

Análise de Dados

Estatística Inferencial

Aplicada à Investigação
em Enfermagem

- ❖ **Objectivos**
- ❖ **Conteúdos Programáticos**
- ❖ **Estratégias de Ensino-Aprendizagem**
- ❖ **Bibliografia**
- ❖ **Material (Máquina de calcular científica)**
- ❖ **Avaliação**



escola superior de
enfermagem
de coimbra

Objectivos:

* Competências:

- Realizar **análises de dados empíricos** (quantitativas e qualitativas);
- Analisar criticamente relatórios de investigação;
- Identificar nos relatórios os aspectos teóricos, metodológicos e empíricos transferíveis;
- Participar de forma tutelada no desenvolvimento de projectos de investigação

* Objectivos Específicos:

- Dominar as bases teóricas e as aplicações fundamentais da **estatística** descritiva e **inferencial**:



Estimação de Parâmetros (médias e proporções)



Testes de Hipóteses (paramétricos e não paramétricos)

Estatística Inferencial

Métodos de extrapolação das conclusões para além do conjunto observado (amostra):

- ✕ Previsões (de acontecimentos no futuro);
- ✕ Estimativas (de parâmetros da população);
- ✕ Testes de hipóteses (determinação da probabilidade das observações realizadas serem resultado do acaso amostral).

**A BASE DA INFERÊNCIA ESTATÍSTICA É A
TEORIA DAS PROBABILIDADES**

INFERÊNCIA ESTATÍSTICA*

*** É necessário fazer uma revisão da Teoria das Probabilidades**
(Ver Fichas de Apoio de Estatística Descritiva e Probabilidades).

INFERÊNCIA ESTATÍSTICA

Geralmente o investigador pretende chegar a conclusões em relação a grandes agregados (Populações**).**

- No entanto, dados as naturais limitações de tempo disponível, de energia, de recursos económicos, etc., é muito raro ser possível estudar individualmente todos os elementos da população (**Censo**). Em vez disso, habitualmente estuda-se conjuntos mais reduzidos de elementos (**Amostragem**)
- Com os dados colhidos da amostra podem calcular-se diversas medidas (**Estatísticas**) que de uma forma sintética a caracterizam em relação às variáveis em estudo.
- A partir destas medidas, as principais preocupações do investigador em Ciências Sociais e Humanas passam a ser:
 - ⊙ Estimar as medidas correspondentes na população (**Parâmetros**);
 - ⊙ Testar hipóteses.

INFERÊNCIA ESTATÍSTICA

A possibilidade de generalização exige que a amostra seja **representativa** !!!

As técnicas de inferência estatística exigem que a amostra seja **aleatória** !!!

Sendo as generalizações feitas a partir de inferências baseadas nos dados amostrais, é evidente que se corre o **risco de erro** !!!



É fundamental que se calcule e se controle esse risco !!!



Esse cálculo e controle baseia-se no conhecimento da distribuição das estatísticas nas amostras possíveis (**Distribuições Amostrais**) relativamente aos respectivos parâmetros (da população).

Distribuições Amostrais

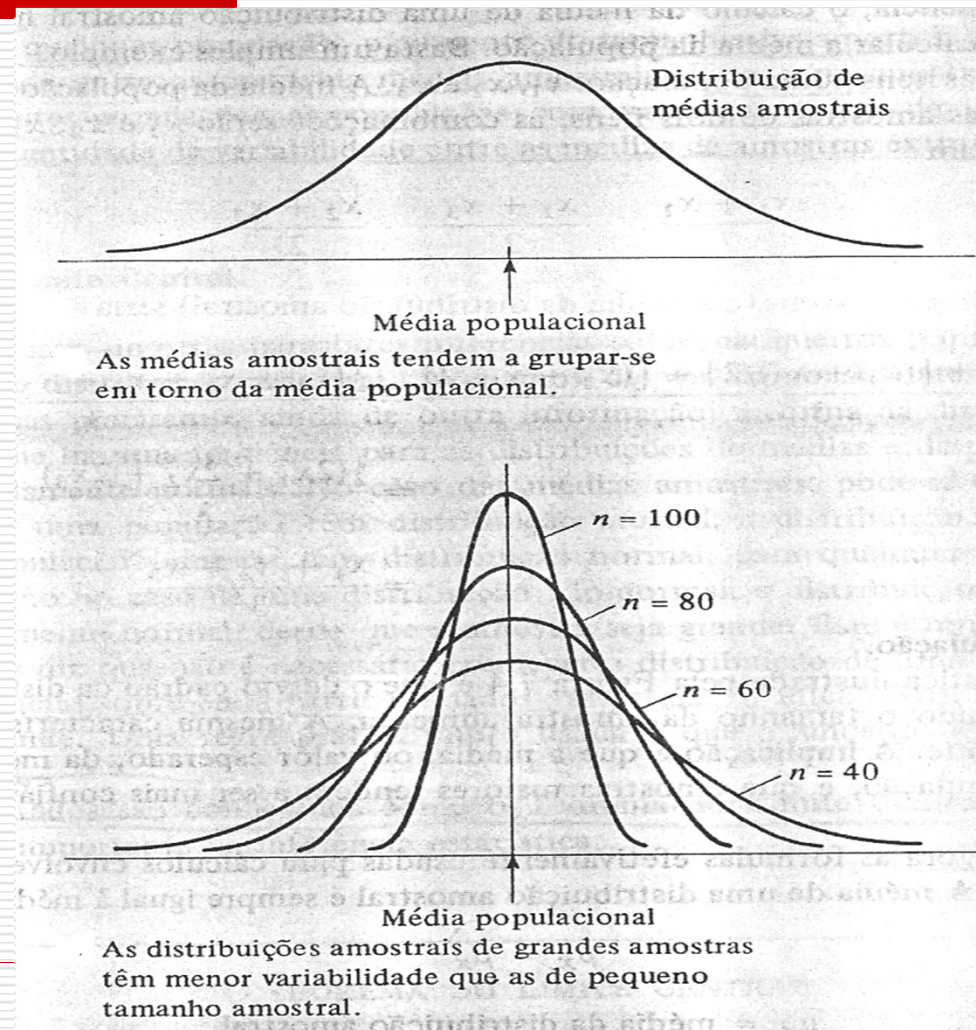
- Uma das realidades da amostragem aleatória é que, quando se extraem repetidas amostras da mesma população, as estatísticas amostrais quase sempre variam de uma amostra para a outra e em relação ao verdadeiro valor do parâmetro populacional. Todavia, distribuem-se na vizinhança do parâmetro.
 - Estas variações resultam de factores aleatórios relacionados com a selecção da amostra.
 - Tendo em conta esta variabilidade amostral, é de esperar uma diferença entre as características da amostra e as da população (**Erro Amostral**).
-

Distribuições Amostrais

- ▲ Sendo a variabilidade das características amostrais (média, proporção, desvio padrão, etc.) resultado de factores do acaso na escolha dos elementos da amostra, estas características são variáveis aleatórias, assumindo determinados valores aos quais corresponde um **distribuição de probabilidades**.
 - ☑ Conhecendo a distribuição de probabilidades da estatística em questão, torna-se possível fazer estimativas do respectivo parâmetro e determinar a probabilidade de determinados desvios amostrais, com o erro sob controlo.
-

Determinantes Fundamentais das distribuições amostrais

- Quanto maior a dispersão da variável (variância) na população, maior a dispersão das estatísticas amostrais em relação ao parâmetro. ▶
- Quanto maior o tamanho da amostra menor a dispersão das estatísticas amostrais em relação ao parâmetro.



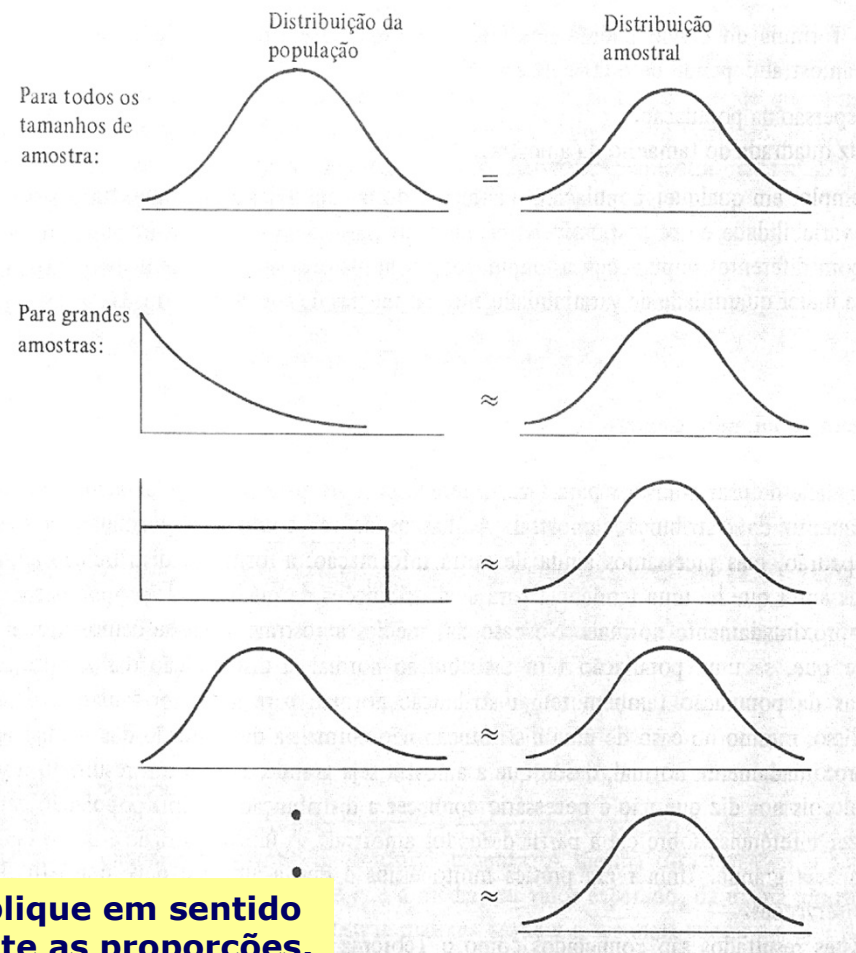
Teorema do Limite Central

A escolha das técnicas de inferência estatística depende do conhecimento da distribuição amostral !!!

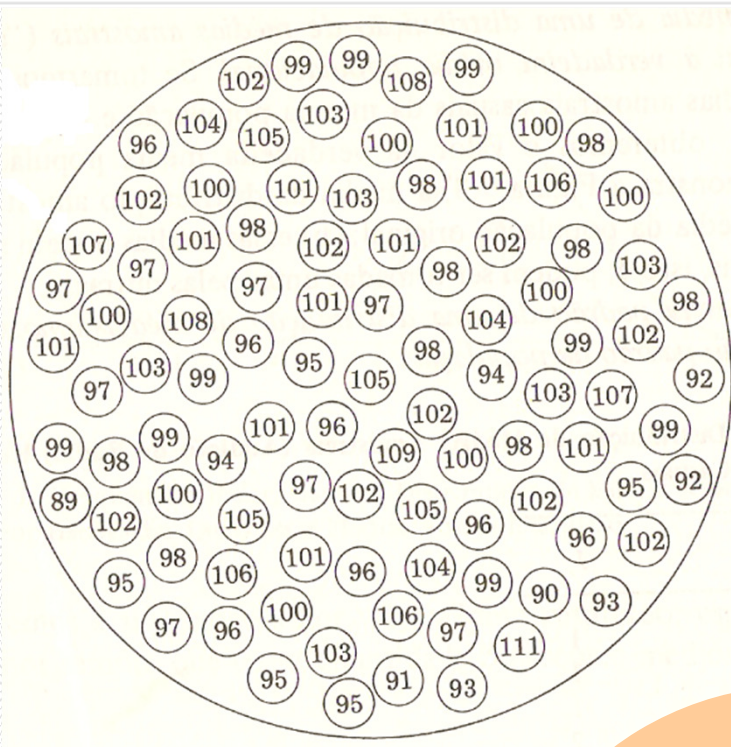
No caso da **média**,

- 1 – Se a variável na população tem uma distribuição normal, a distribuição das médias amostrais será normal para todos os tamanhos de amostra
- 2 – Se a variável na população tem uma distribuição não normal a distribuição das médias amostrais será aproximadamente normal para grandes amostras.

NOTA: Embora o Teorema do Limite Central se aplique em sentido estrito às médias, outras medidas, nomeadamente as proporções, têm idêntico comportamento em grandes amostras ($n > 30$).



