Esperanças e covariâncias

Tiago M. Magalhães

Departamento de Estatística - ICE-UFJF

Juiz de Fora, 19 de novembro de 2024

Roteiro

Roteiro

Introdução

Seja $(X, Y)^\top$ uma variável aleatória (VA) bidimensional e g é uma função destas duas variáveis.

Esperança

Se $(X, Y)^\top$ é uma VA discreta bidimensional, com função de probabilidade conjunta $p(x_i, y_j)$, então

$$\mathbb{E}\left[g(X, Y)\right] = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i, y_j)p(x_i, y_j).$$

Esperança

Se $(X, Y)^\top$ é uma VA discreta bidimensional, com função de probabilidade conjunta $p(x_i, y_j)$, então

$$\mathbb{E}\left[g(X, Y)\right] = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} g(x_i, y_j) p(x_i, y_j).$$

Esperança

Se $(X, Y)^\top$ é uma VA contínua bidimensional, com função densidade de probabilidade (FDP) conjunta $f(x, y)$, então

$$\mathbb{E}\left[g(X, Y)\right] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy.$$

Esperança

Se $(X, Y)^\top$ é uma VA contínua bidimensional, com função densidade de probabilidade (FDP) conjunta $f(x, y)$, então

$$\mathbb{E}\left[g(X, Y)\right] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy.$$

Covariância

A covariância entre X e Y, denotada por $\text{Cov}(X, Y)$,

Covariância

A covariância entre X e Y, denotada por $\text{Cov}(X, Y)$, é definida por

$$\text{Cov}(X, Y) = \mathbb{E}\left\{[X - \mathbb{E}(X)]\,[Y - \mathbb{E}(Y)]\right\}.$$

Covariância

A covariância entre X e Y, denotada por $\text{Cov}(X, Y)$, é definida por

$$\text{Cov}(X, Y) = \mathbb{E}\left\{[X - \mathbb{E}(X)]\,[Y - \mathbb{E}(Y)]\right\}.$$

Alternativamente,

Covariância

A covariância entre X e Y, denotada por $\text{Cov}(X, Y)$, é definida por

$$\text{Cov}(X, Y) = \mathbb{E}\left\{[X - \mathbb{E}(X)]\,[Y - \mathbb{E}(Y)]\right\}.$$

Alternativamente, a covariância entre X e Y pode ser escrita como

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Covariância

A covariância entre X e Y, denotada por $\text{Cov}(X, Y)$, é definida por

$$\text{Cov}(X, Y) = \mathbb{E}\left\{[X - \mathbb{E}(X)]\,[Y - \mathbb{E}(Y)]\right\}.$$

Alternativamente, a covariância entre X e Y pode ser escrita como

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Covariância

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

Covariância

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

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

Covariância

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
- $\text{Cov}(X, X) = \text{Var}(X)$;

Covariância

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
- $\text{Cov}(X, X) = \text{Var}(X)$;
- $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$;

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
- $\text{Cov}(X, X) = \text{Var}(X)$;
- $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$;
- $\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \text{Cov}(X_i, Y_j)$.

Covariância

Sejam a, b, c, d constantes, da definição de covariância, nós temos as seguintes propriedades:

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
- $\text{Cov}(X, X) = \text{Var}(X)$;
- $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$;
- $\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \text{Cov}(X_i, Y_j)$.

Covariância

Da definição e das propriedades da covariância, nós temos alguns resultados.

O primeiro é que, se X e Y são independentes,

Covariância

Da definição e das propriedades da covariância, nós temos alguns resultados.

O primeiro é que, se X e Y são independentes, então

$$\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y),$$

Covariância

Da definição e das propriedades da covariância, nós temos alguns resultados.

O primeiro é que, se X e Y são independentes, então

$$\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y),$$

isto é, se X e Y são independentes, $\text{Cov}(X, Y) = 0$.

Covariância

Da definição e das propriedades da covariância, nós temos alguns resultados.

O primeiro é que, se X e Y são independentes, então

$$\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y),$$

isto é, se X e Y são independentes, $\text{Cov}(X, Y) = 0$.

