確率及び統計／レポート２

学籍番号：　０９５７０７Ｂ
氏名：　大 城 佳 明
提出日：　2010 年 5 月 11 日

1　問題

超幾何分布の確率関数は、次式で定義されている。

$$P(X=k) = \frac{{}_{N_1}C_k {}_{N_2}C_{n-k}}{{}_{N_2}C_n} (N_1, N_2 > 0, N = N_1 + N_2, k = 0, 1, 2, \cdots, n)$$

(1) 確率変数であることを証明せよ。
(2) 期待値を求めよ。
(3) 分散を求めよ。

2　解答

2.1　(1) 確率変数であることを証明せよ。

確率変数であるための条件は

$$0 \leq P(X=k) \leq 1 \quad かつ \quad \sum_{k=0}^{n} P(X=k) = 1$$

の二つを満たす時である。$N_1, N_2 > 0, 0 \leq k$ であるから、${}_{N_1}C_k, {}_{N_2}C_{n-k}, {}_N C_n > 0$ となるので、

$$P(X=k) = \frac{{}_{N_1}C_k {}_{N_2}C_{n-k}}{{}_{N_2}C_{n-k}} \geq 0 \tag{1}$$

が証明された。

次に $\sum_{k=0}^{n} P(X=k) = 1$ であることを示す。

$$(1+x)^a = \sum_{k=0}^{a} {}_aC_k x^k \tag{2}$$

$$(1+x)^b = \sum_{k=0}^{b} {}_bC_k x^k \tag{3}$$

$$(1+x)^{a+b} = \sum_{k=0}^{a+b} {}_{a+b}C_k x^k \tag{4}$$

また

$$(1+x)^a(1+x)^b=(1+x)^{a+b}$$

である。(4) に (1),(2),(3) を代入すると、

$$\begin{aligned}(\sum_{k=0}^{a}{}_aC_kx^k)(\sum_{k=0}^{b}{}_bC_kx^k) &= \sum_{k=0}^{a+b}{}_{a+b}C_kx^k \\ &= ({}_aC_0x^0+{}_aC_0x^1+\cdots+{}_aC_ax^a)({}_bC_0x^0+{}_bC_0x^1+\cdots+{}_bC_bx^a)\end{aligned}$$

ここで、この恒等式の x^n の項だけを見ると、

$$\begin{aligned}{}_{a+b}C_n &= ({}_aC_0x^0\cdot{}_bC_0x^0+{}_aC_1x^1\cdot{}_bC_1x^1+\cdots+{}_aC_{n-1}x^{n-1}\cdot{}_bC_{n-1}x^{n-1}) \\ &= x^n\sum_{k=0}^{n}{}_aC_k{}_bC_{n-k}\end{aligned}$$

であるから、両辺を x^n で割ると

$$_{a+b}C_n=\sum_{k=0}^{n}{}_aC_k{}_bC_{n-k} \tag{5}$$

となる。

$$\begin{aligned}\sum_{k=0}^{n}P(X=k) &= \sum_{k=0}^{n}\frac{{}_{N_1}C_k{}_{N_2}C_{n-k}}{{}_{N_2}C_n} \\ &= \frac{1}{{}_NC_n}({}_{N_1}C_k{}_{N_2}C_{n-k}) \\ &= \frac{1}{{}_NC_n}\cdot{}_{N_1+N_2}C_n \\ &= \frac{{}_NC_n}{{}_NC_n} \\ &= 1 \qquad (6)\end{aligned}$$

(6) より

$$P(X=k)\leq 1 \tag{7}$$

であることにもなる。
したがって、(1),(6),(7) より P (X = k) は確率変数であるという事が証明された。

2.2 (2) 期待値を求めよ。

期待値の定義

$$\begin{aligned} E[X] &= \sum_{k=0}^{n} kP(X=k) \\ &= \sum_{k=1}^{n} kP(X=k) \end{aligned}$$

$$\begin{aligned} kP(X=k) &= k\frac{{}_{N_1}C_{k}\,{}_{N_2}C_{n-k}}{{}_{N_2}C_{n}} \\ &= k\frac{N_1!}{k!(N_1-k)!}\frac{n!(N-n)!}{N!}{}_{N_2}C_{n-k} \\ &= \frac{N_1(N_1-1)!}{(k-1)!(N_1-k)!}\frac{n!(n-1)!(N-n)!}{N(N-1)!}{}_{N_2}C_{n-k} \\ &= \frac{nN_1}{N}\frac{1}{{}_{N-1}C_{n-1}}{}_{N-1}C_{n-1}\,{}_{N-2}C_{n-2} \\ \sum_{k=1}^{n} kP(X=k) &= \frac{nN_1}{N}\frac{1}{{}_{N-1}C_{n-1}}\sum_{k=1}^{n}{}_{N-1}C_{n-1}\,{}_{N-2}C_{n-2} \end{aligned}$$

