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) 確率密度関数 $f(x,\theta)=\theta x^{\theta-1}$ $(0<x<1)$ を持つ母集団から2個の標本を無作為に取り出し、帰無仮説 $H_0:\theta=1$、対立仮説 $H_1:\theta=2$ の検定を行う。棄却域 C を $C=\{(X_1,X_2):\ 4X_2\leq 3X_1\}$ としたとき、この検定の第1種の誤りの起こる確率は $\boxed{}$ 、第2種の誤りの起こる確率は $\boxed{}$ である。

(2) ある銘柄の洗剤を使っている消費者の比率を推定したい。このとき、次の条件の下では、少なくとも $\boxed{}$ 個の標本が必要である。

（条件）母集団比率 p は0.8と0.9の範囲にあることがわかっている。また、標本を十分多くとったときに、少なくとも0.99の確率で標本比率 $\hat{p}$ と母集団比率 p との差を0.02以下とする。

(3) ダイレクトメールにより広告を行い、通信販売をしているある会社で、新しい広告を行うことになり、試験的に3,000通のダイレクトメールを発送することになった。過去の返信率が10%であったとき、この新しい広告の返信率が過去の返信率を上回ると有意水準5%で判断できるのは、今回 $\boxed{}$ 通以上の返信があったときである。

(4) ある品種Aの種子100個と品種Bの種子150個を同一条件下でまき、両品種の発芽率の差を調べたところ、3週間後、品種Aは85個が発芽し、品種Bは141個が発芽した。両品種の発芽率の差を信頼係数95%で区間推定すると、信頼区間は（ $\boxed{}$ ， $\boxed{}$ ）である（小数点以下第4位四捨五入）。

次に、別の品種Cの種子の発芽率を調べるため、13個の種子をまいて実験したところ、4個が発芽した。品種Cの発芽率を信頼係数90%で区間推定すると、信頼区間は（ $\boxed{}$ ， $\boxed{}$ ）である（小数点以下第3位四捨五入）。

(5) 右のデータは、ある工場において原料に含まれる成分Aの含有率 x(%)と、その原料を用いて作った製品中の成分Aの含有率 y(%)を測定したものである。このデータを回帰分析により分析して、y の x に対する回帰式を求めると、y= $\boxed{}$ x+ $\boxed{}$ となる（小数点以下第2位四捨五入）。

測定回数	1回	2回	3回	4回	5回	6回	7回	8回	合計
x	86	82	85	93	95	88	90	94	713
y	91	92	91	86	86	89	88	87	710
x^2	7,396	6,724	7,225	8,649	9,025	7,744	8,100	8,836	63,699
xy	7,826	7,544	7,735	7,998	8,170	7,832	7,920	8,178	63,203
y^2	8,281	8,464	8,281	7,396	7,396	7,921	7,744	7,569	63,052

(6) 母集団のN個の要素$x_1,x_2,\cdots,x_N$を$N_1,N_2,\cdots,N_k$個 $(N_1+N_2+\cdots+N_k=N)$のk組の部分集団（層）に分けて、各層より無作為に標本を抽出して、母集団の性質を調査する方法を層別抽出法という。母集団の平均をm、分散をσ^2とし、各層の平均を$m_1,m_2,\cdots,m_k$、分散を$\sigma_1^2,\sigma_2^2,\cdots,\sigma_k^2$とするとき、次の関係式が成り立つ：

$$\sigma^2=\sum_{i=1}^{k}\frac{N_i}{N}\{\boxed{}\}$$

2．ある合成物の原料については成分Aの含有率が96%以上であることが必要であるとされている。そこで納入された原料に対し、抜取検査を行うことにした。このとき、次の問に答えよ。(20点)

(1) ある原料について、6回の分析を行ったところ、次のような結果が得られた。この原料の成分Aの含有率は96%以上であるといえるか、有意水準5%で検定せよ。

分析結果：96, 93, 95, 97, 94, 92 (%)

(2) 含有率の分析値は分散$\sigma^2=(2.0\%)^2$の誤差を持っていることがわかっているとする。このとき、含有率96%以上の原料は確率95%で合格し、含有率94%以下の原料は確率90%で合格しないようにするためには、納入された原料に対し何回の分析を行い、どのように合否を決めればよいか。

