Investigação Aplicada I

Aula 6

1º Semestre 2016/17

Licenciatura em Ciências Biomédicas Laboratoriais

igrodrigues@ualg.pt; ESSUAlg: gabinete 2.06

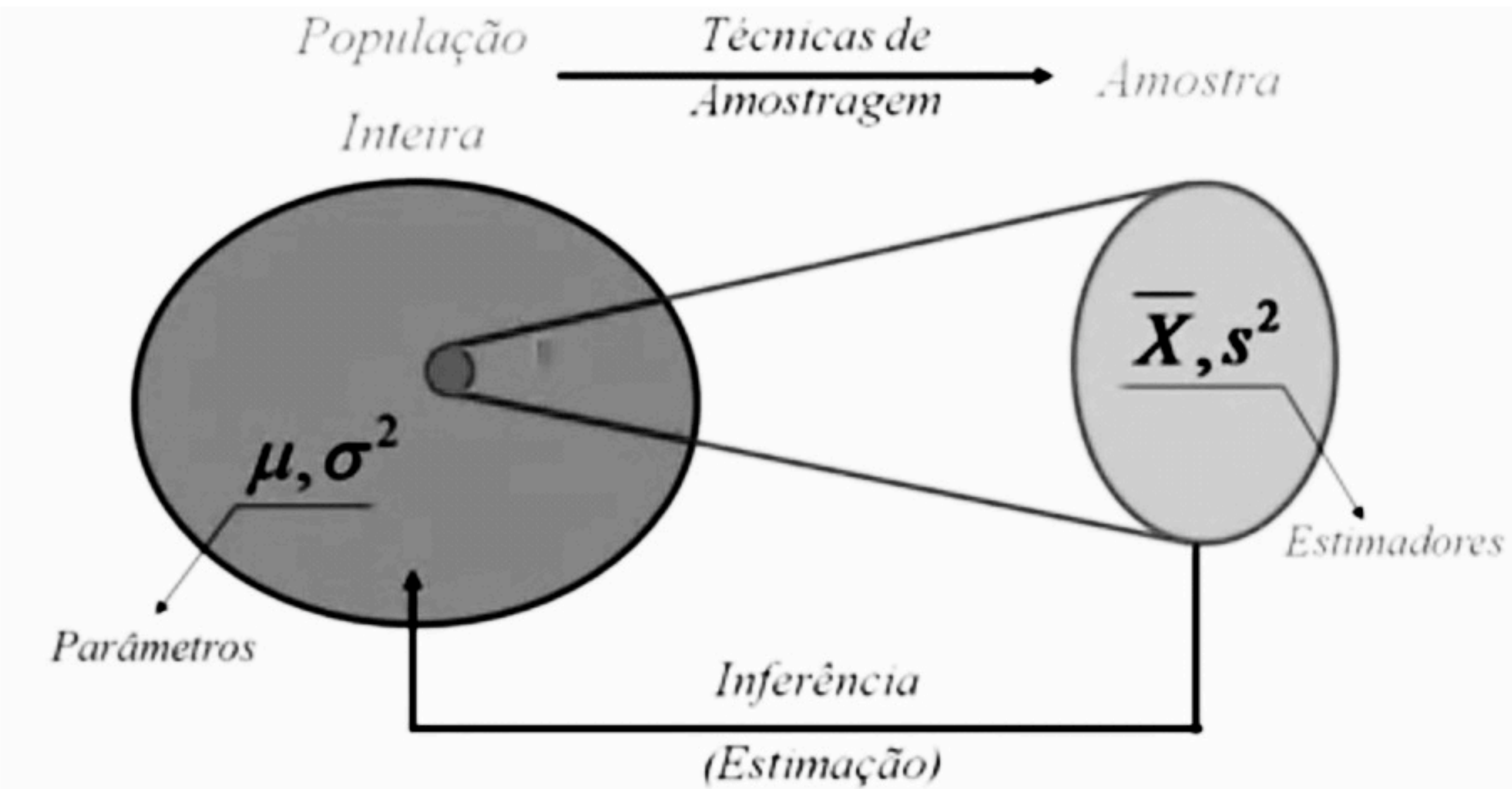
Prof. Inês Rodrigues

Inferência Estatística

- O que se pode dizer acerca de uma população com base numa amostra?
- Perguntas entre grupos de dados
- População: média e variância populacional através da amostra (técnicas de amostragem)
- A partir daqui conseguimos uma estimativa da media e uma estimativa da variância através deste estimadores consegue saber a media e variância da população

Uma amostra é usada para obter resultados representativos de uma população alvo, possíveis de ser generalizados à população em estudo.

Inferência Estatística



Inferência Estatística

Amostragem

- Especificação da população a medir
- Recolha da amostra e realização das medidas
- Calculo das estatísticas, o que de certa forma faz afirmações quanto aos parâmetros da população

Inferência Estatística

Amostragem

- A população está definida sem ambiguidades?
- A variável é claramente observável?
- A amostra é válida?
- O tamanho da amostra é suficiente?

Inferência Estatística – Exercício-Exemplo

- O peso dos recém-nascidos de um determinado hospital em 1998 (população de interesse) pode ser estimado usando observações de uma amostra desses recém-nascidos

- 810 bebés nasceram durante todo o ano de 1998 no hospital

- Foram recolhidos dados referentes **aos primeiros** 81 recém-nascidos do ano, correspondentes ao mês de Janeiro (amostra).

- A média dos pesos foi calculada

Será que esta média da amostra uma boa estimativa da média de pesos da população de interesse?

