Introducción a la unidad

En esta tema se describen los diferentes tipos de distribuciones de probabilidad que existen, las técnicas para el cálculo o asignación de probabilidades aplicable para cada tipo de dato y cada situación, se analizan sus características y la aplicación de una de ellas en las diferentes situaciones que se presentan en el mundo de los negocios.

Una distribución de probabilidades da toda la gama de valores que pueden ocurrir con base en un experimento, y resulta similar a una distribución de frecuencias. Sin embargo, en vez de describir el pasado, define qué tan probable es que suceda algún evento futuro.

Objetivo particular de la unidad

El estudio de esta unidad te permitirá:
Comprender los conceptos de variable aleatoria discreta y continua.

Conocer algunas de las diferentes distribuciones teóricas de probabilidad que existen asociadas al tipo de variable y reconocerá las situaciones en que son aplicables cada una de ellas.

Desarrollar la habilidad para calcular probabilidades haciendo uso de los modelos de distribución binomial, Poisson, exponencial y normal.

Lo que sé

Según cifras publicadas por **Indexmundi**, la esperanza de vida al nacimiento en México en el año 2007 es de 75.63 años.

En este orden de ideas, es posible que para el año 2050 sea de 80 años. ¿Qué opinas de ello?

Para enviar tu respuesta, pulsa el botón **Editar mi envío**; se mostrará un editor de texto en el cual puedes redactar tu información; una vez que hayas concluido, salva tu actividad pulsando el botón **Guardar cambios**.

Temas de la unidad V

1. Variables aleatorias, discretas y continuas
2. Distribuciones de probabilidad de variable discreta
3. Distribución binomial y distribución de Poisson
4. Distribuciones de probabilidad de variable continúa
5. Distribución normal y distribución exponencial
6. Ley de los grandes números.

Resumen de la unidad

Se define el concepto de variable aleatoria y se señalan sus diferentes tipos. Asimismo, se presentan los rasgos que permiten distinguir algunos modelos de distribución probabilística de variables aleatorias, específicamente los modelos binomial, de Poisson, normal y exponencial y se tipifican los mismos a través de las expresiones analíticas de la función de probabilidad y de densidad, su esperanza matemática, su varianza y sus parámetros. Además en el caso de la distribución normal se presenta el concepto de distribución normal estándar y se muestra el manejo de las tablas respectivas, así como el uso de esta distribución por cuanto aproximación al modelo binomial.

Tema 1. Variables aleatorias, discretas y continuas

Objetivo del tema

Distinguir la presencia de variables aleatorias en un fenómeno o situación dada, definirlas y caracterizarlas según sus rasgos esenciales.

Desarrollo

Una **variable** es **aleatoria** si los valores que toma corresponden a los distintos resultados posibles de un experimento; por ello, el hecho de que tome un valor particular es un evento aleatorio.

La variable aleatoria considera situaciones donde los resultados pueden ser de origen cuantitativo o cualitativo, asignando en cualquier caso un número a cada posible resultado.

Por ejemplo, si el experimento consiste en seleccionar a una persona de un colectivo de n de ellas, y lo que nos interesa es el sexo, la variable aleatoria podría tomar los valores 1 si resulta ser un hombre y 2 si resulta ser una mujer. Si lo que nos interesa es la edad, entonces la variable aleatoria tiene tantos posibles valores como edades haya en la población.

En esencia, lo que hace una variable aleatoria es asignar un número a cada posible resultado del experimento.

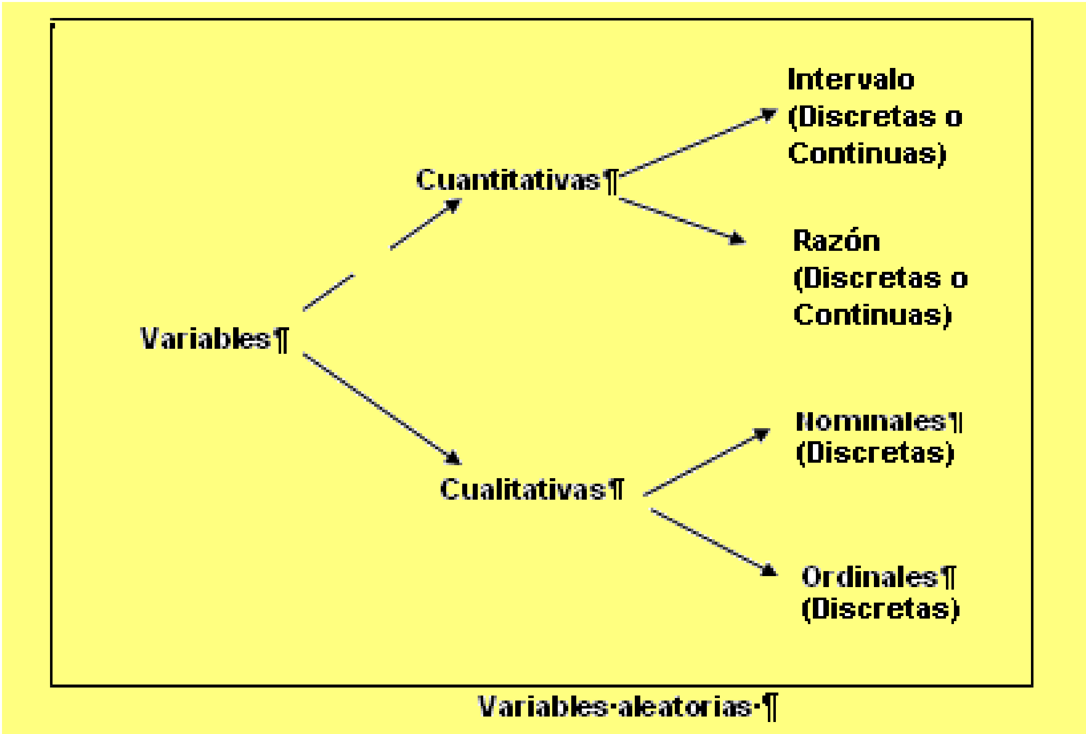
Dependiendo de esta asignación de números las variables aleatorias pueden ser discretas o continuas

- Las **variables discretas o discontinuas** son aquellas que cuantifican la característica de modo tal que el número de posibles resultados se puede contar, esto es, la variable discreta toma un número finito o infinito numerable de posibles valores. Como ejemplo de este tipo de variables tenemos el número de clientes de un banco, el número de hijos de una familia, el número de alumnos en un grupo de la universidad, el número de personas en una población rural, el número de automóviles en una ciudad, etcétera.

- Las **variables continuas** son aquellas que pueden tomar cualquier valor numérico, dentro de un intervalo previamente especificado. Así, por ejemplo, la variable tiempo en una investigación podría medirse en intervalos de horas, o bien, en horas y minutos, o bien en horas, minutos y segundos según sea el requerimiento de la misma.

Desde el punto de vista de la estadística las variables aleatorias también se clasifican de acuerdo a la escala de medición inherente.

Cuando estudiaste el tema de estadística descriptiva tuviste oportunidad de aprender los conceptos de escala nominal, ordinal, de intervalo y de razón. Estas escalas generan precisamente variables aleatorias del mismo nombre. Ocurre que las variables de intervalos y de razón son cuantitativas y pueden ser discretas o continuas. Los casos nominal y ordinal se refieren a cualidades en donde la variable aleatoria al asignar un número a cada resultado asume que tales cualidades son discretas. El cuadro siguiente te proporciona un panorama general de esta situación.



La clasificación de las variables anteriormente expuesta, que parte del punto de vista de la estadística, no es única, pues cada disciplina científica acostumbra hacer alguna denominación para las variables que en ella se manejan comúnmente.

Por ejemplo, en el área de las ciencias sociales es común establecer relaciones entre variables experimentales; por ello, en este campo del conocimiento, las variables se clasifican, desde el punto de vista metodológico, en dependientes e independientes.

La **variable dependiente** es aquella cuyos valores están condicionados por los valores que toma la variable independiente (o las variables independientes) con la que tiene relación.

Por lo tanto, la variable o **las variables independientes** son la causa iniciadora de la acción, es decir, condicionan de acuerdo con sus valores a la variable dependiente.

Ejemplo*:* Consideremos el comportamiento del ahorro de un individuo en una sociedad. El modelo económico que explica su ahorro podría ser:

Ahorro = ingreso – gasto

En este modelo, el ahorro es la variable dependiente y presentará una situación específica de acuerdo con el comportamiento que tengan las variables independientes de la relación.

Un punto importante que debes tener en mente cuando trabajes con variables aleatorias es que no sólo es importante identificarlas y clasificarlas, sino que también deben definirse adecuadamente. Para algunos autores, como Hernández, Fernández y Baptista, su definición deberá establecerse en dos niveles, especificados como nivel conceptual y nivel operacional.

Nivel conceptual*.* Consiste en definir el término o variable con otros términos. Por ejemplo, el término “poder” podría ser definido como “influir más en los demás que lo que éstos influyen en uno”. Este tipo de definición es útil, pero insuficiente para definir una variable debido a que no nos relaciona directamente con la realidad, puesto que, como puede observarse, siguen siendo conceptos.

Nivel operacional. Constituye el conjunto de procedimientos que describen las actividades que un observador realiza para recibir las impresiones sensoriales que indican la existencia de un concepto teórico (conceptual) en mayor o menor grado, es decir, consiste en especificar las actividades u operaciones necesarias que deben realizarse para medir una variable.

Con estas dos definiciones, estás ahora en posibilidad de acotar adecuadamente las variables para un manejo estadístico, de acuerdo con el interés que tengas en ellas, para la realización de un estudio o investigación. Mostraremos a continuación un par de ejemplos de ello.

Ejemplo 1

Variable:	"Ausentismo laboral"
Nivel conceptual:	"El grado en el cual un trabajador no se reportó a trabajar a la hora en la que estaba programado para hacerlo".
Nivel operacional:	"Revisión de las tarjetas de asistencia al trabajo durante el último bimestre".

Ejemplo 2:

Variable:	"Sexo"
Nivel conceptual:	"Condición orgánica que distingue al macho de la hembra".
Nivel operacional:	"Asignación de la condición orgánica: masculino o femenino".

Finalmente, es importante mencionar que a la par que defines una variable aleatoria es importante que le asignes un nombre. Por lo general éste es una letra mayúscula.

ACTIVIDAD 1

Considera 5 situaciones de tu vida cotidiana que consideres dan lugar a experimentos aleatorios, esto es, situaciones en las que no puedes anticipar con toda certeza el resultado.

Establece para cada experimento el conjunto de posibles resultados. Designa la variable aleatoria correspondiente y clasifícala de acuerdo a los criterios que se han revisado en esta unidad.

Realiza tu actividad en un procesador de textos, guárdala en tu computadora y una vez concluida, presiona el botón **Examinar.** Localiza el archivo, ya seleccionado, presiona **Subir este archivo** para guardarlo en la plataforma.

Bibliografía básica

Autor	Capítulo	Páginas
1. Berenson, Levine y Krehbiel.	4. Probabilidad básica y distribuciones de probabilidad. Sección: 4.4 Distribución de probabilidad para una variable aleatoria.	179-186.
2.Hernández, Fernández, Baptista,	6. Formación de hipótesis Sección: Definición conceptual o constitutiva.	145-146
3. Levin y Rubin.	5. Distribuciones de probabilidad Sección: 5.2 Variable aleatoria.	181-187.
4. Lind, Marchal, Mason.	6. Distribuciones de probabilidad discreta. Sección: Variables aleatorias.	194-195

Sitios electrónicos

Sitio	Descripción
http://www.hrc.es/bioest/estadis_21.html	*Variable aleatoria,* hasta la sección Inducción a la probabilidad, publicado por el Hospital Universitario Ramón y Cajal, en donde se presenta el concepto de variable aleatoria y se ilustra con ejemplos en el contexto de la medicina.

Tema 2. Distribuciones de probabilidad discretas

Objetivo del tema

Calcular probabilidades asociadas a variables aleatorias discretas y familiarizarse con los conceptos de esperanza y varianza de las mismas.

Desarrollo

La distribución de probabilidad de una variable aleatoria describe cómo se distribuyen las probabilidades de los diferentes valores de la variable aleatoria. Para una variable aleatoria discreta "**X**", la distribución de probabilidad se describe mediante una **función de probabilidad,** a la que también se conoce como función de densidad, representada por **f(X)**, que define la probabilidad de cada valor de la variable aleatoria.

