

Trabalho06

higor lucas de Araujo Freitas

2024-09-30

```
knitr::opts_chunk$set(echo = TRUE)
```

Parte 1 - Intervalo de Confiança para a Variância

Enunciado

Vamos realizar um exercício para inferir o intervalo de confiança da variância de uma população usando a linguagem de programação R. Este exercício usará uma amostra aleatória e o método baseado na distribuição Chi-quadrado, que é o método padrão para inferir sobre a variância de uma população normalmente distribuída.

Passos:

1. Gerar uma Amostra Aleatória: Começamos gerando uma amostra aleatória de uma distribuição normal com média 0 e desvio padrão 1.
2. Calcular a Variância da Amostra: Usamos a função `var()` para calcular a variância da nossa amostra.
3. Calcular o Intervalo de Confiança da Variância:
4. O intervalo de confiança para a variância σ^2 é baseado na estatística de amostra s^2 (a variância da amostra) e segue uma distribuição chi-quadrado com $n-1$ graus de liberdade, onde n é o tamanho da amostra.

O intervalo de confiança é dado por $[(n-1)s^2/\chi^2(\alpha/2), (n-1)s^2/\chi^2(1-\alpha/2)]$, onde α é o nível de significância (por exemplo, para um intervalo de confiança de 95%, $\alpha = 0.05$) e $\chi^2(\alpha/2)$ e $\chi^2(1-\alpha/2)$ são os valores críticos da distribuição chi-quadrado nos dois extremos do intervalo.

```
# Definir o tamanho da amostra e o nível de confiança
n <- 50 # tamanho da amostra
conf.level <- 0.95 # nível de confiança

# Gerar uma amostra aleatória
set.seed(123) # para reprodutibilidade
amostra <- rnorm(n, mean = 0, sd = 1)

# Calcular a variância da amostra
s2 <- var(amostra)
```

```
# Calcular o intervalo de confiança para a variância
alpha <- 1 - conf.level
chi2_lower <- qchisq(alpha / 2, df = n - 1)
chi2_upper <- qchisq(1 - alpha / 2, df = n - 1)
ci_lower <- ((n - 1) * s2) / chi2_upper
ci_upper <- ((n - 1) * s2) / chi2_lower

# Mostrar o intervalo de confiança
cat("Intervalo de confiança para a variância (", conf.level*100, "%): [", ci_lower, ", ", ci_upper, "]\n")

## Intervalo de confiança para a variância (95%): [0.5981641, 1.331156]
```

Parte 2 - Teste de hipótese para variância de uma população

Enunciado

Vamos fazer um exercício de teste de hipótese para inferência sobre a variância de uma população usando a linguagem de programação R. Este exercício usará um cenário hipotético no qual queremos testar se a variância de uma determinada população é igual a um valor específico, usando uma amostra dessa população.

Passos para o Teste de Hipótese:

1. Formular as Hipóteses Hipótese nula (H_0): A variância da população é igual a σ_0^2 (por exemplo, $\sigma^2 = 1$). Hipótese alternativa (H_a): A variância da população é diferente de σ_0^2 (isto é, $\sigma^2 \neq 1$).
2. Coletar uma Amostra
3. Calcular a Estatística de Teste A estatística de teste é baseada na variância da amostra e segue uma distribuição qui-quadrado com $n-1$ graus de liberdade, onde n é o tamanho da amostra. A estatística de teste é calculada como: $(n-1)s^2 / \sigma_0^2$, onde s^2 é a variância da amostra.
4. Decidir sobre o Nível de Significância: Escolha o nível de significância (α), geralmente 0.05.
5. Determinar o Valor Crítico e a Região Crítica: Encontre os valores críticos da distribuição chi-quadrado com base no nível de significância e nos graus de liberdade.
6. Interpretar Rejeite a hipótese nula se a estatística de teste cair na região crítica.

```
# Definir o tamanho da amostra, a variância sob H0 e o nível de significância
n <- 60 # tamanho da amostra
sigma0_squared <- 1 # variância sob H0
alpha <- 0.05 # nível de significância

# Gerar uma amostra aleatória
set.seed(123) # para reprodutibilidade
# amostra <- rnorm(n, mean = 0, sd = sqrt(sigma0_squared))
amostra <- rnorm(n, mean = 0, sd = sqrt(2)) # mude aqui para verificar saída do teste

# Calcular a variância da amostra
s_squared <- var(amostra)
```

```

# Calcular a estatística de teste
test_statistic <- (n - 1) * s_squared / sigma0_squared

# Determinar os valores críticos
chi2_lower <- qchisq(alpha / 2, df = n - 1)
chi2_upper <- qchisq(1 - alpha / 2, df = n - 1)

# Tomar uma decisão
cat("Estatística de teste: ", test_statistic, "\n")

## Estatística de teste: 97.7966

cat("Região crítica: <", chi2_lower, " ou >", chi2_upper, "\n")

## Região crítica: < 39.66186 ou > 82.11741

if (test_statistic < chi2_lower || test_statistic > chi2_upper) {
  cat("Rejeitamos a hipótese nula: \n a variância da população é diferente de", sigma0_squared, "\n")
} else {
  cat("Não rejeitamos a hipótese nula: \n não há evidências suficientes para afirmar que a variância da
}

## Rejeitamos a hipótese nula:
## a variância da população é diferente de 1

```

Hipótese nula (H_0): A variância da população é diferente a σ_0^2 (por exemplo, $\sigma^2 \neq 1$).

