Teoria do Risco
Aula2

Danilo Machado Pires
danilo.pires@unifal-mg.edu.br

https://atuaria.github.io/portalhalley/index.html

ESPERANÇA DE VARIÁVEIS ALEATÓRIAS

- ...Uma forma de avaliar ganhos em jogos com apostas a dinheiro
- Representa o ponto de equilíbrio da distribuição de seus valores.
- ... Serve como parâmetro para vários modelos probabilísticos.

ESPERANÇA DE VARIÁVEIS ALEATÓRIAS

A esperança de uma variável aleatória X é dada por:

- Variáveis aleatórias discretas:

$$E(X) = \sum_{i=1}^{\infty} x_i P(X = x_i)$$

- Variáveis aleatórias Continuas:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

ESPERANÇA DE VARIÁVEIS ALEATÓRIAS

- É a mais popular medida de posição...

- As medidas de posições mais importantes são média aritmética, mediana e moda.

	Conjunto de Dados	Variável aleatória	
		Discreta	Contínua
Média	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$E(X) = \sum_{i=1}^{\infty} x_i P(X = x_i)$	$E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{dx}$
Mediana	$Md = valor\ central$	$Md: P(X \geq Md) \geq 0.5\ \ e\ \ P(X \leq Md) \geq 0.5$	
Moda	Mo= valor c/ maior freque.	Mo= valor c/ maior probabilidade	

EXEMPLO 1

Em uma cela, há uma passagem secreta que conduz a um porão de onde partem três túneis. O primeiro túnel dá acesso à liberdade em 1 hora; o segundo em 3 horas; o terceiro leva ao ponto de partida em 6 horas. Em média, os prisioneiros que descobrem os túneis conseguem escapar da prisão em quanto tempo?

Túnel		Tempo(horas)	Probabilidade
1		1	$\frac{1}{3}$
2		3	$\frac{1}{3}$
3	1	$6+1$	$\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$
	2	$6+3$	$\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$

$$E(H) = \frac{1 \times 1}{3} + \frac{3 \times 1}{3} + \frac{7 \times 1}{6} + \frac{9 \times 1}{6} = \frac{2+6+7+9}{6} = \frac{24}{6} = 4$$

Esperança de uma função de variáveis aleatórias (V.a. composta)

- Seja X uma variável aleatória e $g(\cdot)$ uma função, ambos com domínio e contradomínio real. O valor esperado do valor da função $g(X)$. Denotado por $E[g(X)]$ é definido por:

$$E[g(X)] = \begin{cases} \sum_j g(x_j)P(X = x_j), & X \; discreto \\ \int_{-\infty}^{\infty} g(x) f_X(x) dx, & X \; contínuo \end{cases}$$

EXEMPLO 2

Segundo determinada tábua de vida o tempo de vida adicional de uma pessoa de 106 é modelado da seguinte forma :

T	0	1	2	3
$P(T)$	0,6751	0,19518	0,121995	0,007681

A expectativa de vida para uma pessoa dessa idade é?

EXEMPLO 2

Segundo determinada tábua de vida o tempo de vida adicional de uma pessoa de 106 é modelado da seguinte forma :

T	0	1	2	3
$P(T)$	0,67514	0,195183	0,1219955	0,0076815

A expectativa de vida para uma pessoa dessa idade é?

$$E(T) = \sum t \times P(t) = 0{,}4622$$

Essa mesma pessoa pretende fazer um seguro de vida que deixa como benefício 1 u.m. Considere a taxa de juros de 3%, calcule o prêmio puro único.

EXEMPLO 3

T	0	1	2	3
$P(T)$	0,67514	0,195183	0,1219955	0,0076815

A expectativa de vida para uma pessoa dessa idade é?

$$E(T) = \sum t \times P(t) = 0{,}4622$$

Essa mesma pessoa pretende fazer um seguro de vida que deixa como benefício 1 u.m. Considere a taxa de juros de 3%, calcule o prêmio puro único.

$$g(T) = v^{T+1}$$

$$\mathrm{A}_{106} = E(v^{T+1}) = \sum v^{t+1} \times P(t) = 0{,}9866602$$

EXEMPLO 4

Um apólice de seguro cobre uma perda aleatória X, com um valor de franquia $d = 0{,}3$. A perda é modelada como um variável aleatória contínua com densidade $f_X(x) = 2x$ para $0 < x < 1$. Calcule o que se pede.

a) Calcule o valor esperado dos excessos de danos em relação a d. Considere que a seguradora é informada de todos os sinistros ocorridos, mesmo aqueles em que o valor ficou abaixo de d.

b) Calcule o valor esperado dos excessos de danos, dado que houve excesso em relação a d. Considere que a seguradora é informada somente dos sinistros superiores a d.

Solução

EXEMPLO 4-Solução

No item $a)$ temos que $Y = g(X) = Max(0; X - d) = (X - d)$, então

$$E(Y) = E[g(X)] = \int_0^d 0\, f_X(x)dx + \int_d^1 (x - d)\, f_X(x)dx$$

$$E(Y) = \int_{0,3}^1 (x - 0{,}3)\, 2xdx \approx 0{,}375667$$

Na prática a seguradora considera que os valores de sinistros menores que a franquia, tem severidade (impacto econômico) 0.

