Revisões sobre Probabilidades, Variáveis Aleatórias e Processos Estocásticos

Desempenho e Dimensionamento de Redes

Prof. Amaro de Sousa (asou@ua.pt)

DETI-UA, 2020/2021

Experiência aleatória

- Numa experiência aleatória, o *espaço de resultados*, S, é o conjunto de todos os resultados possíveis da experiência
- Qualquer subconjunto E do espaço de resultados S designa-se por evento ou *acontecimento*

- Dados dois acontecimentos E e F, podem-se definir outros acontecimentos:

  A *união dos acontecimentos*, $E \cup F$

  A *intersecção dos acontecimentos*, EF

- Quando $EF = \emptyset$ ($\emptyset$ é o conjunto vazio) os acontecimentos dizem-se *mutuamente exclusivos*
- O *complemento de E*, E^c, é o conjunto de elementos de S que não pertencem a E

Probabilidades definidas sobre acontecimentos

- Para cada acontecimento E de S, admite-se a existência de um número $P(E)$ designado por probabilidade de E, se satisfaz as seguintes condições:

  (1) $0 \leq P(E) \leq 1$

  (2) $P(S) = 1$

  (3) Para qualquer conjunto de acontecimentos mutuamente exclusivos $E_1, E_2, E_3, \ldots$

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i)$$

- Corolários:

  $P(E) + P(E^c) = 1$

  $P(E \cup F) = P(E) + P(F) - P(EF)$

Probabilidades condicionadas

- Dados dois acontecimentos E e F, a <u>*probabilidade condicionada de E ocorrer dado que F ocorreu*</u> designa-se por $P(E|F)$ e é definida por

$$P(E|F) = P(EF)/P(F)$$

- Dois acontecimentos E e F dizem-se <u>*acontecimentos independentes*</u> se

$$P(EF) = P(E)P(F)$$

- Se E e F são independentes, então:

$$P(E|F) = P(E) \qquad \text{e} \qquad P(F|E) = P(F)$$

ou seja, se o conhecimento que um acontecimento ocorreu não afetar a probabilidade do outro ocorrer.

- Sejam $F_1, F_2, \ldots, F_n$ acontecimentos mutuamente exclusivos tais que a sua união forma o espaço de resultados *S. Então,*

$$P(E) = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(E \mid F_i)P(F_i)$$

Regra de Bayes

Sejam F_1, F_2, …, F_n acontecimentos mutuamente exclusivos tais que a sua união forma o espaço de resultados S.

Tendo ocorrido o acontecimento E, a probabilidade de F_j (j = 1, 2, …, n) ter ocorrido é dada por:

$$P(F_j \mid E) = \frac{P(EF_j)}{P(E)} = \frac{P(E \mid F_j)P(F_j)}{P(E)} = \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i)P(F_i)}$$

Probabilidades condicionadas – Exemplo 1

Num teste de escolha múltipla, um estudante sabe a resposta certa com probabilidade p e adivinha a resposta com probabilidade $1 - p$. Ao adivinhar a resposta, o estudante acerta com probabilidade $1/m$, sendo m o número de alternativas de escolha múltipla.

Determine a probabilidade de um estudante *(i)* responder corretamente à pergunta e *(ii)* saber a resposta dado que a respondeu corretamente.

Acontecimentos: E – o aluno responde corretamente
F_1 – o aluno sabe a resposta
F_2 – o aluno não sabe a resposta

(i) $P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2)$

$= 1 \times p + 1/m \times (1 - p) =$

$= p + (1 - p)/m$

(ii) $P(F_1|E) = P(E|F_1)P(F_1) \,/\, P(E)$

$= 1 \times p \,/\, [p + (1 - p)/m] =$

$= p\, m \,/\, [1 + (m - 1)\, p]$

Se $p = 50\%$ e $m = 4$, então *(i)* $P(E) = 62.5\%$ e *(i)* $P(F_1|E) = 80\%$

Probabilidades condicionadas – Exemplo 2

Numa ligação sem fios (wireless) entre dois equipamentos, a probabilidade dos pacotes de dados serem recebidos com erros é de 0.1% em condições normais ou de 10% quando há interferências. A probabilidade de haver interferência é de 2%. Os equipamentos têm a capacidade de verificar na receção se os pacotes de dados foram recebidos com erros ou não.

