

# MPEI 2015-2016

# Aula 5

Caracterização de variáveis aleatórias  
(continuação)

Distribuições

# Caracterizando variáveis aleatórias

## Parte 2

# Momentos

- A fdp ou a fda de uma va fornece uma descrição completa dessa va
- Mas em muitos casos não necessitamos de toda a informação
  - Por exemplo, no caso dos “bugs” em módulos de código saber o valor médio e o máximo pode ser suficiente
- Momentos de um va representam um tipo de informação parcial acerca da va de grande utilidade
  - A Média é um exemplo de momento

# Valor esperado

- Valor esperado de N lançamentos de um dado

Ex: 4    1    6    6    5    5    5    3    4    2 ...

$$\text{Média} = (1+1+..1 + 2+2+...+2 + ... ) / N$$

$$= (\text{numDe1s} \times 1 + \text{numDe2s} \times 2 \dots ) / N$$

$$= \text{numDe1s}/N \times 1 + \text{NumDe2s}/N \times 2 + \dots$$

$$= p(1) \times 1 + p(2) \times 2 + p(3) \times 3 \dots + p(6) \times 6$$

$$= \sum_i p(x_i) x_i \quad \text{com } x_i = 1, 2, \dots 6$$

# $E(X)$

- O valor esperado ou expectância de uma variável
  - Designa-se por  $E(X)$
  - No caso discreto:  $E(X) = \sum_i x_i p(x_i)$
  - No caso contínuo:  $E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$
- $E(X)$  é linear
  - $E(aX) = aE(X)$      $E(X+Y) = E(X) + E(Y)$      $E(X+c) = E(X)+c$
  - Sendo  $a$  e  $c$  constantes ( $\in R$ ) e  $X$  e  $Y$  variáveis aleatórias

# Exemplo

$x_i$	$p(x_i)$	$x_i \cdot p(x_i)$
-1	.1	-.1
0	.2	.0
1	.4	.4
2	.2	.4
3	.1	<u>.3</u>
		<b>1.0</b>

$$\mu = E(X) = 1.0$$

# Variância (e desvio padrão)

- $\text{var}(X)=\sigma^2 = \sum_i (x_i - E(X))^2 p(x_i)$

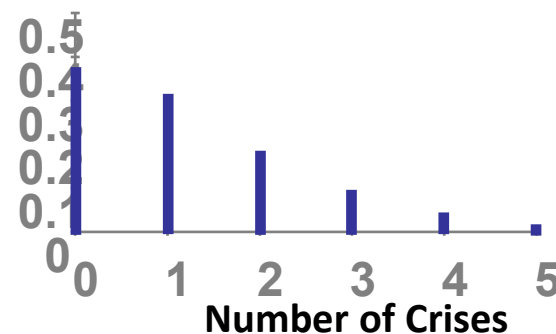
$$= E(X^2) - E^2(X)$$

Demonstra-se facilmente

- A raiz quadrada da variância é o desvio padrão
  - Muitas vezes representado por  $\sigma$



# Exemplo (discreto)



$x_i$	$p(x_i)$	$(x_i - \mu)$	$(x_i - \mu)^2$	$(x_i - \mu)^2 \bullet p(x_i)$
0	.37	-1.15	1.32	.49
1	.31	-0.15	0.02	.01
2	.18	0.85	0.72	.13
3	.09	1.85	3.42	.31
4	.04	2.85	8.12	.32
5	.01	3.85	14.82	.15
				<hr/>
				1.41

# Média e variância - interpretação

- $E(X)$  pode ser interpretado como:
  - Valor médio de  $X$
  - Centro de gravidade da fpm ou fdp
- Desvio padrão / Variância dá uma medida da dispersão da va
  - Pequenos valores indicam va muito concentrada em torno da média
    - Se for zero não temos va aleatória (todos valores iguais à média)

# Média e variância - propriedades

- “shift”

$$E(X+C)=E(X)+C$$

$$\text{var}(X+C)=\text{var}(X)$$

- Multiplicação por um factor de escala

$$E(k X) = k E(X)$$

$$\text{var}(k X) = \mathbf{k^2} \text{var}(X)$$

# Média e variância - propriedades

## %% Matlab

```
X=rand(1, 1e6)*6;
```

```
[mean(X),mean(X+2),mean(X+2)]
```

```
[mean(X),mean(2*X),2*mean(X)]
```

```
[var(X), var(X+2)]
```

```
r=var(2*X)/var(X); % quociente entre var(2X) e var(X)
```

```
k=sqrt(r);
```

```
[var(X), var(2*X),r,k]
```

# Momentos de ordem n

- Os conceitos de média e variância podem ser generalizados
- Momento de ordem n (caso discreto):
- $m_n = E(X^n) = \sum_i x_i^n p_X(x_i)$
- Exemplo (dados) =  $1^2 \times \frac{1}{6} + 2^2 \frac{1}{6} + 3^2 \frac{1}{6} +$   
... =  $\frac{1+2+4+9+16+25+36}{6} = 15,1667$

# Distribuições - Motivação

- As fpm podem assumir as mais variadíssimas formas, mas existe um conjunto de “formas” que aparecem repetidamente em muitos e variados problemas
  - Formam um conjunto de ferramentas base que é muito útil conhecer ...

# Existem muitas distribuições

- Discretas
  - Bernoulli
  - Binomial
  - Poisson
  - Geométrica
  - ...
- Contínuas
  - Uniforme
  - Normal
  - Qui-quadrado
  - T de Student ...
- Ver Wikipedia
  - [https://en.wikipedia.org/wiki/List\\_of\\_probability\\_distributions](https://en.wikipedia.org/wiki/List_of_probability_distributions)

