

## Capítulo 5

# Distribuições Discretas e Contínuas Importantes

AMG, JFO (v10 – 2018)  
adaptado de: *Estatística*,  
Rui Campos Guimarães,  
José A. Sarsfield Cabral

Slide 5.-1

### Conteúdo

<b>5.1 Introdução</b>	<b>5-2</b>
<b>5.2 Distribuições Discretas</b>	<b>5-3</b>
5.2.1 Distribuição Binomial	5-3
5.2.2 Distribuição Geométrica	5-4
5.2.3 Distribuição Binomial Negativa	5-5
5.2.4 Distribuição Hipergeométrica	5-6
5.2.5 Distribuição de Poisson	5-7
5.2.6 Relações entre Distribuições Discretas	5-8
<b>5.3 Distribuições Contínuas</b>	<b>5-9</b>
5.3.1 Distribuição Uniforme	5-9
5.3.2 Distribuição Exponencial Negativa	5-9
5.3.3 Distribuição Normal	5-11
5.3.4 Distribuição Lognormal	5-13
5.3.5 Distribuições “Teóricas” (Qui-quadrado, $t$ de Student e $F$ )	5-14
<b>5.4 Utilização de Tabelas Estatísticas, Calculadoras e Folhas de Cálculo</b>	<b>5-15</b>
5.4.1 Tabelas de Distribuições Discretas	5-15
5.4.2 Tabelas de Distribuições Contínuas	5-16
5.4.3 Utilização de Calculadoras	5-21
5.4.4 Utilização de Folhas de Cálculo	5-22
<b>5.5 Comentários sobre Aproximações, Tabelas e Calculadoras</b>	<b>5-23</b>
<b>5.6 Exercícios</b>	<b>5-24</b>
<b>5.7 Anexos</b>	<b>5-27</b>
5.7.1 Distribuição Multinomial	5-27
5.7.2 Distribuição Hipergeométrica Generalizada	5-28
5.7.3 Distribuição Normal Multivariada	5-28

Slide 5.0

### Resultados de aprendizagem

- Dada a descrição de uma situação envolvendo experiências aleatórias, reconhecer e definir v.a.(s) com as seguintes distribuições: Binomial, (Geométrica,) Binomial Negativa, Hipergeométrica e de Poisson

- Enumerar as suposições em que se baseiam os modelos das seguintes distribuições discretas: Binomial, (Geométrica,) Binomial Negativa, Hipergeométrica e de Poisson
- Calcular probabilidades, com recurso a calculadoras e a tabelas, de v.a.(s) com as seguintes distribuições: Binomial, (Geométrica,) Binomial Negativa, Hipergeométrica e de Poisson
- Calcular valores esperados e variâncias de v.a.(s) com as seguintes distribuições: Binomial, (Geométrica,) Binomial Negativa, Hipergeométrica e de Poisson
- Usar a distribuição Binomial para aproximar a distribuição Hipergeométrica quando as condições de aproximação se verificarem
- Usar distribuição de Poisson para aproximar as distribuições Binomial e Hipergeométrica quando as condições de aproximação se verificarem
- Dada a descrição de uma situação envolvendo experiências aleatórias, reconhecer e definir v.a.(s) com as seguintes distribuições: Uniforme, Exponencial Negativa, Normal e Lognormal
- Enumerar as suposições em que se baseiam os modelos das seguintes distribuições contínuas: Uniforme, Exponencial Negativa, Normal, Lognormal, Qui-quadrado,  $t$  de Student e  $F$
- Calcular probabilidades de v.a.(s) com as seguintes distribuições: Uniforme e Exponencial Negativa
- Transformar v.a.(s) com distribuições Normais ou Lognormais na v.a. Normal Padronizada
- Dada uma probabilidade de uma distribuição Normal padronizada, Qui-quadrado,  $t$  de Student ou  $F$ , obter o valor da v.a. que origina essa probabilidade pela consulta da tabela da respectiva distribuição ou utilizando uma calculadora
- Dada uma desigualdade envolvendo uma v.a. de uma distribuição Normal padronizada, Qui-quadrado,  $t$  de Student ou  $F$ , obter o valor da probabilidade associada consultando tabelas de distribuições ou utilizando uma calculadora
- Calcular valores esperados e variâncias de v.a.(s) com as seguintes distribuições: Uniforme, Exponencial Negativa, Normal e Lognormal
- Identificar e definir distribuições de v.a. que sejam resultantes de combinações lineares de v.a. Normais
- Usar a distribuição Normal para aproximar a distribuição Binomial quando as condições de aproximação se verificarem

Slide 5.1

Slide 5.2

Slide 5.3

## 5.1 Introdução

### Distribuições de probabilidade

⇒ Modelos que descrevem populações reais

(podemos calcular tudo sobre uma população: probabilidades, valores esperados ( $\mu$ ), variâncias ( $\sigma^2$ ), ...)

Exemplo: Modelo Binomial

$$Y \sim B(N, p)$$

$$p(y) = \binom{N}{y} \cdot p^y \cdot q^{N-y}$$

$$\mu = E(Y) = N \cdot p$$

$$\sigma^2 = Var(Y) = N \cdot p \cdot q$$

### Distribuições discretas a estudar

- Binomial
- Geométrica
- Binomial Negativa
- Hipergeométrica
- de Poisson

Porquê estas distribuições discretas?

- pela sua simplicidade
- por modelarem fenómenos aleatórios que ocorrem frequentemente

Slide 5.4

### Distribuições contínuas univariadas a estudar:

- |   |   |   |
|---|---|---|
| <ul style="list-style-type: none"> <li>• Uniforme</li> <li>• Exponencial negativa</li> <li>• Normal</li> <li>• Lognormal</li> </ul> | } | Simplicidade e capacidade de modelizar fenómenos aleatórios muito frequentes  |
| <ul style="list-style-type: none"> <li>• Qui-quadrado</li> <li>• <math>t</math> de Student</li> <li>• <math>F</math></li> </ul>     | } | Indispensáveis à resolução de muitos problemas de inferência estatística (conhecidas como distribuições “teóricas”) |

Slide 5.5

## 5.2 Distribuições Discretas

### 5.2.1 Distribuição Binomial

Lançamento de uma moeda H-T ao ar 100 vezes e contagem do número de Hs

O que caracteriza esta situação?

- cada lançamento só tem 2 resultados possíveis (*H* ou *T*)
- probabilidade mantém-se constante em todos os lançamentos ( $p=0.5$ )
- ou seja, há independência entre os sucessivos lançamentos
- número de repetições da experiência está predeterminado ( $N=100$ )

#### Experiência de Bernoulli:

1. A cada experiência corresponde apenas um de dois resultados possíveis: *sucesso* ou *insucesso*
2. A probabilidade de sucesso ( $p$ ), e logo a de insucesso ( $q = 1 - p$ ), *mantém-se inalterada* ao longo das experiências
3. Os resultados associados a cada experiência são *independentes*

⇒ *Distribuição Binomial* está associada à contagem de sucessos em  $N$  experiências de Bernoulli

$Y$ : variável aleatória que conta o número de sucessos ocorridos no decurso de  $N$  experiências de Bernoulli

⇒ nestas condições  $Y$  segue uma *distribuição Binomial*

$$Y \sim B(N, p)$$

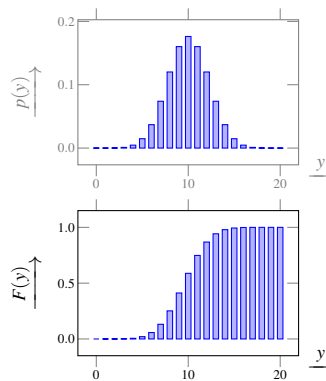
- Função de probabilidade:

$$p(y) = \binom{N}{y} \cdot p^y \cdot q^{N-y}$$

- Valor esperado e variância:

$$\mu = E(Y) = N \cdot p$$

$$\sigma^2 = \text{Var}(Y) = N \cdot p \cdot q$$



#### Notas

- v.a. que conta os insucessos é também Binomial,  $B(N, q)$
- dado que  $q < 1$  então  $\sigma^2 < \mu$

#### Exemplo

De uma linha de produção em série, com uma taxa constante de 10% de peças defeituosas, retiram-se 5 peças (uma por hora)

- a) Qual a probabilidade de se obterem duas peças defeituosas?
- b) Qual o valor esperado e a variância do número de peças defeituosas?

#### Resolução:

$Y$ : Número de peças defeituosas em 5 inspeções

$$Y \sim B(N = 5, p = 0.10)$$

$$p(y = k) = \binom{N}{k} \cdot p^k \cdot q^{N-k}$$

$$a) P(Y = 2) = p(2) = \binom{5}{2} \cdot 0.1^2 \cdot 0.9^{5-2} = 7.29\%$$

$$b) \mu = N \cdot p = 5 \cdot 0.1 = 0.5$$

$$\sigma^2 = N \cdot p \cdot q = 5 \cdot 0.1 \cdot 0.9 = 0.45$$

$$\leftarrow \mu = \sum_{i=1}^5 y_i \cdot p(y_i)$$

$$\leftarrow \sigma^2 = \sum_{i=1}^5 (y_i - \mu)^2 \cdot p(y_i)$$

#### Notas

- DDDDD, DDDDD, DDDDD, DDDDD, DDDDD, DDDDD, DDDDD, DDDDD, DDDDD, DDDDD ⇒  $\binom{5}{2} = 10$  sequências diferentes com 2 Ds e (5-2)  $\bar{D}$ s.
- Porque é que se podem multiplicar as várias probabilidades?

## 5.2.2 Distribuição Geométrica

Lançamento de uma moeda H-T até sair H pela primeira vez

O que caracteriza esta situação?

- cada lançamento só tem 2 resultados possíveis ( $H$  ou  $T$ )
  - probabilidade mantém-se constante em todos os lançamentos ( $p=0.5$ )
  - continua a haver independência entre os sucessivos lançamentos
- ⇒ número de repetições da experiência é desconhecido ( $N=?$ )
- ⇒ pára-se quando sair a primeira vez H num lançamento ( $1^{\circ} H$ )

Continuam a ser *experiências de Bernoulli*, mas agora existe um *critério de paragem*: ocorrer o primeiro sucesso

⇒ *Distribuição Geométrica* conta o número de insucessos em experiências de Bernoulli até que ocorra o primeiro sucesso

Slide 5.9

$Y$ : v.a. que conta o número insucessos até ocorrer o primeiro sucesso

⇒ nestas condições  $Y$  segue uma *distribuição Geométrica*

$$Y \sim G(p)$$

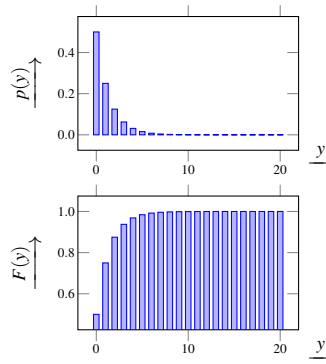
- Função de probabilidade:

$$p(y) = p \cdot q^y$$

- Valor esperado e variância:

$$\mu = E(Y) = \frac{q}{p}$$

$$\sigma^2 = \text{Var}(Y) = \frac{q}{p^2}$$



### Notas

- Dist. Geométrica *não tem memória* ( $P(Y < (k+l) | Y > l) = P(Y < k)$ )
- Pode ser necessária uma sequência de infinita de experiências de Bernoulli
- Número de insucessos =  $Y$  (Número de experiências =  $Y + 1$ )

Slide 5.10

### Exemplo

De uma linha de produção em série, com uma taxa constante de 10% de peças defeituosas, extraem-se peças (uma por hora) até sair a primeira defeituosa

- Qual a probabilidade de se obter a primeira peça defeituosa na terceira extracção?
- Qual o valor esperado e a variância do número de peças extraídas?

### Resolução:

$Y$ : N° de peças não defeituosas até sair a primeira peça defeituosa

$$Y \sim G(p = 0.10)$$

$$p(y = k) = p \cdot q^k$$

$$a) P(Y = 2) = p(2) = 0.1 \cdot 0.9^2 = 8.1\%$$

$$b) \mu = \frac{q}{p} = \frac{0.9}{0.1} = 9 \quad e \quad \sigma^2 = \frac{q}{p^2} = \frac{0.9}{0.1^2} = 90$$

### Nota

A sequência é fixa, a última peça é sempre um  $D$  e as anteriores são todas  $\bar{D} \Rightarrow \bar{D}\bar{D}\bar{D}$ , ou seja: só há uma sequência

Slide 5.11

## 5.2.3 Distribuição Binomial Negativa

Lançamento de uma moeda H-T até sair H pela segunda vez

O que caracteriza esta situação?