Inferência Estatística – Exercício-Exemplo

- O peso dos recém-nascidos de um determinado hospital em 1998 (população de interesse) pode ser estimado usando observações de uma amostra desses recém-nascidos

- 810 bebés nasceram durante todo o ano de 1998 no hospital

- Foram recolhidos dados referentes aos primeiros 81 recém-nascidos do ano, correspondentes ao mês de Janeiro (amostra).

- A média dos pesos foi calculada

Como selecionar uma amostra representativa da população?

Que tipo de amostra foi escolhida para o estudo?

Inferência Estatística

Fazer afirmações sobre os parâmetros da população:

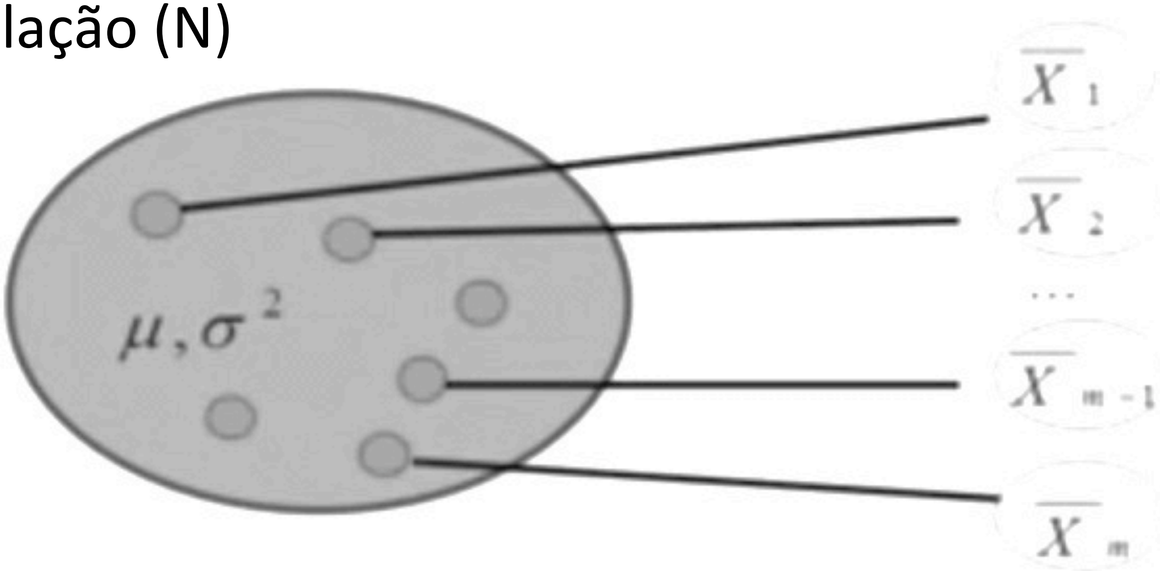
Estimação – Tentativa de estimar os valores dos parâmetros.

Teste de hipóteses - Faz deduções acerca dos parâmetros da população, supondo que eles têm determinados valores e testando depois se os valores observados são consistentes com a hipótese.

Inferência Estatística

Erros da média

População (N) Amostra (n)



Inferência Estatística

Erros da média

$$DP = \sqrt{\frac{\sum_{i=1}^{n}\left(\overline{X}_a - \overline{X}\right)^2}{n-1}}$$

$$\boxed{EPM = \frac{\sigma}{\sqrt{n}} \quad ou \quad EPM = \frac{s}{\sqrt{n}}}$$

