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 $x_1, \ldots, x_n, \ldots$. Seja $p(x_i) = P(X = x_i)$, $i = 1, 2, \ldots, n, \ldots$. 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

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

  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

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

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

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

  $f(x) = 1/(b-a),\ a \le x \le b.$

  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

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

  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

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

  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 = 0,003V². 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,

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

7.3 Variáveis Aleatórias Bidimensionais

- **Definição:**

  Seja (X, Y) uma variável aleatória bidimensional e seja Z = H(X, Y) uma função real de (X, Y). Em consequência, Z será uma variável aleatória (unidimensional) e definiremos E(Z) como se segue:

a) Se Z for uma variável aleatória discreta, com valores possíveis $z_1, z_2, \ldots$ e com $p(z_i) = P(Z = z_i)$, então

$$E(Z) = \Sigma^{\infty}_{i=1} z_i\, p(z_i);$$

b) Se Z for uma variável aleatória contínua com fdp f, definiremos

$$E(Z) = \int_{-\infty}^{+\infty} z f(z)\, dz.$$

- **Comentário:**

  Como no caso unidimensional, o seguinte teorema pode ser demonstrado.

7.3 Variáveis Aleatórias Bidimensionais

- **Teorema:**

  Seja (X, Y) uma variável aleatória bidimensional e façamos Z = H(X, Y).

a) Se (X, Y) for uma variável aleatória discreta e se

$$p(x_i, y_j) = P(X = x_i, Y = y_j),\ i, j = 1,2, \ldots,$$

teremos

$$E(Z) = \Sigma^{\infty}_{j=1} \Sigma^{\infty}_{i=1} H(x_i, y_j)p(x_i, y_j).$$

b) Se (X, Y) for uma variável aleatória contínua com fdp conjunta f, teremos

$$E(Z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} H(x, y)f(x, y)dx\, dy.$$

7.3 Variáveis Aleatórias Bidimensionais

- **Exemplo:**

  Vamos reconsiderar exemplo anterior e achar E(V), em que V = RI. Já vimos que R e I eram variáveis aleatórias com as seguintes fdp *h* e *g*, respectivamente:

  $h(r) = r^2/9$, $0 \leq r \leq 3$ e $g(i) = 2i$, $0 \leq i \leq 1$.

  Também encontramos que a fdp de V era $p(v) = (2/9)v(3-v)$, $0 \leq v \leq 3$. Visto que R e I são variáveis aleatórias independentes, a fdp conjunta de (R, I) é simplesmente o produto das fdp de R e I. Para Calcular E(V), empregando o teorema, teremos

$$E(V) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} H(r,i)\,f(r,i)\,dr\,di = \int_0^1\int_0^3 ri\,\frac{r^2}{9}\,2i\,dr\,di$$

$$= \frac{2}{9}\int_0^3 r^3 dr \int_0^1 i^2 di = \frac{2}{9}\left[\frac{r^4}{4}\right]\Bigg|_0^3 \left[\frac{i^3}{3}\right]\Bigg|_0^1 = \frac{2}{9}\left[\frac{81}{4}\right]\left[\frac{1}{3}\right] = \frac{3}{2}.$$

Empregando diretamente a definição, teremos o mesmo resultado.

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

- **Propriedade III:**

  Seja (X, Y) uma variável aleatória bidimensional com uma distribuição de probabilidade conjunta *f*. Seja $Z = H_1(X, Y)$ e $W = H_2(X, Y)$. Então,

  E(Z+W) = E(Z) + E(W).

- **Demonstração:**

$$E(Z+W) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[H_1(x,y)+H_2(x,y)]\,f(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_1(x,y)\,f(x,y)\,dx\,dy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_2(x,y)\,f(x,y)\,dx\,dy$$

$$= E(Z)+E(W).$$

7.4 Propriedades de Valor Esperado

- **Propriedade IV:**

  Sejam X e Y duas variáveis aleatória quaisquer. Então,

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

- **Demonstração:**

  Decorre diretamente da propriedade anterior, ao se fazer

  $$H_1(X, Y) = X$$

  e

  $$H_2(X, Y) = Y.$$

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.

