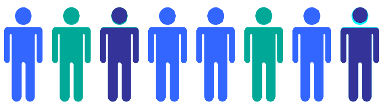
Processo do teste de hipótese

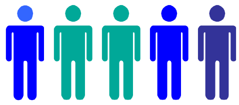
Hipótese de pesquisa: **a idade média da população é 50**

H_0: μ = 50
H_1: μ ≠ 50



População

Selecionar amostra aleatória



amostra

Suponha que: $\overline{X}$ **=20**

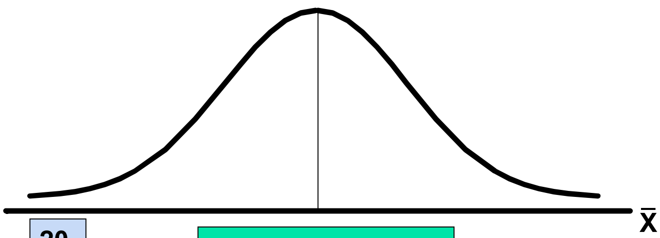
É $\overline{X}$ **= 20** improvável se μ =50?

Sim: Rejeite Ho

Para definir pouco provável devemos recorrer a distribuição de probabilidade da propriedade amostral (no caso a média amostral)

Razão para rejeitar H_0

Distribuição amostral de $\overline{X}$



20

Se é improvável obter uma média amostral com esse valor…

μ = 50
Se H_0 é verdadeira

... Se este fosse a média da população

... Então rejeitamos

Ho: μ = 50.

Nível de significância (α)

- **Define os valores improváveis da estatística amostral se a hipótese nula é verdadeira, ou seja a região de rejeição da distribuição amostral.**
  - Valores típicos para α: 0.01, 0.05, or 0.10 (é uma probabilidade)
- **Deve ser selecionada antes do teste.**
- **Fornece os valores críticos do teste.**

Nível de significância e região de rejeição

Nível de significância $=\alpha$

✧ representa o valor crítico

Região de rejeição: área cinza

H_0: μ = 3
H_1: μ ≠ 3

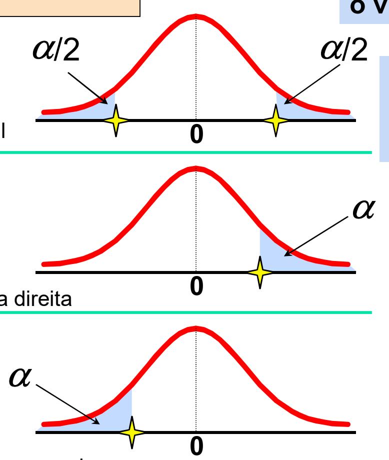
Teste bilateral

H_0: μ = 3
H_1: μ > 3

Teste unilateral a direita

H_0: μ = 3
H_1: μ < 3

Teste unilateral a esquerda



O teste de hipóteses: É um processo de decisão entre duas alternativas para o valor do parâmetro de uma população:

A hipótese nula Ho: contém o sinal de igualdade (não tem algo acontecendo)

A hipótese alternativa: não contém a igualdade (tem alguma coisa) e, em geral, representa a questão de pesquisa a ser testada.

Baseado nos dados amostrais temos que verificar, sob a hipótese nula, o quão provável é o valor amostral. Dado um nível de sgnificância (a priori) para comparação concluímos com a rejeição ou não rejeição da hipótese nula.

