# Modelagem Matemático-computacional
# Aula 4

Nesta aula foram introduzidos conceitos de probabilidade e estatística.

## I. INTRODUÇÃO

Figura 1: Dados são comumente utilizados em exemplos estatísticos. A Estatística foi bastante desenvolvida em teoria de jogos.

A estatística preocupa-se em analisar propriedades de eventos, tais como sua frequência de ocorrência, para prever ou estimar a ocorrência de eventos futuros através da teoria das probabilidades. Quando se fala em Estatística é preciso definir o que é um **experimento estocástico** (também denotado por aleatório). Um experimento aleatório consiste em um tipo de experimento cujos valores dos possíveis resultados são incertos. Um exemplo pode ser o processo de medida de alunos da classe de Modelagem em 2012 na aula de 20 de agosto de 2012.

Como o exemplo anterior mostra, qualquer experimento aleatório requer o máximo de especificação sobre as condições na qual será efetuado. Uma vez definido o experimento com um nível adequado de especificação, pode-se obter os possíveis resultados desse experimento, que constituem o seu respectivo **conjunto universo** $\Omega$. Também denotado por *espaço amostral*, $\Omega$ é o conjunto de todos os *eventos elementares* que um experimento aleatório pode gerar como resultado.

Exemplos de eventos e conjuntos universo são:

1. Moeda arremessada: podemos obter cara(H) ou coroa(T)

   $\Omega_1 = \{H, T\}$

2. Dado arremessado: podemos obter qualquer uma das 6 faces

   $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$

3. Urnas: depende de quais objetos contidos na urna. Ex: bolas coloridas ou numeradas. Faça como exercício: determine $\Omega_3$ para uma urna com 5 bolas brancas (B), 3 pretas (P) e 2 azuis (A).

4. Altura de estudantes de Fiscomp do IFSC no dia 20/08/2012

   $\Omega_4 = [1.4, 2.00]$

Um **evento** nada mais é do que qualquer subconjunto de $\Omega$. Nos três primeiros exemplos anteriores, teríamos:

1. $A_1 = \{H\}$; $B_1 = \{T\}$; $C_1 = \emptyset$ e $D_1 = \{H, T\} = \Omega_1$

2. $A_2 = \{4\}$; $B_2 = \{1, 2, 3\}$; $C_2 = \{4, 6\}$; etc ...

   Note que $\Omega_2$ e $\emptyset$ também são eventos para o lançamento do dado.

3. Urnas: determinar os eventos a partir do exercício proposto.

A um dado evento pode-se atribuir **probabilidade** $\mathbb{P}$, que explicaremos em seções posteriores.

## II. NOÇÕES DE ANÁLISE COMBINATÓRIA

### A. Permutações

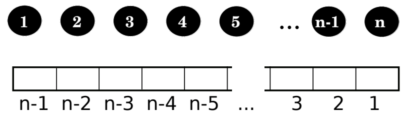


Figura 2: As permutações de $n$ objetos em $n$ posições são $n!$.

Permutações são formas de se arrumar $n$ objetos em $n$ posições, considerando que a mudança da ordem dos objetos implica numa maneira distinta. Um exmeplo simples que leva à dedução de como podemos obter a fórmula da permutação é o seguinte: imagine você tenha uma fileira de $n$ caixas e $n$ bolas numeradas de 1 a n, que deverão ser colocadas em cada uma das $n$ caixas. Inicialmente, todas as caixas estão vazias. Então, você vai colocar escolher uma bola e colocar na primeira caixa, escolher outra e colocar na segunda caixa e assim sucessivamente até que todas as caixas estejam preenchidas com uma bola. Ao preencher a primeira caixa, você tem qualquer uma das $n$ opções de bola. Quando você vai colocar a segunda bola na segunda caixa, sobram $n - 1$ bolas para você escolher e colocar a segunda caixa e, analogamente, restarão $n-2$ bolas de opção para a terceira caixa. Quando você chegar na penúltima caixa, você terá somente 2 opções das bolas que restaram e, finalmente, na última caixa, você terá apenas uma opção (veja a figura 2. Usando o princípio multiplicativo (ver subseção *Princípio Multiplicativo* em caso de dúvida), obtemos a expressão para a permutação das $n$ bolas em $n$ caixas:

