CAPÍTULO III - VARIÁVEIS E VETORES ALEATÓRIOS

- **Conceito e definição**
- **Função distribuição de probabilidade**
- **Função densidade de probabilidade**
- **Revisão: vetores, matrizes e espaços lineares**
- **Vetores aleatórios**
- **Distribuições e densidades conjuntas**
- **Distribuições e densidades condicionais**
- Transformações com variáveis aleatórias
- Médias e esperanças
- Função característica e momentos
- Vetores aleatórios gaussianos
- Funções lineares de vetores gaussianos
- Teorema do Limite Central

Variáveis Aleatórias

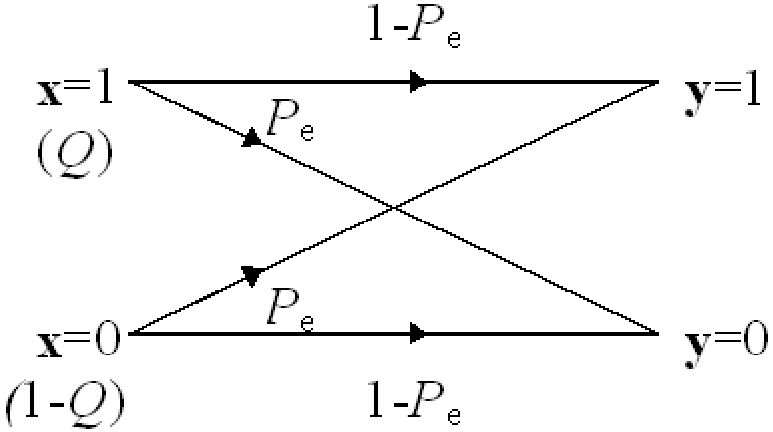
- Uma variável aleatória é um número $\mathbf{x}(\xi)$ atribuído a cada resultado ξ de um experimento. Esse número pode ser o ganho num jogo de azar, a voltagem de uma fonte de sinais aleatórios, o custo de um componente aleatório, ou qualquer outro número que seja de interesse na realização do experimento.
- Mais precisamente, uma variável aleatória é uma função $\mathbf{x}$ cujo domínio é o conjunto S de todos os possíveis resultados de um experimento.
- Usaremos letras em negrito, $\mathbf{x}$, para representar variáveis aleatórias e letras em itálico, x, para nos referir aos valores que essa variável assume.
- A probabilidade de uma variável aleatória $\mathbf{x}$ assumir o valor x será representada, portanto, por $P_{\mathbf{x}}(x)$.

Exemplo: arremesso de um dado

- Ao experimento de se arremessar um dado comum e observar-se o resultado é associada a variável aleatória $\mathbf{x}$, de forma que essa variável aleatória assume o valor igual a 10 vezes o número de pontos resultante. Determine a probabilidade da variável aleatória **x** ser maior que 15, $P(\mathbf{x}>15)$.
- **Solução:** A variável aleatória **x** pode assumir apenas os valores 10, 20, 30, 40, 50 e 60. Podemos admitir que qualquer das faces pode ocorrer com a mesma probabilidade, portanto:
- $P(\mathbf{x}=10)= P(\mathbf{x}=20)= P(\mathbf{x}=30)= P(\mathbf{x}=40)= P(\mathbf{x}=50)= P(\mathbf{x}=60)= 1/6$.
- Assim. $P(\mathbf{x}>15)= 5/6$.

Exemplo: canal binário simétrico

- Num canal binário simétrico, a probabilidade de erro é P_e.
- A probabilidade de se transmitir o dígito binário 1 é Q e a de se transmitir 0 é $1\text{-}Q$.
- Determine a probabilidade de se receber o dígito 0 e de se receber o dígito 1 no receptor.
- **Solução:** Vamos definir $\mathbf{x}$ como sendo a variável aleatória "dígito transmitido", podendo assumir os valores 0 ou 1, e $\mathbf{y}$ como sendo a variável aleatória "dígito recebido".



Exemplo: canal binário simétrico

- Como o canal é simétrico, $P_{y|x}(0|1)=P_{y|x}(1|0)=P_e$ (probabilidade de erro)
- Consequentemente, $P_{y|x}(0|0)=P_{y|x}(1|1)=1-P_e$.
- As probabilidades na transmissão são $P_x(1)=Q$ e $P_x(0)=1-Q$.
- Desejamos calcular $P_y(1)$ e $P_y(0)$. Assim,
- $P_y(1)=P_x(1)P_{y|x}(1|1)+P_x(0)P_{y|x}(1|0)=Q\,(1-P_e)+(1-Q)\,P_e$
- Similarmente,
- $P_y(0)=P_x(1)P_{y|x}(0|1)+P_x(0)P_{y|x}(0|0)=Q\,P_e+(1-Q)\,(1-P_e)$
- Note que, devido aos erros, a probabilidade de se receber o dígito 1 é diferente da de se transmitir o dígito 1 (exceto quando $Q = 0{,}5$).

Exemplo: canal binário simétrico

- Determine qual a probabilidade do dígito transmitido ter sido 1 quando o dígito recebido for 1, ou seja, $P_{\mathbf{x|y}}(1|1)$.
  - Vamos usar o teorema de Bayes:
  - $P_{\mathbf{x|y}}(1|1) = P_{\mathbf{y|x}}(1|1)\, P_{\mathbf{x}}(1) \,/\, P_{\mathbf{y}}(1) = (1 - P_{\mathrm{e}})\, Q \,/[Q\,(1 - P_{\mathrm{e}})+(1 - Q)\, P_{\mathrm{e}}]$
- Alguns exemplos numéricos:
  - $P_{\mathrm{e}} = 0$: $P_{\mathbf{x|y}}(1|1) = 1$
  - $P_{\mathrm{e}} = 1$: $P_{\mathbf{x|y}}(1|1) = 0$
  - $P_{\mathrm{e}} = 0{,}5$: $P_{\mathbf{x|y}}(1|1) = Q$
  - $Q = 0{,}5$: $P_{\mathbf{x|y}}(1|1) = (1 - P_{\mathrm{e}})$
  - $P_{\mathrm{e}} = 0{,}01$, $Q = 0{,}1$: $P_{\mathbf{x|y}}(1|1) = 0{,}916667$

Função distribuição de probabilidade

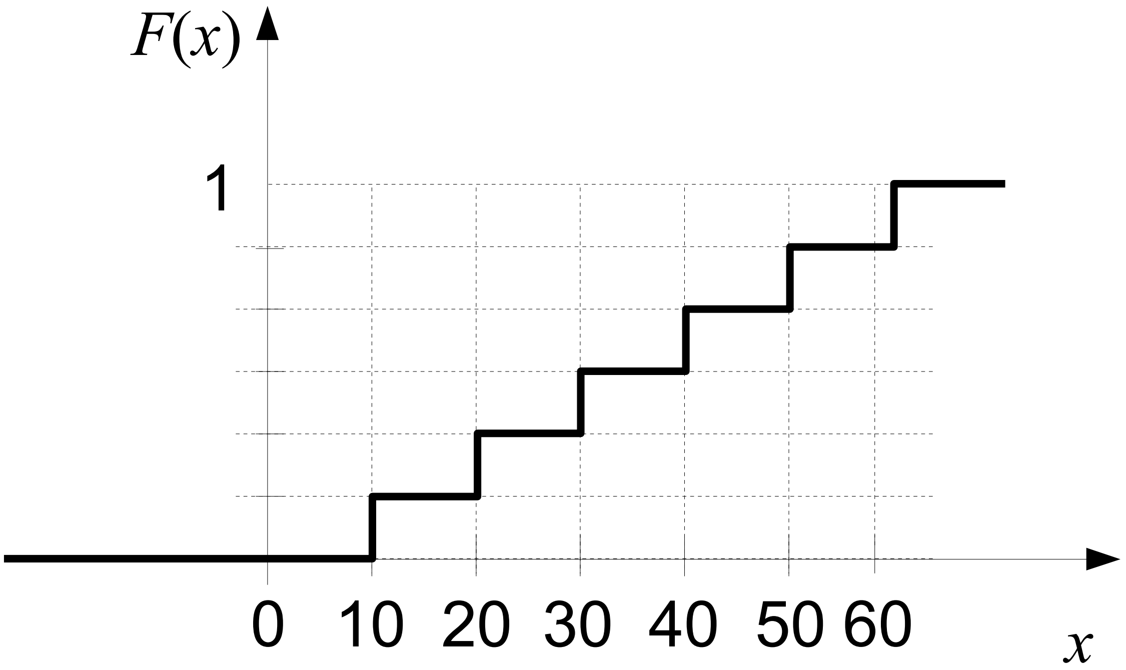
- A função distribuição de probabilidade, ou probabilidade acumulada, de uma variável aleatória $\mathbf{x}$ é a função $F_{\mathbf{x}}(x)$ definida por
  - $F_{\mathbf{x}}(x)=P(\mathbf{x} \leq x)$
- definida para x entre $-\infty$ e $+\infty$.
- Quando não houver risco de ambiguidade, podemos utilizar simplesmente a notação $F(x)$ para representar essa função.
- Propriedades:
  - $F(x) \geq 0$
  - $F(\infty) = 1$
  - $F(-\infty) = 0$
  - $F(x_1) \leq F(x_2)$ para $x_1 < x_2$
- $P(x_1 < \mathbf{x} \leq x_2) = F(x_2) - F(x_1)$

