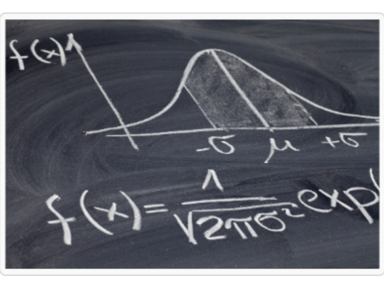
Tema 6. Distribuciones

Introducción

En este tema, conocerás y analizarás el concepto de variable aleatoria y, sobre todo, el de distribución de probabilidad; es importante que revises estos conceptos para que puedas elaborar tablas de distribución en la búsqueda de probabilidades. También entenderás en qué casos se manejan distribuciones discretas y cuándo se recurre a distribuciones continuas; asimismo, adquirirás las herramientas necesarias para interpretar cada una de ellas. Finalmente, estudiarás e identificarás las distribuciones binomial, de Poisson y normal; las dos primeras como distribuciones discretas y la normal como una continua, basada en un intervalo generado a partir de una medición.



Explicación

Tipos y cálculo de distribución

Variable aleatoria

De acuerdo con Zapata (2022), las variables aleatorias son recipientes donde se guardan los resultados de experimentos casuales, de ahí su calificativo. Estas variables, a su vez, se dividen en dos clases: discretas, cuando toman cierto número de valores específicos porque son el resultado de un conteo; y continuas, si surgen de un proceso de medición.

Por ejemplo, si te propones realizar el experimento de lanzar un dado, el espacio muestra sería $S = \{1, 2, 3, 4, 5, 6\}$, así que cada probabilidad de obtener una cara se define de esta manera:

P(A1) = Obtener la cara 1 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$
P(A2) = Obtener la cara 2 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$
P(A3) = Obtener la cara 3 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$
P(A4) = Obtener la cara 4 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$
P(A5) = Obtener la cara 5 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$
P(A6) = Obtener la cara 6 = $\frac{casos\ favorables}{casos\ posibles} = \frac{1}{6}$

No obstante, resulta más sencillo observar los resultados mediante una tabla de distribución de probabilidad, como se aprecia a continuación:

A1	A2	A3	A4	A5	A6
0.17	0.17	0.17	0.17	0.17	0.17

Tabla 1. Distribución de probabilidad del lanzamiento de un dado.

El nombre de esta tabla se debe a que muestra las probabilidades de un evento en particular, con el resultado de la probabilidad de variables aleatorias.

Distribución de probabilidad discreta

Distribución de Bernoulli

El experimento de Bernoulli es un modelo estadístico que consiste en dividir los eventos de un experimento en solo dos resultados posibles: el éxito y el "no éxito" (fracaso). Por ejemplo, si lanzas un dado y pretendes obtener la cara tres, el éxito radicaría en conseguirlo; por consiguiente, cualquier otro resultado se consideraría un "no éxito" o fracaso. Al calcular las probabilidades, se presenta este escenario: $P(A) = obtener\ la\ cara\ 3 = \frac{1}{6}$; este es el éxito, así que el "no éxito" estará dado por $P(1-A) = NO\ obtener\ la\ cara\ 3 = \frac{5}{6}$. Si llamamos p a la probabilidad de éxito y q a la de "no éxito", se puede decir que $q = 1 - p$. Entonces, para solucionar el caso del dado se procede de esta manera:

$$q = 1 - \frac{1}{6} = \frac{5}{6}$$

Distribución binomial

Establecida perfectamente en la teoría de Bernoulli, surge con la necesidad de esquematizar todas las probabilidades de un experimento que se repite una n cantidad de veces. Por ejemplo, tienes un examen de opción múltiple cuya calificación mínima aprobatoria es de 60 puntos y, además, sabes que cuenta con 10 preguntas, cada una con tres posibles respuestas diferentes, pero solo una correcta. ¿Cuál es la probabilidad de aprobar el examen si el experimento corresponde a una distribución binomial?

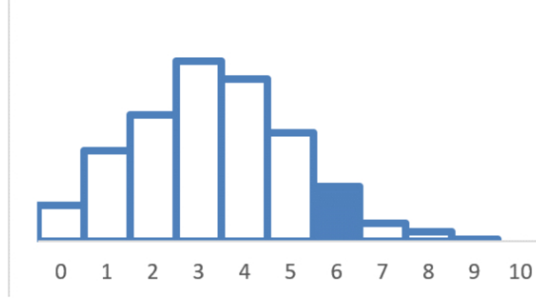
K	$P(x=k)$
0	0.0182
1	0.898
2	0.199
3	0.2614
4	0.2253
5	0.1332
6	0.0547
7	0.0154
8	0.0028
9	0.0003
10	0
Suma	1.00

Tabla 2. Distribución de probabilidad binomial de un examen de 10 preguntas.