7.4 Propriedades de Valor Esperado

- **Propriedade VI:**

  Seja (X, Y) uma variável aleatória bidimensional e suponha-se que X e Y sejam independentes. Então

  E(XY) = E(X) E(Y).

- **Demonstração:**

$$E(XY) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyg(x)h(y)dxdy$$

$$= \int_{-\infty}^{+\infty} xg(x)dx \int_{-\infty}^{+\infty} yh(y)dy = E(X)E(Y).$$

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 σ_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 σ_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 α

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

- **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:

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

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

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

7.7 Expressões Aproximadas da Expectância e da Variância

- **Teorema II:**

  Seja (X, Y) uma variável aleatória bidimensional. Suponha que $E(X) = \mu_x$, $E(Y) = \mu_y$, $V(X) = \sigma_x^2$, $V(Y) = \sigma_y^2$. Seja $Z = H(X, Y)$ [deve-se admitir que as várias derivadas de H existam para (μ_x, μ_y)]. Então, se X e Y forem independentes, ter-se-á

  $$E(Z) \approx H(\mu_x, \mu_y) + [\sigma_x^2 \times \partial^2 H/\partial x^2 + \sigma_y^2 \times \partial^2 H/\partial y^2]/2,$$

  $$V(Z) \approx [\partial H/\partial x]^2 \sigma_x^2 + [\partial H/\partial y]^2 \sigma_y^2,$$

  em que todas as derivadas parciais são calculadas para (μ_x, μ_y).

- **Demonstração:**

  A demonstração envolve o desenvolvimento de H em série de Taylor, no ponto (μ_x, μ_y), para um e dois termos, o abandono do resto, e, a seguir, o cálculo da expectância e da variância de ambos os membros, tal como foi feito na demonstração do teorema anterior.

7.7 Expressões Aproximadas da Expectância e da Variância

- **Exemplo:**

  Admita-se que temos um circuito simples, para o qual a tensão M seja expressa pela Lei de Ohm com M = IR, em que I e R são, respectivamente, a corrente e a resistência do circuito. Se I e R forem variáveis aleatórias independentes, então M será uma variável aleatória, e empregando o teorema, poderemos escrever

$$
\begin{aligned}
E(M) &\approx H(\mu_x, \mu_y) + [\sigma_x^2 \times \partial^2 H/\partial x^2 + \sigma_y^2 \times \partial^2 H/\partial y^2]/2 \\
&\approx E(I)E(R)
\end{aligned}
$$

e

$$
\begin{aligned}
V(M) &\approx [\partial H/\partial x]^2 \sigma_x^2 + [\partial H/\partial y]^2 \sigma_y^2 \\
&\approx [E(R)]^2 V(I) + [E(I)]^2 V(R).
\end{aligned}
$$

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,..., x_n, em (a,b), então vale:

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

- **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 ε 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 $-\infty$ e $(c-\varepsilon)$ e entre $(c+\varepsilon)$ e $+\infty$. 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

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

  Por conseguinte,

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

  Como ε 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*.

7.9 O Coeficiente de Correlação

- **Definição:**

  Seja (X, Y) uma variável aleatória bidimensional. Definiremos ρ_{XY}, o coeficiente de correlação entre X e Y, da seguinte forma

$$\rho_{xy} = \frac{E\{[X - E(X)][Y - E(Y)]\}}{\sqrt{V(X)V(Y)}}.$$

- **Comentários:**

a) Supomos que todas as esperanças matemáticas existam e que ambas V(X) e V(Y) sejam não-nulas. Quando houver dúvida quanto às variáveis, simplesmente escreveremos ρ em vez de ρ_{xy};

b) O numerador de ρ, E{[X-E(X)][Y-E(Y)]}, é denominado ***covariância*** entre X e Y, e algumas vezes é denotado por σ_{xy};

c) O coeficiente de correlação é uma quantidade ***adimensional***;

d) Antes que a definição acima fique realmente expressiva, devemos descobrir exatamente o que ρ mede. Faremos isso pela consideração de algumas ***propriedades*** de ρ.