Função distribuição de probabilidade

- Diremos que uma variável aleatória é do tipo discreto se $F(x)$ tiver a forma de uma "escada" (apresentar descontinuidade num conjunto discreto de pontos, sendo constante nos intervalos entre esses pontos);
- Diremos que uma variável aleatória é do tipo contínuo se $F(x)$ for uma função contínua de x;
- Diremos que uma variável aleatória é do tipo misto se tiver descontinuidades mas não com a forma de uma "escada".

Exemplo: arremesso de um dado

- Seja o experimento de se lançar um dado e observar o resultado. Vamos definir a variável aleatória $\mathbf{x}$ tal que $\mathbf{x}(\text{face } i) = 10\ i$ (i entre 1 e 6). Sendo o dado honesto, $P(\mathbf{x}=10)=P(\mathbf{x}=20)=\ldots=P(\mathbf{x}=60)=1/6$. Assim, a função distribuição de probabilidade terá a forma de uma escada:

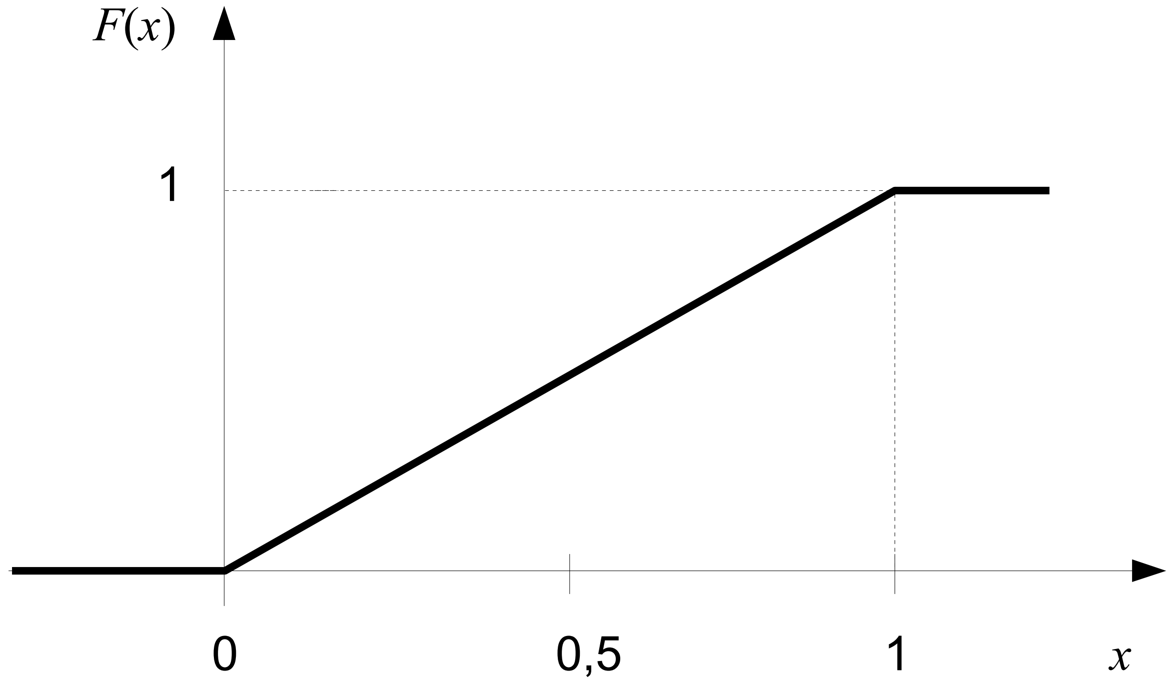


Exemplo: telefonema

- Um telefonema ocorre aleatoriamente no intervalo $(0, 1)$. Neste experimento, os resultados são instantes de tempo entre 0 e 1, e a probabilidade que esteja entre dois instantes t_1 e t_2 (pertencentes ao intervalo especificado) é dada por
  - $P(t_1 \leq \mathbf{x} \leq t_2) = t_2 - t_1$
- Dessa forma, para $x < 0$,
  - $F_{\mathbf{x}}(x) = P(\mathbf{x} \leq x, x < 0) = 0$,
- Para $x > 1$,
  - $F_{\mathbf{x}}(x) = P(\mathbf{x} \leq x, x > 1) = 1$,
- E para $0 \leq x \leq 1$
  - $F_{\mathbf{x}}(x) = P(\mathbf{x} \leq x,\ 0 \leq x \leq 1) = P(0 < \mathbf{x} \leq x,\ 0 \leq x \leq 1) = x$

