

# Apostila de Comunicações Digitais

## Capítulo 1

### Introdução à Probabilidade e Processos Estocásticos

Prof. André Noll Barreto

Universidade de Brasília

Rev. 1.1

Mar/2017

# 1. Noções de Probabilidade

## a. Experimento aleatório

O estudo de probabilidade tem como base a realização de um experimento aleatório, ou seja, um experimento ou observação sobre o qual não se conhece o resultado a priori. O experimento terá uma série de **resultados** (*outcomes*) possíveis  $\xi_i$ , e o conjunto de resultados é chamado de **espaço amostral** (*sample space*)  $\mathcal{S}$ .

Exemplo 1  
Jogar uma moeda: temos dois resultados possíveis, cara ou coroa, que podem ser representados por  $\xi_1$  e  $\xi_2$ , respectivamente. O mesmo vale para gerar um bit aleatório, 0 ou 1.

Exemplo 2  
Jogar um dado: temos seis resultados possíveis, que podem ser representados pelo conjunto  $\mathcal{S} = \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6\}$ . O conjunto  $\mathcal{S}$ , contendo todos os resultados possíveis é chamado de espaço amostral.

Exemplo 3  
Tirar uma carta do baralho: temos 53 resultados possíveis,  $\mathcal{S} = \{\xi_1, \xi_2, \xi_3, \xi_4, \dots, \xi_{53}\}$ , representando as 13 cartas de cada um dos 4 naipes, com  $\xi_1$  a  $\xi_{13}$  representando as cartas de ouros,  $\xi_{14}$  a  $\xi_{26}$  as cartas de copas,  $\xi_{27}$  a  $\xi_{39}$  as cartas de espadas,  $\xi_{40}$  a  $\xi_{52}$  as cartas de paus, além do coringa,  $\xi_{53}$ . Note que os primeiros estudos de probabilidade tinham como objetivo o estudo de jogos de azar nos séculos XVI e XVII.

O espaço amostral pode ser também de tamanho infinito.

Exemplo 4

O número de gols do Flamengo em um ano pode ser considerado como o resultado de um experimento aleatório, com  $\mathcal{S} = \mathbb{N}$ .

Exemplo 5

Um **evento** é qualquer subconjunto do espaço amostral  $\mathcal{S}$ .

Considerando o experimento de jogar um dado, podemos definir diferentes eventos:

- Tirar um número par, representado pelo subconjunto

$$\mathcal{A}_e = \{\xi_2, \xi_4, \xi_6\} \subset \mathcal{S}$$

- Tirar um número ímpar, representado pelo subconjunto

$$\mathcal{A}_o = \{\xi_1, \xi_3, \xi_5\} \subset \mathcal{S}$$

- Tirar um número  $\leq 3$ , representado pelo subconjunto

$$\mathcal{A}_3 = \{\xi_1, \xi_2, \xi_3\} \subset \mathcal{S}$$

Estes eventos são mostrados na Figura 1

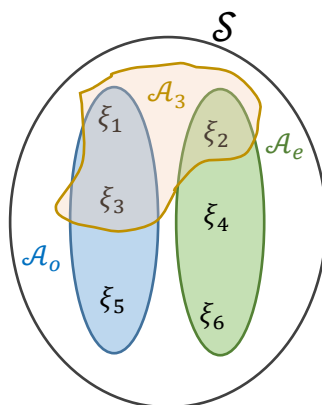


Figura 1. Exemplo de diagrama de Venn, no experimento de jogar os dados

Podemos construir novos eventos fazendo operações entre conjuntos.

Exemplo 6

$$\mathcal{B}_1 = \mathcal{A}_e \cup \mathcal{A}_3 = \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_6\}$$

Note que a operação União  $\cup$  representa o operador Booleano “ou”, ou seja o evento  $\mathcal{B}_1$  significa o conjunto de resultados pares ou menores ou iguais a 3.

$$\mathcal{B}_2 = \mathcal{A}_o \cap \mathcal{A}_3 = \{\xi_1, \xi_3\}$$

Da mesma forma, a operação Interseção  $\cap$  representa o operador Booleano “e”, ou seja o evento  $\mathcal{B}_2$  significa o conjunto de resultados ímpares e menores ou iguais a 3.

$$\mathcal{B}_3 = \bar{\mathcal{A}}_3 = \{\xi_4, \xi_5, \xi_6\},$$

e a operação complemento representa a negação, ou seja o evento  $\mathcal{B}_3$  representa o conjunto dos resultados maiores que 3.

Como todos os subconjuntos são eventos, podemos também dizer que o conjunto vazio  $\emptyset$  e próprio espaço amostral  $\mathcal{S}$  são eventos aleatórios.

Exemplo 7

Se dois eventos  $\mathcal{A}$  e  $\mathcal{B}$  são tais que  $\mathcal{A} \cap \mathcal{B} = \emptyset$ , dizemos que estes eventos são **disjuntos**.

No experimento de jogar os dados, os eventos  $\mathcal{A}_e$  e  $\mathcal{A}_o$  são disjuntos.

## b. Probabilidade

Dada a realização de  $N$  repetições independentes de um experimento aleatório, podemos verificar  $n(\mathcal{A})$  vezes a ocorrência de um evento  $\mathcal{A}$ . A partir disso podemos calcular a frequência relativa deste evento como

$$f(\mathcal{A}, N) = \frac{n(\mathcal{A})}{N} \tag{1}$$

A probabilidade do evento  $P(\mathcal{A})$  deve estar idealmente associada à sua frequência relativa, de modo que

$$P(\mathcal{A}) = \lim_{N \rightarrow \infty} f(\mathcal{A}, N), \quad (2)$$

ou seja, a probabilidade tende à frequência relativa. Claro que, normalmente, não é possível realizar um número muito grande de experimentos e conhecer este valor exatamente, mas, a partir de conhecimento sobre a natureza do experimento podemos definir valores realistas para a probabilidade.

Considerando um dado não viciado, temos que

Exemplo 8 
$$P(\xi_1) = P(\xi_2) = P(\xi_3) = P(\xi_4) = P(\xi_5) = P(\xi_6) = \frac{1}{6}.$$

Considerando a retirada de uma carta de um baralho ideal, e lembrando que  $\xi_{53}$  corresponde ao coringa

$$P(\xi_1) = P(\xi_2) = \dots = P(\xi_{52}) = \frac{1}{54}$$

$$P(\xi_{53}) = \frac{2}{54},$$

Ou seja, resultados distintos podem ter probabilidades diferentes.

A partir da definição de probabilidade pela frequência relativa, temos as seguintes propriedades

$$\begin{aligned} 0 &\leq P(\mathcal{A}) \leq 1 \\ P(\emptyset) &= 0 \\ P(\mathcal{S}) &= 1 \\ P(\bar{\mathcal{A}}) &= 1 - P(\mathcal{A}) \end{aligned} \quad (3)$$

Se um evento  $\mathcal{A}$  é composto por um conjunto de resultados  $\mathcal{A} = \{\xi_1, \xi_2, \dots, \xi_n\}$ , podemos ver então que  $f(\mathcal{A}, N) = \sum_{i=1}^n f(\xi_i, N)$ , e, conseqüentemente,

$$P(\mathcal{A}) = \sum_{i=1}^n P(\xi_i) \quad (4)$$

A partir desta definição, e como podemos visualizar na Figura 2, a probabilidade da união de dois eventos é dada por

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}), \quad (5)$$

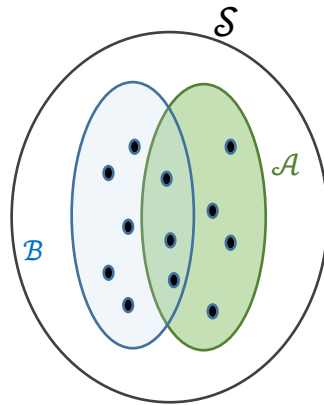


Figura 2. União de 2 eventos

e, se  $\mathcal{A} \cap \mathcal{B} = \emptyset$ , ou seja, são eventos disjuntos,  $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B})$ .

Exemplo 9 Considerando a retirada de uma carta de um baralho ideal, podemos definir o evento tirar uma carta de ouros como  $\mathcal{A}_{ouros} = \{\xi_1, \xi_2, \dots, \xi_{13}\}$ , e

$$P(\mathcal{A}_{ouros}) = P(\xi_1) + P(\xi_2) + \dots + P(\xi_{13}) = \frac{13}{54}.$$

Podemos ainda definir o evento de retirar as cartas numéricas  $\mathcal{A}_{num}$ , com

$$P(\mathcal{A}_{num}) = \frac{36}{54}.$$

Criando agora o evento de cartas de ouros ou numéricas,  $\mathcal{B} = \mathcal{A}_{ouros} \cup \mathcal{A}_{num}$ , temos que

$$P(\mathcal{B}) = P(\mathcal{A}_{ouros}) + P(\mathcal{A}_{num}) - P(\mathcal{A}_{ouros} \cap \mathcal{A}_{num}) = \frac{13}{54} + \frac{36}{54} - \frac{9}{54} = \frac{20}{27}$$

### c. Probabilidade Condicional e Regra de Bayes

Se soubermos que um evento  $\mathcal{A}$  ocorreu, podemos querer saber qual a probabilidade de o evento  $\mathcal{B}$  também ter ocorrido. Esta probabilidade é chamada de probabilidade de  $\mathcal{B}$  dado  $\mathcal{A}$ , e, lembrando da definição de frequência relativa, é definida por

$$P(\mathcal{B}|\mathcal{A}) = \lim_{N \rightarrow \infty} \frac{n(\mathcal{A} \cap \mathcal{B})}{n(\mathcal{A})}. \quad (6)$$

Sabemos ainda que

$$P(\mathcal{A} \cap \mathcal{B}) = \lim_{N \rightarrow \infty} \frac{n(\mathcal{A} \cap \mathcal{B})}{N} = \lim_{N \rightarrow \infty} \frac{n(\mathcal{A})}{N} \frac{n(\mathcal{A} \cap \mathcal{B})}{n(\mathcal{A})}, \quad (7)$$

ou seja,

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B}|\mathcal{A}) \Leftrightarrow P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{A})}. \quad (8)$$

Se soubermos a probabilidade condicional em uma direção, por exemplo,  $P(\mathcal{B}|\mathcal{A})$ , e quisermos achar a condicional na direção inversa, podemos fazer

$$P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} \Leftrightarrow P(\mathcal{A}|\mathcal{B}) = \frac{P(\mathcal{A})P(\mathcal{B}|\mathcal{A})}{P(\mathcal{B})}. \quad (9)$$

Esta expressão é conhecida como **Regra de Bayes**.

Considerando a retirada de uma carta de um baralho ideal, dado que retiramos uma carta de ouros, a probabilidade desta carta ser numérica é

Exemplo 10

$$P(\mathcal{A}_{num}|\mathcal{A}_{ouros}) = \frac{P(\mathcal{A}_{num} \cap \mathcal{A}_{ouros})}{P(\mathcal{A}_{ouros})} = \frac{9/54}{13/54} = \frac{9}{13}$$

e, dado que retiramos uma carta numérica, a probabilidade de ela ser de ouros é

$$P(\mathcal{A}_{ouros}|\mathcal{A}_{num}) = \frac{P(\mathcal{A}_{ouros})P(\mathcal{A}_{num}|\mathcal{A}_{ouros})}{P(\mathcal{A}_{num})} = \frac{13/54 \cdot 9/13}{36/54} = \frac{1}{4}.$$

#### d. Eventos Independentes

Dois eventos  $\mathcal{A}$  e  $\mathcal{B}$  são ditos independentes se

$$P(\mathcal{B}|\mathcal{A}) = P(\mathcal{B}), \quad (10)$$

ou seja, a probabilidade de  $\mathcal{B}$  ocorrer não depende da ocorrência de  $\mathcal{A}$ .

Exemplo 11

Considerando a jogada de um dado, seja o evento  $\mathcal{A}_o$  tirarmos um número ímpar e  $\mathcal{A}_4$  tirarmos um número menor ou igual a 4. Sendo assim

$$P(\mathcal{A}_o|\mathcal{A}_4) = \frac{P(\mathcal{A}_o \cap \mathcal{A}_4)}{P(\mathcal{A}_4)} = \frac{2/6}{4/6} = \frac{1}{2} = P(\mathcal{A}_o),$$

e, portanto, são eventos independentes.

Agora, consideremos os eventos  $\mathcal{A}_o$  e  $\mathcal{A}_3$ , correspondente a tirarmos um número menor ou igual a 3. Neste caso

$$P(\mathcal{A}_3|\mathcal{A}_4) = \frac{P(\mathcal{A}_o \cap \mathcal{A}_3)}{P(\mathcal{A}_3)} = \frac{2/6}{3/6} = \frac{2}{3} \neq P(\mathcal{A}_o),$$

e, portanto, **não são** eventos independentes.

Considerando a retirada de uma carta de um baralho, a probabilidade de tirarmos uma carta de ouros dado que sabemos que tiramos uma carta de copas, é, obviamente

Exemplo 12 
$$P(\mathcal{A}_{ouros}|\mathcal{A}_{copas}) = 0 \neq P(\mathcal{A}_{ouros}),$$
 ou seja, eventos disjuntos não são independentes.

Considerando a jogada de 2 dados, a probabilidade de tirarmos um certo valor  $\xi_i$  no segundo dado independe do resultado obtido no primeiro dado. Neste caso temos realizações ou tentativas (*trials*) independentes de um experimento aleatório.

Sabemos ainda que  $P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A}|\mathcal{B})P(\mathcal{B})$ , e, conseqüentemente, **se  $\mathcal{A}$  e  $\mathcal{B}$  forem independentes**, portanto

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B}) \tag{11}$$

e, também, já que  $P(\mathcal{B}|\mathcal{A}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{A})}$ , da equação acima,  $P(\mathcal{B}|\mathcal{A}) = P(\mathcal{B})$ ,

#### e. Tentativas de Bernouilli

Considere um experimento aleatório em que podem ocorrer apenas dois resultados possíveis, vamos chamar de  $A$  (sucesso) e  $\bar{A}$  (insucesso), ou seja,  $\mathcal{S} = \{A, \bar{A}\}$ , de modo que

$$P(A) = p$$

$$P(\bar{A}) = 1 - p$$

Este experimento é chamado de **tentativa (*trial*) de Bernouilli**.

Uma moeda viciada, tal que  $P(\text{cara}) = \frac{1}{4}$ , é jogada 5 vezes, e queremos a probabilidade de termos a sequência {cara, cara, coroa, cara, coroa}.

Este problema pode ser caracterizado por tentativas de Bernouilli, em que podemos atribuir o evento cara como sendo sucesso, e a probabilidade de ele ocorrer é  $P(A) = p = \frac{1}{4}$ . Queremos então achar  $P(A, A, \bar{A}, A, \bar{A})$ . Como são tentativas independentes, temos que

$$P(A, A, \bar{A}, A, \bar{A}) = P(A)P(A)P(\bar{A})P(A)P(\bar{A}) = p^3(1 - p)^2 = \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 = \frac{9}{1024}$$

De modo geral, dada a realização de  $n$  tentativas independentes, a probabilidade de termos uma certa sequência com  $k$  sucessos em uma ordem específica é  $p^k(1 - p)^{n-k}$ . De análise combinatória sabemos ainda que temos  $\binom{n}{k} = \frac{n!}{k!(n-k)!}$  combinações de  $n$  elementos  $k$  a  $k$ , ou seja  $\binom{n}{k}$  sequências diferentes de  $n$  elementos com  $k$  sucessos. Portanto, a probabilidade de termos uma sequência com  $k$  sucessos em  $n$  tentativas, em qualquer ordem, é dada por

$$P(k \text{ sucessos em } n) = \binom{n}{k} p^k (1-p)^{n-k} \quad (12)$$

No exemplo acima, qual a probabilidade de termos 3 caras em 5 jogadas da moeda viciada?

Exemplo 14

$$P(3 \text{ sucessos em } 5) = \binom{5}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 = \frac{90}{1024}$$

### Canal Binário Simétrico (BSC – Binary Symmetric Channel)

Exemplo 15 Este modelo de canal, que será utilizado diversas vezes ao longo do curso, representa o envio de um bit  $x$ , com espaço amostral  $\mathcal{S}_x = \{0,1\}$  em um canal com erro, de modo que na saída teremos um bit  $y$ , com

$$P(y = 1|x = 0) = P(y = 0|x = 1) = p,$$

representando a probabilidade de erro, como vemos na Figura 3.

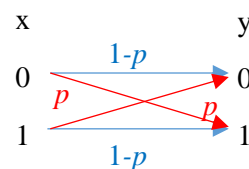


Figura 3. Canal Binário Simétrico

Neste canal, supondo que enviamos um byte (8 bits) com todos os bits iguais a “0”, qual a probabilidade de termos recebido 2 bits iguais a “1”, sabendo que  $p = 0,1$ ?

Supondo que o evento sucesso seja recebermos um bit igual a “1”, temos que

Exemplo 16

$$P(2 \text{ bits iguais a } 1 \text{ recebidos}) = \binom{8}{2} p^2 (1-p)^6 = 28(0,1)^2 (0,9)^6 = 0,1488$$

### Repetidores Digitais Regenerativos

Para se enviar mensagens a longas distâncias, às vezes são utilizados repetidores regenerativos. A mensagem, um bit  $x$ , é modulada e enviada por um enlace a um repetidor, onde o sinal é decodificado e mapeado em um bit  $y_1$ , com probabilidade erro  $p$ . Este bit por sua vez é enviado por um outro enlace, no qual é recebido um bit  $y_2$ , e assim por diante, até chegar no destino, após  $n$  enlaces, em  $y_n$ . Qual a probabilidade de que o bit recebido esteja errado?

Um erro ocorrerá se tivermos nenhum erro em nenhum dos enlaces ou um número par de erros, que assim se cancelarão. Como são todos eventos disjuntos, temos que



$$\begin{aligned}
P(\bar{\mathcal{E}}) &= 1 - P(\mathcal{E}) = P(0 \text{ erros}) + \sum_{k=2,4,\dots,2\lfloor \frac{n}{2} \rfloor} P(k \text{ erros}) \\
&= (1-p)^n + \sum_{k=2,4,\dots,2\lfloor \frac{n}{2} \rfloor} \binom{n}{k} p^k (1-p)^{n-k}
\end{aligned}$$

Considerando que  $p \ll 1$ , podemos ignorar todos as potências de  $p$  maiores que 2 no somatório, e assim,

$$\begin{aligned}
P(\bar{\mathcal{E}}) &\approx (1-p)^n + \binom{n}{2} p^2 \\
&= 1 - np + \binom{n}{2} p^2 - \binom{n}{3} p^3 + \dots + (-1)^n p^n + \frac{n(n-1)}{2} p^2
\end{aligned}$$

Considerando ainda que  $np \ll 1$ , podemos desconsiderar o último termo, e, assim

$$P(\bar{\mathcal{E}}) \approx 1 - np \Rightarrow P(\mathcal{E}) \approx np$$

#### f. Teorema da Probabilidade Total

Suponha  $n$  eventos disjuntos  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ , ou seja,  $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ , se  $i \neq j$ , tais que  $\bigcup_{i=1}^n \mathcal{A}_i = \mathcal{S}$ .

Neste caso, como podemos ver, pela Figura 4,

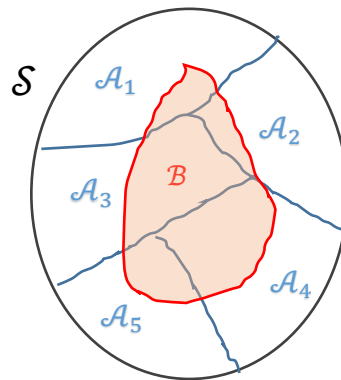


Figura 4. Teorema da Probabilidade Total

Exemplo 17

$$P(\mathcal{B}) = \sum_{i=1}^n P(\mathcal{B} \cap \mathcal{A}_i) = \sum_{i=1}^n P(\mathcal{B}|\mathcal{A}_i)P(\mathcal{A}_i), \quad (13)$$

Considerando um canal binário simétrico, tal que  $P(x=1) = \alpha$ , qual a probabilidade de recebermos um bit igual a "1"?

Podemos visualizar este problema na Figura 5, em que temos um experimento com 4 resultados possíveis, particionado em 2 subconjuntos,  $\mathcal{A}_0$  com  $x=0$ , e  $\mathcal{A}_1$  com  $x=1$ , e queremos a probabilidade de um evento  $\mathcal{B}$ , com  $y=1$ .

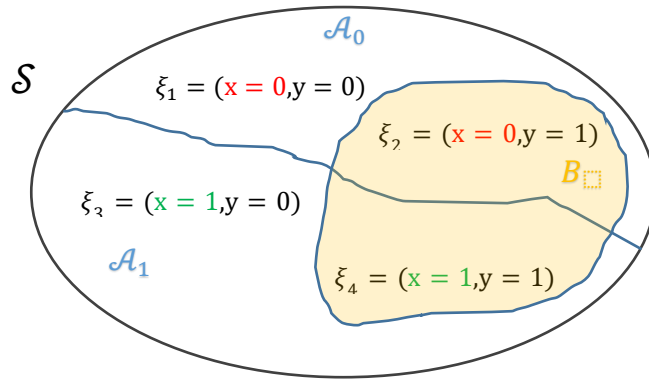


Figura 5. Partição do BSC

Na forma de equações,

$$P(y = 1) = P(y = 1|x = 0)P(x = 0) + P(y = 1|x = 1)P(x = 1) \\ = p(1 - \alpha) + (1 - p)\alpha = p + \alpha - 2\alpha p$$

e, conseqüentemente,  $P(y = 0) = 1 - p - \alpha + 2\alpha p$ .

Se tivermos bits equiprováveis, ou seja  $\alpha = \frac{1}{2}$ , então  $P(y = 1) = \frac{1}{2}$ , e não depende de  $p$ .

Se chamarmos o evento  $\mathcal{E}$  de erro de bit,

$$P(\mathcal{E}) = P(\mathcal{E}|x = 0)P(x = 0) + P(\mathcal{E}|x = 1)P(x = 1) \\ = (1 - \alpha)P(y = 1|x = 0) + \alpha P(y = 0|x = 1) = p,$$

ou seja, a probabilidade de erro independe do bit enviado.

Em um sistema de comunicações, observamos o sinal  $y$  no receptor, e podemos querer saber como é a estatística do sinal no transmissor. Por exemplo, se recebermos um bit igual a "0", qual a probabilidade de termos enviado um bit igual a "1"? Usamos para isso a regra de Bayes (9), e, assim

$$P(x = 1|y = 0) = \frac{P(y = 0|x = 1)P(x = 1)}{P(y = 0)} = \frac{p\alpha}{1 - p - \alpha + 2\alpha p}$$

Exemplo 18

### Código de Repetição

Suponha agora que tenhamos um bit de mensagem  $m$ . Em vez de o enviarmos diretamente pelo canal BSC, iremos repeti-lo  $n$  vezes, e enviar esta sequência repetida pelo canal. Por exemplo, se  $m = 1$ , e tivermos  $n = 3$ , iremos enviar a sequência  $\mathbf{x} = 111$ , e cada bit desta sequência terá uma probabilidade de erro igual a  $p$ , de modo que receberemos uma sequência  $\mathbf{y}$ . Por exemplo, se tivermos um erro no segundo bit, receberíamos  $\mathbf{y} = 101$ .

No receptor iremos detectar o sinal decidindo pela regra da maioria, ou seja, contaremos quantos "0"s e "1"s foram recebidos, e decidiremos pelo que ocorrer mais vezes. Para simplificar podemos considerar  $n$  par, de modo que o número de "0"s e "1"s nunca seja igual.

Deste modo, a probabilidade de erro de detecção  $\mathcal{E}_{rep}$  é a probabilidade de que ocorram mais do que  $\frac{n}{2}$  erros de bit, ou seja, sabendo que a probabilidade de erro de um bit é  $P(\mathcal{E}) = p$ , e que a probabilidade de termos exatamente  $k$  erros em  $n$  é dada por (13), temos que

$$P(\mathcal{E}_{rep}) = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k} p^k (1-p)^{n-k},$$

em que  $\lceil x \rceil$  é a função *ceiling*, ou seja, o menor inteiro maior ou igual a  $x$ .

Por exemplo, considerando  $n = 3$ , temos que

$$P(\mathcal{E}_{rep}|n = 3) = \binom{3}{2} p^2 (1-p) + \binom{3}{3} p^3 = 3p^2 - 2p^3$$

Se  $p \ll 1$ , podemos aproximar por  $P(\mathcal{E}_{rep}|n = 3) \approx 3p^2$ .

Considerando  $p = 10^{-2}$ , temos que  $P(\mathcal{E}_{rep}|n = 3) = 2,98 \times 10^{-4} \approx 3 \times 10^{-4}$ .

Considerando agora  $n = 5$ , temos que

$$P(\mathcal{E}_{rep}|n = 5) = \binom{5}{3} p^3 (1-p)^2 + \binom{5}{4} p^4 (1-p) + \binom{5}{5} p^5 = 10p^3 - 15p^4 + 6p^5 \approx 10p^3$$

e, com  $p = 10^{-2}$ , teríamos  $P(\mathcal{E}_{rep}|n = 5) = 2,98 \times 10^{-4} = 9,8506 \times 10^{-6} \approx 10^{-5}$ .

