- Hipótese nula e alternativa;
- Hipóteses simples e compostas;
- Região crítica e estatística teste;
- Função poder;
- Tipos de erro (I e II);
- P-valor;

No teste de hipóteses estatísticas, identificamos partições do espaço de parâmetros que codificam as hipóteses de interesse.

Definição 40 (Hipótese nula e hipótese alternativa)

Considere o espaço de parâmetros Ω e defina $\Omega_0, \Omega_1 \subset \Omega$ de modo que $\Omega_0 \cup \Omega_1 = \Omega$ e $\Omega_0 \cap \Omega_1 = \emptyset$. Definimos

$$H_0 := \theta \in \Omega_0,$$
$$H_1 := \theta \in \Omega_1.$$

*Dizemos que H_0 é a **hipótese nula** e H_1 é a **hipótese alternativa**.*
Se $\theta \in \Omega_1$, dizemos que rejeitamos a hipótese nula. Por outro lado, se $\theta \in \Omega_0$ dizemos que não rejeitamos ou falhamos em rejeitar H_0.

Exemplo

Suponha que Palmirinha recebeu uma carta da Associação Nacional da Pamonha Gourmet (ANPG), dizendo que a pamonha deve ter, no mínimo, 7 mg/L de concentração de amido. Supondo que a concentração de amido tenha distribuição Normal com parâmetros μ (desconhecido) e σ^2 (conhecido), Palmirinha rabisca num papel:

$$H_0 := \mu \in [7, \infty),$$
$$H_1 := \mu \in (0, 7).$$

Dependendo do tipo de partição do espaço de parâmetros, as hipóteses recebem classificações diferentes.

Definição 41 (Hipótese simples e hipótese compostas)

Dizemos que uma hipótese H_i, é ***simples****, se $\Omega_i = \{\theta_i\}$, isto é, se a partição correspondente é um ponto. Uma hipótese é dita* ***composta*** *se não é simples.*

Exemplo 17 (Hipótese simples sobre a média)

Suponha que estamos estudando o efeito de uma droga na redução da pressão arterial. Modelamos esta redução como uma variável aleatória X com esperança $E[X] =: \theta$. É costumaz testar a hipótese $H_0 : \theta = 0$, que chamamos, especificamente nesse caso, de "hipótese de efeito nulo".

Hipótese unilateral e hipótese bilateral

Em analogia com os intervalos de confiança, também podemos entender as hipóteses como sendo unilaterais ou bilaterais.

Definição 42 (Hipótese unilateral e hipótese bilateral)

Uma hipótese da forma $H_0 : \theta \leq \theta_0$ *ou* $H_0 : \theta \geq \theta_0$ *é dita unilateral ("one-sided"), enquanto hipóteses da forma* $H_0 : \theta \neq \theta_0$ *são ditas bilaterais ("two-sided").*

Observação 20 (Hipóteses bilaterais como consequência de H_0 simples)

Se H_0 *é simples, a hipótese alternativa* H_1 *será, em geral, bilateral.*

Região crítica: exemplo motivador

Exemplo 18 (Teste para a média de uma Normal com variância conhecida)

Suponha que $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ *é uma amostra aleatória de uma Normal com média* μ *e variância* σ^2 *conhecida. Queremos testar a hipótese*

$$
\begin{aligned}
H_0 &: \mu = \mu_0 \\
H_1 &: \mu \neq \mu_0.
\end{aligned}
$$

Intuitivamente, queremos rejeitar H_0 *se* $\bar{X}_n$ *está longe de* μ_0. *Para isso definimos*

$$
S_0 := \left\{\boldsymbol{x} : -c \leq \bar{X}_n - \mu_0 \leq c\right\},
$$

de modo que $S_1 = S_0^C$. *Então, seguimos o procedimento:*

$$
\begin{aligned}
\boldsymbol{X} \in S_1 &\implies \text{rejeitar } H_0, \\
\boldsymbol{X} \in S_0 &\implies \text{não rejeitar } H_0.
\end{aligned}
$$

Região crítica e região de rejeição

Uma maneira mais simples de expressar o procedimento acima é definir $T := |\bar{X}_n - \mu_0|$ e rejeitar H_0 se $T \geq c$.