Exemplo: telefonema

- Gráfico da função distribuição de probabilidade



Função densidade de probabilidade

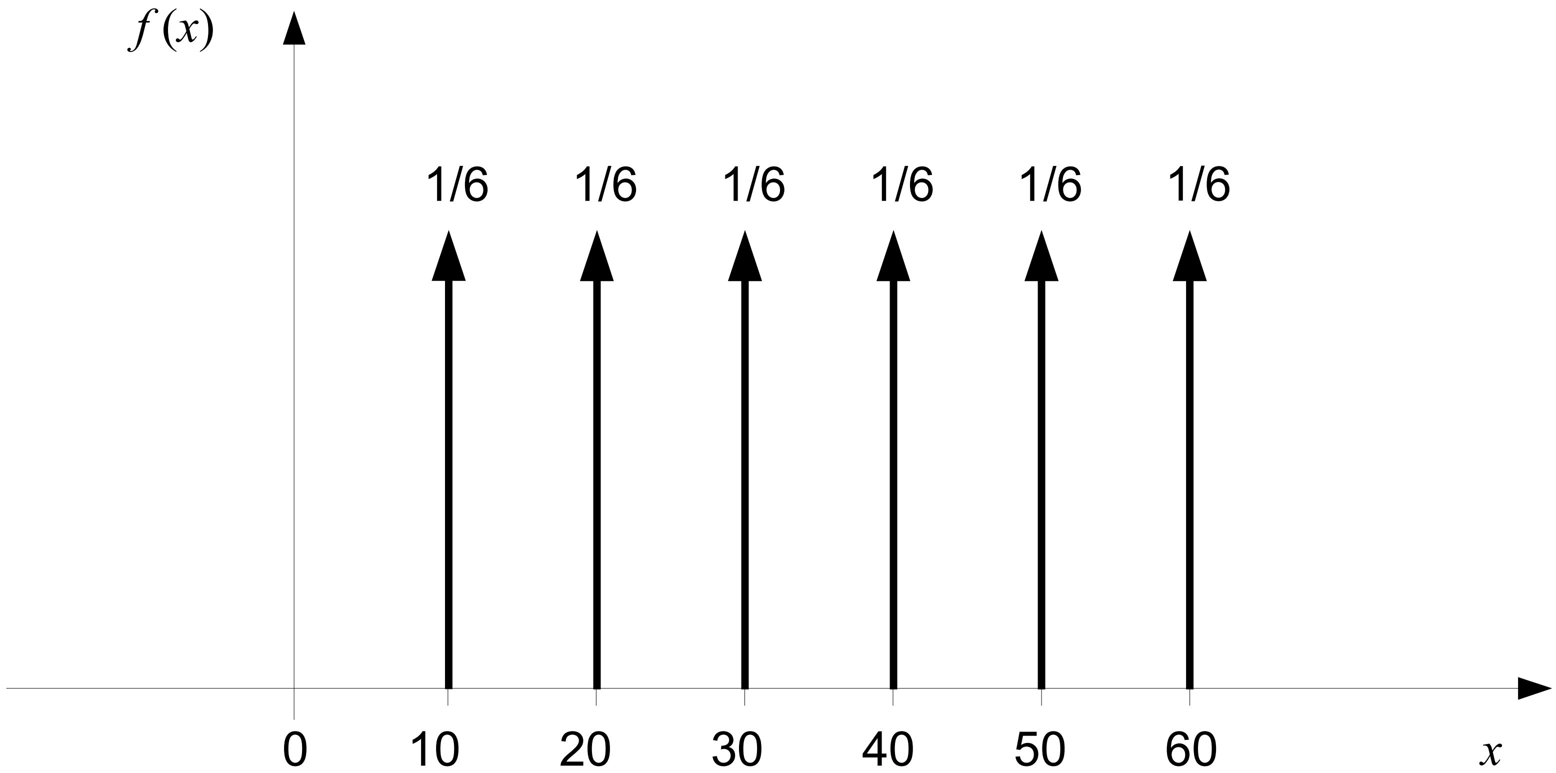
- A derivada de $F(x)$ é chamada de função densidade de probabilidade da variável aleatória $\mathbf{x}$:

$$f(x) = \frac{dF(x)}{dx}$$

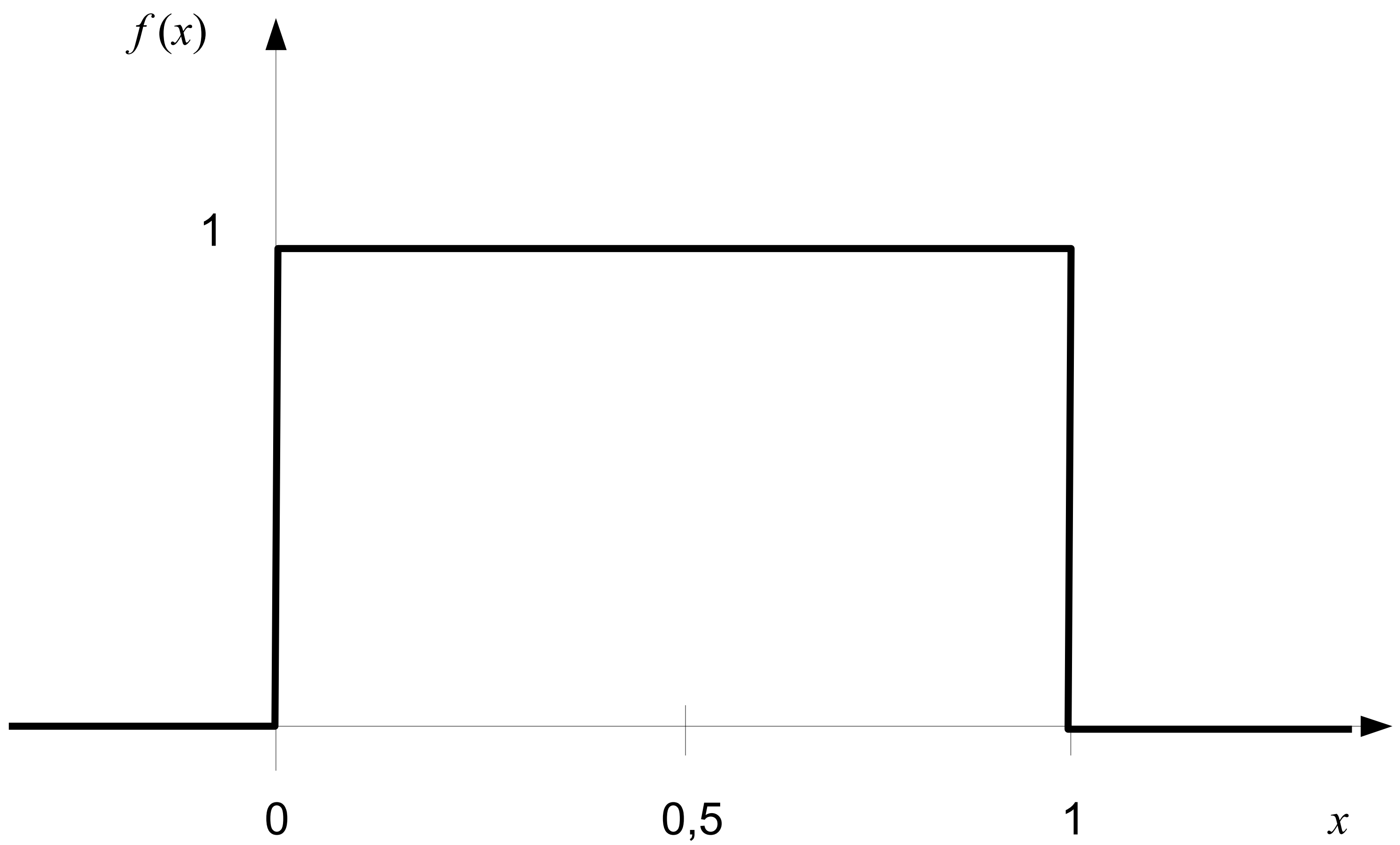
- Se a variável aleatória é do tipo discreto, com os valores x_i tendo probabilidade p_i, então

