数学2（問題）

1．次の各問の$\boxed{\quad\quad}$に入る答のみを、所定の解答用紙に記入せよ。（35点）

（1）ある番組の視聴率を信頼度95％で誤差が3％以内であるように推定したいとき$\boxed{\quad\quad}$人以上のモニターを調査すればよい。

（2）正規母集団N（μ，1）の母平均の検定で、帰無仮説H_0：$\mu=0$、対立仮説H_1：$\mu=1$によって第1種および第2種の誤りをおかす確率をともに0．01以下にしたい。
この検定に要する標本の数は$\boxed{\quad\quad}$以上である。

（3）大きさNの有限母集団から非復元抽出した大きさnの標本を（$X_1, X_2, \cdots, X_n$）とする。
母分散をσ^2とするとき、この母集団の総計推定量

$Z=\dfrac{N}{n}\sum_{k=1}^{n}X_k$ の分散は$V(Z)=\boxed{\quad\quad}\sigma^2$となる。

（4）二項母集団B（p；k）、（ただしpは未知）からとった大きさnの標本を度数によって整理したところ右表の結果を得た。
このときpの最尤推定値は$\boxed{\quad\quad}$である。

標本値	0　1　2　…　j　…　k	合計
度　数	n_0　n_1　n_2　…　n_j　…　n_k	n

（5）母数p（$0<p<1$）の幾何分布に従う母集団から3つの標本X_1, X_2, X_3を抽出し、母数pについての検定を行う。
帰無仮説H_0：$p=1/3$、対立仮説H_1：$p\neq 1/3$、棄却域R：$\max(X_1, X_2, X_3)\geqq 2$
と定めてH_0を検定するとき、有意水準は$\boxed{\quad\quad}$（小数第4位を4捨5入）である。ただし、確率変数Xが母数pの幾何分布に従うとは、非負の整数kにたいして、$P\{X=k\}=p^k(1-p)$となることである。

2．次の問に答えよ。（20点）

（1）区間〔0，θ〕、（$\theta>0$）上の一様分布をもつ母集団から抽出した大きさnの標本$X_1, X_2, \cdots, X_n$に対して統計量　$S=C_1\cdot(X_1+X_2+\cdots+X_n)$、$T=C_2\cdot\max(X_1, X_2, \cdots, X_n)$を考える。
このとき、S，Tともにθの不偏推定量となるように、定数C_1，C_2の値を定めよ。

（2）（1）で求めたθの不偏推定量SとTのどちらがより有効かを示せ。

3．次の問に答えよ。（20点）

（1）母集団分布が指数分布のとき、母平均をμとしてこの母集団からの大きさnの標本に対し、標本平均を$\overline{X}$とするとき、$\left(\dfrac{2n}{\mu}\right)\overline{X}$ が従う確率分布を求めよ。

（2）ある都市で電話の通話時間を15件について調べたところ平均90秒であった。この都市の電話の平均通話時間を信頼度95％で区間推定せよ（小数第2位を4捨5入）。なお、通話時間は指数分布に従うものとする。

4．正規母集団N（μ，σ^2）、（σ^2：未知）からn個の標本を抽出して母平均μを信頼度95％で区間推定するという試行を繰り返し行うことを考える。このとき次の問に答えよ。（25点）