3．区間$[0,\theta]$ $(\theta>0)$上の一様分布を持つ母集団（θは未知とする）から抽出した、大きさnの標本$X_1,X_2,\cdots,X_n$に対して、確率変数Yを$Y=\max(X_1,X_2,\cdots,X_n)$と定義する。このとき、次の問に答えよ。(20点)

(1) $E\left\{(aY-\theta)^2\right\}$を計算せよ。ただし、aは定数とする。

(2) $X_1,X_2,\cdots,X_n$の関数$\delta(X)$を使ってθを推定するとき、$E\left\{(\delta(X)-\theta)^2\right\}$が小さい方の推定量が良い推定量であると考える。

今、$\delta(X)$として、$\delta_1(X)=a^*Y$と$\delta_2(X)=\dfrac{2}{n}\sum_{i=1}^{n}X_i$の2つの推定量を考えるとき、どちらが良い推定量であるか。ここに、a^*は、(1)で求めた$E\left\{(aY-\theta)^2\right\}$でaを動かしたとき、その最小値を与えるaとする。

4．2つの正規母集団$A:N(\mu_A,\sigma_A^2)$, $B:N(\mu_B,\sigma_B^2)$から、それぞれ大きさn_A,n_Bの標本$(X_1,X_2,\cdots,X_{n_A})$, $(Y_1,Y_2,\cdots,Y_{n_B})$を抽出した。このとき、次の問に答えよ。(25点)

(1) $m=\mu_A-\mu_B$の最尤推定量を求めよ。

(2) 2つの分散が等しい$(\sigma_A^2=\sigma_B^2=\sigma^2)$ことが既知であった場合の$\sigma^2$の最尤推定量$\sigma_L^2$を求めよ。さらに、$E(k\cdot\sigma_L^2)=\sigma^2$が成り立つためのkを求めよ。

（付表）必要であれば次の数値を用いよ。

1．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

2．t 分布：$t_{\phi}(\varepsilon)$　（自由度φ、上側確率ε）

φ＼ε	0.100	0.050	0.025	0.010	0.005
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499

3．F 分布：$F_{f_1}^{f_2}(\varepsilon)$　（分母の自由度 f_1、分子の自由度 f_2、上側確率ε）

$\varepsilon = 0.05$

f_1＼f_2	5	6	7	8	9	10	20	30	40
5	5.05	4.95	4.88	4.82	4.77	4.74	4.56	4.50	4.46
6	4.39	4.28	4.21	4.15	4.10	4.06	3.87	3.81	3.77
7	3.97	3.87	3.79	3.73	3.68	3.64	3.44	3.38	3.34
8	3.69	3.58	3.50	3.44	3.39	3.35	3.15	3.08	3.04
9	3.48	3.37	3.29	3.23	3.18	3.14	2.94	2.86	2.83
10	3.33	3.22	3.14	3.07	3.02	2.98	2.77	2.70	2.66
12	3.11	3.00	2.91	2.85	2.80	2.75	2.54	2.47	2.43
14	2.96	2.85	2.76	2.70	2.65	2.60	2.39	2.31	2.27
16	2.85	2.74	2.66	2.59	2.54	2.49	2.28	2.19	2.15
18	2.77	2.66	2.58	2.51	2.46	2.41	2.19	2.11	2.06
20	2.71	2.60	2.51	2.45	2.39	2.35	2.12	2.04	1.99
30	2.53	2.42	2.33	2.27	2.21	2.16	1.93	1.84	1.79
40	2.45	2.34	2.25	2.18	2.12	2.08	1.84	1.74	1.69

以　上

数学２　解答

１．

(1)	第１種の誤りの起こる確率 $\frac{3}{8}$	第２種の誤りの起こる確率 $\frac{23}{32}$
(2)	2, 655	
(3)	328	
(4)	品種Ａと品種Ｂの発芽率の差の信頼区間 （-0.170, -0.010）	品種Ｃの発芽率の信頼区間 （0.11, 0.57）
(5)	$y = -0.5x + 132.9$	
(6)	$\sigma_i^2 + (m_i - m)^2$	