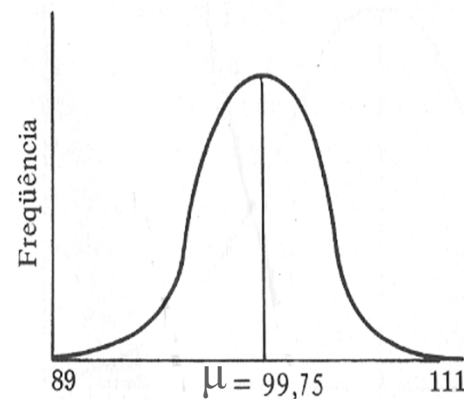
Distribuição das Médias Amostrais



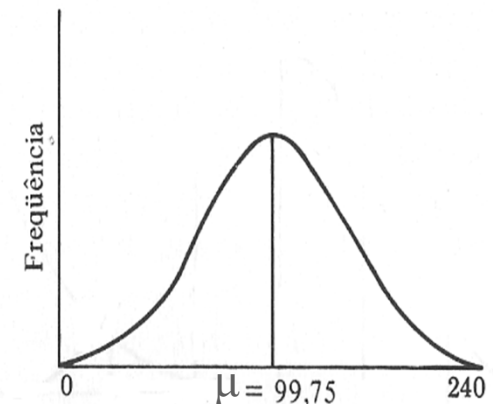
Médias de tempo de uso
da internet por
adolescentes 14-16
anos em amostras de
 $n=500$

**Erro Padrão
da Média
(EPM)**

$\mu = 99,75 \text{ min.}$



Distribuição das $\bar{X}$



Distribuição de X

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

❗ Quando se desconhece σ , $\hat{s}$ é a estimativa menos tendenciosa de σ .

$$\hat{s} = s \sqrt{\frac{n}{n-1}}$$

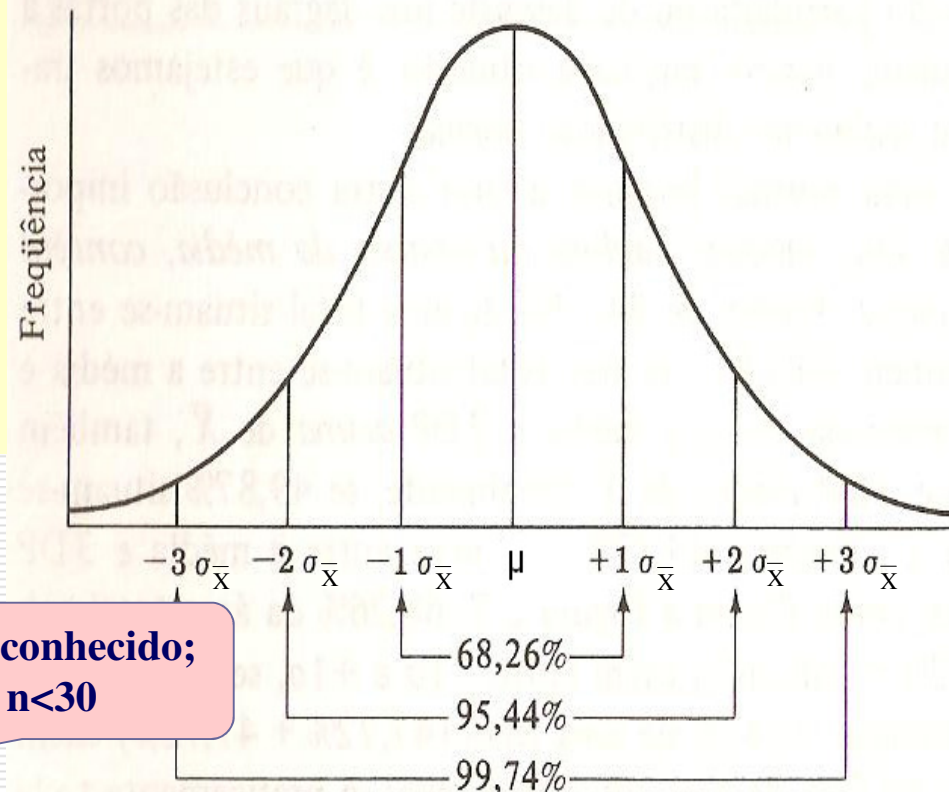
Distribuição das Médias Amostrais

Uma distribuição das médias amostrais é a distribuição de probabilidades que indica quão prováveis são diversas médias de amostras aleatórias retiradas de uma população com média (μ).

$$\bar{X} \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$$

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{ou} \quad t = \frac{\bar{x} - \mu}{\frac{\hat{s}}{\sqrt{n}}}$$

σ desconhecido;
 $n < 30$



$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \text{ou} \quad \sigma_{\bar{x}} \cong \frac{\hat{s}}{\sqrt{n}}; \quad \sigma_{\bar{x}} \cong \frac{s}{\sqrt{n-1}}$$

Problema:

1 – Considerando que o teor de triglicerídeos no plasma dos idosos com mais de 65 anos é uma variável normal com média $\mu = 360$ mg/dl e desvio padrão $\sigma = 53$ mg/dl, calcular a probabilidade de uma amostra aleatória de 120 pessoas deste grupo etário

a) Ter um valor médio ($\bar{x}$) superior a 375 mg/dl.

R: 0,0010

b) Ter um valor médio ($\bar{x}$) entre 350 mg/dl e 370 mg/dl.

R: 0,9616

Distribuição das Proporções Amostrais

Uma distribuição de proporções amostrais é a distribuição de probabilidades que indica quão prováveis são diversas proporções (p) calculadas a partir de amostras aleatórias tendo em conta o tamanho das amostras e a proporção populacional (π).

Erro Padrão da Proporção

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Amostras pequenas:

$$X \sim \beta(n, \pi); \text{ em que } x = np$$

$$P(X = x) = C_x^n \times p^x \times q^{n-x}$$

Amostras grandes ($n > 40$; $np > 5$ e $nq > 5$):

$\approx$ Teorema do Limite Central

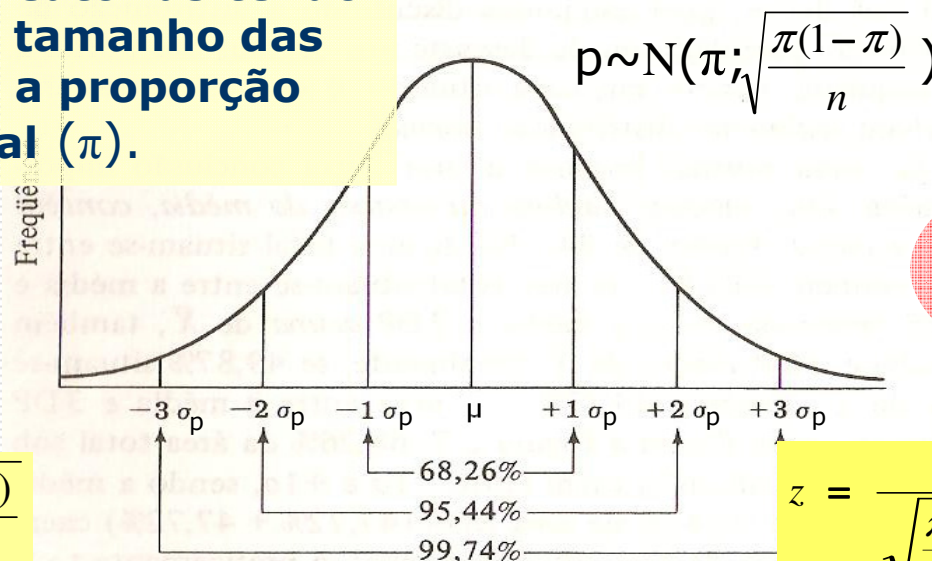
$$p \sim N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right)$$

ou

$$p \sim N\left(\pi, \sqrt{\frac{p(1-p)}{n}}\right)$$

(se π é desconhecido)

$$1 - p = q$$



$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

$$z = \frac{p - \pi}{\sqrt{\frac{pq}{n}}}$$

Problemas:

Numa determinada população, 65% dos elementos são a favor do casamento de homossexuais.

1 – Qual a probabilidade de numa amostra aleatória de 20 indivíduos dessa população

a) Menos de metade ser a favor do casamento de homossexuais?

R: 0,0532

b) Serem a favor do casamento de homossexuais entre 60% e 70%?

R: 0,5170

2 – Qual a probabilidade de numa amostra aleatória de 90 indivíduos dessa população

a) Menos de metade ser a favor do casamento de homossexuais?

R: 0,0014

b) Serem a favor do casamento de homossexuais entre 60% e 70%?

R: 0,6826

Distribuição amostral da frequência de x (n.º de ocorrências)

Para amostras com $n > 40$

Estatística	Média	Erro Padrão
Nº de ocorrências (X)	np	$\sqrt{np(1-p)}$

Amostragem de uma População Finita ($n > 0,05N$)

Corrigir o Erro Padrão !!!

Ex. para a média:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

ESTIMAÇÃO DE PARÂMETROS

ESTIMAÇÃO DE PARÂMETROS

Processo de inferência de um parâmetro populacional (ex.: μ ; π ; ρ ; σ) **apartir de uma estatística amostral – estimador** – (ex.: $\bar{X}$; p ; r ; $\hat{S}$).

Estimador

- Centrado (Não enviesado)
- Eficiente

❖ **Estimação Pontual:** determinação de **um valor** com probabilidade de se situar na proximidade do parâmetro.

❖ **Estimação Intervalar:** determinação de um intervalo de valores que conterá o parâmetro com uma probabilidade predefinida – **grau de confiança** -* (**Intervalo de Confiança**).

* O grau de confiança $(1-\alpha)$ permite determinar o **valor crítico** das estatísticas de inferência (ex.: z_c ; t_c) e com esse valor calcular o **erro máximo provável** da estimativa.

(α : Nível de significância).

ESTIMAÇÃO DA PROPORÇÃO (π)

A proporção amostra (p) situa-se com elevada probabilidade na vizinhança da média populacional (π).

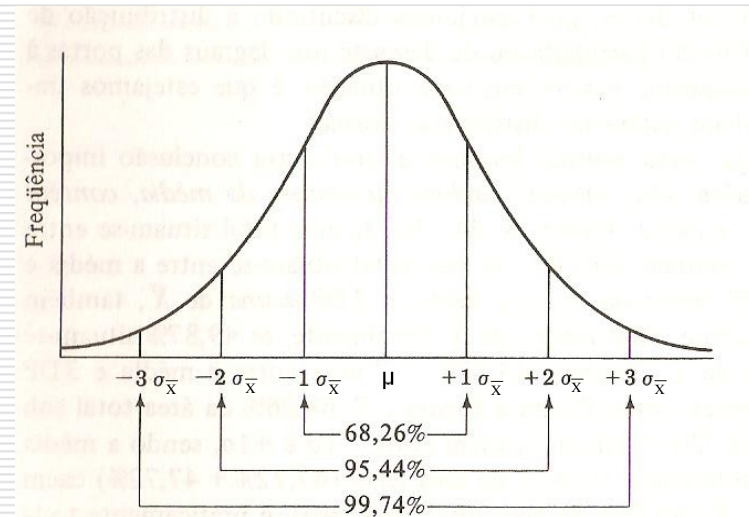
$$z = \frac{p - \pi}{\sqrt{\frac{p(1-p)}{n}}}$$

Pode assumir-se que a proporção de uma amostra aleatória é uma **boa estimativa pontual** da proporção populacional:

$$p \approx \pi$$

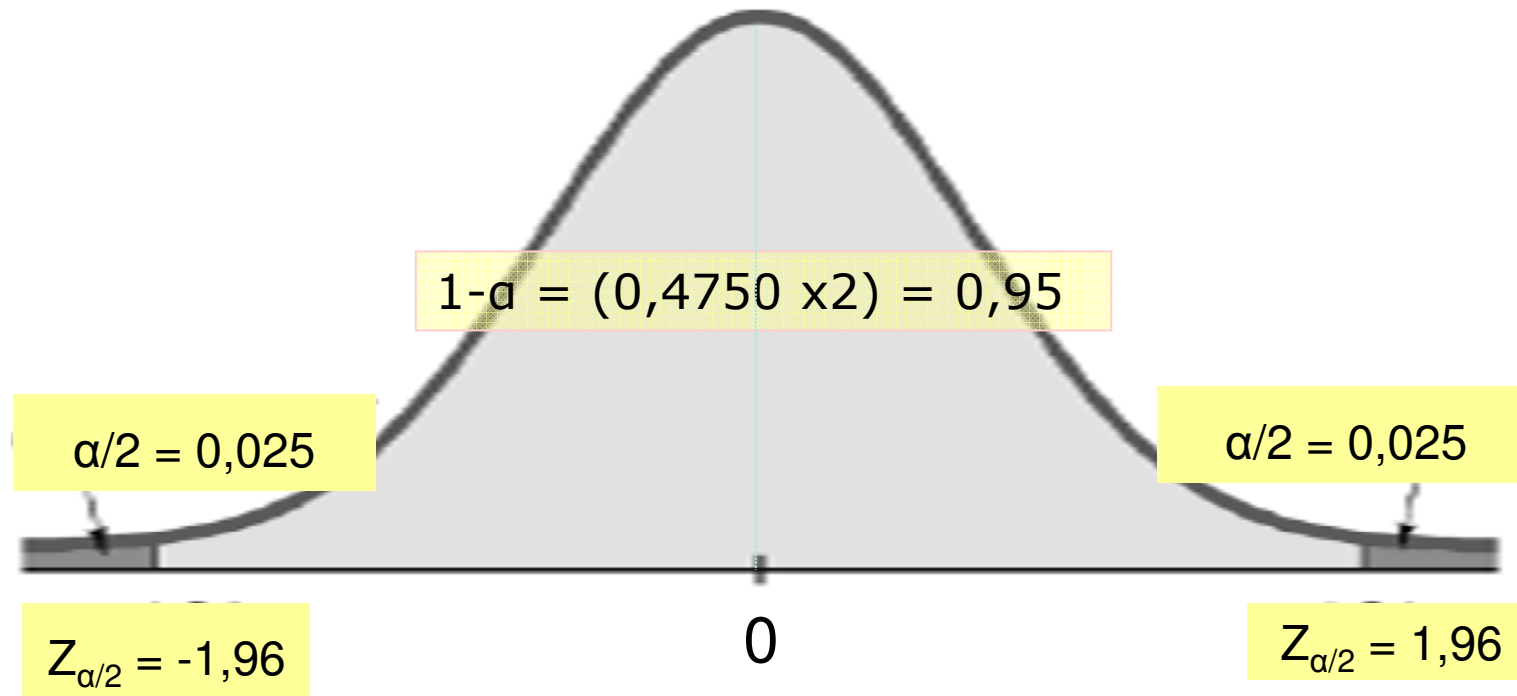
Mas...

uma estimativa pontual não tem controlo sobre o erro. Por isso, é comum realizar estimativa intervalar, definindo um **erro máximo provável** (E) para um determinado **grau de confiança** ($1-\alpha$).



$$p \left(\pi \in \left[p - z_{\frac{1-\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}; p + z_{\frac{1-\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} \right] \right) = 1 - \alpha$$

☒ Originalmente, aplica-se a distribuição binomial. A distribuição normal (Z) é uma boa aproximação para $n > 40$.



