# 7 Caracterização Adicional de Variáveis Aleatórias

## ESQUEMA DO CAPÍTULO

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Parâmetros:**
  1. A relação determinística $ax + by = 0$ possui parâmetros $a$ e $b$, no sentido de que qualquer escolha particular de $a$ e $b$ determina uma função linear específica;
  2. Suponha que X seja uma variável aleatória contínua, com fdp $f(x) = ke^{-kx}$, $x \geq 0$. Este modelo aleatório possui parâmetro $k > 0$, cujo significado será explicado depois;
  3. Etc.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Definição:**

  Seja X uma variável aleatória discreta, com valores possíveis $x_1, ..., x_n, ...$ . Seja $p(x_i) = P(X = x_i)$, $i = 1, 2, ..., n, ...$ . Então, o *valor esperado* de X (ou *esperança matemática* de X), denotado por E(X), é definido como

  $$E(X) = \Sigma^{\infty}_{i=1}\, x_i\, p(x_i),$$

  se a série convergir absolutamente, isto é, se

  $$E(X) = \Sigma^{\infty}_{i=1}\, |x_i|\, p(x_i) < \infty,$$

  Este número é também denominado o *valor médio* de X, ou *expectância* de X.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Comentários:**

a) Se X tomar apenas um número finito de valores, a expressão acima se torna $E(X) = \Sigma^n_{i=1} x_i\, p(x_i)$. Isto pode ser considerado como uma “média ponderada” dos valores possíveis $x_1, ..., x_n$. Se todos esses valores possíveis forem igualmente prováveis, $E(X) = (1/n)\Sigma^n_{i=1} x_i$, a qual representa a média aritmética simples ou usual dos $n$ valores possíveis;

b) Se um dado equilibrado for jogado e a variável aleatória X designar o número de pontos obtidos, então E(X) = (1/6)(1+2+3+4+5+6) = 7/2. Este exemplo simples ilustra nitidamente que E(X) não é o resultado que esperar quando X for observado uma única vez. De fato, na situação anterior, E(X) = 7/2 nem mesmo é um valor possível de X!

c) Devemos notar a semelhança entre a noção de valor esperado, como foi definida acima (especialmente se X puder tomar somente um número finito de valores), e a noção de média de um conjunto de números $z_1, ..., z_n$.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Exemplo:**

  Um fabricante produz peças tais que 10% são defeituosas e 90% são não-defeituosas. Se uma peça defeituosa for produzida, o fabricante perde US$ 1, enquanto uma peça não-defeituosa lhe dá lucro de US$ 5.

  Se X for o lucro líquido por peça, então X será uma variável aleatória cujo valor esperado é calculado como

  $$E(X) = -1\times(0,1) + 5\times(0,9) = \text{US\$ } 4,40.$$

  Suponha-se que um grande número de tais peças seja produzido. Nesse caso, quando o fabricante perder US$ 1 cerca de 10% das vezes e ganhar US$ 5 cerca de 90% da vezes, ele esperará ganhar cerca de US$ 4,40 por peça, a longo prazo.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Teorema:**

  Seja X uma variável aleatória distribuída binomialmente, com parâmetro $p$, baseada em $n$ repetições de um experimento. Então,

  $$E(X) = np.$$

- **Demonstração:**

  Como $P(X = k) = \binom{n}{k} p^k(1-p)^{n-k}$, teremos

  $$E(X) = np\sum_{s=0}^{n-1} \binom{n-1}{s} p^s(1-p)^{n-1-s}.$$

- **Comentário:**

  O resultado acima corresponde à nossa noção intuitiva. Se supusermos que a probabilidade de algum evento A seja, por exemplo, 0,3, então se repetirmos esse experimento, por exemplo, por 100 vezes, deveremos esperar que A ocorra cerca de 100×(0,3) = 30 vezes.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Exemplo:**

