



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

A LOS ASISTENTES A LOS CURSOS

Las autoridades de la Facultad de Ingeniería, por conducto del jefe de la División de Educación Continua, otorgan una constancia de asistencia a quienes cumplan con los requisitos establecidos para cada curso.

El control de asistencia se llevará a cabo a través de la persona que le entregó las notas. Las inasistencias serán computadas por las autoridades de la División, con el fin de entregarle constancia solamente a los alumnos que tengan un mínimo de 80% de asistencias.

Pedimos a los asistentes recoger su constancia el día de la clausura. Estas se retendrán por el periodo de un año, pasado este tiempo la DECFI no se hará responsable de este documento.

Se recomienda a los asistentes participar activamente con sus ideas y experiencias, pues los cursos que ofrece la División están planeados para que los profesores expongan una tesis, pero sobre todo, para que coordinen las opiniones de todos los interesados, constituyendo verdaderos seminarios.

Es muy importante que todos los asistentes llenen y entreguen su hoja de inscripción al inicio del curso, información que servirá para integrar un directorio de asistentes, que se entregará oportunamente.

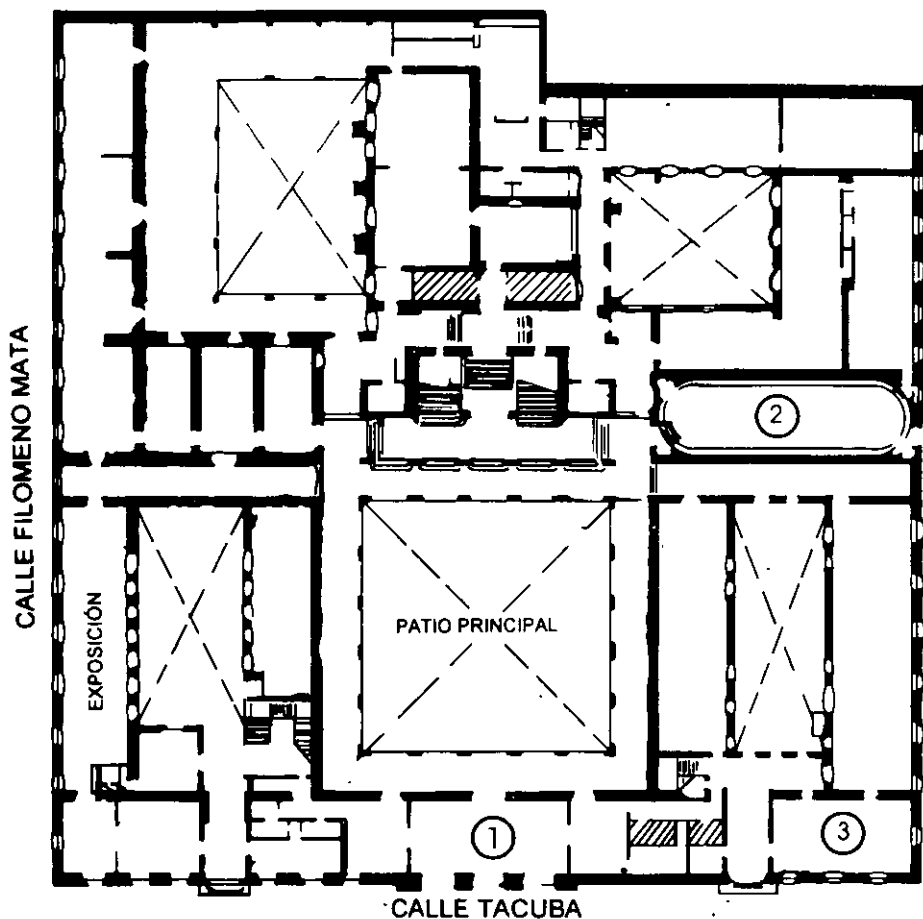
Con el objeto de mejorar los servicios que la División de Educación Continua ofrece, al final del curso deberán entregar la evaluación a través de un cuestionario diseñado para emitir juicios anónimos.

Se recomienda llenar dicha evaluación conforme los profesores impartan sus clases, a efecto de no llenar en la última sesión las evaluaciones y con esto sean más fehacientes sus apreciaciones.

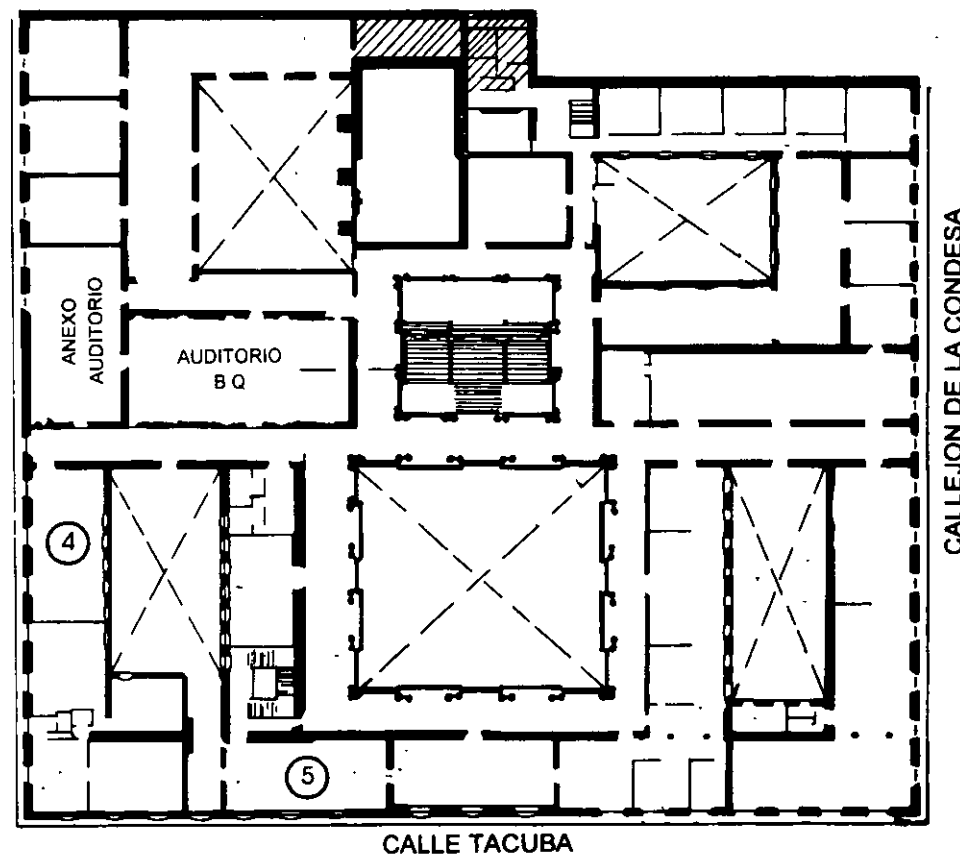
**Atentamente
División de Educación Continua.**



