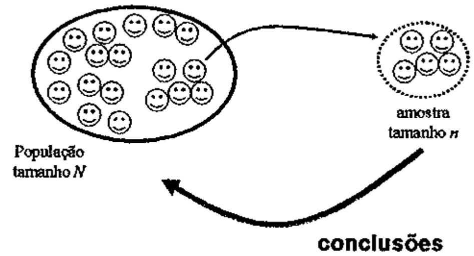
Testes de Hipóteses

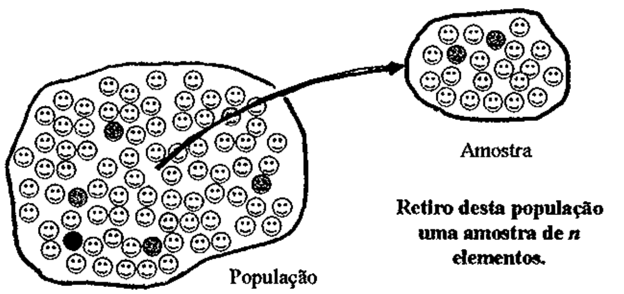
O objetivo da inferência estatística é:
Obter conclusões sobre algumas características de um conjunto de interesse, denominado <u>população</u>, com base na informação oriunda de um conjunto de dados disponíveis, denominado <u>amostra</u>.



Tais conclusões são basicamente obtidas por duas formas:

1-) **Intervalos de confiança** – quando o objetivo é estimar um parâmetro, ou seja, uma característica numérica da população.

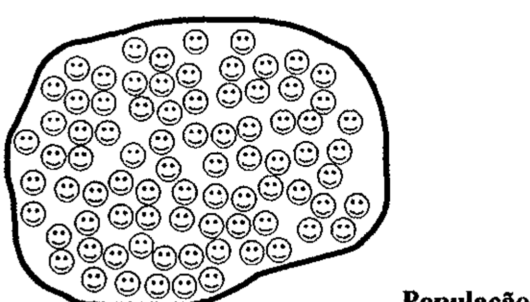
2-) **Testes de hipóteses** – quando há hipóteses sobre características numéricas da população.



Uma hipótese é uma **<u>suposição</u>** sobre um parâmetro populacional desconhecido.

Exemplo:

- Em 1860, após analizar a temperatura da região axilar de aproximadamente 25 mil pessoas, Carl Wunderlich identificou a temperatura média de adultos saudáveis como 37,0° C ou 98,6° F.
- Determinou-se que 37,0° C ou 98,6° F seria uma "temperatura normal" para um indivíduo.
- Wunderlich também estabeleceu que uma temperatura superior a 38,0° C ou 100,4° F seria um "limite superior de normalidade" para a temperatura corporal, sendo que um indivíduo com temperatura maior que este limite seria classificado como portador de febre.



População geral dos adultos saudáveis.

média $\mu_X = 37{,}0°C$

Foi assim estabelecido que na população de adultos saudáveis, a temperatura média na região axilar é 37,0°C.

Em 1992, Mackowiak, Wasserman e Levine perguntaram...

Será que a temperatura média de adultos saudáveis é mesmo 37,0°C ?

Referência: *JAMA*, 268(12):1578-80,1992

Pergunta: $\mu_X \neq 37{,}0°C$ **???**

Hipótese do pesquisador: média $\mu_X \neq 37{,}0°C$

Um teste estatístico de hipóteses é uma **regra** utilizada para decidir quando rejeitar uma hipótese. Esta regra é sempre baseada em uma amostra.

Com base nos resultados de uma amostra aleatória tamanho *n*, tomamos a decisão de rejeitar ou não rejeitar uma hipótese formulada sobre um parâmetro de interesse.

Na prática, consideramos duas hipóteses:

Hipótese alternativa (H_A):

é a "hipótese do pesquisador", aquilo que ele deseja verificar.

Na população de adultos saudáveis, a temperatura média na região axilar é diferente de 37,0°C

Hipótese nula (H_0): é o complemento da alternativa.

Na população de adultos saudáveis, a temperatura média na região axilar é 37,0°C

Hipótese alternativa (H_A): $H_A: \mu_X \neq 37{,}0°C$
Hipótese nula (H_0): $H_0: \mu_X = 37{,}0°C$

Agora uma definição mais completa...
Com base nos resultados de uma amostra aleatória tamanho *n*, tomamos a decisão de **rejeitar** ou **não rejeitar** a hipótese nula.

Estaremos assim sujeitos a dois tipos de erros...

A-) ERRO TIPO I : rejeito H_0, ***mas H_0 é verdadeira***

B-) ERRO TIPO II : não rejeito H_0, ***mas H_0 é falsa***

Decisão	H_0 é verdadeira	H_0 é falsa
Rejeito H_0	Erro tipo I	Sem erro
Não rejeito H_0	Sem erro	Erro tipo II

Exemplo:
Certa droga ("droga 1") vem sendo utilizada no tratamento de uma moléstia. Um pesquisador desenvolve uma nova droga ("droga 2"), que, se mais eficiente, substituirá a droga 1.

