Inferência Estatística II

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Programa de Pós-Graduação em Estatística Aplicada e Biometria
Universidade Federal de Viçosa
Campus UFV - Viçosa

Por que Testes de Hipóteses e Poder importam?

Decidir sob incerteza

No mundo real decidimos com dados imperfeitos: aprovar um tratamento, liberar um lote, avaliar uma política. Testes de hipóteses formalizam essa decisão, controlando o **tamanho** (α) e quantificando a capacidade de detectar efeitos reais via **poder** ($1 - \beta$).

Por que Testes de Hipóteses e Poder importam?

Decidir sob incerteza

No mundo real decidimos com dados imperfeitos: aprovar um tratamento, liberar um lote, avaliar uma política. Testes de hipóteses formalizam essa decisão, controlando o **tamanho** (α) e quantificando a capacidade de detectar efeitos reais via **poder** $(1-\beta)$.

Perguntas-motrizes da aula

- Como controlar erro tipo I e compreender erro tipo II?

Por que Testes de Hipóteses e Poder importam?

Decidir sob incerteza

No mundo real decidimos com dados imperfeitos: aprovar um tratamento, liberar um lote, avaliar uma política. Testes de hipóteses formalizam essa decisão, controlando o **tamanho** (α) e quantificando a capacidade de detectar efeitos reais via **poder** $(1 - \beta)$.

Perguntas-motrizes da aula

- Como controlar erro tipo I e compreender erro tipo II?
- O que determina o **poder** e como aumentá-lo (n, σ, efeito, α)?

Por que Testes de Hipóteses e Poder importam?

Decidir sob incerteza

No mundo real decidimos com dados imperfeitos: aprovar um tratamento, liberar um lote, avaliar uma política. Testes de hipóteses formalizam essa decisão, controlando o **tamanho** (α) e quantificando a capacidade de detectar efeitos reais via **poder** ($1 - \beta$).

Perguntas-motrizes da aula

- Como controlar erro tipo I e compreender erro tipo II?
- O que determina o **poder** e como aumentá-lo (n, σ, efeito, α)?
- Qual a relação entre **teste bilateral** e **intervalo de confiança**?

Por que Testes de Hipóteses e Poder importam?

Decidir sob incerteza

No mundo real decidimos com dados imperfeitos: aprovar um tratamento, liberar um lote, avaliar uma política. Testes de hipóteses formalizam essa decisão, controlando o **tamanho** (α) e quantificando a capacidade de detectar efeitos reais via **poder** ($1 - \beta$).

Perguntas-motrizes da aula

- Como controlar erro tipo I e compreender erro tipo II?
- O que determina o **poder** e como aumentá-lo (n, σ, efeito, α)?
- Qual a relação entre **teste bilateral** e **intervalo de confiança**?
- Como interpretar corretamente o **p-valor**?

Sumário

1. Exemplos
   - Exemplo 1: Teste Bilateral para a Média
   - Exemplo 2: Teste Bilateral para a Média (Prático)
   - Exemplo 3: Teste Bilateral para a Média
   - Exemplo 4: Teste Unilateral para a Média
   - Exemplo 5: Teste unilateral a direita
   - Exemplo 6: Teste unilateral a esquerda
   - Exemplo 7: Teste unilateral para a Proporção Binomial
   - Exemplo 8: Teste Bilateral
   - Exemplo 9: para Duas Amostras Independentes
2. Relação entre Testes de Hipóteses e IC
   - Exemplo 10 - Distribuição Binomial
   - Exemplo 1 - Distribuição Poisson
   - Exemplo 2 (Valor - p)

Exemplo 1: Teste Bilateral para a Média com Variância Conhecida

Considere X uma variável aleatória com média μ e variância finita σ^2. Queremos testar

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0 \tag{1}$$

onde μ_0 é especificado. Sejam $X_1, \ldots, X_n$ uma amostra aleatória i.i.d. da distribuição de X e denotem a média e a variância amostrais por $\bar{X}$ e S^2, respectivamente. Vamos fazer um estudo de sua função poder.