（1）信頼区間の幅の平均を、未知のパラメーターσ、$t_\phi(\varepsilon)$、ガンマ関数$\Gamma(s)$を用いて表わせ。

（2）信頼区間の幅が2σよりも小さくなる確率が95％以上となるためには標本の数nをどの程度の大きさにすればよいか。

必要であれば次の数値を用いよ。

標準正規分布：u（ε）、（上側確率ε）

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

χ^2 分布：$\chi^2_\phi(\varepsilon)$、（自由度φ，上側確率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
7	0.989	1.239	1.690	2.17	14.07	16.01	18.48	20.3
8	1.344	1.646	2.18	2.73	15.51	17.53	20.1	22.0
9	1.735	2.09	2.70	3.33	16.92	19.02	21.7	23.6
10	2.16	2.56	3.25	3.94	18.31	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	19.68	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.12	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.85	31.4	34.2	37.6	40.0
21	8.03	8.90	10.28	11.59	32.7	35.5	38.9	41.4
22	8.64	9.54	10.98	12.34	33.9	36.8	40.3	42.8
23	9.26	10.20	11.69	13.09	35.2	38.1	41.6	44.2
24	9.89	10.86	12.40	13.85	36.4	39.4	43.0	45.6
25	10.52	11.52	13.12	14.61	37.7	40.6	44.3	46.9
26	11.16	12.20	13.84	15.38	38.9	41.9	45.6	48.3
27	11.81	12.88	14.57	16.15	40.1	43.2	47.0	49.6
28	12.46	13.56	15.31	16.93	41.3	44.5	48.3	51.0
29	13.12	14.26	16.05	17.71	42.6	45.7	49.6	52.3
30	13.79	14.95	16.79	18.49	43.8	47.0	50.9	53.7

t分布：$t_\phi(\varepsilon)$、（自由度φ，上側確率ε）

φ ＼ ε	0.050	0.025	0.010	0.005
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055

数学２　解答

1.

（１）全体の視聴率を p，n 人のモニターの視聴率を $\hat{p}$ とすると確率９５％で

$$|p-\hat{p}| \leqq u\left(\frac{0.05}{2}\right)\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}=1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$\hat{p}$ は不明であるが

$$\hat{p}(1-\hat{p})=-\left(\hat{p}-\frac{1}{2}\right)^2+\frac{1}{4}\leqq\frac{1}{4}\qquad(0\leqq\hat{p}\leqq1)$$

であるから、$|p-\hat{p}|\leqq 1.96\sqrt{\dfrac{1}{4n}}\leqq 0.03$　を解けばよい。

従って、　$n\geqq\left(\dfrac{1.96}{0.03}\right)^2\times\dfrac{1}{4}=1067.1\cdots$

求める人数は $\boxed{1068}$ 人以上

（２）第１種の誤りをおかす確率を0.01以下とするため、

$P(\overline{X}\geqq c\mid\mu=0)\leqq 0.01$　すなわち、

$$P(\overline{X}\geqq c\mid\mu=0)=P\left(\frac{\overline{X}-0}{1/\sqrt{n}}\geqq\frac{c}{1/\sqrt{n}}\mid\mu=0\right)$$

$$=\int_{c\sqrt{n}}^{\infty}e^{-1/2x^2}\,dx\leqq 0.01$$

$u(0.01)=2.326$ であるから、$c\sqrt{n}\geqq 2.326$　①

第２の誤りをおかす確率を0.01以下とするため、

$P(\overline{X}<c\mid\mu=1)\leqq 0.01$　すなわち、

$$P(\overline{X}\geqq c\mid\mu=1)=P\left(\frac{\overline{X}-1}{1/\sqrt{n}}\geqq\frac{c-1}{1/\sqrt{n}}\mid\mu=1\right)$$

$$=\int_{\sqrt{n}(c-1)}^{\infty}e^{-1/2x^2}\,dx>0.99$$

$u(0.01)=2.326$ より、$(c-1)\sqrt{n}<-2.326$　②

①－②より、$\sqrt{n}\geqq 2.326\times2=4.652$　$n\geqq 21.641$　答えは $\boxed{22}$

（３）標本平均を$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$とすると

$$V(\overline{X}) = \frac{N-n}{N-1}\frac{\sigma^2}{n}$$

$Z = N\overline{X}$より、$V(Z) = V(N\overline{X}) = N^2 V(\overline{X})$

$$= \boxed{N^2 \frac{N-n}{N-1}\frac{1}{n}}\ \sigma^2$$

（４）尤度函数$L(p) = L(x_1, x_2, \cdots, x_n)$

$$= \prod_{i=1}^{n} P(X = x_i)$$

$$= \{P(X=0)\}^{n_0}\{P(X=1)\}^{n_1}\cdots\{P(X=k)\}^{n_k}$$

$$= \left\{\binom{k}{0} p^0 q^k\right\}^{n_0}\left\{\binom{k}{1} p^0 q^{k-1}\right\}^{n_1}\cdots\left\{\binom{k}{k} p^k q^0\right\}^{n_k}$$