# Distribuições Discretas

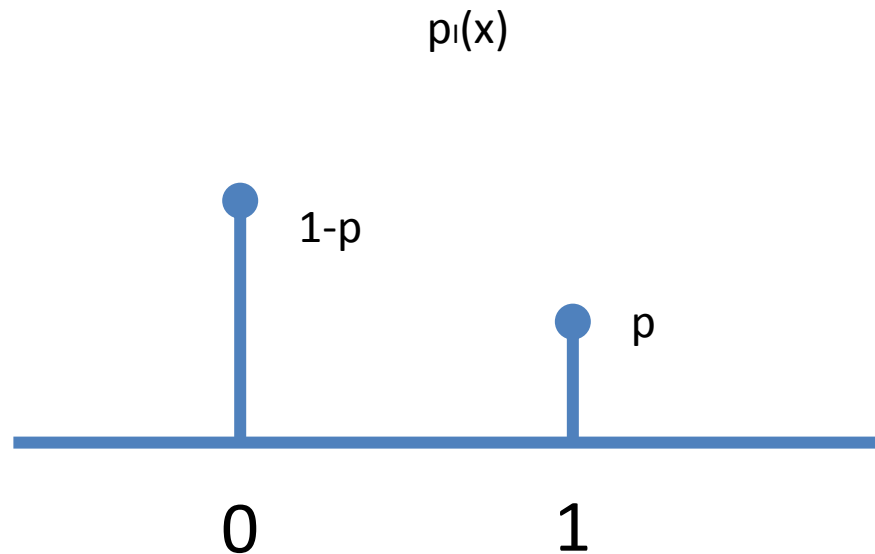


# Variável de Bernoulli

- Distribuição directamente relacionada com as experiências de Bernoulli
- Seja  $A$  um acontecimento relacionado como resultado de uma experiência aleatória
- A variável de Bernoulli define-se como
- $$I_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{caso contrário} \end{cases}$$

# Variável de Bernoulli

- $S_I = \{0,1\}$
- $p = \Pr(A)$
- $p_I(1) = p$
- $p_I(0) = 1-p$
- $E(I) = ?$
- $\text{Var}(I) = ?$



# Variável de Bernoulli

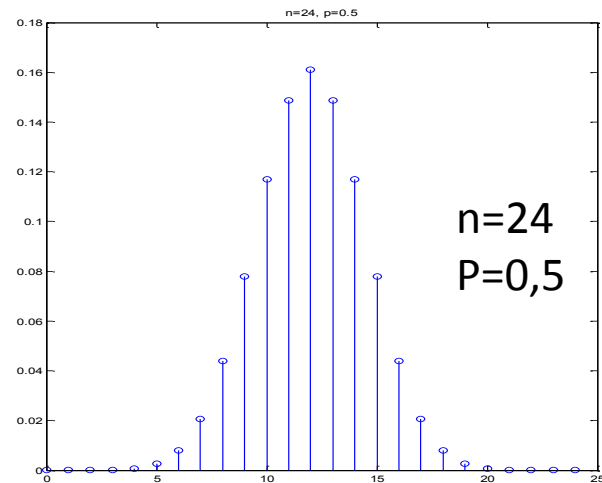
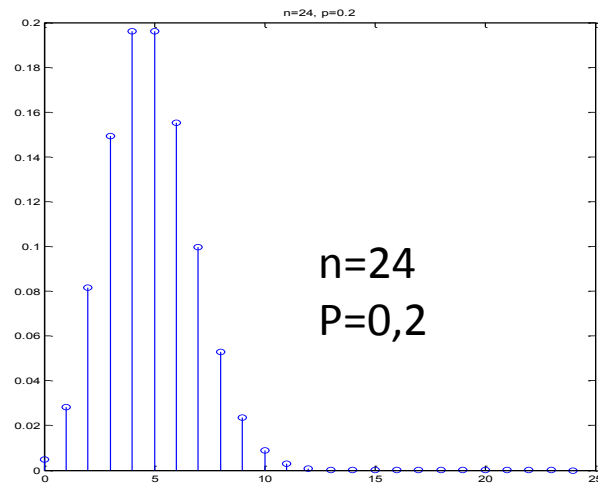
- $E(I) = \sum_i x_i p(x_i)$   
 $= 0 \times (1 - p) + 1 \times p$   
 $= p$
- $Var(I) = E(I^2) - (E(I))^2$
- $E(I^2) = 0^2 \times (1 - p) + 1^2 \times p = p$
- $Var(I) = p - p^2 = p(1 - p)$

# Variável Binomial

- Directamente relacionada com a Lei Binomial
- Seja  $X$  o número de vezes que um acontecimento  $A$  ocorre em  $n$  experiências de Bernoulli
  - Número de sucessos em  $n$  experiências (observações)
- $X = \sum_{j=1}^n I_j \quad \rightarrow \quad S_X = \{0, 1, 2, \dots, n\}$

# Variável Binomial

- $p_X(k) = \Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$



- $F_X(x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1 - p)^{n-k}$

# Variável Binomial – Média e Variância

- Fácil derivar usando o facto de termos  **$n$  variáveis de Bernoulli** independentes, que designamos por  $X_i$

- $$E(X) = E(\sum X_i) = \sum E(X_i) \quad pq ?$$
$$= p + p + \cdots + p = \mathbf{n} p$$

- De forma similar

$$\begin{aligned} \text{Var}(X) &= \\ \text{Var}(\sum X_i) &= \sum \text{Var}(X_i) = \cdots = \mathbf{n} p (1 - p) \end{aligned}$$

# Variável Binomial - Exemplos

- Têm distribuição Binomial, por exemplo:
  - Número de peças defeituosas num lote de um determinado tamanho (ex: 50 peças)
  - Número de respostas certas num exame de verdadeiro falso
  - Número de clientes que efectuaram compras em 100 que entraram numa loja

# V. Binomial - Áreas de aplicação

- A distribuição surge em muitas áreas científico-tecnológicas:
  - Engenharia de produção: Muitas vezes as medidas de controlo de qualidade são baseadas na distribuição binomial
    - O caso Binomial aplica-se a qualquer situação industrial em que o resultado é binário e os resultados de ensaio são independentes e com probabilidades constantes
  - Medicina: Por exemplo os resultados “cura” ou “não cura” são importantes na indústria farmacêutica
  - Indústria Militar: “acerta” “falha” é muitas vezes a interpretação do lançamento de um míssil ou de uma missão
  - **Informática**: “acerto” e “falha” é uma interpretação possível para detectores de SPAM, testes a métodos/funções de um programa, procura de informação na web ...



# Exemplo de aplicação I - Transmissão digital

- Um sistema de transmissão digital envia um pacote de 1 kByte através de canal com ruído sendo a probabilidade de erro de cada bit  $10^{-3}$  (ou seja 1 bit em cada mil).
- Considerando que os erros são independentes, determine:
  - Probabilidade de haver 1 erro ?
  - Probabilidade de haver erro ?

