数　学　２　（問題）

（解答に当り、必要であれば末尾の数表を用いよ。）

１．次の各問の $\boxed{}$ に入る答のみを、所定の解答用紙に記入せよ。　（３０点）

（１）　$X_1, \cdots, X_5$ は区間 $[0,\ \ell]$ 上の一様分布にしたがう母集団からの標本変量とする。$X_{(1)} \leqq \cdots \leqq X_{(5)}$ を対応する順序統計量とし、その標本範囲をRとするとき、$E(R^n) = E\{(X_{(5)} - X_{(1)})^n\} = \boxed{}$ である。

（２）　次の確率密度関数を持つ母集団からのn個の標本変量 $X_1, X_2, \cdots, X_n$ を考える。

$$f(x;\mu,\sigma) = \begin{cases} (1/\sigma)\,exp\{-(x-\mu)/\sigma\} & (x \geqq \mu) \\ 0 & (x < \mu) \end{cases}$$

ここに、$-\infty < \mu < \infty,\ \sigma > 0$ とする。
μ，σ 共に未知とするとき、σ の最尤推定量は $\boxed{}$ である。

（３）　母数 θ を持つ母集団からの統計量Xは次の確率密度関数にしたがう。

$$f(x) = \begin{cases} k\,x^{\theta} & (0 \leqq x \leqq 1) \\ 0 & (x < 0 \ \ or \ \ x > 1) \end{cases}$$

ここに、kは定数である。$x > 0.9$ の区間を棄却域として、仮説 $H_0: \theta = 1$ が検定されるとき、第１種の誤りの起こる確率は $\boxed{①}$ である。
また、対立仮説が $H_1: \theta = 4$ の場合に、第２種の誤りの起こる確率は $\boxed{②}$ である。
（答は四捨五入して小数第２位まで求めよ。）

（４）　２つの正規母集団を考える。１つの母集団から標本数１０の標本を取り出したところ、標本分散は１６であった。もう１つの母集団から標本数９の標本を取り出したとき、標本分散が少なくとも $\boxed{}$ より大きければ２つの母集団の分散は同じでないと言える。
（与えられた数表に基づき有意水準５％の両側検定により、答は四捨五入して小数第１位まで求めよ。）

（５）　x は条件付確率密度関数 $f(x \mid \theta) = 1/\theta \ \ (0 \leqq x \leqq \theta)$ を持つ母集団からの標本とする。θ の確率密度関数は $\xi(\theta) = \theta\,e\,x\,p(-\theta) \ \ (\theta > 0)$ であることが分かっている。
このとき周辺確率密度関数 $f(x) = \boxed{①}$。また x に基づく θ のベイズ推定値は、２乗誤差による損失関数を用いると $E(\theta \mid x) = \displaystyle\int_{\theta > x} \theta\, f(\theta \mid x)\, d\theta$ で与えられる。
$E(\theta \mid x)$ の値は $\boxed{②}$。

２．　正規分布の分散の帰無仮説 $H_0: \sigma^2 = \sigma_0^2$ の有意水準５％の片側検定において、帰無仮説で仮定した分散 σ_0^2 が真の分散の４分の１以下であるときに、この帰無仮説 H_0 が確率９９％以上で棄却されるためには標本数は最低いくつ必要か。ここに、平均、分散は未知とする。
（２０点）

3．　$X_1, X_2, \cdots, X_n$は平均、分散が未知の正規母集団からの標本変量とするとき、次の問に答えよ。

(1) 標本変量分散　$S^2=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2$は分散の最尤推定量であることを示せ。

(2) $S'^2=\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2$は分散の不偏推定量であることを示せ。

