BIOS760 HOMEWORK VI SOLUTION

1. (a) X can be understood as a bivariate random vector (Y, Z) such that

$$P(Y=1, Z=0)=\theta_1, P(Y=0, Z=1)=\theta_2,$$

$$P(Y=0, Z=0)=\theta_3, P(Y=-1, Z=-1)=\theta_4.$$

Thus, by the CLT,

$$\sqrt{n}(\bar{X}_n - E[X_1]) \to_d N(0, Var(X_1)).$$

Since $E[X_1]=\theta_1(1,0)^T+\theta_2(0,1)^T+\theta_3(0,0)^T+\theta_4(-1,-1)^T=(\theta_1-\theta_4, \theta_2-\theta_4)^T$ and

$$Var(X_1)=\theta_1\begin{pmatrix}1\\0\end{pmatrix}(1,0)+\theta_2\begin{pmatrix}0\\1\end{pmatrix}(0,1)+\theta_3\begin{pmatrix}0\\0\end{pmatrix}(0,0)+\theta_4\begin{pmatrix}-1\\-1\end{pmatrix}(-1,-1)-E[X_1]E[X_1]^T$$

$$=\begin{pmatrix}(\theta_1+\theta_4)-(\theta_1-\theta_4)^2 & \theta_4-(\theta_1-\theta_4)(\theta_2-\theta_4)\\ \theta_4-(\theta_1-\theta_4)(\theta_2-\theta_4) & (\theta_2+\theta_4)-(\theta_2-\theta_4)^2\end{pmatrix},$$

we conclude that in large sample, the two coordinates will follow a distribution approximated by a bivariate normal distribution with mean $(\theta_1-\theta_4, \theta_2-\theta_4)^T$ and covariance

$$\begin{pmatrix}(\theta_1+\theta_4)-(\theta_1-\theta_4)^2 & \theta_4-(\theta_1-\theta_4)(\theta_2-\theta_4)\\ \theta_4-(\theta_1-\theta_4)(\theta_2-\theta_4) & (\theta_2+\theta_4)-(\theta_2-\theta_4)^2\end{pmatrix}/n.$$

(b) By the CLT,

$$\sqrt{n}\left\{\begin{pmatrix}\bar{X}_n\\ Z_n\end{pmatrix}-\begin{pmatrix}\lambda\\ \lambda e^{-\lambda}\end{pmatrix}\right\} \to_d N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\lambda & \lambda(1-\lambda)e^{-\lambda}\\ \lambda(1-\lambda)e^{-\lambda} & \lambda e^{-\lambda}(1-\lambda e^{-\lambda})\end{pmatrix}\right).$$

For $p_1(\bar{X}_n)$, the Delta method gives

$$\sqrt{n}(p_1(\bar{X}_n)-p_1(\lambda)) \to_d N(0, \lambda(1-\lambda)^2 e^{-2\lambda}).$$

Consider $g(x,z)=(z, p_1(x))^T$. Since

$$\nabla g=\begin{pmatrix}0 & 1\\ (1-x)e^{-x} & 0\end{pmatrix},$$

we apply the Delta method to $g(\bar{X}_n, Z_n)$ and obtain

$$\sqrt{n}\left\{\begin{pmatrix}Z_n\\ \hat{p}_1\end{pmatrix}-\begin{pmatrix}\lambda \exp -\lambda\\ \lambda e^{-\lambda}\end{pmatrix}\right\} \to_d N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0 & 1\\ (1-\lambda)e^{-\lambda} & 0\end{pmatrix}\right.$$

$$\left.\times\begin{pmatrix}\lambda & \lambda(1-\lambda)e^{-\lambda}\\ \lambda(1-\lambda)e^{-\lambda} & \lambda e^{-\lambda}(1-\lambda e^{-\lambda})\end{pmatrix}\begin{pmatrix}0 & 1\\ (1-\lambda)e^{-\lambda} & 0\end{pmatrix}^T\right)$$

$$=N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\lambda e^{-\lambda}(1-\lambda e^{-\lambda}) & \lambda(1-\lambda)^2 e^{-2\lambda}\\ \lambda(1-\lambda)^2 e^{-2\lambda} & \lambda(1-\lambda)^2 e^{-2\lambda}\end{pmatrix}\right).$$

