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

Sumário

Teste Bilateral para a Média Baseado em Grandes Amostras

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 da distribuição de X e denotem a média e a variância da amostra por $\bar{X}$ e S^2, respectivamente.

Para o teste unilateral, rejeitamos H_0 se $\bar{X}$ for muito grande. Portanto, para as hipóteses (1), usamos a regra de decisão

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

onde h e k são tais que $\alpha = P_{H_0}[\bar{X} \leq h \text{ ou } \bar{X} \geq k]$. Claramente, $h < k$; portanto, temos

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

Uma vez que, pelo menos para amostras grandes, a distribuição de X é simétrica em torno de μ_0, sob H_0, uma regra intuitiva é dividir α igualmente entre os dois termos do lado direito da expressão acima; isto é, h e k são escolhidos de forma que

$$P_{H_0}[\bar{X} \leq h] = \frac{\alpha}{2} \quad \text{e} \quad P_{H_0}[\bar{X} \geq k] = \frac{\alpha}{2}. \tag{4}$$

Sabemos que, para amostras grandes, $(\bar{X} - \mu_0)/(S/\sqrt{n})$ é aproximadamente $N(0, 1)$. Isso e (4) levam à regra de decisão aproximada

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

Substituindo S por σ, 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.

Exemplo 2: Exercício 4.6.2

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

- Se $\mu < \mu_0$, então $a > 0$;

Exemplo 2: 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 $a < 0$;

Exemplo 2: 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,

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

Suponha que desejamos testar

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

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

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

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

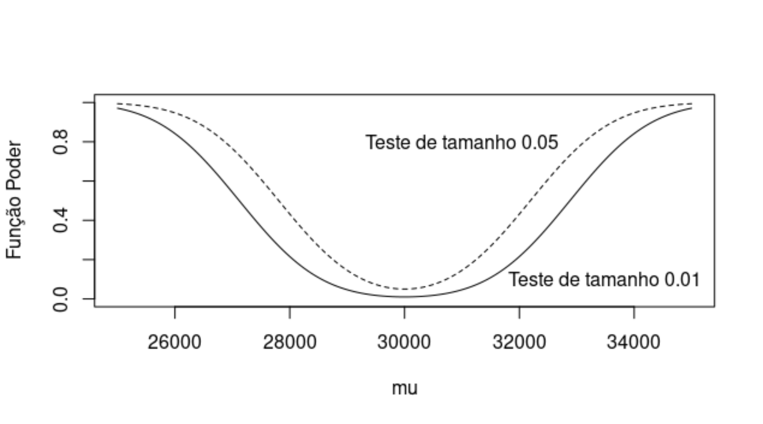

Figura: Função Poder para o teste de hipótese do exemplo

Se assumirmos que X tem uma distribuição normal, então, o seguinte teste tem tamanho exato α para testar $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$:

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_{\alpha/2,n-1}. \tag{8}$$

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).

Distribuição Normal:

- **Hipótese Nula (H0):** A média de uma população é igual a 100.
- **Hipótese Alternativa (H1):** A média de uma população é maior que 100.
- **Tamanho da amostra:** 30
- **Desvio padrão populacional conhecido:** 15
- **Nível de significância:** 0,05
- **Média real sob H1 (Suposição):** 105

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

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

em que Z_α é o valor crítico para o nível de significância α.

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

$$\begin{aligned}
\text{Poder} &= P(Z > 1,645 - \frac{105 - 100}{15/\sqrt{30}}) \\
&= P(Z > 1,645 - 2,738) \\
&= P(Z > -1,093)
\end{aligned}$$

A probabilidade de Z ser maior que $-1,093$ é aproximadamente $0,8628$. Portanto, o poder do teste é de aproximadamente $0,8628$.

Distribuição Binomial

- **Hipótese Nula (H0):** A proporção de sucesso em um experimento binomial é de 0,4.
- **Hipótese Alternativa (H1):** A proporção de sucesso em um experimento binomial é maior que 0,4.
- **Tamanho da amostra:** 100
- **Nível de significância:** 0,01
- **Proporção real sob H1 (Suposição):** 0,55

$$\text{Poder} = 1 - P(X \leq x)$$

Onde x é o valor crítico que corresponde ao nível de significância α. Para $\alpha = 0,01$, x é aproximadamente 52.