Observação: $\text{Cov}(X, Y) = 0$ não implica que X e Y são independentes.

Covariância

Da definição e das propriedades da covariância, nós temos alguns resultados.

O primeiro é que, se X e Y são independentes, então

$$\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y),$$

isto é, se X e Y são independentes, $\text{Cov}(X, Y) = 0$.

Observação: $\text{Cov}(X, Y) = 0$ não implica que X e Y são independentes.

Covariância

No segundo, sejam $X_1, X_2, \ldots, X_n$ uma sequência de VAs, então

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}\left(X_i\right) + 2\sum_{i \neq j}\sum \text{Cov}(X_i, X_j). \qquad (1)$$

Covariância

No segundo, sejam $X_1, X_2, \dots, X_n$ uma sequência de VAs, então

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\,(X_i) + 2\sum_{i \neq j}\sum \mathrm{Cov}(X_i, X_j). \tag{1}$$

Para o caso bidimensional $(X, Y)^\top$,

Covariância

No segundo, sejam $X_1, X_2, \ldots, X_n$ uma sequência de VAs, então

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}\left(X_i\right) + 2\sum_{i\neq j}\sum \text{Cov}(X_i, X_j). \tag{1}$$

Para o caso bidimensional $(X, Y)^\top$, nós temos que (1) resulta em

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

Covariância

No segundo, sejam $X_1, X_2, \ldots, X_n$ uma sequência de VAs, então

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\,(X_i) + 2\sum_{i \neq j}\sum \mathrm{Cov}(X_i, X_j). \tag{1}$$

Para o caso bidimensional $(X, Y)^\top$, nós temos que (1) resulta em

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

Covariância

Agora, se $X_1, X_2, \ldots, X_n$ é uma sequência de VAs independentes duas a duas, (1) resulta em

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right).$$

Covariância

Agora, se $X_1, X_2, \ldots, X_n$ é uma sequência de VAs independentes duas a duas, (1) resulta em

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}\left(X_i\right).$$

Correlação

Seja $(X, Y)^\top$ uma VA bidimensional. Nós definiremos ρ_{xy}, o coeficiente de correlação, entre X e Y, da seguinte forma:

Correlação

Seja $(X, Y)^\top$ uma VA bidimensional. Nós definiremos ρ_{xy}, o coeficiente de correlação, entre X e Y, da seguinte forma:

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

Correlação

Seja $(X, Y)^\top$ uma VA bidimensional. Nós definiremos ρ_{xy}, o coeficiente de correlação, entre X e Y, da seguinte forma:

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

Correlação

- ρ_{xy} é uma quantidade adimensional;

Correlação

- ρ_{xy} é uma quantidade adimensional;
- $-1 \leq \rho_{xy} \leq 1$;

Correlação

- ρ_{xy} é uma quantidade adimensional;
- $-1 \leq \rho_{xy} \leq 1$;
- se X e Y são independentes, então $\rho_{xy} = 0$;

Correlação

- ρ_{xy} é uma quantidade adimensional;
- $-1 \leq \rho_{xy} \leq 1$;
- se X e Y são independentes, então $\rho_{xy} = 0$;
- Se $\rho_{xy}^2 = 1$, então $Y = aX + b$, em que a e b são constantes;

Correlação

- ρ_{xy} é uma quantidade adimensional;
- $-1 \leq \rho_{xy} \leq 1$;
- se X e Y são independentes, então $\rho_{xy} = 0$;
- Se $\rho_{xy}^2 = 1$, então $Y = aX + b$, em que a e b são constantes;
- Se $Y = aX + b$ e $\rho_{xy}^2 = 1$. Se $a > 0$, $\rho_{xy} = 1$, $a < 0$, $\rho_{xy} = -1$.

Correlação

- ρ_{xy} é uma quantidade adimensional;
- $-1 \leq \rho_{xy} \leq 1$;
- se X e Y são independentes, então $\rho_{xy} = 0$;
- Se $\rho_{xy}^2 = 1$, então $Y = aX + b$, em que a e b são constantes;
- Se $Y = aX + b$ e $\rho_{xy}^2 = 1$. Se $a > 0$, $\rho_{xy} = 1$, $a < 0$, $\rho_{xy} = -1$.

