

Investigação Aplicada I

Aula 6

1º Semestre 2020/21

Licenciatura em Ciências Biomédicas Laboratoriais

Sumário:

Inferência Estatística

Testes de Hipóteses – Probabilidade

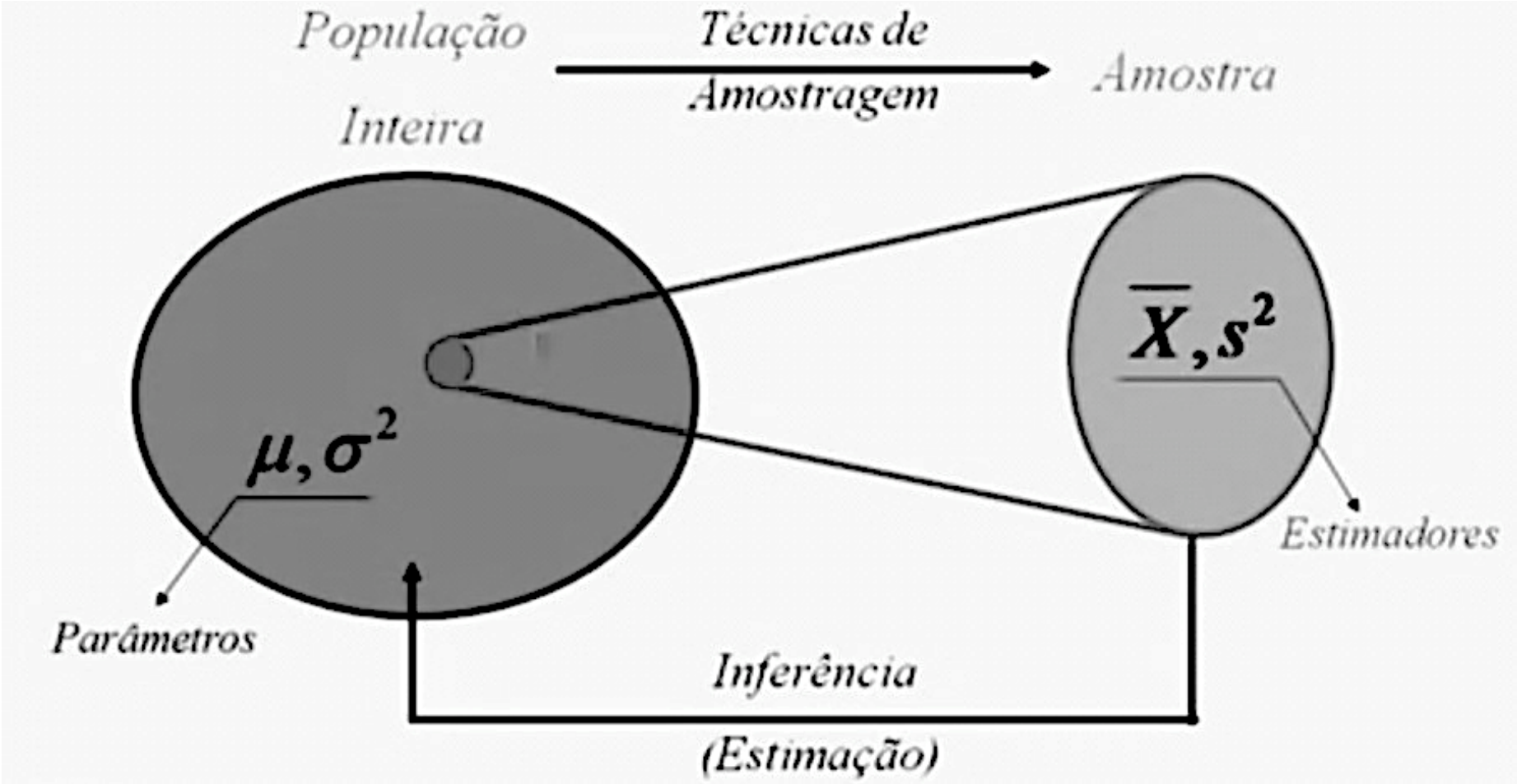
Teste Z

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.