Se rejeitarmos Ho: Concluímos que os dados fornecem evidência de.....(" que tem alguma coisa")

Se não rejeitarmos Ho: Concluímos que os dados fornecem evidência de que não....(não tem o efeito)

Erros tipo I e tipo II

		Resultados possíveis	
		Ho é verdadeira	Ho é falsa (H1 verdadeiro)
Decisão	Rejeitar Ho	Erro tipo I (rejeitar Ho verdadeiro) α	Decisão Correta $1\text{-}\beta$
	Não rejeitar Ho	Decisão Correta $1\text{-}\alpha$	Erro tipo II (deixar de rejeitar Ho falsa) β

α: probabilidade do erro tipo I
β: probabilidade do erro tipo II
Potência (Poder) de um teste: é a probabilidade de rejeitar corretamente H0 = 1 - β

Analogia: teste de hipótese – julgamento

Se usarmos a analogia de um teste de hipótese como um julgamento criminal, podemos associar o veredicto em termos das hipóteses nula e alternativa:

H0: Réu é inocente (igualdade, nada ocorre)
H1: Réu é culpado (desigualdade, algo ocorre)

Então:

Declarar o réu inocente quando ele é realmente culpado é um ERRO TIPO II

Declarar o réu culpado quando ele é realmente inocente é um ERRO TIPO I

Taxa de erro ERRO TIPO I

Rejeitar H0 quando o p-valor é menor que 0,05 (α = 0,05) significa que, para os casos em que H0 é realmente verdade, não queremos rejeitá-la de forma incorreta mais de 5% das vezes.
Assim, quando se utiliza um nível de significância de 5%, há cerca de 5% chance de fazer um erro tipo 1 se H0 é verdadeira.

P (Erro tipo I| H0 é verdadeira) = α

Por isso usamos pequenos valores de α, pois aumentar α aumenta a taxa de erro tipo I.

Taxa de erro tipo II:

Se a hipótese alternativa é verdadeira, qual é a chance de cometermos um erro tipo II, ou seja, não rejeitamos Ho?

- A resposta não é simples.

- Se o verdadeiro parâmetro da população é muito próximo ao valor nulo, será difícil de detectar uma diferença (e rejeitar H0).
- Se a verdadeira parâmetro da população é muito diferente do valor nulo, ele será mais fácil de detectar uma diferença.
- β depende diferença entre o valor estimado e valor da hipótese nula.

Se o Erro tipo I é perigoso, escolha um nível de significância pequeno (por exemplo 0,01).
Objetivo: queremos ser muito cauteloso sobre rejeitar H0, por isso exigimos forte evidências a favor H1 antes de fazê-lo.

Se um erro de tipo 2 é relativamente mais perigoso, escolha um nível de significância mais elevado (por exemplo 0,10).
Objetivo: queremos ser cauteloso sobre deixar de rejeitar H0 quando esta é realmente falsa.

Controlando erros tipo I e tipo II

- **Poder do teste: probabilidade 1- β de se rejeitar uma hipótese nula falsa, que é calculada usando um nível de significância particular α e um valor particular do parâmetro populacional que seja uma altenativa ao assumido na hipótese nula. Ou seja, probabilidade de se apoiar uma hipótese alternativa verdadeira.**

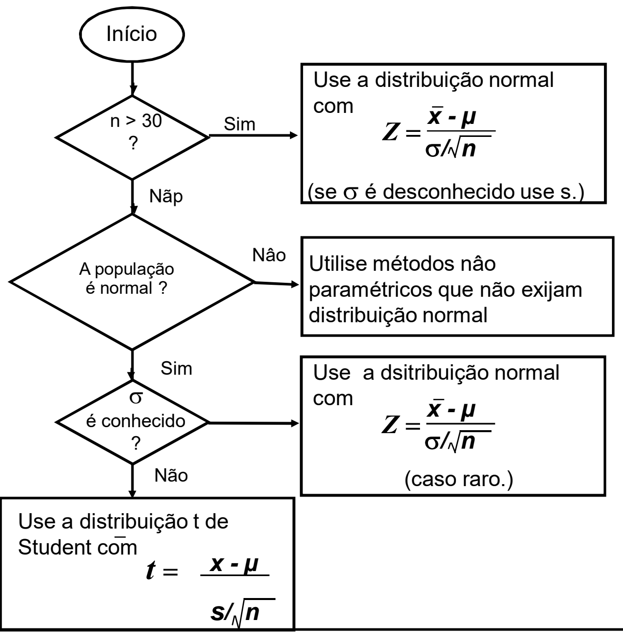
- **Para α fixo, um aumento do tamanho da amostra *n* produz uma redução de β**
- **Para um tamanho de amostra *n fixo*, a redução de α produz um aumento de β. Reciprocamente, um aumento em α produz uma redução em β**
- **Para diminuir ambos α e β, aumente o tamanho da amostra.**