# Exemplo II – segurança de aviões

- Considere que um motor de avião pode falhar com probabilidade  $p$  e que as falhas são independentes entre motores.
- Se um avião se despenha se mais do que 50% dos motores falharem, é mais seguro voar num avião de 4 motores ou de 2 motores ?
- Faz parte do guião Prático 2

# Possível resolução

- O de 2 motores despenha-se se os 2 motores falharem. Qual a probabilidade ?
- $p_2 = p_X(2, n = 2) = \binom{2}{2} p^2 (1 - p)^{2-2} = p^2$
- O de 4 despenha-se se 3 ou 4 falharem. Qual a probabilidade ?
- $p_4 = p_X(3, n = 4) + p_X(4, n = 4)$
- $= \binom{4}{3} p^3 (1 - p)^{4-3} + \binom{4}{4} p^4 (1 - p)^{4-4}$
- $= 4 p^3 (1 - p) + p^4 = 4 p^3 - 3 p^4$
- $\frac{p_4}{p_2} = 4p - 3p^2 = p(4 - 3p)$

...

p	p2	p4	p4/p2		p	p2	p4	p4/p2
0,01	0,0001	0,00000397	0,0397		0,3	0,09	0,0837	0,93
0,02	0,0004	0,00003152	0,0788		0,31	0,0961	0,091458	0,9517
0,03	0,0009	0,00010557	0,1173		0,32	0,1024	0,099615	0,9728
0,04	0,0016	0,00024832	0,1552		0,33	0,1089	0,10817	0,9933
0,05	0,0025	0,00048125	0,1925		0,34	0,1156	0,117126	1,0132
0,06	0,0036	0,00082512	0,2292		0,35	0,1225	0,126481	1,0325
0,07	0,0049	0,00129997	0,2653		0,36	0,1296	0,136236	1,0512
0,08	0,0064	0,00192512	0,3008		0,37	0,1369	0,146387	1,0693
0,09	0,0081	0,00271917	0,3357		0,38	0,1444	0,156934	1,0868
0,1	0,01	0,0037	0,37		0,39	0,1521	0,167873	1,1037
0,2	0,04	0,0272	0,68					
0,3	0,09	0,0837	0,93					
0,4	0,16	0,1792	1,12					
0,5	0,25	0,3125	1,25					
0,6	0,36	0,4752	1,32					
0,7	0,49	0,6517	1,33					
0,8	0,64	0,8192	1,28					
0,9	0,81	0,9477	1,17					

# Exemplo de aplicação III

- According to the U.S. Census Bureau, approximately 6% of all workers in Jackson, Mississippi, are unemployed.
- In conducting a random telephone survey in Jackson, what is the probability of getting two or fewer unemployed workers in a sample of 20?
- De: Business Statistics, Ken Black, 6<sup>th</sup> ed, John Willey & Sons (cap 5)

# Resolução

- 6% desempregado  $\Rightarrow p=0,06$
- Tamanho da amostra é 20  $\Rightarrow n=20$
- 94% têm emprego  $\Rightarrow 1-p=0,94$
- $x$  é o número de sucessos que se pretende
- Qual é a probabilidade de termos 2 ou menos desempregados na amostra de 20 ?
- Neste tipo de problemas o importante e muitas vezes o mais difícil é identificar o  $p$ ,  $n$  e  $x$

# Resolução

$$n = 20$$

$$p = 0,06$$

$$q = 0,94$$

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= 0,2901 + 0,3703 + 0,2246 = 0,8850 \end{aligned}$$

---

$$P(X = 0) = \frac{20!}{0!(20-0)!} (0,06)^0 (0,94)^{20-0} = (1)(1)(0,2901) = 0,2901$$

$$P(X = 1) = \dots$$

$$P(X = 2) = \dots$$

# Binomial para valores de n elevados

- Consideremos o seguinte cenário:
- Num conjunto de programas a probabilidade de encontrar erros ao analisar um conjunto de **1.000 linhas** de código é **p** ( $p < 1$ )
  - Não nos interessa o número de erros, apenas se existe algum ou não
- Se o número total de linhas dos programas for **N x 1000** e os dividirmos em blocos de 1000 linhas a probabilidade de **k** blocos terem erros segue a distribuição **Binomial com parâmetros N e p**
- Se quisermos analisar células de 100 linhas e se considerarmos que a distribuição dos erros é uniforme a probabilidade desce para  $p/10$ 
  - Teríamos então uma Binomial com parâmetros  $10N$  e  $p/10$
  - Conceptualmente temos a forma de cálculo mas basta  $N$  ser um número moderado e  $10N$  começa a ser elevado e os **cálculos complicados [mesmo em computador]**
- Exemplo: Blocos de 100 linhas; 1000 blocos ;  $p=0,98/10$ 
  - Qual a probabilidade de blocos com erro ser inferior ou igual a 100 ?

*P(#blocos de 100 linhas com erro)*

$$= F_x(100) = \sum_{k=0}^{100} \binom{1000}{k} 0,098^k 0,902^{1000-k}$$



...

- As coisas ainda se complicam mais de reduzirmos mais o tamanho dos blocos (100 linhas, 10 linhas ..)
- Será que conseguimos arranjar maneira(s) eficiente(s) de calcular quando o tamanho é muito pequeno ?
  - No limite teremos apenas um bloco que vai ter ou não um erro
    - Número de blocos “infinitesimais” com erro = número de erros