Inferência Estatística

Teste de Hipóteses

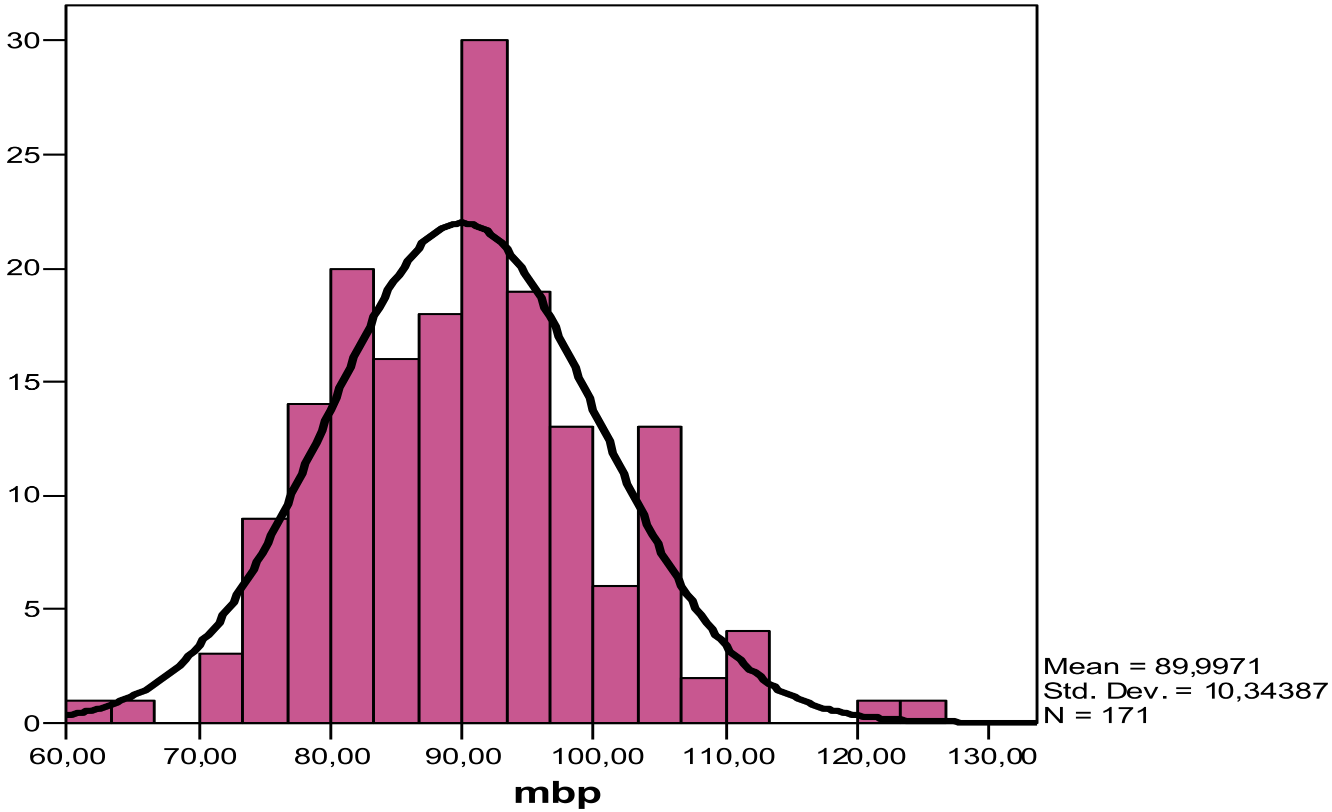
Probabilidade

Percentagem que exprime a ocorrência de um factor dentro de uma série de factores

Para testar hipóteses são estudadas as probabilidades de ocorrência de um determinado parâmetro/acontecimento

As hipóteses são formuladas pelo investigador e são testadas para averiguar se os dados suportam as hipóteses. No entanto nunca podemos provar que a hipótese é verdadeira, porque um estudo NUNCA prova NADA, há sempre a possibilidade da inclusão de erros alheios ao investigador que distorcem os dados.

Inferência Estatística – Distribuição Normal



Inferência Estatística

Teste de Hipóteses

Verificar se dois grupos são iguais ou diferentes para um certo parâmetro

Hipóteses

A = B → **Igualdade (H_0)**

$\left.\begin{array}{l}\text{A melhor que B}\\ \text{A pior que B}\\ \text{A diferente que B}\end{array}\right\}$ **Diferença (H_a)**

Inferência Estatística

Teste de Hipóteses

Hipótese Nula (H_0)
 Propõe que não existe diferença ou relação entre duas variáveis. Se for encontrada uma diferença significativa ou uma relação entre as variáveis então H0 é rejeitada

Hipótese Alternativa (H_a ou H_1)
Hipótese que contradiz H0. Indica a direção da diferença ou relação esperada.

As hipóteses estatísticas são consideradas verdadeiras ou falsas mediante a realização de estatística inferencial.

Testes ou provas de hipóteses

- Definição da hipótese nula e da hipótese alternativa:

  – Hipótese nula

  H_0 – Especifica um hipotético valor ou valores para um parâmetro

  – Hipótese alternativa

  H_1 - Especifica um hipotético valor ou valores para um parâmetro, que será considerado, caso H_0 seja rejeitada

Inferência Estatística

Erros de Inferência

	VERDADE (Realidade não conhecida)	
Conclusões tiradas a partir da amostra (PROTOCOLO)	**H_0 é verdadeira**	**H_o é falsa**
Aceitar H_0	Conclusão correta	**Erro de tipo II Probabilidade β**
Rejeitar H_0	**Erro de tipo I Probabilidade α**	Conclusão correcta

Inferência Estatística

Teste de Hipóteses :

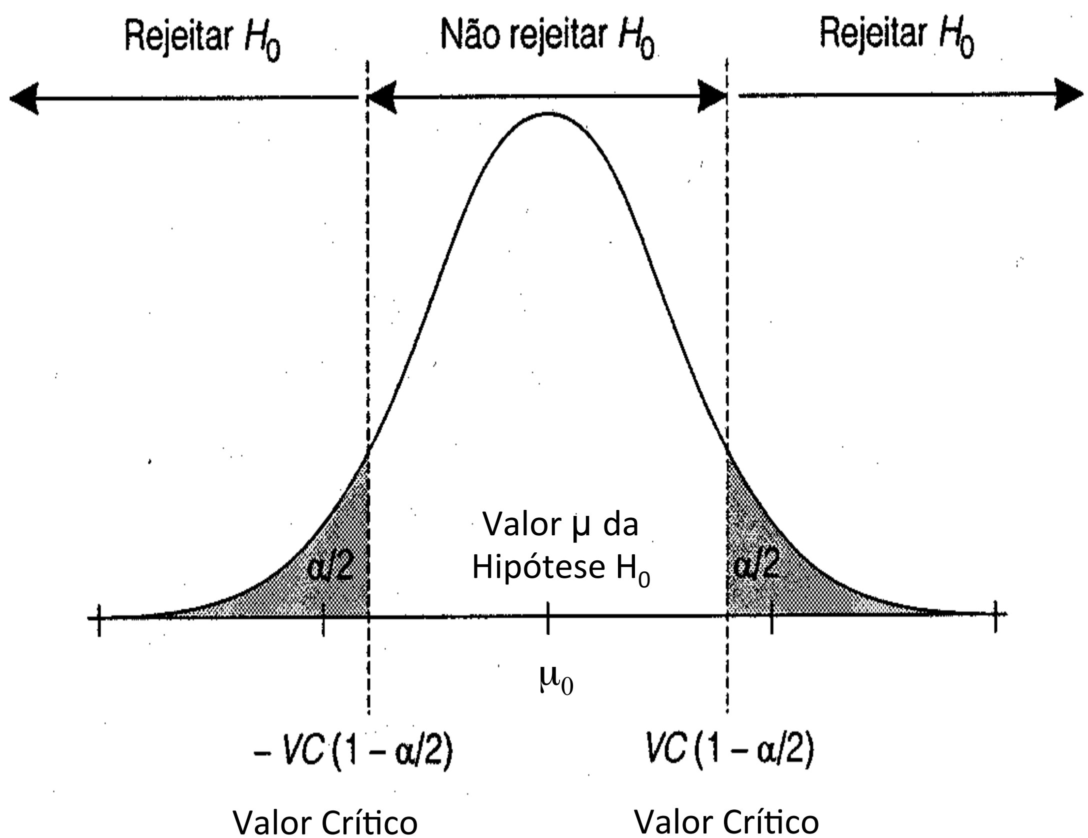
1. Formular as hipóteses de interesse
2. Escolher um nível de significância **α (alpha)**
3. Escolher o teste estatístico apropriado
4. Concluir pelo teste se rejeita ou não $\boldsymbol{H_0}$

