UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA
DIVISIÓN DE ESTUDIOS DE POSGRADO

NOTAS DEL CURSO PROPEDEUTICO DE PROBABILIDAD

Presenta: Torres López Vanessa Cristell
Profesor: Dr. Ignacio Arrollo Fernández
Curso: Probabilidad
Fecha: septiembre de 2024

1. CONCEPTOS BÁSICOS

1.1 INTRODUCCIÓN

La probabilidad es un concepto fundamental en matemáticas y estadística que permite cuantificar la incertidumbre asociada con diferentes eventos. A lo largo del tiempo, han surgido varias interpretaciones y enfoques para definir este concepto. A continuación, se presentan tres de las definiciones más utilizadas en la teoría de probabilidad: la clásica, la frecuentista y la bayesiana. Cada una de ellas ofrece una perspectiva única sobre cómo entender y calcular la probabilidad, dependiendo del contexto y de las suposiciones involucradas.

Definición clásica de probabilidad	• La probabilidad de un evento es el número de resultados favorables al evento dividido por el número total de resultados de un experimento donde todos los resultados son igualmente probables (Ash, 1970)
Definición frecuentista de probabilidad	• La probabilidad de un evento es el límite de su frecuencia relativa conforme el número de ensayos tiende al infinito (Ross, 2014)
Definición bayesiana de probabilidad	• La probabilidad es una medida del grado de creencia o certeza que se tiene sobre la ocurrencia de un evento, la cual puede actualizarse a medida que se obtiene nueva información (Jaynes, 2003).

Ilustración 1. Diferentes definiciones de probabilidad

1.2 ESPACIO DE PROBABILIDAD

Un espacio de probabilidad se define como un trío $(\Omega, \mathscr{F}, P)$ donde:

1. ***Ω es el espacio muestral***: el conjunto de todos los posibles resultados de un experimento aleatorio. Es el conjunto sobre el cual se definen todos los eventos.
2. ***$\mathscr{F}$ es una σ-álgebra (sigma-álgebra) sobre Ω***: es una colección de subconjuntos de Ω que incluye el espacio muestral Ω mismo, es cerrada bajo complementos y bajo uniones numerables de sus elementos. La σ-álgebra define los eventos para los cuales se puede asignar una probabilidad.

3. ***P es la medida de probabilidad***: es una función que asigna a cada conjunto de un espacio de eventos un número entre 0 y 1, cumpliendo ciertas propiedades, como la no negatividad, la aditividad (la probabilidad de la unión de eventos disjuntos es la suma de sus probabilidades) y que la probabilidad del espacio muestral completo es 1. Formalmente, una medida de probabilidad es una función ***P*** definida en un espacio medible (Ω, $\mathscr{F}$) que satisface las siguientes condiciones:

   a. $P(A) \geq 0$ para todo $A \in \mathscr{F}$
   b. $P(\Omega)=1$
   c. Si $A_1, A_2, \ldots \in \mathscr{F}$ son disjuntos, entonces:

   $$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_n)$$

   d. Si A_1 y $A_2 \in \mathscr{F}$ no son disjuntos, entonces

   $$P(A_1 \cup A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2) - P(A_1 \cap A_2)$$

1.3 INDEPENDENCIA

Dos eventos A y B son independientes si y solo si la probabilidad de su intersección es igual al producto de sus probabilidades individuales. Esto se expresa como:

$$P(A \cap B) = P(A)\, P(B)$$

Ejemplo:

Supongamos que estamos lanzando dos dados, y definimos los siguientes eventos:

- *A: Obtener un número par en el primer dado.*
- *B: Obtener un número mayor que 3 en el segundo dado.*

Para estos eventos:

- *La probabilidad de obtener un número par en un dado es $P(A) = \frac{1}{2}$ ya que hay 3 números pares: 2, 4, 6.*
- *La probabilidad de obtener un número mayor que 3 en un dado es $P(B) = \frac{1}{2}$ ya que hay 3 números mayores que 3: 4, 5, 6.*

La probabilidad de que ambos eventos ocurran simultáneamente (el primer dado es par y el segundo dado es mayor que 3) es:

$$P(A \cap B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

1.4 PROBABILIDAD CONDICIONAL

Se define la **probabilidad condicional** como la probabilidad de que un evento ocurra dado que otro evento ya ha ocurrido, se denota por:

$$P(A|B) = \frac{P\ (A\ \cap B\ )}{P\ (B\ )}$$

Ejemplo

Supongamos que tenemos una urna con 3 bolas rojas y 2 bolas azules.

