2
Fundamentos de probabilidade e estatística

2.1
Introdução

A história da estatística pode ser dividida em três fases. De acordo com PESSANHA (2001), a estatística inicialmente não mantinha nenhuma relação com a probabilidade, estando direcionada apenas à contagem. Já, na segunda fase as investigações eram efetuadas com respeito aos fenômenos coletivos, a partir de então, uniu-se o estudo da estatística ao cálculo das probabilidades. Finalmente, na fase que se estende até hoje, o conhecimento científico faz com que a estatística e a probabilidade sejam amplamente utilizadas.

Neste capítulo apresentam-se alguns conceitos sobre probabilidade e estatística. Tais conceitos serão utilizados nos capítulos seguintes, tanto para o estudo da análise de estruturas incorporando incertezas, quanto para a formulação do problema de otimização que são imprescindíveis para o entendimento e o desenvolvimento do trabalho. Conceitos como população, amostra, espaço amostral, eventos, histogramas de freqüência, variáveis aleatórias contínuas e discretas além de outros referentes à estatística serão abordados num primeiro momento. Posteriormente serão apresentados alguns conceitos referentes à probabilidade, tais como: médias, desvio padrão, variância e covariância, funções de densidade de probabilidade, esperança matemática entre outros.

2.2
Estatística

De acordo com SPIEGEL (1978), a estatística trata dos métodos científicos para coleta, organização, resumo, apresentação e análise de dados, bem como da obtenção de conclusões válidas e na tomada de decisões razoáveis baseadas e fundamentadas em tais análises.

2.3
População e amostra

Ao se coletar dados referentes às características de um grupo de objetos ou indivíduos, é muitas vezes impossível ou impraticável se observar todo o grupo, em especial, se este for muito grande. Para todo o grupo, dá-se o nome de população, a uma parte ou parcela do mesmo, dá-se o nome de amostra.

Sendo a amostra uma parte representativa de uma população, conclusões importantes sobre a população podem sem inferidas de sua análise. Estatística indutiva ou inferência estatística é a parte da estatística que trata das condições sob as quais essas inferências são ou não válidas. Como essas inferências não podem ser absolutamente corretas, usa-se dos conceitos da probabilidade para o estabelecimento das conclusões.

Chama-se estatística descritiva ou dedutiva a parte da estatística que procura somente descrever e analisar um certo grupo, sem apontar ou determinar quaisquer conclusões sobre este.

2.4
Espaço amostral, eventos e variáveis aleatórias

Um conjunto que consiste de todos os resultados possíveis de um determinado experimento aleatório é chamado de espaço amostral. Cada resultado é um ponto amostral.

Um evento é um subconjunto do espaço amostral, é um conjunto de resultados possíveis. Se o resultado de um experimento é elemento deste subconjunto, diz-se então, que o evento ocorreu. Por sua vez, um evento que consiste de um único ponto é chamado de evento simples ou evento elementar.

Segundo HART (1982), o conceito de variável aleatória está intimamente ligado às condições do experimento. Se um determinado experimento é realizado repetidamente, com todas as condições precisamente mantidas, e se nessas condições obtém-se resultados idênticos, as variáveis avaliadas são ditas determinísticas. Entretanto, se os resultados numéricos variam, as variáveis avaliadas são ditas aleatórias.

Uma variável aleatória que toma um valor finito é chamada variável aleatória discreta, ao passo que uma variável aleatória que toma um número infinito de valores é uma variável aleatória contínua.

2.5
Distribuições de freqüência

A coleta de dados resultantes de um experimento é apenas o primeiro passo do estudo da estatística. Costuma-se freqüentemente distribuir esses dados em classes ou categorias e determinar o número de indivíduos pertencentes a cada uma das classes, caracterizando-se então, uma freqüência da classe. Um arranjo tabular dos dados por classe, juntamente com as freqüências correspondentes, é denominado distribuição de freqüência ou tabela de freqüência. Essa forma de organização de dados pode ser bastante útil no momento da interpretação dos resultados de um experimento. A tabela (2.1) representa a distribuição de freqüência para o módulo de elasticidade secante de 100 corpos de prova de concreto, sendo eles dispostos em 25 grupos de acordo com um intervalo de classe escolhido, HART (1982).

