## SESIÓN DISTRIBUCIÓN DE PROBABILIDAD

**CONTENIDOS:**

- Variables aleatorias y sus tipos.
- Distribución de probabilidad.
- Distribución acumulada de probabilidad.
- Distribuciones discretas:
  - Uniforme discreta.
  - Bernoulli.
  - Binomial.
  - Poisson.
- Distribuciones continuas:
  - Uniforme continua.
  - Normal.
  - Normal estándar.
  - Chi-cuadrado.
  - T-student.
  - Fischer.
  - Exponencial.
- Recomendaciones para la elección de una distribución.
- Distribución normal:
  - Características de la distribución normal.
  - Forma funcional de la distribución normal.
  - Cálculo de probabilidad con la distribución normal.
  - Distribución normal estándar.

### VARIABLES ALEATORIAS Y SUS TIPOS

Las variables aleatorias son fundamentales en la teoría de la probabilidad y la estadística, ya que permiten cuantificar fenómenos inciertos y analizarlos matemáticamente. Una variable aleatoria es una función que asigna un valor numérico a cada resultado posible de un experimento aleatorio.

Dependiendo de la naturaleza de los valores que puede tomar, las variables aleatorias se dividen en discretas y continuas.

**Variable Aleatoria Discreta**

Una variable aleatoria discreta solo puede tomar un número finito o infinitamente numerable de valores distintos. Se utiliza para modelar situaciones donde los resultados ocurren en valores separados y específicos.

Ejemplo:

El número de caras obtenidas al lanzar una moneda tres veces. En este caso, los valores posibles de la variable aleatoria son {0, 1, 2, 3}, ya que una moneda solo puede salir cara o cruz en cada lanzamiento.

Otras aplicaciones:

- Número de clientes que llegan a un negocio en una hora.
- Cantidad de productos defectuosos en una línea de producción.
- Número de veces que un dado cae en el número 6 en 10 lanzamientos.

**Variable Aleatoria Continua**

Una variable aleatoria continua puede tomar cualquier valor dentro de un intervalo determinado de números reales. Se utiliza para modelar situaciones donde los resultados pueden variar en una escala infinita y continua.

Ejemplo:

El tiempo de espera en una fila. Este valor puede ser 2 minutos, 2.5 minutos, 3.78 minutos o cualquier otro valor dentro del rango de tiempo posible.

Otras aplicaciones:

- Altura de los estudiantes en una escuela.
- Temperatura registrada en una ciudad durante el día.
- Distancia recorrida por un vehículo antes de necesitar mantenimiento.

**Importancia de las Variables Aleatorias en la Ciencia de Datos**

Las variables aleatorias permiten modelar y analizar datos en múltiples disciplinas, incluyendo economía, ingeniería, salud y Machine Learning. La correcta clasificación y manejo de estas variables es clave para seleccionar los modelos matemáticos adecuados y realizar predicciones precisas.

Además, las variables aleatorias están estrechamente relacionadas con las distribuciones de probabilidad, que describen la forma en que los valores de una variable se distribuyen en un conjunto de datos. Algunas distribuciones comunes son:

- Distribución Binomial (para variables discretas).
- Distribución Normal (para variables continuas).
- Distribución de Poisson (para eventos que ocurren en un intervalo de tiempo).

**DISTRIBUCIÓN DE PROBABILIDAD**

La distribución de probabilidad es una herramienta matemática fundamental en la estadística y la ciencia de datos. Describe cómo se asignan las probabilidades a los diferentes valores de una variable aleatoria, permitiendo modelar y analizar fenómenos inciertos. Existen dos tipos principales de distribuciones de probabilidad: discretas y continuas, dependiendo de la naturaleza de la variable aleatoria.

**1. Distribución de Probabilidad para Variables Discretas**

Las variables aleatorias discretas toman valores específicos y contables, como el número de clientes que llegan a una tienda en una hora o el número de veces que una moneda cae en cara en cinco lanzamientos.

