数学２（問題）

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。（３５点）

（１）あるサイコロの１の目が出る確率が１／６といえるか、有意水準１％で検定する。このサイコロを４５００回投げたとき、$\boxed{}$回以上１の目が出ると、１の目が出る確率は１／６といえない。

（２）母集団の確率密度関数が、指数分布

$$f(x;\mu)=\begin{cases}(1/\mu)\,e^{-x/\mu} & (x>0)\\ 0 & (x\leqq 0)\end{cases}$$

に従うとき、大きさｎの標本変量平均$\overline{X}$について統計量$Y=C\overline{X}^2$（Ｃは定数）とする。
ここで、$C=\boxed{}$のとき、Ｙはμ^2の不偏推定量である。

（３）ある池から２０匹の魚をとらえて、これに目印をつけて池に放った。しばらくたって、魚を１匹ずつとらえてはもとに戻し、印のついた魚が得られるまで続けた。これに要した回数（最後の回は計算に入れない）をＸとする。
この試行を１０回繰り返して標本$x_1, x_2, \cdots x_{10}$を得たとする。
このとき、池の中の魚の総数Ｎの最尤推定値は$\boxed{}$である。

（４）（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、標本中の最大値を$X_{(n)}$とする。
確率変数　$Y_n=(n+1)(X_{(n)}-n/(n+1))$　の分布関数を　$G_n(y)$としたとき、

$$\lim_{n\to\infty} G_n(y)=\begin{cases}\boxed{} & (y<1)\\ 1 & (y\geqq 1)\end{cases}$$

である。

（５）（０，θ）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、標本中の最大値を$X_{(n)}$とする。
（$\theta>0$とする。）この時、
①$Z=X_{(n)}/\theta$　の確率密度関数は$\boxed{}$である。
②θに対する信頼係数αの最短の信頼区間は$\boxed{}$である。

２．次の問に答えよ。

（１）確率密度関数が

$$f(x)=\begin{cases}C\,x^{n/2-1}\,e^{-x/2} & (x\geqq 0)\\ 0 & (x<0)\end{cases}$$

であるとき定数Ｃの値をガンマ関数$\Gamma(s)=\int_0^\infty u^{s-1}e^{-u}\,du$を用いて表わせ。

（２）正規母集団　$N(\mu,\sigma^2)$からの標本　$X_1, X_2, \cdots X_n$の標本分散を

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2,\ \text{ただし}\ \overline{X}=\frac{X_1+X_2+\cdots+X_n}{n}\ \text{とする。}$$

$T=kS$が母標準偏差σの不偏推定量であるとき、ｋの値をガンマ関数を用いて表わせ。
（２０点）

３．ある県の中学３年生の数学の学力試験の成績は、平均５０、分散１００であった。いま、無作為に１０００人を選び成績を調べたところ次の表のとおりであった。

成　績	３０未満	３０以上 ４０未満	４０以上 ５０未満	５０以上 ６０未満	６０以上 ７０未満	７０以上
人　数	２５	１５６	２９６	３６６	１２９	２８

この結果から、この試験の成績は正規分布に従っているといえるか。有意水準１％および５％で検定せよ。

（２０点）

必要であれば次の数値を用いよ。
標準正規分布

ε	0.159	0.100	0.050	0.025	0.023	0.010	0.005
$u(\varepsilon)$	1.000	1.282	1.645	1.960	2.000	2.326	2.576

χ^2 分布（自由度ϕ，分布値ε）

ϕ ＼ ε	0.995	0.990	0.975	0.950	0.050	0.025	0.010	0.005
3	0.072	0.115	0.216	0.352	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	12.59	14.45	16.81	18.55

４．指数分布（確率密度関数$f(x)=e^{-x}$　$(x>0)$）に従う母集団から抽出したｎ個の標本X_i　$(i=1,\ 2,\ \cdots,\ n)$を大きさの順に並べなおして　$X_{(1)}<X_{(2)}<\cdots<X_{(n)}$とする。ここで、

$$\begin{cases} Y_1=X_{(1)} \\ Y_2=X_{(2)}-X_{(1)} \\ Y_3=X_{(3)}-X_{(2)} \\ \cdots \\ Y_n=X_{(n)}-X_{(n-1)} \end{cases}$$