# Distribuição de Poisson

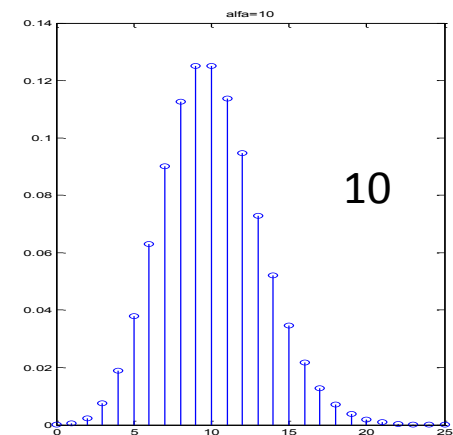
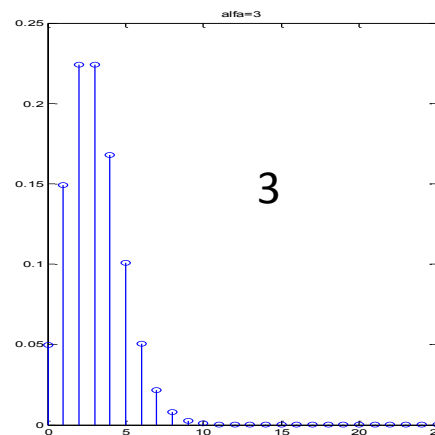
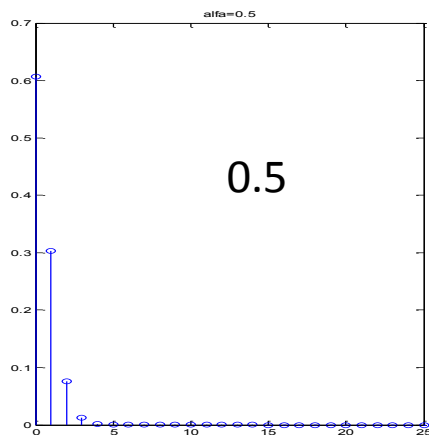
- Considere-se que temos uma variável Binomial,  $n$  cresce e  $p$  decresce por forma a  $np \rightarrow \lambda > 0$
- Para  $n$  grande pode fazer as seguintes aproximações  
 $p \cong \frac{\lambda}{n}$  e  $1 - p \cong 1 - \frac{\lambda}{n}$
- $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \rightarrow \text{Binomial}$
- $\lim_{n \rightarrow \infty} P(X = k) = \lim_{l \rightarrow \infty} \binom{n}{k} p^k (1 - p)^{n-k} =$
- $= \dots = \frac{\lambda^k e^{-\lambda}}{k!}$
- $p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$  é a f.p.m da distribuição de Poisson  
 $k=1,2, \dots$

# Demonstração

- $\lim_{n \rightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k}$
- $= \lim_{n \rightarrow \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$
- $= \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$
- $= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$
- $= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \dots \frac{n-k+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$
- $= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$
- como  $\lim_{n \rightarrow \infty} \left(1 - \frac{\alpha}{n}\right)^n = e^{-\alpha}$ ,  $\lim_{n \rightarrow \infty} \left(1 - \frac{j}{n}\right) = 1$  e o limite do produto é o produto dos limites temos:
- $= \frac{\lambda^k}{k!} \times 1 \times e^{-\lambda} \times 1$

# Distribuição para vários valores do parâmetro

- Uma distribuição do tipo
- $p_X(k) = \Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 1, 2, 3 \dots$   
É designada por distribuição de Poisson
- Tem apenas um parâmetro, o lambda



# Variável de Poisson – Média e Variância

- $E(X) = \lambda$
- $Var(X) = \lambda$

# Variável de Poisson

- A distribuição de Poisson foca-se apenas no número de ocorrências (discreto) num intervalo ou contínuo
  - Esta distribuição não tem um número de experiências ( $n$ ) como na Binomial
  - As ocorrências são independentes das outras ocorrências

# Aproximação de Poisson à distribuição Binomial

- Problemas envolvendo a distribuição Binomial em que **n é grande e o valor de p é pequeno**, gerando desta forma **eventos raros**, são os candidatos à utilização da distribuição de Poisson
- Regra prática (“rule of thumb”) :
  - Se  $n > 20$  e  $np \leq 7$  a aproximação de Poisson é suficiente próxima para ser usada em vez da Binomial

# Aproximação de Poisson à distribuição Binomial

- Procedimento para aproximar a Binomial por Poisson:
  - Calcular a média da Binomial  $\mu = np$
  - Como  $\mu$  é o valor esperado da Binomial , passa a ser o  $\lambda$  ( $=E(X)$ ) de Poisson
  - Usar a fórmula de Poisson (ou uma tabela)

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$



# Aplicação de Poisson

- As distribuições de Poisson surgem em experiências onde se verificam as seguintes propriedades:
  - O número de resultados que ocorrem num determinado intervalo de tempo ou região é independente do número que ocorre em qualquer outro intervalo temporal ou região espacial disjunta
  - A probabilidade que um resultado ocorra durante um intervalo ou região infinitesimal é proporcional ao comprimento do intervalo ou dimensão da região e não depende das ocorrências fora desse intervalo ou região
  - A probabilidade de haver mais que um resultado numa região infinitesimal é desprezável

# Exemplo de aplicação II

- Bank customers arrive randomly on weekday afternoons at an average of 3.2 customers every 4 minutes.
- What is the probability of having more than 7 customers in a 4-minute interval on a weekday afternoon?
- De: Business Statistics, Ken Black, 6<sup>th</sup> ed, John Willey & Sons (cap 5)

# Como resolver

- $\lambda = 3,2$  clientes em 4 minutos
- Consideremos o número de clientes (em intervalos de 4 minutos) como a variável aleatória  $X$
- Pretendemos  $P("X > 7 \text{ clientes /4 minutos}")$
- A solução requer que calculemos para  $x = 8, 9, 10, 11, 12, 13, 14, \dots$ . Até o valor ser aproximadamente zero
  - Ou usemos o complemento e calculemos  $x = 0, 1, 2, 3, 4, 5, 6, 7$
- Depois é só somar as probabilidades
- O resultado, 0,0169, mostra que é pouco provável que um banco que tem em média 3,2 clientes a cada 4 minutos receba 7 clientes num período de 4 minutos
  - TPC: confirmar este valor
- Este tipo de probabilidades são muito úteis para os gestores de Bancos (e outras instituições com atendimento ao público) dimensionarem o número de pessoas e postos de atendimento

# Variável Geométrica

- Seja  $X$  o número de vezes que é necessário repetir uma experiência de Bernoulli até obter um sucesso
  - Prob. Sucesso:  $p$       prob. Falha =  $1-p$
- $p_X(k) = p(1 - p)^{k-1}$ ,  $k = 1, 2, 3, \dots$   
Porque teremos  $k-1$  insucessos e depois sucesso
- $F_X(x) = \sum_{k=0}^{\lfloor x \rfloor} p(1 - p)^{k-1}$