EXEMPLO 4-Solução

No item $b)$ temos que $Y = g(X) = (X - d)|X > d$, então

$$E[g(X)] = E(Y) = \int_{d}^{1-d} (x-d)\, f_Y(y)dy = \int_{d}^{1-d} (x-d)\, f_Y(y)dy$$

$$E(Y) = \int_{d}^{1-d} (x-d)\, \frac{f_X(y+d)}{\bar{F}_X(d)} dy = \frac{1}{\bar{F}_X(d)} \int_{d}^{1-d} (x-d)\, f_X(y+d)dy$$

É necessário migrar a integral em relação a y para x.

$$E(Y) = \frac{1}{\bar{F}_X(d)} \left[\int_{0}^{d} 0\, f_X(x)dx + \int_{d}^{1} (x-d)\, f_X(x)dx \right]$$

$$E(Y) = \frac{1}{0{,}91} \left[\int_{0{,}3}^{1} (x-0{,}3)\, 2xdx \right] \approx 0{,}412821$$

Esperança de variáveis aleatórias

Seja L um valor limite para dentro do domínio de X, e seja Y uma variável aleatória " Valor de X sujeito ao limite L. Então:

$$Y = \begin{cases} X, & X < L \\ L, & X \geq L \end{cases}$$

Logo, para o caso de X se contínuo tem-se que:

$$E(Y) = E(X;L) = \int_{-\infty}^{L} x f_X(x)dx + \int_{L}^{\infty} L f_X(x)dx = \int_{-\infty}^{L} x f_X(x)dx + L\,S_X(L)$$

E no caso de X se discreto, tem-se:

$$E(Y) = E(X;L) = \sum_{i=0}^{x_i<L} x_i P_X(x_i) + \sum_{x_i=L}^{\infty} L P_X(x_i) = \sum_{i=0}^{x_i<L} x_i P_X(x_i) + L\,P(X \geq L)$$

Esperança de variáveis aleatórias

Ao tratar X como uma variável aleatória associada a severidade de sinistros, a modificação que leva a Y, tal que:

$$Y = \begin{cases} X, & X < L \\ L, & X \geq L \end{cases}$$

É conhecida como limite técnico (L) da severidade, ou seja limite máxima de retenção, em que o segurador limita a indenização a um teto, geralmente transferindo os danos que o superarem ao ressegurador.

EXEMPLO 5

Considere uma carteira de seguros com a seguinte função de probabilidade.

S (R$)	0	1000	2000	3000	4000	5000	6000
$P(S)$	0,35	0,31	0,155	0,13	0,04	0,01	0,005

A seguradora determina que irá cobrar dos seus segurados um prêmio baseado no quanto esperam gastar com indenizações, porém esse valor não deve exceder $R\$4000$. Calcule o valor esperado sujeito a esse limite

S (R$)	0	1000	2000	3000	4000	5000	6000
$P(S)$	0,35	0,31	0,155	0,13	0,04	0,01	0,005

Solução:

Seja Y, tal que:

$$Y = \begin{cases} S\,, & S < 4000 \\ 4000\,, & S \geq 4000 \end{cases}$$

$$\Pi_Y = E(Y) = E(S; 4000)$$

$$\Pi_Y = \sum_{s=0}^{3000} s\,P(s) + \sum_{s=4000}^{6000} 4000\,P(s) = R\$1230{,}00$$

Esperança de uma função de variáveis aleatórias

Caso $Y = g(X_1, X_2)$, com X_1 e X_2 variáveis aleatórias contínuas, então:

$$E[g(X_1, X_2)] = \int_{x_1} \int_{x_2} g(X_1, X_2)\, f_{X_1,X_2}(x_1, x_2)\, dx_2 dx_1$$

ou

$$E(Y) = \int_{y} y\, f_Y(y) dy$$

Seja $Y = g(X_1, X_2)$, com X_1 e X_2 variáveis aleatórias discretas, então:

$$E[g(X_1, X_2)] = \sum_{X_1} \sum_{X_2} g(X_1, X_2) P_{X_1,X_2}(x_1, x_2)$$

ou

$$E(Y) = \sum_{y} y\, P_Y(y)$$

EXEMPLO 6

Suponha X e Y duas variáveis aleatórias, tal que: $f_{X,Y}(x,y) = x + y$, $0 < x < 1, \quad 0 < y \leq 1$. Calcule $E[g(X,Y)]$, para:

$$g_1(X,Y) = X^n Y^{\mathrm{k}}$$

$$g_2(X,Y) = 2X + Y$$

$$g_3(X,Y) = X$$

$$g_4(X,Y) = Y$$

Solução:

$$E[g_1(X,Y)] = \int_0^1 \int_0^1 x^n y^k (x+y)\,dx\,dy = \int_0^1 \int_0^1 x^{n+1}\,y^k + x^n y^{k+1} dx\,dy$$

$$E[g_1(X,Y)] = \int_0^1 \left(\frac{x^{n+2}}{n+2} y^k + \frac{x^{n+1}}{n+1} y^{k+1} \right)\Bigg|_{x=0}^{x=1} dy = \int_0^1 \frac{y^k}{n+2} + \frac{y^{k+1}}{n+1} dy$$

$$E[g_1(X,Y)] = \frac{y^{k+1}}{(n+2)(k+1)} + \frac{y^{k+2}}{(n+1)(k+2)}\Bigg|_{y=0}^{y=1}$$

$$E[g_1(X,Y)] = \frac{1}{(n+2)(k+1)} + \frac{1}{(n+1)(k+2)}$$

Solução:

$$E[g_2(X,Y)] = \int_0^1 \int_0^1 (2x+y)(x+y)dxd\,y = \int_0^1 \int_0^1 2x^2 + 3xy + y^2 dx\,dy$$

$$E[g_2(X,Y)] = \int_0^1 \left(\frac{2x^3}{3} + \frac{3x^2y}{2} + xy^2\right)\Bigg|_{x=0}^{x=1} dy = \int_0^1 \frac{2}{3} + \frac{3}{2}y + y^2 dy$$

$$E[g_2(X,Y)] = \frac{2y}{3} + \frac{3y^2}{4} + \frac{y^3}{3}\Bigg|_{y=0}^{y=1} = \frac{21}{12}$$

