数　学　２　　（問題）

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）　平均μが既知の正規母集団$N(\mu,\ \sigma^2)$の標準偏差σを推定する。
大きさｎの標本変量を$(X_1,\ X_2,\ \cdots\cdots,\ X_n)$とするとき、

統計量$T=\dfrac{C}{n}\sum_{i=1}^{n}|X_i-\mu|$　（Ｃは定数）　とする。

ここで、$C=$［　　］のとき、Ｔはσの不偏推定量となる。

（２）　大きさＮの有限母集団から非復元抽出した大きさｎ（ｎ＜Ｎ）の標本変量を
$(X_1,\ X_2,\ \cdots\cdots,\ X_n)$とするとき、
$X_i,\ X_j\ (i\neq j)$の相関係数ρは［　　］である。

（３）　Ｎ組の標本$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots\cdots,\ (x_N,\ y_N)$に対して、ｙのｘに対する回帰直線（最小２乗直線）を、標本平均$\bar{x}$，$\bar{y}$、標本標準偏差σ_x，σ，および標本相関係数ρを用いて表すと
$y=$［　　］$x+$［　　］。

（４）　互いに異なるｎ枚のカードを復元抽出する。各カードを引く確率が等しく１／ｎであるとする。４回目に初めて前回までに引いたカードのいずれかと同じカードを引いたとき、ｎの最尤推定値は［　　］である。

（５）
母集団が密度関数　$f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta} & (x>0)\\ 0 & (x\leqq 0)\end{cases}$　$(\alpha>0,\ \beta>0)$

で表される分布に従っているとき、この母集団からｎ個の標本値$(x_1,\ x_2,\ \cdots\cdots,\ x_n)$が得られたとすると、$\beta$の最尤推定値は［　　］である。ただし、$\alpha$は既知とする。

２．同様の環境下で、等面積の７枚の畑に大麦を栽培した。肥料の濃度を７段階に分けて与えたところそれぞれの畑で、次のような収穫高を得た。肥料の濃度と収穫高との間に相関関係がない（相関係数$\rho=0$）といえるか。有意水準０．０５で検定せよ。

肥料濃度　X	1	2	3	4	5	6	7
収穫高　　Y	3	8	17	26	20	23	22

必要あれば次の数値を用いよ。

分布値 ／ 自由度ϕ	$t_\phi(0.025)$	$t_\phi(0.050)$	$\chi^2_\phi(0.025)$	$\chi^2_\phi(0.050)$	$\chi^2_\phi(0.950)$	$\chi^2_\phi(0.975)$
4	2.776	2.132	11.14	9.49	0.711	0.484
5	2.571	2.015	12.83	11.07	1.145	0.831
6	2.447	1.943	14.45	12.59	1.635	1.237
7	2.365	1.895	16.01	14.07	2.17	1.690

$u(0.025)=1.960$，$u(0.050)=1.645$　（２０点）

３．次の問に答えよ。

（１）　（０，１）上の一様分布に従う母集団から大きさｎの標本を取り出したとき、標本中の最大値と最小値の差をＺとする。確率変数Ｚの確率密度関数ｇ（ｚ）を求めよ。

（２）　大きさ６５の標本の最大値と最小値の間に、母集団の９５％以上が含まれる確率を求めよ。

（２５点）

４．次の問に答えよ。

（１）　自由度ｎの χ^2 分布の積率母関数は $M(\theta)=(1-2\theta)^{-n/2}$ である。この関数を用いて、自由度ｎの χ^2 分布の平均および分散を求めよ。

（２）　正規母集団 $N(\mu,\ \sigma^2)$ からの標本変量を $X_1,\ X_2,\ \cdots\cdots,\ X_n$ とするとき

$$E\left[(kS^2-\sigma^2)^2\right]$$ を最小にするｋの値を求めよ。

ただし、$S^2=\sum_{i=1}^{n}(X_i-\mu)^2/n$　とする。

（２０点）

数学２　解答例

１．（１）確率変数ＸがＮ（μ，σ^2）に従うとき、$Y=\left\{\dfrac{(X-\mu)}{\sigma}\right\}^2$ は、

自由度１の χ^2 分布に従い、Ｙの確率密度関数は

$$f(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}\,y^{-1/2}e^{-y/2} & (y>0)\\ 0 & (y\leqq 0)\end{cases}$$

$|X-\mu|=\sigma\sqrt{Y}$ より

$$E(|X-\mu|)=\sigma E(\sqrt{Y})=\frac{\sigma}{\sqrt{2\pi}}\int_0^{\infty}y^{1/2}\ y^{-1/2}e^{-y/2}dy$$

$$=\frac{\sigma}{\sqrt{2\pi}}\Big[-2e^{-y/2}\Big]_0^{\infty}=\sqrt{2/\pi\sigma}$$