$$P_n = n(n-1)(n-2)...21 = n! \tag{1}$$

### B. Arranjos

Arranjos são formas de se bf ordenar $p$ obejtos escolhidos entre $n$ objetos distintos $pleqn$. Em números, temos:

$$A_p^n = \frac{n!}{(n-p)!} \tag{2}$$

Exemplo: Suponha que tenhamos um supermercado com 4 caixas. Desses 4 caixas, 2 poderão concorrer a um prêmio que dá direito a uma compra grátis, se for o primeiro sorteado, e a 50% de desconto na compra, se for o segundo sorteado. As possíveis formas de se sortear os 2 caixas entre os 4 possíveis é $\frac{4!}{(4-2)!} = 12$. Note que a ordem é importante: se você for sorteado primeiro seu desconto é maior. Uma forma de se visualizar isso, é fazendo a tabela I de possíveis sorteios dos caixas:

Tabela I: Possíveis sorteios de dois caixas de supermercado entre 4.

| C1 | C2 | C3 | C4 |
|---|---|---|---|
| 1º | 2º | | |
| 2º | 1º | | |
| | 1º | 2º | |
| | 2º | 1º | |
| | | 1º | 2º |
| | | 2º | 1º |
| 1º | | 2º | |
| 2º | | 1º | |
| 1º | | | 2º |
| 2º | | | 1º |
| | 1º | | 2º |
| | 2º | | 1º |

### C. Combinação

Combinações são formas de se escolher $p$ objetos a partir de $n$ objetos distintos $p \leq n$. Note que em combinações a ordem de escolha dos objetos não é importante, de modo que uma combinação é equivalente a um arranjo com as formas repetidas descontadas. Matematicamente, temos:

$$C_p^n = \binom{n}{p} = \frac{n!}{p!(n-p)!} \tag{3}$$

No exemplo de sorteio de caixas de supermecado da seção anterior, caso o desconto fosse igual para os dois caixas sorteados, não importaria quem foi sorteado primeiro. Então, as linhas da tabela I como 1º ou 2º aparecem trocados 2 vezes nas mesmas posições, as combinações de sorteios seriam $\frac{4!}{2!(4-2)!} = 6$.

## III. PROBABILIDADE

Formalmente, a probabilidade de ocorrênciade um evento A $\mathbb{P}(A)$ é estimada observando-se o número $N_A$ de ocorrências de A em um número $N$ de realizações no limite:

$$\mathbb{P}(A) = \lim_{N\to\infty} \frac{N_A}{N}, \tag{4}$$

ou seja, é a frequência relativa do evento A quando se repete o experimento um número muito grande de vezes. Essa definição implica em um problema prático, uma vez que é impossível realizar um experimento infinitamente. Considera-se então um número N máximo de realizações praticáveis. Por exemplo, pode-se lançar um dado 20000 vezes e contar em quantos lançamentos obteve-se 4. No caso do dado, podemos ainda considerar que esta estimativa pode ser facilitada através do conceito de **equiprobabilidade**. Na hipótese de equiprobabilidade todos os eventos têm a mesma probabilidade de ocorrência, reduzindo a probabilidade a:

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}, \tag{5}$$

onde $|A|$ e $|\Sigma|$ correspondem aos tamanhos dos conjuntos que representam o evento A e o conjunto universo a ele relacionado, respectivamente. No nosso caso com o dado lançado, a probabilidade de obter a face 4 seria $\mathbb{P}(4) = 1/6$. Sabemos, porém, que esta hipótese não é completamente real, pois o dado pode conter imperfeições que levará a ruídos e imperfeições neste tipo de estimativa. Mais ilustrações do conceito de probabilidade na hipótese de equiprobabilidade nos exemplos iniciais são:

1. $\mathbb{P}(A_1) = \frac{1}{2}$
2. $\mathbb{P}(\{1,2\}) = \frac{2}{6} = \frac{1}{3}$
   $\mathbb{P}(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2}$
   $\mathbb{P}(\{1,2,3,4,5,6\}) = \frac{6}{6} = 1$
   $\mathbb{P}(\emptyset) = 0$

### A. Propriedades de Probabilidade

Uma vez que os eventos são representados por conjuntos é conveniente utilizar a representação de Venn para definir algumas propriedades da probabilidade.

#### 1. Valores de probabilidade

A probabilidade de um evento A é sempre um valor positivo menor que 1:

$$0 \leq \mathbb{P}(A) \leq 1 \tag{6}$$

Dos exemplos introduzidos na última subseção, derivamos que $\mathbb{P}(\Omega) = 1$ e $\mathbb{P}(\emptyset) = 0$

#### 2. Princípio Aditivo

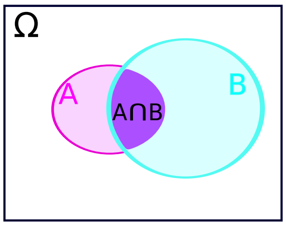


Figura 3: Ilustração do princípio aditivo em Teoria dos conjuntos.

Probabilidade de obter um evento A ou um evento B dentro de um conjunto universo Ω:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B), \tag{7}$$

onde $\mathbb{P}(A \cap B)$, isto é, a probabilidade simultânea de dois eventos é subtraída para evitar contar duas vezes. Note que se a intersecção dos conjuntos é vazia, a probabilidade de A ou B torna-se $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Veja a figura 3

#### 3. Princípio Multiplicativo

Dado um evento A com um número n(A) de possíveis resultados e um outro evento B com um número n(B) de possíveis resultados independentes de A, o produto cartesiano destes eventos pode ser expresso na forma:

$$n(A \times B) = n(A)n(B) \tag{8}$$

Este conceito será generalizado para probabilidade mais adiante.

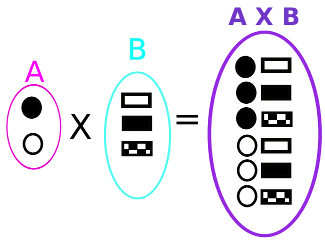


Figura 4: Produto cartesiano de dois grupos A e B, constituídos de 2 e 3 objetos diferentes, respectivamente. Note que o resultado do produto cartesiano são pares de objetos.

#### 4. Complemento

A probabilidade de que um evento B ocorra, tal que B exclui todas as possibilidades de um evento A aconteça, isto é B é complementar a A $B = \overline{A}$, é dada por:

$$\mathbb{P}(\overline{A}) = 1 - (A) \tag{9}$$

## IV. VARIÁVEL ALEATÓRIA

Uma variável aleatória (V.A.) $X$ pode ser entendida como uma *função* que atribui um valor real $x$ a cada evento elementar em um experimento estocástico. Dizemos que estes valores $x$ são as realizações desta variável aleatória. Exemplos de variáveis aleatórias são: altura dos alunos de Modelagem na aula de 27/08/2012, peso das jogadoras da seleção feminina de vôlei, temperatura de uma barra de ferro, valor da face de um dado lançado, etc.

Variáveis aleatórias podem ser de dois tipos: discreta e contínua.

### A. V.A. Discreta e Histogramas

Variáveis aleatórias discretas $X$ são caracterizadas por possuírem um espaço de estados (conjunto universo)Ω enumerável e pela frequência de ocorrência de cada possível resultado de Ω dada por pesos probabilísticos $P_i$, com $\sum_i P_i = 1$. Assim, a probabilidade de uma variável aleatória discreta assumir um certo valor $x_i$ é dada por $\mathbb{P}(X = x_i) = P_i$.