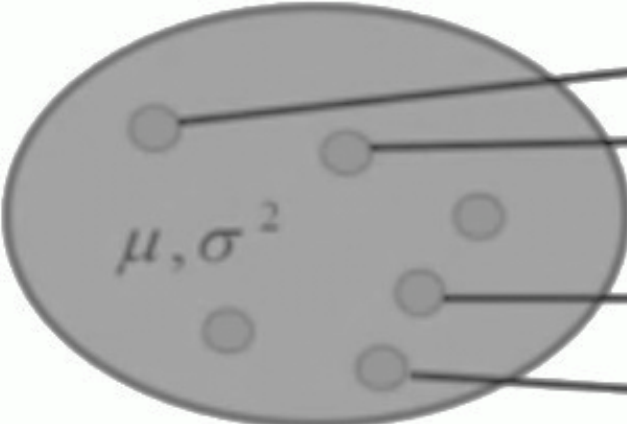
Inferência Estatística



Inferência Estatística

Erros da média

População (N)



Amostra (n)



Inferência Estatística

Amostragem

- Especificação da população a medir
- Recolha da amostra e realização das medidas
- Cálculo 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 – Amostra - 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 bebês 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 é uma boa estimativa da média de pesos da população de interesse?

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

Teste de Hipóteses - Probabilidade

- ✓ Percentagem que exprime a ocorrência de um fator dentro de uma série de fatores
- ✓ 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

Teste de Hipóteses - Probabilidade

Ao realizar um teste de hipóteses, consideram-se sempre duas hipóteses complementares:

- **Hipótese Nula (denominada por H_0)**
- **Hipóteses alternativa (denominada por H_1 ou H_A)**

A hipótese nula assume-se como verdadeira até prova do contrário

No teste verifica-se se os dados amostrais são compatíveis, dentro dos limites da razoabilidade, com a hipótese nula. Não o sendo, aceita-se como sendo mais razoável a hipótese alternativa.

Inferência Estatística

Teste de Hipóteses - Probabilidade

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

Hipóteses

$A = B \rightarrow$ **Igualdade (H_0)**

A maior que B
A menor que B
A diferente que B

Diferença (H_1 ou H_a)

Inferência Estatística

Teste de Hipóteses - Probabilidade

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 H_0 é rejeitada

Hipótese Alternativa (H_a ou H_1)

Hipótese que contradiz H_0 . Indica a direção da diferença ou relação esperada.

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

Inferência Estatística

Teste de Hipóteses - Probabilidade

- 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

Teste de Hipóteses

Existe uma grande variedade de testes de hipóteses, devendo o teste ser aplicado adequadamente, segundo a questão de investigação colocada

Tipos de Testes de Hipóteses mais comuns:

i. Testes de comparação entre grupos

Permitem comparar um parâmetro (média, mediana, frequência) entre dois ou mais grupos

Permite comparar um parâmetro de um grupo com o valor de referência

A hipótese nula geralmente expressa-se por uma igualdade

Inferência Estatística

i. Testes de comparação entre grupos – Exemplos

- Comparar as médias das idades entre um grupo de doentes e um grupo de controlos
→ comparar a média entre os grupos; a H_0 será que os doentes e os controlos têm, em média, a mesma idade.
- Comparar as medianas dos valores de pressão arterial sistólica entre pessoas com menos de 18 anos, pessoas entre 19 e 64 anos e pessoas com mais de 65 anos
→ aplica-se um teste de comparação da mediana entre os 3 grupos; a H_0 será que os grupos têm medianas de valores iguais
- Comparar o peso médio de recém-nascidos em Portugal com o valor de referência
→ A H_0 será que a média do peso dos recém-nascidos em Portugal é igual à do valor de referência
- Comparar as frequências de mulheres entre dois grupos, controlo e doentes
→ a H_0 será que a frequência de mulheres no grupo de controlo é igual à frequência de mulheres no grupo de doentes

Inferência Estatística

ii. Testes de associação entre variáveis

Testa associação entre duas Variáveis Qualitativas

→ H_0 : As variáveis em estudo não estão associadas

Exemplo: Testar se os hábitos tabágicos (ou seja, fumar ou não-fumar) estão associados com a hipertensão (ou seja, ser ou não ser hipertenso)

Inferência Estatística

iii. Testes à correlação entre variáveis

Testa correlação entre duas Variáveis Qualitativas

→ H_0 : As variáveis em estudo não estão correlacionadas

Exemplo: Testar se o peso dos recém-nascidos está correlacionado com o peso das suas mães

Estes testes poderão ainda ser aplicáveis no caso de se ter variáveis qualitativas ordinais, desde que o número de categorias seja suficiente

Inferência Estatística

Como escolher o Teste Hipóteses adequado?