Teste de hipótese para a média (σ conhecido)

- Estatística (da amostra): $\overline{X}$

Estatística teste:

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



Método tradicional: Compare a estatística teste com o valor crítico

- Determine o valor crítico Z para um dado nível de significância α**.**
- Se a estatística de teste estiver no interior da região crítica rejeite H_0 ; caso contrário não rejeite H_0

Método valor p: determine o valor p

- Valor p: probabilidade de obter uma ~~–~~ estatística de teste menor que α.
- Obtenha o valor p de uma tabela ou computador
- Compare o valor p com α
  - se $p < \alpha$, rejeite H_0
  - se $p \geq \alpha$, não rejeite H_0

Procedimento geral de um teste de hipótese

Passos:

1 – Escreva a afirmativa original na forma simbólica.
2 – Escreva o oposto da afirmativa original na forma simbólica.
3 – Expresse Ho e H_1.
4 – Selecione o nível de significância α.
5 – Verifique a distribuição a ser utilizada.
6 – Calcule a estatística teste e depois utilize um método:
Método tradicional: Compare a estatística teste com o valor crítico
Método valor P: determine o valor-P
7 – Tradicional: Estatística teste na região crítica: Rejeite H0
valor P: valor $P < \alpha$, Rejeite H0
8 – Estabeleça a conclusão (ver fraseado final).

Exemplo de teste de hipótese para média

Uma companhia produz papel para impressoras laser. A Largura do papel deve ser 216 mm. A largura verdadeira tem desvio padrão conhecido igual a 0,023 mm (definido pela tecnologia do processo de fabricação). Durante o processo de fabricação a largura pode desviar do valor correto. Um inspetor do controle de qualidade seleciona 50 folhas ao acaso e mede com um instrumento de precisão mostrando uma largura média de 216,007 mm.

Usando um nível de significância α = 0,05, esta amostra demonstra que a média do processo difere das especificações?

Vamos ao teste:

1 – Escreva a afirmativa original na forma simbólica.

Afirmativa original: A média do processo difere das especificações.

$\mu \neq 216$ mm

2 – Escreva o oposto da afirmativa original na forma simbólica.

$\mu = 216$ mm

3 – Expresse H_0 e H_1.

H0: $\mu = 216$ mm

H1: $\mu \neq 216$ mm (teste bilateral)

4 – Selecione o nível de significância α.

$\alpha = 0{,}05$

5 – Verifique a distribuição a ser utilizada.
Teste de médias para sigma conhecido: distribuição normal

ou seja, $z = \pm 1{,}96$ (teste bilateral)

6 – Calcule a estatística teste e depois utilize um método:

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

$$z_{teste} = \frac{0{,}007}{0{,}003253} = 2{,}15$$

Método tradicional: Compare a estatística teste com o valor crítico: $z_{teste} > 1{,}96$
Método valor P: determine o valor-P: 2 x 0,0158 = 0,0316 < alfa

7 – Tradicional: Estatística teste na região crítica: Rejeite H0
valor P: valor P < alfa, Rejeite H0
8 – Estabeleça a conclusão: "Os dados amostrais apóiam a afirmativa de que μ ≠ 216, ou seja, a média do processo difere das especificações."

Teste de uma afirmativa sobre uma Média: σ desconhecido

Amostras grandes

1) **Amostra aleatória**
2) **População com distribuiçao normal ou n > 30**
3) **Se o desvio padrão populacional é desconhecido podemos usar desvio amostral s como estimativa.**

Neste caso a estatística de teste é a Z.