  Uma máquina impressora tem uma probabilidade constante de 0,05 de entrar em pane, em um dia qualquer. Se a máquina não apresentar panes durante a semana, um lucro de $S será obtido. Se 1 ou 2 panes ocorrerem, um lucro de $R será alcançado (R<S). Se 3 ou mais panes ocorrerem, um lucro de -$L será obtido. Seja X o lucro obtido por semana de cinco dias úteis. Os valores possíveis de X são R, S e -L. Seja B o número de panes por semana. Teremos

  $$P(B = k) = \binom{5}{k} (0{,}05)^k(1-0{,}05)^{5-k},\ k = 0, 1, ..., 5.$$

  Verificamos que

  $$\begin{aligned} E(X) &= S\times P(B=0) + R\times P(B=1 \text{ ou } 2) - L\times P(B=3, 4 \text{ ou } 5) \\ &= S\times(0{,}95)^5 + \\ &\quad R\times[5\times(0{,}05)\times(0{,}95)^4 + 10\times(0{,}05)^2\times(0{,}95)^3] - \\ &\quad L\times[10\times(0{,}05)^3\times(0{,}95)^2+5\times(0{,}05)^4\times(0{,}95)^1+(0{,}05)^5] \end{aligned}$$

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Definição:**

  Seja X uma variável aleatória contínua com fdp *f*. O *valor esperado* de X é definido como

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

  Pode acontecer que esta integral imprópria não convirja. Consequentemente, diremos que E(X) existirá se, e somente se,

  $$\int_{-\infty}^{+\infty} |x| f(x)\,dx$$

  for finita.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Exemplo:**
  Seja a variável aleatória X definida como segue. Suponha-se que X seja o tempo (em minutos) durante o qual um equipamento elétrico seja utilizado em máxima carga, em um certo período de tempo especificado. Suponha-se que X seja uma variável aleatória contínua com a seguinte fdp:

$$f(x)=\begin{cases} x/1.500^2, & \text{se } 0\le x\le 1.500,\\ -(x-3.000)/1.500^2, & \text{se } 1.500\le x\le 3.000,\\ 0, & \text{para quaisquer outros valores.}\end{cases}$$

Portanto

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

$$=\frac{1}{(1.500)(1.500)}\left[\int_0^{1.500} x^2dx-\int_{1.500}^{3.000} x(x-3.000)dx\right]=1.500.$$

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Exemplo II:**

  O conteúdo de cinzas (em porcentagem) no carvão, digamos X, pode se considerado como uma variável aleatória contínua com fdp

  $f(x) = (1/4.875)x^2,\ 10 \leq x \leq 25.$

  Portanto,

  $E(X) = (1/4.875) \int_{10}^{25} x^3 dx = 19{,}5.$

  Assim, o conteúdo de cinzas esperado no particular espécime de carvão que está sendo estudado é de 19,5%.

# 7.1 O Valor Esperado de uma Variável Aleatória

- **Teorema:**
  Seja X uniformemente distribuída sobre o intervalo $[a, b]$. Nesse caso,
  $$E(X) = (a+b)/2.$$
- **Demonstração:**
  A fdp de X é dada por
  $$f(x) = 1/(b-a),\ a \leq x \leq b.$$
  Portanto,
  $$E(X) = \int_{-\infty}^{+\infty} x f(x) dx = 1/(b-a) \int_a^b x dx = 1/(b-a)\ x^2/2\Big|_a^b = (a+b)/2.$$
- **Observação:**
  Note que este valor representa o ponto médio do intervalo $[a, b]$, como era de se esperar intuitivamente.

# 7.2 Expectância de uma Função de uma Variável Aleatória

- **Definição:**

  Seja X uma variável aleatória e seja Y = H(X).

  a) Se Y for uma variável aleatória discreta com valores possíveis $y_1, y_2, \ldots$ e se $q(y_i) = P(Y = y_i)$, definiremos

  $$E(Y) = \Sigma^{\infty}_{i=1}\, y_i\, q(y_i).$$

  b) Se Y for uma variável aleatória contínua com fdp $g$, definiremos

  $$E(Y) = \int_{-\infty}^{+\infty} y g(y) dy.$$

- **Comentário:**

  Essas definições são coerentes com as definições anteriores, dadas para o valor esperado de uma variável aleatória. De fato, o que está acima simplesmente representa uma reformulação em termos de Y. No entanto, surge a questão de podermos obter E(Y), sem preliminarmente encontrarmos a distribuição de probabilidade de Y, partindo-se apenas do conhecimento da distribuição de X.