（１）第１種の誤りの起こる確率：

$$P\left[(X_1, X_2) \in C \mid H_0\right] = \int_0^1 \int_0^{\frac{3}{4}x_1} \theta^2 (x_1 x_2)^{\theta-1} dx_2 dx_1 = \int_0^1 \int_0^{\frac{3}{4}x_1} dx_2 dx_1 \ (\because \theta = 1)$$

$$= \int_0^1 \frac{3}{4} x_1 dx_1 = \left[\frac{3}{8} x_1^2\right]_0^1 = \frac{3}{8}$$

第２種の誤りの起こる確率：

$$1 - P\left[(X_1, X_2) \in C \mid H_1\right] = 1 - \int_0^1 \int_0^{\frac{3}{4}x_1} 4 x_1 x_2 dx_2 dx_1 = 1 - 4\int_0^1 \left[\frac{1}{2} x_1 x_2^2\right]_0^{\frac{3}{4}x_1} dx_1$$

$$= 1 - 4\int_0^1 \frac{9}{32} x_1^3 dx_1 = 1 - \frac{9}{8}\left[\frac{1}{4} x_1^4\right]_0^1 = 1 - \frac{9}{32} = \frac{23}{32}$$

（２）標本数ｎが大きい場合、近似的に $Z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$ $(q = 1 - p)$ は標準正規分布に従

うので、$P\left(|Z| < u\left(\frac{0.01}{2}\right)\right) = 0.99$ より $P\left(|\hat{p} - p| < 2.576 \times \sqrt{\frac{pq}{n}}\right) = 0.99$ が満たされる。

また、（条件）より $P(|\hat{p} - p| < 0.02) \geq 0.99$ でなければいけないから、

$2.576 \times \sqrt{\frac{pq}{n}} \leq 0.02$ であるので、$n \geq \left(\frac{2.576}{0.02}\right)^2 pq$ が成立する。

pq=p(1-p)は $0.8 \leq p \leq 0.9$ では p=0.8 のとき最大値を取るから、

$$n \geq \left(\frac{2.576}{0.02}\right)^2 \times 0.8 \times 0.2 = 2654.3104$$

よって、少なくとも標本は 2,655 個以上必要である。

（３）題意より有意水準５％の片側検定であるから、新しい広告の返信率を $\hat{p}$、過去の返信率を p_0 とすれば、$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} \geq u(0.05)$ となればよい。

$p_0 = 0.1,\ u(0.05) = 1.645,\ n = 3000$ であるから、$\dfrac{\hat{p} - 0.1}{\sqrt{\dfrac{0.1 \times (1-0.1)}{3000}}} \geq 1.645$

よって、$\hat{p} \geq 1.645 \times \sqrt{\frac{0.1 \times 0.9}{3000}} + 0.1$

従って、必要な返信数は、

$$3000 \times \hat{p} \geq 3000 \times \left(1.645 \times \sqrt{\frac{0.1 \times 0.9}{3000}} + 0.1\right) = 327.030\cdots \rightarrow 328 \text{ 通以上}$$

（４）前半：品種Ａ，Ｂの母集団発芽率を p_1, p_2、標本数を n_1, n_2、標本発芽率を $\hat{p}_1, \hat{p}_2$ とすると、信頼係数 95%の $p_1 - p_2$ の信頼区間 $\left(\hat{\delta}_L, \hat{\delta}_U\right)$ は、

$$\hat{\delta}_L = \hat{p}_1 - \hat{p}_2 - u\left(\frac{0.05}{2}\right)\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\hat{\delta}_U = \hat{p}_1 - \hat{p}_2 + u\left(\frac{0.05}{2}\right)\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

となる。

$n_1 = 100,\ n_2 = 150,\ \hat{p}_1 = \frac{85}{100} = 0.85,\ \hat{p}_2 = \frac{141}{150} = 0.94,\ u(0.025) = 1.960$ より、

$$\hat{\delta}_L = 0.85 - 0.94 - 1.960 \times \sqrt{\frac{0.85 \times (1-0.85)}{100} + \frac{0.94 \times (1-0.94)}{150}} = -0.1696396\cdots$$

$$\hat{\delta}_U = 0.85 - 0.94 + 1.960 \times \sqrt{\frac{0.85 \times (1-0.85)}{100} + \frac{0.94 \times (1-0.94)}{150}} = -0.0103603\cdots$$