Se a probabilidade de que a diferença seja observada ao acaso for grande (maior que **α***), não se rejeita* H_0

Se a probabilidade de que a diferença seja observada ao acaso for pequena (menos que **α***), rejeita-se* H_0

Inferência Estatística

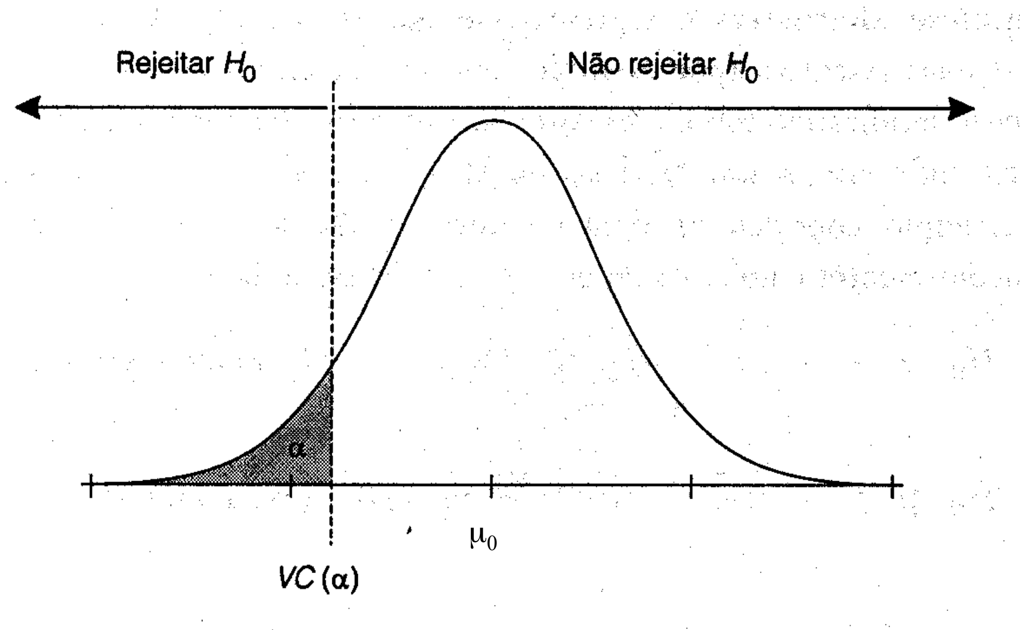
Teste de Hipóteses - Bilateral



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

Inferência Estatística

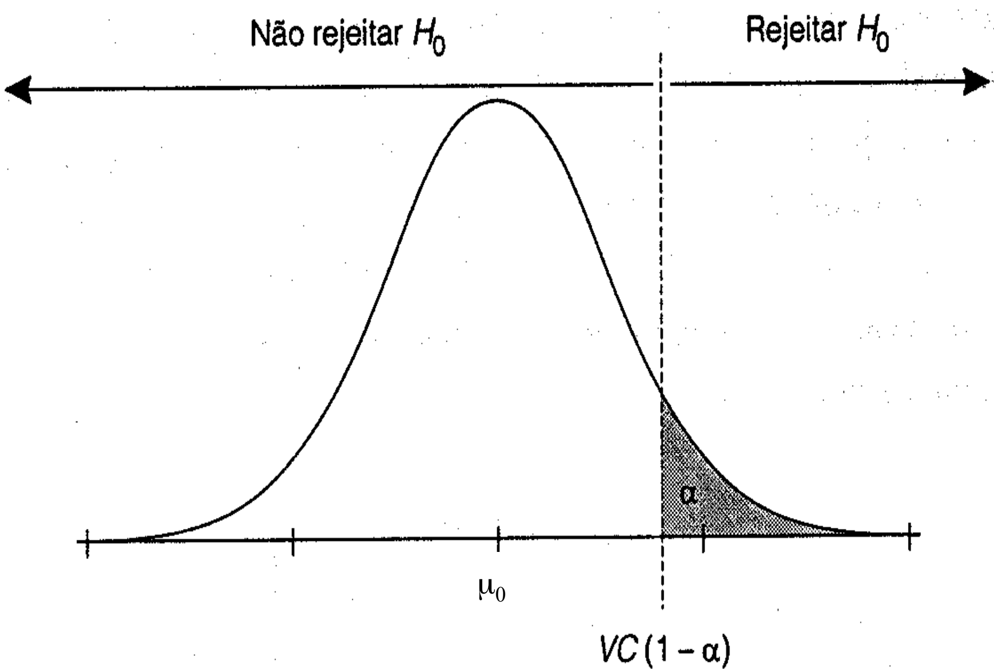
Teste de Hipóteses - Unilateral



H_0: $\mu \geq \mu_0$
H_1: $\mu < \mu_0$

Inferência Estatística

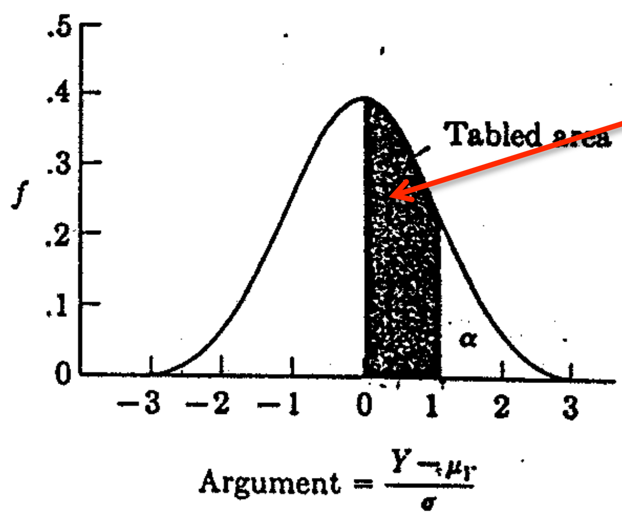
Teste de Hipóteses - Unilateral



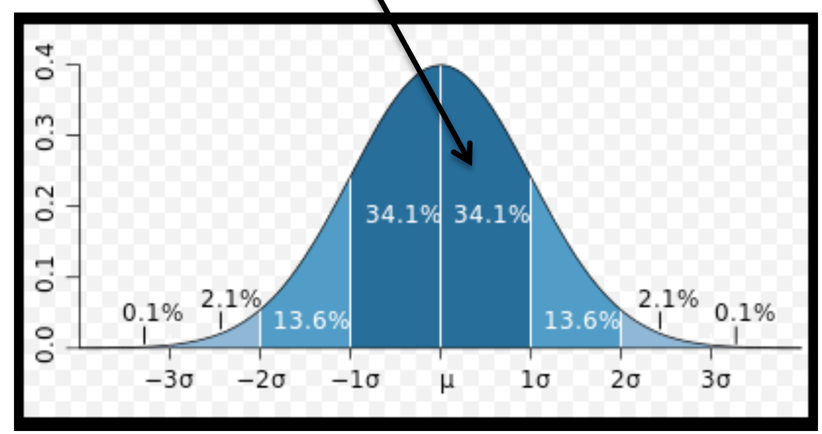
$H_0: \mu \leq \mu_0$
$H_1: \mu > \mu_0$

Inferência Estatística

TESTE Z (uma amostra)



P (0 ≤ Z ≤ 1) = 0,34



- Curva de Gauss padronizada (tabela Z)
- Conhece-se a média e o desvio padrão da população
- Já não se aplica porque normalmente estes valores não são conhecidos