# 7.2 Expectância de uma Função de uma Variável Aleatória

- **Teorema:**

  Seja X uma variável aleatória e seja Y = H(X).

  a) Se X for uma variável aleatória discreta e $p(x_i) = P(X = x_i)$, teremos

  $$E(Y) = E[H(X)] = \Sigma^{\infty}_{i=1} H(x_i)p(x_i).$$

  b) Se X for uma variável aleatória contínua com fdp $f$, teremos

  $$E(Y) = E[H(X)] = \int_{-\infty}^{+\infty} H(x)f(x)dx.$$

- **Comentário:**

  Este teorema torna a avaliação de E(Y) muito mais simples, porque ele quer dizer que não necessitamos achar a distribuição de probabilidade de Y, a fim de avaliarmos E(Y); o conhecimento da distribuição de probabilidade de X é suficiente.

# 7.2 Expectância de uma Função de uma Variável Aleatória

- **Exemplo:**

  Seja V a velocidade do vento (em milhas por hora) e suponha-se que V seja uniformemente distribuída sobre o intervalo [0, 10]. A pressão, digamos W (em libras/pé-quadrado), na superfície da asa de um avião é dada pela relação W = $0{,}003V^2$. Para achar o valor esperado de W, E(W), poderemos proceder de duas maneiras:

a) Empregando o teorema anterior, teremos

$$E(W)=\int_0^{10} H(v)f(v)dv=\int_0^{10} 0{,}003v^2\,\frac{1}{10}\,dv=0{,}1.$$

b) Empregando a definição de E(W), precisaremos primeiramente achar a fdp de W, $g$, e depois calcular $\int_{-\infty}^{+\infty} wg(w)dw$. Para acharmos $g(w)$, observamos que $w = 0{,}003v^2$. é uma função monótona de $v$ para $v \geq 0$. Poderemos aplicar o teorema anteriormente visto no Cap. 5 para encontrar $g(w)$,

$$g(w)=f(v)\left|\frac{dv}{dw}\right|=\frac{1}{10}\left|\sqrt{\frac{1000}{3}}\,\frac{1}{2}\,w^{-1/2}\right|=\frac{1}{2}\sqrt{\frac{10}{3}}w^{-1/2},\, 0\leq w\leq 0{,}3,$$

$$=0,\ \text{para quaisquer outros valores.}$$

Em consequência,

$$E(W)=\int_0^{0,3} wg(w)dw=\int_0^{0,3}\frac{1}{2}\sqrt{\frac{10}{3}}w^{1/2}dw=\frac{1}{2}\sqrt{\frac{10}{3}}\left[\frac{w^{3/2}}{3/2}\right]_0^{3/10}=\frac{1}{3}\left(\frac{3}{10}\right)^{-1/2}\left(\frac{3}{10}\right)^{3/2}=\frac{1}{3}\frac{3}{10}=0{,}1.$$

# 7.4 Propriedades de Valor Esperado

- **Propriedade:**

  Seja X = C, em que C é uma constante, então

  E(X) = C.

- **Demonstração:**

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

$$= C\int_{-\infty}^{+\infty} f(x)dx = C.$$

# 7.4 Propriedades de Valor Esperado

- **Propriedade II:**

  Suponha que C seja uma constante e X seja uma variável aleatória. Então,

  E(CX) = CE(X).