7.9 O Coeficiente de Correlação

- **Teorema:**

$$\rho = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}}.$$

- **Demonstração:**

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

7.9 O Coeficiente de Correlação

- **Teorema II:**

  Se X e Y forem independentes, então $\rho = 0$.

- **Demonstração:**

  Isto decorre imediatamente do teorema anterior, porque

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

  se X e Y forem independentes.

- **Comentário:**

  A recíproca do teorema em geral ***não*** é verdadeira. Isto é, podemos ter $\rho = 0$, e no entanto X e Y não precisam ser independentes. Se $\rho = 0$, diremos que X e Y são ***não correlacionadas***. Portanto, ser não correlacionado e ser independente, em geral, não são equivalentes.

7.9 O Coeficiente de Correlação

- **Exemplo:**

  Sejam X e Y duas variáveis aleatórias quaisquer que tenham a mesma distribuição. Sejam

  $U = X - Y$ e $V = X + Y$.

  Então,

  $E(U) = 0$ e $\sigma(U, V) = E[(X - Y)(X + Y)] = E(X^2 - Y^2) = 0$.

  Portanto, U e V são não correlacionadas.

  Ainda que X e Y sejam independentes, U e V podem ser dependentes, como a seguinte escolha de X e Y indica. Sejam X e Y os números que aparecem no primeiro e no segundo dados, respectivamente, que tenham sido jogados. Nós agora achamos, por exemplo, que $P[V = 4 | U = 3] = 0$ (uma vez que $X - Y = 3$, $X + Y$ não pode ser igual a 4), enquanto $P(V = 4) = 3/36$. Portanto, U e V são dependentes.

7.9 O Coeficiente de Correlação

- **Teorema III:**

  Sejam duas variáveis aleatórias X e Y, então $-1 \leq \rho_{xy} \leq 1$. Isto é, ρ_{xy} toma valores entre -1 e +1, inclusive.

- **Demonstração:**

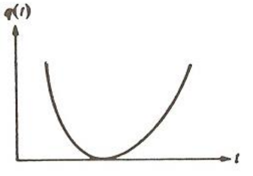
  Considere-se a seguinte função da variável real t

  $$q(t) = E[V + tW]^2,$$

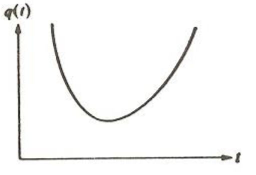
  em que $V = X - E(X)$ e $W = Y - E(Y)$. Visto que $[V + tW]^2 \geq 0$, temos que $q(t) \geq 0$ para todo t. Desenvolvendo, obteremos

  $$q(t) = E[V^2 + 2tVW + t^2W^2] = E(V^2) + 2tE(VW) + t^2E(W^2).$$

  Deste modo, vemos que $q(t)$ é uma expressão quadrática de t. Em geral, se uma expressão quadrática, $q(t) = at^2 + bt + c$, tem a propriedade de que $q(t) \geq 0$ para todo t, isto então significa que seu gráfico toca o eixo dos t em apenas um ponto, ou não toca (vide figuras).



(a)



(b)

7.9 O Coeficiente de Correlação

- **Demonstração (final):**

  Isto por sua vez significa que seu discriminante $b^2 - 4ac$ deve ser ≤ 0, pois se $b^2 - 4ac > 0$ significaria que $q(t)$ teria duas raízes reais distintas. Aplicando-se esta conclusão à função $q(t)$, que estávamos considerando acima, obteremos

  $$4[E(VW)]^2 - 4E(V^2)E(W^2) \leq 0.$$

  Isto acarreta

  $$[E(VW)]^2/E(V^2)E(W^2) \leq 1,$$

  e daí

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

  Portanto, $-1 \leq \rho \leq 1$, como queríamos demonstrar.

7.9 O Coeficiente de Correlação

- **Teorema IV:**

  Suponha que $\rho^2 = 1$. Portanto (com probabilidade 1, no sentido do teorema anteriormente estudado), Y = AX + B, em que A e B são constantes. Em palavras, se o coeficiente de correlação ρ for ± 1, então Y será uma função linear de X (com probabilidade 1).