# Exemplo de aplicação – Helpdesk UA

- Problema:
- Considere o serviço de atendimento via telefone do Helpdesk da UA.
- Supondo que a probabilidade de se conseguir contactar o suporte é  $p=0,1$  (só ao fim de 10 tentativas ☹).
- Determine a probabilidade de necessitar de menos de 3 chamadas até conseguir expor o seu problema ?
- Solução:
- $\Pr(n^{\circ} \text{ chamadas} < 3) = p(1 - p)^{1-1} + p(1 - p)^{2-1} = p(2 - p) = 0,19$

# Variável Geométrica – Média e Variância

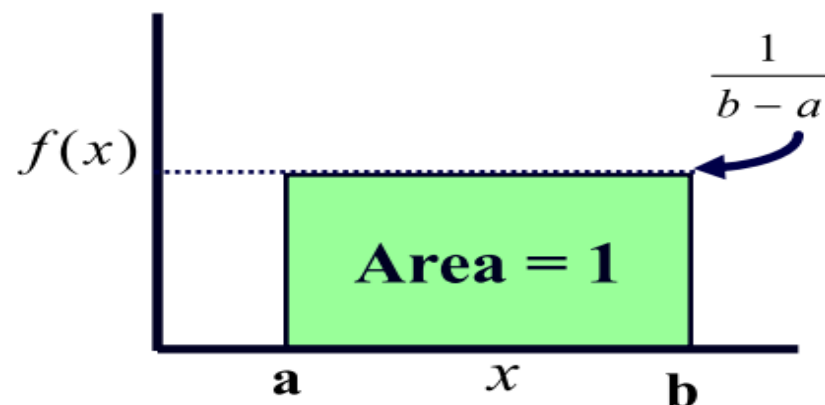
- $E(X) = \frac{1}{p}$
- $Var(X) = (1 - p)/p^2$

# Distribuições contínuas

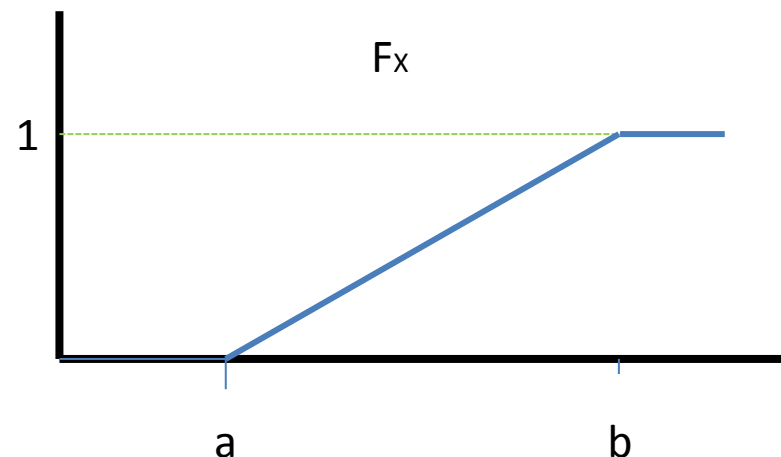
# Variável aleatória uniforme

- $U(a,b)$  é definida por:

- $$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{caso contrário} \end{cases}$$



- $$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$



- $$E(X) = \frac{a+b}{2}$$

- $$Var(X) = \frac{(b-a)^2}{12}$$

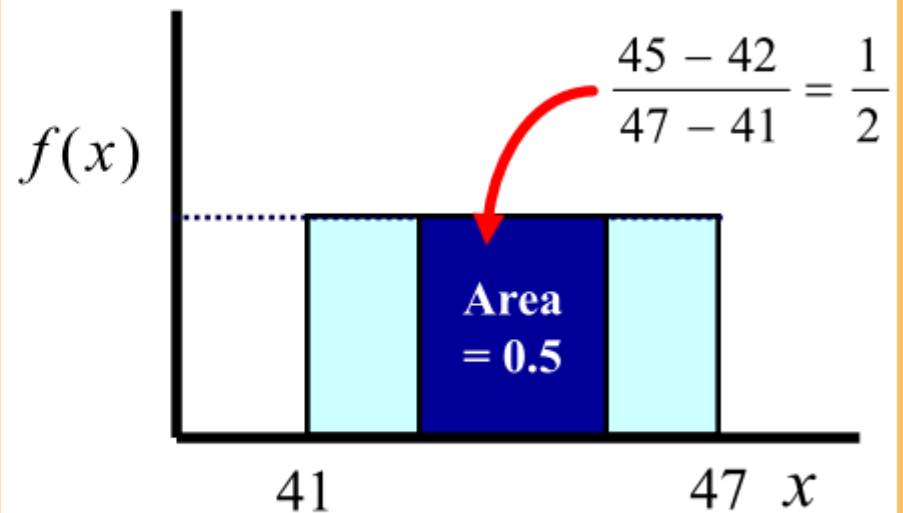


# Exemplo

- $P(42 \leq X \leq 45)$  com  $U(41, 47)$

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$

$$P(42 \leq X \leq 45) = \frac{45 - 42}{47 - 41} = \frac{1}{2}$$

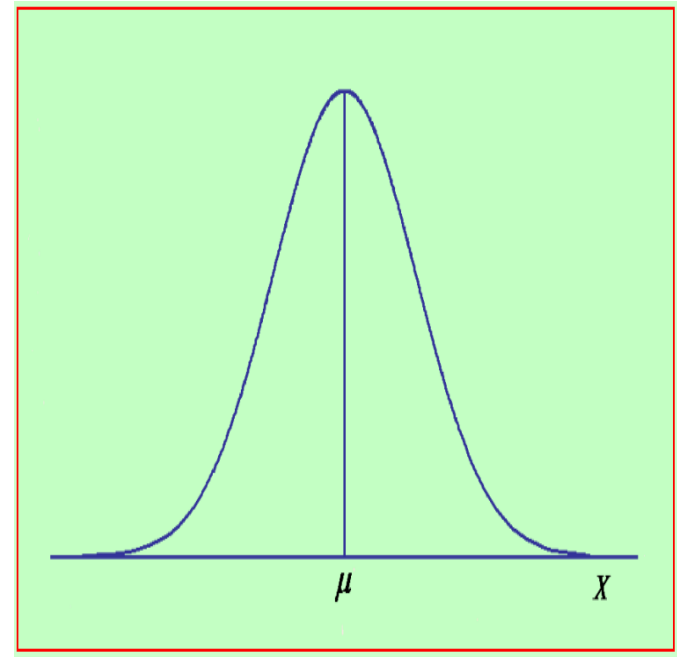


# Variável aleatória Normal (ou Gaussiana)

- Uma va diz-se normal ou Gaussiana se

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

- Frequentemente usa-se a notação  $N(m, \sigma^2)$
- Curva em forma de sino, simétrica em torno de  $m$  e com alargamento  $\sigma$



