Final (2000/2001)

1. $X \sim Bin(n, \theta)$

Set $g(\theta) = P(X \le 2) = \sum_{x=0}^{2} \binom{n}{x} \theta^x (1-\theta)^{n-x}$

$= (1-\theta)^n + n\theta(1-\theta)^{n-1} + \binom{n}{2}\theta^2(1-\theta)^{n-2}$

$X_1, \dots, X_r$ is a random sample of size r from $Bin(n,\theta)$.

$f_X(x;\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$

$= \exp\{\log\binom{n}{x} + x\log\theta + (n-x)\log(1-\theta)\}$

$= \exp\{n\log(1-\theta) + \log\binom{n}{x} + [\log(\frac{\theta}{1-\theta})]x\}$

$\therefore f_X(x;\theta)$ belongs to the exponential family of pdf and $\sum_{i=1}^{r} X_i$ is a complete minimal sufficient statistic for $\theta$.

Let $S = \sum_{i=1}^{r} X_i$.

Note that $S \sim Bin(nr, \theta)$.

The UMVUE of $g(\theta)$ is a function of $S$, say $h(S)$,

$\therefore E(h(S)) = g(\theta) = (1-\theta)^n + n\theta(1-\theta)^{n-1} + \binom{n}{2}\theta^2(1-\theta)^{n-2}$.

Since the expectation function is linear under addition, if we could find $h_0(S)$, $h_1(S)$ and $h_2(S)$ such that $E(h_0(S)) = (1-\theta)^n$, $E(h_1(S)) = n\theta(1-\theta)^{n-1}$ and $E(h_2(S)) = \binom{n}{2}\theta^2(1-\theta)^{n-2}$, then $h(S) = h_0(S) + h_1(S) + h_2(S)$.

$E(h_0(S)) = (1-\theta)^n \Rightarrow \sum_{s=0}^{nr} h_0(s)\binom{nr}{s}\theta^s(1-\theta)^{nr-s} = (1-\theta)^n$

$\Rightarrow \sum_{s=0}^{nr} h_0(s)\binom{nr}{s}\theta^s(1-\theta)^{nr-n-s} = 1$

$\Rightarrow \sum_{s=0}^{nr} h_0(s)\binom{nr}{s}\binom{nr-n}{s}^{-1}\binom{nr-n}{s}\theta^s(1-\theta)^{nr-n-s} = 1$

$\Rightarrow h_0(s)\binom{nr}{s}\binom{nr-n}{s}^{-1} = I_{\{0,\dots,nr-n\}}(s)$

$\Rightarrow h_0(s) = \binom{nr}{s}^{-1}\binom{nr-n}{s} I_{\{0,\dots,nr-n\}}(s)$

$E(h_1(S)) = n\theta(1-\theta)^{n-1} \Rightarrow \sum_{s=0}^{nr} h_1(s)\binom{nr}{s}\theta^s(1-\theta)^{nr-s} = n\theta(1-\theta)^{n-1}$

$\Rightarrow \sum_{s=0}^{nr} h_1(s)\binom{nr}{s}(\frac{1}{n})\theta^{s-1}(1-\theta)^{nr-n-s+1} = 1$

$\Rightarrow \sum_{s=0}^{nr} h_1(s)\binom{nr}{s}(\frac{1}{n})\binom{nr-n}{s-1}^{-1} = I_{\{1,\dots,nr-n\}}(s)$

$\Rightarrow h_1(s) = \binom{nr}{s}^{-1} n \binom{nr-n}{s-1} I_{\{1,\dots,nr-n\}}(s)$

, $E(h_2(S)) = \binom{n}{2}\theta^2(1-\theta)^{n-2} \Rightarrow h_2(s) = \binom{nr}{s}^{-1}\binom{n}{2}\binom{nr-n}{s-2} I_{\{2,\dots,nr-n\}}(s)$.

$\therefore h(S) = \binom{nr}{S}^{-1}\left[\binom{nr-n}{S} I_{\{0,\dots,nr-n\}}(S) + n\binom{nr-n}{S-1} I_{\{1,\dots,nr-n\}}(S) + \binom{n}{2}\binom{nr-n}{S-2} I_{\{2,\dots,nr-n\}}(S)\right]$

is the UMVUE of $g(\theta)$ where $S = \sum_{i=1}^{r} X_i$.