Agora, substituindo os valores:

$$\text{Poder} = 1 - P(X \leq 52)$$

Usando uma calculadora ou software estatístico, encontramos $P(X \leq 52) \approx 0,307$. Portanto, o poder do teste é aproximadamente $1 - 0,307 = 0,693$.

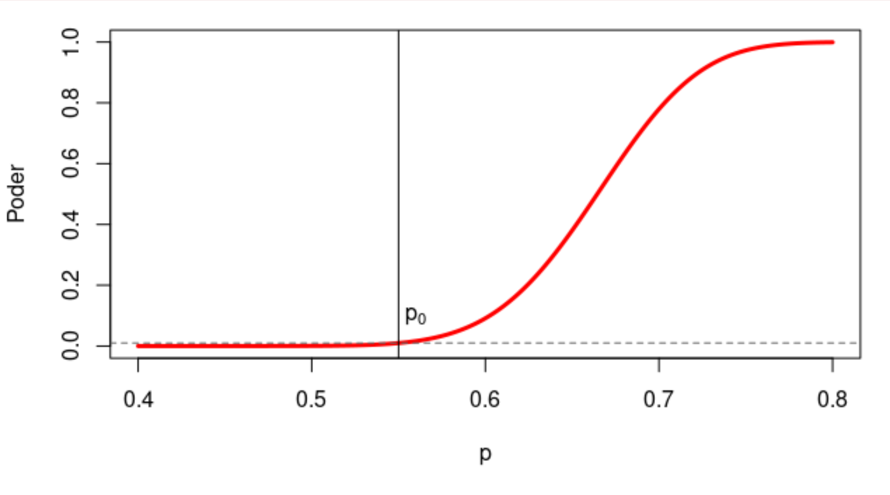

Figura: Função Poder de um teste unilateral usando a distribuição binomial

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 3

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

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

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

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

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

Exemplo 4

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.

O uso de Z_1 fornece melhores probabilidades para cálculos de poder se o verdadeiro valor de p estiver próximo de p_0, enquanto Z_2 é melhor quando H_0 for claramente falsa. No entanto, 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.

Exemplo 5 - 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 $\theta = 0.1$, de modo que a média de Y seja igual a 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.

No entanto, um nível de significância de $\alpha = 0.05$ pode ser alcançado da seguinte maneira. Seja W uma variável aleatória com distribuição de Bernoulli com probabilidade de sucesso igual a

$$P(W = 1) = \frac{\alpha - 0.019}{0.08 - 0.019} = \frac{0.050 - 0.019}{0.080 - 0.019} = \frac{31}{61}.$$

Suponha que W seja selecionado independentemente da amostra. Considere a regra de rejeição: rejeitar H_0 se $\sum_{i=1}^{10} X_i \geq 4$ ou se $\sum_{i=1}^{10} X_i = 3$ e $W = 1$. O nível de significância dessa regra é

$$\begin{aligned} P_{H_0}(Y \geq 4) + P_{H_0}(\{Y = 3\} \cap \{W = 1\}) &= P_{H_0}(Y \geq 4) \\ &\quad + P_{H_0}(Y = 3) P(W = 1) \\ &= 0.019 + 0.061 \times \frac{31}{61} \\ &= 0.05. \end{aligned}$$

Portanto, a regra de decisão tem exatamente um nível de 0.05. O processo de realizar o experimento auxiliar para decidir se rejeita ou não quando $Y = 3$ é às vezes referido como um teste randomizado.