PALACIO DE MINERIA

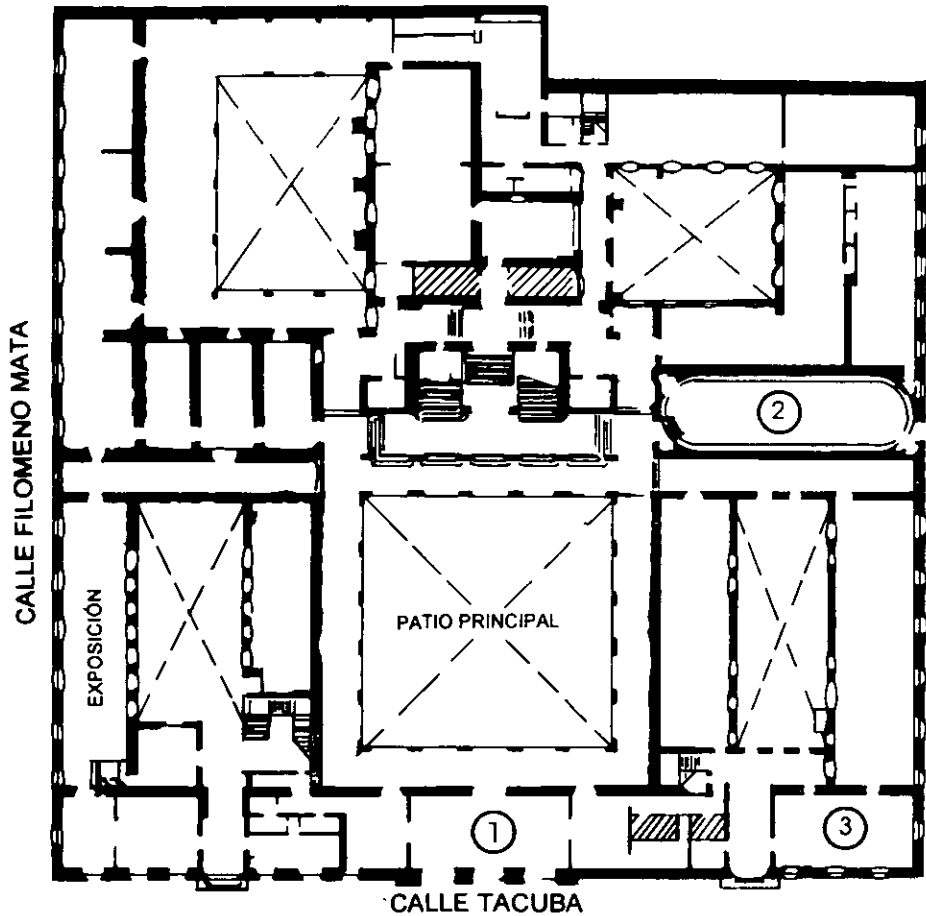


PLANTA BAJA

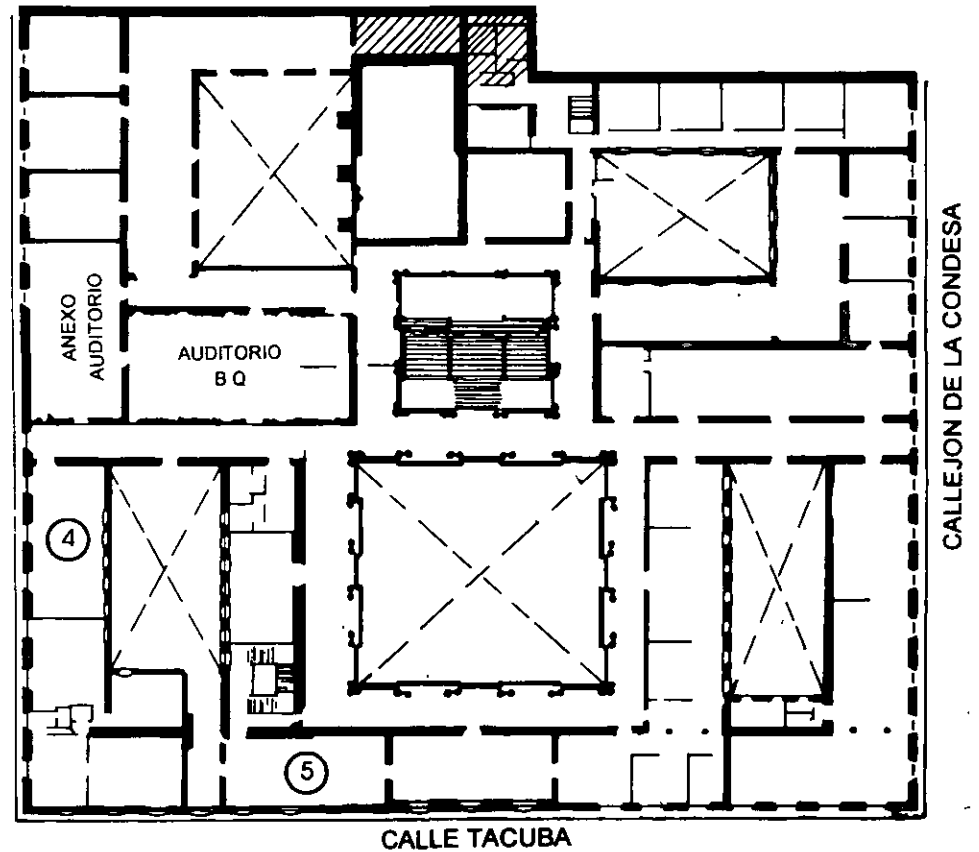


MEZZANINNE

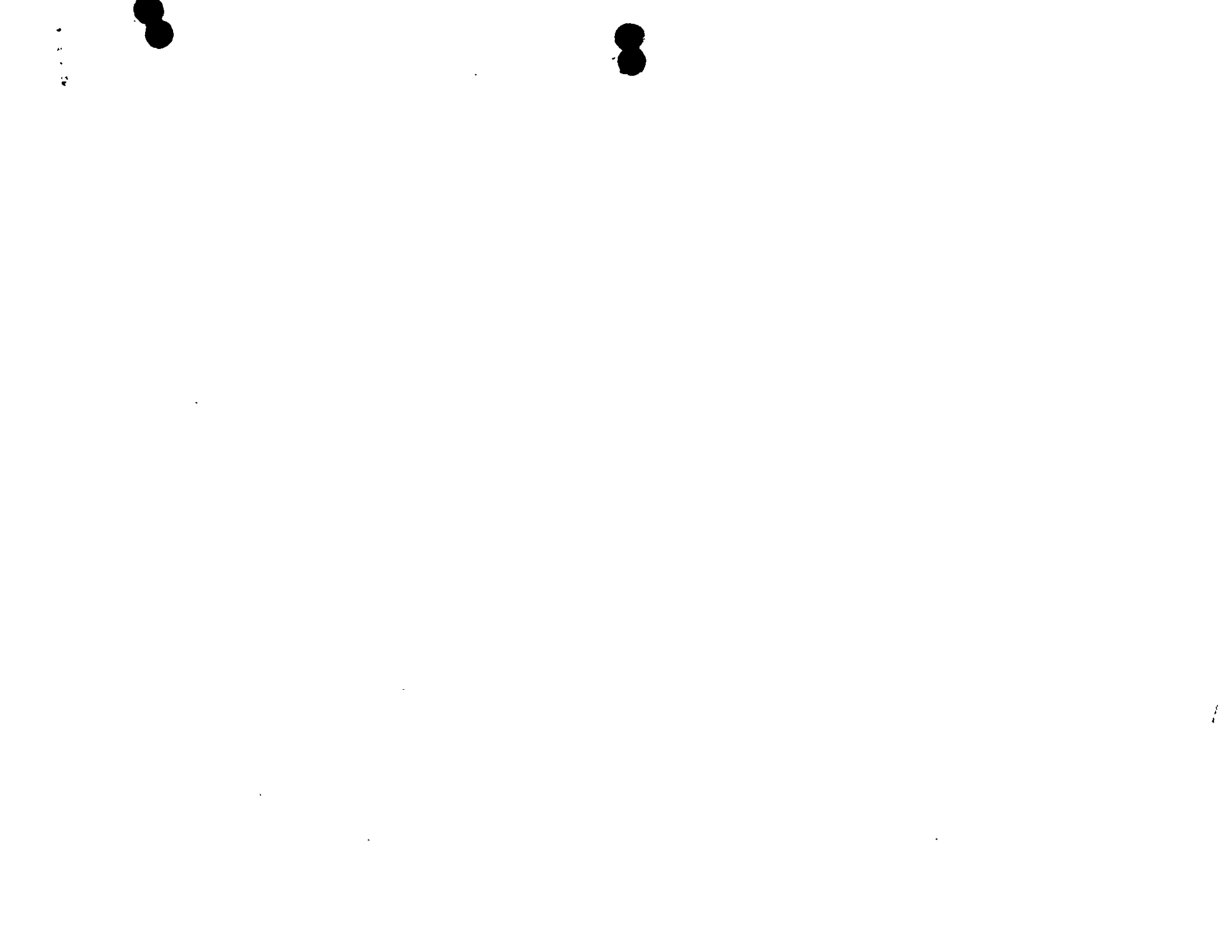
PALACIO DE MINERIA



PLANTA BAJA



MEZZANINNE





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

CURSOS ABIERTOS

ESTADISTICA



TEMA:
**LA ESTADÍSTICA Y LA APLICACIÓN DE SUS PRINCIPALES
HERRAMIENTAS**

M. EN I. ENRIQUE AUGUSTO HERNÁNDEZ RUIZ
PALACIO DE MINERÍA
FEBRERO 2000

LA ESTADÍSTICA Y LA APLICACIÓN DE SUS PRINCIPALES HERRAMIENTAS

M. en I. Enrique Augusto Hernández Ruiz.

• LA ESTADÍSTICA Y LA SIMULACIÓN

EL MÉTODO "PERT"

Esta Técnica de Evaluación y de Revisión de Programas es referida mediante el acrónimo formado por su nombre en inglés (Program Evaluation and Review Technique) y, como anteriormente se afirmó, es similar al de la ruta crítica, pero la diferencia básica consiste en que "PERT" es un método probabilístico que supone el desarrollo de un escenario optimista, uno pesimista y otro esperado; el último de ellos se determina acotado por los dos primeros, aplicando técnicas de simulación.

Obviamente con la programación se trata de establecer el mejor ordenamiento de un conjunto de actividades que deberán ser emprendidas en el futuro, con las cuales se pretende llevar a cabo un proyecto; sin embargo, por ser eventos futuros se tratará inherentemente con incertidumbre respecto de las condiciones en que serán realizadas.

Es precisamente este hecho el que hace que las cosas resulten de manera distinta a la que se supuso de manera determinística, debido a que la ocurrencia de eventos futuros es una cuestión aleatoria, es decir, puede depender del azar; por esta razón es recomendable prepararse para las mejores y las peores condiciones.

El escenario optimista se formulará con la determinación de las mejores condiciones en que pueden ser desarrolladas las actividades del proyecto, o sea, se considerarán los menores costos y mínimas duraciones de las actividades, así como también la mayor disposición de recursos económicos y humanos.

Por el contrario, el escenario pesimista se formará determinando la ocurrencia de las peores condiciones de trabajo en el proyecto, es decir, se supondrá que se alcanzarán los mayores costos y máximas duraciones, además de la menor disposición de los recursos en general.

El escenario esperado resulta ser uno intermedio de los dos anteriores obviamente; su generación se logra simulando el acontecimiento de las situaciones más probables que pueden ocurrir en el futuro, por eso mismo es que este escenario también puede ser llamado "escenario más probable". Un método de simulación muy empleado para este fin es el denominado *Monte Carlo*.

El modelo de Monte Carlo, llamado también método de ensayos estadísticos, es una técnica de simulación de situaciones inciertas que permite definir valores esperados para variables no controlables mediante la selección aleatoria de valores, donde la probabilidad de elegir entre todos los resultados posibles está en estricta relación con sus respectivas distribuciones de probabilidad.

