6. F 検定

（学習内容）

・F 検定の原理を理解する

・F 検定の方法を学ぶ

6.1 ２つの母分散の比に関する検定

（問題）

2 台の機械 A,B の加工精度を比べる。寸法 30.0cm の部品について A から n=11 個、B から m=10 個のデータをサンプリングした。

A：28.9　29.5　30.5　30.4　30.1　30.7　30.9　30.2　29.8　29.3　28.8　(cm)

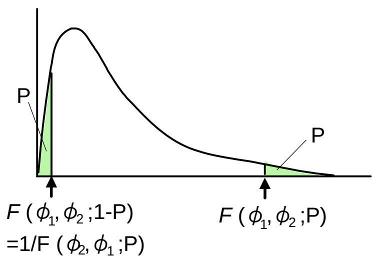
B：29.6　27.9　30.6　29.2　31.3　32.0　28.4　28.0　30.2　31.5　(cm)

機械 A の母集団分布が N(μ_1, σ_1^2)、機械 B の母集団分布が N(μ_2, σ_2^2)に従うとき、σ_1^2 と σ_2^2 の値は等しいか。μ_1, μ_2 は未知。

> $x_{11}, x_{12}, \cdots, x_{1n}$ が互いに独立に N(μ_1, σ_1^2) に従い、
> $x_{21}, x_{22}, \cdots, x_{2m}$ が互いに独立に N(μ_2, σ_2^2) に従うとき、
> $F = \dfrac{V_1/\sigma_1^2}{V_2/\sigma_2^2}$　は自由度(n−1, m−1)の F 分布に従う

（1）F 分布

$$f_{\phi_1,\phi_2}(x) = \frac{\Gamma(\frac{\phi_1+\phi_2}{2})}{\Gamma(\frac{\phi_1}{2})\cdot\Gamma(\frac{\phi_2}{2})}\cdot(\frac{\phi_1}{\phi_2})^{\frac{\phi_1}{2}} x^{\frac{\phi_1-1}{2}} (1+\frac{\phi_1}{\phi_2}x)^{-\frac{\phi_1+\phi_2}{2}}$$



F 分布のグラフ：左右非対称

F 分布表：自由度 ϕ_1, ϕ_2 と確率 P から F(ϕ_1, ϕ_2;P)を求める

　　P=Pr{F≧F(ϕ_1, ϕ_2;P)}

（性質）F(ϕ_1, ϕ_2;1−P)=1/ F(ϕ_2, ϕ_1;P)

P=Pr{F≦F(ϕ_1, ϕ_2;1−P)}=Pr{1/F≧1/F(ϕ_1, ϕ_2;1−P)}

一方、1/F は自由度(m−1, n−1)の F 分布に従うため P=Pr{1/F≧F(ϕ_2, ϕ_1;P)}

これから、1/F(ϕ_1, ϕ_2;1−P)= F(ϕ_2, ϕ_1;P)

よって　　F(ϕ_1, ϕ_2;1−P)= 1/F(ϕ_2, ϕ_1;P)

（例題１）

次の値を求めよ

1) F(12,10; 0.05)　　　2) F(12,10; 0.95)

（解）

1）$F(12,10; 0.05)=2.91$

2）$F(12,10; 0.95)=1/F(10,12; 0.05)=1/2.75=0.364$

(2) χ^2分布とＦ分布の関係

（定義）

「標準正規分布に従う n 個の独立な変数を $u_1, u_2, \cdots u_n$ とするとき、これらの２乗和

$$x = u_1^2 + u_2^2 \cdots + u_n^2$$

は自由度 n の χ^2分布に従う」

「x_i が正規分布 $N(\mu, \sigma^2)$ に従うとき、$u_i=(x_i-\mu)/\sigma$ の標準化を行うと

$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$ は自由度 n の χ^2分布に従う」

μは未知であるため、$\bar{x}$ で置き換えると自由度が１小さくなり

「$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2=\dfrac{S}{\sigma^2}$ は自由度 n-1 の χ^2分布に従う」

→　分散の検定にはχ^2分布を使用

（定義）

「自由度 n_1 のχ^2分布に従う変数を u_1、自由度 n_2 のχ^2分布に従う変数を u_2 とすると、変数

$$x = (u_1/n_1)/(u_2/n_2)$$

は自由度(n_1, n_2)のＦ分布に従う」

χ^2分布に従う変数として、S_1/σ_1^2（自由度 n-1），S_2/σ_2^2（自由度 m-1）を考えると

「$\dfrac{S_1/\sigma_1^2(n-1)}{S_2/\sigma_2^2(m-1)}=\dfrac{V_1/\sigma_1^2}{V_2/\sigma_2^2}$ は自由度(n-1, m-1）のＦ分布に従う」