→求める信頼区間は(-0.170, -0.010)

後半：品種Cの標本数を n、標本発芽率を $\hat{p}$ とすると、$k = n\hat{p} = 4 < 5$ であるので、精密法を適用する。この場合、信頼係数 90%の信頼区間 $(\hat{p}_L, \hat{p}_U)$ は、

$$\hat{p}_L = \frac{n_2}{n_1 F_{n_2}^{n_1}\left(\frac{0.10}{2}\right) + n_2}$$

$$\hat{p}_U = \frac{m_2 F_{m_2}^{m_1}\left(\frac{0.10}{2}\right)}{m_1 F_{m_2}^{m_1}\left(\frac{0.10}{2}\right) + m_2}$$

となる。

ここで、

$$n_1 = 2(n-k+1) = 20,\ n_2 = 2k = 8$$
$$m_1 = 2(k+1) = 10,\ m_2 = 2(n-k) = 18$$

である。

$F_{n_2}^{n_1}(0.05) = F_8^{20}(0.05) = 3.15,\ F_{m_2}^{m_1}(0.05) = F_{18}^{10}(0.05) = 2.41$ より、

$$\hat{p}_L = \frac{8}{20 \times 3.15 + 8} = 0.11267\cdots$$

$$\hat{p}_U = \frac{10 \times 2.41}{10 \times 2.41 + 18} = 0.57244\cdots$$

→求める信頼区間は(0.11,0.57)

（5）求める回帰式を $y-\bar{y}=\beta(x-\bar{x})$ とすると、$\bar{x}=\frac{713}{8}=89.125,\ \bar{y}=\frac{710}{8}=88.75$ で

あり、$\beta=\frac{\sum_{i=1}^{8}x_i y_i - n\bar{x}\cdot\bar{y}}{\sum_{i=1}^{8}x_i{}^2-n\bar{x}^2}=\frac{63203-8\times 89.125\times 88.75}{63699-8\times 89.125^2}=\frac{-75.75}{152.875}$

従って、求める回帰式は $y-88.75=\left(\frac{-75.75}{152.875}\right)(x-89.125)$ より、

$y=-0.5x+132.9$

（6）第 i 層の要素を $\left(x_{i,1},x_{i,2},\cdots,x_{i,N_i}\right)$ とすれば、

$$\sigma^2=\frac{1}{N}\sum_{i=1}^{k}\left\{\sum_{j=1}^{N_i}\left(x_{i,j}-m\right)^2\right\}=\frac{1}{N}\sum_{i=1}^{k}N_i\frac{\sum_{j=1}^{N_i}\left(x_{i,j}-m\right)^2}{N_i}$$

$$=\frac{1}{N}\sum_{i=1}^{k}N_i\frac{\sum_{j=1}^{N_i}\left\{\left(x_{i,j}-m_i\right)^2+\left(m_i-m\right)^2+2\left(x_{i,j}-m_i\right)\left(m_i-m\right)\right\}}{N_i}$$

$$=\frac{1}{N}\sum_{i=1}^{k}N_i\left\{\sigma_i^2+\left(m_i-m\right)^2\right\}=\sum_{i=1}^{k}\frac{N_i}{N}\left\{\sigma_i^2+\left(m_i-m\right)^2\right\}$$

2．

（1） 6回の分析結果の平均 $\bar{x}$ 、分散 s_x^2 は、分析値を $x_1,x_2,\cdots x_6$ とおくと、

$$\bar{x}=\frac{1}{6}\sum_{i=1}^{6}x_i=94.5,\ s_x^2=\frac{1}{6}\sum_{i=1}^{6}x_i^2-\bar{x}^2=2.916$$

ここで、$\bar{x}<96$ であるので、帰無仮説 H_0: $\mu=\left(96\%=\mu_0\right)$ を対立仮説 H_1: $\mu<\mu_0$

に対して検定する。母分散が未知であるので、ｔ－検定（片側検定）を行なう。

今、$t_{n-1}(0.05)=t_4(0.05)=2.015$であるから、

$$\mu_0-t_{n-1}(0.05)\cdot\frac{s_x}{\sqrt{n-1}}=96-2.015\times\sqrt{\frac{2.916}{5}}=94.46<94.5=\bar{x}$$