# Variável Normal

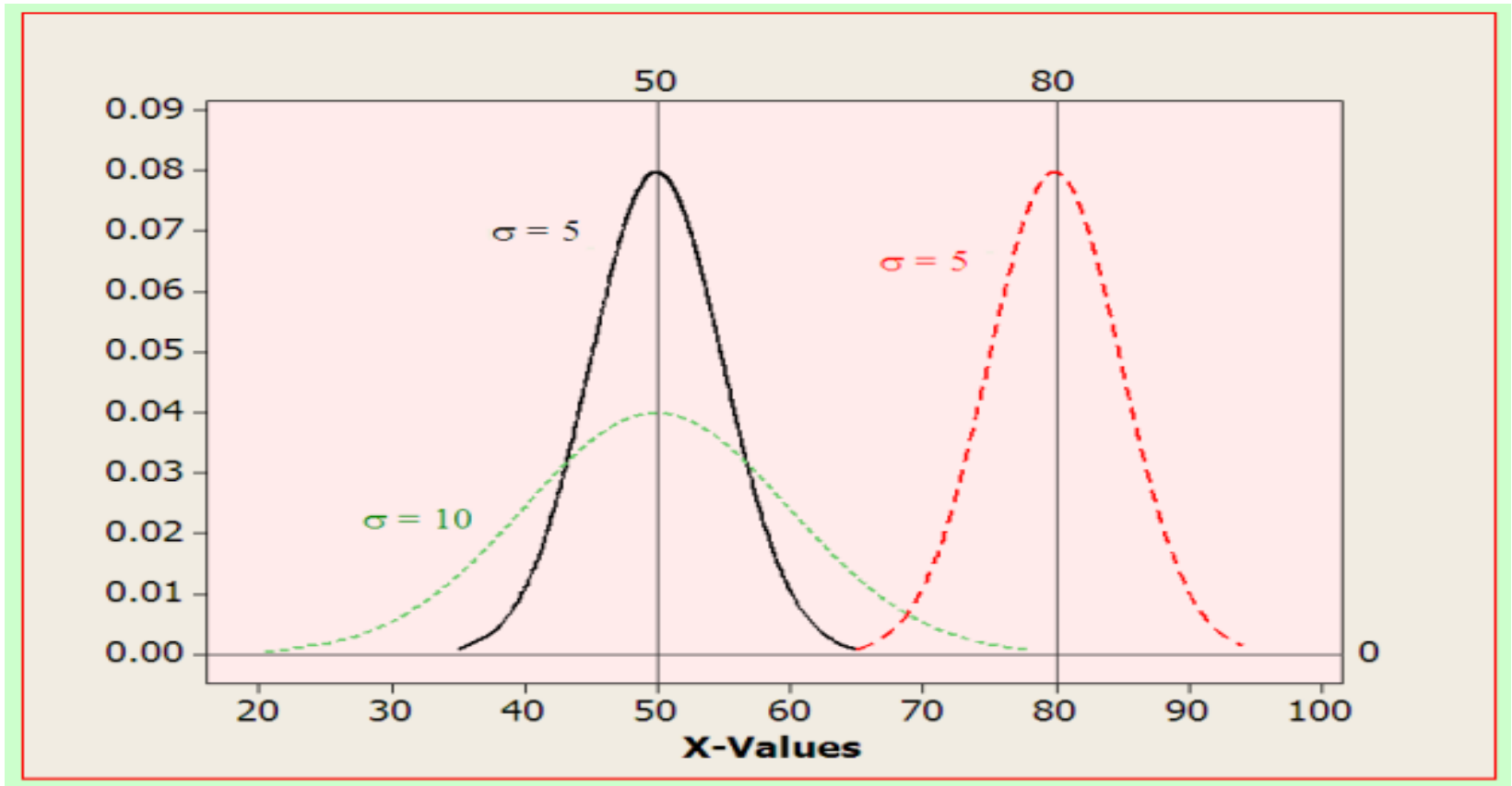
- Função de distribuição acumulada

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

- $E(X) = m$
- $Var(X) = \sigma^2$
- Nota: é muito comum utilizar-se  $\mu$  em vez de  $m$  para representar a média

# Família de curvas

- Variando os 2 parâmetros ...



# Gaussiana normalizada

- Como existe um número infinito de combinações para  $\mu$  e  $\sigma$  pode gerar-se uma família infinita de curvas
  - Sendo pouco prático lidar com esta situação, em especial antes da existência de computadores
- Foi desenvolvido um mecanismo pelo qual qualquer distribuição normal pode ser convertida num distribuição única,
  - a denominada Gaussiana normalizada
    - Com média e variância igual a 1
- A fórmula de conversão é:

$$Z = \frac{x - \mu}{\sigma}$$

Ou seja, subtrair a média e dividir pelo desvio padrão

# Função distribuição acumulada

- A fda da va Gaussiana normalizada,  $N(0,1)$  encontra-se frequentemente tabelada

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t)^2}{2}} dt$$

- Outras tabelas comuns são as de  $Q(x) = 1 - \Phi(x)$
- A fda de  $N(m, \sigma^2)$  pode ser expressa em termos de  $\Phi(x)$

$$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right)$$

# Distribuição Normal

- É muito provavelmente a mais conhecida e utilizada de todas as distribuições
- Adequa-se/ajusta-se a muitas características humanas
  - Altura, peso, velocidade, resultados de testes de inteligência, esperança de vida...
- Também se adequa a muitas outras coisas da natureza
  - Árvores, animais etc têm muitas características que seguem a distribuição normal
- Surge quando vários efeitos acumulados e independentes se sobrepõem