2. $X_1, \ldots, X_n$ is a r.s. of size $n$ from the Negative Binomial distribution with $k$ known and parameter $\theta \in \Omega = (0,1)$.

(a) $f_X(x;\theta) = \binom{k+x-1}{x}\theta^k(1-\theta)^x = \exp\{\log\binom{k+x-1}{x} + k\log\theta + x\log(1-\theta)\}$

$= \exp\{k\log\theta + \log\binom{k+x-1}{x} + x\log(1-\theta)\}$

$\therefore f_X(x;\theta)$ belongs to the exponential family of pdf and $\sum_{i=1}^n X_i$ is a complete minimal sufficient statistic for $\theta$.

Let $S = \sum_{i=1}^n X_i$

Note that $S \sim \text{Neg. Bin}(nk, \theta)$.

The UMVUE of $g(\theta) = \frac{1}{\theta}$ is a function of $S$, say $h(S)$,

$\therefore E(h(S)) = g(\theta) = \frac{1}{\theta} \Rightarrow \sum_{s=0}^{\infty} h(s)\binom{nk+s-1}{s}\theta^{nk}(1-\theta)^s = \frac{1}{\theta}$

$\Rightarrow \sum_{s=0}^{\infty} h(s)\binom{nk+s-1}{s}\theta^{nk+1}(1-\theta)^s = 1$

$\Rightarrow \sum_{s=0}^{\infty} h(s)\binom{nk+s-1}{s}\binom{nk+1+s-1}{s}^{-1}\binom{nk+1+s-1}{s}\theta^{nk+1}(1-\theta)^s = 1$

$\Rightarrow h(s)\binom{nk+s-1}{s}\binom{nk+1+s-1}{s}^{-1} = 1 \Rightarrow h(s) = \binom{nk+s}{s}\binom{nk+s-1}{s}^{-1}$

$\Rightarrow h(s) = \frac{nk+s}{nk} = 1 + \frac{s}{nk}$

$\therefore$ the UMVUE of $g(\theta) = \frac{1}{\theta}$ is $1 + \frac{1}{nk}\sum_{i=1}^n X_i$.

$Var(1 + \frac{1}{nk}\sum_{i=1}^n X_i) = (\frac{1}{nk})^2 \sum_{i=1}^n Var(X_i) = (\frac{1}{nk})^2 \sum_{i=1}^n \left[\frac{k(1-\theta)}{\theta^2}\right] = \frac{1-\theta}{nk\theta^2}$

(b) $\log f_X(x;\theta) = k\log\theta + \log\binom{k+x-1}{x} + x\log(1-\theta)$

$\frac{\partial}{\partial\theta}\log f_X(x;\theta) = \frac{k}{\theta} - \frac{x}{1-\theta}$

$\frac{\partial^2}{\partial\theta^2}\log f_X(x;\theta) = -\frac{k}{\theta^2} - \frac{x}{(1-\theta)^2}$

$E[\frac{\partial^2}{\partial\theta^2}\log f_X(X;\theta)] = E[-\frac{k}{\theta^2} - \frac{X}{(1-\theta)^2}] = -\frac{k}{\theta^2} - \frac{1}{(1-\theta)^2}\frac{k(1-\theta)}{\theta} = \frac{-k}{\theta^2(1-\theta)}$

$\therefore$ the Cramer-Rao lower bound

$= \frac{[g'(\theta)]^2}{-nE[\frac{\partial^2}{\partial\theta^2}\log f_X(X;\theta)]} = \frac{(\frac{-1}{\theta^2})^2}{(-n)(\frac{-k}{\theta^2(1-\theta)})} = \frac{1-\theta}{nk\theta^2} = Var(1 + \frac{1}{nk}\sum_{i=1}^n X_i)$.

$\therefore$ the Cramer-Rao lower bound is attained.