El mecanismo a seguir para realizar tales ensayos estadísticos obedece a estos pasos:

1. Seleccionar un conjunto representativo de proyectos ya realizados para tomar de ellos los parámetros que nos interesan simular en la creación del escenario más probable del proyecto que se pretende realizar.
2. Se formarán clases estadísticas con los datos elegidos, estableciendo un intervalo o amplitud que sea conveniente en las mismas; posteriormente se calculará el valor medio en cada clase, la frecuencia con que se presentó cada clase en el conjunto seleccionado, la frecuencia relativa y la frecuencia relativa acumulada, tal como se ejemplifica en la siguiente tabla:

Valor Medio del Estrato de la Variable "X"	Frecuencia	Frecuencia Relativa P(X)	Frecuencia Relativa Acumulada
200	11	0.1058	0.1058
250	27	0.2596	0.3654
300	34	0.3269	0.6923
350	16	0.1538	0.8462
400	9	0.0865	0.9327
450	5	0.0481	0.9808
500	2	0.0192	1.0000
Suma	104	1.0000	

La frecuencia, frecuencia relativa y frecuencia relativa acumulada deberán cumplir las siguientes condiciones:

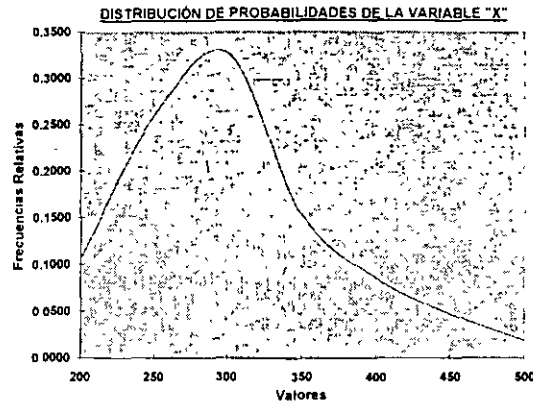
$$n = \sum_{i=1}^k f_i$$

$$fr_i = f_i / n,$$

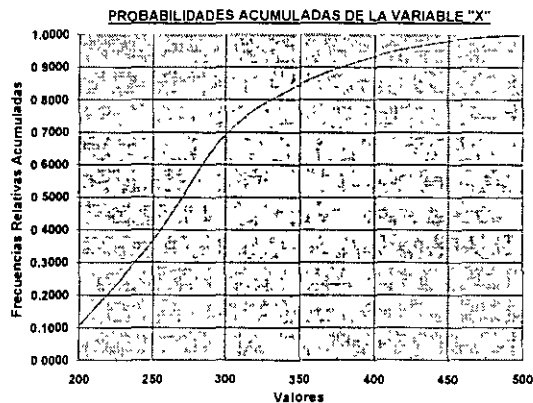
$$fra_i = \sum_{k=1}^i fr_k,$$

donde "n" es el número de elementos que integra el conjunto en estudio, "f_i" el número de elementos del conjunto en estudio que incurren en el estrato "i", "fr_i" y "fra_i" la frecuencia relativa y la frecuencia relativa acumulada que corresponden al estrato "i".

3. Con la frecuencia relativa calculada puede conocerse la distribución de probabilidades de los parámetros tratados, la cual se apreciará en una gráfica como la siguiente:



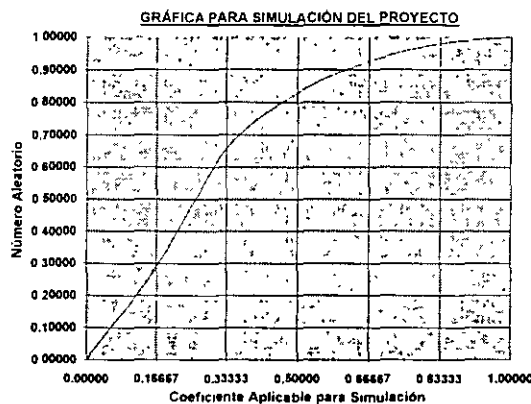
4. La distribución de los datos puede estar concentrada alrededor de cualquier abscisa y para que el proceso de simulación la respete será necesario construir la siguiente gráfica con la frecuencia relativa acumulada que se calculó en la tabla:



5. Como paso inmediato se procederá a "normalizar" el rango empleado en los ejes de las abscisas y de las ordenadas, es decir, a convertir su amplitud de cero a uno con la expresión siguiente:

$$C_j = (V_j - V_{\min}) / (V_{\max} - V_{\min}),$$

donde "V" refiere los valores ubicados en cada eje de la gráfica anterior y "C_i" es cada uno de los valores de los ejes con los que se creará una gráfica "normalizada" como la que se muestra en esta figura:



6. Por último, se generará una serie de números aleatorios que sea lo suficientemente grande para que ésta sea considerada representativa del proyecto en cuanto a su comportamiento; dichos números aleatorios se comprenderán entre el cero y la unidad, y serán ubicados en el eje de las ordenadas de esta última gráfica y se obtendrá el coeficiente que les corresponda sobre las abscisas. Con el coeficiente aplicable para la simulación se aplicará la siguiente expresión, con la cual se calcularán los parámetros buscados con la simulación:

$$S_{Ei} = P_{\min} + C_{Ai} (P_{\max} - P_{\min}),$$

donde "C_{Ai}" es cada coeficiente corregido por la correlación normalizada que se aplicó para el caso específico de cada número aleatorio de la serie generada, "P" corresponde a los valores de los parámetros que se tomaron como base para efectuar el ordenamiento estadístico con que partió el proceso de simulación y S_{Ei} es cada valor resultado de la aplicación de cada "C_{Ai}".

Cabe señalar que en caso de no contar con los datos indicados en el primer paso, los números aleatorios pueden ser aplicados directamente en la expresión del último punto.

Existe la posibilidad de que la cantidad de números aleatorios que se generarán sean determinados mediante la aplicación de conceptos de muestreo aleatorio; incluso estos conceptos pueden apoyar para generar varios escenarios que son posibles y tomar de ellos el más representativo para catalogarlo como el más probable de presentarse en la realidad.

Para concluir con este punto, podemos decir que la simulación aplicada en el método PERT es una manera "virtual" de hacer historia sin correr riesgo, pues en caso de notar deficiencias en puntos específicos del proyecto, es posible tomar medidas pertinentes para evitarlas o disminuirlas sin una repercusión sorpresiva en los recursos.

• EL MUESTREO ALEATORIO

INTRODUCCIÓN

El resultado de un experimento estadístico puede registrarse como un valor numérico o como una representación descriptiva, y es por eso que la estadística se interesa principalmente por el análisis de datos numéricos. En un estudio particular, el número de posibles observaciones puede ser pequeño, grande pero finito, o bien infinito.

Siempre que trabajemos en el muestreo, debemos contar con un plan preciso para delimitar el tamaño de la muestra que deseamos extraer de una población para cumplir con los objetivos de la investigación. Un error muy común consiste en pensar que una muestra debe ser grande para que realmente sea representativa de la población, pero quizá esto no suministre información adecuada sobre el parámetro en cuestión; sin embargo, sí mermará en mucho los recursos económicos que se empleen para llevar al cabo esta actividad.

La totalidad de las observaciones que interesan, sea su número finito o infinito, constituye lo que se llama una "población"; esta palabra considera las observaciones acerca de algo de interés, ya

sean grupos de personas, animales u objetos, y el número de observaciones en la población se define como el tamaño de ésta.

En otros términos, se llamará población al conjunto formado por la totalidad de resultados obtenidos, o posibles, al realizar un experimento cualquiera.

Como ejemplo de una población de tamaño finito podemos citar, entre otros, los números de los naipes de la baraja, las estaturas de los residentes de una ciudad y las longitudes de los peces atrapados en un lago. El experimento de lanzar dados, las observaciones obtenidas al medir la presión atmosférica todos los días, desde el pasado remoto hasta el futuro, o todas las mediciones de la profundidad de un lago en cualquier punto concebible, son ejemplos de poblaciones de tamaño infinito. Algunas poblaciones finitas son tan grandes, que en teoría se supone que son infinitas.

En el campo de la inferencia estadística, interesa lograr conclusiones concernientes a una población cuando es imposible o impráctico observar el conjunto total que forma a la población, y es por eso que se depende de un subconjunto de ésta para poder realizar estudios relativos a la misma. Esto ha conducido al desarrollo de la teoría del muestreo.

A los datos obtenidos al realizar un experimento determinado número de veces se le conocerá como "muestra de la población", por lo que una muestra será entendida como un subconjunto de su población, y para que sean válidas las inferencias que se realicen se deben obtener "muestras representativas" de la citada población.

Con frecuencia, al elegir una muestra se seleccionan los elementos que se consideran más convenientes de la población; pero tal procedimiento puede conducir a inferencias erróneas. Los procedimientos de muestreo que generan inferencias que sobrestimen o subestimen de manera consistente algunas características de la población reciben el nombre de "sesgados".

Para eliminar cualquier posibilidad de sesgo en el procedimiento de muestreo, es deseable recurrir al manejo de "muestras aleatorias", las cuales se seleccionan de modo independiente y al azar, cuyo principal objeto es presentar información representativa acerca de los parámetros de la población que son desconocidos.

Para analizar características específicas de una muestra aleatoria, misma que se considerará representativa de una población, se emplearán los parámetros conocidos como estadísticos, mismos que reciben también el nombre de "medidas de tendencia central". Un estadístico o medida de tendencia central será cualquier función (expresión matemática) que involucre a las variables aleatorias que constituyen una muestra aleatoria.

Los estadísticos más comunes utilizados para determinar el punto medio de un conjunto de datos, dispuestos en orden de magnitud, son la media, la mediana y la moda.

Si X_1, X_2, \dots, X_n constituyen una muestra aleatoria de tamaño "n", donde cada una de ellas tiene la misma probabilidad de ocurrencia, entonces la "media muestral" se define con el estadístico:

$$\mu_x = 1/n \sum_{i=1}^n X_i$$

y en caso de que cada una de estas variables posea su propia y respectiva probabilidad de ocurrencia, el estadístico de la media muestral será:

$$\mu_x = \sum_{i=1}^n P(X_i) X_i.$$

Si X_1, X_2, \dots, X_n constituyen una muestra aleatoria de tamaño "n", dispuesta en orden creciente de magnitud, entonces la "mediana de la muestra" se define con el estadístico siguiente:

$$m_x = X_{(n+1)/2} \text{ si "n" es impar, y}$$

$$m_x = \frac{1}{2} (X_{n/2} + X_{(n/2)+1}) \text{ si "n" es par.}$$

Si X_1, X_2, \dots, X_n , que no son necesariamente diferentes, constituyen una muestra aleatoria de tamaño "n", entonces la "moda muestral" es el valor de la observación que ocurre más a menudo o con la mayor frecuencia. La moda será referida con la letra " M_x ", la cuál puede no existir y cuando existe no es necesariamente única.