- cada lançamento só tem 2 resultados possíveis ( $H$  ou  $T$ )
  - probabilidade mantém-se constante em todos os lançamentos ( $p=0.5$ )
  - continua a haver independência entre os sucessivos lançamentos
  - número de repetições da experiência é desconhecido ( $N=?$ )
- ⇒ pára-se quando sair H, pela segunda vez, num lançamento ( $2^o H$ )

Continuam a ser *experiências de Bernoulli*, mas agora o *critério de paragem* é diferente: ocorrer o segundo sucesso

⇒ *Distribuição Binomial Negativa* conta o número de insucessos em experiências de Bernoulli até que ocorra o segundo sucesso (ou no caso geral, que ocorra o  $r$ -ésimo sucesso)

Slide 5.12

$Y$ : v.a. que conta o número de insucessos até ocorrer o  $r$ -ésimo sucesso

⇒ nestas condições  $Y$  segue uma *distribuição Binomial Negativa*

$$Y \sim BN(r, p)$$

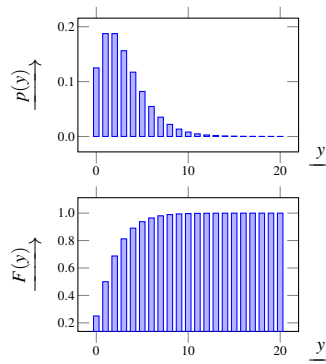
- Função de probabilidade:

$$p(y) = \binom{y+r-1}{y} \cdot p^r \cdot q^y$$

- Valor esperado e variância:

$$\mu = E(Y) = \frac{r \cdot q}{p}$$

$$\sigma^2 = Var(Y) = \frac{r \cdot q}{p^2}$$



### Notas

- Distribuição Geométrica é um caso particular da Bin. Negativa ( $r = 1$ )
- A variável que conta o número de sucessos até ocorrerem  $r$  insucessos também segue uma distribuição Binomial Negativa

Slide 5.13

### Exemplo

De uma linha de produção em série, com uma taxa constante de 10% de peças defeituosas, extraem-se peças (uma por hora) até sair a segunda defeituosa

- Qual a probabilidade de a extracção da segunda peça defeituosa ocorrer à sétima extracção?
- Qual o valor esperado e a variância do número de insucessos até ocorrer o segundo sucesso?

### Resolução:

$Y$ : N° de peças não defeituosas até sair a  $r$ -ésima peça defeituosa

$$Y \sim BN(r = 2, p = 0.10)$$

$$p(y) = \binom{y+r-1}{y} \cdot p^r \cdot q^y$$

$$a) P(Y = 5) = p(5) = \binom{5+2-1}{5} \cdot 0.1^2 \cdot 0.9^5 = 3.5\%$$

$$b) \mu = \frac{r \cdot q}{p} = \frac{2 \times 0.9}{0.1} = 18 \quad e \quad \sigma^2 = \frac{r \cdot q}{p^2} = \frac{0.9}{0.1^2} = 180$$

### Nota

$\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}, \bar{D}\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}, \dots, D\bar{D}\bar{D}\bar{D}\bar{D}\bar{D}\bar{D} \Rightarrow \binom{6}{5}$  sequências com o último  $D$  fixo,  $(2-1)$   $D$ s livres e 5  $\bar{D}$ s livres, ou seja  $\binom{6}{5} = 6$  sequências diferentes

Slide 5.14

## 5.2.4 Distribuição Hipergeométrica

Extracção em bloco de  $N$  peças de uma população finita constituída por  $M$  peças, defeituosas ( $D$ ) e não defeituosas ( $\bar{D}$ ), em proporções  $p$  e  $q$

O que caracteriza esta situação?

- cada extracção só tem 2 resultados possíveis ( $D$  ou  $\bar{D}$ )
- ⇒ probabilidade não se mantém constante em todas as extracções
- ⇒ não há independência entre as sucessivas extracções
- ⇒ número de repetições da experiência está predeterminado ( $N$ )

Deixam de ser *experiências de Bernoulli*

⇒ *Distribuição Hipergeométrica* está associada à contagem do número de sucessos (ou insucessos) em  $N$  experiências que não sejam independentes entre si

Slide 5.15

$Y$ : v.a. que conta o número de elementos de um determinado tipo, existente na proporção  $p$ , entre os  $N$  elementos extraídos em bloco de uma população finita

⇒ nestas condições  $Y$  segue uma *distribuição Hipergeométrica*

$$Y \sim H(M \cdot p, M \cdot q, N)$$

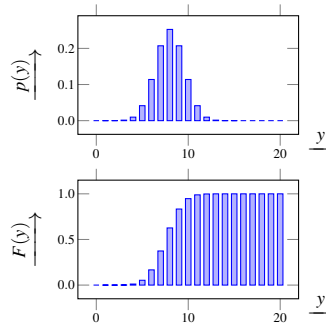
- Função de probabilidade:

$$p(y) = \frac{\binom{M \cdot p}{y} \cdot \binom{M \cdot q}{N-y}}{\binom{M}{N}}$$

- Valor esperado e variância:

$$\mu = E(Y) = N \cdot p$$

$$\sigma^2 = \text{Var}(Y) = N \cdot p \cdot q \cdot \frac{M-N}{M-1}$$



### Notas

- extracção em bloco é equivalente à extracção sem reposição
- semelhanças com a distribuição Binomial...

Slide 5.16

### Exemplo

De uma linha de produção em série, com uma taxa constante de 10% de peças defeituosas, extraem-se 5 peças em bloco (ou seja sem reposição) de entre as 50 produzidas nesse hora

- Qual a probabilidade de se extraírem 2 peças defeituosas?
- Qual o valor esperado e a variância do número de peças defeituosas extraídas?

### Resolução:

$Y$ :  $N^\circ$  de peças defeituosas quando se extraem 5 peças em bloco ( $p = 0.10$ )

$$Y \sim H(M \cdot p = 5, M \cdot q = 45, N = 5)$$

$$p(y) = \frac{\binom{M \cdot p}{y} \cdot \binom{M \cdot q}{N-y}}{\binom{M}{N}}$$

$$a) P(Y = 2) = p(2) = \frac{\binom{5}{2} \cdot \binom{45}{3}}{\binom{50}{5}} = 6.7\%$$

$$b) \mu = N \cdot p = 5 \cdot 0.1 = 0.5$$

$$\sigma^2 = N \cdot p \cdot q \cdot \frac{M-N}{M-1} = 5 \cdot 0.1 \cdot 0.9 \cdot \frac{50-5}{50-1} = 0.413$$

Slide 5.17

## 5.2.5 Distribuição de Poisson

Contagem do número de pessoas que chegam a uma fila de espera de um banco, em que os clientes vão chegando a uma taxa de 3 clientes/hora

O que caracteriza esta situação?

- ocorrência de fenômenos aleatórios ao longo do tempo (ou espaço)
- taxa de ocorrência ( $\lambda$ ) é constante ao longo do tempo (ou espaço)
- chegadas ocorrem uma a uma e nunca em grupos
- trata-se da contagem de um fenômeno “raro” (o “normal” é não chegar nenhum cliente)

⇒ *Distribuição de Poisson* está associada à contagem do número de ocorrências de fenômenos raros que se repetem periodicamente ao longo do tempo (ou espaço)

Slide 5.18

Se o tempo (ou o espaço) se encontrar dividido em pequenos intervalos  $\Delta t$ , a v. a.  $Y$  que conta o número de ocorrências por unidade de tempo (ou espaço) segue uma distribuição de Poisson se se verificarem as seguintes condições:

1. Os números de ocorrências em diferentes intervalos  $\Delta t$  serem independentes entre si
2. A distribuição do número de ocorrências em cada intervalo é a mesma para todos os intervalos
3. A probabilidade  $\Delta P_1$  de se registrar uma ocorrência num intervalo qualquer de dimensão  $\Delta t$  é aproximadamente proporcional ao tamanho do intervalo:  $\Delta P_1 \approx \lambda \Delta t$ ,  $\lambda > 0$ . Quando  $\Delta t$  tende para zero:

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta P_1}{\Delta t} = \frac{dP_1}{dt} = \lambda$$

4. A probabilidade de se registarem duas ou mais ocorrências num intervalo qualquer ( $\Delta P_n$  com  $n \geq 2$ ) é desprezável quando comparada com  $\Delta P_1$ . Quando  $\Delta t$  tende para zero:

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta P_n}{\Delta t} = \frac{dP_n}{dt} = 0 \quad (n \geq 2)$$

Slide 5.19

⇒ nas condições anteriores  $Y$  segue uma distribuição de Poisson

$$Y \sim \text{Poisson}(\lambda)$$

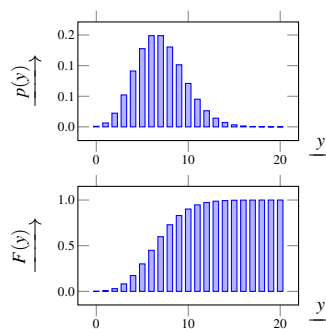
- Função de probabilidade:

$$p(y) = e^{-\lambda} \cdot \frac{\lambda^y}{y!}$$

- Valor esperado e variância:

$$\mu = E(Y) = \lambda$$

$$\sigma^2 = \text{Var}(Y) = \lambda$$



### Notas

- $\lambda$  – Taxa média de ocorrências (número médio de ocorrências por unidade de tempo ou espaço)
- relação de recorrência:  $P(X = r) = (\lambda/r) \cdot P(X = r - 1)$
- $Y$  pode tomar valores inteiros entre 0 e  $+\infty$
- semelhanças com a distribuição Binomial ...

Slide 5.20

### Exemplo

Considere uma máquina de produção de papel. A máquina produz quilómetros de papel em contínuo que vai sendo enrolado em rolos com 2 metros de largura. O número de defeitos por  $m^2$  segue uma distribuição de Poisson de média 0.01 defeitos/ $m^2$

- Qual a probabilidade de não se encontrarem defeitos em 100 metros produzidos?
- Qual a probabilidade de ocorrerem 3 ou menos defeitos em 500 metros produzidos?

### Resolução:

$Y$ : N° de defeitos por  $m^2$

$$Y \sim \text{Poisson}(\lambda = 0.01) \quad p(y) = e^{-\lambda} \cdot \frac{\lambda^y}{y!}$$

$$a) \lambda = 0.01 \text{ defeitos}/m^2 \rightarrow \lambda' = 2 \times 100 \times 0.01 = 2 \text{ defeitos}/200m^2$$

$$P(Y = 0) = p(0) = \frac{e^{-2} \times 2^0}{0!} = 13.53\%$$

$$b) \lambda = 0.01 \text{ defeitos}/m^2 \rightarrow \lambda'' = 2 \times 500 \times 0.01 = 10 \text{ defeitos}/1000m^2$$

$$P(Y \leq 3) = p(0) + p(1) + p(2) + p(3) = 1.03\%$$

Slide 5.21

## 5.2.6 Relações entre Distribuições Discretas

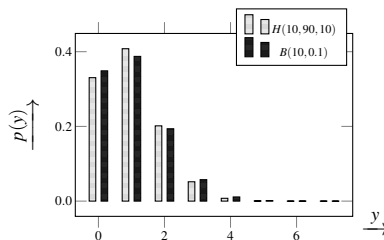
### Relação entre as Distribuições Binomial e Hipergeométrica

Relação entre  $B(N, p)$  e  $H(M \cdot p, M \cdot q, N)$

- têm o mesmo valor esperado ( $N \cdot p$ )
- variâncias distinguem-se pelo factor  $(M - N)/(M - 1)$
- se  $M \gg N$  podemos aproximar  $H(M \cdot p, M \cdot q, N)$  por  $B(N, p)$
- quando  $M \gg N$  a extracção ser em bloco não afecta significativamente as probabilidades de sucesso das sucessivas experiências ( $\Rightarrow$  independência entre as sucessivas experiências)

$\Rightarrow$  Regra prática: quando  $M \geq 10 \times N$  temos que  $(M - N)/(M - 1) \approx 1$   
 $\rightarrow H(M \cdot p, M \cdot q, N) \approx B(N, p)$

Y	$H(10, 90, 10)$	$B(10, 0.1)$
0	0.330	0.349
1	0.408	0.387
2	0.202	0.194
3	0.052	0.057
4	0.008	0.011
5	0.001	0.001
$\mu$	1.000	1.000
$\sigma^2$	0.818	0.900