Como la probabilidad del universo (o evento universal) debe ser igual a 100%, y además cualquier evento que se defina debe estar contenido en el evento universal, cuando hablamos de cómo distribuir las probabilidades nos referimos a cómo es que se reparte este 100% de probabilidad en los diferentes eventos.

Ejemplo 1. Considera el experimento aleatorio que consiste en arrojar un dado dos veces y sumar los resultados de ambas caras. Se desea conocer cuál es la probabilidad de que la suma sea 7.

Solución:

La variable X puede tomar los valores del 2 al 12, inclusive, por lo que se trata por lo que se trata de una variable aleatoria discreta. La siguiente tabla nos permitirá calcular las probabilidades de todos los eventos simples.

Resultado	Segundo dado					
Primer dado	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

En ella vemos que las diagonales, a las que se ha dado diferente color, determinan el mismo valor de la suma para diferentes combinaciones de resultados de cada uno de los dos dados. Por ejemplo, si queremos saber la probabilidad de que la suma sea 7, nos fijaríamos en la diagonal amarilla y observaríamos que hay 6 formas distintas de obtener tal valor, de un total de 36, por lo que la probabilidad es 7/36.

El ejemplo nos permite darnos cuenta además que también podemos calcular fácilmente la probabilidad de que la suma sea menor o igual a 7 y que para ello debemos contar el número de casos que se acumulan desde la diagonal superior izquierda hasta la diagonal amarilla, que corresponden a los valores 2, 3, 4, 5 ,6 y 7. Esto es, se estaría considerando que:

$$P\ (X \leq 7) = P\ (2) + P\ (3) + P\ (4) + P\ (5) + P\ (6) + P\ (7)$$

Para cualquier otro resultado también estaríamos acumulando probabilidades desde la que corresponde al resultado 2 hasta el resultado tope considerado.

De este modo se construye, a partir de la función de probabilidades, otra función, a la que se denomina **función de distribución acumulativa** y que se denota como F(*x*), donde la x indica el valor hasta el cual se acumulan las respectivas probabilidades. Por ejemplo, P ($X \leq 7$) corresponde a F(7).

La tabla siguiente resume la función de probabilidades y la función de distribución acumulativa para el caso del ejemplo

i	Función de probabilidad **P(X = i)**	Función de distribución acumulativa **P(X ≤ i)**
2	1/36	1/36
3	2/36	1/36 + 2/36 = 3/36
4	3/36	3/36 + 3/36 = 6/36
5	4/36	6/36 + 4/36 = 10/36
6	5/36	10/36 + 5/36 = 15/36
7	6/36	15/36 + 6/36 = 21/36
8	5/36	21/36 + 5/36 =26/36
9	4/36	26/36 + 4/36 =30/36
10	3/36	30/36 + 3/36 =33/36
11	2/36	33/36 + 2/36 = 35/36
12	1/36	35/36 +1/36 = 36/36 = 1

Obsérvese que el valor de la función de distribución acumulativa para el último valor de la variable aleatoria acumula precisamente 100%.

Esperanza y varianza

Cuando se trabaja con variables aleatorias, no basta con conocer su distribución de probabilidades. También será importante obtener algunos valores típicos que resuman, de alguna forma, la información contenida en el comportamiento de la variable. De esos valores importan fundamentalmente dos: la esperanza y la varianza.

Esperanza.

Corresponde al valor promedio, considerando que la variable aleatoria toma los distintos valores posibles con probabilidades que no son necesariamente iguales. Por ello se calcula como la suma de los productos de cada posible valor de la variable aleatoria por la probabilidad del respectivo valor. Se le denota como μ

$$\textbf{Esperanza} = \mu = \Sigma\, x\, [P(X=x)]$$

Donde la suma corre para todos los valores x de la variable aleatoria

Varianza

Es el valor esperado o esperanza de las desviaciones cuadráticas con respecto a la media μ. Se denota como σ^2 y se calcula como la suma del producto de cada desviación cuadrática por la probabilidad del respectivo valor.

$$\textbf{Varianza} = \sigma^2 = \Sigma\, (x - \mu)^2\, [P(X=x)]$$

Donde la suma corre para todos los valores x de la variable aleatoria.

Desde luego, la desviación estándar es la raíz cuadrada de la varianza.

Ejemplo 2. Considerando el mismo experimento del ejemplo anterior, determinar la esperanza y varianza de la variable aleatoria respectiva

X	Función de probabilidad P(X = *x*)	*x* P(X = *x*)	$(x - 7)^2$	$(x - 7)^2 P(X = x)$
2	1/36	2/36	25	25/36
3	2/36	6/36	16	32/36
4	3/36	12/36	9	27/36
5	4/36	20/36	4	16/36
6	5/36	30/36	1	5/36
7	6/36	42/36	0	0
8	5/36	40/36	1	5/36
9	4/36	36/36	4	16/36
10	3/36	30/36	9	27/36
11	2/36	22/36	16	32/36
12	1/36	12/36	25	25/36
	Suma	252/36=7		260/36

Podemos decir entonces que al arrojar dos dados y considerar la suma de los puntos que cada uno muestra, el valor promedio será 7 con una desviación estándar de 2.69.

ACTIVIDAD 1

Considera la siguiente situación.

Tres matrimonios, a los que conoceremos como A-B, M-N y P-Q, se han reunido para jugar canasta por una bolsa de $30,000.00. El torneo es de parejas y con este propósito acuerdan que sean A y M quienes en ese orden seleccionen al azar compañero de juego.

La forma en que se determinan las parejas es la siguiente:

- Cada quien, excepto A, escribe su nombre en una papeleta e introduce ésta en una urna.
- La persona cuya papeleta sea seleccionada por A, será la pareja de ésta.
- Si la papeleta seleccionada por A es la de M, entonces M hará pareja con A y ya no extraerá papeleta alguna. En este caso, la segunda papeleta será extraída por P.
- M, o P, según corresponda, extrae la segunda papeleta y de ser la propia, elimina ésta y procede a una nueva extracción.
- La tercera pareja queda automáticamente seleccionada.

Caracterice la variable aleatoria que denota el número de parejas de juego formadas por matrimonios. Tal caracterización debe incluir el nombre de la variable, su tipo (discreto o continuo), su recorrido y su distribución de probabilidades.

Realiza tu actividad en un procesador de textos, guárdala en tu computadora y una vez concluida, presiona el botón **Examinar.** Localiza el archivo, ya seleccionado, presiona **Subir este archivo** para guardarlo en la plataforma.

Bibliografía básica

Autor	Capítulo	Páginas
1. Berenson, Levine y Krehbiel.	4. Probabilidad básica y distribuciones de probabilidad. Sección: 4.4 Distribución de probabilidad para una variable aleatoria.	179-186.
3. Levin y Rubin.	5. Distribuciones de probabilidad Sección: 5.1 ¿Qué es una distribución de probabilidad?	178-181
4. Lind, Marchal, Mason.	6. Distribuciones de probabilidad discreta. Sección:¿Qué es una distribución de probabilidad?	192- 194
	Media, varianza y desviación estándar de una distribución de probabilidad.	195-199.

Sitios electrónicos

Sitio	Descripción
http://www.uam.es/personal_pdi/ciencias/mayter/docencia/sociolog/apuntes.pdf	Rodríguez Mayté, (profesora del curso de Estadística aplicada a las ciencias sociales II de la licenciatura de sociología de la Universidad Autónoma de Madrid), *Variables aleatorias* en el capítulo 3, secciones 3.1 y 3.2, p. 25-27, en el sitio, en donde además de presentar el concepto de variable aleatoria como una función matemática, se introducen los conceptos de variable aleatoria discreta, de función de distribución y de esperanza y varianza de una variable aleatoria discreta, así como algunas de las propiedades matemáticas de la distribución de probabilidades.

Tema 3. Distribución binomial y Distribución Poisson

Objetivo del tema

Caracterizar los modelos de distribución de probabilidades binomial y de Poisson en el caso de variables aleatorias discretas y saber calcular probabilidades haciendo uso de estos modelos.

Desarrollo

Las distribuciones binomial y de Poisson son dos casos de distribuciones de probabilidad de variables aleatorias discretas.

Distribución binomial

La distribución binomial se relaciona con un experimento aleatorio conocido como experimento de Bernoulli el cual tiene las siguientes características:

- El experimento está constituido por un número finito, n, de pruebas idénticas
- Cada prueba tiene exactamente dos resultados posibles. A uno de ellos se le llama arbitrariamente éxito y al otro, fracaso.
- La probabilidad de éxito de cada prueba aislada es constante para todas las pruebas y recibe la denominación de “p”.
- Por medio de la distribución binomial tratamos de encontrar un número dado de éxitos en un número igual o mayor de pruebas.

Puesto que sólo hay dos resultados posibles, la probabilidad de fracaso, a la que podemos denominar q, está dada por la diferencia 1 – p, esto es, corresponde al complemento de la probabilidad de éxito, y como ésta última es constante, entonces también lo es la probabilidad de fracaso.

La probabilidad de “x” éxitos en n intentos está dada por la siguiente expresión:

$$P(x)= nCx\ p^{x}q^{(n-x)}.$$

Esta fórmula nos dice que la probabilidad de obtener “x” número de éxitos en n pruebas (como ya se indicó arriba) está dada por la multiplicación de n combinaciones en grupos de x (el alumno debe recordar el tema de reglas de conteo) por la probabilidad de éxito elevada al número de éxitos deseado y por la probabilidad de fracaso elevada al número de fracasos deseados.

*Con el término **combinaciones** nos referimos al número de formas en que podemos extraer grupos de k objetos tomados de una colección de n de ellos (n ≥ k), considerando que el orden en que se toman o seleccionan no establece diferencia alguna. El símbolo nCk denota el número de tales combinaciones y se lee combinaciones de n objetos tomados en grupos de k. Operativamente,*

$$nCk = n! / [x!\ (n-x)!]$$

donde el símbolo ! indica el factorial del número, de modo que n! = (1)(2)(3)…(n)

A continuación se ofrecen varios ejemplos que nos ayudarán a comprender el uso de esta distribución.

Ejemplo 1. Un embarque de veinte televisores incluye tres unidades defectuosas. Si se inspeccionan tres televisores al azar, indique usted cuál es la probabilidad de que se encuentren dos defectuosos.

Solución:

Podemos verificar si se trata de una distribución binomial checando cada uno de los puntos que caracterizan a esta distribución mediante una lista

Aspecto	Estatus	Observación
Hay un número finito de ensayos	**SI**	Cada televisor es un ensayo y hay 3 de ellos
Cada ensayo tiene sólo dos resultados	**SI**	Cada televisor puede estar defectuoso o no
La probabilidad de éxito es constante	**SI**	La probabilidad de que la unidad esté defectuosa es 3/20
Se desea saber la probabilidad de un cierto número de éxitos	**SI**	Se desea saber la probabilidad de que X=2

Una vez que hemos confirmado que se trata de una distribución binomial aplicamos la expresión $P(x)= nCx\ p^xq^{(n-x)}$, de modo que

$$P(2) = {}_3C_2\ (3/20)^2(17/20)^1 = 3\ (0.0225)(0.85) = 0.057375$$

Ejemplo 2. Una pareja de recién casados planea tener tres hijos. Diga usted cuál es la probabilidad de que los tres hijos sean varones si consideramos que la probabilidad de que el descendiente sea hombre o mujer es igual.

Solución: Verificamos primero si se cumplen los puntos que caracterizan la distribución binomial.

Claramente, es un experimento aleatorio con tres ensayos, y en todos ellos sólo hay dos resultados posibles, cada uno con probabilidad de 0.5 en cada ensayo. Si se define como éxito que el sexo sea masculino, entonces podemos decir que se desea saber la probabilidad de que haya tres éxitos.

Entonces, el experimento lleva a una distribución binomial y

$$P(3) = {}_3C_3\,(1/2)^3(1/2)^0 = (1/2)^3 = 0.0125$$

Ejemplo 3. Se sabe que el 30% de los estudiantes de secundaria en México es incapaz de localizar en un mapa el lugar donde se encuentra Afganistán. Si se entrevista a seis estudiantes de este nivel elegidos al azar:

a) ¿Cuál será la probabilidad de que exactamente dos puedan localizar este país?
b) ¿Cuál será la probabilidad de que un máximo de dos puedan localizar este país?