（２５点）

4．　平均μ、分散σ^2の値が未知のとき、正規母集団$N(\mu, \sigma^2)$からの大きさnの標本値$x_1, \cdots, x_n$を用いて、母平均についての帰無仮説$H_0: \mu=\mu_0$を有意水準εで検定したい。対立仮説を$H_1: \mu\neq\mu_0$とするとき、尤度比検定を用いてt検定を導け。　（２５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≧λ）＝０．０２５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　　$\lambda=F_{f_1}^{f_2}(0.025)$とも書く。

f_1 ＼ f_2	8	9	10
8	4.43	4.36	4.30
9	4.10	4.03	3.96
10	3.85	3.78	3.72

(2) Xがχ^2分布にしたがうとき、P（X≧λ）＝εとなるλを求める表
nは自由度　　$\lambda=\chi^2_n(\varepsilon)$とも書く。

n ＼ ε	0.99	0.95	0.05	0.01
15	5.23	7.26	25.0	30.6
16	5.81	7.96	26.3	32.0
17	6.41	8.67	27.6	33.4
18	7.01	9.39	28.9	34.8
19	7.63	10.12	30.1	36.2
20	8.26	10.85	31.4	37.6

数学２解答例

１．

(1)　各 Xi の確率密度関数 $f(x)$ は題意より $0 \leqq x \leqq l$ で $1/l$，その他で０である。

一般に，R の確率密度関数 $g(r)$ は，$f(x)$ を用いて次のように表わせる。

$$g(r) = 5 \cdot 4 \int_{-\infty}^{\infty} \left(\int_{x(1)}^{r+x(1)} f(t)dt \right)^3 f(x_{(1)}) \cdot f(r+x_{(1)}) dx_{(1)}$$

$$= 20 \int_0^{l-r} (r/l)^3 \cdot (1/l) \cdot (1/l) dx_{(1)} = \frac{20r^3 \cdot (l-r)}{l^5}$$

$$\therefore \quad E(R^n) = \int_0^l r^n \cdot g(r) dr = \int_0^l 20 \left(\frac{r^{n+3}}{l^4} - \frac{r^{n+4}}{l^5} \right) dr$$

$$= 20 \cdot \left\{ \frac{l^n}{n+4} - \frac{l^n}{n+5} \right\} = \boxed{\frac{20 l^n}{(n+4)(n+5)}}$$

(2)　尤度関数 $L(x_1, \cdots, x_n; \mu, \sigma)$ は次のように表わされる。

$$L(x_1, \cdots, x_n; \mu, \sigma) = \prod_{i=1}^{n} f(x_i; \mu, \sigma) = \sigma^{-n} \exp\left\{ -\left(\sum_{i=1}^{n} x_i - n\mu \right) / \sigma \right\}$$

μ，σ をそれぞれ $-\infty < \mu < \infty$，$\sigma > 0$ の範囲で動かして L を最大にする σ を求める。

$$\frac{\partial L}{\partial \sigma} = -n\sigma^{-n-1} \cdot \exp\{-(\sum_{i=1}^{n} x_i - n\mu)/\sigma\}$$

$$+ \sigma^{-n} \cdot (\sum_{i=1}^{n} x_i - n\mu)/\sigma^2 \cdot \exp\{-(\sum_{i=1}^{n} x_i - n\mu)/\sigma\}$$

$$= \sigma^{-n-2} \cdot \exp\{-(\sum_{i=1}^{n} x_i - n\mu)/\sigma\}(\sum_{i=1}^{n} x_i - n\mu - n\sigma) = 0$$

を解いて $\sigma = (\sum_{i=1}^{n} x_i / n) - \mu$

また，$\dfrac{\partial L}{\partial \mu} = n\sigma^{-n-1} \cdot \exp\{-(\sum_{i=1}^{n} x_i - n\mu)/\sigma\} > 0$ なので，L は μ について単調に増加する。一方 $f(x; \mu, \sigma)$ の定義域を考えれば，$\mu \leqq x_1, \cdots, x_n$ であるから μ は $\min(x_1, \cdots, x_n)$ 以上には大きくとれない。L は $\mu = \min(x_1, \cdots, x_n)$ で極大となる。また，$\sum_{i=1}^{n} x_i - n\mu > n\sigma$ では $\partial L / \partial \sigma > 0$，$\sum_{i=1}^{n} x_i - n\mu < n\sigma$ では $\partial L / \partial \sigma < 0$ であるので σ の最尤推定量は $\sigma = \boxed{\left(\sum_{i=1}^{n} Xi \right) / n - \min(X_1, \cdots, X_n)}$ である。$\boxed{\overline{X} - \min Xi}$ も正解とした。

