Detalhamento Matemático: Testes de Razão de Verossimilhança (TRV)

Foco nos Casos 3, 4 e Duas Amostras

Explicação Detalhada para Estudo

November 22, 2025

Contents

1 Introdução

Este documento detalha os passos matemáticos para a construção dos Testes de Razão de Verossimilhança (TRV), focando nos pontos de maior dificuldade: **Casos 3 e 4** (testes para variância de uma normal) e **testes para duas amostras** (comparação de médias e variâncias).

A estatística de Razão de Verossimilhança é definida como:

$$\Lambda = \frac{\sup_{\theta \in \Theta_0}\{L(\theta)\}}{\sup_{\theta \in \Theta}\{L(\theta)\}} \tag{1}$$

Rejeitamos H_0 se $\Lambda < k$, onde k é uma constante associada ao nível de significância α.

2 Caso 3: Teste para Variância com Média Conhecida

2.1 Definição do Problema

Seja $X_1, \ldots, X_n$ uma amostra aleatória de $X \sim N(\mu, \sigma^2)$.

- **Parâmetros:** μ é **conhecido**; σ^2 é **desconhecido**.
- **Hipóteses:** $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$.
- **Espaço Paramétrico Irrestrito (Θ):** $\sigma^2 > 0$.
- **Espaço Paramétrico Restrito (Θ_0):** $\sigma^2 = \sigma_0^2$ (ponto único).

2.2 A Função de Verossimilhança

A densidade conjunta (verossimilhança) é:

$$L(\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

2.3 Passo 1: O Denominador (Máximo sob Θ)

Precisamos encontrar o Estimador de Máxima Verossimilhança (EMV) de σ^2 sem restrições. Aplicando o logaritmo em $L(\sigma^2)$:

$$\ell(\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Derivando em relação a σ^2 e igualando a zero:

$$\frac{d\ell}{d\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

$$\frac{n}{2\hat{\sigma}^2} = \frac{\sum(x_i-\mu)^2}{2(\hat{\sigma}^2)^2} \implies \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

O valor máximo da verossimilhança no denominador é $L(\hat{\sigma}^2)$. Substituindo $\hat{\sigma}^2$ na expressão de L:

$$\begin{aligned}\sup_{\Theta} L &= (2\pi\hat{\sigma}^2)^{-n/2} \exp\left\{ -\frac{1}{2\hat{\sigma}^2} \underbrace{\sum (x_i - \mu)^2}_{n\hat{\sigma}^2} \right\} \\ &= (2\pi\hat{\sigma}^2)^{-n/2} \exp\left\{ -\frac{n\hat{\sigma}^2}{2\hat{\sigma}^2} \right\} = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}\end{aligned}$$

2.4 Passo 2: O Numerador (Máximo sob Θ_0)

Sob H_0, σ^2 é fixo em σ_0^2. Não há maximização a fazer.

$$\sup_{\Theta_0} L = L(\sigma_0^2) = (2\pi\sigma_0^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\}$$

Para simplificar, observe que $\sum (x_i - \mu)^2 = n\hat{\sigma}^2$ (usando o estimador definido acima).

$$\sup_{\Theta_0} L = (2\pi\sigma_0^2)^{-n/2} \exp\left\{ -\frac{n\hat{\sigma}^2}{2\sigma_0^2} \right\}$$

2.5 Passo 3: A Razão Λ

$$\Lambda = \frac{(2\pi\sigma_0^2)^{-n/2} \exp\left\{ -\frac{n\hat{\sigma}^2}{2\sigma_0^2} \right\}}{(2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}}$$

Agrupando os termos semelhantes:

$$\Lambda = \left(\frac{\hat{\sigma}^2}{\sigma_0^2} \right)^{n/2} \frac{\exp\left\{ -\frac{n}{2} \frac{\hat{\sigma}^2}{\sigma_0^2} \right\}}{e^{-n/2}}$$

$$\Lambda = \left(\frac{\hat{\sigma}^2}{\sigma_0^2} \right)^{n/2} \exp\left\{ \frac{n}{2} \left(1 - \frac{\hat{\sigma}^2}{\sigma_0^2} \right) \right\}$$

2.6 Passo 4: A Região Crítica

Seja $u = \frac{\hat{\sigma}^2}{\sigma_0^2}$. Então $\Lambda(u) = (ue^{1-u})^{n/2}$. A função $g(u) = ue^{1-u}$ tem máximo em $u = 1$ e decresce à medida que u se afasta de 1 (tanto para 0 quanto para $+\infty$). Rejeitar H_0 quando $\Lambda < k$ é equivalente a rejeitar quando u é muito pequeno ($u < c_1$) ou muito grande ($u > c_2$).