Grau de Confiança (1- α)	α	$\alpha/2$	Valor Crítico $Z_{\alpha/2}$
90% (0,90)	0,10	0,05	1,645
95% (0,95)	0,05	0,025	1,960
99% (0,99)	0,01	0,005	2,576

Problemas:

1 - Numa amostra aleatória de 180 doentes jovens adultos (18-35 anos) com mais 15 dias de internamento em serviços de medicina, verificou-se uma prevalência de depressão de 21,67%.

a) Determinar o intervalo de 95% de confiança para a proporção de casos de depressão na respectiva população.

R: $p(\pi \in]0,16; 0,28[) = 0,95$

b) Como se poderia estreitar o intervalo mantendo a confiança?

2 - Numa amostra de 140 doentes idosos internados nos serviços de medicina do Hospital de S. Francisco 14 apresentavam úlcera de pressão.

a) Determinar o intervalo de 90% de confiança para a prevalência (em %) de úlceras de pressão na respectiva população.

R: Int. Conf. 90%: 5,83% - 14,17%

ESTIMAÇÃO DA MÉDIA (μ)

A média amostral ($\bar{x}$) situa-se com elevada probabilidade na vizinhança da média populacional (μ).

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$t = \frac{\bar{x} - \mu}{\frac{\hat{s}}{\sqrt{n}}}$$

Logo...

pode assumir-se que a média de uma amostra aleatória é uma **boa estimativa pontual** da média populacional:

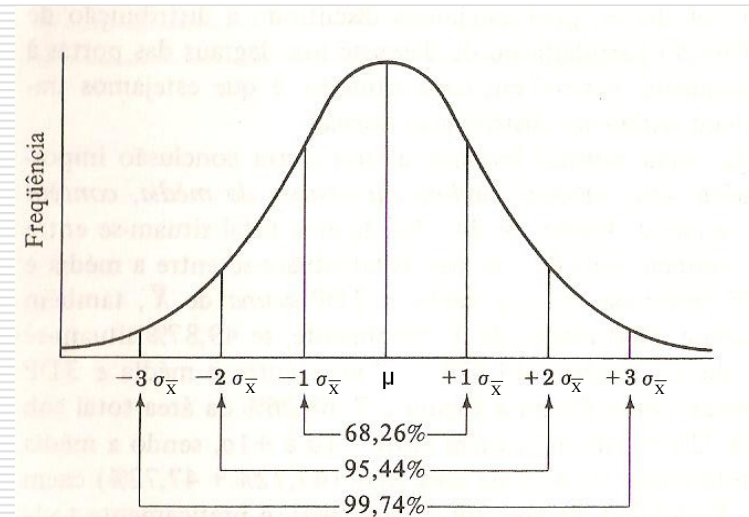
$$\bar{X} \approx \mu$$

Mas...

uma estimativa pontual não tem controlo sobre o erro. Por isso, é comum realizar estimativa intervalar, definindo um **erro máximo provável** (E) para um determinado **grau de confiança** ($1-\alpha$).

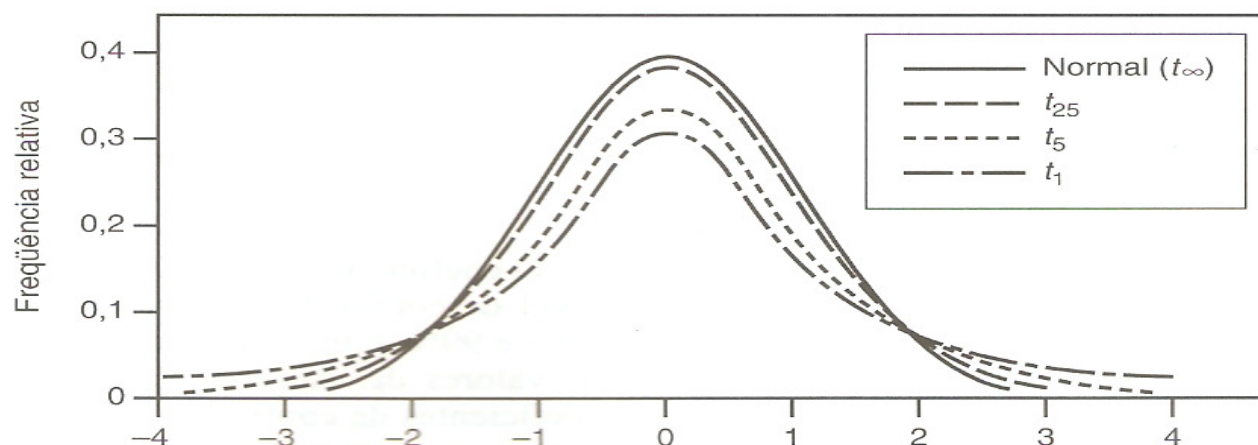
$$p(\mu \in [\bar{x} - E; \bar{x} + E]) = (1 - \alpha)$$

$$p\left(\mu \in \left[\bar{x} - t_{(gl; \alpha/2)} \frac{\hat{s}}{\sqrt{n}}; \bar{x} + t_{(gl; \alpha/2)} \frac{\hat{s}}{\sqrt{n}}\right]\right) = 1 - \alpha$$



⊗ A opção pela distribuição t de Student resulta do facto de na prática quando não se conhece μ não se conhece σ . E t em grandes amostras tende para a distribuição normal (Z).

Relação entre a distribuição t e a distribuição Z



Grau de Confiança (1- α)	α	Tamanho da amostra (n)	Graus de Liberdade (df)*	Valor Crítico $t_{\alpha/2}$
95% (0,95)	0,05	25	24	2,064
95% (0,95)	0,05	61	60	2,000
95% (0,95)	0,05	121	120	1,980
95% (0,95)	0,05	∞	∞	1,96 = z

* Neste caso os graus de liberdade (df) são n-1, mas para outras aplicações calcula-se de modo diverso.

Problemas:

1 – Numa amostra aleatória de 56 jovens do 12^a Ano de Escolaridade da Escola Secundária do Sabugueiro verificou-se um valor médio ($\bar{x}$) de colesterol plasmático de 155,00 mg/100 ml e um desvio padrão (s) de 12,50 mg/100 ml.

a) Determinar o intervalo para a média de colesterol plasmático da população de jovens que frequentam nessa escola o 12^a Ano de Escolaridade, com um nível de confiança de 95%

R: 151,59 mg/100 ml $<\mu<$ 158,41 mg/100 ml com um nível de confiança de 95%

b) Assumindo a média amostral como uma boa estimativa pontual da média populacional, qual o erro máximo provável?

R: 3,41 mg/100 ml

2 – Numa amostra aleatória de 83 grávidas seguidas no C. S. dos Loios, foi estudado o valor da excreção de pregnanodiol às 24 semanas de gestação. Os valores obtidos tinham uma média de 20,50 mg/24 horas e um desvio padrão de 2,8 mg/24 horas.

a) Determinar o intervalo de 99% de confiança para média da população de referência.

R: Int. Conf. 99%:]19,60; 21,40[



Dimensão da amostra

- Recolher e tratar uma amostra grande demais para os resultados que se pretendem obter é um desperdício de recursos.
- Recolher uma amostra cuja dimensão não é suficiente para se poderem tirar conclusões é uma incorrecção.

... Então deve-se antecipadamente determinar o tamanho mínimo da amostra para um determinado **erro máximo aceitável** (E).

A - Determinação do tamanho da amostra para um intervalo de confiança para a proporção:

$$n > Z_c^2 \left(\frac{pq}{E^2} \right)$$

B – Determinação do tamanho da amostra para um intervalo de confiança para a média:

$$n > \left(Z_c \frac{s'}{E} \right)^2$$

TESTES DE HIPÓTESES

TESTES DE HIPÓTESES

A finalidade dos testes de hipóteses é avaliar a veracidade de afirmações sobre parâmetros populacionais (diferenças, desvios, associações) com base em observações realizadas numa amostra.



É muito provável que os resultados amostrais (**estatísticas**) apresentem desvios em relação aos correspondentes parâmetros populacionais !

Até que ponto se pode inferir que os desvios amostrais se devem a variações casuais (**devidos a factores do acaso**) ou a variações reais (**devido à existência real na população**) ?



TESTES

TESTES DE HIPÓTESES

Técnica de inferência estatística que avalia a probabilidade de determinados desvios observados nas distribuições amostrais poderem ser devidos ao acaso ou à sua real existência na população.

ERROS TIPO I E TIPO II

		A Realidade	
		Não há diferença ou relação.	Há diferença ou relação.
Inferência sobre a Realidade	Aceita-se que não há diferença ou relação.	Correcto	Erro tipo II (β)
	Aceita-se que há diferença ou relação.	Erro tipo I (α)	Correcto

α : Nível de significância

$1-\alpha$: Nível de Confiança

$1-\beta$: Poder do teste

ERROS TIPO I E TIPO II

Entre os dois tipos de erros há relação: quando α aumenta β diminui e vice-versa.

	Quando H0 for Verdadeira	Quando H0 for Falsa
Probabilidade de aceitar H0	$1 - \alpha$	β
Probabilidade de rejeitar H0	α	$1 - \beta$

⊠ O modo de reduzir α e β simultaneamente é aumentar o tamanho da amostra.

Hipótese Nula ou Estatística (H0)

- É a hipótese que é submetida a teste.
- Não rejeitar esta hipótese, corresponde a considerar como muito provável que os desvios observados resultam do acaso amostral.

Ex:

“Não existe relação entre a idade dos doentes e a incapacidade funcional após prótese da anca.”

$$\rho = 0$$

“A qualidade de vida dos doentes com reumatismo articular crónico não difere conforme o sexo.”

$$\mu_f = \mu_m$$

“A Intenção de voto é independente do nível socio-económico dos eleitores.”

$$\pi_a = \pi_b = \pi_c$$

Hipótese alternativa ou conceptual (Ha):

Não direccionada (bicaudal):

“Existe relação entre a idade dos doentes e a incapacidade funcional após prótese da anca.”

$$\rho \neq 0$$

“A qualidade de vida dos doentes com reumatismo articular crónico difere conforme o sexo.”

$$\mu_f \neq \mu_m$$

“A Intenção de voto não é independente do nível socio-económico dos eleitores.”

$$\pi_a \neq \pi_b \neq \pi_c$$

Direccionada (unicaudal):

“Existe relação positiva entre a idade dos doentes e a incapacidade funcional após prótese da anca.”

$$\rho > 0$$

“A qualidade de vida dos doentes com reumatismo articular crónico é superior nos doentes do sexo masculino.”

$$\mu_f < \mu_m$$

TESTAR UMA HIPÓTESE

Critério Fundamental:

AMOSTRA ALEATÓRIA

Aplicar um modelo probabilístico adequado à distribuição das variáveis em teste:



Binomial; Normal; Qui-Quadrado; T; F; ...

TESTAR UMA HIPÓTESE (cont.)

TIPOS DE TESTE:

Paramétricos – os mais poderosos – (medição no mínimo intervalar; distribuição normal da população*; homogeneidade da variância).

Não Paramétricos – menos poderosos – (particularmente indicados para amostras pequenas e/ou dados de medição nominal ou ordinal).

É determinada a respectiva estatística teste e, a partir desta, a probabilidade de os desvios observados na amostra ou outros mais extremos ocorrerem por acaso (devidos às condições de aleatoriedade amostral) na condição H_0 .

Esta probabilidade (**p**) pode ser determinada num programa apropriado, designa-se por nível de significância do resultado e é comparada com um valor de **α** crítico, habitualmente:

0,10

0,05

0,01

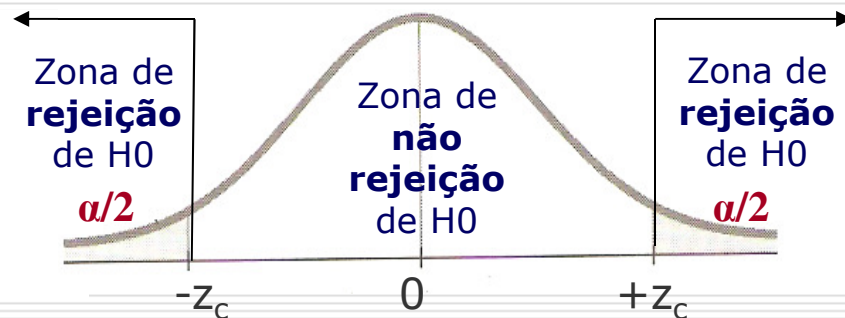
* Discutível para grandes amostras, tendo em conta o Teorema do Limite Central.

TESTAR UMA HIPÓTESE (cont.)