- **Demonstração:**

$$E(CX) = \int_{-\infty}^{+\infty} Cxf(x)dx$$

$$= C\int_{-\infty}^{+\infty} xf(x)dx = CE(X).$$

# 7.4 Propriedades de Valor Esperado

- **Comentários:**

a) Combinando-se as propriedades anteriores, observaremos o seguinte fato importante. Se $Y = aX + b$, em que $a$ e $b$ são constantes, então

$$E(Y) = aE(X) + b.$$

Isso equivale a dizer que o valor esperado de uma função linear é a mesma função linear do valor esperado, mas isso ***não*** será verdadeiro, a menos que se trate de uma função linear. Constitui erro pensar que a propriedade seja válida para outras funções. Por exemplo, $E(X^2) \neq [E(X)]^2$.

b) Em geral, é ***difícil*** obter expressões para $E(1/X)$ ou $E(X^{1/2})$, em termos de $1/E(X)$ ou $[E(X)]^{1/2}$. Contudo, algumas desigualdades, simples de deduzir, estão disponíveis. Por exemplo, se X tomar somente valores positivos e tiver expectância finita, então

$$E(1/X) \geq 1/E(X).$$

Sob as mesmas hipóteses, temos

$$E(X^{1/2}) \leq [E(X)]^{1/2}.$$

# 7.4 Propriedades de Valor Esperado

- **Propriedade V:**
  Sejam $n$ variáveis aleatórias $X_1, ..., X_n$. Então
  $E(X_1 + ... + X_n) = E(X_1) + ... + E(X_n)$.
- **Demonstração:**
  Isto decorre imediatamente da propriedade anterior, pela aplicação da indução matemática.
- **Comentário:**
  Combinando-se esta propriedade com a anterior, obteremos

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

em que os $a_i$ são constantes.

# 7.5 A Variância de uma Variável Aleatória

- Não devemos atribuir a E(X) mais significado do que autorizado, isto é, E(X) significa apenas que se considerarmos um grande número de determinações de X, digamos, $x_1,\ldots x_n$, e calcularmos sua média, esta média estaria próximo de E(X).
- Assim, se temos E(X) = 1.000 horas, como valor esperado de uma variável aleatória X que representa a duração de uma lâmpada, então tanto podemos ter que
  - a maioria das lâmpadas dure um período compreendido entre 900 e 1.100 horas,
  - a população é compreendida por dois tipos de lâmpadas, aproximadamente a metade delas com duração de 1.300 horas e a outra metade, de 700.
- Medidas de variância serão úteis para distinguir as duas situações acima.

# 7.5 A Variância de uma Variável Aleatória

- **Definição:**

  Seja X uma variável aleatória. Definimos a *variância* de X, denotada por V(X) ou $\sigma_X 2$, da seguinte maneira

  $$V(X) = E\{[X-E(X)]^2\}.$$

  A raiz quadrada positiva de V(X) é denominada o *desvio-padrão* de X, e é denotado por $\sigma_X$.

- **Comentários:**

  a) O *número* V(X) é expresso por *unidades quadradas* de X. Este é um bom motivo para considerarmos em seu lugar o desvio-padrão, que será expresso nas *mesmas* unidades;

  b) Outra medida possível é o $E[|X-E(X)|]$, mas $X^2$ é uma função melhor comportada, por exemplo, não apresentando descontinuidade na derivada no ponto zero.

# 7.5 A Variância de uma Variável Aleatória

- **Comentários:**

  c) Se interpretarmos E(X) como o *centro de gravidade* da unidade de massa distribuída sobre uma reta, poderemos interpretar V(X) como o *momento de inércia* dessa massa, em relação a um eixo perpendicular que passe pelo centro de gravidade da massa.

  d) V(X), tal como definida anteriormente, é um caso especial da seguinte noção mais geral. O $k$-ésimo momento da variável aleatória X, em relação à sua expectância, é definida com sendo

  $$\mu_k = E\{[X-E(X)]^k\}.$$

  Para $k = 2$, obteremos a variância.