Slide 5.22

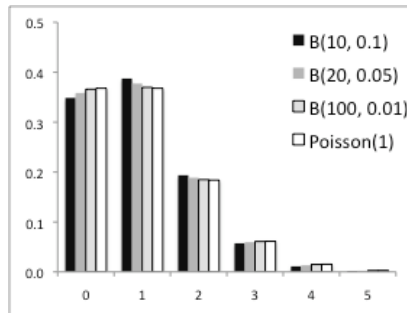
### Relação entre a Distribuição de Poisson e as Distribuições Binomial e Hipergeométrica

Relação entre  $\text{Poisson}(\lambda)$  e  $B(N, p)$

$$\lim_{\substack{N \rightarrow \infty \\ p \rightarrow 0 \\ Np = \lambda}} B(N, p) = \text{Poisson}(\lambda)$$

Regra prática:  $N \geq 20 \rightarrow B(N, p) \approx \text{Poisson}(\lambda = N \cdot p)$

Y	$B(10, 0.1)$	$B(20, 0.05)$	$B(100, 0.01)$	$\text{Poisson}(1)$
0	0.349	0.358	0.366	0.368
1	0.387	0.377	0.370	0.368
2	0.194	0.189	0.185	0.184
3	0.057	0.060	0.061	0.061
4	0.011	0.013	0.015	0.015
5	0.001	0.002	0.003	0.003
$\mu$	1.000	1.000	1.000	1.000
$\sigma^2$	0.900	0.950	0.990	1.000





### Relação entre $Poisson(\lambda)$ e $H(M \cdot p, M \cdot q, N)$

- Distr. Binomial pode aproximar a distr. Hipergeométrica
  - Distr. de Poisson pode aproximar a distr. Binomial
- $\Rightarrow$  Distribuição de Poisson pode aproximar a distribuição Hipergeométrica

Slide 5.23

## 5.3 Distribuições Contínuas

### 5.3.1 Distribuição Uniforme

$X$  segue uma distribuição Uniforme se a sua função densidade de probabilidade estiver uniformemente distribuída (constante) num intervalo finito  $[a, b]$

$$X \sim U(a, b)$$

- Função densidade de probabilidade

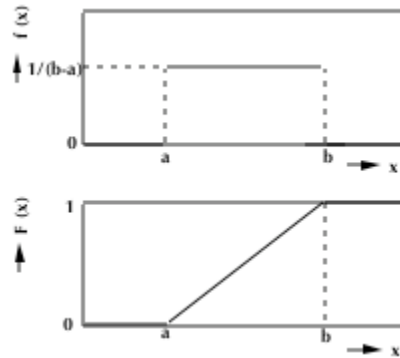
$$f(x) = \begin{cases} k, & x \in [a, b] \\ 0, & x \notin ]a, b[ \end{cases}$$

- Função distribuição

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

- Valor esperado e variância

$$\mu = \frac{a+b}{2} \quad \sigma^2 = \frac{(b-a)^2}{12}$$



Nota – Para que  $f_X(x)$  seja de facto uma função densidade de probabilidade é necessário que  $k = \frac{1}{b-a}$

Slide 5.24

#### Exemplo

O tempo decorrido, em minutos, entre o pedido de uma pizza e a sua entrega ao domicílio varia uniformemente no intervalo  $[15, 35]$  minutos.

- Indique a função densidade de probabilidade,  $f_X(x)$ , e a função distribuição,  $F_X(x)$ , da v.a.  $X$ .
- Calcule a probabilidade de a pizza ser entregue entre os 20 e os 25 minutos.
- Calcule o valor esperado e a variância do tempo de entrega de uma pizza.

**Resolução** –  $X$ : tempo decorrido entre o pedido e a entrega da pizza (min)

$$X \sim U(a = 15, b = 35)$$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & , x \in [a, b] \\ 0 & , x \notin ]a, b[ \end{cases} \quad F_X(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , x > b \end{cases}$$

$$a) \quad f_X(x) = \begin{cases} \frac{1}{20} & , x \in [15, 35] \\ 0 & , x \notin ]15, 35[ \end{cases} \quad F_X(x) = \begin{cases} 0 & , x < 15 \\ \frac{x-15}{20} & , 15 \leq x \leq 35 \\ 1 & , x > 35 \end{cases}$$

$$b) \quad P(20 < X < 25) = F_X(25) - F_X(20) = \frac{25-15}{35-15} - \frac{20-15}{35-15} = 0.25$$

$$c) \quad \mu_X = E(X) = \frac{15+35}{2} = 25 \quad \sigma_X^2 = Var(X) = \frac{(35-15)^2}{12} = 33.33$$

Slide 5.25

### 5.3.2 Distribuição Exponencial Negativa

Relação entre a dist. Exponencial Negativa (EN) e a dist. de Poisson:

$Y$  : número de ocorrências por unidade de tempo (ou espaço)

$X$  : tempo (ou espaço) entre ocorrências sucessivas

$$Y \sim Poisson(\lambda) \quad \longleftrightarrow \quad X \sim EN(\lambda)$$

*Exemplo:* Sejam  $Y$  o número de telefonemas e  $X$  o tempo entre dois telefonemas recebidos ( $X > 0$ ). A função distribuição de  $X$  vem:

$$F_X(x) = \text{probabilidade de se receber pelo menos um telefonema no intervalo } [0, x] = \\ = P(Y \geq 1, \text{ em } [0, x]) = 1 - P(Y = 0, \text{ em } [0, x]) = 1 - e^{-\lambda x} \cdot \frac{(\lambda x)^y}{y!} = 1 - e^{-\lambda x}$$

Vindo a função densidade de probabilidade, para  $X > 0$ :

$$f_X(x) = \frac{dF(x)}{dx} = \lambda \cdot e^{-\lambda x}$$

*Nota* – Função de probabilidade da distribuição de Poisson expressa em função de um intervalo de amplitude  $\lambda x$  em vez de um intervalo unitário (apenas  $\lambda$ )

Slide 5.26

$X$ : tempo (ou espaço) entre ocorrências sucessivas de um processo de Poisson com parâmetro  $\lambda$ ,  $X$  segue uma dist. Exponencial Negativa

$$X \rightsquigarrow EN(\lambda) \quad (\lambda > 0)$$

- Função densidade de probabilidade

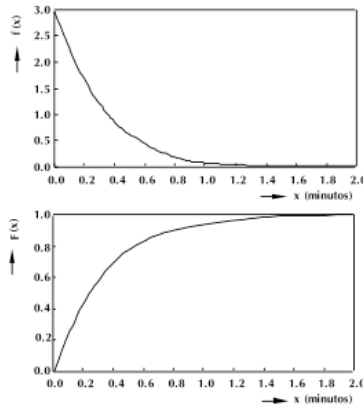
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Função distribuição

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

- Valor esperado e variância

$$\mu = E(X) = \frac{1}{\lambda} \quad \sigma^2 = Var(X) = \frac{1}{\lambda^2}$$



## Notas

- Distribuição Exponencial Negativa não tem memória:  $P(X > s + t | X > s) = P(X > t)$
- $\lambda$  é o  $n^\circ$  médio de ocorrências por unidade de tempo (ou espaço) e  $1/\lambda$  é o tempo (ou espaço) médio entre ocorrências

Slide 5.27

## Exemplo

Num jogo de dardos, a distância entre o local para onde o jogador aponta, o centro do alvo, e o local onde o dardo cai é uma variável aleatória  $X$  com distribuição Exponencial Negativa com  $\lambda = 2$ .

- Indique a função densidade de probabilidade,  $f_X(x)$ , e a função distribuição,  $F_X(x)$ , da v.a.  $X$ .
- Calcule a probabilidade de o dardo cair a uma distância entre 1 e 3.
- Calcule o valor esperado e a variância da distância do local onde cai o dardo ao centro do alvo.

*Resolução:*

$X$ : distância entre o local onde o dardo cai e o centro do alvo

$$X \rightsquigarrow EN(2)$$

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

$$a) f(x) = \begin{cases} 0 & , x < 0 \\ 2e^{-2x} & , x \geq 0 \end{cases} \quad F(x) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-2x} & , x \geq 0 \end{cases}$$

$$b) P(1 < X < 3) = F(3) - F(1) = (1 - e^{-6}) - (1 - e^{-2}) = e^{-2} - e^{-6} = 0.133$$

$$c) \mu_X = E(X) = \frac{1}{2} = 0.5 \quad \sigma_X^2 = Var(X) = \frac{1}{2^2} = 0.25$$

Slide 5.28

### 5.3.3 Distribuição Normal

A distribuição Normal é o resultado:

- da soma de um grande nº de efeitos provocados por causas independentes
- em que não há nenhum efeito que predomine face ao efeito dos restantes

$$X \sim N(\mu, \sigma^2)$$

- Função densidade de probabilidade

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

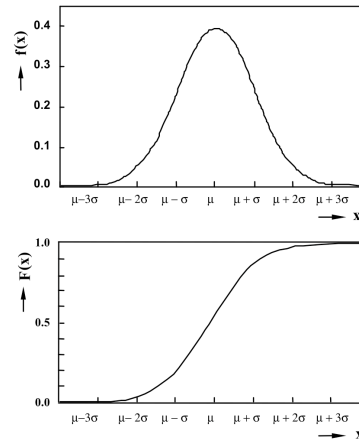
- Função distribuição

$$F(x) = \int_{-\infty}^x f(x) dx$$

$f(x)$  não tem integral  $\Rightarrow$  cálculo por via numérica

- Valor esperado, variância, assimetria e kurtose

$$\mu = E(X) = \mu \quad \sigma^2 = Var(X) = \sigma^2 \quad \gamma_1 = \gamma_2 = 0$$



- Condições muito frequentes de ocorrerem  $\Rightarrow$  Dist. Normal muito utilizada
- $f(x)$  é simétrica em torno de  $x = \mu$ , pontos de inflexão em  $x = \mu \pm \sigma$
- Também conhecida como distribuição Gaussiana ou de Gauss

Slide 5.29

#### Transformação linear de uma v.a. Normal

- $X$  segue uma distribuição Normal com parâmetros  $\mu_X$  e  $\sigma_X^2$
- $V = a + b \cdot X$  é uma transformação linear

$\Rightarrow V$  também segue uma distribuição Normal com parâmetros  $\mu_V$  e  $\sigma_V^2$

$$X \sim N(\mu_X, \sigma_X^2) \xrightarrow{V=a+b \cdot X} V \sim N(\mu_V = a + b \cdot \mu_X, \sigma_V^2 = b^2 \cdot \sigma_X^2)$$

#### Demonstração (apêndice 7.4 do livro recomendado):

- Função geradora de momentos de  $V$ :

$$G(t) = E[e^{t \cdot V}] = E[e^{t \cdot (a+b \cdot X)}] = e^{t \cdot a} \cdot E[e^{(t \cdot b) \cdot X}]$$

- Como  $E[e^{(t \cdot b) \cdot X}] = e^{\mu \cdot (t \cdot b) + \frac{1}{2}[(t \cdot b) \cdot \sigma]^2}$  (ver apêndice 7.3 do livro)

- Vindo  $G(t) = e^{t \cdot a} \cdot e^{\mu \cdot (t \cdot b) + \frac{1}{2}[(t \cdot b) \cdot \sigma]^2} = e^{(a+\mu \cdot b) \cdot t + \frac{1}{2}(b \cdot \sigma \cdot t)^2}$ ,

- A expressão anterior corresponde à função geradora de momentos da distribuição  $N(a + b \cdot \mu, b^2 \cdot \sigma^2)$ .