2.6
Intervalos e limites de classes

O símbolo que define uma classe chama-se intervalo de classe. Na tabela (2.1) tem-se, por exemplo, o intervalo de classe 1,59 – 1,65. Os números extremos desse intervalo são denominados limites de classe. Um intervalo de classe que, ao menos teoricamente, não possui limite superior ou inferior é dito intervalo de classe aberto. Os conceitos de limites de classe são importantes para determinação de histogramas de freqüência.

Tabela 2.1: Exemplo de distribuição de freqüência

Módulo de elasticidade secante no intervalo ($X10^4$ MPa)	Número de observações no intervalo	Freqüência de ocorrências
Abaixo de 1,58	0	0.00
1,58-1,65	1	0,01
1,66-1,72	0	0,00
1,73-1,79	0	0,00
1,80-1,86	0	0,00
1,87-1,92	1	0,01
1,93-1,99	1	0,01
2,00-2,13	4	0,04
2,14-2,20	8	0,08
2,21-2,27	3	0,03
2,28-2,34	10	0,10
2,35-2,41	9	0,09
2,42-2,48	9	0,09
2,49-2,54	16	0,16
2,55-2,61	7	0,07
2,62-2,68	11	0,11
2,69-2,82	6	0,06
2,83-2,89	7	0,07
2,90-3,03	0	0,00
3,04-3,10	0	0,00
3,11-3,16	2	0,02
3,17-3,23	1	0,01
3,24-3,30	1	0,01
3,31-3,37	2	0,02
3,38-3,44	0	0,00
3,45-3,51	1	0,01
Acima de 3,51	0	0,00

2.7 Histogramas de freqüência

Histograma de freqüência é a representação gráfica de uma distribuição de freqüência, ou seja, é a representação gráfica que indica a freqüência relativa de um valor observado de uma variável aleatória entre os limites de uma classe. A figura (2.1) representa o histograma de freqüência para os dados da tabela (2.1).

