



Inferencia **Estadística**

Sesión 3

Variables Aleatorias y sus Tipos

Variables Aleatorias Discretas

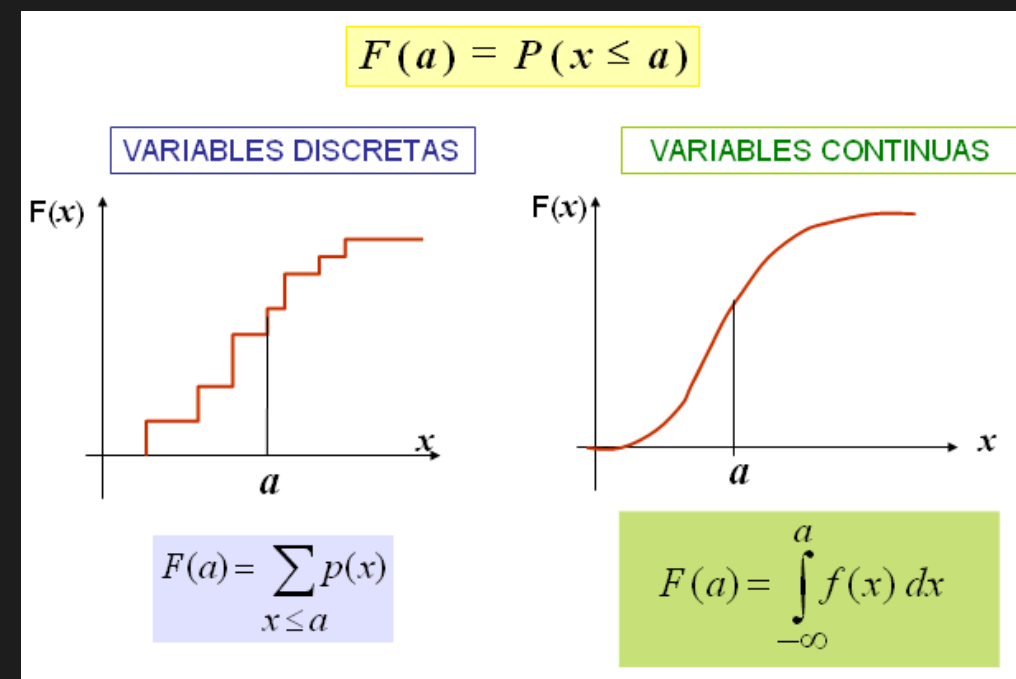
Toman un número finito o infinitamente numerable de valores distintos. Se utilizan para modelar situaciones donde los resultados ocurren en valores separados y específicos.

Ejemplos: Número de caras al lanzar una moneda tres veces (valores posibles: 0, 1, 2, 3), número de clientes que llegan a un negocio en una hora, cantidad de productos defectuosos en una línea de producción.

Variables Aleatorias Continuas

Pueden tomar cualquier valor dentro de un intervalo determinado de números reales. Modelan situaciones donde los resultados pueden variar en una escala infinita y continua.

Ejemplos: Tiempo de espera en una fila, altura de estudiantes en una escuela, temperatura registrada en una ciudad durante el día, distancia recorrida por un vehículo antes de necesitar mantenimiento.



Distribución de Probabilidad

Distribución para Variables Discretas

La función de probabilidad asigna a cada valor χ de la variable X una probabilidad específica: $P(X = \chi)$. Debe cumplir que $0 \leq P(X = \chi) \leq 1$ y la suma de todas las probabilidades es 1.

Ejemplo: Al lanzar un dado justo de seis caras, la probabilidad de cada número es $1/6$.

Distribución para Variables Continuas

Se utiliza la función de densidad de probabilidad (FDP). La probabilidad de que X tome un valor exacto es 0, por lo que se calcula la probabilidad de que X caiga dentro de un intervalo.

Ejemplo: Si la estatura de adultos se distribuye normalmente con media 170 cm y desviación estándar 10 cm, se usa la función de distribución normal.

Importancia

Las distribuciones de probabilidad permiten predecir resultados en experimentos aleatorios, modelar fenómenos inciertos en múltiples disciplinas y realizar inferencias estadísticas para tomar decisiones basadas en datos.

Distribución Acumulada de Probabilidad

$$F(x) = P(X \leq x)$$

Definición

La función de distribución acumulada (CDF) describe la probabilidad acumulada de que una variable aleatoria X tome un valor menor o igual a un valor específico x : $F(x) = P(X \leq x)$.

Ejemplo Discreto

Para el número de caras al lanzar una moneda dos veces ($X = \{0, 1, 2\}$), la CDF sería: $F(0) = 1/4$, $F(1) = 3/4$, $F(2) = 1$.

