INFERÊNCIA ESTATÍSTICA (EST0035)

AULA05: TESTES DE HIPÓTESES (PARTE I)

Frederico Machado Almeida
frederico.almeida@unb.br

Departamento de Estatística
Instituto de Exatas
Universidade de Brasília (UnB)

Testes de Hipóteses

- A parte da estatística conhecida como Inferência Estatística é subdividida em duas subpartes principais: (*i*) Estimação e (*ii*) Testes de hipóteses.
- A primeira foi abordada de forma exaustiva nas aulas anteriores, e nessa aula vamos apresentar uma introdução da segunda parte dessa área de suma importância.
- Nosso principal objetivo nessa parte da disciplina é, desenvolver métodos gerais para testar hipóteses, e de seguida aplicá-los como ferramenta(s) para decidir sobre a veracidade ou não de determinadas afirmações que podem ser feitas sobre as características da população.

Testes de Hipóteses

- Muito mais do que obter estimadores pontuais (ou intervalares), em muitas situações práticas temos interesse em tomar alguma decisão sobre a rejeição ou não de determinadas afirmações, baseando-se em um conjunto de evidências.

  Exemplo 1.1: suponhamos que a duração média da vida de uma especial placa de microcomputador seja igual a 800 horas. Um novo processo de fabricação é proposto, e, deseja-se estabelecer um procedimento de decisão para julgar se o novo processo é melhor do que o processo atual. Em outras palavras, desejamos testar se a duração média das placas produzidas pelo novo processo é maior ou menor que a duração média atual. Por exemplo, formularemos a hipótese de que o novo processo não é melhor do que o atual. Em geral, esperamos que a hipótese seja rejeitada.

Exemplo 1.2: Considere o julgamento de uma pessoa acusada de ter cometido um delito. O processo consiste em apreciar os elementos fornecidos pela acusação e pela defesa e decidir em função deles e da lei. Mas, em princípio, a pessoa é inocente; e a acusação que tem a missão de apresentar provas em contrário. Se não houver evidencia (testemunhas, fatos, etc) nesse sentido, a pessoa continua sendo considerada não culpada.

Observação 1

Observe que, no contexto estatístico, tais evidências serão obtidas por meio de uma amostra aleatória, selecionada da população no qual o pesquisador pretende inferir.

Testes de Hipóteses

Etapas de um teste de hipóteses:

1. Estabelecer as hipóteses (nula e alternativa).
2. Especificar os erros do teste de hipóteses.
3. Fixar o nível de significância α.
4. Transformar a informação amostral em uma estatística de teste.
5. Decidir em rejeitar ou não a hipótese nula.
6. Calcular a magnitude dos possíveis erros a serem cometidos*.
7. Calcular a probabilidade de significância e concluir.

Importante!!

A formulação das hipótese exige um cuidado adicional. Em algumas situações, é `mais fácil` identificar, primeiramente a H_1 e então desenvolver a H_0. Em outras situações o `mais fácil` é desenvolver a H_0 e depois a H_1.

(I) Estabelecendo as Hipóteses

Definição 1 (Hipótese Estatística)

É qualquer afirmação (ou conjectura), a cerca da distribuição de probabilidades de uma ou mais variáveis aleatórias.

- *Uma hipótese estatística pode ser igualmente definida como sendo qualquer afirmação (ou conjetura) sobre determinada característica (parâmetros) da população.*

Em termos de notação, vamos usar H para denotar uma determinada hipótese estatística. Podendo ser: (i) hipótese nula (notação: H_0), que representa ausência de diferença, efeito ou chance e, (ii) hipótese alternativa (notação: H_1 ou H_a), que é aquela que complementa a H_0.

Testes de Hipóteses

- Por foma a fazer uma introdução prática da noção de hipóteses estatísticas, assuma que o interesse central esteja na variável aleatória $X \sim f(x|\theta)$, com $\theta \in \Theta$. Portanto, as hipóteses de interesse podem ser formuladas da seguinte forma:

$$H_0 : \theta \in \Theta_0 \text{ contra } H_1 : \theta \in \Theta_1, \tag{1}$$

com Θ_0 e Θ_1 sendo subconjuntos de Θ, tal que, $\Theta_0 \cup \Theta_1 = \Theta$ e $\Theta_0 \cap \Theta_1 = \emptyset$.

Testes de Hipóteses

- A nomenclatura usada para designar as hipóteses (em nula, H_0), faz referência à ausência de diferença, efeito, chances, etc.

  Exemplo 1.3: Supondo que θ denota a alteração média na pressão sanguínea de um pacientes depois de receber uma medicação. Um experimentador pode estar interessado em testar as seguintes afirmações,

  $$H_0 : \theta = 0 \text{ (ausência de efeito) vs } H_1 : \theta \neq 0 \text{ (presença de efeito).}$$

- A hipótese nula declara que, em média, a medicação não tem efeito na pressão sanguínea. E a hipótese alternativa declara que existe algum efeito (*positivo ou negativo*).

Testes de Hipóteses

Exemplo 1.4: Um segundo exemplo sobre a origem da nomenclaturas das hipóteses H_0 e H_1 é o seguinte: um consumidor pode estar interessado na proporção de itens com defeito que foram produzidos por um fornecedor. Se θ denota a proporção de itens com defeito, talvez o consumidor queira testar as hipóteses, $H_0 : \theta \geq \theta_0$ versus $H_1 : \theta < \theta_0$, onde o valor θ_0 denota a proporção máxima aceitável de itens com defeito. Nesse caso, H_0 declara que a proporção de itens com defeito é inaceitavelmente alta.

- Voltando para o *Exemplo 1.2*, podemos formular as hipóteses de interesse como, H_0 : O réu é inocente contra H_1 : O réu é culpado. Observe que nesse caso, a acusação se concentra em juntar evidência que possam provar que de fato o réu é culpado.

Definição 2 (Classificação das Hipóteses)

Uma hipótese Estatística é dita ser simples, se ela específica complemente a distribuição da variável aleatória, X. Isto é, se $\Theta = \{\theta_0\}$. Nesse caso, a H_0 em (1) pode ser formulada como $H_0 : \theta = \theta_0$. Caso contrário, a H_0 é dita ser composta.

Testes de Hipóteses

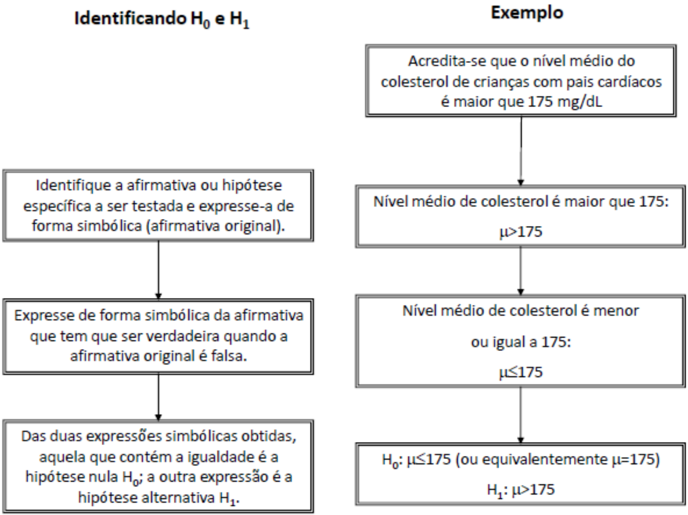

Figura: Estabelecendo as hipóteses (nula e alternativa).

