DISTRIBUCIONES DE PROBABILIDAD

Universidad Nacional Autónoma de México

Maestría en Administración

DISTRIBUCIONES DE PROBABILIDAD

Introducción

Una distribución de probabilidad indica toda la gama de valores quepueden representarse como resultado de un experimento si éste sellevase a cabo. Es decir, describe la probabilidad de que un evento se realice en el futuro, lo implica que se puede diseñar un escenario de acontecimientos futuros considerandolas tendencias actuales de diversos fenómenos naturales.

Una distribución de probabilidades es una lista de las probabilidades de todos los resultados posibles que pudiera resultar si el experimento se hace; es decir, es la suma de todas las funciones en las que interviene la variable aleatoria “x” bajo estudio.

Las distribuciones de probabilidad siempre es la suma de todas las funciones posibles, por tanto su sumatoria siempre tiene que ser igual al espacio muestral; esto es:

$$f(x) = 1 \qquad\qquad f(x) = 100\%$$

Para saber cual es la *f(x)* que corresponde se deberá estudiar los tipos de distribuciones de probabilidad que podemos tener. Para esto, lo importante es definir el tipo de variable que tenemos bajo estudio, y de aquí surge la clasificación de las distribuciones.

Clasificación de las Distribuciones de Probabilidad

a) Distribuciones Discretas:

Se presentan cuando nuestra variable de estudio es discreta; esto es, solo puede asumir valores enteros, sin decimales.

Eventos Independientes
- Distribución Binomial
- Distribución de Poisson

Eventos Dependientes
- Distribución Hipergeométrica

b) Distribuciones Contínuas:

Se presentan cuando nuestra variable de estudio es contínua; esto es, puede asumir valores dentro de un intervalo de valores.

*Distribución Normal

• Aproximación Normal para la Distribución Binomial

* Aproximación Normal para la Distribución de Poisson

DISTRIBUCIÓN BINOMIAL

La distribución Binomial es posiblemente la más importante de las distribuciones discretas.Esta distribución corresponde a la realización de un experimento aleatorioque cumple con las siguientes condiciones:

1. El experimento consta de n ensayos o pruebas idénticas;
2. Los datos recopilados son el resultado de conteos;
3. Cada ensayo sólo puede tener un resultado de dos posibles, que se llaman “éxito” (p) y “fracaso” (q), los cuales son mutuamente excluyentes;
4. La probabilidad p de “éxito” es constante de ensayo a ensayo, es decir, es invariable. La probabilidad de un fracaso es (1 -p) = q
5. Los ensayos son estadísticamente independientes, esto es la ocurrencia de uno no afecta la ocurrencia del otro.
6. Interesa conocer x, el número de éxitos observados en n pruebas

FÓRMULA

$$f(x) = \binom{n}{x} p^{x} q^{n-x} \qquad \text{para } x = 0, 1, 2, \ldots, n$$

En donde:

f (x) = Función de x de la distribución binomial

( ) = Coeficiente binomial (análisis combinatorio)

n = Tamaño de la muestra ó número de ensayos (pruebas)

x = Número de éxitos observados (valor de variable)

p = Probabilidad de éxito en cada ensayo

q = Probabilidad de fracaso que se obtiene por: 1- p

MÉDIA ARITMÉTICA O VALOR ESPERADO DE UNA DISTRIBUCIÓN BINOMIAL

El valor esperado de una variable aleatoria binomial siempre es igual al número de ensayos del experimento multiplicado por la probabilidad de éxito en cualquier ensayo, es decir

$$\mu = n \cdot p$$

VARIANZA Y DESVIACIÓN ESTÁNDAR DE UNA DISTRIBUCIÓN BINOMIAL

La varianza de una variable aleatoria binomial es siempre igual al número de ensayos del experimento multiplicado por la probabilidad de éxito en cualquier ensayo, multiplicado a su vez por la probabilidad de fracaso en cualquier ensayo, es decir

$$\nabla^2 = n \cdot p \cdot q$$

Así mismo la desviación estándar de una variable aleatoria binomial es:

$$\nabla = \sqrt{n \cdot p \cdot q}$$

ASIMETRÍA DE UNA DISTRIBUCIÓN BINOMIAL

1. Cuando $p = 0.50$, esto es $p = q$, cualquier distribución binomial es simétrica, es decir, exhibe cero asimetría.
2. Cuando $p < 0.50$, esto es $p < q$, la distribución es positivamente sesgada (a la derecha).
3. Cuando $p > 0.50$, esto es $p > q$, la distribución es negativamente sesgada (a la izquierda).
4. A medida que el tamaño de la muestra (n) aumenta, la distribución se hace más y más simétrica.

