

Testes de hipótese: discussão

- Como construir um teste que **quase sempre** rejeita H_0 ;
- Significância estatística vs significância prática;
- Rapidinhas.

Um teste esquisito

Suponha que temos X_1, X_2, \dots, X_n vindos de uma distribuição Normal com média θ e variância 1 e queremos testar as hipóteses

$$H_0 : \theta = 0,$$

$$H_1 : \theta = 1.$$

Seguindo o exemplo 9.2.5 de DeGroot, podemos escrever

$$\eta(\mathbf{x}) = \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})},$$

e compor um teste que rejeita H_0 quando $\eta(\mathbf{x}) > c$. Isto é equivalente a construir um teste de tamanho α_0 , de modo que valha

$$\Pr(\bar{X}_n \geq c' \mid \theta = 0) = \alpha_0,$$

o que nos leva a concluir que $c' = \frac{1}{2} + \frac{\log(c)}{n}$ e que $c = \Phi^{-1}(1 - \alpha_0)/\sqrt{n}$.

Qual o problema?

Primeiro, vamos lembrar que, para um teste δ ,

$$\alpha(\delta) := \Pr(\text{Rejeitar } H_0 \mid \theta = 0),$$

$$\beta(\delta) := \Pr(\text{Não rejeitar } H_0 \mid \theta = 1).$$

O problema aqui é que para este teste temos

n	$\alpha(\delta)$	$\beta(\delta)$	c
1	0.05	0.74	0.72
25	0.05	3.97×10^{-4}	2.3×10^{-4}
100	0.05	8×10^{-15}	2.7×10^{-15}

Ou seja, quando temos $n = 100$ observações, os dados podem ser trilhões de vezes mais prováveis sob H_0 e ainda assim vamos rejeitar a hipótese nula.

Soluções

Podemos pensar em duas soluções (complementares) para o problema posto.

Ideia 7 (Ajustando o nível de significância com o tamanho da amostra)

Em várias situações, por exemplo como a mostrada acima, faz sentido ajustar (diminuir) o nível de confiança do teste com o tamanho da amostra de modo a balancear os erros do tipo I e II.

Ideia 8 (Minimizar uma combinação linear das probabilidades de erro)

Poderíamos balancear os erros ao minimizar

$$a\alpha(\delta) + b\beta(\delta).$$

Lehmann (1958)²⁰ propôs a restrição $\beta(\delta) = c\alpha(\delta)$, que tem a vantagem de forçar que ambos os tipos de erro diminuam à medida que obtemos mais dados.

Ver seções 9.2 e 9.8 de DeGroot.

²⁰Lehmann, Erich L. "Significance level and power." The Annals of Mathematical Statistics (1958): 1167-1176.

Relevante?

Suponha que eu estou testando uma nova droga, e o parâmetro θ mede o efeito da droga. Em geral, estamos interessados em testar a hipótese

$$H_0 : \theta \leq 0,$$

$$H_1 : \theta \geq 0.$$

Quando o tamanho de amostra é muito grande, seremos capazes de detectar, com alta probabilidade (poder) se $\theta = 0.000003$ ou $\theta = 0$.

Acontece que uma droga com $\theta = 0.000003$ não oferece nenhuma vantagem prática. Portanto, ao se realizar um teste de hipótese e rejeitar H_0 , não podemos concluir que “a droga funciona”, pelo menos não num sentido médico.

Responda rápido

- a) O que é a função poder de um teste de hipótese e o que esperamos observar em um teste não-enviesado?
- b) Se testarmos uma hipótese um número suficiente de vezes ela eventualmente será rejeitada. Explique esta afirmação e suas consequências.
- c) O que é o p-valor de um teste?
- d) É correto afirmar que uma hipótese nula é falsa se ela for rejeitada? É correto afirmar que uma hipótese alternativa é verdadeira se a nula for rejeitada? Justifique.
- e) Um intervalo de confiança nível de 95% para θ é calculado a partir de n observações. É correto afirmar que o parâmetro verdadeiro θ_0 está dentro deste intervalo com probabilidade 95%? Justifique.
- f) Explique como podemos obter um conjunto de confiança a partir de um teste de hipótese.

O que aprendemos?

- 💡 Rejeição eventual; “Se coletarmos uma quantidade suficiente de dados, podemos rejeitar qualquer hipótese nula”
- 💡 Significância estatística \neq significância prática/científica!

Leitura recomendada

 DeGroot seções 9.2, 9.3 e 9.9;

- **Exercícios recomendados**

- 🔖 DeGroot, seção 9.9: exercícios 2 e 3.