Si sacamos una bola al azar, la probabilidad de que sea roja es $P(R) = \frac{3}{5}$ y la probabilidad de que sea azul es $P(Z) = \frac{2}{5}$

Ahora, si sabemos que la bola sacada es azul, la probabilidad de que sea azul (condicional a que sea una bola) es:

$$P(A|B)$$

Si B representa el evento de que la bola sea azul y A el evento de que sea una bola de cualquier color, entonces:

$$P(A|B) = \frac{P\ (A\ \cap B\ )}{P\ (B\ )} = \frac{P\ (B\ )}{P\ (B\ )} = 1$$

Propiedades de la Probabilidad Condicional:

1. Regla de la Multiplicación: La probabilidad conjunta de dos eventos A y B se puede expresar usando la probabilidad condicional como:

$$P\ (A\ \cap B\ ) = P(A|B)\ P\ (A)$$

a) **Regla de la cadena**. La regla de la cadena se utiliza para calcular la probabilidad conjunta de una secuencia de eventos. Si tenemos una secuencia de eventos $A_1, A_2, \ldots\ A_n$ la probabilidad conjunta de que todos estos eventos ocurran es:

$$P\ (A_1 \cap\ A_2 \ldots\ A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap\ A_2) \cdots P(A_n|A_1 \cap\ A_2 \ldots A_{n-1})$$

2. Teorema de Bayes: La probabilidad condicional puede ser invertida usando el Teorema de Bayes:

$$P(A|B) = \frac{P\ (A,B\ )}{P\ (\ B\ )} = \frac{P(B|A)}{P(B)}$$

Que puede ser expresada en términos de la regla de la suma como:

$$P(A|B) = \frac{P\ (B|A)\ P(A)}{\sum_{i=1}^{n} P(B_i|A)\ P(A)}$$

2. VARIABLES ALEATORIAS

Una **variable aleatoria** es una función X que asigna un valor real a cada resultado en el espacio muestral Ω.

$$X :\ \Omega \rightarrow R$$

donde:

- **Ω** es el espacio muestral, que contiene todos los posibles resultados de un experimento aleatorio.
- **R** es el conjunto de los números reales.

2.1TIPOS DE VARIABLES ALEATORIAS

Variable Aleatoria Discreta:

Una variable aleatoria X es discreta si su conjunto de posibles valores $x_1, x_2, x_3, \ldots$ es finito o contablemente infinito. Para cada valor x_i, existe una probabilidad $P(X) = x_i$ que satisface:

1. $P(X = x_i) \geq 0$ Para todo x_i
2. La suma de las probabilidades es igual a 1:

$$\sum_i P(X = x_i) = 1$$

- Ejemplos:

Número de caras al Lanzar una moneda tres veces

- **Variable Aleatoria**: El número de veces que aparece "cara" al lanzar una moneda justa tres veces.
- **Valores Posibles:** 0, 1, 2, o 3 (puede haber 0 caras, 1 cara, 2 caras o 3 caras).
- **Distribución Probabilística:** Esto puede modelarse con una distribución binomial, donde el número de éxitos (caras) en una serie de ensayos independientes (lanzamientos) es la variable de interés.

Número de llamadas recibidas en una hora en un centro de atención

- **Variable Aleatoria**: El número de llamadas que recibe un centro de atención telefónica en una hora determinada.
- **Valores Posibles:** 0, 1, 2, 3, ...
- **Distribución Probabilística:** Esto puede modelarse con una distribución de Poisson, que es apropiada para contar el número de eventos que ocurren en un intervalo de tiempo fijo.

Número de hijos en una familia

- **Variable Aleatoria**: El número de hijos en una familia seleccionada al azar.
- **Valores Posibles**: 0, 1, 2, 3, 4, ...
- **Distribución Probabilística**: Esto puede modelarse con una distribución de Poisson o una distribución empírica basada en datos censales reales.

Ilustración 2. Ejemplos de variables aleatorias discretas

Variable Aleatoria Continua:

Una variable aleatoria continua X es una función tal que:

1. Puede tomar cualquier valor dentro de un rango continuo de números reales, por ejemplo, en el intervalo [a, b].
2. La probabilidad de que X tome un valor específico es 0, es decir, $P(X = x_i) = 0$ para cualquier valor x_i.
3. La probabilidad de que X tome un valor dentro de un intervalo [a, b] se obtiene integrando la **función de densidad de probabilidad (FDP)** $f_x(x)$ sobre ese intervalo:

$$P(a \leq X \leq b) = \int_a^b f_x\,(x)\,dx$$

Donde $f_x(x) \geq 0$ para todo x, y la integral de $f_x(x)$ en todo el dominio es 1:

$$\int_{-\infty}^{\infty} f_x\,(x)\,dx = 1$$

- Ejemplos:

Tiempo de vida de una bombilla	Altura de una persona	Temperatura en un día dado
• **Variable Aleatoria**: El tiempo (en horas) que la bombilla funciona antes de quemarse. • **Tipo de Distribución:** Esto podría modelarse con una distribución normal si la vida útil promedio es conocida y las variaciones siguen un patrón simétrico alrededor de la media. • **Valores Posibles:** Cualquier número real positivo (el tiempo que puede durar la bombilla).	• **Variable Aleatoria**: La altura de una persona seleccionada al azar de una población. • **Tipo de Distribución:** distribución normal • **Valores Posibles:** Cualquier número real positivo, dentro de un rango razonable.	• **Variable Aleatoria**: La temperatura registrada en un lugar específico a lo largo de un día. • **Tipo de Distribución**: Esto puede modelarse con una distribución normal o log-normal, dependiendo de cómo se distribuyen las temperaturas en esa región y la variabilidad histórica. • **Valores Posibles**: Cualquier número real dentro de un rango definido por las temperaturas extremas de esa región.

Ilustración 3. Ejemplos de variables aleatorias continuas

2.1DISTRIBUCIONES DE PROBABILIDAD

Existen varios tipos de distribuciones de probabilidad**,** que se dividen principalmente en dos grandes categorías: **distribuciones discretas** y **distribuciones continuas**, dependiendo de si las variables aleatorias toman valores discretos o continuos.

[Distribuciones de probabilidad discretas]

• **Distribución Binomial**:

Esta función de probabilidad discreta mide la probabilidad del número de resultados k favorables para un evento en n experimentos que son independientes entre sí. En este caso, también p es una probabilidad e indica la probabilidad de que un resultado sea favorable.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

De la ecuación anterior, tenemos el **coeficiente combinatorio**, que es el número de posibilidades al escoger k objetos de un total de n objetos sin importar el orden.

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Ejemplo.

Supongamos que lanzamos una moneda justa 10 veces. Queremos calcular la probabilidad de obtener exactamente 4 caras.

Aquí, n=10, p=0.5, y queremos encontrar P(X=4). Usamos la fórmula de la función de probabilidad binomial:

$$P(X=4) = \binom{10}{4} 0.5^4 (1-0.5)^{10-4}$$

$$P(X=4) \approx 0.205$$

Esto significa que hay aproximadamente un 20.5% de probabilidad de obtener exactamente 4 caras en 10 lanzamientos de una moneda justa.

• **Distribución de Poisson**:

La Distribución de Poisson es una distribución de probabilidad discreta que modela el número de eventos que ocurren en un intervalo fijo de tiempo, área, volumen u otro tipo de unidad, cuando estos eventos ocurren de manera independiente y a una tasa constante.

La función de probabilidad de la distribución de Poisson está dada por:

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Donde λ es el parámetro de la distribución que representa la tasa promedio de ocurrencia de los eventos en el intervalo dado.

Ejemplo.

Supongamos que en una tienda de comestibles se observa que, en promedio, llegan 3 clientes por hora. Si λ=3, podemos usar la distribución de Poisson para calcular la probabilidad de que lleguen exactamente 5 clientes en una hora.

Usamos la fórmula de la función de probabilidad de Poisson:

$$P(X=5) = \frac{3^5 e^{-3}}{5!} \approx 0.1008$$

Esto significa que hay aproximadamente un 10.08% de probabilidad de que lleguen exactamente 5 clientes en una hora.

• **Distribución Geométrica**:

La Distribución Geométrica es una distribución de probabilidad discreta que modela el número de ensayos necesarios para obtener el primer éxito en una serie de ensayos de Bernoulli.

La función de probabilidad de la distribución geométrica está dada por:

$$P(X = k) = (1 - p)^{k-1}p$$

Ejemplo

Supongamos que la probabilidad de que una máquina funcione correctamente en cualquier momento dado es p=0.2. Si estamos interesados en saber cuántos intentos serán necesarios para que la máquina funcione correctamente por primera vez, podemos usar la distribución geométrica.

La probabilidad de que la máquina funcione correctamente en el tercer intento se calcula usando la fórmula de la función de probabilidad:

$$P(X = 3) = (1 - 0.2)^{3-1}0.2 = 0.128$$

Esto significa que hay una probabilidad del 12.8% de que la máquina funcione correctamente por primera vez en el tercer intento.

Distribuciones de probabilidad continuas

• **Distribución Normal o Gaussiana**:

La Distribución Gaussiana, también conocida como Distribución Normal, es una de las distribuciones de probabilidad más importantes en estadística y probabilidades. Su forma característica es una curva en forma de campana, a menudo llamada "curva de campana de Gauss".

