INFERENCE: Homework $\#5$

Professor Regina

Fan Yang
UNI: fy2232

Problem 1

(a)

$$p - value = P_{H_0}(Y = 6) = \binom{100}{6} p^6 (1-p)^{94} = 0.1232795$$

So $\alpha = 0.1232795$

(b)

$$\begin{aligned} P_{H_A}(\text{fail to reject } H_0) &= P_{H_A}(Y \neq 6) \\ &= 1 - P_{H_A}(Y = 6) \\ &= 1 - \binom{100}{6} p^6 (1-p)^{94} \\ &= 0.8947672 \end{aligned}$$

Problem 2

$$\bar{Y} = 300 \times p = 300p$$
$$s^2 = 300 \times p \times (1-p) = 300p(1-p)$$

$$Pr\left(-Z_{\alpha/2} \leq \frac{Y - np}{\sqrt{np(1-p)}} \leq Z_{\alpha/2}\right) = 1 - \alpha$$

$$Pr\left(\frac{Y}{n} - Z_{\alpha/2}\sqrt{np(1-p)} \leq p \leq \frac{Y}{n} + Z_{\alpha/2}\sqrt{np(1-p)}\right) = 1 - \alpha$$

We estimate p by $\hat{p}_{MLE} = \frac{Y}{n}$

So an approximate confidence interval for p

$$\hat{p} \pm Z_{\alpha/2}\sqrt{n\hat{p}(1-\hat{p})}$$
$$0.25 \pm 1.64 \times 0.025$$

which is $[0.2088787,\ 0.2911213]$

Problem 3

Because $\Gamma(4) = 6$, $\alpha = 4$

$$\bar{X} = \alpha\beta = 4\beta$$
$$s^2 = \alpha\beta^2 = 4\beta^2$$

By central limit theorem,

$$Pr\left(-Z_{\alpha/2} \le \frac{\bar{X}-\mu}{s}\sqrt{n} \le Z_{\alpha/2}\right) = 1-\alpha$$
$$Pr\left(\bar{X} - Z_{\alpha/2}s/\sqrt{n} \le \mu \le \bar{X} + Z_{\alpha/2}s/\sqrt{n}\right) = 1-\alpha$$

and $Z_{\alpha/2} = 1.995393$, $s = 2\beta$

$$Pr\left(\bar{X} - 2Z_{\alpha/2}\beta/\sqrt{n} \le 4\beta \le \bar{X} + 2Z_{\alpha/2}\beta/\sqrt{n}\right) = 1-\alpha$$
$$Pr\left(\frac{\bar{X}}{4+2Z_{\alpha/2}/\sqrt{n}} \le \beta \le \frac{\bar{X}}{4-2Z_{\alpha/2}/\sqrt{n}}\right) = 1-\alpha$$
$$Pr\left(\frac{4\bar{X}}{4+2Z_{\alpha/2}/\sqrt{n}} \le \mu \le \frac{4\bar{X}}{4-2Z_{\alpha/2}/\sqrt{n}}\right) = 1-\alpha$$

The confidence interval is $\dfrac{4\bar{X}}{4.798157} \le \mu \le \dfrac{4\bar{X}}{3.201843}$

or $0.8336534\bar{X} \le \mu \le 1.24928\bar{X}$

Problem 4

$$\begin{aligned}\beta(p) &= P_p(reject\ H_0) = P_p(X_1 \le 3)\\ &= \sum_{i=0}^{3}\binom{10}{i}p^i(1-p)^{10-i}\\ &= (1-p)^{10} + 10p(1-p)^9 + 45p^2(1-p)^8 + 120p^3(1-p)^7\end{aligned}$$

$$\beta(\frac{1}{2}) = \frac{11}{64}$$
$$\beta(\frac{1}{4}) = \frac{134567}{173439} = 0.7758751$$

Problem 5

Likelihood function is

$$f_{12}(x_1, x_2; \theta) = f(x_1; \theta) f(x_2; \theta) = \frac{1}{\theta^2} e^{-(x_1+x_2)/\theta}$$

$y_1 = x_1 + x_2 \quad y_2 = x_2$, which means $x_1 = y_1 - y_2 \quad x_2 = y_2 \qquad 0 < y_2 < y_1 < \infty$