としたとき、次の問に答えよ。

（１）$Y_1,\ Y_2,\ \cdots,\ Y_n$の同時分布の確率密度関数$g(y_1,\ y_2,\ \cdots,\ y_n)$を求めよ。

（２）$E(Y_i)$　$(i=1,\ 2,\ \cdots,\ n)$を求めよ。

（３）$E(X_{(i)})$　$(i=1,\ 2,\ \cdots,\ n)$を求めよ。

（２５点）

数学２　解答

1．

（１）求める回数をｘとおく。

$$H_0 : p = \frac{1}{6} \qquad H_1 : p \neq \frac{1}{6}$$

$$\hat{p} = \frac{x}{4500}$$

$$|u| = \frac{|\hat{p} - p_0|}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{\left|\dfrac{x}{4500} - \dfrac{1}{6}\right|}{\sqrt{\dfrac{1}{4500} \cdot \dfrac{1}{6} \cdot \dfrac{5}{6}}} > 2.576 = u(0.005)$$

これを解くと

$$x > \left(2.576 \times \frac{1}{180} + \frac{1}{6}\right) \times 4500 = 814.4$$

よって求める回数は $\boxed{815}$ 回以上。

（２）$X_1, X_2, \cdots, X_n$ を母集団からの標本変量とする。

$E(X_i) = \mu$、$E(X_i^2) = 2\mu^2 \quad (1 \leqq i \leqq n)$

また、$i \neq j$ のとき $E(X_i X_j) = E(X_i)E(X_j) = \mu^2$

（$\because X_i, X_j$ は独立）

$$E(\overline{X}^2) = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} E(X_i^2) + \frac{1}{n^2}\sum_{i \neq j} E(X_i X_j)$$

$$= \frac{1}{n^2} n \cdot 2\mu^2 + \frac{1}{n^2} n(n-1)\mu^2$$

$$= \frac{n+1}{n}\mu^2$$

$E(Y) = CE(\overline{X}^2)$

従って　$C = \boxed{\dfrac{n}{n+1}}$

（３）Ｘが従う分布の確率密度関数は

$$f(x;N)=\left(1-\frac{20}{N}\right)^{x}\frac{20}{N}$$

$(x_1, x_2, \cdots, x_{10})$ の同時密度関数は

$$\prod_{i=1}^{10} f(x_i;N)=\left(1-\frac{20}{N}\right)^{\sum_{i=1}^{10} x_i}\left(\frac{20}{N}\right)^{10}$$

であるが、これをＮの尤度関数 $\ell(N)$ とおいて

$$\frac{\partial}{\partial N}\log \ell(N)=\sum_{i=1}^{10} x_i \cdot \frac{\frac{20}{N^2}}{1-\frac{20}{N}}-\frac{10}{N}=0$$

これを解けば、

$$N=\boxed{20+2\sum_{i=1}^{10} x_i}$$

（４）$X_{(n)}$ の分布関数は

$$F(X_{(n)})=X_{(n)}^{\ n}$$

$$\begin{aligned}
G_n(y)&=p\{Y_n<y\}\\
&=p\{(n+1)(X_{(n)}-n/(n+1))<y\}\\
&=p\left\{X_{(n)}<\frac{y+n}{n+1}\right\}
\end{aligned}$$

（$y<1$ のとき）
$$=\left(\frac{y+n}{n+1}\right)^{n}$$
$$=\left(1+\frac{y-1}{n+1}\right)^{n}$$

（$y\geqq 1$ のとき）
$$=1$$

（$n\to\infty$ のとき）$\to \boxed{e^{y-1}}$

（５）① $$f_L(x_{(n)})=n\left(\int_{-\infty}^{x_{(n)}} f(x)\,dx\right)^{n-1} f(x_{(n)})$$

$$=n\left(\frac{x_{(n)}}{\theta}\right)^{n-1}\frac{1}{\theta}$$

∴Ｚの確率密度関数は $\boxed{n z^{n-1}}$

② $\int_a^b n z^{n-1} dz = \alpha \quad (0 \leqq a < b \leqq 1)$ からaとbを定めれば、

$\left(\frac{X_{(n)}}{b}, \frac{X_{(n)}}{a} \right)$ は信頼係数αの信頼区間である。

いま、$b = a + h$ とおくと

$$\int_a^{a+h} n z^{n-1} dz = (a+h)^n - a^n = \alpha$$

aで微分すると

$$n(a+h)^{n-1}\left(1 + \frac{dh}{da}\right) - n a^{n-1} = 0$$

$$\therefore \quad \left(1 + \frac{dh}{da}\right) = \left(\frac{a}{a+h}\right)^{n-1}$$