Observação 2

É importante ser um pouco cauteloso na hora de classificar as hipóteses em simples e compostas. Pois, tal classificação deve ser feita analisando a dimensão e como os parâmetros de Θ_0 *foram definidos. Por exemplo, se assumirmos que* $X_1, X_2, \cdots, X_n$ *são cópias* $iid's$ *de* $X \sim N(\mu, \sigma^2)$, *com* $\Theta = \{(\mu, \sigma^2) : \ \mu \in \mathbb{R} \text{ e } \sigma^2\}$ *sendo ambos desconhecidos, considere a seguinte hipótese* $H_0 : \mu = 0$, *nesse caso, temos que,* $\Theta_0 = \{(\mu, \sigma^2) : \ \mu = 0 \text{ e } \sigma^2\}$. *Portanto, como* $0 < \sigma^2 < \infty$ *é desconhecido, segue que a hipótese anterior é composta*

Testes de Hipóteses

- A realização de um bom ensaio de hipóteses parte de uma correta formulação das hipóteses, a qual se obtêm pela análise do problema proposto, a maior parte das vezes através de elementos não estatísticos. De facto, é na natureza da questão que se deve encontrar o modo de formular as hipóteses.
- A recolha dos dados da amostra aleatória e determinante: são os erros de amostragem que impedem que a amostra represente corretamente a população. Mas, na prática, a formulação das hipóteses deve ser anterior à recolha da amostra, para que o procedimento não seja viesado.
- E a informação da amostra que vai ser confrontada com os critérios entretanto estabelecidos para decidir da rejeição ou não da hipótese nula.

Testes de Hipóteses

Definição 3

Chamamos de teste de uma hipótese estatística, a função de decisão, $\phi : \mathcal{X} \to \{0, 1\}$, em que $\phi(\mathbf{x}) = 0$ corresponde a ação de considerar que a H_0 é verdadeira, e $\phi(\mathbf{x}) = 1$ a ação de considerar que a H_1 é verdadeira, com $\mathcal{X}$ denotando o espaço amostral, isto é,
$\mathcal{X} = \{(X_1, X_2, \cdots, X_n) \in \mathbb{R}^n\}$.

- Associado com a Definição 3, temos que, a função de decisão $\phi(\cdot)$ divide o espaço amostral $\mathcal{X}$ em dois subconjuntos,

$$\begin{aligned} RC &= \{(x_1, x_2, \cdots, x_n) \in \mathcal{X} : \ \phi(x_1, x_2, \cdots, x_n) = 1\} \\ RA &= \{(x_1, x_2, \cdots, x_n) \in \mathcal{X} : \ \phi(x_1, x_2, \cdots, x_n) = 0\}, \end{aligned}$$

  onde $RC \cup RA = \mathcal{X}$ e $RC \cap RA = \emptyset$, para algum valor $0 < \alpha < 1$.
- Os subconjuntos RC e RA de $\mathcal{X}$ são igualmente conhecidos como regiões crítica e de aceitação, respectivamente.

Testes de Hipóteses

- Como regra de decisão associada às hipóteses formuladas em (1), temos que, se $t(\mathbf{x})$ é uma quantidade observada da estatística $T(\mathbf{X})$, com $\mathbf{X} = (X_1, X_2, \cdots, X_n)^\top$, então, podemos:

$$\begin{aligned} \text{Rejeitar } H_0 &\text{ se } t(\mathbf{x}) \in RC \\ \text{Não-Rejeitar } H_0 &\text{ se } t(\mathbf{x}) \in RA. \end{aligned}$$

- Tal como nos tópicos anteriores, o intuito da Inferência Estatística é trabalhar com informação amostral, para caraterizar as quantidades da população.
- Portanto, tal como no tópico de estimação, qualquer processo que consiste em usar a informação amostral para refutar ou não determinada afirmação estará sempre a algum tipo de erro.

(II) Especificando os Possíveis Erros

- Uma vez mais, vamos considerar as hipóteses formuladas na expressão (1). Tal que, $H_0 : \theta \in \Theta_0$ contra $H_1 : \theta \in \Theta_1$.
- Portanto, a decisão de rejeitar a H_0 sendo ela verdadeira, significa que estamos cometendo um erro do tipo I. De forma similar, ao rejeitar uma H_1 que é verdadeira, estaremos cometendo um erro do tipo II.
- Um resumo dos dois tipos de erros que podem ocorrer em um teste estatístico de hipóteses são apresentados na Tabela 1.

Decisão	H_0 Verdadeira	H_0 Falsa
Rejeitar H_0	Erro tipo I	**Decisão Correta**
Não Rejeitar H_0	**Decisão Correta**	Erro tipo II

Tabela: Tipos de erros em um teste estatístico.

Testes de Hipóteses

- De forma resumida, os dois tipos de erros encontrados na Tabela acima são tais que,

Erro tipo I	–	ocorre quando se decide rejeitar H_0, sendo ela verdadeira (rejeitar uma hipótese verdadeira).
Erro tipo II	–	ocorre quando se decide em não rejeitar H_0, sendo ela falsa (não rejeitar uma H_0 quando tal) deveria ter sido rejeitada.

Testes de Hipóteses

- Entre os dois tipos de erro apresentados anteriormente, o mais grave é o erro tipo I. Ou seja, podemos concluir que, para o exemplo do julgamento, pode então verificar-se que qualquer decisão tomada pode ser acertada ou errada. Nesse caso, cometer um erro do tipo I, significa condenar um acusado que é inocente. Ou seja, considera-se ser grave privar um indivíduo de liberdade, sendo inocente.
- Ademais, o erro do tipo II seria, atribuir liberdade um indivíduo que é culpado. Ou seja, não condenar um indivíduo culpado.

NOTA: Apesar desses possíveis erros, não é por isso que os tribunais devem ser abolidos! O que deve haver é a preocupação de recolher uma informação mais pertinente e correta de modo a minimizar o risco, isto é, a probabilidade de errar.

Testes de Hipóteses

Erro tipo I: o erro do tipo I (notação: $\alpha(\theta)$), representa a probabilidade de $t(\mathbf{x}) \in RC$ quando ela H_0 é VERDADEIRA. Isto é,

$$\text{Erro tipo I} = \mathbb{P}\left(RH_0 | H_0 \text{ Verdadeira}\right) = \mathbb{P}\left(T(\mathbf{X}) \in RC | \theta \in \Theta_0\right).$$

Erro tipo II: o erro do tipo II (notação: $\beta(\theta)$), representa a probabilidade de $t(\mathbf{x}) \in RA$ quando H_0 ela é FALSA. Isto é,

$$\text{Erro tipo II} = \mathbb{P}\left(NRH_0 | H_0 \text{ Falsa}\right) = \mathbb{P}\left(T(\mathbf{X}) \in RA | \theta \in \Theta_1\right).$$

- Assim, cientes da ocorrência desses tipos de erros, o intuito é selecionar uma região crítica RC, entre várias possibilidades, que minimiza a ocorrência de tais erros.