Solución:

Al igual que en los casos anteriores verificamos si se cumple o no que el experimento lleva a una distribución binomial.

Se trata de un experimento con seis ensayos, en cada uno de los cuales puede ocurrir que el estudiante sepa o no sepa localizar Afganistán en el mapa. Si se define como éxito que sí sepa la localización podemos decir que la probabilidad de éxito es de 0.30. Además, las probabilidades que se desea calcular se refieren al número de éxitos. Concluimos que el experimento es Bernoulli y por lo tanto,

$$P(2) = {}_6C_2\,(0.30)^2(0.70)^4 = 15\,(0.09)(0.2401) = 0.324135$$

Por cuanto hace al inciso (b), la frase « un máximo de dos» significa que X toma los valores cero, uno o dos. Entonces,

$$P(X\leq 2) = P(2) + P(1) + P(0) =$$

$= {}_6C_2\ (0.30)^2(0.70)^4 + {}_6C_1\ (0.30)^1(0.70)^5 + {}_6C_0\ (0.30)^0(0.70)^6 =$

$= 15(0.09)(0.2401) + 6(0.30)(0.16807) + 0.1176 =$

$= 0.661941$

Esperanza y varianza de una variable aleatoria binomial

Consideremos de nueva cuenta el ejemplo 1.

¿Qué pasa con las probabilidades de los otros valores posibles para la variable aleatoria?. Si hacemos los cálculos respectivos tendríamos:

$P(0) = {}_3C_0\ (3/20)^0(17/20)^3 = \quad 0.614125$

$P(1) = {}_3C_1\ (3/20)^1(17/20)^2 = 3\ (0.15)(0.7225) = 0.325125$

$P(3) = {}_3C_3\ (3/20)^3(17/20)^0 = (0.003375) = 0.003375$

Si recordamos que P (2) = 0.057375, entonces podemos confirmar que P(0) + P(1) + P(2) + P(3) = 1.00, lo que era de esperarse puesto que los valores 0, 1, 2 y3 constituyen el universo en el experimento en cuestión.

Con estos valores podemos determinar la esperanza y varianza de la variable aleatoria considerada. Para ello nos es útil acomodar los datos en una tabla recordando que:

$$\mu = \Sigma\, x\, [P(X=x)],\ \text{y que},\ \sigma^2 = \Sigma\, (x - \mu)^2\, [P(X=x)]$$

x	Función de probabilidad P(X = *x*)	*x* P(X = *x*)	$(x - 0.45)^2$	$(x - 0.45)^2$ P(X = x)
0	0.614125	0.000000	0.2025	0.124360
1	0.325125	0.325125	0.3025	0.098350
2	0.057375	0.114750	2.4025	0.137843
3	0.003375	0.010125	6.5025	0.021946
	Suma	0.450000		0.382500

Entonces, la esperanza es 0.45 y la varianza 0.3825.

Si interpretamos las probabilidades anteriores en un sentido frecuentista, diríamos que si consideramos un número grande de realizaciones del experimento, por ejemplo un millón de veces, en aproximadamente 614 125 realizaciones tendremos refrigeradores sin defecto, en 325125 veces encontraremos un refrigerador con defecto, en otras 57375 ocasiones encontraremos dos refrigeradores con defecto y en 3375 veces los tres refrigeradores estarían defectuosos.

Con estos datos podemos elaborar una tabla de distribución de frecuencias y calcular el promedio de refrigeradores defectuosos

Número de refrigeradores defectuosos (x)	Frecuencia (f)	fm
0	614125	0
1	325125	325125
2	57375	114750
3	3375	10125
Total	1000000	450000

Luego,

μ = 450 000 / 1 000 000 = 0.45

Asimismo, podemos calcular la varianza:

σ^2 = [614125 $(0 - 0.45)^2$ + 325125 $(1 - 0.45)^2$ + 57375 $(2 – 0.45)^2$ + 3375 (3 – 0.45)2] / 100

= (124360.313 + 98350.3125 + 137843.438 + 29945.9375) / 100

= 0.3825

Observa que hemos seguido fielmente las lecciones de estadística descriptiva en el cálculo de μ y σ y que hemos llegado a los mismos valores que ya habíamos obtenido. Esto nos proporciona por lo menos un esquema con el cual podemos interpretar la esperanza y varianza, haciendo uso del concepto de frecuencias.

Es importante además, darse cuenta que podemos llegar a estos mismos valores de un modo más sencillo si nos percatamos que:

$\mu = 0.45$ es precisamente el resultado que se obtiene al multiplicar el número de ensayos por la probabilidad de éxito, esto es, 3(0.15)

$\sigma^2 = 0.3825$ es precisamente el resultado que se obtiene al multiplicar el número de ensayos por la probabilidad de éxito y por la de fracaso, esto es, 3(0.15)(0.85)

En otras palabras,

Media y varianza de una variable aleatoria binomial

$\mu = np \qquad \sigma^2 = npq$

Puede ocurrir, como en el caso del ejemplo anterior, que la esperanza da un valor que no coincide con los valores posibles de la variable aleatoria. Por eso se dice que la esperanza es un valor ideal.

Por otra parte, si desglosamos cada uno de los elementos que integran la expresión del cálculo de probabilidades de la distribución binomial y consideramos las expresiones para el cálculo de la media y la varianza, tendremos que:

$nCx = n! / [x! (n-x)!]$

$p^x = p^x$

$q^{n-x} = (1-p)^{n-x}$

Media= np

Varianza = $n\,p\,(1 - p)$

Lo que nos revela que para poder calcular cualquier probabilidad con el modelo binomial o su esperanza o varianza debemos conocer los valores de *n*, el número total de ensayos, y de *p*, la probabilidad de éxito. El valor de *x*, el número de éxitos se establece de acuerdo con las necesidades del problema.

Lo anterior nos permite concluir que la distribución binomial queda completamente caracterizada cuando conocemos los valores de *n* y *p*. Por esta razón a estos valores se les conoce como los **parámetros de la distribución**.

Un error que suele cometerse a propósito de la distribución binomial es considerar que sus parámetros son la esperanza y varianza de la variable aleatoria respectiva. En realidad estos dos valores se expresan en función de los parámetros.

El siguiente ejemplo nos ayudará a entender este concepto.

Ejemplo 1*:* De acuerdo con estudios realizados en un pequeño poblado, el 20% de la población tiene parásitos intestinales. Si se toma una muestra de 1,400 personas, ¿cuántos esperamos que tengan parásitos intestinales?

$$\mu = np = 1,400 \times 0.20 = 280$$

Éste es el número promedio de elementos de la muestra que tendría ese problema.

Usando el teorema de Tchebyshev podríamos considerar que el valor real estaría a dos desviaciones estándar con un 75% de probabilidades y a tres con un 89%. De acuerdo con ello obtenemos la desviación estándar y posteriormente determinamos los intervalos.

Desviación estándar: $\sigma = \sqrt{npq} = \sqrt{1,400 \times 0.20 \times 0.80} = \sqrt{224} = 14.97$

La media más menos dos desviaciones estándar nos daría el intervalo de 250 a 310 personas que tienen problemas.

La media más menos tres desviaciones estándar nos daría el intervalo de 235 a 325 personas que podrían tener problemas.

Teorema de Tchebyshev

El teorema de Tchebyshev señala que la probabilidad de que una variable aleatoria tome un valor contenido en k desviaciones estándar de la media es cuando menos $1 – 1 / k^2$

Distribución de Poisson

Es otra distribución teórica de probabilidad que es discreta y tiene muchos usos en economía y comercio. Se debe al teórico francés Simeón Poisson quien la derivó en 1837como un caso especial (límite) de la distribución binomial.

Se puede utilizar para determinar la probabilidad de un número designado de éxitos cuando los eventos ocurren en un espectro continuo de tiempo y espacio. Es semejante al proceso de Bernoulli,excepto que los eventos ocurren en un espectro continuo, de manera que al contrario del modelo binomial, se tiene un número infinito de ensayos.

Como ejemplo tenemos el número de llamadas de entrada a un conmutador en un tiempo determinado, o el número de defectos en 10 m^2. de tela.
Sólo se requiere conocer el número promedio de éxitos para la dimensión específica de tiempo o espacio de interés.

Este número promedio se representa generalmente por λ (lambda) y la fórmula de una distribución de Poisson es la siguiente:

$$P\left(x/\lambda\right)=\frac{\lambda^{x}\cdot e^{-\lambda}}{x!}$$

En esta fórmula, x representa el número de éxitos cuya probabilidad deseamos calcular; λ. es el promedio de éxitos en un periodo de tiempo o en un cierto espacio; "e" es la base de los logaritmos naturales; y el símbolo de admiración representa el factorial del número que se trate.

Ejemplo 1: El manuscrito de un texto de estudio tiene un total de 40 errores en las 400 páginas de material. Los errores están distribuidos aleatoriamente a lo largo del texto. Calcular la probabilidad de que:

a) Un capítulo de 25 páginas tenga dos errores exactamente.
b) Un capítulo de 40 páginas tenga más de dos errores.
c) Una página seleccionada aleatoriamente no tenga errores.

Solución:

En cada caso debemos establecer primero el número promedio de errores. En el inciso (a) nos referiremos al número promedio por cada 25 páginas, en el inciso (b) por cada 40 páginas y en el (c) por página. Esto lo podemos hacer mediante el procedimiento de proporcionalidad directa o regla de tres.

a) Dos errores en 25 páginas

Datos

$$40 - 400$$
$$\lambda - \quad 25$$
$$\therefore \quad \lambda = 2.5$$
$$x = 2$$
$$e = 2.71828$$

$$p(2/2.5) = \frac{2.5^2 \cdot 2.71828^{-2.5}}{2!} = 0.256 = 25.6\%$$

Existe un 25.6% de probabilidad de que un capítulo de 25 páginas tenga exactamente dos errores.

b) Más de dos errores en 40 páginas

$$40 - 400$$
$$\lambda - \quad 40$$
$$\therefore \quad \lambda = 4$$
$$x = 2$$
$$e = 2.71828$$

$$p(0/4) = \frac{4^0 \cdot 2.71828^{-4}}{0!} = 0.018 = 1.8\%$$

$$P(1/4) = \frac{4^1 \cdot 2.71828^{-4}}{1!} = 0.073 = 7.3\%$$

$$P(2/4) = \frac{4^2 \cdot 2.71828^{-4}}{2!} = 0.146 = 14.6\%$$

$$\therefore P(> 2/4) = 1 - (0.018 + 7.3 + 14.6) = 0.762 = 76.2\%$$

Existe un 76.2% de probabilidad de que un capítulo de 40 páginas tenga más de dos errores.

c) Una página no tenga errores:

Datos

$$40 - 400$$
$$\lambda - \quad 1$$
$$\therefore \quad \lambda = 0.10$$
$$x = 2$$
$$e = 2.71828$$

$$p\left(0/0.10\right) = \frac{0.10^{0} \cdot 2.71828^{-0.10}}{0!} = 0.905 = 90.5\%$$

Existe un 90.5% de probabilidad de que una sola página seleccionada aleatoriamente no tenga errores.

Un aspecto importante de la distribución de Poisson es que su media y varianza son iguales. De hecho,

Media y Varianza de una variable aleatoria Poisson
$\mu = \lambda \qquad \sigma^2 = \lambda$

De acuerdo con lo anterior, para determinar las probabilidades en un modelo de Poisson o calcular su esperanza o varianza debemos conocer el valor de λ, esto es, del número promedio de éxitos. Éste es el parámetro de la distribución.

La distribución de Poisson como una aproximación a la distribución binomial.

En un experimento de Bernoulli, tal como los que acabamos de estudiar en la distribución binomial, puede suceder que el número de ensayos sea muy grande y/o que la probabilidad de acierto sea muy pequeña y los cálculos se vuelven muy laboriosos. En estas circunstancias, podemos usar la distribución de Poisson como una aproximación a la distribución binomial.

Ejemplo 2: Una fábrica recibe un embarque de 1, 000,000 de rondanas. Se sabe que la probabilidad de tener una rondana defectuosa es de .001. Si obtenemos una muestra de 3000 rondanas, ¿cuál será la probabilidad de encontrar un máximo de tres defectuosas?

