Capítulo 3. Esperança Matemática ou Expectância de Variáveis Aleatórias

Variáveis Aleatórias Unidimensionais e Variáveis Aleatórias Bidimensionais

No estudo do primeiro conceito deste capítulo convém inicialmente apresentar um exemplo bastante simples que permite estabelecer a noção básica de expectância ou valor esperado e motivar a definição do mesmo.

Exemplo 1. Dois jogadores, A e B, alternam-se no jogo de lançamento de uma moeda, apostando de acordo com a seguinte regra: se der "cara", A ganha R$ 1,00 (e B perde essa quantia); e se der "coroa", A perde R$ 1,00 (B ganha essa quantia). Suponha-se que uma sequência de n lançamentos – rodadas do jogo – é realizada. Então, denotando por K o resultado "cara", por C o resultado "coroa" e por n_K e n_C , respectivamente, os números de lançamentos em que ocorrem "cara" e "coroa", dentre os n lançamentos realizados (ou seja, as frequências absolutas de "cara" e de "coroa") tem-se:

a) o lucro do jogador A, denotado por L_A , é

$$L_A = n_K .1{,}00 + n_C .(-1{,}00) = n_K - n_C$$

b) o lucro médio de A por rodada do jogo (por lançamento), denotado por m_A é

$$m_A = \frac{1}{n} L_A = \frac{1}{n} n_K - \frac{1}{n} n_C = f_K - f_C$$, onde f_K e f_C denotam as frequências relativas de "caras" e "coroas", respectivamente.

Se o número de lançamentos, n, for suficientemente grande, e a moeda equilibrada, é de se esperar que se verifique, ao menos aproximadamente

$$f_K = \frac{n_K}{n} \to p_K = \frac{1}{2} \text{ e } f_C = \frac{n_C}{n} \to p_C = \frac{1}{2}$$

Portanto, intuitivamente, é de se esperar que o lucro médio (por lançamento) do jogador A após n lançamentos seja zero (ao menos aproximadamente)

$$m_A \simeq p_K - p_C = \frac{1}{2} - \frac{1}{2} = 0$$

isto é, que haja uma espécie de empate, no qual A (e B) não ganhe nem perca dinheiro.

Esse procedimento intuitivo relaciona-se ao conceito de valor esperado, expectância ou média de uma variável aleatória.

A seguir é apresentada a definição de esperança matemática, expectância, valor esperado ou média de uma variável aleatória unidimensional real.

3.1 Expectância de X

Definição:

(I) para X discreta:

$$E(X) = \sum_{x \in R_X} x\, p_X(x)$$

(II) para X contínua:

$$E(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\,dx = \int_{x \in R_X} x\, f_X(x)\,dx$$

Notas:

- Se X é discreta, a expectância pode não existir, caso a expressão que a define seja uma série divergente. Nesse caso, diz-se que a expectância é infinita.
- Se X é contínua, a expectância pode não existir, caso a expressão que a define seja uma integral divergente. Nesse caso, diz-se que a expectância é infinita.
- Se existe, a expectância é um número real; diz-se então que a expectância é finita
- A expectância também pode ser denominada: Expectativa, Esperança Matemática, Valor Esperado ou Média.
- Descritivamente, a expectância é uma medida de posição (ou de localização) da distribuição da variável aleatória.
- A expectância de X é o valor para o qual tende a média aritmética dos valores assumidos por essa variável aleatória, quando o número de observações tende a infinito.
- É usual representar-se os valores numéricos da expectância pela letra grega μ

Exemplo 2. Em certa fábrica, 10% das peças produzidas são defeituosas. Cada peça não defeituosa gera um lucro de R$ 3,00 enquanto cada peça defeituosa acarreta uma perda de R$ 1,00. Se X for a variável aleatória que representa o lucro líquido por peça produzida, determinar o valor esperado desse lucro (ou lucro médio por peça).

Solução

Nessas condições, tem-se

$$E(X) = 3{,}00 . 0{,}9 + (-1{,}00) . 0{,}1 = 2{,}70 - 0{,}10 = 2{,}60$$

Esse exemplo mostra a importância do conceito na Economia.

Exemplo 3. Seja X uma variável aleatória com a seguinte função de probabilidade

$$p(x) = \begin{cases} \frac{1}{4} \ , \ \text{para } x = -1 \\ \frac{1}{2} \ , \ \text{para } x = 0 \\ \frac{1}{4} \ , \ \text{para } x = 1 \end{cases}$$

Determinar o valor esperado (expectância) de X.

Solução

A variável aleatória X é do tipo discreto e o seu valor esperado (ou média) é

$$E(X) = -1\frac{1}{4} + 0\frac{1}{2} + 1\frac{1}{4} = 0$$

Esse exemplo mostra que se a distribuição de X (no caso, a função de probabilidade da variável) for simétrica a sua expectância ou média coincide com o valor da abscissa do eixo de simetria.

Exemplo 4. Seja X uma variável aleatória com função de probabilidade

$$p(x) = \frac{x}{6} \ , \ \text{para } x = 1, 2, 3$$

Determinar o valor esperado (expectância) de X.

Solução

A variável aleatória X é do tipo discreto e a sua função de probabilidade é definida por uma expressão analítica. Logo a expectância (ou média) de X é

$$E(X) = \sum_{x=1}^{3} x\,\frac{x}{6} = \frac{1}{6}\sum_{x=1}^{3} x^2 = \frac{1}{6}(1+4+9) = \frac{14}{6} = \frac{7}{3} = 2{,}33$$

A comparação entre os resultados desse exemplo e do anterior é muito útil para se compreender o conceito de valor esperado, expectância ou média. Note-se que o valor esperado ou expectância de X nesse caso situa-se à direita do valor "central" de X, que é 2, pelo fato de que a distribuição de probabilidade da variável X é crescente com o valor da mesma.

Exemplo 5. Seja X uma variável aleatória com função de densidade de probabilidade

$$f(x) = 1 \text{ , para } 0 < x < 1$$

Determinar a expectância de X.

Solução

A variável aleatória X é do tipo contínuo e a sua expectância (ou média) é

$$E(X) = \int_0^1 x\,1\,dx = \frac{1}{2}x^2\Big|_0^1 = \frac{1}{2} = 0{,}5$$

Note-se, mais uma vez, que a expectância é igual ao valor da abscissa do eixo de simetria da distribuição (nesse caso da função de densidade de probabilidade).

Exemplo 6. Seja X uma variável aleatória com função de densidade de probabilidade

$$f(x) = 2x \text{ , para } 0 < x < 1$$

Determinar a expectância de X.

Solução

A variável aleatória X é do tipo contínuo, então a sua expectância (ou média) é

$$E(X) = \int_0^1 x\,2x\,dx = 2\int_0^1 x^2\,dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3} = 0{,}67$$

Assim como anteriormente, é útil comparar os resultados obtidos nesse exemplo e no anterior. De fato, ao contrário do anterior, no qual a função de densidade da variável aleatória é constante, neste exemplo a função de densidade da variável X é crescente com o valor dela, isto é o valor da função densidade aumenta no seu intervalo de definição quando aumenta o valor x da variável X.

Exemplo 7. Seja X uma variável aleatória com função de densidade de probabilidade

$$f(x) = \alpha e^{-\alpha x}, \text{ para } x > 0, \text{ com } \alpha > 0$$

Determinar a expectância de X.

Solução

A variável aleatória X é contínua e sua expectância (ou média) é expressa por

$$E(X) = \int_0^\infty x\, f(x)\, dx = \int_0^\infty x\, \alpha\, e^{-\alpha x}\, dx$$

Essa integral pode ser resolvida empregando-se o método de integração por partes. Com efeito, fazendo u = x e $dv = \alpha e^{-\alpha x}\, dx$ segue que:

i) du = dx

ii) $dv = \alpha e^{-\alpha x}\, dx = e^{-\alpha x}\, d(\alpha x) \Rightarrow v = -e^{-\alpha x}$

Logo

$$E(X) = \int_0^\infty x\, \alpha\, e^{-\alpha x}\, dx = -x e^{-\alpha x}\Big|_0^\infty + \int_0^\infty e^{-\alpha x} dx = -x e^{-\alpha x}\Big|_0^\infty - \frac{1}{\alpha} e^{-\alpha x}\Big|_0^\infty = x e^{-\alpha x}\Big|_\infty^0 + \frac{1}{\alpha} e^{-\alpha x}\Big|_\infty^0 =$$

$$= 0 - \lim_{x\to\infty} x e^{-\alpha x} + \frac{1}{\alpha} - \lim_{x\to\infty} \frac{1}{\alpha} e^{-\alpha x} = 0 - 0 + \frac{1}{\alpha} - 0 = \frac{1}{\alpha}$$

pois:

i) $\lim_{x\to\infty} x e^{-\alpha x} = 0$

e

ii) $\lim_{x\to\infty} \frac{1}{\alpha} e^{-\alpha x} = \frac{1}{\alpha} \lim_{x\to\infty} e^{-\alpha x} = 0$

Portanto $E(X) = \dfrac{1}{\alpha}$

O próximo exemplo serve para mostrar que nem sempre existe a expectância de uma variável aleatória.

Exemplo 8. Seja X uma variável aleatória com função de densidade de probabilidade

$$f(x) = \frac{1}{\pi(1+x^2)}, \text{ para } -\infty < x < \infty$$

Então a variável aleatória X não possui expectância pois a integral correspondente não é absolutamente convergente. Com efeito,

$$\int_{-\infty}^{\infty} |x| f(x)\, dx = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} dx$$

é divergente pois

$$\frac{1}{\pi}\int_{0}^{\infty} \frac{x}{1+x^2} dx \;=\; \frac{1}{\pi}\int_{-\infty}^{0} \frac{(-x)}{1+x^2} dx \;=\; \infty$$

3.2 Expectância de uma Função de X

Considere-se a variável aleatória unidimensional $g(X)$, onde g(.) é uma função real de uma variável aleatória real X. Prova-se que a expectância de $g(X)$ pode ser calculada da seguinte maneira:

(I) para X discreta:

$$E\left[g(X)\right] = \sum_{x \in R_X} g(x)\, p_X(x)$$

(II) para X contínua:

$$E\left[g(X)\right] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx = \int_{x \in R_X} g(x) f_X(x) dx$$

3.3 Independência de Variáveis Aleatórias

Diz-se que as variáveis aleatórias X e Y são independentes se e somente se

$$P\{X \in A\ ;\ Y \in B\} = P\{X \in A\} P\{Y \in B\}$$

quaisquer que sejam $A \subset R$ e $B \subset R$, tais que $A \neq B$.

3.4 Variável Aleatória Quase Certamente Constante

Uma variável aleatória X é denominada "Quase certamente constante" (ou "*degenerada*") se existe um número real $\boldsymbol{r}$ tal que

$$P\{X = \boldsymbol{r}\} = 1$$

Diz-se, então, que X é "Quase certamente igual a $\boldsymbol{r}$" (ou que X é *degenerada* em $\boldsymbol{r}$)

3.5 Propriedades da Expectância

Sejam X e Y variáveis aleatórias que possuem expectâncias finitas e $\boldsymbol{r}$, a e b constantes reais.

Então,

a) $E(X) = \boldsymbol{r}$, se X é uma variável aleatória quase certamente igual a $\boldsymbol{r}$. Neste caso, escreve-se: $E(\boldsymbol{r}) = \boldsymbol{r}$. É usual entender-se $\boldsymbol{r}$ como se fosse uma constante real.

b) $E(aX) = a\,E(X)$, onde a é qualquer constante real

c) $E(a\,X + b) = a\,E(X) + b$

d) $E(X + Y) = E(X) + E(Y)$

e) Se X e Y são independentes, $E(XY) = E(X)E(Y)$

Nota: As propriedades (b), (c) e (d) caracterizam $E(X)$ como um operador linear.

3.6 Variância de X

Definição:

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

Cálculo:

(I) para X discreta:

$$\text{Var}(X) = \sum_{x \in R_X} [x - E(X)]^2 p_X(x)$$

(II) para X contínua:

$$\text{Var}(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x)\,dx$$

Notas:

- Descritivamente, a variância é uma medida de dispersão da distribuição de X em torno de $E(X)$. Quanto mais dispersa (ou menos concentrada) for a distribuição de X em torno de $E(X)$, maior será $\text{Var}(X)$.
- A variância de X pode não existir. Se $E(X)$ não existe, então $\text{Var}(X)$ também não existe. Entretanto, a existência de $E(X)$ não é suficiente para que $\text{Var}(X)$ exista.
- Se $\text{Var}(X)$ existe, diz-se que X possui variância finita.
- $\text{Var}(X)$ está expressa no quadrado da unidade de medida de X.
- É usual representar-se os valores numéricos da variância por σ^2 (letra grega σ elevada ao quadrado).

3.7 Propriedades da Variância

Sejam X e Y variáveis aleatórias que possuem variância finita, e a, b e c constantes reais.

Então,

a) $\text{Var}(X) \geq 0$ e $\text{Var}(X) = 0$ se e somente se X é quase certamente constante.

b) Var (c) = 0, onde c é uma constante real (ou seja X é quase certamente igual a c)

c) $\text{Var}(aX) = a^2\,\text{Var}(X)$, onde a é uma constante real

d) $\text{Var}(aX + b) = a^2\,\text{Var}(X)$, onde a e b são constantes reais

e) se X e Y são independentes, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

f) $\text{Var}(X) = E\left(X^2\right) - E^2(X)$

Regra mnemônica: A variância de X é igual à média dos quadrados de X menos o quadrado da média de X.

3.8 Desvio Padrão de X

Definição:

$$DP(X) = +\sqrt{Var(X)} = \sigma_X$$

Notas:

- $DP(X) = \sigma_X$ é uma medida de dispersão da distribuição de X em torno de sua expectância.
- $DP(X) = \sigma_X$ está expresso na mesma unidade de medida de X.

3.9 Propriedades do Desvio Padrão

a) $DP(X) \geq 0$ e $DP(X) = 0$ se e somente se X é quase certamente constante

b) $DP(aX) = |a| DP(X)$ onde a é qualquer número real

3.10 Coeficiente de Variação de X

Definição:

$$\gamma(X) = \frac{DP(X)}{E(X)} = \frac{\sigma}{\mu}, \quad \text{para } E(X) = \mu \neq 0$$

Notas:

- $\gamma(X)$ é uma medida relativa de dispersão, adimensional
- $\gamma(X)$ é utilizado para comparar dispersões de distribuições que são expressas em unidades de medida distintas
- É usual expressar $\gamma(X)$ na forma de um percentual

Exemplo 9. Determinar a variância e o desvio padrão da variável aleatória X considerada no exemplo 3.

