μ y σ^2 son la media y la variancia de la población de la cual se seleccionan las muestras aleatorias de tamaño n. Si se emplea un nivel de significancia de α, es posible encontrar dos valores críticos $\bar{x}_1$ y $\bar{x}_2$ tales que el intervalo $\bar{x}_1 < \bar{X} < \bar{x}_2$ defina la región de aceptación y las dos colas de la distribución, $\bar{X} < \bar{x}_1$ y $\bar{X} > \bar{x}_2$, constituyan la región crítica.

La región crítica puede estar dada en términos de los valores z por medio de la transformación

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Por tanto, para un nivel de significancia α en la figura 7-6 se muestra que los valores críticos de la variable aleatoria Z, correspondientes a $\bar{x}_1$ y $\bar{x}_2$, son

$$-z_{\alpha/2} = \frac{\bar{x}_1 - \mu_0}{\sigma/\sqrt{n}}$$

$$z_{\alpha/2} = \frac{\bar{x}_2 - \mu_0}{\sigma/\sqrt{n}}.$$

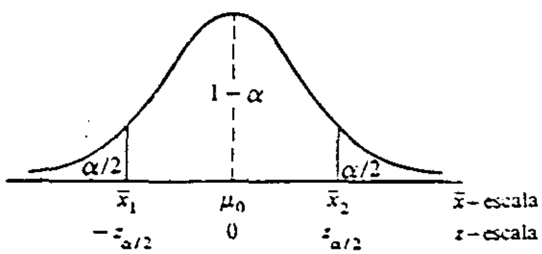


Fig. 7-6 Región crítica para la hipótesis alterna $\mu \neq \mu_0$.

De la población se extrae una muestra aleatoria de tamaño n y se calcula la media de la muestra $\bar{x}$. Si $\bar{x}$ cae en la región de aceptación, $\bar{x}_1 < \bar{X} < \bar{x}_2$, entonces

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

caerá la región $-z_{\alpha/2} < Z < z_{\alpha/2}$ y se concluye que $\mu = \mu_0$. De lo contrario, se rechaza H_0 y se acepta la hipótesis alterna de que $\mu \neq \mu_0$. Generalmente la región crítica se establece más bien en términos de Z que de $\bar{X}$.

El procedimiento de prueba descrito es equivalente a encontrar un intervalo de confianza al $(1 - \alpha)100\%$ para μ y aceptar H_0 si μ_0 está en el intervalo. Si está fuera del intervalo, se rechaza μ_0 en favor de la hipótesis alterna H_1. En

consecuencia, cuando se hacen inferencias sobre la media μ de una población con variancia σ^2, conocida, ya sea por la construcción de un intervalo de confianza o por la prueba de una hipótesis estadística, se emplea la misma estadística $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$.

En general, si se usa una estadística para construir un intervalo de confianza destinado al parámetro θ, como Z, T, X^2, e F, se puede aplicar la misma estadística para probar la hipótesis de que el parámetro es igual a determinado valor θ_0 contra la alternativa apropiada. Naturalmente todas las suposiciones fundamentales hechas en el capítulo 6, relativas al uso de una estadística dada, se aplicarán a las pruebas descritas aquí. Esto significa esencialmente que todas las muestras se toman de poblaciones aproximadamente normales. Sin embargo, una estadística Z puede emplearse para probar hipótesis acerca de medias tomadas de poblaciones no normales cuando $n \geq 30$.

En la tabla 7-1 se enumeran las estadísticas con que se prueba una hipótesis H_0 determinada y se dan las regiones críticas adecuadas para hipótesis alternas H_1 de una y dos colas. Los pasos para probar una hipótesis relativa al parámetro θ de una población contra una hipótesis alterna pueden resumirse como sigue:

1. H_0: $\theta = \theta_0$.
2. H_1: Las alternativas son $\theta < \theta_0$, $\theta > \theta_0$, ó $\theta \neq \theta_0$.
3. Escoja un nivel de significancia igual a α.
4. Seleccione la estadística de prueba apropiada y establezca la región crítica n.
5. Calcule el valor de la estadística partiendo de una muestra aleatoria de tamaño n.
6. Conclusión: Rechace H_0 si la estadística tiene un valor en la región crítica; si no es así, acepte H_0.

Ejemplo 7.1 Un fabricante de artículos deportivos ha inventado una nueva cuerda sintética para pescar, que tiene una resistencia media de 8 kg a la ruptura y una desviación estándar de 0.5 kilogramos. Pruebe la hipótesis de que $\mu = 8$ kg contra la alternativa de que $\mu \neq 8$ kg si se toma una muestra aleatoria de 50 cuerdas y se encuentra que tienen una resistencia media a la ruptura de 7.8 kilogramos. Emplee un nivel de significancia de 0.01.

Solución

1. H_0: $\mu = 8$ kg
2. H_1: $\mu \neq 8$ kg
3. $\alpha = 0.01$.
4. Región crítica: $Z < -2.58$ y $Z > 2.58$, donde

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

Tabla 7-1. Pruebas de medias y variancias

H_0	*Estadística de prueba*	H_1	*Región crítica*
$\mu = \mu_0$	$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$; σ conocida	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$Z < -z_\alpha$ $Z > z_\alpha$ $Z < -z_{\alpha/2}$ o $Z > z_{\alpha/2}$
$\mu = \mu_0$	$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$; $v = n - 1$, σ desconocida	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$T < -t_\alpha$ $T > t_\alpha$ $T < -t_{\alpha/2}$ o $T > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$; σ_1 y σ_2 conocida	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$Z < -z_\alpha$ $Z > z_\alpha$ $Z < -z_{\alpha/2}$ o $Z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$T = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p\sqrt{(1/n_1) + (1/n_2)}}$; $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ pero desconocida $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$T < -t_\alpha$ $T > t_\alpha$ $T < -t_{\alpha/2}$; o $T > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$T' = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$; $v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$; $\sigma_1 \neq \sigma_2$ pero desconocida	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$T' < -t_\alpha$ $T' > t_\alpha$ $T' < -t_{\alpha/2}$ o $T' > t_{\alpha/2}$
$\mu_D = d_0$	$T = \frac{\bar{D} - d_0}{S_d/\sqrt{n}}$; $v = n - 1$, observaciones pareadas	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$T < -t_\alpha$ $T > t_\alpha$ $T < -t_{\alpha/2}$ o $T > t_{\alpha/2}$
$\sigma^2 = \sigma_0^2$	$X^2 = \frac{(n - 1)S^2}{\sigma_0^2}$; $v = n - 1$	$\sigma^2 < \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$X^2 < \chi^2_{1-\alpha}$ $X^2 > \chi^2_\alpha$ $X^2 < \chi^2_{1-\alpha/2}$ o; $X^2 > \chi^2_{\alpha/2}$
$\sigma_1^2 = \sigma_2^2$	$F = S_1^2/S_2^2$; $v_1 = n_1 - 1$ y $v_2 = n_2 - 1$	$\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$F < f_{1-\alpha}(v_1, v_2)$ $F > f_\alpha(v_1, v_2)$ $F < f_{1-\alpha/2}(v_1, v_2)$ y $F > f_{\alpha/2}(v_1, v_2)$

5. Cálculos:

$$\bar{x} = 7.8 \text{ kilogramos} \qquad n = 50$$

$$z = \frac{7.8 - 8}{0.5/\sqrt{50}} = -2.828.$$

6. Conclusión: Se rechaza H_0 y se concluye que la resistencia a la ruptura promedio no es igual a 8 kg sino menor.

Ejemplo 7.2 El tiempo promedio empleado para escribir a los alumnos de una escuela ha sido de 50 minutos con una desviación estándar de 10 minutos. Se está ensayando un nuevo procedimiento de inscripción mediante máquinas computadoras. Si una muestra aleatoria de 12 estudiantes obtiene un tiempo promedio en inscribirse de 42 minutos con una desviación estándar de 11.9 minutos con el nuevo sistema, pruebe la hipótesis de que la media de la población es ahora menor de 50, usando un nivel de significancia de 1) 0.05 y 2) 0.01. Suponga que la población de los tiempos es normal.

