STA 532 Homework6

Jae Hyun Lee, jl914

03 March, 2020

HW6 for STA-532

1.

We want $S_n^2 - \sigma^2)/T_n \to^d N(0,1)$. If we can find $E(S_n^2)$ and $V(S_n^2)$, by CLT, $S_n^2 \to^d N(E(S_n^2), V(S_n^2)))$. Before start, we can know that

$$
\begin{aligned}
&E(Y_i^2) = \mu^2 + \sigma^2, E(Y_i^4) = r \to V(Y_i^2) = r - (\mu^2 + \sigma^2)^2 \\
&\bar{Y} = \sum Y_i/n \to E(\bar{Y}) = \mu, V(\bar{Y}) = \sigma^2/n, \\
&and\ by\ CLT, asymptonically\ \sqrt{n}(\bar{Y} - \mu) \sim N(0,1) \\
&\to n(\bar{Y} - \mu)^2/\sigma^2 \sim \chi^2(1), then\ E(n(\bar{Y} - \mu)^2/\sigma^2) = 1, V(n(\bar{Y} - \mu)^2/\sigma^2) = 2
\end{aligned}
$$

At first,

$$
\begin{aligned}
S_0^2 &= \sum (Y_i - \mu)/n \\
E(S_0^2) &= \frac{1}{n} \sum E[(Y_i - \mu)^2] \\
&= \frac{1}{n} \sum V(Y_i) = \sigma^2 \\
V(S_0^2) &= \frac{1}{n^2} \sum V((Y_i - \mu)^2) \\
&= \frac{1}{2} \sum V(Y_i^2 - 2\mu Y_i + \mu^2) \\
&= \frac{1}{n^2} \sum V(Y_i^2 - 2Y_i\mu) \\
&= \frac{1}{n^2} [\sum V(Y_i^2) + 4\mu \sum V(Y_i)] \\
&= \frac{1}{n}[r - \sigma^4 - 2\mu^2\sigma^2 - \mu^4 + 4\mu^2\sigma^2] \\
&= \frac{1}{n}[r - (\sigma^2 - \mu^2)^2]
\end{aligned}
$$

At second,

$$
\begin{aligned}
S_1^2 &= \sum (Y_i - \bar{Y})^2/n = \sum (Y_i - \mu + \mu - \bar{Y})^2/n \\
&= \frac{1}{n}[\sum (Y_i - \mu)^2 - 2(\bar{Y} - \mu)\sum (Y_i - \mu) + n(\bar{Y} - \mu)^2] \\
&= \sum (Y_i - \mu)^2/n - (\bar{Y} - \mu)^2 = S_0^2 - (\bar{Y} - \mu)^2 \\
E(S_1^2) &= E(S_0^2) - E[(\bar{Y} - \mu)^2] = \sigma^2 - V(\bar{Y}) = \sigma^2 \times \frac{n-1}{n} \\
V(S_1^2) &= V(S_0^2) + V[(\bar{Y} - \mu)^2] \\
V(n(\bar{Y} - \mu)^2/\sigma^2) &= 2 \to V[(\bar{Y} - \mu)^2] = 2\sigma^4/n^2 \\
\to V(S_1^2) &= \frac{1}{n}[r - (\sigma^2 - \mu^2)^2] + 2\sigma^4/n^2
\end{aligned}
$$

Finally,

$$
\begin{aligned}
S^2 = \frac{n}{n-1}S_1^2 \to E(S^2) &= \frac{n}{n-1}E(S_1^2) = \sigma^2 \\
V(S^2) &= \frac{n^2}{(n-1)^2}V(S_1^2) = \frac{nr - n(\sigma^2 - \mu^2)^2 + 2\sigma^4}{(n-1)^2} \\
S^2 \to^d N(E(S^2), V(S^2))
\end{aligned}
$$

Thus S^2 is CAN estimator of σ^2

2.

(a)

$$
\begin{aligned}
\hat{\theta} &= \frac{1}{n}\sum W_i = \frac{1}{n}\sum Y_i/x_i \\
E(\hat{\theta}) &= \frac{1}{n}\sum E(Y_i)/x_i = \frac{1}{n}\sum \theta = \theta \to bias = 0 \\
V(\hat{\theta}) &= V(\frac{1}{n}\sum Y_i/x_i) = \frac{1}{n^2}\sum V(Y_i/x_i) = \frac{1}{n^2}\sum V(Y_i)/x_i^2 = \frac{\sigma^2}{n^2}\sum 1/x_i^2
\end{aligned}
$$