Substituindo u:

$$\frac{\hat{\sigma}^2}{\sigma_0^2} < c_1 \quad \text{ou} \quad \frac{\hat{\sigma}^2}{\sigma_0^2} > c_2$$

$$\frac{\frac{1}{n}\sum (x_i - \mu)^2}{\sigma_0^2} < c_1 \quad \text{ou} \quad \frac{\frac{1}{n}\sum (x_i - \mu)^2}{\sigma_0^2} > c_2$$

Multiplicando por n, obtemos a estatística de teste W:

$$W = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{\sigma_0^2}$$

Sabemos que se $X_i \sim N(\mu, \sigma_0^2)$, então $\frac{X_i - \mu}{\sigma_0} \sim N(0,1)$, e a soma de quadrados de normais padrão é uma Qui-quadrado.

$$W \sim \chi_n^2$$

Portanto, a região crítica é:

$$W < \chi^2_{n,1-\alpha/2} \quad \text{ou} \quad W > \chi^2_{n,\alpha/2}$$

Nota: Aqui usamos χ_n^2 com n graus de liberdade porque a média μ é **conhecida**.

3 Caso 4: Teste para Variância com Média Desconhecida

3.1 Diferença Fundamental

A estrutura é quase idêntica ao Caso 3, mas agora μ é **desconhecido**. Isso muda a maximização no numerador (sob H_0).

3.2 Passo 1: O Denominador (Máximo sob Θ)

Sob Θ (irrestrito), maximizamos em relação a μ e σ^2.

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(Note que agora usamos $\bar{x}$ em vez de μ).

$$\sup_{\Theta} L = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}$$

3.3 Passo 2: O Numerador (Máximo sob Θ_0)

Sob H_0, temos $\sigma^2 = \sigma_0^2$ (fixo), mas μ é livre.

$$L(\mu, \sigma_0^2) = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

Precisamos encontrar o $\hat{\mu}_0$ que maximiza essa expressão. Isso equivale a minimizar $\sum(x_i - \mu)^2$. A soma de quadrados é mínima quando $\mu = \bar{x}$. Logo, $\hat{\mu}_0 = \bar{x}$.

Substituindo na verossimilhança:

$$\sup_{\Theta_0} L = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right\}$$

Usando a notação $n\hat{\sigma}^2 = \sum(x_i - \bar{x})^2$:

$$\sup_{\Theta_0} L = (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{n\hat{\sigma}^2}{2\sigma_0^2}\right\}$$

3.4 Passo 3: A Razão Λ

A expressão algébrica de Λ acaba sendo **idêntica** à do Caso 3:

$$\Lambda = \left[\frac{\hat{\sigma}^2}{\sigma_0^2} \exp\left(1 - \frac{\hat{\sigma}^2}{\sigma_0^2}\right) \right]^{n/2}$$

Onde agora $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$.

3.5 Passo 4: Distribuição e Graus de Liberdade

A região crítica em termos de $\hat{\sigma}^2$ é a mesma:

$$\frac{n\hat{\sigma}^2}{\sigma_0^2} < k_1 \quad \text{ou} \quad \frac{n\hat{\sigma}^2}{\sigma_0^2} > k_2$$

Ou seja:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_0^2} < k_1 \quad \text{ou} \quad \dots$$

Ponto crucial de entendimento: Qual a distribuição de $W = \frac{\sum(x_i - \bar{x})^2}{\sigma_0^2}$? Diferente do Caso 3, aqui temos $\bar{x}$ estimando μ. Isso introduz uma dependência linear entre os termos $(x_i - \bar{x})$, reduzindo a dimensionalidade do vetor de desvios em 1. Sabemos que:

$$S^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$$

$$\frac{(n-1)S^2}{\sigma_0^2} = \frac{\sum(x_i - \bar{x})^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

Portanto, o teste usa a distribuição **Qui-quadrado com $n-1$ graus de liberdade**:

$$W < \chi^2_{n-1,1-\alpha/2} \quad \text{ou} \quad W > \chi^2_{n-1,\alpha/2}$$

4 Duas Amostras: Comparação de Médias (Variâncias Iguais)

4.1 O Problema

$X \sim N(\mu_x, \sigma^2)$ e $Y \sim N(\mu_y, \sigma^2)$. $H_0 : \mu_x = \mu_y = \mu$ vs $H_1 : \mu_x \neq \mu_y$.

4.2 Passo 1: Irrestrito (H_1)

Maximizamos separadamente. $\hat{\mu}_x = \bar{x}$, $\hat{\mu}_y = \bar{y}$. A variância comum σ^2 é estimada pela média dos erros quadráticos totais:

$$\hat{\sigma}_1^2 = \frac{\sum(x_i - \bar{x})^2 + \sum(y_j - \bar{y})^2}{n+m}$$

Supremo: $L_1 \propto (\hat{\sigma}_1^2)^{-(n+m)/2}$.