Quando o teste é executado manualmente, com a ajuda de uma máquina de calcular, determina-se a **estatística teste**, e compara-se esta com o respectivo **valor crítico**.

- ❖ Verifica-se se a estatística teste, se situa na **zona de rejeição de H_0** ou na **zona de não rejeição de H_0** .

Ex. de um teste Z:



Hipótese Bilateral (Sinal de H_a : =)



Hipótese unilateral á direita (Sinal de H_a : >)

- ❖ Em certos casos, calcula-se o valor da probabilidade da distribuição observada ou de uma mais extrema *.

* **Rejeita-se H_0** se a probabilidade (p) for $< \alpha$ (Ex.: $p < 0,05$);

* **Não se Rejeita H_0** se a probabilidade (p) for $\geq \alpha$ (Ex.: $p \geq 0,05$);

TESTAR UMA HIPÓTESE (cont.)

Quando o teste é executado num programa de estatística, este determina o nível de significância resultado (**p**) que corresponde:

- ❖ À probabilidade da hipótese H_0 ser rejeitada quando na realidade é verdadeira (**risco de erro tipo I**).
- ❖ À probabilidade do desvio observado (diferença, associação) ou outros mais extremos serem devidos ao acaso, não existindo realmente na população.

$p < 0,10$		No limiar da significância
$p < 0,05$	*	Significativo
$p < 0,01$	**	Muito significativo
$p < 0,001$	***	

Quanto menor o **p** menor o risco de rejeitar H_0 e não rejeitar H_1 sem erro.

TESTAR UMA HIPÓTESE (Processo)

A partir de amostras...

Aplicar técnicas de inferência estatística que permitem determinar a probabilidade de determinados desvios (diferenças, associações) poderem ser devidos ao acaso amostral ou à sua existência real na população.

Processo:

- 1 – Identificação do problema e definição da hipótese nula (H_0) e da hipótese alternativa (H_a);
- 2 – Extração de uma amostra aleatória representativa da população;
- 3 – Seleção da prova estatística adequada e especificação do nível de significância crítico (α);
- 4 – Cálculo das estatísticas a partir da amostra e aplicação do teste;
- 5 – Comparar o resultado do teste com o(s) valor(es) crítico(s) correspondentes ao nível de significância definido em 3;
- 6 – Decisão de rejeição ou não rejeição de H_0 .

INTERPRETAÇÃO DO RESULTADO DO TESTE

Se o resultado for no sentido da rejeição de H_0 ($p < 0,05$), dizemos que os “desvios” observados são significativos...

... Mas, não se confirma em definitivo a hipótese alternativa (H_a), porque:



Pode ter sido utilizado um modelo não adequado;



A amostra pode não ser representativa;



Podem ter sido cometidos erros sistemáticos;



...

Se o resultado for no sentido da não rejeição de H_0 ($p \geq 0,05$), dizemos que os “desvios” observados não são significativos...

... Mas, não se elimina em definitivo a hipótese alternativa (H_a), porque:



As observações podem ter sido em número insuficiente;



As medições podem não ter sido realizadas com a discriminação necessária;



...

TESTES DE HIPÓTESES UNIVARIADAS

Hipóteses de um valor populacional

PARAMÉTRICO

Pressupostos (para além da **aleatoriedade da amostra**):



Medição no mínimo intervalar da variável

Distribuição normal da variável na população ^(a)

Teste para uma média

Bilateral

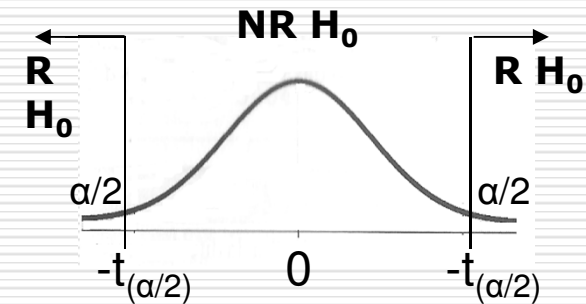
Unilateral (à esquerda / à direita)

(a) Inferência prévia analisando distribuição da amostra (especialmente a simetria) ou, melhor, usando um teste de aderência: Kolmogorov-Smirnov ou Shapiro-Wilks ($n < 50$).

Teste de uma média

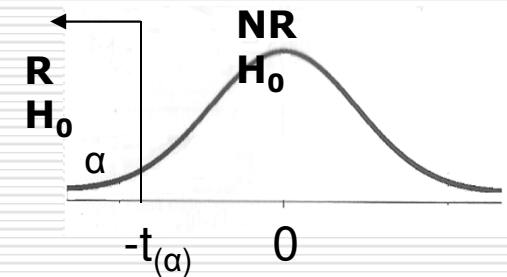
Teste bilateral:

$H_0: \mu = \mu_0$ com $H_1: \mu \neq \mu_0$



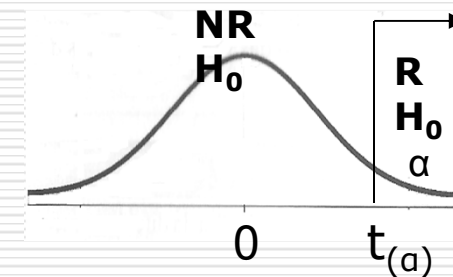
Teste unilateral à esquerda:

$H_0: \mu = \mu_0$ com $H_1: \mu < \mu_0$



Teste unilateral à direita:

$H_0: \mu = \mu_0$ com $H_1: \mu > \mu_0$



$$t_o = \frac{\bar{X} - \mu_0}{\frac{\hat{s}}{\sqrt{n}}}$$

Gl = n - 1



Problema:

- 1 -No sentido de testar a hipótese de que “O teor médio de alumínio no sangue dos doentes hemodializados há mais de 1 ano no centro de hemodiálise da Gerdunha é superior a 50 $\mu\text{g/dl}$ ”, foi realizado um estudo numa amostra aleatória de 35 desses doentes. Nessa amostra, o valor médio observado foi de 52,19 $\mu\text{g/dl}$ e o desvio padrão de 8,23 $\mu\text{g/dl}$.
- a) Assumindo o pressuposto da distribuição normal do teor de alumínio no sangue da população de referência (Teste de Shapiro-Wilks: $p > 0,05$), teste a hipótese enunciada para um nível de significância $\alpha = 0,05$.

R: $H_0: \mu = 50 \mu\text{g/dl}$; $H_a: \mu > 50 \mu\text{g/dl}$

$t_o = 1,54$; $t_c = 1,70$: **Não se rejeita H_0 .**

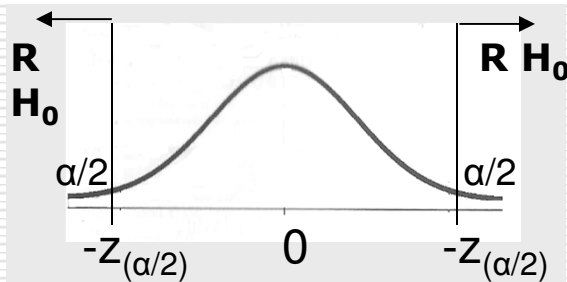
Embora a média de alumínio no sangue na amostra seja de 52,19 $\mu\text{g/dl}$, este valor não constitui prova empírica suficiente, admitindo um erro tipo I máximo de 5%, de que na população de doentes hemodializados há mais de 1 ano no centro de hemodiálise da Gerdunha é superior a 50 $\mu\text{g/dl}$.

TESTES DE HIPÓTESES UNIVARIADAS

Hipóteses de um valor populacional

NÃO PARAMÉTRICOS

Teste para uma proporção

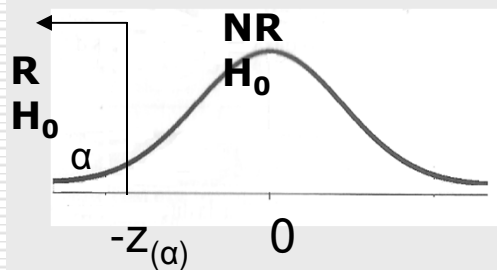


Bilateral $(H_0: \pi = \pi_0 \text{ com } H_1: \pi \neq \pi_0)$

Unilateral

À esquerda $(H_0: \pi = \pi_0 \text{ com } H_1: \pi < \pi_0)$

À direita $(H_0: \pi = \pi_0 \text{ com } H_1: \pi > \pi_0)$



Teste Binomial (amostras pequenas)

$$P(X = x) = C_x^n \times \pi^x \times 1 - \pi^{n-x} \quad \text{em que} \quad C_x^n = \frac{n!}{x!(n-x)!}$$

Teste Z ($n > 40$)*

$$z_o = \frac{p - \pi_0}{\sqrt{\frac{p(1-p)}{n}}}$$

* Com base no Teorema do Limite Central.

Problemas

1 – Testar a hipótese que uma moeda “não é honesta”, assumindo um nível de significância de 0,05.

$$H_0: p(c)=p(k)=0,5; H_1: p(c) \neq p(k)$$

$$\text{Rejeita-se } H_0 \text{ se } p(c \geq x \text{ u } k \leq n-x) < 0,05$$

Exemplo a) Numa experiência de 15 lançamentos, saiu 10 vezes cara (c) e 5 vezes coroa (k).

R: $p(X \geq 10) = 0,1509$; $p(X \leq 5) = 0,1509$; $p = 0,3019$ (**$p > 0,05$**). **Não se rejeita H_0 ; com $\alpha = 0,05$. a experiência não corrobora a hipótese de que a moeda não é honesta,**

Exemplo b) Numa experiência de 20 lançamentos, saiu 15 vezes cara (c).

R: $p(X \geq 15) = 0,0207$; $p(X \leq 5) = 0,0207$; $p = 0,0414$ (**$p < 0,05$**). **Rejeita-se H_0 ; ...**

2 – Para testar a hipótese de que a “taxa de úlceras de pressão nos doentes acamados há mais de 15 dias no Hospital Bom Encontro é inferior a 10%”, realizou-se um estudo por amostragem aleatória de 64 doentes nessa condição.

a) Sabendo que nessa amostra se registou úlcera de pressão em 5 doentes, teste a hipótese para $\alpha = 0,05$.

$$R: H_0: \pi = 0,10; H_a: \pi < 0,10$$

$$z_0 = -0,656; z_c = -1,645: \text{ **Não se rejeita } H_0.**$$

TESTES DE HIPÓTESES BIVARIADAS

A - Hipóteses de diferença entre grupos

TESTES DE HIPÓTESES BIVARIADAS

A - Hipóteses de diferença entre grupos

PARAMÉTRICOS

Pressupostos (para além da **aleatoriedade da amostra**):



Medição no mínimo intervalar da VD



Distribuição normal da VD na população (a)



Homogeneidade das variâncias (homocedasticidade) (b)

Testes de diferenças de médias

Amostras Independentes

Amostras emparelhadas

(a) Inferência prévia analisando distribuição da amostra (especialmente a simetria) ou, melhor, usando um teste de aderência: Kolmogorov-Smirnov ou Shapiro-Wilks ($n < 50$).

(b) Testar, por ex., com o teste de Levene ($F = s^2_M / s^2_m$).

Exemplos:

Teste de diferenças entre duas médias

Grupos independentes (VI nominal dicotômica)

Teste t para amostras independentes:

$$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{(\bar{X}_1 - \bar{X}_2)}}$$

Teste bilateral:

$H_0: \mu_1 = \mu_2$ com $H_1: \mu_1 \neq \mu_2$

Teste unilateral à esquerda:

$H_0: \mu_1 = \mu_2$ com $H_1: \mu_1 < \mu_2$

Teste unilateral à direita:

$H_0: \mu_1 = \mu_2$ com $H_1: \mu_1 > \mu_2$

$$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1)\hat{s}_1^2 + (n_2 - 1)\hat{s}_2^2}{n_1 + n_2 - 2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

ou

$$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

Este teste é...

- ❖ Robusto para a não normalidade da VD na população para amostras grandes ($n > 30$);
- ❖ Existe uma fórmula alternativa para calcular os graus de liberdade associado a t quando as variâncias forem significativamente diferentes (**Ver no SPSS**).

$$Gl = n_1 + n_2 - 2$$

Problema

- 1 – Para testar a hipótese de que “o nível de depressão pós parto é mais elevado nas puérperas primíparas do que nas multíparas”, foram analisados os dados de uma amostra aleatória de 93 puérperas, tendo-se obtido os seguintes resultados:

Paridade	n	Média	DP
Primíparas	54	13,53	6,15
Multíparas	39	10,95	7,46

- a) Assumindo a normalidade da variável na população (primíparas e multíparas) e a homogeneidade das variâncias ($F=1,48$; $p>0,05$), realizar o respetivo teste com um erro máximo tipo I de 0,05.

R: $H_0: \mu_1 = \mu_2$ com $H_1: \mu_1 > \mu_2$

$t_0 = 1,803$; $t_c = 1,671$: **Rejeita H_0 . A diferença entre as médias amostrais é significativa ($p < 0,05$).** (Resultados $p = 0,037$; ...)