- Como a função geradora de momentos de  $V$  define univocamente a sua distribuição, resulta que qualquer v.a. obtida por transformação linear de uma v.a. Normal segue também uma distribuição Normal

Slide 5.30

#### Distribuição Normal Padronizada

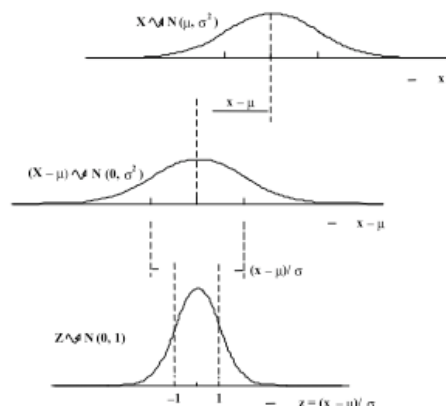
Seja  $X \sim N(\mu, \sigma^2)$  e fazendo

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

$$\Rightarrow Z \sim N(0, 1)$$

$$\mu_Z = E\left(\frac{X - \mu}{\sigma}\right) = \frac{E(X) - \mu}{\sigma} = 0$$

$$\sigma_Z^2 = Var\left(\frac{X - \mu}{\sigma}\right) = \frac{Var(X)}{\sigma^2} = 1$$



## Notas

- Apenas a distribuição Normal Padronizada  $N(0, 1)$  se encontra tabelada
- Possível transformar qualquer distribuição Normal  $X \sim N(\mu, \sigma^2)$  na distribuição Normal Padronizada  $Z \sim N(0, 1)$ , fazendo  $Z = \frac{X - \mu}{\sigma}$
- Trata-se de uma transformação linear (ver slide anterior)

Slide 5.31

- $X_i$  seguem distribuições Normais com parâmetros  $\mu_{X_i}$  e  $\sigma_{X_i}^2$
- $L = \sum_i X_i$  é uma combinação linear de v.a. Normais independentes
- ⇒  $L$  também segue uma distribuição Normal com parâmetros  $\mu_L$  e  $\sigma_L^2$

$$X_i \sim N(\mu_{X_i}, \sigma_{X_i}^2) \xrightarrow{L = \sum_i X_i} V \sim N(\mu_L, \sigma_L^2)$$

Onde:

$$\mu_L = \sum_i \mu_{X_i}$$

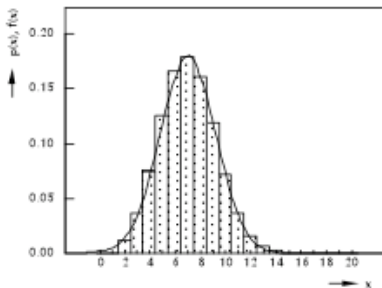
$$\sigma_L^2 = \sum_i \sigma_{X_i}^2$$

Nota: A variância da soma é igual à soma das variâncias, uma vez que as v.a.  $X_i$  são todas independentes entre si

Slide 5.32

## Aproximação da distribuição Binomial pela distribuição Normal

- $Y \sim B(N, p) \longrightarrow X \sim N(N \cdot p, N \cdot p \cdot q)$
- **Regra prática:**  $N \geq 20$ ,  $N \cdot p > 7$  e  $N \cdot q > 7$



Exemplo:

$$B(20, 0.4) \longrightarrow N(20 \cdot 0.4, 20 \cdot 0.4 \cdot (1 - 0.5))$$

$$N = 20 \geq 20$$

$$N \cdot p = 20 \cdot 0.4 = 8 \geq 7$$

$$P(Y = 5) = P_B(y = 5) = 0.0746 \approx P_N(4.5 \leq x \leq 5.5) = 0.0721$$

Notas:

- $P_{Binomial}(X = x) \approx P_{Normal}(x - 0.5 < X < x + 0.5)$
- No próximo capítulo veremos como a distrib. Normal pode ser usada para aproximar outras distribs. (p.e., as dist. Hipergeométrica e de Poisson)

objectivo – facilitar cálculos em situações onde a utilização de uma determinada distribuição apresente dificuldades (p.e., devido a valores muito elevados)

Slide 5.33

## Exemplo

Suponha que a variação relativa diária (%) da cotação de fecho de um dado título segue uma distribuição Normal com  $\mu = 0.2$  e  $\sigma^2 = 1.6^2$

- Calcule a probabilidade de a próxima variação ultrapassar 1%.
- Calcule a probabilidade de a próxima variação estar entre  $\pm 1\%$ .

## Resolução:

$X$ : variação relativa diária (%) da cotação de fecho

$$X \sim N(\mu = 0.2, \sigma^2 = 1.6^2) \quad Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

a)

$$P(X > 1) = P\left(\frac{X - \mu}{\sigma} > \frac{1 - \mu}{\sigma}\right) = P\left(Z > \frac{1 - 0.2}{1.6}\right) = P(Z > 0.5)$$

$$P(Z > 0.5) \stackrel{tabela}{=} 0.3085$$

b)

$$P(-1 < X < 1) = P(X < 1) - P(X < -1) = [1 - P(X > 1)] - [P(X > 1)] =$$

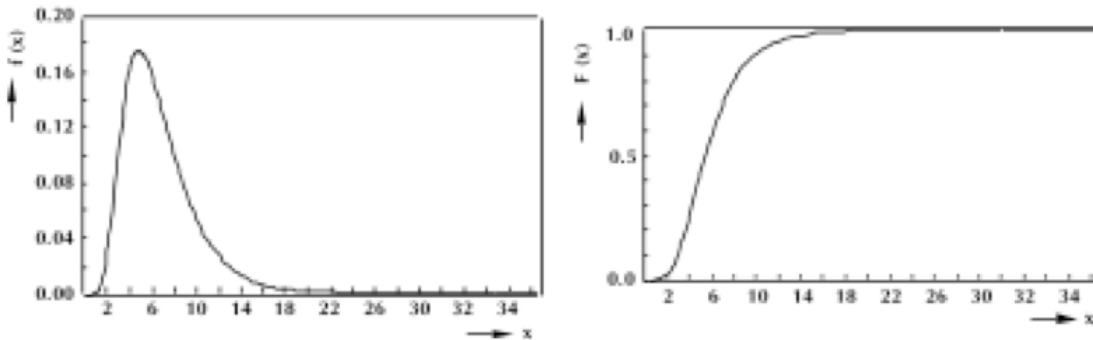
$$= 1 - 2 \cdot P(X > 1) = 1 - 2 \times 0.3085 = 1 - 0.6170 = 0.3830$$

Slide 5.34

### 5.3.4 Distribuição Lognormal

A distribuição *lognormal* é o resultado:

- do *produto* de um grande  $n^\circ$  de efeitos positivos provocados por causas independentes
- em que não há nenhum efeito que predomine face ao efeito global dos restantes



**Transformação**  $V = \ln X$

$$V = \ln X \sim N(\mu_V, \sigma_V^2) \longrightarrow X \sim LN(\mu_V, \sigma_V^2)$$

Slide 5.35

- Valor esperado e variância de  $X$  em função de  $\mu_V$  e  $\sigma_V$

$$\mu_X = E(X) = e^{\mu_V + \sigma_V^2/2} \quad \sigma_X^2 = Var(X) = e^{2 \cdot (\mu_V + \sigma_V^2)} - e^{2 \cdot \mu_V + \sigma_V^2}$$

- Valor esperado e variância de  $V$  em função de  $\mu_X$  e  $\sigma_X$

$$\mu_V = E(V) = \frac{1}{2} \cdot \ln \left( \frac{\mu_X^4}{\sigma_X^2 + \mu_X^2} \right) \quad \sigma_V^2 = Var(V) = \ln \left( \frac{\sigma_X^2}{\mu_X^2} + 1 \right)$$

- Mediana e moda de  $X$

$$\eta_X = e^{\mu_V} \quad \zeta_X = e^{\mu_V - \sigma_V^2}$$

#### Nota

O cálculo de probabilidades de distribuições Lognormais faz-se recorrendo à tabela da distribuição Normal padronizada

Slide 5.36

#### Exemplo

O custo de manutenção de um modelo de automóvel, em euros, no primeiro ano de vida segue uma distribuição Lognormal com valor esperado 80 e variância  $50^2$ , admitindo uma taxa de utilização 15 000 km/ano.

- Calcule a probabilidade do custo ser superior a 125 € no 1º ano.
- Determine a mediana e a moda do custo de manutenção no 1º ano.

**Resolução:**

$$V = \ln X \sim N(\mu_V, \sigma_V^2) \longrightarrow X \sim LN(\mu_V, \sigma_V^2)$$

$$\left| \begin{array}{l} \mu_X = 80 \\ \sigma_X^2 = 50^2 \end{array} \right. \longrightarrow \left| \begin{array}{l} \mu_V = \frac{1}{2} \cdot \ln \left( \frac{80^4}{50^2 + 80^2} \right) = 4.217 \\ \sigma_V^2 = \ln \left( \frac{50^2}{80^2} + 1 \right) = 0.574^2 \end{array} \right.$$

$$a) P(X > 125) = P(\ln X > \ln 125) = P(V > 4.828) = P\left(Z > \frac{4.828 - 4.217}{0.574}\right) = P(Z > 1.064) = 0.1436$$

$$b) \eta_X = e^{\mu_V} = e^{4.217} = 67.83 \quad \zeta_X = e^{\mu_V - \sigma_V^2} = e^{4.21 - 0.574^2} = 48.79$$

Slide 5.37

### 5.3.5 Distribuições “Teóricas” (Qui-quadrado, $t$ de Student e $F$ )

#### Distribuições Qui-quadrado, $t$ de Student e $F$

- Não modelam directamente fenómenos aleatórios
- São indispensáveis à resolução de muitos problemas de inferência estatística
- Funções densidade de probabilidade complexas e que não têm primitiva (expressões apresentadas apenas a título ilustrativo)
- Cálculo de probabilidades por via numérica (usando tabelas ou recorrendo a calculadoras “gráficas”)

#### Notas

Nesta fase pretende-se apenas:

- introduzir e mostrar como surgem estas distribuições
- mostrar como se consultam cada uma das tabelas e se usam as calculadoras “gráficas”

Estas distribuições serão usadas na resolução de exercícios de Inferência Estatística

Slide 5.38

#### Distribuição Qui-quadrado ( $\chi^2$ )

Seja um conjunto de  $GL$  variáveis  $Z_i$ , com  $i = 1, \dots, GL$ , tal que:

- cada v.a.  $Z_i$  segue uma dist. Normal padronizada,  $Z_i \sim N(0, 1)$
- as v.a.  $Z_i$  são mutuamente independentes,  $Z_i \sim IN(0, 1)$

$\Rightarrow X = \sum_i Z_i^2$  segue uma dist. Qui-quadrado com  $GL$  graus de liberdade

$$X = \sum_i Z_i^2 \sim \chi_{GL}^2$$

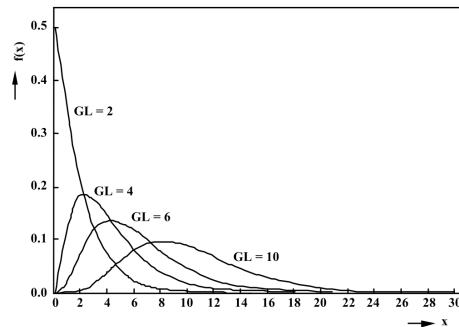
- Função densidade de probabilidade

$$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \quad (x > 0)$$

( $\Gamma(u)$ : função gama)

- Valor esperado e variância

$$\mu_X = E(X) = GL \quad \sigma_X^2 = 2 \cdot GL$$



- $\chi^2$  é uma distribuição assimétrica
- $GL$  aumenta  $\rightarrow$  dist.  $\chi^2$  tende para a dist. Normal

Slide 5.39

#### Distribuição $t$ de Student

- Sejam duas variáveis independentes  $Z \sim N(0, 1)$  e  $V \sim \chi_{GL}^2$

$\Rightarrow X = Z / \sqrt{V/GL}$  segue uma dist.  $t$  de Student com  $GL$  graus de liberdade

$$X = Z / \sqrt{V/GL} \sim t_{GL}$$

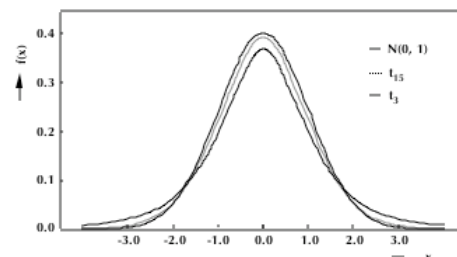
- Função densidade de probabilidade

$$f(x) = \frac{\Gamma(\frac{GL+1}{2})}{\sqrt{GL\pi}\Gamma(\frac{GL}{2})} \left(1 + \frac{x^2}{GL}\right)^{-\frac{GL+1}{2}}$$

- Valor esperado e variância

$$\mu = E(X) = 0, \quad GL > 1$$

$$\sigma^2 = Var(X) = \frac{GL}{GL-2}, \quad GL > 2$$



(para  $GL \leq 1$ ,  $\mu$  não está definido;  
para  $GL \leq 2$ ,  $\sigma^2$  não está definido)

- Forma semelhante à da distribuição Normal mas mais achatada
- $GL$  aumenta  $\rightarrow$   $t$  de Student tende para a Normal Padronizada

Slide 5.40

## Distribuição F

- Sejam duas variáveis independentes  $V_1 \sim \chi_{GL_1}^2$  e  $V_2 \sim \chi_{GL_2}^2$

