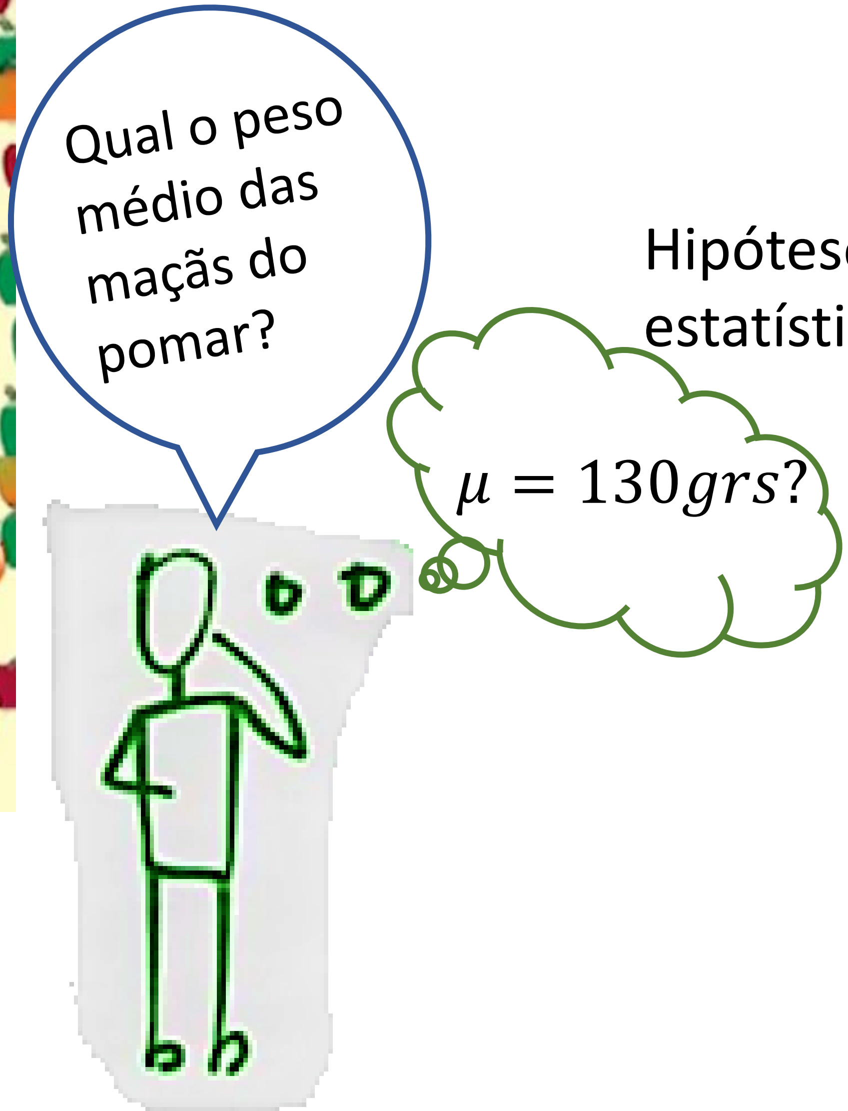
Ensaio de Hipóteses - Paramétricas

Questão

Hipótese estatística



Ensaios de hipóteses paramétricas são **sempre** ensaios sobre os valores dos **<u>parâmetros da população</u>** : média, variância, proporção, …

Ensaio de Hipóteses

8 passos de um Ensaio de Hipóteses:

1. Dados
2. Distribuição da população
3. Hipóteses
4. Escolha da estatística teste
5. Distribuição da estatística teste
6. Regra de decisão
7. Cálculo do valor da estatística teste\valor-p.
8. Teste mais potente – Lema de Neyman-Pearson
9. Decisão estatística

Ensaio de Hipóteses

1. Dados

População | **Amostra aleatória** | **Recolha de informação**

Maçãs de um pomar	10 maçãs	Peso das maçãs
Saltos em comprimento de um atleta	20 saltos	Comprimento dos saltos
População de um concelho	50 residentes no concelho	Prática desporto

Ensaio de Hipóteses

2. Distribuição que melhor representa o comportamento da população é conhecida

População	Questão de interesse	Distribuição
Maçãs de um pomar	Peso médio das maçãs do pomar	Normal Variância conhecida
Saltos em comprimento de um atleta	Variância do Comprimento dos saltos	Normal Variância desconhecida
População de um concelho	Proporção que pratica desporto	Bernoulli

Ensaio de Hipóteses

3. Hipóteses

Hipótese alternativa

São sempre sobre os valores dos parâmetros da população

NÃO SOBRE OS VALORES DAS AMOSTRAS OU DAS ESTATÍSTICAS!

Hipótese nula

É a hipótese que se vai testar
Status Quo\Ausência de Efeito

Ensaio de Hipóteses

3. Hipóteses

Hipótese nula

Hipótese alternativa

É a hipótese que se vai testar
Status Quo\Ausência de Efeito

$$\boldsymbol{\Theta}_0 \cup \boldsymbol{\Theta}_1 = \boldsymbol{\Theta} = \{\mu \in (-\infty, +\infty)\}$$

Questão de interesse

Peso médio das maçãs do pomar

$\boldsymbol{H_0}: \mu = 130\ grs$	$\boldsymbol{H_1}: \mu = 150\ grs$
$\boldsymbol{H_0}: \mu = 130\ grs$	$\boldsymbol{H_1}: \mu \neq 130\ grs$
$\boldsymbol{H_0}: \mu \leq 130\ grs$	$\boldsymbol{H_1}: \mu > 130\ grs$
$\boldsymbol{H_0}: \mu \geq 130\ grs$	$\boldsymbol{H_1}: \mu < 130\ grs$

Ensaio de Hipóteses

3. Hipóteses

Hipótese nula

Hipótese alternativa

Questão de interesse

Variação do comprimento do salto de um atleta

$$\boldsymbol{H_0}: \sigma^2 = 0.5\, mts \qquad \boldsymbol{H_1}: \sigma^2 = 0.7\, mts$$

$$\boldsymbol{\Theta}_0 \cup \boldsymbol{\Theta}_1 = \boldsymbol{\Theta} = \{\sigma^2 \in (0, +\infty)\}$$

$$\boldsymbol{H_0}: \sigma^2 = 0.5\, mts \qquad \boldsymbol{H_1}: \sigma^2 \neq 0.5\, mts$$

$$\boldsymbol{H_0}: \sigma^2 \leq 0.5\, mts \qquad \boldsymbol{H_1}: \mu > 0.5\, mts$$

$$\boldsymbol{H_0}: \mu \geq 0.5\, mts \qquad \boldsymbol{H_1}: \mu < 0.5\, mts$$