- Observe que, em cursos elementares de inferência estatística, uma estimativa $t(\mathbf{x})$ deve ser obtida, e de seguida a região crítica RC deve ser construída. Portanto, a H_0 será rejeitada à favor da H_1 se $t(\mathbf{x}) \in RC$, ou simplesmente,

$$\text{Valor-p} = \mathbb{P}_{H_0}\left(T(\mathbf{X}) \in RC\right) \leq \alpha. \tag{2}$$

  Sendo o lado esquerdo da equação (2) denotando a probabilidade de significância (ou valor-p).
- Assim, uma outra forma de decidir sobre a rejeição ou não da H_0, seria comparar o valor-p com o nível de significância, α.

(III) Estabelecendo o Nível de Significância

Definição 4

Dizemos que uma região crítica RC é de tamanho α (área da região crítica), se

$$\alpha = \sup_{\theta \in \Theta_0} \mathbb{P}\left(T\left(\mathbf{X}\right) \in RC\right) = max\, \mathbb{P}\left(T\left(\mathbf{X}\right) \in RC | \theta \in \Theta_0\right).$$

Observe que, a Definição 4, afirma que, o nível de significância de um teste de hipóteses, é nada mais do que o maior valor da probabilidade do *erro tipo I*.

(IV) Transformando os Dados Amostrais em Estatística de Teste

O processo que consiste em transformar a informação amostral em estatística de teste, requer que:

- Saibamos à priori, qual é a forma da distribuição de probabilidade, onde a amostra aleatória, $X_1, X_2, \cdots, X_n$ foi extraída.
- Conforme apresentado nas aulas anteriores, se $X_1, X_2, \cdots, X_n$ denotam cópias iid's de $X \sim \mathcal{N}(\mu, \sigma^2)$, então, a distribuição da estatística de teste (equivalente a quantidade pivotal), terá uma certa distribuição, a depender do parâmetro que está sendo estimado, e do conhecimento prévio ou não do μ e/ou σ^2.

Testes de Hipóteses

Exemplo 2: Uma pizzaria recebe diariamente encomendas por telefone, que se tem comportado segundo uma lei normal. A empresa está dimensionada para uma procura media diária que não ultrapasse as 200 pizzas, admitindo um desvio-padrão de 15.
Uma campanha promocional realizada nos últimos 9 dias levou a uma procura media de 210 pizzas. O problema consiste em avaliar a necessidade de reforçar a capacidade media de venda, estudando se houve de facto uma alteração significativa na procura diária de pizzas. Pede-se para:

a. Estabelecer as hipóteses de interesse.
b. Obter a estatística de teste e concluir, ao nível de significância de 5%.
c. Construa as regiões crítica e de aceitação.
d. Obtenha a magnitude dos erros tipo I e II, respectivamente. Interprete-os.

Exemplo 3: Considere o caso em que $X_1, X_2, \cdots, X_n$ são cópias iid's de $X \sim \mathcal{N}(\mu, 1)$, e suponha que nosso interesse esteja em testar as seguintes hipóteses: $H_0 : \mu \leq 0$ vs $H_1 : \mu > 0$, onde $\Theta_0 = \{\mu : \mu \leq 0\}$ e $\Theta_1 = \{\mu : \mu > 0\}$. Como $\sigma^2 = 1$ (variância conhecida), segue imediatamente que,

$$z_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{\sigma} \sim \mathcal{N}(0, 1)\,.$$

Portanto, com base nas hipóteses anteriores, RH_0 quando $z_{obs} > z_\alpha$, implicando em, $\sqrt{n}\bar{X}_n > z_\alpha \Longrightarrow \bar{X}_n > z_\alpha/\sqrt{n}$. Ou seja, $RC = \{(x_1, x_2, \cdots, x_n) : \bar{x}_{obs} > z_\alpha/\sqrt{n}\}$, e
$RC = \{(x_1, x_2, \cdots, x_n) : \bar{x}_{obs} \leq z_\alpha/\sqrt{n}\}$. Com $\bar{x}_{obs} = \sum_{i=1}^{n} x_i/n$.

(V) Regra de Decisão

- Com base no Exemplo anterior, podemos definir a seguinte função de decisão, ϕ tal que,

$$\phi(x) = \begin{cases} 1, & \text{se} \quad t(\mathbf{x}) \in RC, \\ 0, & \text{se} \quad t(\mathbf{x}) \in RA. \end{cases}$$

Com $t(\mathbf{x}) = \bar{x}_{obs}$ e $\phi(\mathbf{x})$ denotando a função de decisão, tal que, $\phi(x) : \mathcal{X} \rightarrow \{0, 1\}$.

(V) Regra de Decisão

A decisão em rejeitar ou não uma determinada hipótese, pode ser tomada de quatro formas (ou usando quatro métodos). A saber:

1. Comparando o valor da estatística de teste, com o correspondente valor crítico (aquele obtido em alguma tabela, a depender da distribuição da estatística de teste).

2. Usando as regiões crítica e de aceitação. Nesse caso, a regra de decisão segue a ideia apresentada anteriormente.

3. Por meio da probabilidade de significância (ou valor-p). Aqui, a H_0 será rejeitada à favor da H_1, se o valor-p for menor que o nível de significância.

4. Usando a relação entre os testes de hipóteses, e os intervalos de confiança.

(VI) Função Poder do Teste

- De forma simplificada, a Definição 4 afirma que, o nível de significância α, é dada pela maior probabilidade do *erro tipo I*, obtido assumindo diferentes valores de $\theta \in \Theta_0$.
- No entanto, como nosso interesse é de maximizar a probabilidade de tomar decisões corretas (vide na Tabela 1), queremos então que a probabilidade, $1 - \beta(\theta)$ seja maior possível. Assim,

$$\begin{aligned}
1 - \beta(\theta) &= 1 - \mathbb{P}(\text{Erro tipo II}) \\
&= 1 - \mathbb{P}(\text{Rejeitar } H_1 | \theta \in \Theta_1) \\
&= \mathbb{P}(\text{Não Rejeitar } H_1 | \theta \in \Theta_1) \\
&= \text{DECISÃO CORRETA!}
\end{aligned}$$

Definição 5 (Função Poder)

Seja τ um teste qualquer de uma hipótese nula $H_0 : \theta \in \Theta_0$ contra a hipótese alternativa, $H_1 : \theta \in \Theta_1$. A função $\pi(\theta)$ do teste τ, denotado por $\pi_\tau(\theta)$ é definida como sendo a capacidade do teste rejeitar uma H_0 sendo ela FALSA. Ou não rejeitar H_0, sendo ela VERDADEIRA.

- Do resultado anterior, segue que,

$$\pi_\tau(\theta) = \mathbb{P}(\text{Não Rejeitar } H_1 | \theta \in \Theta_1) = 1 - \beta(\theta). \tag{3}$$

- A expressão (3) mostra que, quando menor for a probabilidade do erro tipo II, maior é a probabilidade de tomar uma decisão correta (função poder).

Testes de Hipóteses

Observação 3

Da mesma forma que o Erro Quadrático Médio desempenha um papel importante na parte de estimação, a função poder, $\pi_\tau(\theta)$ funciona como um padrão ouro, para comparar (ou avaliar) a qualidade dos testes de hipóteses. Ou seja, a ideia da função poder é não rejeitar a H_0, caso ela seja verdadeira.