$$= \prod_{j=1}^{k} \binom{k}{j}^{n_j} p^{jn_j} q^{(k-j)n_j}$$　（ただし$q = 1-p$）

$$\log L = \sum_{j=1}^{k}\left\{n_j \log \binom{k}{j} + j n_j \log p + (k-j) n_j \log(1-p)\right\}$$

両辺をpで微分して

$$\frac{\partial \log L}{\partial p} = \frac{1}{p}\sum_{j=1}^{k} j n_j - \frac{1}{1-p}\sum_{j=1}^{k}(k-j) n_j = 0$$

これをpについて解くと、

$$(1-p)\sum_{j=1}^{k} j n_j = p \sum_{j=1}^{k}(k-j) n_j$$

$$\sum_{j=1}^{k} j n_j = pk\sum_{j=1}^{k} n_j = pkn$$

$$\therefore\quad p = \boxed{\frac{1}{kn}\sum_{j=1}^{k} j n_j}$$

（５）$P(X \geqq n) = \sum_{k=n}^{\infty} p^k (1-p)$

$= p^n (1-p) \sum p^k = p^n$であるから、

$$P(X < n) = 1 - p^n$$

ところで、標本X_1, X_2, X_3について

$$P\{\max(X_1, X_2, X_3) \geqq 2\} = 1 - P\{\max(X_1, X_2, X_3) < 2\}$$

$$= 1 - P\{X < 2\}^3$$

有意水準εは帰無仮説H_0が正しいときにこれを棄却してしまう確率であるから、

$$\varepsilon = 1 - P\{X < 2\}^3 = 1 - (1 - p^2)^3 = \frac{217}{729} \fallingdotseq \boxed{0.298}$$

2. $E(S) = a\sum_{i=1}^{n} E(X_i)$

ここで$E(X_i) = \int_0^\theta \frac{1}{\theta} x\,dx = \frac{1}{\theta}\left[\frac{x^2}{2}\right]_0^\theta = \frac{\theta}{2}$　$(i = 1, \cdots, n)$

よって、$E(S) = C_1 \cdot \frac{\theta}{2} \cdot n = \theta$より、$C_1 = \frac{2}{n}$

次に$U = \max(X_1, X_2, \cdots, X_n)$とおくと、

$$P\{U \leqq x\} = P\{X_1 \leqq x, X_2 \leqq x, \cdots, X_n \leqq x\}$$
$$= P\{X_1 \leqq x\} P\{X_2 \leqq x\} \cdots P\{X_n \leqq x\}$$
$$= \left(\frac{x}{\theta}\right)^n \quad (0 \leqq x \leqq \theta)$$

よって、$E(T) = C_2 \int_0^\theta \left[\left(\frac{x}{\theta}\right)^n\right]' x\,dx = C_2 \cdot \frac{n}{\theta^n}\int_0^\theta x^n\,dx$

$$= C_2 \cdot \frac{n}{\theta^n}\left[\frac{x^{n+1}}{n+1}\right]_0^\theta = C_2 \cdot \frac{n}{n+1}\theta = \theta \text{より}$$

$$C_2 = \frac{n+1}{n}$$

(2) $V(S) = V\left(\frac{2}{n}(X_1 + \cdots + X_n)\right) = \frac{4}{n^2}\sum_{i=1}^{n} V(X_i)$

$$= \frac{4}{n^2} \cdot n \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

$$V(S) = V\left(\frac{n+1}{n} \cdot U\right) = \left(\frac{n+1}{n}\right)^2 V(U)$$

$$= \left(\frac{n+1}{n}\right)^2 \{E(U^2) - (E(U))^2\}$$

$$= \left(\frac{n+1}{n}\right)^2 \left\{ \int_0^\theta x^2 \cdot \frac{n x^{n-1}}{\theta^n} dx - \left(\frac{n}{n+1} \theta \right)^2 \right\}$$

