

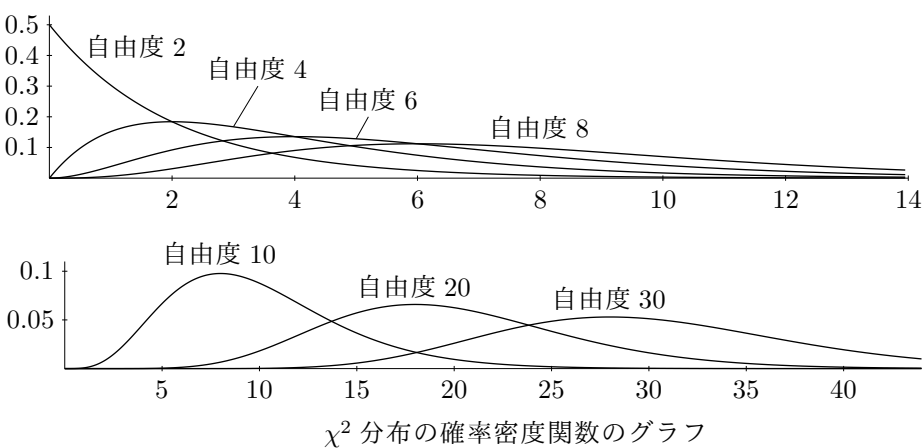
§ 9.1 χ^2 分布

正の自然数 n 及び標準正規分布 $N(0,1)$ に従う互いに独立な n 個の確率変数 $Z_1, Z_2, Z_3, \dots, Z_n$ に対して、確率変数

$$Z_1^2 + Z_2^2 + Z_3^2 + \dots + Z_n^2 = \sum_{k=1}^n Z_k^2$$

が従う確率分布を自由度 n の χ^2 分布（カイ 2 乗分布）という。

χ^2 分布の確率密度関数のグラフを示す。



[定理 9.1] 2 以上の自然数 n に対して n 個の確率変数 $X_1, X_2, X_3, \dots, X_n$ は総て同じ正規分布に従い、互いに独立であるとする。それらの標準偏差を σ とおく。 $X_1, X_2, X_3, \dots, X_n$ の平均を表す確率変数 $\overline{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 及び $X_1, X_2, X_3, \dots, X_n$ の偏差平方和を表す確率変数 $S = \sum_{k=1}^n (X_k - \overline{X})^2$ に対して、確率変数 $\frac{S}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従う。

χ^2 分布を用いる検定が幾つかあるが、本プリントではそのような検定について触れない。

——— 定理の説明

定理 9.1 を説明する。 μ と σ とは定数で $\sigma > 0$ とする。 2 以上の自然数 n に対して、 n 個の確率変数 $X_1, X_2, X_3, \dots, X_n$ は正規分布 $N(\mu, \sigma^2)$ に従い互いに独立であるとする。 n 以下の正の自然数 k に対して、 k 個の確率変数 $X_1, X_2, X_3, \dots, X_k$ の偏差平方和を表す確率変数を S_k とおく：

$$\begin{aligned} S_k &= \sum_{j=1}^k \left(X_j - \frac{1}{k} \sum_{i=1}^k X_i \right)^2 = \sum_{j=1}^k \left\{ X_j^2 - \frac{2X_j}{k} \sum_{i=1}^k X_i + \frac{1}{k^2} \left(\sum_{i=1}^k X_i \right)^2 \right\} \\ &= \sum_{j=1}^k X_j^2 - \frac{2}{k} \sum_{j=1}^k X_j \sum_{i=1}^k X_i + \frac{1}{k^2} \sum_{j=1}^k \left(\sum_{i=1}^k X_i \right)^2 \\ &= \sum_{j=1}^k X_j^2 - \frac{2}{k} \left(\sum_{i=1}^k X_i \right)^2 + \frac{1}{k} \left(\sum_{i=1}^k X_i \right)^2 = \sum_{i=1}^k X_i^2 - \frac{1}{k} \left(\sum_{i=1}^k X_i \right)^2 . \end{aligned}$$