$$f(x) = \sum_i p_i \delta(x - x_i)$$

Exemplo: arremesso de um dado



Exemplo: telefonema



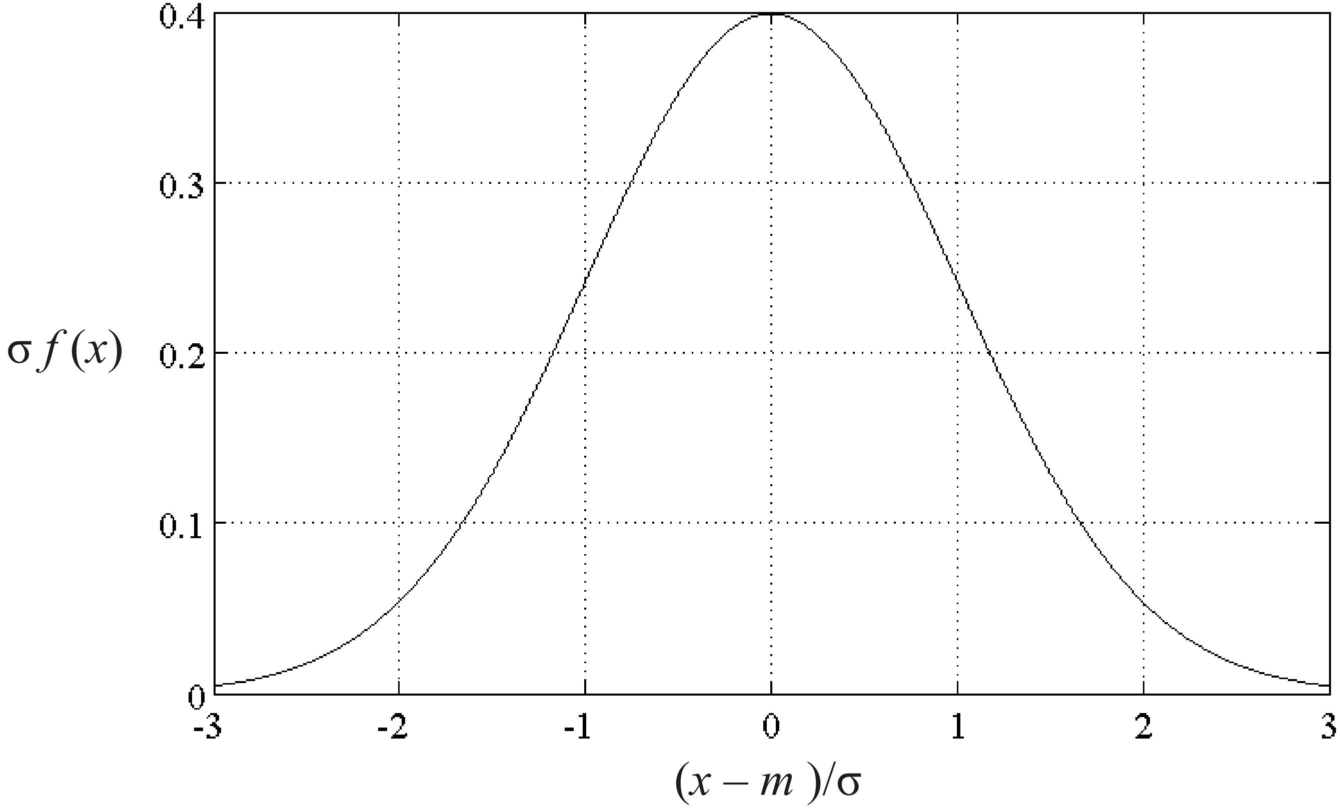
Propriedades

- $f(x) \geqslant 0$
- $F(x) = \int_{-\infty}^{x} f(\alpha)\, d\alpha$
- $\int_{-\infty}^{\infty} f(\alpha)\, d\alpha = 1$
- $P(x_1 < \mathbf{x} \leq x_2) = \int_{]x_1[}^{[x_2]} f(x)\, dx$
- $P(x \leq \mathbf{x} < x + dx) = f(x)\, dx$ (para variáveis contínuas)

Normal ou gaussiana

- Uma variável aleatória **x** é denominada normal ou gaussiana se a sua densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\sigma}\, g\left(\frac{x-m}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-m)^2/2\sigma^2} = N(m, \sigma)$$

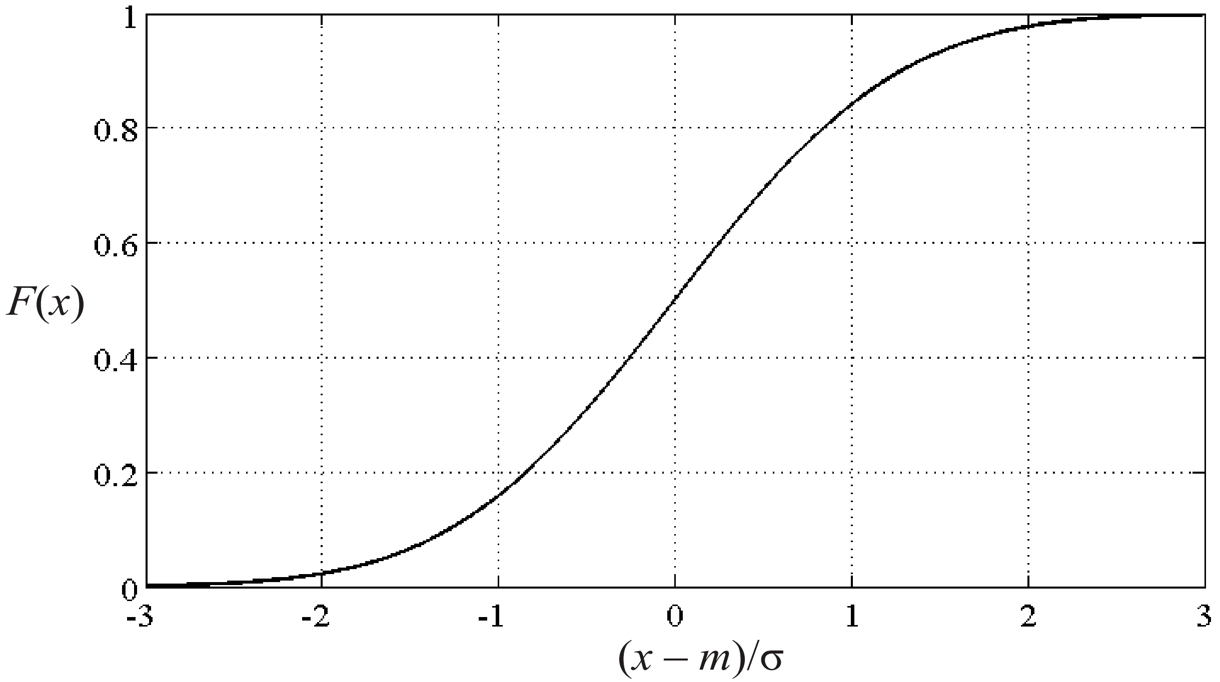


Normal ou gaussiana

- A correspondente função distribuição de probabilidade é dada por

$$F(x) = G\left(\frac{x-m}{\sigma}\right)$$

- onde $G(x)$ é a integral da função $g(x)$, cujos valores podem ser obtidos da tabela fornecida.



Normal ou gaussiana

- Essa tabela lista os valores da função $Q(x)$ na faixa de 0 a 10 em incrementos de $0{,}01$, sendo

$$Q(x)=\frac{1}{\sqrt{2\pi}}\int_{x}^{\infty} e^{-\alpha^2/2}\,d\alpha$$

- Para achar $Q(5{,}36)$, por exemplo, procure a linha que começa com x=$5{,}3$. A coluna correspondente a $0{,}06$ conterá o valor desejado: $4{,}1611\text{e-}08$.

- A função $G(x)$, pode ser obtida por $G(x) = 1 - Q(x)$

Exemplo: distribuição gaussiana

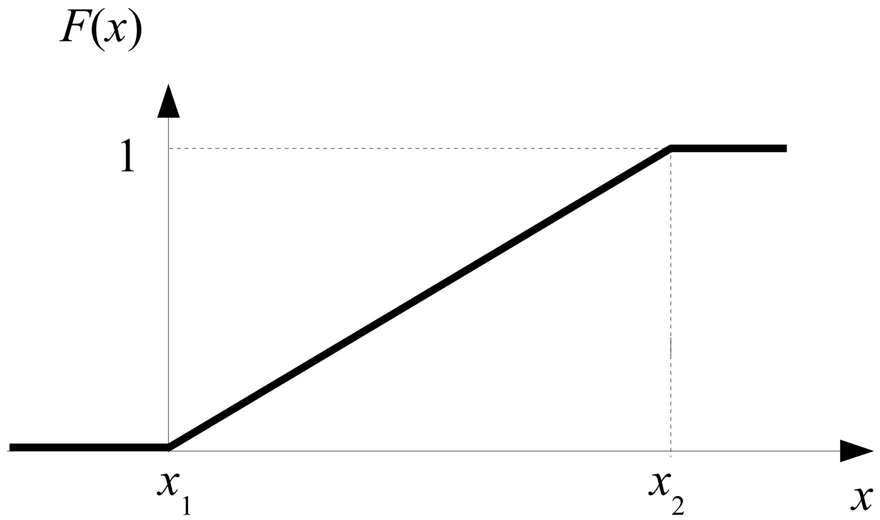
- Uma variável aleatória tem distribuição normal, com $m = 1000$ e $\sigma = 50$. Qual a probabilidade de $\mathbf{x}$ estar entre 900 e 1050?
  - $P(900 \leq \mathbf{x} \leq 1050) = F(1050) - F(900) =$
  - $= G((1050 - 1000)/50) - G((900 - 1000)/50) =$
  - $= 1 - Q((1050 - 1000)/50) - [1 - Q((900 - 1000)/50)] =$
  - $= Q(-2) - Q(1) = 1 - Q(2) - Q(1) =$
  - $= \ 1 - 0{,}022750 - \ 0{,}15866 = \ 0{,}81859$