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

$$= \frac{\theta^2}{n(n+2)}$$

よって、$V(S) - V(T) = \dfrac{\theta^2}{3n} - \dfrac{\theta^2}{n(n+2)}$

$$= \frac{n-1}{3n(n+2)} \theta^2 \geqq 0$$ より

$V(S) \geqq V(T)$ となりTがより有効である。

3．

（1）X_i の確率密度函数は

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

$X = X_1 + X_2 + \cdots + X_n$ の確率密度函数を $f_n(x)$ とすると、

$$f_1(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$$

$$f_2(x) = \int_0^x f_1(t) f_1(x-t) dt$$

$$= \int_0^x \frac{1}{\mu} e^{-\frac{t}{\mu}} \frac{1}{\mu} e^{-\frac{x-t}{\mu}} dt$$

$$= \frac{1}{\mu^2} e^{-\frac{x}{\mu}} x$$

$f_{n-1}(x) = \dfrac{1}{\mu^{n-1}} \dfrac{1}{(n-2)!} x^{n-2} e^{-\frac{x}{\mu}}$ とすると、

$$f_n(x) = \int_0^x \frac{1}{\mu^{n-1}} \frac{1}{(n-2)!} t^{n-2} e^{-\frac{t}{\mu}} \frac{1}{\mu} e^{-\frac{x-t}{\mu}} dt$$

$$= \frac{1}{\mu^n} e^{-\frac{x}{\mu}} \frac{1}{(n-2)!} \int_0^x t^{n-2} dt$$

$$= \frac{1}{\mu^n} e^{-\frac{x}{\mu}} \frac{1}{(n-1)!} \Big[t^{n-1} \Big]_0^x$$

$$= \frac{1}{\mu^n} \frac{1}{(n-1)!} x^{n-1} e^{-\frac{x}{\mu}}$$

よって数学的帰納法により、

$$f_n(x) = \frac{1}{\mu^n} \frac{1}{(n-1)!} x^{n-1} e^{-\frac{x}{\mu}} \qquad (x > 0)$$

$$\frac{2n}{\mu} \overline{X} = \frac{2}{\mu} X = Z \quad \text{とすると、} \quad \frac{dZ}{dX} = \frac{2}{\mu} \quad \text{より}$$

Zの確率密度函数 $g(z)$ は

$$g(z) = \frac{1}{\mu^n} \frac{1}{\Gamma(n)} \left(\frac{\mu}{2} z\right)^{n-1} e^{-\frac{1}{\mu}\frac{\mu}{2}z} \frac{\mu}{2}$$

$$= \frac{1}{2^n} \frac{1}{\Gamma(n)} z^{n-1} e^{-\frac{z}{2}}$$

$$= \frac{1}{2^{\frac{2n}{2}}} \frac{1}{\Gamma\left(\frac{2n}{2}\right)} z^{\frac{2n}{2}-1} e^{-\frac{z}{2}}$$

∴これは自由度 $2n$ の χ^2 分布に従う。

〔別解〕 X_i の確率密度函数は

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

$Y_i = \frac{2}{\mu} X_i$ なる変数変換を行うとき、Y_i の確率密度函数は

$$g(y) = \frac{\mu}{2} f\left(\frac{\mu}{2} y\right) = \frac{1}{2} e^{-y/2} \qquad (y > 0)$$

であり、これは自由度2の χ^2 分布の確率密度函数である。

$Y_1, Y_2, \cdots, Y_n$ は各々独立にこの分布に従う確率変数であることから χ^2 分布の再生性により、

$$Y_1 + Y_2 + \cdots + Y_n = \frac{2}{\mu} (X_1 + X_2 + \cdots + X_n) = \frac{2n}{\mu} \overline{X} \text{は}$$