4.3 Passo 2: Restrito (H_0)

Sob H_0, $\mu_x = \mu_y = \mu$. O estimador de μ é a média ponderada geral:

$$\hat{\mu}_0 = \frac{n\bar{x} + m\bar{y}}{n+m}$$

A variância estimada sob H_0 inclui a variação em relação a essa média comum:

$$\hat{\sigma}_0^2 = \frac{\sum(x_i - \hat{\mu}_0)^2 + \sum(y_j - \hat{\mu}_0)^2}{n+m}$$

Supremo: $L_0 \propto (\hat{\sigma}_0^2)^{-(n+m)/2}$.

4.4 Passo 3: A Razão e a Identidade Algébrica

$$\Lambda = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}\right)^{-(n+m)/2}$$

Rejeitamos para Λ pequeno $\iff$ $\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}$ grande. Vamos decompor o numerador $\sum(x_i - \hat{\mu}_0)^2$:

$$\sum(x_i - \hat{\mu}_0)^2 = \sum(x_i - \bar{x} + \bar{x} - \hat{\mu}_0)^2 = \sum(x_i - \bar{x})^2 + n(\bar{x} - \hat{\mu}_0)^2$$

Fazendo o mesmo para Y e somando:

$$(n+m)\hat{\sigma}_0^2 = \underbrace{\sum(x_i - \bar{x})^2 + \sum(y_j - \bar{y})^2}_{(n+m)\hat{\sigma}_1^2} + \underbrace{n(\bar{x} - \hat{\mu}_0)^2 + m(\bar{y} - \hat{\mu}_0)^2}_{Q}$$

O termo Q mede a distância entre as médias amostrais e a média global. Pode-se provar que:

$$Q = \frac{nm}{n+m}(\bar{x} - \bar{y})^2$$

Assim, a razão é:

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} = 1 + \frac{Q}{(n+m)\hat{\sigma}_1^2} = 1 + \frac{\frac{nm}{n+m}(\bar{x} - \bar{y})^2}{\sum(x_i - \bar{x})^2 + \sum(y_j - \bar{y})^2}$$

4.5 Passo 4: Conexão com o Teste t

O termo variável na razão acima é essencialmente o quadrado da estatística t. Lembrando que $S_p^2 = \frac{\sum(x_i - \bar{x})^2 + \sum(y_j - \bar{y})^2}{n+m-2}$. A estatística T é:

$$T = \frac{\bar{x} - \bar{y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Elevando ao quadrado:

$$T^2 = \frac{(\bar{x} - \bar{y})^2}{S_p^2 \frac{n+m}{nm}} = \frac{\frac{nm}{n+m}(\bar{x} - \bar{y})^2}{S_p^2}$$

Substituindo na razão de verossimilhança:

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} = 1 + \frac{T^2 \cdot S_p^2}{(n+m-2)S_p^2} \cdot \frac{n+m-2}{n+m} \approx 1 + \frac{T^2}{n+m}$$

(Ignorando constantes multiplicativas exatas, a monotonicidade se mantém).

O TRV rejeita quando essa razão é grande $\iff$ T^2 é grande $\iff$ $|T|$ é grande. Isso confirma que o TRV para médias é equivalente ao **Teste t bilateral** usual.

5 Duas Amostras: Comparação de Variâncias

5.1 O Problema

$H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 \neq \sigma_y^2$.

5.2 Passo 1: Irrestrito (H_1)

Maximizando separadamente: $\hat{\sigma}_x^2$ e $\hat{\sigma}_y^2$. $L_1 \propto (\hat{\sigma}_x^2)^{-n/2}(\hat{\sigma}_y^2)^{-m/2}$.

5.3 Passo 2: Restrito (H_0)

Maximizando sob a restrição $\sigma_x^2 = \sigma_y^2 = \sigma^2$:

$$\hat{\sigma}_{comb}^2 = \frac{n\hat{\sigma}_x^2 + m\hat{\sigma}_y^2}{n+m}$$

$L_0 \propto (\hat{\sigma}_{comb}^2)^{-(n+m)/2}$.

5.4 Passo 3: A Razão Λ

$$\Lambda = \frac{(\hat{\sigma}_x^2)^{n/2}(\hat{\sigma}_y^2)^{m/2}}{(\frac{n\hat{\sigma}_x^2 + m\hat{\sigma}_y^2}{n+m})^{(n+m)/2}}$$

Dividindo numerador e denominador por $(\hat{\sigma}_y^2)^{(n+m)/2}$:

$$\Lambda \propto \frac{(\hat{\sigma}_x^2/\hat{\sigma}_y^2)^{n/2}}{(n\frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} + m)^{(n+m)/2}}$$

Seja $F' = \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2}$. A função $\Lambda(F')$ tem a forma $x^a/(c_1x + c_2)^b$. Esta função cresce até um ponto e depois decresce. Rejeitamos H_0 se Λ for pequeno, o que ocorre se F' for **muito pequeno** (próximo de 0) ou **muito grande**.

Isso justifica o uso do teste F clássico:

$$F = \frac{S_x^2}{S_y^2}$$

Onde rejeitamos se $F < F_{\alpha/2}$ ou $F > F_{1-\alpha/2}$.