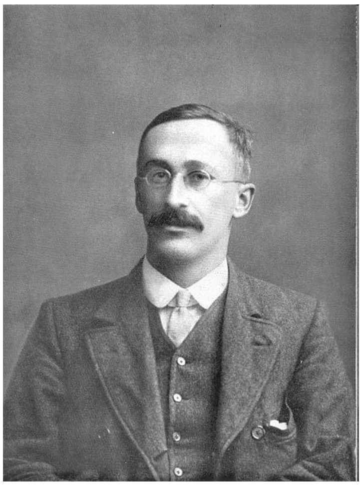
1) **Amostras pequenas (n ≤ 30)**
2) s é desconhecido
3) A população original tem distribuição original essencialmente normal.

Neste caso a estatística de teste é

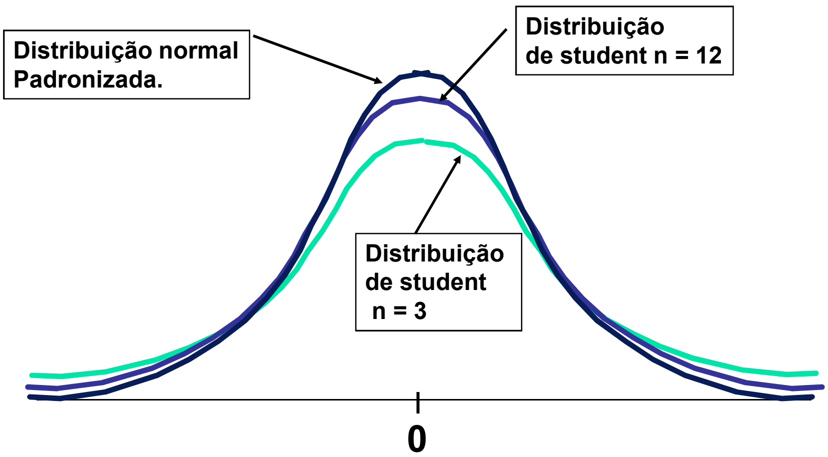
$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Com distribuição distribuição de student .

- *student* é um pseudônimode William Sealy Gosset, que não podia publicar artigos usando seu próprio nome.