Determine: *(i)* a probabilidade de um pacote ser recebido com erros e *(ii)* se um pacote for recebido com erros, qual a probabilidade da ligação estar com interferência.

Acontecimentos: E – o pacote é recebido com erros
F_1 – a ligação está em condições normais
F_2 – a ligação está com interferência

(i) $P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2)$

$= 0.001\times(1 – 0.02) + 0.1\times0.02$

$= 0.00298 = 0.298\%$

(ii) $P(F_2|E) = P(E|F_2)P(F_2) / P(E)$

$= 0.1 \times 0.02 / 0.00298$

$= 0.671 = 67.1\%$

Variáveis aleatórias

- Uma *variável aleatória X é uma função* que atribui um número real a cada ponto do espaço de resultados S de uma experiência aleatória.

- A *função distribuição* (ou função de distribuição cumulativa) da v.a. X é:

$$F(x) = P(X \leq x) \quad , -\infty < x < +\infty$$

- *Propriedades* da função distribuição:

  (1) $0 \leq F(x) \leq 1$ para todo o x

  (2) se $x_1 \leq x_2$ então $F(x_1) \leq F(x_2)$ (função não decrescente)

  (3) $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$

  (4) $P(a < X \leq b) = F(b) - F(a)$, para $a < b$

Variáveis aleatórias discretas

- Uma *variável aleatória X diz-se discreta* se puder tomar, quando muito, um número contável de valores x_1, x_2, …, x_i, …

- Define-se *função probabilidade* (ou função massa de probabilidade) da v.a discreta X por

  $f(x_i) = P(X = x_i)$ para todos os valores de $i = 1,2,3,...$

- Obrigatoriamente, tem de acontecer que: $\sum_{i=1}^{\infty} f(x_i) = 1$

- A *função distribuição* da v.a discreta X é:

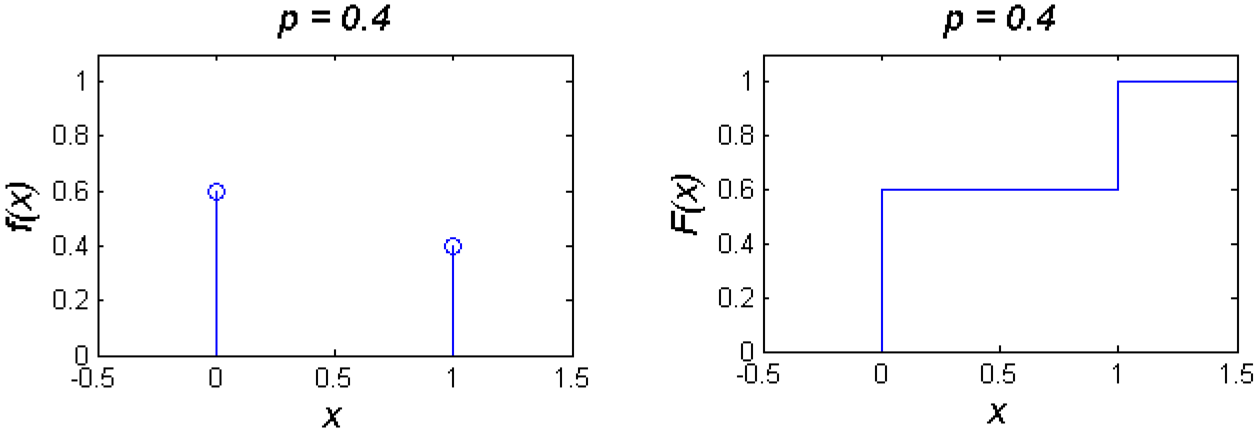
$$F(x) = \sum_{x_i \le x} f(x_i) \qquad , -\infty < x < +\infty$$

Variáveis aleatórias discretas

Variável aleatória de Bernoulli: experiência que pode resultar em sucesso com probabilidade p ou insucesso com probabilidade $1 - p$.