De las tres medidas de tendencia central definidas anteriormente, será la media en la que centraremos nuestra atención, pues servirá para definir otras características de índole estadística que referirá la dispersión que existe de los datos muestrales respecto de su media. Esta información que es referida recibe el nombre de momento de orden "k" con respecto a la media; el cuál, cuando los valores de la muestra tienen la misma probabilidad de ocurrencia, es definido de la siguiente manera:

$$m^k = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_x)^k,$$

pero cuando los valores de dicha muestra poseen distintas probabilidades de ocurrencia, la expresión aplicable será:

$$m_x^k = \sum_{i=1}^n P(X_i) (X_i - \mu_x)^k.$$

En lo sucesivo, será el momento de orden dos con respecto a la media el que nos interesará, el cuál será denominado como varianza de la muestra y se determinará con la siguiente expresión cuando exista la misma probabilidad de ocurrencia en los valores de la muestra:

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_x)^2,$$

y como se ha venido señalando, en caso de que los valores que integran la muestra tengan distinta probabilidad de ocurrencia, la expresión anterior será modificada del siguiente modo:

$$\sigma_x^2 = \sum_{i=1}^n P(X_i) (X_i - \mu_x)^2.$$

A la raíz cuadrada de la varianza se le conocerá con el nombre de desviación estándar, misma que se expresará de la forma siguiente:

$$\sigma_x = (\sigma_x^2)^{1/2}.$$

Adicionalmente es posible determinar de una manera relativa o porcentual la dispersión de los datos analizados en una muestra con respecto de su media, la cual se fundamenta en la determinación de un índice conocido como "coeficiente de variación", mismo que guarda la siguiente equivalencia:

$$v_x = \sigma_x / \mu_x.$$

No obstante; existe una cuarta medida de dispersión que no depende de la media de la muestra, ésta recibe el nombre de "rango de la muestra aleatoria". Si X_1, X_2, \dots, X_n son elementos de una muestra aleatoria, el rango se define como $X_n - X_1$, donde X_n y X_1 son, respectivamente, las observaciones mayor y menor de la muestra.

Con base en lo anteriormente descrito y fundado, cabe destacar que la media es fácil de calcular y emplea toda la información disponible, por esa razón los métodos utilizados en inferencia estadística se basan en la media de la muestra. La única desventaja importante de la media es que puede ser afectada en forma nociva por los valores extremos.

La mediana tiene la ventaja de ser fácil de calcular si el número de observaciones es relativamente pequeño, y no es influida por valores extremos. Al considerar muestras tomadas de poblaciones, las medias muestrales por lo general no varían tanto de una muestra a otra como lo harían las medianas, por consiguiente, la media es más estable que la mediana si se intenta estimar el punto central de una población con base en un valor de muestra. En consecuencia, una media muestral ha de estar probablemente más próxima a la media de la población que la mediana de su muestra.

La moda es la medida menos utilizada de las tres medidas de tendencia central ya referidas. Para conjuntos pequeños de datos su valor es casi inútil, si es que existe. Tiene un valor significativo sólo en el caso de una gran cantidad de datos. Sus dos principales ventajas son que:

1. no requiere cálculo y que,
2. se puede utilizar para evaluar datos cualitativos o cuantitativos.

Sin embargo, las tres medidas de tendencia central definidas no dan por sí solas una descripción adecuada de los datos. Se necesita saber en qué grado las observaciones se apartan del promedio, y es entonces donde cobran relevancia las medidas de dispersión, ya que es posible tener dos conjuntos de observaciones con la misma media o mediana que difieran considerablemente en la variabilidad de sus mediciones con respecto a su respectiva media.

El rango puede ser una medida de variabilidad deficiente, en particular si el tamaño de la muestra o población es grande. Tal medida considera sólo los valores extremos y no expresa nada acerca de la distribución de valores comprendidos entre ellos.

La varianza contrarresta la desventaja del rango, y estas dos medidas de dispersión las complementa la desviación estándar.

Si se toma una población finita o infinita con distribución desconocida, con media " μ " y varianza σ^2/n , la distribución de la media de una muestra aleatoria de tamaño " n " de la misma será aún aproximadamente normal, siempre que el tamaño de la muestra sea muy grande. Este sorprendente resultado es una consecuencia inmediata del siguiente teorema llamado "teorema del límite central":

Teorema del límite central: Si μ_x es la media de una muestra aleatoria de tamaño " n " tomado de una población con media μ y varianza finita σ^2 , entonces la forma límite de la distribución de

$$Z = [\mu_x - \mu] / [\sigma / (n)^{1/2}],$$

cuando $n \rightarrow \infty$, es la distribución normal $n(z; 0, 1)$.

La aproximación normal para " μ_x " será aceptable si $n > 30$, independientemente de la forma de la población. Si $n < 30$, la aproximación es aceptable sólo si dicha población no es muy diferente de una distribución normal y, si se sabe que la población es normal, la distribución muestral de " μ_x " seguirá con exactitud una distribución normal, sin que importe qué tan pequeño sea el tamaño de las muestras.

DETERMINACIÓN DEL TAMAÑO DE UNA MUESTRA

La determinación del tamaño de la muestra incluirá puntos específicos, según sea el caso:

1. cuando se estima la media de la población,
2. cuando se estima la proporción de la población,
3. cuando la población es finita y,
4. cuando se aplican técnicas de muestreo estratificado.

TAMAÑO DE LA MUESTRA AL ESTIMAR LA MEDIA DE LA POBLACIÓN

Al prever el intervalo de confianza resultante de una media muestral y la desviación estándar, es posible aplicar la distribución normal a la delimitación previa de la extensión del intervalo y del grado de confianza que nos brindará. Lo que estamos haciendo es examinar la construcción real del intervalo de confianza antes de que efectuemos el estudio y determinemos la media y la desviación estándar.

La fórmula con que se calcula el tamaño necesario de la muestra para estimar la media de la población es:

$$n = Z^2 \sigma^2 / E^2,$$

donde:

n: Tamaño de la muestra.

- Z: Número de unidades de desviación estándar en la distribución normal que producirá el nivel deseado de confianza.
- σ : Desviación estándar de la población (conocida o estimada a partir de estudios anteriores).
- E: Error, o diferencia máxima entre la media muestra y la media de la población que se está dispuesto a aceptar en el nivel de confianza fijado.

La mayor dificultad al determinar el tamaño de la muestra necesaria para estimar la media de la población consiste en calcular la desviación estándar de la población; después de todo, si tuviéramos un conocimiento completo sobre la población, no habría necesidad de realizar una investigación sobre sus parámetros estadísticos. Si no podemos confiar en los trabajos anteriores, para calcular la desviación estándar de la población, las alternativas incluyen el juicio o el empleo de estudios exploratorios con muestras pequeñas para conocer su valor.

Si lo preferimos, podemos abordar este mismo tipo de problema desde el punto de vista del "error permisible relativo" en vez del "error absoluto". En este caso la desviación estándar " σ " y el error permisible "E" se expresan en función de su porcentaje de la media verdadera de la población connotada como " μ ". La ecuación más apropiada en este caso se parece a la que acabamos de presentar y será:

$$n = Z^2 (\sigma / \mu)^2 / (E / \mu)^2,$$

donde:

- n: Tamaño de la muestra.
- Z: Número de unidades de desviación estándar en la distribución normal que producirá el nivel deseado de confianza.
- σ : Desviación estándar de la población (conocida o estimada a partir de estudios anteriores).
- μ : Media de la población.
- E: Error, o diferencia máxima entre la media muestra y la media de la población que estamos dispuestos a aceptar en el nivel de confianza que hemos indicado.

TAMAÑO DE LA MUESTRA AL ESTIMAR LA PROPORCIÓN DE LA POBLACIÓN.

Determinar el tamaño necesario de la muestra en este caso se parece en principio al procedimiento que seguimos en la sección anterior, salvo que ahora se trata de una proporción y no de una media. La fórmula apropiada es:

$$n = Z^2 P(1 - P) / E^2,$$

donde:

- n: Tamaño necesario de muestra.
- Z: Número de unidades de desviación estándar en la distribución normal, que producirá el grado deseado de confianza.
- P: Proporción de la población que posee la característica de interés.
- E: Error, máxima diferencia entre la media muestral y la media de la población que estamos dispuestos a aceptar en el nivel de confianza señalado.

Al aplicar esta fórmula, primero hay que decir si podemos estimar aproximadamente el valor de la proporción de la población "P"; y en caso de que podamos decir con seguridad que esa proporción difiere mucho de 0.5 en una u otra dirección, estaremos en condiciones de obtener la precisión deseada con un tamaño más pequeño (y menos caro) de la muestra. Como se aprecia en la fórmula, el tamaño será proporcional al producto de P(1 - P) y este producto es mayor cada vez que P=0.5. Observe detenidamente los siguientes productos de P(1 - P):

P	(1-P)	P(1-P)
0.5	0.5	0.25
0.4	0.6	0.24
0.3	0.7	0.21
0.2	0.8	0.16
0.1	0.9	0.09

Como se advierte en la expresión anterior, el producto se vuelve muy pequeño cuando una proporción de la población es sumamente pequeña o demasiado amplia. Así pues, si podemos acortar por lo menos el valor de la proporción de la población, ahorraremos dinero al poder valernos de un tamaño más pequeño de la muestra.

Si tratamos de medir el valor de una proporción de la población pero ignoramos los resultados probables, quizá queramos realizar una encuesta exploratoria con objeto de hacernos una idea aproximada de la proporción. En caso de que la proporción resultante sea muy diferente de .5, plantearemos para conseguir un tamaño menor en la fase principal del estudio.

MUESTREO CON POBLACIONES FINITAS

Hasta ahora hemos supuesto que la muestra será relativamente pequeña en comparación con la población total. Sin embargo, hay casos en que la muestra es 5% o más de la población; entonces hemos de modificar ligeramente el procedimiento. Después de todo si extraemos una muestra de 900 personas de una población de 1 000, tendremos una muy buena idea de la media o proporción de la población. Dicho de otra manera, a medida que el tamaño de la muestra se acerca al de la población, desaparece el error muestral y a la postre tendremos un censo completo de la población. El punto crítico de 5% no es más que una regla práctica, pero suficiente en casi todos los trabajos. En caso de duda, supondremos que la población es finita y aplicaremos las siguientes fórmulas de corrección.