**Función de Probabilidad (FP)**

La función de probabilidad asigna a cada posible valor χ de la variable aleatoria X una probabilidad específica:

$$P(X = x)$$

*Ilustración 1 Función de probabilidad*

Esta probabilidad debe cumplir dos condiciones:

- 0 ≤ P(X = χ) ≤ 1, es decir, cada probabilidad es un numero entre 0 y 1.
- La suma de todas las probabilidades posibles es 1:

$$\sum P(X = x) = 1$$

*Ilustración 2 Suma de probabilidades*

Ejemplo:

Supongamos que un dado justo de seis caras se lanza una vez. La variable aleatoria χ representa el número obtenido. La distribución de probabilidad es:

$$P(X = x) = \begin{cases} \frac{1}{6}, & x = 1, 2, 3, 4, 5, 6 \\ 0, & \text{para otros valores de } x \end{cases}$$

*Ilustración 3 Distribución de probabilidad*

Distribuciones Discretas Comunes:

- Binomial: Modela el número de éxitos en una serie de ensayos independientes con dos posibles resultados (éxito o fracaso).
- Poisson: Modela la cantidad de eventos que ocurren en un intervalo de tiempo o espacio dado.

**2. Distribución de Probabilidad para Variables Continuas**

Las variables aleatorias continuas pueden tomar cualquier valor dentro de un rango. Ejemplos incluyen la altura de una persona, la temperatura de una ciudad o el tiempo de espera en una fila.

**Función de Densidad de Probabilidad (FDP)**

Para variables continuas, la probabilidad de que X tome un valor exacto es 0. En su lugar, se calcula la probabilidad de que X caiga dentro de un intervalo usando la función de densidad de probabilidad (FDP):

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

*Ilustración 4 Función de densidad de probabilidad*

Donde $f(x)$ es la función de densidad, y el área bajo la curva entre $a$ y $b$ representa la probabilidad de que $X$ tome un valor dentro de ese rango.

Ejemplo:

La variable aleatoria X representa la estatura de adultos en una población. Si se distribuye normalmente con media 170 cm y desviación estándar de 10 cm, su FDP sería la función de la distribución normal:

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

*Ilustración 5 Función con datos*

donde $\mu=170$ y $\sigma=10$.

Distribuciones Continuas Comunes:

- Normal: Modela fenómenos naturales como la altura, el peso y el rendimiento en exámenes.
- Exponencial: Modela el tiempo entre eventos en un proceso de Poisson, como el tiempo de espera en una fila.

**Importancia de las Distribuciones de Probabilidad**

Las distribuciones de probabilidad permiten:

- Predecir resultados en experimentos aleatorios.
- Modelar fenómenos inciertos en múltiples disciplinas.
- Realizar inferencias estadísticas y tomar decisiones basadas en datos.

**DISTRIBUCIÓN ACUMULADA DE PROBABILIDAD**

La función de distribución acumulada (CDF, por sus siglas en inglés) es una herramienta fundamental en la teoría de probabilidades y la estadística. Describe la probabilidad acumulada de que una variable aleatoria X tome un valor menor o igual a un valor específico $x$.

**Definición Matemática**

La función de distribución acumulada de una variable aleatoria X se define como:

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

*Ilustración 6 Función de distribución acumulada*

Esto significa que F(x) representa la suma de probabilidades de todos los valores que X puede tomar hasta $x$

Propiedades de la CDF:

1. Monótonamente creciente: $F\,(x)$ nunca disminuye, ya que la probabilidad acumulada solo puede aumentar o permanecer constante.
2. Valores limites:

$$\lim_{x \to -\infty} F(x) = 0 \text{ (la probabilidad acumulada empieza en 0).}$$

*Ilustración 7  Valor limite que empieza en 0*

$$\lim_{x \to \infty} F(x) = 1 \text{ (al considerar todos los valores posibles, la probabilidad total es 1).}$$

*Ilustración 8   Valor límite con probabilidad 1*

3. Para una variable discreta, la CDF tiene saltos en los valores de X.
4. Para una variable continua, la CDF es una función suave y continua.

**Ejemplo con Variable Aleatoria Discreta**

Supongamos que X es el número de caras al lanzar una moneda dos veces. Sus valores posibles son X = {0, 1, 2}, y su distribución de probabilidad es:

$$P(X=0)=\frac{1}{4}, \quad P(X=1)=\frac{2}{4}, \quad P(X=2)=\frac{1}{4}$$

*Ilustración 9 Ejemplo con variable aleatoria discreta*

La CDF se calcula como:

$$F(x)=\begin{cases} 0, & x<0 \\ \frac{1}{4}, & 0\leq x<1 \\ \frac{3}{4}, & 1\leq x<2 \\ 1, & x\geq 2 \end{cases}$$

*Ilustración 10 Calculo de CDF*

Interpretación:

- F(0) = P(X ≤ 0) = P(X = 0) = $\frac{1}{4}$
- F(1) = P(X ≤ 1) = P(X = 0) + P(X = 1) = $\frac{1}{4}+\frac{2}{4}=\frac{3}{4}$
- F(2) = P(X ≤ 2) = 1, porque ya se han acumulado todas las probabilidades.