Ensaio de Hipóteses

3. Hipóteses

Hipótese nula

Hipótese alternativa

Questão de interesse

Prática de desporto pelos residentes de um concelho

$$\boldsymbol{H_0}: \theta = 0.35 \qquad \boldsymbol{H_1}: \theta = 0.4$$

$$\boldsymbol{\Theta}_0 \cup \boldsymbol{\Theta}_1 = \boldsymbol{\Theta} = \{\theta \in [0, 1]\}$$

$$\boldsymbol{H_0}: \theta = 0.35 \qquad \boldsymbol{H_1}: \theta \neq 0.35$$

$$\boldsymbol{H_0}: \theta \leq 0.35 \qquad \boldsymbol{H_1}: \theta > 0.35$$

$$\boldsymbol{H_0}: \theta \geq 0.35 \qquad \boldsymbol{H_1}: \theta < 0.35$$

Ensaio de Hipóteses

3. Hipóteses

Hipótese nula

Hipótese alternativa

Em resumo:

$$\boldsymbol{\Theta}_0 \cup \boldsymbol{\Theta}_1 = \boldsymbol{\Theta} \quad (*)$$

$$\boldsymbol{\Theta}_0 \cap \boldsymbol{\Theta}_1 = \emptyset$$

Muito importante: - a igualdade aparece sempre na H_0

- valores que o parâmetro pode assumir em H_1 são complementares dos valores de H_0 $(*)$

Ensaio de Hipóteses

4. Escolha da estatística teste

Questão de interesse	Informação	Estatística teste
Peso das maçãs do pomar	Peso de cada maçã da amostra	Peso médio na amostra - $\bar{X}$
Variação no comprimento dos saltos	Comprimento de cada salto na amostra	Variância do comprimento dos saltos na amostra - S'^2
Proporção que pratica desporto	Número de pessoas que pratica desporto na amostra	Proporção na amostra que pratica desporto- $\bar{X}$

Ensaio de Hipóteses

5. Distribuição da estatística teste

Estatística teste	Distribuição
Peso médio na amostra - $\bar{X}$	$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ ou $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
Variância do comprimento dos saltos na amostra - S'^2	$\frac{(n-1)S'^2}{\sigma^2} \sim \chi^2_{(n-1)}$
Proporção na amostra que pratica desporto- $\bar{X}$	$\frac{\bar{X}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \dot{\sim} N(0,1)$

Ensaio de Hipóteses

6. Regra de decisão: Rejeitar $\boldsymbol{H_0}$ ou não rejeitar $\boldsymbol{H_0}$

Decisão é tomada tendo por referência a Região de Rejeição

Região de Rejeição\Crítica - W é um subconjunto do espaço amostra X, tal que:

- se $(x_1, x_2, \cdots, x_n) \in W, rejeita-se\ \boldsymbol{H_0}$;

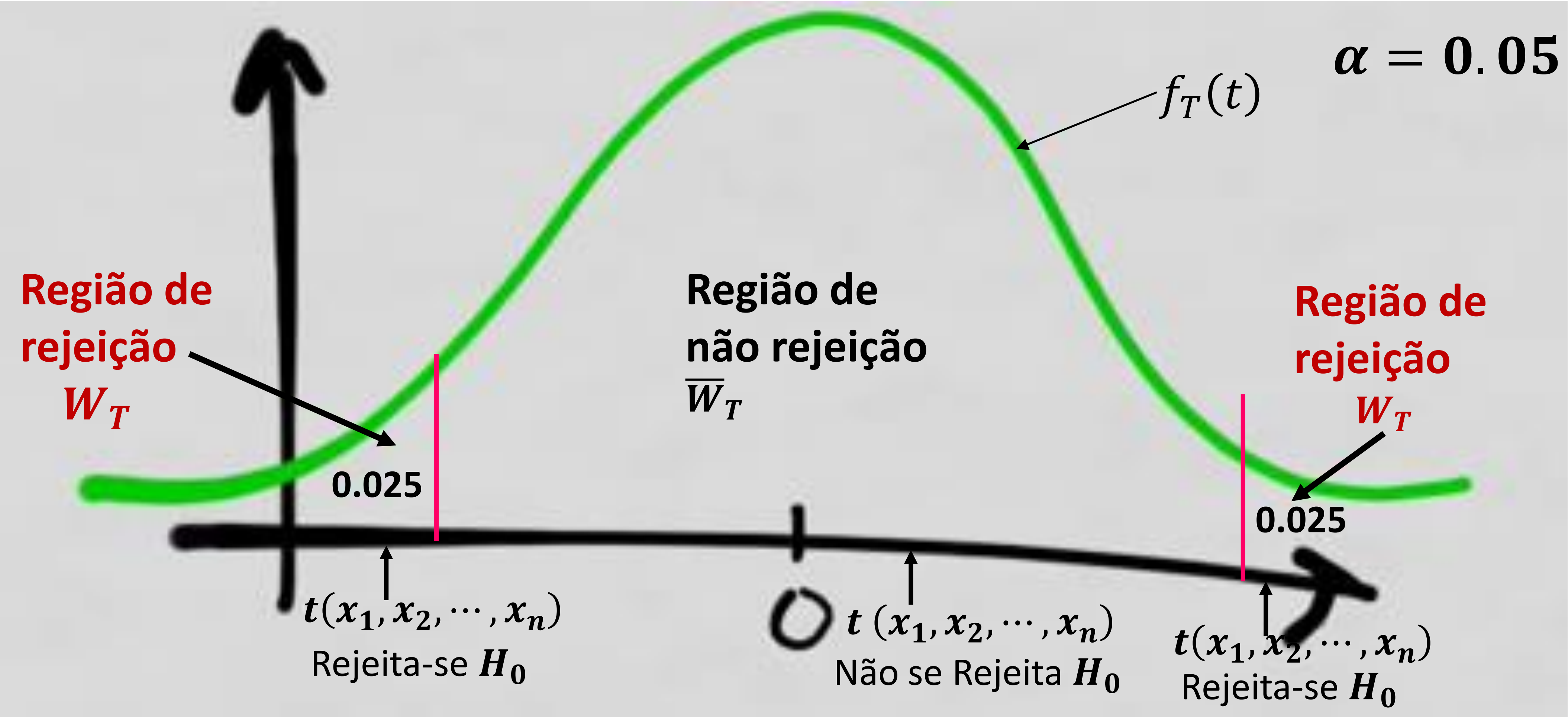
- se $(x_1, x_2, \cdots, x_n) \in \overline{W}, não\ se\ rejeita\ \boldsymbol{H_0}$