- **Demonstração:**

  Considere-se novamente a função $q(t)$, descrita na demonstração do teorema anterior. Daí, $\rho^2 = 1$ acarreta que deve existir ao menos um valor de t, digamos t_0, tal que

  $$q(t_0) = E(V + t_0W)^2 = 0.$$

  Uma vez que $V + t_0W = [X - E(X)] + t_0[Y - E(Y)]$, temos que $E(V + t_0W) = 0$ e, portanto, a variância $(V + t_0W) = E(V + t_0W)^2$.

  Assim, encontramos que a hipótese do teorema leva à conclusão de que a variância de $(V + t_0W) = 0$. Consequentemente, de teorema anterior, podemos concluir que a variável aletória

  $$(V + t_0W) = 0 \text{ (com probabilidade 1).}$$

  Logo, $[X - E(X)] + t_0[Y - E(Y)] = 0$. Reescrevendo isto, encontramos que

  $$Y = AX + B \text{ (com probabilidade 1),}$$

  como queríamos demonstrar.

7.9 O Coeficiente de Correlação

- **Teorema V:**

  Suponha que X e Y sejam duas variáveis aleatórias, para as quais $Y = AX + B$, em que A e B são constantes. Então $\rho^2 = 1$. Se $A > 0$, $\rho = +1$; se $A < 0$, $\rho = -1$.

- **Demonstração:**

  Visto que $Y = AX + B$, teremos $E(Y) = AE(X) + B$ e $V(Y) = A^2V(X)$. Também, $E(XY) = E[X(AX + B)] = AE(X^2) + BE(X)$. Portanto

$$
\begin{aligned}
\rho^2 &= [E(XY) - E(X)E(Y)]^2/V(X)V(Y) \\
&= \{AE(X^2) + BE(X) - E(X)[AE(X)+B]^2\}^2/V(X)A^2V(X) \\
&= \{AE(X^2) + BE(X) - A[E(X)]^2 - BE(X)\}^2/A^2V(X)^2 \\
&= A^2\{E(X^2) - [E(X)]^2\}^2/A^2[V(X)]^2 \\
&= 1.
\end{aligned}
$$

7.9 O Coeficiente de Correlação

- **Comentário:**

  Os dois teoremas anteriores estabelecem a seguinte importante característica do coeficiente de correlação.

  O coeficiente de correlação é uma medida do ***grau de linearidade*** entre X e Y. Valores de ρ próximos de +1 ou -1 indicam um alto grau de linearidade, enquanto valores de ρ próximos de 0 indicam falta de tal linearidade.

  Valores positivos de ρ mostram que Y tende a crescer com o crescimento de X, enquanto valores negativos de ρ mostram que Y tende a decrescer com valores crescentes de X.

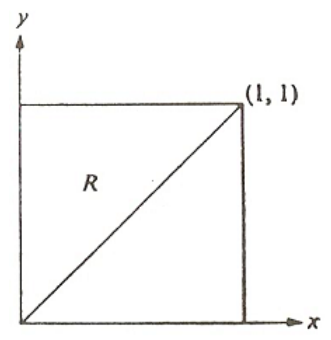
  Existe muito equívoco sobre a interpretação do coeficiente de correlação. Um valor de ρ próximo de zero indica apenas ausência de relação ***linear*** entre X e Y. Ele não elimina a possibilidade de alguma relação ***não-linear***.

7.9 O Coeficiente de Correlação

- **Exemplo:**

  Suponha-se que a variável aleatória bidimensional (X, Y) seja uniformemente distribuída sobre a região triangular R (veja figura)

  $R = \{(x, y) | \ 0 < x < y < 1\}$.

  Consequentemente, a fdp é dada como

  $f(x, y) = 2, \ (x, y) \in R,$

  $\qquad = 0$, para outros quaisquer valores.

