MATH 3423 Final (Solution)

1. (1). $E(X_1 - X_2) = 0$ but $P(X_1 = X_2) < 1$

(2). exp family $\sum X_i$ C-S. ~~$\sum X_i \sim Bin(n, p)$~~

(3). $\log L(\theta) = \sum X_i \log p + (n - \sum X_i) \log(1-p)$

$$\frac{\partial \log L(\theta)}{\partial p} = \frac{\sum X_i}{p} + \frac{\sum X_i - n}{1-p} = 0 \qquad \hat{p} = \frac{\sum X_i}{n}$$

~~$\frac{\bar{X}}{1-\bar{X}} = \frac{\sum X_i}{n - \sum X_i}$ is MLE of $\frac{p}{1-p}$~~ $\frac{\sum X_i}{n}\left(1 - \frac{\sum X_i}{n}\right)$ is MLE of $p(1-p)$

$$\sum_{y=0}^{n} h(y) \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} = p(1-p)$$

$$\sum_{y=0}^{n} h(y) \frac{n(n-1)}{y(n-y)} \frac{(n-2)!}{(y-1)!(n-y-1)!} p^{y-1} (1-p)^{n-y-1} = 1$$

$$h(y) = \begin{cases} \frac{y(n-y)}{n(n-1)} & y = 1, 2, \dots n-1 \\ 0 & y = 0, n \end{cases} = \frac{y(n-y)}{n(n-1)}$$

unbaised estimator is $\frac{\sum X_i}{n}\left(\frac{n - \sum X_i}{n-1}\right)$

so. $\frac{\sum X_i}{n}\left(\frac{n - \sum X_i}{n}\right)$ is not unbaised

$$E\left(\frac{y}{n} \frac{n-y}{n}\right) = \frac{1}{n^2} \sum_{y=0}^{n} y(n-y) \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} = \frac{n-1}{n} p(1-p)$$

$\Rightarrow$ is not unbiased

or use Jenson Inequality $E(Y(1-Y)) < EY(1-EY)$

$Y \sim Bin(n, p)$

(4). $\sum_{y=0}^{n} h(y) \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y-2} = 1$ $\qquad y = \sum_{i=1}^{n} X_i$

$$h(y) = \begin{cases} \frac{(n-y)(n-y-1)}{n(n-1)} & y = 0, 1, 2 \dots n-2 \\ 0 & y = n-1, n. \end{cases} = \frac{(n-y)(n-y-1)}{n(n-1)}$$

UMVUE is $\frac{(n - \sum X_i)(n - \sum X_i - 1)}{n(n-1)}$

$$CRLB = \frac{4(1-p)^2}{\frac{n}{p(1-p)}} = \frac{4p(1-p)^3}{n} \qquad \frac{\partial \log L(p)}{\partial p} = \frac{\sum X_i - np}{p(1-p)} \qquad E\left(\frac{\partial \log L(\theta)}{\partial p}\right)^2 = \frac{np(1-p)}{p^2(1-p)^2} = \frac{n}{p(1-p)}$$

CRLB = var iff $\frac{\partial \log L}{\partial \theta} = A(\theta, n)[h(x) - \tau(\theta)]$ can't achieve

$$= \frac{s}{p} - \frac{n-s}{1-p} = \frac{1}{p(1-p)}(s - np) = \frac{n}{p(1-p)}\left(\frac{s}{n} - p\right)$$

(5). $\sum_{y=0}^{n} h(y) \binom{n}{y} p^{y} (1-p)^{n-y} = p^{m}$.

$$\sum h(y) \cdot \frac{n!\,(y-m)!\,(n-y)!}{y!\,(n-y)!\,(n-m)!} \frac{(n-m)!}{(y-m)!\,(n-y)!} p^{y-m} (1-p)^{n-y} = 1$$

$$h(x) = \begin{cases} \dfrac{y!\,(n-y)!\,(n-m)!}{n!\,(y-m)!\,(n-y)!} & y = m, m+1, \dots n \\ 0 & y = 0, 1, \dots m-1 \end{cases}$$

$$\text{or } h(x) = \begin{cases} \dfrac{\binom{n-m}{y-m}}{\binom{n}{y}} & y = m, m+1, \dots n \\ 0 & y = 0, 1, \dots m-1 \end{cases} = \begin{cases} \dfrac{y(y-1)\cdots(y-m+1)}{n(n-1)\cdots(n-m+1)} & y = m, m+1, \dots, n \\ 0 & y = 0, 1, \dots m-1 \end{cases}$$