Solución*:*

Este ejemplo, desde el punto de vista de su estructura, corresponde a una distribución binomial. Dados los volúmenes y probabilidades que se manejan es conveniente trabajar con la distribución Poisson, tal como se realiza a continuación. Debemos recordar que un máximo de tres defectuosas incluye la probabilidad de encontrar una, dos y tres piezas defectuosas o ninguna

Media: $$\mu = np = 3,000 \times 0.001 = 3$$

$$P\left(x/\mu\right) = \frac{\mu^{x} \cdot e^{-\mu}}{x!}$$

$$P\left(0/3\right) = \frac{3^{0} \cdot 2.71828^{-3}}{0!} = 0.0498 = 5.0\%$$

$$P\left(1/3\right) = \frac{3^{1} \cdot 2.71828^{-3}}{1!} = 0.149 = 14.9\%$$

$$P\left(2/3\right)=\frac{3^{2}\cdot 2.71828^{-3}}{3!}=0.224=22.4\%$$

$$P\left(3/3\right)=\frac{3^{3}\cdot 2.71828^{-3}}{3!}=0.224=22.4\%$$

La probabilidad de encontrar un máximo de tres piezas defectuosas está dado por la suma de las probabilidades arriba calculadas, es decir: 0.647 ó 64.7% aproximadamente.

ACTIVIDAD 1

Considera la siguiente situación.

En el puerto de Balankub hay una sociedad cooperativa de taxis que proporciona servicio desde varias bases a cualquier destino. En la cooperativa desean determinar el número de unidades que deben tener en promedio en la base del aeropuerto. Saben que todos los martes llegan, en el mismo vuelo, cuatro gerentes ejecutivos de cuatro diferentes empresas. Cada uno puede escoger, de manera independiente, ir a la terminal de autobuses si su destino final es A1, ir a la terminal del tren si su destino final es A2 o ir a la terminal del transbordador si su destino final es A3. Si en cierto mes hay cinco días martes, ¿cuál es la probabilidad de que por lo menos en tres de ellos los cuatro ejecutivos escojan el mismo destino?

Para enviar tu respuesta, pulsa el botón **Editar mi envío**; se mostrará un editor de texto en el cual puedes redactar tu información; una vez que hayas concluido, salva tu actividad pulsando el botón **Guardar cambios**.

ACTIVIDAD 2

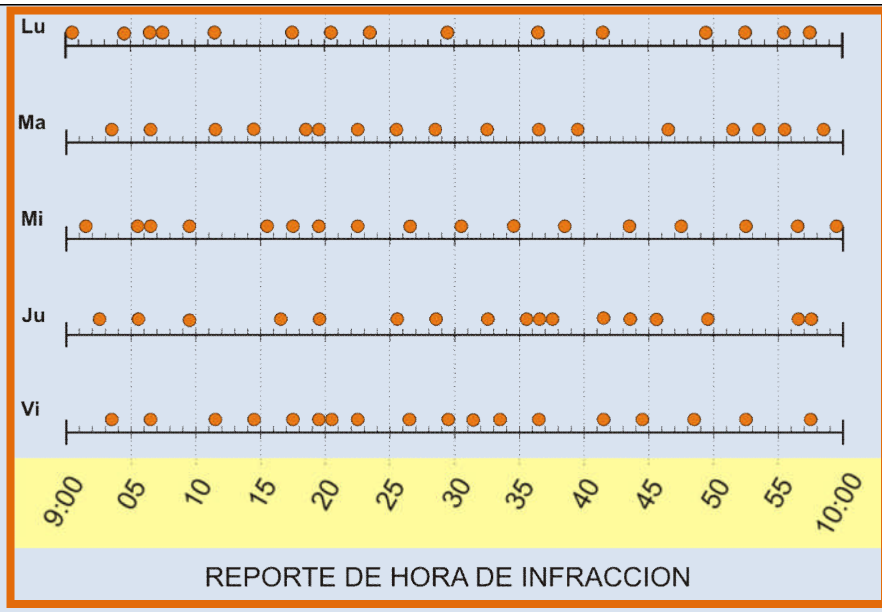
Resuelve los siguientes problemas.

1. En un corporativo con 500 empleados se llevó a cabo una auditoria preliminar de documentos en el área de recursos humanos. Se detectó que en 8 de cada 30 expedientes falta el documento A, que en 6 de cada 24 expedientes falta el documento B y que en uno de cada 50 falta el documento C. Se considera como omisión grave que falte cualquiera de los tres documentos. Se desea saber cuál es la probabilidad de que en 400 expedientes no se detecte omisión alguna.
¿Qué modelo de distribución probabilística aplicarías?. Establece los parámetros del mismo y expresa la relación algebraica que permitiría calcular la probabilidad señalada.

Si consideras que hay dos modelos alternativos, justifica ambos.

2. La carretera que comunica las poblaciones de San Albano y San Miguel tiene un tramo recto de 4.2 kms en el que con frecuencia se registran accidentes por exceso de velocidad, por lo que las autoridades han decidido colocar una cámara-radar de velocidad que envía la información a la computadora de la oficina de tránsito, en la que se registra la hora, la velocidad y número de placa del vehículo. En aquéllos casos en que la velocidad excede el límite establecido se emite la multa correspondiente. Además, se genera un reporte en forma de cinta para mostrar la hora de la infracción

En la figura se muestran tales reportes para los últimos cinco días hábiles entre las 9:00 y las 10:00 am.



a. Se desea determinar la probabilidad de que en un lapso de cinco minutos

- k vehículos excedan la velocidad, con k=0, 1, 2, 3, 4 y 5
- Como máximo tres vehículos excedan la velocidad
- Como mínimo tres vehículos excedan la velocidad

b. ¿Cuál es el valor esperado de vehículos que exceden la velocidad en un lapso de cinco minutos?

c. ¿Cuál es el valor esperado de vehículos que exceden la velocidad en un lapso de una hora?

Realiza tu actividad en un procesador de textos, guárdala en tu computadora y una vez concluida, presiona el botón **Examinar.** Localiza el archivo, ya seleccionado, presiona **Subir este archivo** para guardarlo en la plataforma

Bibliografía básica

Autor	Capítulo	Páginas
1. Berenson, Levine y Krehbiel.	4. Probabilidad básica y distribuciones de probabilidad. Sección: 4.5 Distribución binomial.	186-194.
	4.6 Distribución de Poisson,	194- 197.
3. Levin y Rubin.	5. Distribuciones de probabilidad Sección: 5.4 La distribución binomial.	191- 202
	5.5 La distribución de Poisson.	202- 208
4. Lind, Marchal, Mason.	6. Distribuciones de probabilidad discreta. Sección: Distribución de probabilidad binomial.	200-210
	Distribución de probabilidad de Poisson.	214-217

Sitios electrónicos

Sitio	Descripción
http://www.uam.es/personal_pdi/ciencias/mayter/docencia/sociologia/apuntes.pdf	Rodríguez, Mayté (profesora del curso de Estadística aplicada a las ciencias sociales II, de la licenciatura en sociología de la Universidad Autónoma de Madrid), *Variables aleatorias discretas* en el capítulo 4, secciones 4.1, 4.2 y 4.4, págs. 34 a 41, en donde se caracterizan los modelos de distribución de probabilidad para variables discretas de Bernoulli, Binomial y de Poisson, estableciendo en cada caso, la función de densidad, los parámetros y la esperanza y varianza respectiva; además se presentan ejemplos resueltos y se proponen ejercicios.
http://www.uaq.mx/matematicas/estadisticas/xu4-5.html#t1	Víctor Larios Osorio, (profesor de la Universidad Autónoma de Querétaro) *Unidad 5. Distribuciones de probabilidad, parte 3, la distribución binomial*, del hipertexto Estadística, en donde además de repasar las características generales de las distribuciones de probabilidad para variables discretas, se listan algunos modelos de distribuciones discretas y se desarrolla formalmente, a través del concepto de ensayo Bernoulli, el modelo binomial, incluyendo su esperanza y varianza.

Tema 4. Distribuciones de probabilidad de variable continua

Objetivo del tema

Establecer las preguntas viables que pueden formularse respecto de valores de probabilidad en el caso de variables aleatorias continuas y distinguir los conceptos de función de densidad y función de distribución acumulativa.

Desarrollo

Para comprender la diferencia entre las variables aleatorias discretas y las continuas recordemos que las variables aleatorias continuas pueden asumir cualquier valor dentro de un intervalo de la recta numérica o de un conjunto de intervalos.

Como cualquier intervalo contiene una cantidad infinita de valores, no es posible hablar de la probabilidad de que la variable aleatoria tome un determinado valor; en lugar de ello, debemos pensar en términos de la probabilidad de que una variable aleatoria continua tome un valor dentro de un intervalo dado.

Esto significa que si X es una variable aleatoria continua, entonces por definición $P(X = x) = 0$, cualquiera que sea el valor de x.

Las preguntas de interés tomarán entonces alguna de las siguientes formas básicas:

- $P(X \leq a)$
- $P(X \geq b)$
- $P(c \leq X \leq d)$

donde a, b c y d son números reales

Aquí debe observarse que $P(X \leq a) = P(X < a)$, ya que como se ha hecho notar, $P(X = a) = 0$

Para describir las distribuciones discretas de probabilidades retomamos el concepto de una función de probabilidad **f(x).** Recordemos que en el caso discreto, esta función da la probabilidad de que la variable aleatoria “**x**” tome un valor específico. En el caso continuo, la contraparte de la función de probabilidad recibe el nombre de **función de densidad de probabilidad** que también se representa por **f(x).** Para una variable aleatoria continua, la función de densidad de probabilidad especifica el valor de la función en cualquier valor particular de “**x**” sin dar como resultado directo la probabilidad de que la variable aleatoria tome un valor específico.

Para comprender esto, imagínese que se tiene una variable aleatoria continua relativa a un fenómeno que puede repetirse un número muy grande de veces y que los datos se arreglan en una tabla de distribución de frecuencias con la característica especial de que los intervalos se definen de manera que sean muy finos. A continuación se graficarían los datos de la distribución formando en primera instancia un histograma, luego un polígono de frecuencias y de aquí, como paso subsecuente, una curva suavizada. Al tratar de determinar la probabilidad de que la variable tome valores en un intervalo dado se observaría que en el límite, esto es, entre más finos sean los intervalos, tal probabilidad está dada por el área bajo la curva.

La curva suavizada sería la función de densidad de probabilidad, f(x), de modo que el área entre esta curva y el eje X da la probabilidad. Esto lleva a hacer uso del cálculo integral ya que el área está dada por:

$$P(X \leq a) = \int^{a} f(x)\,dx$$

donde el símbolo $\int$ denota el proceso de integración.

Los valores que se obtienen de P(X ≤ a) para todos los valores posibles a, constituyen la **función de distribución acumulativa** de la variable aleatoria X, misma que se denota como $\mathbf{F_X}$.

Debe ocurrir, para que F_X sea realmente una función de distribución de probabilidades, que $F_X(\infty) = 1$.

En principio parece complicado el manejo de las funciones de distribución de probabilidades en el caso continuo, particularmente si no se manejan las herramientas del cálculo integral. Sin embargo, en el terreno de la contaduría y la administración muchos de los problemas que habrán de enfrentarse hacen referencia a distribuciones de probabilidad muy conocidas, y por lo mismo distribuciones sobre las que se ha trabajado mucho, a grado tal que los valores de la probabilidad están ya tabulados. Uno de estas distribuciones es la distribución normal.

ACTIVIDAD 1

Se ha diseñado un sistema de alerta que emite una señal auditiva para avisar a dependientes invidentes cuando alguien ingresa a un establecimiento comercial. La duración de la señal es una variable aleatoria y se cree que la función de densidad respectiva es $f_X(x) = c$, en el intervalo (0.2, 0.8). Se desea saber cuál debe ser el valor de c para que la función sea efectivamente una función de densidad.

Sugerencia:

Recuerda que para que una función sea de densidad, el área bajo la curva en todo el recorrido de la variable aleatoria debe ser igual a uno. Asigna un valor arbitrario a c (por ejemplo, c= 1) y elabora la grafica respectiva de f(x) = c. Observa la forma de la figura que se genera y establece de qué manera puedes calcularla. Luego, determina cuál debería ser el valor de c para que el área sea uno.