  Por conseguinte, as fdp marginais de X e Y são

  $g(x) = \int_x^1 2dy = 2(1-x), \qquad 0 \leq x \leq 1;$

  $h(y) = \int_0^y 2dx = 2y, \qquad 0 \leq y \leq 1.$

7.9 O Coeficiente de Correlação

- **Exemplo:**

Logo

$$E(X) = \int_0^1 x2(1-\underline{x})dx = 1/3; \quad E(Y) = \int_0^1 y2ydy = 2/3;$$

$$E(X^2) = \int_0^1 x^2 2(1-\underline{x})dx = 1/6; \quad E(Y^2) = \int_0^1 y^2 2ydy = ½;$$

$$V(X) = E(X^2) – [E(X)]^2 = 1/18;$$

$$V(Y) = E(Y^2) – [E(Y)]^2 = 1/18;$$

$$E(XY) = \int_0^1 \int_0^y xy2dxdy = ¼.$$

Portanto,

$$\rho = \{E(XY) – E(X)E(Y)\}/[V(X)V(Y)^{1/2}] = ½.$$

7.9 O Coeficiente de Correlação

- **Teorema VI:**

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

  $$V = AX + B \text{ e } W = CY + D,$$

  em que A, B, C e D são constantes, então

  $$\rho_{vw} = (AC/|AC|)\rho_{xy},$$

  supondo que $A \neq 0$ e $C \neq 0$.

7.10 Valor Esperado Condicionado

- **Definição:**

a) Se (X, Y) for uma variável aleatória ***contínua*** bidimensional, definiremos o ***valor esperado condicionado*** de X, para um dado Y = y, como sendo

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

b) Se (X, Y) for uma variável aleatória ***discreta*** bidimensional, definiremos o ***valor esperado condicionado*** de X, para um dado Y = y_j, como sendo

$$E(X|y_j) = \sum_{i=1}^{\infty} x_i p(x_i|y_j).$$

7.10 Valor Esperado Condicionado

- **Comentários:**

a) A interpretação do valor esperado condicionado é a seguinte. Desde que $g(x|y)$ representa a fdp condicionada de X para um dado Y = y, E(X| y) é o valor esperado de X, condicionado ao evento {Y = y}.

Por exemplo, se (X, Y) representar o esforço de tração e a dureza de um espécime de aço, respectivamente, então E(X| Y = 52,7) será o esforço de tração esperado de um espécime de aço escolhido ao acaso da população com dureza 52,7.

7.10 Valor Esperado Condicionado

- **Comentários (final):**
  b) É importe compreender que, em geral, E(X| *y*) é uma função de *y* e, por isso, é uma ***variável aleatória***. Semelhantemente, E(Y| *x*) é uma função de *x* e, também, é uma variável aleatória [estritamente falando, E(X| *y*) é o valor da variável aleatória E(X| Y)].

  c) Porque E(Y| X) e E(X| Y) são variáveis aleatórias, terá sentido falar de seus valores esperados. Deste modo, poderemos considerar E[E(X| Y)], por exemplo.

7.10 Valor Esperado Condicionado

- **Teorema:**

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

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

- **Demonstração (caso contínuo):**

  Por definição,

  $$E(X|y) = \int_{-\infty}^{+\infty} x g(x|y)dx = \int_{-\infty}^{+\infty} x[f(x, y)/h(y)]dx,$$

  em que f é a fdp conjunta de (X, Y) e h a fdp marginal de Y. Por isso

  $$E[E(X|\ Y)] = \int_{-\infty}^{+\infty} E(X|\ y)h(y)\ dy$$
  $$= \int_{-\infty}^{+\infty} [\int_{-\infty}^{+\infty} x[f(x, y)/h(y)]dx]h(y)dy.$$

  Se existirem todos os valores esperados, será permitido trocar a ordem da integração acima. Por isso

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

7.10 Valor Esperado Condicionado

- **Exemplo:**

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

n:	10	11	12	13	14	15
$P(N = n)$	0,05	0,10	0,10	0,20	0,35	0,20

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

  Para um dado N igual a *n*, X apresentará distribuição binomial. Como N, por sua vez, é uma variável aleatória, procederemos com se segue.