Solução

A função de probabilidade de X é

$$p(x) = \begin{cases} \frac{1}{4}, & \text{para } x = -1 \\ \frac{1}{2}, & \text{para } x = 0 \\ \frac{1}{4}, & \text{para } x = 1 \end{cases}$$

Como já foi visto anteriormente, no exemplo 3, a média de X é E(X) = 0. Logo, a variância de X, calculada pela definição, é

$$V(X)=E\left[\left(X-E(X)\right)^2\right]=E\left[\left(X-0\right)^2\right]=E\left(X^2\right)$$

Portanto, nesse caso

$$V(X)=E\left(X^2\right)=\sum_{x=-1}^{1}x^2\,p(x)=(-1)^2\cdot\frac{1}{4}+0^2\cdot\frac{1}{2}+1^2\cdot\frac{1}{4}=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

Ou seja,

$$V\left(X\right)=\sigma^2=\frac{1}{2}\text{ e }\sigma=\sqrt{\frac{1}{2}}=\frac{1}{\sqrt{2}}=\frac{\sqrt{2}}{2}=\frac{1,4142}{2}=0,7071=0,71$$

Note-se que o cálculo acima foi facilitado pelo fato de que E(X) = 0.

Exemplo 10. Seja Y uma variável aleatória do tipo discreto com a seguinte função de probabilidade

$$p_Y(y)=\frac{1}{5}\qquad\text{para }\ y=-2,-1,0,1,2$$

Determinar a variância e o desvio padrão de Y e compará-los, respectivamente com a variância e o desvio padrão da variável aleatória X do exemplo anterior, interpretando o resultado.

Solução

É fácil verificar que assim como no caso anterior a distribuição de probabilidade é simétrica em relação ao ponto y = 0. Logo a sua expectância (média) é igual a zero

$$E\left(Y\right)=\sum_{y=-2}^{2}y\,p_Y(y)=-2\cdot\frac{1}{5}-1\cdot\frac{1}{5}+0\cdot\frac{1}{5}+1\cdot\frac{1}{5}+2\cdot\frac{1}{5}=0$$

Consequentemente, tem-se

$$V\left(Y\right)=E\left(Y^2\right)$$

Portanto

$$V(Y)=E(Y^2)=\sum_{y=-2}^{2} y^2 p(y)=\frac{1}{5}\sum_{y=-2}^{2} y^2=\frac{1}{5}\left[(-2)^2+(-1)^2+0^2+1^2+2^2\right]=\frac{1}{5}(4+1+0+1+4)$$

donde $V(Y)=\frac{10}{5}=2$ e $\sigma_Y=\sqrt{V(Y)}=\sqrt{2}=1{,}4142=1{,}41$

Comparando as variâncias de X e de Y verifica-se que a de Y é maior, sendo igual a quatro vezes a variância de X. Esse resultado era esperado uma vez que a distribuição de X é mais concentrada em torno da média do que a de Y. Com efeito, não só a variável X possui menos valores que a de Y como também as probabilidades associadas aos seus valores são maiores em torno da média do que os desta última.

Note-se também que, assim como no exemplo anterior, o cálculo acima foi facilitado pelo fato de que a sua expectância é nula: E(Y) = 0. Como, em geral, a expectância de uma variável aleatória é diferente de zero, torna-se mais fácil efetuar o cálculo da variância empregando a expressão $V(X)=E(X^2)-E^2(X)$ (uma das propriedades da variância) como é mostrado nos próximos exemplos.

Exemplo 11. Seja X a variável aleatória que representa o número eventual de pontos obtidos no lançamento de um dado equilibrado. Determinar a variância, o desvio padrão e o coeficiente de variação de X.

Solução

A função de probabilidade de X é

$$p(x)=\frac{1}{6}\ , \text{ para } x=1,2,3,\ \dots\ ,6$$

Como a função de probabilidade tem uma expressão analítica, é possível efetuar o cálculo de forma algébrica, como é mostrado a seguir.

O cálculo da expectância é bastante

$$E(X)=\sum_{x=1}^{6} x\,\frac{1}{6}=\frac{1}{6}\sum_{x=1}^{6} x=\frac{1}{6}\,\frac{(1+6)\cdot 6}{2}=\frac{7}{2}=3{,}5$$

Como se pode verificar, no cálculo algébrico de expressões dessa natureza convém empregar os resultados conhecidos sobre somas de potências de números naturais.

$$E(X^2)=\sum_{x=1}^{6} x^2\,p(x)=\sum_{x=1}^{6} x^2\,\frac{1}{6}=\frac{1}{6}\sum_{x=1}^{6} x^2=\frac{1}{6}\cdot\frac{6\cdot(6+1)\cdot(2\cdot 6+1)}{6}=\frac{91}{6}$$

Portanto

$$V(X) = E(X^2) - E^2(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

Quando a variável aleatória é do tipo discreto pode ser utilizado um dispositivo prático para cálculo da variância, por meio de uma tabela, como é mostrado a seguir.

x	p(x)	x p(x)	x^2 p(x)
1	1/6	1/6	1/6
2	1/6	2/6	4/6
3	1/6	3/6	9/6
4	1/6	4/6	16/6
5	1/6	5/6	25/6
6	1/6	6/6	36/6
Total	1	21/6	91/6

Assim

$$E(X) = \sum_{x=1}^{6} x\, p(x) = \frac{21}{6} = \frac{7}{2} = 3,5$$

Além disso, tem-se

$$E(X^2) = \sum_{x=1}^{6} x^2\, p(x) = \frac{91}{6}$$

Portanto, a variância pode ser calculada por

$$V(X) = E(X^2) - E^2(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

O desvio padrão e o coeficiente de variação são calculados a seguir

$$\sigma = \sqrt{\frac{35}{12}} = \sqrt{2,9167} = 1,7078 = 1,71 \quad \text{e} \quad \gamma = \frac{\sigma}{\mu} = \frac{1,71}{3,5} = 0,4889 = 0,49$$

Exemplo 12. Seja X a variável aleatória considerada no exemplo 5. Determinar a variância, o desvio padrão e o coeficiente de variação de X.

Solução

Nesse caso tem-se

$$E(X)=\int_0^1 x1dx = \frac{1}{2}x^2\Big|_0^1 = \frac{1}{2} = 0,5 \quad \text{como já foi calculado anteriormente}$$

$$E(X^2)=\int_0^1 x^2 1dx = \frac{1}{3}x^3\Big|_0^1 = \frac{1}{3} \quad \text{e então segue}$$

$$V(X)=E(X^2)-E^2(X)=\frac{1}{3}-\left(\frac{1}{2}\right)^2=\frac{1}{3}-\frac{1}{4}=\frac{4-3}{12}=\frac{1}{12}=0,0833=0,08$$

$$\text{e ainda} \quad \sigma=\sqrt{V(X)}=\sqrt{\frac{1}{12}}=\frac{1}{2\sqrt{3}}=\frac{\sqrt{3}}{6}=0,2887=0,29$$

Dos resultados acima decorre diretamente que o coeficiente de variação de X é igual a

$$\gamma=\frac{\sigma}{\mu}=\frac{0,29}{0,5}=0,58$$

Exemplo 13. Calcular a variância, o desvio padrão e o coeficiente de variação da variável aleatória considerada no exemplo 6.

$$E(X)=\int_0^1 x\,2\,x\,dx = 2\int_0^1 x^2\,dx = 2\left(\frac{1}{3}x^3\Big|_0^1\right) = \frac{2}{3} = 0,6667 = 0,67$$

$$E(X^2)=\int_0^1 x^2\,2\,x\,dx = 2\int_0^1 x^3dx = 2\left(\frac{1}{4}x^4\Big|_0^1\right) = \frac{1}{2}x^4\Big|_0^1 = \frac{1}{2} \quad \text{e então segue}$$

$$V(X)=E(X^2)-E^2(X)=\frac{1}{2}-\left(\frac{2}{3}\right)^2=\frac{1}{2}-\frac{4}{9}=\frac{9-8}{18}=\frac{1}{18}=0,0556=0,06$$

$$\text{bem como} \quad \sigma=\sqrt{V(X)}=\sqrt{\frac{1}{18}}=\frac{1}{3\sqrt{2}}=\frac{\sqrt{2}}{6}=0,2357=0,24$$

Consequentemente, segue

$$\gamma = \frac{\sigma}{\mu} = \frac{0,2357}{0,6667} = 0,3535 = 0,35$$

Comparando este valor com o do coeficiente de variação da variável aleatória do exemplo anterior – que é igual a 0,58 – verifica-se que a variabilidade desta é menor do que a daquela. Deve-se ainda registrar que tal comparação pode ser efetuada mesmo quando as naturezas das duas variáveis aleatórias comparadas são distintas pois o coeficiente de variação é uma medida adimensional da variabilidade. Além disso, pode-se também acrescentar que, na prática, verifica-se ser este coeficiente mais estável do que a variância ou o desvio padrão, razão pela qual é amplamente empregado.

3.10 Variáveis Aleatórias Padronizadas

3.10.1 Definição de Variável Aleatória Padronizada

Seja X uma variável aleatória com média $E(X) = \mu$ qualquer e desvio padrão positivo $\sigma = \sqrt{V(X)} = \sqrt{\sigma^2} > 0$. Então, define-se uma variável aleatória transformada de X, denominada variável aleatória X padronizada, expressa por $X^* = \dfrac{X - \mu}{\sigma}$. Destaca-se que a operação que define a variável padronizada X^* em função da variável (original) X é uma transformação linear.

A variável aleatória padronizada X^* possui a seguinte importante propriedade.

3.10.2 Propriedade

A variável aleatória padronizada X^* possui média zero e variância igual a 1, ou seja,

$$E\left(X^*\right) = 0 \quad \text{e} \quad V\left(X^*\right) = 1$$

Os valores de uma variável aleatória padronizada são frequentemente denominados valores ou escores padronizados. Isto é, os valores de X^* são entendidos como unidades padronizadas no sentido de que a unidade padrão, igual a σ, mede os valores das diferenças entre os valores da variável original, X, e a sua média.

As variáveis padronizadas são muito úteis para a comparação de diferentes distribuições.

3.11 Desigualdades de Markov, de Chebychev e de Bienaymé-Chebychev

3.11.1 Desigualdade de Markov

Teorema. Seja X uma variável aleatória que assume apenas valores não negativos e de média $E(X) = \mu$. Então, para qualquer valor $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a} \quad \text{ou} \quad P(X \geq a) \leq \frac{\mu}{a}$$

3.11.2 Desigualdade de Chebychev

Teorema. Seja X uma variável aleatória de média μ e variância σ^2. Tem-se, então, qualquer que seja $\varepsilon > 0$,

$$P\{|X - \mu| > \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

ou, equivalentemente,

$$P\{|X - \mu| \leq \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

3.11.3 Desigualdade de Chebychev (expressão alternativa)

A desigualdade de Chebychev possui uma expressão alternativa, que é mostrada a seguir.

Teorema. Nas mesmas condições do teorema anterior, fazendo $\varepsilon = k\sigma$, tem-se

$$P\{|X - \mu| > k\sigma\} \leq \frac{1}{k^2}$$

qualquer que seja $k > 0$.

Equivalentemente,

$$P\{|X - \mu| \leq k\sigma\} \geq 1 - \frac{1}{k^2}$$

qualquer que seja $k > 0$.

Nota: A desigualdade de Chebychev somente é útil se $\varepsilon^2 > \sigma^2$. Equivalentemente, na sua expressão alternativa, a desigualdade de Chebychev somente é útil se $k > 1$.

A desigualdade de Chebychev constitui uma interessante e importante relação entre a expectância (média) e a variância de uma variável aleatória. Essa desigualdade será muito útil posteriormente, tanto no Cálculo de Probabilidades quanto na Inferência Estatística.

Exemplo 14. O número de unidades de certo produto vendidas em um supermercado durante uma semana é uma variável aleatória com distribuição desconhecida; mas são disponíveis estimativas bastante precisas da sua média e da sua variância, cujos valores são considerados iguais a 800 e 100, respectivamente. Nessas condições, determinar um limite inferior para a probabilidade de que o número de unidades daquele produto vendidas em uma semana esteja compreendido entre 780 e 820.

Solução

Seja X = número de unidades do produto eventualmente vendidas no supermercado no período de uma semana. Sabe-se que $\mu = E(X) = 800$ e que $\sigma^2 = V(X) = 100$. Deseja-se determinar um limite inferior para a probabilidade de ocorrência do seguinte evento:

$$A = (780 \le X \le 820) = (|X - 800| \le 20)$$

Aplicando a desigualdade de Chebychev com $\varepsilon = 20$ segue

$$P(|X - 800| \le 20) \ge 1 - \frac{100}{20^2} \quad \text{portanto} \quad P(|X - 800| \le 20) \ge 0,75$$

Exemplo 15. O número de peças defeituosas fabricadas por certa máquina durante um turno de trabalho possui média igual a 100 e variância 81. Determinar um limite superior para a probabilidade de que em certo dia o número de peças defeituosas fabricadas seja maior que 118 ou menor que 82.

Solução

Seja X = número de peças defeituosas fabricadas em um turno de trabalho

Tem-se que $\mu = E(X) = 100$ e $\sigma^2 = 81$. Deseja-se determinar um limite superior para a probabilidade de ocorrência do evento $A = [(X > 118) \cup (X < 82)] = (|X - 100| > 18)$.

Aplicando a desigualdade de Chebychev com $\varepsilon = 18$ segue

$$P\left(\left|X-100\right| > 18\right) \le \frac{81}{18^2} \quad \text{portanto} \quad P\left(\left|X-100\right| > 18\right) \le 0,25$$

3.12 Momentos de uma Variável Aleatória X

As ideias subjacentes aos conceitos de expectância (média) e de variância de uma variável aleatória podem ser generalizadas. Isso conduz a outro importante conceito – o de momentos – ordinários e centrais – de uma variável aleatória.

3.12.1 Momentos Ordinários de X

Definição:

O momento ordinário de ordem k de X é definido por

$$\mu_k = E\left(X^k\right) \text{ para } k = 0,1,2,3,....$$

O cálculo dos momentos é indicado a seguir.