(3)

$$\int_0^1 f(x) dx = \left[\frac{kx^{\theta+1}}{\theta+1} \right]_0^1 = k/(\theta+1) = 1 \text{ より } k = \theta + 1$$

$\therefore \quad f(x)=(\theta+1)x^{\theta}$

第1種の誤りとはH_0が正しいとき$(\theta=1)$，H_0を正しくないと判定することだから，$x>0.9$の区間を棄却域とすると，その確率は

$$\int_{0.9}^{1} f(x)dx=\int_{0.9}^{1} 2x dx=\boxed{0.19}$$

第2種の誤りとは対立仮説が正しいとき$(\theta=4)$，H_0を正しいとする誤りであるから，その確率は

$$\int_{0}^{0.9} f(x)dx=\int_{0}^{0.9} 5x^4 \mathrm{d}x=0.59049 \Rightarrow \boxed{0.59}$$

(4)

等分散の検定を行なうので，2つの正規母集団を$N(\mu_x,\ \sigma_x^2)$，$N(\mu_y,\ \sigma_y^2)$とし，標本を$(x_1,\ \cdots,\ x_{10})$，$(y_1,\ \cdots,\ y_9)$として不偏分散比を考える。

$$S_x^2=\frac{1}{10}\sum_{i=1}^{10}(x_i-\bar{x})^2=16,\ S_y^2=\frac{1}{9}\sum_{i=1}^{9}(y_i-\bar{y})^2$$

より標本数10の不偏分散は$10 \cdot S_x^2/(10-1)=160/9$

また，標本数9の不偏分散は$9 \cdot S_y^2/(9-1)=(9S_y^2)\ /8$

数表が片側有意水準2.5％のF分布で与えられていること，S_yを大きくして考えることから不偏分散比の分子は標本数9の方を用いる。

即ち，不偏分散比$\{9(S_y^2)/8\}/(160/9)=(81 \cdot S_y^2)/1280$が，$F_9^8(0.025)=4.10$より大なるときに，2つの分散は同じでないと言える。

$\therefore \quad S_y^2 \cdot (81/1280)>4.10$

$\therefore \quad S_y^2>64.790 \Rightarrow \boxed{64.8}$

(5)

xとθの結合密度関数を$g(x,\ \theta)$とおくと，

$0 \leqq x \leqq \theta$，$\theta>0$で$g(x,\ \theta)=\xi(\theta) \cdot f(x \mid \theta)=\theta \cdot e^{-\theta}/\theta=e^{-\theta}$である。

$$\therefore \quad f(x)=\int_{x}^{\infty} g(x,\ \theta)d\theta=[-e^{-\theta}]_x^{\infty}=\boxed{e^{-x}}$$

$$\therefore \quad f(\theta \mid x)=g(x,\ \theta)/f(x)=e^{-\theta}/e^{-x}=e^{x-\theta}$$

$$\therefore \quad E(\theta \mid x)=\int_{x}^{\infty} \theta \cdot f(\theta \mid x)d\theta=\int_{x}^{\infty} \theta \cdot e^{x-\theta}d\theta$$

$$=[\theta \cdot (-e^{x-\theta})]_x^{\infty}+\int_{x}^{\infty} e^{x-\theta}d\theta=\boxed{x+1}$$

なお，$0<\lim_{\theta \to \infty} \theta e^{-\theta}=\lim_{\theta \to \infty}\frac{\theta}{\theta+\theta^2/2+\cdots}<\lim_{\theta \to \infty}\frac{1}{1+\theta/2}=0$を用いた。

2．

一般に平均が未知のとき，真の分散をσ_1^2，標本数をn，標本変量分散をS^2とおくと，仮説H_0：$\sigma^2=\sigma_0^2$は$(ns^2)/\sigma_0^2>\chi^2_{n-1}(0.05)$のとき有意水準5％で棄却される。題意は$\sigma_0^2\leqq\sigma_1^2/4$という見当外れの仮説$H_0$は標本数が増えれば棄却できる筈だが，第2種の誤りの確率1％未満で棄却するには最低いくつの標本を集めれば良いかを求める問題である。

即ち$P\{(ns^2)/\sigma_0^2>\chi^2_{n-1}(0.05)\mid\sigma^2=\sigma_1^2\}\geqq 0.99$ ——①