Uma forma de se construir um modelo para este tipo de V.A. consiste em organizar as possíveis medidas desta V.A. em um **histograma**. Para tanto, obtém-se os possíveis resultados Ω de uma medida desta V.A. contando-se quantos destes resultados apresentaram valores dentro de intervalos de tamanho $\delta$. Quando as contagens são divididas pelo número total de resultados observados, temos o histograma de frequências relativas:

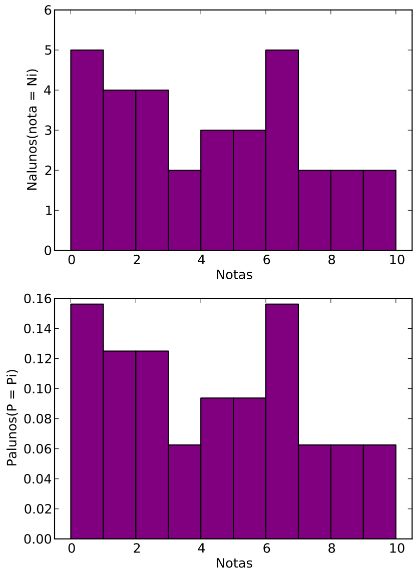


Figura 5: Histograma de contagem e de frequência relativa da distribuição de notas dos alunos de Mecânica Clássica 2.

$$f_X(x = x_i) = \frac{N_{x=x_i}}{N_X} \tag{10}$$

Por exemplo, suponha que foram corrigidas as provas de 32 alunos que cursavam Mecânica Clássica 2, com possíveis notas de 0 a 10 e foram obtidos os resultados $\Omega_{notas} = \{2.50, 4.25, 0.00, 6.80, 6.80, 3.50, 5.50, 5.00, 0.50, 7.50, 1.50, 2.50, 1.00, 7.00, 10.00, 4.75, 4.50, 3.30, 6.50, 8.50, 5.00, 0.00, 6.50, 8.30, 1.00, 2.75, 0.00, 2.00, 6.50, 0.00, 9.50, 1.30\}$. Pode-se construir um histograma com o número de alunos $N_{alunos}$ que obteviveram notas $N_i$ de 0 a 1, de 1 a 2, de 2 a 3, ..., de 9 a 10. Para se obter os pesos probabilísticos associados a cada nota, basta dividir cada uma das contagens do histograma obtido pelo número total de alunos. Veja a figura 5

### B. V.A. Contínua e Função Densidade de Probabilidade

Uma variável aleatória contínua $X$ é caracterizada também pelo seu espaço de estados $\Omega$, que são não enumeráveis. Ao invés de associarmos pesos probabilísticos a cada possível realização da V.A. contínua, utilizamos a **função densidade de probabilidade**(F.D.P) $\rho_X(x)$. A função $\rho_X(x)$ dá a probabilidade de obter um resultado em um dado intervalo real, inexistindo a probabilidade de um resultado para um dado $x = x_i$. Podemos fazer uma analogia entre a função densidade de probabilidade e o histograma de frequência relativa no limite em que os intervalos tendem a 0 e o número de observações $N_X$ tende a infinito:

$$\rho_X(x) = \lim_{\delta \to 0} \lim_{N_X \to \infty} \frac{N_{x=x_i}}{N_X} \tag{11}$$

Existem algumas propriedades que $\rho$ deve obedecer:

1. $\rho_X(x) \geq 0, \forall x \in \Re$
2. $\int_{-\infty}^{\infty} \rho_X(x)dx = 1$
3. $\mathbb{P}(x_1 \leq x \leq x2) = (x_2 - x_1) \int_{x_1}^{x_2} \rho_X(x)dx$
4. $\mathbb{P}(x_i) = 0$

Note que é possível definir uma função densidade de probabilidade para variáveis discretas, basta tomar $p_X(x) = \sum_i P_i \delta(x - x_i)$.