Cálculo:

(I) para X discreta:

$$\mu_k = E\left(X^k\right) = \sum_{x \in R_X} x^k \, p_X(x)$$

(II) para X contínua:

$$\mu_k = E\left(X^k\right) = \int_{-\infty}^{+\infty} x^k \, f_X(x)\,dx = \int_{x \in R_X} x^k \, f_X(x)\,dx$$

Em particular, tem-se

$$\begin{cases} \mu_0 = E\left(X^0\right) = E(1) = 1 \\ \mu_1 = E\left(X^1\right) = E(X) \ \text{(comumente denotado simplesmente por } \mu) \\ \mu_2 = E\left(X^2\right) \\ \mu_3 = E\left(X^3\right) \\ \mu_4 = E\left(X^4\right) \end{cases}$$

Notas:

- Se X é discreta, o momento ordinário de ordem k pode não existir, caso a expressão que o define seja uma série divergente. Nesse caso, diz-se que o momento é infinito.
- Se X é contínua, o momento ordinário de ordem k pode não existir, caso a expressão que o define seja uma integral divergente. Nesse caso, diz-se que o momento é infinito.
- O momento ordinário de ordem 0 (zero) sempre existe e é igual a 1, isto é,

$\mu_0 = E\left(X^0\right) = 1$

- Se existe, o momento ordinário de ordem k é um número real; diz-se então que o momento é finito.
- Prova-se que se existe o momento ordinário de ordem k, então existem todos os momentos ordinários de ordem inferior a k.
- Descritivamente, cada momento ordinário caracteriza um particular aspecto da distribuição da variável aleatória X, sendo que: (i) o primeiro relaciona-se à posição; (ii) o segundo relaciona-se à dispersão; (iii) o terceiro relaciona-se à simetria; e (iv) o quarto relaciona-se à curtose (grau de achatamento).
- O momento ordinário de ordem k de X é o valor para o qual tende a média aritmé. tica dos valores assumidos pela k-ésima potência dessa variável aleatória, quando o número de observações tende a infinito.
- É usual representar-se o valor numérico da expectância pela letra grega μ **.**

3.12.2 Momentos Centrais de X

Definição:

O momento central de ordem k de X é definido por

$$\nu_k = E\left[\left(X - E(X)\right)^k\right] = E\left[\left(X - \mu\right)^k\right] \quad \text{para } k = 0, 1, 2, 3, ...$$

Cálculo:

(I) para X discreta:

$$\nu_k = E\left[\left(X - E(X)\right)^k\right] = \sum_{x \in R_X} \left[x - E(X)\right]^k p_X(x)$$

(II) para X contínua:

$$\mu_k = E\left[\left(X - E(X)\right)^k\right] = \int_{-\infty}^{+\infty}\left[x - E(X)\right]^k f_X(x)\,dx = \int_{x \in R_X}\left[x - E(X)\right]^k f_X(x)\,dx$$

Em particular, tem-se

$$\begin{cases} \nu_0 = E\left[\left(X - E(X)\right)^0\right] = E(1) = 1 \\ \nu_1 = E\left[\left(X - E(X)\right)^1\right] = E(X) - E(X) = 0 \\ \nu_2 = E\left[\left(X - E(X)\right)^2\right] = V(X) \end{cases}$$

Nota: Quanto à existência**,** aplicam-se aos momentos centrais as mesmas observações anteriormente feitas para os momentos ordinários.

Exemplo 16. Determinar a expressão geral dos momentos ordinários da variável aleatória X do exemplo 4.

Solução

A função de probabilidade de X é

$$p(x) = \frac{x}{6}\ , \text{ para } x = 1, 2, 3$$

Logo, a expressão geral dos momentos ordinários de X é

$$\mu_k = E(X^k) = \sum_{x=1}^{3} x^k \frac{x}{6} = \frac{1}{6}\sum_{x=1}^{3} x^{k+1} = 1^{k+1}\frac{1}{6} + 2^{k+1}\frac{1}{6} + 3^{k+1}\frac{1}{6} = \frac{1}{6}(1 + 2^{k+1} + 3^{k+1}) \quad \text{para } k = 0, 1, 2, 3, \ldots$$

Em particular, tem-se:

$$\mu_0 = \frac{1}{6}\left(1 + 2 + 3\right) = 1$$

$$\mu_1 = \frac{1}{6}\left(1 + 2^2 + 3^2\right) = \frac{14}{6} = \frac{7}{3} = \mu = E\left(X\right)$$

$$\mu_2 = \frac{1}{6}\left(1 + 2^3 + 3^3\right) = \frac{36}{6} = 6$$

$$\mu_3 = \frac{1}{6}\left(1+2^4+3^4\right) = \frac{98}{6} = \frac{49}{3}$$

$$\mu_4 = \frac{1}{6}\left(1+2^5+3^5\right) = \frac{276}{6} = 46$$

Consequentemente, a variância (segundo momento central) é obtida por

$$\nu_2 = \mu_2 - \mu_1^2 = 6 - \left(\frac{7}{3}\right)^2 = 6 - \frac{49}{9} = \frac{54-49}{9} = \frac{5}{9}$$

Esse exemplo mostra que a determinação da expressão geral dos momentos de uma variável aleatória pode ser utilizada para simplificar os cálculos na determinação da expectância e da variância.

Exemplo 17. Determinar a expressão geral dos momentos ordinários da variável aleatória X do exemplo 5 e em seguida utilizá-la no cálculo da variância de X.

Solução

A função de densidade de probabilidade de X é

$$f(x) = 1 \text{ , para } 0 < x < 1$$

Então, a expressão geral dos momentos ordinários de X é

$$\mu_k = E(X^k) = \int_0^1 x^k \, 1 \, dx = \frac{1}{k+1} x^{k+1}\Big|_0^1 = \frac{1}{k+1} \qquad \text{para } k = 0,1,2,3,\ldots$$

Em particular, tem-se:

$$\mu_0 = \frac{1}{0+1} = 1$$

$$\mu_1 = \frac{1}{1+1} = \frac{1}{2} = \mu = E\left(X\right)$$

$$\mu_2 = \frac{1}{2+1} = \frac{1}{3}$$

$$\mu_3 = \frac{1}{3+1} = \frac{1}{4}$$

$$\mu_4 = \frac{1}{4+1} = \frac{1}{5}$$

Donde segue

$$V(X) = E(X^2) - E^2(X) = \mu_2 - \mu_1^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$$

Resultados que coincidem com aqueles já obtidos em exemplos anteriores.

Exemplo 18. Seja X uma variável aleatória contínua com função de densidade de probabilidade

$$f(x) = \frac{1}{2}\ , \text{ para } -1 < x < 1$$

Determinar a expressão geral dos momentos ordinários de X.

Solução

A expressão geral dos momentos ordinários de X é

$$\mu_k = E(X^k) = \int_{-1}^{1} x^k \frac{1}{2} dx = \frac{1}{2(k+1)} x^{k+1}\Big|_{-1}^{1} = \frac{1^{k+1} - (-1)^{k+1}}{2(k+1)} \qquad \text{para } k = 0,1,2,3,\ldots$$

portanto

$$\mu_k = E(X^k) = \begin{cases} \dfrac{1}{k+1} & , \quad \text{para } k = 0,2,4,\ldots \\ 0 & , \quad \text{para } k = 1,3,5,\ldots \end{cases}$$

Em particular, os primeiros três momentos são:

$$\mu_0 = \frac{1}{0+1} = 1$$

$$\mu_1 = 0 = \mu = E(X)$$

$$\mu_2 = \frac{1}{2+1} = \frac{1}{3}$$

$$\mu_3 = 0$$

$$\mu_4 = \frac{1}{4+1} = \frac{1}{5}$$

Donde segue

$$V(X) = E(X^2) - E^2(X) = \mu_2 - \mu_1^2 = \frac{1}{3} - 0^2 = \frac{1}{3}$$

3.12.3 Relações entre os Momentos Centrais e os Momentos Ordinários

O cálculo dos momentos centrais fica muito facilitado se for utilizada a seguinte relação existente entre o momento central de ordem k e os momentos ordinários de ordens não superiores a k:

$$\nu_k = E\left[(X - E(X))^k\right] = E\left[(X-\mu)^k\right] = E\left[\sum_{j=0}^{k} C_k^j (-1)^j \mu_1^j X^{k-j}\right] = \sum_{j=0}^{k} (-1)^j C_k^j \ \mu_1^j E\left(X^{k-j}\right)$$

ou seja,

(1) $\quad \nu_k = E\left[(X - E(X))^k\right] = \sum_{j=0}^{k} (-1)^j C_k^j \ \mu_1^j \ \mu_{k-j}$

Para os momentos centrais de ordens mais baixas, os coeficientes $C_k^j \ (j = 0,1,2,3,...,k)$ podem ser obtidos facilmente por recorrência, utilizando-se as seguintes propriedades dos coeficientes binomiais:

(2) $\quad C_k^0 = 1 \quad k = 0,1,2,3,....$

(3) $\quad C_k^j = C_{k-1}^{j-1} + C_{k-1}^j \quad k = 1,2,3,.... \quad$ e $\quad j = 1,2,3,...,k$

A partir do desenvolvimento de (1), para o momento central de ordem 2 tem-se:

$$\nu_2 = E\left[(X - E(X))^2\right] = \sum_{j=0}^{2} (-1)^j C_k^j \ \mu_1^j \ \mu_{k-j} = C_2^0 \ \mu_1^0 \mu_2 - C_2^1 \ \mu_1^1 \mu_1 + C_2^2 \ \mu_1^2 \mu_0$$

e observando que $\mu_0 = 1$, $\mu_1 = \mu = E(X)$ e $\mu_2 = E(X^2)$ bem como que $C_2^0 = C_2^2 = 1$ e $C_2^1 = 2$

então segue

(4) $\quad \nu_2 = E\left[(X - E(X))^2\right] = V(X) = \mu_2 - 2\mu_1^2 + \mu_1^2 = \mu_2 - \mu_1^2 = E(X^2) - E^2(X)$

Também são importantes as relações válidas entre os momentos centrais e os momentos ordinários para os casos do terceiro e do quarto momentos centrais, apresentadas a seguir, por meio de desenvolvimento similar:

$$\nu_3 = E\left[\left(X - E(X)\right)^3\right] = E\left[\left(X - \mu_1\right)^3\right] = \sum_{j=0}^{3} (-1)^j C_3^j \mu_1^j \mu_{3-j}$$

$$\nu_3 = \sum_{j=0}^{3} (-1)^j C_3^j \mu_1^j \mu_{3-j} = C_3^0 \mu_1^0 \mu_3 - C_3^1 \mu_1^1 \mu_2 + C_3^2 \mu_1^2 \mu_1 - C_3^3 \mu_1^3 \mu_0$$

$$\nu_3 = \mu_3 - 3\mu_1 \mu_2 + 3\mu_1^2 \mu_1 - \mu_1^3 = \mu_3 - 3\mu_1 \mu_2 + 3\mu_1^3 - \mu_1^3$$

(5) $$\nu_3 = \mu_3 - 3\mu_1 \mu_2 + 2\mu_1^3$$

$$\nu_4 = E\left[\left(X - E(X)\right)^4\right] = E\left[\left(X - \mu_1\right)^4\right] = \sum_{j=0}^{4} (-1)^j C_4^j \mu_1^j \mu_{4-j}$$

$$\nu_4 = \sum_{j=0}^{4} (-1)^j C_4^j \mu_1^j \mu_{4-j} = C_4^0 \mu_1^0 \mu_4 - C_4^1 \mu_1^1 \mu_3 + C_4^2 \mu_1^2 \mu_2 - C_4^3 \mu_1^3 \mu_1 + C_4^4 \mu_1^4 \mu_0$$

$$\nu_4 = \mu_4 - 4\mu_1 \mu_3 + 6\mu_1^2 \mu_2 - 4\mu_1^4 + \mu_1^4$$

(6) $$\nu_4 = \mu_4 - 4\mu_1 \mu_3 + 6\mu_1^2 \mu_2 - 3\mu_1^4$$

Finalmente, com relação aos momentos ordinários e centrais, convém destacar que:

(i) o primeiro momento ordinário é a média ou expectância da distribuição, consistindo assim em uma medida de posição que indica o “centro” da mesma;

(ii) o segundo momento central é a variância portanto relaciona-se com a variabilidade ou grau de dispersão dos valores da distribuição;

(iii) o terceiro momento central relaciona-se à simetria da distribuição;

(iv) o quarto momento central relaciona-se à curtose (achatamento) da distribuição.

As medidas relativas à assimetria e à curtose serão apresentadas ao final da próxima seção.

3.13 Outras Medidas Características das Distribuições

3.13.1 Outras Medidas de Posição Central

Como foi visto, a média ou expectância de uma variável aleatória unidimensional consiste em uma medida de posição que caracteriza o centro da distribuição de probabilidade da variável. Ela é também conhecida como uma medida de tendência central da distribuição e, com efeito, é considerada a principal medida dessa natureza, sendo a mais utilizada na prática.

Porém, há duas outras medidas de tendência central que também são empregadas na prática. Tais medidas são a moda e a mediana.

1. Moda. A moda de uma variável a aleatória é o valor que é mais provável de ocorrer ou que é o máximo da distribuição de probabilidade da variável. Algumas vezes a distribuição possui mais de um valor mais provável e, neste caso, diz-se que a distribuição é bimodal (se forem só dois valores), trimodal (se existirem três desses valores) ou multimodal. Enfim, a moda de uma variável aleatória X é o valor (ou valores) onde a sua distribuição (função de probabilidade ou função de densidade de probabilidade) tem um máximo local.
2. Mediana. A mediana de uma variável aleatória X é o valor x para o qual se tem a seguinte condição: $P(X \leq x) \leq \frac{1}{2}$ e $P(X > x) > \frac{1}{2}$. Isto é, a mediana consiste no valor que separa a curva da função densidade em duas partes que têm áreas iguais, valendo 0,5 cada uma delas. No caso de uma distribuição discreta, pode não existir uma mediana única.

Além das medidas de posição central, tem-se algumas famílias de medidas de posição que são também muito úteis na descrição de aspectos particulares da distribuição de uma variável. A ideia subjacente à definição dessas medidas consiste em uma generalização da ideia relativa à mediana e equivale a um conjunto de pontos (valores da variável X) que dividem a distribuição de X de forma se tenha frações de área iguais: (i) quartis – que dividem a referida distribuição em quatro partes de áreas iguais; (ii) decis – que dividem a distribuição em dez partes de áreas iguais; e (iii) centis – que dividem a distribuição em cem partes de áreas iguais. Há três quartis, denotados por: $x_{0,25}$, $x_{0,50}$, $x_{0,75}$; há nove decis, denotados por: $x_{0,1}$, $x_{0,2}$, $x_{0,3}$, ... , $x_{0,5}$, ... , $x_{0,9}$; e, equivalentemente, há 99 centis ou percentis, que são indicados por: $x_{0,01}$, $x_{0,02}$, $x_{0,03}$, ... , $x_{0,10}$, ... , $x_{0,50}$, ..., $x_{0,90}$, ... , $x_{0,99}$

3.13.2 Outras Medidas de Dispersão

Assim como existem outras medidas de tendência central (e de posição, de modo mais geral), há também outras medidas de dispersão das distribuições de probabilidade das variáveis aleatórias. Duas das mais comumente empregadas são: (i) o intervalo iterquartílico e (ii) o intervalo semi-interquartílico.