En la tabla 2, se aprecia la distribución de probabilidad de cada acierto en el examen; en este caso, la probabilidad de que apruebes con 60 depende de que aciertes seis reactivos. Si revisas, en la tabla $k = 6$, es decir, k corresponde a la cantidad de preguntas contestadas correctamente; mientras tanto, $P(x = k)$ equivale a la probabilidad de que ocurra exactamente ese resultado. Entonces, la expectativa de aprobar con 60 se calcula así:

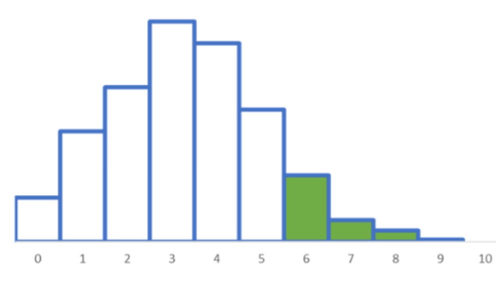
$$P(x = 6) = 0.0547.$$

En el ejemplo anterior, la distribución binomial se representa gráficamente como en la siguiente figura:



Gráfica 1. Representación gráfica para 6 aciertos.

Te habrás percatado de que cada barra corresponde a una posible cantidad de aciertos (de 0 a 10) y su altura está determinada por la probabilidad de que ocurra. Sin embargo, para resolver el ejercicio, se requiere que el alumno apruebe el examen, es decir, que reciba una calificación entre 6 y 10. Observa la siguiente gráfica:



Gráfica 2. Representación gráfica para 6 o más aciertos.

La solución se extrae al sumar las probabilidades entre 6 y 10:

$$P(x \geq 6) = P(X = 6) + P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10) =$$

$$P(x \geq 6) = 0.0547 + 0.0154 + 0.0028 + 0.0003 + 0 = 0.0732 = 7\%$$

De esta manera, la probabilidad de que el alumno apruebe el examen es del 7%. ¡Cuidado! Este porcentaje se determina sin que el individuo haya estudiado. Ahora, ¿cuál es la probabilidad de que el alumno no apruebe?

Se puede llegar a esta información de dos maneras:

a) Sumando las probabilidades:

$$P(X \leq 5) = 0.0182 + 0.0898 + 0.199 + 0.2614 + 0.2253 + 0.1332 = 0.9268 = 93\%$$

b) Usando la fórmula de Bernoulli:

$$q = 1 - p$$

Donde:
$p = Aprobar\ el\ examen = 7\%$.
$q = No\ aprobar\ el\ examen$.

$$q = 1 - 7\% = 93\%$$

Cualquiera de estos dos procedimientos es válido y te servirá para calcular ejercicios de distribuciones de probabilidad.

Distribución Poisson

La distribución de probabilidad de Poisson expresa las ocurrencias totales de un evento en un intervalo determinado, ya sea tiempo, área, volumen o distancia. Se trata de una distribución de probabilidad discreta, ya que sus resultados también se basan en un conteo (1, 2, 3, 4...*n*). Las aplicaciones de esta distribución son muy amplias, por ejemplo, se requieren para modelar situaciones de riesgo, para estimar tiempos de espera en transacciones de la bolsa, para evaluar pérdidas operativas, entre muchas otras.

Revisa el siguiente ejemplo:

a) Si un videojuego en línea recibe en promedio 10 quejas de *lag* o retraso en un día normal, ¿cuál es la probabilidad de que en uno reciba 6 quejas?

Este es un problema basado en distribuciones de probabilidad, así que puedes elaborar una tabla con base en los datos proporcionados. El resultado debe ser similar a este:

K	$P(x=k)$	K	$P(x=k)$
0	0	15	0.0347
1	0.0005	16	0.0217
2	0.0023	17	0.0128
3	0.0076	18	0.0071
4	0.0189	19	0.0037
5	0.0378	20	0.0019
6	0.0631	21	0.0009
7	0.0901	22	0.0004
8	0.1126	23	0.0002
9	0.1251	24	0.0001
10	0.1251	25	0
11	0.1137	26	0
12	0.0948	27	0
13	0.0729	28	0
14	0.0521	Suma	1.00

Tabla 3. Distribución de probabilidad Poisson de reportes de lag.

Contestar la pregunta conlleva un proceso muy simple, pues ya sabes que el promedio de quejas es de diez por día, así que $\mu = 10$; esta información es esencial para identificar el tipo de distribución. Ahora, mediante la tabla, encuentra el valor para $P(x = 6)$; entonces, la probabilidad de que haya seis quejas exactas en un día cualquiera se calcula de esta forma: $P(x = 6) = 0.0631$. En otros términos, existe un 6% de probabilidad de que en un día común se expresen seis inconformidades.