(c) $L(\theta) = f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = \prod_{i=1}^n f_{X_i}(x_i;\theta) = \prod_{i=1}^n \binom{k+x_i-1}{x_i}\theta^k(1-\theta)^{x_i}$

$\log L(\theta) = \sum_{i=1}^n \log\binom{k+x_i-1}{x_i} + nk\log\theta + (\sum_{i=1}^n x_i)\log(1-\theta)$

$\frac{\partial}{\partial\theta}\log L(\theta) = \frac{nk}{\theta} - \frac{\sum_{i=1}^n x_i}{1-\theta}$

$\frac{\partial}{\partial\theta}\log L(\theta)|_{\theta=\hat{\theta}} = 0 \Rightarrow \frac{nk}{\hat{\theta}} - \frac{\sum_{i=1}^n X_i}{1-\hat{\theta}} = 0$

$\Rightarrow nk(1-\hat{\theta}) = \hat{\theta}\sum_{i=1}^n X_i \Rightarrow \hat{\theta} = \frac{nk}{nk + \sum_{i=1}^n X_i}$ (M.L.E. of $\theta$)

3. $X_1, \ldots, X_n$ is a r.s. of size $n$ from $N(\mu, \theta)$ with $\mu$ known.

(a) $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta > \theta_0 \end{cases}$

$f_X(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\{\frac{-1}{2\theta}(x-\mu)^2\} = \exp\{-\frac{1}{2}\log(2\pi\theta) - \frac{1}{2\theta}(x-\mu)^2\}$

$\therefore f_X(x;\theta)$ belongs to the exponential family of p.d.f. where $c(\theta) = -\frac{1}{2\theta}$ and $d(X) = (X-\mu)^2$.

Since $c(\theta) = -\frac{1}{2\theta}$ is increasing, the critical region of the UMP test is in the form

$C_1 = \{\underset{\sim}{x} : \sum_{i=1}^n d(x_i) \geq K\} = \{\underset{\sim}{x} : \sum_{i=1}^n (x_i - \mu)^2 \geq K\}$

Under $H_0: \theta = \theta_0$, $\frac{1}{\theta_0}\sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2_n$.

$\alpha = P(\underset{\sim}{X} \in C_1 \mid H_0)$

$= P(\sum_{i=1}^n (X_i-\mu)^2 \geq K \mid \theta = \theta_0) = P(\frac{1}{\theta_0}\sum_{i=1}^n (X_i-\mu)^2 \geq \frac{1}{\theta_0}K \mid \theta = \theta_0)$

$= P(\chi^2_n \geq \frac{1}{\theta_0}K)$

$\Rightarrow \frac{1}{\theta_0}K = \chi^2_n(\alpha) \Rightarrow K = \theta_0\chi^2_n(\alpha)$

$\therefore$ the UMP test at level of significance $\alpha$ is to reject $H_0$ when $\sum_{i=1}^n (X_i-\mu)^2 \geq \theta_0\chi^2_n(\alpha)$

(b) When $\theta_0 = 4$, $\alpha = 0.05$ and $n = 25$, the power at $\theta = 12$ is

$P(\underset{\sim}{X} \in C_1 \mid \theta = 12)$

$= P(\sum_{i=1}^{25} (X_i-\mu)^2 \geq (4)\chi^2_{25}(0.05) \mid \theta = 12)$

$= P(\frac{1}{12}\sum_{i=1}^{25} (X_i-\mu)^2 \geq \frac{1}{3}\chi^2_{25}(0.05) \mid \theta = 12)$

$= P(\chi^2_{25} \geq \frac{1}{3}\chi^2_{25}(0.05)) = P(\chi^2_{25} \geq \frac{1}{3}(37.652))$

$= P(\chi^2_{25} \geq 12.551)$

$> P(\chi^2_{25} \geq 13.12)$

$= 0.975$

4. $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are two independent random samples with p.d.f.'s $f_1$ and $f_2$ respectively.

$f_1(x;\theta_1) = \frac{1}{\theta_1}\exp\{-\frac{x}{\theta_1}\} I_{(0,\infty)}(x) \qquad \theta_1 \in \Omega = (0,\infty)$

$f_2(x;\theta_2) = \frac{1}{\theta_2}\exp\{-\frac{y}{\theta_2}\} I_{(0,\infty)}(y) \qquad \theta_2 \in \Omega$