1. O intervalo interquartílico representa a diferença entre o terceiro e o primeiro quartis, isto é $x_{0,75} - x_{0,25}$ e o interval semi-interquartílico é $\frac{1}{2}\left(x_{0,75} - x_{0,25}\right)$.
2. O desvio médio de uma variável aleatória X é definido como a expectância do valor absoluto da diferença entre a variável e sua média (expectância), ou seja, a sua expressão é $E\left(\left|X - E(X)\right|\right) = E\left(\left|X - \mu\right|\right)$. Portanto, o cálculo do desvio médio é calculado por

   i) $DM(X) = E\left(\left|X - E(X)\right|\right) = E\left(\left|X - \mu\right|\right) = \sum_{x \in R_X} |x - \mu| p(x)$, se X é discreta

   ii) $DM(X) = E\left(\left|X - E(X)\right|\right) = E\left(\left|X - \mu\right|\right) = \int_{R_X} |x - \mu| f(x)\, dx$, se X é contínua

3.13.3 Medidas de Assimetria e Curtose

Além dessas medidas acima apresentadas, há duas outras importantes medidas características que se relacionam a dois aspectos considerados básicos na descrição de uma distribuição: a assimeria e a curtose ou achatamento.

1. **Assimetria**. Em geral, a distribuição de probabilidade de uma variável aleatória X não é simétrica mas, pelo contrário, se alonga mais para a esquerda ou para a direita do centro da distribuição. Desse modo, diz-se que a distribuição é assimétrica à esquerda ou assimétrica à direita, respectivamente. A medida mais usada para indicar a assimetria é o denominado coeficiente de assimetria, definido por

$$\alpha_3 = \frac{E\left[\left(X - E(X)\right)^3\right]}{\sigma^3} = \frac{E\left[\left(X - \mu\right)^3\right]}{\sigma^3} = \frac{\nu_3}{\sigma^3}$$

   Se o coeficiente de assimetria for zero, a distribuição é simétrica. Se for negativo, a distribuição é assimétrica à esquerda, e, se for positive, a distribuição é assimétrica à direita.

2. **Curtose**. Esse aspecto de uma distribuição se relaciona ao modo como os valores se concentram na região central da distribuição de probabilidade da variável aleatória. Em alguns casos, eles se apresentam com um grau de concentração maior em torno de um ponto central, fazendo com que o gráfico da distribuição tenha uma espécie de pico nesse ponto, enquanto em outros casos os valores da distribuição se encontram mais dispersos em torno desse valor central, fazendo com que o gráfico da distribuição tenha uma espécie de platô ou seja, uma espécie de achatamento em torno desse valor. No primeiro caso, diz-se que a distribuição é leptocúrtica e no segundo caso

diz-se que ela é platicúrtica. A medida mais empregada para indicar a curtose é o denominado coeficiente de curtose, definido por

$$\alpha_4 = \frac{E\left[\left(X - E(X)\right)^4\right]}{\sigma^4} = \frac{E\left[\left(X - \mu\right)^4\right]}{\sigma^4} = \frac{\nu_4}{\sigma^4}$$

Se o coeficiente de curtose for igual a 3, diz-se que a distribuição é mesocúrtica. Se for maior do que 3, ela é classificada como leptocúrtica, e, se for menor do que 3, ela é classificada como platicúrtica.

Exemplo 19. Calcular os coeficientes de assimetria e de curtose para as variáveis aleatórias dos exemplos 16, 17 e 18.

Solução

a) No exemplo 16, a variável aleatória X tem a distribuição dada pela função de probabilidade

$$p(x) = \frac{x}{6} \text{ , para } x = 1, 2, 3$$

Foi visto que a expressão geral dos momentos ordinários é

$$\mu_k = E(X^k) = \frac{1}{6}(1 + 2^{k+1} + 3^{k+1}) \quad \text{para } k = 0, 1, 2, 3, \ldots$$

Já foram calculados os seguintes momentos:

$$\mu_1 = \frac{1}{6}\left(1 + 2^2 + 3^2\right) = \frac{14}{6} = \frac{7}{3} = \mu = E(X)$$

$$\mu_2 = \frac{1}{6}\left(1 + 2^3 + 3^3\right) = \frac{36}{6} = 6$$

$$\mu_3 = \frac{1}{6}\left(1 + 2^4 + 3^4\right) = \frac{98}{6} = \frac{49}{3}$$

$$\mu_4 = \frac{1}{6}\left(1 + 2^5 + 3^5\right) = \frac{276}{6} = 46$$

Consequentemente, a variância (segundo momento central) é obtida por

$$\nu_2 = \mu_2 - \mu_1^2 = 6 - \left(\frac{7}{3}\right)^2 = 6 - \frac{49}{9} = \frac{54-49}{9} = \frac{5}{9}$$

e o desvio padrão é $\sigma = \sqrt{\frac{5}{9}} = \frac{\sqrt{5}}{3}$

Os momentos centrais de ordens 3 e 4 são

$$\nu_3 = \mu_3 - 3\mu_1\,\mu_2 + 2\mu_1^3 = \frac{49}{3} - 3\cdot\frac{7}{3}\cdot 6 + 2\left(\frac{7}{3}\right)^3 = \frac{49}{3} - 42 + \frac{2\cdot 343}{27} = \frac{441 - 42\cdot 27 + 686}{27}$$

$$\nu_3 = \frac{1127-1134}{27} = -\frac{7}{27}$$

$$\nu_4 = \mu_4 - 4\mu_1\,\mu_3 + 6\mu_1^2\,\mu_2 - 3\mu_1^4 = 46 - 4\cdot\frac{7}{3}\cdot\frac{49}{3} + 6\cdot\left(\frac{7}{3}\right)^2\cdot 6 - 3\cdot\left(\frac{7}{3}\right)^4 = 46 - \frac{1372}{9} + \frac{1764}{9} - \frac{343}{9}$$

$$\nu_4 = 46 + \frac{-1372+1764}{9} - \frac{2401}{27} = 46 + \frac{3\cdot 392 - 2401}{27} = 46 + \frac{1176-2401}{27} = 46 + \frac{-1225}{27}$$

$$\nu_4 = \frac{46\cdot 27 - 1225}{27} = \frac{1242-1225}{27} = \frac{17}{27}$$

Portanto, os coeficientes de assimetria e curtose são

$$\alpha_3 = \frac{\nu_3}{\sigma^3} = \frac{-\frac{7}{27}}{\left(\frac{\sqrt{5}}{3}\right)^3} = \frac{-\frac{7}{27}}{\frac{5\sqrt{5}}{27}} = -\frac{7\sqrt{5}}{25} = -\frac{7\cdot 2,2361}{25} = -\frac{15,6527}{25} = -0,6261$$

e

$$\alpha_4 = \frac{\nu_4}{\sigma^4} = \frac{\frac{17}{27}}{\left(\frac{\sqrt{5}}{3}\right)^4} = \frac{\frac{17}{27}}{\frac{25}{81}} = \frac{51}{25} = 2,04$$

Então, a distribuição é assimétrica à esquerda e platicúrtica

b) No exemplo 17, a variável aleatória X tem a distribuição dada pela função de densidade de probabilidade

$$f(x)=1 \text{ , para } 0<x<1$$

Foi visto que a expressão geral dos momentos é

$$\mu_k = E(X^k) = \int_0^1 x^k \, 1 dx = \frac{1}{k+1} \qquad \text{para } k=0,1,2,3,...$$

Já foram calculados os seguintes momentos:

$$\mu_0 = \frac{1}{0+1} = 1$$

$$\mu_1 = \frac{1}{1+1} = \frac{1}{2} = \mu = E(X)$$

$$\mu_2 = \frac{1}{2+1} = \frac{1}{3}$$

$$\mu_3 = \frac{1}{3+1} = \frac{1}{4}$$

$$\mu_4 = \frac{1}{4+1} = \frac{1}{5}$$

Consequentemente, a variância (segundo momento central) é obtida por

$$\nu_2 = \mu_2 - \mu_1^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{4-3}{12} = \frac{1}{12}$$

e o desvio padrão é $\sigma = \sqrt{\frac{1}{12}} = \frac{1}{2\sqrt{3}}$

Os momentos centrais de ordens 3 e 4 são calculados a seguir

$$\nu_3 = \mu_3 - 3\mu_1 \mu_2 + 2\mu_1^3 = \frac{1}{4} - 3 \cdot \frac{1}{2} \cdot \frac{1}{3} + 2 \cdot \frac{1}{2^3} = \frac{1}{4} - \frac{1}{2} + \frac{1}{4} = 0$$

$$\nu_4 = \mu_4 - 4\mu_1 \mu_3 + 6\mu_1^2 \mu_2 - 3\mu_1^4 = \frac{1}{5} - 4 \cdot \frac{1}{2} \cdot \frac{1}{4} + 6 \cdot \left(\frac{1}{2}\right)^2 \cdot \frac{1}{3} - 3 \cdot \left(\frac{1}{2}\right)^4 = \frac{1}{5} - \frac{1}{2} + \frac{1}{2} - \frac{3}{16}$$

$$\nu_4 = \frac{1}{5} - \frac{3}{16} = \frac{16-15}{80} = \frac{1}{80} = 0,0125$$

Portanto, os coeficientes de assimetria e curtose são

$$\alpha_3 = \frac{\nu_3}{\sigma^3} = \frac{0}{\left(\frac{1}{2\sqrt{3}}\right)^3} = 0$$

e

$$\alpha_4 = \frac{\nu_4}{\sigma^4} \frac{\frac{1}{80}}{\left(\frac{1}{2\sqrt{3}}\right)^4} = \frac{\frac{1}{80}}{\frac{1}{144}} = \frac{144}{80} = \frac{9}{5} = 1,8$$

Logo, a distribuição é simétrica e platicúrtica.

c) No exemplo 18, a variável aleatória X tem a distribuição dada pela função de densidade de probabilidade

$$f(x) = \frac{1}{2} \text{ , para } -1 < x < 1$$

Foi visto que a expressão geral dos momentos é

$$\mu_k = E(X^k) = \int_{-1}^{1} x^k \frac{1}{2} dx = \frac{1}{2(k+1)} x^{k+1}\Big|_{-1}^{1} = \frac{1^{k+1} - (-1)^{k+1}}{2(k+1)} \qquad \text{para } k = 0, 1, 2, 3, \ldots$$

portanto

$$\mu_k = E(X^k) = \begin{cases} \dfrac{1}{k+1} & , \quad \text{para } \; k = 0, 2, 4, \ldots \\ 0 & , \quad \text{para } \; k = 1, 3, 5, \ldots \end{cases}$$

Já foram calculados os seguintes momentos:

$$\mu_0 = \frac{1}{0+1} = 1$$

$$\mu_1 = 0 = \mu = E(X)$$

$$\mu_2 = \frac{1}{2+1} = \frac{1}{3}$$

$$\mu_3 = 0$$

$$\mu_4 = \frac{1}{4+1} = \frac{1}{5}$$

Logo, a variância, segundo momento central em relação à média, é

$$V(X) = \nu_2 = \mu_2 - \mu_1^2 = \frac{1}{3} - 0^2 = \frac{1}{3}$$

e o desvio padrão é $\sqrt{\frac{1}{3}} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}$

Os momentos centrais de ordens 3 e 4 são

$$\nu_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 = 0 - 3\cdot 0 \cdot \frac{1}{3} + 2\cdot 0^3 = 0$$

$$\nu_4 = \mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4 = \mu_4 = \frac{1}{5}$$

Portanto, os coeficientes de assimetria e curtose são

$$\alpha_3 = \frac{\nu_3}{\sigma^3} = 0$$

e

$$\alpha_4 = \frac{\nu_4}{\sigma^4} = \frac{\frac{1}{5}}{\left(\frac{1}{3}\right)^2} = \frac{\frac{1}{5}}{\frac{1}{9}} = \frac{9}{5} = 1{,}8$$

Logo, a distribuição é simétrica e platicúrtica.

Além dos tópicos já estudados até aqui, o conceito de valor esperado também possui importantes aplicações em problemas que envolvem funções de variáveis aleatórias. Muitos deles são relacionados a questões no campo econômico. Isso será ilustrado por meio de alguns exemplos apresentados a seguir.

3.14 Aplicações das Propriedades de Expectância e Variância

Exemplo 20. Certa indústria produz determinado tipo de tubo empregado na fabricação de um equipamento. Por contrato, esse tubo deve atender a especificações técnicas, em particular quanto ao seu comprimento, que deve ser igual a 30cm, com tolerância de 0,3 cm, para mais ou para menos. Assim, cada tubo produzido é em seguida examinado pelo setor de controle de qualidade para verificar se atende àquela especificação. O custo de produção de um tubo é R$ 10,00. Se o tubo for produzido dentro das especificações ele é vendido por R$ 20,00. Por outro lado, se o tubo tiver comprimento maior que 30,3 cm ele poderá ter seu comprimento ajustado à especificação, ao custo adicional de R$ 5,00; porém, se o seu comprimento for inferior a 29,7 cm ele deverá ser refugado, com prejuízo total. Admita-se que o comprimento, expresso em centímetros, de um tubo fabricado na indústria considerada é uma variável aleatória com função de densidade de probabilidade

$$f(x) = 1 \text{ , para } 29{,}5 < x < 30{,}5$$

Nessas condições, determine o lucro médio (por tubo) do fabricante.

Solução

Seja Y a variável aleatória que expressa o lucro eventual de um tubo produzido. Então, a expressão do lucro em função de X é

$$L = \begin{cases} 0 - 10{,}00 = -10{,}00 & \text{se } X < 29{,}7 \\ 20{,}00 - 10{,}00 = 10{,}00 & \text{se } 29{,}7 \leq X \leq 30{,}3 \\ 20{,}00 - 15{,}00 = 5{,}00 & \text{se } X > 30{,}3 \end{cases}$$

Além disso, tem-se

$$P(X < 29{,}7) = \int_{29{,}5}^{29{,}7} 1\, dx = x\Big|_{29{,}5}^{29{,}7} = 29{,}7 - 29{,}5 = 0{,}2$$

$$P(29{,}7 \leq X \leq 30{,}3) = \int_{29{,}7}^{30{,}3} 1\, dx = x\Big|_{29{,}7}^{30{,}3} = 30{,}3 - 29{,}7 = 0{,}6$$

$$P(X > 30{,}3) = \int_{30{,}3}^{30{,}5} 1\, dx = x\Big|_{30{,}3}^{30{,}5} = 30{,}5 - 30{,}3 = 0{,}2$$

Portanto, o lucro médio por unidade produzida é

$$E(L) = -10{,}00 . 0{,}2 + 10{,}00 . 0{,}6 + 5{,}00 . 0{,}2 = -2{,}00 + 6{,}00 + 1{,}00 = 5{,}00$$

Ou seja, o lucro médio da indústria é R$ 5,00 por tubo.