1

2

3

4

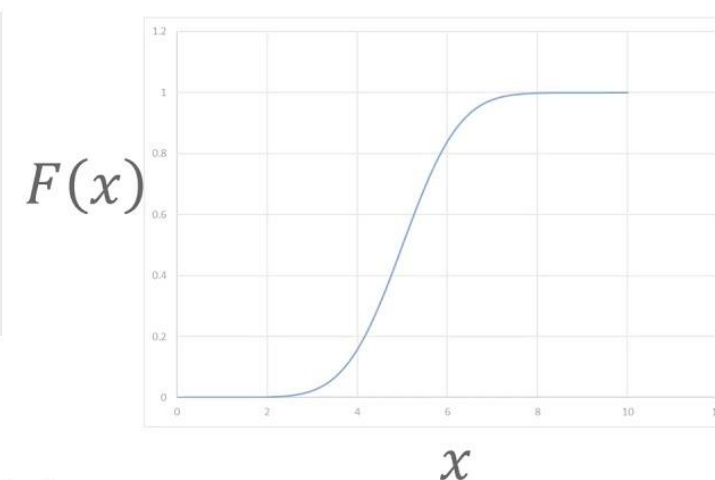
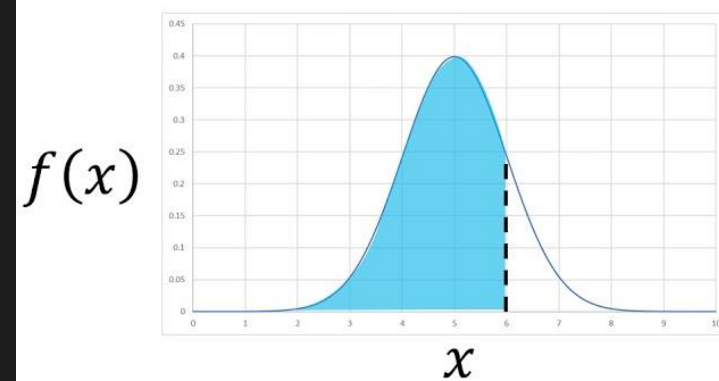
Propiedades

Es monótonamente creciente, comienza en $F(-\infty) = 0$ y termina en $F(\infty) = 1$. Para variables discretas, la CDF tiene saltos en los valores de X , mientras que para variables continuas es una función suave y continua.

Ejemplo Continuo

Si el tiempo de espera sigue una distribución exponencial con $\lambda=1$, la probabilidad de esperar menos de 3 minutos es $F(3) = 1 - e^{-3} = 0.95$, es decir, 95%.

Función de distribución acumulada



$$P(X \leq x) = \int_{-\infty}^x f(u) du = F(x)$$

Distribuciones Discretas



Uniforme Discreta

Todos los resultados tienen la misma probabilidad: $P(X = x) = 1/n$. Ejemplo: Al lanzar un dado justo, $P(X = 3) = 1/6$. Se usa en juegos de azar y muestreo aleatorio.



Binomial

Describe el número de éxitos en n repeticiones independientes de un experimento de Bernoulli. Ejemplo: Probabilidad de obtener 4 caras en 10 lanzamientos de moneda. Se usa en encuestas y control de calidad.



Bernoulli

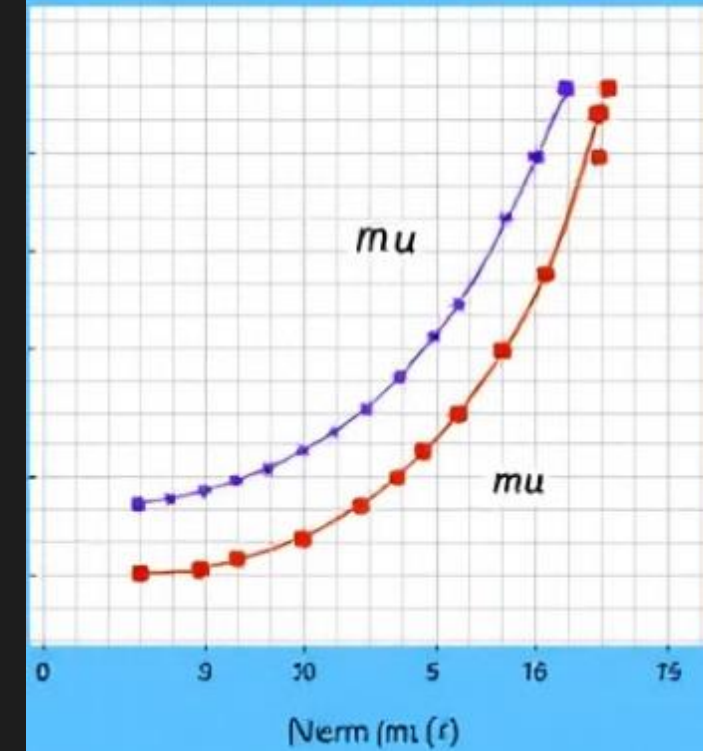
Modela experimentos con dos resultados: éxito (1) o fracaso (0). $P(X = 1) = p$, $P(X = 0) = 1-p$. Ejemplo: Lanzar una moneda sesgada con probabilidad de cara 0.7. Se usa en pruebas de calidad y clasificación binaria.



