Función Generadora de Momentos y Varianza de una Variable Aleatoria

Nexus-Probability

CURSO 1 (PROBABILIDAD I)

PARTE 5 / LECCIÓN 3

1. Introducción

En probabilidad y estadística, el estudio de las funciones generadoras de momentos y la varianza de una variable aleatoria permite analizar sus propiedades y comportamiento. Estas herramientas son esenciales para la inferencia y modelado de datos en diversas aplicaciones matemáticas y actuariales.

2. Funciones Generadoras de Momentos

La función generadora de momentos (FGM) de una variable aleatoria X es una función definida como:

$$M_X(t) = E[e^{tX}]$$

siempre que la esperanza exista en un entorno de $t = 0$. Esta función es útil porque proporciona una manera sistemática de calcular los momentos de la distribución de X, es decir:

$$E[X^n] = M_X^{(n)}(0)$$

Es decir, derivando $M_X(t)$ n veces y evaluando en $t = 0$, obtenemos el n-ésimo momento de X.

3. Varianza

La varianza de una variable aleatoria X mide la dispersión de los valores de X respecto a su media $E[X]$. Se define como:

$$\text{Var}(X) = E[(X - E[X])^2]$$

Utilizando la propiedad de la esperanza, también se puede expresar como:

$$\text{Var}(X) = E[X^2] - (E[X])^2$$

4. Principales Variables Aleatorias y sus Propiedades

Distribución Bernoulli

- Definición: Una variable aleatoria X sigue una distribución Bernoulli con parámetro p si toma valores 1 con probabilidad p y 0 con probabilidad $1-p$.
- FGM:
$$M_X(t) = (1-p) + pe^t$$
- Media: $E[X] = p$
- Varianza: $\text{Var}(X) = p(1-p)$

Distribución Binomial

- Definición: Una variable aleatoria X sigue una distribución Binomial $B(n,p)$ si representa el número de éxitos en n ensayos de Bernoulli independientes.
- FGM:
$$M_X(t) = (1-p+pe^t)^n$$
- Media: $E[X] = np$
- Varianza: $\text{Var}(X) = np(1-p)$

Distribución Poisson

- Definición: Una variable aleatoria X sigue una distribución Poisson con parámetro λ si representa el número de eventos en un intervalo de tiempo fijo con tasa promedio λ.
- FGM:
$$M_X(t) = e^{\lambda(e^t-1)}$$
- Media: $E[X] = \lambda$
- Varianza: $\text{Var}(X) = \lambda$

Distribución Normal

- Definición: Una variable aleatoria X sigue una distribución Normal con media μ y varianza σ^2, denotada como $N(\mu, \sigma^2)$.

- FGM:

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

- Media: $E[X] = \mu$
- Varianza: $\text{Var}(X) = \sigma^2$

Distribución Exponencial

- Definición: Una variable aleatoria X sigue una distribución Exponencial con parámetro λ si modela el tiempo entre eventos en un proceso de Poisson.
- FGM:

$$M_X(t) = \frac{\lambda}{\lambda - t}, \quad \text{para } t < \lambda$$

- Media: $E[X] = \frac{1}{\lambda}$
- Varianza: $\text{Var}(X) = \frac{1}{\lambda^2}$

Ejercicio 1 *Genera 1000 datos aleatorios de una distribución normal con media 50 y varianza 9.*

- *Estima la media y la varianza de los datos generados.*
- *Ajusta un histograma y compara con la distribución teórica.*

Aquí se estudia la varianza en un conjunto de datos generados artificialmente.

- Se generan 1000 valores aleatorios con media 50 y varianza 9.
- Se estima la varianza empírica y se compara con la teórica.
- Se representa un histograma para visualizar la distribución de los datos.

```

# Ejercicio 1: Estimaci n de Varianza en Datos Simulados
samples_var = np.random.normal(50, np.sqrt(9), 1000)
estimated_var = np.var(samples_var)
```

Ejercicio 2 *Dado un conjunto de datos con una relación lineal* $y = 3x + \epsilon$ *donde* $\epsilon \sim N(0, 1)$*:*

- *Simula 500 pares* (x, y) *y ajusta una regresión lineal.*
- *Estima la varianza del error residual y compárala con la varianza teórica.*

Este ejercicio simula una relación lineal con errores aleatorios y analiza la varianza de los residuos.

- Se generan 500 pares (x, y) con una relación $y = 3x + \epsilon$.
- Se ajusta una regresión lineal y se calculan los residuos.
- Se estima la varianza de los residuos y se compara con la teórica.

```
# Ejercicio 2: Regresi n y Varianza de Errores
np.random.seed(42)
n_samples = 500
x = np.random.uniform(0, 10, n_samples)
epsilon = np.random.normal(0, 1, n_samples)
y = 3*x + epsilon
slope, intercept, _, _, _ = stats.linregress(x, y)
residuals = y - (slope*x + intercept)
residual_var = np.var(residuals)
```

Ejercicio 3 *Si $X \sim U(0, 1)$:*

- *Calcula teóricamente su varianza.*
- *Genera 10000 muestras de esta distribución y estima la varianza empírica.*
- *Compara ambos valores y analiza las diferencias.*

Se estudia la varianza de una distribución uniforme en el intervalo [0,1].

- Se calcula la varianza teórica de una distribución uniforme.
- Se generan 10000 muestras y se estima la varianza empírica.
- Se comparan ambos valores y se analiza su concordancia.

```

# Ejercicio 3: Varianza de una Distribuci n Uniforme
samples_uniform = np.random.uniform(0, 1, 10000)
estimated_var_uniform = np.var(samples_uniform)
theoretical_var_uniform = 1/12
```

Ejercicio 4 *Genera dos variables aleatorias independientes $X \sim N(5, 4)$ y $Y \sim N(7, 9)$:*

- *Calcula teóricamente la varianza de $Z = X + Y$.*
- *Simula 10000 valores de Z y estima su varianza empírica.*
- *Compara con el valor teórico.*

Se analiza la varianza de la suma de dos variables normales independientes.

- Se calcula la varianza teórica de $Z = X + Y$, donde $X \sim N(5, 4)$ y $Y \sim N(7, 9)$.
- Se simulan 10000 valores de Z y se estima su varianza empírica.
- Se comparan los valores teóricos y empíricos.

```
# Ejercicio 4: Suma de Variables y su Varianza
samples_x = np.random.normal(5, np.sqrt(4), 10000)
samples_y = np.random.normal(7, np.sqrt(9), 10000)
samples_z = samples_x + samples_y
var_x, var_y = 4, 9
var_z_theoretical, var_z_empirical = var_x + var_y, np.var(
    samples_z)
```

Este ejercicio analiza la varianza en datos de temperatura simulados.

- Se generan datos de temperatura con media 20 y varianza 5.
- Se calcula la varianza en una ventana móvil de 30 días.
- Se grafica la evolución de la varianza y se analiza su estacionariedad.

Estos ejercicios ayudan a comprender el papel de las funciones generadoras y la varianza en el análisis de distribuciones de probabilidad y datos reales.

Ejercicio 5 *Descarga datos históricos de temperatura de una ciudad y:*

- *Calcula la media y la varianza de la temperatura diaria.*
- *Representa la evolución de la varianza en una ventana móvil de 30 días.*
- *Analiza si la varianza es estacionaria o cambia con el tiempo.*

```

# Ejercicio 5: An lisis de Varianza en Datos Reales
np.random.seed(42)
temperaturas = np.random.normal(20, 5, 365)
rolling_var = [np.var(temperaturas[i:i+30]) for i in range(len(
    temperaturas)-30)]
```