昭和59年度（問　題）

1. ある農作物は，豊作であった年の翌年もまた豊作となる確率は75%，不作であった年の翌年がまた不作となる確率は15%であるという。

豊作と不作以外の場合は考えないものとして，十分遠い将来においてこの農作物が豊作である確率はどれくらいと考えられるか。

2. ある学校で，全校の生徒から300人を無差別に選出して学力テストを実施し，その結果得られた男女別の平均点の差により男女間の学力差を推定しようとした。

過去の経験から，学力テストの得点は正規分布に従い，また分散は女子が男子の2倍となっていることがわかっているとすれば，上記推定の信頼区間の幅が最小となるのは300人の内訳が男女それぞれ何人となるときか。

3. 母集団が2つの属性A，Bをもち，それぞれA_1，A_2およびB_1，B_2の2つの階級に分れている。この母集団からn個の標本をとり，次のような2×2分割表を得た。

A＼B	B_1	B_2	計
A_1	a	b	a + b
A_2	c	d	c + d
計	a + c	b + d	n

$(n = a + b + c + d)$

このとき，属性A，Bが独立であるという条件の下で，度数a，b，c，dに対応する最尤推定値を求めよ。

更に，この推定値を使って，χ^2値が次の式で与えられることを示せ。

$$\chi^2 = \frac{n\,(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

4. 正規母集団$N(\mu,\ \sigma^2)$（σ^2既知）の平均値μに対する検定で，帰無仮説を$\mu = \mu_0$，対立仮説を$\mu = \mu_1$としたとき，標本の大きさを十分大きくとれば第2種の過誤の確率をいくらでも小さくすることができることを示せ。

5. A種族から任意に選んだ6人の身長の平均は166.7 cm，分散7.52 (cm)^2

B種族から任意に選んだ10人の身長の平均は168.3 cm，分散7.37 (cm)^2

であった。

A種族とB種族とでは身長の差があるといえるか。また，B種族がA種族より身長が5 cm以上高いといえるか。

ただし，身長は正規分布に従っているものとする。

(注)　$F^5_9(0.05) = 3.48$，$F^9_5(0.05) = 4.77$，$F^1_{14}(0.05) = 4.60$

昭和59年度（解答例）

1.

$H(t)$: t 年が豊作である確率

$F(t)$: t 年が不作である確率

豊作と不作以外は考えなくてよいので

$$H(t)+F(t)=1 \qquad \cdots\cdots①$$

となる。

又，題意より

$$H(t+1)=0.75H(t)+0.85F(t) \qquad \cdots\cdots②$$

$$F(t+1)=0.25H(t)+0.15F(t) \qquad \cdots\cdots③$$

となる。

①より $F(t)=1-H(t)$ となるので，これを②に代入すると

$$H(t+1)=0.85-0.1H(t)$$

が得られる。

これより

$$H(t+1)-\frac{0.85}{1.1}=-0.1\left(H(t)-\frac{0.85}{1.1}\right)$$ と変形できる。

よって

$$H(t)=(-0.1)^t\left(H(0)-\frac{0.85}{1.1}\right)+\frac{0.85}{1.1}$$

$t\to\infty$ とすれば

$$H(t)\to\frac{0.85}{1.1}=\frac{17}{22}$$

したがって十分遠い将来において豊作である確率は $\frac{17}{22}$ と考えられる。

2.

男女の得点分布は正規分布で，分散は女子が男子の2倍となっているので，男女の得点分布を各々 $N(\mu_1, \sigma)$, $N(\mu_2, 2\sigma)$ とおける。

さて，男子 n_1 人，女子 n_2 人をとり，各々の平均値を $\overline{m}, \overline{f}$ とすれば，$\overline{m}, \overline{f}$ の分布

は各々

$N\left(\mu_1, \dfrac{\sigma}{n^1}\right)$, $N\left(\mu_2, \dfrac{2\sigma}{n_2}\right)$ となる。

したがって，$\overline{m}-\overline{f}$ の分布は

$N\left(\mu_1-\mu_2, \dfrac{\sigma}{n_1}+\dfrac{2\sigma}{n_2}\right)$ となる。

この場合信頼度αの信頼区間は

$$P\left\{\frac{|\overline{m}-\overline{f}-(\mu_1-\mu_2)|}{\sqrt{\dfrac{\sigma}{n_1}+\dfrac{2\sigma}{n_2}}} < \varepsilon(\alpha)\right\}=\alpha$$

となる$\varepsilon(\alpha)$ を求めて $2\varepsilon(\alpha)\sqrt{\dfrac{\sigma}{n_1}+\dfrac{2\sigma}{n_2}}$ となる