Teste de diferenças entre duas médias

Grupos emparelhados (medições repetidas)

Teste t para amostras emparelhadas:

Teste bilateral:

$H_0: \mu_a = \mu_d$ com $H_1: \mu_a \neq \mu_d$

Teste unilateral à esquerda:

$H_0: \mu_a = \mu_d$ com $H_1: \mu_a < \mu_d$

Teste unilateral à direita:

$H_0: \mu_a = \mu_d$ com $H_1: \mu_a > \mu_d$

$$t_o = \frac{\sum d}{\sqrt{\frac{n \sum d^2 - (\sum d)^2}{n-1}}}$$

ou

$$t_o = \frac{\bar{d}}{\frac{\hat{s}_d}{\sqrt{n}}}$$

$$Gl = n - 1$$

Problema

- 1 – 10 adolescentes obesas (IMC>30) submeteram-se a um determinado regime de regulação do peso corporal, tendo-se registado os seguintes valores do IMC antes e após 6 meses de iniciar o referido regime:**

Antes	31	30	31	35	34	32	31	34	36	38
Após	32	28	27	29	35	28	24	26	35	33

- a) Com base nestes dados, testar a hipótese de que “com este regime de regulação do peso corporal, ao fim de 6 meses as adolescentes apresentam um IMC inferior ao inicial”.

R: $H_0: \mu_d = \mu_a$ com $H_1: \mu_d < \mu_a$

$T_o = -3,491$; $T_c = -1,833$: Rejeita-se H_0 . O IMC dos adolescentes após 6 meses com esse tipo de regime é significativamente inferior ($p < 0,05$).

Teste de diferenças entre três ou mais médias

Grupos independentes (VI nominal 3 ou + cat.)

ANOVA com um factor

Teste não direccionado:

$H_0: \mu_1 = \mu_2 = \dots = \mu_k$ com $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$

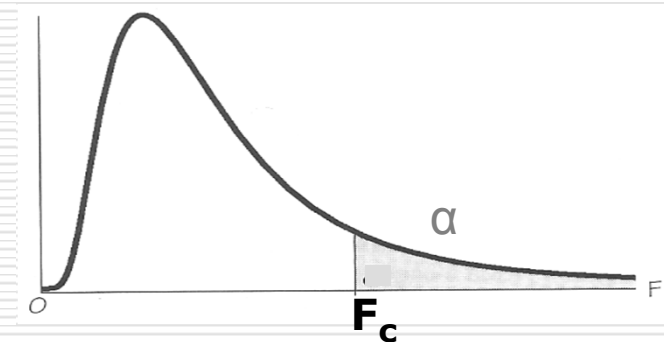
$$F_o = \frac{S_E^2}{S_D^2}$$

$$S_E^2 = \frac{n_1(\bar{X}_1 - \bar{X})^2 + n_2(\bar{X}_2 - \bar{X})^2 + \dots + n_k(\bar{X}_k - \bar{X})^2}{k-1}$$

$$S_D^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \dots + (n_k - 1)S_k^2}{n - k}$$

$$\bar{X} = \frac{n_1 \cdot \bar{X}_1 + n_2 \cdot \bar{X}_2 + \dots + n_k \cdot \bar{X}_k}{n_1 + n_2 + \dots + n_k}$$

GL numerador: (nº grupos-1) ou (k-1)
GL denominador: (n₁-1)+(n₂-1)+(n₃-1)
ou (n-k)



TESTES DE HIPÓTESES BIVARIADAS

A - Hipóteses de diferença entre grupos

NÃO PARAMÉTRICOS **

Pressupostos (para além da **aleatoriedade da amostra**):



A validade de cada teste poderá depender de alguns critérios específicos *

Amostras Independentes

VD ordinal e VI dicotómica (U de Mann-Witney; ...)

VD ordinal e VI com 3 ou + cat. (H de Kruskal-Wallis; ...)

VD nominal e VI dicotómica (Teste Z de diferenças de proporções;
Teste do Qui-quadrado;...)

VD nominal e VI com 3 ou + cat. (Teste do Qui-quadrado)

Amostras Emparelhadas (medidas repetidas)

VD ordinal e 2 medições (Wilcoxon; ...)

VD ordinal e 3 ou + medições (Friedman; ...)

VD nominal dicotómica e 2 medições (McNemar; ...)

...

**** Usados como alternativa aos testes paramétricos quando os respectivos pressupostos não são satisfeitos.**

* Consultar, por exemplo: Siegel, Sidney – *Estatística não paramétrica*. Rio de Janeiro: McGraw Hill, 1975.

Problema

- 1 – De modo a avaliar a eficácia relativa de 3 métodos de redução de peso em adultos obesos, constituíram-se 3 amostras aleatórias. Obtiveram-se os seguintes resultados de evolução do peso ao fim de 6 meses de intervenção:**

	Mét. A	Mét. B	Mét.C
n	42	47	42
Média	-7,2 Kg	-4,9 Kg	-3.2 Kg
DP (n-1)	3,2 Kg	3,6 Kg	3,4 kg

R: $H_0: \mu_A = \mu_B = \mu_C$ com $H_1: \mu_A \neq \mu_B \neq \mu_C$

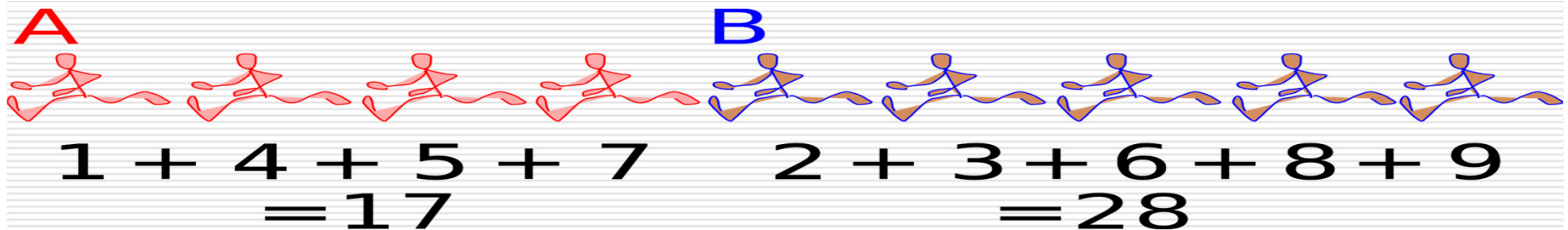
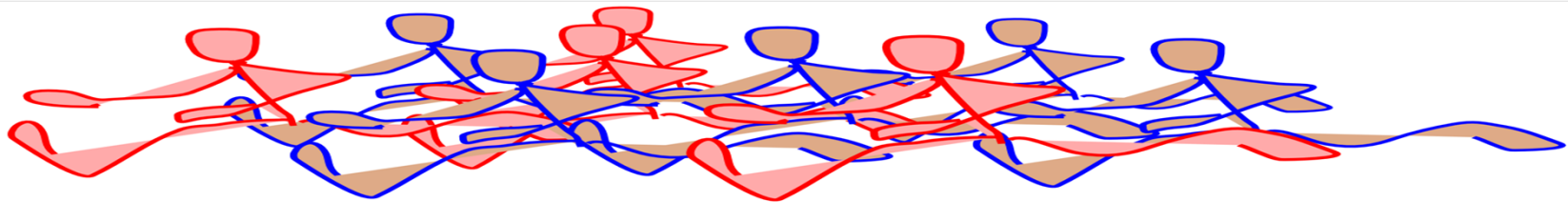
$$s_e^2 = 88,47; s_d^2 = 16,93$$

$F_o = 5,23$; $F_{c(g|2/88; \alpha = 0,05)} = 3,15$: Rejeita-se H_0 . Os dados amostrais constituem prova empírica de que os 3 métodos diferem quanto à sua eficácia, com um erro máximo tipo I de 0,05.

Exemplos:

Teste U de Mann-Witney

Grupos independentes (VI nominal dicotómica)*



Teste U de Mann-Witney *

Grupos independentes (VI nominal dicotômica)

Para $n \leq 20$:

Escolher o **U** de valor inferior e comparar com o valor U crítico da tabela – **Se $U_o < U_c$ rejeita-se H_0 .**

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Para $n > 20$:

$$z_o = \frac{U_o - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 2)}{12}}}$$

R_1 = soma das posições no grupo 1
 R_2 = soma das posições no grupo 2

*** Alternativa não paramétrica para o teste t para amostras independentes. Neste caso, os valores são transformados em postos.**

Problemas

- 1 – Tomando o exemplo dos corredores escolhidos aleatoriamente de 2 equipas, pode afirmar-se que os vermelhos são mais fracos?**

$$n_1=4; R_1=17 \quad U=4 \times 5 + [(4 \times 5)/2] - 17 = 13$$

$$n_2=5; R_2=28 \quad U=4 \times 5 + [(5 \times 6)/2] - 28 = 7 \text{ (valor mais baixo)}$$

R: H0: As equipas não se distinguem.

Ho: $\neq$; p(tuicaada) = 0,278. Não se rejeita H0. Com base nesta amostra e com $\alpha = 0,05$, não se pode afirmar que equipa dos azulados é mais fraca.



- 2 – Considerando os valores da seguinte tabela relativos à glicemia em jejum (mg/dl) de 2 grupos de sujeitos, testar a hipótese de que “a população da qual provém o Grupo A tem valores de glicemia em jejum superiores”.**

G. A (posto)	77 (3)	125 (16)	96 (10)	86 (6)	110 (13)	117 (15)	88 (8)	94 (9)	100 (12)
G. B (posto)	87 (7)	98 (11)	115 (14)	76 (2)	85 (5)	72 (1)	81 (4)		

$$N_1=9 ; R_1=92; U=16$$

$$N_2=7 ; R_2=44; \\ U=47$$

R: H0: Os dois grupos não diferem; H1: O grupo A tem valores mais elevados.

$U_o=16 > U_c=15$: Não se rejeita H0.



Teste Z para diferenças de proporções *

Grupos independentes (VI nominal dicotômica)

Teste bilateral:

H0: $\pi_1 = \pi_2$ com H1: $\pi_1 \neq \pi_2$

Teste unilateral à esquerda:

H0: $\pi_1 = \pi_2$ com H1: $\pi_1 < \pi_2$

Teste unilateral à direita:

H0: $\pi_1 = \pi_2$ com H1: $\pi_1 > \pi_2$

$$z_o = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

* Pode usar-se como alternativa o teste do Qui-quadrado, mas este é exclusivamente bilateral. Ver à frente.

Problema

1 – No Serviço de Medicina do Hospital dos Anjos dos 214 doentes internados no ano transacto, 27 desenvolveram úlceras de pressão, enquanto nos 361 doentes internados no mesmo período no Serviço de Medicina da Clínica do Beato, desenvolveram úlceras de pressão 29.

a) Testar a afirmação do Director da Clínica do Beato de que que “no Serviço de Medicina do Hospital dos Anjos a prevalência de úlceras de pressão é superior à da Clínica do Beato” ($\alpha=0,05$).

R: $H_0: \pi_A = \pi_B$ com $H_1: \pi_A > \pi_B$

$Z_o = 1,787$; $Z_c = 1,645$: Rejeita-se H_0 . A maior prevalência observada nos doentes do Hospital dos Anjos é significativa para $\alpha=0,05$.

b) Testar hipóteses de que “a prevalência de úlceras de pressão no Hospital dos Anjos é diferente da prevalência de úlceras na Clínica do Beato” ($\alpha=0,05$).

R: $Z_o = 1,787$; $Z_c = 1,960$: Não se rejeita H_0 .



Teste do Qui-Quadrado

Diferenças de proporções entre amostras independentes / Relação entre variáveis nominais

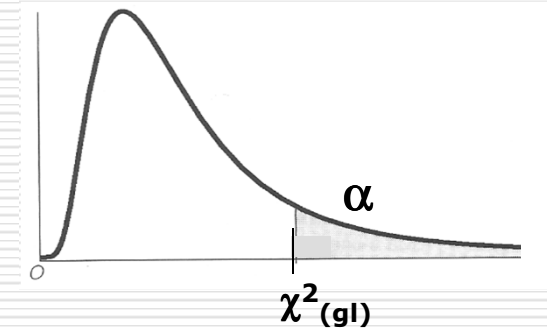
Critérios de Cochran:

Tabelas 2x2 [gl = 1]

- 1 Nenhuma frequência esperada inferior a 5;
- 2 Com $n \gg 40$, utilizar Qui-quadrado (χ^2);
- 3 Com $n > 40$ utilizar Qui-quadrado (χ^2) ou Qui-quadrado com correcção de continuidade (χ_c^2);
- 4 Se $20 \leq n \leq 40$, utilizar Qui-quadrado com correcção de continuidade (χ_c^2);
- 5 Se $n < 20$ ou alguma frequência esperada inferior a 5, utilizar o teste exacto de Fisher.

Tabelas LxC [gl = (L-1)x(C-1)]

Utilizar o Qui-quadrado (χ^2) se no máximo 20% das células com frequência esperada inferior a 5 e nenhuma inferior a 1.



$$\chi^2 = \sum \frac{(F_O - F_E)^2}{F_E}$$

$$\chi_c^2 = \sum \frac{(|F_O - F_E| - 0,5)^2}{F_E}$$



$$F_E = \frac{\text{Total da linha} \times \text{Total da coluna}}{\text{Total dos totais}}$$

$$GL = (L-1)(C-1)$$

Problema

- 1 - Considerar a tabela seguinte relativa aos casos de doentes que desenvolveram úlceras de pressão no Serviço de Medicina do Hospital das Anjos e no Serviço de Medicina da Clínica do Beato.**

	H. Anjos	C. Beato	Totais
C/ Úlcera	27	29	56
S/ Úlcera	187	332	519
Totais	214	361	575

- a) Testar a hipótese de que "a prevalência de úlceras de pressão no Hospital dos Anjos é diferente da prevalência de úlceras na Clínica do Beato ($\alpha=0,05$).

