

III-2-7. F 分布

F 分布は分散の比を確率変数とした確率分布です。t 分布は正規分布する確率と χ^2 分布分布する 確率の積の確率分布でしたが、F 分布は χ^2 分布分布する 2 つの確率の積の分布です。比の分布と いう考え方は χ^2 分布分布のところにも出てきました。今までやってきた、確率の組み合わせと同じ作業をします。違っているのは分散比 (F) が新たな確率変数になるということです。具体的には、 χ^2 分布分布する 2 つの変数があり、それらの分散を、 $\frac{z_1}{n_1}$ 、

$\frac{z_2}{n_2}$ としたとき (n は 自由度、データ数ではないので注意)、

$$F = \frac{\frac{z_1}{n_1}}{\frac{z_2}{n_2}}$$

を F 比として、その確率分布を考えるということです。確率変動するものと確率変動する物の比の確率というのは少し悩ましいのですが、両者が独立していて、一方が何らかの値をとった時に他方が独立的に変動するのだから、両者の確率の積で変動すると考えれば良いでしょう。それぞれは、ある期待値からのへだたりだから、 χ^2 分布します。

それぞれの確率を

$$P(z_1) = \frac{\chi^2_n \left(\frac{n_1-1}{2}\right) e^{-\frac{\chi^2_{n_1}}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n_1}{2}\right)}$$

$$P(z_2) = \frac{\chi^2_n \left(\frac{n_2-1}{2}\right) e^{-\frac{\chi^2_{n_2}}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n_2}{2}\right)}$$

$$P(F) = P(z_1)P(z_2)$$

ということで、F で積分できるように F と直交する平面を考えて、その面積を求めれば良い ということで、今までやっていたことと同じです。f と直交する平面をどう考えるのがポイントです。求積法としての積分の考え方に慣れていればくれば、こういう考え方が自然にできます。F 比を f という変数うにして

$$f = \frac{\frac{z_1}{n_1}}{\frac{z_2}{n_2}} = \frac{z_1}{n_1} \frac{n_2}{z_2}$$

$$z_1 n_2 = f z$$

$$z_2 n_1 = z$$

と置きます。これなら、f の定義そのままだし、f と z は互いに独立しています。

$0 \leq z_1 \leq \infty, 0 \leq z_2 \leq \infty, 0 \leq f \leq \infty, 0 \leq z \leq \infty$ なので確率総和 = 1 の式は

$$\int_0^\infty P(f)df = \int_0^\infty \int_0^\infty P(z_1)P(z_2)dz_1dz_2 = 1$$

となります。

$$\iint P(z_1)P(z_2)dz_1dz_2$$

積分変数を変換します。

$$\frac{dz_1}{df} = \frac{z}{n_2}$$

$$\frac{dz_1}{dz} = \frac{f}{n_2}$$

$$\frac{dz_2}{df} = 0$$

$$\frac{dz_2}{dz} = \frac{1}{n_1}$$

だからヤコビアンは

$$J(z_1, z_2/f, z) = \begin{vmatrix} \frac{z}{n_2} & \frac{f}{n_2} \\ 0 & \frac{1}{n_1} \end{vmatrix} = \frac{z}{n_1 n_2}$$

$$\begin{aligned} \iint P(z_1)P(z_2)dz_1dz_2 &= \iint P(f)P(z) \cdot \frac{z}{n_1 n_2} dz_1 dz_2 \\ &= \iint \frac{z_1^{\frac{n_1}{2}-1}}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{z_1}{2}} \frac{z_2^{\frac{n_2}{2}-1}}{2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{z_2}{2}} \cdot \frac{z}{n_1 n_2} dz_1 dz_2 \\ &= \iint \frac{\left(\frac{fz}{n_2}\right)^{\frac{n_1}{2}-1}}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{fz}{2}} \frac{\left(\frac{z}{n_1}\right)^{\frac{n_2}{2}-1}}{2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{z}{2}} \cdot \frac{z}{n_1 n_2} dz df \\ &= \int \left(\int \frac{\left(\frac{fz}{n_2}\right)^{\frac{n_1}{2}-1}}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} e^{-\frac{fz}{2}} \frac{\left(\frac{z}{n_1}\right)^{\frac{n_2}{2}-1}}{2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} e^{-\frac{z}{2}} \cdot \frac{z}{n_1 n_2} dz \right) df \\ &= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int \left(\frac{z}{n_2}\right)^{\frac{n_1}{2}-1} e^{-\frac{fz}{2}} \left(\frac{z}{n_1}\right)^{\frac{n_2}{2}-1} e^{-\frac{z}{2}} \cdot \frac{z}{n_1 n_2} dz \right) df \\ &= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int \left(\frac{\frac{n_1}{2}-1}{\frac{n_1}{2}}\right) e^{-\frac{fz}{2}} \left(\frac{\frac{n_2}{2}-1}{\frac{n_2}{2}}\right) e^{-\frac{z}{2}} \cdot \frac{z}{n_1 n_2} dz \right) df \end{aligned}$$

$$\begin{aligned}
&= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} n_2^{\frac{n_1}{2}} n_2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int z^{\frac{n_1}{2}-1} z^{\frac{n_1}{2}-1} e^{\frac{fz}{2}} e^{-\frac{z}{n_2}} \cdot z dz \right) df \\
&= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int z^{\frac{n_1+n_2}{2}-2} e^{-\frac{(fz+\frac{z}{n_1})}{2}} \cdot z dz \right) df \\
&= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int z^{\frac{n_1+n_2}{2}-1} e^{-\frac{(f+\frac{1}{n_1})z}{2}} \cdot dz \right) df
\end{aligned}$$