- Este teste é essencial para compreender os próximos testes de distribuição

Inferência Estatística

Teste Z (1 amostra)

Z : Transforma a média das medidas em distância da média verdadeira em erro padrão

Numero **Z** → Tabela Z → Representa a probabilidade da distância da média verdadeira

$$Z = \frac{X - \mu}{\sigma / \sqrt{n}}$$

Inferência Estatística

Teste Z (1 amostras)

Exemplo:

Frequência normal dos batimentos cardíacos por minuto:
u = 69,8
σ = 1,86

Grupo de pacientes tratados com fármacos para a função cardíaca:
X = 70,5

H_0 : a média da amostra é igual á media da população

H_a : a média da amostra é diferente da média da população

Inferência Estatística

Teste Z (1 amostras)

$$Z = \frac{X - \mu}{\textbf{SEM}}$$

Permite saber quanto é que a frequência média da amostra se afasta da frequência média da população

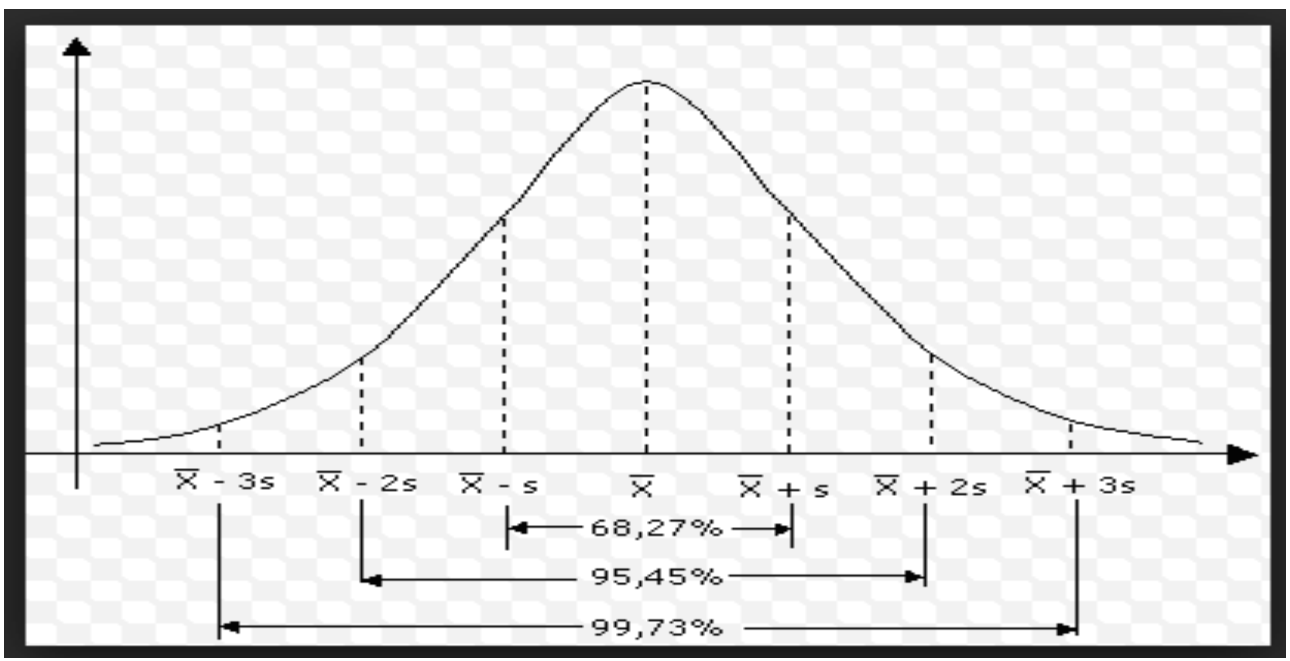
Definir um nivel de significância para o teste: 5% (**α=0,05)**

