2.4　いろいろな確率分布　(教科書 p.86)

正規分布から導かれるいくつかの確率分布で、区間推定や検定に用いられるものを紹介する。

定義 1. **(χ^2 (カイ 2 乗) 分布)** $X_1, X_2, \ldots, X_n$ が独立で、いずれも標準正規分布 $N(0,1)$ に従うとき、

$$Y = {X_1}^2 + {X_2}^2 + \cdots + {X_n}^2$$

は、**自由度 n の χ^2 分布**に従うという。

2. **(t 分布)** Z と Y は独立で、Z は標準正規分布 $N(0,1)$ に、Y は自由度 n の χ^2 分布に従うとき、

$$T = \frac{Z}{\sqrt{Y/n}}$$

は、**自由度 n の t 分布**に従うという。

3. **(F 分布)** X と Y は独立で、X は自由度 m の χ^2 分布に、Y は自由度 n の χ^2 分布に従うとき、

$$W = \frac{X}{m}\bigg/\frac{Y}{n}$$

は、**自由度 (m,n) の F 分布**に従うという。

注意 上記の分布は連続型でその密度関数は教科書 p.133 に述べられている。

定理 A $X_1, X_2, \ldots, X_n$ が独立で、同一の正規分布 $N(\mu, \sigma^2)$ に従うとき、次が成立する。

(1) 標本平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ は正規分布 $N\Big(\mu, \dfrac{\sigma^2}{n}\Big)$ に従う。

(2) 不偏分散 $U^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$ について、$\dfrac{(n-1)U^2}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従う。

(3) $\overline{X}$ と $\dfrac{(n-1)U^2}{\sigma^2}$ は独立。

この定理の証明はこの授業の範囲を超すが、おおよそ以下のように証明できる: $Z_i = \dfrac{X_i - \mu}{\sigma}$ と標準化すると、$Z_1, Z_2, \ldots, Z_n$ は独立で $N(0,1)$ に従う。このとき、$\overline{X} = \sigma\overline{Z} + \mu$, $\dfrac{(n-1)U^2}{\sigma^2} = \displaystyle\sum_{i=1}^{n}(Z_i - \overline{Z})^2$ となるが、$Z_1, Z_2, \ldots, Z_n$ から導かれる独立に $N(0,1)$ に従う確率変数 $S_1, S_2, \ldots, S_n$ があって、

$$\overline{Z} = \frac{1}{\sqrt{n}}S_1, \qquad \sum_{i=1}^{n}(Z_i - \overline{Z})^2 = {S_2}^2 + {S_3}^2 + \cdots + {S_n}^2$$

と表せる。このことにより主張 (1)–(3) は従う。　□

系 1 $X_1, X_2, \ldots, X_n$ を正規母集団 $N(\mu, \sigma^2)$ からの大きさ n の無作為標本とする。

(1) $\dfrac{(n-1)U^2}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従う。
注意: これは教科書 p.101 **1.4 母分散の区間推定**, p.115 **2.4 母分散の検定** に応用される。

(2) $T = \dfrac{\overline{X} - \mu}{\sqrt{U^2/n}}$ とおくと、T は自由度 $n-1$ の t 分布に従う。
注意: これは教科書 p.99 **1.3 母平均の区間推定 (2)**, p.113 **2.3 母平均の検定 (2)** に応用される。

証明: (1) は定理 A (2) そのもの。(2) について $Z = \dfrac{\overline{X} - \mu}{\sqrt{\sigma^2/n}}$, $Y = \dfrac{(n-1)U^2}{\sigma^2}$ とおくと、定理 A より Z と Y は独立で、Z は標準正規分布 $N(0,1)$ に、Y は自由度 $n-1$ の χ^2 分布に従う。ここで、

$$\frac{Z}{\sqrt{Y/(n-1)}} = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}}\Bigg/\sqrt{\frac{(n-1)U^2}{\sigma^2(n-1)}} = \frac{\overline{X} - \mu}{\sqrt{U^2/n}} = T.$$

定義より左辺は自由度 $n-1$ の t 分布に従うから、T も自由度 $n-1$ の t 分布に従うことがわかる。　□

系 2 正規母集団 $N(\mu_1, \sigma^2)$ から大きさ m の、$N(\mu_2, \sigma^2)$ から大きさ n の無作為標本をとり、その標本平均、不偏分散をそれぞれ $\overline{X}$, ${U_1}^2$ と $\overline{Y}$, ${U_2}^2$ とする。