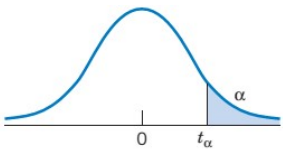


Distribuição de student
n = 3 e n = 12



d.f. \ α	.25	.10	.05	.025	.01	.00833	.00625	.005
1	1.000	3.078	6.314	12.706	31.821	38.204	50.923	63.657
2	.816	1.886	2.920	4.303	6.965	7.649	8.860	9.925
3	.765	1.638	2.353	3.182	4.541	4.857	5.392	5.841
4	.741	1.533	2.132	2.776	3.747	3.961	4.315	4.604
5	.727	1.476	2.015	2.571	3.365	3.534	3.810	4.032
6	.718	1.440	1.943	2.447	3.143	3.287	3.521	3.707
7	.711	1.415	1.895	2.365	2.998	3.128	3.335	3.499
8	.706	1.397	1.860	2.306	2.896	3.016	3.206	3.355
9	.703	1.383	1.833	2.262	2.821	2.933	3.111	3.250
10	.700	1.372	1.812	2.228	2.764	2.870	3.038	3.169
11	.697	1.363	1.796	2.201	2.718	2.820	2.981	3.106
12	.695	1.356	1.782	2.179	2.681	2.779	2.934	3.055
13	.694	1.350	1.771	2.160	2.650	2.746	2.896	3.012
14	.692	1.345	1.761	2.145	2.624	2.718	2.864	2.977
15	.691	1.341	1.753	2.131	2.602	2.694	2.837	2.947
16	.690	1.337	1.746	2.120	2.583	2.673	2.813	2.921
17	.689	1.333	1.740	2.110	2.567	2.655	2.793	2.898
18	.688	1.330	1.734	2.101	2.552	2.639	2.775	2.878
19	.688	1.328	1.729	2.093	2.539	2.625	2.759	2.861
20	.687	1.325	1.725	2.086	2.528	2.613	2.744	2.845
21	.686	1.323	1.721	2.080	2.518	2.601	2.732	2.831
22	.686	1.321	1.717	2.074	2.508	2.591	2.720	2.819
23	.685	1.319	1.714	2.069	2.500	2.582	2.710	2.807
24	.685	1.318	1.711	2.064	2.492	2.574	2.700	2.797
25	.684	1.316	1.708	2.060	2.485	2.566	2.692	2.787
26	.684	1.315	1.706	2.056	2.479	2.559	2.684	2.779
27	.684	1.314	1.703	2.052	2.473	2.552	2.676	2.771
28	.683	1.313	1.701	2.048	2.467	2.546	2.669	2.763
29	.683	1.311	1.699	2.045	2.462	2.541	2.663	2.756
30	.683	1.310	1.697	2.042	2.457	2.536	2.657	2.750
40	.681	1.303	1.684	2.021	2.423	2.499	2.616	2.704
60	.679	1.296	1.671	2.000	2.390	2.463	2.575	2.660
120	.677	1.289	1.658	1.980	2.358	2.428	2.536	2.617
∞	.674	1.282	1.645	1.960	2.326	2.394	2.498	2.576

α é a probabilidade (área) na cauda direita



Teste de afirmativa sobre proporção

- Exemplos de afirmativas:
- 60% dos alunos de administração são meninos.
- A probabilidade de caras para uma moeda honesta é 0.5
- Menos de 25% dos estudantes de administração são fumantes.
- ....

Distribuição Binomial

Seja X uma V.A. com dois resultados possíveis: sim ou não, ligado ou desligado, casado ou solteiro, sucesso ou fracasso,...

X=1 sucesso; P(X=1) = p
X=0 fracasso; P(X=0) = 1 - p = q

$$P(x = k) = \binom{n}{k}(p)^k(1-p)^{n-k} \text{ é a probabilidade de k sucessos}$$

$$\text{média}: \mu = np$$

$$\text{desvio padrão}: \sigma = \sqrt{npq}$$

- Probabilidade de se obter valores maiores ou iguais ao valor observado

$$P(x \geq k) = \sum_{i=k}^{n}\binom{n}{i}(p)^i(1-p)^{n-i}$$

- Probabilidade de se obter valores menores ou iguais ao valor observado

$$P(x \leq k) = \sum_{i=0}^{k}\binom{n}{i}(p)^i(1-p)^{n-i}$$

Hipóteses usadas para testar uma afirmação sobre uma proporção, uma probabilidade ou uma percentagem populacional:

1- São verificadas as condições para um experimento binomial: número fixo de provas independentes com probabilidade constante, e cada prova comporta dois resultados, que designamos “sucesso” e “falha” .

2- a distribuição de probabilidade dos resultados das proporções amostrais é a binomial. No caso em que **np ≥ 5 e nq ≥ 5** a distribuição binomial pode ser aproximada por uma distribuição normal com

$$\text{média}: \mu = np$$

$$\text{desvio padrão}: \sigma = \sqrt{npq}$$

Teste de proporção

$$n = \text{número de provas}$$

$$\hat{p} = \frac{\overline{x}}{n} \quad \text{(proporção amostral)}$$

$$p = \text{proporção populacional (usada na hipótese nula)}$$

$$q = 1 - p$$

$$\text{Estatística de teste } (np \geq 5 \text{ e } nq \geq 5):$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Teste de uma afirmação sobre um desvio padrão ou uma variância

Hipóteses:
* Amostra é aleatória.
* A população deve ter distribuição normal.