Devemos ter em conta:

- Testes de relação entre variáveis quantitativas ou entre duas variáveis qualitativas
- Comparação entre grupos:
 1. Numero de grupos a comparar: 1 grupo (comparar com o valor de referencia), 2 ou mais de 2 grupos;
 2. Nível de mensuração (tipo) da variável a comparar entre grupos, isto é, quantitativa ou qualitativa (ordinal ou nominal)
 3. Relação existente entre grupos: emparelhados ou independentes

Inferência Estatística

Como escolher o Teste Hipóteses adequado?

Considerar:

Determinação do nível de mensuração:

1. Variáveis quantitativas que não seguem uma distribuição normal devem ser analisadas de forma análoga a variáveis ordinais;
2. Variáveis ordinais com poucas categorias (aproximadamente > 5) devem ser analisadas como se fossem nominais para efeito de aplicação do teste.

Inferência Estatística

Como escolher o Teste Hipóteses adequado?

Nível de mensuração / Tipo de Variável					
		Nº de grupos	Quantidade (distribuição normal)	Ordinal ou Quantitativa (sem distribuição normal)	Nominal
		1	t-student 1 amostra	Mediana 1 amostra	Binomial 1 amostra
Amostras Independentes	2		t-student independentes	Mann-Whitney	Qui-quadrado
	3		ANOVA	Kruskal-Walis	Q-Cochran
Amostras Emparelhadas	2		t-student emparelhadas	Wilcoxon	McNemar
	3		ANOVA (medidas repetidas)	Friedman	

Inferência Estatística

Como escolher o Teste Hipóteses adequado?






Nível de mensuração / Tipo de Variável

	Nº de grupos	Quantidade (distribuição normal)	Ordinal ou Quantitativa (sem distribuição normal)	Nominal
	1	t-student 1 amostra	Mediana 1 amostra	Binomial 1 amostra
Amostras Independentes	2	t-student independentes	Mann-Whitney	Qui-quadrado
	3	ANOVA	Kruskal-Walis	Q-Cochran
Amostras Emparelhadas	2	t-student emparelhadas	Wilcoxon	McNemar
	3	ANOVA (medidas repetidas)	Friedman	

Inferência Estatística

Como escolher o Teste Hipóteses adequado?

-  1. Testes à média, compara médias entre dois grupos (t-student independentes)
-  2. Testes à mediana, compara medianas entre dois grupos independentes (Mann-Whitney)
-  3. Testes à frequência/proporção, compra proporções entre dois grupos independentes (teste de independência Qui-quadrado)

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 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 - Probabilidade

A forma de interpretar um teste de hipóteses é por comparação do valor-p obtido com o valor do nível de significância (α) fixado à priori. A regra é bastante simples:

- **Se $p \geq \alpha$, não existe evidência estatística para se rejeitar a hipótese nula;**
- **Se $p \leq \alpha$, existe evidência estatística para se rejeitar a hipótese nula;**

O nível de significância usual é de 0,05 ($\alpha = 0,05$), no entanto outro valor pode ser usados, desde que devidamente justificado

Inferência Estatística

Teste de Hipóteses - Probabilidade

$p \geq \alpha$ & $p \leq \alpha$ – Erros de Inferência

	VERDADE (Realidade não conhecida)	
Conclusões tiradas a partir da amostra (PROTOCOLO)	H₀ é verdadeira	H₀ é falsa
Aceitar H₀	Conclusão correta	Erro de tipo II Probabilidade β
Rejeitar H₀	Erro de tipo I Probabilidade α	Conclusão correcta

Inferência Estatística

Teste de Hipóteses – Unilateral & Bilateral

- A hipótese nula nem sempre exprime uma igualdade.
- Deve estabelecer uma distinção entre testes de hipóteses unilaterais e bilaterais.
- Um teste de hipóteses bilateral envolve uma comparação de igualdade vs não igualdade.
- Um teste unilateral, envolve a colocação da hipóteses sob forma de desigualdade.