(b)

$$
\begin{aligned}
L(\tilde{Y}, \theta) &= \prod \frac{1}{\sqrt{2\pi\sigma^2}} exp\{-\frac{1}{2\sigma^2}(Y_i - \theta x_i)^2\} \\
l(\tilde{Y}, \theta) &= \sum (-\frac{1}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(Y_i - \theta x_i)^2) + C \\
&= -\frac{n}{2}log(\sigma^2) - \sum \frac{1}{2\sigma^2}(Y_i - \theta x_i)^2 + C \\
\frac{d}{d\theta}l(\tilde{Y}, \theta) &= 2\sum x_i(Y_i - \theta x_i) = 0 \\
\to \hat{\theta}_{MLE} &= \frac{\sum x_i Y_i}{\sum x_i^2} \\
E(\hat{\theta}_{MLE}) &= \frac{\sum x_i E(Y_i)}{\sum x_i^2} = \frac{\theta \sum x_i^2}{\sum x_i^2} = \theta \\
V(\hat{\theta}_{MLE}) &= V(\frac{\sum x_i Y_i}{\sum x_i^2}) = \frac{1}{(\sum x_i)^2}V(\sum x_i Y_i) = \frac{1}{(\sum x_i)^2}\sum x_i^2 V(Y_i) = \frac{1}{\sum x_i^2}\sigma^2
\end{aligned}
$$

(c)

Since both estimators are unbiased, MSe is variance of each estimator. Thus, $\frac{1}{n^2}\sum 1/x_i^2, \frac{1}{\sum x_i^2}$. If all $|x_i| = 1$, then $\frac{1}{n^2}\sum 1/x_i^2 = \frac{1}{\sum x_i^2} = 1/n$. Thus if nay $|x_i| \neq 1$ then one estimator is better than the other.

(d)

Both estimator's variance depends on $x_i's$. If x_i has small absolute value $\rightarrow \sum 1/x_i^2, \frac{1}{\sum x_i^2}$, both have large values. We want our estimator has small variance. Thus I would recommend μ_x that has large absolute value and σ_x^2 that has small value so that $x_i's$ have stable large absolute value.

3.

(a)

i.

$$\begin{aligned}
\theta = log\frac{p}{1-p} &\rightarrow p = \frac{e^\theta}{1+e^\theta} \\
P_\theta(y) &= \begin{bmatrix} n \\ y \end{bmatrix} (\frac{p}{1-p})^y(1-p)^n \\
&= \begin{bmatrix} n \\ y \end{bmatrix} exp\{\theta y + nlog(1-p)\} \\
&= \begin{bmatrix} n \\ y \end{bmatrix} exp\{\theta y - nlog(1+e^\theta)\} \\
\rightarrow \theta = log\frac{p}{1-p}, c(y) &= \begin{bmatrix} n \\ y \end{bmatrix}, t(y) = y, A(\theta) = nlog(1+e^\theta)
\end{aligned}$$

ii.

$$\begin{aligned}
P_\theta(y) = e^{-\mu}\mu^y/y! &= \frac{1}{y!}exp\{ylog\mu - \mu\} \\
&= \frac{1}{y!}exp\{y\theta - e^\theta\} \\
\rightarrow \theta = log\mu, c(y) &= \frac{1}{y!}, t(y) = y, A(\theta) = e^\theta
\end{aligned}$$

iii.

$$\begin{aligned}
P_\theta(y) &= \frac{1}{\sqrt{2\pi\sigma^2}}exp\{-\frac{1}{2\sigma^2}(y-\mu)^2\} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}exp\{-\frac{1}{2\sigma^2}(y^2 - 2y\mu + \mu^2)\} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}exp\{-\frac{y^2}{2\sigma^2} + \frac{\mu}{\sigma^2}y - \frac{\mu^2}{2\sigma^2}\} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}exp\{\theta^T t(y) - A(\theta)\} \\
\rightarrow \theta = (-\frac{1}{2\sigma^2}, \mu/\sigma^2), t(y) &= (y^2, y), A(\theta) = -\frac{\mu^2}{2\sigma^2}, c(y) = \frac{1}{\sqrt{2\pi}}
\end{aligned}$$

iV.

$$
\begin{aligned}
P_\theta(y) &= \frac{1}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} \\
&= \frac{1}{\Gamma(a)\Gamma(b)} exp\{(a-1)log(y) + (b-1)log(1-y)\} \\
&= \frac{1}{\Gamma(a)\Gamma(b)} exp\{alog(y) + blog(1-y)\} \times \frac{1}{y(1-y)} \\
&= exp\{alog(y) + blog(1-y) - log\Gamma(a)\Gamma(b)\} \times \frac{1}{y(1-y)} \\
\rightarrow \theta &= (a,b), t(y) = (logy, log(1-y)), c(y) = \frac{1}{y(1-y)}, A(\theta) = log\Gamma(a)\Gamma(b)
\end{aligned}
$$

(b)

$$
\begin{aligned}
logP_\theta(y) &= logC(y) + \theta^T t(y) - A(\theta) \\
l(\theta, y) &= \sum logP_\theta(y_i) = \sum logC(y_i) + \theta^T \sum t(y_i) - nA(\theta) \\
&= logC(\tilde{y}) + \theta^T t(\tilde{y}) - nA(\theta)
\end{aligned}
$$