Como podemos ver, podemos atingir probabilidades de erro cada vez menores apenas aumentando o tamanho do bloco de repetição, em um mesmo canal com erro.

## 2. Variáveis aleatórias

Uma variável aleatória (v.a.) é um mapeamento, não necessariamente biunívoco  $\mathcal{S} \rightarrow \mathbb{R}$ , de cada resultado  $\xi_i$  de um experimento aleatório para um número real  $x_i$ , pertencente a um conjunto  $E \subset \mathbb{R}$ , como vimos na Figura 6.

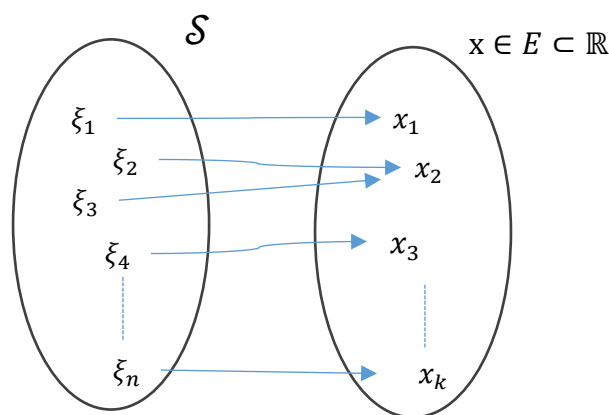


Figura 6. Conceito de variável aleatória

Usaremos a notação  $x$  para representar a variável aleatória, e  $P_x(x) = P(x = x)$  a probabilidade de a variável aleatória  $x$  assumir o valor  $x$ .

Para valores discretos, o valor  $P_x(x)$  é chamado de **função massa de probabilidade**, e temos que, devido ao mapeamento da Figura 6, em que os resultados são disjuntos,

$$P_x(x_i) = \sum_{k|\xi_k \rightarrow x_i} P(\xi_k) \quad (14)$$

e,

$$\sum_{x_i \in E} P(x_i) = 1. \quad (15)$$

Neste exemplo veremos que para um mesmo experimento aleatório podemos ter diferentes mapeamentos em variáveis aleatórias.

Suponha um experimento em que são jogados dois dados. Temos neste caso 36 resultados possíveis,

$$\mathcal{S} = \{(\xi_1, \xi_1), (\xi_1, \xi_2), (\xi_1, \xi_3), \dots, (\xi_2, \xi_1), (\xi_2, \xi_2), \dots, (\xi_6, \xi_6)\},$$

com

$$P((\xi_i, \xi_j)) = \frac{1}{36}$$

Estes eventos podem ser mapeados por exemplo em uma variável aleatória  $y$ , correspondente à soma dos dois dados, tal que, de (14),

$$P_y(2) = P((\xi_1, \xi_1)) = \frac{1}{36}$$

$$P_y(3) = P((\xi_1, \xi_2)) + P((\xi_2, \xi_1)) = \frac{2}{36}$$

$$P_y(4) = P((\xi_1, \xi_3)) + P((\xi_2, \xi_2)) + P((\xi_3, \xi_1)) = \frac{3}{36}$$

$$P_y(5) = P((\xi_1, \xi_4)) + P((\xi_2, \xi_3)) + P((\xi_3, \xi_2)) + P((\xi_4, \xi_1)) = \frac{4}{36}$$

$$P_y(6) = P((\xi_1, \xi_5)) + P((\xi_2, \xi_4)) + P((\xi_3, \xi_3)) + P((\xi_4, \xi_2)) + P((\xi_5, \xi_1)) = \frac{5}{36}$$

$$P_y(7) = P((\xi_1, \xi_6)) + \dots + P((\xi_5, \xi_2)) = \frac{6}{36}$$

$$P_y(8) = \frac{5}{36}; P_y(9) = \frac{4}{36}; P_y(10) = \frac{3}{36}; P_y(11) = \frac{2}{36}; P_y(12) = \frac{1}{36}$$

Alternativamente, o mesmo experimento pode ser mapeado em duas variáveis aleatórias,  $x_1$  e  $x_2$ , correspondentes aos valores do primeiro e do segundo dado, respectivamente, ou seja

$$P_{x_1}(x) = \frac{1}{6}, \quad x = 1,2,3,4,5 \text{ ou } 6$$

$$P_{x_2}(x) = \frac{1}{6}, \quad x = 1,2,3,4,5 \text{ ou } 6$$

Podemos criar ainda uma outra variável aleatória  $z = 6(x_1 - 1) + (x_2 - 1)$ . Neste caso teremos o seguinte mapeamento

$$(\xi_1, \xi_1) \rightarrow z = 0$$

$$(\xi_1, \xi_2) \rightarrow z = 1$$

⋮

$$(\xi_2, \xi_1) \rightarrow z = 6$$

⋮

$$(\xi_6, \xi_6) \rightarrow z = 35$$

e, para cada par de valores, temos um valor de  $z$  diferente, e,

$$P_z(z) = \frac{1}{36}, \quad 0 \leq z \leq 35$$

#### a. Variáveis aleatórias conjuntas

Dadas duas ou mais variáveis aleatórias, podemos definir uma função massa de probabilidade conjunta. Por exemplo, dadas as variáveis aleatórias  $x$  e  $y$ ,

$$P_{x,y}(x, y) = P(x = x \cap y = y)$$

Exemplo 20

No exemplo anterior, por exemplo,  $P_{x_1, x_2}(1,4) = P((\xi_1, \xi_4)) = \frac{1}{36}$ , e  $P_{x_1, y}(1,5) = P((\xi_1, \xi_4)) = \frac{1}{36}$ . Já  $P_{x_1, y}(1,11) = 0$ , pois a soma jamais será igual a "11" se a face do primeiro dado for igual a "1".

Sabemos ainda que

$$\sum_{x_i \in E_x, y_i \in E_y} P(x_i, y_i) = 1. \quad (16)$$

#### b. Função massa de probabilidade condicional e regra de Bayes

Similarmente ao conceito de probabilidade condicional que vimos na Seção 1.c, podemos definir uma função massa de probabilidade condicional

$$P_{x|y}(x|y) = P(x = x|y = y) = \frac{P_{x,y}(x, y)}{P_y(y)}, \quad (17)$$

e a regra de Bayes

$$P_{y|x}(y|x) = \frac{P_{x|y}(x|y)P_y(y)}{P_x(x)}, \quad (18)$$

Considerando que jogamos dois dados, dado que a soma dos dados é igual a 4, ache a probabilidade de o primeiro dado ser igual a 2.

Seguindo a notação do exemplo anterior, queremos

$$P_{x_1|y}(2|4) = \frac{P_{y|x_1}(4|2)P_{x_1}(2)}{P_y(4)},$$

sabendo que

$$P_{y|x_1}(4|2) = \frac{P_{x_1,y}(2,4)}{P_{x_1}(2)} = \frac{1/36}{1/6} = \frac{1}{6}.$$

Portanto,

$$P_{x_1|y}(2|4) = \frac{1/6 \cdot 1/6}{3/36} = \frac{1}{3}.$$

O mesmo resultado poderia ser obtido verificando-se que das três combinações que tem uma soma igual a 4, uma delas tem o primeiro dado igual a 2.

### c. Variáveis aleatórias independentes

Similarmente à definição de eventos independentes, duas variáveis aleatórias  $x$  e  $y$  são ditas independentes se

$$P_{y|x}(y|x) = P_y(y), \tag{19}$$

e, consequentemente,

$$P_{x|y}(x|y) = P_x(x) \tag{20}$$

Exemplo 21

$$P_{x,y}(x, y) = P_x(x)P_y(y) \tag{21}$$

No exemplo anterior, em que são jogados dois dados, as variáveis aleatórias  $x_1$  e  $x_2$  são independentes.

### d. Probabilidade marginal

Se conhecemos a probabilidade conjunta de duas variáveis aleatórias  $x$  e  $y$ , podemos obter a função massa de probabilidade de uma delas,

$$\sum_{x_i \in E} P_{x,y}(x_i, y) = \sum_{x_i \in E} P_{x|y}(x_i|y)P_y(y) = P_y(y) \sum_{x_i \in E} P_{x|y}(x_i|y) \tag{22}$$

Como o somatório na esquerda está considerando a soma das probabilidades de todos os valores possíveis da variável aleatória  $x$ , ele é igual a um, e, portanto

$$P_y(y) = \sum_{x_i \in E} P_{x,y}(x_i, y) \tag{23}$$

Esta é conhecida como a probabilidade marginal de  $y$ .

### Canal Binário Assimétrico

Este modelo de canal é similar ao BSC, porém, com probabilidades de erro diferentes, dependendo se foi enviado o bit 0 ou 1. Ou seja,

$$P(y = 1|x = 0) = p_0$$

Exemplo 22

$$P(y = 0|x = 1) = p_1,$$

como vemos na Figura 6

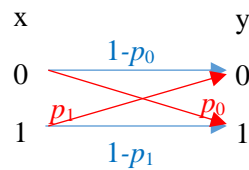


Figura 7. Canal Binário Assimétrico

Vamos criar agora uma variável aleatória  $\epsilon$ , tal que  $\epsilon = 1$ , se ocorrer um erro de bit e  $\epsilon = 0$ , se o bit for transmitido corretamente. Ou seja,

$$P_{\epsilon|x}(1|0) = p_0$$

$$P_{\epsilon|x}(1|1) = p_1$$

Vamos considerar ainda que  $P_x(0) = \alpha$ .

Queremos saber qual a probabilidade de erro, ou seja  $P_\epsilon(1)$ . Temos que

$$P_\epsilon(1) = P_{x,\epsilon}(0,1) + P_{x,\epsilon}(1,1) = P_x(0)P_{\epsilon|x}(1|0) + P_x(1)P_{\epsilon|x}(1|1) = \alpha p_0 + (1 - \alpha)p_1$$

### e. Algumas variáveis aleatórias discretas comuns

#### Variável aleatória de Bernouilli

Uma variável aleatória de Bernouilli é aquela que

$$P_x(x) = \begin{cases} \alpha & ; x = 1 \\ 1 - \alpha & ; x = 0 \\ 0 & ; x \neq 0 \text{ ou } 1 \end{cases} \quad (24)$$

Por exemplo, o evento de erro em um canal binário ou o resultado de se jogar uma moeda são variáveis aleatórias de Bernouilli.

#### Variável aleatória Binomial

O número de sucessos  $k$  em  $n$  realizações de um experimento de Bernouilli com probabilidade de sucesso  $p$  é uma variável aleatória Binomial  $(n, p)$ , de modo que

$$P_x(k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, 2, \dots, n \quad (25)$$

Na Figura 8 podemos ver exemplos da função massa de probabilidade de variáveis aleatórias binomiais.

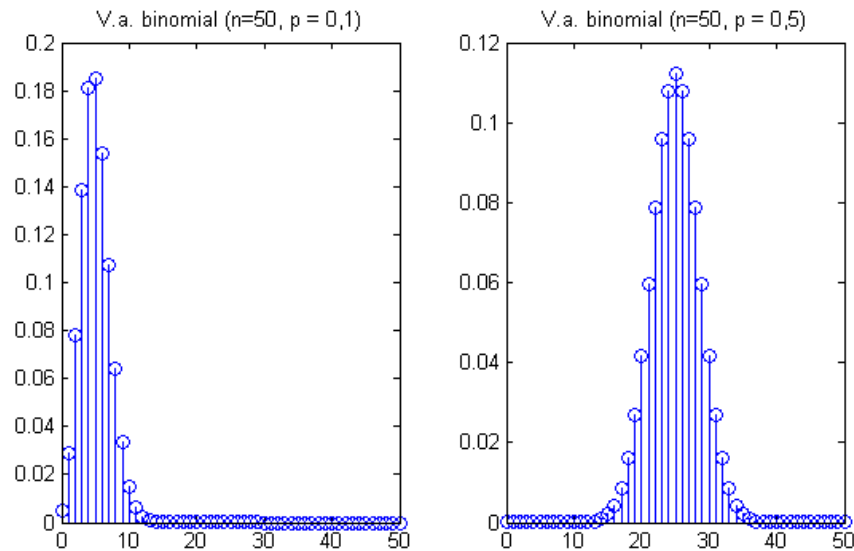


Figura 8. Variável aleatória binomial

#### Variável aleatória de Poisson

Uma variável aleatória bastante comum é a variável aleatória de Poisson, que modela o número de eventos que ocorrem dentro de um dado intervalo, supondo que estes eventos ocorrem independentemente uns dos outros. Estes eventos podem ser, por exemplo, chamadas telefônicas. A v.a. de Poisson com parâmetro  $\lambda$  ( $\lambda \in \mathbb{R}, \lambda > 0$ ) é definida por

$$P_x(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x \in \mathbb{N} \quad (26)$$

Na Figura 9 vemos exemplos de variáveis aleatórias de Poisson com diferentes parâmetros  $\lambda$ .

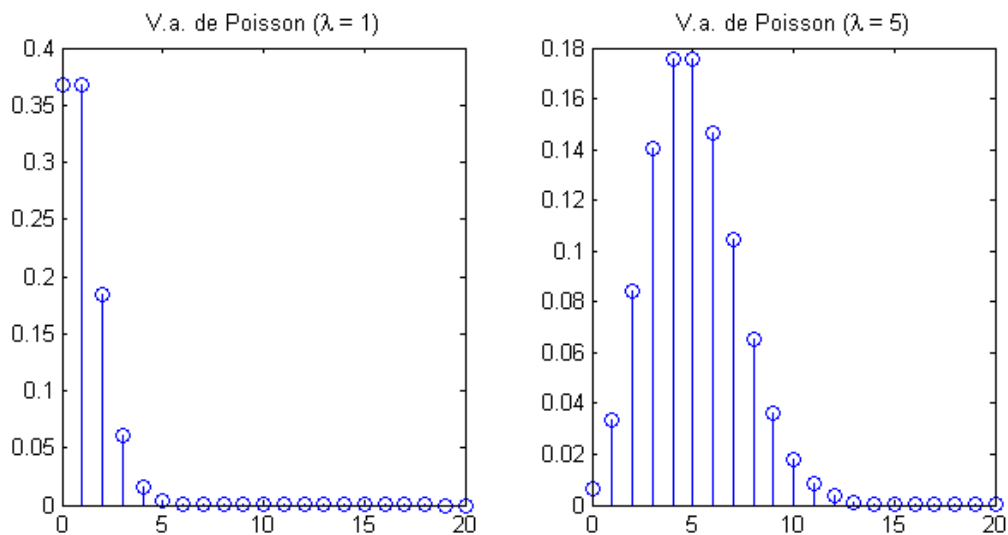


Figura 9. Variável aleatória de Poisson

#### f. Função de Distribuição Cumulativa (CDF – Cumulative Distribution Function)

A função de distribuição cumulativa (CDF) é por definição

$$F_x(x) = \Pr(x \leq x) \quad (27)$$



e, a partir desta definição, temos as seguintes propriedades

$$\begin{aligned} F_x(x) &\geq 0 \\ F_x(\infty) &= 1 \\ F_x(-\infty) &= 0 \\ F_x(x_1) &\leq F_x(x_2), \text{ se } x_1 < x_2 \end{aligned} \tag{28}$$

ou seja, a CDF é uma função não-negativa, não-decrescente.

Considerando a variável aleatória  $y$ , a soma de dois dados, temos por exemplo

$$F_y(0,5) = \Pr(y \leq 0,5) = 0$$

Exemplo 23

$$F_y(3,5) = \Pr(y \leq 3,5) = P_y(2) + P_y(3) = \frac{3}{36}$$

$$F_y(5) = \Pr(y \leq 5) = P_y(2) + P_y(3) + P_y(4) + P_y(5) = \frac{10}{36}$$

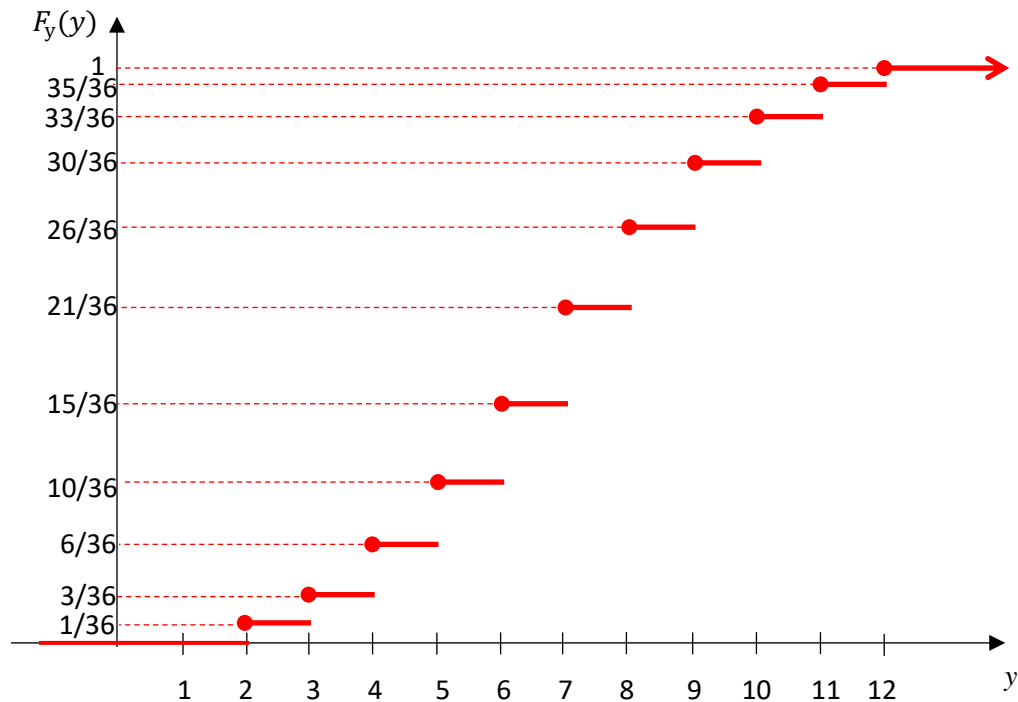


Figura 10. Função de distribuição cumulativa de variável aleatória discreta

Exemplo 24

### g. Variáveis aleatórias contínuas e função densidade de probabilidade

Vimos até agora variáveis aleatórias que podem assumir apenas valores discretos. Existem entretanto inúmeros experimentos aleatórios que podem assumir valores em um intervalo contínuo.

São exemplos de variáveis aleatórias contínuas:

- A temperatura medida em um dado instante;

- A tensão medida em uma resistência;
- A altura e o peso de uma pessoa

Como uma variável aleatória contínua  $x$  pode assumir infinitos valores, a probabilidade de ela assumir um valor específico é tipicamente igual a zero. Neste caso não faz sentido definirmos uma probabilidade  $P_x(x)$ , mas podemos definir uma função de distribuição cumulativa (CDF)  $F_x(x)$ .

A partir da CDF podemos também definir a **função densidade de probabilidade** (PDF – *probability density function*)

$$p_x(x) = \frac{dF_x(x)}{dx} \quad (29)$$

e, conseqüentemente,

$$F_x(x) = \int_{-\infty}^x p_x(\alpha) d\alpha. \quad (30)$$

Por meio da PDF podemos achar a probabilidade de a variável aleatória estar em um certo intervalo  $(x_1, x_2]$ , ou seja,

$$\begin{aligned} \Pr(x_1 < x \leq x_2) &= \Pr(x \leq x_2) - \Pr(x \leq x_1) \\ &= F_x(x_2) - F_x(x_1) \\ &= \int_{-\infty}^{x_2} p_x(x) dx - \int_{-\infty}^{x_1} p_x(x) dx \\ \Pr(x_1 < x \leq x_2) &= \int_{x_1}^{x_2} p_x(x) dx \end{aligned} \quad (31)$$

Como a probabilidade de a variável aleatória assumir um valor específico é tipicamente igual a zero, normalmente não é importante se temos “menor” ou “menor ou igual” dentro de (28).

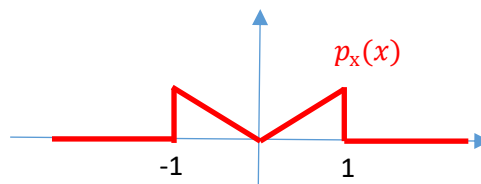
A PDF tem algumas propriedades importantes. Já que a CDF é uma função não decrescente e a PDF é sua derivada, temos que

$$p_x(x) \geq 0. \quad (32)$$

Sabemos ainda que

Exemplo 25

$$F_x(x) = \int_{-\infty}^{\infty} p_x(x) dx = 1 \quad (33)$$



Um experimento aleatório é modelado por uma variável aleatória  $x$ , mostrada na figura abaixo. Qual a probabilidade de que  $x$  seja menor que  $1/2$ ?

Vemos que  $p_x(x) = a|x|$ ,  $|x| \leq 1$ . Sabemos ainda que

$$\int_{-\infty}^{\infty} p_x(x) dx = 2 \int_0^1 ax dx = a = 1$$

e, assim,

$$\Pr\left(x < \frac{1}{2}\right) = \int_{-\infty}^{1/2} p_x(x) dx = \int_{-1}^{1/2} |x| dx = \frac{5}{8}$$

Veremos agora algumas variáveis aleatórias comumente utilizadas.

### Variável Aleatória Uniforme

Uma v.a. uniforme é tal que a PDF é uma constante  $c$  em um dado intervalo, ou seja

$$p_x(x) = \begin{cases} c & , a \leq x \leq b \\ 0 & , \text{caso contrário} \end{cases} \quad (34)$$

Para encontrarmos o valor da constante  $c$

$$\int_{-\infty}^{\infty} p_x(x) dx = \int_a^b c dx = c(b - a) = 1 \Rightarrow c = \frac{1}{b - a} \quad (35)$$

e a CDF é dada por

$$F_x(x) = \begin{cases} 0 & , x \leq a \\ \frac{x - a}{b - a} & , a < x \leq b \\ 1 & , b < x \end{cases} \quad (36)$$

Graficamente, vemos na Figura 11, a PDF e a CDF de uma v.a. uniforme.

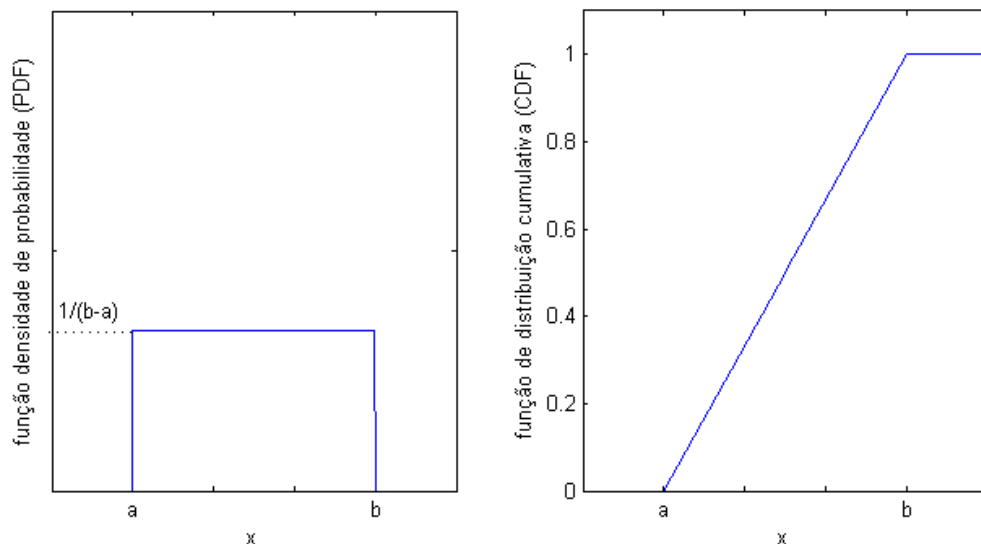


Figura 11. Variável aleatória uniforme

### Variável Aleatória Exponencial

Uma v.a. exponencial com parâmetro  $\lambda$  é definida pela PDF

$$p_x(x) = Ae^{-\lambda x} u(x) \quad (37)$$

e, conseqüentemente, pela CDF

$$F_x(x) = \int_{-\infty}^x p_x(\alpha) d\alpha = \begin{cases} \int_0^x A e^{-\lambda\alpha} d\alpha = \frac{A}{\lambda} (1 - e^{-\lambda x}) & , \text{ se } x \geq 0 \\ 0 & , \text{ se } x < 0 \end{cases} \quad (38)$$

Para encontrarmos o valor da constante  $A$ , sabemos que

$$F_x(\infty) = \frac{A}{\lambda} = 1 \Rightarrow A = \lambda \quad (39)$$

e, portanto,

$$\begin{aligned} p_x(x) &= \lambda e^{-\lambda x} u(x) \\ F_x(x) &= (1 - e^{-\lambda x}) u(x) \end{aligned} \quad (40)$$

Na Figura 12 vemos a PDF e a CDF de uma v.a. exponencial para diferentes valores de  $\lambda$ .