R: $H_0: \pi_A = \pi_B$ com $H_1: \pi_A \neq \pi_B$

$\chi_0^2 = 3,216$; $\chi_c^2 = 3,841$. Não se rejeita H_0 . Os dados não corroboram a hipótese de diferença, com um nível de significância de 0,05.

AMOSTRAS EMPARELHADAS

Exemplos:

Teste dos Sinais*

Grupos emparelhados (duas medições)

Para $n \leq 25$:

Aplicação do **teste Binomial** (**X** número de ocorrências do sinal menos frequente).

$$p(X \leq x) = \sum C_x^n p^x q^{n-x}$$

Para $n > 25$:

$$z = \frac{(x + 0.5) - \frac{n}{2}}{\frac{\sqrt{n}}{2}}$$

n = número de casos com diferença diferente de 0.

* 65% do poder do teste t para amostras emparelhadas (pouco poderoso).

Teste de Postos com Sinais de Wilcoxon*

Grupos emparelhados (duas medições)

Para $n \leq 30$:

Calcular a soma dos postos negativos e a soma dos postos positivos. **T** é a menor das somas.

Comparar com os valores críticos (Tabela)

Para $n > 30$:

$$z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

n = número de casos com diferença diferente de 0.

*** Alternativa não paramétrica para o teste t para amostras emparelhadas (95% do poder).**

Problema

- 1 – 10 adolescentes obesas (IMC>30) submeteram-se a um determinado regime de regulação do peso corporal, tendo-se registado os seguintes valores do IMC antes e após 6 meses de iniciar o referido regime:**

Antes	31	30	31	35	34	32	31	34	36	38
Após	32	28	27	29	35	28	24	26	35	33

P: Não assumindo o pressuposto de normalidade, testar a hipótese ($\alpha=0,05$) de que “com este regime de regulação do peso corporal, ao fim de 6 meses as adolescentes apresentam um IMC inferior ao inicial”.

Respostas

Antes	31	30	31	35	34	32	31	34	36	38	
Após	32	28	27	29	35	28	24	26	35	33	
Diferença	1	2	4	6	1	4	7	8	1	5	
Sinal da diferença	+	-	-	-	+	-	-	-	-	-	-=8 +=2
Posto *	2	4	5,5	8	2	5,5	9	10	2	7	T= 4

* Quando se verificam valores empatados, a ordem desses valores será a média dos postos sucessivos ocupados.

$H_0: Md_{(depois)} = Md_{(antes)}$; $H_1: Md_{(depois)} < Md_{(antes)}$

Teste dos Sinais: $p(x \leq 2) = 0,0547$. Não se rejeita H_0

Teste Wilcoxon ($\alpha = 0,05$; Unilateral): $T_o = 4 < T_c = 11$. Rejeita-se H_0

Teste de McNemar

Grupos emparelhados; Variável nominal (evolução entre dois tempos)

▲ Quando $n < 60$, é mais adequado usar a fórmula com a correção de Yates (χ_c^2).

		ANTES	
		+	-
DEPOIS	-	A	B
	+	C	D

$$\chi^2 = \frac{(A - D)^2}{A + D}$$

ou

$$\chi_c^2 = \frac{(|A - D| - 1)^2}{A + D}$$

Gl=1

Problema

- 1 – Os dados da tabela que se segue apresentam a evolução da atitude de 68 fumadores face à necessidade de iniciar um processo de desabituação tabágica, antes e após uma conferência sobre o tabagismo:

		ANTES	
		Atitude desfavorável	Atitude favorável
DEPOIS	Atitude favorável	11	34
	Atitude desfavorável	20	3

- a) O que se pode concluir em relação à evolução da atitude?

R: H0: Mudança nula; H1: Mudança efectiva

$\chi_0^2=4,571$; $\chi_c^2=3,841$. Rejeita-se H0. A mudança foi significativa ($p<0,05$), no sentido de este tipo de intervenção aumentar proporção de fumadores com atitude favorável à desabituação.

TESTES DE HIPÓTESES BIVARIADAS

B - Hipóteses de relação (associação) entre variáveis

TESTES DE HIPÓTESES BIVARIADAS

B - Hipóteses de relação (associação) entre variáveis

PARAMÉTRICOS

Pressupostos (para além da **aleatoriedade da amostra**):



Medição no mínimo intervalar das variáveis



Distribuição normal das variáveis na população

Teste t de significância da Correlação de Pearson *

Teste bilateral:

$H_0: \rho=0$ com $H_1: \rho \neq 0$

Teste unilateral à esquerda:

$H_0: \rho=0$ com $H_1: \rho < 0$

Teste unilateral à direita:

$H_0: \rho=0$ com $H_1: \rho > 0$

$$r = \frac{S_{xy}}{S_x S_y} = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$t_o = r \sqrt{\frac{n-2}{1-r^2}}$$

Gl=n-2

*** Acresce o pressuposto de Linearidade da relação.**

Análise da Associação entre Duas Variáveis de Medição Intervalar

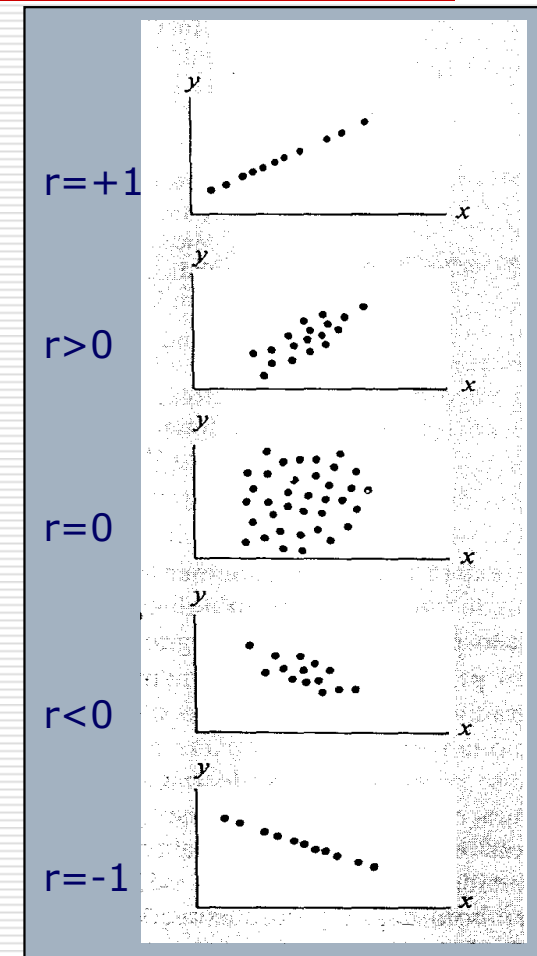
- ❖ **Medida de Associação entre Variáveis Intervalares ou Racionais (Correlação de Bravais-Pearson)**

▲ Relação Linear

$$r = \frac{S_{xy}}{S_x S_y} = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$



Coeficiente de Determinação: r^2
(variância explicada)



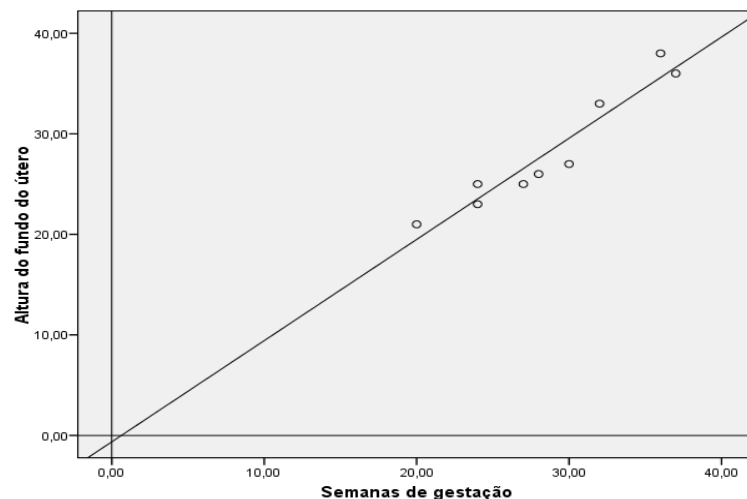
▲ Se os pontos se dispuserem de forma não linear (correlação curvilínea), o r resulta num valor inferior ao real.

Análise da Associação entre Duas Variáveis de Medição Intervalar

❖ Correlação de Bravais-Pearson

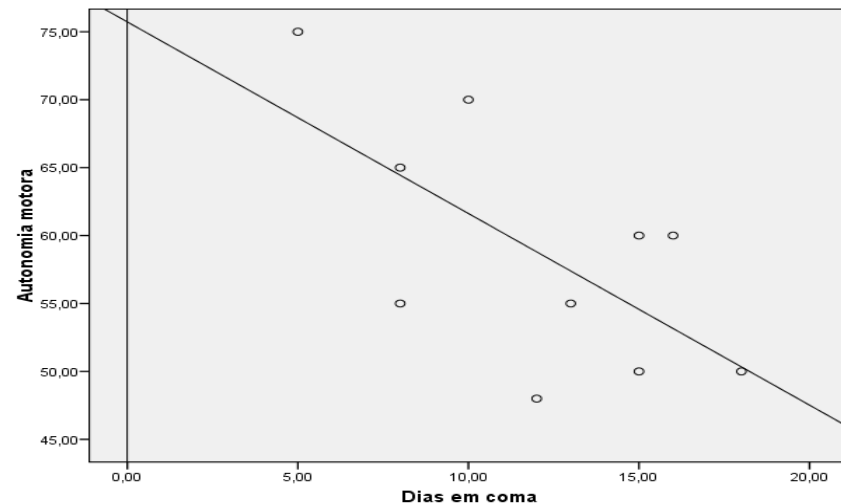
S. G.	A.F.U
20	21
24	23
24	25
27	25
28	26
30	27
32	33
36	38
37	36

$n = 9$
 $r = 0,96$
 $r^2 = 0,92$
 $(r^2\% = 91,52\%)$
 $t=8,69: p=0,000$



D. Coma	A.M.
5	75
8	65
8	55
10	70
16	60
12	48
15	50
13	55
18	50
15	60

$n = 10$
 $r = -0,65$
 $r^2 = 0,43$
 $(r^2\% = 42,52\%)$
 $t=-2,43: p=0,041$



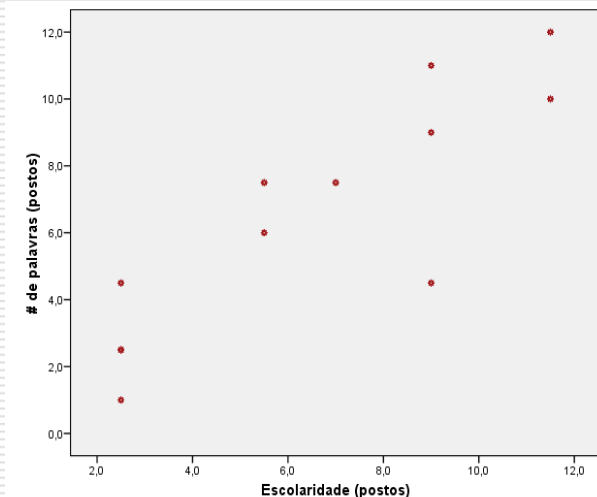
Análise da Associação entre Duas Variáveis

❖ Medida de Associação entre Variáveis Ordinais (Exemplo: Correlação de Spearman)

Doente	Escolaridade	Posto	Palavras	Posto
A	1º Ciclo	2,5	5	1,0
B	Secundário	9,0	8	4,5
C	3º Ciclo	7,0	10	7,5
D	Superior	11,5	12	10,0
E	1º Ciclo	2,5	6	2,5
F	1º Ciclo	2,5	8	4,5
G	2º Ciclo	5,5	10	7,5
H	Superior	11,5	15	12,0
I	Secundário	9,0	14	11,0
J	2º Ciclo	5,5	9	6,0
L	1º Ciclo	2,5	6	2,5
M	Secundário	9,0	11	9,0

$$r_s = 1 - \frac{6 \times 37,5}{12^3 - 12} = 0,87$$

$$r_s = 1 - \frac{6 \sum d^2}{n^3 - n}$$



▲ Com várias observações com ordens empatadas, recomenda-se a utilização de uma fórmula própria.

Problemas

1 – Considerando os seguintes pares de dados relativos às variáveis “hemoglobina glicosilada” (X) e “insulina plasmática” (Y) de uma amostra de 20 doentes diabéticos tipo II,

X (%)	6	7	5	10	8	6	6	9	12	6	8	7	5	9	12	10	9	5	11	8
Y (μU/ml)	20	18	15	6	10	20	23	15	5	15	12	14	20	11	4	8	12	25	6	16

- a) Calcular o coeficiente de correlação das duas variáveis na amostra.
- b) Testar a hipótese de que “a hemoglobina glicosilada está negativamente correlacionada com insulina plasmática nos doentes com diabetes tipo II”.

R(a): $r = -0,893$

R(b): $H_0: \rho = 0$ com $H_1: \rho < 0$

$t_o = -8,418$; $t_{c(gl=18; \alpha=0,05)} = -1,734$. Rejeita-se H_0 . Os dados da amostra corroboram a hipótese de que “a hemoglobina glicosilada está negativamente correlacionada com insulina plasmática nos doentes com diabetes tipo II”.

TESTES DE HIPÓTESES BIVARIADAS

B - Hipóteses de relação (associação) entre variáveis

NÃO PARAMÉTRICOS

Variáveis ordinais (ou que não satisfazem as condições dos testes paramétricos)

**Teste t de significância da Correlação de Spearman
Kendall's Tau**

...