Se $X = 1$ representar um sucesso e $X = 0$ um insucesso, a função probabilidade é:

$$f(i) = p^i (1-p)^{1-i}, i = 0,1$$



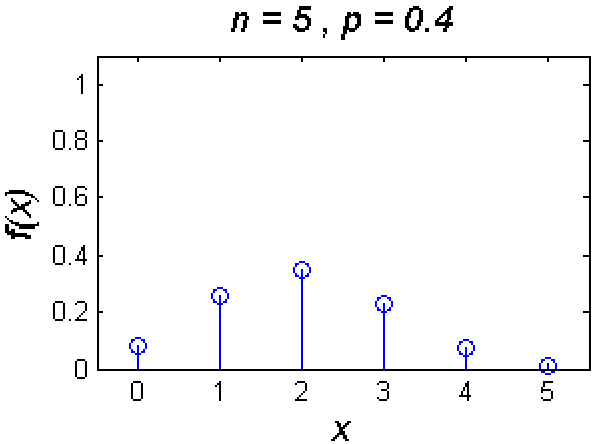
Variáveis aleatórias discretas

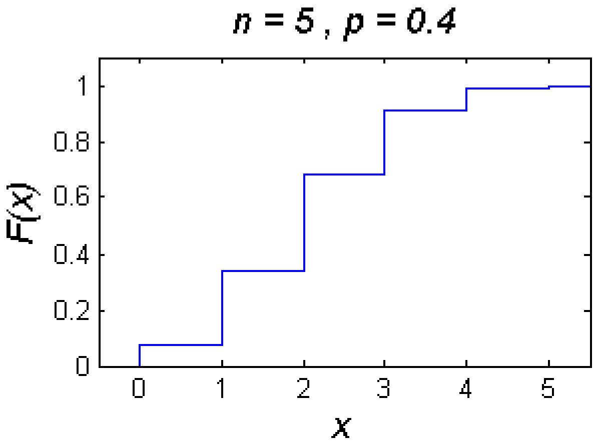
<u>Variável aleatória binomial</u>: conjunto de *n* experiências de Bernoulli independentes, cada uma das quais resulta num sucesso com probabilidade *p* ou num insucesso com probabilidade 1 – *p*.

Se *X* representar o número de sucessos em *n* experiências, a função probabilidade é:

$$f(i) = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, 2, ..., n$$

onde $\binom{n}{i} = \frac{n!}{i!(n-i)!}$



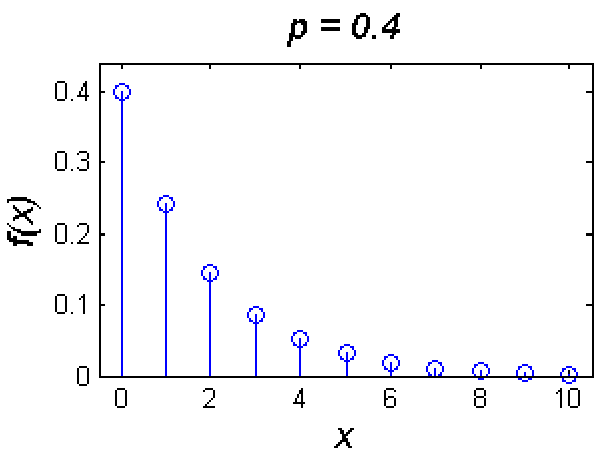


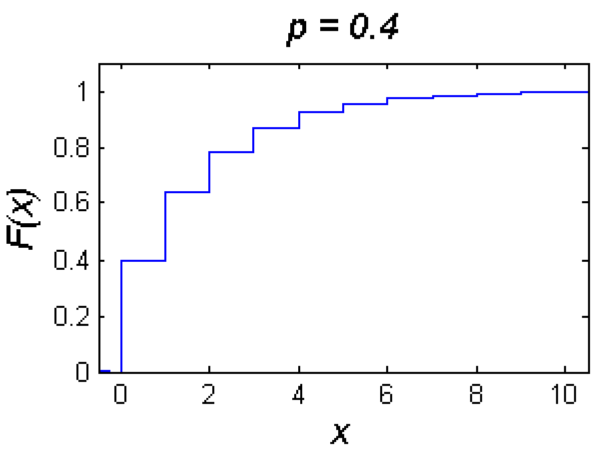
Variáveis aleatórias discretas

Variável aleatória geométrica: são realizadas experiências de Bernoulli independentes com parâmetro *p* (probabilidade de sucesso) até que ocorra um sucesso.

Se *X* representar o número de insucessos antes do sucesso, a função probabilidade é

$$f(i) = (1-p)^i \, p \quad , \; i = 0, 1, 2, \ldots$$





Se *X* representar o número de experiências até ao sucesso, a função probabilidade é

$$f(i) = (1-p)^{i-1} \, p \quad , \; i = 1, 2, \ldots$$

Variáveis aleatórias discretas – Exemplo 3

Numa dada ligação de dados, a probabilidade de erro de bit (BER – *Bit Error Rate*) é 10^{-5} e os erros em diferentes bits são estatisticamente independentes.