Solución

1. H_0: $\mu = 50$ minutos
2. H_1: $\mu < 50$ minutos
3. (1) $\alpha = 0.05$, (2) $\alpha = 0.01$.
4. Región crítica: (1) $T < -1.796$, (2) $T < -2.718$, donde $T = (\bar{X} - \mu_0)/(S/\sqrt{n})$ con $v = 11$ grados de libertad.
5. Cómputos: $\bar{x} = 42$ minutos, $s = 11.9$ minutos y $n = 12$. De donde

$$t = \frac{42 - 50}{11.9/\sqrt{12}} = -2.33.$$

6. Conclusión: Se rechaza H_0 al nivel de significancia 0.05 pero no al nivel 0.01. Esto significa que la media verdadera realmente es menor de 50 minutos pero no lo suficiente para garantizar el alto costo que supone el uso de una computadora.

Ejemplo 7.3 Se efectuó un experimento para comparar el desgaste por abrasión en dos materiales laminados diferentes. Con una máquina para medir el desgaste se probaron 12 piezas del material 1. Diez piezas del material 2 fueron probadas de modo similar. En ambos casos se observó el espesor del desgaste. Las muestras del material 1 dieron un desgaste promedio (codificado) de 85 unidades con una desviación estándar de 4, mientras que las del material 2 dieron un promedio de 81 y una desviación estándar de 5. Pruebe la hipótesis de que los dos tipos de materiales presentan el mismo desgaste al nivel de significancia 0.01. Suponga que las poblaciones son aproximadamente normales con variancias iguales.

Solución Las medias de las poblaciones de los materiales 1 y 2 se representarán, respectivamente, por μ_1 y μ_2 Usando el procedimiento de seis pesos, se tiene

1. H_0: $\mu_1 = \mu_2$ ó $\mu_1 - \mu_2 = 0$.
2. H_1: $\mu_1 \neq \mu_2$ ó $\mu_1 - \mu_2 \neq 0$.
3. $\alpha = 0.10$.
4. Región crítica: $T < -1.725$ y $T > 1.725$, donde

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p\sqrt{(1/n_1) + (1/n_2)}}$$

   con $v = 20$ grados de libertad.

5. Cómputos: $\bar{x}_1 = 85$, $s_1 = 4$, $n_1 = 12$, y $\bar{x}_2 = 81$, $s_2 = 5$, $n_2 = 10$. De donde

$$s_p = \sqrt{\frac{(11)(16) + (9)(25)}{12 + 10 - 2}} = 4.478$$

$$t = \frac{(85 - 81) - 0}{4.478\sqrt{(1/12) + (1/10)}} = 2.07.$$

6. Conclusión: Se rechaza H_0 y se concluye que los dos materiales no representan el mismo desgaste.

Ejemplo 7.4 Para averiguar si existe una diferencia entre un análisis químico y un análisis con rayos X que permita conocer la cantidad de hierro en un material, se emplean cinco muestras de una substancia ferrosa. Cada muestra se parte en dos submuestras y se les aplica ambos dos tipos de análisis. Los datos codificados que indican el análisis de contenido de hierro son los siguientes:

	Muestra				
Análisis	*1*	*2*	*3*	*4*	*5*
Rayos X	2.0	2.0	2.3	2.1	2.4
Químico	2.2	1.9	2.5	2.3	2.4

Suponiendo que la población es normal, pruebe al nivel de significancia 0.05 si los dos métodos de análisis dan, en promedio, el mismo resultado.

Solución Los contenidos promedios de hierro determinados por los análisis químico y de rayos X se representarán respectivamente por μ_1 y μ_2. Se procede como sigue:

1. H_0: $\mu_1 = \mu_2$ ó $\mu_D = 0$.
2. H_1: $\mu_1 \neq \mu_2$ ó $\mu_D \neq 0$.
3. $\alpha = 0.05$.
4. Región crítica: $T < -2.776$ y $T > 2.776$, donde $T = (\bar{D} - d_0)/(S_d/\sqrt{n})$ con $v = 4$ grados de libertad.
5. Cómputos:

Rayos X	*Químico*	d_i	d_i^2
2.0	2.2	−0.2	0.04
2.0	1.9	0.1	0.01
2.3	2.5	−0.2	0.04
2.1	2.3	−0.2	0.04
2.4	2.4	0.0	0.00
		−0.5	0.13

Se encuentra que $\bar{d} = -0.5/5 = -0.1$ y $s_d^2 = [(5)(0.13) - (-0.5)^2]/(5)(4) = 0.02$. Tomando la raíz cuadrada, se tiene $s_d = 0.14142$. De donde

$$t = \frac{-0.1 - 0}{0.14142/\sqrt{5}} = -1.6.$$

6. Conclusión: Se acepta H_0 y se concluye que los métodos de análisis no son significativamente diferentes.

Ejemplo 7.5 Un fabricante de acumuladores para automóvil asegura que la vida de su producto tiene una distribución aproximadamente normal, con una desviación estándar igual a 0.9 años. Si una muestra aleatoria de 10 acumuladores tiene una desviación estándar de 1.2 años, ¿pensaría usted que $\sigma > 0.9$ años? Utilice un nivel de significancia de 0.05.

Solución

1. H_0: $\sigma^2 = 0.81$.
2. H_1: $\sigma^2 > 0.81$.
3. $\alpha = 0.05$.
4. Región crítica: $X^2 > 16.919$, donde $X^2 = (n-1)S^2/\sigma_0^2$ con $v = 9$ grados de libertad.

5. Cómputos: $s^2 = 1.44$, $n = 10$, y

$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16.0.$$

6. Conclusión: Se acepta H_0 y se concluye que no hay razón para dudar que la desviación estándar sea 0.9 años.

7.5 SELECCION DEL TAMAÑO DE LA MUESTRA PARA LA PRUEBA DE MEDIAS

El nivel de significancia para la prueba de hipótesis estadísticas, normalmente lo controla el investigador, mientras que β, o sea la *potencia* de la prueba, definida por $1 - \beta$, está controlada empleando un tamaño adecuado de la muestra. En esta sección se explicará, la selección del tamaño de la muestra para pruebas que contienen una o dos medias de población. Cuando se usa la distribución normal y se conocen la variancia o variancias de la población, es fácil delimitar el tamaño necesario de la muestra para alcanzar la potencia deseada.