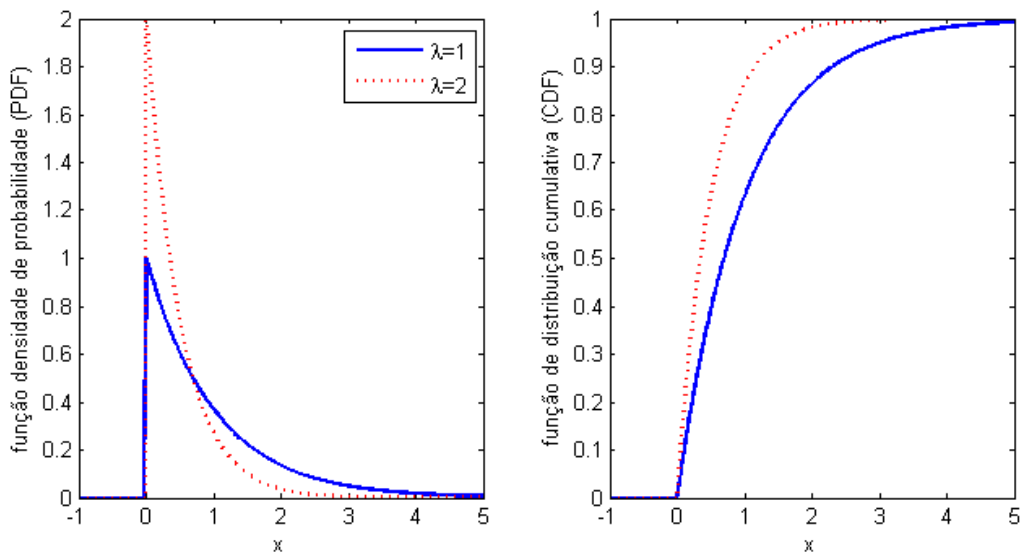


Figura 12. Variável aleatória exponencial

A v.a. exponencial tem ainda a seguinte propriedade

$$\begin{aligned} \Pr(x > \tau + \Delta t | x > \tau) &= \frac{\Pr(x > \tau + \Delta t \cap x > \tau)}{\Pr(x > \tau)} = \frac{\Pr(x > \tau + \Delta t)}{\Pr(x > \tau)} \\ &= \frac{F_x(\tau + \Delta t)}{F_x(\tau)} = \frac{e^{-\lambda(\tau + \Delta t)}}{e^{-\lambda\tau}} = e^{-\lambda\Delta t} = \Pr(x > \Delta t) \end{aligned} \quad (41)$$

#### Exemplo 26

A v.a. exponencial é comumente usada para modelar o tempo de ocorrência de algum evento, e (41) indica que esta é uma variável sem memória, ou seja, neste caso o tempo até o evento acontecer é sempre uma v.a. exponencial, independente de quando começamos a observar.

Sabendo que a duração de uma lâmpada, em anos, é uma variável aleatória exponencial com  $\lambda$ , qual a probabilidade de ela durar ao menos 1 ano.

Neste caso temos que  $t \rightarrow \exp(\lambda = 1)$ , e, portanto,

$$\Pr(t > 1) = 1 - F_t(1) = e^{-\lambda t} = 0,3678$$

Sabendo que a lâmpada já está funcionando há um ano, qual a probabilidade de ela durar mais que três anos no total.

Neste caso, queremos

$$\Pr(t > 3 | t > 1) = \Pr(t > 2) = e^{-2\lambda} = 0,13533$$

### Variável Aleatória Gaussiana ou Normal

Uma variável aleatória Gaussiana (ou normal) padrão é definida pela sua PDF

$$p_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (42)$$

Não existe fórmula fechada para a sua CDF, que é mostrada, junto com a PDF, na Figura 13.

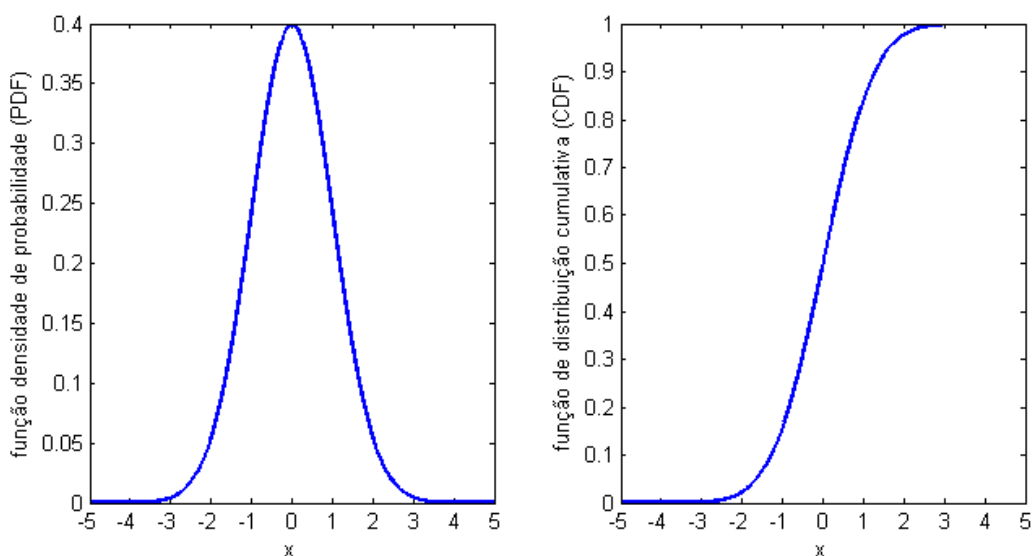


Figura 13. Variável aleatória normal padrão

Comumente estamos interessados em saber a probabilidade de a variável aleatória assumir um valor maior que um dado  $x$ . Para isso não há fórmula fechada, mas utilizamos a função  $Q$ :

$$Q(x) = \Pr(x > x) = 1 - F_x(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{\alpha^2}{2}} d\alpha \quad (43)$$

Uma tabela da função  $Q$  pode ser encontrada por exemplo no Apêndice B. Existem ainda algumas aproximações para esta função, como:

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} \left(1 - \frac{0,7}{x^2}\right) e^{-\frac{x^2}{2}}, \text{ se } x > 2$$

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \text{ se } x \gg 1 \quad (44)$$

Também são comumente encontradas em tabelas e calculadoras as funções de erro e de erro complementar

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\alpha^2} d\alpha$$

$$\text{erfc}(x) = 1 - \text{erf}(x) \quad (45)$$

que podem ser convertidas na função Q com

$$Q(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \quad (46)$$

### Função Densidade de Probabilidade de uma Variável Aleatória Discreta ou Mista

Para variáveis aleatórias discretas, sabemos que

$$F_x(x) = \sum_i P_x(x_i) u(x - x_i) \quad (47)$$

em que  $u(x)$  é a função degrau, como vimos por exemplo na Figura 10. Desta forma, pela definição da PDF, temos que

$$p_x(x) = \sum_i P_x(x_i) \delta(x - x_i) \quad (48)$$

Considerando a variável aleatória  $y$ , a soma de dois dados, temos a seguinte

Exemplo 27

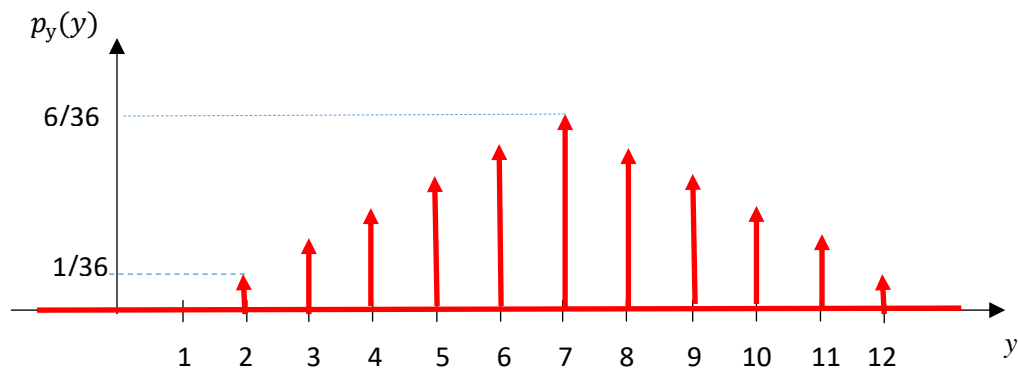


Figura 14. Função de densidade de probabilidade de variável aleatória discreta

Exemplo 28

Podemos ainda ter variáveis aleatórias mistas, ou seja, parte discretas, parte contínuas.

#### Variável Aleatória mista

Suponha que um termômetro consiga medir temperaturas entre  $-10^\circ$  e  $40^\circ$ , mas estamos numa localidade onde a temperatura real  $x$  é uma variável aleatória uniforme entre  $10^\circ$  e  $45^\circ$ . A temperatura medida  $y$  terá uma probabilidade não nula de assumir o valor de  $40^\circ$ , ou seja,  $P_y(40) = \Pr(x > 40) = \frac{1}{7}$ , e a PDF será

$$p_y(y) = \frac{1}{35} \operatorname{rect}\left(\frac{y-25}{30}\right) + \frac{1}{7} \delta(y-40)$$

#### Histograma

Quando não conhecemos a priori a PDF de um experimento aleatório, podemos investigar sua distribuição de probabilidade por meio do histograma. Dividimos todo o intervalo possível de amostras em intervalos de tamanho  $\Delta x$ , chamados de *bins*. Por exemplo, supondo que as amostras começam em  $x_0$ , o intervalo  $k$  será  $I_k = [x_k, x_{k+1})$ , com  $x_k = x_0 + k\Delta x$ . Se fizermos

$N$  realizações independentes do experimento aleatório o histograma corresponderá à contagem do número  $N_k$  de resultados que ocorreram no intervalo  $I_k$ . Pela definição de probabilidade em (2) temos que

$$\Pr(x \in I_k) = \lim_{N \rightarrow \infty} \frac{N_k}{N} = \int_{x_k}^{x_{k+1}} p_x(x) dx \approx \Delta x p_x(x_k) \quad (49)$$

Ou seja, o histograma tende a representar a PDF de uma variável aleatória.

Na Figura 15 podemos ver os histogramas de duas variáveis aleatórias, uma uniforme entre 0 e 1 e outra exponencial, com  $\lambda=1$ , para  $N = 10.000$  e  $N = 100.000$  amostras. Podemos ver que com o aumento do número de amostras o histograma tenda à forma de sua PDF.

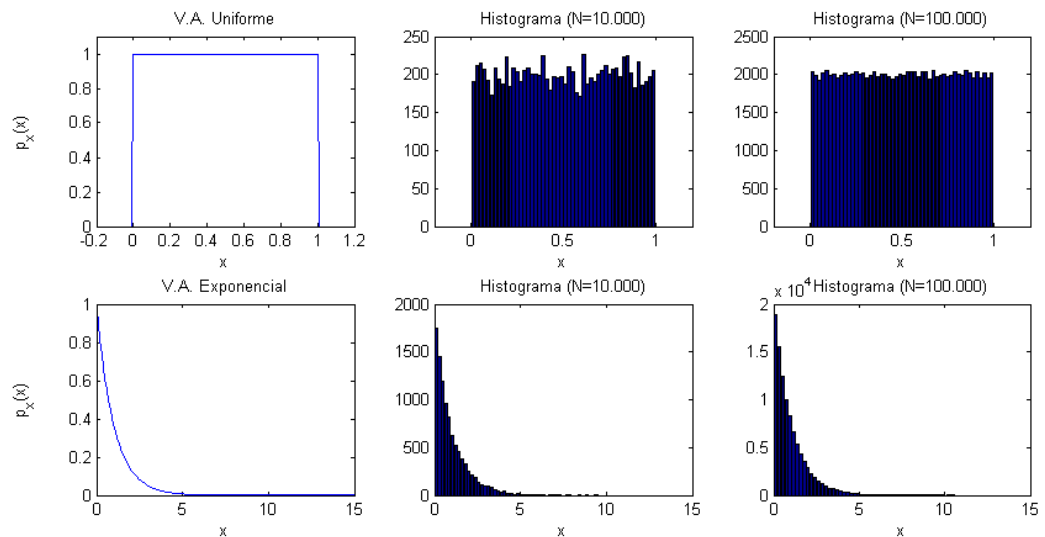


Figura 15. Histograma

#### h. Valor Esperado, Média e Mediana

O valor esperado (ou esperança ou expectância) de uma variável aleatória é a sua média. Ou seja, para uma variável aleatória discreta  $x$ , se forem realizados  $N$  experimentos aleatórios independentes, e o valor  $x_i$  ocorrer  $N_i$  vezes temos que sua **média**, ou seu **valor esperado** é

Exemplo 30

$$\begin{aligned} E\{x\} = \bar{x} &= \lim_{N \rightarrow \infty} \frac{N_1 x_1 + N_2 x_2 + \dots + N_n x_n}{N} \\ &= P_x(x_1) x_1 + P_x(x_2) x_2 + \dots + P_x(x_n) x_n = \sum_i x_i P_x(x_i) \end{aligned} \quad (50)$$

Jogamos um dado, tal que  $P_x(x_i) = \frac{1}{6}$ ,  $x_i = 1; 2; 3; 4; 5; 6$ . A sua média estatística será

$$\bar{x} = \frac{1}{6}(1) + \frac{1}{6}(2) + \frac{1}{6}(3) + \frac{1}{6}(4) + \frac{1}{6}(5) + \frac{1}{6}(6) = 3,5$$

Similarmente, para uma variável aleatória contínua, a média é dada por

$$\bar{x} = \int_{-\infty}^{\infty} xp_x(x)dx \quad (51)$$

Não devemos confundir a **média** com a **mediana**, que é definida como

$$\text{Med}(x) = x, \text{ t.q. } \Pr(x \leq \text{Med}(x)) \geq \frac{1}{2} \text{ e } \Pr(x \geq \text{Med}(x)) \geq \frac{1}{2} \quad (52)$$

Para variáveis aleatórias contínuas, isto equivale a

$$F_x(\text{Med}(x)) = \frac{1}{2} \quad (53)$$

Veremos agora o valor esperado de algumas variáveis aleatórias comuns

#### Variável aleatória de Bernouilli

Dada uma v.a. de Bernouilli com  $\Pr(x = 1) = p$  e  $\Pr(x = 0) = 1 - p$ , temos que

$$\bar{x} = (1 - p)(0) + p(1) = p \quad (54)$$

#### Variável aleatória binomial

Sabemos que uma variável aleatória binomial  $x \rightarrow \text{Binomial}(n, p)$  é a soma de  $n$  variáveis aleatórias de Bernouilli  $y_i$  com parâmetro  $p$ , ou seja,  $x = y_1 + y_2 + \dots + y_n$ .

Veremos posteriormente que o valor esperado é uma operação linear, e, portanto,

$$\bar{x} = E\{x\} = E\{y_1\} + E\{y_2\} + \dots + E\{y_n\} = np \quad (55)$$

#### Variável aleatória de Poisson

A partir da função massa de probabilidade de uma v.a. de Poisson, definida em (26), temos que

$$\bar{x} = \sum_{k=0}^{\infty} \frac{k\lambda^k e^{-\lambda}}{k!} = \lambda \quad (56)$$

#### Variável aleatória uniforme

Para uma variável aleatória contínua uniforme entre  $a$  e  $b$ , temos que

$$\bar{x} = \int_{-\infty}^{\infty} xp_x(x)dx = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b+a}{2} \quad (57)$$

É fácil verificar também que a mediana é

$$\text{Med}(x) = \bar{x}$$

#### Variável aleatória exponencial

Para uma variável aleatória exponencial com parâmetro  $\lambda$ , definida em (40), temos que

$$\bar{x} = \int_0^{\infty} x\lambda e^{-\lambda x} dx = \frac{e^{-\lambda x}}{\lambda} (-\lambda x - 1) \Big|_0^{\infty} = \frac{1}{\lambda} \quad (58)$$

A mediana é dada por



$$F_x(\text{Med}(x)) = 1 - e^{-\lambda \text{Med}(x)} = \frac{1}{2} \Rightarrow e^{-\lambda \text{Med}(x)} = \frac{1}{2} \Rightarrow \text{Med}(x) = \frac{\ln 2}{\lambda} \quad (59)$$

Variável aleatória Gaussiana (normal)

Em (42) vimos a PDF de uma variável aleatória padrão. Podemos ter uma v.a. Gaussiana com parâmetro  $\mu$ ,

$$p_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} \quad (60)$$

que tem média

$$\begin{aligned} \bar{x} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (y + \mu) e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}} dy + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \end{aligned} \quad (61)$$

Como o termo dentro da primeira integral é uma função ímpar, a primeira integral é igual a zero, e no segundo termo, temos a integral de uma PDF, que sabemos que é igual a 1. Portanto, para uma v.a. Gaussiana

$$\bar{x} = \mu \quad (62)$$

Podemos verificar também, que, como a PDF é simétrica em torno de  $\mu$ ,  $F_x(\mu) = \frac{1}{2}$ , e, portanto,

$$\text{Med}(x) = \bar{x} = \mu \quad (63)$$

### i. Valor Esperado de uma Função e Momentos

Dada uma função de uma variável aleatória  $y = g(x)$ , temos que  $y$  também é uma variável aleatória. Se quisermos saber a média de  $y$ , não é necessário conhecer sua PDF, já que

$$E\{y\} = \bar{y} = \int_{-\infty}^{\infty} g(x) p_x(x) dx \quad (64)$$

Pela definição acima, podemos facilmente demonstrar ainda que o valor esperado é uma operação linear, ou seja, dadas as constantes  $a$ ,  $b$  e  $c$ ,

Exemplo 31

$$E\{af(x) + bg(x) + c\} = aE\{f(x)\} + bE\{g(x)\} + c \quad (65)$$

Uma fonte de tensão senoidal gera um sinal  $x(t) = A \cos(2\pi f_c t)$ . Suponha que a tensão desta fonte seja medida em um instante aleatório, sendo o valor medido uma variável aleatória  $x$ . Qual o valor médio e o rms (*root mean square*) desta medida?

Sendo o instante aleatório, podemos considerar que amostraremos um sinal  $x = A \cos \Theta$ , em que a fase aleatória  $\Theta$  é uma v.a. uniforme entre 0 e  $2\pi$ . Sendo assim, a média é

$$\bar{x} = \int_{-\infty}^{\infty} A \cos(\theta) p_{\Theta}(\theta) d\theta = \int_0^{2\pi} \frac{A \cos \theta}{2\pi} d\theta = \frac{A \sin \theta}{2\pi} \Big|_0^{2\pi} = 0$$

O valor rms é igual a  $\text{rms}(x) = \sqrt{E\{x^2\}}$ , com

$$E\{x^2\} = \int_0^{2\pi} \frac{A^2 \cos^2 \theta}{2\pi} d\theta = \frac{A^2}{2\pi} \int_0^{2\pi} \left( \frac{1}{2} + \frac{\cos(2\theta)}{2} \right) d\theta = \frac{A^2}{2}$$

e, portanto,  $\text{rms}(x) = \frac{A}{\sqrt{2}}$ .

### Momentos

Dada uma variável aleatória  $x$ , seu **momento de ordem  $n$**  (ou  $n$ -ésimo momento) é dado por

$$E\{x^n\} = \int_{-\infty}^{\infty} x^n p_x(x) dx \quad (66)$$

e o **momento central de ordem  $n$**  é dado por

$$E\{(x - \bar{x})^n\} = \int_{-\infty}^{\infty} (x - \bar{x})^n p_x(x) dx \quad (67)$$

ou, para uma variável aleatória discreta

$$\begin{aligned} E\{x^n\} &= \sum_{x_i \in \mathcal{S}} x_i^n P_x(x_i) \\ E\{(x - \bar{x})^n\} &= \sum_{x_i \in \mathcal{S}} (x_i - \bar{x})^n P_x(x_i) \end{aligned} \quad (68)$$

Dentre os diferentes momentos centrais, um dos mais utilizados é a **variância**

$$\text{var}(x) = \sigma_x^2 = E\{(x - \bar{x})^2\} \quad (69)$$

Expandindo a equação acima, podemos obter também a variância como

$$\begin{aligned} \text{var}(x) &= E\{x^2 - 2x\bar{x} + \bar{x}^2\} = \overline{x^2} - 2\bar{x}\bar{x} + \bar{x}^2 \\ &= E\{x^2\} - \bar{x}^2 \end{aligned} \quad (70)$$

O valor  $\sigma_x$  é conhecido como desvio padrão e indica o quanto a variável aleatória se afasta de sua média.

O momento de terceira ordem normalizado é chamado de **obliquidade** (*skewness*) e indica a assimetria da PDF:

$$\gamma(x) = \frac{E\{(x - \bar{x})^3\}}{\sigma^3} \quad (71)$$

O momento de quarta ordem normalizado é chamado de **curtose**:

$$\kappa(x) = \frac{E\{(x - \bar{x})^4\}}{\sigma^4} \quad (72)$$

### Função geradora de momentos

A função geradora de momentos é definida como

$$M_x(t) = E\{e^{tx}\}, \text{ com } t \in \mathbb{R} \quad (73)$$

Pela série de Taylor, temos que

$$e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \dots + \frac{t^n x^n}{n!} + \dots \quad (74)$$

$$M_x(t) = 1 + t\bar{x} + \frac{t^2 \overline{x^2}}{2!} + \dots + \frac{t^n \overline{x^n}}{n!} + \dots$$

Derivando-se a função geradora de momentos e considerando  $t = 0$ , temos que

$$\left. \frac{dM_x(t)}{dt} \right|_{t=0} = \bar{x}; \quad \dots; \quad \left. \frac{d^n M_x(t)}{dt^n} \right|_{t=0} = \overline{x^n} \quad (75)$$

Veremos agora o cálculo dos momentos para algumas variáveis aleatórias comuns

#### Variável aleatória de Bernouilli

Para uma v.a. de Bernouilli com  $P_x(1) = p$ ,

$$E\{x^2\} = (1-p)(0^2) + p(1^2) = p \quad (76)$$

e

$$\text{var}(x) = \overline{x^2} - \bar{x}^2 = p - p^2 \quad (77)$$

A função geradora de momentos é

$$M_x(t) = (1-p)e^0 + pe^t = 1-p + pe^t \quad (78)$$

Derivando-se, temos que

$$E\{x\} = \left. \frac{dM_x(t)}{dt} \right|_{t=0} = pe^t|_{t=0} = p$$

$$E\{x^2\} = \left. \frac{d^2 M_x(t)}{dt^2} \right|_{t=0} = pe^t|_{t=0} = p \quad (79)$$

$$E\{x^n\} = \left. \frac{d^n M_x(t)}{dt^n} \right|_{t=0} = pe^t|_{t=0} = p$$

#### Variável aleatória uniforme

Para uma v.a. uniforme entre  $a$  e  $b$ , temos que

$$E\{x^2\} = \int_a^b \frac{x^2}{b-a} dx = \left. \frac{x^3}{3(b-a)} \right|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3} \quad (80)$$

$$\text{var}(x) = \overline{x^2} - \bar{x}^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{b+a}{2}\right)^2 = \frac{(b-a)^2}{12}$$

#### Variável aleatória exponencial

Para uma v.a. exponencial com parâmetro  $\lambda$ , temos que a função geradora de momentos é

$$M_x(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \int_0^\infty \lambda e^{(t-\lambda)x} dx = \left(1 - \frac{t}{\lambda}\right)^{-1} \quad (81)$$

Fazendo a derivada e considerando  $t = 0$ , temos que

$$\begin{aligned}
E\{x\} &= \left. \frac{dM_x(t)}{dt} \right|_{t=0} = -\frac{1}{\lambda}(-1) \left(1 - \frac{t}{\lambda}\right)^{-2} \Big|_{t=0} = \frac{1}{\lambda} \\
E\{x^2\} &= \left. \frac{d^2 M_x(t)}{dt^2} \right|_{t=0} = \frac{1}{\lambda} \left(-\frac{1}{\lambda}\right) (-2) \left(1 - \frac{t}{\lambda}\right)^{-3} \Big|_{t=0} = \frac{2}{\lambda^2}
\end{aligned} \tag{82}$$

e, portanto,

$$\text{var}(x) = \overline{x^2} - \bar{x}^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \tag{83}$$

### Variável Aleatória Gaussiana (Normal)

A v.a. Gaussiana padrão, definida em (42) possui variância igual a 1. Uma v.a. Gaussiana genérica é definida pela seguinte PDF:

$$p_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{84}$$

e sua função geradora de momentos é

$$M_x(t) = e^{t\mu + \frac{\sigma^2 t^2}{2}} \tag{85}$$

Podemos então achar seus momentos como

$$\begin{aligned}
E\{x\} &= \left. \frac{dM_x(t)}{dt} \right|_{t=0} = (\mu + t\sigma^2) e^{t\mu + \frac{\sigma^2 t^2}{2}} \Big|_{t=0} = \mu \\
E\{x^2\} &= \left. \frac{d^2 M_x(t)}{dt^2} \right|_{t=0} = [\sigma^2 + (\mu + t\sigma^2)^2] e^{t\mu + \frac{\sigma^2 t^2}{2}} \Big|_{t=0} = \sigma^2 + \mu^2
\end{aligned} \tag{86}$$

e, portanto,

$$\text{var}(x) = \overline{x^2} - \bar{x}^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \tag{87}$$

Consequentemente, uma v.a. Gaussiana é definida pela sua média e variância, e usamos a notação  $x \rightarrow \mathcal{N}(\mu, \sigma^2)$ .