Exemplo 4: Seja $X_1, X_2, \cdots, X_n \overset{iid}{\sim} \mathcal{N}\left(\mu, \sigma^2 = 25\right)$. Considere a possibilidade de testar as seguintes hipóteses, $H_0 : \mu \leq 17$ contra $H_1 : \mu > 17$. Supondo $\alpha = 0,05$ e $n = 25$, obtenha a função pode, $\pi_\tau(\mu)$, e apresente o respectivo gráfico.

Testes de Hipóteses

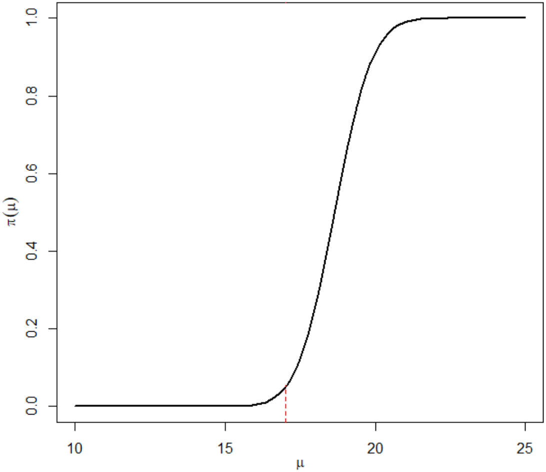

Figura: Função pode para o problema do Exemplo 4.

Testes de Hipóteses

- O gráfico apresentado anteriormente mostra que, para valores de $\mu \geq 20$, o teste quase certamente rejeita a H_0 à favor da H_1 toda vez que H_0 é falsa. Pois, $\pi_\tau(\mu) = 1 - \beta(\mu) = 1 \iff \beta(\mu) = 0$.
- De forma similar, para $\mu < 16$, temos que, $\pi_\theta(\mu) = 0 \iff \beta(\mu) = 1$. O que significa dizer que o teste não rejeita a H_0 como deveria ter rejeitado.
- Por fim, para valores de $16 \leq \mu < 20$, o teste tem uma certa chance de rejeitar e não-rejeitar a H_0.
- Segue da Definição 4 que,

$$\begin{aligned} \alpha &= \sup_{\mu \leq 17} \pi_\tau(\mu) = \sup_{\mu \leq 17} \mathbb{P}_\mu\left(T(\mathbf{X}) \in RC\right) \\ &= \sup_{\mu \leq 17} \mathbb{P}_\mu\left(\bar{X}_n > 17 + \frac{5z_\alpha}{\sqrt{n}}\right) = \sup_{\mu \leq 17} \Phi\left(\mu^* - 18,65\right), \end{aligned}$$

com μ^* um valor pertencente ao espaço Θ_0 (resultado em $\alpha \approx 0,05$).

Observação 4

Conforme a definição da função poder (apresentada anteriormente), é possível concluir que, dada duas regiões de rejeição da H_0 *(regiões críticas), com tamanhos* $\alpha_1 = \alpha_2 = \alpha$*, digamos,* C_{α_1} *e* C_{α_2}*, dir-se-á que a região* C_{α_1} *é melhor que a* C_{α_2}*, se para todo* $\theta \in \Theta_1$*,* $\pi_{C_{\alpha_1}}(\theta) > \pi_{C_{\alpha_2}}(\theta)$*. Ou seja, se para o mesmo tamanho de teste* α*, a região apresentar maior probabilidade de acertos (função poder).*

Testes de Hipóteses para μ

- A construção de um teste de hipóteses para a média μ, requer as mesmas suposições apresentadas no tópico de estimação por intervalos: se a população onde a *va* X é proveniente de uma distribuição normal ou não.
- Em linhas gerais, a abordagem é idêntica com aquela aquela apresentada em cursos introdutórios de Inferência Estatística.

Definição 6 (Modelo Normal)

Assuma que $X_1, X_2, \cdots, X_n$ *denota, cópias iid's de* X, *tal que,* $X \sim \mathcal{N}(\mu, \sigma^2)$, *com* $\mu \in \mathbb{R}$ *sendo desconhecido. Considere as seguintes hipóteses de interesse,* $H_0 : \mu = \mu_0$ *contra* $H_1 : \mu > \mu_0$. *Neste caso,* $\Theta_0 = \{(\mu, \sigma^2) : \mu = \mu_0, \sigma^2 > 0\}$, *e* $\Theta_1 = \{(\mu, \sigma^2) : \mu > \mu_0, \sigma^2 > 0\}$. *Tal como no tópico de estimação, podemos considerar os seguintes casos:*

Testes de Hipóteses para μ

CASO I (Hipóteses do Tipo) $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$.

(i) Supondo σ^2 Conhecido: para esse caso, se α denota o tamanho do teste (ou nível de significância), a H_0 será rejeitada à favor da H_1 se $z_{obs} > z_\alpha$, onde

$$z_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{\sigma} \sim \mathcal{N}\left(0, 1\right).$$

Sendo $\bar{X}_n = \sum_{i=1}^{n} X_i/n$ a média amostral. Portanto, a região crítica é dada por, $\bar{X}_n > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$. Ou seja, $RC = \{\mathbf{x} : \bar{x}_{obs} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\}$. Desta forma, o tamanho do teste é tal que, $\alpha = \mathbb{P}\left(Z > z_\alpha\right)$, e valor-p dado por: valor-p$=\mathbb{P}\left(Z > z_{obs}\right)$.

(ii) Supondo σ^2 Desconhecido: usando a mesma informação de (i), e porque σ^2 é desconhecido, vamos considerar seu melhor estimador pontual, dado por, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$. Desta forma, para um tamanho de teste, α, a H_0 será rejeitada à favor da H_1 se $t_{obs} > t_{(n-1,\alpha)}$, onde

$$t_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{S} \sim t_{n-1}.$$

Portanto, a região crítica é dada por, $\bar{X}_n > \mu_0 + t_{(n-1,\alpha)}\frac{S}{\sqrt{n}}$. Ou seja, depois que a amostra for observada, $RC = \{\mathbf{x} : \bar{x}_{obs} > \mu_0 + t_{(n-1,\alpha)}\frac{s}{\sqrt{n}}\}$.

- Sendo que, o tamanho do teste será tal que, $\alpha = \mathbb{P}_{H_0}\left(T_{n-1} > t_{(n-1,\alpha)}\right)$, com valor-p dado por: valor-p$=\mathbb{P}\left(T_{n-1} > t_{obs}\right)$.

Testes de Hipóteses para μ

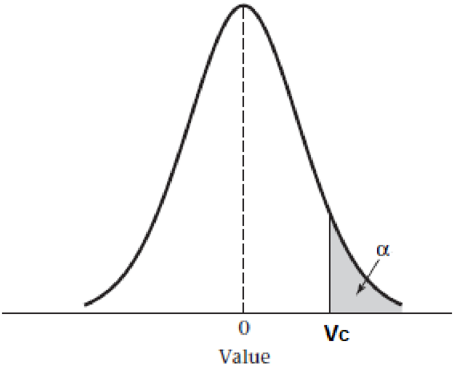

Figura: Região crítica para um teste unilateral à direita. V_c denota o valor crítico (podendo ser t ou z), de área α.