A função $\rho$ dá toda a informação sobre a V.A. Destacamos em seguida, alguns tipos de funções densidade de probabilidade: uniforme, normal e exponencial.

#### 1. Distribuição Uniforme

Na distribuição uniforme a probabilidade de se gerar qualquer ponto em um intevalo contido no espaço amostral $[a, b]$ é proporcional ao tamanho do intervalo e sua função densidade de probabilidade é dada por:

$$\rho_U(x) = \begin{cases} \frac{1}{b-a}, & \text{se} a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases} \tag{12}$$

Graficamente, temos:

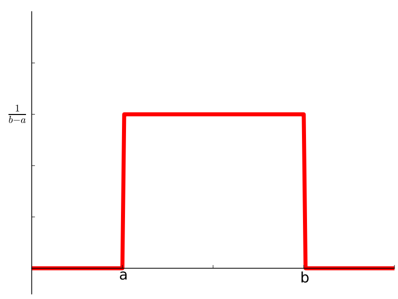


Figura 6: Gráfico de distribuição uniforme no intervalo $[a, b]$.

Na natureza, a variabilidade de indivíduos bem como as imperfeições impedem que sistemas possam ser idealmente modelados por uma distribuição uniforme.

2. Distribuição Normal

Na distribuição normal a densidade de probabilidade está relacionada à curva gaussiana, diferindo desta devido à normalização imposta sobre a densidade de probabilidade, dada por:

$$\rho_N(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \tag{13}$$

onde os parâmetros $\mu$ e $\sigma$ controlam, respectivamente, o ponto de máximo e a abertura da gaussiana. Note que $rho_N(x)$ é par e $\mu$ está relacionado à média, enquanto $\sigma$ refere-se ao desvio padrão. Graficamente, teríamos:

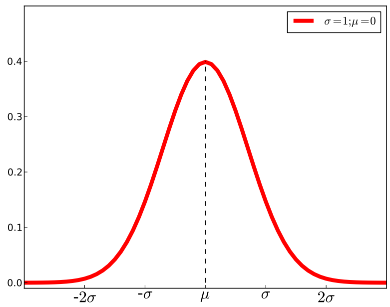


Figura 7: Gráfico de distribuição normal no intervalo $[-\infty, infity]$ com média $\mu$ e desvio padrão $\sigma$.

Muitos sistemas reais são modelados por uma distribuição normal. Um caso físico que se comporta segundo esta distribuição tanto experimental quanto analiticamente é a difusão. Por exemplo, quando um canhão de elétrons lança partículas em um filme que armazena carga, é possível provar que a distribuição de carga no filme segue uma distribuição normal.

3. Distribuição Exponencial

Na distribuição exponencial, existem valores de densidade de probabilidade apenas para valores positivos. Ela é matematicamente dada por:

$$\rho_E(x) = \begin{cases} \alpha e^{-\alpha x}, & \text{se} x \geq x \leq 0 \\ 0, & \text{se} x < 0 \end{cases}, \tag{14}$$

onde $\alpha$ é corresponde ao inverso da posição de decaimento da curva relativa a esta expressão. O gráfico desta distribuição é ilustrado na figura 8.

Um exemplo prático de distribuição exponencial na natureza pode ser a distribuição de calor numa xícara.

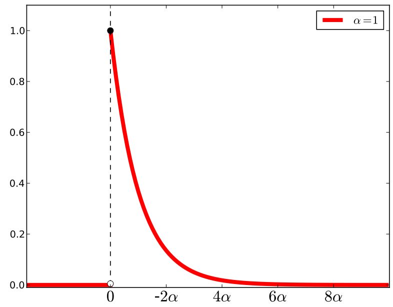


Figura 8: Gráfico de distribuição exponencial no intervalo $[-\infty, infity]$ com parâmetro de decaimento $\alpha$.