よって $E(T)=\dfrac{C}{n}\sum_{i=1}^{n}E(|X_i-\mu|)=\dfrac{C}{n}n\sqrt{2/\pi}\,\sigma=\sqrt{2/\pi}\,C\sigma=\sigma$

$\therefore\ C=\boxed{\sqrt{\pi/2}}$

（２）母集団の各要素を a_1，a_2，………，a_N とし、母平均、母分散を μ，σ^2 とする。

$X_i=a_s$，$X_j=a_t$　（$s\neq t$）となる確率は、

$$P(X_i=a_s,\ X_j=a_t)=\frac{1}{N(N-1)}$$

$$E[(X_i-\mu)(X_j-\mu)]=\sum_{s\neq t}^{N}\frac{(a_s-\mu)(a_t-\mu)}{N(N-1)}$$

$$=\frac{1}{N(N-1)}\sum_{s=1}^{N}(a_s-\mu)\{\sum_{t=1}^{N}(a_t-\mu)-(a_s-\mu)\}$$

$$=\frac{-1}{N(N-1)}\sum_{s=1}^{N}(a_s-\mu)^2=-\frac{\sigma^2}{N-1}$$

$$(\because\sum_{t=1}^{N}(a_t-\mu)=0,\ \frac{1}{N}\sum_{s=1}^{N}(a_s-\mu)^2=\sigma^2)$$

$$\therefore\ \rho=\frac{E[(X_i-\mu)(X_j-\mu)]}{\sigma^2}=\boxed{-\frac{1}{N-1}}$$

（３）求める回帰直線を $y=ax+b$ とすると、a，bは

$S(a,b)=\sum_{i=1}^{N}\{y_i-(ax_i+b)\}^2$　を最小にする値である。

$$\therefore\ \frac{\partial S}{\partial a}=-2\sum_{i=1}^{N}\{y_i-(ax_i+b)\}x_i=0$$

$$\frac{\partial S}{\partial b}=-2\sum_{i=1}^{N}\{y_i-(ax_i+b)\}=0$$

$$\therefore \quad \sum_{i=1}^{N} x_i y_i = a\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i \qquad \cdots\cdots①$$

$$\sum_{i=1}^{N} y_i = a\sum_{i=1}^{N} x_i + b\sum_{i=1}^{N} 1 \qquad \cdots\cdots②$$

②より

$$b = \frac{\sum_{i=1}^{N} y_i - a\sum_{i=1}^{N} x_i}{N}$$

これを①に代入して整理すると

$$a = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{N\sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)^2} = \frac{Cov(x, y)}{\sigma_x^2}$$

$$= \boxed{\frac{\sigma_y}{\sigma_x}\rho}$$

$$b = \frac{\sum_{i=1}^{N} y_i - a\sum_{i=1}^{N} x_i}{N} = \boxed{\bar{y} - \frac{\sigma_y}{\sigma_x}\rho\bar{x}}$$

（4） 4回目に初めて同じカードを引く確率 f（n）は

$$f(n) = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \frac{3}{n}$$

$$= \frac{3(n-1)(n-2)}{n^3} \quad (n \geqq 3)$$

$$f'(n) = \frac{3\{(2n-3)n^3 - (n-1)(n-2)3n^2\}}{n^6}$$

$$= \frac{-3(n^2-6n+6)}{n^4}$$

従って、$n=3+\sqrt{3}$ のとき f（n）は最大となる。（∵題意より $n \geqq 2$）
ここでnは整数であり、f（4）=0.28125, f（5）=0.288 であるので、
n = $\boxed{5}$ で最大となる。

（5）（X_1 , X_2 , ……, X_n ）の確率密度関数をβの尤度関数 l（β）とおいて

$$l(\beta) = \prod_{i=1}^{n} f(x_i ; \alpha, \beta)$$

$$= \frac{1}{\Gamma^{(n)}(\alpha)\beta^{n\alpha}} (\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\frac{1}{\beta}\sum_{i=1}^{n} x_i)$$

$$\log\ l(\beta) = -n\log\Gamma(\alpha) - n\alpha\log\beta$$

$$+ (\alpha-1)\sum_{i=1}^{n} \log x_i - \frac{1}{\beta}\sum_{i=1}^{n} x_i$$

$$\frac{\partial}{\partial \beta} \log l(\beta) = -\frac{n\alpha}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} x_i = 0$$

これを解いて求めるβの最尤推定値は

$$\hat{\beta} = \frac{1}{\alpha n}\sum_{i=1}^{n} x_i = \boxed{\frac{\bar{X}}{\alpha}}$$