Inferência Estatística

Teste de Hipóteses – Unilateral & Bilateral

Exemplos:

1. Pretende-se verificar se a média do número de horas dormido por noite difere entre os trabalhadores de duas empresas diferentes (μ_A e μ_B).

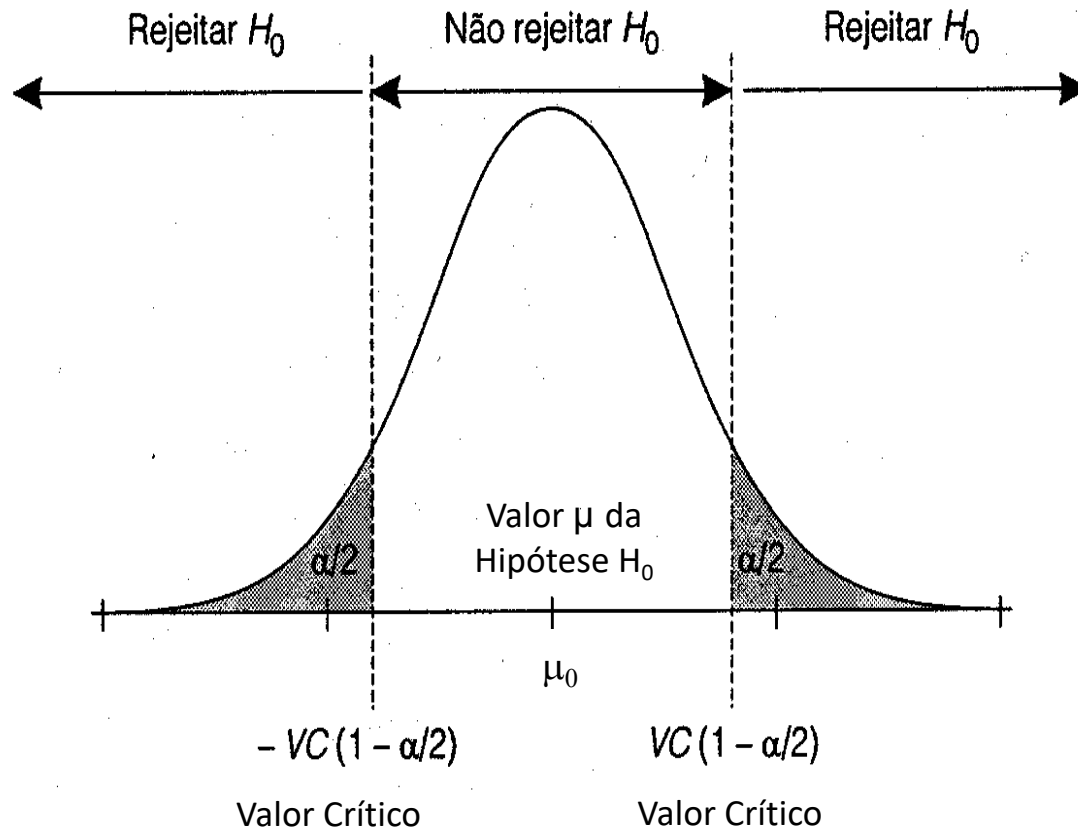
Neste caso, o teste de hipóteses a aplicar será bilateral.

→ A H_0 será a igualdade das médias do número de horas de sono entre os trabalhadores das duas empresas, ou seja $\mu_A = \mu_B$.

→ A H_1 será $\mu_A \neq \mu_B$

Inferência Estatística

Teste de Hipóteses - Bilateral



$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Inferência Estatística

Teste de Hipóteses – Unilateral & Bilateral

Exemplos:

2. Pretende-se verificar se a média do numero de horas dormido por noite difere entre os trabalhadores de duas empresas concorrentes (μ_A e μ_B), sabendo-se que o horário mínimo de trabalho é igual para os trabalhadores de ambas as empresas e que a empresa B pediu aos seus funcionários para trabalharem mais horas, devido a um aumento inesperado do fluxo de trabalho.

