Estadística 1

DISTRIBUCIONES DE PROBABILIDAD

La distribución normal de probabilidad

Las distribuciones de probabilidad continua pueden tomar varias formas…

La media μ localiza el centro de la distribución y la distribución es **simétrica** alrededor de su media μ.

Como **el área bajo la distribución normal de probabilidad es igual a 1**, la simetría implica que el área a la derecha de μ y el área a la izquierda de μ son iguales.

La forma de la distribución está determinada por σ, la desviación estándar de la población. Es decir, valores grandes de σ reducen la altura de la curva y aumentan la dispersión; valores pequeños de σ aumentan la altura de la curva y reducen la dispersión.

Ahora, para hallar la probabilidad de que una variable aleatoria normal x se encuentre en el intervalo de a a b, necesitamos encontrar el área bajo la curva normal entre los puntos a y b.

Existe un número infinitamente grande de distribuciones normales, uno para cada media y desviación estándar diferentes; dado que es sumamente impráctico usamos el proceso de estandarización que nos permite usar la misma tabla para todas las distribuciones normales.

La distribución aleatoria normal estándar

Una variable aleatoria normal x está estandarizada al expresar su valores como el número de desviaciones estándar σ que se encuentra a la izquierda o derecha de su media σ.

La variable aleatoria normal estándar, z, se define como,

$$z = \frac{x - \mu}{\sigma}$$

o bien, lo que es equivalente,

$$x = \mu + z\sigma$$

De la fórmula podemos deducir que:

- Cuando x es menor que la media μ, el valor de Z es NEGATIVO
- Cuando x es mayor que la media μ, el valor de Z es POSITIVO
- Cuando $x = \mu$, el valor de Z es CERO

En las tablas que nos brinda el Departamento tenemos el área bajo la curva normal estándar a la izquierda de un valor especifico de z, por ejemplo z_0, es la probabilidad $P(z \leq z_0)$.

La distribución normal estandarizada porque su media es 0 y su desviación estándar es 1.

Por ejemplo:

- Encuentre la $P(z > 0.5)$:

Sabemos que, por simetría $P(z > 0.5) = 1 - P(z \leq 0.5) = 1 - 0.3085 = 0.6915$

- Encuentre la $P(-1 \leq z \leq 1)$

Que es lo mismo:

$P(-1 \leq z \leq 1) = P(z \leq 1) - P(z \leq -1) = 0.8413 - 0.1587 = 0.6826$

- Encuentre la $P(-2 \leq z \leq 2)$

Que es lo mismo:

$P(-2 \leq z \leq 2) = P(z \leq 2) - P(z \leq -2) = 0.9772 - 0.0228 = 0.9544$

Ahora bien,

1. Sea x una variable aleatoria normalmente distribuida con media de 10 y desviación estándar de 2. Encuentre la probabilidad de que x se encuentre entre 11 y 13.6.

   Debemos, en un primer momento:

   $$P(11 \leq x \leq 13.6)$$

   $$P(x \leq 13.6) - P(x \leq 11)$$

   $$= P(\frac{x-\mu}{\sigma} \leq \frac{13.6-10}{2}) - P(\frac{x-\mu}{\sigma} \leq \frac{11-10}{2})$$

   $$= P(x \leq 1.8) - P(x \leq .5)$$

   $$= 0.9641.0.6915 = 0.2726$$

2. Estudios realizados recientemente demuestran que el uso de gasolina para autos híbridos vendidos en CDMX está normalmente distribuido, con media de 25.5 millas por galón (mpg) y una desviación estándar de 4.5 mpg. ¿Qué porcentaje de compactos recorre 30 mpg?

   Nos piden:

   $$P(x > 30)$$

   Entonces (estandarizamos)

   $$P(\frac{x-\mu}{\sigma} > \frac{30-25.5}{4.5})$$

   Por simetría

   $$1 - P(z \leq 1)$$

   $$= 1 - 0.8413 = 0.1587$$

Distribución geométrica

En teoría de probabilidad y estadística, la distribución geométrica es cualquiera de las dos distribuciones de probabilidad discretas siguientes:

1. Si $X = 1, 2, ...$ es el número necesario para obtener un éxito
2. Si $X = 0, 1, 2, ...$ es el número de fracasos antes del primer éxito

Notación:

Si una variable aleatoria discreta X sigue una distribución geométrica con parámetro $0 < p < 1$ entonces escribiremos $X \backsim Geometrica(p)$ o simplemente $X \backsim Geo(p)$

Función de probabilidad:

Si la variable aleatoria discreta X se usa para modelar el número de fracasos antes de obtener el primer éxito en una sucesión de ensayos independientes Bernoulli en donde cada uno de ellos la probabilidad de éxito es p, entonces la función de probabilidad de $X \backsim Geometrica(p)$ es

$$P[X = x] = p(1 - p)^x$$

Para valores de

$$x = 0, 1, 2, 3, ...$$

Ejemplo 1:

Suponga que cada una de sus llamadas a una estación de radio popular tiene una probabilidad de 0.02 de ser respondida. Asumiendo que las llamadas son independientes, ¿cuál es la probabilidad de que la respondan a la décima llamada?¿Cuál es el número medio de llamadas para conectar?

Sea x el número de llamadas a la estación hasta ser atendido.

Éxito: llamada respondida (p) Fracaso: llamada no respondida (1-p)

Sea $p = 0.02$ y $1 - p = 0.98$

Entonces:

$$f(x) = (1 - p)^{x-1}p$$

$$f(x) = (0.98)^{10-1}(0.02)$$

$$f(x) = 0.0167$$

Para calcular el valor esperado de una distribución geométrica sabemos que:

$$E(x) = \frac{1}{p}$$

Al sustituir:

$$E(x) = \frac{1}{0.02} = 50$$

Ejemplo 2:

La probabilidad de calibrar las llantas de un carro de acuerdo con las especificaciones de la marca es de 0.6. Asumiendo que los intentos de calibración son independientes, ¿cuál es la probabilidad de que a lo más tres intentos de calibración sean requeridos para satisfacer las especificaciones?

sea x = intentos para lograr la calibración.

$p = 0.6$ y $1 - p = 0.4$

Se quiere calcular la:

$$P(X \leq 3) = P(x = 1) + P(x = 2) + P(x = 3)$$

Sabemos que:

$$f(x) = (1-p)^{x-1}p$$

$$f(x=1) = (0.6)(0.4)^0 = 0.6$$
$$f(x=2) = (0.6)(0.4)^1 = 0.24$$
$$f(x=3) = (0.6)(0.4)^2 = 0.096$$

Entonces:

$$P(X \leq 3) = 0.936$$

Distribución Poisson

Media :

$$E(x) = \lambda$$

Varianza:

$$V(x) = \lambda$$

Es decir, tanto el valor esperado como la varianza de una variable aleatoria con distribución de Poisson son iguales a λ

Aproximación a una normal o estandarización

Como consecuencia del TCL, para valores grandes de λ de una variable aleatoria Poisson X puede aproximarse por otra normal dado que el cociente:

$$Y = \frac{X-\lambda}{\sqrt{\lambda}}$$