(c)

$$
\begin{aligned}
\frac{d}{d\theta} P_\theta(y) &= (t(y) - A'(\theta))c(y)exp\{\theta^T t(y) - A(\theta)\} \\
\int \frac{d}{d\theta} P_\theta(y)dy &= \int t(y)c(y)exp\{\theta^T t(y) - A(\theta)\}dy \times A'(\theta) \int c(y)exp\{\theta^T t(y) - A(\theta)\}dy \\
&= E(t(y)) - A'(\theta) \\
l(\theta, y) &= logc(\tilde{y}) + \theta^T t(\tilde{y}) - nA(\theta) \\
\rightarrow \frac{d}{d\theta} \frac{l(\theta, y)}{n} &\propto \frac{1}{n} t(\tilde{y}) - A'(\theta)
\end{aligned}
$$

Now, we can find that $l'(\theta, y)/n \propto \frac{1}{n} t(\tilde{y}) - A'(\theta) \approx \int P'_\theta(y)dy = E(t(y)) - A'(\theta)$. We know that MLE $\hat{\theta}$ that makes $t(\tilde{y}) - A'(\hat{\theta}) = 0$. Then MLE $\hat{\theta}$ also makes $\int P'_\theta(y)dy = E(t(y)) - A'(\hat{\theta}) = 0$. Consequently, if we know expectation of $t(\tilde{y})$ then we can easily find MSE $\hat{\theta}$.

4.

(a)

$$
\begin{aligned}
P(\tilde{y} \mid \theta) &= \theta^{\sum y_i}(1-\theta)^{n-\sum y_i} \\
&= (\frac{\theta}{1-\theta})^{\sum y_i}(1-\theta)^n \\
l(y,\theta) &= \sum y_i log\frac{\theta}{1-\theta} + nlog(1-\theta) \\
&= \sum y_i log\theta + (n - \sum y_i) log(1-\theta) \\
\frac{d}{d\theta} l(y,\theta) &= \sum y_i/\theta - (n - \sum y_i)/(1-\theta) \\
&\to \frac{1}{\hat{\theta}(1-\hat{\theta})}\sum y_i = \frac{n}{1-\hat{\theta}} \to \hat{\theta} = \sum y_i/n
\end{aligned}
$$

(b)

$$
\begin{aligned}
P(\tilde{y} \mid \psi) =& (\frac{e^\psi}{1+e^\psi})^{\sum y_i}(1+e^\psi)^{\sum y_i - n} \\
l(\tilde{y},\psi) &= \sum y_i log(\frac{e^\psi}{1+e^\psi}) + (n - \sum y_i) log\frac{1}{1+e^\psi} \\
&= \sum y_i loge^\psi - \sum y_i log(1+e^\psi) + \sum y_i log(1+e^\psi) - nlog(1+e^\psi) \\
&= \psi \sum y_i - nlog(1+e^\psi) \\
\frac{d}{d\psi} l(\tilde{y},\psi) &= \sum y_i - e^\psi \frac{n}{1+e^\psi} \\
\to e^{\hat{\psi}}/(1+e^{\hat{\psi}}) &= \sum y_i/n = \hat{\theta} \\
\to e^{\hat{\psi}} &= \frac{\sum y_i}{n}(1+e^{\hat{\psi}}) \\
\to e^{\hat{\psi}}(1 - \sum y_i/n) &= \sum y_i/n \\
\to e^{\hat{\psi}} &= \frac{\sum y_i/n}{1-\sum y_i/n} \to \hat{\psi} = log\frac{\sum y_i/n}{1-\sum y_i/n} = log\frac{\hat{\theta}}{1-\hat{\theta}}
\end{aligned}
$$

5.

$$
\begin{aligned}
&P_1 = \{f_\theta(y) : \theta \in \Theta\} \\
&P_2 = \{g_\psi(y) : \psi \in \Psi\} \\
&and\ we\ assume\ h(\theta)\ exist\ which\ is\ 1-1\ function\ mapping\ \Theta \to \Psi
\end{aligned}
$$

Let $\hat{\theta}$ be MLE based on P_1, then $\hat{\psi} = h(\hat{\theta})$ based on P_2.

Proof are as below:

Let likelihood function based on $P_1 = l_1(\theta, y), P_2 = l_2(\psi, y) = l_2(h(\theta), y)$.

$$
\begin{aligned}
&\frac{d}{d\theta} l_1(\theta,y) = l_1'(\theta,y) \to l_1'(\hat{\theta},y) = 0\ at\ MLE\ \hat{\theta} \\
&\frac{d}{d\theta} l_2(h(\theta),y) = l_2'(h(\theta),y)h'(\theta)
\end{aligned}
$$

By change of variable,

$$l_1(\theta, y) = l_2(h(\theta), y) \mid h'(\theta) \mid$$

$$l_1'(\theta, y) = \frac{d}{d\theta} l_2'(h(\theta), y) \mid h'(\theta) \mid$$

$$l_1'(\hat{\theta}, y) = \frac{d}{d\theta} l_2'(h(\hat{\theta}), y) \mid h'(\hat{\theta}) \mid = 0$$

$$and$$

$$\frac{d}{d\psi} l_2(\hat{\psi}, y) = 0 \rightarrow \hat{\psi} = h(\hat{\theta})$$