Parte 3 - Teste de hipótese para comparação da variância de duas populações

Enunciado

Vamos realizar um exercício aplicando um teste de hipótese para comparar as variâncias de duas populações. Usaremos o teste F, que é comumente usado para este propósito. O teste F compara as razões das variâncias amostrais de duas populações independentes e assume que as populações seguem uma distribuição normal.

Passos

1. Formule as hipóteses:

Hipótese nula (H_0): As variâncias das duas populações são iguais.

Hipótese alternativa (H_a): As variâncias das duas populações são diferentes. Colete amostras aleatórias de cada uma das duas populações.

2. Calcule as variâncias amostrais de cada uma das amostras.
3. Determine a estatística de teste F calculando a razão das variâncias amostrais, dividindo a maior variância pela menor.

4. Escolha um nível de significancia, tipicamente $\alpha = 0.05$.
5. Determine os valores criticos para o teste F com base nos graus de liberdade das duas amostras.
6. Faça uma decisão comparando o valor F calculado com os valores críticos. Se o valor F calculado estiver fora dos limites definidos pelos valores críticos, rejeite a hipótese nula.

```
# Definir os dados da amostra
# amostra1 <- c(12, 15, 14, 16, 15, 13, 14, 15, 13, 14)
amostra1 <- c(14, 16, 17, 18, 19, 19, 15, 13, 11, 12)
# amostra2 <- c(11, 13, 14, 13, 16, 14, 13, 15, 12, 13)
amostra2 <- c(12, 11, 14, 17, 15, 13, 19, 15, 13, 14)

# Calcular as variâncias amostrais
var1 <- var(amostra1)
var2 <- var(amostra2)

# Calcular a estatística de teste F
F_valor <- var1 / var2

# Obter os graus de liberdade
n1 <- length(amostra1) - 1
n2 <- length(amostra2) - 1

# Calcular os valores criticos
F_critico1 <- qf(0.025, n1, n2) # limite inferior (use 0.025 para um teste bilateral com alpha = 0.05)
F_critico2 <- qf(0.975, n1, n2) # limite superior

# Tomar uma decisao
cat("Valor F calculado: ", F_valor, "\n")

## Valor F calculado: 1.48503

cat("Valores criticos do teste F: [", F_critico1, ",", F_critico2, "]\n")

## Valores criticos do teste F: [ 0.2483859 , 4.025994 ]

if(F_valor < F_critico1 | F_valor > F_critico2) {
  cat("Rejeitamos a hipotese nula: as variancias sao diferentes.\n")
} else {
  cat("Nao rejeitamos a hipotese nula: \n nao ha evidencias suficientes para afirmar que as variancias ")
}
```

```
## Nao rejeitamos a hipotese nula:
## nao ha evidencias suficientes para afirmar que as variancias sao diferentes.
```

Hipotese nula (H0): As variancias das duas populacoes sao diferentes

Parte 4 - Teste de Brown-Forsythe-Levene

Enunciado

Suponha que você é um analista de dados em uma empresa que quer verificar se três máquinas diferentes produzem componentes com variações consistentes em suas medidas. Para isso, você coleta uma amostra

de 10 medidas de cada máquina. Sua tarefa é aplicar o teste de Brown-Forsythe-Levene para verificar se há diferenças significativas nas variâncias das medidas entre as três máquinas. Use um nível de significância de 0.05.

```
# Instalar pacotes necessários se ainda não estiverem instalados
#if (!require(car)) install.packages("car")
#if (!require(dplyr)) install.packages("dplyr")

# Carregar os pacotes necessários
library(car)
```

```
## Carregando pacotes exigidos: carData
```

```
library(dplyr)
```

```
##
```

```
## Anexando pacote: 'dplyr'
```

```
## O seguinte objeto é mascarado por 'package:car':
```

```
##
```

```
## recode
```

```
## Os seguintes objetos são mascarados por 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## Os seguintes objetos são mascarados por 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
# Simular dados de amostra para as três máquinas
```

```
set.seed(123) # Para reprodutibilidade
```

```
maquina1 <- rnorm(15, mean = 50, sd = 8)
```

```
maquina2 <- rnorm(15, mean = 50, sd = 4)
```

```
maquina3 <- rnorm(15, mean = 50, sd = 3)
```

```
# Combinar os dados em um data frame
```

```
dados <- data.frame(
  Medidas = c(maquina1, maquina2, maquina3),
  Maquina = factor(rep(c("Maquina1", "Maquina2", "Maquina3"), each = 15))
)
```

```
# Aplicar o teste de Brown-Forsythe-Levene
```

```
resultado_bfl <- leveneTest(Medidas ~ Maquina, data = dados, center = median)
```

```
resultado_bfl
```

```
## Levene's Test for Homogeneity of Variance (center = median)
```

```
## Df F value Pr(>F)
```

```
## group 2 3.7226 0.03248 *
```

```
## 42
```

```
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```