7.10 Valor Esperado Condicionado

- **Exemplo (final):**

  Temos $E(X) = E[E(X|\ N)]$. Contudo, $E(X|\ N) = 0{,}10N$, porque para N dado, X tem distribuição binomial. Consequentemente,

$$
\begin{aligned}
E(X) &= E(0{,}10N) = 0{,}10E(N) \\
&= 0{,}10[10(0{,}05) + 11(0{,}10) + 12(0{,}10) + 13(0{,}20) + 14(0{,}35) + 15(0{,}20)] \\
&= 1{,}33.
\end{aligned}
$$

7.10 Valor Esperado Condicionado

- **Teorema II:**

  Suponha que X e Y sejam variáveis aleatórias independentes. Então

  $$E(X|\,Y) = E(X)$$

  e

  $$E(Y|\,X) = E(Y).$$

- **Demonstração:**

  Deixada como exercício.

7.10 Valor Esperado Condicionado

- **Exemplo:**

  Suponha que o fornecimento energético (KW) a uma companhia hidrelétrica, durante um período especificado, seja uma variável aleatória X, a qual admitiremos ter uma distribuição uniforme sobre [10, 30]. A demanda de potência é também uma variável aleatória Y, que admitiremos ser uniformemente distribuída sobre [10, 20] (note que em média mais potência é fornecida, E(X) = 20, do que é demandada, E(Y) = 15).

  Para cada KW fornecido, a companhia realiza um lucro de US$ 0,03. Se a demanda exceder a oferta, a companhia obterá potência adicional de outra fonte, realizando um lucro de US$ 0,01 por KW desta potência fornecida. Qual será o lucro esperado, durante o especificado período considerado?

7.10 Valor Esperado Condicionado

- **Exemplo (continuação):**

  Seja T este lucro. Teremos

  $$T = \begin{cases} 0{,}03Y, & \text{se } Y < X, \\ 0{,}03\,X + 0{,}01(Y - X), & \text{se } Y > X. \end{cases}$$

  Para calcular E(T), o escreveremos como $E[E(T|\,X)]$.

7.10 Valor Esperado Condicionado

- **Exemplo (continuação):**

  Teremos

$$E(T\mid x)=\begin{cases}\int_{10}^{x}(0{,}03y)\frac{1}{10}\,dy+\int_{x}^{20}\left[0{,}02x+0{,}01y\right]\frac{1}{10}\,dy, & \text{se } 10<x<20,\\ \int_{10}^{20}0{,}03y\,\frac{1}{10}\,dy, & \text{se } 20<x<30,\end{cases}$$

$$=\begin{cases}\frac{1}{10}\left[0{,}015(x^2-100)+0{,}02x(20-x)+0{,}005(20^2-x^2)\right], & \text{se } 10<x<20,\\ \frac{9}{20}, & \text{se } 20<x<30,\end{cases}$$

$$=\begin{cases}0{,}05+0{,}04x-0{,}001x^2, & \text{se } 10<x<20,\\ 0{,}45, & \text{se } 20<x<30,\end{cases}$$

7.10 Valor Esperado Condicionado

- **Exemplo (final):**

Logo, da equação 7.6

$$E\left[\underbrace{E(T \mid X)}_{H(X)}\right] = \int_{-\infty}^{+\infty} H(x)g(x)dx, \text{com } g(x) = 1/20,$$

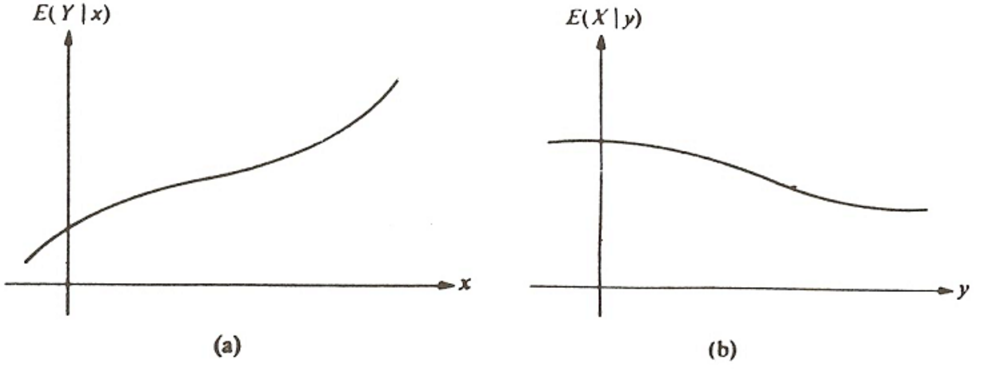
temos

$$E[E(T \mid X)] = \frac{1}{20}\int_{10}^{20}(0{,}05 + 0{,}04x - 0{,}001x^2)dx + \frac{1}{20}\int_{20}^{30} 0{,}45dx$$

