

数 学 2 (問題)

(解答に当たり、必要であれば末尾の数表を用いよ。)

1. 次の各問の に入る答のみを、所定の解答用紙に記入せよ。 (35点)

(1) 確率分布が $f(x; p) = p(1-p)^x$ ($x=0, 1, 2, \dots, 0 < p < 1$) で与えられている母集団からの大きさ n の標本を (x_1, \dots, x_n) とするとき、この標本より求めた p の最尤推定値は である。

(2) 平均 μ 、分散 σ^2 を持つ母集団から、 $2N-1$ 個の標本変数 $X_1, X_2, \dots, X_{2N-1}$ をとったとき、

$$\text{cov}\left(\sum_{i=1}^N X_i, \sum_{j=N}^{2N-1} X_j\right) = \text{ } \text{ である。}$$

(3) 箱の中に24個の玉が入っている。 H_0 : 黄色20個、緑4個、 H_1 : 黄色4個、緑20個のどちらかである。復元抽出法で3個玉を取り出し、第1種の誤りの確率10%以下で H_0 を検定する。次の3つの棄却域を定める方法のうち、最良と思われる方法の第2種の誤りの確率は である。
方法1: 少なくとも1個が緑、 方法2: 2個以上が緑、 方法3: 3個とも緑。

(4) 「1」という札が1枚、「2」という札が2枚、 \dots 、「 N 」という札が N 枚の有限母集団から、 N 個の標本を取るとき札の値の標本平均の分散は である。

(5) 母集団の確率密度関数が、一様分布

$$f(x; \alpha) = \begin{cases} \frac{1}{\alpha} & (\beta - \frac{1}{2}\alpha \leq x \leq \beta + \frac{1}{2}\alpha, \quad \alpha > 0, \quad \alpha, \beta \text{ は未知}) \\ 0 & (\text{その他の領域}) \end{cases}$$

にしたがうとき、 n 個の標本変数 (X_1, \dots, X_n) によって母数 α を推定する。

α の最尤推定量を $\hat{\alpha} = \hat{\alpha}(X_1, \dots, X_n)$ とすれば、これは α の不偏な推定量ではないが、 $\times \hat{\alpha}$ は α の不偏推定量である。

2. X_1, X_2, \dots, X_n は互いに独立でいずれも母平均 θ の不偏推定量とする。

(1) $Y = \sum_{i=1}^n a_i X_i$ が θ の不偏推定量であるために必要な a_1, \dots, a_n の条件を求めよ。

(2) 各 X_i の分散が母分散 σ^2 であるとする。 $Y = \sum_{i=1}^n a_i X_i$ が θ の不偏推定量であるとき、 $V(Y)$ を最小とする (a_1, \dots, a_n) および $V(Y)$ の最小値を求めよ。 (20点)

3. 生まれたばかりの実験動物15匹中、8匹には飼料Aを与え、残りの7匹には飼料Bを与えて生育した。一定期間を経た後に体重を計ったら、前者は

46.0, 46.2, 47.1, 45.0, 48.7, 47.6, 46.8, 48.6

であり、後者は

48.6, 49.0, 47.5, 51.0, 50.3, 49.0, 49.7

であった。飼料の違いによる影響があると言って良いか否かを有意水準0.01で検定せよ。

(標本は正規母集団からのものと考えてよい。) (25点)

4. X_1, X_2, \dots, X_n をある母集団からの標本変量とする。その母集団の母平均を μ 、母分散を σ^2 とし、各 X_i を正規化した $X_i^* = \frac{X_i - \mu}{\sigma}$ の積率母関数を $\phi(t)$ とおく。

また、標本変量平均 $\bar{X} = \frac{X_1 + \dots + X_n}{n}$ を正規化して $Y_n = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$ とおく。

- (1) Y_n の積率母関数を $\phi_n(t)$ とするとき、 $\phi_n(t)$ を $\phi(t)$ で表わせ。
- (2) $n \rightarrow \infty$ のとき、 $\phi_n(t)$ の極限を求めよ。
- (3) 相当多数の者が受験したある数学の試験の平均点が m 、標準偏差が s であった場合を例として (2) の結果を言葉で説明せよ。