Muitas vezes queremos saber a probabilidade de a variável aleatória assumir um valor maior que um certo  $x$ :

$$\Pr(x > x) = \int_x^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\alpha-\mu)^2}{2\sigma^2}} d\alpha \tag{88}$$

e, substituindo-se  $\frac{\alpha-\mu}{\sigma} = y$ , podemos reescrever

$$\Pr(x > x) = \int_{\frac{x-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = Q\left(\frac{x-\mu}{\sigma}\right), \tag{89}$$

com a função  $Q$  definida em (43)

### j. Variáveis Aleatórias Conjuntas

Dadas duas variáveis aleatórias  $x$  e  $y$ , podemos definir a CDF conjunta

$$F_{x,y}(x, y) = \Pr(x \leq x \cap y \leq y) \tag{90}$$

e a PDF conjunta

$$p_{x,y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{x,y}(x, y). \quad (91)$$

Assim, a probabilidade em um certo intervalo é dada por

$$\Pr(x_1 < x \leq x_2 \cap y_1 < y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} p_{x,y}(x, y) dy dx, \quad (92)$$

e, sabemos ainda que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{x,y}(x, y) dy dx = 1. \quad (93)$$

Se quisermos achar a **PDF marginal**, ou seja, a PDF de apenas uma das variáveis aleatórias, temos que

$$p_x(x) = \int_{-\infty}^{\infty} p_{x,y}(x, y) dy. \quad (94)$$

Da mesma forma, a CDF marginal é dada por

$$F_x(x) = F_{x,y}(x, \infty). \quad (95)$$

Assim como temos uma probabilidade condicional para variáveis aleatórias discretas, podemos ter uma PDF condicional

$$p_{y|x}(y|x) = \frac{p_{x,y}(x, y)}{p_x(x)}, \quad (96)$$

e, conseqüentemente, a Regra de Bayes

$$p_{x|y}(x|y) = \frac{p_{x,y}(x, y)}{p_y(y)} = \frac{p_{y|x}(y|x)p_x(x)}{p_y(y)}. \quad (97)$$

Similarmente à lei da probabilidade total para v.a.s discretas, podemos achar a PDF marginal com

$$p_y(y) = \int_{-\infty}^{\infty} p_{x,y}(x, y) dx = \int_{-\infty}^{\infty} p_x(x) p_{y|x}(y|x) dx. \quad (98)$$

Variáveis aleatórias contínuas são **independentes** se

Exemplo 32

$$p_{x|y}(x|y) = p_x(x) \Leftrightarrow p_{y|x}(y|x) = p_y(y) \Leftrightarrow p_{x,y}(x, y) = p_x(x)p_y(y). \quad (99)$$

Variável aleatória Gaussiana complexa circularmente simétrica e distribuição de Rayleigh

Considere uma v.a. complexa  $z = x + jy$ , em que  $x, y \rightarrow \mathcal{N}(0, \sigma^2)$ , e  $x, y$  são independentes. Sendo assim

$$p_{x,y}(x, y) = p_x(x)p_y(y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

A PDF conjunta pode ser vista na Figura 16, de dois ângulos diferentes. Na figura da direita a PDF é vista com níveis de cores diferentes, e podemos visualizar o conceito de circularmente simétrica.

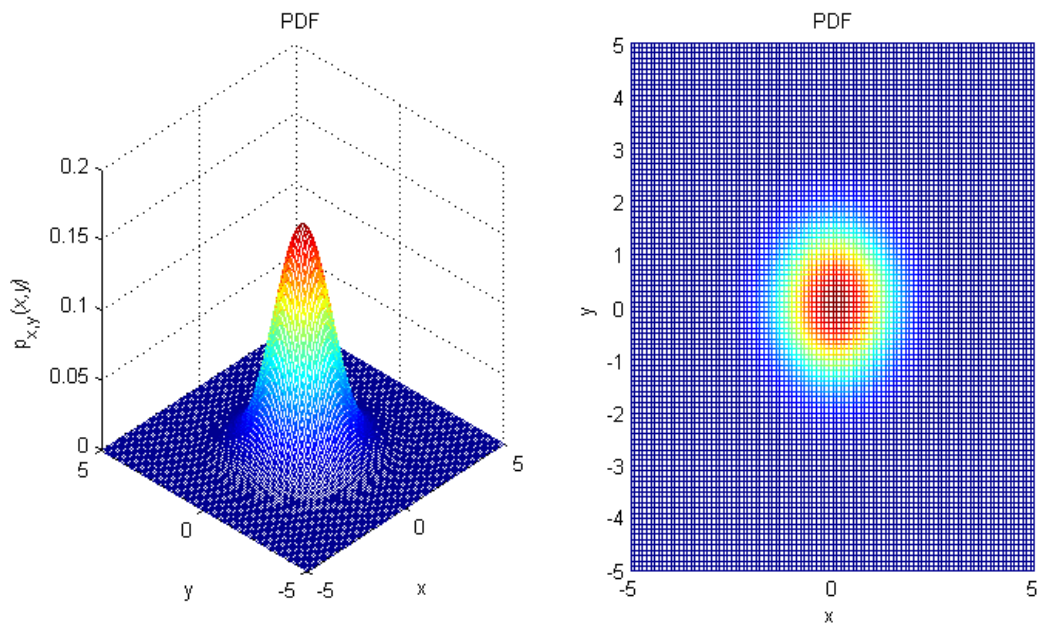


Figura 16. V.a. Gaussiana complexa circularmente simétrica

Neste caso, é de interesse acharmos a a PDF da amplitude  $r = |z| = \sqrt{x^2 + y^2}$  e da fase  $\theta = \angle z = \text{atan}\frac{y}{x} + \pi u(-x)$ , ou seja, qual  $p_r(r)$  e  $p_\theta(\theta)$ .

Podemos considerar uma área infinitesimal de tamanho  $d\theta dr$  em torno de  $r$  e  $\theta$ . A probabilidade desta área é dada pela área multiplicada pela densidade de probabilidade neste ponto. Neste caso,

$$\Pr(r \in dr, \theta \in d\theta) = p_{r,\theta}(r, \theta) dr d\theta$$

Podemos representar isto também em coordenadas Euclidianas, como visto na Figura 17.

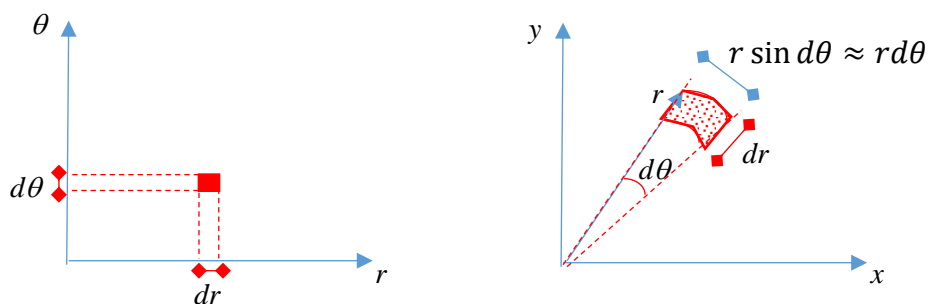


Figura 17. Probabilidade da amplitude de uma v.a. complexa

Em coordenadas Euclidianas, temos que

$$\Pr(r \in dr, \theta \in d\theta) = p_{x,y}(x, y) r d\theta dr$$

Igualando as duas equações, temos que

$$p_{r,\theta}(r, \theta) dr d\theta = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} r d\theta dr$$

$$\Leftrightarrow p_{r,\theta}(r, \theta) = \frac{r}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \text{ para } r \geq 0 \text{ e } 0 \leq \theta < 2\pi$$

e, considerando a probabilidade marginal

$$p_r(r) = \int_0^{2\pi} \frac{r}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} d\theta, \text{ para } r \geq 0,$$

ou seja,

$$p_r(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} u(r).$$

Esta distribuição é conhecida como **distribuição de Rayleigh**, que pode ser vista na Figura 18.

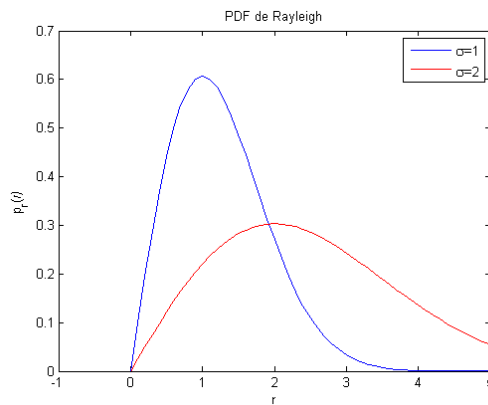
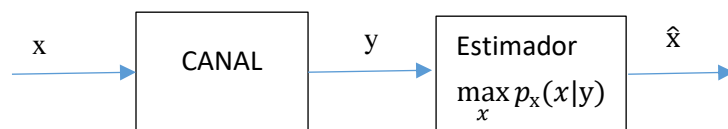


Figura 18. Distribuição de Rayleigh

Um caso de particular interesse de variáveis aleatórias conjuntas em comunicações é quando enviamos um sinal aleatório  $x$  em um canal de comunicações, como vimos na Figura 19. Na saída, devido a distorções e ruído, recebemos um sinal aleatório  $y$ , que tem uma probabilidade conjunta com  $x$ , dependendo da caracterização do sistema.



Exemplo 33

Figura 19. Canal de Comunicações

Um sinal aleatório  $x$  com distribuição exponencial e média unitária passa por um canal ruidoso, cujo ruído  $n$  é v.a. com distribuição uniforme entre  $-1$  e  $+1$ . Se observarmos um determinado valor  $y$  na saída, qual a PDF condicional do sinal  $x$ ?

Temos que

$$p_x(x) = e^{-x}u(x)$$

$$y = x + n \Rightarrow p_{y|x}(y|x) = \frac{1}{2} \text{rect}\left(\frac{y-x}{2}\right)$$

e queremos descobrir

$$p_{x|y}(x|y) = \frac{p_{y|x}(y|x)p_x(x)}{p_y(y)}$$

Falta encontrar

$$p_y(y) = \int_{-\infty}^{\infty} p_x(x)p_{y|x}(y|x)dx = \int_0^{\infty} \frac{e^{-x}}{2} \text{rect}\left(\frac{y-x}{2}\right) dx$$

Lembrando que a função rect na integral é igual a 1 para  $y - 1 \leq x \leq y + 1$ , temos que

$$p_y(y) = \begin{cases} 0 & , y < -1 \\ \int_0^{y+1} \frac{e^{-x}}{2} dx = \frac{1}{2}(1 - e^{-y-1}) & , -1 \leq y < 1 \\ \int_{y-1}^{y+1} \frac{e^{-x}}{2} dx = \frac{e^{-z}}{2}(e - e^{-1}) & , y \geq 1 \end{cases}$$

e, portanto,

$$p_{x|y}(x|y) = \begin{cases} e^{-x}u(x)\text{rect}\left(\frac{y-x}{2}\right)\frac{1}{1 - e^{-y-1}} & , -1 \leq y < 1 \\ e^{-x}u(x)\text{rect}\left(\frac{y-x}{2}\right)\frac{e^y}{e - e^{-1}} & , y > 1 \end{cases}$$

Os conceitos vistos acima para duas variáveis aleatórias conjuntas podem ser facilmente expandidos para mais de duas v.a.s.

### Valor esperado

Dada uma certa função de duas variáveis aleatórias  $g(x,y)$ , o valor esperado desta função é dado por

$$E\{g(x,y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)p_{x,y}(x,y)dxdy \quad (100)$$

Como vimos, o valor esperado é uma função linear, e, portanto

$$\text{se } g(x,y) = ax + by \Rightarrow E\{g(x,y)\} = aE\{x\} + bE\{y\}. \quad (101)$$

Se  $x$  e  $y$  forem independentes, e  $g(x,y) = g_1(x)g_2(y)$  temos que

$$\begin{aligned} E\{g_1(x)g_2(y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x)g_2(y)p_x(x)p_y(y)dxdy \\ &= \int_{-\infty}^{\infty} g_1(x)p_x(x)dx \int_{-\infty}^{\infty} g_2(y)p_y(y)dy \\ &= E\{g_1(x)\}E\{g_2(y)\} \end{aligned} \quad (102)$$

Como caso particular, temos que se  $x$  e  $y$  forem independentes,

$$E\{xy\} = E\{x\}E\{y\} \quad (103)$$



Se multiplicarmos uma variável aleatória por um escalar  $a$ , temos o seguinte impacto nas suas média e variância:

Se tivermos a **soma**  $z$  de duas variáveis aleatórias  $x$  e  $y$  temos que

$$\begin{aligned} \text{se } z = x + y &\Rightarrow \\ \bar{z} &= \bar{x} + \bar{y} \\ \sigma_z^2 &= E\{(z - \bar{z})^2\} = E\{(x - \bar{x} + y - \bar{y})^2\} \\ &= \sigma_x^2 + \sigma_y^2 - 2E\{(x - \bar{x})(y - \bar{y})\} \end{aligned} \quad (104)$$

Ainda, se  $x$  e  $y$  forem **independentes**

$$\begin{aligned} E\{(x - \bar{x})(y - \bar{y})\} &= E\{(x - \bar{x})\}E\{(y - \bar{y})\} = 0 \\ \Rightarrow \sigma_z^2 &= \sigma_x^2 + \sigma_y^2 \end{aligned} \quad (105)$$

No Exemplo 33, qual a variância do sinal  $y$  na saída do sistema?

Exemplo 34

Podemos considerar que o sinal  $x \rightarrow \text{Exp}(\lambda = 1)$  e o ruído  $n \rightarrow \text{Unif}(a = -1; b = 1)$  são independentes. Desta forma a variância do sinal  $y = x + n$  é dada por

$$\sigma_y^2 = \sigma_x^2 + \sigma_n^2 = \frac{1}{\lambda^2} + \frac{(b - a)^2}{12} = 1 + \frac{1}{3} = \frac{4}{3}$$

Exemplo 35

#### Distorção em sistema PCM

Em um sistema PCM a distorção no sinal reconstruído pode vir de duas causas independentes. A primeira é o erro de quantização e a segunda é devida ao possível erro dos bits transmitidos.

Ou seja, uma sequência de amostras  $m_k$  é inicialmente quantizada, gerando a sequência quantizada  $\tilde{m}_k$ , e esta sequência é mapeada em  $n$  bits/amostra  $b_0^{(k)} b_1^{(k)} \dots b_{n-1}^{(k)}$ . Estes bits são enviados em um canal simétrico binário com probabilidade de erro  $p$  de modo que recebemos  $\hat{b}_0^{(k)} \hat{b}_1^{(k)} \dots \hat{b}_{n-1}^{(k)}$ , com  $\Pr(\hat{b}_i^{(k)} \neq b_i^{(k)}) = p$ . A partir destes bits é que será reconstruída a mensagem, obtendo-se a estimativa  $\hat{m}_k$ .

Esta estimativa é dada por

$$\hat{m}_k = m_k + q_k + \epsilon_k = m_k + e_k,$$

em que  $q_k$  é o erro de quantização e  $\epsilon_k$  é o erro causada por erros de bit.

Queremos saber qual a potência da distorção total  $e_k$ .

Podemos ignorar o índice  $k$  para simplificar. O erro de quantização é uma variável aleatória uniforme entre  $-\frac{m_p}{L}$  e  $\frac{m_p}{L}$ , em que  $m_p$  é a amplitude de pico do sinal e  $L = 2^n$  é o número de níveis de quantização.

Como veremos mais tarde, a potência de um sinal é dada por  $P_x = E\{x^2\}$ . Para o erro de quantização, temos, portanto, que

$$P_q = E\{q^2\} = \sigma_q^2 = \frac{1}{3} \frac{m_p^2}{2^{2n}}$$

Já a distorção devida ao erro de bits depende da posição do bit na palavra. Se for o bit menos significativo (bit  $i = n$ ), o erro vai ser igual a um intervalo de quantização, ou seja,  $\epsilon_n = \pm \Delta v = \pm \frac{2m_p}{L} = \pm \frac{2m_p}{2^n}$ . Porém, se for o bit mais significativo (bit  $i = 1$ ), o erro vai ser bem maior,  $\epsilon_1 = \pm m_p$ . Para bits em outras posições, podemos generalizar como  $\epsilon_i = \pm \frac{2m_p}{2^i}$ . É razoável supor também que  $\bar{\epsilon} = 0$ .

Supondo agora que a probabilidade de erro  $p \ll 1$ , é improvável que tenhamos mais de um erro de bit em cada amostra, e, desta forma

$$\sigma_\epsilon^2 = E\{\epsilon^2\} = \sum_{i=1}^n \epsilon_i^2 P_\epsilon(\epsilon_i) = p \sum_{i=1}^n \epsilon_i^2 = p \sum_{i=1}^n \frac{4m_p^2}{2^{2i}} = \frac{4m_p^2 p (2^{2n} - 1)}{3(2^{2n})}$$

Portanto,

$$P_e = E\{e^2\} = \sigma_e^2 = \sigma_q^2 + \sigma_\epsilon^2 = \frac{m_p^2}{3(2^{2n})} (1 + 4p(2^{2n} - 1))$$

#### k. Lei dos Grandes Números

Sejam  $n$  realizações independentes de um experimento aleatório, de modo que em cada experimento tenhamos como resultado uma variável aleatória  $x_i$ . Ou seja, temos  $n$  variáveis aleatórias i.i.d. (independentes, identicamente distribuídas)  $x_0, x_1, x_2, \dots, x_{n-1}$ , com  $E\{x_i\} = \mu$  e  $\text{var}(x_i) = \sigma^2 < \infty$ .

Seja a média amostral de  $n$  realizações

$$\tilde{x}_n = \frac{1}{n} \sum_{i=0}^{n-1} x_i \quad (106)$$

A lei dos grandes números indica que esta média amostral tende à média estatística à medida que o número de realizações tende a infinito, ou seja, se realizarmos um número muito grande de experimentos, teremos uma média amostral muito próxima à média estatística (ou valor esperado). Em notação matemática, temos

$$\tilde{x}_n \xrightarrow{n \rightarrow \infty} \mu \quad (107)$$

Esta convergência pode ser expressa de duas formas. Pela **lei fraca dos grandes números** temos uma convergência em probabilidade, em que, para qualquer  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \Pr(|\tilde{x}_n - \mu| < \epsilon) = 1. \quad (108)$$

Pela **lei forte dos grandes números** a convergência se dá com quase certeza, ou seja,

$$\Pr\left(\lim_{n \rightarrow \infty} |\tilde{x}_n - \mu| < \epsilon\right) = 1 \quad (109)$$

A lei dos grandes números pode ser visualizada na

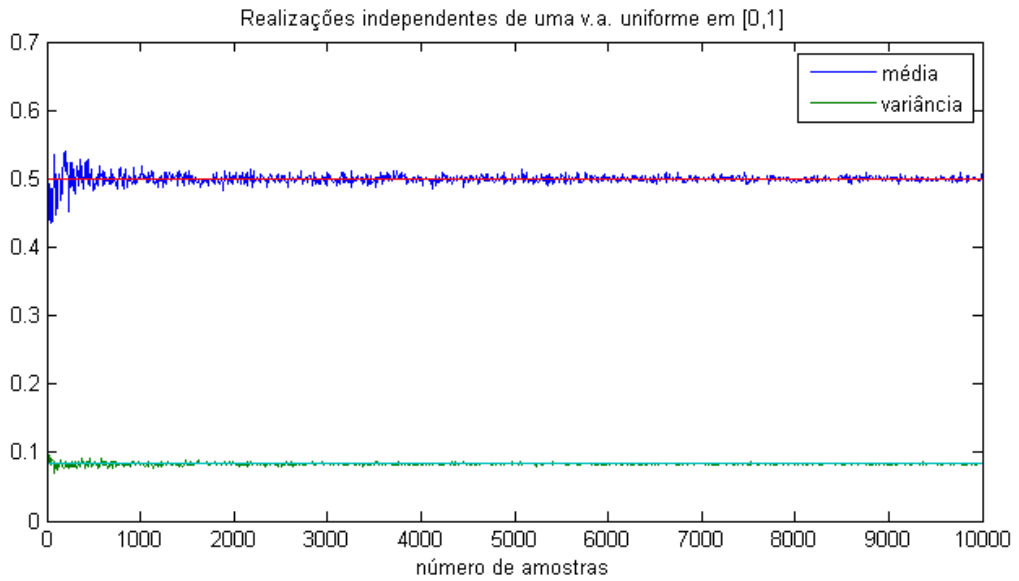


Figura 20. Lei dos grandes números

### I. Soma de duas variáveis aleatórias

Dadas duas variáveis aleatórias  $x$  e  $y$ , é muitas vezes necessário sabermos a função de distribuição de probabilidade da soma delas. Considerando  $z = x + y$ , temos que a sua CDF é

$$F_z(\zeta) = \Pr(z < \zeta) = \Pr(x \leq \infty; y \leq \zeta - x) \quad (110)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\zeta-x} p_{x,y}(x, y) dy dx$$

A partir da CDF, podemos obter a PDF

$$p_z(\zeta) = \frac{dF_z(\zeta)}{d\zeta} = \int_{-\infty}^{\infty} p_{x,y}(x, \zeta - x) dx \quad (111)$$

Se  $x$  e  $y$  forem independentes, temos que

Exemplo 36

$$p_z(\zeta) = \int_{-\infty}^{\infty} p_x(x) p_y(\zeta - x) dx = p_x(\zeta) * p_y(\zeta) \quad (112)$$

#### Soma de variáveis aleatórias uniformes

Se  $x_1$  e  $x_2$  forem duas variáveis aleatórias independentes uniformes em  $[-1,1]$ , e  $y = x_1 + x_2$ , temos que

$$p_y(y) = p_{x_1}(y) * p_{x_2}(y) = \frac{1}{2} \text{rect}\left(\frac{y}{2}\right) * \frac{1}{2} \text{rect}\left(\frac{y}{2}\right) = \frac{1}{2} \Delta\left(\frac{y}{4}\right).$$

Na Figura 21 podemos ver a PDF da soma de até 4 v.a.'s uniformes independentes, e, para a soma de 4 v.a.s, comparamos a PDF da soma com a PDF de uma Gaussiana com a mesma média e variância. Podemos ver que as curvas são bastante aproximadas, e, à medida que aumentamos o número de v.a.s somadas, a PDF vai ficando cada vez mais próxima de uma Gaussiana.

Na mesma figura, à direita, vemos o histograma da soma de várias v.a.s uniformes independentes (com 10.000 amostras de cada), e vemos que a forma do histograma tende à forma da PDF obtida pela convolução, como era de se esperar.

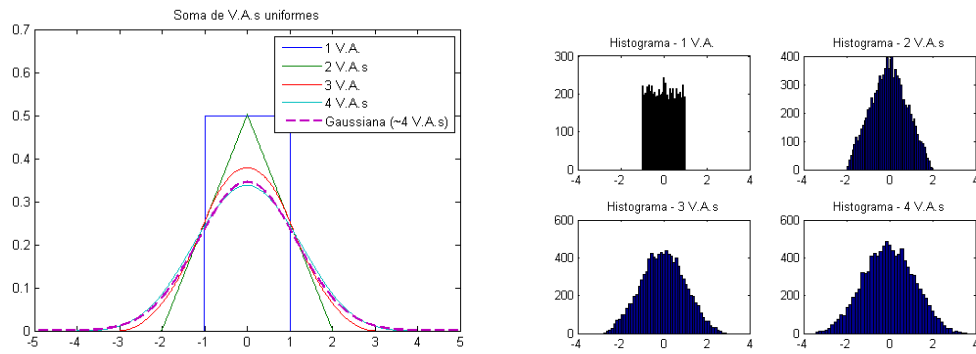


Figura 21. PDF e histograma da soma de variáveis aleatórias uniformes

### Soma de variáveis aleatórias exponenciais

Agora, se  $x_1$  e  $x_2$  forem duas variáveis aleatórias exponenciais com  $\lambda=1$ , e  $y = x_1 + x_2$ , temos que

$$p_y(y) = \int_{-\infty}^{\infty} e^{-\tau} u(\tau) e^{-(y-\tau)} u(y-\tau) d\tau = \int_0^y e^{-y} d\tau = ye^{-y} u(y).$$

Na Figura 22 podemos ver agora a soma de v.a.s exponenciais comparadas com a PDF de uma Gaussiana. Assim como no caso de v.a.s uniformes podemos ver que as curvas se aproximam de uma Gaussiana à medida que aumentamos o número de v.a.s somadas, embora mais lentamente.

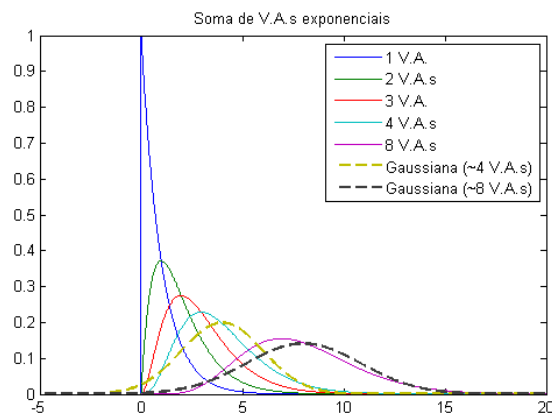


Figura 22. PDF da soma de variáveis aleatórias exponenciais

Uma propriedade interessante da variável aleatória Gaussiana é que a soma de duas variáveis aleatórias Gaussianas também é uma Gaussiana, com média e variância dadas pelas propriedades em (104).