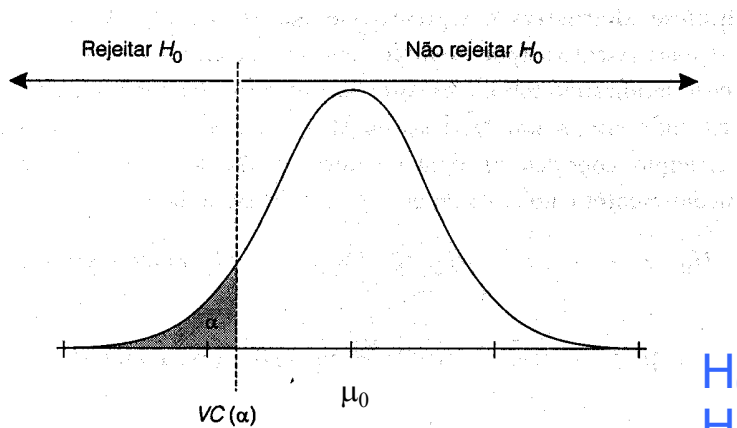
A questão colocada será se os funcionários da empresa B dormem menos horas que os funcionários da empresa A ($\mu_A > \mu_B$).

$$\rightarrow H_0: \mu_A \leq \mu_B$$

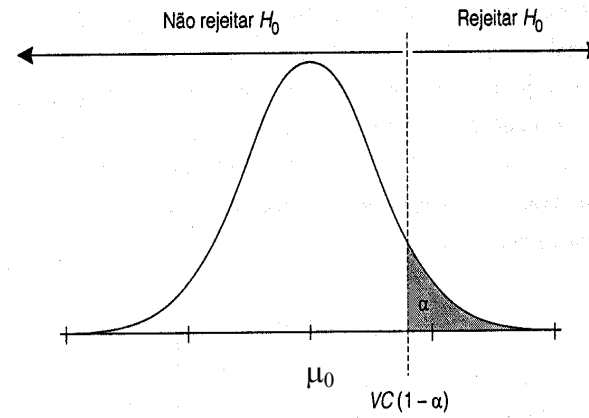
$$\rightarrow H_1: \mu_A > \mu_B$$

Inferência Estatística

Teste de Hipóteses - Unilateral



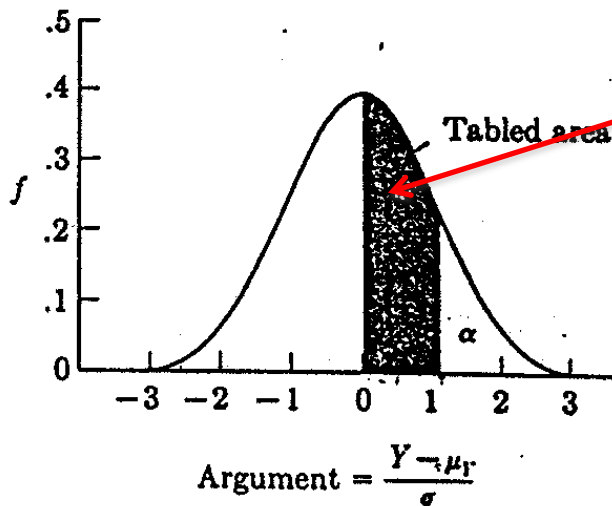
$$H_0: \mu \geq \mu_0$$
$$H_1: \mu < \mu_0$$



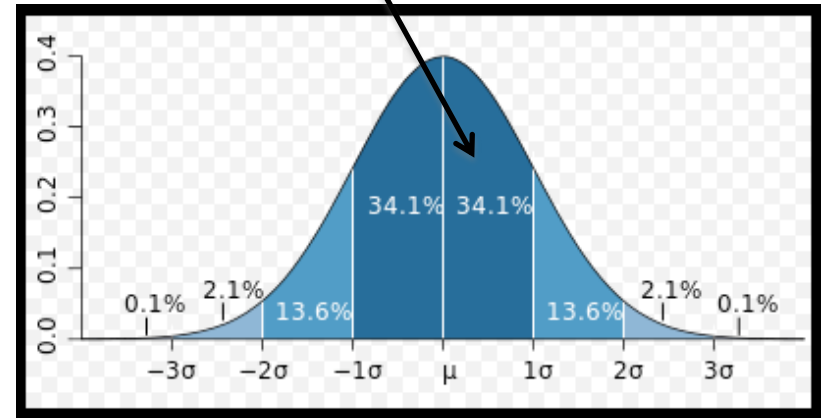
$$H_0: \mu \leq \mu_0$$
$$H_1: \mu > \mu_0$$