よって、H_0は棄却されないので、含有率は96%未満であるとは認められず、96%以上であることを否定することはできない。

（２）帰無仮説H_0: $\mu=96\%(=\mu_0)$（または$\mu\geq 96\%$）、対立仮説H_1: $\mu<\mu_0$として、ｎ回の分析による分析値の平均値$\bar{x}$について、H_0の下では、分散σ^2が既知の場合、標本変量平均$\bar{X}$は$N\left(\mu_0,\frac{\sigma^2}{n}\right)$に従う。

したがって、

$$\bar{x}<\mu_0-u(\varepsilon)\cdot\frac{\sigma}{\sqrt{n}}$$ のときH_0を棄却し、

$$\bar{x}>\mu_0-u(\varepsilon)\cdot\frac{\sigma}{\sqrt{n}}$$ のときH_0を採択する

ようにすればよい。

ここで、$\mu_0=96\%$の原料の合格の確率が95%であることより、危険率$\varepsilon=1-0.95=0.05$である。

また、含有率94%の原料の合格する確率は$1-0.90=0.10$であるから、このとき

$$P\left(\bar{X}>\mu_0-u(\varepsilon)\cdot\frac{\sigma}{\sqrt{n}}\right)=0.10,\quad \varepsilon=0.05$$

とならなければならない。

ここで、$\bar{X}$の分布は平均$\mu=94\%$、標準偏差$\sigma/\sqrt{n}$の正規分布にしたがっているので、これを標準化して、$U=\frac{\bar{X}-94}{\sigma/\sqrt{n}}$とおけば、上式は

$$P\left(\frac{\bar{X}-94}{\sigma/\sqrt{n}}>\frac{\mu_0-94-u(\varepsilon)\cdot\sigma/\sqrt{n}}{\sigma/\sqrt{n}}\right)=0.10,\quad \varepsilon=0.05$$

となる。これからnを定めればよいから、

$$u(0.10)=\frac{\mu_0-94-u(0.05)\cdot\sigma/\sqrt{n}}{\sigma/\sqrt{n}}=\frac{\mu_0-94}{\sigma/\sqrt{n}}-u(0.05)$$

を解けばよい。$\mu_0=96,\ \sigma=2.0$ を代入すれば、

$$n=\left(\frac{u(0.10)+u(0.05)}{\mu_0-94}\right)^2\cdot\sigma^2=\left(\frac{1.282+1.645}{96-94}\right)^2\times 2.0^2=8.567\cdots$$

よって、9回分析を行い、その分析値の平均が、

$\mu_0-u(0.05)\cdot\sigma/\sqrt{n}=96-1.645\cdot\dfrac{2.0}{\sqrt{9}}=94.90\cdots(\%)$ 以下のとき不合格とすればよい。

3.

（1）確率変数Yの分布関数を $G(y)$ とする。

$0\le y\le\theta$ のとき、$G(y)=P(Y\le y)=\left(\dfrac{y}{\theta}\right)^n$

よって、確率密度関数 $g(y)$ は $g(y)=G'(y)=\dfrac{n}{\theta^n}y^{n-1}$

また、$y>\theta$ のとき $G(y)=1$、$y<0$ のとき $G(y)=0$ であるから、$y<0,\ \theta<y$ では $g(y)=0$。

したがって、

$$\begin{aligned}
E\left\{(aY-\theta)^2\right\}&=\int_0^\theta (ay-\theta)^2\frac{n}{\theta^n}y^{n-1}dy\\
&=\int_0^\theta\left(a^2y^2-2ay\theta+\theta^2\right)\frac{n}{\theta^n}y^{n-1}dy\\
&=\int_0^\theta\left(\frac{a^2n}{\theta^n}y^{n+1}-\frac{2an}{\theta^{n-1}}y^n+\frac{n}{\theta^{n-2}}y^{n-1}\right)dy\\
&=\left[\frac{1}{n+2}\cdot\frac{a^2n}{\theta^n}y^{n+2}-\frac{1}{n+1}\cdot\frac{2an}{\theta^{n-1}}y^{n+1}+\frac{1}{n}\cdot\frac{n}{\theta^{n-2}}y^n\right]_0^\theta\\
&=\frac{a^2n}{\theta^n}\cdot\frac{\theta^{n+2}}{n+2}-\frac{2an\theta^2}{n+1}+\theta^2
\end{aligned}$$