Testes de Hipóteses para μ

CASO II (Hipóteses do Tipo) $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$.

(i) Supondo σ^2 Conhecido: para esse caso, se α denota o tamanho do teste (ou nível de significância), a H_0 será rejeitada à favor da H_1 se $z_{obs} < -z_\alpha$, onde

$$z_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{\sigma} \sim \mathcal{N}(0,1).$$

Portanto, a região crítica é dada por, $\bar{X}_n < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$. Ou, $RC = \{\mathbf{x} : \bar{x}_{obs} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}\}$. Com valor-p dado por: valor-p$=\mathbb{P}\left(Z < z_{obs}\right)$.

Testes de Hipóteses para μ

(ii) Supondo σ^2 Desconhecido: segue a mesma ideia do CASO I, que,

$$t_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{S} \sim t_{n-1}.$$

Portanto, rejeitamos a H_0 à favor da H_1 se $t_{obs} < -t_{(n-1,\alpha)}$, com região crítica dada por, $RC = \{\mathbf{x} : \bar{x}_{obs} < \mu_0 - t_{(n-1,\alpha)}\frac{s}{\sqrt{n}}\}$.

- Sendo que, o tamanho do teste será tal que, $\alpha = \mathbb{P}\left(T_{n-1} > t_{(n-1,\alpha)}\right)$, com calor-p dado por: valor-p$=\mathbb{P}\left(T_{n-1} > t_{obs}\right)$.

Testes de Hipóteses para μ

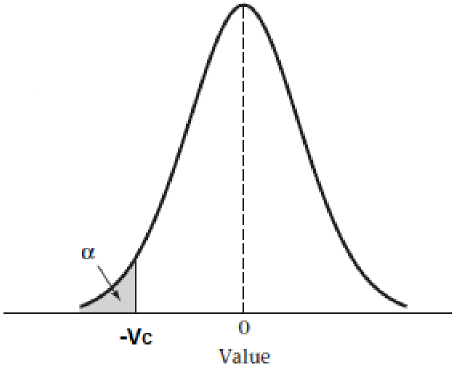

Figura: Região crítica para um teste unilateral à esquerda. $-V_c$ denota o valor crítico (podendo ser t ou z), de área α.

Testes de Hipóteses para μ

CASO III (Hipóteses do Tipo) $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$.

(*i*) Supondo σ^2 Conhecido: para esse caso, se α denota o tamanho do teste (ou nível de significância), a H_0 será rejeitada à favor da H_1 se $|z_{obs}| > z_{\alpha/2}$, onde

$$z_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \overset{sob\, H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{\sigma} \sim \mathcal{N}\left(0, 1\right).$$

Portanto, a região crítica é dada por, $\bar{X}_n < \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ ou $\bar{X}_n > \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. Ou, seja $RC = \{\mathbf{x} : \bar{x}_{obs} < \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ ou $\bar{x}_{obs} > \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}$. Por ser um teste bilateral, o valor-p é dado por: valor-p$=\mathbb{P}\left(|Z| > z_{obs}\right)$.

Testes de Hipóteses para μ

(*ii*) Supondo σ^2 Desconhecido: segue a mesma ideia dos casos anteriores. Ou seja, em primeiro lugar vamos calcular a variância amostral S^2. Assim,

$$t_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S} \overset{sob\,H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{S} \sim t_{n-1}.$$

Portanto, rejeitamos a H_0 à favor da H_1 se $|t_{obs}| > t_{(n-1,\alpha/2)}$, com região crítica dada por,

$$RC = \left\{\mathbf{x} : \bar{x}_{obs} < \mu_0 - t_{(n-1,\alpha/2)}\frac{S}{\sqrt{n}} \text{ ou } \bar{x}_{obs} > \mu_0 + t_{(n-1,\alpha/2)}\frac{S}{\sqrt{n}}\right\}.$$

Por ser um teste bilateral, o valor-p é dado por: valor-p$=\mathbb{P}\left(|T_{n-1}| > t_{obs}\right)$.

Sendo que, o o nível de significância é tal que, $\alpha = \mathbb{P}\left(T_{n-1} > t_{(n-1,\alpha)}\right)$, com

Testes de Hipóteses para μ

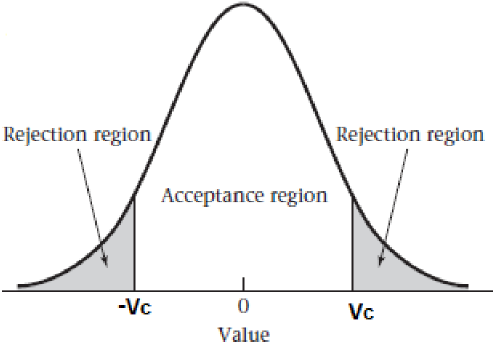

Figura: Região crítica para um teste bilateral. $-V_c$ e V_c denotam os valores críticos (podendo ser t ou z), de área $\alpha = \alpha/2 + \alpha/2$.

Testes de Hipóteses para μ

- Observe que, as regiões críticas das hipóteses formuladas nos CASOS I-III, forram obtidas facilmente sob a suposição de que a variável aleatória X segue uma distribuição normal.

Definição 7 (Resultado Assintótico)

Assuma que $X_1, X_2, \cdots, X_n$ são cópias iid's de uma variável aleatória $X \sim F$, onde F é uma função de distribuição (discreta ou contínua), caracterizada por uma média $\mathbb{E}(X) = \mu < \infty$ e variância, $Var(X) = \sigma^2 > 0$. Isto é, $\Theta = \{(\mu, \sigma^2) : \mu < \infty \text{ e } 0 < \sigma^2 < \infty\}$. Segue que, $\bar{X}_n = \sum_{i=1}^{n} X_i / n$ e $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$ são estimadores não viesados e consistentes para μ e σ^2, respectivamente.

Testes de Hipóteses para μ

- Assim, segue do TCL que, $\bar{X}_n \overset{a}{\sim} \mathcal{N}\left(\mu, \sigma^2/n\right)$. Desta forma, como σ^2 é uma quantidade desconhecida, então,

$$z_{obs} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S} \overset{sob\ H_0}{=} \frac{\sqrt{n}\left(\bar{X}_n - \mu_0\right)}{S} \overset{a}{\sim} \mathcal{N}\left(0,1\right).$$

- Portanto, as regiões críticas consideradas nos casos anteriores podem ser replicadas aqui, porém, agora considerando apenas quantis da distribuição normal, sob a suposição de que $n \to \infty$.
- Considerando as hipóteses do CASO I: $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$, podemos obter a seguinte função poder (essa ideia pode ser replicada para os casos anteriores):

$$\pi_\tau\left(\mu\right) \;=\; \mathbb{P}\left(T\left(\mathbf{X}\right) \in RC | \mu > \mu_0\right) = \Phi\left(-z_\alpha - \frac{\sqrt{n}\left(\mu_0 - \mu^*\right)}{S}\right),$$

com $\mu^* > \mu_0$, tal que, $\mu^* \in \Theta_1$.