Ahora bien, ¿cuál será la probabilidad de que ocurran más de seis quejas? Como en el caso anterior, recurre a la tabla, aunque necesitarás otros valores:

$$P(x > 6) = P(x = 7) + P(x = 8) + \cdots .. P(x = 28)$$

Si simplificas esta ecuación y te apegas a los conceptos de Bernoulli, la solución se obtiene a partir de este procedimiento:

$$q = 1 - (P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5) + P(x = 6) =$$

$$q = 1 - (0 + 0.0005 + 0.0023 + 0.0076 + 0.0189 + 0.0378 + 0.0631) = 1 - 0.1302 = 0.8698 = 87\%$$

En conclusión, el videojuego tiene un 87% de probabilidad de recibir más de seis quejas en un día.

Distribución normal

Sin lugar a duda, la distribución de probabilidad más conocida y usada es la normal. De acuerdo con Rodó (2022), también se le denomina gaussiana o campana de Gauss debido a su gráfica, la cual simula una campana por su peculiar simetría. Además, forma parte de la distribución continua, a diferencia de la binomial y de Poisson que son discretas; esto significa que se le asignan valores a partir de un proceso de medición. Examina el siguiente ejemplo.

a) El tiempo de espera para entrar al servidor de un videojuego sigue una distribución normal, con una media de 3 minutos y una desviación estándar de un minuto. ¿Cuál será la probabilidad de que un usuario espere más de cuatro minutos para jugar?

Imagen 1. Campana de Gauss con m= 3 minutos.

En la imagen 1, se observa una gráfica semejante a una campana, donde el valor medio corresponde al promedio o media aritmética, en este caso, $\mu = 3\ minutos$. No obstante, se intenta determinar la probabilidad de que el jugador espere más de 4 minutos.

Imagen 2. Campana de Gauss, cola derecha.

En la imagen 2, por tanto, interesa el área marcada desde el minuto cuatro hacia la derecha, es decir, la probabilidad de que sea mayor o igual a esa cantidad.

Se puede llegar al valor de la probabilidad de tres maneras: primera, si se aplican conceptos de cálculo integral (área bajo la curva); segunda, con un software (simulador) matemático para ahorrar trabajo; tercera, mediante tablas para la curva normal tipificada.

Entonces, ¿qué será más probable?, ¿que el jugador espere menos de dos o más de cuatro minutos?

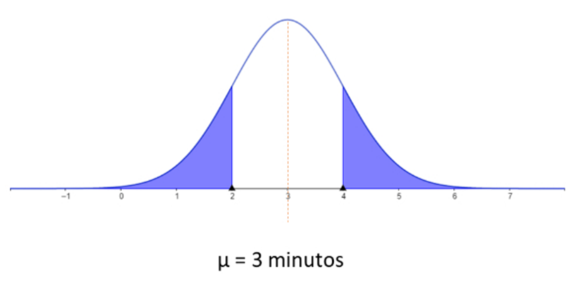


Imagen 3. Campana de Gauss, dos colas.

En este caso, las probabilidades son idénticas, ya que la campana de Gauss tiene forma simétrica. En otras palabras, existe los mismos valores en ambos extremos (izquierdo o derecho), los cuales equivalen a 0.5, ya que al sumarlos obtendrás el 1 o el 100% de probabilidad.

Cierre

A lo largo de este tema, estudiaste el concepto de las variables aleatorias discreta y continua. Esto te permitirá comprender los conceptos de distribuciones de probabilidad y, además, reconocer sus diferentes clases, como las distribuciones continuas o discretas, de acuerdo con la variable que representan. También observaste y revisaste sus tres modelos posibles: dos para las discretas (distribuciones binomial y Poisson) y una para las continuas (distribución normal). Finalmente, resolviste varios ejemplos básicos, ya que la práctica resulta fundamental para el cálculo y para la resolución de problemas aplicativos a la vida diaria.

Checkpoint

Asegúrate de:

- Entender los conceptos de las variables aleatorias discreta y continua para identificar el tipo de distribución que representan.
- Comprender los tipos de distribuciones para identificarlas en problemas específicos donde intervengan distribuciones de probabilidad.

Bibliografía

- Rodó, P. (2022). *Distribución normal*. Recuperado de https://economipedia.com/definiciones/distribucion-normal.html
- Zapata, F. (2022). *Distribución binomial: concepto, ecuación, características*. Recuperado de https://www.lifeder.com/distribucion-binomial/