Definição 43 (Região crítica)

O conjunto

$$S_1 := \{ \boldsymbol{x} : |\bar{X}_n - \mu_0| \geq c \},$$

é chamado de ***região crítica*** *do teste.*

Analogamente, considere a estatística $T = r(\boldsymbol{X})$ e tome $R \subseteq \mathbb{R}$. Então podemos definir

Definição 44 (Região de rejeição)

Se $R \subseteq \mathbb{R}$ é tal que dizemos que "rejeitamos H_0 se $T \in R$", então R é chamada uma ***região de rejeição*** *para a estatística T e o teste associado.*

Dividindo o espaço amostral e o espaço de parâmetros

Começamos com uma observação:

Observação 21 (Correspondência entre região crítica e região de rejeição)

Podemos relacionar os conceitos de região crítica e região de rejeição notando queremos

$$S_1 := \{\mathbf{x} : r(\mathbf{x}) \in R\}.$$

Ideia 4 (Dividindo o espaço amostral e o espaço de parâmetros)

Suponha que temos um modelo estatístico dado pela distribuição $f(x \mid \theta)$, com $x \in \mathcal{X}$ e $\theta \in \Omega$. Desta forma, uma amostra aleatória $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ mora em $\mathcal{X}^n$. Para formular uma hipótese estatística, estabelecemos uma partição do espaço de parâmetros Ω em Ω_0 e Ω_1 disjuntos. Isto, por sua vez, induz uma partição $S_0, S_1 \in \mathcal{X}^n$. Estes objetos, embora, relacionados, ***não são a mesma coisa****. Por exemplo, nós observamos se $\boldsymbol{X} \in S_0$ ou $\boldsymbol{X} \in S_1$, mas raramente "observamos" se $\theta \in \Omega_0$ ou $\theta \in \Omega_1$.*

Nossa capacidade de rejeitar H_0 depende do valor de $\theta \in \Omega$. Esta dependência é capturada pela função poder.

Definição 45 (Função poder)

Seja δ um procedimento de aceitação/rejeição como visto anteriormente. A ***função poder*** *é definida como*

$$\pi(\theta \mid \delta) := \Pr\left(\boldsymbol{X} \in S_1 \mid \theta\right) = \Pr\left(T \in R \mid \theta\right), \theta \in \Omega. \tag{28}$$

Idealmente, queremos $\pi(\theta \mid \delta) = 1$ para $\theta \in \Omega_1$ (por quê?).

A função poder de um teste para a média da Normal

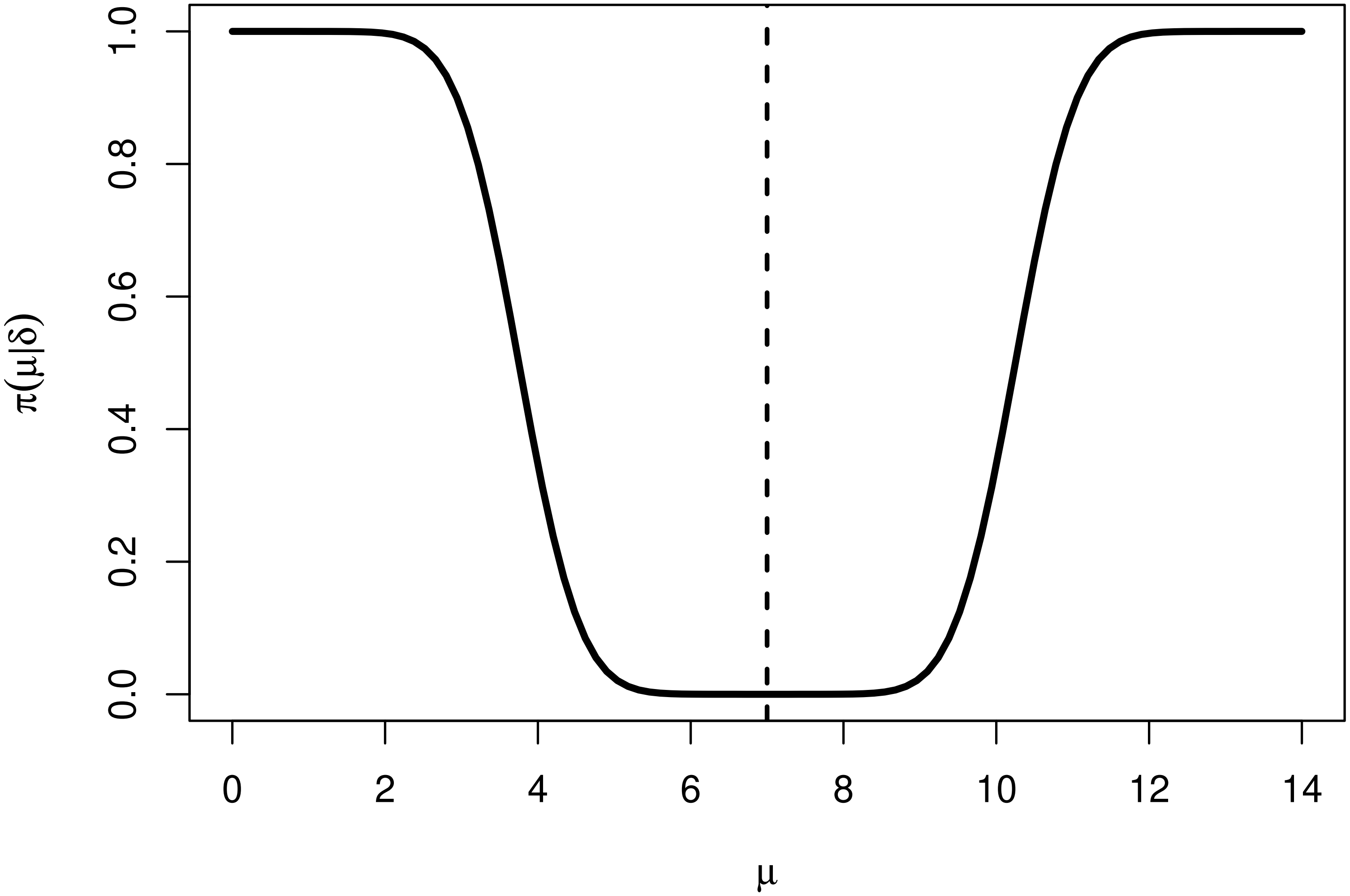
Considere a situação em que $X_1, X_2, \ldots, X_n$ vêm de uma Normal com média μ, desconhecida, e variância σ^2, conhecida.