$\Rightarrow X = \frac{V_1/GL_1}{V_2/GL_2}$  segue uma distribuição F com  $GL_1$  e  $GL_2$  graus de liberdade

$$X = \frac{V_1/GL_1}{V_2/GL_2} \sim F_{GL_1, GL_2}$$

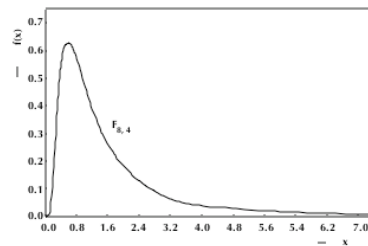
- Função densidade de probabilidade

$$f(x) = \frac{\Gamma\left(\frac{GL_1+GL_2}{2}\right)}{\Gamma\left(\frac{1}{2}GL_1\right)\Gamma\left(\frac{1}{2}GL_2\right)} \cdot \frac{\left(\frac{GL_1}{GL_2}\right)^{\frac{GL_1}{2}} x^{\left(\frac{GL_1}{2}-1\right)}}{(GL_2 + GL_1 \cdot x)^{\frac{GL_1+GL_2}{2}}}$$

- Valor esperado e variância

$$\mu = E(X) = \frac{GL_2}{GL_2 - 2}$$

$$\sigma^2 = Var(X) = \frac{2GL_2(GL_1 + GL_2 - 2)}{GL_1(GL_2 - 2)^2(GL_2 - 4)}$$



$$\begin{cases} f(x) & \text{para } x > 0 \\ E(X) & \text{para } GL_2 > 2 \\ Var(X) & \text{para } GL_2 > 4 \end{cases}$$

Fora dos valores indicados para  $GL_2$  a média e a variância não estão definidas

Slide 5.41

## 5.4 Utilização de Tabelas Estatísticas, Calculadoras e Folhas de Cálculo

### 5.4.1 Tabelas de Distribuições Discretas

#### Tabela da distribuição Binomial

Valores tabelados correspondem às probabilidade acumuladas  $F(y) = \sum_{u=0}^y \left[ \binom{N}{u} \cdot p^u \cdot q^{N-u} \right]$  para diferentes combinações de  $N$  e  $p$

N	y ↓	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	← p
2	0	.9025	.81	.7225	.64	.5625	.49	.4225	.36	.3025	.25	2
	1	.9975	.99	.9775	.96	.9375	.91	.8775	.84	.7975	.75	1
	2	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	0
	p →	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	y ↑
N	y ↓	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	← p
3	0	.8574	.729	.6141	.512	.4219	.343	.2746	.216	.1664	.125	3
	1	.9928	.972	.9393	.896	.8438	.784	.7183	.648	.5748	.5	2
	2	.9999	.999	.9966	.992	.9844	.973	.9571	.936	.9089	.875	1
	3	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	0
	p →	0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	y ↑
N	y ↓	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	← p
4	0	.8145	.6561	.522	.4096	.3164	.2401	.1785	.1296	.0915	.0625	4
	1	.986	.9477	.8905	.8192	.7383	.6517	.563	.4752	.391	.3125	3
	2	.9995	.9963	.988	.9728	.9492	.9163	.8735	.8208	.7585	.6875	2
	3	1.	.9999	.9995	.9984	.9961	.9919	.985	.9744	.959	.9375	1
	4	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	0
p →		0.95	0.90	0.85	0.80	0.75	0.70	0.65	0.60	0.55	0.50	y ↑

Nota: Os valores tabelados correspondem aos valores da função distribuição ( $F(y)$ )

Slide 5.42

#### Como utilizar:

- Procurar tabela correspondente ao valor de  $N$
- Procurar coluna correspondente ao valor de  $p$ 
  - Se  $p < 0.5 \rightarrow$  procurar  $y$  à esquerda
  - Se  $p > 0.5 \rightarrow$  procurar  $y$  à direita
  - Se  $p = 0.5 \rightarrow$  tanto faz
- Valor tabelado corresponde à função probabilidade,  $p(y)$ , quando  $Y \sim B(N, p)$

#### Exemplos

- Seja  $Y \sim B(N = 3, p = 0.30)$ , calcule a probabilidade de  $Y = 1$ :  

$$p(y = 1) = F(1) - F(0) = 0.784 - 0.343 = 0.441$$
- Seja  $Y \sim B(N = 2, p = 0.80)$ , calcule a probabilidade de  $Y = 0$ :  

$$p(y = 0) = F(0) = 0.040$$

Slide 5.43

## Tabela da distribuição de Poisson

Valores tabelados correspondem às probabilidades acumuladas  $F(y) = \sum_{u=0}^y \left( \frac{e^{-\lambda} \cdot \lambda^u}{u!} \right)$  para diferentes valores do parâmetro  $\lambda$

y ↓	λ									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066	0.3679
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725	0.7358
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371	0.9197
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865	0.9810
4	1.0000	1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977	0.9963
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9997	0.9994
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999

y ↓	λ									
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	0.3329	0.3012	0.2725	0.2466	0.2231	0.2019	0.1827	0.1653	0.1496	0.1353
1	0.6990	0.6626	0.6268	0.5918	0.5578	0.5249	0.4932	0.4628	0.4337	0.4060
2	0.9004	0.8795	0.8571	0.8335	0.8088	0.7834	0.7572	0.7306	0.7037	0.6767
3	0.9743	0.9662	0.9569	0.9463	0.9344	0.9212	0.9068	0.8913	0.8747	0.8571
4	0.9946	0.9923	0.9893	0.9857	0.9814	0.9763	0.9704	0.9636	0.9559	0.9473
5	0.9990	0.9985	0.9978	0.9968	0.9955	0.9940	0.9920	0.9896	0.9868	0.9834
6	0.9999	0.9997	0.9996	0.9994	0.9991	0.9987	0.9981	0.9974	0.9966	0.9955
7	1.0000	1.0000	0.9999	0.9999	0.9997	0.9996	0.9994	0.9992	0.9989	0.9989
8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9998	0.9998

...

Nota: Os valores tabelados correspondem aos valores da função distribuição ( $F(y)$ )

Slide 5.44

### Como utilizar

1. Procurar tabela/coluna correspondente ao valor de  $\lambda$
2. Procurar linha correspondente ao valor de  $Y$
3. Valor da tabela corresponde à função distribuição,  $F(y)$ , quando  $Y \rightsquigarrow \text{Poisson}(\lambda)$

### Exemplos:

Seja  $Y \rightsquigarrow \text{Poisson}(1.8)$ . calcule as probabilidades de:

- $p(y \leq 2) = F(2) = 0.7306$
- $p(y = 2) = F(2) - F(1) = 0.7306 - 0.4628 = 0.2678$
- $p(y \leq 10) \approx p(\leq 8) = F(8) = 0.9999$
- $p(y = 10) \approx 0.0000$

Atenção: os valores da tabela correspondem à função distribuição

Slide 5.45

## 5.4.2 Tabelas de Distribuições Contínuas

- Nas distribuições contínuas o recurso a tabelas assume maior importância já que algumas funções densidade de probabilidade não são integráveis analiticamente
- Tabelas podem ser utilizadas em duas situações distintas:
  - conhecemos o valor da v.a. ( $x$ ) e queremos determinar a probabilidade ( $\alpha$ )

$$X \rightsquigarrow N(0, 1^2) \quad P(X > 0.51) \stackrel{(tabela)}{=} 0.3050 = 30.5\%$$

- conhecemos a probabilidade ( $\alpha$ ) e queremos determinar o valor da v.a. ( $x$ )

$$X \rightsquigarrow N(0, 1^2) \quad P(X > x) = 0.0630 \stackrel{(tabela)}{\rightarrow} x = 1.53$$

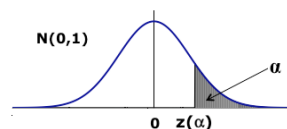
- As tabelas aparecem de forma compacta e sob variadas formas:
  - em distribuições simétricas só metade dos valores estão tabelados
    - $P(X > x) = P(X < -x)$
  - as probabilidades podem corresponder a áreas à direita ou à esquerda
    - $P(X > x) = 1 - P(X < x)$
- ⇒ ter em atenção a legenda da tabela que se está a consultar

Slide 5.46



## Distribuição Normal Padronizada Probabilidades associadas à cauda direita da distribuição $N(0, 1)$

Os valores tabelados correspondem à área  $\alpha$  na figura:  
 $P[Z \geq z(\alpha)] = \alpha$



$z(a) = a + b$ $b \rightarrow$ $a \downarrow$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

Slide 5.47

**Como utilizar** (para valores de  $z$  positivos e valores de  $\alpha$  de caudas à direita):

$z(\alpha)$  conhecido  $\rightarrow$  calcular  $\alpha$

$\alpha$  conhecido  $\rightarrow$  calcular  $z(\alpha)$

1. Decompor  $z(\alpha)$  em  $a + b$
2. Do valor de  $a \rightarrow$  linha
3. Do valor de  $b \rightarrow$  coluna
4. Ler o valor de  $\alpha$  na tabela

1. Procurar  $\alpha$  na tabela
2. Da linha de  $\alpha \rightarrow$  valor de  $a$
3. Da coluna de  $\alpha \rightarrow$  valor de  $b$
4. Obter  $z(\alpha)$  fazendo  $z(\alpha) = a + b$

Nota:  $z = a + b \rightarrow z = N.NN$  vem  $a = N.N$  e  $b = 0.0N$

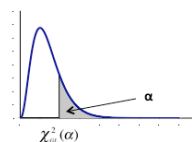
Exemplos — Seja  $Z \sim N(0, 1^2)$ , calcule:

- $P(Z > 1.52) \xrightarrow{\text{tabela}} 0.0643$
- $P(Z < -1.96) = P(Z > 1.96) \xrightarrow{\text{tabela}} 0.0250$
- $P(Z > x) = 0.1515 \xrightarrow{\text{tabela}} x = 1.03$
- $P(Z < x) = 0.0500 \xrightarrow{\text{tabela}} x = -1.645$
- $P(Z > 3.5) \xrightarrow{\text{tabela}} \approx 0$

Slide 5.48

## Distribuição Qui-quadrado ( $\chi^2$ )

Na tabela registam-se os valores  $\chi_{GL}^2(\alpha)$  tais que:  $\alpha = \int_{\chi_{GL}^2(\alpha)}^{+\infty} f(x) dx$



GL ↓	α												
	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	0.000	0.000	0.001	0.004	0.016	0.102	0.455	1.323	2.71	3.84	5.02	6.63	7.88
2	0.010	0.020	0.051	0.103	0.211	0.575	1.386	2.77	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	1.213	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	20.84	25.34	30.43	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	21.75	26.34	31.53	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	22.66	27.34	32.62	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	23.57	28.34	33.71	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.65	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.1	118.5	124.3	129.6	135.8	140.2

Slide 5.49

**Como utilizar** (para valores de  $\alpha$  de caudas à direita, com GL conhecido):

$\chi^2(\alpha)$  conhecido  $\rightarrow$  calcular  $\alpha$

1. Obter linha a partir de GL
2. Procurar valor  $\chi^2(\alpha)$  na linha
3. Obter coluna a partir de  $\chi^2(\alpha)$
4. Ler  $\alpha$  no topo da coluna

$\alpha$  conhecido  $\rightarrow$  calcular  $\chi^2(\alpha)$

1. Obter linha a partir de GL
2. Obter coluna a partir de  $\alpha$
3. Ler o valor de  $\chi^2(\alpha)$  na intersecção da linha e da coluna

Notas:

- a distribuição  $\chi^2$  não é simétrica
- tabela tem poucos valores de  $\alpha$   $\rightarrow$  apenas os habitualmente usadas na construção de Intervalos de Confiança (Cap. 10) e realização de Testes de Hipóteses (Cap. 11)

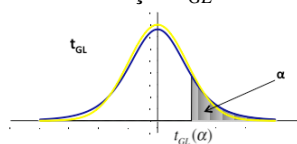
Exemplos — Seja  $X \sim \chi_{10}^2$ , calcule:

- $P(X > 9.34) \xrightarrow{\text{tabela}} 0.500$
- $P(X < x) = 0.010 \xrightarrow{\text{tabela}} x = 2.56$

Slide 5.50

## Distribuição $t$ de Student Probabilidades associadas à cauda direita de distribuições $t_{GL}$

Na tabela registam-se os valores  $t_{GL}(\alpha)$  tais que  $\alpha = \int_{t_{GL}(\alpha)}^{+\infty} f(x) dx$



GL ↓	$\alpha$								
	0.25	0.20	0.15	0.10	0.05	0.025	0.010	0.005	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.460
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
$\infty$	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Slide 5.51

**Como utilizar** (para valores de  $\alpha$  de caudas à direita, com  $GL$  conhecido):

$t(\alpha)$  conhecido  $\rightarrow$  calcular  $\alpha$

$\alpha$  conhecido  $\rightarrow$  calcular  $t(\alpha)$

1. Obter linha a partir de  $GL$
  2. Procurar valor  $t(\alpha)$  na linha
  3. Obter coluna a partir de  $t(\alpha)$
  4. Ler  $\alpha$  no topo da coluna
1. Obter linha a partir de  $GL$
  2. Obter coluna a partir de  $\alpha$
  3. Ler o valor de  $t(\alpha)$  na intersecção da linha e da coluna