Regra de decisão (bicaudal)

Para o teste **bilateral**, rejeitamos H_0 quando $\bar{X}$ estiver *muito distante* de μ_0. Assim, para as hipóteses (1), usamos a regra de decisão

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \bar{X} \leq h \quad \text{ou} \quad \bar{X} \geq k, \tag{2}$$

onde $h < k$ são tais que

$$\alpha \; = \; P_{H_0}[\bar{X} \leq h \quad \text{ou} \quad \bar{X} \geq k] = P_{H_0}[\bar{X} \leq h] + P_{H_0}[\bar{X} \geq k]. \tag{3}$$

Aproximação assintótica e regiões críticas

Pelo **TCL** e por **Slutsky**, sob H_0 temos, para n grande,

$$T_n = \frac{\sqrt{n}\,(\bar{X} - \mu_0)}{S} \xrightarrow{d} N(0,1).$$

Logo, uma escolha natural é dividir α igualmente entre as duas caudas da *distribuição assintótica de* T_n:

$$P_{H_0}\left[T_n \leq -z_{1-\alpha/2}\right] = \frac{\alpha}{2} \quad \text{e} \quad P_{H_0}\left[T_n \geq z_{1-\alpha/2}\right] = \frac{\alpha}{2}. \tag{4}$$

Isso leva à regra (nível assintótico α):

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \ \text{ se } \left|\frac{\bar{X} - \mu_0}{S/\sqrt{n}}\right| \geq z_{1-\alpha/2}. \tag{5}$$

Aproximação assintótica e regiões críticas

Equivalentemente, em termos de $\bar{X}$,

$$h \approx \mu_0 - z_{1-\alpha/2}\,\frac{S}{\sqrt{n}} \quad \text{e} \quad k \approx \mu_0 + z_{1-\alpha/2}\,\frac{S}{\sqrt{n}}.$$

Observação: Se X é exatamente Normal, então $T_n \sim t_{(n-1)}$ sob H_0 (para qualquer n) e podemos usar $t_{1-\alpha/2,(n-1)}$ no lugar de $z_{1-\alpha/2}$.

Substituindo S por σ e dado que $Z_{1-\alpha/2} = -Z_{\alpha/2}$, segue facilmente que a função poder aproximada é

$$\begin{aligned}\gamma(\mu) &= P_\mu(\bar{X} \leq \mu_0 - |z_{\alpha/2}|\sigma/\sqrt{n}) + P_\mu(\bar{X} \geq \mu_0 + |z_{\alpha/2}|\sigma/\sqrt{n}) \\ &= \Phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - |z_{\alpha/2}|\right) + 1 - \Phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + |z_{\alpha/2}|\right),\end{aligned}$$

em que $\Phi(z)$ é a função de distribuição acumulada de uma variável aleatória normal padrão. Observe que a derivada da função poder é

$$\gamma'(\mu) = \frac{\sqrt{n}}{\sigma}\left(\phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + |z_{\alpha/2}|\right) - \phi\left(\frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - |z_{\alpha/2}|\right)\right)$$

em que $\phi(z)$ é a função de densidade de probabilidade de uma variável aleatória normal padrão.

Exercício 4.6.2

Considere $a = \dfrac{\sqrt{n}(\mu_0 - \mu)}{\sigma}$ e notem que,

- Se $\mu < \mu_0$, então $a > 0$;

Exercício 4.6.2

Considere $a = \dfrac{\sqrt{n}(\mu_0 - \mu)}{\sigma}$ e notem que,

- Se $\mu < \mu_0$, então $a > 0$;
- Se $\mu > \mu_0$, então $a < 0$;

Exercício 4.6.2

Considere $a = \dfrac{\sqrt{n}(\mu_0 - \mu)}{\sigma}$ e notem que,

- Se $\mu < \mu_0$, então $a > 0$;
- Se $\mu > \mu_0$, então $a < 0$;

Podemos reescrever então a derivada da função poder como

$$\gamma'(\mu) = \frac{\sqrt{n}}{\sigma}\left(\phi\left(|z_{\alpha/2}| + a\right) - \phi\left(|z_{\alpha/2}| - a\right)\right),$$

uma vez que $\phi(x) = \phi(-x)$.