TESTES DA RAZÃO DE VEROSSIMILHANÇA

Testes da Razão de Verossimilhança

- Nessa parte da aula vamos fazer uma extensão dos testes de hipóteses apresentados anteriormente.
- Os testes de razão de verossimilhança são baseados na função de verossimilhança previamente abordado nas aulas anteriores (veja: Slides AULA INF04-GRAD).
- Em linhas gerais, assuma que $X_1, X_2, \cdots, X_n$ são cópias iid's de $X \sim f(x|\theta)$, com $\theta \in \Theta$. Segue da independência das variáveis aleatórias que,

$$L(\theta|\mathbf{X}) = \prod_{i=1}^{n} f(x_i|\theta), \tag{4}$$

com $\mathbf{X} = (X_1, X_2, \cdots, X_n)^\top$.

Testes da Razão de Verossimilhança

- Supondo que estamos interessados em testar as seguintes hipóteses,

$$H_0 : \theta \in \Theta_0 \text{ vs } H_1 : \theta \in \Theta_1.$$

- Sendo $\Theta = \Theta_0 \cup \Theta_1$ denota o espaço paramétrico de θ, a amostra aleatória pode ser proveniente de duas distribuições, $f_0(\theta|x) = f(\theta_0|x)$ para $\theta_0 \in \Theta_0$ e $f_1(\theta|x) = f(\theta_1|x)$ para $\theta_1 \in \Theta_1$.

Definição 8 (Teste da Razão de Verossimilhança)

Seja $X_1, X_2, \cdots, X_n$ cópias iid's de $X \sim f(x|\theta)$, com $\theta \in \Theta$. Assuma que nosso interesse seja de testar as hipóteses apresentadas anteriormente. Então, a estatística do Teste da Razão de Verossimilhança (TRV) é dada por:

$$\Lambda = \frac{\sup\limits_{\theta \in \Theta_0} L(\theta|\mathbf{X})}{\sup\limits_{\theta \in \Theta} L(\theta|\mathbf{X})} = \frac{L\left(\hat{\theta}_0|\mathbf{X}\right)}{L\left(\hat{\theta}_n|\mathbf{X}\right)}. \tag{5}$$

Testes da Razão de Verossimilhança

- O TRV, como o próprio nome sugere, é obtido fazendo o quociente das funções de verossimilhança baseadas no modelo restrito (sob H_0), com aquela obtida para o modelo irrestrito (sob H_1).
- Sob H_1, o modelo é dito ser irrestrito, porque o EMV obtido é baseado em todo espaço paramétrico.
- Como regra de decisão, podemos afirmar que: para alguma constante k fixada, tal que $0 \leq k \leq 1$, o TRV rejeita H_0 à favor da H_1 se $\Lambda \leq k$, onde

$$\alpha = \mathbb{P}_{\theta_0}\left(\Lambda \leq k\right).$$

- Como $0 < \Lambda \leq 1$, então, o TRV não rejeita H_0 se $\Lambda \approx 1$.

Testes da Razão de Verossimilhança

- Ou seja, para algum k fixo, o TRV rejeita a H_0 se a razão de verossimilhança em (5) for pequena. Quando isso acontece, significa que é muito provável que a *aa* tenha sido extraída de $f_1\left(x|\theta\right)$, com $\theta \in \Theta_1$ e não da $f_0\left(x|\theta\right),\ \theta \in \Theta_0$.
- O item anterior afirma que, a razão de duas funções de verossimilhança $L\left(\theta_0|\mathbf{X}\right)$ e $L\left(\theta_1|\mathbf{X}\right)$ será pequena se existirem pontos no espaço paramétrico, para qual a amostra observada seja mais provável de ter sido proveniente de Θ_1, do que qualquer outro ponto, $\theta \in \Theta_0$.

Exemplo 5: Seja $X_1, X_2, \cdots, X_n$ uma amostra aleatória de $X \sim \mathcal{N}(\theta, 1)$. Considere as seguintes hipóteses $H_0 : \theta = \theta_0$ contra $H_1 : \theta \neq \theta_0$. Observe que, de forma similar aos casos anteriores, aqui θ_0 é um valor fixado pelo experimentador.

Neste caso, a função de verossimilhança é dada por:

$$L(\theta|\mathbf{X}) = \prod_{i=1}^{n} f(X_i|\theta) = (2\pi)^{-n/2} \exp\left(-\sum_{i=1}^{n} (X_i - \theta)^2 / n\right).$$

Note que, para o modelo restrito (ou seja, sob H_0), é possível observar que estamos perante uma hipótese simples, i.e., $\Theta_0 = \{\theta_0\}$ com $\sigma^2 = 1$, então a função de verossimilhança no modelo é dada por,

$$L(\theta_0|\mathbf{X}) = (2\pi)^{-n/2} \exp\left(-\sum_{i=1}^{n} (X_i - \theta_0)^2 / n\right).$$

Testes da Razão de Verossimilhança

- No caso da maximização irrestrita, o EMV em Θ é $\hat{\theta}_n = \bar{X}_n$, e portanto, obtemos

$$L\left(\hat{\theta}_n|\mathbf{X}\right) = L\left(\bar{X}_n|\mathbf{X}\right) = (2\pi)^{-n/2} \exp\left(-\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2/n\right).$$

- A estatística do TRV é então dada por:

$$\begin{aligned}\Lambda &= \frac{\sup\limits_{\theta\in\Theta_0} L(\theta|\mathbf{X})}{\sup\limits_{\theta\in\Theta} L(\theta|\mathbf{X})} = \frac{(2\pi)^{-n/2}\exp\left(-\sum\limits_{i=1}^{n}(X_i-\theta_0)^2/n\right)}{(2\pi)^{-n/2}\exp\left(-\sum\limits_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2/n\right)}.\\ &= \exp\left[-\frac{n}{2}\left(\bar{X}_n-\theta_0\right)^2\right].\end{aligned}$$

Testes da Razão de Verossimilhança

Desta forma, a H_0 será rejeitada à favor da H_1 se $\Lambda \leq k$. Isto é,

$$\Lambda \leq k \iff \exp\left[-\frac{n}{2}\left(\bar{X}_n - \theta_0\right)^2\right] \leq k,$$

resultando na seguinte região crítica
$C_\alpha = \{(x_1, x_2, \cdots, x_n) : |\bar{x} - \theta_0| > \sqrt{-\frac{2\log(k)}{n}}\}$.

Portanto, como $0 \leq k \leq 1$, a região crítica vai mapear valores em $\mathbb{R}^+ = (0, \infty)$.

Comentário: O TRV para o presente exemplo, engloba exatamente aqueles testes rejeitam a H_0 se a média amostral diferir do valor hipotético θ_0 em mais de $k^* = \sqrt{-\frac{2\log(k)}{n}}$ unidades.

Testes da Razão de Verossimilhança

Teorema 1

Seja $X_1, X_2, \cdots, X_n$ cópias iid$'$s de $X \sim f(x|\theta)$, com $\theta \in \Theta$. Sob as condições de regularidade usuais, se $\theta \in \Theta_0$ (sob H_0 verdadeira), então, a distribuição da estatística $-2\log\Lambda \sim \chi^2_q$, com q denotando a diferença entre do número de parâmetros em Θ e Θ_0, respectivamente.

Exemplo 6: Seja $X_1, X_2, \cdots, X_n$ cópias iid's de X com distribuição de probabilidades dada por: $f(x|\theta) = e^{-(x-\theta)}$, para $x \geq \theta$ e zero, caso contrário. Mostre que a região crítica do teste é dada por:
$C_\alpha = \{(x_1, x_2, \cdots, x_n) : x_{(1)} > \theta - \log(k)/n\}$.