→　分散の比の検定にはＦ分布を使用

検定統計量は　V_1/V_2

(3) 母分散の比の検定手順

1）帰無仮説 H_0：$\sigma_1^2 = \sigma_2^2$　　　対立仮説 H_1：$\sigma_1^2 \neq \sigma_2^2$

2）有意水準αを定める　$\alpha=0.05$

3）棄却域Ｒを求める　$F_0 \leqq F(10,9;1-0.025)=1/F(9,10;0.025)=1/3.78=0.265$

$F_0 \geqq F(10,9;0.025)=3.96$

4）データ $x_{11}, x_{12}, \cdots, x_{1n}$、$x_{21}, x_{22}, \cdots, x_{2m}$から検定統計量 F_0の値を計算する

$S_1=\Sigma(x_{1i}-\bar{x}_1)^2=5.116$

$S_2=\Sigma(x_{2i}-\bar{x}_2)^2=19.94$

$V_1=S_1/(n-1)=5.116/10=0.5116$

$V_2=S_2/(m-1)=19.94/9=2.216$

$F_0=V_1/V_2=0.5116/2.216=0.2309$　　　F_0は自由度(10,9)のＦ分布に従う

5) F_0が棄却域にあれば有意と判定し、H_0を棄却

$F_0=0.2309 \leqq F(10,9;1-0.025)=0.265$　となりH_0は棄却

(4) 母分散の比の推定手順

点推定：　$\sigma_1^2/\sigma_2^2=V_1/V_2=0.5116/2.216=0.231$

区間推定：

$F(\phi_1, \phi_2;1-\alpha/2) \leqq (V_1/\sigma_1^2)/(V_2/\sigma_2^2) \leqq F(\phi_1, \phi_2;\alpha/2)$

$V_1/V_2 \cdot 1/F(\phi_1, \phi_2;\alpha/2) \leqq \sigma_1^2/\sigma_2^2 \leqq V_1/V_2 \cdot 1/F(\phi_1, \phi_2;1-\alpha/2)$

$V_1/V_2 \cdot 1/F(\phi_1, \phi_2;\alpha/2) \leqq \sigma_1^2/\sigma_2^2 \leqq V_1/V_2 \cdot F(\phi_2, \phi_1;\alpha/2)$

$0.5116/2.216/F(10,9;0.025) \leqq \sigma_1^2/\sigma_2^2 \leqq 0.5116/2.216 \cdot F(9,10;0.025)$

$0.0583 \leqq \sigma_1^2/\sigma_2^2 \leqq 0.8727$

6.2 t検定の予備検定

ｔ検定では、等分散を仮定する場合（Studentのｔ検定）と仮定しない場合（Welchのｔ検定）で以降の処理が異なる

ｔ検定の予備検定として、Ｆ検定を用いる

この場合、有意水準は20% を用いる

（仮説の検証ではなく分類に用いるため、5%では厳しすぎる）

（問題）

Aから$n=10$個、Bから$m=9$個のデータをサンプリングした。

A：78　80　79　83　82　85　78　74　76　84

B：81　84　82　88　86　83　78　84　89

Aの母集団分布が$N(\mu_1, \sigma_1^2)$、Bの母集団分布が$N(\mu_2, \sigma_2^2)$に従うとき、μ_1とμ_2は等しいか。

1) 帰無仮説H_0：$\sigma_1^2=\sigma_2^2$　　　対立仮説H_1：$\sigma_1^2 \neq \sigma_2^2$

2) 有意水準α　$\alpha=0.20$

3) 棄却域R　　$F_0 \leqq F(9,8;1-0.10)=1/F(8,9;0.10)=1/2.47=0.405$

$F_0 \geqq F(9,8;0.10)=2.56$

4) データ$x_{11}, x_{12}, \cdots, x_{1n}$、$x_{21}, x_{22}, \cdots, x_{2m}$から検定統計量$F_0$の値を計算する

$S_1=\Sigma(x_{1i}-\bar{x}_1)^2=114.9$

$S_2=\Sigma(x_{2i}-\bar{x}_2)^2=94.89$

$V_1=S_1/(n-1)=114.9/9=12.77$

$V_2=S_2/(m-1)=94.89/8=11.86$

$F_0=V_1/V_2=12.77/11.86=1.08$　　　F_0は自由度(9,8)のＦ分布に従う

5）F_0が棄却域にあれば有意と判定し、H_0を棄却

$F_0=1.08 \leqq F(9,8;0.10)=2.56$　なので有意でない

→　$\sigma_1^2 = \sigma_2^2$　として母平均の差の検定を行う