

Introducción a la Probabilidad y la Estadística

Martes y Jueves Aula B17
Dra Ana Georgina Flesia

Covarianza

Definición

Sea (X, Y) un vector aleatorio con densidad conjunta f_{XY} y esperanzas finitas entonces

$$\begin{aligned} Cov(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (X - E(X))(y - E(Y))f_{X,Y}(x, y)dxdy \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Covarianza

Proposición:

Si X_1, \dots, X_n son variables aleatorias con varianzas finitas y $a_i \in \mathbb{R} \forall i = 1, \dots, n$, entonces

1.

$$E \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i E (X_i)$$

2. Si Y_1, \dots, Y_m son variables aleatorias con varianzas finitas, $b_j \in \mathbb{R}, \forall j = 1, \dots, m$ entonces:

$$\text{Cov} \left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j \right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov} (X_i, Y_j)$$

Covarianza

Proposición:

1.

$$V \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j \text{Cov}(X_i, X_j)$$

2. Si X_1, \dots, X_n son variables aleatorias independientes entonces

$$V \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 V(X_i)$$

Sumas

Proposición:

- Si X_1, \dots, X_n son variables aleatorias independientes e idénticamente distribuidas (i.i.d.) con $E(X) = \mu$ y $V(X) = \sigma^2$ entonces:

$$E(\bar{X}_n) = \mu \quad \text{y} \quad V(\bar{X}_n) = \frac{\sigma^2}{n}$$

Sumas

Definición:

Diremos que X_1, \dots, X_n es una **muestra aleatoria** si son variables aleatorias independientes e idénticamente distribuidas.

Proposición:

Si X_1, \dots, X_n es una muestra aleatoria de la distribución normal, o sea $X_i \sim N(\mu, \sigma^2)$, $\forall i = 1, \dots, n$. Entonces

$$\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2) \quad \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{1}{n}\sigma^2\right)$$

$$\sqrt{n} \left(\frac{\bar{X} - \mu}{\sigma} \right) \sim N(0, 1)$$

Sumas

Proposición:

Si X_1, \dots, X_n son v.a. independientes y con distribución normal, o sea que $X_i \sim N(\mu_i, \sigma_i^2)$, $\forall i = 1, \dots, n$. Entonces

$$\sum_{i=1}^n a_i X_i \sim N\left(E\left(\sum_{i=1}^n a_i X_i\right), V\left(\sum_{i=1}^n a_i X_i\right)\right)$$

donde

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i) = \sum_{i=1}^n a_i \mu_i$$

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i) = \sum_{i=1}^n a_i^2 \sigma_i^2$$

Sumas

Proposición:

- ▶ Si X e Y son discretas con función de masa conjunta $p(x, y)$, y sea $Z = X + Y$
- ▶ Notemos que $Z = z$ cuando $X = x$ y $Y = z - x$, por lo cual

$$p_Z(z) = \sum_{x=-\infty}^{\infty} p(x, z - x)$$

- ▶ Si X e Y son independientes, entonces $p(x, y) = p_X(x)p_Y(y)$ y

$$p_Z(z) = \sum_{x=-\infty}^{\infty} p_X(x)p_Y(z - x)$$

Esta suma se llama convolución de las funciones p_X y p_Y

Sumas

Proposición:

Sean X_1, \dots, X_r variables aleatorias independientes geométricamente distribuidas con parámetro p . Entonces $X_1 + \dots + X_r$ tiene distribución binomial negativa de parámetros r, p .

Proposición:

Sean X_1, \dots, X_r variables aleatorias independientes con distribución binomial con parámetro n_i, p . Entonces $X_1 + \dots + X_r$ tiene distribución binomial de parámetros $n_1 + \dots + n_r, p$.

Sumas

Proposición:

Sean X_1, \dots, X_r variables aleatorias independientes con distribución binomial negativa con parámetro n_i, p . Entonces $X_1 + \dots + X_r$ tiene distribución binomial negativa de parámetros $n_1 + \dots + n_r, p$.