Inferência Estatística

Teste Z (1 amostras)

$$Z = \frac{X - \mu}{\textbf{SEM}}$$

Permite saber quanto é que a frequência média da amostra se afasta da frequência média da população



Inferência Estatística

Teste Z (1 amostras)

SE:

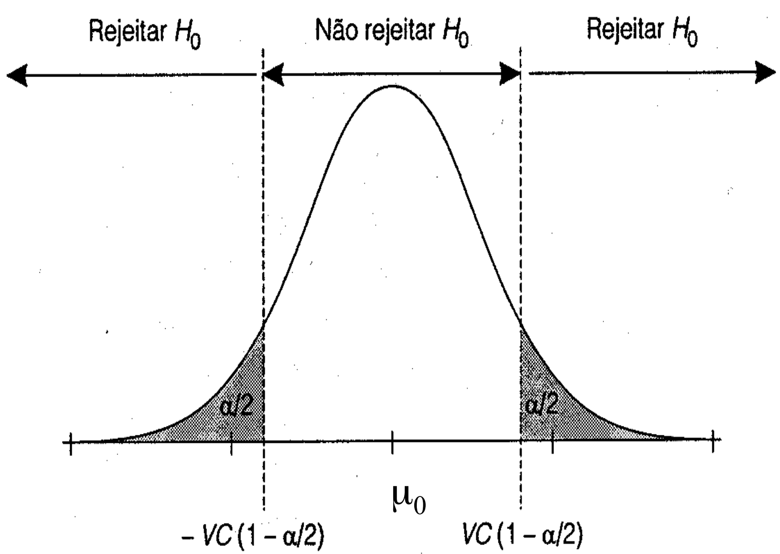
Z < -1,96 → rejeita-se H_0 → p< **α /2**

Z > 1,96 → rejeita-se H_0 → p< **α /2**

-1,96 $\leq$ Z $\leq$ 1,96 → aceita-se H_0 → p> **α /2**

P < α → rejeita-se H_0

P > α → aceita-se H_0



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

Inferência Estatística

Teste Z (1 amostras)

Mesmo Exemplo:

Frequência normal dos batimentos cardíacos por minuto:
u = 69,8
σ = 1,86

Grupo de pacientes tratados com fármacos para a função cardíaca:
X = 70,5

Hipótese diferentes!

H_0 : a média da amostra é menor ou igual á media da população

H_a : a média da amostra é maior que a média da população

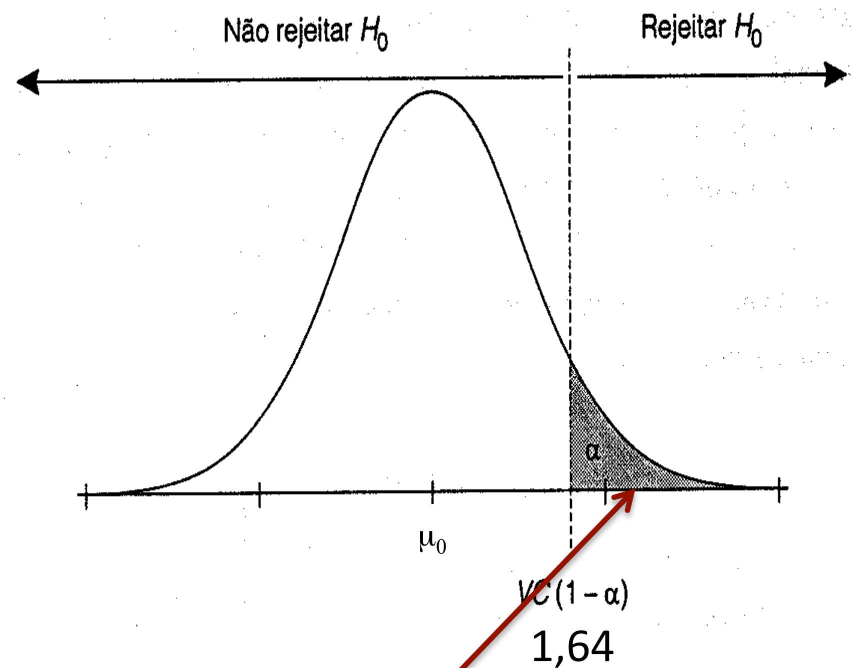
Inferência Estatística

Teste Z (1 amostras)

u = 69,8
σ = 1,86
SEM = 0,26

X = 70,5

***H₀* : X ≤ u**
***Hₐ* : X > u**



H_0: $\mu \leq \mu_0$
H_1: $\mu > \mu_0$

Z = (70,5 – 69,8)/ 0,26 = 2,69

Então: **Z** > 1,64 → Rejeita-se H_0

Inferência Estatística

Teste Z (2 amostras)

Exemplo:

Grupo Placebo (a)	**Grupo tratado (b)**
Xa = 8,5	**Xb = 10,1**
σ a = 2	σ b = 2
n = 36	n = 36

H_0 : Xa igual Xb
H_a : Xa diferente de Xb

Que nível de significância escolhemos? Queremos ser muito ou pouco rigorosos??