(c) Let $Y_1, ..., Y_n$ and $Z_1, ..., Z_n$ be i.i.d $N(0,1)$. Then X_n has the same distribution as $\sqrt{n}\bar{Y}_n/\sqrt{(Z_1^2 + ... + Z_n^2)/n}$. Since $\sqrt{n}\bar{Y}_n \to_d N(0,1)$ and $(Z_1^2+...+Z_n^2)/n \to_p 1$, from the Slutsky theorem, $X_n \to_d N(0,1)$.

2. (a) $n\bar{Y}_n \sim \chi_n^2 = Gamma(2, n/2)$, then $E(n\bar{Y}_n) = n, Var(n\bar{Y}_n) = 2n$, So $E(\bar{Y}_n) = 1, Var(\bar{Y}_n) = 2/n$, by CLT, $\sqrt{n}(\bar{Y}_n - 1) \to_d N(0, \sigma^2)$, where $\sigma^2 = 2$.

(b) From the Delta method,

$$\sqrt{n}(\bar{Y}_n^r - 1) \to N(0, 2r^2)$$

so $V(r) = \sqrt{2}r$.

(c) Since

$$\frac{\sqrt{n}\{\bar{Y}_n^{1/3} - (1 - 2/(9n))\}}{\sqrt{2/9}} = \sqrt{n}(\bar{Y}_n^{1/3} - 1)/\sqrt{2/9} + \sqrt{2/9}/\sqrt{n},$$

from the Slutsky theorem, it converges in distribution $N(0, V(1/3)^3/(2/9)) = N(0,1)$.

(d) For $n = 5$, I plot the normal plots. The y-axis is the normal probabilities at the quartiles of the distributions (a) and (c). I.e, for $0 < p < 1$, the normal plot for (a) is the curve with points

$$(p, \Phi((\chi_n^{2^{-1}}(p) - n)/\sqrt{2n});$$

while the normal plot for (c) is the curve with points

$$(p, \Phi(\{(\chi_n^{2^{-1}}(p)/n)^{1/3} - (1 - 2/(9n))\}/\sqrt{2/9n})).$$

If the approximation is accurate, the curve should be the diagonal line. The plot shows that (c) is more accurate.

Splus(R) codes are:

```
p <- (1:99)/100 plot.760.hw4 <- function(n, p){
    y1 <- pnorm((qchisq(p,n)-n)/sqrt(2*n))
    y2 <- pnorm(((qchisq(p,n)/n)^(1/3)-(1-2/(9*n)))*sqrt(n)/sqrt(2/9))
    plot(p,p, type="n", ylab="Normal probabilities of quartiles",
         xlab="Percentiles",
         xlim=c(0,1), ylim=c(0,1))
    lines(p, y1, type="l", lty=1)
    lines(p, y2, type="l", lty=2)
    legend(0.2, 0.8, c("(a)", "(c)"), lty=1:2)
```