$Y = \sum X_i$

2. (1) δ is known. $f(x) = \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\delta)}$ $1\{x \geq \delta\}$. exp family

$T = \sum X_i$ is C.S. $X_i - \delta \sim exp.(\frac{1}{\theta})$ $E(X_i - \delta) = \theta$

$\frac{T}{n} - \delta$ is UMVUE of θ $T \sim G(n, \frac{1}{\theta}) + n\delta$

$$P_r(X_i > 1) = e^{\frac{1}{\theta}(\delta - 1)} \quad (\delta < 1)$$
$$= e^{-\frac{1}{\theta}(1-\delta)}$$

$\sum X_i - n\delta = Y \sim G(n, \frac{1}{\theta})$

$$\int_0^\infty h(y) \frac{(\frac{1}{\theta})^n}{\Gamma(n)} y^{n-1} e^{-\frac{1}{\theta} y} dy = e^{-\frac{1}{\theta}(1-\delta)}$$

$$h(y) = \begin{cases} \left(\dfrac{y-1+\delta}{y}\right)^{n-1} & y \geq 1-\delta \\ 0 & 0 < y < 1-\delta \end{cases}$$

$y = \sum X_i - n\delta$

$h(\sum X_i - n\delta)$ is UMVUE of $P(X_i > 1)$

(2) θ is known $f = \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\delta)} 1\{x \geq \delta\}$

joint pdf $f_X = \frac{1}{\theta^n} e^{-\frac{1}{\theta}(\Sigma X_i - n\delta)} 1\{X_{(1)} \geq \delta\}$

$X_{(1)}$ is ~~suff~~

pdf of $X_{(1)}$ $f_{X_{(1)}} = n(1-F)^{n-1} f = n\frac{1}{\theta} e^{-\frac{n}{\theta}(x-\delta)}$

$1-F = Pr(X_i > x)$
$= Pr(X_i - \delta > x - \delta)$
$= \exp\left(-\frac{1}{\theta}(x-\delta)\right)$

$E g(X_{(1)}) = 0 \Rightarrow \int_\delta^\infty g(x) e^{-n\frac{1}{\theta}x} dx = 0$

differentiate w.r.t. δ

$g(\delta) e^{-\frac{n}{\theta}\delta} = 0$ $g(x) = 0$ a.s. $X_{(1)}$ is complete

$X_{(1)} - \delta \sim \exp\{\frac{n}{\theta}\}$ $E(X_{(1)} - \delta) = \frac{\theta}{n}$

$X_{(1)} - \frac{\theta}{n}$ is UMVUE of δ

$P(X_i > 1) = e^{-\frac{1}{\theta}(1-\delta)}$ when $\delta < 1$

$\int_\delta^\infty h(x) \frac{n}{\theta} e^{-\frac{n}{\theta}(x-\delta)} dx = e^{-\frac{1}{\theta}(1-\delta)}$ for any δ

$e^{\frac{1}{\theta}} \cdot \frac{n}{\theta} \int_\delta^\infty h(x) e^{-\frac{n}{\theta}x} dx = e^{-\frac{(n-1)}{\theta}\delta}$

differentiate w.r.t δ

$\int_\delta^\infty h(x) \frac{n}{\theta} e^{-\frac{n}{\theta}(x-\delta)} e^{\frac{1}{\theta}(1-\delta)} dx = 1$
$\Rightarrow \int_\delta^\infty h(x) \frac{n}{\theta} e^{-\frac{(n-1)}{\theta}(x-\delta)} e^{-\frac{1}{\theta}(x-1)} dx = 1$
$\Rightarrow \int_\delta^\infty h(x) \frac{n}{n-1} e^{-\frac{x-1}{\theta}} \frac{n-1}{\theta} e^{-\frac{n-1}{\theta}(x-\delta)} dx = 1$
$\Rightarrow h(x) = \frac{n-1}{n} e^{-\frac{1}{\theta}(1-x)}$

$e^{\frac{1}{\theta}} \frac{n}{\theta} h(\delta) e^{-\frac{n}{\theta}\delta} = e^{-\frac{n-1}{\theta}\delta} \cdot \frac{(n-1)}{\theta}$