Correlação

Resultado

Se ρ_{xy} for o coeficiente de correlação entre X e Y,

Correlação

Resultado

Se ρ_{xy} for o coeficiente de correlação entre X e Y, e se $V = aX + b$ e $W = cY + d$,

Correlação

Resultado

Se ρ_{xy} for o coeficiente de correlação entre X e Y, e se $V = aX + b$ e $W = cY + d$, nas quais a, b, c, d, $a \neq 0$ e $c \neq 0$ são constantes,

Correlação

Resultado

Se ρ_{xy} for o coeficiente de correlação entre X e Y, e se $V = aX + b$ e $W = cY + d$, nas quais a, b, c, d, $a \neq 0$ e $c \neq 0$ são constantes, então

$$\rho_{vw} = \frac{ac}{|ac|}\rho_{xy}.$$

Correlação

Resultado

Se ρ_{xy} for o coeficiente de correlação entre X e Y, e se $V = aX + b$ e $W = cY + d$, nas quais a, b, c, d, $a \neq 0$ e $c \neq 0$ são constantes, então

$$\rho_{vw} = \frac{ac}{|ac|}\rho_{xy}.$$

Exemplo

Seja $(X, Y)^\top$ uma VA bidimensional, com FDP conjunta dada por:

$$
\begin{aligned}
f(x, y) &= 2\,\mathbb{I}_{(0,1)}(x)\,\mathbb{I}_{(x,1)}(y) \\
&= 2\,\mathbb{I}_{(0,1)}(y)\,\mathbb{I}_{(0,y)}(x).
\end{aligned}
$$

Encontre ρ_{xy}.

Exemplo

Seja $(X, Y)^\top$ uma VA bidimensional, com FDP conjunta dada por:

$$
\begin{aligned}
f(x, y) &= 2\, \mathbb{I}_{(0,1)}(x)\, \mathbb{I}_{(x,1)}(y) \\
&= 2\, \mathbb{I}_{(0,1)}(y)\, \mathbb{I}_{(0,y)}(x).
\end{aligned}
$$

Encontre ρ_{xy}.

Roteiro

1. Esperanças e covariâncias
2. Valor esperado condicionado
3. Regressão da média
4. Bibliografia

Valor esperado condicionado

Tal como nós definimos o valor esperado de uma variável aleatória X, nós também podemos definir o **valor esperado condicionado** de uma VA.

Valor esperado condicionado

Tal como nós definimos o valor esperado de uma variável aleatória X, nós também podemos definir o **valor esperado condicionado** de uma VA.

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA discreta bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y_j$,

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA discreta bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y_j$, como

$$\mathbb{E}(X|Y = y_j) = \sum_{i=1}^{+\infty} x_i \, p_{X|Y}(x_i|y_j).$$

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA discreta bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y_j$, como

$$\mathbb{E}(X|Y = y_j) = \sum_{i=1}^{+\infty} x_i \, p_{X|Y}(x_i|y_j).$$

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA contínua bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y$,

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA contínua bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y$, como

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{+\infty} x\, f_{X|Y}(x|y)\, dx.$$

Valor esperado condicionado

Se $(X, Y)^\top$ for uma VA contínua bidimensional, nós definiremos o valor esperado condicionado de X, para um dado $Y = y$, como

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{+\infty} x\, f_{X|Y}(x|y)\, dx.$$

Valor esperado condicionado

De maneira geral, se $(X, Y)^\top$ for uma VA bidimensional e g é uma função de X, nós temos que

$$\begin{cases} \mathbb{E}\left[g(X)|Y = y_j\right] = \sum_{i=1}^{+\infty} g(x_i)\, p_{X|Y}(x_i|y_j), & \text{caso discreto,} \\ \mathbb{E}\left[g(X)|Y = y\right] = \int_{-\infty}^{+\infty} g(x)\, f_{X|Y}(x|y) dx, & \text{caso contínuo.} \end{cases}$$

Valor esperado condicionado

De maneira geral, se $(X, Y)^\top$ for uma VA bidimensional e g é uma função de X, nós temos que

$$\begin{cases} \mathbb{E}\left[g(X)|Y = y_j\right] = \sum_{i=1}^{+\infty} g(x_i)\, p_{X|Y}(x_i|y_j), & \text{caso discreto,} \\ \mathbb{E}\left[g(X)|Y = y\right] = \int_{-\infty}^{+\infty} g(x)\, f_{X|Y}(x|y)dx, & \text{caso contínuo.} \end{cases}$$

Observação

Em geral, $\mathbb{E}(X|Y = y)$ é uma função de y e, por isso, é uma VA.