La función de densidad de probabilidad de la distribución normal se define como:

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

• **Distribución Exponencial**:

La distribución exponencial se utiliza para modelar el tiempo entre eventos en un proceso de Poisson, es decir, para modelar el tiempo hasta el primer evento en una serie de eventos que ocurren a una tasa constante.

La función de densidad de probabilidad que la define es:

$$f(x) = \lambda e^{-\lambda x}$$

• **Distribución Uniforme Continua**:

La distribución uniforme continua describe una variable aleatoria que tiene la misma probabilidad de tomar cualquier valor dentro de un intervalo especificado. Es una distribución en la que todos los resultados posibles son igualmente probables.

La función de densidad de probabilidad que define a la distribución uniforme continua es:

$$f(x) = \frac{1}{b-a} \;\; para\; a \leq x \leq b$$

2.2. DENSIDAD MARGINAL

La densidad marginal de una variable en un espacio de n dimensiones es la función de densidad que se obtiene al efectuar la sumatoria (en caso discreto) o integrar (en caso continuo) la función de densidad conjunta sobre las variables restantes.

Densidad marginal en el caso discreto

Para una función de probabilidad conjunta discreta $P(X_1, X_2, \dots, X_n)$ la probabilidad marginal de una variable específica X_i se calcula sumando la probabilidad conjunta sobre todas las demás variables. La probabilidad marginal de X_i se define como:

$$P(X_i = x_i) = \sum_{x_1} \sum_{x_2} \dots \sum_{x_i} \sum_{x_{i+1}} \dots \sum_{x_n} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

Caso bidimensional

Consideremos en caso bidimensional con dos variables aleatorias discretas X y Y. La probabilidad marginal de X se calcula sumando las probabilidades conjuntas para todos los valores posibles de Y:

$$p_x(x) = P(X = x) = \sum_{y \in \Omega Y} p_{xy}(x, y)$$

De manera similar, la probabilidad marginal de Y se calcula sumando las probabilidades conjuntas para todos los valores posibles de X:

$$p_y(y) = P(Y = y) = \sum_{x \in \Omega X} p_{xy}(x, y)$$

Densidad marginal en el caso continuo

Para una función de densidad conjunta $f_{X_1, X_2, \dots X_n}(x_1, x_2, \dots x_n)$ que describe la distribución conjunta de las variables aleatorias $X_1, X_2, \dots X_n$, la densidad marginal de una variable X_i se obtiene integrando la función de densidad conjunta sobre todas las demás variables.

$$f_{X_i}(x_i) = P(x_i \epsilon B) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1, \dots, dx_n$$

Caso bidimensional

Si tenemos dos variables aleatorias X y Y con una función de densidad conjunta $f_{X,Y}(x, y)$, la densidad marginal de X se calcula integrando sobre y:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

De manera similar, la densidad marginal de Y se obtiene integrando sobre x:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

2.3. VALOR ESPERADO

Sea X una variable aleatoria discreta que toma valores en un conjunto finito **Ω**, entonces el valor esperado de X se define como:

$$\mathbb{E}|x| = \sum_{x \epsilon \Omega} x \, P(X = x)$$

En particular si X es constante, X = c para todo xϵΩ.

Para el caso de una variable aleatoria continua se tiene que

$$\mathbb{E}|x| = \int_{-\infty}^{\infty} x \, f_X \, (x) \, dx$$

Algunas propiedades del Valor Esperado

1. Si $\mathbb{E}|x| = \infty$ o $\mathbb{E}|x| = -\infty$ entonces el valor esperado de x no existe.
2. Si $\mathbb{E}|x| = b$, entonces $\mathbb{E}|x|$ minimiza a $\mathbb{E}|(x-b)^2|$
3. Sean X, Y variables aleatorias con expectaciones finitas, entonces:
   - Si g(x) ≥ h(x) para todo x∈ℝ entonces $\mathbb{E}|g(x)| \geq \mathbb{E}|h(x)|$.
   - $\mathbb{E}|aX + bY + c| = \alpha\mathbb{E}|X| + \beta\mathbb{E}|Y| + c$ para cualesquiera a,b,c ∈ℝ.
4. Sean X, Y variables aleatorias que cumplen:
   $$\mathbb{E}|XY| = \mathbb{E}|X|\mathbb{E}|Y|$$
   O bien,
   $\mathbb{E}|g(X)h(Y)| = \mathbb{E}|g(X)|\mathbb{E}|h(Y)|$ para cualesquiera funciones reales g, h.