ここで、問題 1 の (5) の式を使う。k = 0, 1, 2..., n − 1 であるから k − 1 = i と置くと i = 0, 1, 2..., n − 2 となるので、

$$\begin{aligned} \sum_{k=1}^{n} kP(X=k) &= \frac{nN_1}{N}\frac{1}{{}_{N-1}C_{n-1}}{}_{N_1+N_2-1}C_{n-1} \\ &= \frac{nN_1}{N}\frac{{}_{N-1}C_{n-1}}{{}_{N-1}C_{n-1}} \\ &= \frac{nN_1}{N} \end{aligned}$$

したがって、期待値は $\frac{nN_1}{N}$ である。

2.3　(3) 分散を求めよ。

分散の定義

$$V[X] = \sum_{k=0}^{n}(k - E[X])^2 P(X = k)$$

より

$$\begin{aligned} V[X] &= \sum_{k=0}^{n}(k - \frac{nN_1}{N})^2 P(X = k) \\ &= \sum_{k=0}^{n}\{k^2 - 2k\frac{nN_1}{N} + (\frac{nN_1}{N})^2\}P(X = k) \\ &= \sum_{k=0}^{n} k^2 P(k) - 2\frac{nN_1}{N}\sum_{k=0}^{n} kP(k) + (\frac{nN_1}{N})^2 \sum_{k=0}^{n} kP(k) \end{aligned} \tag{8}$$

となる。それぞれの値は

$$\sum_{k=0}^{n} kP(k) = \frac{nN_1}{N} \tag{9}$$

$$\sum_{k=0}^{n} P(k) = 1 \tag{10}$$

より、(9),(10) を (8) に代入すると

$$V[X] = \sum_{k=0}^{n} k^2 P(k) - (\frac{nN_1}{N})^2 \tag{11}$$

$\sum_{k=0}^{n} k^2 P(k)$ の値を求める。

$$\begin{aligned} \sum_{k=0}^{n} k^2 P(k) &= \sum_{k=2}^{n}\{k(k-1)P(k) + kP(k)\} \\ &= \sum_{k=2}^{n} k(k-1)P(k) + \frac{nN_1}{N} \\ k(k-1)P(k) &= k(k-1)\frac{N_1!}{k!(N_1-k)!}\frac{n!(N-n)!}{N!}{}_{N_2}Cn - k \\ &= \frac{N_1(N_1-1)(N_1-2)!}{(k-2)!(N_1-k)!}\frac{n(n-1)(n-2)!(N-n)!}{N(N-1)(N-2)!}{}_{N_1}C_{n-1} \\ &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\frac{1}{{}_{N-2}C_{n-2}}{}_{N_1}C_{k-2}{}_{N_2}C_{n-k} \\ \sum_{k=2}^{n} k(k-1)P(k) &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\frac{1}{{}_{N-2}C_{n-2}}\sum_{k=2}^{n}{}_{N_1}C_{k-2}{}_{N_2}C_{n-k} \end{aligned}$$

ここでもまた問題１の (4) の式を使いたいので $k-2=j$ と置くことで $j = 0, 1, 2, ..., n-2$ となるので

$$\begin{aligned} \sum_{k=2}^{n} k(k-1)P(k) &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\frac{1}{{}_{N-2}C_{n-2}}\sum_{j=0}^{n}{}_{N_1}C_{j}{}_{N_2}C_{n-j+2} \\ &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\frac{1}{{}_{N-2}C_{n-2}}({}_{N_1+N_2-2}C_{n-2}) \\ &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)}\frac{{}_{N-2}C_{n-2}}{{}_{N-2}C_{n-2}} \\ &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)} \end{aligned} \tag{12}$$

よって (12) を (11) の式に代入すると、

$$V[X] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)} - (\frac{nN_1}{N})^2 \tag{13}$$

よって (9),(10),(13) を (8) に代入してまとめると、

$$\begin{aligned} V[X] &= \frac{n(n-1)N_1(N_1-1)}{N(N-1)} + \frac{nN_1}{N} - (\frac{nN_1}{N})^2 \\ &= \frac{nN_1}{N}\{\frac{(n-1)(N_1-1)}{N-1} + 1 - \frac{nN_1}{N} \\ &= \frac{nN_1}{N}\{\frac{(n-1)(N_1-1) + (N-nN_1)(N-1)}{N(N-1)} \\ &= \frac{nN_1}{N}\{\frac{N^2 - nN - NN_1 + nN_1}{N(N-1)} \\ &= \frac{nN_1}{N}\frac{(N-N_1)(N-n)}{N(N-1)} \end{aligned}$$

したがって、分散のは $\frac{nN_1(N-N_1)(N-n)}{N^2(N-1)}$ である。