$h(x) = \frac{n-1}{n} e^{-\frac{1}{\theta}(1-x)}$ ~~$x = X_{(1)}$~~

UMVUE of $P(X_i > 1)$ when $\delta < 1$ is $h(X_{(1)})$

3\. (1). joint pdf $f_X = \theta^n \exp\{\theta \cdot (-\Sigma X_i)\}$

reject H_0 when $-\Sigma X_i \le k$ or $\Sigma X_i \ge k$

$\Sigma X_i \sim G(n, \theta)$ $\quad 2\theta \Sigma X_i \sim G(n, \frac{1}{2}) = \chi^2(2n)$

$$k = \frac{\chi^2(2n,\alpha)}{2\theta_0} = \frac{\chi^2(20, 0.05)}{2} = 31.41/2 = 15.705$$

(2) $P_{\theta=0.3}(\Sigma X_i \ge k) = P_{\theta=0.3}(\Sigma X_i \ge 15.705) = P(0.6\Sigma X_i \ge 9.423) \approx 0.9769$

$2\theta \Sigma X_i \sim \cancel{G(n,} \chi^2(2n) = \chi^2(20)$

(3). $\theta = \mu = \lambda$.

$$\log L = \log \lambda^{2n} e^{-\lambda(\Sigma X_i + \Sigma Y_i)} = 2n \log \lambda - \lambda(\Sigma X_i + \Sigma Y_i)$$

$$\hat{\lambda} = \frac{2n}{\Sigma X_i + \Sigma Y_i}$$

$\cancel{\Sigma}$ $\theta \neq \mu$. $\quad L = \theta^n e^{-\theta \Sigma X_i} \mu^n e^{-\mu \Sigma Y_i}$

$$\frac{\partial \log L}{\partial \theta} = n\frac{1}{\theta} - \Sigma X_i = 0 \quad \hat{\theta} = \frac{n}{\Sigma X_i} \quad \hat{\mu} = \frac{n}{\Sigma Y_i}$$

$$\lambda(x,y) = \frac{2^{2n} (\Sigma X_i)^n (\Sigma Y_i)^n}{(\Sigma X_i + \Sigma Y_i)^{2n}}$$

(4). $C_1 = \{(x,y) \mid \lambda(x,y) \le k\}$

$\lambda(x,y) \le k \iff T \le k_1$ or $T \ge k_2$ $\quad T = \dfrac{\Sigma X_i}{\Sigma X_i + \Sigma Y_i}$

when $\theta = \mu = \lambda$. $\quad 2\lambda \Sigma X_i \sim \chi^2(2n) \quad 2\lambda \Sigma Y_i \sim \chi^2(2n)$

$$\Rightarrow \Sigma y_i / \Sigma x_i \sim F(2n, 2n)$$

$$T \le k_1 \Rightarrow \frac{\Sigma y_i}{\Sigma x_i} \ge F(2n, 2n; \alpha/2) = F(20, 20, 0.05) = 2.12$$

$$T \ge k_2 \Rightarrow \frac{\Sigma y_i}{\Sigma x_i} \le F(2n, 2n, 1-\alpha/2) = \frac{1}{F(2n, 2n, \alpha/2)} = 0.47$$

(5) $C_1 = \{-2\log \lambda(x) \ge k'\}$ $\qquad k' = \chi^2(1, \alpha)$

$\Rightarrow 2\{2n \log(\Sigma x_i + \Sigma y_i) - 2n \log 2 - n \log(\Sigma x_i) - n \log(\Sigma y_i)\} \ge \chi^2_{1,\alpha}$

when $\alpha = 0.05$, $\Sigma x_i = 100$, $\Sigma y_i = 50$, $n = 50$, $\chi^2_{obs} = 11.78 > \chi^2_{1,0.05} = 3.84$

4. (1) $p = \frac{820}{900} = 0.9111$ $\theta_{03} = p^3 = 0.7563$ $\theta_{02} = 3p^2(1-p) = 0.2214$ $\theta_{01} = 3p(1-p)^2 = 0.02160$

$\theta_{00} = 0.0007$ by Pearson goodness-of-fit test

Expected values : $e_{x=0}$ $e_{x=1}$ $e_{x=2}$ $e_{x=3}$

0.2107 6.4790 66.4099 226.9004

expected freq. ≤ 5, should be grouped

$$Q = \frac{(7-6.6897)^2}{6.6897} + \frac{(65-66.4099)^2}{66.4099} + \frac{(228-226.9004)^2}{226.9004} = 0.04965$$