4. Distribuição Delta de Dirac

A delta de Dirac é uma importante distribuição que simplifica a representação de vários sistemas físicos como, por exemplo, uma carga puntual no espaço contínuo. Matematicamente, define-se a $\delta(t)$ de uma variável $t$ na forma:

$$\delta(t) = \begin{cases} 0, & \text{se} t \neq 0 \\ 0, & \text{nao definido em} t = 0 \end{cases}. \tag{15}$$

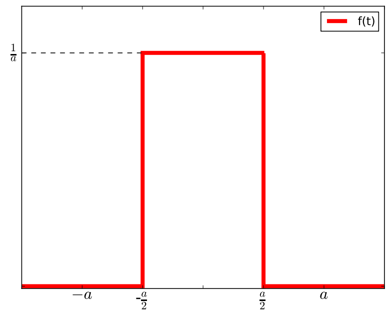


Figura 9: Uma das possíveis representações gráficas para uma delta de Dirac consiste em uma função tipo poço de potencial bastante encontrada em Mecânica Quântica. Note que ela obedece às condições da delta de Dirac no limite $\lim_{a\to 0} f(t)$.

Uma forma pictórica de se representar uma delta de Dirac é ilustrada na figura 9. A distribuição delta de Dirac ainda pode ser expressa através de funções, como a gaussiana e a sinc. Veja a figura 10.

A delta de Dirac obedece às propriedades:

1. $\int_{-\infty}^{\infty} \delta(t)dt = 1$

2. $f(t)(t - t_0) = f(t_0)\delta(t - t_0)$

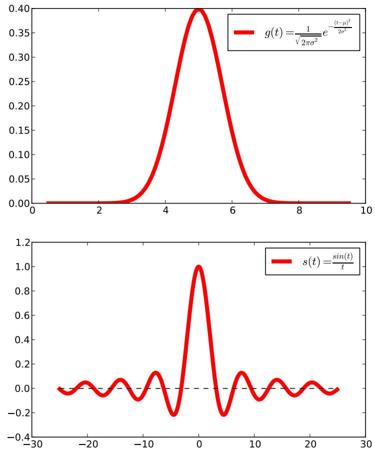

Figura 10: As funções gaussiana e sinc podem ser usadas como uma possível representação para a delta de Dirac pois convergem para ela em dados limites.

3. $\int_{-\infty}^{\infty} f(t)\delta(t-t_0) = f(t_0)$
4. $\int_{-\infty}^{\infty} \delta(at)dt = \frac{1}{|a|}$
5. $t\delta(t) = 0$
6. $\int_{-\infty}^{\infty} f(t)\delta(g(t))dt = \sum_{i=raizes} \frac{f(x_i)}{|g'(x_i)|}$

### C. Momentos estatísticos de uma variável aleatória

Momentos estatísticos são uma forma de caracterizar uma variável aleatória $X$ que obedece uma dada densidade de probabilidade $\rho_X$. Em particular, a *média* e o *desvio padrão* (que também estão relacionados aos momentos estatísticos) são duas medidas que permitem obter muita informação sobre o sistema estudado.

O momento estatístico $\langle X^k \rangle$ de ordem $k$ de uma V.A. com densidade de probabilidade $\rho_X$ é dada por:

$$\langle X^k \rangle = \int_{-\infty}^{\infty} x^k \rho_X(x)dx. \tag{16}$$

Em termos dos momentos estatísticos, a média de $X$ é dada por:

$$\mu_X = \langle X \rangle = \int_{-\infty}^{\infty} x\rho_X(x)dx. \tag{17}$$

Outro importante momento estatístico é a *variância* (ou primeiro momento central) de $X$:

$$\mathbb{V}(X) = \int_{-\infty}^{\infty} (x - \langle X \rangle)^2 \rho_X(x)dx, \tag{18}$$

cuja raiz quadrada dá o desvio padrão da V.A. $X$, ou seja $\sqrt{\mathbb{V}(X)} = \sigma_X$. Assim, podemos definir um momento central de ordem $k$ como segue:

$$\langle (X - \langle X \rangle)^k \rangle = \int_{-\infty}^{\infty} (x - \langle X \rangle)^k \rho_X(x)dx. \tag{19}$$