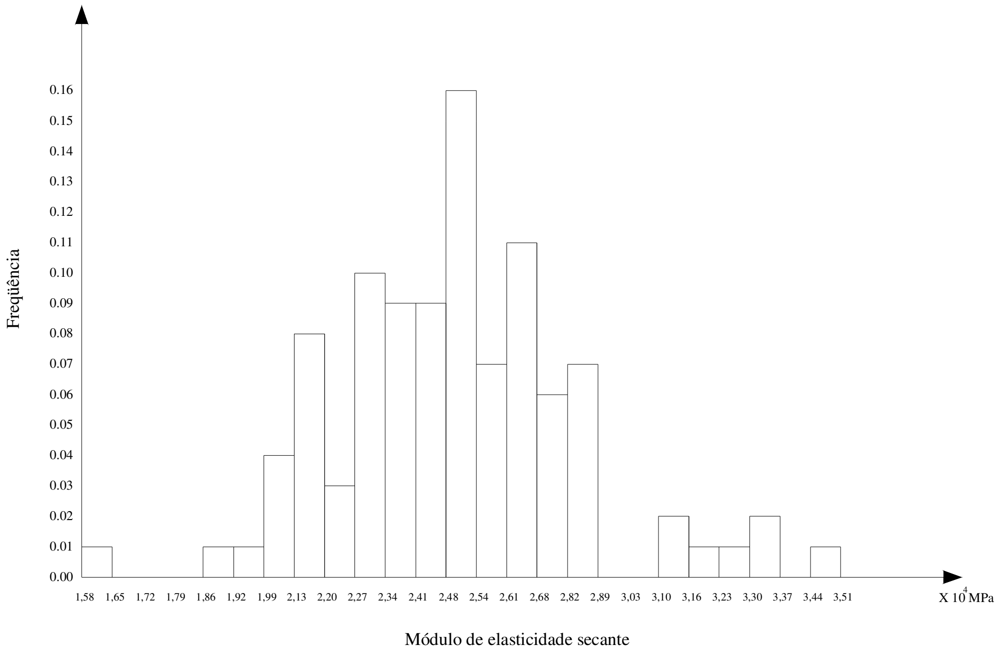

Figura 2.1: Histograma de freqüência**.**

2.8
Distribuições de freqüência relativa e acumulada

A freqüência relativa de uma classe é definida como a relação entre a freqüência dessa classe e o total de todas as classes.

O histograma de freqüência acumulada é o gráfico que representa a freqüência de valores observados de uma variável aleatória que são iguais ou menores que um valor prescrito.

Duas ou mais variáveis aleatórias podem ser observadas simultaneamente durante um dado experimento. Para este caso, o conceito de histograma de freqüência pode ser estendido para múltiplas variáveis aleatórias.

2.9
Medidas de tendência central

As médias são valores típicos ou representativos de um conjunto de dados. Como esses valores típicos tendem a se localizar em um ponto central, dentro de um conjunto de dados ordenados segundo suas grandezas, as médias são denominadas medidas de tendência central.

Vários tipos de médias podem ser definidos, sendo que neste trabalho será amplamente utilizado o conceito de média aritmética ou abreviadamente, média.

Além da média aritmética tem-se, entre outras, a moda e a média harmônica. Cada uma delas apresenta características próprias, sendo o seu uso atribuído de acordo com o fim desejado.

A média aritmética ou simplesmente média, como será tratada a partir deste ponto, para n números x_j, é definida por:

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{2.1}$$

Para mais de uma variável aleatória defini-se o conceito de vetor das médias como:

$$\bar{\mathbf{X}} = \begin{Bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ . \\ . \\ \bar{X}_n \end{Bmatrix} \tag{2.2}$$

2.10 Probabilidade

Quando se realizam experimentos aleatórios depara-se com incertezas quanto à ocorrência, ou não, de determinado evento. A probabilidade é a medida da chance de ocorrência desse evento. Atribui-se um número entre zero e 1 para essa chance, se o evento ocorrer com certeza diz-se que sua probabilidade é de 100% ou 1. Quando se tem a certeza de que ele não ocorrerá, diz-se que sua probabilidade é zero. Na grande maioria das vezes trabalha-se com valores intermediários.

2.11
Eventos independentes

Caso a probabilidade de ocorrência de B não seja afetada pela ocorrência de A, diz-se que A e B são eventos independentes. Ou seja: $P(B \mid A) = P(B)$.

2.12
Distribuições discretas de probabilidade

Seja X uma variável aleatória discreta e $x_1, x_2, ..., x_n$, os valores que ela pode assumir, e supondo probabilidades de ocorrência para esses valores dadas por:

$$P(X = x_k) = p(x_k) \qquad k = 1,2... \tag{2.3}$$

Introduzindo-se o conceito de distribuição discreta de probabilidade dada por:

$$P(X = x) = p(x) \tag{2.4}$$

Pode-se dizer que para $x = x_k$ a expressão (2.4) se reduz a (2.3). Para os demais valores de x tem-se que $p(x) = 0$. De modo geral, $p(x)$ será uma função de densidade de probabilidade se:

$$p(x) \geq 0 \tag{2.5}$$

$$\sum_{i=1}^{n} p(x_i) = 1 \tag{2.6}$$

2.13
Distribuições contínuas de probabilidade

Para definição de distribuição contínua de probabilidade, segundo SPIEGEL (1978) é necessário se falar da probabilidade de uma variável aleatória contínua X estar compreendida entre dois valores diferentes.