Nível de Significância Observado (p-valor)

Muitos estatísticos não gostam de testes randomizados na prática, pois o uso deles implica que dois estatísticos podem fazer as mesmas suposições, observar os mesmos dados, aplicar o mesmo teste e, no entanto, tomar decisões diferentes. Portanto, eles costumam ajustar seu nível de significância para evitar a aleatoriedade. 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

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

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

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

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,

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

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 6 (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

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } T \geq k, \tag{10}$$

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

$$\text{valor-p} = P(H_0(\bar{X} \leq \bar{x})).$$

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

$$\text{Rejeitar } H_0 \text{ em favor de } H_1 \text{ se } \bar{X} \leq l \text{ ou } \bar{X} \geq k. \tag{11}$$

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 $2 \times (0,0248) = 0,0496$.

Como nota final sobre valores-p para hipóteses bilaterais, suponha que a estatística de teste possa ser expressa em termos de uma estatística de teste t. Nesse caso, o valor-p pode ser encontrado de forma equivalente da seguinte maneira. Se d for o valor realizado da estatística de teste t, então o valor-p é

$$\text{valor-p} = P(H_0[|t| \geq |d|]), \tag{12}$$

em que, sob H_0, t tem uma distribuição t com $n-1$ graus de liberdade.

Como nota final sobre valores-p para hipóteses bilaterais, suponha que a estatística de teste possa ser expressa em termos de uma estatística de teste t. Nesse caso, o valor-p pode ser encontrado de forma equivalente da seguinte maneira. Se d for o valor realizado da estatística de teste t, então o valor-p é

$$\text{valor-p} = P(H_0[|t| \geq |d|]), \tag{12}$$

em que, sob H_0, t tem uma distribuição t com $n-1$ graus de liberdade.

Nessa discussão sobre valores-p, lembre-se de que a boa ciência dita que as hipóteses devem ser conhecidas antes que os dados sejam coletados.

Explicação Detalhada:

O valor-p é uma medida estatística que ajuda a avaliar a força das evidências contra uma hipótese nula em um teste de hipóteses. Ele é calculado com base nos dados observados e na distribuição da estatística de teste sob a suposição de que a hipótese nula é verdadeira.

Explicação Detalhada:

O valor-p é uma medida estatística que ajuda a avaliar a força das evidências contra uma hipótese nula em um teste de hipóteses. Ele é calculado com base nos dados observados e na distribuição da estatística de teste sob a suposição de que a hipótese nula é verdadeira.

Ele não é estritamente uma probabilidade no sentido tradicional. Embora seja comumente interpretado como uma probabilidade, é uma medida de evidência estatística em vez de uma probabilidade direta de um evento.

Explicação Detalhada:

O valor-p é uma medida estatística que ajuda a avaliar a força das evidências contra uma hipótese nula em um teste de hipóteses. Ele é calculado com base nos dados observados e na distribuição da estatística de teste sob a suposição de que a hipótese nula é verdadeira.

Ele não é estritamente uma probabilidade no sentido tradicional. Embora seja comumente interpretado como uma probabilidade, é uma medida de evidência estatística em vez de uma probabilidade direta de um evento.

O valor-p é, portanto, uma medida de quão consistentes os dados observados são com a hipótese nula. É a probabilidade de obter resultados tão extremos quanto os observados, assumindo que a hipótese nula seja verdadeira.

Aqui estão as etapas para entender o valor-p:

- Formulação das Hipóteses:
  A primeira etapa é estabelecer duas hipóteses: a hipótese nula (H_0) e a hipótese alternativa (H_1).

Aqui estão as etapas para entender o valor-p:

- Formulação das Hipóteses:
  A primeira etapa é estabelecer duas hipóteses: a hipótese nula (H_0) e a hipótese alternativa (H_1).
- Coleta de Dados:
  Em seguida, você coleta dados relevantes para o teste de hipóteses.

Aqui estão as etapas para entender o valor-p:

- Formulação das Hipóteses:
  A primeira etapa é estabelecer duas hipóteses: a hipótese nula (H_0) e a hipótese alternativa (H_1).
- Coleta de Dados:
  Em seguida, você coleta dados relevantes para o teste de hipóteses.
- Cálculo da Estatística de Teste:
  Você calcula uma estatística de teste com base nos dados. A escolha da estatística depende do tipo de teste e da pergunta de pesquisa.