Exemplo 19 (Função poder no teste para média da Normal (σ^2 conhecida))

Lembrando que $T = |\bar{X}_n - \mu_0|$, e tomando δ como o procedimento descrito acima, escrevemos

$$\begin{aligned}
\pi(\mu \mid \delta) &= \Pr(T \in R \mid \mu), \\
&= \Pr\left(\bar{X}_n \geq \mu_0 + c \mid \mu\right) + \Pr\left(\bar{X}_n \leq \mu_0 - c \mid \mu\right), \\
&= \left\{1 - \Phi\left(\sqrt{n}\frac{\mu_0 + c - \mu}{\sigma}\right)\right\} + \Phi\left(\sqrt{n}\frac{\mu_0 - c - \mu}{\sigma}\right).
\end{aligned}$$

Pelo poder da pamonha...



Tipos de Erro

Quando testamos uma hipótese, nunca estamos livres de cometer um erro. É conveniente classificar os possíveis erros em duas categorias.

Definição 46 (Tipos de erros)

Nome	*Erro cometido*
Erro tipo I	*Rejeitar H_0 quando ela é **verdadeira**.*
Erro tipo II	*Falhar em rejeitar H_0 quando ela é **falsa**.*

Isto nos leva a concluir que

Situação	Quantidade	Interpretação
$\theta \in \Omega_0$	$\pi(\theta \mid \delta)$	Pr(Erro tipo I)
$\theta \in \Omega_1$	$1 - \pi(\theta \mid \delta)$	Pr(Erro tipo II)

Balanceando um teste

Idealmente, gostaríamos de um teste δ para o qual as probabilidades de erro fossem as menores possíveis. Infelizmente, em geral, diminuir o erro tipo I implica aumentar o erro tipo II.
Em geral, precisamos encontrar um equilíbrio entre os tipos de erros.

Ideia 5 (Encontrando um balanço entre erro tipo I e tipo II)

Tome $0 < \alpha_0 < 1$. Nós construímos o procedimento δ^* de modo que

$$\pi(\theta \mid \delta^*) \leq \alpha_0, \; \forall\, \theta \in \Omega. \tag{29}$$

Então, entre todos os testes que satisfazem (29), buscamos o teste que tenha $\pi(\theta \mid \delta^*)$ máxima em $\theta \in \Omega_1$.

Tamanho de um teste

Definição 47 (Tamanho/nível de um teste)

*Dizemos que um teste, δ, tem **tamanho** ou **nível de significância** $\alpha(\delta)$, com*

$$\alpha(\delta) := \sup_{\theta \in \Omega_0} \pi(\theta \mid \delta).$$

Um teste que atende à condição anterior (29) tem que tamanho?