2．母相関係数をρとして、対立仮説　$H_1 : \rho \neq 0$に対して、
帰無仮説　$H_0 : \rho = 0$を検定する。

標本平均　$E(X) = \frac{1}{n}\sum X_i = 4$，$E(Y) = \frac{1}{n}\sum Y_i = 17$

標本分散　$S(X)^2 = \frac{1}{n}\sum (X_i - \bar{X})^2 = 4$，

$$S(Y)^2 = \frac{1}{n}\sum (Y_i - \bar{Y})^2 = 61.14$$

標本共分散　$Cov(X, Y) = \frac{1}{n}\sum (X_i - \bar{X})(Y_i - \bar{Y}) = 12.857143$

従って、標本相関係数　$r(X, Y)$は

$$r(X, Y) = \frac{Cov(X, Y)}{S(X)S(Y)} = 0.8221321$$

帰無仮説　$H_0 : \rho = 0$のもとで、nを標本数とすると統計量

$$T = \frac{\sqrt{n-2}\,r}{\sqrt{1-r^2}}$$ は、自由度$n-2$のt分布に従う。

いま、$n=7$，$r=0.8221321$ だから、Tの実現値は、

$$T = \frac{\sqrt{7-2}\cdot 0.8221321}{\sqrt{1-0.8221321^2}}$$
$$= 3.2291475 > 2.571 = t_5(0.025)$$

よって、H_0は棄却される。肥料濃度と収穫高との間に関係がないとはいえない。

3．（1）最小値をX_1，最大値をX_nとすると、
これらの同時確率密度関数$F(x_1, x_n)$は

$$F(x_1, x_n) = \begin{cases} \frac{n!}{(n-2)!}(x_n - x_1)^{n-2} & (0 \leqq x_1 \leqq x_n \leqq 1) \\ 0 & (\text{それ以外}) \end{cases}$$

ここで、$Y = X_1$，$Z = X_n - X_1$，これらの同時確率密度関数を$G(y, z)$とする。

$X_1 = Y$，$X_n = Y + Z$より、$\frac{\partial x_1}{\partial y} = 1$，$\frac{\partial x_1}{\partial z} = 0$，$\frac{\partial x_n}{\partial y} = 1$，$\frac{\partial x_n}{\partial z} = 1$

よって　$\left|\frac{\partial(x_1, x_n)}{\partial(y, z)}\right| = 1$

従って、$y \geqq 0,\ z \geqq 0,\ y+z \leqq 1$のとき

$$G(y,\ z) = F(x_1,\ x_n)\left|\frac{\partial(x_1,\ x_n)}{\partial(y,\ z)}\right|$$

$$= \frac{n!}{(n-2)!}(x_n - x_1)^{n-2} \cdot 1$$

$$= n(n-1)z^{n-2}$$

それ以外のとき、$G(y,\ z) = 0$

zの確率密度関数$g(z)$は、

$0 \leqq z \leqq 1$のとき

$$g(z) = \int_0^{1-z} n(n-1)z^{n-2}\,dy = n(n-1)z^{n-2}(1-z)$$

$z<0,\ z>1$のとき　$g(z)=0$

(2)　$\mathrm{Prob}(z \geqq 0.95)$

$$= \int_{0.95}^{1} 65 \cdot 64 \cdot z^{63}(1-z)\,dz = 0.8424$$

よって求める確率は84．2％である。

4．(1) $M'(\theta) = n(1-2\theta)^{-n/2-1}$

$M''(\theta) = n(n+2)(1-2\theta)^{-n/2-2}$

平均$E(X) = M'(0) = n$

分散$V(X) = E(X^2) - [E(X)]^2 = M''(0) - [M'(0)]^2$

$= n(n+2) - n^2 = 2n$

(2) $\displaystyle\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} = \frac{nS^2}{\sigma^2}$ は自由度nのχ^2分布に従う確率変数であるから、

自由度nのχ^2分布の平均はn、分散は2nであることより、

$$E\left(\frac{nS^2}{\sigma^2}\right) = n \quad \therefore E(S^2) = \sigma^2$$

$$V\left(\frac{nS^2}{\sigma^2}\right) = E\left(\frac{n^2 S^4}{\sigma^4}\right) - \left[E\left(\frac{nS^2}{\sigma^2}\right)\right]^2 = \frac{n^2}{\sigma^4}E(S^4) - n^2 = 2n$$

$$\therefore E(S^4) = \sigma^4\frac{2+n}{n}$$

$$E[(kS^2-\sigma^2)^2] = k^2E(S^4) - 2k\sigma^2E(S^2) + \sigma^4$$

$$= k^2\sigma^4\frac{2+n}{n} - 2k\sigma^4 + \sigma^4$$

よって、$k = \dfrac{n}{n+2}$ のとき$E[(kS^2-\sigma^2)^2]$は最小となる。