Supóngase que se desea probar la hipótesis

$$H_0:\quad \mu = \mu_0,$$
$$H_1:\quad \mu > \mu_0,$$

con un nivel de significancia α cuando la variancia σ^2 es conocida. Para una alternativa específica, digamos $\mu = \mu_0 + \delta$, la potencia de la prueba está dada por

$$1 - \beta = P\left[\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \mid H_1 \text{ es verdadera}\right]$$
$$= P\left[\frac{\bar{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}} > z_\alpha - \frac{\delta}{\sigma/\sqrt{n}} \mid \mu = \mu_0 + \delta\right].$$

Según la hipótesis alterna $\mu = \mu_0 + \delta$, la estadística

$$\frac{\bar{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}}$$

es una variable normal estándar. Por lo tanto,

de lo que se concluye que

$$-z_\beta = z_\alpha - \frac{\delta\sqrt{n}}{\sigma}$$

y por tanto

$$n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{\delta^2},$$

resultando que tambiém es verdadero cuando la hipótesis alterna es $\mu < \mu_0$.

En el caso de una prueba de dos colas se obtiene la potencia $1 - \beta$ para determinada alternativa cuando

$$n \simeq \frac{(z_{\alpha/2} + z_\beta)^2\sigma^2}{\delta^2}.$$

Ejemplo 7.6 Supóngase que se desea probar la hipótesis

$$H_0:\quad \mu = 68,$$
$$H_1:\quad \mu > 68,$$

con un nivel de significancia $\alpha = 0.05$ cuando $\sigma = 5$. Encuentre el tamaño requerido de la muestra si la potencia de la prueba es de 0.95 cuando la media verdadera es 69.

Solución Como $\alpha = \beta = 0.05$, se tiene $z_\alpha = z_\beta = 1.645$. Para alternativa: $\mu = 69$, se toma $\delta = 1$ y después

$$n = \frac{(1.645 + 1.645)^2(25)}{1} = 270.6.$$

Por lo tanto, se necesitan 271 observaciones si la prueba se realiza para rechazar la hipótesis nula en 95% de las veces cuando de hecho el valor de μ es 69.

Se puede aplicar un procedimiento similar para determinar el tamaño de la muestra $n = n_1 = n_2$ que requiere una potencia de la prueba en la cual se comparan las medias de dos poblaciones. Por ejemplo, supóngase que se desea probar la hipótesis cuando σ_2 y σ_1 son conocidas. Para una alternativa específica, digamos $x\,\mu_1 - \mu_2 = \delta$, la potencia de la prueba está dada por

$$1 - \beta = P\left[\frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} > z_{\alpha/2} \,|\, \mu_1 - \mu_2 = \delta\right].$$

Por lo tanto,

$$\begin{aligned}\beta &= P\left[-z_{\alpha/2} < \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < z_{\alpha/2} \,|\, \mu_1 - \mu_2 = \delta\right]\\ &= P\left[-z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < \frac{\bar{X}_1 - \bar{X}_2 - \delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < z_{\alpha/2}\right.\\ &\qquad\qquad \left. - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} \,|\, \mu_1 - \mu_2 = \delta\right].\end{aligned}$$

Según la hipótesis alterna, $\mu_1 - \mu_2 = \delta$, la estadística

$$\frac{\bar{X}_1 - \bar{X}_2 - \delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

es una variable normal estándar. Por lo tanto,

$$\beta = P\left[-z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}} < Z < z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}\right],$$

de la cual se concluye que

$$-z_\beta \simeq z_{\alpha/2} - \frac{\delta}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

y por tanto

$$n \simeq \frac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{\delta^2}.$$

Para la prueba de una cola, la expresión del tamaño requerido de la muestra cuando $n = n_1 = n_2$ la hallamos en

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{\delta^2}.$$

Cuando se ignora la variancia de la población (o variancias en las situaciones de dos muestras), la selección del tamaño de la muestra no es directa. Al probar la hipótesis $\mu = \mu_0$ si el verdadero valor es $\mu = \mu_0 + \delta$, la estadística

$$\frac{\bar{X} - (\mu_0 + \delta)}{S/\sqrt{n}}$$

no sigue la distribución t, como se cabría esperar sino la *distribución t no central.* Sin embargo, hay tablas o gráficas basadas en esta última para determinar el tamaño apropiado de la muestra si se dispone de alguna estimación de σ o si δ es un múltiplo de σ. La tabla IX (cf. Tablas estadísticas) proporciona los tamaños de la muestra necesarios para controlar los valores de α y β en diferentes valores de

$$\Delta = \frac{|\delta|}{\sigma} = \frac{|\mu - \mu_0|}{\sigma}$$

en las pruebas de una y dos colas. En el caso de la prueba t de dos muestras, en la cual se desconocen las variancias pero se suponen que son iguales, se obtienen los tamaños de las muestras $n = n_1 = n_2$ que se requieren para controlar los valores de α y β en diferentes valores de

$$\Delta = \frac{|\delta|}{\sigma} = \frac{|\mu_1 - \mu_2|}{\sigma}$$

contenidos en la tabla X.

Ejemplo 7.7 Al comparar el rendimiento de dos catalizadores en el resultado de una reacción química, se efectúa una prueba t de dos muestras con $\alpha = 0.05$. Se considera que las variancias de los resultados son iguales para los dos catalizadores. ¿De qué tamaño deben ser las muestras de cada catalizador para probar la hipótesis

$$\begin{aligned}H_0&: \mu_1 = \mu_2,\\ H_1&: \mu_1 \neq \mu_2,\end{aligned}$$

si es esencial detectar una diferencia de 0.8σ entre los catalizadores, con una probabilidad de 0.9?

Solución En la tabla X, con $\alpha = 0.05$ para una prueba de dos colas, $\beta = 0.1$, y

$$\Delta = \frac{|0.8\sigma|}{\sigma} = 0.8,$$

se encuentra que el tamaño requerido de la muestra es $n = 34$.

7.6 PRUEBAS DE PROPORCIONES

Las pruebas de hipótesis relativas a proporciones son útiles muchos campos. Los políticos quieren saber qué fracción de los votantes los apoyarán en las próximas elecciones. Todos los fabricantes desean conocer la proporción de piezas defectuosas contenidas en un embarque. El jugador se basa en un conocimiento de la proporción de resultados que estime favorables.

Estudiaremos el problema de probar la hipótesis de que la proporción de éxitos en un experimento binomial sea igual a determinado valor. Es decir, se va a probar la hipótesis nula H_0 de que $p = p_0$, donde p es el parámetro de la distribución binomial. La hipótesis alterna puede ser una de las alternativas usuales unilaterales o bilaterales: $p < p_0$, $p > p_0$, ó $p \neq p_0$.