Inferência Estatística

TESTE Z (uma amostra)



$$P(0 \leq Z \leq 1) = 0,34$$



- 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:

$$\mu = 69,8$$

$$\sigma = 1,86$$

Grupo de pacientes tratados com fármacos para a função cardíaca:

$$\bar{X} = 70,5$$

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)

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

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

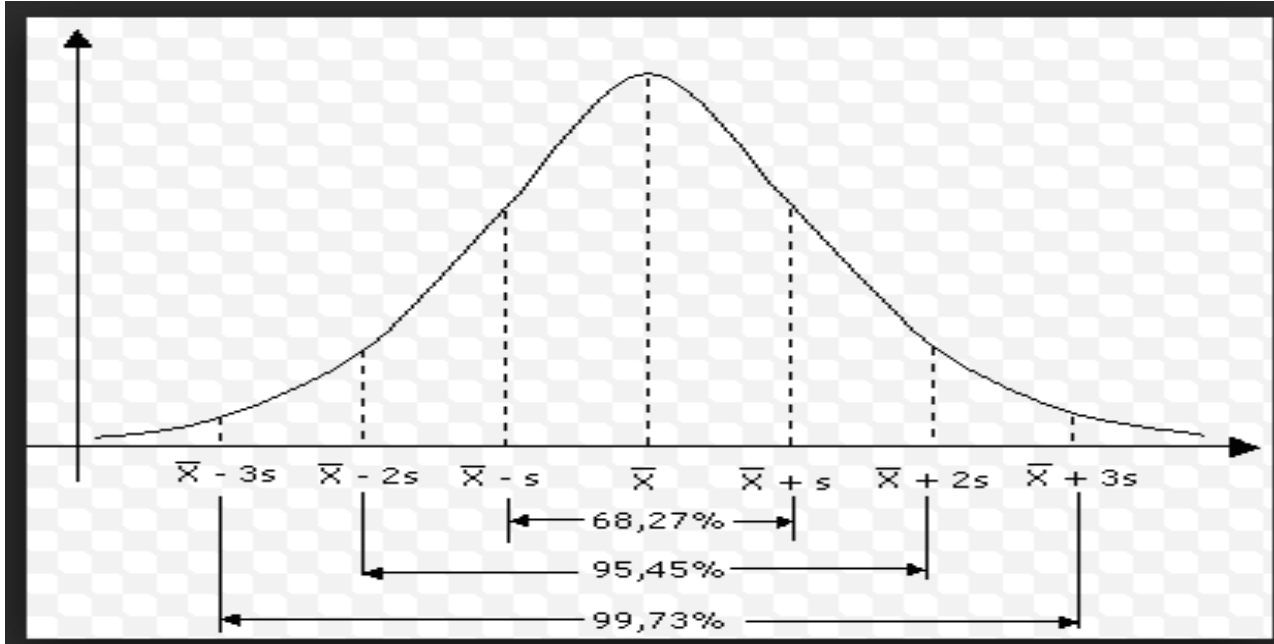
Definir um nível de significância para o teste: 5% (**$\alpha=0,05$**)

Inferência Estatística

Teste Z (1 amostras)

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

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



Inferência Estatística

Teste Z (1 amostras)

SE:

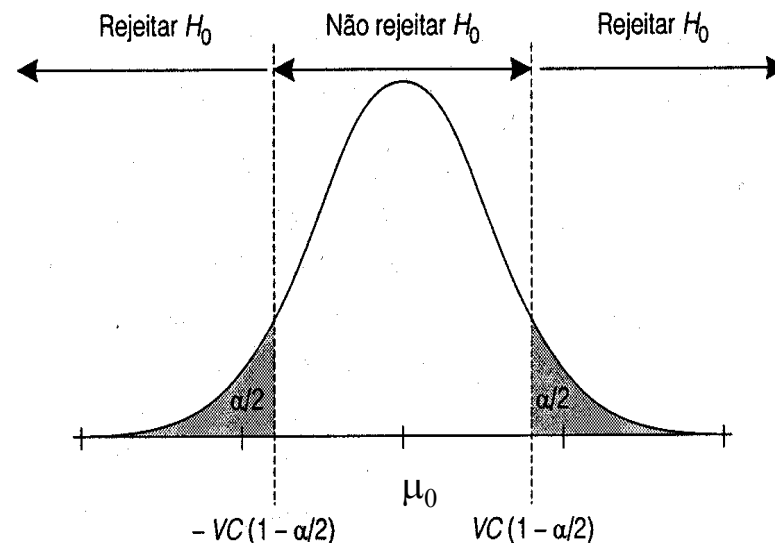
$Z < -1,96 \rightarrow$ rejeita-se $H_0 \rightarrow p < \alpha / 2$