**Ejemplo con Variable Aleatoria Continua**

Si X es el tiempo de espera en una fila y sigue una distribución exponencial con parámetro λ=1, la CDF se define como:

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

*Ilustración 11 Ejemplo con variable aleatoria continua*

Interpretación:

Si queremos conocer la probabilidad de esperar menos de 3 minutos, calculamos:

$$F(3) = 1 - e^{-1(3)} = 1 - e^{-3} \approx 0.95$$

*Ilustración 12 Interpretación con variable de 3 minutos*

Esto significa que hay un 95% de probabilidad de que el tiempo de espera sea menor a 3 minutos.

**Importancia de la Función de Distribución Acumulada**

- Permite calcular probabilidades para intervalos de valores de manera sencilla.
- Es útil para comparar distribuciones y analizar la probabilidad de eventos en datos reales.
- Se utiliza en modelos estadísticos y aprendizaje automático para realizar inferencias sobre datos.

**DISTRIBUCIONES DISCRETAS**

Las distribuciones de probabilidad discretas describen situaciones en las que las variables aleatorias pueden tomar un conjunto finito o infinito numerable de valores. Se utilizan ampliamente en estadística, ciencia de datos y modelado probabilístico para analizar fenómenos que ocurren en el mundo real.

**Distribución Uniforme Discreta**

La distribución uniforme discreta se aplica cuando todos los resultados de un experimento tienen la misma probabilidad de ocurrir.

Fórmula:

$$P(X = x) = \frac{1}{n}$$

*Ilustración 13 Función distribución uniforme discreta*

donde **n** es el número total de posibles resultados.

Ejemplo:

Si se lanza un dado justo de 6 caras, cada número tiene una probabilidad de P(X = x) = $\frac{1}{6}$

Así, la probabilidad de obtener un 3 es:

$$P(X = 3) = \frac{1}{6}$$

*Ilustración 14 Ejemplo distribución uniforme discreta*

Características:

- Se usa en juegos de azar, simulaciones y muestreo aleatorio.
- La probabilidad es constante para todos los valores posibles.

**Distribución de Bernoulli**

La distribución de Bernoulli modela experimentos con dos posibles resultados: éxito (1) o fracaso (0).

Fórmula:

$$P(X = x) = p^x (1 - p)^{1-x}, \quad x \in \{0, 1\}$$

*Ilustración 15 Fórmula distribución de Bernoulli*

Donde p es la probabilidad de éxito y 1 – p la probabilidad d fracaso.

Ejemplo:

Si lanzamos una moneda sesgada donde la probabilidad de que salga cara es 0.7 y de cruz es 0.3, entonces:

$$P(X = 1) = 0.7, \quad P(X = 0) = 0.3$$

*Ilustración 16 Ejemplo distribución de Bernoulli*

Características:

- Se usa en pruebas de calidad, encuestas y clasificación binaria en machine learning.
- Es la base de la distribución binomial.

**Distribución Binomial**

La distribución binomial describe el número de éxitos en **n** repeticiones de un experimento de Bernoulli, donde cada prueba es independiente.

Fórmula:

$$P(X = k) = C(n, k) \cdot p^k \cdot (1 - p)^{n-k}$$

*Ilustración 17 Fórmula distribución binomial*

Donde:

- C (n, k) es el coeficiente binominal $\frac{n!}{k!(n-k)!}$
- n es el número total de ensayos.
- k es el número de éxitos deseados.
- p es la probabilidad de éxito en cada ensayo.

Ejemplo:

Si lanzamos 10 veces una moneda equilibrada, la probabilidad de obtener exactamente 4 caras es:

$$P(X = 4) = C(10, 4) \cdot 0.5^4 \cdot 0.5^6$$

*Ilustración 18 Ejemplo distribución binomial*

Características:

- Se usa en encuestas, genética y control de calidad.
- Modela situaciones donde hay un número fijo de intentos y cada uno es independiente.

**Distribución de Poisson**

La distribución de Poisson se usa para modelar el número de eventos que ocurren en un intervalo de tiempo o espacio, suponiendo que ocurren de manera aleatoria e independiente.