Then the Jacobian is

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

$$\begin{aligned} f(y_1, y_2) &= f(x_1, x_2)|J| \\ &= \frac{1}{\theta^2} e^{-y_1/\theta} \end{aligned}$$

$$\begin{aligned} E(Y_2) = E(X_2) &= \int_0^\infty x f(x; \theta)\ dx \\ &= \int_0^\infty x/\theta e^{-x/\theta}\ dx \\ &= \theta \\ Var(Y_2) = Var(X_2) &= E(X_2^2) - E^2(X_2) \\ &= \int_0^\infty x^2 f(x; \theta)\ dx - \theta^2 \\ &= \int_0^\infty x^2/\theta e^{-x/\theta}\ dx - \theta^2 \\ &= \theta^2 \end{aligned}$$

$$\begin{aligned} f(y_1) &= \int_0^{y_1} f(y_1, y_2)\ dy_2 \\ &= \int_0^{y_1} \frac{1}{\theta^2} e^{-y_1/\theta}\ dy_2 \\ &= \frac{y_1}{\theta^2} e^{-y_1/\theta} \end{aligned}$$

$$\begin{aligned} f(y_2|y_1) &= \frac{f(y_1, y_2)}{f(y_1)} \\ &= \frac{1/\theta^2 e^{-y_1/\theta}}{y_1/\theta^2 e^{-y_1/\theta}} \\ &= \frac{1}{y_1} \end{aligned}$$

$$E(Y_2|y_1) = \int_0^{y_1} \frac{y_2}{y_1}\, dy_2 = \frac{y_1}{2} = \phi(y_1)$$

Since $\phi(y_1) = \frac{y_1}{2}$,

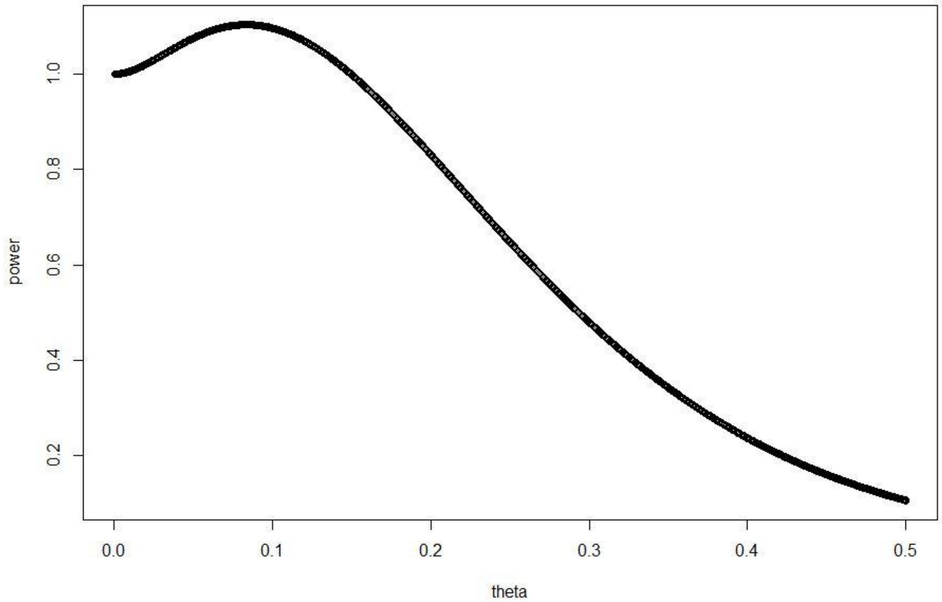
$$\begin{aligned} Var(\phi(Y_1)) &= Var(Y_1/2) \\ &= \frac{Var(X_1) + Var(X_2)}{4} \\ &= \frac{\theta^2}{2} \end{aligned}$$

Problem 6

$$\begin{aligned} Y = X_1 + X_2 + ... + X_{12} &\sim Poisson(n\theta) \\ \beta(\theta) = P_\theta(reject\ H_0) &= P_\theta(Y \leq 2) \\ &= e^{-n\theta} \sum_{i=0}^{2} \frac{(n\theta)^i}{i!} \\ &= e^{-n\theta}(1 + n\theta + n^2\theta^2) \end{aligned}$$

$$\beta(\theta) = \begin{cases} 43e^{-6}, & \theta = 1/2 \\ 21e^{-4}, & \theta = 1/3 \\ 13e^{-3}, & \theta = 1/4 \\ 7e^{-2}, & \theta = 1/6 \\ 3e^{-1}, & \theta = 1/12 \end{cases}$$