$Z > 1,96 \rightarrow$ rejeita-se $H_0 \rightarrow p < \alpha / 2$

$-1,96 \leq Z \leq 1,96 \rightarrow$ aceita-se $H_0 \rightarrow p > \alpha / 2$

$P < \alpha \rightarrow$ rejeita-se H_0

$P > \alpha \rightarrow$ 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:

$$\mu = 69,8$$

$$\sigma = 1,86$$

Grupo de pacientes tratados com fármacos para a função cardíaca:

$$\bar{X} = 70,5$$

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)

$$u = 69,8$$

$$\sigma = 1,86$$

$$\text{SEM} = 0,26$$

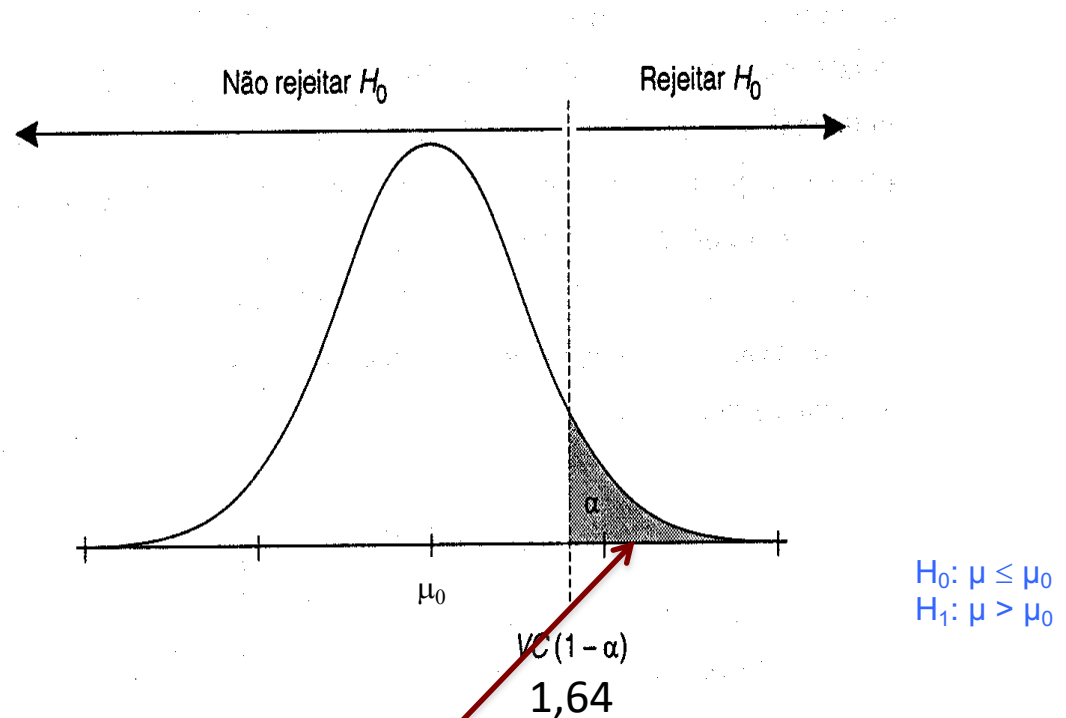
$$X = 70,5$$

$$H_0 : X \leq u$$

$$H_a : X > u$$

$$Z = (70,5 - 69,8) / 0,26 = 2,69$$

Então: $Z > 1,64 \rightarrow$ Rejeita-se H_0



Inferência Estatística

Teste Z (2 amostras)

Exemplo:

Grupo Placebo (a)

$$X_a = 8,5$$

$$\sigma_a = 2$$

$$n = 36$$

Grupo tratado (b)

$$X_b = 10,1$$

$$\sigma_b = 2$$

$$n = 36$$

H_0 : X_a igual X_b

H_a : X_a diferente de X_b

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

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**