Qualquer função que satisfaça às três propriedades apresentadas abaixo será uma função de densidade de probabilidade.

$$p(x) \geq 0 \tag{2.7}$$

$$\int_{-\infty}^{+\infty} p(x)dx = 1 \tag{2.8}$$

$$p(x_a < x < x_b) = \int_{x_a}^{x_b} p(x)dx \tag{2.9}$$

Para variáveis aleatórias contínuas representa-se a distribuição de probabilidade como:

$$P(x_a) = \int_{-\infty}^{x_a} p(x)dx = P[X \leq x_a] \tag{2.10}$$

onde $P(x_a)$ é a probabilidade da variável X ter um valor igual ou menor que x_a.

Quando se trabalha com mais de uma variável aleatória contínua, tem-se uma função de densidade de probabilidade múltipla. Esta função deve obedecer às condições a seguir:

$$p(x_1, x_2, ..., x_n) \geq 0 \tag{2.11}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}...\int_{-\infty}^{+\infty} p(x_1, x_2, ..., x_n)dx_1 dx_2 ...dx_n = 1 \tag{2.12}$$

$$\begin{aligned} & p[a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, ..., a_n \leq x_n \leq b_n] = \\ & = \int_{a_1}^{b_1}\int_{a_2}^{b_2}...\int_{a_n}^{b_n} p(x_1, x_2, ..., x_n)dx_1 dx_2 ...dx_n \end{aligned} \tag{2.13}$$

A função de probabilidade múltipla é definida como:

$$\begin{aligned} P[x_a, x_b, ..., x_n] &= \int_{-\infty}^{x_a}\int_{-\infty}^{x_b}...\int_{-\infty}^{x_n} p(x_1, x_2, ..., x_n)dx_1 dx_2 ...dx_n \\ &= P[x_1 \leq x_a, x_2 \leq x_b, ..., x_n \leq x_n] \end{aligned} \tag{2.14}$$

onde $x_a, x_b, ..., x_n$ são valores conhecidos.

A distribuição de probabilidade de uma variável aleatória contínua obtida a partir de uma função de densidade de probabilidade múltipla é chamada de função de densidade marginal. Para ilustrar, suponha-se que a probabilidade de x_j situar-se entre a_j e b_j é dada por:

$$P\lfloor a_j \le x_j \le b_j \rfloor = P\lfloor a_j \le x_j \le b_j \;\; e \;\; -\infty < x_k < \infty \;\; para \;\; todo \;\; k \ne j \rfloor$$
$$= \int_{a_j}^{b_j} \left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} p(x_1, x_2, \ldots, x_n) dx_1\, dx_2 \ldots dx_{j-1} dx_{j+1} \ldots dx_n \right] dx_j = \int_{a_j}^{b_j} p(x_j) dx_j \quad (2.15)$$

onde:

$$p(x_j) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} p(x_1, x_2, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_n) dx_1\, dx_2 \ldots dx_{j-1} dx_{j+1} \ldots dx_n \quad (2.16)$$

2.14 Esperança matemática

O conceito de esperança matemática, valor esperado ou simplesmente esperança de uma variável aleatória tem grande importância em problemas de probabilidade e estatística. Para variáveis aleatórias discretas X que podem assumir valores $x_1, x_2, \ldots, x_n$, defini-se a esperança de X como:

$$E\langle X \rangle = x_1 P(X = x_1) + \ldots + x_n P(X = x_n) = \sum_{j=1}^{n} x_j P(X = x_j) \quad (2.17)$$

Para uma variável aleatória contínua X com função de densidade de probabilidade $p(x)$, define-se a esperança de X como:

$$E\langle X \rangle = \int_{-\infty}^{+\infty} x p(x) dx \quad (2.18)$$

A esperança de uma função *g(x)* com variável aleatória X é dada por:

$$E\langle g(X) \rangle = \int_{-\infty}^{+\infty} g(x) p(x) dx \quad (2.19)$$

A seguir são apresentadas algumas propriedades importantes da esperança matemática.