(20点)

(解答に当り、必要であれば以下の数表を用いよ。)

数表

(1) X が F 分布にしたがうとき、 $P(X \geq \lambda) = 0.05$ となる λ を求める表

f_1 は分母の自由度、 f_2 は分子の自由度 $\lambda = F_{f_1, f_2}(0.05)$ とも書く。

$f_2 \backslash f_1$	5	6	7	8	9
5	5.05	4.95	4.88	4.82	4.77
6	4.39	4.28	4.21	4.15	4.10
7	3.97	3.87	3.79	3.73	3.68
8	3.69	3.58	3.50	3.44	3.39
9	3.48	3.37	3.29	3.23	3.18

(2) X が t 分布にしたがうとき、 $P(X \geq \lambda) = \varepsilon$ となる λ を求める表

n は自由度 $\lambda = t_n(\varepsilon)$ とも書く。

$\varepsilon \backslash n$	0.01	0.005
5	3.365	4.032
6	3.143	3.707
7	2.998	3.499
8	2.896	3.355
9	2.821	3.250
10	2.764	3.169
11	2.718	3.106
12	2.681	3.055
13	2.650	3.012
14	2.624	2.977
15	2.602	2.947

数学 2 解答例

1. (1) 尤度関数 $L(p)$ は次式で表わされる。

$$L(p) = p^n (1-p)^{x_1} (1-p)^{x_2} \cdots (1-p)^{x_n}$$

両辺の対数を取って

$$\log L(p) = n \log p + \left(\sum_{i=1}^n x_i \right) \log(1-p)$$

$\log L(p)$ の極値を求める。

$$\frac{d \log L(p)}{dp} = \frac{n}{p} - \frac{1}{1-p} \sum_{i=1}^n x_i = 0 \quad \text{より} \quad p = \frac{n}{n + \sum_{i=1}^n x_i} = \frac{1}{1 + \bar{x}} \quad \text{①}$$

$$\frac{d^2 \log L(p)}{dp^2} = -\frac{n}{p^2} - \frac{1}{(1-p)^2} \sum_{i=1}^n x_i < 0 \quad (\because \sum_{i=1}^n x_i > 0)$$

従って①の p は最大値を与える。

$$(2) \quad E\left(\sum_{i=1}^n X_i\right) = E\left(\sum_{j=N}^{2N-1} X_j\right) = N\mu$$

$$\begin{aligned} & E\left\{\left(\sum_{i=1}^N X_i\right)\left(\sum_{j=N}^{2N-1} X_j\right)\right\} = E\left\{\left(X_1 + \cdots + X_N\right)\left(X_N + X_{N+1} + \cdots + X_{2N-1}\right)\right\} \\ & = E\left\{X_N^2 + \sum_{i=1}^{N-1} E(X_N X_i) \sum_{j=N+1}^N \sum_{k=N+1}^{2N-1} E(X_i X_j)\right\} = E(X_N^2) + \sum_{i=1}^{N-1} E(X_N) E(X_i) + \sum_{i=1}^N \sum_{j=N+1}^{2N-1} E(X_i) E(X_j) \\ & = \sigma^2 + \mu^2 + (N-1)\mu^2 + N(N-1)\mu^2 = \sigma^2 + N^2\mu^2 \end{aligned}$$

$$\begin{aligned} & \text{cov}\left(\sum_{i=1}^N X_i, \sum_{j=N}^{2N-1} X_j\right) = E\left\{\left(\sum_{i=1}^N X_i\right)\left(\sum_{j=N}^{2N-1} X_j\right)\right\} - E\left(\sum_{i=1}^N X_i\right) E\left(\sum_{j=N}^{2N-1} X_j\right) \\ & = \sigma^2 + N^2\mu^2 - (N\mu)(N\mu) = \boxed{\sigma^2} \end{aligned}$$