Observação

Em geral, $\mathbb{E}(X|Y = y)$ é uma função de y e, por isso, é uma VA. Estritamente falando, $\mathbb{E}(X|Y = y)$ é o valor da VA $\mathbb{E}(X|Y)$.

Observação

Em geral, $\mathbb{E}(X|Y = y)$ é uma função de y e, por isso, é uma VA. Estritamente falando, $\mathbb{E}(X|Y = y)$ é o valor da VA $\mathbb{E}(X|Y)$.

Valor esperado condicionado

Um propriedade extremamente importante do valor esperado condicional é dada por

$$\mathbb{E}(X) = \mathbb{E}\left[\mathbb{E}(X|Y)\right]. \tag{2}$$

Semelhantemente, $\mathbb{E}(Y) = \mathbb{E}\left[\mathbb{E}(Y|X)\right]$.

Valor esperado condicionado

Um propriedade extremamente importante do valor esperado condicional é dada por

$$\mathbb{E}(X) = \mathbb{E}\left[\mathbb{E}(X|Y)\right]. \tag{2}$$

Semelhantemente, $\mathbb{E}(Y) = \mathbb{E}\left[\mathbb{E}(Y|X)\right]$.

Valor esperado condicionado

O resultado em (2) também continua válido para situações como

- $\mathbb{E}[g(X)] = \mathbb{E}\left\{\mathbb{E}[g(X)|Y]\right\}$;

Valor esperado condicionado

O resultado em (2) também continua válido para situações como

- $\mathbb{E}[g(X)] = \mathbb{E}\left\{\mathbb{E}[g(X)|Y]\right\}$;
- $\mathbb{E}(XY) = \mathbb{E}\left[\mathbb{E}(XY|Y)\right]$.

Valor esperado condicionado

O resultado em (2) também continua válido para situações como

- $\mathbb{E}[g(X)] = \mathbb{E}\left\{\mathbb{E}[g(X)|Y]\right\}$;
- $\mathbb{E}(XY) = \mathbb{E}\left[\mathbb{E}(XY|Y)\right]$.

Valor esperado condicionado

Um outra propriedade extremamente importante do valor esperado condicional é dada por

$$\text{Var}(X) = \mathbb{E}\left[\text{Var}(X|Y)\right] + \text{Var}\left[\mathbb{E}(X|Y)\right].$$

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia.

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia. Se N for o número de peças da remessa,

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia. Se N for o número de peças da remessa, a distribuição de probabilidade da variável aleatória N,

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia. Se N for o número de peças da remessa, a distribuição de probabilidade da variável aleatória N, será dada assim:

N	10	11	12	13	14	15
$\mathbb{P}(N)$	0,05	0,10	0,10	0,20	0,35	0,20

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia. Se N for o número de peças da remessa, a distribuição de probabilidade da variável aleatória N, será dada assim:

N	10	11	12	13	14	15
$\mathbb{P}(N)$	0,05	0,10	0,10	0,20	0,35	0,20

Com $\mathbb{E}(N) = 13{,}3$.

Exemplo

Suponham que remessas, contendo um número variável de peças, chegue cada dia. Se N for o número de peças da remessa, a distribuição de probabilidade da variável aleatória N, será dada assim:

N	10	11	12	13	14	15
$\mathbb{P}(N)$	0,05	0,10	0,10	0,20	0,35	0,20

Com $\mathbb{E}(N) = 13{,}3$.

Exemplo

A probabilidade de que qualquer peça em particular seja defeituosa é a mesma para todas as peças e igual a 0,10.