Suponha $\mu < \mu_0$

Nesse caso,

$$
\begin{aligned}
|z_{\alpha/2}| + a > |z_{\alpha/2}| - a &\Rightarrow -\frac{(|z_{\alpha/2}| + a)^2}{2} < -\frac{(|z_{\alpha/2}| - a)^2}{2} \\
&\Rightarrow e^{-\frac{(|z_{\alpha/2}| + a)^2}{2}} < e^{-\frac{(|z_{\alpha/2}| - a)^2}{2}} \\
&\Rightarrow \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \left[e^{-\frac{(|z_{\alpha/2}| + a)^2}{2}} - e^{-\frac{(|z_{\alpha/2}| - a)^2}{2}} \right] < 0 \\
&\Rightarrow \gamma'(\mu) < 0
\end{aligned}
$$

Suponha $\mu > \mu_0$

Nesse caso,

$$
\begin{aligned}
|z_{\alpha/2}| + a < |z_{\alpha/2}| - a &\Rightarrow -\frac{(|z_{\alpha/2}| + a)^2}{2} > -\frac{(|z_{\alpha/2}| - a)^2}{2} \\
&\Rightarrow e^{-\frac{(|z_{\alpha/2}| + a)^2}{2}} > e^{-\frac{(|z_{\alpha/2}| - a)^2}{2}} \\
&\Rightarrow \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \left[e^{-\frac{(|z_{\alpha/2}| + a)^2}{2}} - e^{-\frac{(|z_{\alpha/2}| - a)^2}{2}} \right] > \\
&\Rightarrow \gamma'(\mu) > 0
\end{aligned}
$$

Conclusões (Teste bilateral, grandes amostras)

- **Pressupostos:** $X_1, \dots, X_n$ i.i.d., $E[X] = \mu$, $V(X) = \sigma^2 < \infty$; n grande.
- **Estatística:** $T_n = \dfrac{\sqrt{n}(\bar{X} - \mu_0)}{S} \overset{H_0}{\overset{d}{\to}} N(0,1)$.
- **Regra (nível α):** rejeitar H_0 se $|T_n| \geq z_{1-\alpha/2}$.
- **Função poder:** $\gamma(\mu_0) = \alpha$; $\gamma'(\mu) < 0$ se $\mu < \mu_0$ e $\gamma'(\mu) > 0$ se $\mu > \mu_0$; **consistente**: $\gamma(\mu) \to 1$ para $\mu \neq \mu_0$ quando $n \to \infty$.
- **Casos especiais:** se X é Normal, usar t_{n-1} no lugar de z; se σ é conhecido, o teste Z é exato.
- **Prática:** para n pequeno ou caudas pesadas, preferir t, transformações, bootstrap ou testes robustos.

Exemplo 2: Teste Bilateral para a Média com Variância Conhecida

Suponha que desejamos testar

$$H_0 : \mu = 30,000 \text{ versus } H_1 : \mu \neq 30,000. \tag{6}$$

Suponha que $n = 20$ e $\alpha = 0.01$. Então, a regra de rejeição se torna

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \frac{\bar{X} - 30,000}{\sigma/\sqrt{20}} \geq |z_{\frac{0.01}{2}}|. \tag{7}$$

A próxima Figura exibe a curva da função poder para este teste quando $\sigma = 5000$. Para comparação, a curva da função poder para o teste com nível $\alpha = 0.05$ também é apresentada. Veja também shiny da função poder!

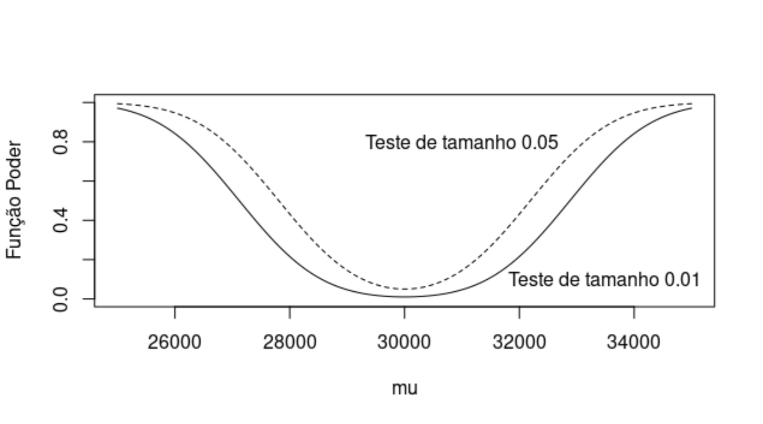

Figura: Função Poder para o teste de hipótese do exemplo

Exemplo 3: Teste Bilateral para a Média com Variância Desconhecida

Se X é Normal, então sob H_0 temos

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}.$$

Logo, o teste bicaudal de tamanho exato α é:

$$\text{Rejeitar } H_0 \text{ se } \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \geq t_{1-\alpha/2,\, n-1}.$$

Ele também possui uma curva da função poder em forma de ‘bacia” semelhante à Figura anterior, embora não seja tão fácil de mostrar; veja Lehmann (1986).