# 7.5 A Variância de uma Variável Aleatória

- **Teorema:**

  $V(X) = E(X^2) - [E(X)]^2.$

- **Demonstração:**

  Desenvolvendo $E(X^2) - [E(X)]^2$ e empregando as propriedades já estabelecidas para o valor esperado, obteremos

  $$\begin{aligned} V(X) &= E\{[X - E(X)]^2\} \\ &= E\{X^2 - 2XE(X) + [E(X)]^2\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \quad (E(X) \text{ é uma constante}) \\ &= E(X^2) - [E(X)]^2. \end{aligned}$$

# 7.6 Propriedades da Variância de uma Variável Aleatória

- **Propriedade I:**
  Se C for uma constante, então
  $V(X + C) = V(X).$
- **Demonstração:**

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

# 7.6 Propriedades da Variância de uma Variável Aleatória

- **Propriedade II:**

  Se C for uma constante, então

  $V(CX) = C^2V(X).$

- **Demonstração:**

$$
\begin{aligned}
V(Y) &= E[Y^2] - [E(Y)]^2 \\
V(CX) &= E[(CX)^2] - [E(CX)]^2 \\
&= C^2E[X^2] - C^2[E(X)]^2 \\
&= C^2\{E(X^2) - [E(X)]^2\} \\
&= C^2V(X).
\end{aligned}
$$

# 7.6 Propriedades da Variância de uma Variável Aleatória

- **Propriedade III:**

  Se (X, Y) for uma variável aleatória bidimensional, e se X e Y forem independentes, então

  $$V(X + Y) = V(X) + V(Y).$$

- **Demonstração:**

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

# 7.6 Propriedades da Variância de uma Variável Aleatória

- **Propriedade IV:**

  Sejam $X_1, \ldots, X_n$, *n* variáveis aleatórias independentes. Então

  $$V(X_1 + \ldots + X_n) = V(X_1) + \ldots + V(X_n).$$

- **Demonstração:**

  Decorre da propriedade anterior, por indução matemática.

# 7.6 Propriedades da Variância de uma Variável Aleatória

- **Propriedade V:**

  Seja X uma variável aleatória com variância finita. Então, para qualquer número real $\alpha$

  $$V(X) = E[(X - \alpha)^2] - [E(X) - \alpha]^2$$

- **Demonstração:**

  Exercício 7.36.

# 7.7 Expressões Aproximadas da Expectância e da Variância

- **Teorema:**

  Seja X uma variável aleatória com $E(X) = \mu$ e $V(X) = \sigma^2$. Suponha $Y = H(X)$. Então

  $$E(Y) \approx H(\mu) + H''(\mu)\sigma^2/2,$$

  $$V(Y) \approx [H'(\mu)]^2\sigma^2.$$

- **Demonstração:**

  A fim de estabelecer a primeira equação, desenvolveremos a função H em série de Taylor, próximo de , $x = \mu$, até dois termos. Deste modo

  $$Y = H(\mu) + (X-\mu)H'(\mu) + (X-\mu)^2H''(\mu)/2 + R_1,$$

  em que $R_1$ é o resto. Se abandonarmos o resto e tomarmos o valor esperado de ambos o membros, teremos (lembrando que $E(X-\mu)=0$)

  $$E(Y) \approx H(\mu) + H''(\mu)\sigma^2/2.$$

  A fim de estabelecer a segunda equação, desenvolveremos H em série de Taylor até um termo, para $x = \mu$, abandonamos o resto e tomamos a variância de ambos os membros, tendo

  $$V(Y) \approx [H'(\mu)]^2\sigma^2.$$

# 7.7 Expressões Aproximadas da Expectância e da Variância

- **Exemplo:**

  Sob certas condições, a tensão superficial de um líquido (dina/centímetro) será dada pela fórmula $S = 2(1-0{,}005T)^{1,2}$, na qual T é a temperatura do líquido (em graus centígrados).