Exemplo

A probabilidade de que qualquer peça em particular seja defeituosa é a mesma para todas as peças e igual a 0,10. Se X for o número de peças defeituosas que chegue cada dia, qual será o valor esperado de X?

Exemplo

A probabilidade de que qualquer peça em particular seja defeituosa é a mesma para todas as peças e igual a 0,10. Se X for o número de peças defeituosas que chegue cada dia, qual será o valor esperado de X?

Roteiro

Regressão da média

Como nós já salientamos, $\mathbb{E}(X|Y = y)$ é o valor da VA $\mathbb{E}(X|Y)$ e é uma função de y. O gráfico dessa função de y é conhecido como a **curva de regressão** (da média) de X em Y.

Regressão da média

Como nós já salientamos, $\mathbb{E}(X|Y = y)$ é o valor da VA $\mathbb{E}(X|Y)$ e é uma função de y. O gráfico dessa função de y é conhecido como a **curva de regressão** (da média) de X em Y. Para cada valor fixado y, $\mathbb{E}(X|Y = y)$ será o valor esperado da VA (unidimensional) $X|Y = y$.

Regressão da média

Como nós já salientamos, $\mathbb{E}(X|Y = y)$ é o valor da VA $\mathbb{E}(X|Y)$ e é uma função de y. O gráfico dessa função de y é conhecido como a **curva de regressão** (da média) de X em Y. Para cada valor fixado y, $\mathbb{E}(X|Y = y)$ será o valor esperado da VA (unidimensional) $X|Y = y$.

Teorema

Seja $(X, Y)^\top$ uma VA bidimensional e suponham que

$$\mathbb{E}(X) = \mu_X,\ \mathbb{E}(Y) = \mu_Y,\ \text{Var}(X) = \sigma_X^2 \text{ e } \text{Var}(Y) = \sigma_Y^2.$$

Teorema

Seja $(X, Y)^\top$ uma VA bidimensional e suponham que

$$\mathbb{E}(X) = \mu_X,\ \mathbb{E}(Y) = \mu_Y,\ \text{Var}(X) = \sigma_X^2 \text{ e } \text{Var}(Y) = \sigma_Y^2.$$

Teorema

Seja ρ o coeficiente de correlação entre X e Y. Se a regressão de Y em X for linear,

Teorema

Seja ρ o coeficiente de correlação entre X e Y. Se a regressão de Y em X for linear, nós teremos

$$\mathbb{E}(Y|X=x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X).$$

Teorema

Seja ρ o coeficiente de correlação entre X e Y. Se a regressão de Y em X for linear, nós teremos

$$\mathbb{E}(Y|X = x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Se a regressão de X em Y for linear,

Teorema

Seja ρ o coeficiente de correlação entre X e Y. Se a regressão de Y em X for linear, nós teremos

$$\mathbb{E}(Y|X=x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X).$$

Se a regressão de X em Y for linear, nós teremos

$$\mathbb{E}(X|Y=y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y).$$

Teorema

Seja ρ o coeficiente de correlação entre X e Y. Se a regressão de Y em X for linear, nós teremos

$$\mathbb{E}(Y|X=x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X).$$

Se a regressão de X em Y for linear, nós teremos

$$\mathbb{E}(X|Y=y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y).$$

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$,

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$, então nós verificamos que $\mathbb{E}(X|Y = y)$ não depende de y.

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$, então nós verificamos que $\mathbb{E}(X|Y = y)$ não depende de y. Se $\rho \neq 0$, o sinal de ρ determina o sinal da declividade da reta de regressão;

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$, então nós verificamos que $\mathbb{E}(X|Y = y)$ não depende de y. Se $\rho \neq 0$, o sinal de ρ determina o sinal da declividade da reta de regressão;
- Se ambas as funções de regressão forem lineares,

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$, então nós verificamos que $\mathbb{E}(X|Y = y)$ não depende de y. Se $\rho \neq 0$, o sinal de ρ determina o sinal da declividade da reta de regressão;
- Se ambas as funções de regressão forem lineares, nós verificamos que as retas de regressão se interceptam no "centro" da distribuição, (μ_X, μ_Y).