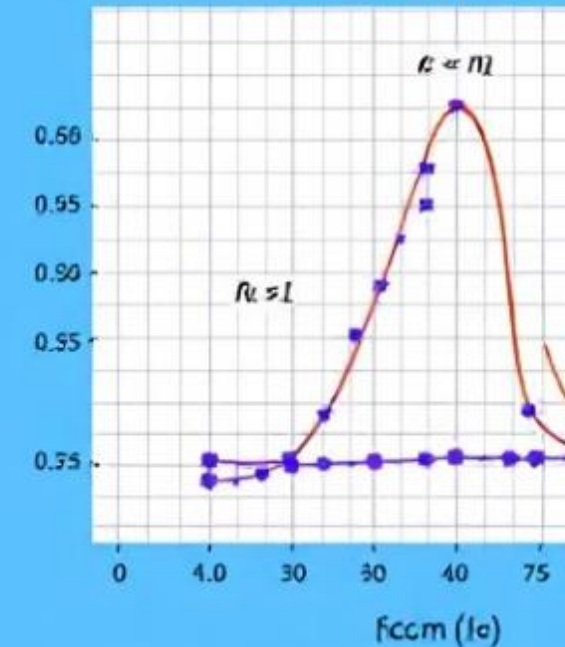
Poisson

Modela el número de eventos en un intervalo de tiempo o espacio. Ejemplo: Si un call center recibe 5 llamadas por hora, la probabilidad de recibir 3 llamadas es $P(X = 3) = (5^3 \cdot e^{-5})/3! = 0.14$.