Solução:

$$E[g_3(X,Y)] = \int_0^1 \int_0^1 x(x+y)dx\,dy = \int_0^1 \int_0^1 x^2 + xy\,dx\,dy$$

$$E[g_3(X,Y)] = \int_0^1 \left(\frac{x^3}{3} + \frac{x^2 y}{2}\right)\Bigg|_{x=0}^{x=1} dy = \int_0^1 \frac{1}{3} + \frac{y}{2} dy$$

$$E[g_3(X,Y)] = \frac{y}{3} + \frac{y^2}{4}\Bigg|_{y=0}^{y=1} = \frac{7}{12} = E(X)$$

Solução:

$$E[g_4(X,Y)] = \int_0^1 \int_0^1 y(x+y)dx\,dy = \int_0^1 \int_0^1 yx + y^2\,dx\,dy$$

$$E[g_4(X,Y)] = \int_0^1 \left(\frac{yx^2}{2} + \frac{y^3}{3}\right)\Bigg|_{x=0}^{x=1} dy = \int_0^1 \frac{y}{2} + \frac{1}{3}\,dy$$

$$E[g_4(X,Y)] = \frac{y^2}{4} + \frac{y}{3}\Bigg|_{y=0}^{y=1} = \frac{7}{12} = E(Y)$$

Solução:

A partir do exemplo dado é possível notar dentre outras coisas que

$$E(Y) = \int_0^1 \int_0^1 y f_{X,Y}(x,y)dxdy = \int_0^1 y\left[\int_0^1 f_{X,Y}(x,y)dx\right]dy = \int_0^1 y f_Y(y)dy$$

$$E(X) = \int_0^1 \int_0^1 x f_{X,Y}(x,y)dydx = \int_0^1 x\left[\int_0^1 f_{X,Y}(x,y)dy\right]dx = \int_0^1 x f_X(x)dx$$

Propriedades do valor esperado*

1) Seja um vetor aleatória $X_1, X_2, \ldots, X_n$ (contínua ou discreta), temos.

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

2) Seja o vetor $X_1, X_2, \ldots, X_n$ for composto por variáveis aleatórias independentes então a seguinte propriedade também é valida.

$$E\left(\prod_{i=1}^{n} X_i\right) = E(X_1)E(X_2) \ldots E(X_n)$$

Demonstração: (sem perda de generalidade)

1)

$$E(X_1 \pm X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X_1 \pm X_2) f_{X_1,X_2}(x_1,x_2)dx_1dx_2 =$$

*Teorema de Fubini

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f_{X_1,X_2}(x_1,x_2)dx_1dx_2 \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f_{X_1,X_2}(x_1,x_2)dx_1dx_2$$

Sobe condições de regularidade é feita uma troca da ordem das integrais

$$\int_{-\infty}^{\infty} x_1 \left[\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2\right]dx_1 \pm \int_{-\infty}^{\infty} x_2 \left[\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_1\right]dx_2$$

$$= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)dx_1 \pm \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2)dx_2$$

$$E(X_1 \pm X_2) = E(X_1) \pm E(X_2)$$

DEMONSTRAÇÃO: (sem perda de generalidade)

2)

$$E\left(\prod_{i=1}^{n} X_i\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}\left(\prod_{i=1}^{n} x_i\right) f_{X_1,\dots,X_n}(x_1,\dots,x_n)dx_1 \dots dx_n$$

A partir da pressuposição de independência temos:

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}(x_n x_{n-1}\, x_{n-2} \dots x_1) f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_n}(x_n) dx_1 \dots dx_n$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}(x_n x_{n-1}\, x_{n-2} \dots) f_{X_2}(x_2) \dots f_{X_n}(x_n)\left[\int_{-\infty}^{\infty} x_1\, f_{X_1}(x_1) dx_1\right] dx_2 \dots dx_n$$

$$E\left(\prod_{i=1}^{n} X_i\right) = E(X_1)E(X_2)\dots E(X_n).$$

Variância probabilística

- A variância da variável aleatória (e seu desvio padrão) mede satisfatoriamente a dispersão dos dados ao redor do valor esperado.

- A definição matemática da variância de uma variável aleatória X é tal que:

$$var(X) = E[(X - \mu_X)^2] = \sigma^2$$

- A raiz quadrada da variância é denominada de desvio-padrão e representado por σ.

Variância probabilística

- Pode se calcular a variância também por:

$$var(X) = E(X^2) - E(X)^2$$

- Como a esperança a variância apresenta certas propriedades, destaca-se neste texto as seguintes:

1) $var(X + k) = var(X)$

2) $var(kX) = k^2 var(X)$

DEMONSTRAÇÃO:

1)

$$var(X+k) = E\{[(X+k) - E(X+k)]^2\}$$

$$var(X+k) = E\{[X+k-E(X)-k]^2\}$$

$$var(X+k) = E\{[X-E(X)]^2\} = var(X)$$

2)

$$var(kX) = E[(kX)^2] - E(kX)^2 = k^2E(X^2) - k^2E(X)^2$$

$$var(kX) = k^2[E(X^2) - E(X)^2] = k^2var(X)$$

Variância probabilística

- Seja duas variáveis aleatórias X_1 e X_2, então:

$$var(X_1 + X_2) = var(X_1) + var(X_2) + 2cov(X_1, X_2)$$

Demonstração:

$$var(X_1 + X_2) = E[(X_1 + X_2)^2] - [E(X_1 + X_2)]^2$$

$$var(X_1 + X_2) = E(X_1^2 + 2X_1X_2 + X_2^2) - [E(X_1) + E(X_2)]^2$$

$$var(X_1 + X_2) = \boldsymbol{E\left(X_1^2 + 2X_1X_2 + X_2^2\right)} - [E(X_1)^2 + 2E(X_1)E(X_2) + E(X_2)^2]$$

$$var(X_1 + X_2) = \boldsymbol{E\left(X_1^2\right)} + \boldsymbol{E(2X_1X_2)} + \boldsymbol{E\left(X_2^2\right)} - E(X_1)^2 - 2E(X_1)E(X_2) - E(X_2)^2$$

$$var(X_1 + X_2) = \left[\boldsymbol{E\left(X_1^2\right)} - E(X_1)^2\right] + \left[\boldsymbol{E\left(X_2^2\right)} - E(X_2)^2\right] + [\boldsymbol{2E(X_1X_2)} - 2E(X_1)E(X_2)]$$

$$var(X_1 + X_2) = var(X_1) + var(X_2) + 2cov(X_1, X_2)$$

Consequentemente

$$var(X_1 - X_2) = var(X_1) - var(X_2) + 2cov(X_1, X_2)$$

Variância de uma variáveis aleatórias

A partir da propriedade anterior é possível perceber que se X e Y são independentes temos que:

$$var(X_1 + X_2) = var(X_1) + var(X_2)$$

Logo:

Seja um vetor aleatória $(X_1, X_2, \ldots, X_k)$ de variáveis aleatória (contínua ou discreta), independentes:

$$var(X_1 + X_2 + \ldots + X_n) = var(X_1) + var(X_2) + \cdots + var(X_n).$$

Variância de uma variáveis aleatórias

Adicionalmente tem-se que para X e Y independendes tem-se:

$$var(XY) = E(X^2)E(Y^2) - E(X)^2E(Y)^2$$

Demonstração:

$$var(XY) = E[(XY)^2] - E(XY)^2 = E(X^2)E(Y^2) - [E(X)E(Y)]^2$$

$$var(XY) = E(X^2)E(Y^2) - E(X)^2E(Y)^2$$

- A probabilidade condicional satisfaz todas as condições para ser, ela mesma, uma probabilidade.

- Logo, pode-se, em tese, repetir todo o desenvolvimento feito para o valor esperado, usando probabilidades condicionadas a algum evento.

Valor esperado Condicional

Para o caso de modelos condicionais tem-se que a esperança condicional de $Y|X = x$ é então definida no caso de variáveis continuas como:

Variáveis aleatórias discretas

$$E(Y|X = x) = \sum_{i=1}^{\infty} y_i P(y_i|x)$$

Variáveis aleatórias Contínuas

$$E(Y|X = x) = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x) dy$$

Teorema

Sendo X e Y variáveis no mesmo espaço de probabilidade, temos

$$E[E(Y|X)] = E(Y)$$

Desde que as esperanças existam.

DEMONSTRAÇÃO

Seja X e Y variáveis aleatórias continuas definidas no mesmo espaço de probabilidades então:

$$E[E(Y|X)] = E[h(X)]$$

Pois $h(X)$ é uma função de X já que :

$$E(Y|X) = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|x) dy = h(X)$$

DEMONSTRAÇÃO

$$E[h(X)] = \int_{-\infty}^{\infty} h(X) f_X(x) dx$$

$$E[h(X)] = \int_{-\infty}^{\infty} E(Y|X) f_X(x) dx$$

$$E[h(X)] = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y\, f_{Y|X}(y|x) dy \right] f_X(x) dx$$

$$E[h(X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \left[f_{Y|X}(y|x) f_X(x) \right] dy dx$$

$$E[h(X)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\, f_{X,Y}(x,y) dy dx = E(Y)$$

$$E[h(X)] = \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx \right] d_y$$

$$E[h(X)] = \int_{-\infty}^{\infty} y [f_Y(y)] d_y = E(Y)$$

$$E[E(Y|X)] = E(Y)$$

Variância condicional

- E a variância condicional de X, dado que $\mathrm{Y} = \mathrm{y}$, é o valor esperado dos desvios da variável em relação à esperança condicional, ou seja:

$$var(X|Y = y\,) = E\{[X - E(X|Y = y\,)\,]^2\} = E(X^2|Y)\ - E(X|Y)^2$$

supondo a existência das esperanças envolvidas.

- Adicionalmente, tem-se também que:

$$E[var(X|Y)] = var(X) - var\big(E(X|Y)\big)$$

DEMONSTRAÇÃO

Seja $E[var(X|Y)] = E[E(X^2|Y=y) - [E(X|Y=y)]^2]$

$$E[var(X|Y)] = E[E(X^2|Y)] - E\left[\left(E(X|Y)\right)^2\right]$$

- Usando da propriedade de valor esperado

$$E[E(X^2|Y)] = E(X^2)$$

- Então

$$E[var(X|Y)] = E(X^2) - E(X)^2 - E[E(X|Y)^2] + E(X)^2$$

$$E[var(X|Y)] = var(X) - E[E[X|Y]^2] + E[E(X|Y)]^2$$

$$E[var(X|Y)] = var(X) - var\big(E(X|Y)\big)$$

Covariância e Correlação entre variáveis aleatórias

- Dado X e Y duas variáveis aleatórias no mesmo espaço de probabilidade. A covariância entre X e Y é definida por:

$$Cov(X,Y) = \sigma_{X,Y} = E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big]$$

ou

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

DEMONSTRAÇÃO

$$E\{[X - E(X)][Y - E(Y)]\} = E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$E\{[X - E(X)][Y - E(Y)]\} = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

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

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

desde que as esperanças, presentes na expressão existam.