### m. Teorema Central do Limite

Como vimos na Seção 1.m, a soma de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.)<sup>1</sup> tende a uma v.a. Gaussiana. Mais especificamente, se tivermos  $n$  v.a.s i.i.d.  $x_1, x_2, \dots, x_{n-1}$  com média  $\mu$  e variância  $\sigma^2$ , podemos considerar a soma normalizada

$$y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{x_i - \mu}{\sigma}. \quad (113)$$

O Teorema Central do Limite diz que

$$\lim_{n \rightarrow \infty} Pr(y_n \leq y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}} d\alpha, \quad (114)$$

ou seja, a soma normalizada tende a uma v.a. Gaussiana,  $y_n \xrightarrow{n \rightarrow \infty} \mathcal{N}(0,1)$ .

Alternativamente, podemos considerar a soma não normalizada

$$z_n = \sum_{i=1}^n x_i = \sigma\sqrt{n}(y_n - n\mu) \quad (115)$$

e, para valores grandes de  $n$ , podemos aproximá-la por uma Gaussiana,

$$z_n \approx \mathcal{N}(n\mu, n\sigma^2). \quad (116)$$

#### Exemplo 38

Um pacote de  $N_{bytes}$  bytes é enviado em um canal BSC com probabilidade de erro  $p$ . Sabendo que podemos corrigir até  $t$  erros por meio de um código corretor de erros, qual a probabilidade de erro de pacotes?

Podemos considerar a variável de Bernoulli  $x_i = 1$ , se ocorrer um erro na transmissão do bit  $i$  e  $x_i = 0$ , caso não ocorra um erro, com  $1 \leq i \leq 8N_{bytes}$ .

O número de erros em um pacote é dado então por

$$v = \sum_{i=1}^{8N_{bytes}} x_i,$$

que é uma variável aleatória binomial. A probabilidade de termos mais que  $t$  erros é dada por

$$Pr(v > t) = \sum_{m=t+1}^{8N_{bytes}} \binom{8N_{bytes}}{m} p^m (1-p)^{8N_{bytes}-m}.$$

Porém, para valores grandes de  $t$  e  $N_{bytes}$  este cálculo fica inviável, mesmo numericamente. Por exemplo, considerando  $N = 128$  e  $t = 30$ , esta conta causa *overflow* em uma representação de ponto flutuante tipo *double*. Podemos, porém, achar uma aproximação pelo Teorema Central do Limite. Ou seja, podemos considerar  $v$  aproximadamente uma variável aleatória Gaussiana. Se tivermos  $p = 0,01$ , sabemos que

---

<sup>1</sup> Para variáveis aleatórias não identicamente distribuídas, o Teorema Central do Limite também vale, desde que seja satisfeitas algumas condições, o que foge de nosso escopo.

$$E\{v\} = \bar{v} = 128(8)\bar{x}_i = 1024p = 10,24$$

$$\text{var}(v) = \sigma_v^2 = 1024\sigma_{x_i}^2 = 1024(p - p^2) = 10,1376$$

e, portanto,

$$\Pr(v > t) \approx Q\left(\frac{t - \bar{v}}{\sigma_v}\right) = Q(6,206) = 1,925 \times 10^{-10}$$

## n. Correlação

Uma análise importante em probabilidade e estatística é a da relação entre variáveis aleatórias distintas. Por exemplo, podemos investigar estatisticamente se há uma relação entre o consumo de cigarros e a incidência de câncer na população. Em comunicações, a existência de dependência entre variáveis é fundamental, já que contamos que há uma relação de dependência entre o sinal enviado e o sinal recebido em um receptor, a fim de podermos estimar ou detectar o sinal que foi enviado.

Começamos definindo a **covariância** entre duas v.a.s  $x$  e  $y$

$$\begin{aligned}\sigma_{xy} &= E\{(x - \bar{x})(y - \bar{y})\} \\ &= E\{xy\} - \bar{x}\bar{y}\end{aligned}\tag{117}$$

Dizemos que duas v.a.s são **descorrelatadas**

$$\text{se } \sigma_{xy} = 0 \Leftrightarrow E\{xy\} = \bar{x}\bar{y}\tag{118}$$

A medida comumente utilizada para se determinar a correlação entre v.a.s distintas é o **coeficiente de correlação** de Pearson

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \text{ com } -1 \leq \rho_{x,y} \leq 1\tag{119}$$

O coeficiente de correlação indica a dependência linear entre duas variáveis aleatórias. Suponhamos por exemplo que  $y = ax$ , com  $a \in \mathbb{R}$  uma constante qualquer. Neste caso, teremos que

$$\begin{aligned}\sigma_y^2 &= a^2 \sigma_x^2 \Rightarrow \sigma_y^2 = |a| \sigma_x \\ \sigma_{xy} &= E\{xax\} - \bar{x}a\bar{x} = a\bar{x}^2 - a\bar{x}^2 = a\sigma_x^2 \\ \Rightarrow \rho_{xy} &= \frac{a\sigma_x^2}{|a|\sigma_x^2} = \text{sign}(a),\end{aligned}\tag{120}$$

ou seja, no caso de uma dependência linear total, a coeficiente de correlação será  $\rho_{xy} = \pm 1$ , com +1 indicando uma dependência linear direta, e -1 uma inversa.

Exemplo 59  
Em particular, se  $x$  e  $y$  forem v.a.s independentes, temos que  $E\{xy\} = E\{x\}E\{y\} \Leftrightarrow \sigma_{xy} = 0$ . Consequentemente, como vimos em (118) elas são também descorrelatadas. Notem, porém, que o inverso nem sempre é verdade, ou seja, duas variáveis aleatórias descorrelatadas não são necessariamente independentes.

Considere duas variáveis aleatórias  $z_1$  e  $z_2$ , representando o resultado de dois dados independentes. Considere ainda que temos duas v.a.s, representando a soma,  $x = z_1 + z_2$ , e a diferença entre elas,  $y = z_1 - z_2$ . Vamos calcular a covariância entre elas.

Sabemos que

$$P_x(2) = P_x(12) = P_y(-5) = P_y(5) = 1/36$$

$$P_x(3) = P_x(11) = P_y(-4) = P_y(4) = 2/36$$

$$P_x(4) = P_x(10) = P_y(-3) = P_y(3) = 3/36$$

$$P_x(5) = P_x(9) = P_y(-2) = P_y(2) = 4/36$$

$$P_x(6) = P_x(8) = P_y(-1) = P_y(1) = 5/36$$

$$P_x(7) = P_y(0) = 6/36$$

e, portanto,  $E\{x\} = 7$  e  $E\{y\} = 0$ .

Para acharmos a covariância, devemos agora encontrar  $E\{xy\}$ , que pode ser achada como

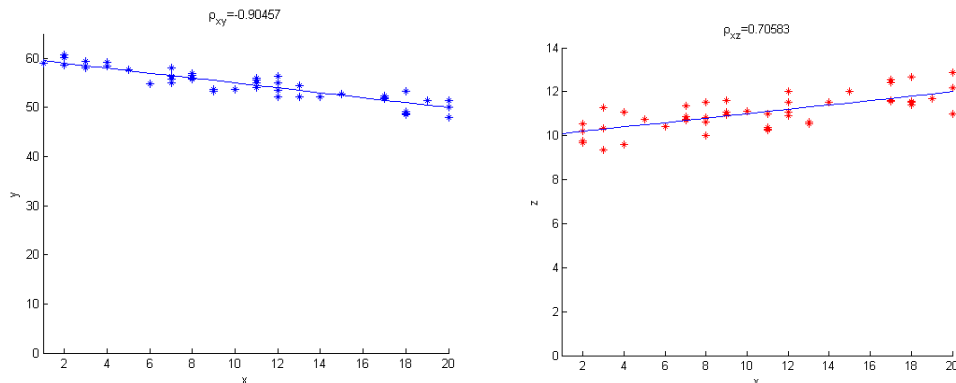
$$\begin{aligned} E\{xy\} &= \sum_{\zeta_1=1}^6 \sum_{\zeta_2=1}^6 (\zeta_1 + \zeta_2)(\zeta_1 - \zeta_2) P_{z_1 z_2}(\zeta_1, \zeta_2) \\ &= \frac{1}{36} \sum_{\zeta_1=1}^6 \sum_{\zeta_2=1}^6 (\zeta_1 + \zeta_2)(\zeta_1 - \zeta_2) = \frac{1}{36} \sum_{\zeta_1=1}^6 \sum_{\zeta_2=1}^6 (\zeta_1^2 - \zeta_2^2) = 0 \end{aligned}$$

Portanto,  $\sigma_{xy} = E\{xy\}E\{x\}E\{y\} = 0$  e as v.a.s são descorrelatadas. Entretanto, elas são claramente dependentes, por exemplo  $P_{y|x}(-5|2) = 0 \neq P_y(-5)$ .

Uma concepção equivocada do significado da correlação é achar que o fato de duas variáveis terem uma correlação implica uma ser a causa da outra, como podemos ver no exemplo abaixo.

Exemplo 40

Suponhamos que coletamos os dados de três eventos aleatórios em variáveis aleatórias  $x$ , representado a colocação do Flamengo no final do Campeonato Brasileiro,  $y$ , o consumo de cerveja per capita no Brasil (em litros/habitante/ano) e  $z$ , a taxa de suicídios média (suicídios/100.000 habitantes)<sup>2</sup> no mesmo ano. Os dados e os coeficientes de correlação medidos estão na Figura 23.



<sup>2</sup>Os dados são fictícios

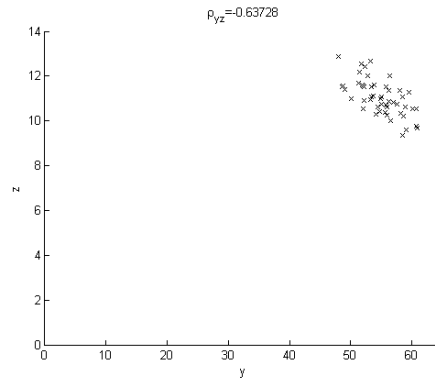


Figura 23. Exemplo de correlação

Vemos que existe uma correlação negativa entre as v.a.s  $y$  e  $z$ , o que poderia levar a crer que o aumento no consumo de cerveja ajudaria a reduzir a taxa de suicídios. Porém, analisando-se os gráficos da correlação entre  $x$  e  $y$  e entre  $x$  e  $z$ , podemos ver que a causa tanto do aumento do consumo de cerveja como na diminuição dos suicídios é o bom desempenho do Flamengo no campeonato e a reação de sua torcida.

Exemplo 41

Podem ser encontrados inúmeros exemplos de correlações espúrias, algumas delas podem ser vistas no site <http://www.tylervigen.com/spurious-correlations>, como por exemplo a grande correlação entre o número de filmes com o Nicolas Cage e o número de pessoas afogadas em piscinas nos EUA a cada ano, com um coeficiente de correlação  $\rho = 0,67$ .

#### o. Estimação LMS

Suponha que queiramos estimar uma variável aleatória  $x$  a partir da observação de uma outra variável aleatória  $y$ , ou seja queremos obter a estimativa  $\hat{x} = f(y)$ . Qual seria o melhor estimador  $f(\cdot)$ ? A abordagem mais simples é escolhermos um estimador linear  $\hat{x} = ay$ , e o critério para escolhermos o parâmetro  $a$  é minimizarmos o erro quadrático médio, ou seja, queremos

$$a = \underset{a}{\operatorname{argmin}} \overline{\epsilon^2} = \underset{a}{\operatorname{argmin}} E\{(x - \hat{x})^2\} \quad (121)$$

Este é o critério MSE (*mean square error*), e o estimador linear que satisfaz este critério é o estimador LMS (*least mean squares*).

Para simplificar, vamos supor que  $\bar{x} = \bar{y} = 0$ . Supondo um estimador linear  $\hat{x} = ay$ , temos que

$$\overline{\epsilon^2} = E\{(x - ay)^2\} = E\{\overline{x^2} + a^2\overline{y^2} - 2aE\{xy\}\} \quad (122)$$

Vamos chamar  $E\{xy\} = R_{xy}$ . Queremos minimizar o erro quadrático em relação a  $a$ , ou seja, queremos

$$\begin{aligned} \frac{d\overline{\epsilon^2}}{da} &= 2a\overline{y^2} - 2E\{xy\} = 0 \\ \Rightarrow a &= \frac{E\{xy\}}{\overline{y^2}} = \frac{R_{xy}}{R_{yy}} \end{aligned} \quad (123)$$



Neste caso, o erro é dado por

$$\epsilon = x - ay = x - \frac{R_{xy}}{R_{yy}}y \Rightarrow y\epsilon = xy - \frac{R_{xy}}{R_{yy}}y^2 \quad (124)$$

e, conseqüentemente,

$$E\{y\epsilon\} = E\{xy\} - \frac{R_{xy}}{R_{yy}}E\{y^2\} = R_{xy} - R_{xy} = 0. \quad (125)$$

Isto é conhecido como o princípio da ortogonalidade, que é uma condição para que o estimador linear tenha o erro quadrático mínimo.

É de interesse sabermos também qual o erro quadrático que podemos atingir caso utilizemos o estimador óptimo. Teremos então

$$\begin{aligned} E\{\epsilon^2\} &= E\left\{\left(x - \frac{R_{xy}}{R_{yy}}\right)\epsilon\right\} = E\{x\epsilon\} = E\{x(x - ay)\} \\ &= R_{xx} - aR_{xy} \end{aligned} \quad (126)$$

Caso  $\bar{x} \neq 0$ , podemos trabalhar com variáveis auxiliares  $x' = x - \bar{x}$  e  $y' = y - \bar{y}$ , de modo que  $\bar{x}' = \bar{y}' = 0$ . Assim,

$$\begin{aligned} \hat{x}' &= ay' \\ a &= \frac{E\{(x - \bar{x})(y - \bar{y})\}}{E\{(y - \bar{y})^2\}} = \frac{\sigma_{xy}}{\sigma_y^2} \\ \hat{x} &= \hat{x}' + \bar{x} = \frac{\sigma_{xy}}{\sigma_y^2}(y - \bar{y}) + \bar{x} \end{aligned} \quad (127)$$

Podemos estender a análise anterior se tivermos mais de uma variável de observação, ou seja, suponhamos que observamos as variáveis aleatórias  $y_1, y_2, \dots, y_n$ , todas com média nula. O estimador linear óptimo será dado por

$$\hat{x} = a_1y_1 + a_2y_2 + \dots + a_ny_n \quad (128)$$

Com

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} R_{y_1y_1} & R_{y_1y_2} & \dots & R_{y_1y_n} \\ R_{y_2y_1} & R_{y_2y_2} & \dots & R_{y_2y_n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{y_ny_1} & R_{y_ny_2} & \dots & R_{y_ny_n} \end{bmatrix}^{-1} \begin{bmatrix} R_{xy_1} \\ R_{xy_2} \\ \vdots \\ R_{xy_n} \end{bmatrix} \quad (129)$$

o erro médio quadrático dado por

$$E\{\epsilon^2\} = R_{xx} - (a_1R_{xy_1} + a_2R_{xy_2} + \dots + R_{xy_n}). \quad (130)$$

Queremos estimar o ganho de canal  $h$ , em que dada uma entrada  $x_k$  no instante  $k$ , temos na saída a resposta  $y_k = hx_k + n_k$ , em que  $n_k$  é um componente de ruído. Para estimarmos o canal baseado na observação de  $y$ , enviamos no instante  $t = 1$  um valor conhecido  $x_1 = -1$ , e no instante  $t = 2$ , enviamos  $x_2 = 2$ . Suponhamos que conhecemos a PDF do canal, que é uma variável aleatória uniforme em  $[0,1]$ . Sabemos ainda que no instante  $t = 1$ , o ruído é uma v.a. Gaussiana  $n_1 \rightarrow \mathcal{N}(0;0,2)$ , e no instante  $t = 2$ , temos que  $n_2 \rightarrow \mathcal{N}(0;0,1)$ . Qual

o estimador linear óptimo do valor do canal  $h$ , baseado nas observações  $y_1$  e  $y_2$ ?

Inicialmente, vamos obter os momentos da variável do canal  $h$

$$\bar{h} = \frac{1}{2}; \quad \sigma_h^2 = \frac{1}{12},$$

e, conseqüentemente,  $\bar{y}_1 = -\bar{h} + \bar{n}_1 = -\frac{1}{2}$ , e  $\bar{y}_2 = 2\bar{h} + \bar{n}_2 = 1$ . Para simplificar, trabalharemos com as variáveis  $h' = h - \bar{h}$ ,  $y'_1 = y_1 - \bar{y}_1$ , e  $y'_2 = y_2 - \bar{y}_2$ . Calculamos então as correlações, lembrando que o canal  $h$  e o ruído  $n$  são independentes, assim como são independentes amostras do ruído em instantes diferentes.

$$R_{y_1 y_1} = E\{(y_1 - \bar{y}_1)^2\} = \sigma_{y_1}^2 = \sigma_h^2 + \sigma_{n_1}^2 = \frac{1}{12} + \frac{2}{10} = \frac{17}{60}$$

$$R_{y_2 y_2} = E\{(y_2 - \bar{y}_2)^2\} = \sigma_{y_2}^2 = 4\sigma_h^2 + \sigma_{n_2}^2 = \frac{4}{12} + \frac{1}{10} = \frac{26}{60}$$

$$R_{y_1 y_2} = E\{(y_1 - \bar{y}_1)(y_2 - \bar{y}_2)\} = -2\sigma_h^2$$

$$R_{h y_1} = E\{(h - \bar{h})(y_1 - \bar{y}_1)\} = E\{(h - \bar{h})[-(h - \bar{h}) + (n_1 - \bar{n}_1)]\} = -\sigma_h^2 = -\frac{1}{12}$$

$$R_{h y_2} = E\{(h - \bar{h})(y_2 - \bar{y}_2)\} = E\{(h - \bar{h})[2(h - \bar{h}) + (n_2 - \bar{n}_2)]\} = 2\sigma_h^2 = \frac{1}{6}$$

Com estes valores podemos achar os coeficientes

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 17/60 & -1/6 \\ -1/6 & 26/60 \end{bmatrix}^{-1} \begin{bmatrix} -1/12 \\ 1/6 \end{bmatrix} = \begin{bmatrix} -0,0877 \\ 0,3509 \end{bmatrix}.$$

Ou seja, podemos obter uma estimativa do canal

$$\hat{h} = -0,0877y_1 + 0,3509y_2$$

Exemplo 43

Sabemos que na digitalização usando um codificador DPCM (*Differential Pulse Code Modulation*), em vez de quantizarmos a amostra da mensagem  $m_k$ , como no PCM, quantizamos a diferença  $\epsilon_k = m_k - \hat{m}_k$ , em que  $\hat{m}_k$  é uma estimativa da amostra atual  $m_k$ . Suponhamos que esta estimativa seja linear, e obtida a partir das duas amostras anteriores, ou seja,  $\hat{m}_k = f(m_{k-1}, m_{k-2}) = a_1 m_{k-1} + a_2 m_{k-2}$ .

Sabemos que, pela lei dos grandes números, as correlações podem ser obtidas medindo-se uma amostra de voz típica, de modo que, com  $N$  amostras disponíveis teremos

$$R_n = E\{m_k m_{k-n}\} \approx \frac{1}{N} [m_0 m_n + m_1 m_{1+n} + \dots + m_N m_{N+n}]$$

Suponhamos ainda que, foram feitas estas medições em amostras de voz típicas e obtemos os seguintes valores

$$R_1 = 0,825 \overline{m^2}, \text{ e } R_2 = 0,562 \overline{m^2},$$

e que sabemos que as amostras de voz têm média  $\bar{m} = 0$ . Podemos então obter os coeficientes  $a_1$  e  $a_2$  ótimos pelo critério LMS, lembrando que

$$\begin{aligned} R_{m_k m_k} &= R_{m_{k+1} m_{k+1}} = R_{m_{k+2} m_{k+2}} = \overline{m^2}, \\ R_{m_k m_{k+1}} &= R_{m_{k+1} m_{k+2}} = R_{m_{k+2} m_{k+1}} = R_1 = 0,825 \overline{m^2}, \\ R_{m_k m_{k+2}} &= R_2 = 0,562 \overline{m^2}, \end{aligned}$$

e resolvendo o sistema

$$\begin{aligned} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= \begin{bmatrix} R_{m_{k+1} m_{k+1}} & R_{m_{k+1} m_{k+2}} \\ R_{m_{k+2} m_{k+1}} & R_{m_{k+2} m_{k+2}} \end{bmatrix}^{-1} \begin{bmatrix} R_{m_k m_{k+1}} \\ R_{m_k m_{k+2}} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0,825 \\ 0,825 & 1 \end{bmatrix}^{-1} (\overline{m^2})^{-1} \begin{bmatrix} 0,825 \\ 0,562 \end{bmatrix} \overline{m^2} = \begin{bmatrix} 1,1314 \\ -0,3714 \end{bmatrix} \end{aligned}$$

ou seja,  $\hat{m}_k = 1,1314 m_{k-1} - 0,3714 m_{k-2}$

O erro quadrático médio é dado por

$$E\{\epsilon^2\} = R_{m_k m_k} - a_1 R_{m_k m_{k+1}} - a_2 R_{m_k m_{k+2}} = 0,2753$$

Sabemos ainda que, dependendo da distribuição de probabilidade da amplitude da mensagem, temos que  $\overline{m^2} = \alpha m_p^2$ , em que  $m_p$  é a amplitude de pico da mensagem. Lembremos que, para o PCM, quantizamos a mensagem  $m_k$ , e o erro de quantização é

$$E\{q_{PCM}^2\} = \frac{1}{3} \frac{m_p^2}{2^{2n}} = \frac{1}{3} \frac{\alpha \overline{m^2}}{2^{2n}} = K \overline{m^2}$$

Para o DPCM, quantizamos a diferença  $\epsilon_k$ , em vez de  $m_k$ . Podemos supor que o mesmo vale para o erro de quantização, ou seja, que

$$E\{q_{DPCM}^2\} = K \overline{\epsilon^2}$$

Temos, portanto, uma redução no erro quadrático médio de

$$\frac{E\{q_{DPCM}^2\}}{E\{q_{PCM}^2\}} = \frac{\overline{\epsilon^2}}{\overline{m^2}} = 0,2753 = -5,6\text{dB}$$

### 3. Processos Estocásticos

Processos estocásticos<sup>3</sup> ou aleatórios são sinais aleatórios  $x(t)$ , ou seja, o resultado de um experimento aleatório é uma forma de onda  $x(t)$ , como vimos na Figura 24, em que o resultado de um experimento  $\xi_i$  é mapeado em uma forma de onda  $x_i(t)$ . Na parte de baixo desta figura vemos ainda exemplo de diferentes formas de onda possíveis resultantes deste experimento.

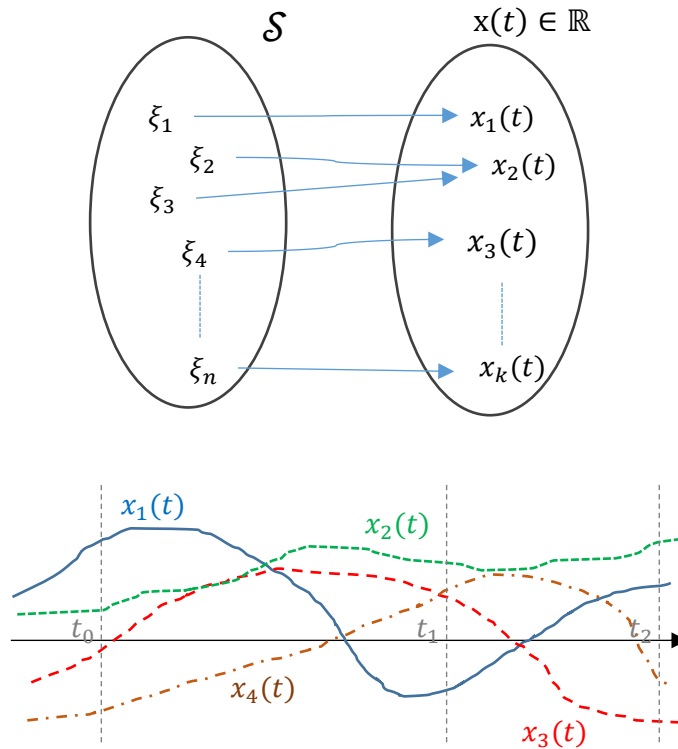


Figura 24. Conceito de processos aleatórios.

Ainda na Figura 24, podemos ver que ao amostrarmos o sinal em um dado instante  $t_i$ , o valor observado é uma variável aleatória  $x_i = x(t_i)$ , cuja distribuição de probabilidade depende das probabilidades de as diversas formas de onda ocorrerem. Para conhecermos integralmente um processo aleatório, caso amostramos o sinal em qualquer conjunto de instantes  $t_0, t_1, t_2, \dots, t_{N-1}$ , teremos um conjunto de variáveis aleatórias  $\mathbf{x} = [x_0, x_1, x_2, \dots, x_{N-1}]^T$ , e devemos conhecer as suas CDF e PDF conjuntas

$$F_{\mathbf{x}}(x_0, x_1, \dots, x_{N-1}) = \Pr(x_0 \leq x_0; x_1 \leq x_1; \dots; x_{N-1} \leq x_{N-1}) \quad (131)$$

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{\partial^n F_{\mathbf{x}}(\mathbf{x})}{\partial x_0 \partial x_1 \dots \partial x_{N-1}}$$

Na maioria dos casos, conhecer esta distribuição para todas as combinações possíveis é bastante difícil, além de ser difícil de se tratar. Veremos a seguir algumas caracterizações que descrevem parcialmente um certo processo, e que são suficientes na maioria dos casos.