Realiza tu actividad en un procesador de textos, guárdala en tu computadora y una vez concluida, presiona el botón **Examinar.** Localiza el archivo, ya seleccionado, presiona **Subir este archivo** para guardarlo en la plataforma.

Bibliografía básica

Autor	Capítulo	Páginas
1. Berenson, Levine y Krehbiel.	4. Probabilidad básica y distribuciones de probabilidad. Sección: 4.7 Distribución normal.	198- 219.
3. Levin y Rubin.	5. Distribuciones de probabilidad Sección: 5.6 La distribución normal: distribución de una variable aleatoria continua.	209-222
4. Lind, Marchal, Mason.	7. Distribuciones de probabilidad normal. Sección: La familia de distribuciones de probabilidad normal.	227- 229

Sitios electrónicos

Sitio	Descripción
http://www.uaq.mx/matematicas/estadisticas/xu4-5.html#t1	Víctor Larios Osorio, (profesor de la Universidad Autónoma de Querétaro) *Unidad 5. Distribuciones de probabilidad parte 4, distribuciones de probabilidad para variables continuas*, del hipertexto Estadística, en donde además de presentar las características generales de las distribuciones de probabilidad para variables aleatorias continuas, se ilustran los conceptos de función de densidad y de distribución mediante graficaciones animadas; asimismo, se listan algunos modelos de distribuciones continuas.

Tema 5. Distribución normal y distribución exponencial

Objetivo del tema

Caracterizar los modelos de distribución de probabilidades normal y exponencial en el caso de variables aleatorias continuas, así como calcular probabilidades haciendo uso de estos modelos.

Desarrollo

Distribución normal

Esta distribución de probabilidad también es conocida como "Campana de Gauss" por la forma que tiene su gráfica y en honor del matemático que la desarrolló. Tal vez dé la impresión al alumno de ser un tanto complicada, pero no debe preocuparse por ello, pues para efectos del curso, no es necesario usarla de manera analítica, sino comprender intuitivamente su significado.

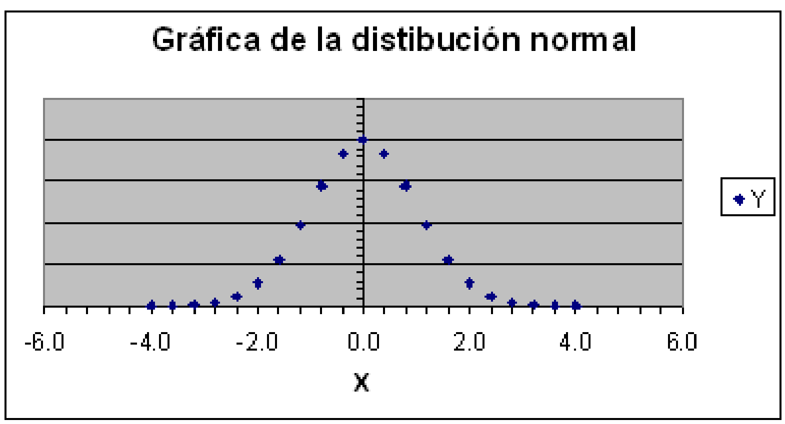
De cualquier manera se dará una breve explicación de la misma para efectos de una mejor comprensión del tema. Su función aparece a continuación.

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

En esta función todos los términos son conocidos por el estudiante. La "y" es la ordenada de las coordenadas rectangulares cartesianas que el alumno ya conoce y que representa la altura sobre el eje "x"; x es la abscisa en este sistema de coordenadas; π = 3.14159; "e" corresponde a la base de los logaritmos naturales que el estudiante ya tuvo ocasión de utilizar en la distribución de Poisson. Los símbolos μ y σ corresponden a la media y a la desviación estándar.

Podemos decir que ésta es la expresión de la ecuación normal, de la misma manera que y=mx+b es la expresión de la ecuación de la recta (en su forma cartesiana), por lo que así como podemos asignar distintos valores a m (la pendiente) y b (la ordenada al origen), para obtener una ecuación particular (p. ej. y=4x+2), de la misma manera podemos sustituir μ y σ por cualquier par de valores para obtener un caso particular de la función normal. Si lo hacemos de esa manera, por ejemplo, dándole a la media un valor de cero y a la desviación estándar un valor de 1, podemos ir asignando distintos valores a "x" (en el rango de –4 a 4, por ejemplo) para calcular los valores de "y". Una vez que se ha completado la tabla es fácil graficar en el plano cartesiano. Obtendremos una curva de forma acampanada. A continuación se muestran tanto los puntos como la gráfica para estos valores.

X	Y
-4.0	0.00013
-3.6	0.00061
-3.2	0.00238
-2.8	0.00792
-2.4	0.02239
-2.0	0.05399
-1.6	0.11092
-1.2	0.19419
-0.8	0.28969



Curva normal

-0.4	0.36827
0.0	0.39894
0.4	0.36827
0.8	0.28969
1.2	0.19419
1.6	0.11092
2.0	0.05399
2.4	0.02239
2.8	0.00792
3.2	0.00238
3.6	0.00061
4.0	0.00013

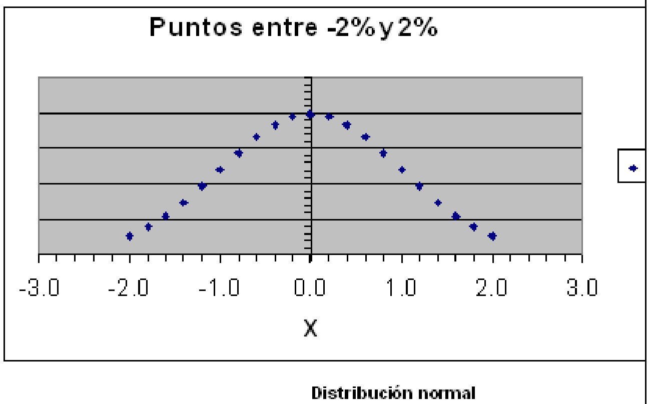
Es importante mencionar que el área que se encuentra entre la curva y el eje de las abscisas es igual a la unidad ó 100%. La curva normal es simétrica en relación con la media. Esto quiere decir que la parte de la curva que se encuentra a la derecha de la curva es como una imagen reflejada en un espejo de la parte que se encuentra a la izquierda de la misma. Esto es importante, pues el área que se encuentra a la izquierda de la media es igual a la que se encuentra a la derecha de la misma y ambas son iguales a 0.5 o el 50%.

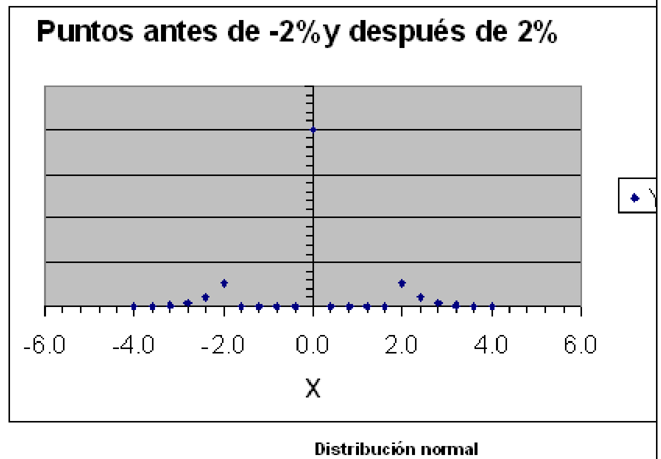
Para trabajar con la distribución normal debemos unir los conceptos de área bajo la curva y de probabilidad. La probabilidad de un evento es proporcional al área bajo la curva normal que cubre ese mismo evento. Un ejemplo nos ayudará a entender estos conceptos.

Con base en la Figura vamos a suponer “Curva normal”, vamos a suponer que el rendimiento de las acciones en la Bolsa de Valores en un mes determinado tuvo una media de 0% con una desviación estándar de 1%. (Esto se asimila a lo dicho sobre nuestra gráfica de una distribución normal con una media de cero y una desviación estándar de 1). De acuerdo con esta información, es mucho más probable encontrar acciones que ganen entre –2% y 2%, que acciones que ganen más o menos (ver las siguientes gráficas).

El valor exacto de las áreas y, por tanto, de las probabilidades de ocurrencia requeriría de herramientas de cálculo integral, pues se trata de calcular áreas bajo las curvas. Dado que se necesitaría calcular una integral distinta para cada pareja de valores de la media y de la desviación estándar, el trabajo sería muy laborioso. Afortunadamente existe una alternativa, que consiste en el uso de una tabla que contiene las áreas bajo la curva normal y que se encuentra en la mayoría de los libros de estadística.

Al final de esta sección se encuentra una tabla de la distribución normal para el caso en que se tiene una media igual a 0 y una desviación estándar igual a 1, a la que se conoce como **distribución normal estándar**. En los próximos párrafos aprenderemos a utilizar la tabla de la distribución normal estándar.





La figura "Puntos menores a -2% y 2%", nos muestra el **área** que hay entre los valores -2 y 2 y además nos enseña que al ser la campana simétrica, tal área es el doble de la que hay entre los valores 0 y 2. En general, el área entre los valores –z y z es el doble de la que hay entre los valores 0 y z. ¿De qué manera nos puede ayudar la tabla a encontrar el valor de tal área?

Al examinar la tabla de la distribución normal, (el alumno puede consultar la que aparece en el apéndice de esta unidad o la de cualquier libro de estadística), podemos observar que la columna de la extrema izquierda tiene, precisamente el encabezado de "Z". Los valores de la misma se van incrementando de un décimo en un décimo a partir de 0.0 y hasta 4.2 (en nuestra tabla, en otras puede variar). El primer renglón de la tabla también tiene valores de "Z" que se incrementan de un centésimo en un centésimo de .00 a .09. Este arreglo nos permite encontrar los valores del área bajo la curva para valores de "Z" de 0.0 a 4.29.

Así podemos ver que para Z=1 (primera columna del cuerpo de la tabla y renglón de 1.0), el área es de 0.34134. Esto quiere decir, que entre la media y una unidad de z a la derecha tenemos el 34.134% del área de toda la curva.

Por el mismo procedimiento podemos ver que para un valor de Z=1.96 (renglón de Z=1.9 y columna de Z=.06), tenemos el 0.47500 del área. Esto quiere decir que entre la media y una Z de 1.96 se encuentra el 47.5% del área bajo la curva normal. De esta manera, para cualquier valor de Z se puede encontrar el área bajo la curva.

En el caso en z=2, la tabla nos da un valor de 0.47725, por lo que el área encerrada bajo la curva entre los valores -2 y 2 es 2(0.47725) = 0.9545

La manera en que este conocimiento de la tabla de la distribución normal puede aplicarse a situaciones más relacionadas con nuestras profesiones se puede ver en el siguiente ejemplo.

Ejemplo 1*:* Una empresa tiene 2000 clientes. Cada cliente debe en promedio $7000 con una desviación estándar de $1000. La distribución de los adeudos de los clientes es aproximadamente normal. Diga usted cuantos clientes esperamos que tengan un adeudo entre $7000 y $8,500.

Solución:

Nos percatamos de que valores como 7000 o 1000 no aparecen en la tabla de la distribución normal. Es allí donde interviene la variable Z porque nos permite convertir los datos de nuestro problema en números que podemos utilizar en la tabla. Lo anterior lo podemos hacer con la siguiente fórmula:

$$Z = \frac{x_i - \mu}{\sigma}$$

En nuestro caso, nos damos cuenta de que buscamos el área bajo la curva normal entre la media, 7000, y el valor de 8500. Sustituyendo los valores en la fórmula obtenemos lo siguiente:

$$z = \frac{8,500 - 7,000}{1,000} = 1.5$$

Buscamos en la tabla de la normal el área bajo la curva para Z=1.5 y encontramos 0.43319. Esto quiere decir que aproximadamente el 43.3% de los saldos de clientes están entre los dos valores señalados.

En caso de que el cálculo de Z arroje un número negativo significa que estamos trabajando a la izquierda de la media. El siguiente ejemplo ilustra esta situación.

Ejemplo 2: En la misma empresa del ejemplo anterior deseamos saber qué proporción de la población estará entre $6,500.00 y $7,000.00.