COVARIÂNCIA E CORRELAÇÃO ENTRE VARIÁVEIS ALEATÓRIAS

- ...a covariância poderia ser definida como o valor esperado do produto dos desvios das variáveis, em relação às suas médias.

- A covariância é uma medida de dependência linear entre as variáveis e pode assumir valores de qualquer sinal.

- Se a covariância for zero não há dependência linear entre as variáveis, mas pode haver outro tipo de relação entre elas e, portanto, não pode-se dizer que são independentes.

COVARIÂNCIA E CORRELAÇÃO ENTRE VARIÁVEIS ALEATÓRIAS

Corolário: Se as variáveis aleatórias X e Y, forem independentes, $E(XY)$ fatora no produto das esperanças, e a covariância se anula.

Demonstração:

Seja X e Y independentes então :

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$
$$Cov(X,Y) = E(X)E(Y) - E(X)E(Y) = 0$$

A recíproca do corolário não é verdadeira.

COVARIÂNCIA E CORRELAÇÃO ENTRE VARIÁVEIS ALEATÓRIAS

Proposição: Se X e Y são variáveis aleatórias de valor real e a, b, c e d constantes, então verificamos as seguintes propriedades:

(II) $Cov(X,Y) = Cov(Y,X)$

(II) $Cov(X,X) = var(X)$

(III) $Cov(aX + b, cY + d) = ac\, Cov(X,Y)$

Demonstração:

(I)

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$
$$= E[(Y - \mu_Y)(X - \mu_X)] = Cov(Y,X)$$

(II)

$$Cov(X,X) = E(XX) - \mu_X \mu_X$$
$$Cov(X,X) = E(X^2) - {\mu_X}^2 = var(x)$$

(III)

$$E(aX + b) = a\mu_X + b$$
$$E(cY + d) = c\mu_Y + d$$

$$Cov(aX + b, cY + d) = E\left[\left(aX + b - E(aX + b)\right)\left(cY + d - E(cY + d)\right)\right]$$

$$Cov(aX + b, cY + d) = E[(aX + b - a\mu_X - b)(cY + d - c\mu_Y - d)]$$

$$Cov(aX + b, cY + d) = E[(aX - a\mu_X)(cY - c\mu_Y)]$$

$$Cov(aX + b, cY + d) = acE[(X - \mu_X)(Y - \mu_Y)]$$

$$Cov(aX + b, cY + d) = ac\ Cov(X, Y)$$

COVARIÂNCIA E CORRELAÇÃO ENTRE VARIÁVEIS ALEATÓRIAS

O coeficiente de correlação, denotado por $\rho[X, Y]$ ou $\rho_{X,Y}$ entre as variáveis aleatórias X e Y é denotado por:

$$Corr(X, Y) = \rho_{X,Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

Dado que σ_X e σ_Y existem, e $\sigma_Y > 0$ e $\sigma_X > 0$.

$$-1 \leq \rho_{X,Y} \leq 1$$

Exemplo 7

O centro acadêmico de uma faculdade de administração fez um levantamento da remuneração dos estágios os alunos, em salários mínimos, com relação ao ano que estão cursando, as probabilidades de cada caso são apresentadas a baixo:

Salário\Ano	2	3	4	5	$P(Sal. = x)$
2	2/25	2/25	1/25	0	5/25
3	2/25	5/25	2/25	2/25	11/25
4	1/25	2/25	2/25	4/25	9/25
$P(Ano = y)$	5/25	9/25	5/25	6/25	1

Calcule $E(Y|X)$ e a covariância entre X e Y.

Y	$P(Y\|X=2)=\dfrac{P(Y,X=2)}{P(X=2)}$	$P(Y\|X=3)=\dfrac{P(Y,X=3)}{P(X=3)}$	$P(Y\|X=4)=\dfrac{P(Y,X=4)}{P(X=4)}$
2	$\dfrac{\frac{2}{25}}{\frac{5}{25}}=\dfrac{2}{5}$	$\dfrac{\frac{2}{25}}{\frac{11}{25}}=\dfrac{2}{11}$	$\dfrac{\frac{1}{25}}{\frac{9}{25}}=\dfrac{1}{9}$
3	$\dfrac{\frac{2}{25}}{\frac{5}{25}}=\dfrac{2}{5}$	$\dfrac{\frac{5}{25}}{\frac{11}{25}}=\dfrac{5}{11}$	$\dfrac{\frac{2}{25}}{\frac{9}{25}}=\dfrac{2}{9}$
4	$\dfrac{\frac{1}{25}}{\frac{5}{25}}=\dfrac{1}{5}$	$\dfrac{\frac{2}{25}}{\frac{11}{25}}=\dfrac{2}{11}$	$\dfrac{\frac{2}{25}}{\frac{9}{25}}=\dfrac{2}{9}$
5	$\dfrac{0}{\frac{5}{25}}=0$	$\dfrac{\frac{2}{25}}{\frac{11}{25}}=\dfrac{2}{11}$	$\dfrac{\frac{4}{25}}{\frac{9}{25}}=\dfrac{4}{9}$
	1	**1**	**1**

$$E(Y|X=x)=\sum_{i=1}^{\infty} y_i P(y_i|x)$$

$E[Y\|X]$	$2\left(\frac{2}{5}\right)+3\left(\frac{2}{5}\right)+4\left(\frac{1}{5}\right)+5(0)=\frac{14}{5}$	$2\left(\frac{2}{11}\right)+3\left(\frac{5}{11}\right)+4\left(\frac{2}{11}\right)+5\left(\frac{2}{11}\right)=\frac{37}{11}$	$2\left(\frac{1}{9}\right)+3\left(\frac{2}{9}\right)+4\left(\frac{2}{9}\right)+5\left(\frac{4}{9}\right)=\frac{36}{9}$
X	2	3	4