Determine: *(i)* a probabilidade de um pacote de dados de 100 Bytes ser recebido sem erros e *(ii)* a probabilidade de um pacote de dados de 1000 Bytes ser recebido com 2 ou mais erros.

O número de erros num pacote é uma variável aleatória binomial em que a probabilidade de sucesso é o BER e o número de experiências de Bernoulli é o número de bits do pacote:

(i) $$f(0)=\binom{n}{0}p^0(1-p)^{n-0}=\binom{100\times 8}{0}\times\left(1-10^{-5}\right)^{100\times 8}=0.992=99.2\%$$

(ii) $$1-f(0)-f(1)=1-\binom{n}{0}p^0(1-p)^{n-0}-\binom{n}{1}p^1(1-p)^{n-1}$$
$$=1-\left(1-10^{-5}\right)^{8000}-8000\times 10^{-5}\left(1-10^{-5}\right)^{7999}=3.034\text{E-}3=0.3\%$$

Variáveis aleatórias contínuas

- Uma *variável aleatória X diz-se contínua* se existir uma função não negativa $f(x)$ tal que para qualquer conjunto de números reais B:

$$P(X \in B) = \int_B f(x)dx \qquad \int_{-\infty}^{+\infty} f(x)dx = 1$$

$f(x)$ é a *função densidade de probabilidade* da v.a contínua X

- Resulta então que: $$P\{a \le X \le b\} = \int_a^b f(x)dx$$

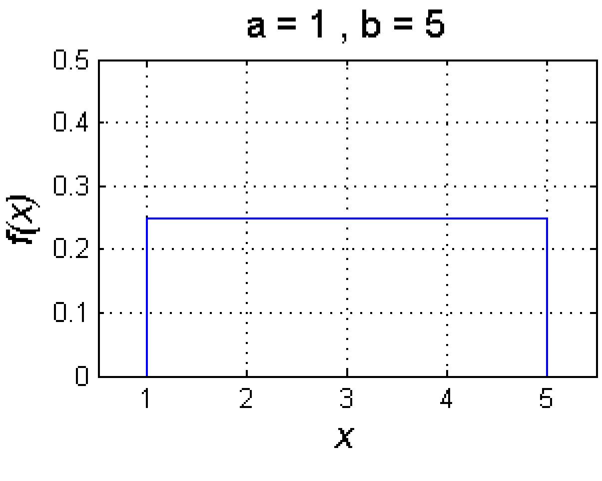
- A *função distribuição* da v.a contínua X é:

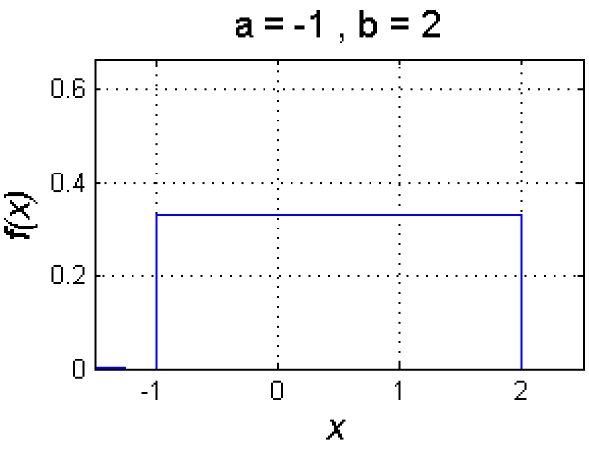
$$F(x) = P(X \in [-\infty, x]) = \int_{-\infty}^{x} f(y)dy$$

Exemplos de variáveis aleatórias contínuas

Variável aleatória com Distribuição Uniforme: uma v.a. diz-se uniformemente distribuída no intervalo [a,b] se a função densidade de probabilidade for dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & , a < x < b \\ 0 & , cc \end{cases}$$