NOTA: Estas generalizaciones se pueden notar claramente al trazar la gráfica de la distribución binomial.

DISTRIBUCIÓN DE PROBABILIDAD HIPERGEOMÉTRICA

Cuando el muestreo se realiza sin reposición para cada uno de los elementos que se toman de una población finita de elementos, no se puede aplicar el proceso de Bernoulli debido a que existe un cambio sistemático en la probabilidad de éxitos al ir extrayendo elementos de la población, en este caso la distribución discreta de probabilidad apropiada resulta ser la distribución hipergeométrica.

CARACTERÍSTICAS DE LA DISTRIBUCIÓN HIPERGEOMÉTR1CA

Suponga que se obtiene una muestra de elementos de una población y se determina si cada uno de ellos tiene o no una característica determinada. Los datos obtenidos son de tipo “acierto” y “falla” como en el caso de la distribución binomial.

Si el número de elementos de la población es grande en comparación con el de la muestra, la probabilidad de seleccionar un elemento con una determinada característica en un solo ensayo es igual a la proporción de “éxito” p de elementos con esa característica en la población. Dado que la población es grande comparada con el tamaño de la muestra, esta probabilidad permanecerá constante (para propósitos prácticos) de ensayo a ensayo y el número de aciertos sigue una distribución de probabilidad binomial.

Sin embargo, si el número de elementos en la población es pequeño en relación con el tamaño de la muestra, es decir

$$n > (0.05)(a + b) \text{ ó } n > (0.05)(n)$$

La probabilidad de un acierto en un ensayo dado depende de los resultados precedentes. Entonces el número a de aciertos sigue lo que se conoce como una distribución de probabilidad hipergeométrica.

FÓRMULA

Supóngase que vamos a muestrear una población finita sin reposición en la que los elementos de la población están clasificados únicamente en dos clases: aciertos y fallos. Y si estamos interesados en tomar una muestra aleatoria de n elementos, formado a “éxitos” o aciertos y b “fracasos” o fallos y nos interesa conocer la probabilidad de obtener x aciertos y n - x fallos, entonces está dada por:

$$f(x) = \frac{\binom{a}{x}\binom{b}{\ }}{\binom{N}{\ }} \qquad \text{para } x = 0, 1, 2, \ldots\ldots\ldots, n$$

f (x) = Función de probabilidad de x de la distribución hipergeométrica;

a = Número de elementos de la población finita considerados aciertos;

b = Número de elementos de la población finita considerados fallos;

a + b = N = Número de elementos de la población finita;

x = El valor de la variable discreta;

n = Tamaño de la muestra;

( ) = Análisis combinatorio.

MEDIA ARITMÉTICA O VALOR ESPERADO DE LA DISTRIBUCIÓN HIPERGEOMÉTRICA

$$\mu = n\left(\frac{a}{N}\right)$$

VARIANZA Y DESVIACIÓN ESTÁNDAR DE LA DISTRIBUCIÓN HIPERGEOMÉTRICA

Varianza $$\sigma^2 = n \cdot p \cdot q \left(\frac{N - n}{N - 1}\right)$$

Desviación Estándar $$\sigma = \sqrt{n \cdot p \cdot q \left(\frac{N - n}{N - 1}\right)}$$

DISTRIBUCIÓN DE PROBABILIDAD DE POISSON

La distribución recibe su nombre en honor de Simeon Poisson (1781-1840) quien la estudió y dio a conocer en 1837. El definió lo que ahora recibe el nombre de variable aleatoria de Poisson, como el número de acontecimientos raros de un evento específico que ocurren en una unidad de tiempo o espacio especificado.