Notas: $W \cup \overline{W} = \mathbb{R}^n$ e $W \cap \overline{W} = \emptyset$

Atenção: como $W \subset \mathbb{R}^n$ pode ser complicado dizer se uma amostra particular pertence ou não a W. A utilização de uma estatística teste $\mathrm{T}(X_1, X_2, \cdots, X_n)$ e definição de uma região de rejeição W_T resolve esta dificuldade.

Ensaio de Hipóteses

Definido um **nível de significância α**, define-se a Região de Rejeição W_T com base na distribuição por amostragem da estatística teste T.



Ensaio de Hipóteses

Quando tomamos uma decisão podemos cometer dois tipos de erros:

Erro tipo 1: rejeitar $\boldsymbol{H_0}$ e $\boldsymbol{H_0}$ ser verdadeira

Erro tipo 2: não rejeitar $\boldsymbol{H_0}$ e $\boldsymbol{H_0}$ ser falsa

Situação / Decisão tomada	$\boldsymbol{H_0}$ verdadeira	$\boldsymbol{H_0}$ falsa
Rejeitar $\boldsymbol{H_0}$	**Erro de 1ª espécie - $\boldsymbol{\alpha}$**	**Decisão correcta - $\boldsymbol{\beta}$**
Não rejeitar $\boldsymbol{H_0}$ (Aceitar)	**Decisão correcta**	**Erro de 2ª espécie** $\boldsymbol{1-\beta}$

Ensaio de Hipóteses

Considere−se $\boldsymbol{H_0}: \theta = \theta_0$ contra $\boldsymbol{H_A}: \theta = \theta_1$

Dimensão do ensaio\erro 1ª espécie

$$P(rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ verdadeira) = P\big((X_1, X_2, \cdots, X_n) \in W_T|\theta = \theta_0\big) = \alpha$$

Erro 2ª espécie

$$P(não\ rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in \overline{W_T}|\theta = \theta_1\big) = 1 - \beta$$

$$P(rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in W_T|\theta = \theta_1\big) = \beta$$