  Suponha que T seja uma variável aleatória contínua com a seguinte fdp

  $$f(t) = 3.000t^{-4},\ t \geq 10,$$
  $$= 0, \text{ para quaisquer outros valores.}$$

  Daí,

  $$E(T) = \int_{10}^{\infty} 3.000t^{-3}dt = 15 \text{ graus},$$

  e

  $$V(T) = E(T^2) - (15)^2 = \int_{10}^{\infty} 3.000t^{-2}dt - 225 = 75 \text{ graus}^2.$$

# 7.7 Expressões Aproximadas da Expectância e da Variância

- **Exemplo (final):**

  Para calcularmos E(S) e V(S), teremos que calcular as seguintes integrais:

  $$E(S) = \int_{10}^{\infty} (1-0{,}005t)^{1{,}2} t^{-4} dt \text{ e } V(S) = \int_{10}^{\infty} (1-0{,}005t)^{2{,}4} t^{-4} dt.$$

  Em lugar de calcular essas expressões, obteremos aproximações de E(S) e V(S). Para utilizarmos as expressões aproximadas, teremos que calcular H'(15) e H''(15), em que $H(t) = 2(1-0{,}005t)^{1{,}2}$. Teremos

  $$H'(t) = 2{,}4(1-0{,}005t)^{0{,}2}(-0{,}005) = -0{,}012(1-0{,}005t)^{0{,}2}.$$

  Portanto, H(15) = 1,82, H'(15) = 0,01. Semelhantemente,

  $$H''(t) = -0{,}0024(1-0{,}005t)^{-0{,}8}(-0{,}005) = 0{,}000012(1-0{,}005t)^{-0{,}8}.$$

  Logo, $H''(15) = 0{,}000012(0{,}925)^{-0{,}8} = 0^{+}$. Portanto, teremos

  $$E(S) \approx H(15) + H''(15)75/2 = 1{,}82 \text{ d/cm},$$

  $$V(S) \approx [H'(15)]^2 75 = 0{,}87 \text{ (d/cm)}^2.$$

# 7.8 A Desigualdade de Tchebycheff

- Veremos como a **média** e o **desvio padrão** de uma variável aleatória nos permitem fazer algumas estimativas sobre probabilidades envolvendo essa variável;
- **Desigualdade de Jensen:**

  Seja $f:(a,b)\rightarrow R$ duas vezes diferenciável e $f''(x)\geq 0$ (função convexa) em $(a,b)$, então, para quaisquer $x_1$, $x_2,\ldots, x_n$, em $(a,b)$, então vale:

$$f[(x_1+x_2+\cdots+x_n)/n] \leq [f(x_1)+f(x_2)+\cdots+f(x_{1n})]/n$$

- **Desigualdade de Jensen II:**

  Se $X$ é uma variável aleatória e $f$ é uma função convexa, então:

$$f(\mathrm{E}[X]) \leq \mathrm{E}[f(X)]$$

# 7.8 A Desigualdade de Tchebycheff

- **Desigualdade de Markov:**
  Seja $X$ uma variável aleatória não negativa. Então vale, para todo $x > 0$:

  $$P(X \geq x) \leq E[X]/x$$

- **Demonstração:**
  Defina a variável aleatória auxiliar, $Y = x$, se $X \geq x$ e $Y = 0$, caso contrário, ou seja, se $X < x$.

  Logo deve ser verdade que $X \geq Y$ e que $E[X] \geq E[Y]$.

  Desde que $Y$ assume apenas os valores 0 e $x$, com probabilidades $P(Y = 0) = P(X < x)$ e $P(Y = x) = P(X \geq x)$. Portanto:

  $$\begin{aligned} E[X] \geq E[Y] &= 0 \cdot P(Y = 0) + x \cdot P(Y = x) \\ &= 0 \cdot P(X < x) + x \cdot P(X \geq x) \\ &= x \cdot P(X \geq x) \end{aligned}$$

# 7.8 A Desigualdade de Tchebycheff

- **Desigualdade de Tchebycheff:**

  Seja X uma variável aleatória, com $E(X) = \mu$, e seja $c$ um número real qualquer. Então, se $E(X - c)^2$ for finita e $\varepsilon$ for qualquer número positivo, teremos

  $$P[|X - c| \geq \varepsilon] \leq E(X - c)^2/\varepsilon^2.$$

  As seguintes formas *equivalentes* são imediatas

a) Se considerarmos o evento complementar, obteremos
   $P[|X - c| < \varepsilon] \geq 1 - E(X - c)^2/\varepsilon^2$.