- Determinação do Valor-p
  Com a estatística de teste em mãos, você calcula o Valor-p. Esse cálculo envolve a probabilidade de obter uma estatística de teste tão extrema quanto a observada, supondo que a hipótese nula seja verdadeira. Essa probabilidade é chamada de Valor-p.

- Determinação do Valor-p
  Com a estatística de teste em mãos, você calcula o Valor-p. Esse cálculo envolve a probabilidade de obter uma estatística de teste tão extrema quanto a observada, supondo que a hipótese nula seja verdadeira. Essa probabilidade é chamada de Valor-p.
- Interpretação do Valor-p
  O valor *p* é interpretado comparando-o a um nível de significância pré-definido, geralmente denotado como α. Se o valor *p* for menor ou igual a α, é comum rejeitar a hipótese nula em favor da hipótese alternativa. Quanto menor o valor *p*, mais fortes são as evidências contra a hipótese nula.

Exemplo 7

Suponha que um fabricante afirma que a média de vida útil de suas lâmpadas é de 1000 horas. Você suspeita que as lâmpadas, na verdade, têm uma vida útil média diferente.

- Hipóteses

$$H_0 : \mu = 1000 \text{ horas}$$
$$H_1 : \mu \neq 1000 \text{ horas}$$

Exemplo 7

Suponha que um fabricante afirma que a média de vida útil de suas lâmpadas é de 1000 horas. Você suspeita que as lâmpadas, na verdade, têm uma vida útil média diferente.

- Hipóteses

$$H_0 : \mu = 1000 \text{ horas}$$
$$H_1 : \mu \neq 1000 \text{ horas}$$

- Coleta de Dados

Você coleta uma amostra de 30 lâmpadas e calcula a média da vida útil das lâmpadas da amostra. Suponha que você obtém uma média amostral de 980 horas.

- Cálculo da Estatística de Teste

Para este exemplo, a estatística de teste é o valor da média amostral subtraído do valor hipotético (1000) dividido pelo erro padrão. Neste caso:

$$\text{Estatística de Teste} = \frac{980 - 1000}{\text{erro padrão}}$$

- Cálculo da Estatística de Teste

Para este exemplo, a estatística de teste é o valor da média amostral subtraído do valor hipotético (1000) dividido pelo erro padrão. Neste caso:

$$\text{Estatística de Teste} = \frac{980 - 1000}{\text{erro padrão}}$$

- Determinação do Valor-p

Você calcula o Valor-p, que é a probabilidade de obter uma estatística de teste tão extrema quanto a observada (no caso, -20) se a média real fosse 1000 horas, sob a suposição da distribuição de médias amostrais.

- Interpretação do Valor-p

Suponhamos que o Valor-p calculado seja $0,03$. Isso significa que, se a hipótese nula ($\mu = 1000$) fosse verdadeira, haveria uma probabilidade de 3% de obter uma média amostral tão extrema quanto 980 horas. Como esse Valor-p é menor que um nível de significância (α) típico de $0,05$, você pode optar por rejeitar a hipótese nula. Isso sugere que as lâmpadas podem ter uma vida útil média diferente de 1000 horas.

- Interpretação do Valor-p

Suponhamos que o Valor-p calculado seja $0,03$. Isso significa que, se a hipótese nula ($\mu = 1000$) fosse verdadeira, haveria uma probabilidade de 3% de obter uma média amostral tão extrema quanto 980 horas. Como esse Valor-p é menor que um nível de significância (α) típico de $0,05$, você pode optar por rejeitar a hipótese nula. Isso sugere que as lâmpadas podem ter uma vida útil média diferente de 1000 horas.

Lembre-se de que o Valor-p não fornece a magnitude da diferença; ele simplesmente indica se as evidências observadas são consistentes com a hipótese nula. A interpretação do Valor-p deve ser feita em conjunto com o contexto e a pergunta de pesquisa. Quanto menor o Valor-p, mais forte é a evidência contra a hipótese nula.

Referências I