<sup>3</sup> A palavra “estocástico”, vem da palavra grega  $\sigma\tau\omicron\chi\omicron\zeta$  (stochos) que significa “alvo, objetivo”

Antes, é importante definirmos o conceito de **autocorrelação** de um processo<sup>4</sup> em dois instantes  $t_1$  e  $t_2$

$$\begin{aligned} R_x(t_1, t_2) &= E\{x^*(t_1)x(t_2)\} = E\{x_1^*x_2\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^*x_2 p_{x_1x_2}(x_1^*, x_2) dx_1 dx_2, \end{aligned} \quad (132)$$

ou seja, é a correlação entre as variáveis aleatórias representando o sinal amostrado nos instantes  $t_1$  e  $t_2$ .

### a. Processos Estacionários

Uma categoria importante de processos estocásticos são os chamados processos estacionários. Como veremos, processos estacionários representam uma grande parte dos processos aleatórios considerados em comunicações.

Um processo é chamado de **estacionário** se, dados dois vetores de variáveis aleatórias

$$\begin{aligned} \mathbf{x} &= [x_0, x_1, \dots, x_{N-1}]^T, & x_i &= x(t_i) = x(t_0 + \tau_i) \\ \mathbf{x}' &= [x'_0, x'_1, \dots, x'_{N-1}]^T, & x'_i &= x(t_i + T_0) = x(t_0 + \tau_i + T_0), \end{aligned} \quad (133)$$

ambos têm a mesma densidade de probabilidade conjunta, ou seja, se

$$p_{\mathbf{x}}(x_0, x_1, \dots, x_{N-1}) = p_{\mathbf{x}'}(x_0, x_1, \dots, x_{N-1}). \quad (134)$$

O vetor  $\mathbf{x}'$  representa amostras deslocadas de  $T_0$  em relação às amostras obtidas em  $\mathbf{x}$ , e esta condição indica que a estatística destas amostras não depende de quando as amostras ocorrem, mas apenas do intervalo entre as amostras  $\tau_i$ . Ou seja, as propriedades estatísticas do sinal não mudam com o tempo.

#### Exemplo 44

A temperatura em um dado local não é um processo estacionário, já que a sua distribuição de probabilidade depende de diversas condições, como hora do dia (de noite é em média mais frio) ou estação do ano.

Estritamente falando, não existe processo estacionário, já que todo sinal existente tem um começo e um fim, e, portanto, as suas propriedades estatísticas mudam nestes pontos. Apesar disso, alguns processos com durações muito longas podem ser considerados, pelo menos no intervalo de observação de interesse.

Supondo que  $x(t)$  seja um processo estacionário, podemos mostrar que, como a densidade de probabilidade independe do instante de amostragem,  $p_{x(t)}(x) = p_{x(t+T_0)}(x)$

$$E\{x(t)\} = \int_{-\infty}^{\infty} x p_{x(t)}(x) dx = \int_{-\infty}^{\infty} x p_{x(t+T_0)}(x) dx = E\{x(t + T_0)\}, \quad (135)$$

ou seja, a média é igual em todos os instantes.

Da mesma forma, pela definição de estacionariedade em (134), podemos achar a autocorrelação

$$\begin{aligned} R_x(t_1, t_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^* x_2 p_{x(t_1)x(t_2)}(x_1^*, x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^* x_2 p_{x(t_1+T_0)x(t_2+T_0)}(x_1^*, x_2) dx_1 dx_2 \\ &= R_x(t_1 + T_0, t_2 + T_0) \end{aligned} \quad (136)$$

<sup>4</sup> Para generalizar, consideramos variáveis aleatórias complexas, e  $x^*$  é o complexo conjugado de  $x$ .

ou seja, a autocorrelação depende apenas da diferença entre os instantes de amostragem  $\tau = t_2 - t_1 = (t_2 + T_0) - (t_1 + T_0)$ , e, para processos estacionários, escrevemos

$$\begin{aligned} R_x(\tau) &= R_x(t_2 - t_1) \triangleq R_x(t_1, t_2) \\ &= E\{x^*(t)x(t + \tau)\} \end{aligned} \quad (137)$$

Uma propriedade da autocorrelação é que

$$\begin{aligned} R_x(-\tau) &= E\{x^*(t)x(t - \tau)\} = E\{x^*(t' + \tau)x(t')\} \\ &= R_x^*(\tau) \end{aligned} \quad (138)$$

### Processos Estacionários no Sentido Amplo

A condição de estacionariedade em (134) é normalmente bastante restritiva e difícil de se demonstrar, e muitas vezes desnecessária. Usualmente, estamos interessados em uma classe de processos chamados de **estacionários em sentido amplo** (WSS – *Wide Sense Stationary*). Estes são processos que satisfazem as seguintes condições:

$$\begin{aligned} E\{x(t)\} &= C \\ R_x(t_1, t_1 + \tau) &= R_x(\tau). \end{aligned} \quad (139)$$

Já havíamos visto que todos processos estacionários satisfazem a estas condições, mas de um processo estacionário em sentido amplo nada mais é exigido, ou seja, todo processo estritamente estacionário é também estacionário no sentido amplo, mas o inverso nem sempre é verdade.

#### Exemplo 45

Vamos verificar se o processo  $x(t) = A \cos(2\pi f_c t + \Theta)$ , com  $\Theta$  uma variável aleatória uniforme entre 0 e  $2\pi$ , é um processo estacionário no sentido amplo.

Primeiro vamos verificar a primeira condição

$$E_{\Theta}\{x(t)\} = \int_0^{2\pi} \frac{1}{2\pi} A \cos(2\pi f_c t + \theta) d\theta = 0,$$

em que a integral acima pode ser facilmente resolvida verificando-se que temos a integral em um período de uma senoide em  $\theta$ , conseqüentemente com valor nulo.

Verificando a segunda condição,

$$\begin{aligned} R_x(t_1, t_2) &= E_{\Theta}\{A^2 \cos(2\pi f_c t_1 + \theta) \cos(2\pi f_c t_2 + \theta)\} \\ &= \int_0^{2\pi} \frac{1}{2\pi} \frac{A^2}{2} \cos(2\pi f_c (t_1 + t_2 + 2\theta)) d\theta \\ &\quad + \int_0^{2\pi} \int_0^{2\pi} \frac{1}{2\pi} \frac{A^2}{2} \cos(2\pi f_c (t_2 - t_1)) d\theta \\ &= \frac{A^2}{2} \cos(2\pi f_c (t_2 - t_1)) = \frac{A^2}{2} \cos(2\pi f_c \tau) = R_x(\tau) \end{aligned}$$

Vemos, portanto, que as duas condições são satisfeitas e o processo é estacionário no sentido amplo.

### Ergodicidade

Inicialmente, suponhamos que ao observarmos uma realização do processo aleatório, obtemos a forma de onda  $x(t)$ . A partir desta observação, podemos obter a média temporal  $f(\overline{x(t)})$  de alguma função desta forma de onda

$$\overline{f(x(t))} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} f(x(t)) dt \quad (140)$$

Um processo aleatório é dito **ergódico** em relação a uma certa estatística  $E\{f(x(t))\}$  se esta média estatística for igual à média temporal<sup>5</sup> definida acima, para uma realização qualquer deste processo.

Dado o processo aleatório usado no exemplo anterior,  $x(t) = A \cos(2\pi f_c t + \Theta)$ , com  $\Theta$  uma variável aleatória uniforme entre 0 e  $2\pi$ , vamos verificar se este processo é ergódico em relação à média e à autocorrelação.

Exemplo 46

Vamos verificar a média temporal da média,

$$\overline{x(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A \cos(2\pi f_c t + \theta) dt = 0 = E\{x(t)\},$$

onde o resultado acima é igual a zero pois a integral é limitada, enquanto o denominador  $T$  tende a infinito. Portanto, o processo é ergódico em relação à média.

Vamos verificar a média temporal da autocorrelação,

$$\begin{aligned} \overline{R_x(\tau)} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} A^2 \cos(2\pi f_c t_1 + \theta) \cos(2\pi f_c t_2 + \theta) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{A^2}{2} \cos(2\pi f_c (2t + \tau + 2\theta)) dt + \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{A^2}{2} \cos(2\pi f_c \tau) dt \\ &= \frac{A^2}{2} \cos(2\pi f_c \tau) = R_x(\tau) \end{aligned}$$

Portanto, este processo também é ergódico em relação à autocorrelação.

Processos ergódicos são importantíssimos, pois, considerando esta ergodicidade, podemos estimar a média e a variância estatística a partir de uma realização qualquer deste experimento. Obviamente, é impossível medirmos qualquer processo durante um tempo infinito, mas se o intervalo de observação for grande o suficiente, as médias temporais se aproximam bastante dos valores esperados estatísticos.

## b. Densidade Espectral de Potência

Vamos inicialmente relembrar o conceito de **densidade espectral de energia** de um sinal determinístico  $y(t)$  com energia finita e transformada de Fourier  $Y(f)$ ,

$$\Psi_y(f) = |Y(f)|^2. \quad (141)$$

, e, sabemos que, pelo teorema de Parseval, a energia pode ser obtida tanto no domínio do tempo como no domínio da frequência

<sup>5</sup> Em inglês, há uma diferença de nomenclatura entre a média estatística (*mean*) e a média temporal ou amostral (*average*).

$$E_y = \int_{-\infty}^{\infty} |y(t)|^2 dt = \int_{-\infty}^{\infty} \Psi_y(f) df \quad (142)$$

e a energia em uma certa faixa de frequência entre  $f_1$  e  $f_2$  é dada por

$$E_{y,[f_1, f_2]} = \int_{-f_2}^{-f_1} \Psi_y(f) df + \int_{f_1}^{f_2} \Psi_y(f) df. \quad (143)$$

Para sinais de potência, ou seja, sinais com energia infinita, usamos a **densidade espectral de potência** (PSD – *power spectral density*). Sinais de potência têm usualmente duração infinita, e a fim de termos uma definição de energia, consideramos um sinal truncado  $x_T(t) = x(t) \text{rect}\left(\frac{t}{T}\right)$  com duração  $T$ . A potência é então a razão entre a energia e o tempo, e a densidade espectral de potência é a razão entre a densidade espectral de potência e o tempo, no limite,

$$S_x(f) = \lim_{T \rightarrow \infty} \frac{\Psi_{x_T}(f)}{T}. \quad (144)$$

Processos estocásticos estacionários no sentido amplo são uma abstração matemática, e, como comentado anteriormente, têm duração infinita, e, conseqüentemente, têm uma densidade espectral de potência. Por serem aleatórios, esta densidade espectral de potência é definida pelo valor esperado

$$S_x(f) = E \left\{ \lim_{T \rightarrow \infty} \frac{\Psi_{x_T}(f)}{T} \right\} = E \left\{ \lim_{T \rightarrow \infty} \frac{|x_T(f)|^2}{T} \right\} \quad (145)$$

Assim como para a densidade espectral de energia, a potência em uma certa faixa de frequência entre  $f_1$  e  $f_2$  é dada por

$$P_{x,[f_1, f_2]} = \int_{-f_2}^{-f_1} S_x(f) df + \int_{f_1}^{f_2} S_x(f) df. \quad (146)$$

**Para sinais reais**, o módulo da transformada de Fourier, e, conseqüentemente, a densidade espectral de potência, é uma função par, e

$$P_{x,[f_1, f_2]} = 2 \int_{f_1}^{f_2} S_x(f) df. \quad (147)$$

Pela definição em (145), é comumente difícil obter-se a PSD. Este cálculo pode ser facilitado, considerando que

$$\begin{aligned} S_x(f) &= \lim_{T \rightarrow \infty} \frac{1}{T} E\{|x_T(f)|^2\} \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \int_{-\frac{T}{2}}^{\frac{T}{2}} x^*(t) e^{j2\pi ft} dt \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t') e^{-j2\pi ft'} dt' \right\} \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} E\{x^*(t)x(t')\} e^{j2\pi f(t-t')} dt dt' \end{aligned} \quad (148)$$

Considerando  $t' = t + \tau$ , e lembrando o conceito de autocorrelação, temos que

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j2\pi f\tau} d\tau = \mathcal{F}\{R_x(\tau)\}. \quad (149)$$

Este resultado é conhecido como teorema de Wiener-Khinchin, e mostra que podemos calcular a densidade espectral de potência de um sinal a partir de sua autocorrelação.

Da mesma forma, podemos obter a autocorrelação como



$$R_x(\tau) = \mathcal{F}^{-1}\{S_x(f)\} = \int_{-\infty}^{\infty} S_x(f) e^{j2\pi f\tau} df \quad (150)$$

Para  $\tau = 0$ , temos que

$$R_x(0) = E\{|x|^2\} = \int_{-\infty}^{\infty} S_x(f) df = P_x, \quad (151)$$

em que  $P_x$  é a potência do sinal.

Ou seja, a potência de um sinal aleatório é dada por  $E\{|x|^2\}$ .

Dado novamente o processo aleatório  $x(t) = A \cos(2\pi f_c t + \Theta)$ , com  $\Theta$  uma variável aleatória uniforme entre 0 e  $2\pi$ , a autocorrelação já foi encontrada e

Exemplo 47

$$R_x(\tau) = \frac{A^2}{2} \cos(2\pi f_c \tau)$$

e, conseqüentemente,

$$S_x(f) = \mathcal{F}\{R_x(\tau)\} = \frac{A^2}{4} [\delta(f + f_c) + \delta(f - f_c)].$$

A potência pode ser achada como

$$\begin{aligned} P_x &= \int_{-\infty}^{\infty} S_x(f) df = \frac{A^2}{4} \int_{-\infty}^{\infty} [\delta(f + f_c) + \delta(f - f_c)] df = \frac{A^2}{2} \\ &= R_x(0) = \frac{A^2}{2} \cos(0) = \frac{A^2}{2} \end{aligned}$$

Exemplo 48

#### Sinal AM

Suponha um sinal aleatório  $m(t)$  modulado em amplitude, tal que  $\varphi(t) = m(t) \cos(2\pi f_c t + \Theta)$ , com  $\Theta$  uma variável aleatória uniforme entre 0 e  $2\pi$ . A autocorrelação de  $\varphi(t)$  é

$$R_\varphi(\tau) = E\{m(t) \cos(2\pi f_c t + \Theta) m(t + \tau) \cos(2\pi f_c (t + \tau) + \Theta)\}.$$

Considerando que  $m(t)$  e  $\Theta$  são independentes, temos que

$$\begin{aligned} R_\varphi(\tau) &= E\{m(t)m(t + \tau)\} E\left\{\frac{1}{2} [\cos(2\pi f_c (2t + \tau) + 2\Theta) + \cos(2\pi f_c \tau)]\right\} \\ &= \frac{1}{2} R_m(\tau) \cos(2\pi f_c \tau). \end{aligned}$$

Lembrando que uma multiplicação no tempo é uma convolução na frequência, temos que a PSD é dada por

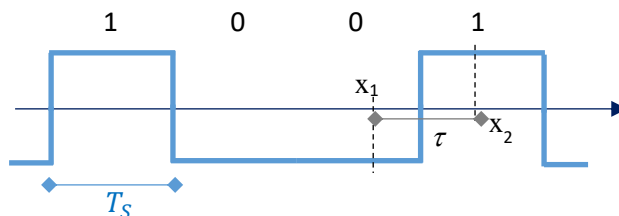
$$\begin{aligned} S_\varphi(f) &= \frac{1}{2} S_m(f) * \left[ \frac{1}{2} \delta(f - f_c) + \frac{1}{2} \delta(f + f_c) \right] \\ &= \frac{1}{4} [S_m(f - f_c) + S_m(f + f_c)], \end{aligned}$$

ou seja, o espectro da mensagem é o espectro da mensagem deslocado para  $f_c$ , e a potência do sinal modulado

$$P_\varphi = R_\varphi(0) = \frac{1}{2}R_m(0) = \frac{1}{2}P_m$$

$$= \int_{-\infty}^{\infty} S_\varphi(f) df = \frac{1}{2} \int_{-\infty}^{\infty} S_m(f) df = \frac{1}{2}P_m$$

Exemplo 4.9 Consideramos um esquema de codificação digital binário em que bits 0 são representados por um nível  $x(t) = -1$  e bits 1 por um nível  $x(t) = +1$ . Cada bit tem uma duração  $T_s$ , e a transmissão começa em um instante aleatório  $\alpha$ .



O sinal transmitido pode ser escrito como

$$x(t) = \sum_n a_n \text{rect}\left(\frac{t - nT_s - \alpha}{T_s}\right), \quad a_n = \pm 1$$

Considerando  $x_1 = x(t_1)$ , e  $x_2 = x(t_2) = x(t_1 + \tau)$ , e sabendo que  $x_1, x_2 \in \{-1, +1\}$ , a autocorrelação é dada por

$$R_x(\tau) = E\{x_1 x_2\} = \sum_{x_1 = \pm 1} \sum_{x_2 = \pm 1} x_1 x_2 P_{x_1, x_2}(x_1, x_2)$$

$$= P_{x_1, x_2}(-1, -1) - P_{x_1, x_2}(-1, +1) - P_{x_1, x_2}(+1, -1) + P_{x_1, x_2}(+1, +1)$$

Pela simetria do sistema, podemos considerar que  $P_{x_1, x_2}(-1, -1) = P_{x_1, x_2}(+1, +1)$ , e  $P_{x_1, x_2}(-1, +1) = P_{x_1, x_2}(+1, -1)$ , e, conseqüentemente

$$R_x(\tau) = 2[P_{x_1, x_2}(+1, +1) - P_{x_1, x_2}(-1, +1)]$$

$$= 2P_{x_1}(1)[P_{x_2|x_1}(+1|+1) - P_{x_2|x_1}(-1|+1)]$$

Considerando os bits equiprováveis, ou seja,  $P_{x_1}(1) = \frac{1}{2}$ , e, lembrando que  $P_{x_2|x_1}(+1|+1) = 1 - P_{x_2|x_1}(-1|+1)$ , temos que

$$R_x(\tau) = 1 - 2P_{x_2|x_1}(-1|+1)$$

Precisamos agora encontrar  $P_{x_2|x_1}(-1|+1)$ , que é a probabilidade de que haja uma mudança de nível de +1 para -1 entre os instantes  $t_1$  e  $t_1 + \tau$ , dado que o nível no instante inicial nível era +1.

É fácil verificar que  $E\{x_i\} = 0$ . Se  $\tau > T_s$ , estaremos amostrando em instantes pertencentes a bits diferentes, e, supondo que eles são independentes,

$$R_x(\tau) = E\{x_1 x_2\} = E\{x_1\}E\{x_2\} = 0.$$

Já se  $0 < \tau \leq T_s$ , só ocorrerá mudança de nível se houver transição de bit neste intervalo, e, se nesta transição o bit passar de +1 para -1. A probabilidade do nível passar de +1 para -1 é  $\Pr(b_{k+1} = 0 | b_k = 1) = \frac{1}{2}$ . Já a probabilidade de haver uma transição neste intervalo é maior quanto for maior o intervalo, e, chamando de  $I_k$  o intervalo do bit  $k$ , temos que  $\Pr(t_1 + \tau \in I_{k+1} | t_1 \in I_k) = \frac{\tau}{T_s}$ . Portanto,

$$P_{x_2|x_1}(-1|+1) = \Pr(t_1 + \tau \in I_{k+1} | t_1 \in I_k) \Pr(b_{k+1} = 0 | b_k = 1) = \frac{\tau}{2T_s}$$

e, lembrando que, como se trata de um sinal real  $R_x(\tau) = R_x(-\tau)$ , temos consequentemente,

$$\begin{aligned} R_x(\tau) &= 1 - \frac{|\tau|}{T_s}, \quad |\tau| \leq T_s \\ &= \Delta\left(\frac{t}{2T_s}\right). \end{aligned}$$

A densidade espectral de potência é dada então por

$$S_x(f) = \mathcal{F}\{R_x(\tau)\} = T_s \text{sinc}^2(\pi f T_s)$$

### c. Transmissão de Processos Aleatórios em Sistemas Lineares

Grande parte dos elementos em sistemas de telecomunicações consistem em sistemas lineares. Dentro desta categoria estão incluídos filtros de transmissão e recepção, assim como canais de multipercursos comuns tanto em sistemas sem fio como sistemas cabeados.

Em um sistema linear, dada uma entrada determinística  $x(t)$ , a saída é dada por

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} x(t - \tau)h(\tau)d\tau, \quad (152)$$

em que  $h(t)$  é a resposta ao impulso do sistema linear.

No domínio da frequência, a análise é mais simples, e considerando as transformadas de Fourier, temos que

$$Y(f) = X(f)H(f). \quad (153)$$

Para sinais aleatórios, temos que considerar os valores esperados, autocorrelação e densidade espectral de potência. Neste caso, supondo que temos na entrada um processo estacionário no sentido amplo  $x(t)$ , teremos na saída um processo aleatório  $y(t)$ , com média dada por

$$\begin{aligned} E\{y(t)\} &= E\{h(t) * x(t)\} = \int_{-\infty}^{\infty} h(\tau)E\{x(t - \tau)\}d\tau \\ &= \bar{x} \int_{-\infty}^{\infty} h(\tau)d\tau = \bar{x}H(0). \end{aligned} \quad (154)$$

A autocorrelação do processo na saída, considerando um processo real, é encontrada por

$$\begin{aligned}
R_y(\tau) &= E\{y(t)y(t+\tau)\} = E\left\{\int_{-\infty}^{\infty} h(\alpha)x(t-\alpha)d\alpha \int_{-\infty}^{\infty} h(\beta)x(t+\tau-\beta)d\beta\right\} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha)h(\beta)R_x(\tau+\alpha-\beta)d\beta d\alpha \\
&= \int_{-\infty}^{\infty} h(\alpha) \int_{-\infty}^{\infty} h(\beta)R_x(\tau+\alpha-\beta)d\beta d\alpha \\
&= \int_{-\infty}^{\infty} h(\alpha) [R_x(\tau+\alpha) * h(\tau+\alpha)]d\alpha = h(-\tau) * h(\tau) * R_x(\tau)
\end{aligned} \tag{155}$$

e, conseqüentemente, a densidade espectral de potência na saída é dada por

$$\begin{aligned}
S_y(f) &= \mathcal{F}\{R_y(\tau)\} = H(-f)H(f)S_x(f) \\
&= |H(f)|^2 S_x(f)
\end{aligned} \tag{156}$$

Exemplo 50 Um processo aleatório tem função de autocorrelação  $R_x(\tau) = \frac{1}{1+\tau^2}$ . Se este sinal passar por um derivador, qual a potência na saída deste derivador?

Sabemos que a densidade espectral de potência do sinal é

$$S_x(f) = \mathcal{F}\{R_x(\tau)\} = \pi e^{-2\pi|f|}$$

Um derivador é um sistema linear com  $H(f) = j2\pi f$ , e, portanto, na saída temos o sinal  $y$ , com

$$S_y(f) = |H(f)|^2 S_x(f) = 4\pi^3 f^2 e^{-2\pi|f|}$$

A potência na saída será

$$P_y = \int_{-\infty}^{\infty} S_y(f)df = 8\pi^3 \int_0^{\infty} f^2 e^{-2\pi f} df = 2$$

#### d. Ruído Branco Aditivo Gaussiano

Uma das principais fontes de distorção em sistema de comunicações é o ruído branco Gaussiano aditivo (AWGN – *additive white Gaussian noise*).

Seja  $w(t)$  um processo aleatório estacionário representando o ruído, e tenhamos enviado o sinal  $x(t)$ , o sinal recebido será

$$y(t) = x(t) + w(t), \tag{157}$$

daí o nome aditivo.

O ruído será branco se sua densidade espectral de potência for constante

$$S_w(f) = \frac{N_0}{2}, \tag{158}$$

em analogia à luz branca, que contém componentes de todas as frequências do espectro da luz visível.

A sua autocorrelação é, portanto, dada por

$$R_w(\tau) = \mathcal{F}^{-1}\{S_w(f)\} = \frac{N_0}{2} \delta(\tau), \quad (159)$$

ou seja, amostras do ruído em instantes distintos são sempre decorrelatadas, não importa quão pequeno for este intervalo.

A potência do ruído branco é dada por

$$P_w = \int_{-\infty}^{\infty} \frac{N_0}{2} df = \infty, \quad (160)$$

o que, obviamente, é uma impossibilidade física. Porém, em sistemas reais, estamos interessados no sinal contido em uma determinada banda, e podemos considerar um **ruído branco limitado em banda**, com densidade espectral de potência

$$S_n(f) = \frac{N_0}{2} \text{rect}\left(\frac{f}{2B}\right), \quad (161)$$

cuja potência é dada por

$$P_n = \int_{-B}^B \frac{N_0}{2} df = N_0 B. \quad (162)$$

Este ruído apresenta agora a autocorrelação

$$R_n(\tau) = \mathcal{F}^{-1}\{S_n(f)\} = N_0 B \text{sinc}(2\pi B\tau). \quad (163)$$

Lembrando do comportamento de um processo aleatório em sistemas lineares, o ruído branco limitado em banda equivale a aplicarmos um ruído branco a um filtro passa-baixa, como vemos na Figura 25.

