知能システム論第 3 回課題

37186305
航空宇宙工学専攻修士一年
荒居秀尚

2018 年 10 月 18 日

1 宿題 1

超幾何分布の分散が

$$V(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)} \tag{1.1}$$

で与えられることを示す。

分散は

$$V(X) = E[X(X-1)] + E[x] - \{E[X]\}^2 \tag{1.2}$$

と表されるので $E[X(X-1)]$ を求めることを考える。

$$E[X(X-1)] = \frac{1}{{}_N\mathrm{C}_n} \sum_{x=0}^{n} x(x-1) {}_M\mathrm{C}_x \cdot {}_{N-M}\mathrm{C}_{n-x} \tag{1.3}$$

$$= \frac{1}{{}_N\mathrm{C}_n} \sum_{x=2}^{n} x(x-1) {}_M\mathrm{C}_x \cdot {}_{N-M}\mathrm{C}_{n-x} \tag{1.4}$$

ここで

$${}_M\mathrm{C}_x = \frac{M(M-1)}{x(x-1)} {}_{M-2}\mathrm{C}_{x-2} \tag{1.5}$$

より、

$$E[X(X-1)] = \frac{M(M-1)}{{}_N\mathrm{C}_n} \sum_{x=2}^{n} {}_{M-2}\mathrm{C}_{x-2} \cdot {}_{N-M}\mathrm{C}_{n-x} \tag{1.6}$$

$$= \frac{M(M-1)}{{}_N\mathrm{C}_n} \sum_{x'=0}^{n-2} {}_{M-2}\mathrm{C}_{x'} {}_{N-M}\mathrm{C}_{n-x'-2} \tag{1.7}$$

$$= \frac{n(n-1)M(M-1)}{N(N-1)} \sum_{x=0}^{n-2} \frac{{}_{M-2}\mathrm{C}_x \cdot {}_{N-M}\mathrm{C}_{n-x-2}}{{}_{N-2}\mathrm{C}_{n-2}} \tag{1.8}$$

ここで、

$$\sum_{x=0}^{n} \frac{{}_M\mathrm{C}_x \cdot {}_{N-M}\mathrm{C}_{n-x}}{{}_N\mathrm{C}_n} = 1 \tag{1.9}$$

の式で M, N, n をそれぞれ 1 ずつ下げたとして結果が変わらないことから帰納的に

$$\sum_{x=0}^{n-2} \frac{{}_{M-2}\mathrm{C}_x \cdot {}_{N-M}\mathrm{C}_{n-x-2}}{{}_{N-2}\mathrm{C}_{n-2}} = 1 \tag{1.10}$$

となるため、

$$E[X(X-1)] = \frac{n(n-1)M(M-1)}{N(N-1)} \tag{1.11}$$

である。これを用いて式 (1.2) に $E[X]$ と共に代入することで式 (1.1) を得る。

2　宿題 2

確率変数 X の確率密度関数が $f(x)$ と表される、という条件は、任意の実数 x に対して、

$$P(X \leq x) = \int_{-\infty}^{x} f(u)du \tag{2.1}$$

が成り立つことを意味する。また、$X = t(Y)$ の関係と式 (2.1) から

$$P(X \leq t(y)) = \int_{-\infty}^{t(y)} f(u)du \tag{2.2}$$

と書ける。一方同様に、確率変数 Y の確率密度関数は $g(y)$ のため、

$$P(Y \leq y) = \int_{-\infty}^{y} g(u)du \tag{2.3}$$

ここで、$X = t(Y)$ と $t(\cdot)$ より

$$P(X \leq t(y)) = P(Y \leq y) \tag{2.4}$$

式 (2.4) と式 (2.2)、式 (2.3) より、

$$\int_{-\infty}^{y} g(u)du = \int_{-\infty}^{t(y)} f(u')du' \tag{2.5}$$

両辺を y で微分して

$$g(y) = f(t(y))\frac{dt(y)}{dy} \tag{2.6}$$

を得る。

3 宿題 3

$E[X^2]$ を求めることを考える。

$$E[X^2] = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha+1}(1-x)^{\beta-1} dx \tag{3.1}$$
$$= \frac{1}{B(\alpha,\beta)} \left\{ \left[x^{\alpha+1} \left(-\frac{(1-x)^\beta}{\beta} \right) \right]_0^1 - \int_0^1 (\alpha+1)x^\alpha \left(-\frac{(1-x)^\beta}{\beta} \right) dx \right\} \tag{3.2}$$
$$= \frac{\alpha+1}{\beta} \frac{1}{B(\alpha,\beta)} \int_0^1 x^\alpha (1-x)^\beta dx \tag{3.3}$$
$$= \frac{\alpha+1}{\beta} \frac{1}{B(\alpha,\beta)} \left\{ \left[x^\alpha \left(-\frac{(1-x)^{\beta+1}}{\beta+1} \right) \right]_0^1 - \int_0^1 \alpha x^{\alpha+1} \left(-\frac{(1-x)^{\beta+1}}{\beta+1} \right) dx \right\} \tag{3.4}$$
$$= \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}(1-2x+x^2) dx \tag{3.5}$$
$$= \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \frac{1}{B(\alpha,\beta)} (1 - 2E[X] + E[X^2]) \tag{3.6}$$