Notas:

- tabela com a mesma estrutura da tabela da distribuição  $\chi^2$
- a distribuição  $t$  de Student é simétrica  $\rightarrow$  apenas caudas à direita
- valores da linha para  $GL = \infty$  são iguais aos da distribuição  $N(0, 1^2)$

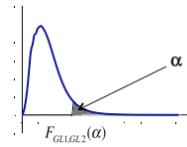
Exemplos — Seja  $X \sim t_{12}$ , calcule:

- $P(X > 1.083) \xrightarrow{\text{tabela}} 0.15$
- $P(X > x) = 0.0250 \xrightarrow{\text{tabela}} x = 2.179$
- $W \sim t_{120}, P(W \leq w) = 0.0250 \xrightarrow{\text{tabela}} w = 1.980$

Slide 5.52

## Distribuição $F$ Probabilidades associadas à cauda direita de distribuições $F_{GL_1, GL_2}$

Na tabela registam-se os valores  $F_{GL_1, GL_2}(\alpha)$  tais que  $\alpha = \int_{F_{GL_1, GL_2}(\alpha)}^{+\infty} f(x) dx$



$\alpha = 10\%$																				
$GL_1 \rightarrow$	1	2	3	4	5	6	7	8	9	10	11	12	14	16	18	20	25	30	40	50
1	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60	60	61	61	61	62	62	62	62	63	63
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.40	9.41	9.42	9.43	9.44	9.44	9.45	9.46	9.47	9.47
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.22	5.20	5.20	5.19	5.18	5.17	5.16	5.15	5.14
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.91	3.90	3.88	3.86	3.85	3.84	3.83	3.82	3.80	3.77
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.28	3.27	3.25	3.23	3.22	3.21	3.19	3.17	3.16	3.12
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.92	2.90	2.88	2.86	2.85	2.84	2.81	2.80	2.78	2.73
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.68	2.67	2.64	2.62	2.61	2.59	2.57	2.56	2.54	2.48
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.52	2.50	2.48	2.45	2.44	2.42	2.40	2.38	2.36	2.31
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.40	2.38	2.35	2.33	2.31	2.30	2.27	2.25	2.23	2.17
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.30	2.28	2.26	2.23	2.22	2.20	2.17	2.16	2.13	2.07
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.23	2.21	2.18	2.16	2.14	2.12	2.10	2.08	2.05	1.99
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.17	2.15	2.12	2.09	2.08	2.06	2.03	2.01	1.99	1.92
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.12	2.10	2.07	2.04	2.02	2.01	1.98	1.96	1.93	1.86
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.07	2.05	2.02	2.00	1.98	1.96	1.93	1.91	1.89	1.82
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.04	2.02	1.99	1.96	1.94	1.92	1.89	1.87	1.85	1.77
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	2.01	1.99	1.95	1.93	1.91	1.89	1.86	1.84	1.81	1.74
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.98	1.96	1.93	1.90	1.88	1.86	1.83	1.81	1.78	1.71
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.95	1.93	1.90	1.87	1.85	1.84	1.80	1.78	1.75	1.68
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.93	1.91	1.88	1.85	1.83	1.81	1.78	1.76	1.73	1.66
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.91	1.89	1.86	1.83	1.81	1.79	1.76	1.74	1.71	1.63
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.90	1.87	1.84	1.81	1.79	1.78	1.74	1.72	1.69	1.61
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86	1.83	1.80	1.78	1.76	1.73	1.70	1.67	1.59
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.87	1.84	1.81	1.78	1.76	1.74	1.71	1.69	1.66	1.57
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83	1.80	1.77	1.75	1.73	1.70	1.67	1.64	1.55
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.84	1.82	1.79	1.76	1.74	1.72	1.68	1.66	1.63	1.54
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.83	1.81	1.77	1.75	1.72	1.71	1.67	1.65	1.61	1.52
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.82	1.80	1.76	1.74	1.71	1.70	1.66	1.64	1.60	1.51
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.81	1.79	1.75	1.73	1.70	1.69	1.65	1.63	1.59	1.50
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.80	1.78	1.75	1.72	1.69	1.68	1.64	1.62	1.58	1.49
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77	1.74	1.71	1.69	1.67	1.63	1.61	1.57	1.48
$GL_2 \uparrow$																				

$\alpha = 10\%$																				
$GL_1 \rightarrow$	1	2	3	4	5	6	7	8	9	10	11	12	14	16	18	20	25	30	40	50
32	2.9	2.5	2.3	2.1	2.0	2.0	1.9	1.9	1.8	1.805	1.78	1.758	1.722	1.694	1.671	1.652	1.616	1.59	1.556	1.535
34	2.86	2.47	2.25	2.12	2.02	1.96	1.90	1.86	1.82	1.79	1.77	1.75	1.71	1.68	1.66	1.64	1.60	1.58	1.54	1.45
36	2.85	2.46	2.24	2.11	2.01	1.94	1.89	1.85	1.81	1.78	1.76	1.73	1.70	1.67	1.65	1.63	1.59	1.56	1.53	1.43
38	2.84	2.45	2.23	2.10	2.01	1.94	1.88	1.84	1.80	1.77	1.75	1.72	1.69	1.66	1.63	1.61	1.58	1.55	1.52	1.42
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.74	1.71	1.68	1.65	1.62	1.61	1.57	1.54	1.51	1.41
42	2.83	2.43	2.22	2.08	1.99	1.92	1.86	1.82	1.78	1.75	1.73	1.71	1.67	1.64	1.62	1.60	1.56	1.53	1.50	1.40
44	2.82	2.43	2.21	2.08	1.98	1.91	1.86	1.81	1.78	1.75	1.72	1.70	1.66	1.63	1.61	1.59	1.55	1.52	1.49	1.39
46	2.82	2.42	2.21	2.07	1.98	1.91	1.85	1.81	1.77	1.74	1.71	1.69	1.65	1.63	1.60	1.58	1.54	1.52	1.48	1.38
48	2.81	2.42	2.20	2.07	1.97	1.90	1.85	1.80	1.77	1.73	1.71	1.69	1.65	1.62	1.59	1.57	1.54	1.51	1.47	1.37
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76	1.73	1.70	1.68	1.64	1.61	1.59	1.57	1.53	1.50	1.46	1.36
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66	1.62	1.59	1.56	1.54	1.50	1.48	1.44	1.33
70	2.78	2.38	2.16	2.03	1.93	1.86	1.80	1.76	1.72	1.69	1.66	1.64	1.60	1.57	1.55	1.53	1.49	1.46	1.42	1.31
80	2.77	2.37	2.15	2.02	1.92	1.85	1.79	1.75	1.71	1.68	1.65	1.63	1.59	1.56	1.53	1.51	1.47	1.44	1.40	1.28
90	2.76	2.36	2.15	2.01	1.91	1.84	1.78	1.74	1.70	1.67	1.64	1.62	1.58	1.55	1.52	1.50	1.46	1.43	1.39	1.27
100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69	1.66	1.64	1.61	1.57	1.54	1.52	1.49	1.45	1.42	1.38	1.26
125	2.75	2.35	2.13	1.99	1.89	1.82	1.77	1.72	1.68	1.65	1.62	1.60	1.56	1.53	1.50	1.48	1.44	1.41	1.36	1.23
150	2.74	2.34	2.12	1.98	1.89	1.81	1.76	1.71	1.67	1.64	1.61	1.59	1.55	1.52	1.49	1.47	1.43	1.40	1.35	1.22
200	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.58	1.54	1.51	1.48	1.46	1.41	1.38	1.34	1.20
300	2.72	2.32	2.10	1.96	1.87	1.79	1.74	1.69	1.65	1.62	1.59	1.57	1.53	1.49	1.47	1.45	1.40	1.37	1.32	1.18
500	2.72	2.31	2.09	1.96	1.86	1.79	1.73	1.68	1.64	1.61	1.58	1.56	1.52	1.49	1.46	1.44	1.39	1.36	1.31	1.16
$GL_2 \uparrow$																				

Slide 5.53

**Como utilizar** (para valores de  $\alpha$  de caudas à direita, com  $GL$  conhecido):

$f(\alpha)$  conhecido  $\longrightarrow$  calcular  $\alpha$

$\alpha$  conhecido  $\longrightarrow$  calcular  $f(\alpha)$

- Para cada tabela
  - Obter linha a partir de  $GL_1$
  - Obter coluna a partir de  $GL_2$
  - Procurar valor  $f(\alpha)$  na tabela
  - Retirar  $\alpha$  da tabela com  $f(\alpha)$
- Procurar tabela correspondente ao valor de  $\alpha$
  - Obter linha a partir de  $GL_1$
  - Obter coluna a partir de  $GL_2$
  - Ler o valor de  $\chi^2$  na intersecção da linha e da coluna

Notas:

- há uma tabela para cada valor de  $\alpha$
- $F_{GL_1, GL_2}(1 - \alpha) = \frac{1}{F_{GL_2, GL_1}(\alpha)}$

Exemplos — Seja  $X \sim F_{5,15}$ , calcule:

- $P(X > x) = 0.10 \xrightarrow{\text{tabela}} x = 2.27$
- $W \sim F_{20,10}, P(W > w) = 0.90 \xrightarrow{\text{tabela}} w = 1/1.94 = 1/1.94 = 0.515$

Slide 5.54

### 5.4.3 Utilização de Calculadoras

- A maior parte das calculadoras “gráficas” possui funções específicas para o cálculo de probabilidades de várias distribuições específicas, desde funções de probabilidade, de densidade e de distribuição a funções “inversas”.
- Funções de probabilidade (PDF) permitem calcular probabilidades pontuais para distribuições discretas; Funções de distribuição (CDF) permitem calcular probabilidades em intervalos para distribuições discretas e contínuas; Funções “inversas” (INV) permitem obter o valor da v.a. que deixa uma determinada probabilidade à esquerda.
- As mais básicas costumam ter pelo menos a função de distribuição da Normal, enquanto as mais avançadas chegam a ter as funções inversas da Qui-quadrado e da F.
- Devem consultar o manual das vossas calculadoras para conhecerem quais as funções que existem e como se usam. (não venham perguntar como funciona a vossa calculadora).
- Os princípios gerais de funcionamento são semelhantes em todos os modelos e apresentam-se de seguida, juntamente com alguns exemplos de funções específicas para as calculadoras TI.

Slide 5.55

#### Princípios gerais:

$DISTpdf(x, PARAM)$  – função de probabilidade (para variáveis discretas) e densidade de probabilidade (para variáveis contínuas) da distribuição  $DIST$  com parâmetros  $PARAM$ , devolve a probabilidade/densidade quando a v.a. é igual a  $x$ .

$DISTcdf(x, PARAM)$  – função de distribuição (acumulada) para variáveis discretas e contínuas da distribuição  $DIST$  com parâmetros  $PARAM$ , devolve a probabilidade da v.a. ser igual ou inferior a  $x$ .

$invDIST(prob, PARAM)$  – devolve o valor  $x$  da v.a. com distribuição  $DIST$  e parâmetros  $PARAM$  que deixa a probabilidade  $prob$  à esquerda.

$DIST$  : binomial, Poisson, norm, ...

$PARAM$  : valor dos parâmetros para a distribuição em causa.

Nota – no cálculo de probabilidades de v.a. contínuas devem usar as distribuições acumuladas ( $cdf$ ) e não as funções densidade ( $pdf$ )

Slide 5.56

#### Alguns exemplos para as TI:

- $Y \sim B(14, 0.8)$ ,  $P(Y = 10) = ?$   
 $\rightarrow binompdf(14, 0.8, 10) = 0.170$
- $Z \sim N(0, 1^2)$ ,  $P(-1.23 < Z < 2.45) = ?$   
 $\rightarrow normalcdf(-1.23, 2.45) = 0.8835$
- $X \sim N(19, 2.5^2)$ ,  $P(17 < X < 24) = ?$   
 $\rightarrow normalcdf(17, 24, 19, 2.5) = 0.7654$
- $X \sim N(19, 2.5^2)$ ,  $P(X > 24) = ?$   
 $\rightarrow normalcdf(24, 1E99, 19, 2.5) = 0.0228$
- $X \sim N(500, 100^2)$ ,  $P(X < x) = 0.9$   
 $\rightarrow x = invnorm(0.9, 500, 100) = 628.2$

[http://calculator.mga.edu/calc\\_topics.html](http://calculator.mga.edu/calc_topics.html)

Notas:

- acesso ao menu com as distribuições:  $2ND + VARS \rightarrow DISTR$
- introdução de  $+\infty$ :  $E99$  (valor suficientemente grande para ser considerado infinito:  $E99 = 10^{99}$ )

Slide 5.57

- Outras distribuições funcionam de forma análoga, p.e.  $tcdf(\dots)$ ,  $invt(\dots)$ ,  $chi2cdf(\dots)$ ,  $fcdf(\dots)$ ; (nota: nem todos os modelos de calculadoras têm todas as distribuições)
- As calculadoras usam, habitualmente, as caudas esquerdas das distribuições (probabilidades iguais ou inferiores a um dado valor), ao contrário das tabelas apresentadas anteriormente que usam as caudas direitas (probabilidades iguais ou superiores a um dado valor);
- Não esquecer que em situações onde uma determinada função não esteja disponível poderá ser possível introduzir directamente a expressão da função (p.e., binomial, uniforme, exponencial, ...);

- Na resolução de exercícios envolvendo distribuições Normais será sempre necessário apresentar a passagem para a Normal padronizada.