- (3) 第一種の誤りの確率を α , 第2種の誤りの確率を β とする。

$$\text{方法 1} \quad \begin{cases} \alpha = P\{\text{少なく共1個緑} \mid H_0\} = 1 - P\{\text{3個共黄色} \mid H_0\} = 1 - \left(\frac{20}{24}\right)^3 = 0.421 \\ \beta = P\{\text{3個共黄色} \mid H_1\} = \left(\frac{4}{24}\right)^3 = 0.005 \end{cases}$$

$$\text{方法 2} \quad \begin{cases} \alpha = P\{\text{緑が2個以上} \mid H_0\} = {}_3C_2 \left(\frac{4}{24}\right)^2 \left(\frac{20}{24}\right) + {}_3C_3 \left(\frac{4}{24}\right)^3 = 3 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right) + \left(\frac{1}{6}\right)^3 = \frac{2}{27} \\ \beta = P\{\text{緑が1個以下} \mid H_1\} = {}_3C_1 \left(\frac{20}{24}\right) \cdot \left(\frac{4}{24}\right)^2 + {}_3C_3 \left(\frac{4}{24}\right)^3 = \frac{2}{27} = 0.074 \end{cases}$$

$$\text{方法 3} \quad \begin{cases} \alpha = P\{\text{3個共緑} \mid H_0\} = \left(\frac{4}{24}\right)^3 = 0.005 \\ \beta = P\{\text{少なく共1個黄色} \mid H_1\} = 1 - \{ \text{3個共緑} \mid H_1 \} = 1 - \left(\frac{20}{24}\right)^3 = 0.421 \end{cases}$$

α が10%以下は方法2と方法3, そのうち β の小さいのは方法2なので3つ
のうち最良なのは方法2である。

求める答は $\beta = \boxed{\frac{2}{27}}$ (0.074等小数3位まで正しければ正解とした。)

$$(4) \text{ 母平均は } \frac{1+2 \times 2(\text{個})+\cdots+N \times N(\text{個})}{1+2+\cdots+N} = \frac{2 \cdot N(N+1)(2N+1)}{N(N+1) \cdot 6} = \frac{2N+1}{3}$$

$$\text{母分散 } \sigma^2 \text{は } \frac{1^2+2^2 \times 2(\text{個})+\cdots+N^2 \times N(\text{個})}{N(N+1)/2} - \left(\frac{2N+1}{3}\right)^2$$

$$= \frac{\{N(N+1)/2\}^2}{N(N+1)/2} - \left(\frac{2N+1}{3}\right)^2 = \frac{N^2+N}{2} - \frac{4N^2+4N+1}{9} = \frac{N^2+N-2}{18}$$

$$\text{標本平均の分散 } \sigma_x^2 = \frac{\sigma^2}{N} \left(\frac{N(N+1)}{2} - N \right) = \frac{N^2+N-2}{18N} \cdot \frac{N^2+N-2N}{N^2+N-2} = \boxed{\frac{N-1}{18}}$$

(5) $a \geq \max(x_1, \dots, x_n) - \min(x_1, \dots, x_n)$ である。

尤度関数 $\ell(a) = \left(\frac{1}{a}\right)^n \leq \left(\frac{1}{\max(x_1, \dots, x_n) - \min(x_1, \dots, x_n)}\right)^n$ となるから

$$\hat{a} = \max(x_1, \dots, x_n) - \min(x_1, \dots, x_n)$$

$$E(\hat{a}) = E\{\max(x_1, \dots, x_n)\} - E\{\min(x_1, \dots, x_n)\}$$

ここで, X_i が $1/a$ の一様分布のとき $\max(x_1, \dots, x_n)$ の密度関係は $n\left(\frac{x}{a}\right)^{n-1} \cdot \frac{1}{a}$

又 $\min(x_1, \dots, x_n)$ の密度関数は $n\left(1 - \frac{x}{a}\right)^{n-1} \frac{1}{a}$

$$\therefore E(\hat{a}) = \int_0^a xn \frac{x^{n-1}}{a^n} dx - \int_0^a xn \frac{(a-x)^{n-1}}{a^n} dx = \left[\frac{n}{a^n} \cdot \frac{x^{n+1}}{n+1} \right]_0^a - \left\{ \left[\frac{-n}{a^n} \cdot \frac{x(a-x)^n}{n} \right]_0^a \right.$$

$$\left. + \int_0^a \frac{(a-x)^n}{a^n} dx \right\} = \frac{n}{n+1} a - \left[\frac{-n}{a^n} \cdot \frac{(a-x)^{n+1}}{n+1} \right]_0^a = \frac{n-1}{n+1} a$$

従って $\boxed{\frac{n-1}{n+1}} \times \hat{a}$ は a の不偏推定量となる。

2.