Exemplo: Testes à Normalidade

$$\begin{cases} H_0: \text{Os diâmetros dos leucócitos seguem distribuição Normal} \\ H_1: \text{Os diâmetros dos leucócitos não seguem distribuição Normal} \end{cases}$$

$$\begin{cases} H_0: \text{Os diâmetros das células tumorais seguem distribuição Normal} \\ H_1: \text{Os diâmetros das células tumorais não seguem distribuição Normal} \end{cases}$$

Tests of Normality

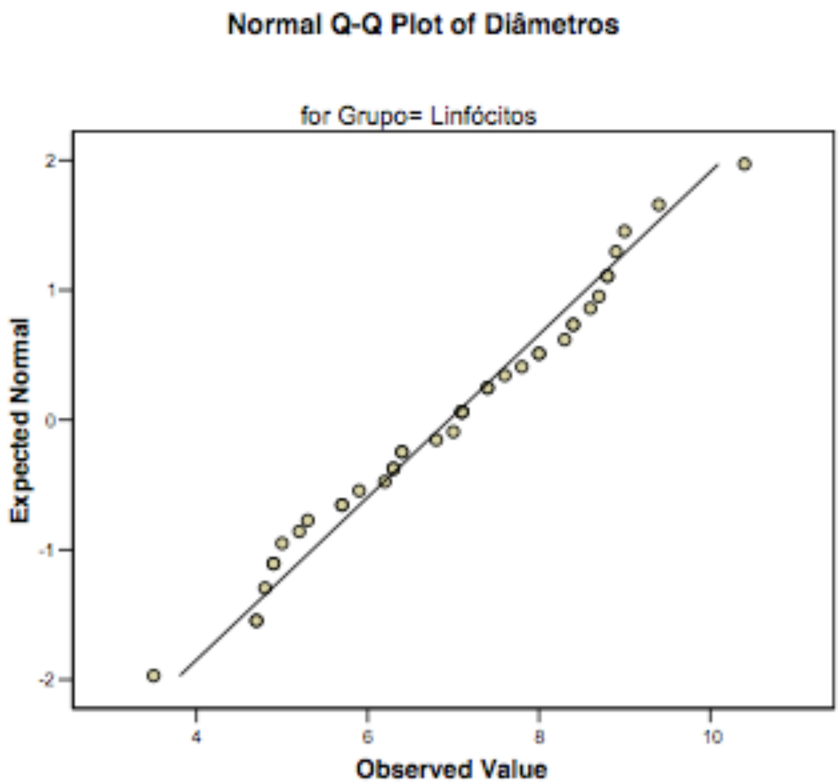
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Grupo	Statistic	df	Sig.	Statistic	df	Sig.
Diâmetros	Linfócitos	.076	40	.200*	.978	40	.605
	Células tumorais	.066	50	.200*	.987	50	.860

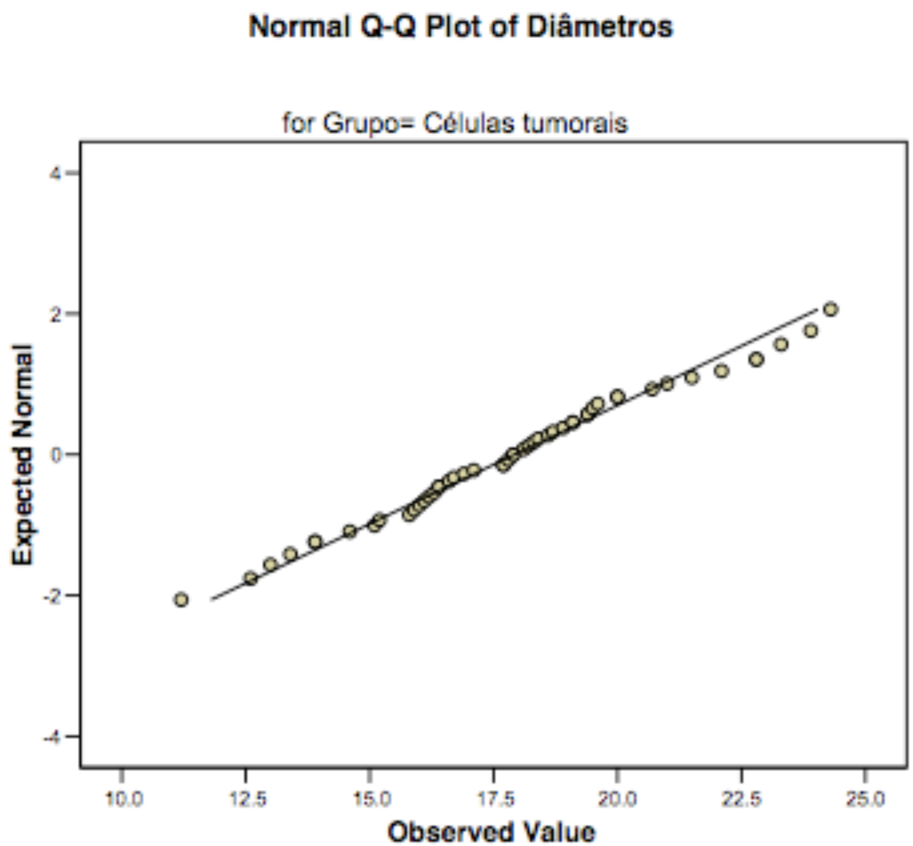
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Aceito a Hipotese H_0