#### Questão final

Como são calculadas nas calculadoras “gráficas” as probabilidades de distribuições cujas expressões não têm primitiva (p.e., Normal)?

Slide 5.58

### 5.4.4 Utilização de Folhas de Cálculo

- Funções disponíveis nas folhas de cálculo são semelhantes às das calculadoras gráficas e funcionam de modo semelhante (DIST – funções de probabilidade, de densidade de probabilidade e de distribuição, acumuladas; INV – funções inversas);
- Para além de funções que usam as caudas esquerdas das distribuições (probabilidade de ser igual ou inferior a um determinado valor) há também funções equivalentes que usam as caudas direitas (RT, *Right Tail* – probabilidade de ser igual ou superior a um determinado valor);
- Tal como acontece nas calculadoras “gráficas”, probabilidades de distribuições com expressões sem primitiva são calculadas por integração numérica;
- De seguida apresentam-se as funções existentes no *Excel*, assim como alguns exemplos.

Slide 5.59

#### Funções do Excel

Distribution	PDF/CDF	Inverse	Right Tail (RT)	
Binomial	BINOM.DIST	BINOM.INV		
Negative Binomial	NEGBINOM.DIST			
Hypergeometric	HYPGEOM.DIST			
Poisson	POISSON.DIST			
Uniform	(*)	(*)		
Exponential	EXPON.DIST	EXPON.INV		
Normal	NORM.DIST	NORM.INV		
Standard Normal	NORM.S.DIST	NORM.S.INV		
Lognormal	LOGNORM.DIST	LOGNORM.INV		
Chi-square	CHISQ.DIST	CHISQ.INV	CHISQ.DIST.RT	CHISQ.INV.RT
Student's t	T.DIST	T.INV	T.DIST.RT	T.INV.RT
F	F.DIST	F.INV	F.DIST.RT	F.INV.RT
Beta	BETA.DIST	BETA.INV		
Gamma	GAMMA.DIST	GAMMA.INV		
Weibull	WEIBULL.DIST			

(\*) - Use analytic expression directly in Excel

Usar sempre as funções apresentadas em cima em situações onde existam funções com nomes parecidos (p.e., usar BINOM.DIST e não BINOMDIST).

Slide 5.60

#### Exemplos

$NORM.DIST(x, \mu, \sigma, 0)$  → densidade de probabilidade de uma v.a. Normal com média  $\mu$  e desvio padrão  $\sigma$  no ponto  $x$

$NORM.DIST(x, \mu, \sigma, 1)$  → probabilidade de uma v.a. Normal com média  $\mu$  e desvio padrão  $\sigma$  ser inferior  $x$  (cauda esquerda)

$NORM.INV(prob, \mu, \sigma)$  → valor da v.a. Normal com média  $\mu$  e desvio padrão  $\sigma$  que deixa  $prob$  à esquerda (cauda esquerda)

Outras distribuições funcionam de forma análoga:

- $T.DIST(x, GL, 0/1)$  (cauda esquerda),  $T.DIST.2T(x, GL)$  (duas caudas),  $T.DIST.RT(x, GL)$  (cauda direita),  $T.INV(prb, GL)$ ,  $T.INV.2T(prb, GL)$
- $CHISQ.DIST(x, GL, 0/1)$  (cauda esquerda),  $CHISQ.INV(prb, GL)$ ,  $CHISQ.DIST.RT(x, GL)$  (cauda direita),  $CHISQ.INV.RT(prb, GL)$
- $F.DIST(x, GL_1, GL_2, 0/1)$  (cauda esquerda),  $F.INV(prb, GL_1, GL_2)$ ,  $F.DIST.RT(x, GL_1, GL_2)$  (cauda direita),  $F.INV.RT(prb, GL_1, GL_2)$
- $EXP.DIST(x, \lambda, 0/1)$  (cauda esquerda)
- não existem funções para a distribuição uniforme e para a distribuição exponencial inversa → usar directamente expressões em fórmulas

Slide 5.61

## 5.5 Comentários sobre Aproximações, Tabelas e Calculadoras

### Comentários sobre Aproximações, Tabelas e Calculadoras

#### Contexto

- Situações onde cálculos envolvendo Distribuições de Probabilidade colocam dificuldades:
  - distribuições com expressões que não admitem primitiva (p.e., Normal, t,  $\chi^2$ , F);
  - amostras muito grandes em distribuições cujas expressões tenham combinações ou factoriais.

⇒ Nestas situações recorre-se a *aproximações* entre distribuições, a *tabelas* estatísticas e/ou a *calculadoras* “gráficas”.

⇒ *Recomendação*: utilização conjunta de calculadoras e tabelas, como forma de verificação e validação dos resultados obtidos

Nota : Com o poder de cálculo disponível actualmente (calculadoras “gráficas” e computadores) o recurso a aproximações de distribuições e à utilização de tabelas perdeu muita importância

Slide 5.62

#### Aproximações entre distribuições discretas

- *objectivo* – facilitar os cálculos (i.e., utilizar distribuições cujo cálculo da função probabilidade seja mais simples)
- necessário verificar validade das condições de aproximação:
  - aproximação da Hipergeométrica pela Binomial

$$H(M \cdot p, M \cdot q, N) \xrightarrow{M \geq 10 \cdot N} B(N, p)$$

- aproximação da Binomial pela *Poisson* (só tem interesse quando  $B(N, p)$  for assimétrica:  $N \cdot p \leq 7$  ou  $N \cdot q \leq 7$ , senão aproximar a Binomial pela Normal)

$$B(N, p) \xrightarrow{N \geq 20} \text{Poisson}(\lambda = N \cdot p)$$

- aproximação da Hipergeométrica pela *Poisson*

$$H(M \cdot p, M \cdot q, N) \xrightarrow{\substack{M \geq 10 \cdot N \\ N \geq 20}} \text{Poisson}(\lambda = N \cdot p)$$

- *Nota* – no próximo capítulo veremos aproximações envolvendo distribuições contínuas.

Slide 5.63

#### Utilização de tabelas estatísticas

- *objectivo* – simplificar e acelerar o cálculo de probabilidades;
- rápido, fácil e preciso para valores tabelados (*requerem prática*);
- dificuldades e pouca precisão para valores não tabelados (necessário recorrer a *interpolações* entre os valores tabelados mais próximos)

#### Interpolação

- O que fazer quando uma tabela estatística não tem o valor procurado?

$$X \rightsquigarrow \chi^2_{10} \quad P(X > x) = 0.15 \xrightarrow{(tabela)} x = ?? \quad (\text{não está tabelado})$$

⇒ interpolação linear (ou de grau superior) entre os valores tabelados mais próximos ou aproximar por um valor tabelado muito próximo

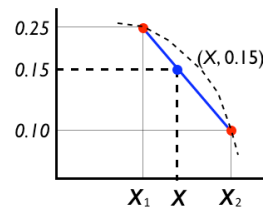
Slide 5.64



### Exemplo

$$P(X > x_1) = 0.25 \xrightarrow{(tabela)} x_1 = 12.55 \xrightarrow{(interp.linear)} x = x_1 + (y - y_1) \cdot \frac{x_2 - x_1}{y_2 - y_1}$$
$$P(X > x_2) = 0.10 \xrightarrow{(tabela)} x_2 = 15.99$$

$$x = 12.55 + (0.15 - 0.25) \cdot \frac{15.99 - 12.55}{0.10 - 0.25} = 14.84$$



### Utilização de calculadoras “gráficas”:

- Integram numericamente as funções densidade de probabilidade;
- Para valores não tabelados, a precisão nas calculadoras é  *muito maior*  do que nas tabelas;
- Todas as calculadoras “gráficas” (desde as mais simples às mais sofisticadas)  *requerem prática*  na sua utilização.

Slide 5.65

## 5.6 Exercícios

1. Admita três caixas com a seguinte composição:

Caixa A: 3 peças azuis, 4 brancas e 2 vermelhas

Caixa B: 5 peças azuis, 2 brancas e 7 vermelhas

Caixa C: 4 peças azuis, 4 brancas e 4 vermelhas

Se seleccionar ao acaso uma caixa e dela retirar três peças, qual a probabilidade de obter pelo menos uma peça azul? (Resp.: 0.76)

2. Suponha que um aluno resolveu tentar a sua sorte num exame composto por 12 perguntas de escolha múltipla. Para cada pergunta são dadas 5 respostas alternativas, sendo apenas uma delas correcta (não há penalizações por respostas erradas).

i) Calcule a probabilidade de o aluno ser aprovado (i.e., ter 6 ou mais respostas certas) se responder às perguntas ao acaso. (Resp.: 0.0193)

ii) Considere agora que o aluno é capaz de identificar, correctamente, para cada pergunta duas das respostas como estando claramente erradas. Qual a probabilidade de o aluno ser aprovado se para cada pergunta escolher ao acaso uma das três respostas restantes? (Resp.: 0.1777)

Slide 5.66

3. Dois amigos, o senhor X e o senhor Y, foram à pesca. Sabendo que o número de peixes pescados por cada um numa tarde de pescaria segue uma distribuição de *Poisson* com  $\lambda = 2$  e  $\lambda = 5$  respectivamente. Calcule:

i) A probabilidade de os dois em conjunto trazerem 10 ou mais peixes. (Resp.: 0.1695)

ii) A probabilidade de o senhor X pescar mais peixes que o senhor Y. (Resp.: 0.083)

Nota – A soma de duas v.a. de *Poisson* ainda é uma v.a. de *Poisson*, já a diferença entre duas v.a. de *Poisson* segue uma distribuição desconhecida ...

Slide 5.67

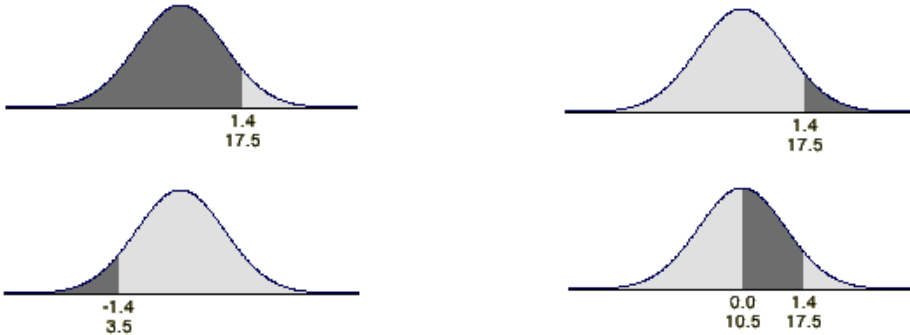


- 4 Considere uma v.a.  $X$  normalmente distribuída, com valor esperado  $\mu = 10.5$  e desvio padrão  $\sigma = 5$ . Determine, utilizando a tabela da distribuição Normal, a probabilidade de  $X$  ser inferior a 17.5?

$$P(X < 17.5) = P\left(\frac{X - \mu}{\sigma} < \frac{17.5 - 10.5}{5}\right) = P(Z < 1.4) =$$

$$= 1 - P(Z > 1.4) = 1 - 0.0808 = 0.9192$$

*Sugestão:* o desenho da curva da Normal ajuda a confirmar o resultado obtido



Slide 5.68

5. Seja  $X_1 \sim N(1, 0.5^2)$ , calcule:

- a)  $P(X_1 < 1.1)$ ,  $P(-0.1 < X_1 < 1.5)$  e  $P(X_1 = 1)$  (sol.: 57.9%, 82.7% e 0%)  
b)  $k$  tal que  $P(X_1 < k) = 25\%$

$$P(Z < z) = 0.25 \Leftrightarrow P(Z > -z) \xrightarrow{(tabela)} -z = 0.675$$

$$(k - 1)/0.5 = z \longrightarrow k = 0.5 \times z + 1 = 0.6625$$

- Seja  $X_2 \sim \chi_{10}^2$ , calcule:

- c)  $P(3.94 < X_2 < 18.31)$  e  $P(X_2 > 3.0)$  (sol.: 90.0% e  $\approx 98\%$ )  
d)  $k$  tal que garante  $P(X_2 > k) = 25\%$  (sol.: 12.55)

- Seja  $X_3 \sim t_5$ , calcule:

- e)  $P(X_3 > 1.156)$  e  $P(|X_3| < 4.032)$  (sol.: 15.0% e 99%)  
f)  $k$  tal que garante  $P(X_3 > k) = 5\%$  (sol.: 2.015)

- Seja  $X_4 \sim F_{10,5}$ , calcule:

- g)  $k_1$  tal que garante  $P(X_4 > k_1) = 10\%$  (sol.: 3.30)  
h)  $k_2$  tal que garante  $P(X_4 > k_1) = 90\%$  (sol.: 0.397)

Slide 5.69

6. A amostra dos resultados dos exames referentes à disciplina de Estatística nos últimos anos revelou que os resultados seguem uma distribuição Normal com média 10.5 e desvio padrão 4. Sabe-se que estão inscritos para o presente exame 16 alunos.