Exemplo 4: Teste Unilateral para a Média com Variância Conhecida

Distribuição Normal:

- **Hipótese Nula (H0):** A média de uma população é igual a 100 e **Hipótese Alternativa (H1):** A média de uma população é maior que 100.
- $n = 30, \ \sigma = 15$ e $\alpha = 0,05$
- **Média real sob H1 (Suposição):** 105

Para calcular a função poder, usamos a distribuição normal padrão (Z) e a fórmula:

$$\gamma(\mu) = P(Z > Z_{1-\alpha} - \frac{\mu - \mu_0}{\sigma/\sqrt{n}})$$

em que $Z_{1-\alpha}$ é o valor crítico para o nível de significância α.

Para $\alpha = 0,05$, $Z_{0,95} \approx 1,645.$
Agora, substituindo os valores:

$$\begin{aligned}\gamma(105) &= P(Z > 1,645 - \frac{105 - 100}{15/\sqrt{30}}) \\ &= P(Z > 1,645 - 1,826) \\ &= P(Z > -0,181)\end{aligned}$$

A probabilidade de Z ser maior que $-0,181$ é aproximadamente $0,5718$. Portanto, o poder do teste é de aproximadamente $0,572$. Logo, com essa amostra, temos cerca de $57{,}2\%$ de chance de rejeitar H_0 quando a média verdadeira é 105. O erro tipo II é $\beta \approx 0{,}428$.

Exemplo 5: Teste Z unilateral para a média (σ conhecido)

Testar

$$H_0: \ \mu \leq \mu_0 \quad \text{vs} \quad H_1: \ \mu > \mu_0,$$

com $\mu_0 = 100$, $n = 36$, $\sigma = 12$, $\alpha = 0{,}05$. Rejeitamos H_0 se

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}.$$

Como $z_{1-\alpha} = z_{0{,}95} \approx 1{,}645$, o ponto crítico em termos de $\bar{X}$ é

$$x_c = \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}} = 100 + 1{,}645 \cdot \frac{12}{6} = 100 + 3{,}29 = 103{,}29.$$

Logo, rejeite H_0 se $\bar{X} > 103{,}29$. Observação: $\alpha = P_{\mu=\mu_0}(\bar{X} > x_c)$.

Cálculo do Poder do Teste

O poder do teste é a probabilidade de rejeitar H_0 quando a verdadeira média μ é maior que μ_0. Assumindo que a verdadeira média é $\mu = 105$:

$$\begin{aligned}\gamma(\mu = 105) &= P_{\mu=105}(\bar{X} > x_c) = P(\bar{X} > 103,29) \\ &= P(Z > \frac{(103,29 - 105) \times 6}{12}) = P(Z > -0,855) \\ &= 0,804\end{aligned}$$

Vejam os gráficos de poder cliquem aqui!

Exemplo 6: Teste unilateral para a Média (σ conhecido)

Vamos testar a seguinte hipótese nula e alternativa:

$$H_0 : \mu = \mu_0 = 50 \quad \text{contra} \quad H_1 : \mu < 50$$

Em que $\mu_0 = 50$ é a média sob H_0, e μ é a média populacional desconhecida. Temos uma amostra de tamanho $n = 36$, com desvio padrão populacional $\sigma = 8$, e o nível de significância $\alpha = 0.05$.

Exemplo 6: Teste unilateral para a Média (σ conhecido)

Vamos testar a seguinte hipótese nula e alternativa:

$$H_0 : \mu = \mu_0 = 50 \quad \text{contra} \quad H_1 : \mu < 50$$

Em que $\mu_0 = 50$ é a média sob H_0, e μ é a média populacional desconhecida. Temos uma amostra de tamanho $n = 36$, com desvio padrão populacional $\sigma = 8$, e o nível de significância $\alpha = 0.05$.