よって$n_1+n_2=300$ のもとで $\dfrac{\sigma}{n_1}+\dfrac{2\sigma}{n_2}$ が最小となる n_1, n_2 を求めればよい。

したがって

$f(n_1)=\dfrac{1}{n_1}+\dfrac{2}{300-n_1}$ なる n_1 の関数を考えて $f(n_1)$ が最小となるように

すればよい。

$$f'(n_1)=\frac{-1}{n_1^2}+\frac{2}{(300-n_1)^2}=\frac{n_1^2+600n_1-90000}{n_1^2(300-n_1)^2}=0$$

となる n_1は $n_1=-300\pm\sqrt{300^2+90000}=-300\pm\sqrt{180000}$

$=-300\pm300\sqrt{2}$

$n_1>0$ より $n_1=300(\sqrt{2}-1)=124.264\cdots\cdots$

さて，$f(124)=0.01942815\cdots\cdots$

$f(125)=0.01942857\cdots\cdots$

$f(124)<f(125)$ であるので

求めるものは $n_1=124, n_2=176$.

よって，男子124人，女子176人のとき信頼区間は最小となる。

3.

AとBは独立だから

$$P(Ai\cap Bj)=P(Ai)P(Bj)\quad (i,j=1,2)$$

又，$P(A_1)=p$，$P(B_1)=q$ とおくと

$P(A_2)=1-p$，$P(B_2)=1-q$ となる。

よって尤度関数は

$$L(p,q)=\{P(A_1\cap B_1)\}^a\cdot\{P(A_1\cap B_2)\}^b\cdot\{P(A_2\cap B_1)\}^c\cdot\{P(A_2\cap B_2)\}^d$$

$$=(pq)^a\{p(1-q)\}^b\cdot\{(1-p)q\}^c\cdot\{(1-p)(1-q)\}^d$$

$$=p^{a+b}\cdot(1-p)^{c+d}\cdot q^{a+c}\cdot(1-q)^{b+d}$$

$L(p,q)$ を最大にするためには

$$p^{a+b}(1-p)^{c+d}=p^{a+b}(1-p)^{n-a-b}$$

$$q^{a+c}(1-q)^{b+d}=q^{a+c}(1-q)^{n-a-c}$$

の各々を最大にすればよい。

$$\begin{cases}L_1(p)=p^{a+b}(1-p)^{n-a-b}\\ L_2(q)=q^{a+c}(1-q)^{n-a-c}\end{cases}\quad \text{とおく。}$$

$$\log L_1(p)=(a+b)\log p+(n-a-b)\log(1-p)$$

$$\frac{d\log L_1(p)}{dp}=\frac{a+b}{\hat{p}}-\frac{n-a-b}{1-\hat{p}}=\frac{a+b-n\hat{p}}{\hat{p}(1-\hat{p})}=0$$

$$\therefore\ n\hat{p}=a+b\qquad \therefore\ \hat{p}=\frac{a+b}{n}$$

同様に $$\hat{q}=\frac{a+c}{n}$$

よって
$$1-\hat{p}=\frac{c+d}{n}$$
$$1-\hat{q}=\frac{b+d}{n} \qquad \text{となる。}$$

ところで度数a, b, c, d に対応する期待値を $\alpha, \beta, \gamma, \delta$ とすれば，

$$\alpha=npq,\ \ \beta=np(1-q),\ \ \gamma=n(1-p)q,\ \ \delta=n(1-p)(1-q)$$

であるから a, b, c, d に対応する最尤推定値 $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}$ は

$$\hat{\alpha}=n\hat{p}\hat{q}=\frac{(a+b)(a+c)}{n},\ \ \hat{\beta}=n\hat{p}(1-\hat{q})=\frac{(a+b)(b+d)}{n}$$

$$\hat{\gamma}=n(1-\hat{p})\hat{q}=\frac{(c+d)(a+c)}{n},\ \ \hat{\delta}=n(1-\hat{p})(1-\hat{q})=\frac{(c+d)(b+d)}{n}$$

となる。

次に

$$\chi^2=\frac{(a-\hat{\alpha})^2}{\hat{\alpha}}+\frac{(b-\hat{\beta})^2}{\hat{\beta}}+\frac{(c-\hat{\gamma})^2}{\hat{\gamma}}+\frac{(d-\hat{\delta})^2}{\hat{\delta}} \quad \text{であり，}$$

$$a-\hat{\alpha}=\frac{a(a+b+c+d)-(a+b)(a+c)}{n}=\frac{ad-bc}{n}$$

$$b-\hat{\beta}=-\frac{ad-bc}{n},\ \ c-\hat{\gamma}=-\frac{ad-bc}{n},\ \ d-\hat{\delta}=\frac{ad-bc}{n} \quad \text{となるから，}$$

$$\therefore\ \chi^2=\frac{(ad-bc)^2}{n}\cdot\frac{(c+d)(b+d)+(c+d)(a+c)+(a+b)(b+d)+(a+b)(a+c)}{(a+b)(c+d)(a+c)(b+d)}$$

$$=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

4.