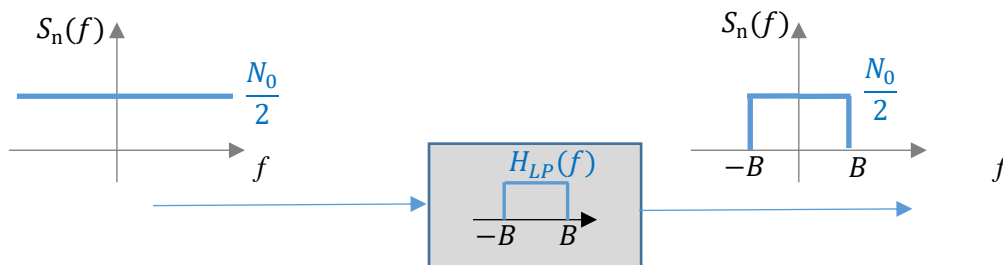


Figura 25. Ruído branco limitado em banda

O ruído branco é **Gaussiano** quando sua densidade de probabilidade em qualquer instante for uma Gaussiana com média nula, e, conseqüentemente, variância igual à potência, ou seja,

$$p_{n(t)}(x) = \frac{1}{\sqrt{2\pi BN_0}} e^{-\frac{x^2}{2BN_0}} \quad (164)$$

### Ruído Térmico

A principal fonte de ruído em qualquer circuito elétrico, e, no nosso caso, em um receptor de comunicações, é o chamado ruído térmico. Este ruído, conhecido como ruído de Johnson-Nyquist, é uma tensão aleatória que ocorre em qualquer condutor elétrico devido ao movimento aleatório dos elétrons causado pelo calor. Este ruído pode ser considerado como branco e Gaussiano, e tem uma densidade espectral de potência

$$S_w(f) = \frac{N_0}{2} = 2kTR, \quad (165)$$

em que  $R$  é a resistência do circuito, em ohms,  $T$ , a temperatura em kelvin, e  $k = 1,380649 \times 10^{-23} \text{ J/K}$  é a constante de Boltzmann.

Um resistor real pode ser representado então como possuindo, além de uma resistência, uma fonte de tensão com densidade espectral de potência  $S_w(f)$ , como visto na Figura 26.

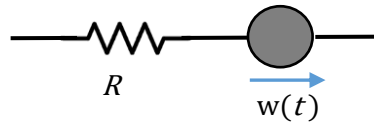


Figura 26. Circuito equivalente de um resistor real

Dado o circuito abaixo, qual o valor rms da tensão  $v_o(t)$  no capacitor, considerando uma temperatura ambiente de  $20^\circ\text{C}$ ?

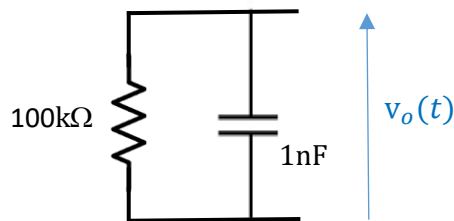


Figura 27. Circuito Resistor-Capacitor

Podemos representar este circuito com o circuito equivalente da Figura 26.

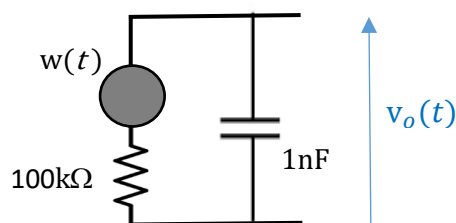


Figura 28. Circuito Resistor-Capacitor equivalente

Sabendo que a impedância de um capacitor é dada por  $1/j2\pi fC$ , a resposta na frequência deste circuito é dada por

$$H(f) = \frac{V_o(f)}{W(f)} = \frac{1/j2\pi fC}{R + 1/j2\pi fC} = \frac{1}{1 + j2\pi fRC}$$

e, portanto, a densidade espectral de potência do sinal  $v_o(t)$  é dada por

$$S_o(f) = |H(f)|^2 S_w(f) = \left| \frac{1}{1 + j2\pi fRC} \right|^2 \frac{N_0}{2} = \frac{2kTR}{1 + 4\pi^2 R^2 C^2 f^2}$$

Por definição  $v_{rms} = \sqrt{E\{v_o^2\}}$ , com

$$E\{v_o^2\} = \int_{-\infty}^{\infty} S_o(f) df = \frac{kT}{2\pi^2 RC^2} \int_{-\infty}^{\infty} \frac{1}{\left(\frac{1}{2\pi RC}\right)^2 + f^2} df$$

$$= \frac{kT}{\pi C} \operatorname{atan}(2\pi RCf) \Big|_{-\infty}^{\infty} = \frac{kT}{C}.$$

Lembrando que  $T(K) = 273,15 + T(^{\circ}C) = 293,15$ , temos que

$$v_{rms} = \sqrt{\frac{kT}{C}} = \sqrt{\frac{1,3806 \times 10^{-23} (293,15)}{10^{-9}}} = 2,01 \times 10^{-6} V$$

### Ruído Branco em Banda Passante

Grande parte dos sistemas de comunicação são sistemas em banda passante, ou seja, ocupam uma faixa do espectro. Já sabemos que qualquer sinal em banda passante em torno de  $f_c$  pode ser representado pela sua envoltória complexa. O mesmo vale para um ruído branco em banda passante  $n_p(t)$ , tal que

$$S_{n_p}(f) = \begin{cases} N_0/2 & , |f| - f_c \leq \frac{B}{2} \\ 0 & , c. c. \end{cases} \quad (166)$$

que pode ser representado como

$$\begin{aligned} n_p(t) &= n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \\ &= \operatorname{Re}\{n(t)e^{j2\pi f_c t}\} \end{aligned} \quad (167)$$

em que

$$n(t) = n_I(t) + jn_Q(t) \quad (168)$$

é a envoltória complexa do ruído.

Podemos mostrar que  $n_I(t)$  e  $n_Q(t)$  são processos aleatórios independentes com média nula, e que

$$\begin{aligned} S_{n_I}(f) &= S_{n_Q}(f) = N_0, & |f| \leq B/2 \\ S_n(f) &= 2N_0, & |f| \leq B/2 \end{aligned} \quad (169)$$

O espectro do ruído em banda passante pode ser visto na Figura 29, e a dedução destas expressões pode ser encontrada no Apêndice 7.a.

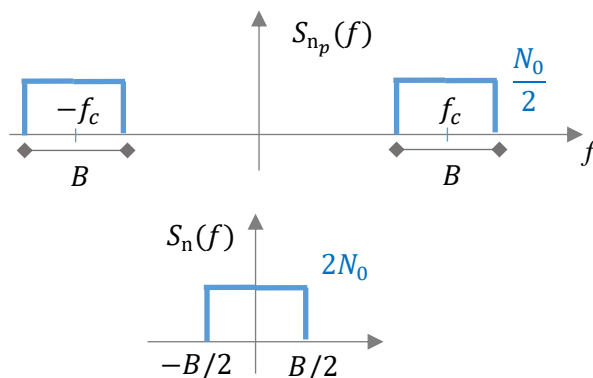


Figura 29. Ruído branco em banda passante

### e. Soma de Processos Aleatórios

Suponha agora que tenhamos o processo estocástico  $z(t) = x(t) + y(t)$ , em que  $x(t)$  e  $y(t)$  são processos estacionários reais no sentido amplo. A autocorrelação de  $z(t)$  é dada por

$$R_z(\tau) = E\{(x(t) + y(t))(x(t + \tau) + y(t + \tau))\} \quad (170)$$

$$= R_x(\tau) + R_{xy}(\tau) + R_{yx}(\tau) + R_y(\tau)$$

em que

$$R_{xy}(\tau) = E\{x(t)y(t + \tau)\} \quad (171)$$

é a função de **correlação cruzada** entre  $x(t)$  e  $y(t)$ .

Os processos  $x(t)$  e  $y(t)$  são ditos **descorrelatados** se

$$R_{xy}(\tau) = E\{x(t)\}E\{y(t + \tau)\} = \bar{x}\bar{y}. \quad (172)$$

A autocorrelação da soma de processos descorrelatados é portanto

$$R_z(\tau) = R_x(\tau) + R_y(\tau) + 2\bar{x}\bar{y} \quad (173)$$

Os processos  $x(t)$  e  $y(t)$  são ditos **ortogonais** ou **incoerentes** se  $R_{xy}(\tau) = 0$ .

A autocorrelação da soma de processos ortogonais é portanto

$$R_z(\tau) = R_x(\tau) + R_y(\tau) \quad (174)$$

Consequentemente, lembrando que a transformada de Fourier é linear,

$$S_z(f) = S_x(f) + S_y(f) \quad (175)$$

#### Exemplo 52

Dado o circuito da Figura 30, em que a fonte  $x_1(t)$  gera um sinal com densidade espectral de potência  $S_1(f) = 6 \text{ V}^2/\text{Hz}$ , e a fonte  $x_2(t)$  gera um sinal com  $S_2(f) = \frac{2}{1+(2\pi f)^2} \text{ V}^2/\text{Hz}$ , ambos com média nula.

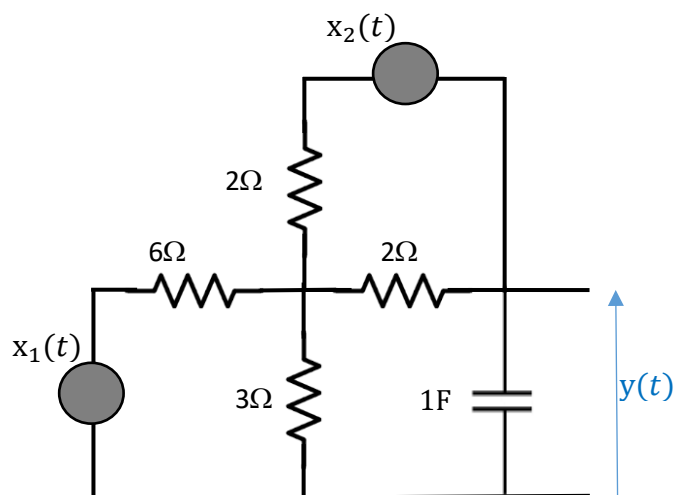


Figura 30. Exemplo de circuito com duas fontes

Pelo teorema da superposição sabemos que a tensão na saída é a soma dos sinais devidos às duas fontes independentes, ou seja



$$y(t) = y_1(t) + y_2(t),$$

com

$$Y_1(f) = X_1(f)H_1(f), \quad Y_2(f) = X_2(f)H_2(f).$$

Sabemos que a resposta à entrada  $x_1(t)$  é obtida com a fonte  $x_2(t)$  em curto, e vice-versa. Teremos então os seguintes circuitos na Figura 31.

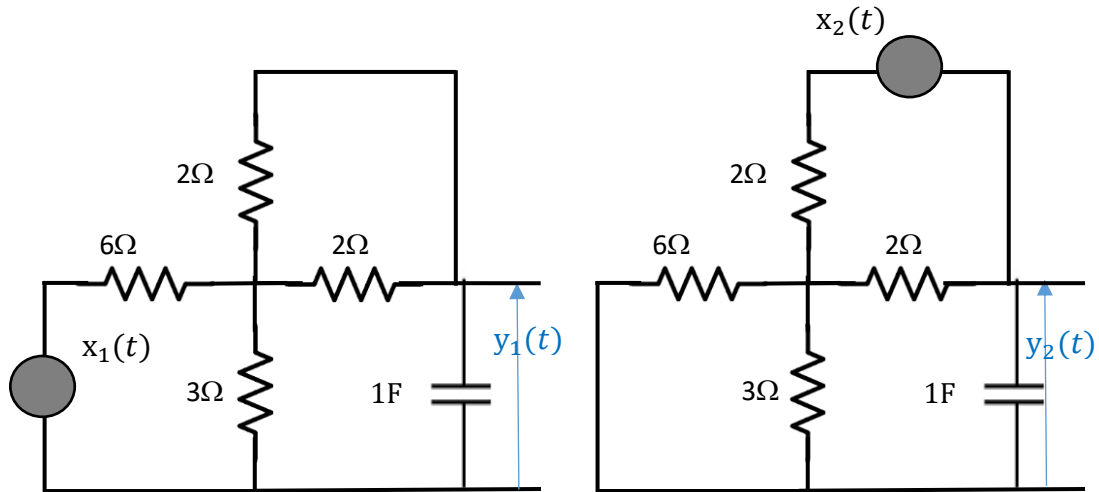


Figura 31. Circuitos separados pelo princípio da superposição

As funções de transferência podem ser obtidas como

$$H_1(f) = \frac{Y_1(f)}{X_1(f)} = \frac{1}{9j(2\pi f) + 3}$$

$$H_2(f) = \frac{Y_2(f)}{X_2(f)} = \frac{1}{6j(2\pi f) + 2}$$

Supondo que os dois sinais  $x_1(t)$  e  $x_2(t)$  são independentes, eles são, conseqüentemente decorrelatados. Supondo ainda que  $\bar{x} = \bar{y} = 0$ , eles são ortogonais e, assim

$$\begin{aligned} S_y(f) &= S_{y_1}(f) + S_{y_2}(f) = |H_1(f)|^2 S_{x_1}(f) + |H_2(f)|^2 S_{x_2}(f) \\ &= \frac{6}{81(2\pi f)^2 + 9} + \frac{2}{(1 + (2\pi f)^2)(4 + 36(2\pi f)^2)} \end{aligned}$$

Poderíamos encontrar a potência pela integral da densidade espectral de potência, ou, neste caso, mais facilmente pela autocorrelação

$$\begin{aligned} R_y(\tau) &= \mathcal{F}^{-1}\{S_y(f)\} = R_{y_1}(\tau) + R_{y_2}(\tau) \\ &= \frac{1}{9}e^{-\frac{|\tau|}{3}} + \frac{3 - e^{-|\tau|}}{32} \end{aligned}$$

A potência é dada por

$$P_y = R_y(0) = \frac{1}{9} + \frac{1}{16} = \frac{25}{144}$$

e o valor rms é

$$y_{rms} = \sqrt{P_y} = \frac{5}{12}$$

#### f. Filtragem Ótima (Filtro de Wiener)

Uma das funções de filtros é eliminarmos os componentes de ruído fora da banda de interesse do sinal, ou seja, frequências que não contém informação útil podem ser desconsideradas. Suponha que queiramos enviar a mensagem  $m(t)$ , cuja densidade espectral de potência é  $S_m(f)$  em um canal com ruído aditivo  $n(t)$ , com densidade espectral de potência  $S_n(f)$ . Suponha ainda que o sinal passe por um canal com resposta impulsional  $h(t)$ . Na recepção teremos, portanto, o sinal

$$y(t) = m(t) * h(t) + n(t) \quad (176)$$

, e queremos filtrá-lo por um filtro  $G(f)$  de modo que tenhamos a melhor estimativa possível  $\hat{m}(t)$  da mensagem original. Teremos na saída do filtro a estimativa

$$\begin{aligned} \hat{m}(t) &= y(t) * g(t) = m(t) * h(t) * g(t) + n(t) * g(t) \\ &= m(t) + \epsilon(t) = m(t) + m_d(t) + n_o(t) \end{aligned} \quad (177)$$

em que  $\epsilon(t)$  é o erro total,

$$m_d(t) = m(t) * h(t) * g(t) - m(t) = m(t)(h(t) * g(t) - 1) \quad (178)$$

é a distorção residual após o filtro no sinal, com densidade espectral de potência

$$S_d(f) = S_m(f) |H(f)G(f) - 1|^2. \quad (179)$$

e  $n_o(t)$  é o ruído filtrado com

$$S_{n_o}(f) = |G(f)|^2 S_n(f) \quad (180)$$

Se não tivermos ruído, a solução óbvia é fazermos  $G(f) = 1/H(f)$ , e assim teremos na saída, no domínio da frequência  $\hat{M}(f) = M(f)H(f)G(f) = M(f)$ , ou seja, exatamente o sinal de entrada.

Porém, considerando também a presença de ruído, devemos minimizar a potência (ou o erro quadrático médio) da distorção total. Devemos então escolher o filtro  $g(t)$  que minimize o erro quadrático médio

$$\begin{aligned} MSE &= E\{|\epsilon(t)|^2\} = E\{(m_d(t) + n_o(t))^2\} \\ &= E\{|m_d(t)|^2\} + E\{|n_o(t)|^2\} \end{aligned} \quad (181)$$

onde a última linha vem do fato de que o sinal e o ruído são considerados independentes.

Podemos também analisar no domínio da frequência,

$$\begin{aligned} P_\epsilon &= \int_{-\infty}^{\infty} S_\epsilon(f) df = \int_{-\infty}^{\infty} (S_d(f) + S_{n_o}(f)) df \\ &= \int_{-\infty}^{\infty} (S_m(f) |H(f)G(f) - 1|^2 + |G(f)|^2 S_n(f)) df. \end{aligned} \quad (182)$$

Para minimizar  $P_\epsilon$  podemos minimizar  $S_\epsilon(f)$ . Para isso podemos fazer a derivada de Wirtinger em relação a  $G(f)$  e torna-la igual a zero

$$\begin{aligned} \frac{\partial S_\epsilon(f)}{\partial G(f)} &= H(f)(G(f)H(f) - 1)^* S_m(f) + G^*(f) S_n(f) \\ &= (|H(f)|^2 S_m(f) + S_n(f)) G^*(f) - H(f) S_m(f) = 0 \end{aligned} \quad (183)$$

e, conseqüentemente, lembrando que  $S_n(f)$  e  $S_m(f)$  são reais,

$$G(f) = \frac{H^*(f)S_m(f)}{|H(f)|^2 S_m(f) + S_n(f)}, \quad (184)$$

que é conhecido com filtro de deconvolução de Wiener.

Podemos ver que, se não há ruído, ou seja,  $S_n(f) = 0$ , o filtro óptimo será  $G(f) = 1/H(f)$ , como intuído anteriormente.

Se for usado o filtro óptimo, e substituindo (184) em (183) podemos ver que a potência da distorção é dada

$$\begin{aligned} P_n &= \int_{-\infty}^{\infty} \frac{|H(f)|^2 S_m^2(f) S_n(f) + S_m(f) S_n^2(f)}{(|H(f)|^2 S_m(f) + S_n(f))^2} df \\ &= \int_{-\infty}^{\infty} \frac{S_m(f) S_n(f)}{|H(f)|^2 S_m(f) + S_n(f)} df. \end{aligned} \quad (185)$$

Exemplo 53 Um sinal com densidade espectral de potência  $S_m(f) = \frac{2}{1+(2\pi f)^2}$  passa por um canal com ruído branco aditivo, com densidade espectral de potência  $S_n(f) = \frac{N_0}{2}$ . Ache o filtro de recepção óptimo.

Podemos considerar  $H(f) = 1$  em (184), e assim

$$G(f) = \frac{S_m(f)}{S_m(f) + S_n(f)} = \frac{2}{1 + (2\pi f)^2} \frac{1}{\frac{2}{1 + (2\pi f)^2} + \frac{N_0}{2}} = \frac{4}{4 + N_0 + N_0(2\pi f)^2}$$

Na Figura 32 podemos ver a densidade espectral de potência deste sinal e os filtros correspondentes. Vemos que, quanto maior for a potência do ruído, mais importante será a distorção causada pelo mesmo, e, portanto, mais estreito deverá ser o filtro a fim de se filtrá-lo.

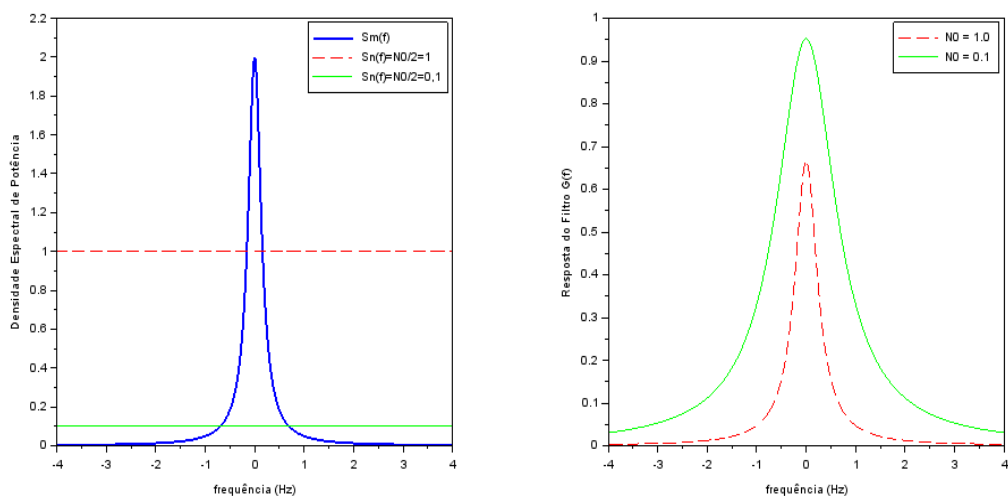


Figura 32. Exemplo de filtro de Wiener

No domínio do tempo, o filtro pode ser escrito como

$$g(t) = \mathcal{F}^{-1}\{G(f)\} = \frac{2}{N_0 B} e^{-\frac{4+N_0}{N_0}|t|}$$

que, como  $g(t) \neq 0$  para  $t < 0$ , é um filtro não causal e, portanto, não realizável. Porém ele pode ser implementado aproximadamente por meio de um truncamento e atraso.

## 4. Exercícios

### Exercício 1. ([1] Ex. 8.1-1)

Uma rede de comunicações tem 26 nós. Cada nó falha independentemente com probabilidade  $p = 0,05$ . Qual a probabilidade de que não mais que 5 nós falhem?

### Exercício 2. ([1] Ex. 8.1-2)

Uma senha bancária tem 4 dígitos de 0 a 9. Qual a probabilidade de que a soma dos dígitos seja igual a 7?

### Exercício 3. ([1] Ex. 8.1-6 - modificado)

Em um pacote de 2 bytes, cada bit tem a probabilidade de erro de 0,04, independentemente dos outros bits. Qual a probabilidade de ocorrerem 4 erros de bit ou mais?

### Exercício 4. ([1] Ex. 8.1-8, 8.1-9)

Um sistema de transmissão consiste em 10 enlaces concatenados serialmente. Se um enlace falhar o sistema inteiro falha.

- Sabendo que todos os enlaces são independentes e que a probabilidade de cada um estar funcionando é de 98%, qual a probabilidade de falha do sistema?
- Se queremos que o sistema funcione com 99% de probabilidade, qual deve ser a confiabilidade, ou seja, a probabilidade de funcionamento de cada enlace?
- Repita o item (a), considerando que o primeiro enlace tenha uma probabilidade de falha de 3%, e todos os outros tenham probabilidade de falha de 1%.
- A fim de se aumentar a confiabilidade, é instalado um outro sistema de transmissão redundante em paralelo com o primeiro, de modo que, se um falhar, a transmissão ainda possa ser realizada. Usando os dados do item (a), qual a confiabilidade do sistema agora?

### Exercício 5. ([1] Ex. 8.1-19)

Em uma rede com 12 enlaces em cascata, um dos enlaces falhou. Um técnico deve testar os enlaces um a um, a fim de verificar qual deles está defeituoso.

- Qual a probabilidade de ele encontrar o enlace defeituoso no primeiro teste?
- Qual a probabilidade de ele encontrar o enlace defeituoso em até 5 testes?
- Qual o valor esperado do número de testes a ser realizado?

### Exercício 6. ([1] Ex. 8.2-1, modificado)

É dado um canal binário assimétrico, tal que  $P_{y|x}(0|1) = 0,1$  e  $P_{y|x}(1|0) = 0,2$ , em que  $x$  é o bit transmitido e  $y$  o recebido. Suponha que  $P_x(0) = 0,4$ .

- Determine  $P_y(0)$  e  $P_y(1)$ .
- Qual a probabilidade de que em uma sequência de 10 bits qualquer na entrada não ocorra nenhum bit 1 na saída?
- Qual a probabilidade de que ao menos 5 bits iguais a 0 ocorram na saída para uma sequência de 10 bits?
- Qual a probabilidade de erro deste canal?

Exercício 7. ([1] Ex. 8.2-2)

Um canal simétrico binário tem probabilidade de erro  $p$ . A probabilidade de transmissão de um bit igual a 1 é  $q$ . Se o receptor receber um bit igual 1, qual a probabilidade de que o bit enviado também foi 1?

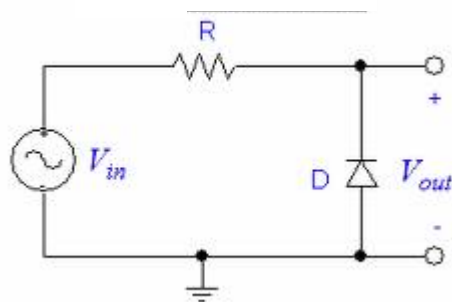
Exercício 8. ([1] Ex. 8.2-4)

A PDF de um certo sinal é dada por  $p_x(x) = C|x|e^{-K|x|}$ , com  $C$  e  $K$  constantes positivas.

- Ache  $C$  em função de  $K$ .
- Ache a probabilidade de  $-1 < x \leq 2$ .

Exercício 9. ([1] Ex. 8.2-6)

Um sinal aleatório  $V_{in}$ , Gaussiano com média nula, é aplicado na entrada de um circuito retificador de meia onda, como na figura abaixo.



Supondo um diodo ideal, determine a função de distribuição cumulativa (CDF) e a função de densidade de probabilidade da saída  $V_{out}$ .