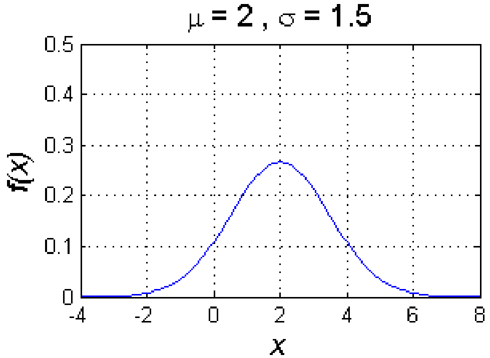




Exemplos de variáveis aleatórias contínuas

<u>Variável aleatória com Distribuição Gaussiana (ou Normal)</u>: Uma v.a. X tem uma distribuição Gaussiana com média μ e desvio padrão σ se a função densidade é dada por:

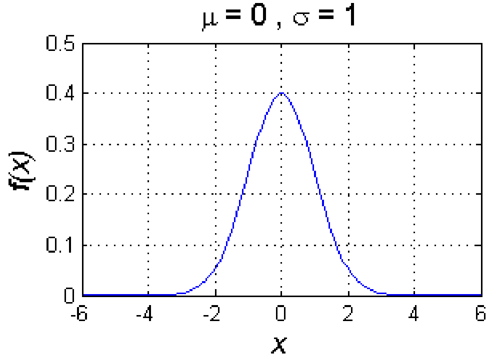
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$



Designa-se por distribuição Gaussiana (ou Normal) padrão à distribuição Gaussiana com média 0 e desvio padrão 1.

Neste caso:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



Média de uma variável aleatória

- *Média* ou *valor esperado* de uma v.a. *X*:

$$E[X] = \begin{cases} \sum_{j=1}^{\infty} x_j f_X(x_j) & se\ X\ discreta \\ \int_{-\infty}^{+\infty} x f_X(x) dx & se\ X\ continua \end{cases}$$

- Propriedades importantes:

$$E[cX] = cE[X]$$

$$E\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i E[X_i]$$

- Média da v.a. $Y = g(X)$:

$$E[g(X)] = \begin{cases} \sum_{j=1}^{\infty} g(x_j) f_X(x_j) & se\ X\ discreta \\ \int_{-\infty}^{+\infty} g(x) f_X(x) dx & se\ X\ continua \end{cases}$$

Variância de uma variável aleatória

- *Variância* de uma v.a. X:

$$Var[X] = E\left[\left(X - E[X]\right)^2\right] = E\left[X^2\right] - E[X]^2$$

- Propriedades importantes:

$Var[X] \geq 0$

$Var[cX] = c^2\, Var[X]$

$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i]$ se X_i forem independentes

Exemplo 4

Uma ligação de dados de 10 Mbps suporta um fluxo de pacotes cujo tamanho é 100 Bytes com probabilidade 10%, 500 Bytes com probabilidade 50% e 1500 Bytes com probabilidade 40%. Considere a variável aleatória *X* representativa do tempo de transmissão dos pacotes.

Determine: *(i)* o tempo médio (i.e., $E[X]$) de transmissão dos pacotes *(ii)* o segundo momento (i.e., $E[X^2]$) do tempo de transmissão dos pacotes e *(iii)* a variância (i.e., $Var[X]$) do tempo de transmissão dos pacotes.

(i) $$E[X] = \sum_{j=1}^{\infty} x_j f_X(x_j) = \frac{100 \times 8}{10^7} \times 0.1 + \frac{500 \times 8}{10^7} \times 0.5 + \frac{1500 \times 8}{10^7} \times 0.4$$

$$= 0.688 \times 10^{-3} \text{ seg} = 0.688 \text{ mseg}$$

(ii) $$E[X^2] = \sum_{j=1}^{\infty} (x_j)^2 f_X(x_j) = \left(\frac{100 \times 8}{10^7}\right)^2 \times 0.1 + \left(\frac{500 \times 8}{10^7}\right)^2 \times 0.5 + \left(\frac{1500 \times 8}{10^7}\right)^2 \times 0.4$$

$$= 6.5664 \times 10^{-7} \text{ seg}^2$$

Exemplo 4 - continuação

Uma ligação de dados de 10 Mbps suporta um fluxo de pacotes cujo tamanho é 100 Bytes com probabilidade 10%, 500 Bytes com probabilidade 50% e 1500 Bytes com probabilidade 40%. Considere a variável aleatória *X* representativa do tempo de transmissão dos pacotes.

Determine: *(i)* o tempo médio (i.e., $E[X]$) de transmissão dos pacotes *(ii)* o segundo momento (i.e., $E[X^2]$) do tempo de transmissão dos pacotes e *(iii)* a variância (i.e., $Var[X]$) do tempo de transmissão dos pacotes.