(1) $E(Y) = \theta$ となる a_1, \dots, a_n の条件を求めればよい

$$E(Y) = a_1 E(X_1) + \cdots + a_n E(X_n) = a_1 \theta + \cdots + a_n \theta (\because E(X_i) = \theta)$$

$$= (a_1 + \cdots + a_n) \theta$$

従って, 求める条件は $a_1 + \cdots + a_n = 1$ ——①

$$\begin{aligned}
(2) \quad V(Y) &= a_1^2 V(X_1) + \cdots + a_n^2 V(X_n) \quad (\because \text{各} X_i \text{は互いに独立}) \\
&= a_1^2 \sigma^2 + \cdots + a_n \sigma^2 \quad (\because V(X_i) = \sigma^2) \\
&= (a_1^2 + \cdots + a_n^2) \sigma^2 \text{ —— ②}
\end{aligned}$$

Yがθの不偏推定量であるとき V(Y)の最小値を求めるには
①の条件の下で②を最小にすればよい。

ところで

$$\sum_{i=1}^n a_i^2 = \sum_{i=1}^n \left(a_i - \frac{1}{n} \right)^2 + \frac{1}{n} \quad (\because a_1 + \cdots + a_n = 1) \text{ —— ③}$$

③式より $a_1 = \cdots = a_n = \frac{1}{n}$ のとき、②は最小分散 $\frac{\sigma^2}{n}$ を持つ。

[別解1] $V(Y) = a_1^2 \sigma^2 + \cdots + a_n^2 \sigma^2 \text{ —— ④}$

①の条件の下で、④を最小にする。ラグランジュ乗数法で

$$g(a_1, \dots, a_n, \lambda) \equiv a_1^2 \sigma^2 + \cdots + a_n^2 \sigma^2 + \lambda(1 - a_1 - \cdots - a_n)$$

gを各変数で偏微分してgの極値を求める。

$$\left\{ \begin{array}{l} \frac{\partial g}{\partial a_1} = 2a_1 \sigma^2 - \lambda = 0 \\ \vdots \\ \frac{\partial g}{\partial a_n} = 2a_n \sigma^2 - \lambda = 0 \\ \frac{\partial g}{\partial \lambda} = 1 - a_1 - \cdots - a_n = 0 \text{ を解くと} \end{array} \right. \quad \text{⑤} \quad \left\{ \begin{array}{l} \lambda = \frac{2}{n/\sigma^2} \\ a_1 = \frac{\lambda}{2\sigma^2} = \frac{1}{n} \\ \vdots \\ a_n = \frac{\lambda}{2\sigma^2} = \frac{1}{n} \end{array} \right.$$

ところで④式は下に凸であるから⑤の a_1, \dots, a_n は④の最小値を与える。

⑤を④に代入して最小値は $V(Y) = \frac{\sigma^2}{n}$

[別解2] コーシー・シュバルツの不等式を使う場合

$a_1 + \cdots + a_n = 1$ で $a_1^2 + \cdots + a_n^2$ の最小を求めるには

各 $a_i \geq 0$ を仮定して考えてよい。

⊙ $\exists i$ で $a_i < 0$ とすると, $a_i' = 0$ としてみると $a_i > |a_i|$ となる j がある筈だから
 a_j の代わりに $a_j' = a_j + a_i$ を用い, a_i の代わりに a_i' を用いても
 ①の関係式は満たされる。
 このとき
 $|a_j'| < |a_j|$, $0 = |a_i'| < |a_i|$ だから $(a_j')^2 + (a_i')^2 < (a_j)^2 + (a_i)^2$
 従って a_i のうち負のものがあつたらそれは②の最小値を与えない。

$a_i \geq 0$ を仮定するとコーシー・シュバルツの不等式から

$$1 = \left(\sum_{i=1}^n 1 \cdot a_i \right)^2 \leq \left(\sum_{i=1}^n 1^2 \right) \cdot \left(\sum_{i=1}^n a_i^2 \right) = n \cdot \sum_{i=1}^n a_i^2$$