TAMAÑO DE LA MUESTRA AL ESTIMAR LA MEDIA DE UNA POBLACIÓN FINITA

$$n = \sigma^2 / [(E^2 / Z^2) + (\sigma^2 / N)],$$

donde:

- n: Tamaño de la muestra.
- N: Tamaño de la población.
- Z: Número de unidades de desviación estándar en la distribución normal que producirá el nivel deseado de confianza.
- σ : Desviación estándar de la población (conocida o estimada a partir de estudios anteriores).

E: Error, o diferencia máxima entre la media muestra y la media de la población que se está dispuesto a aceptar en el nivel de confianza establecido.

TAMAÑO DE LA MUESTRA AL ESTIMAR LA PROPORCIÓN DE UNA POBLACIÓN FINITA.

La expresión que será empleada en esta caso será la siguiente:

$$n = P(1-P) / \{(E^2 / Z^2) + [P(1-P) / N]\},$$

n: Tamaño necesario de muestra.

N: Tamaño de la población.

Z: Número de unidades de desviación estándar en la distribución normal, que producirá el grado deseado de confianza.

P: Proporción de la población que posee la característica de interés.

E: Error, o diferencia máxima entre la media muestra y la media de la población que se está dispuesto a aceptar en el nivel de confianza establecido.

TAMAÑO DE LA MUESTRA EN UN MUESTREO ESTRATIFICADO DE UNA POBLACIÓN FINITA

Para llevar a cabo este tipo de muestreo, habrá que considerar que se trata con una población sumamente grande que es dividida a su vez en un determinado número de estratos o clases, y que de acuerdo con el teorema del límite central de probabilidad, dicha población y sus estratos obedecen a una distribución aproximada a la normal, cuya función es:

$$F(Z) = e^{-Z^2/2} / (2\pi)^{1/2},$$

donde:

Z: Variable aleatoria cuyo valor dependerá del grado de confianza que se espera en la muestra.

F(Z): Ordenada de la variable aleatoria Z.

e: Número equivalente al número real 2.718281828459.

π : Número equivalente al número real 3.14159265359.

Habrá que tener presente, que en este caso, según los postulados de la probabilidad, se cumple lo siguiente:

$$\mu_x = \mu_x^*,$$

$$\sigma_x^{*2} = [\sigma_x^2 / n] [(N - n) / (N - 1)],$$

donde:

μ_x : Media estadística de todo el estrato o clase de la población.

μ_x^* : Media de la muestra del estrato o clase de la población.

σ_x^2 : Varianza de todo el estrato o clase de la población.

σ_x^{*2} : Varianza de la muestra del estrato o clase de la población.

N: Número total de elementos en el estrato o clase de la población.

n: Tamaño de la muestra del estrato o clase de la población.

El tamaño de la muestra de cada estrato o clase se determinará con base en los fundamentos de probabilidad ya expuestos, y sustituyéndolos en la expresión aplicable para determinar el tamaño de una muestra estimando la media de una población infinita:

$$n = Z^2 \sigma_x^2 / E^2,$$

de donde se desprende la siguiente fórmula:

$$n = N Z^2 (\sigma_x)^2 / [E^2 + Z^2 (\sigma_x)^2],$$

donde:

n: Tamaño de la muestra del estrato o clase de la población.

N: Número total de elementos en el estrato o clase de la población.

Z: Variable aleatoria cuyo valor dependerá del grado de confianza que se espera en la muestra.

σ_x^2 : Varianza del estrato o clase de la población.

E Error entre el estrato y su muestra.

El error o desviación existente entre todo el estrato y su muestra se interpretará como la diferencia entre la media estadística del estrato y la media de la muestra ($\mu_x - \mu_x^*$), por lo que la expresión anterior será equivalente a la siguiente:

$$n = N Z^2 (\sigma_x)^2 / [(\mu_x - \mu_x^*)^2 + Z^2 (\sigma_x)^2],$$

Al momento de aplicar esta última expresión, podría pensarse que es necesario conocer el valor de la media de la muestra del estrato o clase, pero esto no es posible si no ha procedido la acción de muestreo. Sin embargo, este error o diferencia puede establecerse en términos porcentuales relativos a la media de todo el estrato o clase; o bien como un valor absoluto, por lo que la igualdad expresada puede ser aplicada de la siguiente manera:

$$n = N Z^2 (\sigma_x)^2 / [(0.10 \mu_x)^2 + Z^2 (\sigma_x)^2],$$

El valor que corresponde a la variable aleatoria "Z" se determinará de acuerdo con el grado de confianza que convenga aplicar al caso. A continuación se refieren los valores de la variable aleatoria "Z" con diferentes niveles de confianza, que van del 90 al 99% :

NIVEL DE CONFIANZA (%)	VALOR APLICABLE DE "Z"
90	1.645
91	1.695
92	1.750
93	1.810
94	1.880
95	1.960
96	2.055
97	2.170
98	2.330
99	2.575

• ÁRBOLES DE PROBABILIDAD Y DE DECISIÓN

AXIOMAS DEL COMPORTAMIENTO RACIONAL

Cada individuo tiene una estructura de preferencias que servirá para seleccionar las mejores alternativas de decisión, siempre y cuando dicha estructura quede acotada por lo que se considera un comportamiento racional establecido por seis axiomas.

Antes de referir los seis axiomas mencionados, habrá que definir lo que es una lotería. Una lotería es un ensayo probabilístico caracterizado por un conjunto mutuamente excluyente y colectivamente exhaustivo de resultados posibles: $r_1, r_2, r_3, \dots, r_m$ y sus probabilidades respectivas: $p_1, p_2, p_3, \dots, p_m$, donde la suma de todas ellas sea equivalente a la unidad.

Al analizar cualquier situación deberemos considerar todas las consecuencias posibles y se dirá que una lista de consecuencias completa es *colectivamente exhaustiva*. Por otro lado, si ocurre una consecuencia, no debe haber sucedido al mismo tiempo otra de ellas de la lista, es decir, será imposible que se presenten combinaciones de ellas; cuando se cumpla esta condición para cualquier consecuencia de una lista, se dirá que éstas son *mutuamente excluyentes*.

Las loterías serán representadas de la siguiente manera:

$$L[r_1, r_2, r_3, \dots, r_m; p_1, p_2, p_3, \dots, p_m].$$

Por ejemplo, $L[60, -20; 0.7, 0.3]$ representa la lotería donde es factible ganar 60 unidades con probabilidad del 70%, o perder 20 unidades con probabilidad del 30%.

Una lotería puede poseerse y venderse; poseerla implica que se tiene la obligación de jugarlo se quiera o no. Sin embargo, en el momento en que dicha lotería proporciona un resultado, ésta deja de existir.

El deseo por una lotería dependerá de sus consecuencias y probabilidades de ocurrencia, así como de los aspectos personales; tales aspectos incluyen los recursos de la persona en juego, sus necesidades y sus actitudes ante el riesgo.

La lotería del ejemplo anterior podría ser valuada en 40 unidades por alguien que ha incrementado su patrimonio asumiendo riesgos y una persona más conservadora tal vez la valuaría en 10 unidades, pero para una persona que actúa conforme a los promedios, esta lotería representaría un valor de 36 unidades, valor que se obtendría siguiendo el siguiente razonamiento:

$$V_E = \sum_{i=1}^m (r_i p_i).$$

Para el caso de $L[60, -20; 0.7, 0.3]$:

$$\begin{aligned} V_E &= (60)(0.70) + (-20)(0.30) \\ V_E &= 42 - 6 \\ V_E &= 36 \end{aligned}$$

Para analizar oportunidades de decisión existe una estructura de preferencias tal que si nos presentan dos resultados cualesquiera: r_j y r_k , siempre se podrá decir que se prefiere r_j , que se prefiere r_k , o bien que se es indiferente entre ambas. La notación que se utilizará para referir lo anterior será la siguiente:

$$\begin{aligned} r_j > r_k &.- \text{ se prefiere } r_j \\ r_j < r_k &.- \text{ se prefiere } r_k \\ r_j r_k &.- \text{ hay indiferencia ente } r_j \text{ y } r_k \end{aligned}$$

Una vez mencionado lo anterior, se procederá a enunciar los axiomas del comportamiento racional, indicando que aceptar que éstos normen nuestras preferencias es equivalente a convenir que un decisor racional nunca violará estas reglas de selección de alternativas.

Axioma 1: Comparación de loterías con consecuencias idénticas. Si se tienen dos loterías con dos resultados que son los mismos para ambas, se elegirá aquella cuya probabilidad de obtener el mejor resultado sea mayor. Por ejemplo, si $L_1[120, -50; 0.8, 0.2]$ y $L_2[120, -50; 0.6, 0.4]$ son dos loterías posibles de acceder, se preferirá L_1 .

Axioma 2: Cuantificación de preferencias. Sean "a" y "b" dos valores tal que $a > r_j > b$, entonces para r_j podrá especificarse un número "q(r_j)" entre cero y uno tal que se sea indiferente entre poseer r_j con certeza y poseer la lotería $L[a,b; q(r_j), 1-q(r_j)]$. Por ejemplo, supóngase que se tiene una lotería con dos posibles resultados: se pueden ganar 800 unidades o perder 100 y, al mismo tiempo, se tiene la posibilidad de obtener con certeza 700 unidades. Si se considera que quien tomará la decisión sigue un comportamiento lineal ante el riesgo, el número "q(r_j)" se determinará planteando que:

$$\begin{aligned} &L[800, -100; q(r_j), 1-q(r_j)] \\ &[800][q(r_j)] + [-100][1-q(r_j)] = 700 \\ &q(r_j) = 0.8889 \end{aligned}$$

Nota.- Cuando el decisor sigue un comportamiento diferente al lineal ante el riesgo, como ocurre en la mayor parte de los casos, habrá que construir una función de utilidad.