La estadística apropiada en la cual se basa el criterio de decisión es la variable aleatoria binomial X, aunque se puede usar en la misma forma la estadística $\hat{P} = X/n$. Los valores de X que están alejados de la media $\mu = np_0$ conducirán al rechazo de la hipótesis nula. Para probar la hipótesis

$$H_0:\quad p = p_0,$$
$$H_1:\quad p < p_0,$$

se recurre a la distribución binomial con $p = p_0$ y $q = 1 - p_0$ para determinar $P(X \leq x \mid H_0$ es verdadera). El valor x es el número de éxitos en la muestra de tamaño n. Si $P(X \leq x \mid H_0$ es verdadera) $< \alpha$, la prueba es significativa al nivel α y H_0 se rechaza en favor de H_1. Similarmente, para probar la hipótesis

$$H_0:\quad p = p_0,$$
$$H_1:\quad p > p_0,$$

se encuentra $P(X \geq x \mid H_0$ es verdadera) y se rechaza H_1 en favor de H_0 si esta probabilidad es menor que α. Finalmente, para probar la hipótesis

$$H_0:\quad p = p_0,$$
$$H_1:\quad p \neq p_0,$$

al nivel de significancia α, H_0 se rechaza cuando $x < np_0$ y $P(X \leq x \mid H_0$ es verdadera) $< \alpha/2$ o cuando $x > np_0$ y $P(X \geq x \mid H_0$ es verdadera) $< \alpha/2$.

Los pasos para probar una hipótesis acerca de una proporción contra las diferentes alternativas se resumen a continuación:

1. H_0: $p = p_0$.
2. H_1: Las alternativas son $p < p_0$, $p > p_0$, ó $p \neq p_0$.
3. Se escoge un nivel de significancia igual a α.
4. Región crítica:
   a) Todos los valores x tales que $P(X \leq x \mid H_0$ es verdadera) $< \alpha$ para la alternativa $p < p_0$.
   b) Todos los valores x tales que $P(X \geq x \mid H_0$ es verdadera) $< \alpha$ para la alternativa $p > p_0$.
   c) La unión de todos los valores x tales que $P(X \leq x \mid H_0$ es verdadera) $< \alpha/2$ cuando $x < np_0$, con todos los valores x tales que $P(X \geq x \mid H_0$ es verdadera) $< \alpha/2$ cuando $x > np_0$, para la alternativa $p \neq p_0$.
5. Cómputos: Encuentre x y calcule la probabilidad adecuada.
6. Conclusión: Rechace H_0 si x cae en la región crítica; si no es así acepte H_0.

Ejemplo 7.8 Un cazador afirma que derriba el 80% de los pájaros a los que dispara. ¿Acepta usted esa afirmación si, en un día dado, él derriba 9 de las 15 aves a las que disparó? Utilice un nivel de significación de 0.05.

Solución Se sigue el procedimiento de seis pasos:

1. H_0: $p = 0.8$.
2. H_1: $p \neq 0.8$.
3. $\alpha = 0.05$.
4. Región crítica: Todos los valores de x tales que $P(X \leq x \mid H_0$ es verdadera) < 0.025.
5. Cómputos: Se tiene $x = 9$ y $n = 15$. Por lo tanto, usando la tabla II,

$$P(X \leq 9 \mid p = 0.8) = \sum_{x=0}^{9} b(x; 15, 0.8)$$
$$= 0.0611 > 0.025.$$

6. Conclusión: acepte H_0 y llegará a la conclusión de que no hay razón para dudar de la afirmación del cazador.

En la sección 4.3 se vio que, cuando n es pequeño, las probabilidades binomiales se pueden obtener de la formula binomial real o de la tabla II. Para n, grandes, se requieren procedimientos de aproximación. Cuando el valor supuesto p_0 está muy cercano a 0 ó 1, se puede emplear la distribución de Poisson con parámetro $\mu = np_0$. Si n es grande, generalmente se prefiere la aproximación a la curva normal, que es bastante exacta en la media que p_0 no esté extremadamente cercano a 0 ni a 1. Al emplear la aproximación normal, el criterio de decisión se basa en la variable aleatoria

$$Z = \frac{\hat{P} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{X - np_0}{\sqrt{np_0 q_0}}.$$

Por tanto, en una prueba de dos colas al nivel de significancia α, la región crítica es $Z < -z_{\alpha/2}$ y $Z > z_{\alpha/2}$. Para la alternativa unilateral $p < p_0$, la región crítica es $Z < -z_\alpha$, y para la alternativa $p > p_0$, la región crítica es $Z > z_\alpha$.

Para probar una hipótesis acerca de una proporción, se aplica la aproximación a la curva normal y se procede como sigue:

1. H_0: $p = p_0$.
2. H_1: Las alternativas son $p < p_0$, $p > p_0$, ó $p \neq p_0$.
3. Escoja un nivel de significancia igual a α.
4. Región crítica:
   a) $Z < -z_\alpha$ para la alternativa $p < p_0$.
   b) $Z > z_\alpha$ para la alternativa $p > p_0$.
   c) $Z < -z_{\alpha/2}$ y $Z > z_{\alpha/2}$ para la alternativa $p \neq p_0$.
5. Cómputos: Encuentre x a partir del tamaño de la muestra n, y después calcule

$$z = \frac{x - np_0}{\sqrt{np_0 q_0}}.$$

6. Rechace H_0 si z cae en la región crítica; de lo contrario, acéptela H_0.

Ejemplo 7.9 Una compañía manufacturera ha encontrado que el 90% de los artículos producidos mediante cierto proceso no están defectuosos. Se está considerando la conveniencia de introducir una modificación en él se cree disminuirá la proporción de artículos defectuosos a menos del 10% actual. En un experimento de 100 artículos producidos con el nuevo proceso se encuentran 5 defectuosos. ¿Es esta evidencia suficiente para afirmar que el procedimiento modificado es mejor? Use un nivel de significancia de 0.05.

Solución Siguiendo el procedimiento de seis pasos:

1. H_0: $p = 0.9$.
2. H_1: $p > 0.9$.
3. $\alpha = 0.05$.
4. Región crítica: $Z > 1.645$.

5. Cómputos: $x = 95$, $n = 100$, $np_0 = (100)(0.95) = 95$, y

$$z = \frac{95 - 90}{\sqrt{(100)(0.90)(0.10)}} = 1.67.$$

6. Conclusión: Se rechaza H_0 y se concluye que el proceso modificado ha reducido la proporción de artículos defectuosos.

7.7 PRUEBA DE LA DIFERENCIA ENTRE DOS PROPORCIONES

Frecuentemente surgen situaciones donde se desea probar la hipótesis de que dos proporciones son iguales. Por ejemplo, quizás se trate de probar que la proporción de pediatras en una ciudad es igual a la que hay en otra ciudad. Una persona puede decidirse a dejar de fumar sólo si se le convence de que la proporción de fumadores que contraen cáncer pulmonar excede a la de los no fumadores.

En general, se desea probar la hipótesis nula

$$H_0: \quad p_1 = p_2 = p$$

contra alguna alternativa idónea. Los parámetros p_1 y p_2 son las proporciones de las dos poblaciones de lo que se investiga. La estadística en que se basa el criterio de decisión es la variable aleatoria $\hat{P}_1 - \hat{P}_2$. De dos poblaciones binomiales se extraen al azar muestras independientes de tamaños n_1 y n_2 y se calcula la proporción de éxitos $\hat{P}_1$ y $\hat{P}_2$ para las dos muestras. En la sección 6.6 se señaló que la estadística

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{(p_1 q_1/n_1) + (p_2 q_2/n_2)}} = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{pq[(1/n_1) + (1/n_2)]}}$$

tiene una distribución normal estándar cuando H_0 es verdadera y n_1 y n_2 son grandes. Para calcular un valor de Z, se debe estimar el parámetro p que aparece en el radical. Fusionamos los datos de ambas muestras y escribimos

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}.$$

donde x_1 y x_2 son el número de éxitos de las dos muestras. Substituyendo $\hat{p}$ por p, la estadística Z toma la forma

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{p}\hat{q}[(1/n_1) + (1/n_2)]}},$$

donde $\hat{q} = 1 - \hat{p}$. Las regiones críticas para la hipótesis alterna apropiada se establecen antes de usar los puntos críticos de la curva normal estándar.