Fórmula:

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

*Ilustración 19 Fórmula distribución de Poisson*

Donde λ es la tasa de ocurrencia del evento en el intervalo analizado.

Ejemplo:

Si un call center recibe en promedio 5 llamadas por hora, la probabilidad de recibir exactamente 3 llamadas en una hora es:

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

*Ilustración 20 Ejemplo distribución de Poisson*

Características:

- Se usa en análisis de tráfico, fallos en sistemas y ocurrencia de eventos raros.
- La variable aleatoria no tiene un límite superior, ya que teóricamente pueden ocurrir infinitos eventos.

**DISTRIBUCIONES CONTINUAS**

Las distribuciones continuas describen variables aleatorias que pueden tomar infinitos valores dentro de un intervalo. Son fundamentales en estadística y ciencia de datos para modelar fenómenos en los que las mediciones pueden asumir valores fraccionarios o cualquier punto dentro de un rango.

**Distribución Uniforme Continua**

La distribución uniforme continua se aplica cuando todos los valores dentro de un intervalo tienen la misma densidad de probabilidad.

Fórmula:

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

*Ilustración 21 Fórmula distribución uniforme continua*

Donde:

- a y b son los límites inferior y superior del intervalo.
- La probabilidad de cualquier subintervalo es proporcional a su longitud.

Ejemplo:

Si el tiempo de espera en un semáforo está distribuido uniformemente entre 0 y 60 segundos, entonces:

$$f(x) = \frac{1}{60}, \quad 0 \leq x \leq 60$$

*Ilustración 22 Ejemplo distribución continua*

Características:

- Se usa en simulaciones, distribución equitativa de recursos y tiempos de espera.
- No tiene sesgo, ya que todos los valores son igualmente probables.

**Distribución Normal**

La distribución normal, o de Gauss, es la más importante en estadística debido a su aplicación en numerosos fenómenos naturales. Tiene forma de campana y está definida por su media **μ** y desviación estándar **σ**.

Fórmula:

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

*Ilustración 23 Fórmula distribución normal*

Ejemplo:

Las alturas de las personas en una población suelen seguir una distribución normal con una media y una dispersión específicas.

Características:

- Se usa en procesos biológicos, físicos y económicos.
- Simétrica alrededor de la media.
- Regla empírica: Aproximadamente el 68% de los valores caen dentro de 1 desviación estándar, el 95% dentro de 2, y el 99.7% dentro de 3.

**Distribución Normal Estándar**

La distribución normal estándar es una normal con media μ=0 y desviación estándar σ=1. Se utiliza para calcular probabilidades mediante la tabla Z.

Fórmula:

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

*Ilustración 24 Fórmula distribución normal estándar*

Uso:

Permite calcular probabilidades sin necesidad de conocer μ y σ, utilizando la transformación estándar:

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

*Ilustración 25 transformación estándar*

**Distribución Chi-Cuadrado**

La distribución chi-cuadrado se usa para modelar la suma de los cuadrados de variables normales estándar independientes.

Uso:

- Pruebas de hipótesis para determinar si una distribución sigue un patrón esperado.
- Análisis de varianza y evaluación de independencia en tablas de contingencia.

Ejemplo:

Se usa en la prueba de independencia de chi-cuadrado para analizar si dos variables categóricas están relacionadas.

**Distribución t-Student**

La distribución t-Student se usa cuando el tamaño de la muestra es pequeño y la varianza poblacional es desconocida.

Uso:

- Pruebas t para comparar medias de dos muestras pequeñas.
- Intervalos de confianza para medias con muestras reducidas.

Ejemplo:

Comparar el rendimiento académico de dos grupos de estudiantes con una muestra reducida.

Características:

Similar a la distribución normal, pero con colas más largas, lo que refleja mayor incertidumbre en muestras pequeñas.

**Distribución F de Fisher**

La distribución F de Fisher modela la razón de dos varianzas muestrales y se usa en análisis de varianza.

Uso:

- ANOVA (Análisis de Varianza) para comparar múltiples medias.
- Evaluación de relaciones entre variables en regresión lineal.

Ejemplo:

Comparar la varianza de rendimiento de tres grupos de empleados en una empresa.

Características:

- Asimétrica, con valores positivos.
- Usada en comparación de modelos estadísticos.

**Distribución Exponencial**

La distribución exponencial modela el tiempo entre eventos en un proceso de Poisson.