Vamos rejeitar H_0 se $\bar{X} < x_c < 50$, tal que $\alpha = P_{\text{Sob } H_0}(\bar{X} < x_c)$

$$Z_\alpha = -1,645$$

Cálculo do Valor Crítico sob H_0

$$
\begin{aligned}
0,05 = P_{\textsf{Sob } H_0}(\bar{X} < x_c) &= P(Z < \frac{(x_c - 50) \times 6}{8}) \\
\Rightarrow \frac{(x_c - 50) \times 6}{8} &= -1,645 \\
\Rightarrow x_c &= 47,81
\end{aligned}
$$

Cálculo do Valor Crítico sob H_0

$$\begin{aligned}
0,05 = P_{\textsf{Sob } H_0}(\bar{X} < x_c) = P(Z < \frac{(x_c - 50) \times 6}{8})\\
\Rightarrow \frac{(x_c - 50) \times 6}{8} = -1,645\\
\Rightarrow x_c = 47,81
\end{aligned}$$

Cálculo do Poder do Teste

O poder do teste é a probabilidade de rejeitar H_0 quando a verdadeira média μ é menor que μ_0. Vamos assumir que a verdadeira média é $\mu = 47$:

$$\begin{aligned}
\gamma(47) &= P_{\mu=47}(\bar{X} < x_c) = P(\bar{X} < 47,81)\\
&= P(Z < \frac{(47,81 - 47) \times 6}{8}) = P(Z < 0,6075)\\
&= 0,728
\end{aligned}$$

Vejam os gráficos de poder cliquem aqui!

Exemplo 7: Teste unilateral para a Proporção Binomial

Vamos testar a seguinte hipótese nula e alternativa:

$$H_0 : p = p_0 = 0,4 \quad \text{contra} \quad H_1 : p > 0,4$$

O tamanho da amostra é $n = 100$, o nível de significância é $\alpha = 0,01$, e assumimos que a proporção real sob H_1 é $p = 0,55$.

Exemplo 7: Teste unilateral para a Proporção Binomial

Vamos testar a seguinte hipótese nula e alternativa:

$$H_0 : p = p_0 = 0,4 \quad \text{contra} \quad H_1 : p > 0,4$$

O tamanho da amostra é $n = 100$, o nível de significância é $\alpha = 0,01$, e assumimos que a proporção real sob H_1 é $p = 0,55$.

Vamos rejeitar H_0 se $\bar{p} > p_c > 0,4$, tal que $\alpha = P_{\text{Sob } H_0}(\bar{p} > p_c)$

$$Z_\alpha = 2,33$$

Cálculo do Valor Crítico sob H_0

$$0,01 = P_{\textsf{Sob } H_0}(\bar{p} > p_c) = P(Z > \frac{(p_c - 0,4) \times 10}{\sqrt{0,04 \times (1 - 0,04)}})$$

$$\Rightarrow \frac{(p_c - 0,4) \times 10}{\sqrt{0,4 \times (1 - 0,4)}} = 2,33 \Rightarrow p_c = 0,5141$$

Cálculo do Valor Crítico sob H_0

$$0,01 = P_{\text{Sob } H_0}(\bar{p} > p_c) = P(Z > \frac{(p_c - 0,4) \times 10}{\sqrt{0,04 \times (1 - 0,04)}})$$

$$\Rightarrow \frac{(p_c - 0,4) \times 10}{\sqrt{0,4 \times (1 - 0,4)}} = 2,33 \Rightarrow p_c = 0,5141$$

Cálculo do Poder do Teste

O poder do teste é a probabilidade de rejeitar H_0 quando a verdadeira proporção p é maior que $0,4$. Assumindo que a verdadeira proporção é $p = 0,55$:

$$\begin{aligned}\gamma(p = 0,55) &= P_{p=0,55}(\bar{p} > p_c) = P(\bar{p} > 0,5141) \\ &= P(Z > \frac{(0,5141 - 0,55) \times 10}{\sqrt{0,55 \times (1 - 0,55)}}) = P(Z > -0,722) \\ &= 0,7648\end{aligned}$$

Vejam os gráficos de poder cliquem aqui!