(a) $\begin{cases} H_0: \frac{\theta_1}{\theta_2} = \Delta_0 \\ H_1: \frac{\theta_1}{\theta_2} \neq \Delta_0 \end{cases}$

$$L(\theta_1,\theta_2) = f_{\underset{\sim}{X},\underset{\sim}{Y}}(\underset{\sim}{x},\underset{\sim}{y};\theta_1,\theta_2) = \prod_{i=1}^{m} f_{X_i}(x_i;\theta_1)\prod_{j=1}^{n} f_{Y_j}(y_j;\theta_2)$$

$$= \Big[\prod_{i=1}^{m}\Big(\frac{1}{\theta_1}\Big)\exp\Big\{-\frac{x_i}{\theta_1}\Big\}\Big]\Big[\prod_{j=1}^{n}\Big(\frac{1}{\theta_2}\Big)\exp\Big\{-\frac{y_j}{\theta_2}\Big\}\Big] = \theta_1^{-m}\theta_2^{-n}\exp\Big\{-\sum_{i=1}^{m}\frac{x_i}{\theta_1} - \sum_{j=1}^{n}\frac{y_j}{\theta_2}\Big\}$$

The likelihood ratio test is to reject $H_0$ when

$$\lambda(\underset{\sim}{x},\underset{\sim}{y}) = \frac{\sup\{L(\theta_1,\theta_2) : (\theta_1,\theta_2)\in\Theta_0\}}{\sup\{L(\theta_1,\theta_2) : (\theta_1,\theta_2)\in\Theta\}} \le k$$

where $\Theta = \{(\theta_1,\theta_2) : \theta_1\in\Omega, \theta_2\in\Omega\}$

$\Theta_0 = \{(\theta_1,\theta_2) : \frac{\theta_1}{\theta_2} = \Delta_0, \theta_1\in\Omega, \theta_2\in\Omega\}$

For the numerator, $\left(\frac{\theta_1}{\theta_2} = \Delta_0 \Rightarrow \theta_1 = \theta_2\Delta_0\right)$

$$L(\theta_1,\theta_2) = L(\theta_2\Delta_0,\theta_2) = (\theta_2\Delta_0)^{-m}(\theta_2)^{-n}\exp\Big\{-\sum_{i=1}^{m}\frac{x_i}{\theta_2\Delta_0} - \sum_{j=1}^{n}\frac{y_j}{\theta_2}\Big\}$$

$$= \Delta_0^{-m}\theta_2^{-(m+n)}\exp\Big\{-\frac{1}{\theta_2}\Big(\sum_{i=1}^{m}\frac{x_i}{\Delta_0} + \sum_{j=1}^{n} y_j\Big)\Big\}$$

$$\log L(\theta_2\Delta_0,\theta_2) = -m\log\Delta_0 - (m+n)\log\theta_2 - \frac{1}{\theta_2}\Big(\sum_{i=1}^{m}\frac{x_i}{\Delta_0} + \sum_{j=1}^{n} y_j\Big)$$

$$\frac{\partial}{\partial\theta_2}\log L(\theta_2\Delta_0,\theta_2) = -\frac{m+n}{\theta_2} + \frac{1}{\theta_2^2}\Big(\sum_{i=1}^{m}\frac{x_i}{\Delta_0} + \sum_{j=1}^{n} y_j\Big)$$

$$\frac{\partial}{\partial\theta_2}\log L(\theta_2\Delta_0,\theta_2)\Big|_{\theta_2=\hat\theta_{02}} = 0 \Rightarrow -\frac{m+n}{\hat\theta_{02}} + \frac{1}{\hat\theta_{02}^2}\Big(\sum_{i=1}^{m}\frac{x_i}{\Delta_0} + \sum_{j=1}^{n} y_j\Big) = 0$$

$$\Rightarrow \hat\theta_{02} = \frac{1}{m+n}\Big(\frac{1}{\Delta_0}\sum_{i=1}^{m} x_i + \sum_{j=1}^{n} y_j\Big)$$

For the denominator,

$$\log L(\theta_1,\theta_2) = -m\log\theta_1 - n\log\theta_2 - \sum_{i=1}^{m}\frac{x_i}{\theta_1} - \sum_{j=1}^{n}\frac{y_j}{\theta_2}$$

$$\frac{\partial}{\partial\theta_1}\log L(\theta_1,\theta_2) = -\frac{m}{\theta_1} + \frac{1}{\theta_1^2}\sum_{i=1}^{m} x_i$$

$$\frac{\partial}{\partial\theta_2}\log L(\theta_1,\theta_2) = -\frac{n}{\theta_2} + \frac{1}{\theta_2^2}\sum_{j=1}^{n} y_j$$