Axioma 3: Cuantificación de incertidumbre. Para cada evento "E" existe una cantidad $p(E)$ entre cero y uno tal que so es indiferente entre la lotería $L[a,b; p(E), 1-p(E)]$, y otra lotería donde se tendrá "a" si "E" ocurre y "b" si éste no sucede.

Axioma 4: Transitividad. Sea una lotería con r_1, r_2 y r_3 como resultados. Si $r_1 > r_2$ y $r_2 > r_3$, entonces $r_1 > r_3$. Así mismo, si $r_1 r_2$ y $r_2 r_3$, implica que $r_1 r_3$.

Axioma 5: Sustitución de resultados. Si una oportunidad de decisión se modifica reemplazando un resultado r_j con otro resultado r_k , y si se es indiferente entre estos dos resultados, entonces también habrá indiferencia entre la posesión de la oportunidad original y la posesión de la modificada.

Axioma 6: Equivalencia entre la situación real y de conjetura. Si se tienen determinados resultados que dependen de la ocurrencia de un evento "E" y además un orden de preferencias por ellos, después de que acontece "E" nuestro orden de preferencias por esos resultados deberá seguir siendo el mismo.

FUNCIONES DE UTILIDAD

El análisis de decisiones consta de tres etapas: estructurar las oportunidades de decisión, cuantificar los impactos y la incertidumbre y, establecer los criterios de decisión y determinar la mejor opción. Pero cada individuo tiene su propia estructura de preferencias para seleccionar las mejores decisiones, siempre y cuando dicha estructura no esté fuera de lo que se considera un comportamiento racional establecido por los seis axiomas expuestos anteriormente.

Para cuantificar la preferencia de un resultado " r_i " que está dentro de un intervalo $[B, A]$, donde "B" es lo que se prefiere menos y "A" lo que se prefiere más, se buscará la probabilidad " $q(r_i)$ " que hace equivalente en preferencias a " r_i " con la lotería donde con probabilidad " $q(r_i)$ " se obtendrá "A" y con probabilidad " $1-q(r_i)$ " resultará "B". La preferencia del resultado " r_i " es la probabilidad de obtener "A", es decir, " $q(r_i)$ ".

Supóngase que se desea cuantificar la preferencia que se tiene por la lotería: $L[r_1, r_2, r_3, \dots, r_m; p_1, p_2, p_3, \dots, p_m]$; para ello se cuantificará primero las preferencias de " r_1 ", " r_2 ", " r_3 ", ..., " r_m ". Asíumase que éstas son, respectivamente " $q(r_1)$ ", " $q(r_2)$ ", " $q(r_3)$ ", ..., " $q(r_m)$ ", de manera que:

- " r_1 " es equivalente a la lotería $L[A, B; q(r_1), 1-q(r_1)]$,
- " r_2 " es equivalente a la lotería $L[A, B; q(r_2), 1-q(r_2)]$,
- " r_3 " es equivalente a la lotería $L[A, B; q(r_3), 1-q(r_3)]$,
- ... y,
- " r_m " es equivalente a la lotería $L[A, B; q(r_m), 1-q(r_m)]$.

De manera que la probabilidad de tener "A" (preferencia de "A") y la probabilidad de tener "B" son:

$$P(A) = p_1q(r_1) + p_2q(r_2) + p_3q(r_3) + \dots + p_mq(r_m)$$

$$P(B) = 1 - P(A)$$

Si se observa, la preferencia de una cantidad se obtiene sustituyendo cada " r_i " por su preferencia " $q(r_i)$ " en la lotería y calculando su valor esperado " V_E " que ya fue inducido anteriormente.

Se dice que las probabilidades de indiferencia o cualquier transformación lineal positiva de la forma " $u(r_i) = aq(r_i) + b$ ", donde " a " debe ser mayor que cero, constituyen una función de utilidad.

Una persona que cumpla con el primer axioma del comportamiento racional preferirá un resultado con una probabilidad de indiferencia más grande que un resultado con una probabilidad de indiferencia menor.

Debido a que la utilidad de un resultado es una transformación lineal positiva de su probabilidad de indiferencia, puede aseverarse que una persona siempre preferirá un resultado cuya utilidad sea más grande que otro con utilidad menor y, que la utilidad de una lotería es el valor esperado de las utilidades de los resultados posibles de la lotería.

Por ejemplo, supóngase que se desea cuantificar las preferencias que un decisor siente por las cantidades que varían de -10 u.m. (unidades monetarias) a 80 u.m. La pregunta será: ¿cuál es la mínima cantidad por la que estará dispuesto a vender esta lotería, en la que con probabilidad de 60% puede ganar 80 u.m. y con probabilidad de 1-0.60 perder 10 u.m.? Supóngase que el decisor responde que la cantidad es 30 u.m.

Seguidamente, con esa información se formarán dos loterías: $L_1[80, 30; 0.6, 0.4]$ y $L_2[30, -10; 0.6, 0.4]$, y de nuevo se le preguntará al decisor por las mínimas cantidades por las que estará dispuesto a cambiar las loterías si es que son atractivas para él, o las máximas cantidades que pagaría por no tener que jugarlas si no son de su agrado. Asíumase que esas cantidades son 50 y 5 u.m. respectivamente para ambas loterías, cantidades que serán empleadas posteriormente.

Como la cantidad que prefiere más es 80 u.m., se le asignará una utilidad equivalente a la unidad, y como la que se prefiere menos es -10 u.m. se le asignará una utilidad de cero. En este momento se aceptará que la utilidad de una lotería es su utilidad esperada.

La utilidad esperada de la primera lotería se calcula sustituyendo sus valores monetarios por sus utilidades y calculando el valor esperado de la siguiente manera:

$$V_E = (1) (0.60) + (0) (0.40) \\ V_E = 0.60.$$

Como la primera lotería es equivalente a 30 u.m., la utilidad de 30 u.m. y de la lotería deben ser iguales, es decir, 0.60.

A continuación se sustituirá en las últimas dos loterías las cantidades por sus utilidades respectivas, de manera que las utilidades de la segunda y tercera loterías son:

$$V_{E1} = (1) (0.60) + (0.60) (0.40) \\ V_{E1} = 0.84,$$

$$V_{E2} = (0.60) (0.40) + (0) (0.40) \\ V_{E2} = 0.24.$$

Como el decisor definió que 50 u.m. es equivalente a " L_1 ", su utilidad es 0.84 y, como 5 u.m. es equivalente a " L_2 ", su utilidad es 0.24.

La función utilidad será la curva que se trace a través de estos cinco puntos: [-10,0.00], [5,0.24], [30,0.60], [50,0.84] y [80,1.00], donde en el eje horizontal se representarán las unidades monetarias y en el vertical las utilidades. Esta función utilidad se podrá utilizar para analizar las oportunidades de decisión, ya que un resultado de la aceptación de los seis axiomas es que se elegirá aquella alternativa que maximice la utilidad esperada.

ÁRBOLES DE PROBABILIDAD

La incertidumbre la medimos con números denominados probabilidades que varían de cero a uno, donde por convención es aceptado que la unidad representa la probabilidad de lo que acontece con certeza y el cero la probabilidad del evento imposible. Todos los eventos imposibles tienen probabilidad cero, aunque no todos los eventos con probabilidad cero son imposibles.

La teoría de la probabilidad es un conjunto de deducciones derivadas de los axiomas formulados por Kolmogorov, que son los siguientes:

1. Una probabilidad es un número entre cero y uno asignado a una consecuencia, mismo que representa su posibilidad de ocurrencia.
2. La suma de las probabilidades que les corresponden a las consecuencias mutuamente excluyentes y colectivamente exhaustivas debe ser equivalente a la unidad.
3. La probabilidad de una consecuencia compuesta por consecuencias mutuamente excluyentes es la suma de sus probabilidades.

Con fundamento en la fuente del conocimiento de la probabilidad, ésta puede dividirse en dos tipos:

- Subjetiva o a priori y,
- Objetiva, estadística o a posteriori.

La primera se formula con suposiciones hechas por el decisor y la segunda se fundamenta en el análisis de hechos consumados. Cuando se cuenta con un conjunto de datos ordenado por clases estadísticas, la frecuencia con que se presentó cada una de ellas determinará su probabilidad de ocurrencia en el futuro, por lo que la frecuencia relativa calculada en un estudio estadístico, se convertirá automáticamente en una probabilidad.

En lo sucesivo se conocerá con el nombre de *experimento aleatorio* a cualquier acción que dé origen a un resultado cualquiera que dependerá del azar, y se llamará evento a cada uno de estos resultados posibles. La probabilidad de un evento es una cantidad que carece de unidades, comprendida entre el cero y la unidad, incluyendo estos límites, y suele expresarse como una magnitud porcentual (porcentaje); dicha cantidad referirá el número de veces que ocurrirá el evento al repetir cien veces un experimento aleatorio.

Supóngase que se desea conocer la probabilidad de que en un mes cualquiera se presente un temblor de grado 7 en la escala Richter en la ciudad de México, sabiendo que los movimientos de esta magnitud ocurren al menos cada 50 años (periodo de retorno). Dicha probabilidad se determinará del siguiente modo:

$$P_m(T_7) = (1 \text{ mes}) / [(50 \text{ años}) (12 \text{ meses} / \text{año})]$$

$$P_m(T_7) = 1 / [(50) (12)]$$

$$P_m(T_7) = 0.167\%$$

Si, en cambio, se desea conocer la probabilidad de que en un año cualquiera se presente el mismo evento, ésta quedará establecida de una forma análoga:

$$P_a(T_7) = (1 \text{ año}) / (50 \text{ años})$$

$$P_a(T_7) = 1 / 50$$

$$P_a(T_7) = 2.00\%$$

En el mismo sentido, la probabilidad de que se presente en un día cualquiera será:

$$P_d(T_7) = (1 \text{ día}) / [(50 \text{ años}) (12 \text{ meses / año}) (30 \text{ días / mes})]$$

$$P_m(T_7) = 1 / [(50) (12) (30)]$$

$$P_m(T_7) = 0.006\%$$

Por su puesto, si conocemos que un evento de esa índole se presentó hace 49 años, la probabilidad de que se presente en el año en curso será mucho más alta, aunque sin llegar a ser seguro, y se determinará acumulando la probabilidad que correspondió a cada uno de esos 49 años, es decir:

$$P_a'(T_7) = (49 \text{ años}) / (50 \text{ años})]$$

$$P_m(T_7) = 49 / 50$$

$$P_m(T_7) = 98.00\%$$

La dificultad que representa tomar una decisión se incrementará a medida que el número de eventos que pueden ocurrir se incrementa. La incertidumbre asociada a una consecuencia cambia al variar nuestro conocimiento, por lo que las probabilidades estarán condicionadas por dicho conocimiento.