TESTES DE HIPÓTESES ÓTIMOS

Testes Mais Poderosos

- Nas aulas anteriores apresentamos a parte introdutória dos testes de hipóteses, nomeadamente: (*i*) sua definição, (*ii*) Classificação, (*iii*) Regras de decisão, (*iv*) Tipos de erros que podem ocorrer durante a execução de um teste de hipóteses, e (*v*) O poder do teste.
- Os testes apresentados até o momento eram em geral, formados por uma H_0 simples (ou composta), mas com uma H_1 composta (em todos os casos).
- Ou seja, $H_0 : \theta \in \Theta_0$ contra $H_1 : \theta \in \Theta_1$, com $\Theta_0 \cup \Theta_1 = \Theta$ e $\Theta_0 \cap \Theta_1 = \emptyset$. Como consequência,

$$\alpha = \sup_{\theta \in \Theta_0} \mathbb{P}\left(T\left(\mathbf{X}\right) \in C_\alpha\right).$$

Testes Mais Poderosos

- A ideia que norteia os testes ótimos (testes mais poderosos) é a minimização dos erros tipo I e II, respectivamente. Com principal enfoque no erro do tipo II (que está diretamente associado com a função poder).
- Os testes mais poderosos (*notação:* TMP) visam essencialmente obter **a melhor região crítica** de tamanho α, entre varias outras regiões críticas existentes em $\mathcal{X}$.
- Para esse tipo de testes, as duas hipóteses (H_0 e H_1), o modelo será completamente especificado. Ou seja,

$$H_0 : \theta = \theta^{'} \text{ contra } H_1 : \theta = \theta^{''}, \tag{6}$$

com $\Theta_0 = \{\theta : \theta = \theta^{'}\}$ e $\Theta_1 = \{\theta : \theta = \theta^{''}\}$.

Testes Mais Poderosos

Definição 9

Seja C_α *um subconjunto de* $\mathcal{X}$*. Então,* C_α *é dita ser* **a melhor região crítica** *de tamanho* α *para testar as hipóteses simples,*
$H_0 : \theta = \theta^{'}$ *contra* $H_1 : \theta = \theta^{''}$*, se:*

a. $\mathbb{P}_{\theta'}\left(T\left(\mathbf{X}\right) \in C_\alpha\right) = \alpha.$
b. *Para todo subconjunto* A_α *de* $\mathcal{X}$ *segue que,*
   $\mathbb{P}_{\theta'}\left(T\left(\mathbf{X}\right) \in A_\alpha\right) = \alpha \Longrightarrow \mathbb{P}_{\theta''}\left(T\left(\mathbf{X}\right) \in C_\alpha\right) \geq \mathbb{P}_{\theta''}\left(T\left(\mathbf{X}\right) \in A_\alpha\right)$

- A Definição 9 afirma que, se existem múltiplos subconjuntos $A_\alpha \in \mathcal{X}$, todos com o tamanho α, assuma que existe igualmente um subconjunto C_α com o mesmo tamanho de teste, porém, com maior poder estatístico que os A_α's então, C_α é a melhor região crítica do teste de tamanho α.

Exemplo 6: Considere a variável X com distribuição Binom(n, *theta*), com $n = 5$ e $\theta = \mathbb{P}(X = 1)$. Seja $f(x|\theta)$, com $\theta \in \Theta$ a *fp* de X. E considere as seguintes hipóteses, $H_0 : \theta = 1/2$ contra $H_1 : \theta = 3/4$. Com base na informação do enunciado, temos os resultados da tabela abaixo:

x	0	1	2
$f(x;1/2)$	1/32	5/32	10/32
$f(x;3/4)$	1/1024	15/1024	90/1024
$f(x;1/2)/f(x;3/4)$	32/1	32/3	32/9
x	3	4	5
$f(x;1/2)$	10/32	5/32	1/32
$f(x;3/4)$	270/1024	405/1024	243/1024
$f(x;1/2)/f(x;3/4)$	32/27	32/81	32/243

Figura: Probabilidades calculadas para cada ponto em $\mathcal{X} = \{0, 1, \cdots, 5\}$, para $\Theta = \{\theta : \theta = 1/2, 3/4\}$.

Testes Mais Poderosos

- Observe que, com base na Tabela anterior, fixando um um nível de significância $1/32 = \alpha = \mathbb{P}_{\theta=1/2}(X = 0) = \mathbb{P}_{\theta=1/2}(X = 5)$, podemos definir duas regiões críticas, digamos, $A_1 = \{x : x = 0\}$ e $A_2 = \{x : x = 5\}$, tal que, $A_1, A_2 \in \mathcal{X} = \{0, 1, 2, 3, 4, 5\}$ são potencias melhores regiões críticas com tamanho $\alpha = 1/32$.

- No entanto, ter o mesmo tamanho de teste é uma condição necessária, mas não suficiente. Desta forma, vamos calcular seus respectivos poderes de teste. Assim, para A_1, note que,

$$\pi_{A_1}(\theta = 3/4) \;=\; \mathbb{P}_{\theta=3/4}(X \in A_1) = \mathbb{P}_{\theta=3/4}(X = 0) = 1/1024 < \alpha.$$

- Portanto, para a região crítica A_1, a probabilidade de rejeitar uma H_0 sendo ela verdadeira é maior que a potencia do teste (não rejeitar H_0 sendo ela verdadeira/decisão correta). Portanto, A_1 não é a melhor região crítica para o teste.

Testes Mais Poderosos

- De forma idêntica, para a região crítica A_2, temos que, $1/32 = \alpha = \mathbb{P}_{\theta=1/2}(X = 5)$. Cujo seu poder é dado por:

$$\pi_{A_2}(\theta = 3/4) = \mathbb{P}_{\theta=3/4}(X \in A_2) = \mathbb{P}_{\theta=3/4}(X = 5) = 243/1024.$$

- Ou seja $\pi_{A_2}(\theta = 3/4) = 243/1024 > \alpha$. Então, como essa é a única região crítica com tamanho α maior que o poder de teste, concluímos que A_2 é a melhor região crítica para o *TMP*.
- Ou seja, a ideia de encontrar uma região crítica $C_\alpha = A_2$ que seja melhor que as demais, pode ser resolvido procurando um C_α onde $f(x|H_0)$ seja relativamente menor que $f(x|H_1)$.
- O ponto anterior é verdadeiro, porque a escolha da melhor região crítica pode ser feita olhando pelos pontos de x, associados a um mesmo tamanho de teste, que apresentam um maior rácio $f(x|H_0)/f(x|H_1)$.