# Distribuição Normal e a Binomial

- Demonstra-se a Binomial de média  $m=np$  e  $\sigma = np(1 - p)$  com  $m$  não muito pequeno e  $n$  elevado pode ser aproximada por:

$$\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(k-m)^2}{2\sigma^2}}$$

- Ou seja a distribuição normal
  - Desde que usemos  $m=np$  e variância igual  $np(1-p)$



# Distribuição exponencial

- Surge frequentemente em problemas envolvendo filas de espera e fiabilidade
  - Exemplos:
    - Tempo até um computador avariar
    - Tempo entre chegada de utentes à urgência de um Hospital
- É uma va não negativa (prob. 0 para  $x < 0$ )
- Está relacionada com a distribuição (discreta) de Poisson
  - Se o número de acontecimentos que ocorrem num intervalo seguem distribuição de Poisson, o tempo entre eles segue exponencial

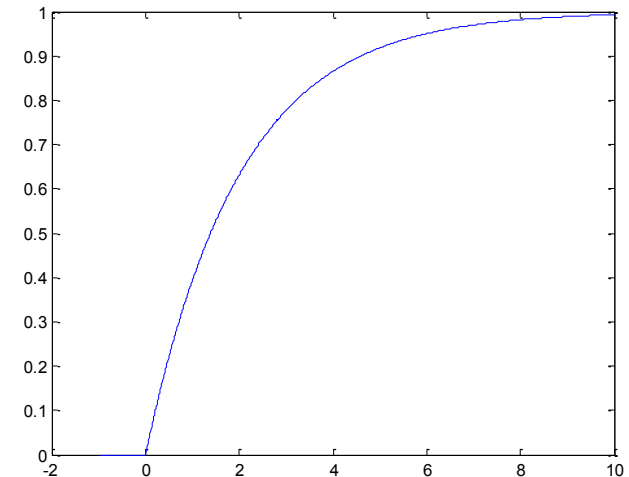
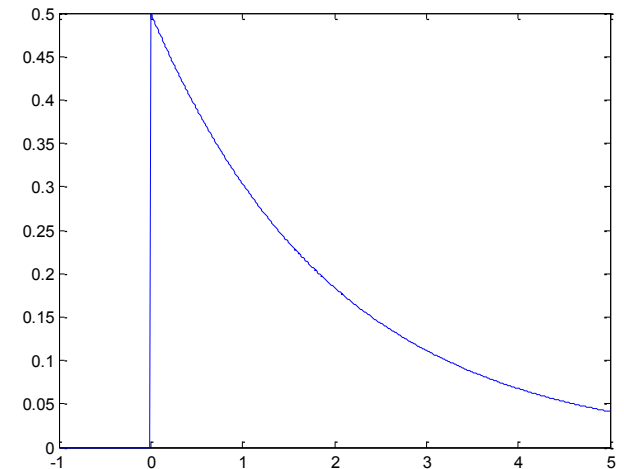
# Distribuição exponencial

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

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

- $E(x) = \frac{1}{\lambda}$

- $\text{Var}(x) = \frac{1}{\lambda^2}$



# Outras distribuições

# Distribuição dos primeiros dígitos

- Em 1881, um matemático e astrónomo americano, Simon Newcomb, percebeu que as primeiras páginas dos livros de logaritmos das bibliotecas estavam mais gastas que o resto, intrigado, investigou o assunto e...
- percebeu que em amostras aleatórias de dados reais os números de 1 à 9 no primeiro dígito de um número não obedeciam a distribuição mais intuitiva, de  $1/9$ , porém os números menores apareciam com maior frequência, o dígito 1 aparece quase  $1/3$  das vezes.
- Mais tarde, em 1938, o físico **Frank Benford** após uma investigação mais profunda chegou à mesma conclusão que Newcomb, indo mais além aplicando a fórmula em uma variedade de números

# Lei/Distribuição de Benford

- Função probabilidade →

$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(\frac{d+1}{d}\right) = \log_{10}\left(1 + \frac{1}{d}\right).$$

d	P(d)	Valor Relativo de P(d)
1	30.1%	
2	17.6%	
3	12.5%	
4	9.7%	
5	7.9%	
6	6.7%	
7	5.8%	
8	5.1%	
9	4.6%	

- A Lei/Distribuição de Benford, também conhecida como a "Lei dos Primeiros Dígitos", é uma ferramenta muito poderosa e muito simples que aponta suspeitas de fraudes, erros de digitação etc
- Mais info:
  - [https://pt.wikipedia.org/wiki/Lei\\_de\\_Benford](https://pt.wikipedia.org/wiki/Lei_de_Benford)
  - <http://gigamatematica.blogspot.pt/2011/07/lei-de-benford.html>

# Génese da Lei de Zipf

- George Kingsley **Zipf**, linguista da Universidade de Harvard, analisou a obra monumental de James Joyce, *Ulisses*, e contou as palavras distintas, ordenando-as por frequência.
- Verificou-se que:
  - a palavra mais comum surgia 8000 vezes;
  - a décima, 800 vezes;
  - a centésima, 80 vezes;
  - a milésima, 8 vezes.