Variável aleatória uniforme

- Uma variável aleatória $\mathbf{x}$ é denominada uniforme entre x_1 e x_2 se sua densidade é constante no intervalo (x_1, x_2), e zero nos outros intervalos:

$$f(x)=\begin{cases} \dfrac{1}{x_2-x_1} & x_1\leq x\leq x_2 \\ 0 & \text{outros pontos} \end{cases}$$

- A correspondente função distribuição será, portanto, uma rampa, como a mostrada no exemplo 4:



Exemplo:

- A função densidade de probabilidade da amplitude de um certo sinal é dada por

$$f_{\mathbf{x}}(x) = x\ e^{-x}\ u(x)$$

- Determine a probabilidade de que $1 < \mathbf{x} < 2$.

$$P(1 < \mathbf{x} < 2) = \int_1^2 f_{\mathbf{x}}(x)\,dx = \int_1^2 x\ e^{-x} dx =$$
$$= -e^{-x}(x+1)\Big|_1^2 = e^{-1}(1+1) - e^{-2}(2+1) =$$
$$= 0{,}330$$

Revisão: vetores, matrizes e espaços lineares

- Uma matriz A $(n \times k)$ é definida como uma tabela de n linhas e k colunas da seguinte forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

- O elemento a_{ij} é o elemento localizado na intersecção da linha i com a coluna j.

- A transposta de uma matriz A $(n \times k)$ é a matriz A^T, obtida trocando-se as linhas por colunas em A, ou seja,

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{nk} \end{bmatrix}$$

Revisão: vetores, matrizes e espaços lineares

- Chamamos uma matriz $(k \times 1)$, com apenas uma coluna, de vetor, com a notação a seguir:

$$\underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

- A transposta de um vetor é denotada por $\underline{b}^T$, e é uma matriz $(1 \times k)$ com uma única linha.
- Duas matrizes, A e B, são iguais se e somente se elas têm as mesmas dimensões e $a_{ij}=b_{ij}$ para todo i e j.
- A soma de duas matrizes A e B $(n \times k)$, é uma nova matriz C $(n \times k)$ cujos elementos são dados por $c_{ij}=a_{ij}+b_{ij}$ para todo i e j. $C = A+B$ só é definido se A e B têm as mesmas dimensões.
- O produto escalar de uma matriz por uma constante c é uma nova matriz B, $B = c\ A$, onde $b_{ij}=c\ a_{ij}$. A multiplicação é comutativa.

Revisão: vetores, matrizes e espaços lineares

- O produto de duas matrizes $A\ (n\times k)$ e $B\ (k\times m)$ é uma terceira matriz $C\ (n\times m)$, $C = A\ B$ cujos elementos são dados por

$$c_{ij}=\sum_{l=1}^{k} a_{il} b_{lj}$$

- Note que o número de colunas da primeira matriz deve ser igual ao número de linhas da segunda, e que o produto não é comutativo, nem mesmo para matrizes quadradas.

- Note, também, que dados 2 vetores, $\underline{x}$ e $\underline{y}$ de mesma dimensão n, $\underline{x}\ \underline{y}^{\mathrm{T}} = A\ (n\times n)$, com $a_{ij} = x_i y_j$, e que

$$\underline{x}^T \underline{y}=\sum_{i=1}^{n} x_i y_i$$

Revisão: vetores, matrizes e espaços lineares

- $A(B+C)=AB+AC$ $(AB)C=A(BC)$ $AB\neq BA$ em geral
- $(AB)^T = B^T A^T$
- sendo $\underline{x}$ $(k\times 1)$, $\underline{y}$ $(k\times 1)$ e A $(k\times k)$: $\underline{x}^T A \underline{y}=\sum_{i=1}^{k}\sum_{j=1}^{k} x_i a_{ij} y_j$
- Dada uma matriz A $(k\times k)$, sua inversa A^{-1} é definida como sendo uma matriz $(k\times k)$ tal que $A\,A^{-1}=A^{-1}A=I$, onde I é a matriz diagonal com elementos unitários em sua diagonal e elementos nulos nas outras posições. A matriz I é denominada matriz identidade pois $C\,I=I\,C = C$.
- $(A^{-1})^{-1}= A$
- $\det(A)=|A|=\sum_{i=1}^{n} a_{ij} C_{ij}$

  onde C_{ij} é o cofator de a_{ij} definido por
  $C_{ij}=(-1)^{i+j}|M_{ij}|$
  sendo M_{ij} a matriz obtida de A eliminando-se a sua i-ésima linha e a sua j-ésima coluna.

Revisão: vetores, matrizes e espaços lineares

- Se A é uma matriz real $n \times n$, então $\det(A)$ pode ser interpretado como o hiper-volume do paralelepípedo formado pelos vetores coluna dessa matriz no espaço $\mathbf{R}^n$, com um sinal ($+$ ou $-$) que depende da orientação relativa desses vetores.
- Quando A não admite inversa ela é denominada singular. Esta situação ocorre sempre que as linhas (ou colunas) de A forem linearmente dependentes, o que implica em $\det(A) = 0$.
- Sendo $B = A^{-1}$, temos que $b_{ij} = (-1)^{i+j} \dfrac{|A_{ji}|}{|A|}$

  sendo A_{ji} a matriz obtida de A eliminando-se a j-ésima linha e a i-ésima coluna, e $|A_{ji}|$ o seu determinante.
- $(AB)^{-1} = B^{-1}A^{-1}, \quad (A^{-1})^T = (A^T)^{-1}$.

Vetores aleatórios

- Dadas duas variáveis aleatórias, $\mathbf{x}$ e $\mathbf{y}$, definidas como anteriormente, os eventos $\{\mathbf{x} \leq x\}$ e $\{\mathbf{y} \leq y\}$ têm probabilidades $P(\mathbf{x} \leq x) = F_{\mathbf{x}}(x)$ e $P(\mathbf{y} \leq y) = F_{\mathbf{y}}(y)$, onde $F_{\mathbf{x}}$ e $F_{\mathbf{y}}$ são as funções distribuição de probabilidade das variáveis aleatórias $\mathbf{x}$ e $\mathbf{y}$.
- O produto $\{\mathbf{x} \leq x\}\{\mathbf{y} \leq y\} = \{\mathbf{x} \leq x, \mathbf{y} \leq y\}$ que consiste de todos os resultados ξ do experimento tais que $\mathbf{x}(\xi) \leq x$ e $\mathbf{y}(\xi) \leq y$ é também um evento. A probabilidade desse evento é função de x e y e é denominada função de distribuição conjunta das variáveis aleatórias $\mathbf{x}$ e $\mathbf{y}$.
- O conceito anterior pode ser estendido para um número n qualquer de variáveis aleatórias, sendo a coleção dessas variáveis aleatórias $\mathbf{x}_i$ $(i = 1,2,\ldots,\mathrm{n})$ denominada de vetor aleatório $\underline{\mathbf{x}} \in \mathbf{R}^{\mathrm{n}}$:

$$\underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_\mathbf{n} \end{bmatrix}^T$$

Vetores aleatórios

- Exemplo: sorteia-se uma pessoa da população e associa-se ao resultado três variáveis aleatórias: sua idade, seu peso e sua altura.
- Obtém-se, assim, um vetor aleatório:

$$\underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix}^T$$

- Note que essas três variáveis, ao serem analisadas de forma conjunta, nos permitem determinar relações entre elas:
  - Pessoas mais velhas são mais altas?
  - Pessoas mais altas são mais pesadas?