Estatística de teste para testar hipóteses sobre σ ou σ^2 :

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2}$$

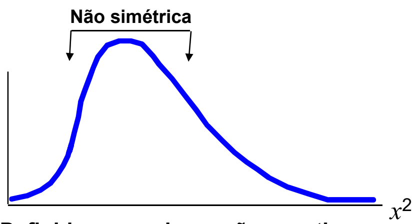
onde: n = tamanho da amostra

s^2 = variância amostral

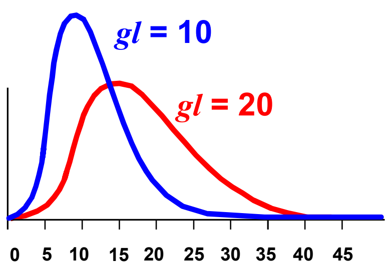
σ^2 = variância populacional (dada na hipótese nula)

distribuição de χ^2 com $gl = n - 1$ graus de liberdade

Relembrando:
Propriedades da distribuição de χ^2

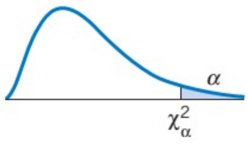


Definida para valores não negativos



Existe uma distribução diferente para cada grau de liberdade α

α / d.f.	.99	.975	.95	.90	.50	.10	.05	.025	.01
1	.0002	.001	.004	.02	.45	2.71	3.84	5.02	6.63
2	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21
3	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34
4	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28
5	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09
6	.87	1.24	1.64	2.20	5.35	10.64	12.59	14.45	16.81
7	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48
8	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09
9	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67
10	2.56	3.24	3.94	4.87	9.34	15.99	18.31	20.48	23.21
11	3.05	3.81	4.57	5.58	10.34	17.28	19.68	21.92	24.72
12	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22
13	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69
14	4.66	5.62	6.57	7.79	13.34	21.06	23.68	26.12	29.14
15	5.23	6.26	7.26	8.55	14.34	22.31	25.00	27.49	30.58
16	5.81	6.90	7.96	9.31	15.34	23.54	26.30	28.85	32.00
17	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41
18	7.01	8.23	9.39	10.86	17.34	25.99	28.87	31.53	34.81
19	7.63	8.90	10.12	11.65	18.34	27.20	30.14	32.85	36.19
20	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57
21	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93
22	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98
25	11.52	13.11	14.61	16.47	24.34	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64
27	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96
28	13.56	15.30	16.93	18.94	27.34	37.92	41.34	44.46	48.28
29	14.26	16.04	17.71	19.77	28.34	39.09	42.56	45.72	49.59
30	14.95	16.78	18.49	20.60	29.34	40.26	43.77	46.98	50.89
40	22.16	24.42	26.51	29.05	39.34	51.81	55.76	59.34	63.69
50	29.71	32.35	34.76	37.69	49.33	63.17	67.50	71.42	76.15
60	37.48	40.47	43.19	46.46	59.33	74.40	79.08	83.30	88.38
70	45.44	48.75	51.74	55.33	69.33	85.53	90.53	95.02	100.43
80	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33
90	61.75	65.64	69.13	73.29	89.33	107.57	113.15	118.14	124.12
100	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81



$$H_0 : \sigma^2 = \sigma_0{}^2$$

$$H_1 : \sigma^2 \neq \sigma_0{}^2, \quad H_0 : \sigma^2 > \sigma_0{}^2, \quad H_0 : \sigma^2 < \sigma_0{}^2$$

Bilateral: rejeite Ho se $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha/2}$ ou $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{\alpha/2}$

Unilateral a direita: rejeite Ho se $\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{\alpha}$

Unilateral a esquerda: rejeite Ho se $\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha}$

Exemplo

- A Stewart Aviation Products Company vem fabricando altímetros de aviões com erros distribuídos normalmente com média 0 (obtida por calibragem) e desvio padrão de 43,7 pés. Após a instalação da nova produção de equipamentos selecionaram-se aleatoriamente 30 altímetros da nova linha de produção. Esta amostra acusou erros com desvio padrão de 54,7 pés. Ao nível de significância de 0,05, teste a afirmação de que os novos altímetros tem desvio padrão diferente do valor anterior de 43,7 pés.