自由度 $2n$ の χ^2 分布に従う。

（２）（１）より $\frac{2n}{\mu}\overline{X}$ は自由度２ｎの χ^2 分布に従うことから、信頼係数を（１－α）とすると

$$P\left\{\chi^2_{2n}\left(1-\frac{\alpha}{2}\right) \leqq \frac{2n}{\mu}\overline{X} \leqq \chi^2_{2n}\left(\frac{\alpha}{2}\right)\right\} \geqq 1-\alpha \quad \text{より}$$

$$\frac{2n\overline{X}}{\chi^2_{2n}\left(\frac{\alpha}{2}\right)} \leqq \mu \leqq \frac{2n\overline{X}}{\chi^2_{2n}\left(1-\frac{\alpha}{2}\right)}$$ が求める信頼区間である。

$n=15,\ \overline{X}=90,\ \alpha=0.05$ より

$$\frac{2\times15\times90}{\chi^2_{30}(0.025)} \leqq \mu \leqq \frac{2\times15\times90}{\chi^2_{30}(0.975)}$$

$$57.4\text{秒} \leqq \mu \leqq 160.8\text{秒}$$

４．

（１）μ の信頼区間を（μ_l, μ_u）とすると、

$$\mu_l = \overline{x} - t_{n-1}\left(\frac{\varepsilon}{2}\right)\frac{s}{\sqrt{n-1}}$$

$$\mu_u = \overline{x} + t_{n-1}\left(\frac{\varepsilon}{2}\right)\frac{s}{\sqrt{n-1}}$$

ここに

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

であるから信頼区間の幅Ｌは

$$L = \frac{2t_{n-1}(\varepsilon/2)}{\sqrt{n(n-1)}}\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

$Y = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2$ が自由度 $n-1$ の χ^2 分布に従うことを用いる。

$L = \alpha\sigma Y^{1/2}$（α は定数 $= \frac{2t_{n-1}(\varepsilon/2)}{\sqrt{n(n-1)}}$）であるから、

$$E[L] = \alpha\sigma E[Y^{1/2}]$$

$$E\left[Y^{1/2}\right]=\int_0^{\infty}\frac{x^{(n-1)/2-1}\,e^{-x/2}}{\Gamma\left((n-1)/2\right)2^{(n-1)/2}}\,x^{1/2}\,dx$$

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

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

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

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

$$\therefore E\left[L\right]=\frac{2t_{n-1}(\varepsilon/2)}{\sqrt{n(n-1)}}\cdot\sigma\cdot\frac{2^{1/2}\,\Gamma(n/2)}{\Gamma\left((n-1)/2\right)}$$

$$=\frac{2\sigma t_{n-1}(\varepsilon/2)}{\sqrt{n(n-1)/2}}\cdot\frac{\Gamma(n/2)}{\Gamma\left((n-1)/2\right)}$$

（２）$P\{L\leqq 2\sigma\}$

$=P\{\alpha\sigma Y^{1/2}\leqq 2\sigma\}$

$=P\left\{Y^{1/2}\leqq\dfrac{2}{\alpha}\right\}=P\left\{Y\leqq\dfrac{4}{\alpha^2}\right\}\geqq 0.95$となれば良いが、

Ｙは自由度ｎ－１のχ^2分布に従うから、求める条件は

$\dfrac{4}{\alpha^2}\geqq\chi^2{}_{n-1}(0.05)$ すなわち、$\dfrac{n(n-1)}{t_{n-1}{}^2(0.025)}\geqq\chi^2{}_{n-1}(0.05)$

これを計算すると

n	8	9	10	11	12
$\chi^2{}_{n-1}(0.05)$	14.07	15.51	16.92	18.31	19.68
$\dfrac{n(n-1)}{t_{n-1}{}^2(0.025)}$	10.01	13.54	17.58	22.16	27.25

∴ｎは１０以上必要である。