Exemplo 8: Teste Bilateral para a Média

Vamos testar:

$$H_0 : \mu = \mu_0 = 50 \quad \text{contra} \quad H_1 : \mu \neq 50$$

com $n = 36$, desvio-padrão populacional conhecido $\sigma = 10$ e nível $\alpha = 0{,}05$.

Exemplo 8: Teste Bilateral para a Média

Vamos testar:

$$H_0 : \mu = \mu_0 = 50 \quad \text{contra} \quad H_1 : \mu \neq 50$$

com $n = 36$, desvio-padrão populacional conhecido $\sigma = 10$ e nível $\alpha = 0{,}05$.

Regra de decisão e quantis

Rejeitar H_0 se $\bar{X} > k$ ou $\bar{X} < h$, com

$$\frac{\alpha}{2} = P_{H_0}(\bar{X} > k) = P_{H_0}(\bar{X} < h), \quad z_{1-\alpha/2} = 1{,}96.$$

Crítico do lado direito (k)

$$0{,}025 = P_{H_0}(\bar{X} > k) = P\left(Z > \frac{k-50}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{(k-50)\cdot 6}{10}\right)$$

$$\Rightarrow \frac{(k-50)\cdot 6}{10} = 1{,}96 \ \Rightarrow \ k = 50 + 1{,}96 \cdot \frac{10}{6} = 53{,}27.$$

Crítico do lado direito (k)

$$0{,}025 = P_{H_0}(\bar{X} > k) = P\left(Z > \frac{k-50}{\sigma/\sqrt{n}}\right) = P\left(Z > \frac{(k-50)\cdot 6}{10}\right)$$

$$\Rightarrow \frac{(k-50)\cdot 6}{10} = 1{,}96 \ \Rightarrow \ k = 50 + 1{,}96 \cdot \frac{10}{6} = 53{,}27.$$

Crítico do lado esquerdo (h)

$$0{,}025 = P_{H_0}(\bar{X} < h) = P\left(Z < \frac{h-50}{\sigma/\sqrt{n}}\right) = P\left(Z < \frac{(h-50)\cdot 6}{10}\right)$$

$$\Rightarrow \frac{(h-50)\cdot 6}{10} = -1{,}96 \ \Rightarrow \ h = 50 - 1{,}96 \cdot \frac{10}{6} = 46{,}73.$$

Função poder (bicaudal)

Para $\mu \in \mathbb{R}$,

$$\gamma(\mu) \;=\; P_\mu(\bar{X} > k) \;+\; P_\mu(\bar{X} < h).$$

Função poder (bicaudal)

Para $\mu \in \mathbb{R}$,

$$\gamma(\mu) \ = \ P_\mu(\bar{X} > k) \ + \ P_\mu(\bar{X} < h).$$

Exemplo numérico: $\mu = 54$

Com $k = 53{,}27$, $h = 46{,}73$ e $\sigma/\sqrt{n} = 10/6 \approx 1{,}667$,

$$\begin{aligned}\gamma(54) = P\left(Z > \frac{53{,}27 - 54}{1{,}667}\right) + P\left(Z < \frac{46{,}73 - 54}{1{,}667}\right) \\ = P(Z > -0{,}438) + P(Z < -4{,}36).\end{aligned}$$

Como $P(Z < -4{,}36) \approx 0$, conclui-se

$$\gamma(54) \approx P(Z > -0{,}438) \approx 0{,}67.$$

Vejam os gráficos de poder cliquem aqui!

Exemplo 9: para Duas Amostras Independentes

Considere amostras aleatórias independentes de $N(\mu_1, \sigma^2)$ e $N(\mu_2, \sigma^2)$, respectivamente. Definimos $n = n_1 + n_2$ como o tamanho combinado da amostra e S_p^2 como o estimador combinado da variância comum, dado por

$$S_c^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n - 2}.$$

A um nível de significância $\alpha = 0.05$, rejeitamos $H_0 : \mu_1 = \mu_2$ em favor da alternativa unilateral $H_1 : \mu_1 > \mu_2$ se

$$T = \frac{\bar{X} - \bar{Y} - 0}{S_c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{0.05, n-2},$$

pois, sob H_0, T segue uma distribuição t com $n-2$ graus de liberdade.