$$Cov(X,Y) = E(XY) - \mu_X \mu_Y$$

$$\begin{aligned} \boldsymbol{E(XY)} \\ &= \left[(2\times 2)\frac{2}{25} + (2\times 3)\frac{2}{25} + (2\times 4)\frac{1}{25} + (2\times 5)0\right] \\ &+ \left[(3\times 2)\frac{2}{25} + (3\times 3)\frac{5}{25} + (3\times 4)\frac{2}{25} + (3\times 5)\frac{2}{25}\right] \\ &+ \left[(4\times 2)\frac{1}{25} + (4\times 3)\frac{2}{25} + (4\times 4)\frac{2}{25} + (4\times 5)\frac{4}{25}\right] = 11,32 \end{aligned}$$

$$E(X) = 2\left(\frac{5}{25}\right) + 3\left(\frac{11}{25}\right) + 4\left(\frac{9}{25}\right) = 3,16$$

$$E(Y) = 2\left(\frac{5}{25}\right) + 3\left(\frac{9}{25}\right) + 4\left(\frac{5}{25}\right) + 5\left(\frac{6}{25}\right) = 3,48$$

$$\boldsymbol{Cov(X,Y) = 11,32 - (3,16)(3,48) = 0,3232}$$

$$E(X) = 2\left(\frac{5}{25}\right) + 3\left(\frac{11}{25}\right) + 4\left(\frac{9}{25}\right) = 3{,}16$$

$$E(X^2) = 2^2\left(\frac{5}{25}\right) + 3^2\left(\frac{11}{25}\right) + 4^2\left(\frac{9}{25}\right) = 10{,}52$$

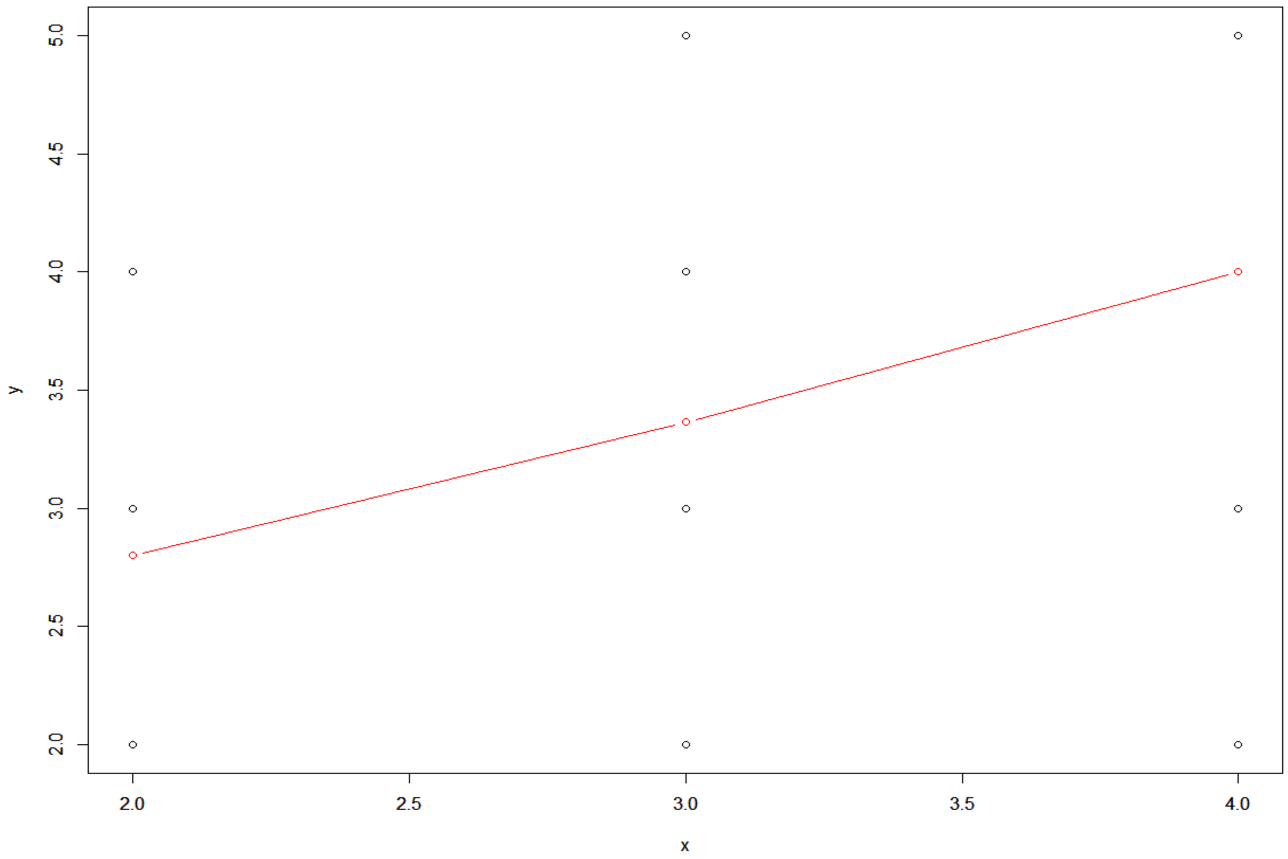
$$\boldsymbol{\sigma_X = \sqrt{var(x)} = \sqrt{10,52 - 3,16^2} = 0,73102}$$

$$E(Y) = 2\left(\frac{5}{25}\right) + 3\left(\frac{9}{25}\right) + 4\left(\frac{5}{25}\right) + 5\left(\frac{6}{25}\right) = 3{,}48$$

$$E(Y^2) = 2^2\left(\frac{5}{25}\right) + 3^2\left(\frac{9}{25}\right) + 4^2\left(\frac{5}{25}\right) + 5^2\left(\frac{6}{25}\right) = 13{,}24$$

$$\boldsymbol{\sigma_Y = \sqrt{var(x)} = \sqrt{10,52 - 3,16^2} = 1,0628}$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{0{,}3232}{(0{,}73102)(1{,}0628\ )} = 0{,}4159973$$