Si se quiere probar la hipótesis de que dos proporciones son iguales cuando las muestras son grandes, se procede con los seis pasos siguientes:

1. H_0: $p_1 = p_2$.
2. H_1: Las alternativas son $p_1 < p_2$, $p_1 > p_2$, ó $p_1 \neq p_2$.
3. Seleccione un nivel de significancia igual a α.
4. Región crítica:
   a) $Z < -z_\alpha$ para la alternativa $p_1 < p_2$.
   b) $Z > z_\alpha$ para la alternativa $p_1 > p_2$.
   c) $Z < -z_{\alpha/2}$ y $Z > z_{\alpha/2}$ para la alternativa $p_1 \neq p_2$.
5. Cómputos: Calcule $\hat{p}_1 = x_1/n_1$, $\hat{p}_2 = x_2/n_2$, y $\hat{p} = (x_1 + x_2)/(n_1 + n_2)$ y después encuentre

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}[(1/n_1) + (1/n_2)]}}.$$

6. Conclusión: Rechace H_0 si z cae en la región crítica; de lo contrario, acéptela H_0.

Ejemplo 7.10 Para decidir si se construye una planta química propuesta, se hace una encuesta entre los residentes de una ciudad y del condado circundante. El sitio de la construcción está dentro de los límites de la ciudad y por eso muchos habitantes del condado piensan que una gran proporción de los residentes la apoyarán. Para averiguar si hay una diferencia significativa en la proporción de unos y otros se realiza una encuesta. Si 120 de 200 residentes de la ciudad y 240 de 500 habitantes del condado estan a favor de la construcción, ¿se puede considerar que la proporción de residentes de la ciudad que apoyan la construcción de la planta química es mayor que la de los habitantes del condado? Utilice un nivel de significancia de 0.025.

Solución Sean p_1 y p_2, respectivamente, las proporciones de habitantes de la ciudad y del condado que respaldan el proyecto. Siguiendo el procedimiento de seis pasos:

1. H_0: $p_1 = p_2$.
2. H_1: $p_1 > p_2$.

3. $\alpha = 0.025$.
4. Región crítica: $Z > 1.96$.
5. Cómputos:

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{120}{200} = 0.60$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{240}{500} = 0.48$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{120 + 240}{200 + 500} = 0.51.$$

Por lo tanto,

$$z = \frac{0.60 - 0.48}{\sqrt{(0.51)(0.49)[(1/200) + (1/500)]}} = 2.9.$$

6. Conclusión: Se rechaza H_0 y se concluye que la proporción de residentes de la ciudad que apoyan el proyecto es mayor que la de los habitantes del condado.

7.8 PRUEBA DE BONDAD DE AJUSTE

En este capítulo nos hemos ocupado de la prueba de hipótesis estadísticas acerca de parámetros individuales de la población, tales como μ, σ^2, y p. Ahora estudiaremos una prueba para determinar si una población tiene una distribución teórica especificada. La prueba se basa en el grado de ajuste que hay entre la frecuencia de ocurrencia de las observaciones en una muestra observada y las frecuencias esperadas que se obtienen de una distribución hipótetica.

El lanzamiento de un dado nos servirá de ejemplo. Supongamos que el dado está balanceado, lo cual equivale a probar la hipótesis de que la distribución de los resultados es uniforme. Suponga que el dado se tira 120 veces, anotando cada resultado. Teóricamente, si el dado está balanceado, se puede esperar que cada lado salga 20 veces. Los resultados se dan en la tabla 7-2. Al comparar las

Tabla 7-2. Frecuencias observadas y esperadas de 120 lanzamientos de un dado

	Caras					
	1	*2*	*3*	*4*	*5*	*6*
Observadas	20	22	17	18	19	24
Esperadas	20	20	20	20	20	20

frecuencias observadas con las frecuencias esperadas correspondientes, se debe decidir si estas discrepancias se deben probablemente a las fluctuaciones del muestreo (y por lo tanto el dado estará balanceado); de lo contrario el dado no lo está y la distribución de los resultados no es uniforme. Es una práctica común llamar celda a cada resultado posible de un experimento. Así pues, en este ejemplo se tienen seis celdas. La estadística apropiada en que se basa el criterio de decisión para un experimento con k celdas la define el siguiente

> TEOREMA 7.1 *La prueba de* bondad de ajuste *entre las frecuencias observadas y esperadas se basa en la cantidad*
>
> $$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$
>
> *donde* χ^2 *es un valor de la variable aleatoria* X^2 *cuya distribución muestral se aproxima bastante a la distribución ji cuadrada. Los símbolos* o_i *y* e_i *representan, respectivamente, las frecuencias observadas y las esperadas para la i-ésima celda.*

Si las frecuencias observadas están cercanas a sus correspondientes frecuencias esperadas el valor χ^2 será pequeño, lo cual indica un buen ajuste. Si observadas difieren considerablemente, el valor χ^2 será grande y el ajuste pobre. Un buen ajuste lleva a la aceptación de H_0, mientras que un ajuste, deficiente conduce a un rechazo. Por lo tanto, la región crítica caerá en el extremo derecho de la distribución ji cuadrada. Para un nivel de significancia igual a α, el valor crítico χ^2_α se encuentra de la tabla VI y en consecuencia $X^2 > \chi^2_\alpha$ constituye la región crítica. Este criterio de decisión no debe usarse si las frecuencias esperadas no son igual a 5, por lo menos.

El número de grados de libertad asociados con la distribución ji cuadrada aquí empleada depende de dos factores: el número de celdas en el experimento y el número de cantidades que se obtengan de los datos observados que se necesitan para calcular las frecuencias esperadas. A este número se llega por medio del siguiente

> TEOREMA 7.2 *El número de grados de libertad en una prueba de bondad de cantidades obtenidas, tomadas de los datos observados que se emplean en el cálculo de las frecuencias esperadas.*

La frecuencia total es la única cantidad proporcionada por los datos observados, al calcular las frecuencias esperadas para el resultado de lanzar un dado. Así pues, según la definición, el valor calculado χ^2 tiene $6 - 1 = 5$ grados de libertad.

En la tabla 7-2 se observa que el valor χ^2 es

$$\begin{aligned}\chi^2 &= \frac{(20-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(17-20)^2}{20} \\ &\quad + \frac{(18-20)^2}{20} + \frac{(19-20)^2}{20} + \frac{(24-20)^2}{20} \\ &= 1.7.\end{aligned}$$

Usando la tabla VI, se encuentra que $\chi^2_{0.05} = 11.070$ para $v = 5$ grados de libertad. Como 1.7 es menor que el valor crítico, H_0 no se rechaza y se concluye que la distribución es uniforme. En otras palabras, el dado está balanceado.