Normal

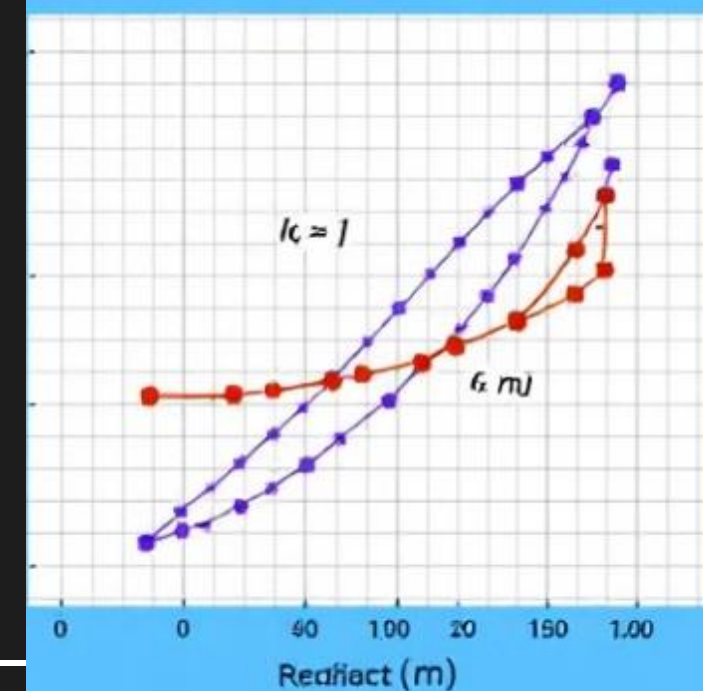


Binomial

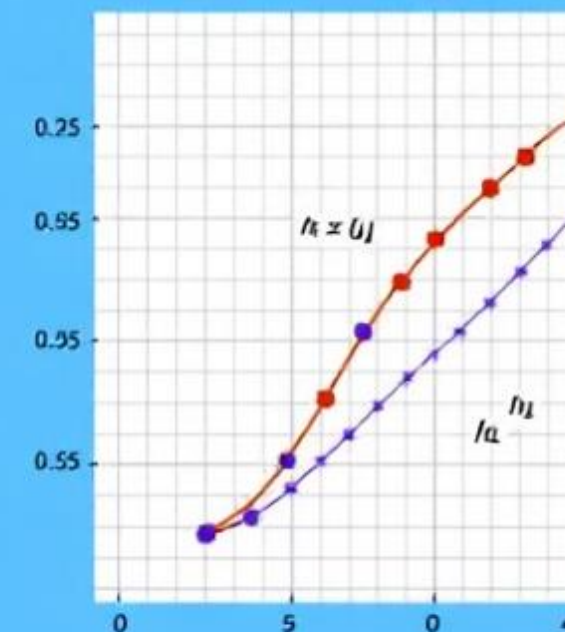


ability Distributions

Binomial



Poisson



Distribuciones Continuas

1

Uniforme Continua

Todos los valores dentro de un intervalo $[a,b]$ tienen la misma densidad de probabilidad: $f(x) = 1/(b-a)$. Ejemplo: Tiempo de espera en un semáforo distribuido uniformemente entre 0 y 60 segundos. Se usa en simulaciones y distribución equitativa de recursos.

2

Normal

La más importante en estadística, con forma de campana. Definida por su media μ y desviación estándar σ . Se usa en procesos biológicos, físicos y económicos. Es simétrica alrededor de la media.

3

Chi-Cuadrado, t-Student y F de Fisher

Utilizadas en pruebas de hipótesis, análisis de varianza y evaluación de independencia. La t-Student se usa con muestras pequeñas, mientras que la F de Fisher modela la razón de dos varianzas muestrales.

4

Exponencial

Modela el tiempo entre eventos en un proceso de Poisson: $f(x) = \lambda e^{(-\lambda x)}$. Ejemplo: Si en un banco llegan clientes cada 5 minutos en promedio, el tiempo entre llegadas sigue una distribución exponencial con $\lambda = 1/5$.

Recomendaciones para la Elección de una Distribución

1 Discreta vs. Continua

Para datos finitos o contables (número de clientes), use distribuciones discretas. Para valores que pueden tomar infinitos valores dentro de un intervalo (alturas), use distribuciones continuas.

3 Pruebas de Bondad de Ajuste

Verifique si una distribución se ajusta a los datos mediante pruebas como Kolmogorov-Smirnov, Chi-cuadrado o Shapiro-Wilk (para normalidad).

2 Contexto del Problema

Elija según la naturaleza del fenómeno: Binomial para procesos de éxito/fracaso, Poisson para cantidad de eventos en un intervalo, Normal para datos con tendencia central, Exponencial para tiempos entre eventos.

4 Visualización y Ajuste

Analice los datos gráficamente con histogramas, QQ-Plots y boxplots. Si conoce la distribución subyacente, estime sus parámetros; si no, considere métodos no paramétricos.

Distribución Normal: Características

Forma de Campana

Su gráfica tiene forma de campana y es completamente simétrica alrededor de la media. La mayor parte de los valores se agrupan cerca de la media.

Propiedades Adicionales

Es simétrica, asintótica (nunca toca el eje X) y tiene su máximo en la media.



Media, Mediana y Moda

En una distribución normal, la media (μ), la mediana y la moda coinciden, ubicándose en el centro de la distribución.

Regla 68-95-99.7

El 68% de datos están dentro de 1σ de la media, el 95% dentro de 2σ , y el 99.7% dentro de 3σ .