(iii) **1ª alternativa:** $Var[X] = E\left[\left(X - E[X]\right)^2\right]$

$$Var[X] = \left(\frac{100\times 8}{10^7} - E[X]\right)^2 \times 0.1 + \left(\frac{500\times 8}{10^7} - E[X]\right)^2 \times 0.5 + \left(\frac{1500\times 8}{10^7} - E[X]\right)^2 \times 0.4$$

$$= 1.833\times 10^{-7} \text{ seg}^2$$

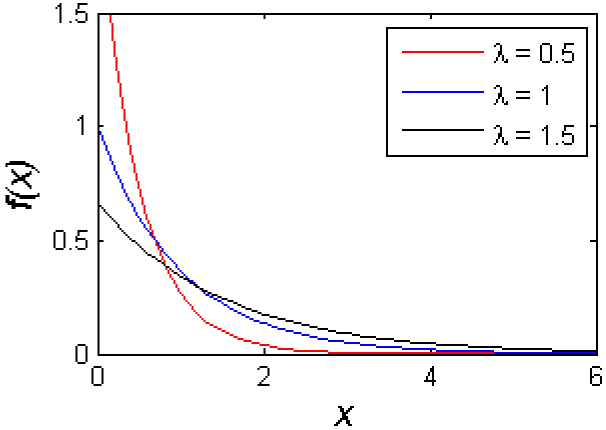
2ª alternativa: $Var[X] = E[X^2] - E[X]^2$

$$Var[X] = 6.5664\times 10^{-7} - \left(0.688\times 10^{-3}\right)^2 = 1.833\times 10^{-7} \text{ seg}^2$$

Distribuição exponencial

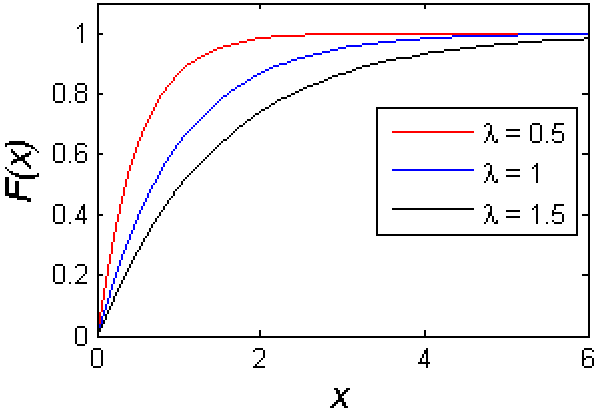
- Uma v. a. *X* tem uma distribuição exponencial com parâmetro λ, $\lambda > 0$, se a sua função densidade de probabilidade for:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



- A função de distribuição é dada por:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Distribuição exponencial

- A média e a variância de uma distribuição exponencial são:

$$E[X] = \frac{1}{\lambda} \qquad Var[X] = \left(\frac{1}{\lambda}\right)^2$$

- A distribuição exponencial não tem memória, isto é,

$$P\{X > s + t \mid X > t\} = P\{X > s\}$$

- Se X_1 e X_2 são v. a. independentes e exponencialmente distribuídas com médias $1/\lambda_1$ e $1/\lambda_2$ respetivamente, então

$$P\{X_1 < X_2\} = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

Distribuição exponencial – Exemplo 5

Uma ligação de dados de 10 Mbps suporta um fluxo de pacotes cujo tamanho é exponencialmente distribuído com média de 1000 Bytes. Considere a variável aleatória X representativa do tempo de transmissão dos pacotes.

Determine: *(i)* o tempo médio (i.e., $E[X]$) de transmissão dos pacotes *(ii)* a variância (i.e., $Var[X]$) do tempo de transmissão dos pacotes e *(iii)* o segundo momento (i.e., $E[X^2]$) do tempo de transmissão dos pacotes.