He aquí otro ejemplo, la hipótesis de que la distribución de frecuencia de la vida de los acumuladores, dada en la tabla 2-2, puede aproximarse por medio de la distribución normal. Las frecuencias esperadas para cada clase (celda), listadas en la tabla 7-3, se obtienen de una curva normal que tiene las mismas media y desviación estándar que la muestra que nos ocupa. De los datos de la tabla 2-1 se deduce que la muestra de 40 acumuladores tiene una media $\bar{x} = 3.4125$ y una desviación estándar $s = 0.703$. Estos valores se usarán para μ y σ en el cálculo de los valores z correspondientes a los límites de clase. Por ejemplo, el valor z correspondiente a los límites de la cuarta clase son

$$z_1 = \frac{2.95 - 3.4125}{0.703} = -0.658$$

$$z_2 = \frac{3.45 - 3.4125}{0.703} = 0.053.$$

De la tabla IV se deduce que el área entre $z_1 = -0.658$ y $z_2 = 0.053$ es

$$\begin{aligned}\text{área} &= P(-0.658 < Z < 0.053) \\ &= P(Z < 0.053) - P(Z < -0.658) \\ &= 0.5211 - 0.2552 \\ &= 0.2659.\end{aligned}$$

Tabla 7-3. Frecuencias observadas y esperadas de la vida de los acumuladores suponiendo normalidad

Limites de clase	o_i	e_i
1.45-1.95	2	0.6
1.95-2.45	1 (7)	2.7 (10.1)
2.45-2.95	4	6.8
2.95-3.45	15	10.6
3.45-3.95	10	10.3
3.95-4.45	5 (8)	6.1 (8.3)
4.45-4.95	3	2.2

Por tanto, la frecuencia esperada para la cuarta clase es

$$e_4 = (0.2659)(40) = 10.6.$$

La frecuencia esperada para el primer intervalo de clase se obtiene usando el área total bajo la curva normal hacia la izquierda del límite 1.95. Para el último intervalo de clase, se emplea el área total hacia la derecha del límite inferior del intervalo, el cual es de 4.45. Todas las demás frecuencias esperadas se determinan por el método descrito para la cuarta clase. Nótese que cuando las frecuencias esperadas son menores de 5, se han combinado clases adyacentes en la tabla 7-3. En consecuencia, el número total de intervalos se reduce de 7 a 4. El valor χ^2 está dado entonces por

$$\chi^2 = \frac{(7-10.1)^2}{10.1} + \frac{(15-10.6)^2}{10.6} + \frac{(10-10.3)^2}{10.3} + \frac{(8-8.3)^2}{8.3}$$

$$= 2.797.$$

El número de grados de libertad para esta prueba será $4 - 3 = 1$, ya que se requieren tres cantidades, la frecuencia total, la media y la desviación estándar de los datos observados, para encontrar las frecuencias esperadas. Como el valor calculado χ^2 es menor que $\chi^2_{0.05} = 3.841$ para 1 grado de libertad, no hay razón para rechazar la hipótesis nula y concluir que la distribución normal proporciona un buen ajuste para la distribución de la vida de los acumuladores.

7.9 PRUEBA DE INDEPENDENCIA

El procedimiento de la prueba ji cuadrada, expuesto en la sección 7.8, también puede emplearse para probar la hipótesis de independencia de dos variables. Ahora, supóngase que se desea estudiar la relación existente entre la

Tabla 7-4. Tabla de contingencia 2 × 3

	Protestantes	*Católicos*	*Judíos*	*Total*
Costa este	182	215	203	600
Costa oeste	154	136	110	400
Total	336	351	313	1000

religión y la región geográfica. Se seleccionan al azar dos grupos de personas, uno de la costa este y otro de la costa oeste de un pais, y se clasifica a cada persona como protestante, católico o judío. Las frecuencias observadas se presentan en la tabla 7-4, la cual se llama *tabla de contingencias*.

A una tabla de contingencias con r filas y c columnas se denomina tabla $r \times c$. El símbolo $r \times c$ se lee "r por c". A los totales de las filas y las columnas de la tabla 7-4 se les llama *frecuencias marginales*. Para probar la hipótesis H_0 de independencia entre la religión de una persona y el lugar donde vive, se debe encontrar primero las frecuencias esperadas para cada celda de la tabla 7-4, en la hipótesis de H_0 es verdadera.

Definamos los siguientes eventos:

P: un individuo seleccionado de la muestra que es protestante.

C: un individuo seleccionado de la muestra que es católico.

J: un individuo seleccionado de la muestra que es judío.

E: un individuo seleccionado de la muestra que vive en la costa este.

O: un individuo seleccionado de la muestra que vive en la costa oeste.

Con las frecuencias marginales se pueden listar las siguientes probabilidades:

$$P(P) = \frac{336}{1000}, \quad P(C) = \frac{351}{1000}, \quad P(J) = \frac{313}{1000}, \quad P(E) = \frac{600}{1000},$$

$$P(W) = \frac{400}{1000}.$$

Ahora, [illegible] es verdadera y las dos variables son independientes, se tendrá

$$P(P \cap E) = P(P)P(E) = \left(\frac{336}{1000}\right)\left(\frac{600}{1000}\right)$$

$$P(P \cap W) = P(P)P(W) = \left(\frac{336}{1000}\right)\left(\frac{400}{1000}\right)$$

$$P(C \cap E) = P(C)P(E) = \left(\frac{351}{1000}\right)\left(\frac{600}{1000}\right)$$

$$P(C \cap W) = P(C)P(W) = \left(\frac{351}{1000}\right)\left(\frac{400}{1000}\right)$$

$$P(J \cap E) = P(J)P(E) = \left(\frac{313}{1000}\right)\left(\frac{600}{1000}\right)$$

$$P(J \cap W) = P(J)P(W) = \left(\frac{313}{1000}\right)\left(\frac{400}{1000}\right)$$

Las frecuencias esperadas se obtienen multiplicando la probabilidad de cada celda por el número total de observaciones. Así, el número esperado de protestantes que viven en la costa este será

$$\left(\frac{336}{1000}\right)\left(\frac{600}{1000}\right)(1000) = \frac{(336)(600)}{1000} = 202$$

cuando H_0 es verdadera. La fórmula general para obtener la frecuencia esperada de cualquier celda está dada por

$$e = \frac{RC}{T},$$

donde R y C son los totales correspondientes a renglones y columnas y T es el gran total de todas las frecuencias observadas. En la tabla 7-5 se anotan, entre paréntesis y al lado del valor real observado, las frecuencias esperadas para cada celda. Nótese que las frecuencias esperadas en cualquier fila o columna se suman al total marginal correspondiente. En el ejemplo sólo se necesitan calcular las dos frecuencias esperadas en la fila superior de la tabla 7-5 y después encontrar las otras mediante resta. Usando los tres totales marginales y el gran total para llegar a las frecuencias esperadas, se han perdido 4 grados de libertad, dejando un total de dos. Una fórmula sencilla para proporcionar el número correcto de grados de libertad está dada por $v = (r - 1)(c - 1)$. Así, en este ejemplo, $v = (2 - 1)(3 - 1) = 2$ grados de libertad.