2.4. MOMENTOS DE UNA VARIABLE ALEATORIA

Los momentos de una variable aleatoria son una extensión del concepto de valor esperado que proporcionan información más detallada sobre la distribución de la variable. Los momentos ayudan a describir características como la tendencia central, la dispersión y la forma de la distribución.

Momento de orden k

El k-ésimo momento de una variable aleatoria X, denotado por $\mathbb{E}|XY|$, es el valor esperado de la k-ésima potencia de X. Se expresa como:

$$\mathbb{E}|X^k| = \sum_i x_i^k \, P(X = x_i)$$

En el caso de las variables aleatorias continuas, el momento de orden k se expresa como:

$$\mathbb{E}|X^k| = \int_{-\infty}^{\infty} x^k \, f_X\,(x)\, dx$$

Momento central de orden k

El momento central de orden k se refiere al valor esperado de la k-ésima potencia de la desviación de la variable aleatoria respecto a su valor esperado. Para una variable aleatoria X con valor esperado $\mathbb{E}|X|$, el momento central de orden k se define como:

$$\beta_k \; = \; \mathbb{E}|(X-m)^k|$$

Si m es finito y $\mathbb{E}|X|$ existe.

Si X es discreto:

$$\beta_k \; = \sum_{x \epsilon \Omega} (X-m)^k f_X(x)$$

Si X es continuo:

$$\beta_k \; = \int_{-\infty}^{\infty} (X-m)^k \, f_X\,(x)\, dx$$

Nótese que si k = 1, entonces el momento central obtenido representa la desviación estándar de X, además, cuando x=m, entonces $\beta_1 \; = 0$

$$\beta_1 \; = \; \mathbb{E}|(X-m)|$$

Si k = 2, entonces el momento central obtenido representa la varianza de X.

$$\beta_2 \; = \; \mathbb{E}|(X-m)^2|$$

Teorema Binomial

Cuando se calcula un momento central de orden k, es necesario evaluar expresiones de la forma $|(X-m)^k|$, donde X es una variable aleatoria y $\mathbb{E}|X|$ es su valor esperado. El teorema binomial permite expandir estas expresiones de forma más manejable.

El teorema binomial aplicado a $\beta_k \; = \; \mathbb{E}|(X-m)^k|$ proporciona la siguiente expansión:

$$\beta_k = \mathbb{E}|(X-m)^k| = \sum_{k=0}^{n} \binom{n}{k} (-m)^{n-k} \alpha_k$$

Para el caso particular de la varianza (k=2), el momento central de segundo orden, se necesita calcular $\beta_2 = \mathbb{E}|(X-m)^2|$. Utilizando el teorema binomial:

$$\mathbb{E}|(X-m)^2| = X^2 - 2m\mathbb{E}|X| + (m)^2$$

Simplificando:

$$Var\, X = \mathbb{E}|X^2| - (\mathbb{E}|X|)^2$$

Momentos centrales conjuntos

Los momentos centrales conjuntos para dos variables aleatorias X y Y son extensiones de los momentos centrales de una sola variable aleatoria, que describen cómo las desviaciones de las variables respecto a sus medias están relacionadas conjuntamente. Se definen como:

$$\alpha_{jk} = \mathbb{E}\left|x_1^j\, x_2^k\right| \text{ para todo j, k >0}$$

Así mismo, los momentos centrales conjuntos

$$\beta_{jk} = \mathbb{E}\left|(X_1 - m_1)^j\, (X_2 - m_2)^k\right|$$

Para el caso específico de β_{11}

$$\beta_{11} = \mathbb{E}|(X_1 - m_1)\,(X_2 - m_2)|$$
$$= \mathbb{E}|X_1 X_2| - \mathbb{E}|X_1|\mathbb{E}|X_2|$$
$$= Cov(X_1, X_2)$$

Coeficiente de correlación

Si $\mathbb{E}|X_1^2|$ y $\mathbb{E}|X_2^2|$ son finitos, las varianzas γ_1^2 y γ_2^2 respectivas son mayores que cero, entonces:

$$\rho(X_1, X_2) = \frac{Cov\,(X_1, X_2)}{\gamma_1^2\, \gamma_2^2}$$

Nota: Si X_1 y X_2 son independientes, entonces $\rho(X_1, X_2) = 0$. Lo contrario no siempre se cumple.

Enlace GitHub:

https://github.com/CristellTL/Probabilidad_Ash-1970-

Referencias

Ash, R. B. (1970). *Basic probability theory*. Dover Publications. ISBN 978-0-486-46628-6.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Ross, S. M. (2014). *Introduction to probability and statistics for engineers and scientists* (5th ed.). Academic Press.