Solución:

Como en el caso anterior, nos damos cuenta de que nos piden el valor de un área entre la media y otro número. Volvemos a calcular el valor de Z

$$z = \frac{6,500 - 7,000}{1,000} = -0.5$$

Este valor de Z no significa un área negativa; lo único que indica es que el área buscada se encuentra a la izquierda de la media.

Aprovechando la simetría de la curva buscamos el área bajo la curva en la tabla para Z=0.5 (positivo, la tabla no maneja números negativos) y encontramos que el área es de 0.19146. Es decir que la proporción de saldos entre los dos valores considerados es de aproximadamente el 19.1%.

No siempre el área que se necesita bajo la curva normal se encuentra entre la media y cualquier otro valor. Frecuentemente son valores a lo largo de toda la curva. Por ello, es buena idea hacer un pequeño dibujo de la curva de distribución normal para localizar el o las áreas que se buscan. Esto facilita mucho la visualización del problema y, por lo mismo, su solución. A continuación se presenta un problema en el que se ilustra esta técnica.

Ejemplo 3: Una pequeña población recibe, durante la época de sequía, la dotación de agua potable mediante pipas que surten del líquido a la cisterna del pueblo una vez a la semana. El consumo semanal medio es de 160 metros cúbicos con una desviación estándar de 20 metros cúbicos. Indique cuál será la probabilidad de que el suministro sea suficiente en una semana cualquiera si se surten:

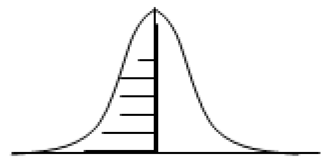
a) 160 metros cúbicos.
b) 180 metros cúbicos.
c) 200 metros cúbicos.
d) Indique asimismo cual será la probabilidad de que se acabe el agua si una semana cualquiera surten 190 metros cúbicos.

Solución*:*

a) 160 metros cúbicos.

El valor de Z en este caso sería: $z = \frac{160 - 160}{20} = 0.0$ Esto nos puede desconcertar un poco; sin embargo, nos podemos dar cuenta de que si se surten

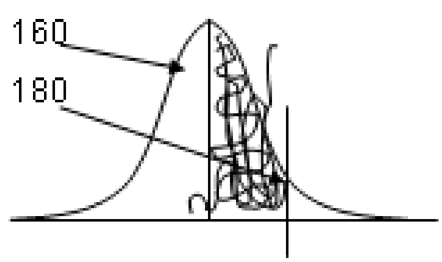
160 metros cúbicos el agua alcanzará si el consumo es menor que esa cifra. La media está en 160. Por ello el agua alcanzará en toda el área de la curva que se muestra rayada. Es decir, toda la mitad izquierda de la curva. El área de cada una de las mitades de la curva es de 0.5, por tanto la probabilidad buscada es también de 0 .5.



b) 180 metros cúbicos

El valor de Z es $$z = \frac{180 - 160}{20} = 1.0$$

El área que se busca es la que está entre la media, 160, y 180. Se marca con una curva en el diagrama Si buscamos el área bajo la curva en la tabla de la normal, para z=1.0, encontraremos el valor de 0.34134. Sin embargo, debemos agregarle toda la mitad izquierda dela curva (que por el diseño de la tabla no aparece). Ese valor, como ya se comentó es de .5. Por tanto, el valor buscado es de .5 más 0.34134. Por ello la probabilidad de que el agua alcance si se surten 180 metros cúbicos es de 0.84134, es decir, aproximadamente el 84%.



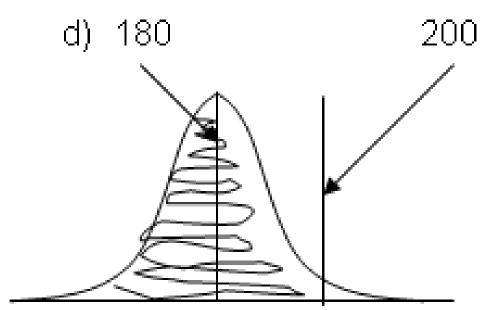
z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.0	**0.34134**	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214

c) 200 metros cúbicos

El área buscada se señala en el dibujo. Incluye la primera mitad de la curva y parte de la segunda mitad (la derecha), la que se encuentra entre la media y 200. Ya sabemos que la primera mitad de la curva tiene un área de 0.5. Para la otra parte tenemos que encontrar el valor de Z y buscar el área correspondiente en la tabla.

$$Z=(200-160)/20=2$$

lo que nos lleva a un valor en tablas de 0.47725. Al sumar las dos partes nos queda .97725. Es decir, si se surten 200 metros cúbicos hay una probabilidad de casi 98% de que el agua alcance.



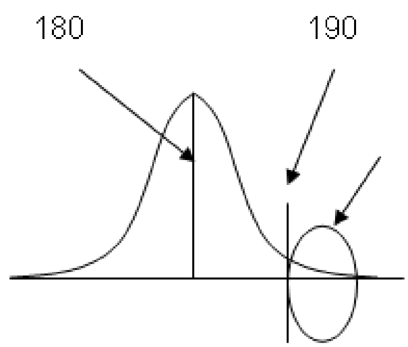
d) Indique asimismo cual será la probabilidad de que se acabe el agua si una semana cualquiera surten 190 metros cúbicos.

La probabilidad de que se termine el agua en estas condiciones se encuentra representada en la siguiente figura. El cálculo de la probabilidad de que ocurra se encuentra a continuación.

La probabilidad de que falte el agua está representada por el área en la cola de la distribución, después del 190. La tabla no nos da directamente ese valor. Para obtenerlo debemos de calcular Z para 190 y el valor del área entre la media y 190 restársela a .5 que es el área total de la parte derecha de la curva.

$$Z= (190-160)/20= 1.5.$$

El área para Z= 1.5 es de 0.43319. Por tanto, la probabilidad buscada es 0.5000 – 0.43319 = 0.06681 o aproximadamente el 6.7%.



Búsqueda de Z cuando el área bajo la curva es conocida

Frecuentemente el problema no es encontrar el área bajo la curva normal mediante el cálculo de Z y el acceso a la tabla para buscar el área ya mencionada. Efectivamente, a veces debemos enfrentar el problema inverso. Conocemos dicha área y deseamos conocer el valor de la variable que lo verifica. El siguiente problema ilustra esta situación.

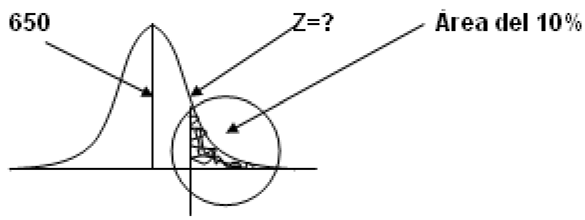
Ejemplo 4: Una universidad realiza un examen de admisión a 10,000 aspirantes para asignar los lugares disponibles. La calificación media de los estudiantes es de 650 puntos sobre 1000 y la desviación estándar es de 100 puntos; las calificaciones siguen una distribución normal. Indique usted qué calificación mínima deberá de tener un aspirante para ser admitido si:

a) Se aceptará al 10% de los aspirantes con mejor calificación.

b) Se aceptará al 5% de aspirantes con mejor calificación.

Solución:

a) Si hacemos un pequeño esquema de la curva normal, los aspirantes aceptados representan el 10% del área que se acumula en la cola derecha de la distribución. El siguiente esquema nos dará una mejor idea.



El razonamiento que se hace es el siguiente: Si el área que se busca es el 10% de la cola derecha, entonces el área que debemos de buscar en la tabla es lo más cercano posible al 40%, esto es 0.4000 (esto se busca en el cuerpo de la tabla, no en los encabezados que representan el valor de Z). Este es el valor de **0.39973** y se encuentra en el renglón donde aparece un valor para Z de 1.2 y en la columna de 0.08. Eso quiere decir que el valor de Z que más se aproxima es el de 1.28. No importa si al valor de la tabla le falta un poco o se pasa un poco; la idea es que sea el más cercano posible.

z	**0**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	**0.39973**	0.40147

Si ya sabemos el valor de Z, calcular el valor de la calificación (es decir “x”) es un problema de álgebra elemental y se trabaja despejando la fórmula de Z, tal como a continuación se indica.

Partimos de l a relación

$$Z= (X-650) / 100.$$

Observa que ya sustituimos los valores de la media y de la desviación estándar. Ahora sustituimos el valor de Z y nos queda:

$$1.28 =(X - 650)/100.$$

A continuación despejamos el valor de x.

$$1.28(100) = X - 650$$

$$128 + 650= X$$

$$X =778$$

En estas condiciones los aspirantes comenzarán a ser admitidos a partir de la calificación de 778 puntos en su examen de admisión.

b) El razonamiento es análogo al del inciso a. Solamente que ahora no buscamos que el área de la cola derecha sea el 10% del total sino solamente el 5% del mismo. Esto quiere decir que debemos buscar en la tabla en complemento del 5%, es decir 45% ó 0.45000. Vemos que el valor más cercano se encuentra en el renglón de Z de 1.6. y en la centésima 0.04

z	**0**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
1.6	0.44520	0.44630	0.44738	0.44845	**0.44950**	0.45053	0.45154	0.45254	0.45352	0.45449

Esto nos indica que el valor de z que buscamos es el de 1.64. El despeje de x se lleva a efecto de manera análoga al inciso anterior, tal como a continuación se muestra.

1.64 =(X - 650)/100.

1.64(100) = X - 650

164 + 650= X

X =814.

En caso de que se desee mayor precisión se puede recurrir a interpolar los valores (por ejemplo, en este caso entre 1.64 y 1.65) o buscar valores más precisos en paquetes estadísticos de cómputo.

Distribución exponencial

En una distribución de Poisson los eventos ocurren en un espectro continuo de tiempo o espacio. Se considera entonces que son eventos sucesivos de modo tal que la longitud o tiempo que transcurre entre cada realización del evento es una variable aleatoria, cuya distribución de probabilidades recibe precisamente el nombre de distribución exponencial.

Ésta se aplica cuando estamos interesados en el tiempo o espacio hasta el «primer evento», el tiempo entre dos eventos sucesivos o el tiempo hasta que ocurra el primer evento después de cualquier punto aleatoriamente seleccionado. Así, se presentan dos casos:

a) La probabilidad de que el primer evento ocurra dentro del intervalo de interés. Su fórmula es: $P\left(T \leq t\right) = 1 - e^{-\lambda}$

b) La probabilidad de que el primer evento *no* ocurra dentro del intervalo de interés. Su fórmula es:

$$P\left(T > t\right) = e^{-\lambda}$$

Donde λ es la tasa promedio de eventos por unidad de tiempo o longitud, según se trate. Como en el caso de la distribución de Poisson su parámetro es precisamente esta tasa promedio.

Ejemplo 1: Un departamento de reparaciones recibe un promedio de 15 llamadas por hora. A partir de este momento, cuál es la probabilidad de que:

a) En los siguientes 5 minutos no se reciba ninguna llamada.

b) Que la primera llamada ocurra dentro de esos 5 minutos.

c) En una tabla indicar las probabilidades de ocurrencia de la primera llamada en el minuto 1, 5, 10, 15, y 30.

Solución*:*

a) 1er. evento no ocurra:

Como la tasa promedio está expresada en llamadas por hora y en la pregunta se hace referencia a un periodo de 5 minutos, primero debemos hacer compatibles las unidades. Para ello, establecemos una relación de proporcionalidad directa

$$15 - 60$$
$$\lambda - \quad 5$$
$$\therefore \quad \lambda = 1.25$$

$$P\left(T > t\right) = e^{-\lambda}$$

$$p\left(t > 5\right) = 2.71828^{-1.25} = 0.286 = 28.6\%$$

$$e = 2.71828$$

Dice 0.286 = 28.6%   Debe decir 0.287 = 28.7%

b) 1ª llamada en 5 minutos:

$$P(T \le t) = 1 - e^{-\lambda}$$

$$P(T \le 5) = 1 - 2.71828^{-1.25} = 1 - 0.287 = 0.713 = 71.3\%$$

c) Diversas probabilidades:

Espacio tiempo	λ	Probabilidad ocurra	Probabilidad No ocurra
1 minuto	0.25	0.221	0.779
5 minutos	1.25	0.713	0.287
10 minutos	2.50	0.918	0.082
15 minutos	3.75	0.976	0.024
30 minutos	7.50	0.999	0.001

Examen de autoevaluación del tema

Si X es una variable aleatoria con distribución normal con media 15 y desviación estándar 3, determinar la probabilidad de que

Escribe en los espacios en blanco tu respuesta. Una vez hayas concluido obtendrás tu calificación de manera automática.

a) $X \geq 14$=________
b) $X \leq 17$=________
c) $13 \leq X \leq 14$=_______
d) $14 \leq X \leq 17$=_______
e) $X \geq 16$=_______

ACTIVIDAD 1

Contesta las siguientes preguntas.