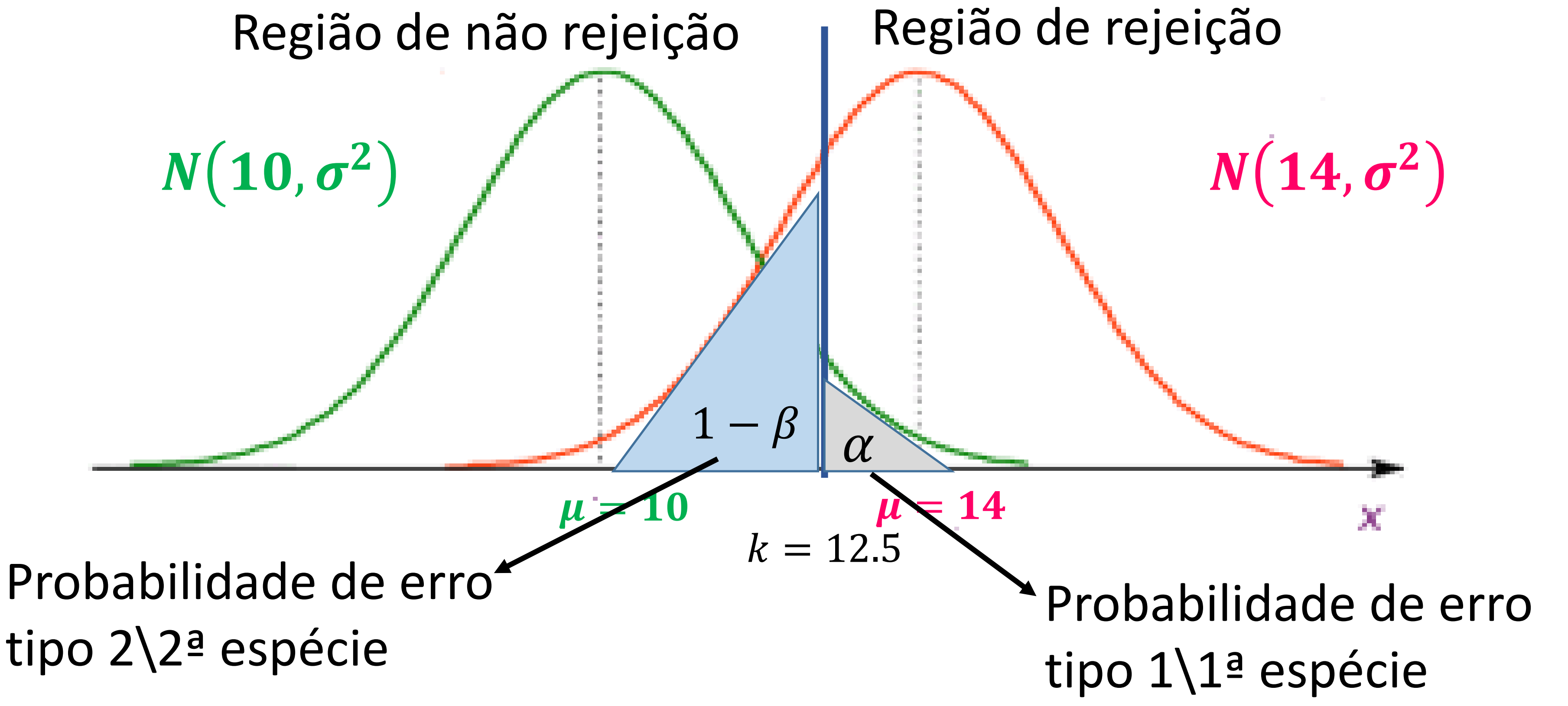
Potência do ensaio\teste

Ensaio de Hipóteses

Seja uma população $X \sim N(\mu, \sigma^2 = 4)$

Pretende-se testar $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu = 14$

$$W = \{\bar{x} : \bar{x} > k\}$$



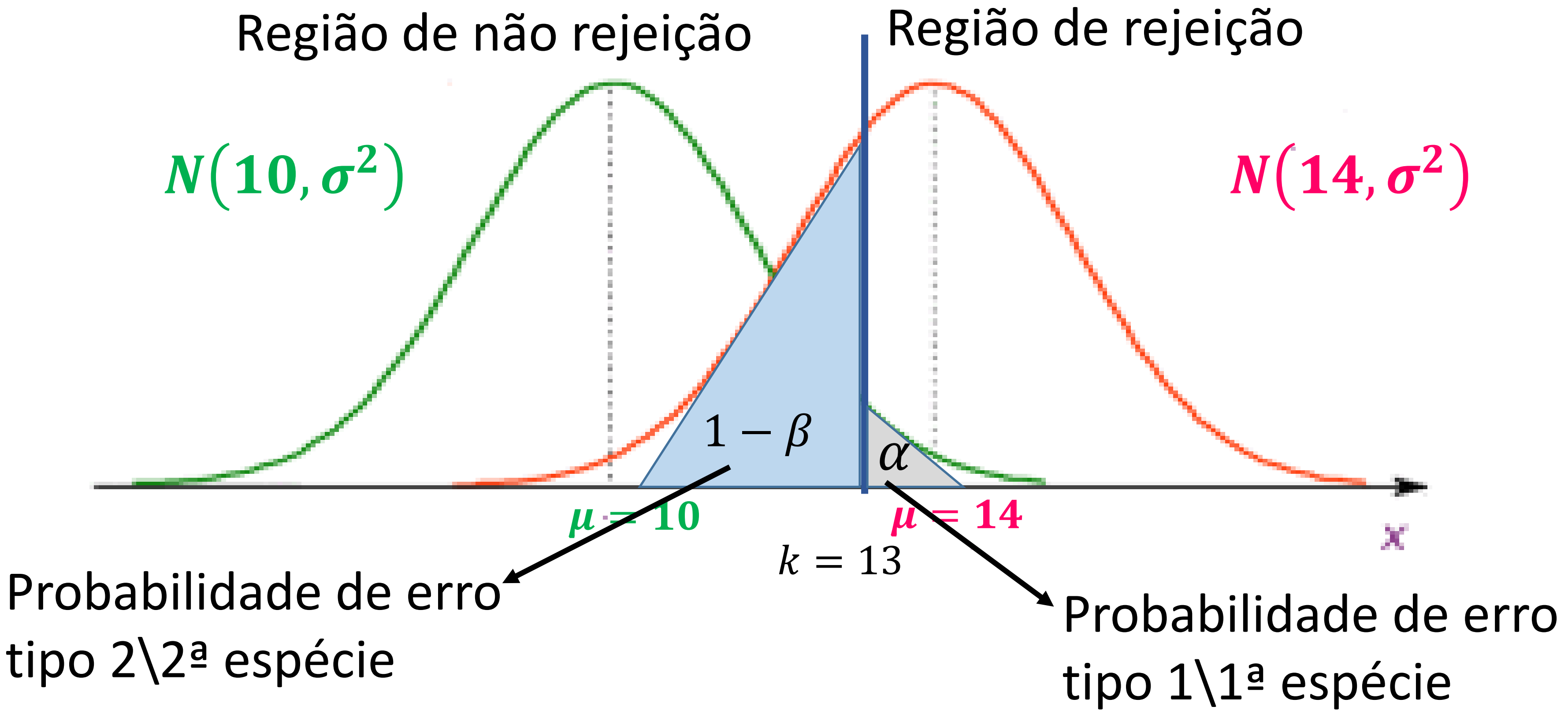
Ensaio de Hipóteses

Seja uma população $X \sim N(\mu, \sigma^2 = 4)$

Pretende-se testar $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu = 14$

$W = \{\bar{x} : \bar{x} > k\}$

Se se reduz α, aumenta $1 - \beta$



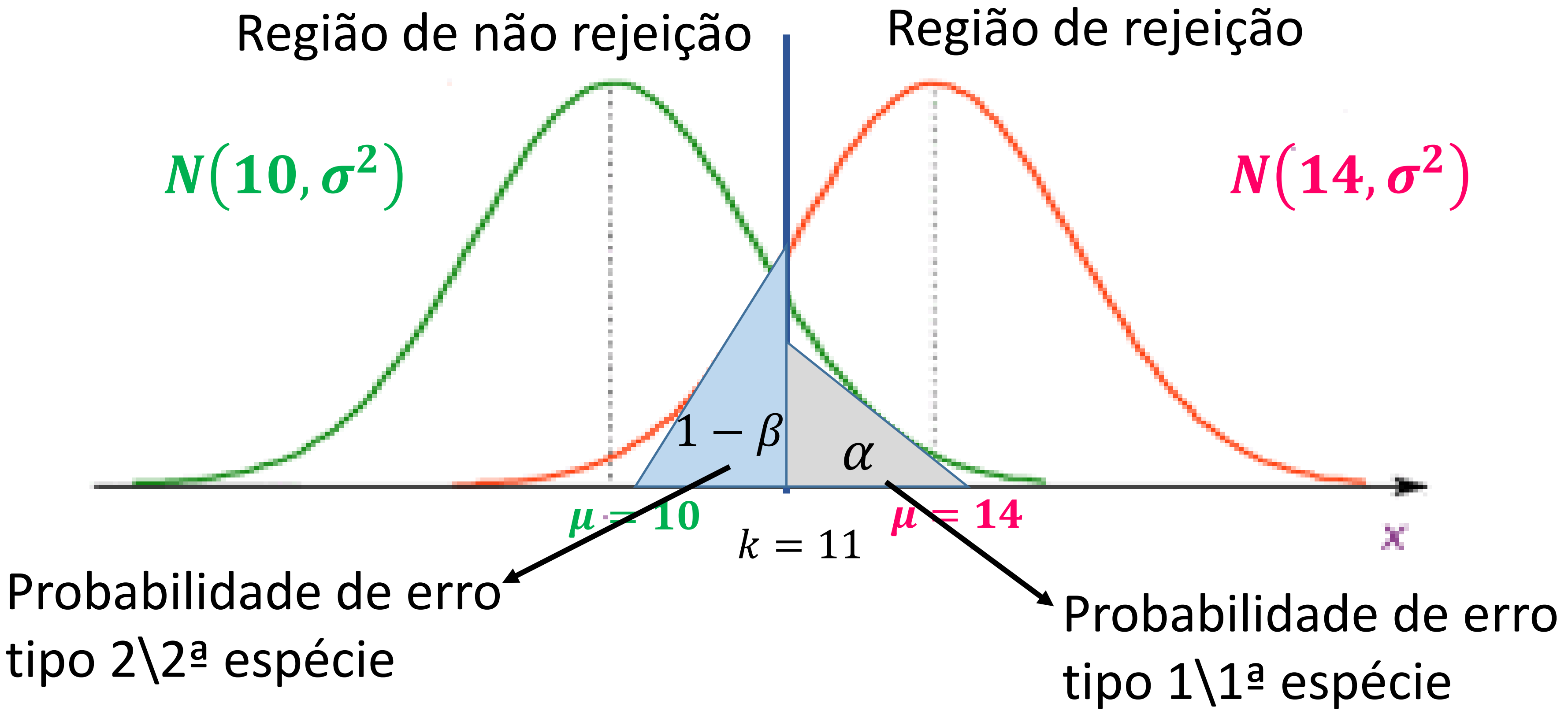
Ensaio de Hipóteses

Seja uma população $X \sim N(\mu, \sigma^2 = 4)$

Pretende−se testar $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu = 14$

$$W = \{\bar{x} : \bar{x} > k\}$$

Se se reduz $1 - \beta$, aumenta α



Ensaio de Hipóteses

8. Teste mais potente – Lema de Neyman-Pearson

Ideia importante: A redução das duas probabilidades (ou de uma delas, supondo a outra fixa) só se consegue aumentando a dimensão da amostra.

Na impossibilidade de minimizar ambos os erros, o Lema Neyman-Pearson define o **teste mais potente** com base em dois critérios:

1. Fixar a probabilidade de erro tipo 1 (α) porque atribui a este erro maior importância. Em geral $\alpha = 0{,}1, 0.05, 0.01$. Apenas se rejeita H_0 se houver forte evidência estatística contra esta hipótese.

2. Minimizar a probabilidade de erro tipo 2 $(1 - \beta)$ ou maximizar a potência do ensaio (β).

A região de rejeição assim definida corresponde ao teste mais potente de dimensão α.

Ensaio de Hipóteses

Regra intuitiva: No teste de médias, variâncias ou proporções a região de rejeição está do lado da alternativa quando se utiliza a estatística "natural".

Exemplos:

$\boldsymbol{H_0}: \mu \leq 130\ grs\ contra\ \boldsymbol{H_1}: \mu > 130 \Rightarrow W_T = \{(x_1, x_2, \cdots, x_n): \bar{x} > k\}$

$\boldsymbol{H_0}: \sigma^2 \geq 0.5\ mts$ contra $\boldsymbol{H_1}: \sigma^2 < 0.5\ mts \Rightarrow W_T = \left\{(x_1, x_2, \cdots, x_n):\ s'^2 < k\right\}$