が成り立ち, 等号の成立するのは $a_1 = \dots = a_n$ ⑥

のときである。

①, ⑥より求めた解 $a_1 = \dots = a_n = \frac{1}{n}$ を②に代入すると最小の分散 $\frac{\sigma^2}{n}$ を得る。

3. 飼料 A の方の標本の平均を \bar{x}_a , 標本分散を S_a^2 , 標本数を n_a 標本変量平均を \bar{X}_a とし
 飼料 B について同様に \bar{x}_b , S_b^2 , n_b , \bar{X}_b とすると

$$\bar{x}_a = 47.0, \bar{S}_a^2 = 1.4375, n_a = 8, \bar{x}_b = 49.3, \bar{S}_b^2 = 1.1371, n_b = 7$$

飼料 A で育成した動物の体重の母集団は $N(\mu_a, \sigma_a^2)$ に従い, 飼料 B の母集団は $N(\mu_b, \sigma_b^2)$

に従う (題意より) と考えて良いが, まず $H_0: \sigma_a^2 = \sigma_b^2$ が成り立つか否かを調べる。

飼料 A の不偏分散は $\frac{n_a S_a^2}{n_a - 1} = \frac{8 \times 1.4375}{7} \doteq 1.643$, 飼料 B の不偏分散 $\frac{7 \times 1.1371}{6} \doteq 1.327$

従って不偏分散比の分子は大きい方を取ると, $F_0 = \frac{\frac{n_a S_a^2}{n_a - 1}}{\frac{n_b S_b^2}{n_b - 1}}$ は自由度 $(n_a - 1, n_b - 1)$ の

F 分布に従い $F_0 > F_{n_b - 1}^{\alpha} \left(\frac{\epsilon}{2} \right)$ ならば H_0 を棄却する。

問題の数表から $\epsilon/2 = 0.05$ とすると $F_0 = \frac{1.643}{1.327} \doteq 1.238 < F_6^7(0.05) = 4.21$ となり H_0 は
 棄却されない。即ち $\sigma_a^2 = \sigma_b^2$ と考えて良いことが分った。

従って, 平均値の差の検定は t 検定を用いることとする。

帰無仮説を $H_1: \mu_a - \mu_b = 0; \mu_a = \mu_b$ とおき, 対立仮説を $H_2: \mu_a \neq \mu_b$ とする。

$T = \frac{\bar{X}_a - \bar{X}_b}{\sqrt{\frac{n_a S_a^2 + n_b S_b^2}{n_a + n_b - 2} \cdot \left(\frac{1}{n_a} + \frac{1}{n_b} \right)}}$ は自由度 $n_a + n_b - 2$ の t-分布に従い,

$|T| > t_{n_a + n_b - 2} \left(\frac{\epsilon}{2} \right)$ ならば H_1 を棄却する。

$$\text{標本から } |T| = \left| \frac{47.0 - 49.3}{\sqrt{\frac{8 \times 1.4375 + 7 \times 1.1371}{8 + 7 - 2} \left(\frac{1}{8} + \frac{1}{7} \right)}} \right| \doteq \frac{2.2161}{0.633} \doteq 3.63$$

一方, $\varepsilon = 0.01$ とするとき $\frac{\varepsilon}{2} = 0.005$ であるから, 数表より $t_{13}\left(\frac{\varepsilon}{2}\right) = t_{13}(0.005) = 3.012$
 $|T| > t_{13}\left(\frac{\varepsilon}{2}\right)$ となり, 有意水準 0.01 という低い水準でも H_1 は棄却される。
 即ち飼料の違いによる影響があると言ってよい。

4.