(i) $$E[X] = \frac{1000 \times 8}{10^7} = 8 \times 10^{-4} = 0.8\,\text{mseg}$$

Capacidade da ligação

$$E[X] = \frac{1}{\mu} \quad \Leftrightarrow \quad \mu = \frac{1}{E[X]} = \frac{1}{8 \times 10^{-4}} = 1250\,\text{pacotes/s}$$

(ii) $$Var[X] = \left(\frac{1}{\mu}\right)^2 = \left(8 \times 10^{-4}\right)^2 = 6.4 \times 10^{-7}\ \text{seg}^2$$

(iii) $$Var[X] = E[X^2] - E[X]^2 \quad \Leftrightarrow \quad E[X^2] = Var[X] + E[X]^2$$

$$E[X^2] = 6.4 \times 10^{-7} + \left(8 \times 10^{-4}\right)^2 = 1.28 \times 10^{-6}\ \text{seg}^2$$

Distribuição exponencial – Exemplo 6

Uma pessoa entra num banco e encontra os 2 empregados do banco ocupados a servir clientes. Não existem outros clientes no banco, pelo que a pessoa começará a ser atendida logo que um dos clientes que se encontra a ser atendido deixe o banco.

O empregado A serve os clientes segundo uma distribuição exponencial à taxa λ_A = 12 clientes por hora e o empregado B serve os clientes segundo uma distribuição exponencial à taxa λ_B = 8 clientes por hora. Determine a probabilidade da pessoa que entrou ser o segundo cliente a ser servido.

Evento A – o empregado A termina de servir o seu cliente antes do empregado B

Evento B – o empregado B termina de servir o seu cliente antes do empregado A

Evento C – a pessoa que entrou é o segundo cliente a ser servido

$$P(C) = P(C \mid A) \times P(A) + P(C \mid B) \times P(B) =$$

$$= \frac{\lambda_A}{\lambda_A + \lambda_B} \times \frac{\lambda_A}{\lambda_A + \lambda_B} + \frac{\lambda_B}{\lambda_A + \lambda_B} \times \frac{\lambda_B}{\lambda_A + \lambda_B} =$$

$$= \frac{12}{12+8} \times \frac{12}{12+8} + \frac{8}{12+8} \times \frac{8}{12+8} = 0.52 = 52\%$$

Processos estocásticos

- Um *processo estocástico* $\{X(t),\ t \in T\}$ é um conjunto de variáveis aleatórias: para cada $t \in T$, $X(t)$ é uma variável aleatória.

- O índice t é frequentemente interpretado como tempo. Nesta interpretação, $X(t)$ é o *estado* do processo no instante t.

- O conjunto T é o *conjunto de índices* do processo.

  (1) se T é um conjunto contável, designa-se o processo estocástico como sendo *em tempo discreto*

  (2) se T é um intervalo da reta real, designa-se o processo estocástico como sendo *em tempo contínuo*

- O *espaço de estados* é o conjunto de todos os valores que as variáveis aleatórias $X(t)$ podem tomar.

Exemplos de processos estocásticos

Considere um sistema com uma fila de espera e um servidor. A este sistema chegam clientes para serem servidos.

Atrasos sofridos por cada cliente

na fila de espera

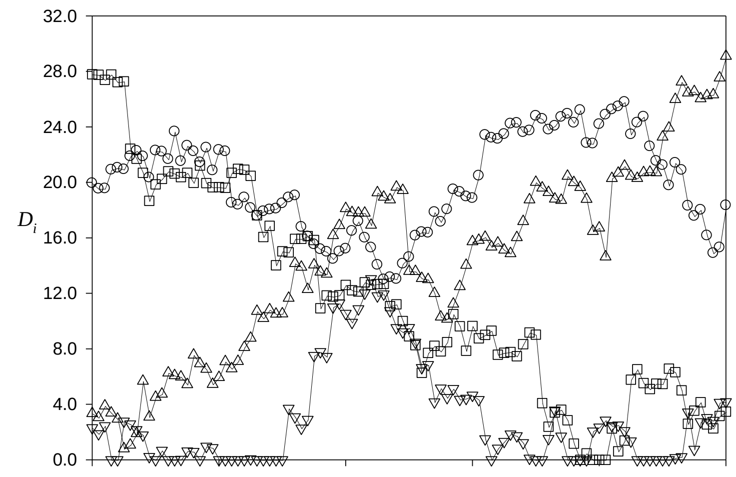
(1) é um processo estocástico em

tempo discreto (1º cliente,

2º cliente, etc.)

(2) o estado é uma variável contínua

(o tempo de espera é um valor real)



O número de clientes em espera

(1) é um processo estocástico

em tempo contínuo

(2) o estado é uma variável discreta

(0 clientes, 1 cliente, 2 clientes, etc.)