Relação entre Testes de Hipóteses e IC

Existe uma relação entre testes bilaterais e intervalos de confiança. Considere o teste t bilateral. Aqui, usamos a regra de rejeição com "se e somente se" substituindo "se". Portanto, em termos de aceitação, temos Aceitar H_0, se, e somente se,

$$\mu_0 - t_{\alpha/2,n-1}S/\sqrt{n} < \bar{X} < \mu_0 + t_{\alpha/2,n-1}S/\sqrt{n}.$$

Relação entre Testes de Hipóteses e IC

Existe uma relação entre testes bilaterais e intervalos de confiança. Considere o teste t bilateral. Aqui, usamos a regra de rejeição com "se e somente se" substituindo "se". Portanto, em termos de aceitação, temos Aceitar H_0, se, e somente se,

$$\mu_0 - t_{\alpha/2,n-1}S/\sqrt{n} < \bar{X} < \mu_0 + t_{\alpha/2,n-1}S/\sqrt{n}.$$

Isso pode ser facilmente demonstrado como "Aceitar H_0 se, e somente se",

$$\mu_0 \in \left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right).$$

Ou seja, aceitamos H_0 ao nível de significância α se e somente se μ_0 está no intervalo de confiança de $(1-\alpha)100\%$ para μ. De forma equivalente, rejeitamos H_0 ao nível de significância α se, e somente se, μ_0 não está no intervalo de confiança de $(1-\alpha)100\%$ para μ. Isso é válido para todos os testes de hipóteses bilaterais.

Exemplo 10 - Distribuição Binomial

Suponha que X segue uma distribuição binomial com parâmetros 1 e p. Considere o teste de hipótese $H_0 : p = p_0$ contra $H_1 : p < p_0$. Seja $X_1, \ldots, X_n$ uma amostra aleatória da distribuição de X, e seja $\hat{p} = \frac{X}{n}$. Para testar H_0 versus H_1, utilizamos uma das seguintes estatísticas:

$$Z_1 = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \leq c \text{ ou } Z_2 = \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1 - \hat{p})/n}} \leq c.$$

Se o tamanho da amostra n for grande, tanto Z_1 quanto Z_2 têm distribuições normais aproximadas, desde que $H_0 : p = p_0$ seja verdadeira. Portanto, se c for definido como -1.645, o nível de significância aproximado é $\alpha = 0.05$. Ambos os métodos fornecem resultados numéricos semelhantes.

Com uma hipótese alternativa bilateral, Z_2 fornece uma melhor relação com o intervalo de confiança para p. Ou seja, $|Z_2| < z_{\alpha/2}$ é equivalente a p_0 estar no intervalo

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right),$$

que é o intervalo que fornece um intervalo de confiança aproximado de $(1-\alpha)100\%$ para p, conforme discutido na aula de Intervalos de Confiança. Vejam os gráficos de poder cliquem aqui!

Exemplo - Distribuição Poisson

Seja $X_1, X_2, \ldots, X_{10}$ uma amostra aleatória de tamanho $n = 10$ de uma distribuição de Poisson com média θ. A região crítica para testar $H_0 : \theta = 0.1$ contra $H_1 : \theta > 0.1$ é dada por $Y = \sum_{i=1}^{10} X_i \geq 3$. A estatística Y segue uma distribuição de Poisson com média 10θ. Portanto, com $\theta = 0.1$, a média de Y é 1, o nível de significância do teste é

$$P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - 0.920 = 0.080.$$

Por outro lado, se a região crítica definida por $\sum_{i=1}^{10} X_i \geq 4$ for usada, o nível de significância é

$$\alpha = P(Y \geq 4) = 1 - P(Y \leq 3) = 1 - 0.981 = 0.019.$$

Por exemplo, se um nível de significância de aproximadamente $\alpha = 0.05$ for desejado, a maioria dos estatísticos usaria um desses testes, ou seja, eles ajustariam o nível de significância para o teste mais conveniente.

Nível de Significância Observado (p-valor)

Muitos estatísticos não gostam de testes randomizados na prática. Na verdade, muitos estatísticos relatam o que são comumente chamados de níveis de significância observados ou valores-p (para valores de probabilidade). Um exemplo geral é suficiente para explicar os níveis de significância observados.

Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$, em que tanto μ quanto σ^2 são desconhecidos. Considere, primeiro, as hipóteses unilaterais $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$, em que μ_0 é especificado. Escreva a regra de rejeição como

$$\text{Rejeitar } H_0 \text{ em favor de } H_1, \text{ se } \bar{X} \geq k, \tag{8}$$

em que $\bar{X}$ é a média da amostra.

Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição $N(\mu, \sigma^2)$, em que tanto μ quanto σ^2 são desconhecidos. Considere, primeiro, as hipóteses unilaterais $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$, em que μ_0 é especificado. Escreva a regra de rejeição como

$$\text{Rejeitar } H_0 \text{ em favor de } H_1, \text{ se } \bar{X} \geq k, \tag{8}$$

em que $\bar{X}$ é a média da amostra.

Anteriormente, especificamos um nível e, em seguida, resolvemos para k. Na prática, no entanto, o nível não é especificado. Em vez disso, uma vez que a amostra é observada, o valor realizado $\bar{x}$ de $\bar{X}$ é calculado e fazemos a pergunta: O valor $\bar{x}$ é suficientemente grande para rejeitar H_0 em favor de H_1?

Para responder a isso, calculamos o valor-p, que é a probabilidade,

$$\text{valor-p} = P_{\text{Sob } H_0}(\bar{X} \geq \bar{x}).$$

Observe que este é um "nível de significância" baseado nos dados, e chamamos isso de nível de significância observado ou valor-p.

A hipótese H_0 é rejeitada em todos os níveis maiores ou iguais ao valor-p. Por exemplo, se o valor-p for $0,048$ e o nível nominal α for $0,05$, então H_0 será rejeitada; no entanto, se o nível nominal α for $0,01$, então H_0 não será rejeitada. Em resumo, o experimentador define as hipóteses; o estatístico seleciona a estatística de teste e a regra de rejeição; os dados são observados e o estatístico relata o valor-p para o experimentador; e o experimentador decide se o valor-p é suficientemente pequeno para justificar a rejeição de H_0 em favor de H_1. O próximo exemplo fornece uma ilustração numérica.

Exemplo (Valor - p)

Considere os dados de Darwin do Exemplo 4.5.5 do livro do Hogg (Edição 8). Os dados são um design emparelhado sobre as alturas de plantas de Zea mays cruzadas e autofertilizadas. Em cada um dos 15 vasos, uma planta cruzada e uma autofertilizada foram cultivadas. Os dados de interesse são as 15 diferenças emparelhadas, (cruzada - autofertilizada). Como no Exemplo, deixe X_i denotar a diferença emparelhada para o i-ésimo vaso. Deixe μ ser a verdadeira diferença média. As hipóteses de interesse são $H_0 : \mu = 0$ versus $H_1 : \mu > 0$. A regra de rejeição padronizada é

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } T \geq k, \tag{9}$$

em que $T = \frac{\bar{X}}{S/\sqrt{15}}$, $\bar{X}$ e S são, respectivamente, a média amostral e o desvio padrão das diferenças.

A hipótese alternativa afirma que, em média, as plantas cruzadas são mais altas do que as plantas autofertilizadas. Do Exemplo 4.5.5, a estatística do teste t tem o valor 2,15. Deixando $t(14)$ denotar uma variável aleatória com a distribuição t com 14 graus de liberdade e usando R, o valor-p para o experimento é

$$P[t(14) > 2.15] = 1 - pt(2.15, 14) = 1 - 0.9752 = 0.0248.$$

Na prática, com esse valor-p, H_0 seria rejeitada em todos os níveis maiores ou iguais a 0,0248.

Suponha que as hipóteses sejam $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$. Obviamente, o valor-p observado neste caso é

$$\text{valor-p} = P_{\text{Sob } H_0}(\bar{X} \geq \bar{x}).$$

Para a hipótese bilateral $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, nossa regra de rejeição "não especificada" é

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \bar{X} \leq l \text{ ou } \bar{X} \geq k. \tag{10}$$

Para o valor-p, calculamos cada um dos valores-p de um lado, pegamos o menor valor-p e o dobramos. Como ilustração, no exemplo de Darwin, suponha que as hipóteses sejam $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$. Então, o valor-p é $2 \times (0,0248) = 0,0496$.

Referências I

Hogg, RV, J McKean e AT Craig (2019). *Introduction to Mathematical Statistics*.