となるような最小のnを求めればよい。

ここで，$\sigma_1^2=K\sigma_0^2(K\geqq 4)$とすると，

$$P\{(ns^2)/\sigma_1^2\geqq\chi^2_{n-1}(0.99)\}=0.99=P\{(ns^2)/\sigma_0^2\geqq K\chi^2_{n-1}(0.99)\}$$

であるから，$\chi^2_{n-1}(0.05)\leqq K\chi^2_{n-1}(0.99)$ ——②

となるようなK，nを求めれば，確率99％以上で①が満たされる。ところで②式は$K=4$で成立すれば全ての$K\geqq 4$で成り立つから，$K=4$で考えれば十分である。

即ち$\chi^2_{n-1}(0.05)<4\chi^2_{n-1}(0.99)$を満たす最小の$n$を求めればよい。

$$\chi^2_{18}(0.05)=28.9>4\chi^2_{18}(0.99)=4\times 7.01=28.04$$

$$\chi^2_{19}(0.05)=30.1>4\chi^2_{19}(0.99)=4\times 7.63=30.52$$

であるから$n-1$は最低19である。従って標本数は最低20必要である。

3．(1)

尤度関数$L(\mu,\ \sigma^2)$は次の式で表わされる。

$$L(\mu,\ \sigma^2)=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right)$$

これを最大とするμとσ^2を求めるために対数をとって偏微分して得た次の連立方程式を解く。

$$\begin{cases}\dfrac{\partial}{\partial\mu}\log L(\mu,\sigma^2)=\dfrac{\partial}{\partial\mu}\left\{-\dfrac{n}{2}\log 2\pi\sigma^2-\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}=0\\[2ex]\dfrac{\partial}{\partial(\sigma^2)}\log L(\mu,\sigma^2)=\dfrac{\partial}{\partial(\sigma^2)}\left\{-\dfrac{n}{2}\log 2\pi\sigma^2-\dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right\}=0\end{cases}$$

$$\therefore\quad\frac{-1}{\sigma^2}\cdot n\mu+\frac{1}{\sigma^2}\sum_{i=1}^{n}X_i=0\qquad\cdots\cdots①$$

$$\frac{-n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(X_i-\mu)^2=0\qquad\cdots\cdots②$$

①より

$$\mu=\frac{1}{n}\sum_{i=1}^{n}X_i\equiv\overline{X} \qquad \cdots\cdots\cdots\cdots\cdots③$$

③を②に代入して整理すると

$$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=0 \qquad \cdots\cdots\cdots\cdots\cdots④$$

また

$$\frac{\partial^2}{\partial\mu^2}\log L(\mu,\sigma^2)=-n/\sigma^2<0,$$

$$\frac{\partial^2}{\partial(\sigma^2)^2}\log L(\mu,\sigma^2)=\frac{-n}{2(\sigma^2)^2}+\frac{-1}{2(\sigma^2)^3}\sum_{i=1}^{n}(X_i-\overline{X})^2<0$$

であるので，③，④で与えられる極値は最大値を示す。従って題意は示された。

(2)

$$\begin{aligned}S^2&=\{(X_1-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}/n\\&=\{(X_1^2-2X_1\overline{X}+\overline{X}^2)+\cdots+(X_n^2-2X_n\overline{X}+\overline{X}^2)\}/n\\&=(X_1^2+\cdots+X_n^2)/n-\overline{X}^2\end{aligned}$$

であるから

$$ES^2=(EX_1^2+\cdots+EX_n^2)/n-E\overline{X}^2$$

ところで

$$\begin{aligned}E(\overline{X}^2)&=E\{(X_1+\cdots+X_n)/n\}^2=\frac{1}{n^2}\sum_{i=1}^{n}EX_i^2+\frac{1}{n^2}\sum_{i\neq j}E(X_iX_j)\\&=(EX_1^2+\cdots+EX_n^2)/n^2+n(n-1)\mu^2/n^2\end{aligned}$$

($\because EX_i=\mu$，X_i，X_j は独立だから $E(X_iX_j)=EX_iEX_j=\mu^2$)

$\therefore\quad ES^2=(EX_1^2+\cdots+EX_n^2)(n-1)/n^2-n(n-1)\mu^2/n^2$

$$\begin{aligned}\therefore\quad ES'^2&=\{n/(n-1)\}ES^2\\&=(EX_1^2+\cdots+EX_n^2)/n-n\mu^2/n\\&=\{(EX_1^2-\mu^2)+\cdots+(EX_n^2-\mu^2)\}/n\\&=\sigma^2\end{aligned}$$

($\because EX_i=\mu$ だから $(EX_i^2-\mu^2)=\sigma^2$)

従って題意は示された。

4.