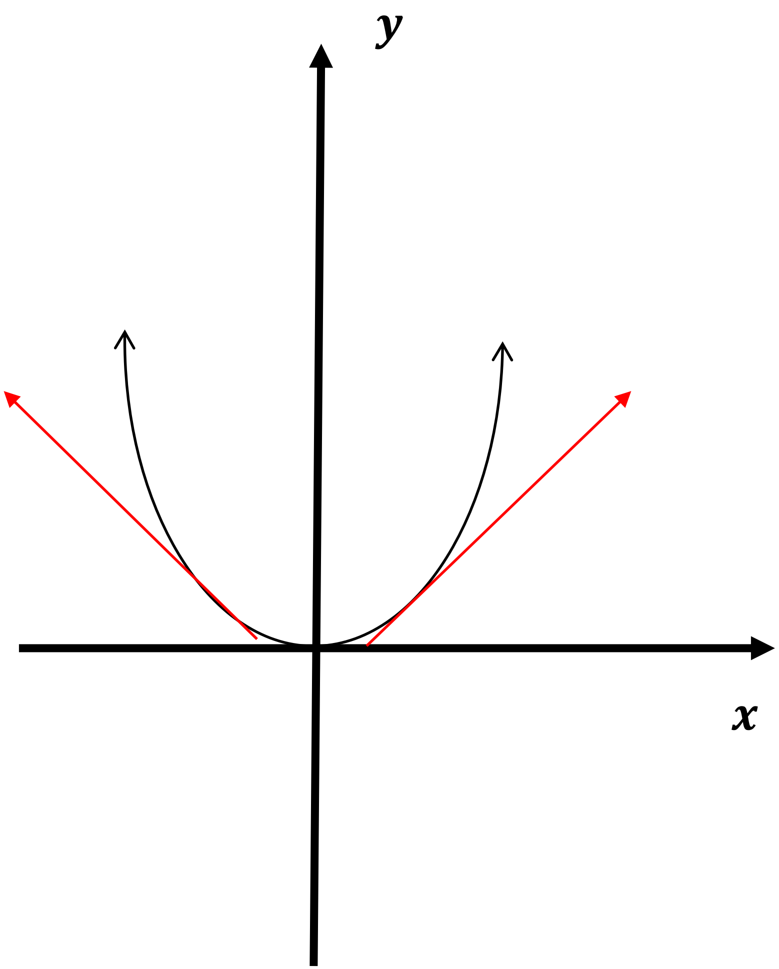
O valor esperado condicional $E(Y|X = x)$ é chamado de curva de regressão.

DESIGUALDADE DE JENSEN

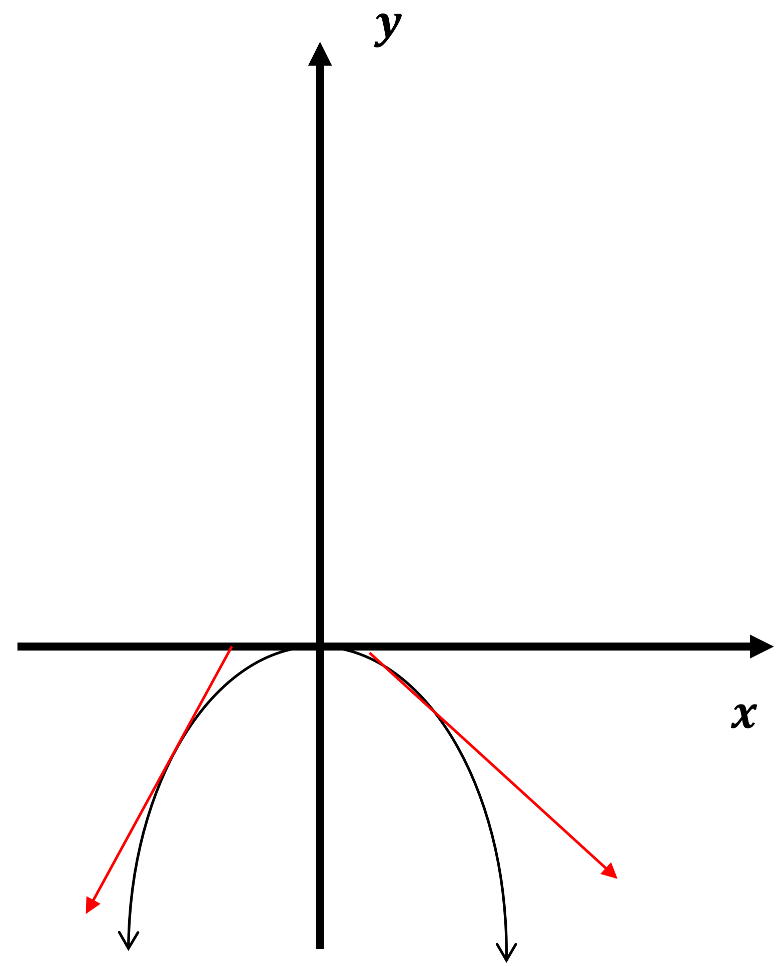
- Algumas proposições envolvendo a média, como desigualdades, auxiliam no estudo do comportamento das variáveis aleatórias.