$\boldsymbol{H_0}: \theta = 0.35$ contra $\boldsymbol{H_1}: \theta = 0.4 \Rightarrow W_T = \{(x_1, x_2, \cdots, x_n): \bar{x} > k\}$

Ensaio de Hipóteses

Exemplo: Seja uma população $X \sim N(\mu, \sigma^2 = 4)$ e uma amostra $n = 9$

Pretende−se testar $\boldsymbol{H_0}: \mu_0 = 10$ contra $\boldsymbol{H_1}: \mu_1 = 14$ $\quad\quad \alpha = 0.05$

$\mu_1 > \mu_0 \Rightarrow W_T = \{\bar{x} > k\}$ (Lema Neyman Pearson – regra intuitiva)

$$k: \mathrm{P}(\bar{X} > k | \mu = \mu_0) = \alpha \Leftrightarrow \mathrm{P}(\bar{X} > k | \mu = 10) = 0.05 \Leftrightarrow \mathrm{P}(\bar{X} \leq k | \mu = 10) = 0.95$$

$$\Rightarrow k = invnorm\left(0.95, 10, {}^{2}/_{3}\right) = 11.097$$

ou

$$k: \mathrm{P}\left(\frac{\bar{X} - \mu_0}{{}^{\sigma}/_{\sqrt{n}}} > \frac{k - 10}{{}^{2}/_{3}}\right) = 0.05 \Leftrightarrow \mathrm{P}\left(Z > \frac{k - 10}{{}^{2}/_{3}}\right) = 0.05 \Leftrightarrow \mathrm{P}\left(Z \leq \frac{k - 10}{{}^{2}/_{3}}\right) = 0.95$$

$$\Rightarrow \frac{k - 10}{{}^{2}/_{3}} = invnorm(0.95, 0, 1) = 1.645 \Rightarrow k = 10 + \frac{2}{3} * 1.645 = 11{,}1$$

Ensaio de Hipóteses

Probabilidade de erro de 2ª espécie:

$$1-\beta = P(não\ rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in W_T|\theta = \theta_1\big)$$

$$= P(\bar{X} \leq 11.1|\mu = 14) = \text{normal cdf}(-10000, 11.097, 14{,}2/3) \approx 0$$

Potência do ensaio: $\beta \approx 1 - 0 = 1$

$$\beta = P(rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in W_T|\theta = \theta_1\big)$$

$$= P(\bar{X} > 11.1|\mu = 14) = 1 - \text{normal cdf}(11.1, 100000{,}14{,}2/3) \approx 1$$

$$\beta = P(rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in W_T|\theta = \theta_1\big)$$

$$= P(\bar{X} > 11.1|\mu = 14) = \mathrm{P}\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} > \frac{11.1 - 14}{2/3}\right) = \mathrm{P}(Z < -4{,}35)$$

$$= \mathrm{P}(Z \geq 4{,}35) = 1$$

Ensaio de Hipóteses

Retome-se o exemplo anterior: população $X \sim N(\mu, \sigma^2 = 4), (X_1, X_2, \cdots, X_n)$

Pretende−se testar $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu > 10$ $\quad\alpha = 0.05$