Variáveis nominais (ou organizadas em categorias para efeitos de análise estatística)

**Teste Qui-quadrado (χ^2)
Teste exacto de Fisher**

Exemplos:

Teste de Significância de Correlação de Spearman

Variáveis ordinais (ou que não satisfazem as condições dos testes paramétricos)

Teste t de significância da Correlação de Spearman

$$r_s = 1 - \frac{6 \sum d^2}{n^3 - n}$$

Para $n > 10$:

$$t_o = r_s \sqrt{\frac{n - 2}{1 - r_s^2}}$$

Problema

1 – Considerando os seguintes pares de dados relativos às variáveis “hemoglobina glicosilada” (X) e “insulina plasmática” (Y) de uma amostra de 20 doentes diabéticos tipo II,

X (%)	6	7	5	10	8	6	6	9	12	6	8	7	5	9	12	10	9	5	11	8
Y (μ U/ml)	20	18	15	6	10	20	23	15	5	15	12	14	20	11	4	8	12	25	6	16

Supondo que **não se assume a normalidade** das variáveis na população,

- Calcular o coeficiente de correlação das duas variáveis na amostra.
 - Testar a hipótese de que “a hemoglobina glicosilada está negativamente correlacionada com insulina plasmática nos doentes com diabetes tipo II”.
-

Resposta

X	Y	RankX	RankY	Dif.	Dif ²
6	20	5,5	17,0	11,50	132,25
7	18	8,5	15,0	6,50	42,25
5	15	2,0	12,0	10,00	100,00
10	6	16,5	3,5	-13,00	169,00
8	10	11,0	6,0	-5,00	25,00
6	20	5,5	17,0	11,50	132,25
6	23	5,5	19,0	13,50	182,25
9	15	14,0	12,0	-2,00	4,00
12	5	19,5	2,0	-17,50	306,25
6	15	5,5	12,0	6,50	42,25
8	12	11,0	8,5	-2,50	6,25
7	14	8,5	10,0	1,50	2,25
5	20	2,0	17,0	15,00	225,00
9	11	14,0	7,0	-7,00	49,00
12	4	19,5	1,0	-18,50	342,25
10	8	16,5	5,0	-11,50	132,25
9	12	14,0	8,5	-5,50	30,25
5	25	2,0	20,0	18,00	324,00
11	6	18,0	3,5	-14,50	210,25
8	16	11,0	14,0	3,00	9,00
					2466,00

$$r_s = 1 - \frac{6 \sum d^2}{n^3 - n}$$

$$r_s = 1 - \frac{6 \times 2466}{20^3 - 20} = -0,854$$

$$t_o = r_s \sqrt{\frac{n-2}{1-r_s^2}} = -6,96$$

R: Rejeita-se H0: A amostra fornece evidência empírica suficiente para se poder afirmar, com um erro tipo I máximo de 5%, que "a hemoglobina glicosilada está negativamente correlacionada com insulina plasmática nos doentes com diabetes tipo II".

Teste do Qui-Quadrado

Diferenças de proporções entre amostras independentes / Relação entre variáveis nominais

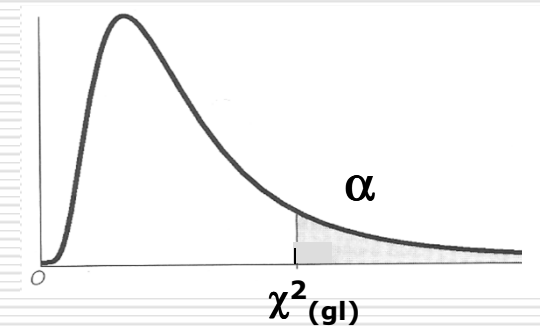
Critérios de Cochran:

Tabelas 2x2 [gl = 1]

- 1 Nenhuma frequência esperada inferior a 5;
- 2 Com $n \gg 40$, utilizar Qui-quadrado (χ^2);
- 3 Com $n > 40$ utilizar Qui-quadrado (χ^2) ou Qui-quadrado com correcção de continuidade (χ_c^2);
- 4 Se $20 \leq n \leq 40$, utilizar Qui-quadrado com correcção de continuidade (χ_c^2);
- 5 Se $n < 20$ ou alguma frequência esperada inferior a 5, utilizar o teste exacto de Fisher.

Tabelas LxC [gl = (L-1)x(C-1)]

Utilizar o Qui-quadrado (χ^2) se no máximo 20% das células com frequência esperada inferior a 5 e nenhuma inferior a 1.



$$\chi^2 = \sum \frac{(F_O - F_E)^2}{F_E}$$

$$\chi_c^2 = \sum \frac{(|F_O - F_E| - 0,5)^2}{F_E}$$



$$F_E = \frac{\text{Total da linha} \times \text{Total da coluna}}{\text{Total dos totais}}$$

$$GL = (L-1)(C-1)$$

Análise da Associação entre Duas Variáveis Nominais

❖ Exemplos:

Características do parto * Paridade					
			Paridade		Total
			Múltipara	Primipara	
Características do parto	cesariana	Fo	2	8	10
		Fe	4,44	5,56	10,00
		%	8,33	26,67	18,52
		Resíduos (Fo-Fe)	-2,44	2,44	
	via vaginal	Fo	22	22	44
		Fe	19,56	24,44	44,00
		%	91,67	73,33	81,48
		Resíduos (Fo-Fe)	2,44	-2,44	
Total	Fo	24	30	54	
	Fe	24,00	30,00	54,00	
	%	100,00	100,00	100,00	

$$V_{\text{Cramer}} = \sqrt{\frac{\chi^2}{n[\min.(C, L) - 1]}} \quad \Theta$$

$$\chi^2 = \sum \frac{(Fo - Fe)^2}{Fe}$$

$$Fe_{ij} = \frac{\sum L_i \times \sum C_j}{n}$$

$$\Phi = 0,235$$

$$\text{Odds R\acute{a}tio} = 0,25$$

Θ Em tabelas de L=2 e C=2 (2x2), designa-se por Φ (phi).

$$\Phi = \sqrt{\frac{\chi^2}{n}}$$

Problema

- 1 – Vários estudos tem evidenciado a elevada prevalência de lombalgias nas crianças em idade escolar e tem-se colocado a hipótese de que esse fenómeno se relaciona com o modo habitual de transportar o material escolar. Para verificar essa hipótese, realizou-se um estudo amostral, cujos dados se apresentam tabelados:

	Modo habitual de transporte			
	Mochila com rodas	Mochila nos dois ombros	Outro	Total
C/ lombalgias	12 (18,46%) -	19 (20,21%) -	34 (29,31%) +	65 (23,64%)
S/ lombalgias	53 (81,54%) +	75 (79,79%) +	82 (70,69%) -	210 (76,33%)
Total	65 (100,00%)	94 (100,00%)	116 (100,00%)	275 (100,00%)

R: Pela aplicação do teste de Qui-Quadrado, verifica-se que a evidência não é suficiente para sustentar a hipótese de relação entre as duas variáveis ($\chi_0^2=3,60 < \chi_c^2_{(gl=2; \alpha=0,05)}=5,991$).

Teste exacto de Fisher *

Relação entre variáveis nominais tabelas de 2x2

	Grupo A	Grupo B
-	A	B
+	C	D

$$\sum_{x=x_0}^{x=0} p(X = x_i)$$

$$p(X = x_i) = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{n!A!B!C!D!}$$

Exemplo:

	Grupo A	Grupo B	
-	9	1	10
+	5	4	9
	14	5	19

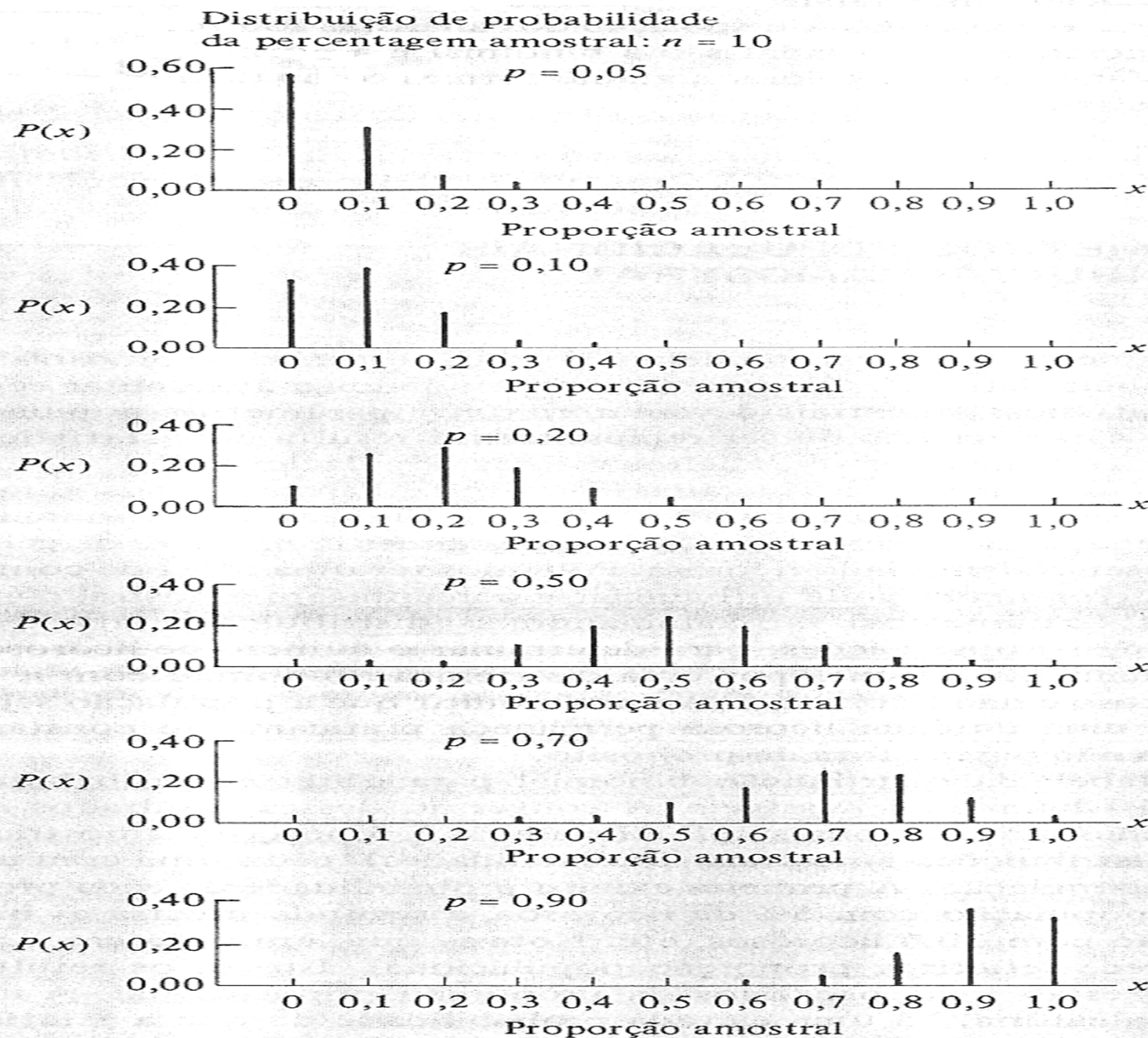
Situação observada (X=1)

	Grupo A	Grupo B	
-	10	0	10
+	4	5	9
	14	5	19

Situação mais extrema (X=0)

R: $p(X=1)+p(X=0)=0,0108+0,1084=0,1192$
 Relação não significativa ($p>0,05$).

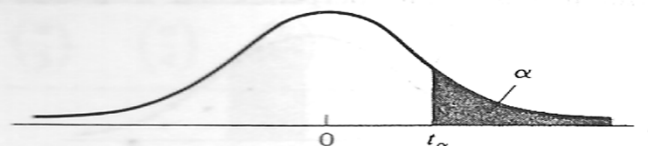
* Alternativa para o teste Qui-quadrado em tabelas de 2x2. Baseia-se na distribuição hipergeométrica.



Dispersão amostral relacionada com a dispersão da variável na população

TABELAS *

Distribuições t-Student



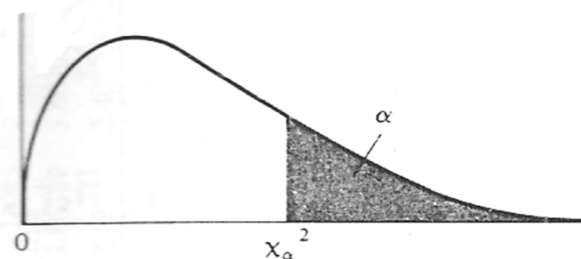
(Ou $\alpha/2$ nos intervalos de confiança e testes bilaterais)

Esta tabela dá os valores de t_α que correspondem a uma área α na cauda direita (superior) e a um número específico de graus de liberdade.

Graus de liberdade	Na cauda superior área α							
	.1	.05	.025	.01	.005	.0025	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	1.371	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Fonte: De Ronald A. Fisher: *Statistical Methods for Research Workers*, 14th ed., copyright © 1970 University of Adelaide.