(a) Se c é uma constante, então:

$$E\langle cX\rangle = cE\langle X\rangle \equiv c\,\bar{X} \tag{2.20}$$

(b) Se X e Y são variáveis aleatórias, então:

$$E(X+Y) = E(X) + E(Y) \tag{2.21}$$

(c) Se X e Y são variáveis aleatórias independentes, pode-se escrever:

$$E\langle XY\rangle = E\langle X\rangle E\langle Y\rangle \tag{2.22}$$

2.15 Variância e desvio padrão

Segundo SPIEGEL (1978), o grau ao qual os dados numéricos tendem a dispersar em torno de um valor médio é chamado de variação ou dispersão dos dados. Dispõe-se de várias medidas de dispersão ou de variação.

A variância é utilizada como medida da dispersão. Se os valores medidos tendem a se concentrar próximos da média a variância é pequena. Porém, no caso em que tendem a se afastar da média, a variância é grande.

Representa-se a variância de uma variável aleatória X por $Var(X)$, sendo esta definida por:

$$Var(X) = E\left\langle \left(X - \bar{X} \right)^2 \right\rangle \tag{2.23}$$

Para o caso de variáveis aleatórias discretas com funções de densidade de probabilidade $p(x)$ iguais, pode-se descrever a variância como:

$$Var(X) = \frac{1}{n}\sum_{j=1}^{n}\left(x_j - \bar{X}\right)^2 \qquad (2.24)$$

Caso a variável aleatória seja contínua com função de densidade de probabilidade $p(x)$, a variância de X é dada por:

$$Var(X) = E\left\langle\left(X - \bar{X}\right)^2\right\rangle \equiv \int_{-\infty}^{+\infty}\left(X - \bar{X}\right)^2 p(x)dx \qquad (2.25)$$

Algumas propriedades da variância são apresentadas a seguir:

(a) Para uma variável aleatória X:

$$Var(X) = E\left\langle\left(X - \bar{X}\right)^2\right\rangle = E\langle X^2\rangle - \bar{X}^2 = E\langle X^2\rangle - \left[E\langle X\rangle\right]^2 \qquad (2.26)$$

(b) Se c é uma constante, então:

$$Var(cX) = c^2 Var(X) \qquad (2.27)$$

(c) Se X e Y são variáveis aleatórias independentes:

$$Var(X + Y) = Var(X) + Var(Y) \qquad (2.28)$$

$$Var(X - Y) = Var(X) + Var(Y) \qquad (2.29)$$

A raiz quadrada da variância é o desvio padrão. O desvio padrão é representado por s.

$$s = +\sqrt{Var(X)} \qquad (2.30)$$

2.16
Variáveis aleatórias padronizadas

Sendo X uma variável aleatória com média $\bar{X}$ e desvio padrão s, definimos uma variável aleatória padronizada X^* como:

$$X^* = \frac{X - \bar{X}}{s} \tag{2.31}$$

Uma propriedade importante de X^* é que sua média é zero e sua variância é 1. Salienta-se ainda que X^* é uma quantidade adimensional.

2.17 Covariância

Quando duas ou mais variáveis aleatórias são observadas durante um experimento é possível se descrever a correlação existente entre essas variáveis. Esta correlação é chamada de covariância e é expressa para o caso de variáveis aleatórias contínuas da seguinte maneira.

$$Cov(X,Y) = E\left\langle \left(X - \bar{X} \right)\left(Y - \bar{Y} \right) \right\rangle \tag{2.32}$$

Onde:

$$\bar{X} = E\langle X \rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xp(x,y)dxdy \tag{2.33}$$

$$\bar{Y} = E\langle Y \rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yp(x,y)dxdy \tag{2.34}$$

Assim:

$$Cov(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left(X - \bar{X} \right)\left(Y - \bar{Y} \right) p(x,y)dxdy \tag{2.35}$$

Para variáveis aleatórias discretas pode-se escrever simplesmente:

$$Cov(X,Y) = \frac{1}{n}\sum_{j=1}^{n} \left(x_j - \bar{X} \right)\left(y_j - \bar{Y} \right) \tag{2.36}$$