$$\begin{cases} \frac{\partial}{\partial\theta_1}\log L(\theta_1,\theta_2)\big|_{(\theta_1,\theta_2)=(\hat\theta_1,\hat\theta_2)} = 0 \\ \frac{\partial}{\partial\theta_2}\log L(\theta_1,\theta_2)\big|_{(\theta_1,\theta_2)=(\hat\theta_1,\hat\theta_2)} = 0 \end{cases}$$

$$\Rightarrow \begin{cases} -\frac{m}{\hat\theta_1} + \frac{1}{\hat\theta_1^2}\sum_{i=1}^{m} x_i = 0 \\ -\frac{n}{\hat\theta_2} + \frac{1}{\hat\theta_2^2}\sum_{j=1}^{n} y_j = 0 \end{cases} \Rightarrow \begin{cases} \hat\theta_1 = \frac{1}{m}\sum_{i=1}^{m} x_i \\ \hat\theta_2 = \frac{1}{n}\sum_{j=1}^{n} y_j \end{cases}$$

$$\therefore \lambda(\underset{\sim}{x},\underset{\sim}{y}) = \frac{L(\hat\theta_{02}\Delta_0,\hat\theta_{02})}{L(\hat\theta_1,\hat\theta_2)} = \frac{\Delta_0^{-m}\hat\theta_{02}^{-(m+n)}\exp\Big\{-\frac{1}{\hat\theta_{02}}\Big(\sum_{i=1}^{m}\frac{x_i}{\Delta_0} + \sum_{j=1}^{n} y_j\Big)\Big\}}{\hat\theta_1^{-m}\hat\theta_2^{-n}\exp\Big\{-\frac{1}{\hat\theta_1}\sum_{i=1}^{m} x_i - \frac{1}{\hat\theta_2}\sum_{j=1}^{n} y_j\Big\}}$$

$$= \frac{\Delta_0^{-m}(m+n)^{m+n}\Big[\frac{1}{\Delta_0}\sum_{i=1}^{m} x_i + \sum_{j=1}^{n} y_j\Big]^{-(m+n)}}{m^m\Big(\sum_{i=1}^{m} x_i\Big)^{-m} n^n\Big(\sum_{j=1}^{n} y_j\Big)^{-n}} \le k$$

4. (a) (cont'd)

$$\Leftrightarrow \frac{\left(\frac{1}{\Delta_0}\sum_{i=1}^m x_i\right)^m \left(\sum_{j=1}^n y_j\right)^n}{\left(\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j\right)^{m+n}} \le k' \Leftrightarrow \left(\frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j}\right)^m \left(\frac{\sum_{j=1}^n y_j}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j}\right)^n \le k'$$

$$\Leftrightarrow \left(\frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j}\right)^m \left(1 - \frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j}\right)^n \le k'$$

Since $0 \le \frac{1}{\Delta_0}\sum_{i=1}^m x_i \Big/ \left(\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j\right) \le 1$, $\lambda(\underset{\sim}{x}, \underset{\sim}{y}) \le k$

$$\Leftrightarrow \frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j} \le k_1 \quad \text{or} \quad \frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j} \ge k_2$$

Under $H_0: \frac{\theta_1}{\theta_2} = \Delta_0$,

$\frac{1}{\Delta_0}\sum_{i=1}^m X_i \sim Gamma(m, \frac{1}{\theta_2})$ and $\sum_{j=1}^n Y_j \sim Gamma(n, \frac{1}{\theta_2})$ are independent.

$$\therefore \frac{\frac{1}{\Delta_0}\sum_{i=1}^m X_i}{\frac{1}{\Delta_0}\sum_{i=1}^m X_i + \sum_{j=1}^n Y_j} \sim Beta(m, n)$$

$\therefore$ the likelihood ratio test at level of significance $\alpha$ is to reject $H_0$ when

$$\frac{\frac{1}{\Delta_0}\sum_{i=1}^m X_i}{\frac{1}{\Delta_0}\sum_{i=1}^m X_i + \sum_{j=1}^n Y_j} \le Beta(m, n; 1-\frac{\alpha}{2})$$

or

$$\frac{\frac{1}{\Delta_0}\sum_{i=1}^m X_i}{\frac{1}{\Delta_0}\sum_{i=1}^m X_i + \sum_{j=1}^n Y_j} \ge Beta(m, n; \frac{\alpha}{2})$$