Exemplo 21. A demanda diária de certo produto alimentício rapidamente perecível em uma lanchonete é uma variável aleatória com a seguinte função de probabilidade:

$$p(x)=\frac{1+x}{10} \text{ , para } x=0,1,2,3$$

Todos os dias são adquiridas e colocadas à venda 3 unidades desse produto; aquelas que não forem vendidas até o final do dia são perdidas. Sendo o custo de aquisição de cada unidade do referido produto igual a R$ 5,00 , determinar o valor a ser fixado para o seu preço de venda unitário, de modo que a lanchonete obtenha um lucro diário esperado igual a R$ 10,00.

Solução

O custo diário de aquisição das 3 unidades do produto é R$ 15,00. Denotando por **a** o valor do preço unitário de venda, a ser determinado, e sendo X a variável aleatória que representa o número eventual de unidades do produto vendidas em um dia, então o lucro diário, Y, é expresso por $Y = a\,X - 15$.

Portanto, o valor esperado do lucro diário é

$$E(Y) = E(a\,X-15) = a\,E(X)-15$$

O valor esperado do número de unidades vendidas é

$$E(X) = \sum_{x=0}^{3} x\,\frac{x+1}{10} = 0\frac{1}{10}+1\frac{2}{10}+2\frac{3}{10}+3\frac{4}{10} = \frac{0+2+6+12}{10} = \frac{20}{10} = 2$$

Logo

$$E(Y) = a\,2-15 = 10, \text{donde } 2a = 25 \text{ e assim } a=\frac{25}{2}=12,5 \,.$$

Ou seja, o preço unitário a ser estabelecido para a venda é: a = R$ 12,5

O problema anterior foi resolvido nas condições em que o produto é adquirido de modo a atender a demanda máxima, sem a preocupação de otimizar a decisão quanto à quantidade comprada, de modo a evitar prejuízo. Isso tem reflexos na determinação do preço de venda mas exige liberdade na fixação deste, o que nem sempre é possível. Esse tipo de problema, quando há preocupação em minimizar o prejuízo decorrente da perda de unidades não vendidas requer uma abordagem diferente, como é mostrado no próximo exemplo.

Exemplo 22. Um comerciante adquire, semanalmente, no início da semana, unidades de um produto perecível para comercialização ao longo da semana. Cada unidade desse produto é adquirida ao valor de 1,50 unidades monetária (u.m.) e é vendida ao preço de 2,00 u.m. Todas as unidades não vendidas durante cada semana são descartadas ao final da mesma, com prejuízo integral. Seja X a variável aleatória que representa a demanda eventual do produto durante uma semana. A distribuição de probabilidade de X é apresentada na tabela a seguir.

x	p(x)
0	0,1
1	0,1
2	0,3
3	0,4
4	0,1
Total	1,0

Nessas condições, determinar o número de unidades do produto a serem encomendadas semanalmente pelo comerciante de modo a maximizar o lucro semanal esperado.

Solução

Seja X a variável aleatória que representa a demanda eventual do produto durante uma semana. É fácil determinar o lucro com cada unidade que é vendida e o prejuízo de cada uma não vendida. Nesse sentido, verifica-se que cada unidade vendida fornece um lucro igual a

$$2{,}00 - 1{,}50 = 0{,}50 \text{ u.m.}$$

e cada unidade não vendida acarreta um prejuízo de 1,50 u.m. Assim, o lucro é função da demanda, que é uma variável aleatória, e também do número **r** de unidades encomendadas, que é uma constante, determinada pelo comerciante.

Para determinar o valor ótimo da encomenda semanal pode-se recorrer a uma pesquisa do tipo exaustiva, isto é, examinando cada uma das possibilidades.

Denotando por Y o lucro por unidade e por **r** a quantidade encomendada do produto, tem-se a análise mostrada a seguir, para r = 1, 2, 3, 4 .

		r = 1		r = 2		r = 3		r = 4	
x	p(x)	y	y p(x)	y	y p(x)	y	y p(x)	y	y p(x)
0	0,1	- 1,50	- 0,15	- 3,00	- 0,30	- 4,50	- 0,45	- 6,00	- 0,60
1	0,1	0,50	0,05	- 1,00	- 0,10	- 2,50	- 0,25	- 4,00	- 0,40
2	0,3	0,50	0,15	1,00	0,30	- 0,50	- 0,15	- 2,00	- 0,60
3	0,4	0,50	0,20	1,00	0,40	1,50	0,60	0,00	0,00
4	0,1	0,50	0,05	1,00	0,10	1,50	0,15	2,00	0,20
total	1,0		0,30		0,40		- 0,10		- 1,40

Portanto, verifica-se que o lucro unitário esperado é máximo quando r = 2.

Embora a pesquisa exaustiva de todas as possibilidades de valores de **r** conduza a um raciocínio direto para a solução do problema, ela não é a melhor forma de resolver esse tipo de problema. Com efeito, é fácil perceber que se a demanda fosse uma variável aleatória que admitisse muitos valores, seria muito trabalhosa uma solução desenvolvida por meio desse raciocínio direto.

Prova-se que esse tipo de problema pode ser resolvido pelo exame da função de distribuição da demanda, procurando o maior valor x* de X tal que $F(x^*) \le \frac{a}{a+b}$ onde a é o lucro líquido unitário e b a perda líquida unitária e tomando-se como máximo o valor r = x* + 1. Nesse caso, tem-se

$$\frac{a}{a+b} = \frac{0,50}{0,50+1,50} = \frac{0,50}{2,00} = 0,25 \text{ ou seja } F(x) \le 0,25$$

Os valores da função de distribuição acumulada nos pontos em que p(x) é definida são mostrados na tabela a seguir

x	p(x)	F(x)
0	0,1	0,1
1	0,1	0,2
2	0,3	0,5
3	0,4	0,9
4	0,1	1,0
Total	1,0	

Verifica-se facilmente que x* = 1, pois F(1) = 0,20 e F(2) = 0,50 portanto o ponto de máximo é r = x* + 1 = 1+1 = 2 (como também foi obtido com o uso do trabalhoso método de pesquisa exaustiva anteriormente empregado).

O conceito de expectância pode ser estendido ao caso de variáveis aleatórias bidimensionais, aonde também possui grande utilidade. Isso será visto a seguir.

3.15 Expectância de Função e Momentos de Variáveis Aleatórias Bidimensionais

3.15.1 Definição de Expectância de Função de Variável Aleatória Bidimensional

Sejam (X,Y) uma variável aleatória bidimensional e h(X,Y) uma função real de duas variá-veis. Define-se a expectância de h(X,Y) por

(i) $$E\left[h(X,Y)\right]=\sum_{(x,y)\in R^2_{XY}}\sum h(x,y)\ p_{XY}(x,y),\ \text{para (X,Y) discreta}$$

(ii) $$E\left[h(X,Y)\right]=\iint_{R^2_{XY}} h(x,y)\ f_{XY}(x,y)\ dx\,dy,\ \text{para (X,Y) contínua}$$

Notas:

- A expectância pode não existir, caso a expressão que a define seja uma série ou uma integral divergente; nesse caso diz-se que a expectância é infinita.

- Se existe, a expectância é um número real; diz-se então que a expectância é finita.

Analogamente ao que ocorreu antes com as aplicações da expectância de funções de variáveis aleatórias unidimensionais, constituídas pelos momentos, também são muito importantes as aplicações da expectância de funções de variáveis aleatórias bidimensionais que originam os momentos – ordinários e centrais – dessas variáveis, que são apresentados a seguir

3.15.2 Momentos Conjuntos ou Mistos de Variáveis Aleatórias Bidimensionais

a) Momentos Conjuntos Ordinários

Definição

Seja (X, Y) uma variável aleatória bidimensional. Define-se o momento conjunto (ou misto) ordinário de ordens i e j dessa variável por

$$\mu_{i,j}=E\left(X^i\ Y^j\right)\quad \text{para } i,j=0,1,2,3,\ldots$$

Notas:

- $\mu_{i,0}=E\left(X^i\ Y^0\right)=E\left(X^i\right)$ para $i=0,1,2,3,\ldots$

- $\mu_{0,j}=E\left(X^0\ Y^j\right)=E\left(Y^j\right)$ para $j=0,1,2,3,\ldots$

- $\mu_{1,1} = E\left(X^1\ Y^1\right) = E\left(X\,Y\right)$

- $\mu_{i,j} = E\left(X^i\ Y^j\right)$ para $i, j = 0, 1, 2, 3, ...$

Cálculo

i) Se (X, Y) é discreta

$$\mu_{i,j} = E\left(X^i\ Y^j\right) = \sum_{(x,y)\in R^2_{XY}} \sum x^i\, y^j\, p_{XY}(x,y) \qquad \text{para } i, j = 0, 1, 2, 3, ...$$

ii) Se (X, Y) é contínua

$$\mu_{i,j} = E\left(X^i\ Y^j\right) = \int_{R^2_{XY}} \int x^i y^j\ f_{XY}(x,y)\ dx\, dy \qquad \text{para } i, j = 0, 1, 2, 3, ...$$

b) Momentos Conjuntos Centrais

Definição

Seja (X, Y) uma variável aleatória bidimensional. Define-se o momento conjunto (ou misto) central de ordens i e j dessa variável por

$$\nu_{i,j} = E\left\{\left[X - E(X)\right]^i \left[Y - E(Y)\right]^j\right\} \qquad \text{para } i, j = 0, 1, 2, 3, ...$$

Nota

- O momento central de ordens 1 e 1, $\nu_{1,1}$, é denominado covariância entre X e Y

$$\nu_{1,1} = E\left\{\left[X - E\left(X\right)\right]\left[Y - E\left(Y\right)\right]\right\} = E\left(X\,Y\right) - E\left(X\right)E\left(Y\right)$$

Cálculo

i) Se (X, Y) é discreta

$$\nu_{i,j} = E\left\{\left[X - E\left(X\right)\right]^i \left[Y - E\left(Y\right)\right]^j\right\} = \sum_{(x,y)\in R^2_{XY}} \sum \left[x - E\left(X\right)\right]^i \left[y - E\left(Y\right)\right]^j p_{XY}(x,y) \text{ para } i, j = 0, 1, 2, 3, ...$$

ii) Se (X, Y) é contínua

$$\nu_{i,j} = E\left\{\left[X - E\left(X\right)\right]^i \left[Y - E\left(Y\right)\right]^j\right\} = \int_{R^2_{XY}} \int \left[x - E(X)\right]^i \left[y - E(Y)\right]^j\ f_{XY}(x,y)\ dx\, dy \text{ para } i, j = 0, 1, 2, 3, ...$$

3.16 Covariância entre X e Y para Variáveis Aleatórias Bidimensionais

3.16.1 Definição

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

3.16.2 Propriedades da Covariância

a) $\mathrm{Cov}(X,Y)$ está expressa no produto das unidades de medida de X e de Y

b) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

c) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

d) $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$

e) Se X e Y são independentes, então $\mathrm{Cov}(X,Y) = 0$

Nota: A recíproca não é verdadeira: existem variáveis aleatórias cuja covariância é zero e que não são independentes.

f) Se $U = aX + b$ e $V = cY + d \Rightarrow \mathrm{Cov}(U,V) = a\,c\,\mathrm{Cov}(X,Y)$

g) $\mathrm{Cov}(X+Y, Z+W) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(X,W) + \mathrm{Cov}(Y,Z) + \mathrm{Cov}(Y,W)$

3.16.3 Variância da Soma e da Diferença de duas Variáveis Aleatórias

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

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

Um outro importante conceito envolvendo variáveis aleatórias bidimensionais e que se relaciona à covariância é o de coeficiente de correlação linear, o qual será mostrado a seguir.

3.17 Coeficiente de Correlação Linear (ou Coeficiente de Correlação de Pearson)

3.17.1 Definição

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

3.17.2 Propriedades do Coeficiente de Correlação Linear

a) ρ_{XY} é adimensional

b) $-1 \leq \rho_{XY} \leq 1$

c) Se X e Y são independentes $\rho_{XY} = 0$

Notas:

- A recíproca não é verdadeira: existem variáveis aleatórias cujo coeficiente de correlação linear é igual a zero e que não são independentes.

- Se $\rho_{XY} = 0$, as variáveis aleatórias X e Y são ditas "não correlacionadas"

d) $U = aX + b$ e $V = cY + d \;\Rightarrow\; \rho_{UV} = \begin{cases} \rho_{XY} & \text{se } ac > 0 \\ -\rho_{XY} & \text{se } ac < 0 \end{cases}$

e) $\rho_{XY} = -1$ se e somente se existem números reais $a < 0$ e b tais que

$$P\{Y = aX + b\} = 1$$

f) $\rho_{XY} = +1$ se e somente se existem números reais $a > 0$ e b tais que

$$P\{Y = aX + b\} = 1$$

3.17.3 Coeficiente de Determinação

ρ_{XY}^2 denomina-se Coeficiente de Determinação e mede o grau de dispersão das determinações de (X,Y) em torno de uma linha reta.

Exemplo 23. Sejam X e Y variáveis aleatórias com a seguinte distribuição conjunta de probabilidades:

X \ Y	2	5
1	1/8	2/8
3	3/8	2/8

Pede-se:

a) A covariância entre X e Y;
b) o coeficiente de correlação linear entre X e Y;
c) verificar se X e Y são independentes.
Solução

A partir da tabela que apresenta a função de probabilidade conjunta de X e de Y podem ser determinadas as distribuições marginais, como é mostrado a seguir.

X	Y		$p_X(x)$
	2	5	
1	1/8	2/8	3/8
3	3/8	2/8	5/8
$p_Y(y)$	4/8	4/8	1

Então:

a) Cálculo da covariância de X e Y

$$p_X(x) = \begin{cases} 3/8 & x = 1 \\ 5/8 & x = 3 \end{cases} \qquad p_Y(y) = \begin{cases} 1/2 & y = 2 \\ 1/2 & y = 5 \end{cases}$$

$$E(XY) = (1)(2)\left(\frac{1}{8}\right) + (1)(5)\left(\frac{2}{8}\right) + (3)(2)\left(\frac{3}{8}\right) + (3)(5)\left(\frac{2}{8}\right)$$

$$= \frac{2}{8} + \frac{10}{8} + \frac{18}{8} + \frac{30}{8} = \frac{15}{2}$$

$$E(X) = (1)\left(\frac{3}{8}\right) + (3)\left(\frac{5}{8}\right) = \frac{9}{4} \qquad E(Y) = (2)\left(\frac{1}{2}\right) + (5)\left(\frac{1}{2}\right) = \frac{7}{2}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \left(\frac{15}{2}\right) - \left(\frac{9}{4}\right)\left(\frac{7}{2}\right) = -\frac{3}{8}$$

b) Cálculo do coeficiente de correlação linear de X e Y

$$E(X^2) = (1^2)\left(\frac{3}{8}\right) + (3^2)\left(\frac{5}{8}\right) = \frac{48}{8} = 6$$

$$E(Y^2) = (2^2)\left(\frac{1}{2}\right) + (5^2)\left(\frac{1}{2}\right) = \frac{29}{2}$$

$$Var(X) = E(X^2) - E^2(X) = 6 - \left(\frac{9}{4}\right)^2 = \frac{15}{16}$$

$$Var(Y) = E(Y^2) - E^2(Y) = \frac{29}{2} - \left(\frac{7}{2}\right)^2 = \frac{9}{4}$$

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{-\left(\frac{3}{8}\right)}{\sqrt{\frac{15}{16}\frac{9}{4}}} = -\frac{1}{\sqrt{15}} = -0,2582$$

c) $\rho_{XY} \neq 0 \Rightarrow$ X e Y não são independentes.