$\mu_1 > \mu_0 \Rightarrow W = \{\bar{x} > k\}$ (Lema Neyman Pearson – regra intuitiva)

Cálculo de k faz-se exactamente do mesmo modo $\Rightarrow W_T = \{\bar{x} > k(n)\}$

Com $k(n) = 10 + 1.645 * \frac{2}{\sqrt{n}}$

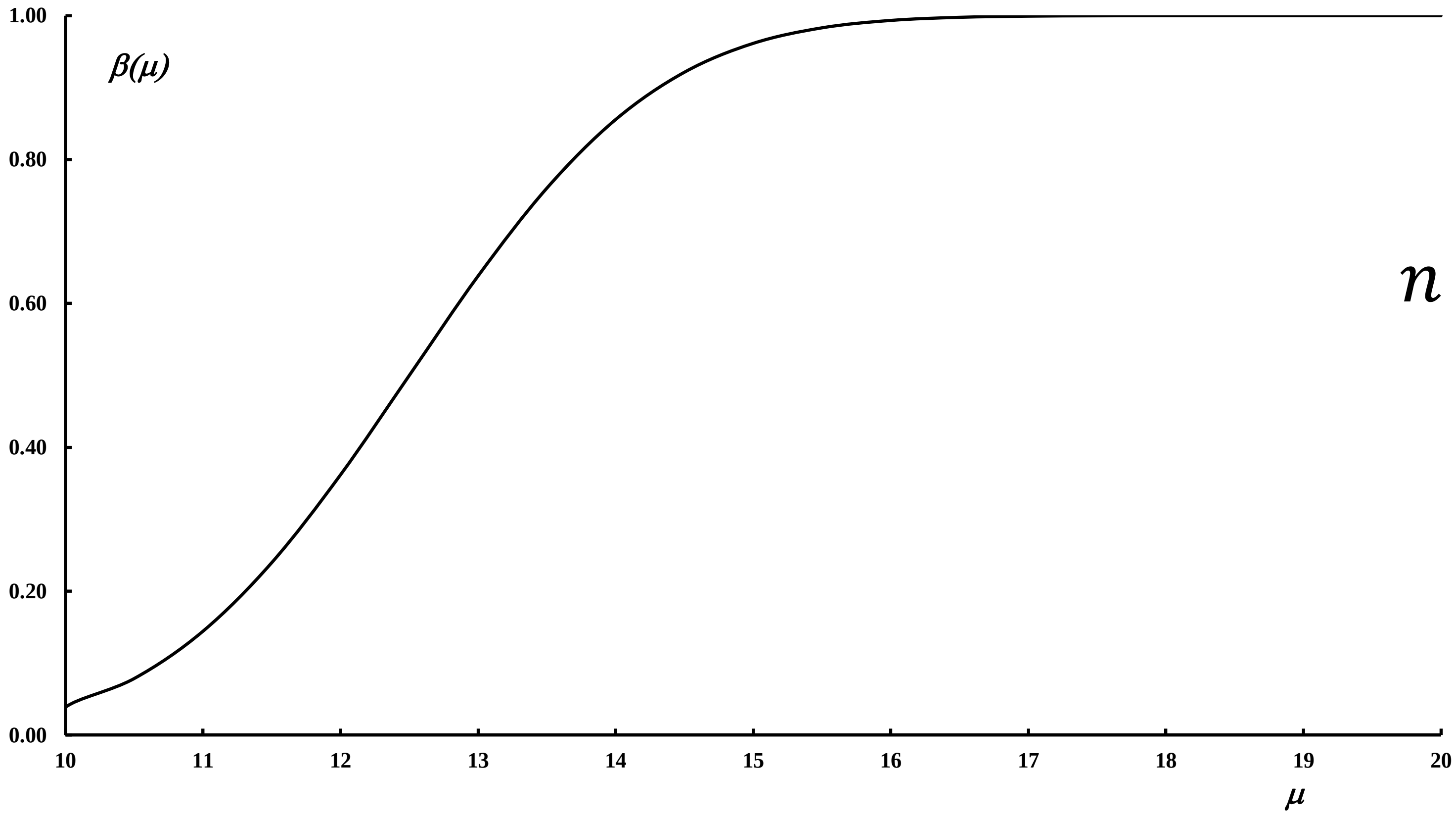
Potência do ensaio $-$ $\beta(\boldsymbol{\mu}) = P(rejeitar\boldsymbol{H_0}|\boldsymbol{H_0}\ falsa) = P\big((X_1, X_2, \cdots, X_n) \in W_T | \theta > \theta_0\big)$

$$= P(\bar{X} > k(n)|\mu > 10) \; = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{k(n) - \boldsymbol{\mu}}{2/\sqrt{n}} |\mu > 10\right)$$

$$= 1 - P\left(Z > \frac{k(n) - \boldsymbol{\mu}}{2/\sqrt{n}} |\mu > 10\right) = 1 - \boldsymbol{\Phi}\left(\frac{k(n) - \boldsymbol{\mu}}{2/\sqrt{n}} |\mu > 10\right)$$