これを $E[X^2]$ についてまとめると

$$E[X^2] = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \tag{3.7}$$

これと $V(X) = E[X^2] - \{E[X]\}^2$ を用いると

$$V(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{3.8}$$

となる。

4 宿題 4

Python3.6.6 を用いた。

```python
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

if __name__ == "__main__":
    np.random.seed(1234)
    data = []
    for i in range(1, 11):
        size = [10000, i]
        x = np.random.normal(size=size)
        y = np.sum(np.square(x), axis=1)
        data.append(y)

    plt.figure(figsize=[10, 8])
    sns.distplot(data[0], kde_kws={"color": "k", "label": "dimension1"})
    sns.distplot(data[1], kde_kws={"color": "r", "label": "dimension2"})
```

```
    sns.distplot(data[2], kde_kws={"color": "g", "label": "dimension3"})
    sns.distplot(data[3], kde_kws={"color": "b", "label": "dimension4"})
    sns.distplot(data[4], kde_kws={"color": "y", "label": "dimension5"})
    sns.distplot(data[5], kde_kws={"color": "cyan", "label": "dimension6"})
    sns.distplot(data[6], kde_kws={"color": "purple", "label": "dimension7"})
    sns.distplot(data[7], kde_kws={"color": "orange", "label": "dimension8"})
    sns.distplot(data[8], kde_kws={"color": "gray", "label": "dimension9"})
    sns.distplot(data[9], kde_kws={"color": "brown", "label": "dimension10"})
    plt.title("Histogram␣of␣chi␣square␣distribution")
    plt.xlabel("value")
    plt.ylabel("freq")
    plt.savefig("hist.png")
```

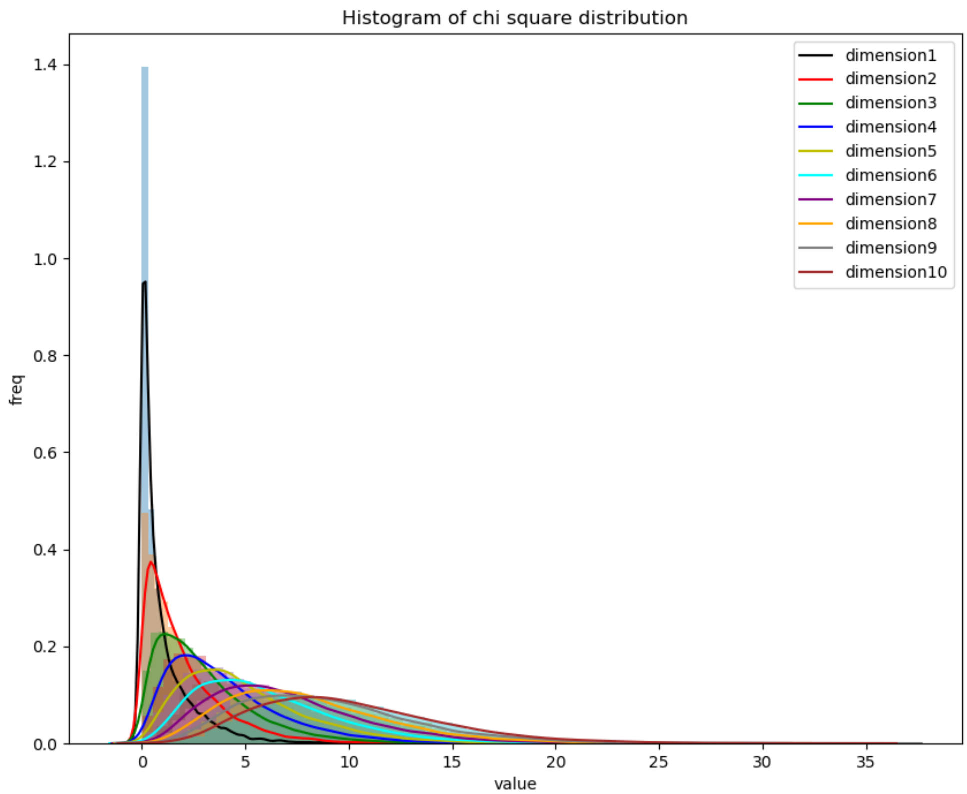

図 4.1: 自由度が 1-10 のカイ二乗分布のヒストグラム