$\chi^2(3-1-1, 0.05) = 3.84 < Q$ $\Rightarrow$ Can't reject H_0

(2) (i) $$\frac{\bar{X}-\bar{Y}}{\sqrt{\bar{Z}(1-\bar{Z})(\frac{1}{n}+\frac{1}{n})}} = \frac{0.9-0.8}{\sqrt{0.85 \times 0.15 \times (\frac{1}{900}+\frac{1}{900})}} = 5.94$$

H_0 $p_1 = p_2$

H_1 $p_1 \neq p_2$

$Z \sim N(0,1)$ $z_{0.025} = 1.96$

reject H_0

(ii) when $p_1 = p_2 = p$

$$L = p^{\sum X_i + \sum Y_i}(1-p)^{2n-\sum X_i - \sum Y_i}$$ $$\frac{\partial \log L}{\partial p} = \frac{\sum X_i + \sum Y_i}{p} - \frac{2n - \sum X_i - \sum Y_i}{1-p} = 0$$

$$\hat{p} = \frac{\sum X_i + \sum Y_i}{2n}$$

when $p_1 \neq p_2$ $$\frac{\partial \log L}{\partial p_1} = 0 = \frac{\sum X_i}{p_1} - \frac{n - \sum X_i}{1-p_1}$$ $\hat{p}_1 = \frac{\sum X_i}{n}$

$\hat{p}_2 = \frac{\sum Y_i}{n}$

$$\lambda(X,Y) = \frac{\left(\frac{\bar{X}+\bar{Y}}{2}\right)^{\sum X_i + \sum Y_i}\left(1-\frac{\bar{X}+\bar{Y}}{2}\right)^{2n-\sum X_i - \sum Y_i}}{(\bar{X})^{\sum X_i}(1-\bar{X})^{n-\sum X_i}(\bar{Y})^{\sum Y_i}(1-\bar{Y})^{n-\sum Y_i}}$$

$-2\log\lambda = 35.88$ $\chi^2(1, 0.05) = 3.841$

reject H_0

$$\text{Or} = 2\left\{810 \log\frac{810*1800}{1530*900} + 720*\log\frac{720*1800}{1530*900} + 90\log\frac{90*1800}{270*900} + 180\log\frac{180*1800}{270*900}\right\} = 35.88$$

(iii)

	chocolate	cheese	
sold	810 (820)	720	1530 (1540)
not sold	90 (80)	180	270 (260)
	900	900	1800

$$G = \frac{\left(810 - \frac{1530\times 900}{1800}\right)^2}{\frac{1530\times 900}{1800}} + \frac{\left(720 - \frac{1530\times 900}{1800}\right)^2}{\frac{1530\times 900}{1800}} + \frac{\left(90 - \frac{270\times 900}{1800}\right)^2}{\frac{270\times 900}{1800}} + \frac{\left(180 - \frac{270\times 900}{1800}\right)^2}{\frac{270\times 900}{1800}}$$

$$= 2.647 + 2.647 + 15 + 15 = 35.294 > \chi^2(1, .05) = 3.84$$

reject H_0.

recommendation is $P_1 \neq P_2$. chocolate is more popular

$$\text{OR}: G = 1800\left(\frac{810^2}{1530 * 900} + \frac{720^2}{1530 * 900} + \frac{90^2}{270 * 900} + \frac{180^2}{270 * 900} - 1\right)$$

$$= 35.2941$$

(iii) $$G = 1800\left(\frac{820^2}{1540\times 900} + \frac{720^2}{1540\times 900} + \frac{80^2}{260\times 900} + \frac{180^2}{260\times 900} - 1\right)$$

$$= 44.9550$$

(ii) $$2\left\{820 \log\frac{820\times 1800}{1540\times 900} + 720 * \log\frac{720\times 1800}{1540\times 900} + 80 * \log\frac{80 * 1800}{260 * 900} + 180 \log\frac{180 * 1800}{260 * 900}\right\} = 45.9689$$

(i) $$\frac{\frac{820}{900} - \frac{720}{900}}{\sqrt{\frac{1540}{1800}\left(1 - \frac{1540}{1800}\right)\left(\frac{1}{900} + \frac{1}{900}\right)}} = 6.7049$$