Algumas propriedades importantes da covariância são apresentadas a seguir:

(a) Para variáveis aleatórias X e Y:

$$Cov(X,Y) = E\langle XY\rangle - E\langle X\rangle E\langle Y\rangle \equiv E\langle XY\rangle - \bar{X}\,\bar{Y} \tag{2.37}$$

(b) Se *X* e *Y* são variáveis aleatórias independentes:

$$Cov(X,Y) = 0 \tag{2.38}$$

A matriz de covariância para um número *n* de variáveis aleatórias *X* pode ser representada da seguinte maneira

$$\mathbf{S_x} = \begin{bmatrix} Var(x_1) & Cov(x_1,x_2)\ .\ .\ . & Cov(x\ \ ,x_n) \\ Cov(x_2,x_1) & Var(x_2)\ \ \ .\ .\ . & Cov(x_2,x_n) \\ . & . & \\ . & . & \\ Cov(x_n,x_1) & Cov(x_1,x_2)\ .\ .\ . & Var(x_n) \end{bmatrix} \tag{2.39}$$

2.18
Coeficiente de correlação e coeficiente de variação

Pode-se desejar uma medida de dependência das variáveis aleatórias. Para isso, utiliza-se o conceito de coeficiente de correlação ρ. Este coeficiente, dado por:

$$\rho = \frac{Cov(X,Y)}{s(X)\,s(Y)} \tag{2.40}$$

é um valor adimensional que varia entre *–1 e 1*.

Já o coeficiente de variação é dado pela relação entre o desvio padrão e a média da variável aleatória.

2.19
Funções de densidade de probabilidade

Apontaram-se anteriormente algumas condições para definição de funções de densidade de probabilidade. Porém, muitas funções matemáticas satisfazem essas condições. Apresentam-se a seguir funções amplamente difundidas e utilizadas na engenharia bem como algumas de suas propriedades.

Uma função de densidade de probabilidade uniforme com dois parâmetros (*a* e *b*) é definida por:

$$p(x) = \left(\frac{1}{b-a}\right) \qquad a \le x \le b \tag{2.41}$$

Para valores fora desse intervalo $p(x) = 0$.

A média de *X* é dada por:

$$\bar{X} = \int_{-\infty}^{+\infty} xp(x)dx = \left(\frac{a+b}{2}\right) \tag{2.42}$$

E a variância de *X* é dada por:

$$Var(X) = s_x^2 = \int_{-\infty}^{+\infty} \left(x - \bar{X}\right)^2 p(x)dx = \left(\frac{a-b}{12}\right)^2 \tag{2.43}$$

Uma função de densidade de probabilidade normal com dois parâmetros (*a* e b) é definida por:

$$p(x) = \frac{1}{a\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-b}{a}\right)^2\right\} \tag{2.44}$$

Onde: $-\infty \le x \le \infty$ e $a > 0, b > -\infty$

A média de *X* é *b* e a variância de *X* é a^2.

Pode-se reescrever a função de densidade de probabilidade normal utilizando-se o conceito de variável aleatória padronizada. Desta forma tem-se:

$$p\left(x^*\right) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{X^*}{2}\right\} \tag{2.45}$$

Uma função de densidade de probabilidade log-normal com dois parâmetros (a e *b*) é definida por:

$$p(x) = \frac{1}{ax\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{\ln x - b}{a} \right)^2 \right\} \tag{2.46}$$

Onde: $-\infty \leq x \leq \infty$ e $a > 0, b > -\infty$

A média de X é:

$$\bar{X} = \exp\left\{ b + \frac{a^2}{2} \right\} \tag{2.47}$$

E a variância de X é:

$$Var(X) = s_x^2 = \left(\exp\{2b + a^2\} \right)\left(\exp\{a^2\} - 1 \right) \tag{2.48}$$