Al tener conocimiento de una nueva información habrá una nueva probabilidad, conocida como probabilidad condicional. Dicha probabilidad se calcula como el cociente que resulta de dividir la probabilidad de que ocurran simultáneamente, el evento cuya nueva probabilidad se quiere obtener y el evento condicionante, entre la probabilidad del evento condicionante, es decir:

$$P(\alpha/\beta) = P(\alpha \bullet \beta) / P(\beta)$$

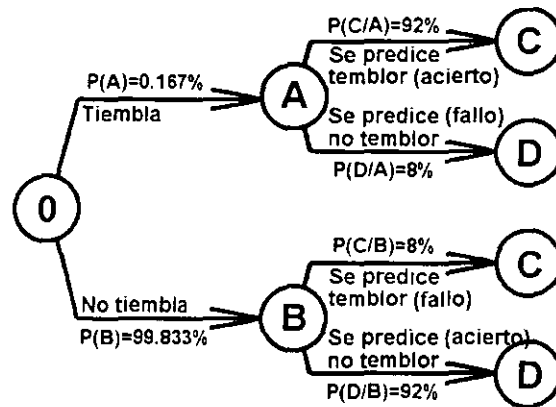
$$P(\alpha \bullet \beta) = [P(\alpha)] [P(\beta)]$$

Este concepto fundamentará la formulación de árboles de probabilidad, los cuales son la representación gráfica que esquematiza secuencialmente situaciones fallibles, es decir, que pueden ocurrir o no como una dependencia del azar.

Un árbol de probabilidad se conformará con nodos de incertidumbre que dependerán de otros y, a su vez, otros dependerán de ellos. Su principal característica es que, sobre un nodo de incertidumbre, no se tiene ningún control, es decir, no hay certeza sobre su acontecimiento.

Los nodos de incertidumbre se representarán gráficamente mediante el uso de circunferencias y las relaciones que entre ellos existan con líneas rectas o con un solo quiebre.

Con base en el ejemplo anterior donde se determinaron probabilidades de ocurrencia de un sismo con un periodo de retorno de 50 años, supóngase ahora que además existe una "alarma sísmica", un aparato que puede predecir temblores con 1 mes de anticipación y que acierta en sus pronósticos el 92% de las ocasiones. Las probabilidades correspondientes serán determinadas con el apoyo de un árbol de probabilidad como el que a continuación se expone:



Nótese que la predicción de que temblará será un "acierto" del aparato en caso de que efectivamente tiemble en el mes que se está analizando, pero será un "fallo" si no tiembra; en cambio, el aparato incurrirá en un "fallo" si predice que no temblará y en realidad ocurre un sismo, pero será un "acierto" si hay ausencia del mismo. Lo anterior induce que el "acierto" o "fallo" de la alarma sísmica dependerá de que tiemble o no, por tal motivo éstas ramas del árbol de probabilidad dependen de las que señalan los hechos de temblar o no.

En la gráfica anterior el evento "A" representa la ocurrencia de un temblor, el evento "B" la carencia del mismo, el evento "C" el pronóstico del aparato afirmando que temblará y el evento "D" el pronóstico del aparato negando la presencia de un sismo. También, en la misma, está reflejado lo que se ha expuesto en el párrafo anterior: si el aparato pronostica que temblará y, transcurrido el mes, efectivamente tiembra, entonces el aparato habrá acertado (92%), pero si no tiembra, el aparato habrá fallado (1-0.92=8%); en cambio, si el aparato pronostica que no temblará y en realidad tiembra, el aparato habrá incurrido en un fallo (8%), pero si no ocurre el temblor, entonces el aparato habrá acertado (92%).

Los eventos "A" y "B" son complementarios ya que son los únicos dos posibles resultados del experimento aleatorio (temblará o no temblará), por lo que ambos cumplen con el primero y el segundo de los axiomas de Kolmogorov. Igualmente, los eventos "C" y "D" (el aparato pronostica que temblará, o bien, pronostica que no temblará) cumplen con el mismo teorema.

La probabilidad de que ocurran simultáneamente los eventos "A" y "C" se calculará multiplicando las probabilidades que se ubican sobre las ramas correspondientes del árbol, es decir:

$$P(A \bullet C) = (0.00167) (0.92)$$

$$P(A \bullet C) = 0.154\%$$

y de la misma manera se podrán calcular las demás probabilidades:

$$P(A \bullet D) = (0.00167) (0.08)$$

$$P(A \bullet D) = 0.013\%$$

$$P(B \bullet C) = (0.99833) (0.08)$$

$$P(B \bullet C) = 7.987\%$$

$$P(B \bullet D) = (0.99833) (0.92)$$

$$P(B \bullet D) = 91.846\%$$

Nótese que igualmente estos cuatro eventos compuestos cumplen con los primeros dos axiomas de Kolmogorov, ya que:

$$P(A \bullet C) + P(A \bullet D) + P(B \bullet C) + P(B \bullet D) = 100\%$$

$$0.00154 + 0.00013 + 0.07987 + 0.91846 = 1.00.$$

Aplicando el tercer axioma de Kolmogorov, mismo que sostiene que la probabilidad de una consecuencia compuesta por consecuencias mutuamente excluyentes es la suma de sus probabilidades, la probabilidad del evento "C" y la probabilidad del evento "D" se calcularán como sigue:

$$P(C) = P(A \bullet C) + P(B \bullet C)$$

$$P(C) = 0.00154 + 0.07987$$

$$P(C) = 8.141\%$$

$$P(D) = P(A \bullet D) + P(B \bullet D)$$

$$P(D) = 0.00013 + 0.91846$$

$$P(D) = 91.859\%$$

De igual forma, se siguen cumpliendo los dos primeros axiomas de Kolmogorov, ya que:

$$P(C) + P(D) = 100\%$$

$$0.08141 + 0.91859 = 1.00.$$

Con estos valores calculados será posible determinar la probabilidad condicional de que tiemble dado que el aparato predijo un temblor aplicando la expresión expuesta con antelación a este caso:

$$P(A/C) = P(A \bullet C) / P(C)$$

$$P(A/C) = 0.00154 / 0.08141$$

$$P(A/C) = 1.892\%$$

La probabilidad condicional de que tiemble dado que el aparato negó que fuera a temblar será:

$$P(A/D) = P(A \bullet D) / P(D)$$

$$P(A/D) = 0.00013 / 0.91859$$

$$P(A/D) = 0.014\%$$

La probabilidad condicional de que no se presente un sismo dado que el aparato predijo que sí temblaría se calculará de la misma manera, es decir:

$$P(B/C) = P(B \bullet C) / P(C)$$

$$P(B/C) = 0.07987 / 0.08141$$

$$P(B/C) = 98.108\%$$

Por último, la probabilidad condicional de que no haya un sismo dado que la “alarma sísmica” pronosticó que no habrá movimiento telúrico alguno se calculará de manera análoga a las anteriores:

$$P(B/D) = P(B \cdot D) / P(D)$$

$$P(B/D) = 0.91846 / 0.91859$$

$$P(B/D) = 99.986\%.$$

Debe tenerse presente que el evento condicionante es la indicación del aparato, que de hecho es el primer suceso que ocurre cronológicamente hablando, posteriormente se presentará el sismo o no.

Si se efectúa por un lado la sumatoria de P(A/C) y de P(B/C) y, por el otro, la suma de P(A/D) y de P(B/D), se tendrá que:

$$P(A/C) + P(B/C) = 100\%$$

$$0.01892 + 0.98108 = 1.00,$$

$$P(A/D) + P(B/D)$$

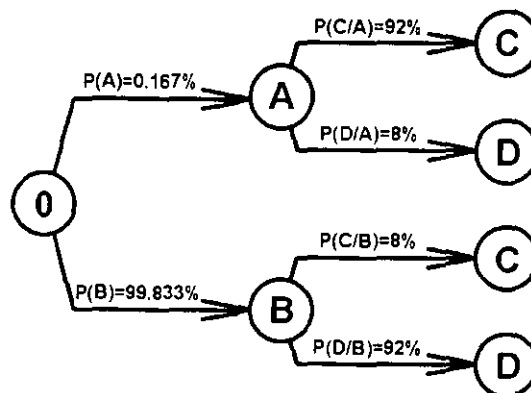
$$0.00014 + 0.99986 = 1.00,$$

por lo que se continúan cumpliendo los dos primeros axiomas de Kolmogorov.