$$(1) \phi(t) = E(e^{tX_i^*})$$

$$\phi_n(t) = E(e^{tY_n}) = E\left(e^{t\sqrt{n}(\bar{x} - \mu)/\sigma}\right) = E\left[\exp\left\{t\left(\frac{X_1 + \dots + X_n - n\mu}{n}\right) \cdot \frac{\sqrt{n}}{\sigma}\right\}\right]$$

$$= E\left[\exp\left\{\frac{t}{\sqrt{n}}\left(\frac{X_1 - \mu}{\sigma} + \dots + \frac{X_n - \mu}{\sigma}\right)\right\}\right]$$

$$= E\left[\exp\left(\frac{t}{\sqrt{n}} X_1^*\right)\right] \cdots E\left[\exp\left(\frac{t}{\sqrt{n}} X_n^*\right)\right] = \phi^n\left(\frac{t}{\sqrt{n}}\right) \quad (\because X_i^* \text{ は独立})$$

$$(2) \text{ ところで } \phi(t) = 1 + E(X_i^*)t + \dots + \frac{E(X_i^{*n})}{n!} t^n + \dots$$

$$= 1 + E(X_i^*)t + \frac{E(X_i^{*2})}{2!} t^2 + \vartheta(t^3) \quad \left(\text{ここに } \vartheta(t^3) \text{ は } t^3 \text{ と同位の無限小を意味する}\right)$$

$$E(X_i^*) = E\left\{\frac{X_i - \mu}{\sigma}\right\} = 0 \quad (\because E(X_i) = \mu)$$

$$E(X_i^{*2}) = V(X_i^*) + \{E(X_i^*)\}^2 = \left(\frac{1}{\sigma}\right)^2 E(X_i - \mu)^2 = 1$$

$$\therefore \phi_n(t) = \phi^n\left(\frac{t}{\sqrt{n}}\right) = \left\{1 + 0 \cdot \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \vartheta\left(\frac{t^3}{\sqrt{n}^3}\right)\right\}^n$$

$$\lim_{n \rightarrow \infty} \phi_n(t) = \lim_{n \rightarrow \infty} \left\{1 + \frac{t^2}{2} \cdot \frac{1}{n} + \frac{\vartheta(t^3)}{n\sqrt{n}}\right\}^n$$

$$= \lim_{n \rightarrow \infty} \left[1 + \frac{1}{n} \left\{\frac{t^2}{2} + \frac{\vartheta(t^3)}{\sqrt{n}}\right\}\right]^n = e^{\frac{t^2}{2}}$$

$$\textcircled{\ast} \quad y = \left[1 + \frac{1}{h} \left\{ \frac{t^2}{2} + \frac{\vartheta(t^3)}{\sqrt{h}} \right\} \right]^h \text{とおき対数を取ると } \log y = \frac{\log \left[1 + \frac{1}{h} \left\{ \frac{t^2}{2} + \frac{\vartheta(t^3)}{\sqrt{h}} \right\} \right]}{h^{-1}}$$

ロピタルの定理より

$$\begin{aligned} \lim_{h \rightarrow \infty} \log y &= \lim_{h \rightarrow \infty} \frac{\frac{d}{dh} \log \left[1 + \frac{t^2}{2h} + \frac{\vartheta(t^3)}{h^{3/2}} \right]}{\frac{d}{dh} (h^{-1})} = \lim_{h \rightarrow \infty} \frac{-\frac{t^2}{2h^2} - \left(\frac{3}{2} \right) \frac{\vartheta(t^3)}{h^{5/2}}}{1 + \frac{t^2}{2h} + \frac{\vartheta(t^3)}{h^{3/2}}} \cdot \frac{1}{-\frac{1}{h^2}} \\ &= \lim_{h \rightarrow \infty} \frac{\frac{t^2}{2} + \frac{3}{2} \frac{\vartheta(t^3)}{\sqrt{h}}}{1 + \frac{t^2}{2h} + \frac{\vartheta(t^3)}{h^{3/2}}} = \frac{t^2}{2} \quad \therefore h \rightarrow \infty \text{のとき } y \rightarrow e^{\frac{t^2}{2}} \end{aligned}$$

- (3) $\phi_n(t)$ の極限は $N(0, 1)$ の積率母関数である。個々の受験者の得点の確率分布は分らなくても、相当程度受験者が多い場合 $\mu = m, \frac{\sigma^2}{n} = S^2$ としてこの数学の試験の平均点の確率分布は $N(m, S^2)$ で近似できるという事を示している。