Proposición:

Sean X_1, \dots, X_r variables aleatorias independientes con distribución Poisson con parámetro λ_i . Entonces $X_1 + \dots + X_r$ tiene distribución Poisson con parámetro $\lambda_1 + \dots + \lambda_r$.

Sumas

Ejemplo:

El número de imperfecciones que tiene una placa fotográfica sigue la distribución de Poisson de parámetro $0,1$ imperfecciones por cm^2 .

1. Si de tal placa se toma una muestra de $30\ cm^2$, ¿cuál es la probabilidad de que esa muestra contenga exactamente tres irregularidades?
2. Si ahora se toma en forma independiente 5 muestras de $30\ cm^2$, ¿cuál es la probabilidad de que exactamente dos de ellas contengan a lo sumo una irregularidad?

Sumas

1. Si Y es el número imperfecciones que tiene una placa fotográfica sigue la distribución de Poisson de parámetro 0,1 imperfecciones por cm^2 .
2. Si tenemos $30cm^2$, entonces el número de imperfecciones X en toda la placa es la suma de las imperfecciones en cada centímetro cuadrado, $X = Y_1 + \dots + Y_{30}$, una suma de variables aleatorias independientes Poisson con el mismo parámetro 0.1.
3. Por lo cual X también es Poisson, con parámetro $\lambda = 0.1 + \dots + 0.1 = 30 \times 0.1 = 3$. Esto quiere decir que podemos esperar una media de 3 imperfecciones por placa de $30 cm^2$.
4. Ahora la probabilidad de que la placa contenga **exactamente 3** irregularidades es

$$P(X = 3) = \frac{\lambda^3}{3!} e^{-\lambda} = \frac{27}{6} e^{-3} = 0.2240$$

Sumas

1. Si ahora tomamos 5 muestras independientes de 30 cm^2 cada una, la variable Z que mide el número de placas con a lo sumo una irregularidad es una variable binomial, con parámetro $n = 5$ y probabilidad $p = P(X \leq 1)$, donde X es la variable Poisson que mide las irregularidades por placa.

$$p = P(X \leq 1) = e^{-\lambda} + \frac{\lambda^2}{2!}e^{-\lambda} = e^{-3}(1 + \frac{9}{2}) = 0.2738$$

2. Por lo cual la probabilidad de que exactamente dos tengan a lo sumo una irregularidad es

$$P(Z = 2) = \binom{5}{2} 0.2738^2 (1 - 0.2738)^{5-2}$$

Sumas

Proposición:

Si X e Y son variables aleatorias continuas con densidad conjunta $f(x, y)$ y sea $Z = X + Y$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z-x)dx \quad -\infty < z < \infty$$

Si X e Y son variables independientes entonces $X + Y$ tiene densidad

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx \quad -\infty < z < \infty$$

Si X e Y son independientes y no negativas entonces $X + Y$ tiene densidad

$$f_{X+Y}(z) = \int_0^z f_X(x)f_Y(z-x)dx \quad -\infty < z < \infty$$

- ▶ Notemos que $Z \leq z$ cuando $X = x$ y $Y \leq z - x$, por lo cual

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) dy dx$$

- ▶ Haciendo el cambio de variable en la integral interior $y = v - x$ se obtiene

$$F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^z f(x, v - x) dv dx = \int_{-\infty}^z \int_{-\infty}^{\infty} f(x, v - x) dx dv$$

- ▶ Diferenciando con respecto a z queda

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z - x) dx$$

- ▶ Si X e Y son independientes, entonces $f(x, y) = f_X(x)f_Y(y)$ y

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x) dx$$

Esta integral se llama convolución de las funciones f_X y f_Y

Sumas

Ejemplo:

La vida útil de un cierto tipo de fusibles puede suponerse con distribución exponencial con $\lambda = 1/3$, donde las mediciones son en cientos de horas.