- Elas fornecem limites, em função da variância e de momentos, para a probabilidade da variável aleatória exceder um certo valor de interesse.

DESIGUALDADE DE JENSEN



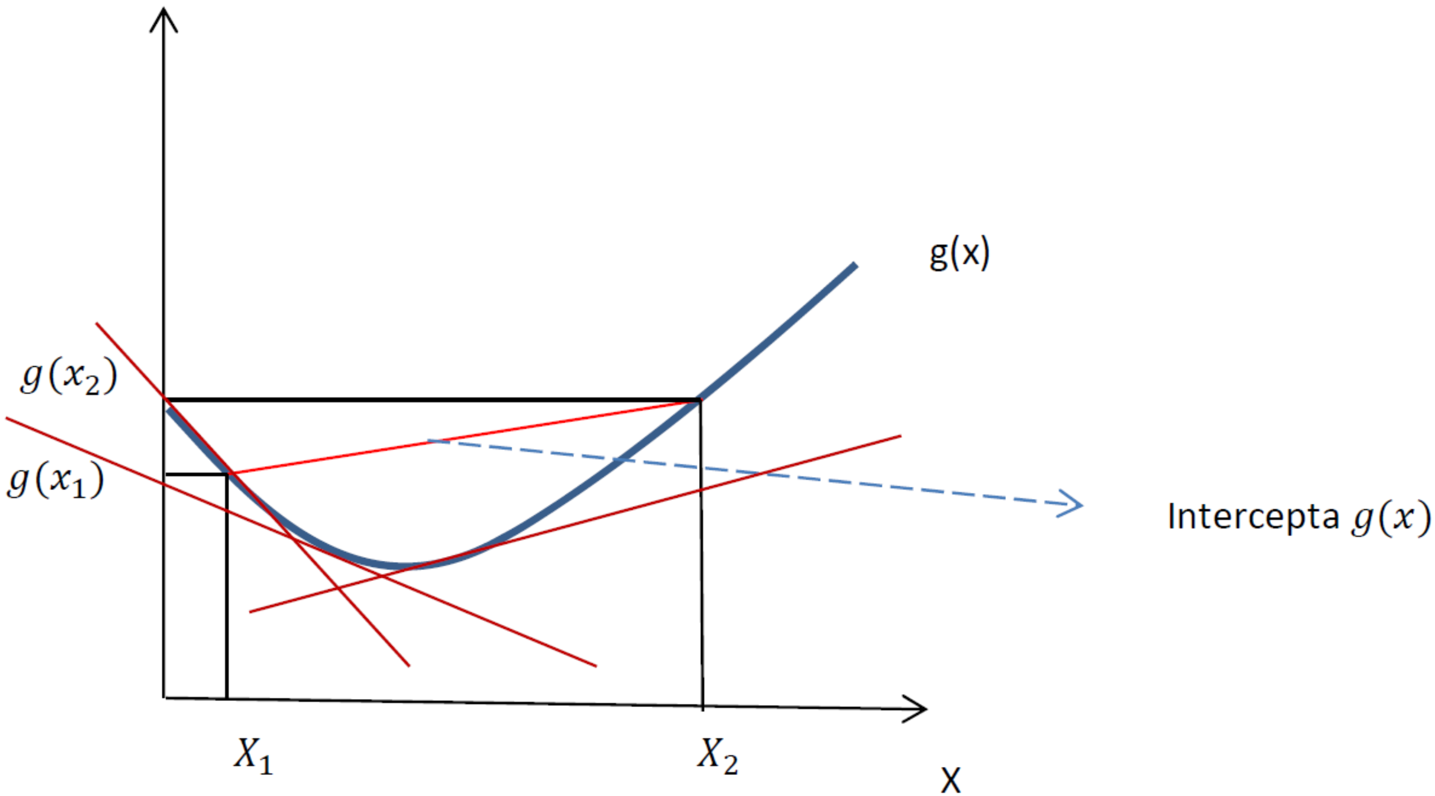
$$g''(x) > 0$$



$$g''(x) < 0$$

DESIGUALDADE DE JENSEN

- Função Convexa



- A função é côncava se $-g(x)$ for convexa

DESIGUALDADE DE JENSEN

- Teorema:

  - Seja X uma variável aleatória com média $E(X)$ e seja $g(.)$uma função convexa. Então:

$$E[g(X)] \geq g\big(E(X)\big) \quad \rightarrow \quad g''(x) > 0$$

$$E[g(X)] \leq g\big(E(X)\big) \quad \rightarrow \quad g''(x) < 0$$

DESIGUALDADE DE JENSEN

Um dentre os muitos resultados que podem ser obtidos mediante a desigualdade de Jensen é o relacionado a correlação, uma vez que através das desigualdades de Jensen e Cauchy-Schwarz pode-se demonstrar que a correlação não pode exceder 1 em valor absoluto.

A desigualdade de Cauchy-schwarz estabelece que dadas as variáveis aleatórias X e Y com $\mathrm{E}(\mathrm{X}^2)$ e $\mathrm{E}(\mathrm{Y}^2)$ finitas, então:

$$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$$

Referências

- Magalhães, M. N. Lima, A. C. P. **Noções de Probabilidade e Estatística**, Editora USP: SAO Paulo, 2001.
- JAMES,B. R.; **Probabilidade: Um Curso em nível intermediário**, IMPA, Rio de Janeiro, 3 ed., 2004.
- PIRES,M.D.;COSTA,L.H.;FERREIRA,L.;MARQUES,R. **Teoria do risco atuarial: Fundamentos e conceitos.** Curitiba, CRV 2020.