The plot of $\beta(\theta)$ is below. significance level is $\beta_{H_0}(\theta) = 43e^{-6} = 0.1065863$



Problem 7

(a)

$$f_4(y) = 4F_x^3(y)f_x(y) = 4(\frac{y}{\theta})^3\frac{1}{\theta}$$
$$F_4(y) = F_x^4(y) = (\frac{y}{\theta})^4$$

$$P_{H_0}(Y_4 \geq c) = 0.05$$
$$P_{H_0}(Y_4 \leq c) = 0.95$$
$$F_4(c) = 0.95$$
$$(\frac{c}{\theta})^4 = 0.95$$
$$c = 0.95^{1/4} = 0.9872585$$

(b)

$$\begin{aligned}\beta(\theta) &= P_\theta(reject\ H_0) = P_\theta(Y_4 \geq c)\\ &= 1 - P_\theta(Y_4 \leq c)\\ &= 1 - F_4(c)\\ &= 1 - (\frac{c}{\theta})^4\end{aligned}$$

Problem 8

$$P_{H_0}(nS^2/\sigma_0^2 \geq c) = 0.025$$
$$P_{H_0}((n-1)S^2/\sigma_0^2 \geq \frac{n-1}{n}c) = 0.025$$

$$(n-1)S^2/\sigma_0^2 \sim \chi^2(n-1)$$

$$\begin{aligned}\frac{n-1}{n}c &= \chi^2_{0.975}(n-1)\\ \frac{12}{13}c &= 23.33666\\ c &= 25.28139\end{aligned}$$

Problem 9 (#10 on page 529)

$$Pr\left(-t_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma}\sqrt{n} \leq t_{\alpha/2}\right) = 1-\alpha$$
$$Pr\left(\bar{X} - t_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{X} + t_{\alpha/2}\sigma/\sqrt{n}\right) = 1-\alpha$$

The length of this interval is $2 \times t_{\alpha/2}\sigma/\sqrt{n}$

Then the squared length of this interval is $4t^2_{\alpha/2}(n-1)\frac{\sigma^2}{n}$

$$\begin{aligned}4t^2_{\alpha/2}\frac{\sigma^2}{n} &< \sigma^2/2\\ n &> 8t^2_{\alpha/2}(n-1)\\ n &= 24\end{aligned}$$

Problem 10 (#13 on page 529)

(a)

The likelihood function is

$$\begin{aligned} p(\theta|x_1,...,x_n) &\propto p(\theta)p(x_1,...,x_n|\theta) \\ &\propto exp\left\{-\frac{(\theta-\mu)^2}{2v^2}\right\} exp\left\{-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\} \\ &\propto exp\left\{-\frac{(\theta-\mu)^2}{2v^2}-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\} \end{aligned}$$

which is also a normal distribution with mean $\frac{\sigma^2\mu+nv^2\bar{x}_n}{\sigma^2+nv^2}$ and variance $\frac{\sigma^2 v^2}{\sigma^2+nv^2}$

$$\mu^{`} = \frac{\sigma^2\mu+nv^2\bar{X}_n}{\sigma^2+nv^2}$$

$$\sigma^{2^{`}} = \frac{\sigma^2 v^2}{\sigma^2+nv^2}$$

Then the interval should be $\mu^{`} - Z_{\alpha/2}\sigma^{`} \leq \theta \leq \mu^{`} + Z_{\alpha/2}\sigma^{`}$

which is $\mu^{`} - 1.959964\sigma^{`} \leq \theta \leq \mu^{`} + 1.959964\sigma^{`}$

(b)

As $v^2 \to \infty$, $\mu^{`} = \frac{\sigma^2\mu+nv^2\bar{X}_n}{\sigma^2+nv^2} \to \bar{X}_n$ and $\sigma^{2^{`}} = \frac{\sigma^2 v^2}{\sigma^2+nv^2} \to \frac{\sigma^2}{n}$

then the interval becomes $\bar{X}_n - Z_{\alpha/2}\frac{\sigma^2}{n} \leq \theta \leq \bar{X}_n + Z_{\alpha/2}\frac{\sigma^2}{n}$, which is the same as the confidence interval of θ