Vamos ao teste:

1 – Escreva a afirmativa original na forma simbólica.

Afirmativa original: O desvio padrão dos altímetros é diferente de 43,7 pés

$\sigma \neq 43{,}7$ pés

2 – Escreva o oposto da afirmativa original na forma simbólica.

$\sigma = 43{,}7$ pés

3 – Expresse H_0 e H_1.

H0: $\sigma = 43{,}7$

H1: $\sigma \neq 43{,}7$ (teste bilateral)

4 – Selecione o nível de significância α.

$\alpha = 0{,}05$

5 – Verifique a distribuição a ser utilizada e as condições de seu uso

Temos: amostra aleatória de 30 elementos de população (são os erros de mensuração de altimetros) distribuida segundo uma N(0, 43,7). Logo a distribuição amostral da variância (desvio-padrão ao quadrado) é a χ^2

Para α =0.05 bilateral temos e gl=n-1=29 temos as regiões críticas:

$\chi^2_{.975}$ = 16,04 e $\chi^2_{.025}$ = 45,72

6 – Calcule a estatística teste e depois utilize um método:

$$16{,}04 < \chi^2 = \frac{(n-1)\,s^2}{\sigma^2} = \frac{(30-1)(54{,}7)^2}{(43{,}7)^2} = 45{,}437 < 45{,}72$$

Não rejeita-se Ho.

7 – Estabeleça a conclusão: não há evidências amostrais
de que a variância da população (ou desvio padrão) tenha se alterado.

Resumindo

O objetivo do teste de hipótese é fornecer um método que permita verificar se os dados amostrais trazem evidências que apóiam ou não uma hipótese formulada.

A hipótese testada diretamente é chamada hipótese nula:

$$H_0 : \theta = \theta_0$$

A hipótese alternativa:

$$H_1 : \theta < \theta_0 \qquad H_1 : \theta \neq \theta_0 \qquad H_1 : \theta > \theta_0$$

Teste de hipótese: A idéia e sempre a mesma…

1. Escreva a afirmativa original em símbolos matemáticos utilize =, ≠, >, <, ≤, ≥.
2. Escreva o oposto da afirmativa original em símbolos matemáticos.
3. A hipótese nula, H_0, é (entre 1 e 2) a afirmativa que contém a igualdade.
4. A hipótese alternativa H_1, é a outra, porém utilize apenas os símbolos <, ≠ ou >.
5. Selecione o nível de significância (α)
6. Calcule a Estatística teste.
7. Determine o *valor-P:* Probabilidade associada à estatística teste.
8. Se *valor-P* < α rejeitar H_0; caso contrário, não rejeitar H_0 *(ou compare a estatística de teste com o valor crítico associado a α)*
9. Retome a afirmativa original e estabeleça uma conclusão.

Resumo: Testes de hipóteses para uma amostra

Parâmetro	Requisitos: Amostra Aleatória Simples e ...	Distribuição e Estatística de Teste
Proporção	$np \geq 5$ e $nq \geq 5$	Normal: $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$
Média	σ conhecido e população normalmente distribuída *ou* σ conhecido e $n > 30$	Normal: $z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\dfrac{\sigma}{\sqrt{n}}}$
	σ desconhecido e população normalmente distribuída *ou* σ desconhecido e $n > 30$	t-Student: $t = \dfrac{\bar{x} - \mu_{\bar{x}}}{\dfrac{s}{\sqrt{n}}}$
	População não-normal e $n \leq 30$	Use um método não-paramétrico ou *bootstrap*.
Desvio Padrão ou Variância	População normalmente distribuída	Qui-quadrado: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$