PROCESO DE POISSON

Un proceso de Poisson es el acontecimiento de una serie de eventos de un tipo dado, en forma aleatoria, en un tiempo o espacio, tal que:

1. El número de acontecimientos dentro de un tiempo o espacio especificado es igual a cualquier entero entre 0 e infinito;
2. El número de acontecimientos dentro de una unidad de tiempo o espacio es independiente del de cualquier otra unidad (que no se traslape):
3. La probabilidad de acontecimientos es la misma en todas esas unidades.

EVENTOS QUE ACONTECEN AL AZAR EN EL TIEMPO

Las clases de eventos que Poisson tenía en mente se encuentran en todas las actividades humanas, tales como:

1. Número de personas que solicitan servicio en un banco, parada de autobuses, en un consultorio médico, en la oficina de correos o en un restaurante.
2. Las llegadas de carros al taller de lavado, las entradas a carreteras troncales o a gasolineras, en estacionamientos, semáforos o casetas de cuotas.
3. Llegada de aviones a aeropuertos, de embarcaciones a puertos, de camiones a terminales.
4. Llamadas telefónicas a puestos centrales cuando se requieren los servicios de ambulancias, bomberos o policía.
5. Número de accidentes industriales, automovilísticos, deportivos, etc.
6. Número de clientes que llegan a una tienda de autoservicio, a una caja registradora de un supermercado, a una tintorería, a una peluquería, a una tortillería, etc.

7. Reclamaciones en oficinas de compañías de seguros, de paquetes a la oficina de correos, de órdenes en un almacén, de productos semiterminados, de solicitudes de devolución en el departamento de quejas, de víctimas de suicidios en el servicio forense. En todos estos casos y en muchos más, las condiciones del proceso de Poisson pueden aplicarse muy bien.

EVENTOS DISPERSOS EN FORMA ALEATORIA EN EL ESPACIO

1. Véase el "espacio" como una línea y considérese dichos eventos como la aparición de defectos en un rollo de alambre forrado, de fugas en un oleoducto o nombres mal escritos en un directorio telefónico, baches en un camino o errores de tipografía en un libro.
2. Véase el "espacio" como área y contemple el número de nacimientos en una ciudad, de burbujas en una ventana de vidrio, delitos en una ciudad, los defectos en una alfombra, las casas de agricultores en un municipio, los incendios en un bosque ocasionados por rayos, pensemos en minas, meteoritos o yerbas en un campo.
3. Véase el espacio como volumen y considérese eventos como hallar bacterias en un caldo, pasitas en un pan o semillas de maleza en un paquete de semillas de flores, etc.

FÓRMULA

$$f(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} \qquad \text{para } x = 0, 1, 2, \ldots\ldots\ldots, \infty$$

En donde:

f (x) = Función de x de la distribución de probabilidad de Poisson

μ = Valor esperado (media aritmética del número de ocurrencias (éxitos) en un intervalo de tiempo determinado)

x = Número de acontecimientos por unidad de tiempo o espacio (valor de la variable)

! = Factorial de un número

e = 2.71828 (base de logaritmos naturales)

MEDIA ARITMÉTICA O VALOR ESPERADO DE LA DISTRIBUCIÓN DE POISSON

$$\mu = n \cdot p$$

VAR1ANZA Y DESVIACIÓN ESTÁNDAR DE UNA DISTRIBUCIÓN DE PO1SSON

Varianza $\nabla^2 = n \cdot p$

Desviación Estándar $\nabla = \sqrt{n \cdot p}$ ó $\nabla = \sqrt{\mu}$

APROXIMACIÓN DE LA BINOMIAL POR LA DISTRIBUCIÓN DE PROBABILIDAD DE POISSON

Este uso está particularmente justificado en el caso de un proceso Bernoulli en el que el número de ensayos, n. es grande y la probabilidad de éxito en cualquier intento, p, es pequeño, es decir:

$$n > 30 \qquad y \qquad p \ [\ 0.05$$

y de preferencia que el valor esperado sea también pequeño, es decir:

$$\mu \ [\ 5$$

DISTR1BUCIÓN DE PROBABILIDAD NORMAL

Esta distribución de probabilidad fue descrita primeramente en 1753 por Abraham de Moivre (1667-1 754), pero fue popularizada por Adolphe Quetelet (1796 - 1574) quien la utilizó para analizar el concepto de "el hombre promedio", y por Carl Friedrich Gauss (1777-1855) que la usó para describir errores de medición en astronomía.