Distribuições Qui-quadrado



0.10	0.05	0.025	0.01	0.005	P g.l.
2.71	3.84	5.02	6.63	7.88	1
4.61	5.99	7.38	9.21	10.60	2
6.25	7.81	9.35	11.34	12.84	3
7.78	9.49	11.14	13.28	14.86	4
9.24	11.07	12.83	15.09	16.75	5
10.64	12.59	14.45	16.81	18.55	6
12.02	14.07	16.01	18.48	20.3	7
13.36	15.51	17.53	20.1	22.0	8
14.68	16.92	19.02	21.7	23.6	9
15.99	18.31	20.5	23.2	25.2	10
17.28	19.68	21.9	24.7	26.8	11
18.55	21.0	23.3	26.2	28.3	12
19.81	22.4	24.7	27.7	29.8	13
21.1	23.7	26.1	29.1	31.3	14
22.3	25.0	27.5	30.6	32.8	15
23.5	26.3	28.8	32.0	34.3	16
24.8	27.6	30.2	33.4	35.7	17
26.0	28.9	31.5	34.8	37.2	18
27.2	30.1	32.9	36.2	38.6	19
28.4	31.4	34.2	37.6	40.0	20
29.6	32.7	35.5	38.9	41.4	21
30.8	33.9	36.8	40.3	42.8	22
32.0	35.2	38.1	41.6	44.2	23
33.2	36.4	39.4	43.0	45.6	24
34.4	37.7	40.6	44.3	46.9	25
35.6	38.9	41.9	45.6	48.3	26
36.7	40.1	43.2	47.0	49.6	27
37.9	41.3	44.5	48.3	51.0	28
39.1	42.6	45.7	49.6	52.3	29
40.3	43.8	47.0	50.9	53.7	30
51.8	55.8	59.3	63.7	66.8	40
63.2	67.5	71.4	76.2	79.5	50
74.4	79.1	83.3	88.4	92.0	60
85.5	90.5	95.0	100.4	104.2	70
96.6	101.9	106.6	112.3	116.3	80
107.6	113.1	118.1	124.1	128.3	90
118.5	124.3	129.6	135.8	140.2	100

Valores de F para o nível de significância (α) de 0,05

		(gl para o numerador)							
gl		1	2	3	4	5	6	8	12
(gl para o denominador)	1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91
	11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79
	12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69
	13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18
	25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16
	26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13
	28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12
	29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09
	40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00
	60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92
	120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83
	∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.75

FONTE: R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, 4th ed, Oliver & Boyd, Edinburgh, Table V, com a permissão dos autores e do editor.

TABELA Probabilidades Associadas ao U de Mann-Whitney.

(Referem-se a uma H_a unicaudal. Para uma H_a bicaudal, dobrar as probabilidades.)

$n_2 = 3$			
$n_1 \backslash U$	1	2	3
0	.250	.100	.050
1	.500	.200	.100
2	.750	.400	.200
3		.600	.350
4			.500
5			.650

$n_2 = 4$				
$n_1 \backslash U$	1	2	3	4
0	.200	.067	.028	.014
1	.400	.133	.057	.029
2	.600	.267	.114	.057
3		.400	.200	.100
4		.600	.314	.171
5			.429	.243
6			.571	.343
7				.443
8				.557

$n_2 = 5$					
$n_1 \backslash U$	1	2	3	4	5
0	.167	.047	.018	.008	.004
1	.333	.095	.036	.016	.008
2	.500	.190	.071	.032	.016
3	.667	.286	.125	.056	.028
4		.429	.196	.095	.048
5		.571	.286	.143	.075
6			.393	.206	.111
7			.500	.278	.155
8			.607	.365	.210
9				.452	.274
10				.548	.345
11					.421
12					.500
13					.579

$n_2 = 6$						
$n_1 \backslash U$	1	2	3	4	5	6
0	.143	.036	.012	.005	.002	.001
1	.286	.071	.024	.010	.004	.002
2	.428	.143	.048	.019	.009	.004
3	.571	.214	.083	.033	.015	.008
4		.321	.131	.057	.026	.013
5		.429	.190	.086	.041	.021
6		.571	.274	.129	.063	.032
7			.357	.176	.089	.047
8			.452	.238	.123	.066
9			.548	.305	.165	.090
10				.381	.214	.120
11				.457	.268	.155
12				.545	.331	.197
13					.396	.242
14					.465	.294
15					.535	.350
16						.409
17						.469
18						.531

TABELA Teste U de Mann-Witney

(Tabela válida para: $\alpha = 0,05$ se $H_a \rightarrow$ unicaudal
 $\alpha = 0,10$ se $H_a \rightarrow$ bicaudal)

$n_1 \backslash n_2$	9	10	11	12	13	14	15	16	17	18	19	20
1											0	0
2	1	1	1	2	2	2	3	3	3	4	4	4
3	3	4	5	5	6	7	7	8	9	9	10	11
4	6	7	8	9	10	11	12	14	15	16	17	18
5	9	11	12	13	15	16	18	19	20	22	23	25
6	12	14	16	17	19	21	23	25	26	28	30	32
7	15	17	19	21	24	26	28	30	33	35	37	39
8	18	20	23	26	28	31	33	36	39	41	44	47
9	21	24	27	30	33	36	39	42	45	48	51	54
10	24	27	31	34	37	41	44	48	51	55	58	62
11	27	31	34	38	42	46	50	54	57	61	65	69
12	30	34	38	42	47	51	55	60	64	68	72	77
13	33	37	42	47	51	56	61	65	70	75	80	84
14	36	41	46	51	56	61	66	71	77	82	87	92
15	39	44	50	55	61	66	72	77	83	88	94	100
16	42	48	54	60	65	71	77	83	89	95	101	107
17	45	51	57	64	70	77	83	89	96	102	109	115
18	48	55	61	68	75	82	88	95	102	109	116	123
19	51	58	65	72	80	87	94	101	109	116	123	130
20	54	62	69	77	84	92	100	107	115	123	130	138

FONTE: *Bulletin of the Institute of Educational Research at Indiana University*, 1, Nº 2.

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078	.0037	.0016	.0006	.0002	.0001	.0000	.0000	.0000	.0000
	1	.9556	.8503	.7166	.5767	.4449	.3294	.2338	.1586	.1024	.0625	.0357	.0188	.0090	.0038	.0013	.0004	.0001	.0000	.0000
	2	.9962	.9743	.9262	.8520	.7564	.6471	.5323	.4199	.3164	.2266	.1529	.0963	.0556	.0288	.0129	.0047	.0012	.0002	.0000
	3	.9998	.9973	.9879	.9667	.9294	.8740	.8002	.7102	.6083	.5000	.3917	.2898	.1998	.1260	.0706	.0333	.0121	.0027	.0002
	4	1.0000	.9998	.9988	.9953	.9871	.9712	.9444	.9037	.8471	.7734	.6836	.5801	.4677	.3529	.2436	.1480	.0738	.0257	.0038
	5	1.0000	1.0000	.9999	.9996	.9987	.9962	.9910	.9812	.9643	.9375	.8976	.8414	.7662	.6706	.5551	.4233	.2834	.1497	.0444
	6	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9984	.9963	.9922	.9848	.9720	.9510	.9176	.8665	.7903	.6794	.5217	.3017
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039	.0017	.0007	.0002	.0001	.0000	.0000	.0000	.0000	.0000
	1	.9428	.8131	.6572	.5033	.3671	.2553	.1691	.1064	.0632	.0352	.0181	.0085	.0036	.0013	.0004	.0001	.0000	.0000	.0000
	2	.9942	.9619	.8948	.7969	.6785	.5518	.4278	.3154	.2201	.1445	.0885	.0498	.0253	.0113	.0042	.0012	.0002	.0000	.0000
	3	.9996	.9950	.9786	.9437	.8862	.8059	.7064	.5941	.4470	.3633	.2604	.1737	.1061	.0580	.0273	.0104	.0029	.0004	.0000
	4	1.0000	.9996	.9971	.9896	.9727	.9420	.8939	.8263	.7396	.6367	.5230	.4059	.2936	.1941	.1138	.0563	.0214	.0050	.0004
	5	1.0000	1.0000	.9998	.9988	.9958	.9887	.9747	.9502	.9115	.8555	.7799	.6846	.5722	.4482	.3215	.2031	.1052	.0381	.0058
	6	1.0000	1.0000	1.0000	.9999	.9996	.9987	.9964	.9915	.9819	.9648	.9368	.8936	.8309	.7447	.6329	.4967	.3428	.1869	.0572
	7	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9993	.9983	.9961	.9916	.9832	.9681	.9424	.8999	.8322	.7275	.5695	.3366
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020	.0008	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	1	.9288	.7748	.5995	.4362	.3003	.1960	.1211	.0705	.0385	.0195	.0091	.0038	.0014	.0004	.0001	.0000	.0000	.0000	.0000
	2	.9916	.9470	.8591	.7382	.6007	.4628	.3373	.2318	.1495	.0898	.0498	.0250	.0112	.0043	.0013	.0003	.0000	.0000	.0000
	3	.9994	.9917	.9661	.9144	.8343	.7297	.6089	.4826	.3614	.2539	.1658	.0994	.0536	.0253	.0100	.0031	.0006	.0001	.0000
	4	1.0000	.9991	.9944	.9804	.9511	.9012	.8283	.7334	.6214	.5000	.3786	.2666	.1717	.0988	.0489	.0196	.0056	.0009	.0000
	5	1.0000	.9999	.9994	.9969	.9900	.9747	.9464	.9006	.8342	.7461	.6386	.5174	.3911	.2703	.1657	.0856	.0339	.0083	.0006
	6	1.0000	1.0000	1.0000	.9997	.9987	.9957	.9888	.9750	.9502	.9102	.8505	.7682	.6627	.5372	.3993	.2618	.1409	.0530	.0084
	7	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9986	.9962	.9909	.9805	.9615	.9295	.8789	.8040	.6997	.5638	.4005	.2252	.0712
10	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9992	.9980	.9954	.9899	.9793	.9596	.9249	.8658	.7684	.6126	.3698
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.9139	.7361	.5443	.3758	.2440	.1493	.0860	.0464	.0233	.0107	.0045	.0017	.0005	.0001	.0000	.0000	.0000	.0000	.0000
	2	.9885	.9298	.8202	.6778	.5256	.3828	.2616	.1673	.0996	.0547	.0274	.0123	.0048	.0016	.0004	.0001	.0000	.0000	.0000
	3	.9990	.9872	.9500	.8791	.7759	.6496	.5138	.3823	.2660	.1719	.1020	.0548	.0260	.0106	.0035	.0009	.0001	.0000	.0000
	4	.9999	.9984	.9901	.9672	.9219	.8497	.7515	.6331	.5044	.3770	.2616	.1662	.0949	.0473	.0197	.0064	.0014	.0001	.0000
	5	1.0000	.9999	.9986	.9936	.9803	.9527	.9051	.8338	.7384	.6230	.4956	.3669	.2485	.1503	.0781	.0328	.0099	.0016	.0001
11	6	1.0000	1.0000	.9999	.9991	.9965	.9894	.9740	.9452	.8980	.8281	.7340	.6177	.4862	.3504	.2241	.1209	.0500	.0128	.0010
	7	1.0000	1.0000	1.0000	.9999	.9996	.9984	.9952	.9877	.9726	.9453	.9004	.8327	.7384	.6172	.4744	.3222	.1798	.0702	.0115
	8	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9995	.9983	.9955	.9893	.9767	.9536	.9140	.8507	.7560	.6242	.4557	.2639	.0861
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9990	.9975	.9940	.9865	.9718	.9437	.8926	.8031	.6513	.4013	.1013
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.8981	.6974	.4922	.3221	.1971	.1130	.0606	.0302	.0139	.0059	.0022	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	2	.9848	.9104	.7788	.6174	.4552	.3127	.2001	.1189	.0652	.0327	.0148	.0059	.0020	.0006	.0001	.0000	.0000	.0000	.0000
	3	.9984	.9815	.9306	.8389	.7133	.5696	.4256	.2963	.1911	.1133	.0610	.0293	.0122	.0043	.0012	.0002	.0000	.0000	.0000
	4	.9999	.9972	.9841	.9496	.8854	.7897	.6683	.5328	.3971	.2744	.1738	.0994	.0501	.0216	.0076	.0020	.0003	.0000	.0000
	5	1.0000	.9997	.9973	.9883	.9657	.9218	.8513	.7535	.6331	.5000	.3669	.2465	.1487	.0782	.0343	.0117	.0027	.0003	.0000

TABELA A-8 Valores Críticos de T para o Teste de Postos com Sinais de Wilcoxon

n	α			
	0,005 unilateral 0,01 bilateral	0,01 unilateral 0,02 bilateral	0,025 unilateral 0,05 bilateral	0,05 unilateral 0,10 bilateral
5	*	*	*	1
6	*	*	1	2
7	*	0	2	4
8	0	2	4	6
9	2	3	6	8
10	3	5	8	11
11	5	7	11	14
12	7	10	14	17
13	10	13	17	21
14	13	16	21	26
15	16	20	25	30
16	19	24	30	36
17	23	28	35	41
18	28	33	40	47
19	32	38	46	54
20	37	43	52	60
21	43	49	59	68
22	49	56	66	75
23	55	62	73	83
24	61	69	81	92
25	68	77	90	101
26	76	85	98	110
27	84	93	107	120
28	92	102	117	130
29	100	111	127	141
30	109	120	137	152

NOTAS:

1. * indica que não é possível obter um valor na região crítica.
2. Rejeitar a hipótese nula se a estatística de teste T é no máximo igual ao valor crítico da tabela. Não rejeitar a hipótese nula em caso contrário.

De *Some Rapid Approximate Statistical Procedures*, Copyright © 1949, 1964 Lederle Laboratories Division of American Cyanamid Company. Reproduzido com autorização de American Cyanamid Company.