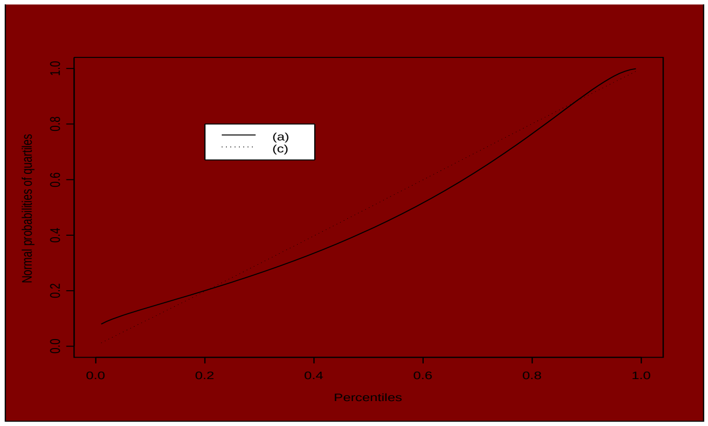

Figure 1: Normal plot for Problem 3(d)

```
} n<- 5 plot.760.hw4(n,p)
```

3. Note
$$\frac{\sqrt{n}(\bar{X}_n - \bar{p}_n)}{\sqrt{n^{-1}\sum_{i=1}^n p_i(1-p_i)}} = \sum_{i=1}^n \frac{X_i - p_i}{\sqrt{\sum_{i=1}^n p_i(1-p_i)}}$$
and $E[|X_i - p_i|^3] = (1-p_i)^3 p_i + p_i^3(1-p_i) \le p_i(1-p_i)$. Then we can verify the Liaponov's condition:
$$\sum_{i=1}^n \frac{E[|X_i - p_i|^3]}{\{\sum_{i=1}^n p_i(1-p_i)\}^{3/2}} \le \frac{1}{\sqrt{\sum_{i=1}^n p_i(1-p_i)}} \to 0.$$
The result follows from the Liaponov's CLT. The two examples can be $p_1 = ... = p_n = p \in (0,1)$ and $p_1 = 1/2, p_2 = p_3 = ... = 0$ respectively. For the latter,
$$\frac{\sqrt{n}(\bar{X}_n - \bar{p}_n)}{\sqrt{n^{-1}\sum_{i=1}^n p_i(1-p_i)}} = 2X_1 - 1.$$

4. (a) By Chebyshev's THM, $P(|\bar{X}_n - \mu| > \epsilon) \le \frac{Var(\bar{X}_n)}{\epsilon^2} = \frac{\sum_{i=1}^n \sigma_i^2/n^2}{\epsilon^2} \to 0$, with $\sum_{i=1}^n \sigma_i^2 = o(n^2)$.

(b) Use Lindeberg-Feller CLT, with $max_{i\leq n}\sigma_i^2/\sum_{i=1}^n \sigma_i^2 \to 0$, show that Lindeberg condition holds:

$$
\begin{aligned}
& (1/n\bar{\sigma}_n^2)\sum_{i=1}^n E[|X_i-\mu|^2 I(|X_i-\mu|^2 \geq \delta\sum_{i=1}^n \sigma_i^2)] \\
= \;& (1/n\bar{\sigma}_n^2)\sum_{i=1}^n E[\sigma_i^2\epsilon_i^2 I(\sigma_i^2\epsilon_i^2 \geq \delta\sum_{i=1}^n \sigma_i^2))] \\
\leq \;& (1/n\bar{\sigma}_n^2)\sum_{i=1}^n \sigma_i^2 E[\epsilon_i^2 I(\epsilon_i^2 \geq \delta(\sum_{i=1}^n \sigma_i^2/max_{i\leq n}\sigma_i^2))] \\
= \;& E[\epsilon_1^2 I(\epsilon_1^2 \geq \frac{\delta}{max_{i\leq n}\sigma_i^2/\sum_{i=1}^n \sigma_i^2})] \to 0
\end{aligned}
$$

Then $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\bar{\sigma}_n} \to_d N(0,1)$. With $\bar{\sigma}_n^2 \to \sigma_0^2$, from Slutsky Theorem, $\sqrt{n}(\bar{X}_n-\mu) \to_d N(0,\sigma_0^2)$.

(c) $max_{i\leq n}\sigma_i^2/\sum_{i=1}^n \sigma_i^2 = 1/\sum_{i=1}^n (i/n)^r \to 0$, but $\bar{\sigma}_n^2 = (1/n)A\sum_{i=1}^n i^r \to \infty$.