Hipótese do pesquisador:
A droga 2 é mais eficiente.
Hipóteses:
H_0: A droga 1 é mais eficiente ou tão eficiente quanto a droga 2.
H_A: A droga 2 é mais eficiente.

Qual é o erro mais grave ?

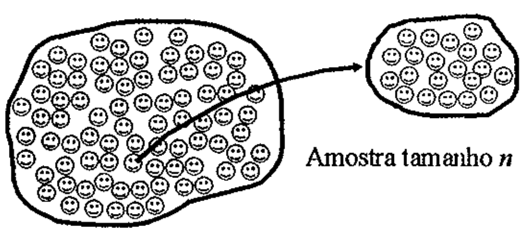
- A probabilidade de se cometer um erro tipo I é chamada *nível de significância do teste* e é denotada por α.
- A probabilidade de se cometer um erro tipo II é denotada por β.
- Na área da saúde, a quantidade $1-\beta$ é geralmente chamada *poder* (ou *potência*) do teste.
- O *nível de significância do teste* (α) é fixado antes da coleta dos dados.
  Na área da saúde, é muito comum fixar $\alpha = 5\%$.
- A probabilidade de se cometer um erro tipo II (β) é geralmente usada para o cálculo do tamanho amostral.
  Escolhas comuns: $\beta = 5\%$, 10% ou 20%.

Voltando ao exemplo...

Antes de retirarmos a amostra da população, vamos fixar o nível de significância do teste. A maioria dos estudos da área da saúde adota $\alpha = 0{,}05$.

Se, na população, a temperatura média na região axilar segue uma distribuição normal com média μ_X e variância σ_X^2, ou seja, $X \sim N(\mu_X ; \sigma_X^2)$, sabemos que $T = \frac{(\bar{X}-\mu_X)\sqrt{n}}{s_X}$ segue uma distribuição *t* de Student com $n-1$ graus de liberdade.

Se a hipótese nula for verdadeira, $T_0 = \frac{(\bar{X}-37)\sqrt{n}}{s_X}$ segue uma distribuição *t* de Student com $n-1$ graus de liberdade.

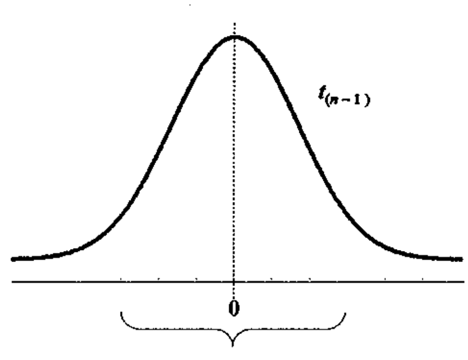


População geral dos adultos saudáveis.

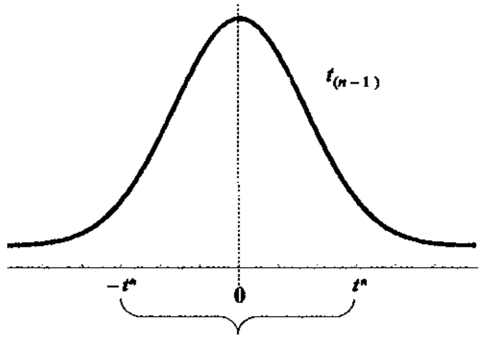
A partir de uma amostra tamanho n, vamos calcular

$$t_0 = \frac{(\bar{x}-37)\sqrt{n}}{s_X}$$

Se a hipótese nula for verdadeira, este valor de t_0 calculado da amostra será o resultado de uma distribuição *t* de Student com $n-1$ gl.



Se a hipótese nula é verdadeira, é mais provável que encontremos em uma amostra tamanho *n* valores de t_0 situados nesta região "central" da curva, onde há maior densidade.



Assim, vamos rejeitar H_0 se encontrarmos em uma amostra tamanho n um valor de t_0 maior que t^* ou menor que $-t^*$, de forma que t^* é determinado da seguinte maneira...

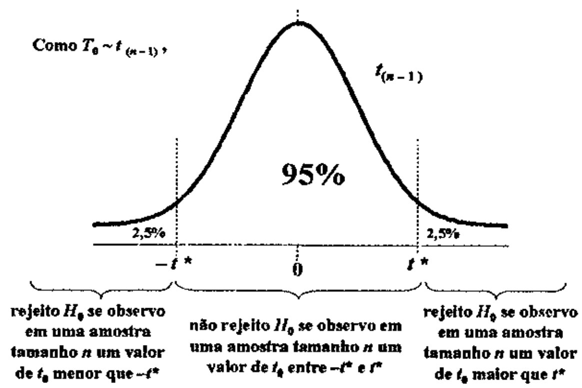
... se fixarmos o nível de significância em 5%, então

Nível de significância = P(erro tipo I) = P(rejeitar H_0 | H_0 é verdadeira) = 5%.