Exemplo 24. Seja (X,Y) uma variável aleatória bidimensional do tipo contínuo com a seguinte densidade de probabilidade conjunta:

$$f_{XY}(x,y) = 3x \quad \text{para } 0 < x < 1 \text{ e } 0 < y < x$$

Pede-se:

a) Determinar a covariância entre X e Y
b) Verificar se X e Y são independentes.

Solução

A função de densidade de probabilidade conjunta de X e de Y é

$$f_{XY}(x,y) = 3x \quad \text{para } 0 < x < 1 \text{ e } 0 < y < x$$

a) $$f_X(x) = \int_0^x 3x\,dy = 3x\int_0^x dy = 3x\left[y\right]_0^x = 3x^2 \text{ para } 0 < x < 1$$

$$f_Y(y) = \int_y^1 3x\,dx = 3\int_y^1 x\,dx = 3\left[\left(x^2/2\right)\right]_y^1 = (3/2)\left(1-y^2\right) \text{ para } 0 < y < 1$$

$$E(X) = \int_0^1 x\,3x^2\,dx = 3\int_0^1 x^3\,dx = 3\left[\left(x^4/4\right)\right]_0^1 = 3/4$$

$$E(Y) = \int_0^1 (3/2)y\left(1-y^2\right)dy = (3/2)\left(\int_0^1 y\,dy - \int_0^1 y^3\,dy\right) =$$

$$= (3/2)\left\{\left[y^2/2\right]_0^1 - \left[y^4/4\right]_0^1\right\} = (3/4)-(3/8) = 3/8$$

$$E(XY) = \int_0^1 \int_0^x (xy)(3x)\,dydx = 3\int_0^1 x^2 \int_0^x y\,dydx = 3\int_0^1 x^2 \left[y^2/2\right]_0^x dx =$$

$$= (3/2)\int_0^1 x^4 dx = (3/2)\left[x^5/5\right]_0^1 = 3/10$$

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = (3/10) - (3/4)(3/8) = 3/10 - 9/32 = (48-45)/160$$

Ou seja, $\mathrm{Cov}(X,Y) = 3/160$

b) Como $\mathrm{Cov}(X,Y) \neq 0$, X e Y não são independentes.

Exemplo 25. Sejam X e Y duas variáveis aleatórias tais que $Y = X^2$. Considere-se, ainda, que X tem a seguinte função de probabilidade:

$$p_X(x) = \frac{1}{4}\ ,\ \text{para}\ x = -2, -1, 1, 2$$

Determinar:

a) a função de probabilidade conjunta de X e de Y;
b) a covariância de X e Y;
Solução

A função de probabilidade marginal de X é

$$p_X(x) = \frac{1}{4}\ ,\ \text{para}\ x = -2, -1, 1, 2$$

e as duas variáveis aleatórias X e Y são relacionadas por $Y = X^2$.

Então:

a) A distribuição marginal de Y é

$$p_Y(y) = \frac{1}{2},\ \text{para}\ y = 1, 4$$

b) A distribuição conjunta de X e Y é expressa por

$$p_{XY}(x,y)=\frac{1}{4}\text{ , para }(x,y)\in\{(-2,4),(-1,1),(1,1),(2,4)\}$$

c) A covariância entre X e Y é calculada a seguir

$$E(XY)=(-2)4\cdot\frac{1}{4}+(-1)1\cdot\frac{1}{4}+1\cdot1\cdot\frac{1}{4}+2\cdot4\cdot\frac{1}{4}=\frac{1}{4}(-8-1+1+8)=0$$

$$E(X)=\frac{1}{4}\left[(-2)+(-1)+1+2\right]=0 \quad \text{e} \quad E(Y)=\frac{1}{2}(1+4)=\frac{5}{2}$$

donde

$$Cov(X,Y)=E(XY)-E(X)E(Y)=0-0.\frac{5}{2}=0$$

ou seja, $Cov(X,Y)=0$

d) Portanto, o coeficiente de correlação é nulo

$$\rho_{XY}=0$$

Isso mostra que o coeficiente de correlação (linear) pode ser nulo mesmo para variáveis aleatórias interrelacionadas (não linearmente).

Exemplo 26. Seja (X, Y) uma variável aleatória do tipo contínuo com a seguinte função de densidade de probabilidade:

$$f_{XY}(x,y) = x + y \text{ , para } 0 < x < 1 \text{ e } 0 < y < 1$$

a) Verificar que a função acima atende as condições de uma função de densidade de probabilidade conjunta;
b) as funções de densidade de probabilidade marginais de X e de Y;
c) as funções de densidade de probabilidade condicionadas associadas à variável aleatória (X,Y);
d) calcular a covariância e o coeficiente de correlação.
Solução

A função de probabilidade conjunta de (X,Y) é

$$f_{XY}(x,y) = x + y \text{ , para } 0 < x < 1 \text{ e } 0 < y < 1$$

a) sim, pois $f_{XY}(x,y) \ge 0$, para todo $(x,y) \in R^2_{XY}$ e $\iint f_{XY}(x,y)\,dx\,dy = 1$ pois

$$\iint_{R^2_{XY}} f_{XY}(x,y)\,dx\,dy = \int_0^1\int_0^1 (x+y)\,dx\,dy = \int_0^1\left(\int_0^1 x\,dx + y\int_0^1 dx\right)dy =$$

$$= \int_0^1\left[\left(\frac{1}{2}x^2\Big|_0^1\right) + \left(y\,x\Big|_0^1\right)\right]dy = \int_0^1\left(\frac{1}{2}+y\right)dy = \frac{1}{2}\int_0^1 dy + \int_0^1 y\,dy = \frac{1}{2}y\Big|_0^1 + \frac{1}{2}y^2\Big|_0^1 = \frac{1}{2}+\frac{1}{2} = 1$$

b)

$$f_X(x) = \int_0^1 (x+y)\,dy = x\,y\Big|_0^1 + \frac{1}{2}y^2\Big|_0^1 = x + \frac{1}{2}\text{ , para } 0 < x < 1$$

$$f_Y(y) = \int_0^1 (x+y)\,dx = \frac{1}{2}x^2\Big|_0^1 + y\,x\Big|_0^1 = \frac{1}{2} + y\text{ , para } 0 < y < 1$$

c)

$$f_{X|Y}(x\,|\,y) = \frac{x + y}{y + \frac{1}{2}}\text{ , para } 0 < x < 1\text{ , com } 0 < y < 1$$

$$f_{Y|X}(y\,|\,x) = \frac{x + y}{x + \frac{1}{2}}\text{ , para } 0 < y < 1\text{ , com } 0 < x < 1$$

d) a expressão geral dos momentos ordinários mistos

$$\mu_{ij} = E(X^i\,Y^j) = \int_0^1\int_0^1 x^i\,y^j\,(x+y)\,dx\,dy = \int_0^1\left(\int_0^1 x^{i+1}\,y^j\,dx + \int_0^1 x^i\;y^{j+1}\,dx\right)dy =$$

$$= \int_0^1\left(y^j\int_0^1 x^{i+1}\,dx + y^{j+1}\int_0^1 x^i\,dx\right)dy = \int_0^1\left(y^j\,\frac{x^{i+2}}{i+2}\Big|_0^1 + y^{j+1}\,\frac{x^{i+1}}{i+1}\Big|_0^1\right)dy =$$

$$= \int_0^1\left(\frac{1}{i+2}y^j + \frac{1}{i+1}y^{j+1}\right)dy = \frac{1}{i+2}\int_0^1 y^j\,dy + \frac{1}{i+1}\int_0^1 y^{j+1}\,dy =$$

$$= \frac{1}{i+2}\frac{1}{j+1}y^{j+1}\Big|_0^1 + \frac{1}{i+1}\frac{1}{j+2}y^{j+2}\Big|_0^1$$

$$\mu_{ij} = \frac{1}{i+2}\frac{1}{j+1} + \frac{1}{i+1}\frac{1}{j+2} = \frac{1}{(i+2)(j+1)} + \frac{1}{(i+1)(j+2)}$$

Portanto, tem-se:

$$\mu_{ij} = E(X^iY^j) = \frac{1}{(i+2)(j+1)} + \frac{1}{(i+1)(j+2)} \qquad \text{para } i, j = 0,1,2,3,...$$

A partir da expressão acima são calculadas facilmente as expectâncias, as variâncias e o momento conjunto de ordens 1 e 1

$$E(X) = \mu_{10} = \frac{7}{12} \text{ e } E(Y) = \mu_{01} = \frac{7}{12}$$

$$E(X^2) = \mu_{20} = \frac{5}{12} \text{ e } E(Y^2) = \mu_{02} = \frac{5}{12}$$

$$V(X) = E(X^2) - E^2(X) = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144} = V(Y)$$

$$E(XY) = \mu_{11} = \frac{1}{3}$$

$$\text{Cov}(X,Y) = E(XY - E(X)E(Y) = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{1}{3} - \frac{49}{144} = \frac{48-49}{144} = \frac{-1}{144}$$

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{-1}{144}$$

$$\rho_{XY} = \text{Cor}(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} = \frac{\frac{-1}{144}}{\sqrt{\frac{11}{144}\frac{11}{144}}} = \frac{-1}{11}$$

3.18 Expectâncias e Variâncias Condicionadas para Variáveis Bidimensionais

3.18.1 Definições

Seja (X,Y) uma variável aleatória bidimensional, do tipo discreto com função de probabilidade conjunta $p_{XY}(x,y)$, ou do tipo contínuo com função de densidade de probabilidade conjunta $f_{XY}(x,y)$. Conforme vimos anteriormente, as distribuições de probabilidade das variáveis aleatórias condicionadas (Y|X = x) e (X|Y = y) são obtidas a partir da distribuição conjunta de (X,Y), como segue:

Tipo discreto: $p_{X|Y}(x|y) = \dfrac{p_{XY}(x,y)}{p_Y(y)}$ e $p_{Y|X}(y|x) = \dfrac{p_{XY}(x,y)}{p_X(x)}$

Tipo contínuo: $f_{X|Y}(x|y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$ e $f_{Y|X}(y|x) = \dfrac{f_{XY}(x,y)}{f_X(x)}$

Admitindo-se que as variáveis (Y|X = x) e (X|Y = y) possuam variâncias finitas, podemos calcular as expectâncias e variâncias dessas duas variáveis aleatórias, denominadas <u>expectâncias e variâncias condicionadas</u>:

Tipo discreto:

$$E(X|Y=y)=E(X|y)=\sum_x x\, p_{X|Y}(x|y)$$

$$E(Y|X=x)=E(Y|x)=\sum_y y\, p_{Y|X}(y|x)$$

$$V(X|Y=y)=V(X|y)= E(X^2|y) - E^2(X|y) =$$

$$= \sum_x x^2\, p_{X|Y}(x|y) - \left[\sum_x x\, p_{X|Y}(x|y)\right]^2$$

$$V(Y|X=x)=V(Y|x)= E(Y^2|x) - E^2(Y|x) =$$

$$= \sum_y y^2\, p_{Y|X}(y|x) - \left[\sum_y y\, p_{Y|X}(y|x)\right]^2$$

Tipo contínuo:

$$E(X|Y=y)=E(X|y)= \int x\, f_{X|Y}(x|y)\, dx$$

$$E(Y|X=x)=E(Y|x)=\int y\, f_{Y|X}(y|x)\; dy$$

$$V(X|Y=y)=V(X|y)=E(X^2|y)-E^2(X|y)=$$
$$=\int x^2 f_{X|Y}(x|y)\,dx-\left[\int x\, f_{X|Y}(x|y)\,dx\right]^2$$

$$V(Y|X=x)=V(Y|x)=E(Y^2|x)-E^2(Y|x)=$$
$$=\int y^2 f_{Y|X}(y|x)\,dy-\left[\int y\, f_{Y|X}(y|x)\,dy\right]^2$$

3.18.2 Variáveis Aleatórias Associadas às Expectâncias e Variâncias Condicionadas

a) **Expectâncias condicionadas aleatórias**

i) Represente-se por E(X|Y) uma variável aleatória que assume o valor E(X|y) quando a variável aleatória Y assume o valor y, para qualquer y. Essa variável aleatória denomina-se expectância condicionada aleatória de X dado Y.

ii) Analogamente, represente-se por E(Y|X) uma variável aleatória que assume o valor E(Y|x) quando a variável aleatória X assume o valor x, para qualquer x. Essa variável aleatória denomina-se expectância condicionada aleatória de Y dado X.

b) **Variâncias condicionadas aleatórias**

i) Represente-se por V(X|Y) uma variável aleatória que assume o valor V(X|y) quando a variável aleatória Y assume o valor y, para qualquer y. Essa variável aleatória denomina-se variância condicionada aleatória de X dado Y.

ii) Analogamente, represente-se por V(Y|X) uma variável aleatória que assume o valor V(Y|x) quando a variável aleatória X assume o valor x, para qualquer x. Essa variável aleatória denomina-se variância condicionada aleatória de Y dado X.

3.18.3 Teoremas Básicos sobre Expectâncias e Variâncias Condicionadas Aleatórias

Com relação às expectâncias e variâncias condicionadas, são importantes os três teoremas a seguir apresentados.

Teorema 1. Lei das Expectâncias Iteradas

i) $E\left[E\left(X|Y\right)\right]=E(X)$

ii) $E\left[E\left(Y|X\right)\right] = E(Y)$

Teorema 2. Decomposição da Variância

i) $V(X) = E\left[V\left(X|Y\right)\right] + V\left[E\left(X|Y\right)\right]$

ii) $V(Y) = E\left[V\left(Y|X\right)\right] + V\left[E\left(Y|X\right)\right]$

Teorema 3. Regressão Linear

Seja (X,Y) uma variável aleatória bidimensional, tal que

$E\left(X\right) = \mu_x \quad E\left(Y\right) = \mu_y \quad V\left(X\right) = \sigma_X^2 \quad V\left(Y\right) = \sigma_Y^2$ e $\mathrm{Corr}\left(X,Y\right) = \rho$

Então:

i) Se a regressão de X em Y for linear, isto é, se a função de regressão tiver a forma funcional linear, a referida função é expressa em termos dos parâmetros inicialmente dados da seguinte forma

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

ii) Se a regressão de Y em X for linear, isto é, se a função de regressão tiver a forma funcional linear, a referida função é expressa em termos dos parâmetros inicialmente dados da seguinte forma

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

Exemplo 27. Seja (X, Y) uma variável aleatória bidimensional referente com a seguinte função de probabilidade conjunta:

X \ Y	2	5
1	1/8	2/8
3	3/8	2/8

a) Determinar as expectâncias condicionadas E(X|y) e E(Y|x) e as distribuições das variáveis aleatórias E(X|Y) e E(Y|X).

b) Aplicar a lei das expectâncias iteradas a E(X|Y) e E(Y|X).