Distribuições e densidades conjuntas

- Para duas variáveis aleatórias $\mathbf{x}$ e $\mathbf{y}$, define-se uma função distribuição de probabilidade conjunta $F_{\mathbf{xy}}(x,y)$ como

$$P(\mathbf{x}\leq x \text{ e } \mathbf{y}\leq y)=F_{\mathbf{xy}}(x,y)$$

- e a densidade conjunta como

$$f_{\mathbf{xy}}(x,y)=\frac{\partial^2}{\partial x\,\partial y}F_{\mathbf{xy}}(x,y)$$

- Segue-se portanto que

$$P(\mathbf{x}\leq x \text{ e } \mathbf{y}\leq y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f(\alpha,\beta)\,d\beta\,d\alpha$$

- e assim,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\alpha,\beta)\,d\beta\,d\alpha=1$$

Distribuições e densidades conjuntas

- Para o caso de mais de duas variáveis aleatórias, podemos definir funções distribuição e densidade de probabilidade para esse vetor como

$$F_{\mathbf{x}}(x_1, x_2, \ldots, x_n) = F_{\mathbf{x}}(\underline{x}) = P(\mathbf{x_1} \leq x_1, \mathbf{x_2} \leq x_2, \ldots, \mathbf{x_n} \leq x_n)$$

$$f_{\underline{\mathbf{x}}}(\underline{x}) = \frac{\partial^n F_{\underline{\mathbf{x}}}(\underline{x})}{\partial x_1 \partial x_2 \ldots \partial x_n}$$

- No nosso exemplo de idade/peso/altura essas funções nos permitem determinar as relações entre essa características das pessoas.

Distribuições e densidades conjuntas

- Propriedades:
  - $F_{\underline{\mathbf{x}}}(\infty,\infty,\ldots,\infty)=1$
  - $\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} f_{\mathbf{x}}(\underline{x})\,dx_1\,dx_2\ldots dx_n=1$
  - $f_{\underline{\mathbf{x}}}(\underline{x})\geq 0, \quad \underline{x}\in R^n$
  - $F_{\underline{\mathbf{x}}}$ é uma função não decrescente de todas as variáveis;
  - dada uma região D do espaço n-dimensional, a probabilidade do ponto com coordenadas $\mathbf{x_1},\mathbf{x_2},\ldots,\mathbf{x_n}$ estar em D é

$$P(\underline{\mathbf{x}}\in D)=\int\underset{D}{\cdots}\int f(x_1,x_2,\cdots,x_n)\,dx_1\,dx_2\cdots dx_n$$

Distribuições e densidades conjuntas

- Uma sequência finita de n variáveis aleatórias estará estatisticamente determinada se a sua função de distribuição conjunta for conhecida.
- Pode-se afirmar, também, que uma sequência infinita de variáveis aleatórias estará estatisticamente determinada se a função de distribuição conjunta de quaisquer n delas for conhecida para qualquer n.
- Quando estamos usando duas ou mais variáveis aleatórias, as distribuições e densidades individuais de cada uma delas podem ser obtidas a partir da densidade conjunta.
- Essas distribuições e densidades passam, então, a se chamar distribuições e densidades marginais, podendo ser calculadas por

$$F_{\mathbf{x_i}}(x_i) = F_{\underline{\mathbf{x}}}(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty)$$

$$f_{\mathbf{x_i}}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{x}}(x_1, \cdots, x_n)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n$$

Distribuições e densidades conjuntas

- De forma análoga, pode-se obter a distribuição marginal de apenas k das n variáveis aleatórias. Por exemplo, para $n=4$,

$$F_{\mathbf{x}_1\mathbf{x}_3}(x_1, x_3) = F_{\underline{\mathbf{x}}}(x_1, \infty, x_3, \infty)$$

$$f_{\mathbf{x}_1\mathbf{x}_3}(x_1, x_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{x}}(x_1, x_2, x_3, x_4)\, dx_2\, dx_4$$

Distribuições e densidades condicionais

- A distribuição condicional de uma variável aleatória $\mathbf{y}$, dado um evento M, pode ser definida por

$$F_{\mathbf{y}|M}(y|M)=P(\mathbf{y}\leq y\ |M)=\frac{P(\mathbf{y}\leq y\ ,M)}{P(M)}\ ,\quad P(M)>0$$

- vamos expressar M em termos da variável aleatória $\mathbf{x}$.

- Supondo-se, inicialmente,

$$M=\{\mathbf{x}\leq x\}$$

temos que

$$P\{M\}=P\{\mathbf{x}\leq x\}=F_{\mathbf{x}}(x)$$

e, portanto,

$$F_{\mathbf{y}}(y\ |\mathbf{x}\leq x)=\frac{P(\mathbf{y}\leq y\ ,\mathbf{x}\leq x)}{P(\mathbf{x}\leq x)}=\frac{F_{\mathbf{xy}}(x,y)}{F_{\mathbf{x}}(x)}$$

e

$$f_{\mathbf{y}}(y\ |\mathbf{x}\leq x)=\frac{\partial F_{\mathbf{xy}}(x,y)/\partial y}{F_{\mathbf{x}}(x)}=\frac{\int_{-\infty}^{x} f_{\mathbf{xy}}(\xi,y)\,d\xi}{\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{\mathbf{xy}}(\xi,y)\,d\xi\,dy}$$

Distribuições e densidades condicionais

- Admitindo-se, agora,

$$M=\{x_1<\mathbf{x}\le x_2\} \qquad P(M)=F_{\mathbf{x}}(x_2)-F_{\mathbf{x}}(x_1)$$

e então:

$$F_{\mathbf{y}}(y\,|\,x_1<\mathbf{x}\le x_2)=\frac{P(\mathbf{y}\le y\ ,x_1<\mathbf{x}\le x_2)}{P(x_1<\mathbf{x}\le x_2)}=$$

$$=\frac{F_{\mathbf{xy}}(x_2,y)-F_{\mathbf{xy}}(x_1,y)}{F_{\mathbf{x}}(x_2)-F_{\mathbf{x}}(x_1)}=\frac{\int_{-\infty}^{y}\int_{x_1}^{x_2} f_{\mathbf{xy}}(x,\zeta)\,dx\,d\zeta}{\int_{-\infty}^{\infty}\int_{x_1}^{x_2} f_{\mathbf{xy}}(x,\xi)\,dx\,d\xi}$$

$$f_{\mathbf{y}}(y\,|\,x_1<\mathbf{x}\le x_2)=\frac{\int_{x_1}^{x_2} f_{\mathbf{xy}}(x,y)\,dx}{\int_{-\infty}^{\infty}\int_{x_1}^{x_2} f_{\mathbf{xy}}(x,y)\,dx\,dy}$$

Distribuições e densidades condicionais

- E para $M=\{\mathbf{x}=x\}$:

$$F_{\mathbf{y}}(y\,|\mathbf{x}=x)=F_{\mathbf{y}|\mathbf{x}}(y\,|x)=\lim_{\Delta x\to 0} F_{\mathbf{y}}(y\,|x<\mathbf{x}\le x+\Delta x) =$$

$$=\lim_{\Delta x\to 0}\frac{F_{\mathbf{xy}}(x+\Delta x,y)-F_{\mathbf{xy}}(x,y)}{F_{\mathbf{x}}(x+\Delta x)-F_{\mathbf{x}}(x)}=\frac{\partial F_{\mathbf{xy}}(x,y)/\partial x}{dF_{\mathbf{x}}(x)/dx}=\frac{\int_{-\infty}^{y} f_{\mathbf{xy}}(x,\zeta)\,d\zeta}{f_{\mathbf{x}}(x)}$$

- e

$$f_{\mathbf{y}|\mathbf{x}}(y|x)=\frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{x}}(x)}=\frac{f_{\mathbf{xy}}(x,y)}{\int_{-\infty}^{\infty} f_{\mathbf{xy}}(x,y)\,dy}$$