$$= \text{US\$ } 0{,}43.$$

7.11 Regressão da Média

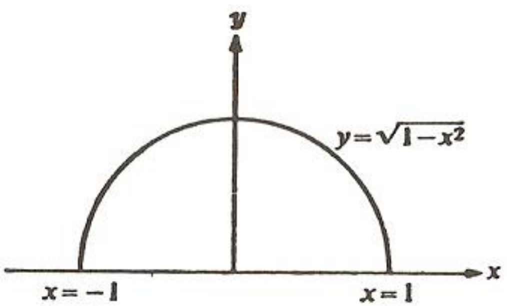
- $E(X|\, y)$ é o valor da variável aleatória $E(X|\, Y)$ e é *uma função de y*;
- O gráfico desta função de y é conhecido como a *curva de regressão* (da média) de X em Y;
- Analogamente, o gráfico da função de x, $E(Y|\, x)$ é denominado a *curva de regressão* (da média) de Y em X;
- Para cada valor fixado y, $E(X|\, y)$ será o valor esperado da variável aleatória (unidimensional) cuja distribuição de probabilidade é dada por $p(x_i|\, y_i)$, caso discreto, ou $g(x|\, y)$, caso contínuo (ver figura);
- Em geral, este valor esperado dependerá de y.



7.11 Regressão da Média

- **Exemplo:**

Suponha que (X, Y) seja uniformemente distribuída sobre o semicírculo indicado na ***figura***. Nesse caso, $f(x, y) = 2/\pi$, $(x, y) \in$ semicírculo. Portanto



$$g(x) = \int_0^{\sqrt{1-x^2}} \frac{2}{\pi} dy = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1;$$

$$h(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{2}{\pi} dx = \frac{4}{\pi}\sqrt{1-y^2}, \quad 0 \le y \le 1.$$

Logo

$$g(x \mid y) = \frac{1}{2\sqrt{1-y^2}}, -\sqrt{1-y^2} \le x \le \sqrt{1-y^2};$$

$$h(y \mid x) = \frac{1}{\sqrt{1-x^2}}, \quad 0 \le y \le \sqrt{1-x^2}.$$

7.11 Regressão da Média

- **Exemplo (final):**

  Em consequência

$$E(Y \mid x) = \int_0^{\sqrt{1-x^2}} y h(y \mid x) dy = \int_0^{\sqrt{1-x^2}} y \frac{1}{\sqrt{1-x^2}} dy$$

$$= \frac{1}{\sqrt{1-x^2}} \frac{y^2}{2} \Bigg|_0^{\sqrt{1-x^2}} = \frac{1}{2}\sqrt{1-x^2}.$$

Semelhantemente

$$E(X \mid y) = \int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} x g(x \mid y) dx = \int_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} x \frac{1}{2\sqrt{1-y^2}} dx$$

$$= \frac{1}{2\sqrt{1-y^2}} \frac{x^2}{2} \Bigg|_{-\sqrt{1-y^2}}^{+\sqrt{1-y^2}} = 0.$$

7.11 Regressão da Média

- **Exemplo II:**

  Suponha que (X, Y) seja uniformemente distribuída sobre o triângulo indicado na ***figura***.