Note que existe uma relação binívoca entre $\rho_X(x)$ e seus momentos estatísticos: conhecendo $\rho_X$ é possível obter $\langle X^k \rangle$ para todo $x$ e, por outro lado, conhecendo todos os $\langle X^k \rangle$, é possível recuperar $\rho_X$, sem perda de informação.

### D. Transformações de Varáveis Aleatórias

Dadas duas V.A. $X_1$ e $X_2$, podemos combinálas para obter uma nova V.A. $Z$ dada por qualquer operação ou função matemática que envolva alguma ou ambas $X_1$ e $X_2$. Exemplos:

- $Z = 2X_1$
- $Z = X_1 - X_2$
- $Z = a + bX_1$, onde $a, b \in \Re$
- $Z = f(X_1)$

Uma importante transformação que pode ser feita sobre uma variável aleatória X consiste em torná-la adimensional. Para tanto, basta deslocá-la pela sua média e normalizá-la pelo seu desvio padrão. Esse procedimento, denominado *estandardização*, é comumente empregado em análise de dados onde as escalas de medidas variam dentro do conjunto de dados que se está analisando. Matematicamente, temos:

$$Z = \frac{X - \mu_X}{\sigma_X}. \tag{20}$$

É possível provar que as variáveis transformadas estarão distribuídas entre $[-2, 2]$.

## V. ESTATÍSTICA MULTIVARIADA

### A. Momentos conjuntos

Vimos na seção anterior como são definidos os momentos de uma variável aleatória quando esta é considerada isoladamente. Podemos querer entender como duas variáveis se relacionam entre si, isto é, como a dependência de duas variáveis aleatórias reflete sua variação conjunta. Para clarificar esta idéia de dependência de variáveis aleatórias, basta pensar em peso e altura de atletas; espera-se que quanto maior a altura do atleta maior o seu peso de modo que há uma relação direta entre estas duas grandezas (variáveis aleatórias). Qualitativamente, uma forma de se visualizar este relacionamento consiste em fazer um gráfico com as duas grandezas de interesse cada uma em um dos eixos x ou y.

Quantativamente, uma medida que dá o grau de relacionamento entre duas variáveis aleatórias $X$ e $Y$ é dada pelo momento conjunto:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\rho_{X,Y}(x,y)dxdy, \tag{21}$$

que é denominada *correlação* de X e Y.

Observe que a densidade de probabilidade que aparece na integral é uma densidade de probabilidade conjunta e, somente sob hipótese de independência das duas V.A.'s não se pode dizer que $\rho_{X,Y}(x,y) = \rho_X(x)\rho_Y(y)$.

Pode-se definir a correlação após transformações nas variáveis X e Y, transladando-as para suas respectivas médias, ou ainda fazendo a estandardização. No primeiro caso, temos a *covariância* de X e Y, dada por:

$$E[(X-\mu_X)(Y-\mu_y)] = \iint_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)\rho_{X,Y}(x,y)dxdy, \tag{22}$$

e no segundo, temos a *correlação de Pearson*:

$$E\left[\frac{X-\mu_X}{\sigma_X}\frac{Y-\mu_y}{\sigma_Y}\right] = \iint_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)\rho_{X,Y}(x,y)dxdy. \tag{23}$$

Neste último caso, os valores para a correlação de Pearson estarão compreendidos no intervalo $[-1,1]$.

## Referências

[1] COSTA, L. da F., CESAR Jr, R.M. Shape analysis and classification: theory and pratice, Boca Raton, FL, CRC Press; 2001.

[2] MAIA, L. P. Notas de aula do curso Mecânica Estatística A, 2012.