Solução.

A distribuição conjunta e as distribuições marginais de X e de Y são mostradas na tabela abaixo.

	Y		
X	2	5	$p_X(x)$
1	1/8	2/8	3/8
3	3/8	2/8	5/8
$p_Y(y)$	4/8	4/8	1

As distribuições condicionadas associadas a (X|Y) são mostradas a seguir.

$$p_{X|Y}(x|2)=\begin{cases}\frac{1}{4}, \text{ se } x=1\\ \frac{3}{4}, \text{ se } x=3\end{cases} \qquad p_{X|Y}(x|5)=\begin{cases}\frac{2}{4}=\frac{1}{2}, \text{ se } x=1\\ \frac{2}{4}=\frac{1}{2}, \text{ se } x=3\end{cases}$$

$$p_{Y|X}(y|1)=\begin{cases}\frac{1}{3}, \text{ se } y=2\\ \frac{2}{3}, \text{ se } y=5\end{cases} \qquad p_{Y|X}(y|3)=\begin{cases}\frac{3}{5}, \text{ se } y=2\\ \frac{2}{5}, \text{ se } y=5\end{cases}$$

Então, as expectâncias condicionadas de (X|Y=y) são:

$$E(X|Y=2)=E(X|2)=1.\frac{1}{4}+3.\frac{3}{4}=\frac{1+9}{4}=\frac{10}{4}=\frac{5}{2}$$

$$E(X|Y=5)=E(X|5)=1.\frac{1}{2}+3.\frac{1}{2}=\frac{1+3}{2}=\frac{4}{2}=2$$

e as expectâncias condicionadas de (Y|X=x) são:

$$E(Y|X=1)=E(Y|1)=2.\frac{1}{3}+5.\frac{2}{3}=\frac{2+10}{3}=\frac{12}{3}=4$$

$$E(Y|X=3)=E(Y|3)=2.\frac{3}{5}+5.\frac{2}{5}=\frac{6+10}{5}=\frac{16}{5}=3,2$$

Portanto, as variáveis aleatórias associadas às expectâncias condicionadas são:

$$E(X|Y)=\begin{cases}\frac{5}{2}\text{ , se } y=2 \text{ ; com } P(Y=2)=p_Y(2)=\frac{1}{2}\\ 2\text{ , se } y=5 \text{ ; com } P(Y=5)=p_Y(5)=\frac{1}{2}\end{cases}$$

$$E(Y|X)=\begin{cases}4 \quad \text{, se } x=1 \text{ ; com } P(X=1)=p_X(1)=\frac{3}{8}\\ \frac{16}{5}\text{ , se } x=3 \text{ ; com } P(X=3)=p_X(3)=\frac{5}{8}\end{cases}$$

Aplicando agora a lei das expectativas iteradas, tem-se:

$$E\left[E\left(X|Y\right)\right]=\frac{5}{2}\frac{1}{2}+2\frac{1}{2}=\frac{5}{4}+1=\frac{5+4}{4}=\frac{9}{4}$$

$$E\left[E\left(Y|X\right)\right]=4\frac{3}{8}+\frac{16}{5}\frac{5}{8}=\frac{3}{2}+2=\frac{3+4}{2}=\frac{7}{2}$$

Deve-se observar que E[E(X|Y)] = E(X) e E[E(Y|X)] = E(Y) , pois

$$E(X)=1\frac{3}{8}+3\frac{5}{8}=\frac{3+15}{8}=\frac{18}{8}=\frac{9}{4}$$
$$E(Y)=2\frac{1}{2}+5\frac{1}{2}=\frac{2+5}{2}=\frac{7}{2}$$

Exemplo 28. Seja (X, Y) uma variável aleatória com a seguinte função de densidade de probabilidade conjunta:

$$f_{XY}(x,y) = \frac{1}{2}\text{ , para } 0<x<2 \text{ e } 0<y<x$$

a) Determinar as funções de densidade de probabilidade marginais de X e de Y;
b) determinar as funções de densidade de probabilidade condicionadas associadas à variável aleatória (X, Y);
c) determinar as expectâncias condicionadas E(X | y) e E(Y | x);
d) aplicar a lei das expectâncias iteradas a E(X | Y) e E(Y | X).

Solução

A função de densidade de probabilidade conjunta é

$$f_{XY}(x, y) = \frac{1}{2} \quad , \text{ para } 0 < x < 2 \ \text{ e } \ 0 < y < x$$

Então segue:

a) as funções de densidade marginais de X e de Y são:

$$f_X(x) = \int_0^x \frac{1}{2} dy = \frac{1}{2}\int_0^x dy = \frac{1}{2}\left(y\Big|_0^x\right) = \frac{1}{2}x \ \text{ para } \ 0 < x < 2$$

$$f_X(x) = \frac{1}{2}x, \ \text{ para } 0 < x < 2$$

$$f_Y(y) = \int_y^2 \frac{1}{2} dx = \frac{1}{2}\int_y^2 dx = \frac{1}{2} x\Big|_y^2 = \frac{1}{2}(2-y) \ , \text{ para } 0 < y < 2$$

$$f_Y(y) = \frac{1}{2}(2-y), \ \text{ para } 0 < y < 2$$

b) as funções de densidade condicionais associadas a (X, Y) são:

$$f_{X|Y}(x|y) = \frac{\frac{1}{2}}{\frac{1}{2}(2-y)} = \frac{1}{2-y}, \ \text{ para } y < x < 2, \text{ com } 0 < y < 2$$

$$f_{X|Y}(x|y) = \frac{1}{2-y}, \ \text{ para } y < x < 2, \text{ com } 0 < y < 2$$

$$f_{Y|X}(y|x) = \frac{\frac{1}{2}}{\frac{1}{2}x} = \frac{1}{x}, \ \text{ para } 0 < y < x \ , \text{ com } 0 < x < 2$$

$$f_{Y|X}(y|x) = \frac{1}{x}, \ \text{ para } 0 < y < x \ , \text{ com } 0 < x < 2$$

c) as expectâncias condicionadas são calculadas a seguir:

c-i) pelo método direto

$$E(X|Y=y) = E(X|y) = \mu_X(y) = \int_y^2 x \frac{1}{2-y}\, dx = \frac{1}{2-y}\int_y^2 x\, dx = \frac{1}{2-y}\,\frac{1}{2}\, x^2\Big|_y^2 =$$

$$= \frac{1}{2}\,\frac{1}{2-y}\left(4-y^2\right) = \frac{1}{2}\,\frac{(2+y)(2-y)}{2-y} = \frac{2+y}{2} \quad , \quad \text{com } 0 < y < x$$

$$E(X|Y=y) = E(X|y) = \mu_X(y) = \frac{1}{2}(2+y) = \frac{2+y}{2} \text{ , com } 0 < y < 2$$

e

$$E(Y|X=x) = E(Y|x) = \mu_Y(x) = \int_0^x y\frac{1}{x} dy = \frac{1}{x}\int_0^x y\, dy = \frac{1}{x}\,\frac{1}{2}\, y^2\Big|_0^x =$$

$$= \frac{1}{2x} x^2 = \frac{x}{2} \text{ , } \quad \text{com } \ 0 < x < 2$$

$$E(Y|X=x) = E(Y|x) - \mu_Y(x) = \frac{1}{2} x \text{ , com } 0 < x < 2$$

c-ii) pelo método alternativo

Examinando a expressão da função de densidade de (X|y) verifica-se que essa variável aleatória tem distribuição uniforme no intervalo (y , 2). Portanto, tem-se, diretamente,

$$E(X|Y=y) = E(X|y) = \mu_X(y) = \frac{1}{2}(2+y) = \frac{2+y}{2} \text{ , com } 0 < y < 2$$

Examinando a expressão da função de densidade de (Y|x) verifica-se que essa variável aleatória tem distribuição uniforme no intervalo (0 , x). Portanto, tem-se, diretamente,

$$E(Y|X=x) = E(Y|x) = \mu_Y(x) = \frac{1}{2} x \text{ , com } 0 < x < 2$$

d) Lei das expectâncias iteradas

Considerando a variável aleatória E(X|Y), tem-se:

i) pelo método direto

$$E[E(X|Y)] = \int_0^2 \frac{2+y}{2}\, f_Y(y)\, dy = \int_0^2 \frac{2+y}{2}\,\frac{1}{2}\,(2-y)\, dy = \frac{1}{4}\int_0^2 (4-y^2)\, dy$$

$$= \frac{1}{4}\left(4\int_0^2 dy - \int_0^2 y^2\, dy\right) = \frac{1}{4}\left(4\,y\Big|_0^2 - \frac{1}{3}y^3\Big|_0^2\right) = \frac{1}{4}\left(8-\frac{8}{3}\right) = 2 - \frac{2}{3} = \frac{4}{3}$$

$$E(X) = \int_0^2 x\,\frac{1}{2}x\,dx = \frac{1}{2}\int_0^2 x^2\,dx = \frac{1}{2}\,\frac{1}{3}\,x^3\Big|_0^2 = \frac{8}{6} = \frac{4}{3}$$

ii) pelo método alternativo

$$E(X|y) = \frac{1}{2}(2+y) \quad \text{donde} \quad E(X|Y) = \frac{1}{2}(2+Y)$$

portanto

$$E\left[E(X|Y)\right] = E\left[\frac{1}{2}(2+Y)\right] = \frac{1}{2}\left[2+E(Y)\right]$$

mas

$$E(Y) = \int_0^2 y\,\frac{1}{2}(2-y)\,dy = \frac{1}{2}\int_0^2 y(2-y)\,dy = \frac{1}{2}\left(\int_0^2 2y\,dy - \int_0^2 y^2\,dy\right) =$$

$$= \frac{1}{2}\left(y^2\Big|_0^2 - \frac{1}{3}y^3\Big|_0^2\right) = \frac{1}{2}\left(4-\frac{8}{3}\right) = 2-\frac{4}{3} = \frac{2}{3}$$

Logo

$$E\left[E(X|Y)\right] = \frac{1}{2}\left[2+E(Y)\right] = \frac{1}{2}\left[2+\frac{2}{3}\right] = 1+\frac{1}{3} = \frac{4}{3} = E(X)$$

Considerando a variável aleatória E(Y|X), tem-se:

i) pelo método direto

$$E[E(Y|X)] = \int_0^2 \frac{x}{2}\, f_X(x)\,dx = \int_0^2 \frac{x}{2}\,\frac{1}{2}\,x\,dx = \frac{1}{4}\int_0^2 x^2\,dx = \frac{1}{4}\,\frac{1}{3}\,x^3\Big|_0^2 = \frac{8}{12} = \frac{2}{3}$$

$$E(Y) = \int_0^2 y\,\frac{1}{2}(2-y)\,dy = \frac{1}{2}\int_0^2 y(2-y)\,dy = \frac{1}{2}\left(\int_0^2 2y\,dy - \int_0^2 y^2\,dy\right) =$$

$$= \frac{1}{2}\left(y^2\Big|_0^2 - \frac{1}{3} y^3\Big|_0^2 \right) = \frac{1}{2}\left(4-\frac{8}{3}\right) = 2-\frac{4}{3} = \frac{2}{3}$$

ii) pelo método alternativo

$$E(Y|x) = \frac{1}{2}x \quad \text{donde} \quad E(Y|X) = \frac{1}{2}X$$

portanto

$$E[E(Y|X)] = E\left[\frac{1}{2}X\right] = \frac{1}{2}\,E(X)$$

mas

$$E(X) = \int_0^2 x\,\frac{1}{2}x\,dx = \frac{1}{2}\int_0^2 x^2\,dx = \frac{1}{2}\,\frac{1}{3}\,x^3\Big|_0^2 = \frac{8}{6} = \frac{4}{3}$$

$$E[E(Y|X)] = \frac{1}{2}\frac{4}{3} = \frac{2}{3} = E(Y)$$

Os conceitos de expectância e de variância também possuem muita aplicação prática no estudo de combinações lineares de variáveis aleatórias. Este assunto será tratado a seguir.

3.19 Combinação Linear de Variáveis Aleatórias

Considerem-se n variáveis aleatórias quaisquer $X_1, X_2, X_3, \ldots, X_n$ e n números reais $a_1, a_2, a_3, \ldots, a_n$. Faça-se

$$Y = \sum_{i=1}^{n} a_i X_i \qquad (1)$$

A variável aleatória Y denomina-se combinação linear de $X_1, X_2, X_3, \ldots, X_n$.

Média (Expectância) e Variância da Combinação Linear

Adotando a notação:

$$\mu_Y = E(Y) \qquad \sigma_Y^2 = V(Y) \qquad \mu_i = E(X_i) \qquad \sigma_i^2 = V(X_i) \quad i = 1, 2, 3, \ldots, n$$

$$\sigma_{ij} = Cov(X_i, X_j) \qquad \rho_{ij} = \rho_{X_i X_j} = Cor(X_i, X_j) \quad i, j = 1, 2, 3, \ldots, n$$

a média (expectância) de Y é dada por

$$\mu_Y = E(Y) = \sum_{i=1}^{n} a_i\,\mu_i \tag{2}$$

e a variância de Y é dada por

$$\sigma_Y^2 = V(Y) = E\left[\left(Y - \mu_Y\right)^2\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{Cov}\left(X_i, X_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\,\sigma_{ij}$$

Notando que

$$\begin{aligned} \sigma_{ij} &= \operatorname{Cov}(X_i, X_j) = \operatorname{Cov}(X_j, X_i) = \sigma_{ji} \qquad i, j = 1, 2, 3, \ldots, n \quad \text{e} \\ \sigma_{ii} &= \operatorname{Cov}(X_i, X_i) = V(\,X_i) = \sigma_i^2 \qquad i = 1, 2, 3, \ldots, n \end{aligned} \tag{3}$$

A expressão acima pode ser reescrita como

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2\, V(X_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j>i}}^{n} a_i\, a_j \operatorname{Cov}\left(X_i, X_j\right) = \sum_{i=1}^{n} a_i^2\,\sigma_i^2 + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j>i}}^{n} a_i\, a_j\,\sigma_{ij} \tag{4}$$

Lembrando a definição de coeficiente de correlação linear, temos:

$$\rho_{ij} = \rho_{X_i X_j} = \frac{\operatorname{Cov}(X_i, X_j)}{\sqrt{V(X_i)\,V(X_j)}} = \frac{\sigma_{ij}}{\sigma_i\,\sigma_j} \quad \text{o que implica} \quad \sigma_{ij} = \rho_{ij}\,\sigma_i\,\sigma_j \tag{5}$$

Assim, podemos reescrever a expressão (4), alternativamente, na forma:

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2\,\sigma_i^2 + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j>i}}^{n} a_i\, a_j\,\sigma_i\,\sigma_j\,\rho_{ij} \tag{6}$$

Se as variáveis $X_1, X_2, X_3, \ldots, X_n$ forem não correlacionadas (linearmente), o que evidentemente ocorre no caso mais particular de serem independentes, verifica-se, tanto de (4) quanto de (6), que a expressão da variância se reduz ao caso em que só tem o primeiro termo, isto é:

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2\,\sigma_i^2 \tag{7}$$

Exemplo 29. Sejam X_1 e X_2 duas variáveis aleatórias, a_1 e a_2 dois números reais e Y uma combinação linear dessas duas variáveis, isto é, $Y = \sum_{i=1}^{2} a_i\,X_i = a_1\,X_1 + a_2\,X_2$. Determinar as expressões da expectância e da variância de Y.