Tabla 7-5. Frecuencias observadas y esperadas

	Protestantes	*Católicos*	*Judíos*	*Total*
Costa este	182 (202)	215 (211)	203 (187)	600
Costa oeste	154 (134)	136 (140)	110 (126)	400
Total	336	351	313	1000

Para probar la hipótesis nula de independencia, se usa el siguiente criterio de decisión:

PRUEBA DE INDEPENDENCIA *Se calcula*

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i},$$

donde la suma se extiende sobre todas las celdas en la tabla de contingencia $r \times c$. Si $\chi^2 > \chi^2_\alpha$, se rechaza la hipótesis nula de independencia al nivel de significancia α; de otra forma, se acepta la hipótesis nula. El número de grados de libertad es

$$v = (r - 1)(c - 1).$$

Aplicando este criterio al ejemplo en cuestión, se encuentra que

$$\chi^2 = \frac{(182 - 202)^2}{202} + \frac{(215 - 211)^2}{211} + \frac{(203 - 187)^2}{187}$$
$$+ \frac{(154 - 134)^2}{134} + \frac{(136 - 140)^2}{140} + \frac{(110 - 126)^2}{126}$$
$$= 8.556.$$

De la tabla VI se deduce que $\chi^2_{0.05} = 5.991$ para $v = (2 - 1)(3 - 1) = 2$ grados de libertad. La hipótesis nula se rechaza al nivel de significancia 0.05 y se concluye que la creencia religiosa y la región donde se vive no son independientes.

La estadística ji cuadrada para probar la independencia también es aplicable cuando se prueba la hipótesis de que k poblaciones binomiales tienen el mismo parámetro p.

Por lo tanto, esto es una extensión de la prueba presentada en la sección 7.7, de la diferencia entre dos proporciones para la diferencia entre k proporciones. Por lo tanto, si queremos probar la hipótesis

$$H_0: \quad p_1 = p_2 = \cdots = p_k = p$$

contra la hipótesis alterna de que las proporciones de las poblaciones no son iguales, lo cual es equivalente a probar que el número de éxitos o fracasos es independiente de la muestra escogida. Para efectuar esta prueba, se seleccionan primero muestras aleatorias independientes de tamaños $n_1, n_2, \ldots, n_k$ de k poblaciones y los datos se arreglan un una tabla de contingencia $2 \times k$ (tabla 7-6). Las frecuencias esperadas de celda se calculan como antes y se substituyen, junto con las frecuencias observadas, en la fórmula ji cuadrada para independencia con $\nu = k - 1$ grados de libertad. Seleccionando una región crítica apropiada, se puede obtener ahora una conclusión respecto a H_0.

Recuérdese que la estadística en que se basa la decisión tiene una distribución que sólo se aproxima por medio de la distribución ji cuadrada. Los valores calculados χ^2 dependen de las frecuencias de las celdas y consecuentemente son discretos. La distribución ji cuadrada continua parece aproximar muy bien la distribución muestral discreta de X^2 viendo que el número de grados de libertad sea mayor que 1. En una tabla de contingencia 2×2, donde sólo hay 1 grado de libertad, se aplica una corrección llamada *corrección por continuidad de Yates*. La fórmula corregida se convierte en

$$\chi^2 \text{ (corregida)} = \sum_i \frac{(|o_i - e_i| - 0.5)^2}{e_i}.$$

Si las frecuencias de celda esperadas son grandes, los resultados corregidos y no corregidos son casi los mismos. Cuando las frecuencias esperadas se encuentran entre 5 y 10, debe aplicarse la corrección de Yates. Para frecuencias esperadas menores de 5, debe usarse la prueba exacta de Fisher-Irwin. Una explicación de esta prueba puede ser

Tabla 7-6. k muestrales binomiales independientes

	Muestra			
	1	*2*	...	*k*
Exitos	x_1	x_2	...	x_k
Fracasos	$n_1 - x_1$	$n_2 - x_2$	...	$n_k - x_k$

DIRECTORIO DE ASISTENTES AL CURSO

PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES

1 9 8 5

NOMBRE Y DIRECCION	EMPRESA O INSTITUCION
1. LUIS ALCANTARA MENCHINELLY Rancho La Cuchilla N° 93 Fracc. Haciendas Coyoacán Deleg. Coyoacán 04970 México, D.F. 594 81 15	DISTRIBUIDORA CONASUPO Av. Tamaulipas 150-12° piso Col. Condesa Deleg. Hidalgo México, D.F. 553 99 55 exts. 243 y 298
2. EFREN AMADOR MARTINEZ Río de Guadalupe 331 San Pedro el Chico Deleg. Gustavo A. Madero 07480 México, D.F. 760 60 20	VIDRIO PLANO DE MEXICO, S.A. Ex-hacienda de la Santa Cruz San Juan Ixhuatepec Edo. de México 586 02 53
3. ABEL ARRONIZ TOLEDO Vicente Suárez 147-3° piso Col. Condesa México, D.F.	CONGRESO DEL TRABAJO Flores Magón 44-5° piso Col. Guerrero México, D.F.
4. MIGUEL A. BARCENAS SARABIA Sur 107 N° 1303-7 Col. Aeronáutica Militar Deleg. Venustiano Carranza 15970 México, D.F. 768 03 01	FES CUAUTITLAN Cuautitlán Izcalli
5. SANDRA ELVIA CAPETILLO FERREIRO Monte Albán 518-101-A Col. Narvarte Deleg. Benito Juárez 03600 México, D.F. 519 51 34	S C T Xola y Av. Universidad Col. Narvarte Deleg. Benito Juárez México, D.F. 519 51 34
6. MA. DEL ROCIO CRAMONA CALTZONZI Eje Lázaro Cárdenas 817-C-303 Col. Narvarte Deleg. Benito Juárez México, D.F. 696 10 55	DINA CAMIONES, S.A. DE C.V. Margaritas 433 Col. Agrícula Deleg. Alvaro Obregón México, D.F. 658 51 00

7. ADRIANA CASTAÑEDA ALVAREZ
Paricutín 16
Fracc. Los Volcanes
Cuernavaca, Mor.
3 26 82

DIRECCION DE EVALUACION ACADEMICA, UAEM
Chamilpa

8. ANTONIO CASTILLO MARTINEZ
Michoacán 605
Col. Guadalupe Chalma
Deleg. Gustavo A. Madero
México, D.F.

DEPARTAMENTO DEL DISTRITO FEDERAL
División del Norte 3330
Col. Cd. Jardín
Deleg. Coyoacán
México, D.F.
544 01 20

9. MARGARITA CHICAS FRANCO
Oriente 160 N° 238
Col. Moctezuma
México, D.F.

CENTRO DE CONTROL TOTAL DE CALIDADES
Puebla N° 282
Col. Roma
Deleg. Cuauhtémoc
06700 México, D.F.
525 40 10

10. ENRIQUE DEL CASTILLO FRAGOSO
Cuitlahuac 74
Col. Aragón
Deleg. Gustavo A. Madero
07000 México, D.F.
781 53 10

ENEP, ACATLAN
Av. Alcanfores y San Juan Totoltepec
Naucalpan de Juárez
Edo. de México
373 23 99