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

（2）まず、a を動かしたとき、$E\left\{(aY-\theta)^2\right\}$の最小値を与える$a\left(=a^*\right)$を求める。

（1）より、

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

よって、$a^*=\dfrac{n+2}{n+1}$、そのときの最小値は$\left(\dfrac{\theta}{n+1}\right)^2$

したがって、$E\left\{(\delta_1(X)-\theta)^2\right\}=E\left\{\left(a^*Y-\theta\right)^2\right\}=\left(\dfrac{\theta}{n+1}\right)^2$

一方、$E(\delta_2(X))=E\left(\dfrac{2}{n}\sum_{i=1}^{n}X_i\right)=\dfrac{2}{n}\sum_{i=1}^{n}E(X_i)=\dfrac{2}{n}\sum_{i=1}^{n}\dfrac{\theta}{2}=\theta$

$\left(\because\ X_i\text{は}[0,\theta]\text{上の一様分布に従うから}E(X_i)=\dfrac{\theta}{2}\right)$

よって、

$$
\begin{aligned}
E\left\{(\delta_2(X)-\theta)^2\right\} &= E\left[\left\{\delta_2(X)-E(\delta_2(X))\right\}^2\right] \\
&= V(\delta_2(X)) \\
&= V\left(\frac{2}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{4}{n^2}\sum_{i=1}^{n} V(X_i) \qquad (\because X_1, X_2, \cdots, X_n \text{は互いに独立}) \\
&= \frac{4}{n^2}\sum_{i=1}^{n} \frac{\theta^2}{12} \qquad \left(\because X_i \text{は}[0,\theta]\text{上の一様分布に従うから} V(X_i)=\frac{\theta^2}{12}\right) \\
&= \frac{\theta^2}{3n}
\end{aligned}
$$

したがって、

$$
\begin{aligned}
E\left\{(\delta_2(X)-\theta)^2\right\} - E\left\{(\delta_1(X)-\theta)^2\right\} &= \frac{\theta^2}{3n} - \left(\frac{\theta}{n+1}\right)^2 = \frac{\theta^2}{3n(n+1)^2}\left(n^2-n+1\right) \\
&= \frac{\theta^2}{3n(n+1)^2}\left\{\left(n-\frac{1}{2}\right)^2 + \frac{3}{4}\right\} > 0
\end{aligned}
$$

$$
\therefore\ E\left\{(\delta_1(X)-\theta)^2\right\} < E\left\{(\delta_2(X)-\theta)^2\right\}
$$

したがって、$\delta_1(X)$の方が$\delta_2(X)$より良い推定量である。

4.

（1） $m=\mu_A-\mu_B$ より $\mu_A=m+\mu_B$

このとき、m, μ_B の尤度関数 $L(m, \mu_B)$ は次のとおりとなる。

$$
\begin{aligned}
L(m, \mu_B) &= \left(2\pi\sigma_A^2\right)^{-\frac{n_A}{2}} \cdot \left(2\pi\sigma_B^2\right)^{-\frac{n_B}{2}} \\
&\qquad \times \exp\left\{-\frac{1}{2\sigma_A^2}\sum_{i=1}^{n_A}(X_i-m-\mu_B)^2 - \frac{1}{2\sigma_B^2}\sum_{i=1}^{n_B}(Y_i-\mu_B)^2\right\}
\end{aligned}
$$

ここで、m の最尤推定量を求めるために、この尤度関数を最大にする m, μ_B を求める。

そこで、

$$\frac{\partial}{\partial m}\log L(m,\mu_B)=0 \cdots(a)$$

$$\frac{\partial}{\partial \mu_B}\log L(m,\mu_B)=0\cdots(b)$$

を解くことを考える。

(a)より、 $\frac{1}{\sigma_A^2}\sum_{i=1}^{n_A}(X_i-m-\mu_B)=0\cdots(a)'$

(b)より、 $\frac{1}{\sigma_A^2}\sum_{i=1}^{n_A}(X_i-m-\mu_B)+\frac{1}{\sigma_B^2}\sum_{i=1}^{n_B}(Y_i-\mu_B)=0\cdots(b)'$