$w = \left(\frac{f}{n_2} + \frac{1}{n_1}\right) \frac{z}{2}$ とおきかえると、 $z = \frac{2}{\left(\frac{f}{n_2} + \frac{1}{n_1}\right)} w$ だから

$$\begin{aligned}
&= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int \left(\frac{2}{\left(\frac{f}{n_2} + \frac{1}{n_1}\right)} w \right)^{\frac{n_1+n_2}{2}-1} e^{-w} \frac{2}{\left(\frac{f}{n_2} + \frac{1}{n_1}\right)} \cdot dw \right) df \\
&= \int \left(\frac{(f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int \left(\frac{2}{\left(\frac{f}{n_2} + \frac{1}{n_1}\right)} w \right)^{\frac{n_1+n_2}{2}} e^{-w} \cdot dw \right) df \\
&= \int \left(\frac{2^{\frac{n_1+n_2}{2}} (f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right)} \int \left(\frac{1}{\left(\frac{f}{n_2} + \frac{1}{n_1}\right)} w \right)^{\frac{n_1+n_2}{2}} e^{-w} \cdot dw \right) df \\
&= \int \left(\frac{2^{\frac{n_1+n_2}{2}} (f)^{\frac{1}{n_1}-1}}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right) \left(\frac{f}{n_2} + \frac{1}{n_1}\right)^{\frac{n_1+n_2}{2}}} \int w^{\frac{n_1+n_2}{2}} e^{-w} \cdot dw \right) df
\end{aligned}$$

となりますが、 w の関数で表された

$$\int w^{\frac{n_1+n_2}{2}} e^{-w} \cdot dw$$

だけを無限積分すると、以下のような、ガンマ関数です、

$$\int_0^\infty w^{\frac{n_1+n_2}{2}} e^{-w} \cdot dw = \Gamma\left(\frac{n_1+n_2}{2}\right)$$

ですから、

$$\int P(f)df = \int \left(\frac{2^{\frac{n_1+n_2}{2}} (f)^{\frac{1}{n_1}-1} \Gamma\left(\frac{n_1+n_2}{2}\right)}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right) \left(\frac{f}{n_2} + \frac{1}{n_1}\right)^{\frac{n_1+n_2}{2}}} \right) df$$

となりますが、 $P(f)$ を $f=0$ から $f=f$ を積分したものを、 f で微分したものが、 f の確率密度関数ですから、

$$P(f) = \frac{2^{\frac{n_1+n_2}{2}} (f)^{\frac{1}{n_1}-1} \Gamma\left(\frac{n_1+n_2}{2}\right)}{2^{\frac{n_1}{2}} \left(\frac{1}{n_2}\right)^{\frac{1}{n_1}} \left(\frac{1}{n_1}\right)^{\frac{1}{n_2}} \Gamma\left(\frac{n_1}{2}\right) 2^{\frac{n_2}{2}} \Gamma\left(\frac{n_2}{2}\right) \left(\frac{f}{n_2} + \frac{1}{n_1}\right)^{\frac{n_1+n_2}{2}}}$$

となります。この式を整理すれば、もっと簡単に書き表せますが、そこまでやるのはこの講義の主旨ではないので、結論として、 β 関数を使って書き表したFの確率密度関数を示します。

$$P(f) = \frac{n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{\beta(n_1, n_2)} \cdot \frac{f^{\frac{1}{n_1}-1}}{(n_1 f + n_2)^{\frac{n_1+n_2}{2}}}$$

こんな面倒なことをしてF関数を導出する必要は全くありません。覚えたとしても、一生使うことはないでしょう。ここでわざわざ計算過程を描いたのは、途中で Γ 関数が出てきて、答えが出てしまうところが面白かったからです。知識として、頭にとどめておかなくてはならないのは、**F比はデータから求めた分散の比で、分母も分子も χ^2 分布するという**ことです。ですから、**F比の確率密度は、分母・分子の確率の積になります。**

また、最後に示した式からわかるように、F分布は、**分母分子の自由度で変わります。**

F比を使った検定は、結構たくさんあります。もっとも、重要なのは、分散はいくつかの部分分散に分けることが出来て、それを使って、多要因分散分析が出来ます。そうした、王王的な展開を考えると、現時点で大切なことは、F分布が何の確率分布なのか、どのように牡蛎率密度関数が作られているのかを理解することです。