11. RAMON DOMINGUEZ CASTILLO
Dios Fuego N° 55 - Secc. Los Parques
Cuautitlán Izcalli
873 71 21

INDUSTRIAL VINICOLAS DOMECQ, S.A. DE C.V.
Av. México 135
Col. Del Carmen
Deleg. Coyoacán
México, D.F.
658 33 99

12. MIGUEL FERRER CAMPOS
Parque Central Edif. 6-4
Unidad Adolfo López Mateos
Edo. de México
397 68 32

NACIONAL DE COBRE, S.A.
Poniente 134 N° 719
Col. Industrial Vallejo
Deleg. Atzcapotzalco
México, D.F.
567 11 44

13. CARLOS AUGUSTO GALINDO GALINDO
Cabo Catoche 40
Col. Gabriel Hernández
Deleg. Gustavo A. Madero
México, D.F.
757 35 04

U N A M
Av. Universidad
Deleg. Coyoacán
México, D.F.
550 52 15 ext. 3725

14. ARTURO JIMENEZ MAYEN
Av. 531 N° 89
Unidad Aragón
México, D.F.
551 12 21

PERIODICO EL FINANCIERO
Milton 66
Col. Anzures
México, D.F.
254 32 58

15. RICARDO JIMENEZ NOVELO
Ret. 15 de Fray Servando T. de Mier 9
Col. Jardín Balbuena
Deleg. Venustiano Carranza
15900 México, D.F.
784 78 55

INDUSTRIAS VINICOLAS DOMECQ, S.A.
Av. México 337
Col. Del Carmen
Deleg. Coyoacán
México, D.F.
658 33 99

16. RUBEN LEYVA GUZMAN
Calle Elena N° 191-2
Col. Nativitas
Deleg. Benito Juárez
03500 México, D.F.

IMPRESOS Y CAJAS, S.A.
Arenal 42
Col. Tránsito
Deleg. Venustiano Carranza
México, D.F.
552 72 66

17. EDUARDO MAURICIO LOBREROS LOPEZ
Av. Hidalgo 514
Col. La Libertad
72130 Puebla, Pue.
48 03 31

SRIA. DE COMUNICACIONES Y TRANSPORTES
13 Poniente N° 2306
Col. La Piedad
72000 Puebla, Pue.
48 79 86

18. JOSE ESTEBAN LICONA LOPEZ
Oriente 95 N° 3709
Col. La Joya
Deleg. Gustavo A. Madero
México, D.F.
551 05 45

HOSPITAL A.B.C.
Sur 136 - Esq. Observatorio
Col. América
18901 México, D.F.
277 50 00

19. MA. IRMA LOPEZ SOLACHE
Adolfo Prieto 1385
Col. Del Valle
Deleg. Benito Juárez
México, D.F.

LAB. CENTRAL DE CONTROL DE LA D.G.C.O.H.
Av. División del Norte 3330
Cd. Jardín
Deleg. Coyoacán
México, D.F.
544 01 20 ext. 20

20. ERASMO MONROY CRUZ
Ayuntamiento 159-5
Centro
Deleg. Cuauhtémoc
06040 México, D.F.
512 49 12

ABARROTES LA CARMELITA

21. PEDRO MARROQUIN SUAREZ
San Francisco 340-603
Col. Del Valle
Deleg. Benito Juárez
03100 México, D.F.
687 76 83

FOMENTO INDUSTRIAL SOMEX
Reforma 211-9° piso
Col. Cuauhtémoc
Deleg. Benito Juárez
México, D.F.
591 06 99

22. FELIX LUIS MORALES CASTRO
Cerrada Valle de México N° 1
Secc. Cumbria
Cuautitlán Izcalli
Edo. de México
873 12 34

TELEFONOS DE MEXICO, S.A.
20 de Noviembre 206
Cuautitlán de Romero Rubio
Edo. de México
872 03 13

23. DAVID MORALES MORALES
Calle 17 Mz. 7 Lte. 1
Col. Oriental Rodeo
Deleg. Iztacalco
México, D.F.
558 37 84

GOBIERNO DEL ESTADO DE TLAXCALA
Palacio de Gobierno, P.B. Ciudad
Edo. de Tlaxcala
2 03 66 ext. 144

24. ANTONIO ORTIZ FERNANDEZ
Torres Adalid 1558-104
Col. Narvarte
Deleg. Benito Juárez
03020 México, D.F.

HOSPITAL LA RAZA, IMSS
Vallejo y Jacarandas,
Atzcapotzalco, D.F.

24. JUAN CARLOS F. PAIZANNI HERRERA
Plutarco Elías Calles 660-Edif. Q-2
Col. Iztacalco
02990 México, D.F.
696 74 83

COLEGIO DE BACHILLERES
Plantel 1 - Rosario
México, D.F.
382 71 65

25. XAVIER PATIÑO MONROY
Igualdad 28
Centro
México, D.F.
588 27 59

CECYT N° 10
Loreto Fabela
Deleg. Venustiano Carranza
México, D.F.

26. EVELIA PEREZ RAMOS
Javier Sorondo 260
Col. Villa de Cortez
Deleg. Benito Juárez
03530 México, D.F.
590 54 45

S C T
Av. Universidad y Xola
Cuerpo-C Basamento
Col. Narvarte
Deleg. Benito Juárez
México, D.F.
519 51 34

27. DELFINO RAMIREZ ZUÑIGA
Chihuahua N° 191-501-A
Col. Roma
México, D.F.
531 64 06

PEMEX
Marina Nacional 329
Col. Verónica Anzures
México, D.F.
531 64 06

28. FERNANDO REYES PEREZ
Andes 75
Lomas Verdes IV-Sección
Naucalpan, Edo. de México
393 65 13

CONSTRUCTORA METRO, S.A. DE C.V.
Altadena 23
Col. Nápoles
Deleg. Benito Juárez
México, D.F.
563 85 23

29. MIGUEL ANGEL ROJAS FIGUEROA
Carruajes 17
Villas de la Hacienda
Atizapán, Edo. de México
C.P. 54500

AEROMEXICO
Reforma 445-11° piso
Col. Cuauhtémoc
Deleg. Cuauhtémoc
México, D.F.
514 21 64

30. RICARDO ROSAS LEYVA
Edif. E-10
Unidad San Simón
Deleg. Cuauhtémoc
México, D.F.
658 51 00

DINA CAMIONES, S.A.
Margaritas 433
Col. Agrícola
Deleg. Alvaro Obregón
México, D.F.
658 51 00

31. PATRICIA GPE. RUIZ PIÑA
Xochicalco 21
Col. Narvarte
Deleg. Benito Juárez
03020 México, D.F.
519 25 69

CASTRO ANGOITIA Y ASOCIADOS, S.A.
Benjamín Franklin 222-301
Col. Escandón
11800 México, D.F.
516 03 39

32. JAIME SALGADO RODRIGUEZ
Av. Deportes 168
Las Arboledas
54500 Edo. de México
379 01 74

DEPARTAMENTO DEL DISTRITO FEDERAL
Dirección General de Construcción y
Operación Hidráulica
San Antonio Abad 231
Col. Obrera
México, D.F.
588 37 66

33. BENJAMIN VARGAS DIAZ
Valle Mota 6-1
Valle Aragón
Edo. de México

S C T
Dirección General de Obras Marítimas
Providencia 807-3° piso
Col. Del Valle
México, D.F.
523 48 53