# F-test

Mengye Liu

4/13/2022

The Two-Sample F-Test tests the null hypothesis that two samples come from two independent populations having the equal variances.

Take two sided test as an example.

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad v.s. \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

1.$\mu_1$ and $\mu_2$ are given.

That is $X_i \sim N(\mu_1, \sigma_1^2), Y_j \sim N(\mu_2, \sigma_2^2)$ and $\mu_1$ and $\mu_2$ are known, $i = 1, 2, 3, \ldots, n; j = 1, 2, 3, \ldots, m$

Then the test statistics is

$$F_0 = \frac{S_1^2}{S_2^2} \sim F_{n,m}, \quad \text{under } H_0$$

where

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_1)^2$$

$$S_2^2 = \frac{1}{m}\sum_{j=1}^{m}(Y_j - \mu_2)^2$$

2. $\mu_1$ and $\mu_2$ are not given.

Then the test statistics is

$$F_0 = \frac{S_1^2}{S_2^2} \sim F_{n-1,m-1}, \quad \text{under } H_0$$

where $S_1^2$ and $S_2^2$ are the sample variance for $X$ and $Y$ respectively.

That is,

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \bar{Y})^2$$

Assume we already have the normality assumption. Then test(two sided F test)

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad v.s. \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

```r
# import data (a little bit different from 2.25)
Type1 = c(65,82,82,67,57,59,66,75,77)
Type2 = c(64,56,71,69,83,74,59,82,65)
S1 = var(Type1)
S2 = var(Type2)
n1 = length(Type1)
n2 = length(Type2)
a = S1/S2
b = S2/S1
result = data.frame(c(a,b),c(a,b)>1)
colnames(result) = c('ratio','>1')
result
```

```
##      ratio    >1
## 1 0.997789 FALSE
## 2 1.002216  TRUE
```

Then get P-value

```r
pf(a, n1-1, n2-1) + pf(b, n1-1, n2-1, lower.tail = FALSE)
```

```
## [1] 0.9975791
```

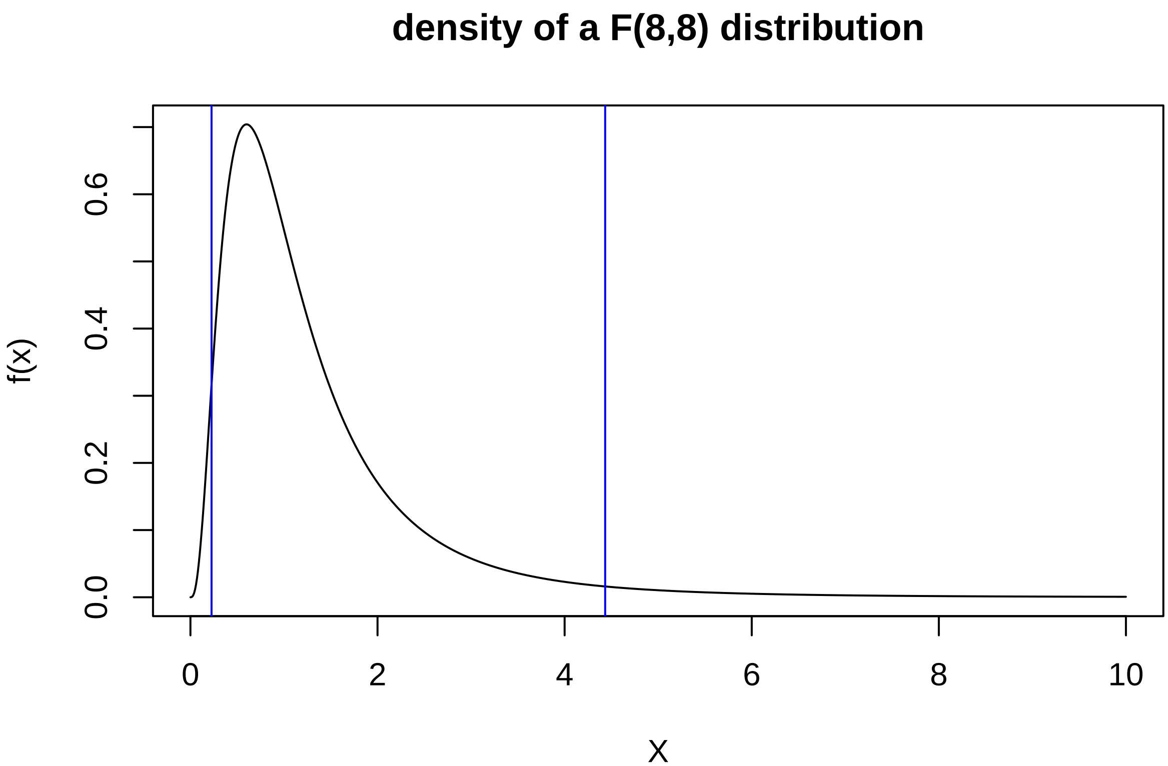
Using the critical value

```r
bounds = c(qf(0.975, n1-1, n2-1), qf(0.025, n1-1, n2-1))
names(bounds) = c('upper','lower')
bounds
```

```
##     upper     lower
## 4.4332599 0.2255676
```

Check $S_1^2/S_2^2 >$ upper or $<$ lower. If yes, then reject null hypothesis.

```r
# Plot a F distribution
input = seq(0, 10, 1/100)
y = df(input, df1=n1-1, df2=n2-1)
plot(input, y, xlab="X",ylab="f(x)", main = "density of a F(8,8) distribution", type="l")
abline(v=qf(c(0.025, 0.975), df1=n1-1, df2=n2-1), col=c("blue","blue"))
```




```
var.test(Type1, Type2, alternative = "two.sided")
```

```
##
##  F test to compare two variances
##
## data:  Type1 and Type2
## F = 0.99779, num df = 8, denom df = 8, p-value = 0.9976
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2250689 4.4234580
## sample estimates:
## ratio of variances
##           0.997789
```

For one sided F-test,

- $H_a : \sigma_1^2 > \sigma_2^2$ then check $S_1^2/S_2^2 > F_{1-\alpha,n_1-1,n_2-1}$. Or $p = P(F_{n_1-1,n_2-1} > S_1^2/S_2^2) < \alpha$.
- $H_a : \sigma_1^2 < \sigma_2^2$ then check $S_1^2/S_2^2 < F_{\alpha,n_1-1,n_2-1}$.Or $p = P(F_{n_1-1,n_2-1} < S_1^2/S_2^2) < \alpha$.