Week\_06

230701195

2025-03-29

### F - test

# Problem 1: F-Test for comparing Variances (Two-Tailed Test)  
  
# A study tests whether the variances of two manufacturing processes are the same. Sample 1 has 25 observations with variance s1^2=100, and sample 2 has 30 observations with variance s2^2 = 80. Test at α = 0.05.  
  
# Given values  
s1\_squared <- 100  
s2\_squared <- 80  
n1 <- 25  
n2 <- 30  
alpha <- 0.05  
  
# Calculate the F-statistic  
F\_stat <- s1\_squared / s2\_squared  
  
# Degrees of freedom  
df1 <- n1 - 1  
df2 <- n2 - 1  
  
# Critical F value for two-tailed test at alpha = 0.05  
F\_critical <- qf(1 - alpha / 2, df1, df2)  
  
# Output results  
F\_stat

## [1] 1.25

F\_critical

## [1] 2.154006

# Decision  
if(F\_stat > F\_critical | F\_stat < 1 / F\_critical) {  
 print("Reject the null hypothesis")  
} else {  
 print("Fail to reject the null hypothesis")  
}

## [1] "Fail to reject the null hypothesis"

# Problem 2: F-Test for comparing Variances (Right-Tailed Test)  
  
# A company compares the variances in processing times between two production lines. Line 1 has 40 samples with variance s1^2 = 120, and Line 2 has 35 samples with variance s2^2 = 100. Test if Line 1 has a significantly greater variance than Line 2 at α = 0.01.  
  
# Given values  
s1\_squared <- 120  
s2\_squared <- 100  
n1 <- 40  
n2 <- 35  
alpha <- 0.01  
  
# Calculate the F-statistic  
F\_stat <- s1\_squared / s2\_squared  
  
# Degrees of freedom  
df1 <- n1 - 1  
df2 <- n2 - 1  
  
# Critical F value for right-tailed test at alpha = 0.01  
F\_critical <- qf(1 - alpha, df1, df2)  
  
# Output results  
F\_stat

## [1] 1.2

F\_critical

## [1] 2.218146

# Decision  
if(F\_stat > F\_critical) {  
 print("Reject the null hypothesis")  
} else {  
 print("Fail to reject the null hypothesis")  
}

## [1] "Fail to reject the null hypothesis"

# Problem 3: F-Test for comparing Variances (Left-Tailed Test)  
  
# A chemist compares the variability in pH levels between two chemical solutions. Solution 1 has 50 samples with variance s12=0.25s\_1^2 = 0.25s12=0.25, and Solution 2 has 60 samples with variance s22=0.20s\_2^2 = 0.20s22=0.20. Test if the variance of Solution 1 is smaller than Solution 2 at α = 0.05.  
  
s1\_squared <- 0.25  
s2\_squared <- 0.20  
n1 <- 50  
n2 <- 60  
alpha <- 0.05  
  
# Calculate the F-statistic  
F\_stat <- s1\_squared / s2\_squared  
  
# Degrees of freedom  
df1 <- n1 - 1  
df2 <- n2 - 1  
  
# Critical F value for left-tailed test at alpha = 0.05  
F\_critical <- qf(alpha, df1, df2)  
  
# Output results  
F\_stat

## [1] 1.25

F\_critical

## [1] 0.6318137

# Decision  
if(F\_stat < F\_critical) {  
 print("Reject the null hypothesis")  
} else {  
 print("Fail to reject the null hypothesis")  
}

## [1] "Fail to reject the null hypothesis"

### Chi-squared test

# Problem 1: Chi-Square Goodness of Fit Test  
  
# A die is rolled 60 times, and the number of times each number appeared is recorded as follows:   
# 1: 8,   
# 2: 12,   
# 3: 10,   
# 4: 11,   
# 5: 9,   
# 6: 10.  
# Test if the die is fair at α = 0.05.  
  
# Observed frequencies  
observed <- c(8, 12, 10, 11, 9, 10)  
  
# Expected frequencies (equal for a fair die)  
expected <- rep(10, 6) # 60 rolls, 6 outcomes, so expected is 60 / 6 = 10 for each outcome  
  
# Perform the chi-square goodness of fit test  
chisq\_test <- chisq.test(observed, p = rep(1/6, 6), rescale.p = TRUE)  
  
# Output results  
chisq\_test

##   
## Chi-squared test for given probabilities  
##   
## data: observed  
## X-squared = 1, df = 5, p-value = 0.9626

# Decision  
if(chisq\_test$p.value < 0.05) {  
 print("Reject the null hypothesis: The die is not fair.")  
} else {  
 print("Fail to reject the null hypothesis: The die is fair.")  
}

## [1] "Fail to reject the null hypothesis: The die is fair."

# Problem 2: Chi-Square Test for Independence (No Association)  
  
# A study investigates if there's a relationship between region (North, South, East, West) and product preference (A, B, C). The data is:  
  
# A B C  
# North 10 20 30  
# South 15 25 40  
# East 20 30 50  
# West 25 35 45  
  
# Test at α = 0.05.  
  
  
# Data  
data <- matrix(c(10, 20, 30, 15, 25, 40, 20, 30, 50, 25, 35, 45), nrow = 4, byrow = TRUE)  
  
# Perform the chi-square test for independence  
chisq\_test <- chisq.test(data)  
  
# Output results  
chisq\_test

##   
## Pearson's Chi-squared test  
##   
## data: data  
## X-squared = 2.1362, df = 6, p-value = 0.9068

# Decision  
if(chisq\_test$p.value < 0.05) {  
 print("Reject the null hypothesis: There is an association between region and product preference.")  
} else {  
 print("Fail to reject the null hypothesis: There is no association between region and product preference.")  
}

## [1] "Fail to reject the null hypothesis: There is no association between region and product preference."

# Problem 3: Chi-Square Goodness of Fit Test with Custom Expected Values  
  
# A bag contains colored marbles:  
# Red: 40, Blue: 35, Green: 25.  
# You expect the proportions to be 1:2:3 (Red:Blue:Green). Test if the distribution matches the expected proportions at α = 0.05.  
  
# Observed frequencies  
observed <- c(40, 35, 25)  
  
# Total number of marbles  
total <- sum(observed)  
  
# Expected proportions  
expected\_proportions <- c(1, 2, 3) / sum(c(1, 2, 3))  
  
# Expected frequencies  
expected <- expected\_proportions \* total  
  
# Perform the chi-square goodness of fit test  
chisq\_test <- chisq.test(observed, p = expected\_proportions, rescale.p = TRUE)  
  
# Output results  
chisq\_test

##   
## Chi-squared test for given probabilities  
##   
## data: observed  
## X-squared = 45.25, df = 2, p-value = 1.493e-10

# Decision  
if(chisq\_test$p.value < 0.05) {  
 print("Reject the null hypothesis: The distribution does not match the expected proportions.")  
} else {  
 print("Fail to reject the null hypothesis: The distribution matches the expected proportions.")  
}

## [1] "Reject the null hypothesis: The distribution does not match the expected proportions."