Exercício 10. ([1] Ex. 8.2-7)

Para uma variável aleatória com PDF  $p_x(x) = \frac{1}{C\sqrt{2\pi}} e^{-x^2/32} u(x)$

- Encontre  $C$  e esboce  $p_x(x)$ .
- Determine  $P(x \geq 1)$ .
- Como poderíamos gerar esta v. a.  $x$  a partir de uma outra v.a.  $y$  Gaussiana?

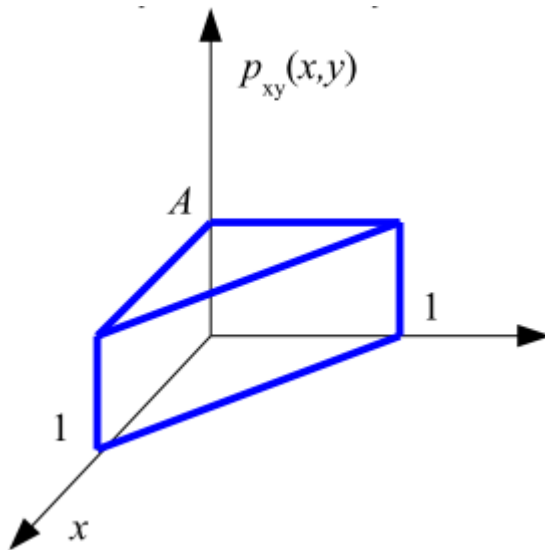
Exercício 11. ([1] Ex. 8.2-9, 8.3-5, 8.4-1)

A PDF conjunta de duas v.a.s é dada por  $p_{x,y}(x,y) = Axye^{-x^2}e^{-y^2/2}u(x)u(y)$

- Encontre  $A$ .
- encontre  $p_x(x)$  e  $p_{y|x}(y|x)$ .
- $x$  e  $y$  são independentes?
- ache a média e a variância de  $x$ .
- determine a covariância e o coeficiente de correlação entre  $x$  e  $y$ .

Exercício 12. ([1] Ex. 8.2-10, 8.3-7, 8.4-2, 8.5-1)

A PDF conjunta de duas variáveis aleatórias é dada pela figura abaixo



- determine A.
- encontre  $p_x(x)$  e  $p_y(y)$
- encontre  $p_{y|x}(y | x)$
- x e y são independentes?
- encontre a média e a variância de y
- ache o coeficiente de correlação entre x e y
- encontre o estimador LMS de y a partir de x,  $\hat{y} = f(x)$ .

Exercício 13. ([1] Ex. 8.3-2)

Determine a média e a variância da variável aleatória com PDF  $p_x(x) = Ce^{-2|x-2|}$ .

Exercício 14. ([1] Ex. 8.3-4, 8.5-4, modificado)

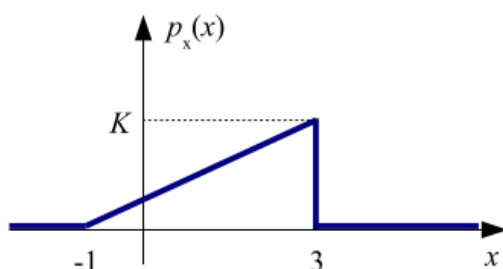
Sejam  $x_i = \pm 1$  com  $0 \leq i \leq 8$  variáveis aleatórias independentes binárias equiprováveis. Uma nova variável y é dada por

$$y = \sum_{i=1}^8 x_i.$$

- Determine a média e a variância de y.
- Encontre o estimador LMS de y a partir de  $x_1, x_2, x_3$  e  $x_4$ .
- ache o mínimo erro quadrático médio no item (b).

Exercício 15. ([1] Ex. 8.3-6)

Determine a média e a variância da variável aleatória com PDF mostrada na figura abaixo.



Exercício 16. ([1] Ex. 8.5-7)

Dados  $x = \cos \Theta$  e  $y = \sin \Theta$ , em que  $\Theta$  é uma variável aleatória uniforme entre 0 e  $2\pi$ , determine se  $x$  e  $y$  são descorrelatadas e independentes.

Exercício 17. ([1] Ex. 8.6-2)

Um sinal binário  $x$  pode assumir dois valores,  $x=0$  ou  $x=3$ , com  $P_x(0) = 0,4$ . A este sinal é adicionado um ruído  $n$  com densidade de probabilidade exponencial simétrica, como no Exercício 8, com média nula e variância 2. Determine a PDF da amplitude do sinal resultante  $y = x + n$ .

Exercício 18. ([1] Ex. 8.7-1)

Em um canal de comunicações binário são enviados pacotes de dados de 256 bytes cada. A probabilidade de erro de cada bit é de  $10^{-3}$ , encontre a probabilidade aproximada de que menos que 10% dos bits estejam em erro.

Exercício 19.

Para a transmissão de dados digitais é usual a fragmentação de uma mensagem em diversos pacotes, que, por motivos diversos, podem não chegar todos ao receptor. Este tipo de situação é tratado com um modelo de apagamento, ou seja, além de erro de bit podemos ter o caso de não termos informação nenhuma sobre este bit, caso que chamamos de apagamento (*erasure*). Chamando de  $A$  o evento de apagamento, o bit na saída do canal pode ter três valores diferentes, **0,1** ou **A**. Sendo  $b$  o bit enviado e  $r$  o valor recebido, e supondo as seguintes probabilidades:

$$\begin{aligned}Pr(b = 0) &= 0,3 \\Pr(r = A) &= 0,05 \\Pr(r = 1 \mid b = 0) &= 0,01 \\Pr(r = 0 \mid b = 1) &= 0,02\end{aligned}$$

- i. Qual a probabilidade de erro?
- ii. Sabendo que o bit recebido foi 1, qual a probabilidade de ter sido enviado também o bit 1?
- iii. Criamos agora uma variável aleatória  $y = \begin{cases} 2r - 1, & r = 0 \text{ ou } r = 1 \\ 0, & r = A \end{cases}$ . Calcule a média e a variância de  $y$ .
- iv. Suponhamos agora que os bits sejam agrupados em pacotes de 1000 bytes e que sejam acrescentados 500 bytes de redundância por meio de um código corretor de erros. O código permite uma detecção correta desde que  $f + 2e < 200$ , em que  $e$  representa o número de erros e  $f$  o número de apagamentos. Qual a probabilidade aproximada de erro na detecção do pacote?

Exercício 20. ([1] Ex. 9.1-1, 9.1-5, modificado)

Considere o processo aleatório  $x(t) = \cos(2\pi f_c t + \Theta)$  em que  $\Theta$  é uma variável aleatória com distribuição uniforme entre 0 e  $A$ .

- a) apenas observando o conjunto de sinais possíveis, determine e explique se este é um processo estacionário.
- b) encontre  $\overline{x(t)}$  e  $R_x(t_1, t_2)$  e verifique se satisfazem os critérios de estacionariedade no sentido amplo.
- c) o processo é ergódico?

Exercício 21. ([1] Ex. 9.1-9)



Repita o Exercício 1, considerando agora que  $a$  é uma v.a. uniforme entre  $-1$  e  $1$ , e  $\Theta$  é uma v.a. uniforme entre  $0$  e  $2\pi$ , ambas independentes.

Exercício 22. ([1] Ex. 9.1-10)

Ache a potência média do processo aleatório do Exercício 2.

Exercício 23. ([1] Ex. 9.2-2)

Diga se as funções abaixo podem representar uma função de autocorrelação de um processo aleatório estacionário real. Explique.

a)  $\frac{1}{\tau^2+9}$    b)  $\frac{1}{\tau^2-1}$    c)  $e^{-\tau}u(t)$    d)  $N_0\delta(\tau) + \frac{1}{\tau^2+4}$    e)  $\text{sen}(2\pi f_0\tau)\text{f}\cos(2\pi f_0\tau)$

Exercício 24. ([1] Ex. 9.2-3)

Diga se as funções abaixo podem representar uma densidade espectral de potência de um processo aleatório estacionário real. Explique.

a)  $\frac{(2\pi f)^2}{(2\pi f)^4+2}$    b)  $\frac{1}{(2\pi f)^2-2}$    c)  $\frac{(2\pi f)^2}{(2\pi f)^3+2}$    d)  $\delta(f - f_0) + \frac{1}{(2\pi f)^2+2}$   
 e)  $\delta(f + f_0) - \delta(f - f_0)$    f)  $\cos(2\pi(f + f_0)) + j\text{sen}(2\pi(f - f_0))$    g)  $e^{-(2\pi f)^2}$

Exercício 25. ([1] Ex. 9.2-13)

Um sinal  $x(t)$  muda abruptamente sua amplitude em instantes aleatórios. Acontecem em média  $\beta$  alterações de amplitude a cada segundo, e a probabilidade de que não ocorra nenhuma alteração em  $\tau$  segundos é dada por  $P_0(\tau) = e^{-\beta\tau}$ . Os valores das mudanças de amplitude são independentes das mudanças anteriores, e as amplitudes são distribuídas aleatoriamente com PDF  $p_x(x)$ , cuja média é nula. Mostre que

$$R_x(\tau) = \overline{x^2} e^{-\beta|\tau|} \quad \text{e} \quad S_x(f) = \frac{2\beta\overline{x^2}}{\beta^2 + (2\pi f)^2}$$

Exercício 26. ([1] Ex. 9.3-2)

Dois processos aleatórios  $x(t)$  e  $y(t)$  são dados por  $x(t) = A\cos(2\pi f_0 t + \phi)$  e  $y(t) = B\text{sen}(2\pi n f_0 t + n\phi + \psi)$ , em que  $n$  é um inteiro  $n \neq 0$  e  $\phi$  é uma v.a. Uniformemente distribuída entre  $0$  e  $2\pi$ . Mostre que os dois processos são incoerentes.

Exercício 27. ([1] Ex. 9.3-3)

Se  $x(t)$  e  $y(t)$  forem dois processos aleatórios incoerentes, e

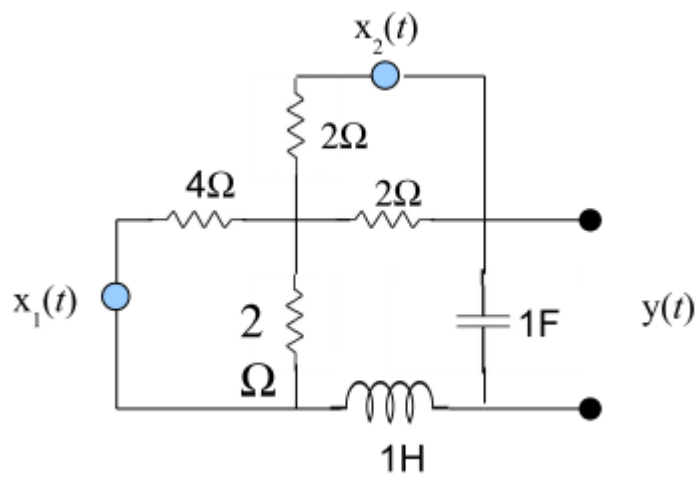
$$u(t) = 2x(t) - y(t)$$

$$v(t) = x(t) + 3y(t)$$

encontre  $R_u(\tau)$ ,  $R_v(\tau)$ ,  $R_{uv}(\tau)$  e  $R_{vu}(\tau)$  em termos de  $R_x(\tau)$  e  $R_y(\tau)$ .

Exercício 28. ([1] Ex. 9.4-1, modificado)

Considere o circuito abaixo



em que  $x_1$  e  $x_2$  são processos estocásticos independentes com densidade espectral de potência

$$S_{x_1} = \frac{2\alpha}{\alpha^2 + (j2\pi f)^2} \text{ e } S_{x_2} = K.$$

Ache a densidade espectral de potência da saída  $y(t)$ .

## 5. Resolução dos Exercícios

### Exercício 1

Chamando de  $n_{fail}$  o número de falhas

$$\Pr(n_{fail} \leq 5) = \Pr\left(\bigcup_{i=0}^5 n_{fail} = i\right)$$

Como são eventos disjuntos

$$\begin{aligned}\Pr(n_{fail} \leq 5) &= \sum_{i=0}^5 \Pr(n_{fail} = i) = \sum_{i=0}^5 \binom{26}{i} p^i (1-p)^{26-i} \\ &= (1-p)^{26} + 26p(1-p)^{25} + 325p^2(1-p)^{24} + 2600p^3(1-p)^{23} \\ &\quad + 14950p^4(1-p)^{22} + 65780p^5(1-p)^{21} = 0,278\end{aligned}$$

### Exercício 2

Uma senha bancária tem 4 dígitos de 0 a 9. Qual a probabilidade de que a soma dos dígitos seja igual a 7?

Temos  $N = 10^4 = 10000$  possibilidades de senhas. Chamando os dígitos de  $a_1 a_2 a_3 a_4$ , temos as seguintes condições para que a soma seja igual a 7

$$a_1 \leq 7, a_2 \leq 7 - a_1, a_3 \leq 7 - a_1 - a_2, a_4 = 7 - a_1 - a_2 - a_3$$

O número de combinações é

$$\begin{aligned}N_7 &= \sum_{a_1=0}^7 \sum_{a_2=0}^{7-a_1} 8 - a_1 - a_2 \\ &= (1 + 2 + 3 + 4 + 5 + 6 + 7 + 8) + (1 + 2 + 3 + 4 + 5 + 6 + 7) + \dots + 1 \\ &= 8(1) + 7(2) + 6(3) + \dots + 1(8) = 120\end{aligned}$$

e

$$\Pr(a_1 + a_2 + a_3 + a_4 = 7) = \frac{120}{10000} = 0,012$$

### Exercício 19

a)

$$\begin{aligned}\Pr(\varepsilon) &= \Pr(b = 0)\Pr(r = 1 \mid b = 0) + \Pr(b = 1)\Pr(r = 0 \mid b = 1) \\ \Pr(b = 1) &= 1 - \Pr(b = 0) = 0,7\end{aligned}$$

$$\Pr(\varepsilon) = (0,3)(0,01) + (0,7)(0,02) = 0,017$$

b)

$$Pr(b = 1 | r = 1) = \frac{Pr(r = 1 | b = 1)Pr(b = 1)}{Pr(r = 1)}$$

$$Pr(r = 1 | b = 1) = 1 - Pr(r = 0 | b = 1) - Pr(r = A) = 0,93$$

$$Pr(r = 1) = Pr(b = 0)Pr(r = 1 | b = 0) + Pr(b = 1)Pr(r = 1 | b = 1) = (0,3)(0,01) + (0,7)(0,93) = 0,654$$

$$Pr(b = 1 | r = 1) = \frac{0,93(0,7)}{0,654} = 0,9954$$

c)

$$Pr(y = -1) = Pr(r = 0) = Pr(b = 0)Pr(r = 0 | b = 0) + Pr(b = 0)Pr(r = 0 | b = 1) = 0,296$$

$$Pr(y = 0) = Pr(r = A) = 0,05$$

$$Pr(y = 1) = Pr(r = 1) = 0,654$$

$$E\{y\} = \sum_y y Pr(y = y) = -0,296 + 0,654 = 0,358$$

$$E\{y^2\} = \sum_y y^2 Pr(y = y) = 0,296 + 0,654 = 0,950$$

$$\text{var}\{y\} = E\{y^2\} - E^2\{y\} = 0,822$$

d) podemos criar a variável aleatória  $z_i = 1$  se houver um apagamento no  $i$ -ésimo bit,  $z_i = 2$  se houver um erro e  $z_i = 0$  se a transmissão for correta.

Desta forma temos que  $z = f + 2e = \sum_{i=1}^N z_i$ .

Queremos saber  $Pr(z < 200)$  e podemos usar o teorema do limite central.

$z$  é uma v.a Gaussiana com média  $\bar{z} = N\bar{z}_i$  e variância  $\sigma_z^2 = N\sigma_{z_i}^2$ .

$$\bar{z}_i = 1 \times 0,05 + 2 \times 0,017 + 0 \times Pr(C) = 0,084$$

$$E\{z_i^2\} = 1 \times 0,05 + 4 \times 0,017 = 0,118$$

$$\sigma_{z_i}^2 = E\{z_i^2\} - \bar{z}_i^2 = 0,1109$$

$$N = 1500 (8) = 12000$$

Portanto

$$\bar{z} = 1008, \sigma_z^2 = 1331$$

$$Pr(\varepsilon) = Pr(z > 200) \approx Q\left(\frac{200 - \bar{z}}{\sigma_z}\right) = Q\left(\frac{-200 - 1008}{\sqrt{1331}}\right) \approx 1$$

## 6. Referências

Este texto se baseia em grande parte no livro texto

- [1] B.P. Lathi e Z. Ding, *Sistemas de Comunicações Analógicas e Digitais Modernos*, 4ª Ed., Editora LTC, 2012, Caps. 8 e 9

Boas referências em probabilidade e processos estocásticos são

- [2] J.A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*, Cambridge University Press, 2006
- [3] J.P.A. Albuquerque, J.M.P. Fortes, W.A. Finamore, *Probabilidade, Variáveis Aleatórias e Processos Estocásticos*, Interciência, 2008
- [4] A. Papoulis e S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 4ª Edição, 2002

## 7. Apêndices

### a. Equivalente em Banda Base de Processos Aleatórios em Banda Passante

Dado um sinal aleatório em banda-passante  $x_p(t)$ , ou seja um sinal cuja densidade espectral de potência está limitada a um certo intervalo de frequência, ou seja,

$$S_{x_p}(f) = 0, \text{ para } ||f| - f_c| > \frac{B}{2} \quad (186)$$

Sabemos que qualquer sinal em banda passante pode ser escrito em termos de sua envoltória complexa, e o mesmo vale para sinais aleatórios, ou seja,

$$x(t) = x_Q(t) \cos(2\pi f_c t) - x_I(t) \sin(2\pi f_c t), \quad (187)$$

em que

$$\begin{aligned} x_I(t) &= x_p(t) \cos(2\pi f_c t) + \tilde{x}_p(t) \sin(2\pi f_c t) \\ x_Q(t) &= \tilde{x}_p(t) \cos(2\pi f_c t) - x_p(t) \sin(2\pi f_c t) \end{aligned} \quad (188)$$

são os componentes em fase e em quadratura, respectivamente, e  $\tilde{x}_p(t) = \mathcal{H}\{x_p(t)\}$  representa a transformada de Hilbert de  $x_p(t)$ .

Podemos ainda representar o sinal como

$$x_p(t) = \text{Re}\{x(t)e^{j2\pi f_c t}\} \quad (189)$$

em que

$$x(t) = x_I(t) + jx_Q(t) \quad (190)$$

é a envoltória complexa do sinal.

Primeiramente, supondo que  $x_{BP}(t)$  tem média zero, como a transformada de Hilbert é uma operação linear,  $\mathcal{H}\{x_{BP}(t)\}$  também tem média zero, e, conseqüentemente,  $x_I(t)$  e  $x_Q(t)$  também têm média zero.

Para acharmos a densidade espectral de potência de  $x_I(t)$ , achamos primeiro sua autocorrelação

$$\begin{aligned} R_{x_I}(\tau) &= E\{x_I(t + \tau)x_I(t)\} \\ &= E\left\{\left[x_p(t + \tau) \cos(2\pi f_c(t + \tau)) + \tilde{x}_p(t + \tau) \sin(2\pi f_c(t + \tau))\right] \right. \\ &\quad \left. \times \left[x_p(t) \cos(2\pi f_c t) + \tilde{x}_p(t) \sin(2\pi f_c t)\right]\right\} \\ &= R_{x_p}(\tau) \cos(2\pi f_c(t + \tau)) \cos(2\pi f_c t) \\ &\quad + R_{x_p \tilde{x}_p}(\tau) \cos(2\pi f_c(t + \tau)) \sin(2\pi f_c t) \\ &\quad + R_{\tilde{x}_p x_p}(\tau) \sin(2\pi f_c(t + \tau)) \cos(2\pi f_c t) \\ &\quad + R_{\tilde{x}_p}(\tau) \sin(2\pi f_c(t + \tau)) \sin(2\pi f_c t) \end{aligned} \quad (191)$$

Sabemos que a transformada de Hilbert equivale a uma filtragem, ou seja,

$$y(t) = \mathcal{H}\{x(t)\} = \tilde{x}(t) \Leftrightarrow Y(f) = \tilde{X}(f) = X(f)(-j \text{sgn}(f)) = H_{Hil}(f)X(f) \quad (192)$$

Para processos aleatórios temos que, se  $y(t)$  é obtido pela filtragem de  $x(t)$  por um filtro  $H(f)$ , sabemos que

$$\begin{aligned}
S_{yx}(f) &= H(f)S_x(f) \\
S_{xy}(f) &= H^*(f)S_x(f) \\
S_y(f) &= |H(f)|^2 S_x(f)
\end{aligned} \tag{193}$$

e, portanto, para a transformada de Hilbert,

$$\begin{aligned}
S_{\tilde{x}_p}(f) &= S_{x_p}(f) \Leftrightarrow R_{\tilde{x}_p}(\tau) = R_{x_p}(\tau) \\
S_{\tilde{x}_p x_p}(f) &= H_{Hil}(f)S_x(f) \Leftrightarrow R_{\tilde{x}_p x_p}(\tau) = \tilde{R}_{x_p}(\tau) \\
S_{x_p \tilde{x}_p}(f) &= H_{Hil}^*(f)S_x(f) = -H_{Hil}(f)S_x(f) \Leftrightarrow R_{x_p \tilde{x}_p}(\tau) = -\tilde{R}_{x_p}(\tau)
\end{aligned} \tag{194}$$

Substituindo estes resultados em (191), e com algumas manipulações trigonométricas, podemos verificar que

$$R_{x_I}(\tau) = R_{x_p}(\tau) \cos(2\pi f_c \tau) + \tilde{R}_{x_p}(\tau) \sin(2\pi f_c \tau) \tag{195}$$

A partir da função de autocorrelação, podemos então obter a densidade espectral de potência

$$\begin{aligned}
S_{x_I}(f) &= \mathcal{F}\{R_{x_I}(\tau)\} \\
&= \frac{1}{2} \left\{ S_{x_p}(f - f_0) + S_{x_p}(f + f_0) + \right. \\
&\quad \left. -jS_{x_p}(f - f_0)(-j)\text{sgn}(f - f_0) + jS_{x_p}(f + f_0)(-j)\text{sgn}(f + f_0) \right\} \\
&= \frac{1}{2} \left\{ S_{x_p}(f - f_0) + S_{x_p}(f + f_0) + \right. \\
&\quad \left. -S_{x_p}(f - f_0)\text{sgn}(f - f_0) + S_{x_p}(f + f_0)\text{sgn}(f + f_0) \right\} \\
&= \begin{cases} S_{x_p}(f - f_0) + S_{x_p}(f + f_0) & , |f| < f_c \\ 0 & , c. c. \end{cases}
\end{aligned} \tag{196}$$

ou seja,  $x_I(t)$  é um processo aleatório em banda base, e pode ser escrito de outra forma como

$$S_{x_I}(f) = S_{x_p}^+(f + f_0) + S_{x_p}^-(f + f_0) \tag{197}$$

em que  $S_{x_p}^+(f) = S_{x_p}(f)u(f)$  é a parte positiva do espectro de  $x_p(t)$  e  $S_{x_p}^-(f) = S_{x_p}(f)u(-f)$  sua parte negativa.

Usando a mesma abordagem, podemos ver também que

$$\begin{aligned}
S_{x_Q}(f) &= S_{x_I}(f) \\
S_{x_I x_Q}(f) &= -S_{x_Q x_I}(f) = \begin{cases} j \left( S_{x_p}(f + f_0) - S_{x_p}(f - f_0) \right) & , |f| < f_c \\ 0 & , c. c. \end{cases}
\end{aligned} \tag{198}$$

### Ruído Branco em Banda Passante

Um ruído branco em banda passante  $n_p(t)$  é tal que

$$S_{n_p}(f) = \begin{cases} \frac{N_0}{2} & , |f - f_c| < \frac{B}{2} \\ 0 & , c. c. \end{cases} \tag{199}$$

Como o espectro é simétrico em torno de 0, portanto  $S_{n_p}^+(f + f_0) = S_{n_p}^-(f - f_0)$  e

$$S_{n_I n_Q}(f) = 0 \Leftrightarrow R_{n_I n_Q}(\tau) = 0, \tag{200}$$

ou seja, seus componentes em fase  $n_I(t)$  e quadratura  $n_Q(t)$  são decorrelatados, e podemos ver que

$$S_{n_I}(f) = S_{n_Q}(f) = \begin{cases} N_0 & , |f| < \frac{B}{2} \\ 0 & , c. c. \end{cases} \quad (201)$$

Podemos ver ainda que para a envoltória complexa

$$\begin{aligned} R_n(\tau) &= E\{n(t+\tau)n^*(t)\} \\ &= \left\{ \left( n_I(t+\tau) + jn_Q(t+\tau) \right) \left( n_I(t) - jn_Q(t) \right) \right\} \\ &= R_{n_I}(\tau) + R_{n_Q}(\tau) - jR_{n_I n_Q} + jR_{n_Q n_I} = R_{n_I}(\tau) + R_{n_Q}(\tau) \end{aligned} \quad (202)$$

e, portanto,

$$S_n(f) = S_{n_I}(f) + S_{n_Q}(f) = \begin{cases} 2N_0 & , |f| < \frac{B}{2} \\ 0 & , c. c. \end{cases} \quad (203)$$