Fórmula:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

*Ilustración 26 Fórmula distribución exponencial*

Donde λ es la tasa de ocurrencia de los eventos.

Ejemplo:

Si en un banco llegan clientes en promedio cada 5 minutos, el tiempo entre llegadas sigue una distribución exponencial con $\lambda = \frac{1}{5}$

Características:

- No tiene memoria: La probabilidad de ocurrencia de un evento no depende de cuánto tiempo ha pasado.
- Se usa en modelado de tiempos de espera y fallos de sistemas.

**RECOMENDACIONES PARA LA ELECCIÓN DE UNA DISTRIBUCIÓN**

Seleccionar la distribución de probabilidad adecuada es clave para modelar correctamente los datos y realizar inferencias precisas. A continuación, se presentan algunos criterios fundamentales para elegir la distribución más adecuada en cada situación.

**Discreta vs. Continua**

- Datos discretos: Si los valores posibles son finitos o contables (por ejemplo, número de clientes en una tienda), se usa una distribución discreta (Binomial, Poisson, Bernoulli).
- Datos continuos: Si los valores pueden tomar infinitos valores dentro de un intervalo (por ejemplo, alturas de personas), se usa una distribución continua (Normal, Exponencial, Chi-cuadrado).

**Contexto del Problema**

Es fundamental elegir una distribución que se ajuste a la naturaleza del fenómeno estudiado:

- Distribución Binomial: Para procesos de éxito/fracaso en un número fijo de ensayos.
- Distribución Poisson: Para modelar la cantidad de eventos en un intervalo de tiempo o espacio (por ejemplo, número de llamadas en un call center).

- Distribución Normal: Para datos que siguen un comportamiento natural con tendencia central y dispersión (como el peso de personas).
- Distribución Exponencial: Para modelar tiempos entre eventos en un proceso de Poisson (por ejemplo, tiempo entre llegadas de clientes).

**Pruebas de Bondad de Ajuste**

Para verificar si una distribución específica se ajusta a los datos, se pueden aplicar pruebas estadísticas:

- Prueba de Kolmogorov-Smirnov (KS): Evalúa la diferencia entre la distribución empírica de los datos y la distribución teórica.
- Prueba de Chi-cuadrado: Compara la frecuencia observada con la esperada para determinar si los datos siguen una distribución dada.
- Prueba de Shapiro-Wilk: Especialmente útil para verificar si los datos siguen una distribución normal.

**Visualización de Datos**

Antes de elegir una distribución, es recomendable analizar los datos gráficamente:

- Histograma: Para observar la forma de la distribución.
- Gráficos QQ-Plot: Para comparar los cuantiles de los datos con los de una distribución teórica.
- Boxplot: Para detectar asimetrías y valores atípicos.

**Ajuste Paramétrico vs. No Paramétrico**

- Si se conoce la distribución subyacente: Se pueden estimar sus parámetros (media, varianza) para modelar los datos.
- Si no se conoce la distribución: Se pueden emplear métodos no paramétricos como estimadores de densidad Kernel para aproximar la distribución de los datos.

**DISTRIBUCIÓN NORMAL**

La distribución normal, también conocida como distribución de Gauss, es una de las distribuciones de probabilidad más importantes en estadística y ciencia de datos. Se utiliza para modelar una gran variedad de fenómenos naturales, como alturas de personas, puntuaciones en exámenes y errores de medición. Además, es la base de numerosos métodos estadísticos e inferenciales.

**Características de la distribución normal.**

Las características de la función normal son fundamentales para entender su comportamiento y aplicaciones. A continuación, se describen sus propiedades clave:

**1. 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, mientras que los valores extremos son menos frecuentes.

**2. Media, Mediana y Moda**

- En una distribución normal, la media ($\mu$), la mediana y la moda coinciden, ubicándose en el centro de la distribución.
- Esto significa que la curva es perfectamente simétrica, sin sesgo hacia la derecha o la izquierda.

**3. Regla Empírica (Regla 68-95-99.7)**

- Aproximadamente el 68% de los datos se encuentran dentro de una desviación estándar ($\sigma$) de la media.
- Aproximadamente el 95% de los datos se encuentran dentro de dos desviaciones estándar ($2\sigma$) de la media.
- Aproximadamente el 99.7% de los datos se encuentran dentro de tres desviaciones estándar ($3\sigma$) de la media.