Já uma função log-normal com três parâmetros (a, b e c) é definida por:

$$p(x) = \frac{1}{(x-c)a\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(\frac{\ln(x-c) - b}{a} \right)^2 \right\} \tag{2.49}$$

Onde: $c \leq x \leq +\infty$ $a > 0$, $b < +\infty$ e $c > -\infty$

A média de X é:

$$\bar{X} = c + \exp\left\{ b + \frac{a^2}{2} \right\} \tag{2.50}$$

E a variância de X é:

$$Var(X) = s_x^2 = \left(\exp\{2b + a^2\} \right)\left(\exp\{a^2\} - 1 \right) \tag{2.51}$$

Muitas vezes os problemas possuem múltiplas variáveis. Uma função de densidade de probabilidade normal para múltiplas variáveis é definida por:

$$p(x_1, x_2, ..., x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\mathbf{S_x}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j,k=1}^{n} \left(x_j - \bar{X}_j \right) \mathbf{S}_{\mathbf{x}_{jk}}^{-1} \left(x_k - \bar{X}_k \right) \right\} \tag{2.52}$$

onde $\mathbf{S_x}$ é a matriz de covariância e $\mathbf{S}^{-1}_{\mathbf{x}_{jk}}$ é o termo *jk* da sua inversa. $|\mathbf{S_x}|^{\frac{1}{2}}$ representa a raiz quadrada do determinante da matriz de covariância.

2.20 Combinação linear de variáveis aleatórias

Segundo HART (1982), expressar propriedades estatísticas de uma variável aleatória em função das propriedades estatísticas de outra variável aleatória pode ser bastante interessante e proveitoso.

Em seguida serão considerados dois vetores de variáveis aleatórias $\mathbf{X}$ e $\mathbf{Y}$, com relação de dependência linear definida como:

$$\mathbf{Y} = \mathbf{C}\,\mathbf{X} \tag{2.53}$$

onde $\mathbf{C}$ representa uma matriz de constantes. Representando este sistema de equações na forma de somatório:

$$Y_j = \sum_{k=1}^{n} C_{jk} X_k \qquad j = 1,2,...,n \tag{2.54}$$

e calculando-se a esperança nos dois lados da equação (2.54):

$$E\langle Y_j \rangle = \bar{Y}_j = E\left\langle \sum_{k=1}^{n} C_{jk} X_k \right\rangle \tag{2.55}$$

pode-se escrever que:

$$\bar{Y}_j = \sum_{k=1}^{n} C_{jk}\, \bar{X}_k \qquad j = 1,2,...,n \tag{2.56}$$

Ou, na forma matricial:

$$\bar{\mathbf{Y}} = \mathbf{C}\,\bar{\mathbf{X}} \tag{2.57}$$

A transformação matricial representada na equação (2.57) relaciona os valores médios das n variáveis aleatórias $\mathbf{Y}$ com os valores médios das n variáveis aleatórias $\mathbf{X}$.

A matriz de covariância de $\mathbf{Y}$ também pode ser obtida da matriz de covariância de $\mathbf{X}$. Para demonstrar essa afirmativa, HART(1982) considerou duas variáveis aleatórias, Y_j e Y_k.

Da equação (2.53) pode-se escrever que:

$$Y_j = \sum_{l=1}^{n} C_{jl} X_l \tag{2.58}$$

$$Y_k = \sum_{h=1}^{n} C_{kh} X_h \tag{2.59}$$

Usando o conceito de matriz de covariância é possível se escrever que:

$$Cov(Y_j, Y_k) = \sum_{l=1}^{n} \sum_{h=1}^{n} C_{jl} C_{kh} Cov(X_l X_h) \tag{2.60}$$

Ou na forma matricial:

$$\mathbf{S_y} = \mathbf{C S_x C^T} \tag{2.61}$$

Onde $\mathbf{S_y}$ e $\mathbf{S_x}$ são respectivamente as matrizes de covariância das variáveis aleatórias $\mathbf{Y}$ e $\mathbf{X}$.