信頼区間の長さを $l = \frac{X_{(n)}}{a} - \frac{X_{(n)}}{a+h}$ とおくと、

$$\frac{dl}{da} = -\frac{X_{(n)}}{a^2} + \frac{X_{(n)}}{(a+h)^2}\left(1 + \frac{dh}{da}\right)$$

$$= -\frac{X_{(n)}}{a^2} + \frac{X_{(n)}}{(a+h)^2}\left(\frac{a}{a+h}\right)^{n-1}$$

$$= \frac{X_{(n)}}{a^2}\left\{\left(\frac{a}{a+h}\right)^{n+1} - 1\right\} < 0$$

従って l は $a + h = 1$ のとき最小になる。

$$1 - a^n = \alpha$$

$$\therefore \quad a = (1-\alpha)^{1/n}$$

従って求める信頼区間は $\boxed{\left(X_{(n)}, \frac{X_{(n)}}{(1-\alpha)^{1/n}}\right)}$

2. (1) $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} C x^{n/2-1} e^{-x/2} dx$

ここで $\frac{x}{2} = u$ とおいて置換積分を行うと、

$$= C\,2^{n/2} \int_0^{\infty} u^{n/2-1} e^{-u} du = C\,2^{n/2}\,\Gamma\left(\frac{n}{2}\right) = 1$$

$$\therefore \quad C = \frac{1}{2^{n/2}\,\Gamma\left(\frac{n}{2}\right)}$$

（２）$X_1, X_2, \cdots, X_n$ は独立でどれも $N(\mu, \sigma^2)$ に従うから、

$$Y = \frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

は自由度 $n-1$ の χ^2 分布に従う。よって

$$E(\sqrt{Y}) = \int_0^\infty \sqrt{y}\ \frac{1}{2^{(n-1)/2}\,\Gamma\left(\frac{n-1}{2}\right)}\ y^{(n-1)/2-1}\ e^{-y/2}dy$$

$$= \frac{1}{\sqrt{2}\ \Gamma\left(\frac{n-1}{2}\right)}\int_0^\infty \left(\frac{x}{2}\right)^{n/2-1} e^{-y/2}dy$$

ここで $t = \dfrac{y}{2}$ とおいて

$$= \frac{\sqrt{2}}{\Gamma\left(\frac{n-1}{2}\right)}\int_0^\infty t^{n/2-1}\ e^{-t}dt$$

$$= \frac{\sqrt{2}}{\Gamma\left(\frac{n-1}{2}\right)}\Gamma\left(\frac{n}{2}\right)$$

$$E(T) = E(kS) = kE(S) = kE\left(\frac{\sigma}{\sqrt{n}}\sqrt{Y}\right) = k\frac{\sigma}{\sqrt{n}}E(\sqrt{Y})$$

$$= k\frac{\sigma}{\sqrt{n}}\ \frac{\sqrt{2}}{\Gamma\left(\frac{n-1}{2}\right)}\Gamma\left(\frac{n}{2}\right) = \sigma$$

$$\therefore\quad k = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\frac{2}{n}}}$$

3．仮説「数学の学力試験の成績は正規分布N（５０，１００）に従う。」を検定する。正規分布表より理論度数を計算すると、つぎのとおりとなる。

成　績	３０未満	３０以上 ４０未満	４０以上 ５０未満	５０以上 ６０未満	６０以上 ７０未満	７０以上
理論度数	２３	１３６	３４１	３４１	１３６	２３

適合度検定により $\chi^2 = \Sigma \dfrac{(\text{実現値}-\text{理論度数})^2}{\text{理論度数}}$ は自由度５の χ^2 分布に従うと考えてよい。

$$\chi^2 = \frac{(25-23)^2}{23} + \frac{(156-136)^2}{136} + \frac{(296-341)^2}{341} + \frac{(366-341)^2}{341} + \frac{(129-136)^2}{136} + \frac{(25-23)^2}{23} = 12.33$$