Observações

- É possível que uma das regressões da média seja linear, enquanto a outra não o seja;
- Se a regressão de X em Y, por exemplo, for linear e, se $\rho = 0$, então nós verificamos que $\mathbb{E}(X|Y = y)$ não depende de y. Se $\rho \neq 0$, o sinal de ρ determina o sinal da declividade da reta de regressão;
- Se ambas as funções de regressão forem lineares, nós verificamos que as retas de regressão se interceptam no "centro" da distribuição, (μ_X, μ_Y).

Princípio dos mínimos quadrados

Embora as funções de regressão não precisem ser lineares, nós poderemos estar interessados em aproximar a curva de regressão com uma função linear.

Isto é usualmente feito recorrendo-se ao **princípio dos mínimos quadrados**.

Princípio dos mínimos quadrados

Embora as funções de regressão não precisem ser lineares, nós poderemos estar interessados em aproximar a curva de regressão com uma função linear. Isto é usualmente feito recorrendo-se ao **princípio dos mínimos quadrados**.

Princípio dos mínimos quadrados

No princípio dos mínimos quadrados, duas constantes a e b são escolhidas de modo que

$$\mathbb{E}\left\{[\mathbb{E}(Y|X) - (aX + b)]^2\right\}$$

se torne mínima.

Princípio dos mínimos quadrados

No princípio dos mínimos quadrados, duas constantes a e b são escolhidas de modo que

$$\mathbb{E}\left\{[\mathbb{E}(Y|X) - (aX+b)]^2\right\}$$

se torne mínima. A reta $y = ax + b$ é denominada a aproximação de mínimos quadrados à curva de regressão $\mathbb{E}(Y|X = x)$.

Princípio dos mínimos quadrados

No princípio dos mínimos quadrados, duas constantes a e b são escolhidas de modo que

$$\mathbb{E}\left\{[\mathbb{E}(Y|X) - (aX+b)]^2\right\}$$

se torne mínima. A reta $y = ax + b$ é denominada a aproximação de mínimos quadrados à curva de regressão $\mathbb{E}(Y|X = x)$.

Teorema

Se $y = ax + b$ for a aproximação de mínimos quadrados a $\mathbb{E}(Y|X = x)$ e

se $\mathbb{E}(Y|X = x)$ for de fato uma função linear de x,

Teorema

Se $y = ax + b$ for a aproximação de mínimos quadrados a $\mathbb{E}(Y|X = x)$ e se $\mathbb{E}(Y|X = x)$ for de fato uma função linear de x, isto é,

$$\mathbb{E}(Y|X = x) = a'x + b',$$

então $a = a'$ e $b = b'$.

Teorema

Se $y = ax + b$ for a aproximação de mínimos quadrados a $\mathbb{E}(Y|X = x)$ e se $\mathbb{E}(Y|X = x)$ for de fato uma função linear de x, isto é,

$$\mathbb{E}(Y|X = x) = a'x + b',$$

então $a = a'$ e $b = b'$.

Observação: de forma análoga, os resultados valem para $\mathbb{E}(X|Y = y)$.

Teorema

Se $y = ax + b$ for a aproximação de mínimos quadrados a $\mathbb{E}(Y|X = x)$ e se $\mathbb{E}(Y|X = x)$ for de fato uma função linear de x, isto é,

$$\mathbb{E}(Y|X = x) = a'x + b',$$

então $a = a'$ e $b = b'$.

Observação: de forma análoga, os resultados valem para $\mathbb{E}(X|Y = y)$.

Roteiro

Bibliografia

Meyer, P. L. (2017), *Probabilidade: aplicações à Estatística*, 2 edn, Livros Técnicos e Científicos Editora Ltda, Rio de Janeiro.

Ross, S. (2010), *Probabilidade: um curso moderno com aplicações*, 8 edn, Bookman, Porto Alegre.

Ross, S. (2019), *A first course in probability*, 10 edn. global edn, Pearson Education, Malaysia.

Obrigado!

✉ tiago.magalhaes@ufjf.br

🖱 ufjf.br/tiago_magalhaes

🌍 Departamento de Estatística, Sala 319