# Lei de Zipf

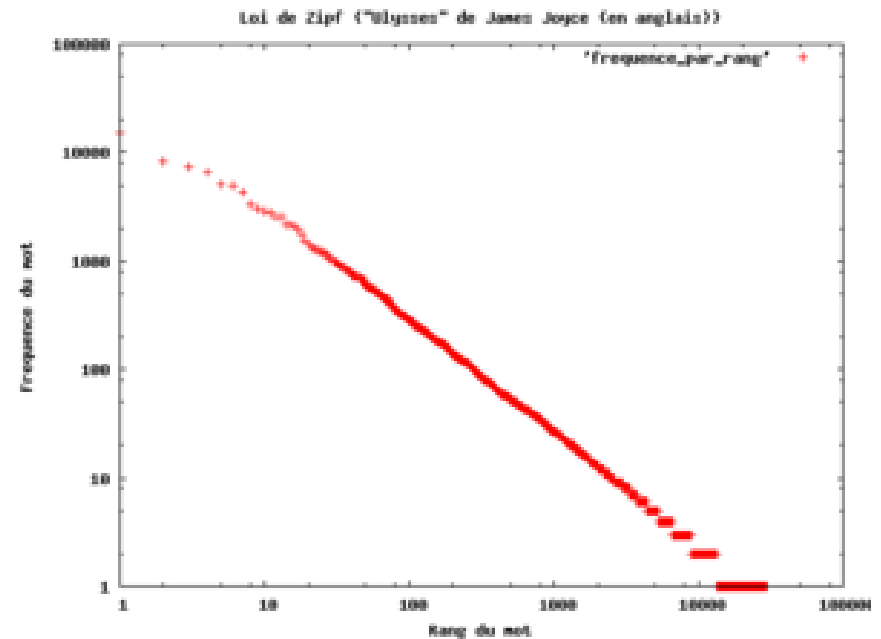
- **A Lei de Zipf**
  - formulada na década de 1940 por [Zipf](#), na sua obra *Human Behaviour and the Principle of Least-Effort* ("Comportamento Humano e o Princípio do Menor Esforço"),
- é uma lei empírica que rege a dimensão, importância ou frequência dos elementos de uma lista ordenada
- Trata-se de uma lei de potências sobre a distribuição de valores de acordo com o  $n^{\circ}$  de ordem numa lista.
  - Numa lista, o membro  $n$  teria uma relação de valor com o  $1^{\circ}$  da lista segundo  $1/n$ .
- Mais info: [https://pt.wikipedia.org/wiki/Lei\\_de\\_Zipf](https://pt.wikipedia.org/wiki/Lei_de_Zipf)

# Lei de Zipf

- A lei de Zipf prevê que num dado texto, a probabilidade de ocorrência  $p(n)$  de uma palavra esteja ligada à sua ordem  $n$  na ordem das frequências por uma lei da forma: onde  $K$  é uma constante:

$$p(n) = \frac{K}{n}$$

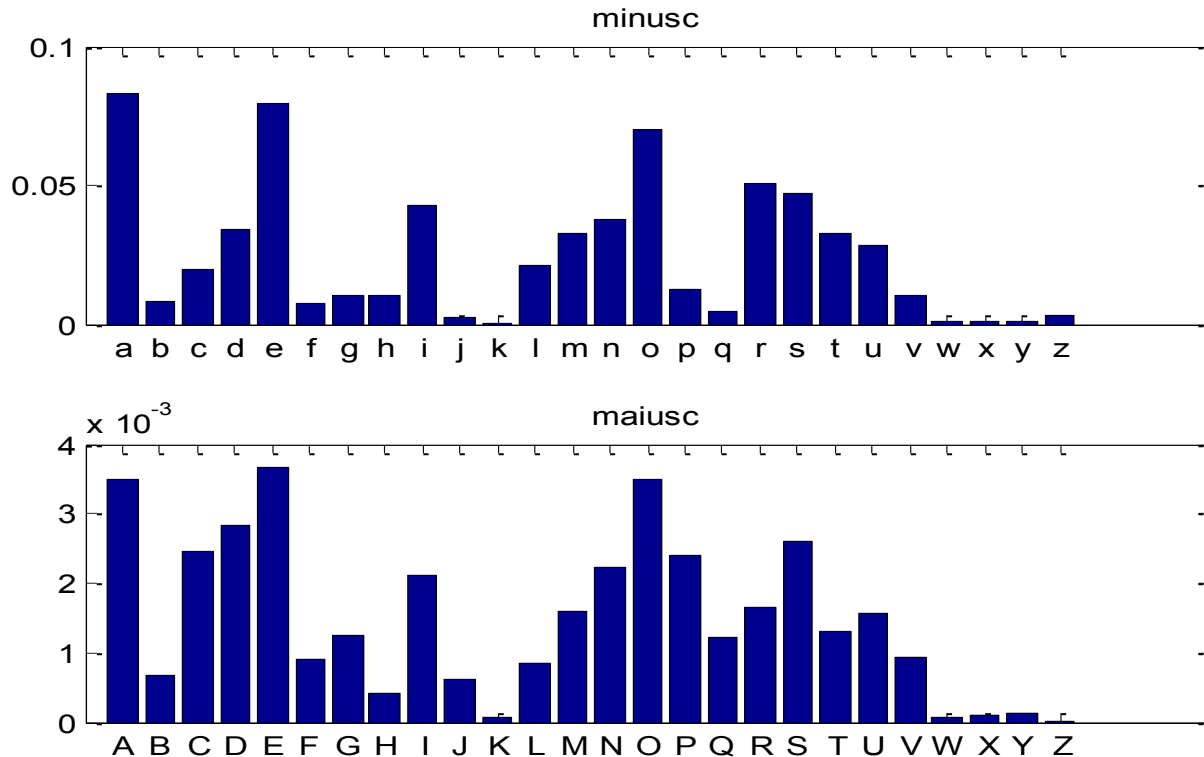
- $K$  depende da língua



Frequência das palavras em função da ordem na versão original de [Ulysses](#) de [James Joyce](#).  
De: Wikipedia



# Distribuição das letras em Português ...



<i>char</i>	<i>prob.</i>
<i>a</i>	0.083
<i>b</i>	0.008
<i>c</i>	0.020
<i>d</i>	0.034
<i>e</i>	0.079
<i>f</i>	0.007
<i>g</i>	0.010
<i>h</i>	0.011
<i>i</i>	0.043

.....

- Probabilidades estimadas usando o texto pg21209.txt do projecto Gutenberg

# Mais informação

- Material online
  - Slides relativos aos cap. 5 e 6 do livro “Business Statistics”, Ken Black, 4ed
    - [http://business.uni.edu/slides/ECON-1011\\_Luk/ch05.pdf](http://business.uni.edu/slides/ECON-1011_Luk/ch05.pdf)
    - [http://business.uni.edu/slides/ECON-1011\\_Luk/ch06.pdf](http://business.uni.edu/slides/ECON-1011_Luk/ch06.pdf)
  - Lectures:
    - <http://www.stat.berkeley.edu/~stark/SticiGui/Text/randomVariables.htm>
  - Wikipedia
- Capítulo 3 do livro “Probabilidades e Processos Estocásticos”, F. Vaz

Aula(s) seguintes

Como ter vectores com estas  
distribuições ?

Geração de números aleatórios

Geração de variáveis aleatórias

Simulação