

Momentos, Variância, Covariância e Funções Características

Cristian Coletti

Primeiro Quadrimestre 2026

Desigualdade de Chebyshev

Aplicada à v.a. $X - E(X)$ com $\alpha = 2$, a desigualdade de Markov (2.28) torna-se:

Desigualdade de Chebyshev: Para todo $t > 0$

$$P(|X - E(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2}. \quad (2.29)$$

Variância da Soma de Duas V.A.'s

$$\begin{aligned}\text{Var}(X + Y) &= E[(X + Y - E(X) - E(Y))^2] \\ &= \text{Var}(X) + \text{Var}(Y) + 2E[(X - E(X))(Y - E(Y))].\end{aligned}$$

Se definirmos

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y),$$

então

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

Covariância e Independência

$\text{Cov}(X, Y)$ é a **covariância** de X e Y . Note que $\text{Cov}(X, Y)$ nada mais é que o produto escalar em L^2 de $X - E(X)$ e $Y - E(Y)$.

Se X e Y são independentes, pelo Corolário 2.10,

$$\text{Cov}(X, Y) = 0,$$

logo

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Não Correlação vs. Independência

A recíproca não é verdadeira: existem exemplos de v.a.'s com covariância nula que não são independentes. Isso não é surpreendente: $\text{Cov}(X, Y) = 0$ significa $E(XY) = E(X)E(Y)$, enquanto independência requer $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ para todo par de funções de Borel limitadas g, h .

Matriz de Covariância

Se $X = (X_1, \dots, X_m)$ é uma v.a. m -dimensional, sua **matriz de covariância** é a matriz $m \times m$ C com elementos

$$c_{ij} = \text{Cov}(X_i, X_j) = E[(X_i - E(X_i))(X_j - E(X_j))].$$

C é simétrica, com as variâncias na diagonal e as covariâncias fora dela.

Formulação Matricial

Elegantemente,

$$C = E[(X - E(X))(X - E(X))^*] \quad (2.31)$$

onde $X - E(X)$ é um vetor coluna. Sob uma transformação linear $Y = AX$, a matriz de covariância de Y é

$$C_{AX} = ACA^*. \quad (2.32)$$

Definição Positiva da Matriz de Covariância

A matriz de covariância é sempre positiva definida: para todo $\xi \in \mathbb{R}^m$,

$$\langle C\xi, \xi \rangle = \sum_{i,j=1}^m c_{ij} \xi_i \xi_j \geq 0.$$

De fato,

$$\langle C\xi, \xi \rangle = E [\langle \xi, X - E(X) \rangle^2] \geq 0. \quad (2.33)$$

Exemplo 2.24: A "Reta" de Regressão

Considere uma v.a. real Y e uma v.a. m -dimensional X , ambas de quadrado integrável. Qual função afim-linear de X melhor aproxima Y ? Queremos minimizar

$$E[(\langle a, X \rangle + b - Y)^2].$$

Suponha que a matriz de covariância C de X é estritamente positiva definida.

Após manipulações, queremos minimizar

$$S(a) = \langle Ca, a \rangle + \tilde{b}^2 + \text{Var}(Y) - 2\langle a, R \rangle,$$

onde $R_i = \text{Cov}(X_i, Y)$. O diferencial é $DS(a) = 2Ca - 2R$. Para $DS(a) = 0$,

$$a = C^{-1}R.$$

Escolha b tal que $\tilde{b} = 0$:

$$b = E(Y) - \langle a, E(X) \rangle.$$

Caso Unidimensional ($m = 1$)

Se $m = 1$, $C = \text{Var}(X)$ e $R = \text{Cov}(X, Y)$. Então

$$a = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}, \quad b = E(Y) - aE(X). \quad (2.35)$$

A função $x \mapsto ax + b$ é a **reta de regressão** de Y sobre X .

Interpretação da Covariância

O coeficiente angular a tem o mesmo sinal da covariância. Portanto:

- Covariância positiva: reta crescente.
- Covariância negativa: reta decrescente.

V.A.'s independentes (não correlacionadas) têm reta de regressão horizontal.

Melhor Aproximação Constante

Seja Y uma v.a. m -dimensional de quadrado integrável. Qual vetor $b \in \mathbb{R}^m$ minimiza $E(|Y - b|^2)$?

Considere

$$b \mapsto E(|Y - b|^2) = E(|Y|^2) - 2\langle b, E(Y) \rangle + |b|^2.$$

O diferencial é $b \mapsto -2E(Y) + 2b$ e se anula em $b = E(Y)$. Logo, a esperança é a constante que melhor aproxima Y em L^2 .

Coeficiente de Correlação

A covariância é sensível a mudanças de escala. Mais útil é o **coeficiente de correlação**:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}.$$

É invariante sob mudanças de escala. Por (2.30), $-1 \leq \rho_{X,Y} \leq 1$.

Interpretação do Coeficiente de Correlação

- Valores próximos de 0 indicam "quase independência".
- Valores próximos de 1 indicam forte dependência "no mesmo sentido".
- Valores próximos de -1 indicam forte dependência "em sentido oposto".

Definição

Seja X uma v.a. m -dimensional. Sua **função característica** é $\phi : \mathbb{R}^m \rightarrow \mathbb{C}$ definida por

$$\phi(\theta) = E(e^{i(\theta, X)}) = E(\cos(\theta, X)) + iE(\sin(\theta, X)). \quad (2.36)$$

Sempre definida, pois $|e^{i(\theta, X)}| = 1$.

Propriedades Básicas

$$|\phi(\theta)| \leq 1 \quad \text{para todo } \theta \in \mathbb{R}^m, \quad \phi(0) = 1.$$

Se μ é a lei de X ,

$$\phi(\theta) = \int_{\mathbb{R}^m} e^{i(\theta, x)} d\mu(x). \quad (2.37)$$

A função característica depende apenas da lei de X .

Se μ e ν são probabilidades em $\mathbb{R}^m$,

$$\phi_{\mu*\nu}(\theta) = \phi_{\mu}(\theta)\phi_{\nu}(\theta). \quad (2.38)$$

Pois se X e Y são independentes com leis μ e ν ,

$$\phi_{X+Y}(\theta) = E(e^{i(\theta,X)})E(e^{i(\theta,Y)}) = \phi_{\mu}(\theta)\phi_{\nu}(\theta).$$

$$\phi_{-X}(\theta) = \overline{\phi_X(\theta)}. \quad (2.39)$$

Logo, se X é simétrica ($X \sim -X$), então ϕ_X é real. A recíproca? Veremos.

Transformação Afim

Se $Y = AX + b$, com $A \in \mathbb{R}^{d \times m}$ e $b \in \mathbb{R}^d$, então para $\theta \in \mathbb{R}^d$

$$\phi_Y(\theta) = e^{i(\theta, b)} \phi_X(A^* \theta). \quad (2.40)$$

Exemplo 2.25 (a): Binomial $B(n, p)$

Para $\theta \in \mathbb{R}$,

$$\phi(\theta) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} e^{i\theta k} = (1 - p + pe^{i\theta})^n.$$

Exemplo 2.25 (b): Geometria

$$\phi(\theta) = \sum_{k=0}^{\infty} p(1-p)^k e^{i\theta k} = p \sum_{k=0}^{\infty} ((1-p)e^{i\theta})^k = \frac{p}{1 - (1-p)e^{i\theta}}.$$