El desarrollo de este proceso puede sintetizarse como la aplicación del teorema de Bayes, el cual afirma que si se conocen las probabilidades *a priori* P(H_i) de un conjunto de tamaño “n” de eventos mutuamente excluyentes y colectivamente exhaustivos; si se conocen además las probabilidades P(α/H_i) donde “α” es un evento que se verifica cuando alguno de los eventos “H_i” ocurre; y se pretenden determinar las probabilidades de que al ocurrir “α” se verifique uno de los “H_i”, esto es, se quieren calcular las probabilidades P(H_i/α) conocidas como probabilidades *a posteriori*, se deberá aplicar la siguiente expresión:

$$P(H_i/\alpha) = [P(H_i) P(\alpha/H_i)] / [\sum_{i=1}^n \{ P(H_i) P(\alpha/H_i) \}].$$

En el ejemplo anterior que fue planteado, el teorema de Bayes se aplicó de la siguiente manera:



- $n=2$ (probabilidades a priori de la ocurrencia del sismo: presencia o ausencia del temblor),
- $P(H_1) = P(A) = 0.167\%$ (probabilidad de que se presente un sismo con periodo de retorno equivalente a 50 años en un mes cualquiera),
- $P(H_2) = P(\bar{B}) = 99.833\%$ (complemento de la probabilidad del evento "A"),
- $P(\alpha/H_1) = P(C/A) = 92\%$ (probabilidad de acierto del aparato),
- $P(\alpha/H_2) = P(C/B) = 8\%$ (probabilidad de fallo del aparato),

por lo tanto:

$$P(H_1/\alpha) = [P(H_1) P(\alpha/H_1)] / [\{ P(H_1) P(\alpha/H_1) \} + \{ P(H_2) P(\alpha/H_2) \}]$$

$$P(H_1/\alpha) = P(A/C) = [(0.00167) (0.92)] / [\{ (0.00167) (0.92) \} + \{ (0.99833) (0.08) \}]$$

$$P(H_1/\alpha) = P(A/C) = 1.892\%$$

$$P(H_2/\alpha) = [P(H_2) P(\alpha/H_2)] / [\{ P(H_1) P(\alpha/H_1) \} + \{ P(H_2) P(\alpha/H_2) \}]$$

$$P(H_2/\alpha) = P(B/C) = [(0.99833) (0.08)] / [\{ (0.00167) (0.92) \} + \{ (0.99833) (0.08) \}]$$

$$P(H_2/\alpha) = P(B/C) = 98.108\%$$

Del mismo modo se aplicó:

- $n=2$ (probabilidades a priori de la ocurrencia del sismo: presencia o ausencia del temblor),
- $P(H_1) = P(A) = 0.167\%$ (probabilidad de que se presente un sismo con periodo de retorno equivalente a 50 años en un mes cualquiera),
- $P(H_2) = P(B) = 99.833\%$ (complemento de la probabilidad del evento "A"),
- $P(\beta/H_1) = P(D/A) = 8\%$ (probabilidad de fallo del aparato),
- $P(\beta/H_2) = P(D/B) = 92\%$ (probabilidad de acierto del aparato),

por lo que:

$$P(H_1/\beta) = [P(H_1) P(\beta/H_1)] / [\{ P(H_1) P(\beta/H_1) \} + \{ P(H_2) P(\beta/H_2) \}]$$

$$P(H_1/\beta) = P(A/D) = [(0.00167) (0.08)] / [\{ (0.00167) (0.08) \} + \{ (0.99833) (0.92) \}]$$

$$P(H_1/\beta) = P(A/D) = 0.014\%$$

$$P(H_2/\beta) = [P(H_2) P(\beta/H_2)] / [\{ P(H_1) P(\beta/H_1) \} + \{ P(H_2) P(\beta/H_2) \}]$$

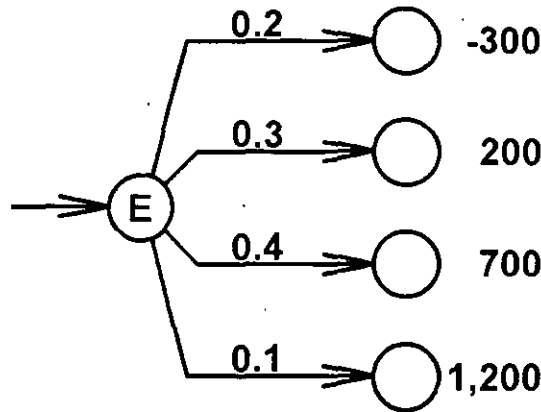
$$P(H_2/\beta) = P(B/D) = [(0.99833) (0.92)] / [\{ (0.00167) (0.08) \} + \{ (0.99833) (0.92) \}]$$

$$P(H_2/\beta) = P(B/D) = 99.986\%$$

En lo sucesivo, un nodo de incertidumbre con "m" ramas deberá ser entendido como la representación gráfica de la lotería: $L[r_1, r_2, r_3, \dots, r_m; p_1, p_2, p_3, \dots, p_m]$; por lo tanto será posible determinar su valor esperado. El valor esperado de un nodo de incertidumbre se define como la suma de los productos que resultan al multiplicar cada consecuencia por su respectiva probabilidad como anteriormente fue definido, es decir:

$$V_E = \sum_{i=1}^m (r_i p_i)$$

El valor esperado de un nodo de incertidumbre es una cantidad equivalente que podrá sustituir al nodo de incertidumbre en cuestión; esta acción servirá para jerarquizar y seleccionar las mejores alternativas que sean expresadas en un árbol de probabilidad o de decisión. Por ejemplo, si en una de las ramas de un árbol de probabilidad se expresó la lotería: $L[-300, 200, 700, 1,200; 0.2, 0.3, 0.40, 0.1]$, y se desea determinar su valor esperado para sustituir el nodo de incertidumbre en dicho árbol, deberá hacerse lo siguiente:



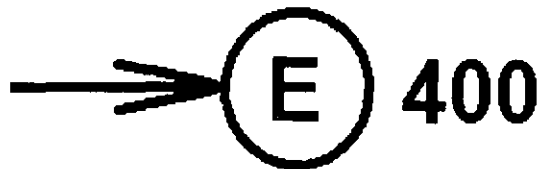
reemplazando las cuatro ramas del nodo "E" de este ejemplo por su valor esperado, el cual se calculará de la siguiente manera:

$$V_E = (-300)(0.20) + (200)(0.30) + (700)(0.40) + (1,200)(0.10)$$

$$V_E = -60 + 60 + 280 + 120$$

$$V_E = 400,$$

por lo que la representación gráfica de esa sección del árbol de probabilidad quedará ahora como sigue:

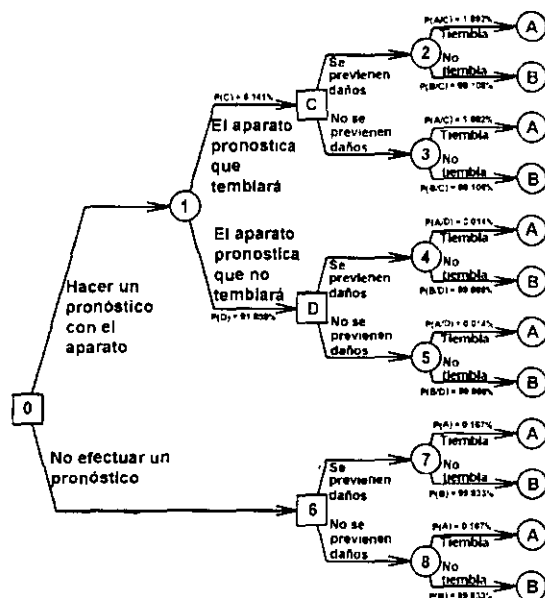


ÁRBOLES DE DECISIÓN

Un árbol de decisión será similar a un árbol de probabilidad. Consiste en un diagrama donde se presentan alternativas seguidas por sus respectivas consecuencias y se compone por nodos de decisión y por nodos de incertidumbre.

Un nodo de decisión se caracteriza por que en él hay la libertad de elegir, mientras que uno de incertidumbre no se sabe si sucederá, situación que ya fue expuesta. Los nodos de decisión se representarán gráficamente mediante cuadrados.

Ejemplo de un árbol de decisión factible para la situación planteada en el punto inmediato anterior (la de la ocurrencia de un sismo y de la "alarma sísmica" que lo puede predecir) es el siguiente:



En el árbol de decisión expresado en el gráfico anterior puede mandarse o no hacer una prueba para que el aparato pronostique si temblará o no; después de conocer el pronóstico se deberá tomar la decisión de llevar a cabo o no acciones para prevenir los daños del temblor (recuérdese que se dispondrá de al menos un mes); y posteriormente se sabrá si el sismo existió o no.

En los nodos de incertidumbre de este árbol se emplearon las probabilidades $P(A)$, $P(B)$, $P(C)$, $P(D)$, $P(A/C)$, $P(B/C)$, $P(A/D)$ y $P(B/D)$, donde el evento "A" representa la ocurrencia de un temblor, el evento "B" la carencia del mismo, el evento "C" el pronóstico del aparato afirmando que temblará y el evento "D" el pronóstico del aparato negando la presencia de un sismo. La nomenclatura "A/C" indica la ocurrencia de un temblor dado que el aparato pronosticó que sí podría temblar, "B/C" expresa la ausencia del sismo dado que el aparato pronosticó que efectivamente temblaría, "A/D" señala la presencia de un temblor dado que el pronóstico del aparato negó que ocurriera y, por último, "B/D" manifiesta que no tembló dado que el aparato dijo en su pronóstico que no se presentaría el sismo.

Evidentemente, cada una de las ramas finales del árbol proporcionará resultados distintos, es decir, diferentes beneficios o perjuicios según sea el caso. El hecho de disponer de la alarma sísmica genera costos de operación y mantenimiento, además de los de adquisición (estos costos dejarán de existir si se desea no hacer pronósticos); sin embargo, el disponer y confiar en este equipo podría hacer que en un momento dado se ahorren recursos para la prevención del desastre y se destinen hacia otros rubros, aunque el estar preparados contra los daños hará que éstos no sean tan cuantiosos como si no se estuviera preparado. En fin, habrá que valorar cada una de las ramas para poder tomar la mejor decisión, bien sea en términos financieros, económicos, sociales o políticos.

El propósito de elaborar un árbol como el anterior es orientar nuestras decisiones hacia la obtención de los mayores y mejores beneficios; o bien, de los menores perjuicios. Esto se puede lograr, en caso de estar llevando a cabo inversiones de recursos, sustituyendo los nodos de incertidumbre por sus respectivos valores esperados como ya fue explicado, y así poder jerarquizar las distintas alternativas a las que el decisor puede acceder al paso del tiempo.