$(b)'-(a)'$ より、 $\frac{1}{\sigma_B^2}\sum_{i=1}^{n_B}(Y_i-\mu_B)=0$

よって、 $\mu_B=\frac{1}{n_B}\sum_{i=1}^{n_B}Y_i=\overline{Y}$

これを $(a)'$ に代入して、求める m の最尤推定量は $m=\frac{1}{n_A}\sum_{i=1}^{n_A}X_i-\overline{Y}=\overline{X}-\overline{Y}$

（２） $\sigma_A^2=\sigma_B^2=\sigma^2$ として、尤度関数を作ると（簡単のため $s=\sigma^2$ とおく）、

$$L(s,\mu_A,\mu_B)=(2\pi s)^{-\frac{n_A+n_B}{2}}\exp\left\{-\frac{1}{2s}\left(\sum_{i=1}^{n_A}(X_i-\mu_A)^2+\sum_{i=1}^{n_B}(Y_i-\mu_B)^2\right)\right\}$$

が得られる。

ここで、 $s(=\sigma^2)$ の最尤推定量を求めるために、この尤度関数の対数を取り、s で偏微分すると、

$$\frac{\partial}{\partial s}\log L(s,\mu_A,\mu_B)=\frac{\partial}{\partial s}\left\{-\frac{n_A+n_B}{2}\log(2\pi s)-\frac{1}{2s}\left(\sum_{i=1}^{n_A}(X_i-\mu_A)^2+\sum_{i=1}^{n_B}(Y_i-\mu_B)^2\right)\right\}$$

$$=-\frac{n_A+n_B}{2s}+\frac{1}{2s^2}\left(\sum_{i=1}^{n_A}(X_i-\mu_A)^2+\sum_{i=1}^{n_B}(Y_i-\mu_B)^2\right)$$

を得る。

そこで、$\frac{\partial}{\partial s}\log L(s, \mu_A, \mu_B) = 0$ とおいて s を求めると、

$$s = \frac{1}{n_A + n_B}\left\{\sum_{i=1}^{n_A}(X_i - \mu_A)^2 + \sum_{i=1}^{n_B}(Y_i - \mu_B)^2\right\}$$

また、μ_A, μ_B の最尤推定量は、

$\frac{\partial}{\partial \mu_A}\log L(s, \mu_A, \mu_B) = 0$, $\frac{\partial}{\partial \mu_B}\log L(s, \mu\ \ , \mu_B) = 0$ を解くことにより

$\mu_A = \overline{X}$, $\mu_B = \overline{Y}$ となるので、求める最尤推定量 σ_L^2 は

$\sigma_L^2 = \frac{1}{n_A + n_B}\left\{\sum_{i=1}^{n_A}(X_i - \overline{X})^2 + \sum_{i=1}^{n_B}(Y_i - \overline{Y})^2\right\}$ により与えられる。

次に、$s_A^2 = \sum_{i=1}^{n_A}(X_i - \overline{X})^2$, $s_B^2 = \sum_{i=1}^{n_B}(Y_i - \overline{Y})^2$ とおくと、$\frac{s_A^2}{\sigma^2}$, $\frac{s_B^2}{\sigma^2}$ はそれぞれ自由度 $n_A - 1$, $n_B - 1$ の χ^2 分布に従うので、$E\left(\frac{s_A^2}{\sigma^2}\right) = n_A - 1$, $E\left(\frac{s_B^2}{\sigma^2}\right) = n_B - 1$ である。

よって、$E(s_A^2) = \sigma^2(n_A - 1)$, $E(s_B^2) = \sigma^2(n_B - 1)$ となる。

一方、$\sigma_L^2 = \frac{1}{n_A + n_B}(s_A^2 + s_B^2)$ だから、

$$E(\sigma_L^2) = \frac{1}{n_A + n_B}\left\{\sigma^2(n_A - 1) + \sigma^2(n_B - 1)\right\} = \frac{n_A + n_B - 2}{n_A + n_B}\sigma^2 \text{ である。}$$

したがって、$E(k \cdot \sigma_L^2) = \sigma^2$ が成り立つための k は $k = \frac{n_A + n_B}{n_A + n_B - 2}$ で与えられる。