(b) Under $H_0: \frac{\theta_1}{\theta_2} = \Delta_0$,

$2\left(\frac{1}{\theta_2\Delta_0}\right)\sum_{i=1}^m X_i \sim \chi^2_{2m}$ and $2\left(\frac{1}{\theta_2}\right)\sum_{j=1}^n Y_j \sim \chi^2_{2n}$ are independent,

$$\therefore \frac{2\left(\frac{1}{\theta_2}\right)\sum_{j=1}^n Y_j/(2n)}{2\left(\frac{1}{\theta_2\Delta_0}\right)\sum_{i=1}^m X_i/(2m)} \sim F_{2n,2m} \quad \text{i.e.} \quad \frac{\sum_{j=1}^n Y_j/n}{\frac{1}{\Delta_0}\sum_{i=1}^m X_i/m} \sim F_{2n,2m}$$

$$\lambda(\underset{\sim}{x}, \underset{\sim}{y}) \le k \Leftrightarrow \frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j} \le k_1 \quad \text{or} \quad \frac{\frac{1}{\Delta_0}\sum_{i=1}^m x_i}{\frac{1}{\Delta_0}\sum_{i=1}^m x_i + \sum_{j=1}^n y_j} \ge k_2$$

$$\Leftrightarrow \frac{1}{1 + \left(\sum_{j=1}^n y_j\right)/\left(\frac{1}{\Delta_0}\sum_{i=1}^m x_i\right)} \le k_1 \quad \text{or} \quad \frac{1}{1 + \left(\sum_{j=1}^n y_j\right)/\left(\frac{1}{\Delta_0}\sum_{i=1}^m x_i\right)} \ge k_2$$

$$\Leftrightarrow \left(\frac{1}{n}\sum_{j=1}^n y_j\right)\Big/\left(\frac{1}{m\Delta_0}\sum_{i=1}^m x_i\right) \le k_1' \quad \text{or} \quad \left(\frac{1}{n}\sum_{j=1}^n y_j\right)\Big/\left(\frac{1}{m\Delta_0}\sum_{i=1}^m x_i\right) \ge k_2'$$

$\therefore$ the likelihood ratio test at level of significance $\alpha$ reduces to reject $H_0$ when

$$\frac{\frac{1}{n}\sum_{j=1}^n Y_j}{\frac{1}{m\Delta_0}\sum_{i=1}^m X_i} \le F_{2n,2m}(1-\frac{\alpha}{2}) \quad \text{or} \quad \frac{\frac{1}{n}\sum_{j=1}^n Y_j}{\frac{1}{m\Delta_0}\sum_{i=1}^m X_i} \ge F_{2n,2m}(\frac{\alpha}{2})$$

5. $X_1, \ldots, X_{100}$ is a r.s. from $Bin(1, P)$.
Now, $\sum_{i=1}^{100} x_i = 60$ is observed.

(a) $\begin{cases} H_0: P = \frac{1}{2} \\ H_1: P \neq \frac{1}{2} \end{cases}$

$$L(P) = f_X(\underset{\sim}{x}; P) = \prod_{i=1}^{100} f_{X_i}(x_i; P) = \prod_{i=1}^{100} P^{x_i}(1-P)^{1-x_i}$$
$$= P^{\sum_{i=1}^{100} x_i}(1-P)^{100-\sum_{i=1}^{100} x_i}$$

The likelihood ratio test is to reject $H_0$ when

$$\lambda(\underset{\sim}{x}) = \frac{\sup\{L(P) : P \in \Theta_0\}}{\sup\{L(P) : P \in \Theta\}} \le k \quad \text{where } \Theta = (0,1) \text{ and } \Theta_0 = \{\tfrac{1}{2}\}$$

The numerator $(\Theta_0 = \{\frac{1}{2}\})$,

$$\sup\{L(P) : P \in \Theta_0\} = L(\tfrac{1}{2}) = (\tfrac{1}{2})^{\sum_{i=1}^{100} x_i}(1-\tfrac{1}{2})^{100-\sum_{i=1}^{100} x_i} = (\tfrac{1}{2})^{100}$$