$\mu_1>\mu_0$ なるとき採用されるべき棄却領域Rは $\bar{x}>\lambda_n(\alpha)$ となる。ただし，$\lambda_n(\alpha)$は第 1 種の過誤の確率αと標本の大きさnとにより定められる。

すなわち，

$$P(\overline{X} > \lambda_n(\alpha) \mid \mu = \mu_0) = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}(\lambda_n(\alpha)-\mu_0)/\sigma}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt = \alpha$$

これより $\sqrt{n}(\lambda_n(\alpha)-\mu_0)/\sigma = k$（$n$ に無関係）とおくことができる。

$$P(\overline{X} \geqq \lambda_n(\alpha) | \mu = \mu_1) = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}(\lambda_n(\alpha)-\mu_1)/\sigma}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{\sqrt{n}(\lambda_n(\alpha)-\mu_0)+\sqrt{n}(\mu_0-\mu_1)}{\sigma}}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{k+\frac{\sqrt{n}(\mu_0-\mu_1)}{\sigma}}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

$\mu_0-\mu_1 < 0$ なる故に α を一定にしておけば，

$$P(\overline{X} \geqq \lambda_n(\alpha) \mid \mu = \mu_1) \longrightarrow 1 \ (n \to \infty)$$

第 2 種の過誤の確率は，

$$1-P(\overline{X} \geqq \lambda_n(\alpha) \mid \mu = \mu_1)$$

なる故に n を十分大きくとればいくらでも小さくすることができる。

又 $\mu_1 < \mu_0$ の場合も同様である。

5.

$\overline{x}_1 = 166.7$，$n_1 = 6$，$s_1^2 = 7.52$ とおくと，不偏分散は

$$\mu_1^2 = \frac{6}{5} s_1^2 = 9.024$$

$\overline{x}_2 = 168.3$，$n_2 = 10$，$s_2^2 = 7.37$ とおくと，不偏分散は

$$\mu_2^2 = \frac{10}{9} s_2^2 = 8.189$$

2 つの母集団の分散を σ_1^2，σ_2^2 とし，帰無仮説を $\sigma_1^2/\sigma_2^2 = 1$ とすると $y = \dfrac{\mu_1^2}{\mu_2^2}$

は，自由度（5，9）の F 一分布に従う。

$$y = \frac{\mu_1^{\,2}}{\mu_2^{\,2}} = \frac{9.024}{8.189} = 1.1 < 3.48$$

よって，この帰無仮説は棄却できない。

そこで $\sigma_1^2 = \sigma_2^2$ と仮定する。

分散の等しい，正規母集団 $N(m_i,\ \sigma^2)(i=1,2)$ からそれぞれ大きさ $n_i(i=1,2)$ の標本を選び，

$$y = \frac{[(\bar{x}_1-\bar{x}_2)-(m_1-m_2)]^2}{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \cdot \frac{n_1+n_2-2}{n_1 s_1^{\,2}+n_2 s_2^{\,2}}$$

とおくと，これは自由度が $(1,\ n_1+n_2-2)$ の F 分布に従う。

まず，$m_1-m_2=0$ という仮説をたてる。

すると，

$$y = \frac{(166.7-168.3)^2}{\left(\frac{1}{6}+\frac{1}{10}\right)} \cdot \frac{6+10-2}{6\times7.52+10\times7.37} = \frac{1.6^2\times60\times14}{16\times118.82} = 1.13$$

< 4.60 ……自由度（1，14）の F 分布の5％点。

この仮説は棄却できない。つまり差がないということを否定できない。

次に $m_2-m_1=5$ という仮説をたてる。

$$y = \frac{(166.7-168.3+5)^2}{\frac{1}{6}+\frac{1}{10}} \times \frac{10+6-2}{6\times7.52+10\times7.37} = \frac{3.4^2\times60\times14}{16\times118.82} = 5.11$$

> 4.60 ……自由度（1，14）の F 分布5％点。

95％の信頼率で B 種族の方が A 種族より5 cm以上高いとは言えない。