Si un sistema tiene dos fusibles con vida útil independiente ensamblados de tal forma que uno se activa cuando el otro falla, por lo cual el sistema falla cuando ambos lo hacen, encuentra la probabilidad de que el sistema dure menos de 1 (ciento de hora).

Sumas

Resolución

Sea V la vida útil del sistema, X e Y la vida útil de cada fusible, entonces $V = X + Y$, la suma de dos variables continuas no negativas independientes con distribución exponencial.

Por lo tanto

$$\begin{aligned}f_V(z) &= \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx \\&= \lambda^2 e^{-\lambda z} \int_0^z dx \\&= \lambda^2 z e^{-\lambda z} \quad 0 < z\end{aligned}$$

Sumas

Resolución

$$\begin{aligned}P(V < 1) &= \int_0^1 \lambda^2 z e^{-\lambda z} dz = uv - \int vu' \\&= \lambda^2 \left[\left[z \frac{1}{-\lambda} e^{-\lambda z} \right]_0^1 - \frac{1}{-\lambda} \int_0^1 e^{-\lambda z} dz \right] \\&= \lambda^2 \left[-\frac{1}{\lambda} e^{-\lambda} - \frac{1}{-\lambda} \left[\frac{1}{-\lambda} e^{-\lambda z} \right]_0^1 \right] \\&= \lambda^2 \left[-\frac{1}{\lambda} e^{-\lambda} - \frac{1}{\lambda^2} e^{-\lambda} + \frac{1}{\lambda^2} \right] \\&= 1 - e^{-\lambda} (\lambda + 1) \\&= 1 - \frac{4}{3} e^{\frac{1}{3}}\end{aligned}$$

Sumas

Ejemplo

Se diseña un ascensor de carga cuyo límite es 1000 kg. El peso de cada caja sigue una distribución normal con un peso medio de 32 kg y un desvío estándar de 10 kg.

1. ¿Cuál es la probabilidad de que un grupo de 30 cajas exceda el límite de carga?
2. Se toma una muestra aleatoria simple de 5 cajas ¿cuál es la probabilidad de que el mínimo de la muestra sea superior a 30kg e inferior a 33kg?

Resolución

1. Para que el grupo exceda el límite de carga la suma de los pesos de las cajas debe exceder 1000kg.
2. Como el peso de cada caja es una variable normal independiente con los mismos parámetros μ y σ^2 , la variable $X = \sum_{i=1}^{30} X_i$ también es normal con parámetros $\mu_X = 30 \times \mu = 960$ y $\sigma_X^2 = 30 \times \sigma^2 = 3000$.
3. Entonces estandarizando calculamos la probabilidad

$$\begin{aligned}P(X > 1000) &= P((X - \mu_X)/\sigma_X > (1000 - \mu_X)/\sigma_X) \\&= P(Z > (1000 - 960)/54.7723) \\&= P(Z > 0.7303) \\&= 0.2326\end{aligned}$$

Resolución

- Si las variables X_1, X_2, X_3, X_4, X_5 representan el peso de las cinco cajas elegidas al azar, y $K = \min(X_1, X_2, X_3, X_4, X_5)$ con distribución

$$F_K(k) = P(K \leq k) = 1 - P(K \geq k) = 1 - P(X_1 \geq k)^5 = 1 - [1 - F(k)]^5$$

- Entonces

$$\begin{aligned}P(30 < K < 33) &= F_K(33) - F_K(30) \\&= 1 - [1 - F(33)]^5 - (1 - [1 - F(30)]^5) \\&= [1 - \Phi((30 - \mu)/\sigma)]^5 - [1 - \Phi((33 - \mu)/\sigma)]^5 \\&= (1 - \Phi(-0.2))^5 - (1 - \Phi(0.1))^5 \\&= (1 - 0.4920)^5 - (1 - 0.5040)^5 \\&= 0.0038\end{aligned}$$