La gerencia de recursos humanos de un corporativo aplica a un grupo de solicitantes de empleo una prueba de aptitud. La calificación promedio obtenida por los solicitantes es de 78 puntos con una desviación estándar de 13.

a) ¿Cuál es la probabilidad de que si se selecciona al azar a uno de tales solicitantes, éste tenga una calificación
   - superior a 85 puntos?
   - menor a 75 puntos?
   - entre 70 y 90 puntos?

b) ¿Entre qué valores se encuentra el 80% de la población que excluye al 10% más apto y al 10% menos apto?

c) ¿Cuál es la calificación máxima del 25% menos apto?

Realiza tu actividad en un procesador de textos, guárdala en tu computadora y una vez concluida, presiona el botón **Examinar.** Localiza el archivo, ya seleccionado, presiona **Subir este archivo** para guardarlo en la plataforma

Bibliografía básica

Autor	Capítulo	Páginas
1. Berenson, Levine y Krehbiel.	4. Probabilidad básica y distribuciones de probabilidad. 4.7 Distribución normal.	198- 219.
	4.8 Verificación de la suposición de normalidad.	219- 230
3. Levin y Rubin.	5. Distribuciones de probabilidad 5.6 La distribución normal: distribución de una variable aleatoria continua.	209-222
4. Lind, Marchal, Mason.	7. Distribuciones de probabilidad normal. La familia de distribuciones de probabilidad normal.	227- 229
	Distribución de probabilidad normal estándar.	229- 243

Sitios electrónicos

Sitio	Descripción
http://www.uaq.mx/matematicas/estadisticas/xu4-5.html#t1	Larios O Víctor (profesor de la Universidad Autónoma de Querétaro), *Unidad 5. Distribuciones de probabilidad, parte 5, La distribución normal*, del hipertexto Estadística, en donde además de presentar las características generales de la distribución normal, se demuestran de manera rigurosa las principales propiedades de la curva de distribución normal.
http://www.itch.edu.mx/academic/industrial/sabaticorita/_private/01UNIDAD%20%20V.htm	Luna Gándara MC Rita (profesora de la licenciatura en ingeniería industrial del Instituto Tecnológico de Chihuahua), *Unidad V. Distribuciones de probabilidad continuas*, apuntes del curso de probabilidad y estadística, en donde se presenta por medio de ejemplos el cálculo de probabilidades para el caso de la distribución normal, mediante el uso de la tabla respectiva.
http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/dis_normal.htm	García Cebrián María José, *La distribución normal*, material que forma parte del proyecto Descartes 2D del ministerio de educación, política social y deporte del gobierno de España, en donde por medio de graficaciones

	interactivas se muestra el uso de las tablas de probabilidad de la distribución normal, así como el proceso de estandarización.

Tema 6. Ley de los grandes números

Objetivo del tema

Aplicar la distribución normal para calcular probabilidades de variables aleatorias que siguen un modelo binomial.

Desarrollo

La ley de los grandes números sugiere que **la probabilidad de una desviación significativa** de un valor de probabilidad determinado empíricamente, a partir de uno determinado teóricamente, es menor cuanto más grande sea el número de repeticiones del experimento.

Esta ley forma parte de lo que en la probabilidad se conoce como teoremas de límites, uno de los cuales es el teorema de De Moivre-Laplace según el cual, la distribución binomial – que se presenta en múltiples casos en los que se requiere conocer la probabilidad de ocurrencia de un número determinado de éxitos en una muestra aleatoriamente seleccionada – puede aproximarse por la distribución normal si el número de ensayos es suficientemente grande y donde el error en la aproximación disminuye en la medida en que la probabilidad de éxito se acerca a 0.5.

Desde el punto de vista de las operaciones, si lo que deseamos es calcular la probabilidad de que una variable aleatoria binomial con parámetros n y p tome valores entre a y b, entonces debemos:

- Determinar la media y desviación estándar de la variable binomial, esto es, calcular los valores de $\mu = np$, y $\sigma = (npq)^{1/2}$
- Reexpresar la probabilidad deseada en el contexto binomial por la probabilidad deseada en el contexto de la distribución normal, incorporando una corrección por finitud, esto es, si nuestra pregunta original es

determinar el valor de $P(a \leq X \leq b)$ entonces, buscaremos aplicar la distribución normal para calcular

$$P([a - 0.5 - np] / (npq)^{1/2} < Z < [b + 0.5 - np] / (npq)^{1/2}),$$

donde los sumandos 0.5 y -0.5 constituyen la corrección por finitud

- Emplear la tabla de la distribución normal

Veamos un ejemplo.

Ejemplo 1. Se arroja una moneda legal 200 veces. Se desea saber la probabilidad de que aparezca sol más de 110 veces pero menos de 130.

Solución:

El hecho de que la moneda sea legal significa que la probabilidad de que el resultado sea sol es igual a la probabilidad de que salga águila, de modo que tanto la probabilidad de éxito como de fracaso es 0.5, y esta probabilidad no cambia de ensayo a ensayo. Podemos decir entonces que estamos en presencia de un experimento Binomial, con lo que podemos plantear que la solución al problema es:

P(sol mas de 110 veces pero menos de 130) = P(sol 111 veces) + P(sol 112 veces) + P(sol 113 veces) + … + P(sol 129 veces)

$$= \Sigma\ {}_{\mathbf{200}}\mathbf{C}_{\mathbf{i}}\ (0.5)^{i}\ (0.5)^{200-i}$$

donde i recorre los valores desde 111 hasta 129.

El problema es que al hacer las operaciones, incluso con una calculadora tendríamos dificultades. Es aquí donde es útil aplicar la distribución normal como aproximación a la distribución binomial.

Como n=200 y p=0.5, entonces la media es $\mu = 200(0.5) = 100$, en tanto que la varianza es $\sigma^2 = 200(0.5)(0.5) = 50$, de modo que la desviación estándar es $\sigma = 7.07$.

En consecuencia,

$$P(111 \le X \le 129) = P\left[(110.5 - 100) / 7.07 \le Z \le (129.5 - 100) / 7.07 \right]$$
$$= P(10.5 / 7.07 \le Z \le 29.5 / 7.07)$$
$$= P(1.49 \le Z \le 4.17)$$
$$= 0.5 - 0.4319$$
$$= 0.0681$$

Cuando el número de ensayos es grande pero el valor dela probabilidad de éxito se acerca a cero o a uno, esto es se aleja de 0.5, es mejor emplear la distribución de Poisson como aproximación a la binomial

DISTRIBUCIÓN NORMAL ESTÁNDAR

ÁREA BAJO LA CURVA

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0.7	0.2580	0.2611	0.2642	0.2673	0.2703	0.2733	0.2763	0.2793	0.2823	0.2852

	4	5	4	0	5	7	7	5	0	4
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41308	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807

2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861
3.0	0.49865	0.49869	0.49874	0.49878	0.49882	0.49886	0.49889	0.49893	0.49896	0.49900
3.1	0.49903	0.49906	0.49910	0.49913	0.49916	0.49918	0.49921	0.49924	0.49926	0.49929
3.2	0.49931	0.49934	0.49936	0.49938	0.49940	0.49942	0.49944	0.49946	0.49948	0.49950
3.3	0.49952	0.49953	0.49955	0.49957	0.49958	0.49960	0.49961	0.49962	0.49964	0.49965
3.4	0.49966	0.49968	0.49969	0.49970	0.49971	0.49972	0.49973	0.49974	0.49975	0.49976
3.5	0.49977	0.49978	0.49978	0.49979	0.49980	0.49981	0.49981	0.49982	0.49983	0.49983
3.6	0.49984	0.49985	0.49985	0.49986	0.49986	0.49987	0.49987	0.49988	0.49988	0.49989
3.7	0.49989	0.49990	0.49990	0.49990	0.49991	0.49991	0.49992	0.49992	0.49992	0.49992
3.8	0.49993	0.49993	0.49993	0.49994	0.49994	0.49994	0.49994	0.49995	0.49995	0.49995
3.9	0.49995	0.49995	0.49996	0.49996	0.49996	0.49996	0.49996	0.49996	0.49997	0.49997
4.0	0.49997	0.49997	0.49997	0.49997	0.49997	0.49997	0.49998	0.49998	0.49998	0.49998
4.1	0.49998	0.49998	0.49998	0.49998	0.49998	0.49998	0.49998	0.49998	0.49999	0.49999
4.2	0.49999	0.49999	0.49999	0.49999	0.49999	0.49999	0.49999	0.49999	0.49999	0.49999

ACTIVIDAD 1

La gerencia de un banco está interesada en determinar la probabilidad de errores en las operaciones de depósito. Si se auditan 5 000 de estas operaciones, ¿cuál es la probabilidad de encontrar entre 10 y 15 operaciones con error?

a. Si se sabe que la probabilidad de cometer un error es de 0.005.

b. Si se sabe que la probabilidad de cometer un error es de 0.3.

Justifica el uso el uso de las distribuciones normal o de Poisson como aproximación a la distribución real.

Para enviar tu respuesta, pulsa el botón **Editar mi envío**; se mostrará un editor de texto en el cual puedes redactar tu información; una vez que hayas concluido, salva tu actividad pulsando el botón **Guardar cambios**.

ACTIVIDAD 2

Considera la siguiente situación:

En una ciudad de tamaño medio se han establecido 3 academias particulares, sin que desde el punto de vista de la investigación de mercados haya mayores diferencias entre ellas. Se calcula que para el siguiente año hay 800 alumnos que podrían matricularse de los cuales finalmente lo harán 450. Las autoridades de una de las academias desean saber cuál debe ser la capacidad de su escuela para que sea suficiente en el 80% de los cursos

Si tú fueses consultor, qué respuesta darías. Justifica ésta tan ampliamente como puedas.

Para enviar tu respuesta, pulsa el botón **Editar mi envío**; se mostrará un editor de texto en el cual puedes redactar tu información; una vez que hayas concluido, salva tu actividad pulsando el botón **Guardar cambios**.

Bibliografía básica

Autor	Capítulo	Páginas
4. Lind, Marchal, Mason.	7. Distribuciones de probabilidad normal. Sección: Aproximación normal a la binomial.	243-248

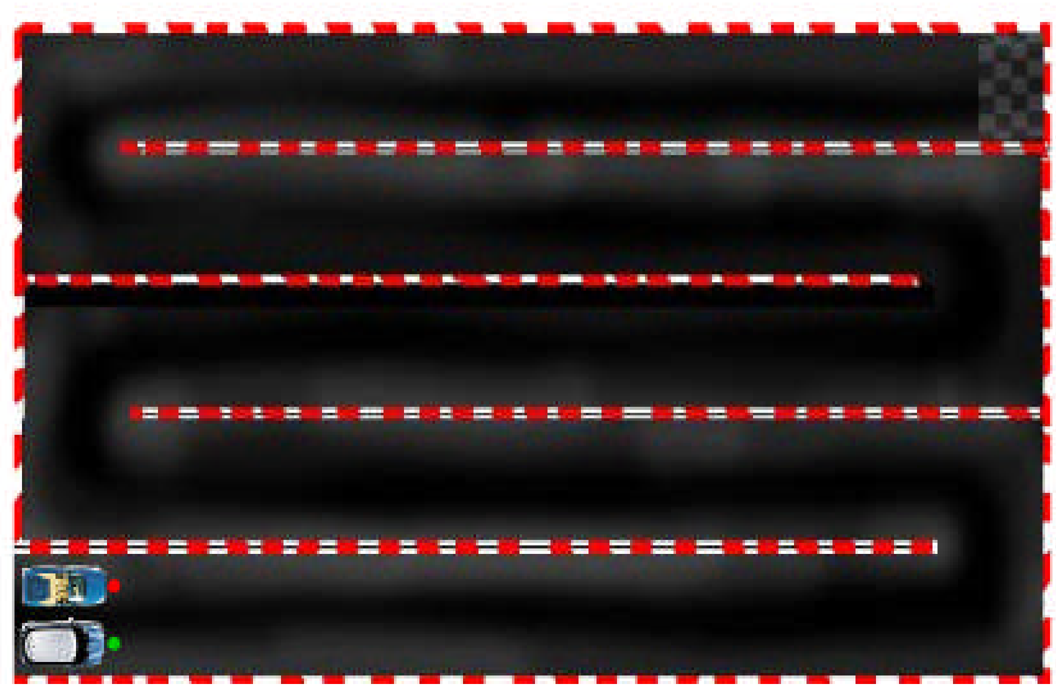
Sitios electrónicos

Sitio	Descripción
http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/aplic_normal.htm	García Cebrián, José María, *Otras aplicaciones de la distribución normal*, publicado como parte del proyecto Descartes 2D del ministerio de educación, política social y deporte del gobierno de España, en donde se muestra y ejemplifica el empleo de la distribución normal para aproximar una distribución binomial.