b) Escolhendo $c = \mu$, obteremos
   $P[|X - \mu| \geq \varepsilon] \leq V(X)/\varepsilon^2$.

c) Escolhendo $c = \mu$ e $\varepsilon = k\sigma$, em que $\sigma^2 = V(X) > 0$, obteremos
   $P[|X - \mu| \geq k\sigma] \leq k^{-2}$.

# 7.8 A Desigualdade de Tchebycheff

- **Demonstração:**
  Consideremos
  $$P[|X-c|\geq \varepsilon]=\int_{x:|x-c|\geq\varepsilon} f(x)dx.$$
  O limite da integral diz que estamos integrando entre -∞ e ($c$-ε) e entre ($c$+ε) e +∞. Ora, $|x - c| \geq \varepsilon$ é equivalente a $(x - c)^2/\varepsilon^2 \geq 1$. Consequentemente, a integral acima é
  $$\leq \int_{x:|x-c|\geq\varepsilon} \frac{(x-c)^2}{\varepsilon^2} f(x)dx.$$
  Esta integral é, por sua vez,
  $$\leq \int_{-\infty}^{+\infty} \frac{(x-c)^2}{\varepsilon^2} f(x)dx, \text{que é igual a } \frac{1}{\varepsilon^2}\mathrm{E}[\mathrm{X} - \mathrm{c}]^2,$$
  como se queria demonstrar.
- **Observação:**
  A forma b) vem da igualdade de Markov, desde que $(X-\mu)^2$ é uma variável aleatória não negativa.

# 7.8 A Desigualdade de Tchebycheff

- **Comentários:**

a) É importante compreender que o resultado acima é ***notável***, pois muito pouco é suposto a respeito do comportamento probabilístico da variável aleatória X.

b) Como poderíamos suspeitar, informação adicional sobre a distribuição da variável aleatória X nos permitirá ***melhorar*** a desigualdade deduzida. Por exemplo, se C = 3/2, teremos a desigualdade de Tchebycheff

$$P[|X - \mu| \geq 3\sigma/2] \leq 4/9 = 0{,}44.$$

Entretanto, se soubermos que X é uniformemente distribuída sobre $(1-1/3^{1/2}, 1+1/3^{1/2})$, teremos que E(X) = 1, V(X) = 1/9 e, portanto,

$$\begin{aligned} P[|X - \mu| \geq 3\sigma/2] &= P[|X - 1| \geq 1/2] \\ &= 1 - P[|X - 1| < 1/2] \\ &= 1 - P[1/2 < X < 3/2] = 1 - 3^{1/2}/2 \\ &= 0{,}134. \end{aligned}$$

# 7.8 A Desigualdade de Tchebycheff

- **Teorema:**

  Suponha que $V(X) = 0$. Então, $P[X = \mu] = 1$, em que $\mu = E(X)$. Isto é, $X = \mu$, com probabilidade 1.

- **Demonstração:**

  Da desigualdade de Tchebycheff, encontramos que

  $$P[|X - \mu| \geq \varepsilon] = 0, \text{ para qualquer } \varepsilon > 0.$$

  Por conseguinte,

  $$P[|X - \mu| < \varepsilon] = 1, \text{ para qualquer } \varepsilon > 0.$$

  Como $\varepsilon$ pode ser escolhido arbitrariamente pequeno, o teorema fica ***demonstrado***.

- **Comentários:**

  a) Este teorema mostra que variância nula implica que toda probabilidade está concentrada em um *único* ponto, a saber, em $E(X)$;

  b) Se $E(X) = 0$, então $V(X) = E(X^2)$, e por isso, nesse caso, $E(X^2) = 0$ acarreta a *mesma* conclusão;

  c) É no sentido acima que dizemos uma variável X é *degenerada*.