Ensaio de Hipóteses

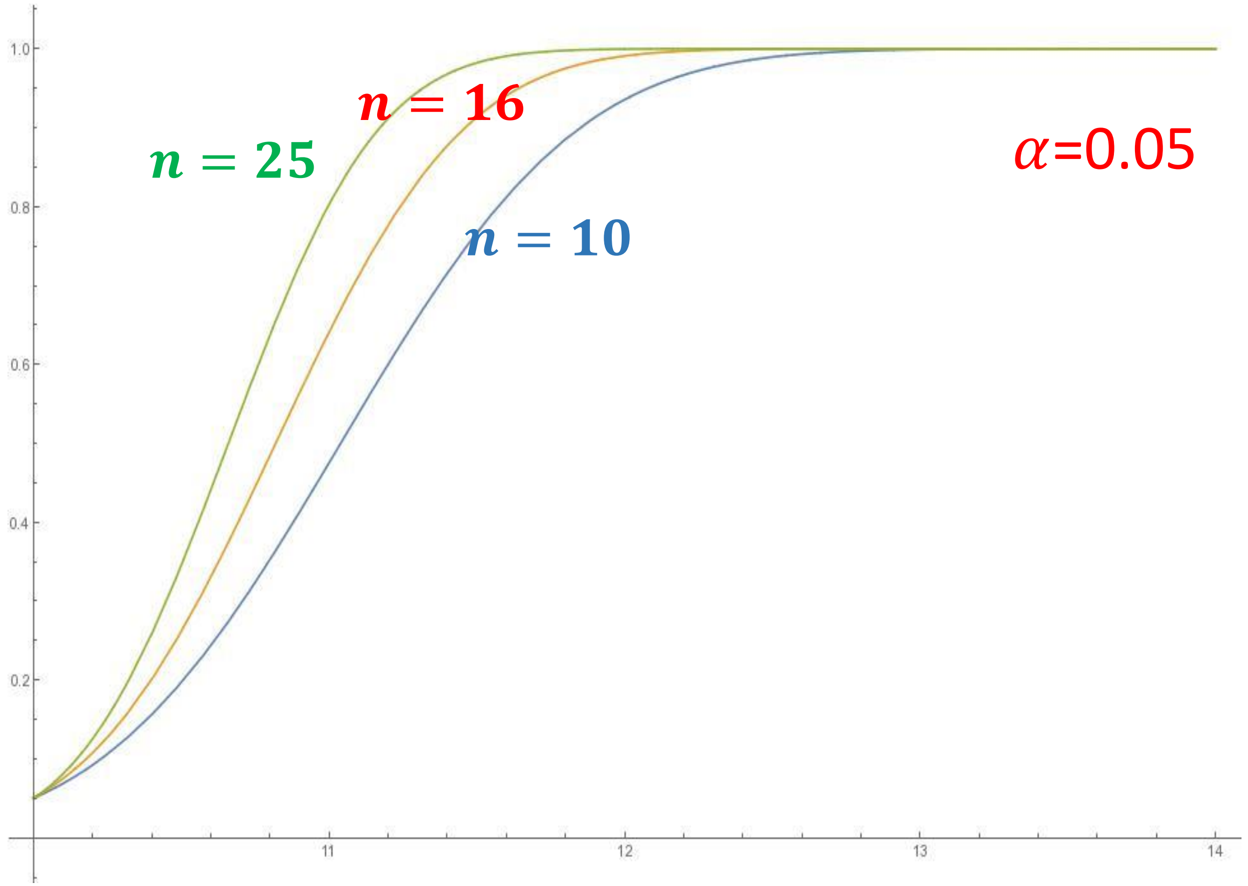
O gráfico que se segue ilustra a função potência $\beta(\mu)$ para $\mu > 10$



n fixo

Ensaio de Hipóteses

O que acontece à função potência quando, mantendo-se α constante, aumenta a dimensão da amostra?



Notas:

. Fixado α, $k(n)$ é função da dimensão da amostra n;

- Para cada $\mu > 10$, quanto maior a dimensão da amostra mais elevada é a potência do teste.

Ensaio de Hipóteses

Notas:

- Para testar $\boldsymbol{H_0}: \theta \leq \theta_0$ contra $\boldsymbol{H_1}: \theta > \theta_0$ procede-se como para testar $\boldsymbol{H_0}: \theta = \theta_0$ contra $\boldsymbol{H_1}: \theta > \theta_0$

- Para testar $\boldsymbol{H_0}: \theta \geq \theta_0$ contra $\boldsymbol{H_1}: \theta < \theta_0$ procede-se como para testar $\boldsymbol{H_0}: \theta = \theta_0$ contra $\boldsymbol{H_1}: \theta < \theta_0$

- Em ambos os casos está-se a escolher a pior situação.

Ensaio de Hipóteses

Testes bilaterais: $\boldsymbol{H_0}: \theta = \theta_0$ contra $\boldsymbol{H_1}: \theta \neq \theta_0$

Numa situação destas é fácil entender que **não existe, em geral, um teste UMP**.

Se se retomar o exemplo anterior $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu \neq 10$, tal corresponderia a ter simultaneamente 2 testes, quando:

- $\mu > 10$ ter-se-ia a região crítica $W_T = \{\bar{x} > k\}$
- $\mu < 10$ a região crítica $W_T = \{\bar{x} < k\}$.

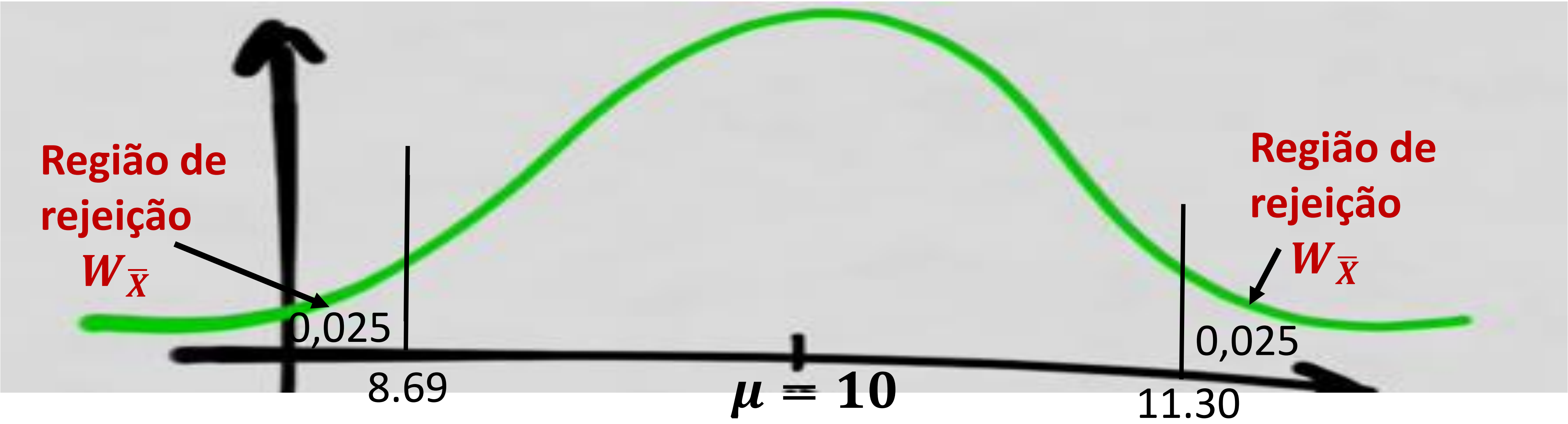
Isto não pode acontecer simultaneamente.