Forma Funcional de la Distribución Normal

1

Función de Densidad

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

2

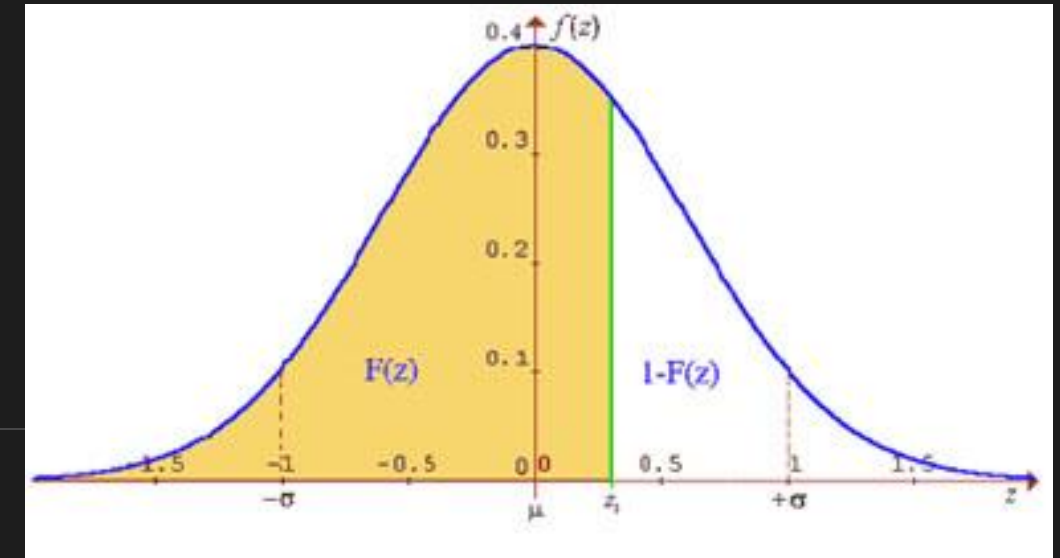
Parámetros

μ es la media, σ es la desviación estándar

3

Interpretación

Valores cercanos a μ tienen mayor probabilidad



La función de densidad de probabilidad de la distribución normal describe la probabilidad relativa de cada valor dentro de la distribución. Cuando x está cerca de la media (μ), la función tiene su valor máximo, indicando que los valores más comunes se encuentran alrededor de la media.

A medida que x se aleja de la media, la función disminuye exponencialmente, mostrando que valores extremos son menos frecuentes. La curva es simétrica respecto a la media, lo que significa que la probabilidad de obtener un valor por encima o por debajo de la media es la misma.

Cálculo de Probabilidad con la Distribución Normal

1

Estandarizar

Convertir a distribución normal estándar usando
 $Z = (X - \mu) / \sigma$

2

Consultar

Buscar probabilidad en tabla Z o usar software estadístico

3

Interpretar

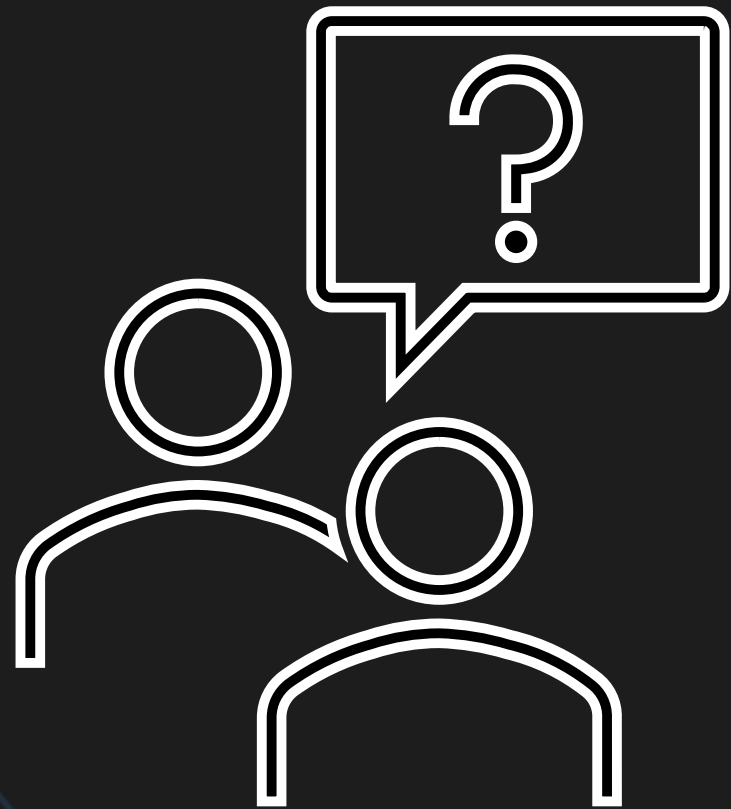
Calcular $P(X \leq a)$, $P(X \geq a)$ o $P(a \leq X \leq b)$ según necesidad

Ejemplo: Para calcular $P(X \leq 70)$ con $\mu = 65$ y $\sigma = 5$, primero estandarizamos: $Z = (70 - 65) / 5 = 1.00$. Luego buscamos en la tabla Z: $P(Z \leq 1.00) = 0.8413$, lo que significa que hay un 84.13% de probabilidad de que X sea menor o igual a 70.

Para la distribución normal estándar ($\mu=0$, $\sigma=1$), podemos usar directamente las tablas Z. Por ejemplo, $P(Z \leq 1.96) = 0.975$, lo que significa que el 97.5% de los valores están por debajo de $Z=1.96$. Esta propiedad es útil en pruebas de hipótesis con nivel de significancia del 5%.

Preguntas

Sección de preguntas



A background network diagram consisting of numerous small blue nodes connected by thin, light blue lines, creating a complex web-like structure. The nodes are more densely packed in some areas and more sparse in others, with some nodes appearing slightly brighter than others.

Inferencia **Estadística**

Continúe con las
actividades