Examen de autoevaluación del tema

Una vez que hemos estudiado esta unidad te invitamos a que participes en este rally, para que compruebes el logro del objetivo que nos planteamos.

Responde las siguientes preguntas, una vez que termines obtendrás tu calificación de manera automática.



Pregunta 1

a b c d e

1. Las empresas mexicanas están aprovechando la condición de costo de mano de obra más barato para realizar trabajos en el extranjero; para ello utilizan los servicios de empresas de contratación locales para la resolución de todos los aspectos legales. Las encuestas realizadas por el Banco de Comercio Exterior indican que el 20% de las empresas mexicanas utilizan a este tipo de empresas. Si el banco selecciona al azar a un grupo de 15 empresas

mexicanas. ¿Calcule la probabilidad de que exactamente cinco de ellas estén empleando a estas empresas locales?

a. 0.1032
b. 0.1058
c. 0.1028
d. 0.1035
e. 0.1038

Pregunta 2

a b c d e

1. 2. De acuerdo con la situación anterior, calcule la probabilidad de que el número de empresas mexicanas que contratan empresas locales en el extranjero se ubique entre seis y nueve.

a. 0.0651
b. 0.0609
c. 0.0631
d. 0.0607
e. 0.0609

Pregunta 3

a	b	c	d	e

3. El Banco Nacional de México sabe por su experiencia que durante los días lunes, entre las 9:00 y las 10:00, se presentan a la ventanilla de atención a clientes un promedio de 2.8 clientes cada 4 minutos, número que la cajera puede atender con eficiencia. Con el propósito de verificar si el número de cajeras es el adecuado, calcule la probabilidad de que se presente un total de cuatro clientes en un intervalo de cuatro minutos.

a. 0.1568
b. 0.1557
c. 0.1535
d. 0.1678
e. 0.1456

Pregunta 4

a	b	c	d	e

4. Un auxiliar de contador puede cometer 1.2 errores por cada 200 declaraciones fiscales. ¿Calcule la probabilidad de que al seleccionar una de las declaraciones elaboradas por él no se encuentre algún error?

a. 0.9600
b. 0.9940
c. 0.3012
d. 0.5990
e. 0.7890

Pregunta 5

a b c d e

5. Un banco recibe en promedio a 3.2 clientes cada 4 minutos. ¿Cuál es la probabilidad de obtener exactamente 10 clientes en los próximos 8 minutos?

a. 0.0538

b. 0.0635

c. 0.0535

d. 0.0525

e. 0.0528

Pregunta 6

a b c d e

6. Una persona presentará el examen de conocimientos y dominio de la lengua inglesa, denominado GMAT; sus resultados tienen un valor medio de 494 puntos con desviación estándar de 100; La persona desea conocer la probabilidad de obtener 700 puntos. Considere que los resultados siguen una distribución normal estándar.

a. 0.0255

b. 0.0204

c. 0.0197

d. 0.0199

e. 0.0193

Pregunta 7

a	b	c	d	e

7. Una empresa de seguros está considerando incluir entre los riesgos cubiertos, una enfermedad denominada Túnel Carpiano, la cual aparece en manos y muñecas, provocada por los esfuerzos realizados con estas partes del cuerpo durante tiempos prolongados. Se estima que el costo de tratamiento de estas afecciones es alrededor de $30,000 pesos al año por trabajador lesionado, con una desviación estándar de $9,000.00. La aseguradora supone que la afección está normalmente distribuida y desea estimar los costos en que puede incurrir. Calcule la probabilidad de que el costo de atención se encuentre entre $15,000 y $ 45,000.

a. 0.9050
b. 0.0950
c. 0.9152
d. 0.9070
e. 0.9030

Pregunta 8

a b c d e

8. Una empresa de automóviles menciona en su publicidad que sustituirá por una unidad nueva los autos que presenten cualquier tipo de falla en el tren motriz durante los primeros 80,000 kilómetros. Si la empresa sabe que el valor medio del kilometraje sin fallas es de 80,000 kilómetros y la desviación estándar de 10 000 kilómetros, ¿cuál debería ser el kilometraje garantizado para no tener que reponer más del 10% de los autos?

a. 65 600

b. 68 300

c. 67 200

d. 68 450

e. 63 600

Pregunta 9

a b c d e

9. Una distribución normal tiene una media de 4.9 y una desviación estándar de 1.2. ¿Qué porcentaje del área bajo la curva es mayor que 6?

a. 0.1685

b. 0.1814

c. 0.1797

d. 0.1788

e. 0.1750

Pregunta 10

a	b	c	d	e

10. Se aplica un examen de Matemáticas a 4000 estudiantes próximos a egresar del ciclo de educación media superior. Si en experiencias previas ha ocurrido que la calificación promedio es de 6.7 con una desviación estándar de 3.1, y bajo el supuesto de que las calificaciones se distribuyen de manera normal, ¿cuál es el número de estudiantes que podría esperarse en esta ocasión obtuviesen una calificación superior a 9.0?

a. 229
b. 918
c. 230
d. 770
e. 3082

LO QUE APRENDÍ DE LA UNIDAD

En esta unidad has entrado en contacto con conceptos tales como variable aleatoria, distribución de probabilidades y esperanza matemática. Con ellos puedes desarrollar una explicación e interpretación formal del concepto de esperanza de vida al nacimiento, entre otras muchas aplicaciones de las ideas que aquí se han presentado.

En particular puedes explicar qué aspectos inciden en el incremento en el valor de la esperanza de vida y discute con tus compañeros ¿Por qué crees que esto ocurre?

Pulse el botón **Colocar un nuevo tema de discusión aquí**.
Escriba en el apartado **Asunto** el título de su aportación, redacte su comentario en el área de texto y de clic en el botón **Enviar al foro**.

Glosario de la unidad

Distribución binomial

Es la distribución de probabilidades en un experimento binomial, de modo que la variable aleatoria, discreta, se refiere al número de éxitos en n ensayos.

Distribución normal

Describe el comportamiento estadístico de muchas situaciones del mundo real en donde:

- Hay simetría alrededor de un punto medio.
- Se miden poblaciones grandes.
- El 95% de los datos se concentra en un intervalo de longitud 1.96 veces la desviación estándar alrededor de la media.

Distribución de Poisson

Es la distribución de probabilidades en un experimento Poisson donde la variable aleatoria, discreta, se refiere al número de veces que ocurre un cierto resultado A.

Esperanza matemática

Es un atributo de las variables aleatorias y por lo tanto de la respectiva distribución de probabilidades que describe lo que ocurre en el centro de la distribución, aunque no siempre existe. En analogía con los indicadores de estadística descriptiva se puede decir que corresponde a la media.

Experimento binomial

Es un experimento que se caracteriza porque:

- Hay un número finito de ensayos.
- En cada ensayo sólo hay dos posibles resultados, a saber, éxito o fracaso.
- Los ensayos son independientes, de modo que la probabilidad de éxito permanece constante en cada ensayo.

Experimento Poisson

Es un experimento que se caracteriza porque:

- Un cierto resultado A puede ocurrir un número infinito pero numerable de veces a lo largo de un intervalo (de tiempo, longitud o área).
- La probabilidad de que ocurra el resultado A es proporcional al tamaño del intervalo.
- La probabilidad de que ocurra el resultado A en una fracción del intervalo es independiente de la probabilidad de que ocurra en otra fracción del intervalo.
- La probabilidad de que ocurra más de una vez el resultado A en una fracción muy pequeña del intervalo es prácticamente nula.
- Se conoce el valor de la tasa promedio de ocurrencia del resultado A.

Función de distribución

Es aquélla que acumula los valores de la función de probabilidad. Se denota como F(x) y corresponde al valor de P(X ≤ x).

Función de probabilidad discreta

Es aquélla que asocia una probabilidad a cada posible valor de una variable aleatoria discreta.

Variable aleatoria

Es una aplicación que asocia un número real a cada elemento del espacio muestral de un experimento.

Variable aleatoria continua

Es una variable aleatoria que hace referencia a un experimento cuyos universo de resultados se encuentra en un intervalo del tipo (a,b) y la variable puede tomar cualesquier valor en ese intervalo.

Variable aleatoria discreta

Es una variable aleatoria que hace referencia a los distintos resultados de un experimento tomando un número finito o infinito numerable de posibles valores. Intuitivamente, una variable es discreta si deja huecos entre cada par de posibles valores.

MESOGRAFÍA

Bibliografía básica

1. Berenson, Mark L., David M. Levine, y Timothy C. Krehbiel, (2001), *Estadística para administración*, 2ª edición, México, Prentice Hall, 734 páginas.
2. Hernández Sampieri, R., C. Fernández Collado, Lucio P Baptista, (2006), *Metodología de la investigación*, 4ª edición, México: McGraw Hill Interamericana, 850 páginas.
3. Levin, Richard I. y David S. Rubin, (2004), *Estadística para administración y economía*, 7a. Edición, México, Pearson Educación Prentice Hall, 826 páginas más anexos.
4. Lind, Douglas A., William G., Marchal, Robert D. Mason, (2004), *Estadística para administración y economía*, 11ª edición, Bogotá, Alfaomega grupo editor, 830 págs.

Sitios electrónicos

- García Cebrián María José, *La distribución normal*, material que forma parte del proyecto Descartes 2D del ministerio de educación, política social y deporte del gobierno de España, en el sitio: http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/dis_normal.htm.
- García Cebrián, José María, *Otras aplicaciones de la distribución normal*, publicado como parte del proyecto Descartes 2D del ministerio de educación, política social y deporte del gobierno de España, en el sitio: http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/aplic_normal.htm
- Larios O. Víctor, (profesor de la Universidad Autónoma de Querétaro), *Unidad 5. Distribuciones de probabilidad,* del hipertexto Estadística, en el sitio:

http://www.uaq.mx/matematicas/estadisticas/xu4-5.html#t1.

- Luna Gándara MC Rita (profesora de la licenciatura en ingeniería industrial del Instituto Tecnológico de Chihuahua), *Unidad V. Distribuciones de probabilidad continuas*, apuntes del curso de probabilidad y estadística, en el sitio:
  http://www.itch.edu.mx/academic/industrial/sabaticorita/_private/01UNIDAD%20%20V.htm.
- Hospital Universitario Ramón y Cajal, *Variable aleatoria,* hasta la sección Inducción a la probabilidad, en el sitio:
  http://www.hrc.es/bioest/estadis_21.html
- Rodríguez Mayté *(*profesora del curso de Estadística aplicada a las ciencias sociales II de la licenciatura de sociología de la Universidad Autónoma de Madrid), *Variables aleatorias,* en el capítulo 3, secciones 3.1 y 3.2, p. 25-27, en el sitio:
  http://www.uam.es/personal_pdi/ciencias/mayter/docencia/sociolog/apuntes.pdf
- Rodríguez Mayté (profesora del curso de Estadística aplicada a las ciencias sociales II, de la licenciatura en sociología de la Universidad Autónoma de Madrid) *Variables aleatorias discretas,* en el capítulo 4, secciones 4.1, 4.2 y 4.4, págs. 34 a 41, en el sitio:
  http://www.uam.es/personal_pdi/ciencias/mayter/docencia/sociologia/apuntes.pdf