Si el tipo de variable aleatoria cuantitativa que se presenta es de naturaleza continua, es decir, que pueden tomar valores en todos los puntos de una escala y sin interrupciones entre valores posibles. Considérese las características medidas en unidades de dinero, tiempo, distancia o peso, etc., la distribución que se va a utilizar es la llamada Distribución de Probabilidad Normal.

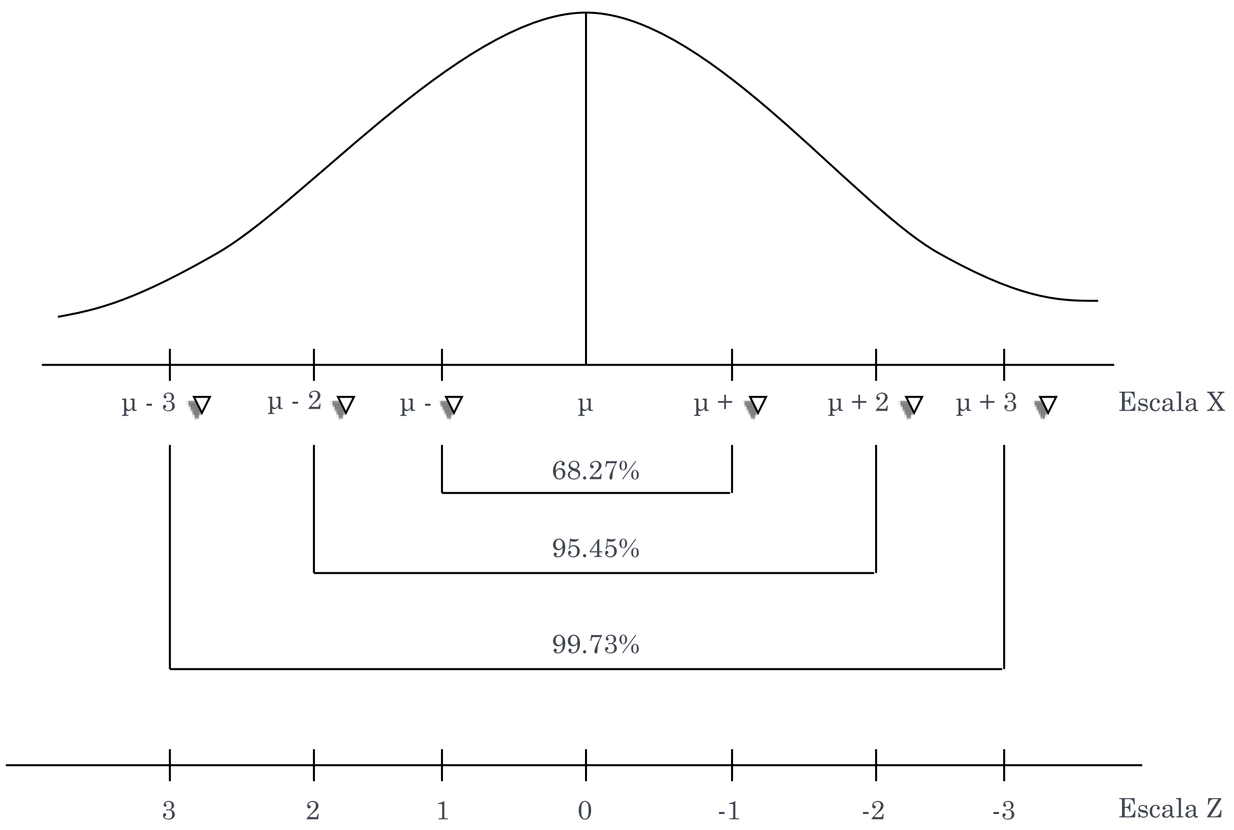
CARACTERÍSTICAS DE LA DISTRIBUCIÓN DE PROBABILIDAD NORMAL

Todos los tipos de distribución normal se representan en curvas simétricas, cada una de ellas con su respectiva media y desviación estándar. Para poder estudiar este tipo de distribuciones fue creada la distribución normal estandarizada, llamada así por que utiliza las puntuaciones estándar Z. Esta distribución estandarizada sigue las siguientes características:

1. La curva normal tiene perfil de campana y presenta un solo pico en el centro exacto de la distribución. La media aritmética, la mediana y la moda de la distribución son iguales y están en el punto central. De esta forma la mitad del área baja la curva se halla a un lado (o por encima del valor central) de ese punto. y la otra mitad, al otro lado (o por debajo). Es decir 50% de un lado y 50% del otro lado.
2. Por lo anterior podemos decir que es una curva simétrica perfecta con una media igual a cero y una desviación estándar de uno, se denomina Curva normal estándar ($\mu = 0$ y $= 1$).

∇

3. Está completamente definida a ± 3 ∇ de ahí se extiende a ± infinito, esto quiere decir que jamás toca el eje de las x.
4. Esta curva tiene su punto máximo en la media y tiene puntos de inflexión en ± 1 . ∇
5. Las probabilidades de que a cualquier observación quede a una, dos y tres desviaciones estándar ( ) a partir de ∇ la media (μ) son 68.27%, 95.45% y 99.73%, respectivamente.



FÓRMULA DE LA DISTRIBUCIÓN NORMAL ESTÁNDAR

$$Z = \frac{X - \mu}{\nabla}$$

En donde:

Z = Puntuaciones estándar

x = Valor observado

μ = Media poblacional de la distribución

∇ = Desviación estándar poblacional de la distribución

LA APROXIMACIÓN NORMAL PARA LA DISTRIBUCIÓN BINOMIAL Y DE POISSON

CARACTERÍSTICAS

La distribución normal de probabilidad da una aproximación bastante buena a la distribución binomial y de Poisson cuando el número de ensayos es grande. Una regla práctica nos dice que podemos usar la aproximación normal a la binomial y de Poisson cuando:

Aproximación normal para la distribución binomial	Aproximación normal para la distribución de Poisson
$n > 30$	$n > 30$
$p > 0.05$	p [0.05
$n\,p > 5$	$n\,p > 5$
$n\,q > 5$	$n\,q > 5$

FACTOR DE CORRECIÓN POR CONTINUIDAD

Puesto que la distribución normal maneja variables continuas y la distribución binomial y de Poisson variables discretas, se hace necesario corregir las variables discretas que manejan el modelo binomial y de Poisson y esto se logra restando o sumando ½ al valor de la variable aleatoria discreta.

FÓRMULA DE LA APROXIMACIÓN NORMAL PARA LA DISTRIBUCIÓN BINOMIAL Y DE POISSON

$$Z = \frac{(X \pm \tfrac{1}{2}) - \mu}{\nabla}$$

En donde:

x = Valor de la variable aleatoria discreta que se comporta de acuerdo al modelo binomial o de Poisson.

$\mu = n \cdot p$ = La media poblacional de una distribución binomial o de Poisson.

$\nabla = \sqrt{n \cdot p \cdot q}$ = La desviación estándar poblacional de una distribución binomial.

$\nabla = \sqrt{n \cdot p}$ = La desviación estándar poblacional de una distribución de Poisson.

Z = Valor de x en unidades de desviación estándar ya corregida.

½ = Factor de corrección por continuidad (F.C.C.)

TABLA PARA SABER CUANDO SUMAR O RESTAR EL

FACTOR DE CORRECCIÓN POR CONTINUIDAD

$P(X < X_0) = X - ½$
$P(X [X_0) = X + ½$
$P(X > X_0) = X + ½$
$P(X \mu X_0) = X - ½$
$P(X = X_0) = X \pm ½$

En donde:

X_0 = Cualquier valor mayor o igual a 0 ($X_0 = 0, 1, \ldots\ldots\ldots, n$)

X = Valor de la variable aleatoria bajo estudio.