$\chi_5{}^2(0.01) = 15.09$ であるから、有意水準１％では仮説は正しいといえるが、

$\chi_5{}^2(0.05) = 11.07$ であるから、有意水準５％では仮説は正しいとは言えない。

4．（１）$X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ の同時分布の確率密度関数は

$$f(x_{(1)}, x_{(2)}, \cdots, x_{(n)}) = \begin{cases} n!\exp(-(x_{(1)} + x_{(2)} + \cdots + x_{(n)})) & (0 < x_{(1)} < x_{(2)} < \cdots < x_{(n)}) \\ 0 & (\text{その他}) \end{cases}$$

ここで　$Y_1 = X_{(1)}$
$Y_2 = X_{(2)} - X_{(1)}$
$Y_3 = X_{(3)} - X_{(2)}$
…
$Y_n = X_{(n)} - X_{(n-1)}$ と変換すると、

$X_{(1)} < X_{(2)} < \cdots < X_{(n)} \Leftrightarrow Y_1 > 0,\ Y_2 > 0,\ \cdots,\ Y_n > 0$

また　$X_{(1)} = Y_1$
$X_{(2)} = Y_1 + Y_2$
$X_{(3)} = Y_1 + Y_2 + Y_3$
…
$X_{(n)} = Y_1 + Y_2 + Y_3 + \cdots + Y_n$

従って　$x_{(1)} + x_{(2)} + \cdots + x_{(n)} = ny_1 + (n-1)y_2 + \cdots + y_n$

また

$$\frac{\partial(x_{(1)}, x_{(2)}, \cdots, x_{(n)})}{\partial(y_1, y_2, \cdots, y_n)} = \begin{vmatrix} \frac{\partial x_{(1)}}{\partial y_1} & \frac{\partial x_{(1)}}{\partial y_2} & \cdots\cdots & \frac{\partial x_{(1)}}{\partial y_n} \\ \frac{\partial x_{(2)}}{\partial y_1} & \cdots\cdots & & \\ \cdots\cdots & & & \\ \frac{\partial x_{(n)}}{\partial y_1} & \cdots\cdots & & \frac{\partial x_{(n)}}{\partial y_n} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & \cdots\cdots & 0 \\ 1 & 1 & 0 & \cdots\cdots & \\ 1 & 1 & 1 & \cdots\cdots & \\ \cdots\cdots & & & & \\ 1 & 1 & 1 & \cdots\cdots & 1 \end{vmatrix} = 1$$

従って、変換のヤコビアン＝1。

以上により$Y_1, Y_2, Y_3, \cdots, Y_n$の同時分布の確率密度関数は

$g(y_1, y_2, \cdots, y_n)$

$$= \begin{cases} n!\exp(-(ny_1 + (n-1)y_2 + \cdots + y_n)) & (y_1 > 0,\ y_2 > 0,\ \cdots,\ y_n > 0) \\ 0 & (\text{それ以外}) \end{cases}$$

（2） $g(y_1, y_2, \cdots, y_n)$

$$= \begin{cases} n\exp(-ny_1)\ \{(n-1)\exp(-(n-1)y_2)\}\ \cdots\exp(-y_n) & (y_1 > 0,\ y_2 > 0,\ \cdots,\ y_n > 0) \\ 0 & (\text{それ以外}) \end{cases}$$

であるから$Y_1, Y_2, \cdots Y_n$は互いに独立で、Y_iの確率密度関数は

$$g_i(y_i) = \begin{cases} (n-i+1)\exp(-(n-i+1)y_i) & (y_i > 0) \\ 0 & (y_i \leqq 0) \end{cases}$$

これは平均 $\dfrac{1}{n-i+1}$ の指数分布の確率密度関数だから

$$E(Y_i) = \frac{1}{n-i+1} \qquad (i = 1,\ 2,\ \cdots,\ n)$$

（3） $E(X_{(i)}) = E(Y_1 + Y_2 + \cdots + Y_i)$

$$= \sum_{k=1}^{i} E(Y_k) = \sum_{k=1}^{i} \frac{1}{n-k+1}$$