- De forma análoga:

$$f_{\mathbf{x}|\mathbf{y}}(x|y)=\frac{f_{\mathbf{xy}}(x,y)}{f_{\mathbf{y}}(y)}$$

Distribuições e densidades condicionais

- Outras relações úteis, que podem ser obtidas das anteriores, são:

$$f_{\mathbf{x}|\mathbf{y}}(x|y)=\frac{f_{\mathbf{y}|\mathbf{x}}(y|x)\,f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)}$$

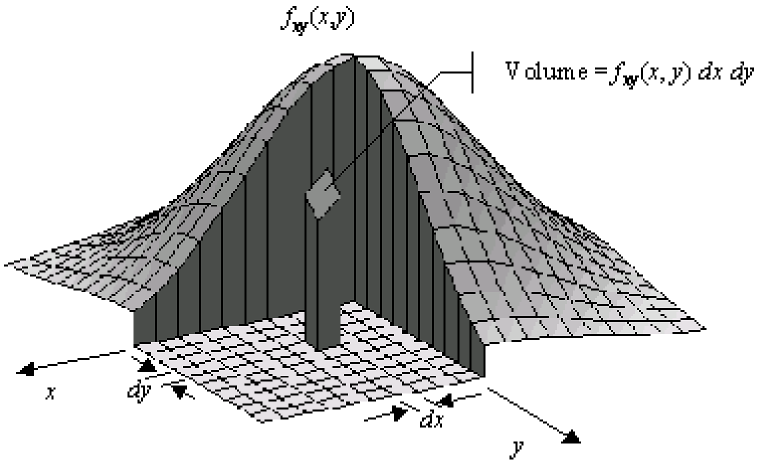
$$f_{\mathbf{y}}(y)=\int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{x}}(y|x)\,f_{\mathbf{x}}(x)\,dx$$

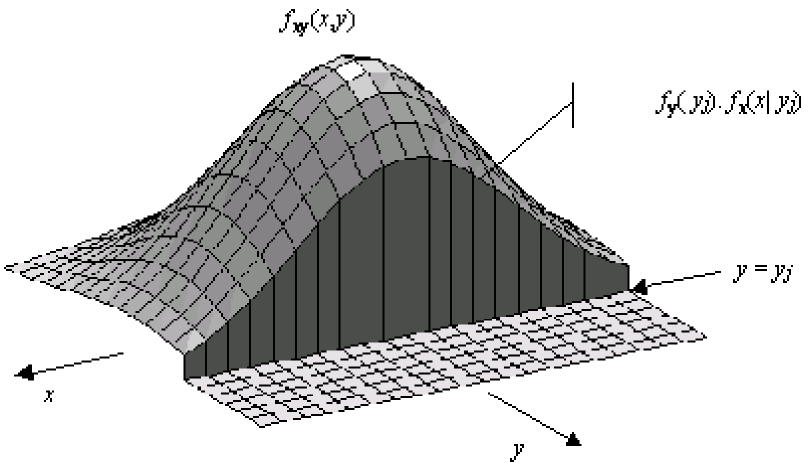
- No caso de mais de duas variáveis aleatórias temos:

$$f_{\mathbf{x}_1\cdots\mathbf{x}_k|\mathbf{x}_{k+1}\mathbf{x}_n}(x_1,\cdots,x_k|x_{k+1},\cdots,x_n)=\frac{f(x_{1,}\cdots,x_k,x_{k+1},\cdots,x_n)}{f(x_{k+1},\cdots,x_n)}$$

$$f(x_1,\cdots,x_n)=f(x_n|x_{n-1},\cdots,x_1)\,f(x_{n-1}|x_{n-2},\cdots,x_1)\ldots f(x_2|x_1)\,f(x_1)$$

Distribuições e densidades condicionais





Variáveis aleatórias independentes

- As variáveis aleatórias $\mathbf{x}_1, \ldots, \mathbf{x}_n$ são ditas independentes se os eventos $\{\mathbf{x}_1 \leq x_1\}, \ldots, \{\mathbf{x}_n \leq x_n\}$ são independentes para quaisquer $x_1, \ldots, x_n$.
- Tem-se, então, que

$$F_{\underline{\mathbf{x}}}(x_1, \cdots, x_n) = F_{\mathbf{x}_1}(x_1) \ldots F_{\mathbf{x}_n}(x_n)$$

- e

$$f_{\mathbf{x}}(x_1, \cdots, x_n) = f_{\mathbf{x}_1}(x_1) \ldots f_{\mathbf{x}_n}(x_n)$$

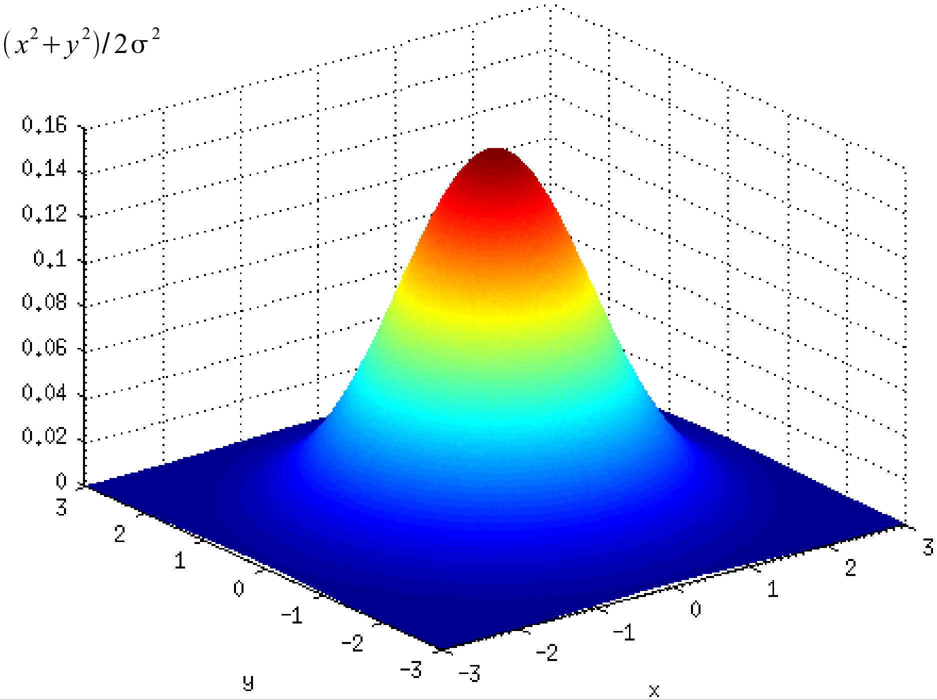
Exemplo: gaussianas e Rayleigh

- Considere duas variáveis aleatórias gaussianas independentes $\mathbf{x}$ e $\mathbf{y}$, com mesma densidade:

$$f_{\mathbf{x}}(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2} \quad \text{e} \quad f_{\mathbf{y}}(y)=\frac{1}{\sigma\sqrt{2\pi}}e^{-y^2/2\sigma^2}$$

- Como elas são independentes:

$$f_{\mathbf{xy}}(x,y)=f_{\mathbf{x}}(x)\,f_{\mathbf{y}}(y)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}$$