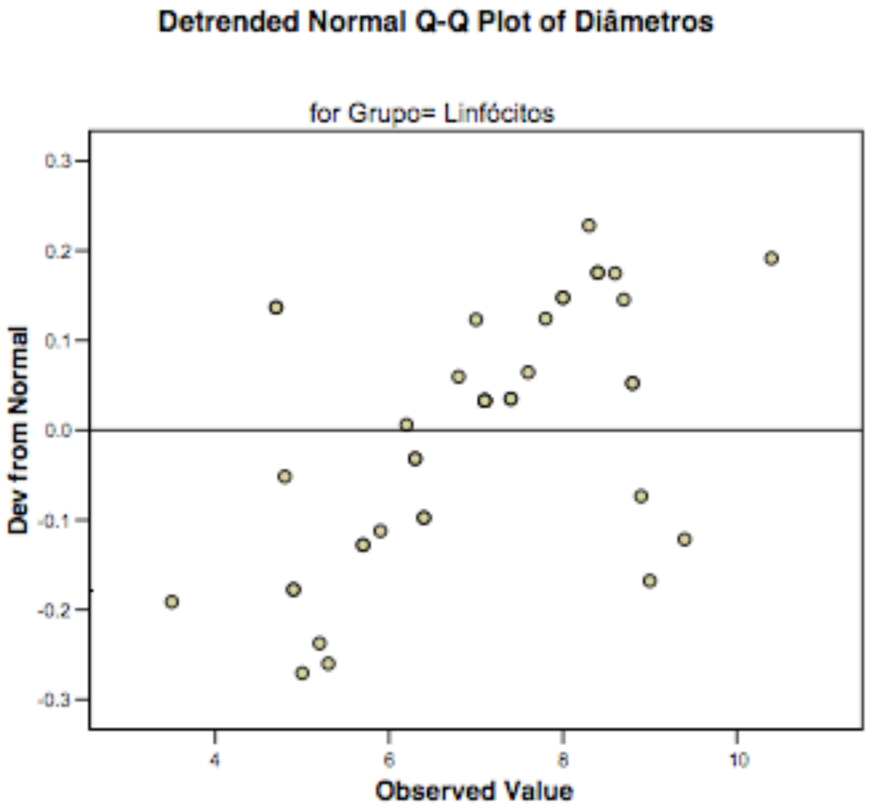
Exemplo: Testes à Normalidade

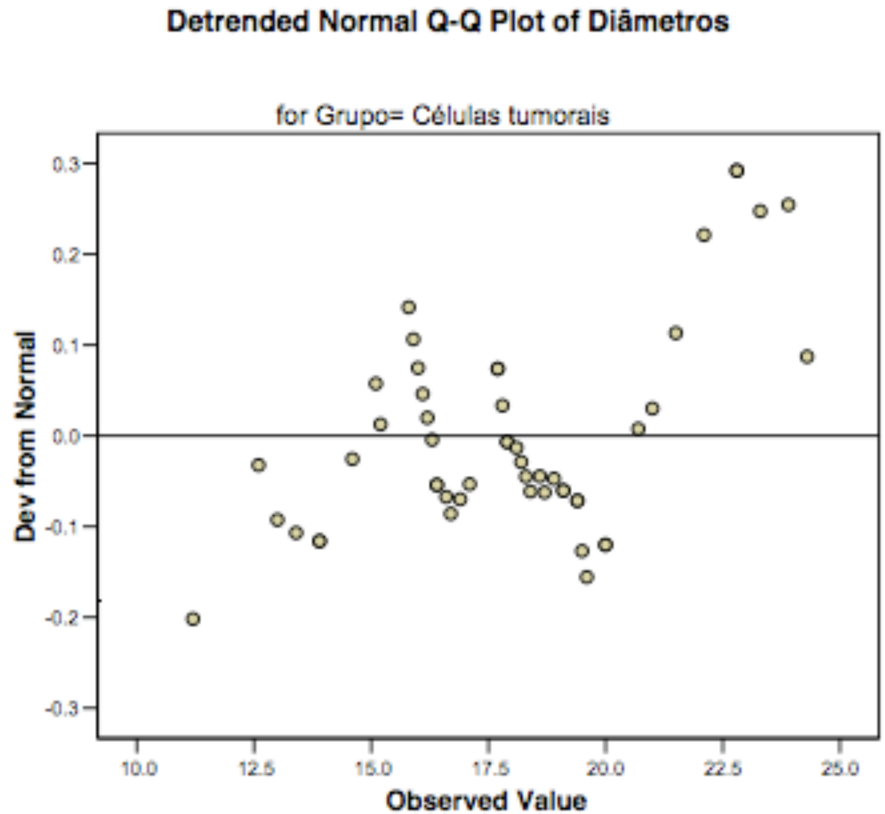




Aceito a Hipotese H_0

Exemplo: Testes à Normalidade





Aceito a Hipotese H_0

Diferença das médias

Quando substituímos o desvio padrão da população pelo desvio padrão da amostra estamos a introduzir erro.

Na maioria dos casos não se conhece o desvio padrão da população e este é substituído pelo desvio padrão da amostra

Para reduzir este erro, fazem-se testes estatísticos para a ***distribuição t***