- a) Calcule a probabilidade de um aluno ter uma classificação entre 10 e 13 valores, inclusivé. (Resp.: 0.372)

*Sugestão:* tenha em atenção o arredondamento das notas

- b) Calcule a probabilidade de haver pelo menos 60% de aprovações. (Resp.: 0.523)

*Sugestão:* recorra a uma distribuição Binomial e a uma distribuição Normal ...

Slide 5.70

7. Numa determinada ilha existem três modelos de automóveis disponíveis para compra, um de cada uma das seguintes marcas: *Onda*, *Sabe* e *Fia-Te*. Segundo os representantes das referidas marcas, e tendo em conta as características da rede de estradas local, o consumo dos referidos modelos é caracterizado pelas seguintes distribuições:

- *Onda*:  $N(10, 1^2)$  (litros aos cem)
- *Sabe*:  $N(11, 1.5^2)$  (litros aos cem)
- *Fia-Te*:  $N(8, 1.25^2)$  (litros aos cem)

A família do senhor J.G., residente na ilha, possui um automóvel da marca *Fia-Te*, dois da marca *Onda* e três da marca *Sabe*. O referido senhor tem por hábito escolher à Segunda-feira de manhã o automóvel que vai utilizar ao longo da semana.

- a) Sabendo que na última semana o automóvel utilizado gastou em média mais de 10 litros aos cem, calcule a probabilidade de ter sido um *Onda* o escolhido. (Resp.: 0.303)
- b) Qual a probabilidade do consumo de um automóvel da marca *Fia-Te* ser pelo menos 1 litro inferior ao consumo de um automóvel da marca *Onda*? (Resp.: 0.734)

Slide 5.71

8. Um determinado programa televisivo, transmitido semanalmente, é um sucesso de audiências. Uma das razões para este sucesso é a atribuição de um prémio monetário a um dos telespectadores. A atribuição do prémio é feita através da seleção aleatória de um número da lista telefónica nacional e ligando-se para esse número. Se quem atender o telefone for capaz de citar uma frase publicitária que é anunciada durante o programa, ganha o prémio monetário em jogo nessa sessão. Caso contrário é sorteado novo número. São realizadas no máximo 10 tentativas. A Produção, a partir de estudos de audiências, atribui a probabilidade  $p=10\%$  de uma chamada telefónica ser dirigida a um telespectador do programa (e que, portanto, conhece a frase) e considera-a constante durante o período em que são feitas as chamadas telefónicas.

- a) A variável aleatória  $X$ : “número de chamadas telefónicas realizadas num programa até entregar o prémio (não considerando o limite de 10 chamadas por programa)”, segue que distribuição? (Resp.:  $X \sim G(p = 0.1)$ )
- b) Calcule a probabilidade de se entregar um prémio num programa, ou seja, considerando agora o limite de 10 chamadas por programa. (Resp.: 0.6513)
- c) Calcule o valor esperado do número de chamadas realizadas num programa até entregar o prémio, considerando o limite de 10 chamadas por programa. (Resp.: 3.03 chamadas)

Slide 5.72

Já terminou a terceira série do programa de televisão que distribuíu prémios monetários a quem soubesse a frase publicitária (centenas de edições, portanto). Analisando os prémios distribuídos, concluiu-se que estes seguiam uma distribuição normal com média 650 € e desvio-padrão 10 €.

- d) Qual é a probabilidade de numa das edições do programa, escolhida ao acaso, se distribuir mais do que 660 €? (Resp.: 0.1587)
- e) Se escolhermos 15 edições ao acaso, qual é a probabilidade de encontrar duas ou mais edições com um prémio acima dos 660 €? (Resp.: 0.7133)
- f) Qual é o valor do prémio que, tomando um conjunto grande de edições, será entregue em pelo menos 25% das edições? (Resp.: 658.8 €)
- g) Qual é a probabilidade de a soma dos prémios de duas edições do programa, escolhidas ao acaso, exceder 1310 €? (Resp.: 0.2398)

Slide 5.73

9. Considere que os salários anuais dos trabalhadores de uma dada empresa são uniformemente distribuídos entre 1.000 € e 4.000 €.

- a) Calcule a probabilidade de um trabalhador ganhar mais de 2.500 €/ano. (Resp.: 0.5)
- b) Calcule o valor do salário que é auferido por 20% dos trabalhadores (i.e., 20% dos trabalhadores têm um salário igual ou superior a esse valor). (Resp.: 3400 €)

10. Os acidentes que ocorrem numa empresa de construção civil são independentes e seguem uma distribuição de Poisson com média igual a 6 por mês. Considere a variável aleatória  $T$  que representa o intervalo de tempo entre acidentes consecutivos. Suponha que o estaleiro está em laboração contínua.

- a) Qual é a probabilidade de não ocorrerem acidentes na próxima semana? (Resp.: 0.2230)
- b) Qual é a probabilidade de ocorrerem acidentes em, pelo menos, 3 das próximas 6 semanas? (Resp.: 0.9749)
- c) Calcule o tempo entre acidentes que é ultrapassado em 70% das situações. (Resp.: 0.05944 meses ou 0.2378 semanas)

Slide 5.74

11. Num determinado tipo de compressores,  $D$  é o diâmetro de um cilindro e  $d$  o do pistão. As peças são fabricadas em equipamentos distintos e, assim, ambos os diâmetros são independentes. Admita que  $D$  e  $d$  seguem distribuições Normais com parâmetros:  $\mu_D = 70.0mm$ ,  $\sigma_D = 0.019mm$ ,  $\mu_d = 69.96mm$  e  $\sigma_d = 0.006mm$ .

- a) Qual é a probabilidade de o pistão não entrar no cilindro?
- b) Admita que a folga máxima deverá ser de  $0.07mm$ . Qual é a probabilidade de haver compressores rejeitados por não cumprirem aquela especificação?
- c) Admita agora que as peças são marcadas de acordo com o valor do respectivo diâmetro. Nestas condições, a montagem pistão/cilindro permite que em cada compressor os diâmetros das duas peças estejam correlacionados, de tal forma que o coeficiente de correlação  $\rho_{Dd}$  toma o valor 0.88. Neste caso, qual será a percentagem de compressores rejeitados por não cumprirem a especificação definida em b).

Slide 5.75

## 5.7 Anexos

### 5.7.1 Distribuição Multinomial

- A distribuição Multinomial representa uma generalização da distribuição Binomial, quando numa sucessão de experiências independentes há, em cada prova, mais de dois resultados possíveis.
- As hipóteses subjacentes à distribuição Multinomial são análogas às consideradas no caso da distribuição Binomial.
- Considerem-se então  $n$  experiências aleatórias em que:
  - Em cada experiência aleatória, existem  $k$  resultados possíveis,  $A_i$  (com  $i = 1, 2, \dots, k$ ), mutuamente exclusivos.
  - As probabilidades associadas a cada um dos  $A_i$  que representaremos por  $p_i$  (com  $i = 1, 2, \dots, k$ ), mantêm-se constantes de experiência para experiência, verificando-se que:  $\sum_{i=1}^k p_i = 1$
  - As  $n$  experiências aleatórias são independentes.
- Nestas condições, as variáveis  $X_i$ , com  $i = 1, \dots, k$ , definidas como

$X_i$  - número de vezes, em  $n$ , em que ocorre  $A_i$  (com  $i = 1, 2, \dots, k$ )

têm distribuição Multinomial:

$$(X_1, X_2, \dots, X_k) \rightsquigarrow M(n, p_1, p_2, \dots, p_k)$$

Slide 5.76

- Função de probabilidade conjunta é dada por:

$$p_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) = P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) \\ = \frac{n!}{x_1! \cdot x_2! \cdot \dots \cdot x_k!} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \cdot p_k^{x_k}$$

com:

- $x_i \geq 0$
- $x_1 + x_2 + \dots + x_k = n$
- $p_1 + p_2 + \dots + p_k = 1$

- Valor esperado dos vários  $X_i$ :

$$\mu_{X_i} = E(X_i) = n \cdot p_i \quad (i = 1, 2, \dots, k)$$

- Variância dos vários  $X_i$ :

$$\sigma_{X_i}^2 = \text{Var}(X_i) = n \cdot p_i \cdot (1 - p_i) \quad (i = 1, 2, \dots, k)$$

- Covariância entre os vários  $X_i$ :

$$\gamma_{X_i, X_j} = \text{Cov}(X_i, X_j) = -n \cdot p_i \cdot p_j \quad (i \neq j)$$

Slide 5.77

## 5.7.2 Distribuição Hipergeométrica Generalizada

- Esta distribuição corresponde, tal como o nome indica, a uma generalização da distribuição Hipergeométrica.
- Considere-se então uma população finita constituída por  $M$  elementos distribuídos por  $k$  categorias mutuamente exclusivas e exaustivas de dimensão  $M_i$  (com  $i = 1, 2, \dots, k$ ).
- Nestas condições a proporção de elementos de cada uma das categorias é dada por:

$$p_i = \frac{M_i}{M}$$

verificando-se que:  $\sum_{i=1}^k M_i = M$  e, portanto,  $\sum_{i=1}^k p_i = 1$

- Admita-se agora que desta população se retira uma amostra de  $N$  elementos sem reposição e definam-se  $k$  v.a. do seguinte modo:

$X_i$  - número de elementos na amostra pertencentes à categoria  $i$  ( $i = 1, 2, \dots, k$ )

- As  $X_i$  variáveis aleatórias assim definidas seguem uma distribuição Hipergeométrica Generalizada.

- Função de probabilidade conjunta é dada por:

$$\begin{aligned} p_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) &= P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) \\ &= \frac{\binom{M_1}{x_1} \cdot \binom{M_2}{x_2} \cdot \dots \cdot \binom{M_k}{x_k}}{\binom{M}{N}} \end{aligned}$$

com:

- $x_i \geq 0$
- $x_1 + x_2 + \dots + x_k = n$
- $p_1 + p_2 + \dots + p_k = 1$

- Valor esperado dos vários  $X_i$ :

$$\mu_{X_i} = E(X_i) = n \cdot p_i \quad (i = 1, 2, \dots, k)$$

- Variância dos vários  $X_i$ :

$$\sigma_{X_i}^2 = Var(X_i) = n \cdot p_i \cdot (1 - p_i) \cdot \frac{M - N}{M - 1} \quad (i = 1, 2, \dots, k)$$

- Covariância entre os vários  $X_i$ :

$$\gamma_{X_i, X_j} = Cov(X_i, X_j) = -n \cdot p_i \cdot p_j \cdot \frac{M - N}{M - 1} \quad (i \neq j)$$

Slide 5.78

Slide 5.79

## 5.7.3 Distribuição Normal Multivariada

- Esta distribuição corresponde, tal como o nome indica, a uma generalização da distribuição Normal.

$$X \sim N(\mu, \Sigma) \quad (\text{ou } X \sim N_k(\mu, \Sigma))$$

- Função densidade de probabilidade conjunta é dada por:

$$f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) = f_X(X) = \frac{1}{\sqrt{(2\pi)^k \cdot |\Sigma|}} \cdot e^{-\frac{1}{2} \cdot (X - \mu)^T \Sigma^{-1} (X - \mu)}$$

com

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix} = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_k] \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,k} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k,1} & \sigma_{k,2} & \cdots & \sigma_{k,k} \end{bmatrix}$$

Slide 5.80