$k \geq 2$ とする。

$$\begin{aligned} S_k - S_{k-1} &= \sum_{i=1}^k X_i^2 - \frac{1}{k} \left(\sum_{i=1}^k X_i \right)^2 - \sum_{i=1}^{k-1} X_i^2 + \frac{1}{k-1} \left(\sum_{i=1}^{k-1} X_i \right)^2 \\ &= X_k^2 + \frac{1}{k-1} \left(\sum_{i=1}^{k-1} X_i \right)^2 - \frac{1}{k} \left(\sum_{i=1}^k X_i \right)^2 \\ &= X_k^2 + \frac{1}{k-1} \left(\sum_{i=1}^{k-1} X_i \right)^2 - \frac{1}{k} \left(\sum_{i=1}^{k-1} X_i + X_k \right)^2 \\ &= X_k^2 + \frac{1}{k-1} \left(\sum_{i=1}^{k-1} X_i \right)^2 - \frac{1}{k} \left\{ \left(\sum_{i=1}^{k-1} X_i \right)^2 + 2X_k \sum_{i=1}^{k-1} X_i + X_k^2 \right\} \\ &= \frac{1}{k(k-1)} \left\{ \left(\sum_{i=1}^{k-1} X_i \right)^2 - 2(k-1)X_k \sum_{i=1}^{k-1} X_i + (k-1)^2 X_k^2 \right\} \\ &= \frac{1}{k(k-1)} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\}^2 \\ &= \left[\frac{1}{\sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\} \right]^2 , \end{aligned}$$

σ^2 で割ると

$$\frac{S_k}{\sigma^2} - \frac{S_{k-1}}{\sigma^2} = \left[\frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\} \right]^2 .$$

確率変数 Z_{k-1} を次のように定める：

$$Z_{k-1} = \frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\} .$$

このとき

$$\frac{S_k}{\sigma^2} - \frac{S_{k-1}}{\sigma^2} = Z_{k-1}^2 .$$

確率変数 $X_1, X_2, X_3, \dots, X_k$ は正規分布 $N(\mu, \sigma^2)$ に従い互いに独立なので、

$$\begin{aligned} E[Z_{k-1}] &= E \left[\frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\} \right] \\ &= \frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} E[X_i] - (k-1)E[X_k] \right\} \\ &= \frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} \mu - (k-1)\mu \right\} \\ &= 0 , \end{aligned}$$

$$\begin{aligned} V[Z_{k-1}] &= V \left[\frac{1}{\sigma \sqrt{k(k-1)}} \left\{ \sum_{i=1}^{k-1} X_i - (k-1)X_k \right\} \right] \\ &= \frac{1}{\sigma^2 k(k-1)} \left\{ \sum_{i=1}^{k-1} V[X_i] + (k-1)^2 V[X_k] \right\} \\ &= \frac{1}{\sigma^2 k(k-1)} \left\{ \sum_{i=1}^{k-1} \sigma^2 + (k-1)^2 \sigma^2 \right\} = \frac{1}{\sigma^2 k(k-1)} \{ \sigma^2(k-1) + \sigma^2(k-1)^2 \} \\ &= 1 . \end{aligned}$$

定理 4.1.2 より Z_{k-1} は標準正規分布 $N(0,1)$ に従う。 証明は略しますが確率変数 $Z_1, Z_2, Z_3, \dots, Z_{n-1}$ は互いに独立である。 $\frac{S_k}{\sigma^2} - \frac{S_{k-1}}{\sigma^2} = Z_{k-1}^2$ より、

$$\sum_{k=2}^n \left(\frac{S_k}{\sigma^2} - \frac{S_{k-1}}{\sigma^2} \right) = \sum_{k=2}^n Z_{k-1}^2 ,$$

$$\frac{S_2}{\sigma^2} - \frac{S_1}{\sigma^2} + \frac{S_3}{\sigma^2} - \frac{S_2}{\sigma^2} + \frac{S_4}{\sigma^2} - \frac{S_3}{\sigma^2} + \dots + \frac{S_n}{\sigma^2} - \frac{S_{n-1}}{\sigma^2} = Z_1^2 + Z_2^2 + Z_2^2 + \dots + Z_{n-1}^2 ,$$

$S_1 = 0$ なので

$$\frac{S_n}{\sigma^2} = \sum_{k=1}^{n-1} Z_k^2 .$$

故に確率変数 $\frac{S_n}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従う。