Decidimos rejeitar H_0 nas situações onde $T_0 > t^*$ e $T_0 < -t^*$.

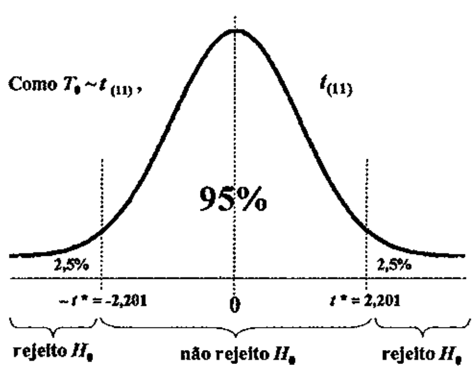
Assim, P(não rejeitar H_0 | H_0 é verdadeira) = 95%

$P(-t^* < T_0 < t^*) = 95\%$



$-t^*$ e t^* são chamados valores críticos de T_0.

Vamos retirar da população uma amostra tamanho $n = 12$.



Da tabela da distribuição t de Student, temos que $t^* = 2{,}201$.

Na amostra tamanho $n = 12$, observamos as seguintes temperaturas (em °C):

35,6	35,8	36,3	35,5	35,9	36,9
37,1	36,2	37,2	36,1	37,6	37,7

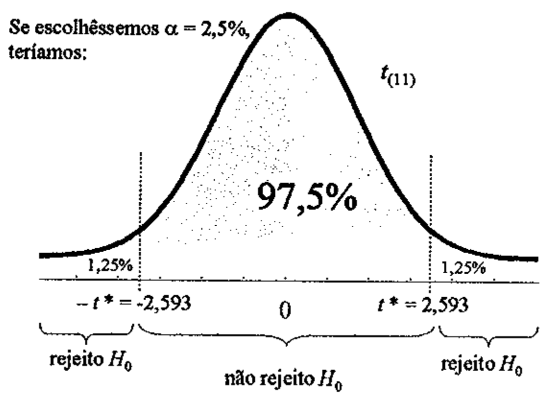
Temos então uma média amostral $\bar{x} = 36{,}5$°C e um desvio padrão amostral $s_X = 0{,}775$ °C.

$$\text{Assim, } t_0 = \frac{(\bar{x} - 37)\sqrt{n}}{s_X} = \frac{(36{,}5 - 37)\sqrt{12}}{0{,}775} = -2{,}23.$$

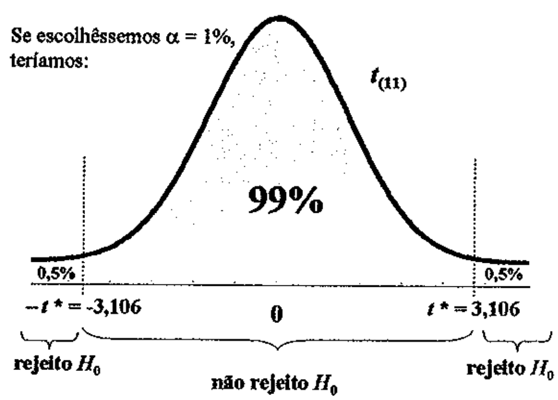
Como $t_0 < -2{,}201$, rejeitamos H_0; ou seja, temos evidências de que a temperatura média na região axilar na população de adultos saudáveis é diferente de 37,0°C.

Esta conclusão é valida para um nível de significância $\alpha = 5\%$.

Uma nota...



Como $t_0 = -2{,}23$, nós não rejeitaríamos H_0 para $\alpha = 2{,}5\%$.



Como $t_0 = -2{,}23$, também não rejeitaríamos H_0 para $\alpha = 1\%$.

Temos $t_0 = -2{,}23$.

α	$-t^*$	t^*	Decisão
0,05	-2,2010	2,2010	Rejeito H_0
0,048	-2,2243	2,2243	Rejeito H_0
0,047	-2,2364	2,2364	Não rejeito H_0
0,04	-2,3281	2,3281	Não rejeito H_0
0,03	-2,4907	2,4907	Não rejeito H_0
0,025	-2,5931	2,5931	Não rejeito H_0
0,02	-2,7181	2,7181	Não rejeito H_0
0,01	-3,1058	3,1058	Não rejeito H_0

Nota-se que rejeitaríamos H_0 se escolhêssemos um valor maior ou igual a 0,048 para α.

O menor valor de α que nos levaria a rejeitar a hipótese nula é chamado "p valor".