Testes Mais Poderosos

Teorema 2 (Neyman-Pearson)

Seja $X_1, X_2, \cdots, X_n$ *uma amostra aleatória de tamanho n de uma variável aleatória* $X \sim f(x|\theta)$, $\theta \in \Theta$. *Onde* $f(x|\theta)$ *pode ser uma fdp/fp. Vimos anteriormente que, como os* X_i*'s são iid's segue que,*

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} f(x_i|\theta), \ \mathbf{x} = (x1, x2, \cdots, x_n)^\top. \tag{7}$$

Seja θ' *e* θ'' *dois valores fixos e distintos, e seja k uma constante positiva. Se* C_α *denota a região crítica, e* $R(\mathbf{x}) = \frac{L(\theta'|\mathbf{x})}{L(\theta''|\mathbf{x})}$ *a razão de verossimilhança, então,*

- a) $R(\mathbf{x}) \leq k$, *se* $t(\mathbf{x}) \in C_\alpha$.
- b) $R(\mathbf{x}) \geq k$, *se* $t(\mathbf{x}) \in C_\alpha^c$.
- c) $\alpha = \mathbb{P}_{H_0}[T(\mathbf{X}) \in C_\alpha]$.

Obs.: Os detalhes da prova podem ser encontrados em Casella & Berger, 1999; pag.347.

Testes Mais Poderosos

- Observe que, nos casos e que $f\left(x|\theta\right)$ denota uma função de probabilidade, a função de decisão baseada no resultado do Teorema 2, é dada por:

$$\phi\left(x\right)=\begin{cases} 1, & \text{se} \quad R\left(\mathbf{x}\right)<k, \\ \gamma, & \text{se} \quad R\left(\mathbf{x}\right)=k, \\ 0, & \text{se} \quad R\left(\mathbf{x}\right)>k. \end{cases}$$

- Em linhas gerais, a noção de um *TMP* indica que, uma boa região crítica é aquela que apresenta um tamanho de teste α, com um poder superior a probabilidade do erro tipo I, no ponto $\theta_0 \in \Theta_0$.
- Ou seja, o teorema anterior indica que a probabilidade de se cometer um erro do tipo I, deve ser sempre menor que a probabilidade de se tomar uma decisão correta.

Exemplo 7: Seja $X_1, X_2, \cdots, X_n$ uma *aa* proveniente de uma distribuição $\mathcal{N}(\theta, 1)$, com $\theta \in \mathbb{R}$. Assuma que nosso interesse é testar as seguintes hipóteses, $H_0 : \theta = \theta' \;\; (\theta' = 0)$ contra $H_1 : \theta = \theta'' \;\; (\theta'' = 1)$. Ou seja, ambas as hipóteses são simples.

Solução: Como a *aa* foi extraída de uma população normal, a *fdp* da variável aleatória X é dada por, $f(x|\theta) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x-\theta)^2\right\}$, resultando na seguinte função de verossimilhança,

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} f(x_i|\theta) = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right\}.$$

O próximo passo seria obter as funções de verossimilhanças restritas às duas hipóteses em questão.

Testes Mais Poderosos

Desta forma, segue que,

$$\text{Sob } H_0: \qquad L\left(\theta'|\mathbf{x}\right) = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\}.$$

$$\text{Sob } H_1: \qquad L\left(\theta''|\mathbf{x}\right) = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} (x_i - 1)^2\right\}.$$

Com base no resultado anterior, a estatística que define a razão das duas funções de verossimilhança é dada por:

$$R\left(\mathbf{x}\right) = \frac{L\left(\theta'|\mathbf{x}\right)}{L\left(\theta''|\mathbf{x}\right)} = \frac{(2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\}}{(2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} (x_i - 1)^2\right\}} = \exp\left\{-\sum_{i=1}^{n} x_i + n/2\right\}.$$

Para algum $k > 0$, rejeitamos H_0 à favor da H_1 se $R\left(\mathbf{x}\right) \leq k$,

Testes Mais Poderosos

- Resultando na seguinte região crítica $C_\alpha = \left\{ \mathbf{x} : \sum_{i=1}^{n} x_i \geq k^* \right\}$, onde $k^* = n/2 - \log(k)$. Observe que, a região crítica pode ser igualmente apresentada da seguinte forma $C_\alpha = \{\mathbf{x} : \bar{x}_{obs} \geq k^{**}\}$, com $k^{**} = 1/2 - (1/n)\log(k)$.

- Como os dados são provenientes de uma distribuição normal, segue das aulas anteriores que, $\bar{X}_n \sim \mathcal{N}(\theta, 1/n)$. Ou seja, a quantidade k^{**} pode ser obtida facilmente usando os quantis da distribuição normal, usando o fato de que (com α dado),

$$\alpha = \mathbb{P}_{H_0}\left(\bar{X}_n \geq k^{**}\right).$$

- De igual forma, calculado o valor de k^{**}, podemos obter o pode do teste, por forma a concluirmos que de fato, C_α é a melhor região crítica para as hipóteses apresentadas anteriormente.

Exemplo 8: Seja $X_1, X_2, \cdots, X_n$ uma *aa* de uma variável aleatória $X \sim Ber(\theta)$, com $\theta \in \Theta = \{\theta : 0 < \theta < 1\}$. Considere as seguintes hipóteses $H_0 : \theta = \theta'$ contra $H_1 : \theta = \theta''$, com $\theta > \theta'$. Observe que, para esse problema, a função de verossimilhança é dada por:

$$L(\theta|\mathbf{x}) = \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n - \sum_{i=1}^{n} x_i}.$$

Consequentemente, segue do Teorema de Neyman Pearson que,

$$R(\mathbf{x}) = \frac{L(\theta'|\mathbf{x})}{L(\theta''|\mathbf{x})} = \frac{(\theta')^{\sum_{i=1}^{n} x_i} (1-\theta')^{n-\sum_{i=1}^{n} x_i}}{(\theta'')^{\sum_{i=1}^{n} x_i} (1-\theta'')^{n-\sum_{i=1}^{n} x_i}} = \left(\frac{1-\theta'}{1-\theta''}\right)^n \left[\frac{\theta'(1-\theta'')}{\theta''(1-\theta')}\right]^{\sum_{i=1}^{n} x_i},$$

desta forma, rejeitamos H_0 à favor da H_1 se para algum $k > 0$, $R(\mathbf{x}) \leq k$, fornecendo a seguinte região crítica, $C_\alpha = \{\mathbf{x} : \sum_{i=1}^{n} x_i \leq k^*\}$.

Testes Mais Poderosos

- Assim, C_α é a melhor região crítica para testar as hipóteses $H_0 : \theta = \theta'$ contra $H_1 : \theta = \theta''$, com $\theta > \theta'$ com $k^* = \dfrac{\log\left(\frac{1-\theta''}{1-\theta'}\right)^n}{\log\left(\frac{\theta'(1-\theta'')}{\theta''(1-\theta')}\right)}$.
- Tal como no caso anterior, k^* é uma constante que pode ser obtida usando a distribuição de probabilidade da estatística $T(\mathbf{X}) = \sum_{i=1}^{n} X_i \sim Binom(n, \theta)$. Ou seja, para algum α dado, segue que,

$$\alpha = \mathbb{P}_{H_0}\left(\sum_{i=1}^{n} X_i \leq k^*\right).$$

- Por fim, o poder do teste pode ser igualmente obtido, e de seguida comparado com o nível α, supondo que os valores θ', θ'' e n dados.

Testes Mais Poderosos

- É importante clarificar que, os TMP não pressupõem apenas testar hipóteses paramétricas. Ou seja, além da suposição de que tais hipóteses devem ser simples, e as variáveis devem ser $iid's$, podemos igualmente obter um TMP testando distribuição (com o mesmo suporte nas duas hipóteses).