Para definir a região crítica recorre-se a uma regra intuitiva que consiste em considerar uma região de rejeição nas duas abas da distribuição da estatística teste - $W_T = \{\bar{x} < k_1 \cup \bar{x} > k_2\}$ com $k_1, k_2: P(\bar{x} < k_1) = P(\bar{x} > k_2) = \alpha/2$.

Ensaio de Hipóteses

Testes bilaterais: $\boldsymbol{H_0}: \theta = \theta_0$ contra $\boldsymbol{H_1}: \theta \neq \theta_0$

Retome-se o exemplo anterior $X \sim N(\mu, \sigma^2 = 4)$, $n = 9$ e consider-se o ensaio $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu \neq 10$ $\qquad \alpha = 0.05 \Leftrightarrow \frac{\alpha}{2} = 0.025$



Ensaio de Hipóteses

Valor-p

Num teste de hipóteses fixada a dimensão α, rejeita-se (ou não) $\boldsymbol{H_0}$ sem se ter em conta se a estatística teste T está longe ou perto do valor crítico k.

O **valor-p** é uma forma alternativa de reportar o resultado do teste que permite ultrapassar esta limitação.

Definição: Seja $T(x_1, x_2, \cdots, x_n) = t_{obs}$. O **valor-p** é a **probabilidade**, p_{obs}, **de observar um valor de** t_{obs} **tão ou mais desfavorável para** H_0, **admitindo** H_0 **verdadeira**

O **valor-p mede a evidência que os dados fornecem a favor de** H_0. Qto menor for o valor-p menor é a consistência dos dados com H_0, logo mais se rejeita H_0. Ex: se $p_{obs} = 0.0001$ rejeita-se, se $p_{obs} = 0{,}25$ não se rejeita. E se $p_{obs} = 0.052$?

Ensaio de Hipóteses

Cálculo do Valor-p

1. Obter a distribuição da estatística teste assumindo que $\boldsymbol{H_0}$ (ou o seu valor limite no caso de uma hipótese composta) é verdadeira.

2. Definir o acontecimento mais improvável do que o observado (+ para o lado da alternativa, depende de $\boldsymbol{H_1}$).

3. Calcular a sua probabilidade.

Exemplo: $X \sim N(\mu, \sigma^2 = 4),\ n = 16, T(x_1, x_2, \cdots, x_n) = t_{obs}$=11

1. $\boldsymbol{H_0}: \mu = 10\ (\geq 10)$ contra $\boldsymbol{H_1}: \mu < 10 \Rightarrow T = \bar{X} \sim N(10, {}^4/_{16})$

Os casos tão ou mais desfavoráveis para $\boldsymbol{H_0}$ correspondem a observar uma média amostral igual ou mais “afastada” de 10 do que o valor observado, 11.

$p_{obs} = P(\bar{X} \leq 11 | \mu = 10) = 0{,}98 \Rightarrow$ não se rejeita $\boldsymbol{H_0}$

Ensaio de Hipóteses

Cálculo do Valor-p

Exemplo: $X \sim N(\mu, \sigma^2 = 4),\ n = 16, T(x_1, x_2, \cdots, x_n) = t_{obs}$=11

2. $\boldsymbol{H_0}: \mu = 10\ (\leq 10)$ contra $\boldsymbol{H_1}: \mu > 10\ \Rightarrow T = \bar{X} \sim N(10, {}^4/_{16})$

Os casos tão ou mais desfavoráveis para $\boldsymbol{H_0}$ correspondem a observar uma média amostral igual ou mais "afastada" de 10 do que o valor observado, 11.

$$p_{obs} = P(\bar{X} \geq 11|\mu = 10) = 1 - P(\bar{X} < 11|\mu = 10) = 1 - 0{,}98 \approx 0.02$$

$\Rightarrow$ rejeita-se $\boldsymbol{H_0}$ para um nível de significância $\alpha = 0.05$

Ensaio de Hipóteses

Cálculo do Valor-p

Exemplo: $X \sim N(\mu, \sigma^2 = 4),\ n = 16, T(x_1, x_2, \cdots, x_n) = t_{obs}$=11

2. $\boldsymbol{H_0}: \mu = 10$ contra $\boldsymbol{H_1}: \mu \neq 10 \Rightarrow T = \bar{X} \sim N(10, {}^4\!/_9)$

Os casos tão ou mais desfavoráveis para $\boldsymbol{H_0}$ correspondem a observar uma média amostral mais "afastada" de 10 do que o valor observado, 11 quer para valores inferiores quer para valores superiores.

$$p_{obs} = P(|\bar{X}| \geq 11|\mu = 10) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq -2\right) + P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq 2\right) = 0.046$$

$\Rightarrow$ rejeita-se $\boldsymbol{H_0}$

Ensaio de Hipóteses

Populações normais $X \sim N(\mu, \sigma^2)$ – testes de médias

1. Testes de medias com variância σ^2conhecida

Estatística teste: $\bar{X} \sim N\left(\mu_0, {}^{\sigma^2}/_{n}\right)$ ou $\frac{\bar{X}-\mu_0}{\sigma/_{\sqrt{n}}} \sim N(0,1)$

Aplicar a regra intuitiva para o tipo de ensaio escolhido

1. Testes de medias com variância σ^2 desconhecida

Estatística teste: $\frac{\bar{X}-\mu_0}{S'/_{\sqrt{n}}} \sim t(n-1)$

Aplicar a regra intuitiva para o tipo de ensaio escolhido

Ensaio de Hipóteses

Populações Bernoulli $X \sim B(1,\theta)$ – teste à proporção amostral

Estatística teste: $\bar{X} \dot{\sim} N\left(\theta_0, {}^{\theta_0(1-\theta_0)}/_{n}\right)$ ou $\dfrac{\bar{X}-\theta_0}{\sqrt{\dfrac{\theta_0(1-\theta_0)}{n}}} \dot{\sim} N(0,1)$

Aplicar a regra intuitiva para o tipo de ensaio escolhido

Populações Poisson $X \sim Po(\lambda)$ – teste à média

Estatística teste: $\dfrac{\bar{X}-\lambda_0}{\sqrt{\dfrac{\lambda}{n}}} \dot{\sim} N(0,1)$

Aplicar a regra intuitiva para o tipo de ensaio escolhido