- Esta propiedad es clave en análisis estadísticos, ya que permite hacer inferencias sobre la probabilidad de ocurrencia de ciertos valores.

**Propiedades Adicionales**

- Simetría: La distribución es completamente simétrica en torno a la media.
- Asintótica: La curva nunca toca el eje X, extendiéndose infinitamente en ambas direcciones.
- Máxima en la Media: El punto más alto de la curva se encuentra en la media, donde hay mayor concentración de valores.

## FORMA FUNCIONAL DE LA DISTRIBUCIÓN NORMAL

La distribución normal, también conocida como distribución de Gauss, se define mediante una función de densidad de probabilidad (PDF, por sus siglas en inglés), la cual describe la probabilidad relativa de cada posible valor de la variable aleatoria.

**Función de Densidad de Probabilidad (PDF) de la Distribución Normal**

La función de la distribución normal es:

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

*Ilustración 27 Función de la distribución normal*

Donde:

- μ es la media.
- σ es la desviación estándar.
- π y e son constantes matemáticas.

Esta función describe la probabilidad relativa de cada valor dentro de la distribución.

**Interpretación de la Función**

- Cuando x está cerca de la media (μ), la función tiene su valor máximo, lo que indica 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**

El cálculo de probabilidades en una distribución normal se basa en estandarizar los valores a la distribución normal estándar y luego utilizar tablas o software estadístico para obtener la probabilidad deseada.

**Pasos para calcular probabilidades con la distribución normal**

1. Convertir a la distribución normal estándar
   Se usa la fórmula de estandarización:

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

*Ilustración 28 Fórmula de estandarización*

Donde:

- X es el valor de la variable aleatoria.
- μ es la media de la distribución.
- σ es la desviación estándar.
- Z es el puntaje Z correspondiente.

2. Buscar la probabilidad en la tabla Z
    - Una vez obtenido el valor de Z, se consulta en una tabla Z o se usa software estadístico para encontrar P(Z ≤ z), que representa la probabilidad acumulada hasta ese punto.
    - Alternativamente, se pueden usar herramientas como Excel, Python (usando scipy.stats.norm.cdf(z)) o calculadoras científicas.

Interpretar el resultado

- Si se busca P(X ≤ a), el resultado de la tabla Z da la respuesta directa.
- Si se busca P(X ≥ a), se usa:

$$P(X \geq a) = 1 - P(X \leq a)$$

*Ilustración 29 Fórmula p(x ≥ a )*

- Para probabilidades entre dos valores P(a≤X≤b), se calcula

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a)$$

*Ilustración 30 Fórmula para probabilidades entre dos valores*

Ejemplo:

Calcular P(X ≤ 70) para una normal con μ = 65 y σ = 5.

Paso 1: Estandarizar X = 70 usando la fórmula

$$Z = \frac{70 - 65}{5} = \frac{5}{5} = 1$$

*Ilustración 31 Ejemplo fórmula con datos*

Paso 2: Buscar en la tabla Z el valor P (Z ≤ 1.00)

- De la tabla, P (Z ≤ 1.00) = 0.8413
- Esto significa que hay un 84.13% de probabilidad de que X sea menor o igual a 70.

Paso 3: Interpretación

Si quisiéramos calcular P (Z ≥ 70), restamos:

$$P(X \geq 70) = 1 - 0.8413 = 0.1587$$

*Ilustración 32 Fórmula para interpretación*

Es decir, hay un 15.87% de probabilidad de que X sea mayor que 70.

**DISTRIBUCIÓN NORMAL ESTÁNDAR**

La distribución normal estándar es una normal con media $\mu=0$ y desviación estándar $\sigma=1$, lo que permite usar tablas Z de manera directa.

Ejemplo:

Calcular $P(Z \leq 1.96) = 0.975$.

- De la tabla Z, $P(Z \leq 1.96) = 0.975$.
- Esto significa que el 97.5% de los valores están por debajo de $Z=1.96$.
- Si quisiéramos calcular $P(Z \geq 1.96)$, restamos:

$$P(Z \geq 1.96) = 1 - 0.975 = 0.025$$

*Ilustración 33 Fórmula para calcular P(Z ≥ 1.96),*

Es decir, solo el 2.5% de los valores están por encima de $Z=1.96$.

Aplicación:

Esta propiedad es útil en pruebas de hipótesis, donde el valor crítico en una prueba de significancia del 95% es precisamente $Z=1.96$.