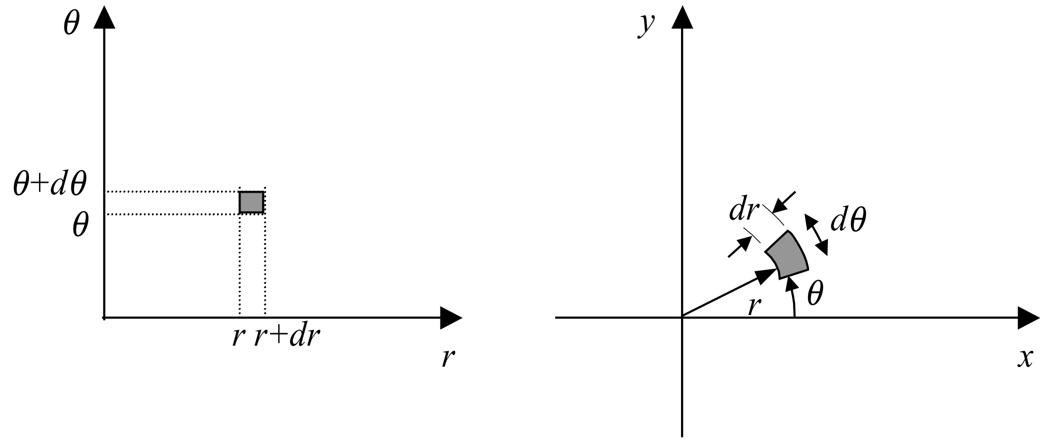


Exemplo: gaussianas e Rayleigh

- Os pontos no plano x, y podem também ser descritos em coordenadas polares (r, θ), com

$$r=\sqrt{x^2+y^2} \quad \text{e} \quad \theta=\tan^{-1}\frac{y}{x}$$

- Assim, podemos definir as variáveis aleatórias $\mathbf{r}$ e $\boldsymbol{\theta}$ e calcular sua densidade de probabilidade:



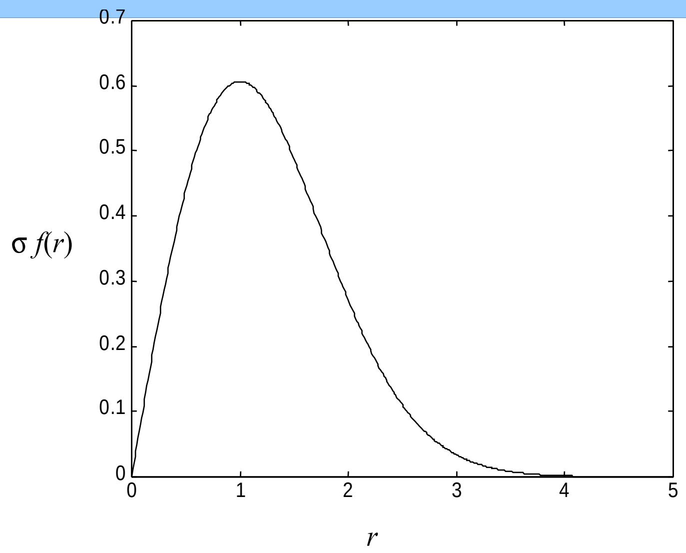
$$f_{\mathbf{r}\boldsymbol{\theta}}(r,\theta)\,dr\,d\theta=f_{\mathbf{xy}}(x,y)\,dr\;r\,d\theta \Rightarrow f_{\mathbf{r}\boldsymbol{\theta}}(r,\theta)=\frac{r}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}=\frac{r}{2\pi\sigma^2}e^{-r^2/2\sigma^2}$$

Exemplo: gaussianas e Rayleigh

- Dessa forma:

$$f_{\mathbf{r}}(r)=\int_{-\infty}^{\infty} f_{\mathbf{r}\boldsymbol{\theta}}(r,\theta)\, d\theta=\frac{r}{\sigma^2}e^{-r^2/2\sigma^2} \quad r\geq 0$$

  - Distribuição de Rayleigh.



- e

$$f_{\boldsymbol{\theta}}(\theta)=\int_{-\infty}^{\infty} f_{\mathbf{r}\boldsymbol{\theta}}(r,\theta)\, dr=\frac{1}{2\pi} \quad 0\leq\theta\leq 2\pi$$

  - Distribuição uniforme.

- Note que **r** e **θ** são, também, independentes.

Exemplo: detecção de pulsos binários

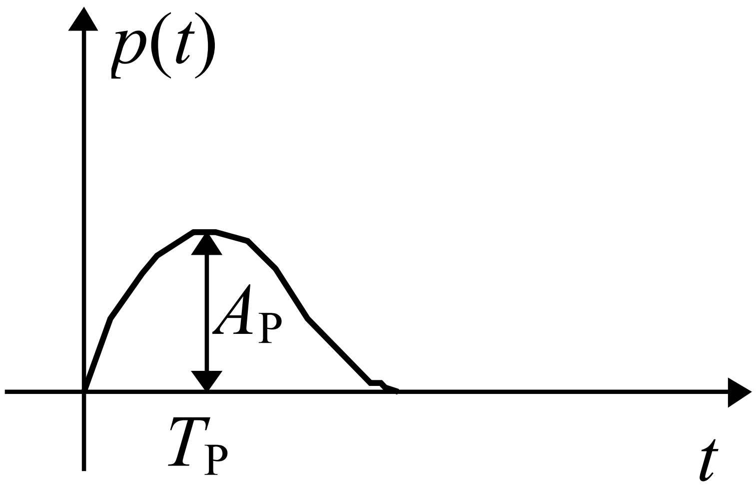
- Num certo canal binário, mensagens 1 e 0 são transmitidas com igual probabilidade usando um pulso positivo e um pulso negativo, respectivamente.
- O pulso recebido correspondente a 1 é $p(t)$, mostrado abaixo, e o correspondente pulso para 0 é $-p(t)$. Seja A_p a amplitude de $p(\mathrm{t})$ no instante $t = T_p$.
- Devido ao ruído no canal, os pulsos recebidos serão $(+/-)\ p(t) + \mathbf{n}(t)$.
- Para detectar os pulsos no receptor, cada pulso é amostrado no instante de seu pico. Na ausência de ruído, a saída do amostrador seria A_p ou $-A_\mathrm{p}$.
- Devido ao ruído, a saída amostrada será $(+/-)\ A_\mathrm{p}+\mathbf{n}$, onde $\mathbf{n}$, a amplitude do ruído no instante da amostragem, é uma variável aleatória.

Exemplo: detecção de pulsos binários

- Para ruído gaussiano, a função densidade de probabilidade de **n** é:

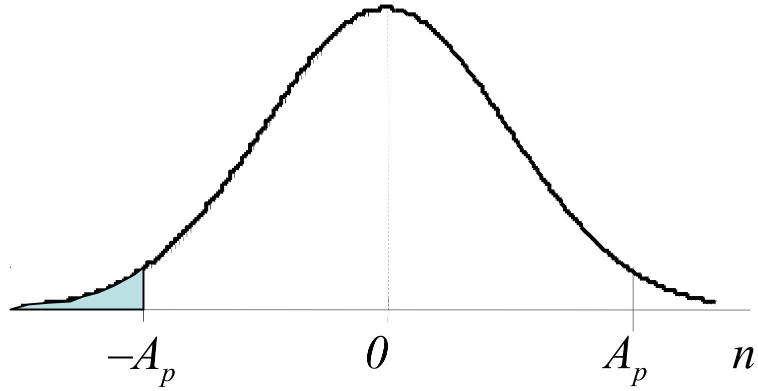
$$f_n(n) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{n^2}{2\sigma_n^2}\right)$$

- Devido à simetria, a detecção ótima é zero, ou seja, a mensagem será detectada como 1 ou 0 dependendo do valor da amostra ser positivo ou negativo.
- Determine o valor da probabilidade de erro neste canal simétrico.



Exemplo: detecção de pulsos binários

- $P_{\mathbf{y}|\mathbf{x}}(0|1) = P(\mathbf{n} < -A_P) = F_{\mathbf{n}}(-A_P) = Q(A_P/\sigma_{\mathbf{n}})$



- Da mesma forma, $P_{\mathbf{y}|\mathbf{x}}(1|0) = P(\mathbf{n} > A_P) = 1 - F_{\mathbf{n}}(A_P) = Q(A_P/\sigma_{\mathbf{n}})$
- Portanto, $P_e = P_{\mathbf{y}|\mathbf{x}}(0|1)\, P_{\mathbf{x}}(1) + P_{\mathbf{y}|\mathbf{x}}(1|0)\, P_{\mathbf{x}}(0) = Q(A_P/\sigma_{\mathbf{n}})$