The denominator $(\Theta = (0,1))$,

$$\log L(P) = \sum_{i=1}^{100} x_i \log P + (100 - \sum_{i=1}^{100} x_i)\log(1-P)$$
$$\frac{\partial}{\partial P}\log L(P) = \frac{1}{P}\sum_{i=1}^{100} x_i - \frac{1}{1-P}(100 - \sum_{i=1}^{100} x_i)$$
$$\frac{\partial}{\partial P}\log L(P)\Big|_{P=\hat{P}} = 0 \Rightarrow \hat{P} = \frac{1}{100}\sum_{i=1}^{100} x_i$$

$$\therefore \lambda(\underset{\sim}{x}) = \frac{\sup\{L(P) : P \in \Theta_0\}}{\sup\{L(P) : P \in \Theta\}} = \frac{(\frac{1}{2})^{100}}{(\frac{1}{100}\sum_{i=1}^{100} x_i)^{\sum_{i=1}^{100} x_i}(1 - \frac{1}{100}\sum_{i=1}^{100} x_i)^{100-\sum_{i=1}^{100} x_i}} \le k$$

$$\Leftrightarrow (\bar{x})^{\bar{x}}(1-\bar{x})^{1-\bar{x}} \ge k'$$

$$\Leftrightarrow \bar{x} \le k_1 \text{ or } \bar{x} \ge k_2 \Leftrightarrow \sum_{i=1}^{100} x_i \le k_1' \text{ or } \sum_{i=1}^{100} x_i \ge k_2'$$

Under $H_0: P = \frac{1}{2}$, $\sum_{i=1}^{100} X_i \mathrel{\dot\sim} N(100(\frac{1}{2}), 100(\frac{1}{2})(1-\frac{1}{2})) \sim N(50, 25)$

$\therefore$ the likelihood ratio test at level of significance $\alpha = 0.1$ is to reject $H_0$ when $\sum_{i=1}^{100} X_i \le k_1'$ or $\sum_{i=1}^{100} X_i \ge k_2'$ where ($k_1'$ and $k_2'$ are integers)

$$P(\sum_{i=1}^{100} X_i \le k_1' | H_0) = \frac{\alpha}{2} = 0.05 \quad \text{and} \quad P(\sum_{i=1}^{100} X_i \ge k_2' | H_0) = \frac{\alpha}{2} = 0.05$$

$$\Rightarrow P\left(\frac{\sum_{i=1}^{100} X_i - 50}{\sqrt{25}} \le \frac{k_1' + 0.5 - 50}{\sqrt{25}} \Big| H_0\right) = 0.05 \text{ and } P\left(\frac{\sum_{i=1}^{100} X_i - 50}{\sqrt{25}} \ge \frac{k_2' - 0.5 - 50}{\sqrt{25}}\right) = 0.05$$

$$\Rightarrow P(Z \le \tfrac{1}{5}(k_1' - 49.5)) \approx 0.05 \quad \text{and} \quad P(Z \ge \tfrac{1}{5}(k_2' - 50.5)) \approx 0.05$$

$$\Rightarrow \tfrac{1}{5}(k_1' - 49.5) \approx -1.645 \quad \text{and} \quad \tfrac{1}{5}(k_2' - 50.5) \approx 1.645$$

$$\Rightarrow k_1' \approx 41.275 \ (\therefore k_1' = 41) \quad \text{and} \quad k_2' \approx 58.725 \ (\therefore k_2' = 59)$$

Now, $\sum_{i=1}^{100} x_i = 60 \ge 59$

$\therefore H_0$ is rejected at level of significance $\alpha = 0.1$.

5(b) $\begin{cases} H_0: P = \frac{1}{2} \\ H_1: P \neq \frac{1}{2} \end{cases}$

The chi-square goodness-of-fit test is to reject $H_0$ when

$$\sum_{i=1}^{2} \frac{(n_i - n\theta_{0i})^2}{n\theta_{0i}} \geq \chi^2_{2-1}(0.1) = \chi^2_1(0.1) = 2.71$$

where $n_1 = \sum_{i=1}^{100} x_i$, $n_2 = n - n_1$, $n = 100$ and $\theta_{0i} = \frac{1}{2}$

Now, $\sum_{i=1}^{100} x_i = 60$

$$\therefore \frac{(60 - 100(\frac{1}{2}))^2}{100(\frac{1}{2})} + \frac{((100-60) - 100(\frac{1}{2}))^2}{100(\frac{1}{2})} = 2 + 2 = 4 \geq 2.71$$

$\therefore H_0$ is rejected at level of significance $\alpha = 0.1$