尤度関数 $L(x_1,\cdots,x_n;\mu,\sigma^2)=\left(\frac{1}{2\pi\sigma^2}\right)^{n/2}\exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$

のμ，σ^2の最尤推定量は本年問題3(1)より$\mu=\bar{x}$，$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$であるから

$$\max_{\Omega} L(x_1, \cdots, x_n ; \mu, \sigma^2)=\frac{\exp(-n/2)}{(2\pi)^{n/2}}\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{-n/2}$$

また，$\mu=\mu_0$のときのσ^2の最尤推定量は$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2$であるから

$$\max_{H_0} L(x_1, \cdots, x_n ; \mu, \sigma^2)=\frac{\exp(-n/2)}{(2\pi)^{n/2}}\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2\right]^{-n/2}$$

従って，尤度比λは

$$\lambda=\frac{\max_{H_0} L}{\max_{\Omega} L}=\left[\ \{\sum_{i=1}^{n}(x_i-\mu_0)^2\}/\{\sum_{i=1}^{n}(x_i-\bar{x})^2\}\ \right]^{-n/2}$$

と表わせるから，$P(\lambda\leqq\lambda_0 \mid H_0)\leqq\varepsilon$となる$\lambda_0$に対し，標本値による尤度比$\lambda$が$\lambda\leqq\lambda_0$を満たせば$H_0$は棄却される。

$\lambda\leqq\lambda_0 \Leftrightarrow \lambda^{n/2}\leqq\lambda_0^{n/2} \Leftrightarrow \lambda_0^{-2/n}\leqq\lambda^{-2/n}$であるから $C=\lambda_0^{-2/n}$ とおくと，棄却域$\lambda\leqq\lambda_0$は，

$$\text{棄却域}\quad \left\{\sum_{i=1}^{n}(x_i-\mu_0)^2\right\}/\left\{\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}\geqq C \text{——①}$$

と同値である。

①式の左辺の分子は$\sum_{i=1}^{n}(x_i-\mu_0)^2=\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu_0)^2$と変形できるから

①式は

$$n(\bar{x}-\mu_0)^2/\left\{\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}\geqq C-1$$

と変形できる。従って更に次のように変形できる。

$$\text{棄却域}\quad n(\bar{x}-\mu_0)^2\cdot(n-1)/\left\{\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}\geqq(C-1)\cdot(n-1) \text{——②}$$

ここで$U=\{(\bar{x}-\mu_0)\sqrt{n}\}/\sigma$，$V^2=\sum_{i=1}^{n}(X_i-\bar{X})^2/\sigma^2$とおくと，$H_0$の下で$\bar{X}$の平均は$\mu_0$，分散は$\sigma^2/n$である。よって$U$は$N(0,1)$に従う。

また，一般にV^2は自由度$n-1$のχ^2分布に従い，$\bar{X}$(従ってU)と互いに独立である。

$t=U\sqrt{n-1}/\sqrt{V}$とおくと，

$$t=\{\sqrt{n}\,(\overline{X}-\mu_0)\sqrt{n-1}\}\Big/\Big[\sum_{i=1}^{n}(X_i-\overline{X})^2\Big]^{1/2} \text{——③}$$

となり，t は自由度 $n-1$ のステューデントの t 分布に従う。

②式の左辺は③式の t を２乗したものだから②式の棄却域は

棄却域　$t^2 \geqq (C-1)\cdot(n-1)$

と同値になる。

従って，t分布により

$$P(\,|\,t\,|\geqq t_0)=\varepsilon \text{ or } P(t\geqq t_0)=\frac{\varepsilon}{2}$$

となる t_0 を定めておけば棄却域は $|\,t\,|\geqq t_0$ で与えられ，t 検定が導かれた。