Solução

Então, tem-se:

$$\mu_Y = E(Y) = \sum_{i=1}^{2} a_i\,\mu_i = a_1\,\mu_1 + a_2\,\mu_2$$

e

$$\sigma_Y^2 = V(Y) = \sum_{i=1}^{2} a_i^2\,\sigma_i^2 + 2\sum_{i=1}^{2}\sum_{\substack{j=1\\ j>i}}^{2} a_i\,a_j\,\sigma_{ij} = a_1^2\,\sigma_1^2 + a_2^2\,\sigma_2^2 + 2\,a_1\,a_2\,\sigma_{12}$$

Exemplo 30. Sejam X_1, X_2 e X_3 três variáveis aleatórias, e a_1, a_2 e a_3 três números reais. Seja Y uma combinação linear dessas três variáveis, isto é,

$$Y = \sum_{i=1}^{3} a_i\,X_i = a_1\,X_1 + a_2\,X_2 + a_3\,X_3.$$

Determinar a expectância e a variância de Y.

Solução

Então, tem-se:

$$\mu_Y = E(Y) = \sum_{i=1}^{3} a_i\,\mu_i = a_1\,\mu_1 + a_2\,\mu_2 + a_3\,\mu_3$$

e

$$\sigma_Y^2 = V(Y) = \sum_{i=1}^{3} a_i^2\,\sigma_i^2 + 2\sum_{i=1}^{3}\sum_{\substack{j=1\\ j>i}}^{3} a_i\,a_j\,\sigma_{ij} =$$

$$= a_1^2\,\sigma_1^2 + a_2^2\,\sigma_2^2 + a_3^2\,\sigma_3^2 \;+\; 2\left(a_1\,a_2\,\sigma_{12} + a_1\,a_3\,\sigma_{13} + a_2\,a_3\,\sigma_{23}\right)$$

Aplicação de combinações lineares de duas variáveis – carteira de investimentos

Exemplo 31. Uma carteira de investimentos deverá ser formada apenas pelos ativos A e B. Sejam X e Y, respectivamente, os retornos de A e B para uma aplicação durante certo período de tempo (por exemplo um mês). Sabendo-se que a variância de X é igual a 25, a variância de Y é igual a 16 e o coeficiente de correlação linear entre X e Y é igual a 0,5, determinar a proporção dos recursos da carteira que devem ser investidos em cada um desses ativos, de sorte a minimizar a variância do retorno (também denominada risco) da carteira.

Solução

Sejam:

R = retorno da carteira e $V(R) = \sigma_R^2$ a variância do retorno da carteira

α = proporção dos recursos investidos no ativo A; sendo $0 \leq \alpha \leq 1$

O retorno da carteira é uma combinação linear dos retornos de A e de B, podendo ser expresso por:

$$R = \alpha X + (1-\alpha) Y \quad ; \quad \text{para} \quad 0 \leq \alpha \leq 1$$

Empregando a expressão da variância de uma combinação linear de duas variáveis aleatórias, tem-se:

$$V(R) = \sigma_R^2 = \alpha^2 \sigma_1^2 + (1-\alpha)^2 \sigma_2^2 + 2\alpha(1-\alpha)\sigma_1 \sigma_2 \rho_{12} =$$

$$= 25\alpha^2 + 16(1-\alpha)^2 + 2\alpha(1-\alpha)\sqrt{25}\sqrt{16} \; . \, 0{,}5 =$$

$$= 25\alpha^2 + 16 - 32\alpha + 16\alpha^2 + 20\alpha - 20\alpha^2 =$$

$$= 21\alpha^2 - 12\alpha + 16 \quad ; \quad \text{para} \quad 0 \leq \alpha \leq 1$$

Se não levarmos em consideração a restrição $0 \leq \alpha \leq 1$ vemos que a expressão anterior é um trinômio do segundo grau em α , não negativo (uma vez que $\sigma_R^2 \geq 0$), cujo coeficiente de α^2 é positivo. Portanto, esse trinômio do segundo grau apresenta um mínimo em todo o campo real, correspondente ao valor de α em que se anula a primeira derivada da expressão da variância do retorno da carteira, σ_R^2. Após a obtenção desse valor, deve-se verificar se ele pertence ou não ao intervalo [0,1], para se concluir qual é a solução adequada.

No caso, tem-se:

$$\frac{d\sigma_R^2}{d\alpha} = \frac{d}{d\alpha}\left(21\alpha^2 - 12\alpha + 16\right) = 42\,\alpha - 12 = 0$$

logo

$$\alpha = \frac{12}{42} = 0{,}2857 = 0{,}286$$

Assim, em todo o campo real, o mínimo do trinômio corresponde a um valor de α que pertence ao intervalo [0,1]. Portanto, deve-se investir 28,6% dos recursos no ativo A e 71,4% no ativo B, para que o risco de mercado da carteira seja mínimo.

Exemplo 32. Admita-se, no exemplo anterior, que se tenha: $\sigma_1^2 = 1$, $\sigma_2^2 = 9$ e $\rho_{12} = 0,6$

Determinar a proporção dos recursos da carteira que devem ser investidos em cada um desses ativos, de sorte a minimizar a variância do retorno.

Solução

Neste caso,

$$V(R) = \sigma_R^2 = 1\alpha^2 + 9(1-\alpha)^2 + 2\alpha(1-\alpha)\sqrt{1}\,\sqrt{9}\ .\ 0,6 =$$

$$= 1\alpha^2 + 9 - 18\alpha + 9\alpha^2 + 3,6\alpha - 3,6\alpha^2 =$$

$$= 6,4\alpha^2 - 14,4\alpha + 9 \ \ ; \ \text{para} \ \ 0 \le \alpha \le 1$$

do que resulta

$$\frac{d\sigma_R^2}{d\alpha} = \frac{d}{d\alpha}\left(6,4\alpha^2 - 14,4\alpha + 9\right) = 12,8\ \alpha - 14,4 = 0$$

ou seja

$$\alpha = \frac{14,4}{12,8} = 1,125$$

Assim, verifica-se que $\alpha > 1$ e, portanto, o trinômio $\sigma_R^2 = 6,4\alpha^2 - 14,4\alpha + 9$ é decrescente para $\alpha < 1,125$. Portanto, o valor de $\alpha \in [0,1]$ que minimiza σ_R^2 é igual a 1. Logo deve-se investir todos os recursos da carteira no ativo A.

Com relação ao tópico combinações lineares de variáveis aleatórias, são importantes também aquelas que envolvem apenas variáveis aleatórias independentes. Neste caso, a expectância da combinação linear não se altera mas a variância sim, como foi visto anteriormente, tendo apenas o primeiro termo da equação (6), ou seja, a sua expressão é aquela da equação (7) da seção 3.19. Embora este seja um caso mais restrito, devido a sua importância em certas aplicações, o assunto será apresentado com algum detalhe a seguir.

3.20 Combinação Linear de Variáveis Aleatórias Independentes

Cálculos da Expectância e da Variância

Sejam $X_1, X_2, X_3, ..., X_n$ n variáveis aleatórias independentes e $a_1, a_2, a_3, ..., a_n$ n números reais. A variável aleatória Y definida por

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n = \sum_{i=1}^{n} a_i X_i$$

é denominada combinação linear das variáveis $X_1, X_2, X_3, ..., X_n$.

Sejam $E(X_i) = \mu_i$ (com $i = 1, 2, 3, ..., n$) as médias e $V(X_i) = \sigma_i^2$ (com $i = 1, 2, 3, ..., n$) as variâncias das variáveis aleatórias componentes da combinação linear Y. Então, pelas propriedades anteriormente vistas da expectância e da variância, demonstra-se que a expectância e a variância da combinação linear Y, são, respectivamente:

(i) $E(Y) = \sum_{i=1}^{n} a_i E(X_i)$ ou seja $\mu_Y = \sum_{i=1}^{n} a_i \mu_i$

(ii) $V(Y) = \sum_{i=1}^{n} a_i^2 V(X_i)$ ou seja $\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2$

Dois casos particulares importantes são vistos a seguir.

a) Se todas as variáveis componentes da combinação linear possuem a mesma expectância e a mesma variância:

$$E(X_i) = \mu \text{ (com } i = 1, 2, 3, ..., n) \text{ e } V(X_i) = \sigma^2 \text{ (com } i = 1, 2, 3, ..., n)$$

Nesse caso:

(i) $E(Y) = \sum_{i=1}^{n} a_i \mu$ ou seja $\mu_Y = \left(\sum_{i=1}^{n} a_i\right)\mu$

(ii) $V(Y) = \sum_{i=1}^{n} a_i^2 \sigma^2$ ou seja $\sigma_Y^2 = \left(\sum_{i=1}^{n} a_i^2\right)\sigma^2$

b) Se todas as variáveis componentes da combinação linear possuem a mesma expectância e a mesma variância e, além disso, todos os coeficientes da combinação linear são iguais:

$E(X_i)=\mu$ (com $i=1,2,3,\ldots,n$), $V(X_i)=\sigma^2$ (com $i=1,2,3,\ldots,n$) e, ainda, os coeficientes $a_i=a$ (para $i=1,2,3,\ldots,n$)

Nesse caso:

(i) $E(Y)=\sum_{i=1}^{n} a\mu$ ou seja $\mu_Y = n\,a\,\mu$

(ii) $V(Y)=\sum_{i=1}^{n} a^2\sigma^2$ ou seja $\sigma_Y^2 = n\,a^2\,\sigma^2$

Esses resultados serão muito úteis no desenvolvimento posterior.

Aplicações importantes do último caso particular. Soma e Média Aritmética.

Como aplicações mais importantes desse último caso tem-se as duas seguintes combinações lineares, frequentemente encontradas no emprego da Estatística:

1) Soma das variáveis aleatórias $S = X_1+X_2+X_3+\ldots+X_n=\sum_{i=1}^{n} X_i$.

Para a variável aleatória soma tem-se:

(i) $E(S)=\sum_{i=1}^{n} \mu$ ou seja $\mu_S = n\,\mu$

(ii) $V(S)=\sum_{i=1}^{n} \sigma^2$ ou seja $\sigma_S^2 = n\,\sigma^2$

Nota: a soma corresponde ao caso particular em que a combinação linear tem todos os coeficientes iguais a 1.

2) Média aritmética das variáveis aleatórias $\bar{X}=\frac{1}{n}(X_1+X_2+X_3+\ldots+X_n)=\frac{1}{n}\sum_{i=1}^{n} X_i$

Para a variável aleatória média aritmética tem-se:

(i) $E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}n\mu = \mu$ ou seja $\mu_{\bar{X}} = \mu$

(ii) $V(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}n\sigma^2 = \frac{1}{n}\sigma^2$ ou seja $\sigma^2_{\bar{X}} = \frac{1}{n}\sigma^2$

Nota: a média aritmética corresponde ao caso particular em que a combinação linear tem todos os coeficientes iguais a 1/n.

A maioria das principais aplicações do tópico aqui apresentado será vista posteriormente, a partir do próximo capítulo. Entretanto, será a seguir apresentado um dos mais importantes resultados do Cálculo de Probabilidades – a lei dos grandes números (na sua forma mais simples, chamada lei fraca dos grandes números).

3.20 Lei Fraca dos Grandes Números

Seja $\{X_n\} = \{X_1, X_2, X_3, ..., X_n, ...\}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, cada uma com média finita $E(X_i) = \mu$. Então, para qualquer $\varepsilon > 0$,

$$P\left\{\left|\frac{X_1 + X_2 + X_3 + ... + X_n}{n} - \mu\right| > \varepsilon\right\} \to 0 \quad \text{quando } n \to \infty$$

Demonstração

A demonstração a seguir apresentada requer o pressuposto adicional de que as variáveis aleatórias tenham variância finita $V(X_i) = \sigma^2$. Nessas condições,

$$E\left(\frac{X_1 + X_2 + X_3 + ... + X_n}{n}\right) = \mu \quad \text{e} \quad V\left(\frac{X_1 + X_2 + X_3 + ... + X_n}{n}\right) = \frac{\sigma^2}{n}$$

Portanto, aplicando-se a desigualdade de Chebychev resulta

$$P\left\{\left|\frac{X_1 + X_2 + X_3 + ... + X_n}{n} - \mu\right| > \varepsilon\right\} \le \frac{\sigma^2}{n\varepsilon^2}$$

logo

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1+X_2+X_3+...+X_n}{n}-\mu\right|>\varepsilon\right\} \le \lim_{n\to\infty}\frac{\sigma^2}{n\,\varepsilon^2}$$

ou seja

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1+X_2+X_3+...+X_n}{n}-\mu\right|>\varepsilon\right\} \le 0$$

Por outro lado, a probabilidade acima deve também ser maior ou igual a zero (axioma do Cálculo de Probabilidades) e assim

$$\lim_{n\to\infty} P\left\{\left|\frac{X_1+X_2+X_3+...+X_n}{n}-\mu\right|>\varepsilon\right\} = 0$$

Como será visto posteriormente, isso corresponde ao conceito de convergência em probabilidade. Ou seja, nas condições estabelecidas, diz-se que a média aritmética das variáveis aleatórias convergente em probabilidade para a média (expectância) das variáveis.

Portanto, $\left\{\bar{X}_n\right\}=\left\{\frac{1}{n}\sum_{i=1}^{n}X_i\right\}\xrightarrow{p}\mu$ (converge em probabilidade para μ).