(1) $\dfrac{{U_1}^2}{{U_2}^2}$ は自由度 $(m-1, n-1)$ の F 分布に従う。
注意: これは教科書 p.117 **2.5 等分散の検定** に応用される。

(2) $T = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{U^2(1/m + 1/n)}}$, $U^2 = \dfrac{(m-1){U_1}^2 + (n-1){U_2}^2}{m+n-2}$ とおくと、T は自由度 $m+n-2$ の t 分布に従う。 注意: これは教科書 p.119 **2.6 母平均の差の検定** に応用される。

証明: (1) は ${U_1}^2, {U_2}^2$ が独立だから、定義と定理 A(2) から示される。(2) は $\overline{X} - \overline{Y}$ が $N\Big(\mu_1 - \mu_2, \dfrac{\sigma^2}{m} + \dfrac{\sigma^2}{n}\Big)$ に、$\dfrac{(m+n-2)U^2}{\sigma^2}$ が自由度 $m+n-2$ の χ^2 分布に従うことに注意すれば、系 1(2) と同様に示される。 □

1.3 母平均の区間推定 (2) (教科書 p.99)

系 1(2) を用いて、母平均の区間推定で、母分散が未知の場合を導く。

$0 < \alpha < 1$ である値 α と自由度 n の t 分布に従う確率変数 T について $P(T \geq k) = \alpha$ を満たす k の値を $t_n(\alpha)$ と書き、**t 分布の上側 α 点**という。(教科書 p.87 のグラフを参照のこと。) その値は教科書 p.168 の t 分布表から読み取る。

定理 (母平均の区間推定 (母分散が未知の場合)) 正規母集団 $N(\mu, \sigma^2)$ から大きさ n の無作為標本の標本平均と不偏分散の実現値をそれぞれ $\overline{x}$, u^2 とすると、母平均 μ の $100(1-\alpha)\%$ 信頼区間は

$$\overline{x} - t_{n-1}(\alpha/2)\sqrt{\frac{u^2}{n}} \leq \mu \leq \overline{x} + t_{n-1}(\alpha/2)\sqrt{\frac{u^2}{n}}. \tag{1}$$

証明: 正規母集団 $N(\mu, \sigma^2)$ から大きさ n の無作為標本の標本平均を $\overline{X}$, 不偏分散を U^2 とすると、$T = \dfrac{\overline{X} - \mu}{\sqrt{U^2/n}}$ は自由度 $n-1$ の t 分布に従うので、

$$P\left(-t_{n-1}(\alpha/2) \leq \frac{\overline{X} - \mu}{\sqrt{U^2/n}} \leq t_{n-1}(\alpha/2)\right) = 1 - \alpha.$$

括弧内の不等式を μ について解くと、$\overline{X} - t_{n-1}(\alpha/2)\sqrt{\dfrac{U^2}{n}} \leq \mu \leq \overline{X} + t_{n-1}(\alpha/2)\sqrt{\dfrac{U^2}{n}}$ となるが、これは μ がこの区間に含まれる確率が $1-\alpha$ であることを示している。この $\overline{X}$, U^2 にその実現値 $\overline{x}$, u^2 を代入することで (1) 式を得る。 □

例題 1 (問題文は教科書 p.100 を見よ。)

解: t 分布表から $t_{7-1}(0.025) = 2.447$ より

$$\overline{x} \pm t_{n-1}(\alpha/2)\sqrt{\frac{u^2}{n}} = 11.12 \pm 2.447\sqrt{\frac{7.527}{7}} = 11.12 \pm 2.537\cdots = \left\{ \begin{array}{r} 13.657\cdots \\ 8.582\cdots \end{array} \right.$$

従って、$8.58 \leq \mu \leq 13.66$. □

§ 1 いろいろな検定, § 3 回帰分析 (教科書 p.125)

この教科書では 5 章に χ^2 分布を用いる **1.1 適合度の検定**, **1.2 独立性の検定** が、さらに §3 で**回帰分析** (回帰直線に関する推測統計) についてその概略が述べられています。興味のある方は必要に応じて文献を参照しつつ勉強してください。その証明まで書かれている参考文献としては、例えば、稲垣宣生 著 数理統計学 裳華房, 2003 がありますが、その本に書かれている内容は数理科学科で 3 年次、4 年次に学ぶような内容に相当します。また、実際のデータに対し統計解析を行う際は煩雑な計算が必要となります。教科書では R という統計解析のためのフリーソフトウェアを用いる方法が紹介されています。