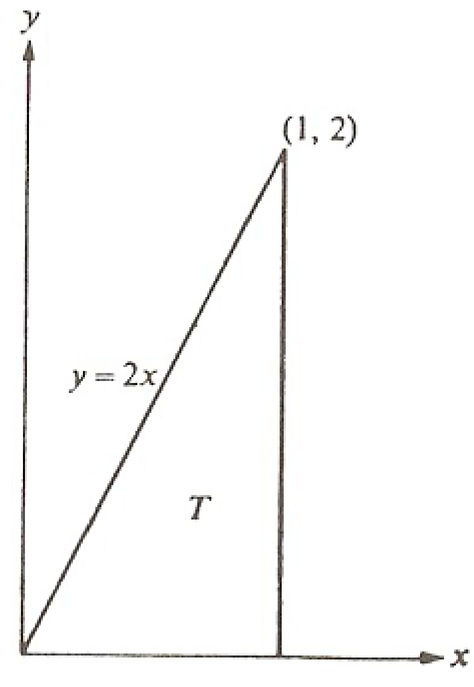
  Nesse caso, $f(x, y) = 1$, $(x, y) \in$ T. As seguinte expressões para as fdp marginal e condicionada são facilmente ***encontradas***

  $g(x) = 2x, \quad 0 \le x \le 1;$

  $h(y) = (2-y)/2, \quad 0 \le y \le 2;$

  $g(x|y) = 2/(2-y), \quad y/2 \le x \le 1;$

  $h(y|x) = 1/(2x), \quad 0 \le y \le 2x;$



7.11 Regressão da Média

- **Exemplo II :**

  Portanto

  $$E(Y|\,x) = \int_0^{2x} y h(y|x)\,dy = \int_0^{2x} y(1/2x)\,dy$$
  $$= x.$$

  Semelhantemente

  $$E(X|\,y) = \int_{y/2}^{1} x g(x|y)\,dx = \int_{y/2}^{1} x2/(2-y)\,dx$$
  $$= y/4 + 1/2.$$

  Por conseguinte, ambas as regressões de Y em X e de X em Y são lineares.

7.11 Regressão da Média

- **Teorema:**

  Seja (X, Y) uma variável aleatória bidimensional e suponha-se que

  $E(X) = \mu_x$; $E(Y) = \mu_y$; $V(X) = \sigma_x^2$; $V(Y) = \sigma_y^2$.

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

  $E(Y|\, x) = \mu_y + \rho(\sigma_y/\sigma_x)(x - \mu_x)$.

  Se a regressão de X em Y for linear, teremos

  $E(X|\, y) = \mu_x + \rho(\sigma_x/\sigma_y)(y - \mu_y)$.

- **Demonstração:**

  (deixada como exercício)

7.11 Regressão da Média

- **Comentários:**
  a) Como foi sugerido pela exposição acima, é possível que ***uma*** das regressões da média seja linear, enquanto a outra não o seja;
  b) Observe-se que papel decisivo representado pelo coeficiente de correlação nas expressões acima. ***Se*** a regressão de X em Y, por exemplo, for linear, e se $\rho = 0$, então verificamos que $E(X|\, y)$ não depende de y. Observe-se também que o sinal algébrico de ρ determina o sinal da declividade da rede de regressão.
  c) Se ambas as funções de regressão forem lineares, verificaremos, pela resolução simultânea das equações anteriores que as retas de regressão se interceptam no "centro" da distribuição, (μ_x, μ_y).

7.11 Regressão da Média

- **Observações:**
  1. As funções de regressão não precisam ser lineares, mas ainda assim podemos estar interessados em ***aproximar*** a curva de regressão com uma função linear.
  2. A aproximação é feita usualmente recorrendo-se ao ***princípio dos mínimos quadrados***, o que neste contexto é assim enunciado

     Escolham-se as constantes a e b de modo que a esperança abaixo se torne mínima

     $$E[E(Y|X) - (aX + b)]^2.$$

     Semelhantemente, escolhem-se constantes c e d, de modo que a esperança abaixo se torne mínima

     $$E[E(X|Y) - (cY + d)]^2.$$
  3. As retas $y = ax + b$ e $x = cy + d$ são denominadas as ***aproximações de mínimos quadrados*** às correspondentes curvas de regressão $E(Y|x)$ e $E(X|y)$, respectivamente.

7.11 Regressão da Média

- **Teorema II:**

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

  $$E(Y|\, x) = a'x + b',$$

  então $a = a'$ e $b = b'$. Enunciado análogo vale para a regressão de X em Y.

- **Demonstração:**

  (deixada como exercício)