Problem 11 (#22 on page 529)

(a)

$$\begin{aligned} Pr(Y_n/\theta < y) &= Pr(Y_n < y\theta) \\ &= Pr^n(X_i < y\theta) \\ &= (\frac{y\theta}{\theta})^n \\ &= y^n \end{aligned}$$

so the pdf of Y_n/θ is $f(y) = ny^{n-1} \qquad y \geq \frac{max\{x_1,\ldots,x_n\}}{\theta}$

(b)

$$\begin{aligned} E(Y_n/\theta) &= \int_0^1 y \times ny^{n-1}\ dy = \frac{n}{n+1} \\ E(Y_n) &= \frac{n}{n+1}\theta \\ bias &= E(Y_n) - \theta = -\frac{1}{n+1}\theta \end{aligned}$$

(d)

$$\begin{aligned} Pr(a \leq \theta \leq b) &= \gamma \\ Pr(\frac{Y_n}{b} \leq \frac{Y_n}{\theta} \leq \frac{Y_n}{a}) &= \gamma \\ F(\frac{Y_n}{a}) - F(\frac{Y_n}{b}) &= \gamma \\ (\frac{Y_n}{a})^n - (\frac{Y_n}{b})^n &= \gamma \end{aligned}$$

We can set $a = \frac{Y_n}{(\frac{1+\gamma}{2})^{1/n}}$ and $b = \frac{Y_n}{(\frac{1-\gamma}{2})^{1/n}}$, then $Pr(a \leq \theta \leq b) = \gamma$

Problem 12 (#9 on page 622)

Use Neymann-Pearson Lemma

$$
\begin{aligned}
C &= \{\underset{\sim}{x} : \frac{f_0(x)}{f_1(x)} \le k\} \\
&= \{\underset{\sim}{x} : \frac{1}{\frac{1}{\sqrt{2\pi}} exp\{-\frac{x^2}{2}\}} \le k\}
\end{aligned}
$$

Size α test:

$$
P_{H_0}\left(\frac{1}{\frac{1}{\sqrt{2\pi}} exp\{-\frac{x^2}{2}\}} \le k\right) = \alpha
$$

$$
P_{H_0}\left(\frac{1}{\sqrt{2\pi}} exp\{-\frac{x^2}{2}\} \ge \frac{1}{k}\right) = \alpha
$$

$$
P_{H_0}\left(x \le \sqrt{2ln\frac{\sqrt{2\pi}}{k}}\right) = \alpha
$$

So $\sqrt{2ln\frac{\sqrt{2\pi}}{k}} = \alpha = 0.01$ thus $k = 2.506503$

$$
\begin{aligned}
C &= \{\underset{\sim}{x} : \frac{f_0(x)}{f_1(x)} \le 2.506503\} \\
&= \{\underset{\sim}{x} : \frac{1}{\frac{1}{\sqrt{2\pi}} exp\{-\frac{x^2}{2}\}} \le 2.506503\} \\
&= \{\underset{\sim}{x} : x \le 0.01\}
\end{aligned}
$$

When $x < 0$ or $x > 0$, $\frac{f_0(x)}{f_1(x)} = 0 < k$, so the criterion region should be modified to

$$
C = \{\underset{\sim}{x} : x \le 0.01 \text{ or } x > 1\}
$$

The power when H_1 is true is

$$
\begin{aligned}
P_{H_1}(x \le 0.01) + P_{H_1}(x > 1) &= \Phi(0.01) + 1 - \Phi(1) \\
&= 0.6626446
\end{aligned}
$$

Problem 13 (#11 on page 622)

Use Neymann-Pearson Lemma

$$\begin{aligned}
C &= \left\{ \frac{L(\theta_0)}{L(\theta_1)} \le k \right\} \\
&= \left\{ \frac{exp\{-\frac{\sum(X_i-\theta_0)^2}{2}\}}{exp\{-\frac{\sum(X_i-\theta_1)^2}{2}\}} \le k \right\} \\
&= \left\{ exp\{-\frac{\sum(X_i-\theta_0)^2}{2} + \frac{\sum(X_i-\theta_1)^2}{2}\} \le k \right\} \\
&= \left\{ exp\{\frac{\theta_0 \sum X_i - \theta_1 \sum X_i - n\theta_0^2 + n\theta_1^2}{2}\} \le k \right\} \\
&= \left\{ (\theta_0 - \theta_1) \sum X_i - n(\theta_0^2 - \theta_1^2) \le 2lnk \right\} \\
&= \left\{ \bar{X_n} \ge k_1 \right\}
\end{aligned}$$