Observação 22 (Tamanho de um teste com H_0 simples)

Se H_0 é simples, então $\alpha(\delta) = \pi(\theta_0 \mid \delta)$.

Exemplo 20 (Teste para o parâmetro de uma uniforme)

Suponha que $X_1, X_2, \ldots, X_n$ *tem distribuição Uniforme em* $[0, \theta]$, *com* θ *desconhecido, e que aventamos as seguintes hipóteses:*

$$H_0 : 3 \leq \theta \leq 4,$$
$$H_1 : \theta < 3 \text{ ou } \theta > 4$$

Lembre que $\hat{\theta}_{EMV} = \max\{X_1, X_2, \ldots, X_n\}$ *e suponha que temos um teste* δ *da forma*

Condição	*Ação*
$\hat{\theta}_{EMV} \notin (2.9, 4)$	*Rejeitar* H_0
$\hat{\theta}_{EMV} \in (2.9, 4)$	*Falhar em rejeitar* H_0.

- Qual a região de rejeição para δ?
- Como escrever $\pi(\theta \mid \delta)$?
- Qual o tamanho de δ?

Construindo um teste com o tamanho adequado

Em geral, sempre conseguimos construir um teste que tenha o tamanho desejado.

Observação 23 (Construindo um teste de tamanho α_0)

Se $T = r(\boldsymbol{X})$ é uma estatística, podemos quase sempre encontrar c tal que valha

$$\sup_{\theta \in \Omega_0} \Pr(T \geq c \mid \theta) \leq \alpha_0, \tag{30}$$

ou seja, encontrar c tal que δ tenha tamanho (ou nível de significância) α_0.

Começamos com uma observação:

Observação 24 (Testes são decisões binárias)

Um teste de hipótese reduz a informação contida nos dados a uma decisão binária: rejeitar ou não H_0. Se observamos $T = c + 10^{-10}$ ou $T = c + 10^{10}$, tomamos a mesma decisão de rejeitar H_0 ao nível α_0.

Ao invés disso, podemos reportar o maior nível de significância que ainda levaria à rejeição de H_0.

Definição 48 (O p-valor)

Para cada t, seja δ_t o teste que rejeita H_0 se $T \geq t$. Então, quando $T = t$, o ***p-valor*** *vale*

$$p(t) := \sup_{\theta \in \Omega_0} \pi(\theta \mid \delta_t) = \sup_{\theta \in \Omega_0} \Pr\left(T \geq t \mid \theta\right), \tag{31}$$

ou seja, o p-valor é o tamanho do teste δ_t.

Exemplo: tá me enganando, parceiro?

Vamos voltar a uma pergunta não respondida lá no início do curso. Suponha que você encontre um "artista" de rua, que joga uma moeda e pede para as pessoas apostarem se vai dar cara ou coroa. Conhencendo estatística e probabilidade, você decide (i) observar o jogo à distância para coletar alguns dados (ii) fazer algumas contas para ver se vale a pena apostar.

Pergunta 4 (Esta moeda é justa? cont. I)

Suponha que uma moeda tenha sido lançada dez vezes, obtendo o seguinte resultado:

$$KKKCKCCCKC$$

a) *Esta moeda é justa?*

b) *Quanto eu espero ganhar se apostar R$ 100,00 que é justa?*

Hoje vamos dar uma resposta parcial à pergunta a).

O que aprendemos?

- Hipóteses nula e alternativa, simples e composta;
- Região crítica e região de rejeição;
- Função poder;
  "O poder de um teste é a probabilidade de rejeitarmos H_0 caso ela seja falsa"
- Erro tipo I e tipo II;
  - Tipo I: Rejeitar erroneamente H_0;
  - Tipo II: Falhar em rejeitar H_0 quando ela é falsa.
- P-valor;
  "O p-valor pode ser interpretado como a probabilidade, sob H_0, de observarmos uma estatística tão ou mais extrema do que aquela observada"

Leitura recomendada

- De Groot seção 9.1;
- * De Groot seções 9.2 e 9.3.
- Próxima aula: De Groot, seção 9.1 (razão de verossimilhanças);
- **Exercícios recomendados**
  - De Groot.

    Seção 9.1: 3, 8 e 13.

    * Seção 9.1: 19 e 21.