STAT 245 HW2 Solution

Yuguan Wang

January 2023

Q1

Using Delta method, we have

$$\sqrt{n}(f(\bar{X}) - f(p)) \xrightarrow{d} N(0, [f'(p)]^2 p(1-p)),$$

which suggests setting $f'(p)\sqrt{p(1-p)} = 1$. Thus, $f(p) = 2\arcsin(\sqrt{p})$, and $f(\bar{X}) \overset{d}{\approx} N(g(p), 1/n)$.

So the $1-\alpha$ CI can be constructed as $g^{-1}\left(g(\bar{X}) \pm z_{\alpha/2}/\sqrt{n}\right)$.

Q2

Since $\bar{X} \overset{d}{\approx} N(\mu, 1/n)$, the 95% CI for μ is $\bar{X} \pm z_{0.025}/\sqrt{n}$. So we want $2z_{0.025}/\sqrt{n} \leq 0.5 \Rightarrow n \geq 62$.

Q3

(a) $E[Z^2] = 1$ and $E[Z^4] = 3$.

(b) Let $Z_1, \cdots, Z_n \overset{i.i.d.}{\sim} N(0,1)$, then $Y \overset{d}{=} \sum_{i=1}^n Z_i^2$. Therefore $E[Y] = \sum_{i=1}^n E[Z_i^2] = n$ and $\mathrm{Var(Y)} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{Var(Z_i^2)} = \mathrm{n(E[Z_1^4] - (E[Z_1^2])^2) = 2n}$.

Q4

$$\begin{aligned}
E(\hat{\theta} - \theta)^2 &= E(\hat{\theta} - E[\hat{\theta}] + E[\hat{\theta}] - \theta)^2 \\
&= E(\theta - E[\hat{\theta}])^2 + 2E(\hat{\theta} - E[\hat{\theta}]) \cdot (E[\hat{\theta}] - \theta) + E(E[\hat{\theta}] - \theta)^2 \\
&= \mathrm{Var}(\hat{\theta}) + (\mathrm{E}[\hat{\theta}] - \theta)^2.
\end{aligned}$$

Q5

(a) The likelihood function is

$$L(\mu, \sigma^2|X) = \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right).$$

Then

$$l(\mu, \sigma^2) = \sum_{i=1}^{n} -\frac{1}{2}\log(\sigma^2) - \frac{(X_i - \mu)^2}{2\sigma^2}.$$

Setting $\frac{\partial l}{\partial \mu} = 0$ gives $\hat{\mu} = \bar{X}$.

Setting $\frac{\partial l}{\partial \sigma^2} = 0$ gives

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$$

Therefore

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

and $c = 1/n$.

(b)

$$\begin{aligned}
E[\hat{\sigma}_c^2] &= E\left(c\sum_{i=1}^{n}(X_i - \bar{X})^2\right) \\
&= c\sigma^2 E\left(\sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2\right) \\
&= c\sigma^2(n-1). \\
\mathrm{Var}(\hat{\sigma}_{\mathrm{c}}^2) &= c^2\sigma^4 \mathrm{Var}\left(\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{X_i} - \bar{\mathrm{X}})^2/\sigma^2\right) \\
&= 2c^2\sigma^4(n-1).
\end{aligned}$$

When $c = \frac{1}{n-1}$, $\hat{\sigma}_c^2$ is unbiased.

(c) Using the formula from Q3, $\mathrm{MSE}(\sigma^2) = (\mathrm{c}\sigma^2(\mathrm{n}-1)\sigma^2)^2 + 2\mathrm{c}^2\sigma^4(\mathrm{n}-1)$. Take the derivative over σ^2 and set it to zero, we can get

$$2(n-1)c\sigma^4 - 2\sigma^4(n-1) + 4(n-1)c\sigma^4 = 0$$

which implies $c = \frac{1}{n+1}$.

Q6

(a) From the question, the statistical model is $X_1, \cdots, X_{180} \overset{i.i.d.}{\sim} \text{Poisson}(10\lambda)$ and $Y_1, \cdots, Y_{20} \overset{i.i.d.}{\sim} \text{Poisson}(20\lambda)$, so the likelihood function is

$$L(\lambda) = \prod_{i=1}^{180} \frac{e^{-10\lambda}(10\lambda)^{X_i}}{X_i!} \prod_{i=1}^{20} \frac{e^{-20\lambda}(20\lambda)^{Y_i}}{Y_i!}.$$

Then

$$\frac{\partial}{\partial\lambda}(\log L(\lambda)) = -1800 + \frac{1}{\lambda}\sum_{i=1}^{180} X_i - 400 + \frac{1}{\lambda}\sum_{i=1}^{20} Y_i.$$

So

$$\hat{\lambda} = \frac{\sum_{i=1}^{180} X_i + \sum_{i=1}^{20} Y_i}{2200} \approx 0.1577.$$

(b) By addivity of Poisson,

$$\sum_{i=1}^{180} X_i + \sum_{i=1}^{20} Y_i \sim \text{Poisson}(2200\lambda) \sim \sum_{\text{i}=1}^{2200} \text{Z}_\text{i},$$

where $Z_i \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$. By CLT, we have

$$\sqrt{2200}(\bar{Z} - \lambda) \xrightarrow{d} N(0, \lambda),$$

thus $\hat{\lambda} \overset{d}{\approx} N(\lambda, \lambda/2200)$.