Under H_1, $\bar{X_n} \sim N(1, \frac{1}{\sqrt{n}})$

$$\begin{aligned}
P_{\theta=1}(\bar{X_n} \ge k_1) &= 0.95 \\
P(Z \ge (k_1 - 1)\sqrt{n}) &= 0.95 \\
P(Z \le k_1\sqrt{n} - \sqrt{n}) &= 0.05 \\
k_1 * 4 - 4 &= -1.644854 \\
k_1 &= 0.5887866
\end{aligned}$$

Under H_0, $\bar{X_n} \sim N(\theta_0, \frac{1}{\sqrt{n}})$

$$\begin{aligned}
P_{H_0}(\bar{X_n} \ge k_1) &= P(Z \ge (k_1 - \theta_0)\sqrt{n}) \\
&\le P_{\theta_0=0}(Z \ge k_1\sqrt{n}) \\
&\le P_{\theta_0=0}(Z \ge 0.5887866 * 4) \\
&\le P_{\theta_0=0}(Z \ge 2.355146) \\
&= 0.009257715
\end{aligned}$$

Problem 14 (#12 on page 622)

Use Neymann-Pearson Lemma

$$\begin{aligned} C &= \left\{ \frac{L(\theta_0)}{L(\theta_1)} \le k \right\} \\ &= \left\{ \frac{\theta_0^n (\prod X_i)^{\theta_0 - 1}}{\theta_1^n (\prod X_i)^{\theta_1 - 1}} \le k \right\} \\ &= \left\{ (\frac{\theta_0}{\theta_1})^n (\prod X_i)^{\theta_0 - \theta_1} \le k \right\} \end{aligned}$$

Because $\theta_0 < \theta_1$, $(\frac{\theta_0}{\theta_1})^n (\prod X_i)^{\theta_0 - \theta_1}$ is a decreasing function of $\prod X_i$. So,

$$\begin{aligned} C &= \left\{ \prod X_i \ge k_1 \right\} \\ &= \left\{ \sum \log X_i \ge k_2 \right\} \end{aligned}$$

Under H_0, $f(x|\theta) = 1$, $-2\sum \log X_i \sim \chi^2(2n)$

$$\begin{aligned} P_{H_0}(\sum \log X_i \ge k_2) &= 0.05 \\ P_{H_0}(-2\sum \log X_i \le -2k_2) &= 0.05 \\ P(\chi^2 \le -2k_2) &= 0.05 \\ -2k_2 &= 7.961646 \\ k_2 &= -3.980823 \end{aligned}$$

$$C = \left\{ \sum \log X_i \ge -3.980823 \right\}$$

Problem 15 (#13 on page 622)

Use Neymann-Pearson Lemma

$$\begin{aligned} C &= \left\{ \frac{L(\theta_0)}{L(\theta_1)} \le k \right\} \\ &= \left\{ \frac{\frac{1}{2^{\theta_0/2}\Gamma(\theta_0/2)} (\prod X_i)^{\theta_0/2 - 1} e^{-\sum X_i/2}}{\frac{1}{2^{\theta_1/2}\Gamma(\theta_1/2)} (\prod X_i)^{\theta_1/2 - 1} e^{-\sum X_i/2}} \le k \right\} \\ &= \left\{ 2^{\frac{-\theta_0 + \theta_1}{2}} \frac{\Gamma(\theta_1/2)}{\Gamma(\theta_0/2)} (\prod X_i)^{\frac{\theta_0 - \theta_1}{2}} \le k \right\} \end{aligned}$$

Because $\theta_0 < \theta_1$, $2^{\frac{-\theta_0 + \theta_1}{2}} \frac{\Gamma(\theta_1/2)}{\Gamma(\theta_0/2)} (\prod X_i)^{\frac{\theta_0 - \theta_1}{2}}$ is a decreasing function of $\prod X_i$. So,

$$
\begin{aligned}
C &= \left\{\prod X_i \geq k_1\right\} \\
&= \left\{\sum \log X_i \geq k_2\right\}
\end{aligned}
$$