20200130 Yujun Kim IE541 Homework 4

#1 $X_1, \cdots, X_n \sim \text{Uniform}[0, \theta]$

$Y = \max\{X_1, \cdots, X_n\}$.

By example 9.12 of textbook, $\hat{\theta}_n = Y$.

To show $\hat{\theta}_n$ is consistent, (i.e. $\hat{\theta}_n \xrightarrow{P} \theta$),

$$P(|\hat{\theta}_n - \theta| < \varepsilon) = P(\theta - \varepsilon < \hat{\theta}_n = Y < \theta + \varepsilon)$$

$(\because Y = \max\{X_i\} \leq \theta)$

$$= P(Y < \theta + \varepsilon) - P(Y < \theta - \varepsilon) = 1 - P(\theta - \varepsilon)$$

$$P(Y < c) = P(X_1 < c, \cdots, X_n < c) = P(X_1 < c) \cdots P(X_n < c).$$

$$\leq \left(\frac{c}{\theta}\right)^n \quad \text{for } c > 0.$$

$$\left(\because P(X_i < c) = \begin{cases} \frac{c}{\theta} & \text{if } c \in [0, \theta] \\ 0 & \text{otherwise} \end{cases} \leq \frac{c}{\theta} \quad \text{for } c > 0\right)$$

Take $\varepsilon < |\theta|$. Then, $\theta - \varepsilon > 0$.

$$P(|\hat{\theta}_n - \theta| < \varepsilon) = 1 - \left(\frac{\theta - \varepsilon}{\theta}\right)^n \longrightarrow 1 \quad \text{as } n \to \infty.$$

Thus, $\hat{\theta}_n \xrightarrow{P} \theta$.

#2 $X_1, \dots, X_n \sim \text{Poisson}(\lambda)$.

① Method of moments estimator

$$\alpha_1(\hat{\lambda}_n) = \mathbb{E}_{\hat{\lambda}_n}[X] \quad (\text{for } X \sim \text{Poisson}(\hat{\lambda}_n))$$
$$= \hat{\lambda}_n$$
$$\hat{\alpha}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

By $\alpha_1(\hat{\lambda}_n) = \hat{\alpha}_1$, we have $\hat{\lambda}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

② Maximum Likelihood Estimator.

$$f(x;\lambda) = e^{-\lambda}\frac{\lambda^x}{x!} . \quad \mathcal{L}_n(\lambda) = \prod f(X_i;\lambda) = e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^{n} X_i}}{\prod_{i=1}^{n}(X_i!)}$$

$$\ell_n(\lambda) = -n\lambda + \left(\sum_{i=1}^{n} X_i\right)\log\lambda - \underbrace{\sum \log(X_i!)}_{\text{we can ignore constant.}}$$

$$\frac{\partial \ell_n}{\partial \lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n} X_i = 0 . \Rightarrow \hat{\lambda}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

③ Fisher Information

$$s(X;\lambda) = \frac{\partial}{\partial\lambda}\log f(X;\lambda) = \frac{\partial}{\partial\lambda}\log\left(e^{-\lambda}\frac{\lambda^X}{X!}\right) = -1 + \frac{X}{\lambda}$$

$$\frac{\partial}{\partial\lambda}s(X;\lambda) = -\frac{X}{\lambda^2}$$

$$I(\theta) = -\mathbb{E}_\lambda\left[\frac{-X}{\lambda^2}\right] = \frac{1}{\lambda} . \qquad I_n(\theta) = \frac{n}{\lambda} .$$

#3 (a) $\beta(\theta) = \mathbb{P}(Y > c) = 1 - \mathbb{P}(Y \le c) = 1 - \mathbb{P}(X_1 < c, \cdots, X_n < c)$

$$= 1 - \prod_{i=1}^{n} \mathbb{P}(X_i \le c) = \begin{cases} 1 - \left(\frac{c}{\theta}\right)^n & \text{if } c \in (0, \theta) \\ 1 & \text{if } c \le 0 \\ 0 & \text{if } c \ge \theta \end{cases}$$

(b) $\alpha = \sup_{\theta \in \Theta_0} \beta(\theta) = \beta(\frac{1}{2}) = 0.05$

$\to 1 - (2c)^n = 0.05$ (given $c \in (0, \frac{1}{2})$)

$\to 0.95 = (2c)^n$

$\to \sqrt[n]{0.95} = 2c$

$\to c = \frac{1}{2}\sqrt[n]{0.95} \in (0, \frac{1}{2})$. Hence, such c works.

(c). By theorem 10.12,

p-value $= \mathbb{P}_\theta(Y \ge 0.48)$, with $\theta = \frac{1}{2}$, $n = 20$.

$$= 1 - \left(\frac{0.48}{0.5}\right)^{20} = 1 - 0.96^{20} \approx 0.557$$

(or by def, $1 - (2c)^n = \alpha \to c = \frac{1}{2}\sqrt[n]{1-\alpha} < Y = 0.48$

$\to \alpha > 1 - 0.96^{20}$).

(d) By theorem 10.12

p-value $= \mathbb{P}_\theta(Y \ge 0.52) = 0$. ($\because X_i \sim \text{Uniform}(0, \frac{1}{2})$).

#4 By $X_i \sim \text{Binomial}(p)$, $i=1,\cdots,n=1919$,

$H_0: \theta = \theta_0 := \frac{1}{2}$, $H_1: \theta > \frac{1}{2}$. θ: probability of death week after.

$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is asymptotically normal.

Reject H_0 if $Y \geq c$

$$\hat{\theta} = \frac{997}{1919}, \quad \hat{se} = \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}} = 0.0114$$

$$w = \frac{\hat{\theta} - \theta_0}{\hat{se}} = 1.7134$$

$$\text{p-value} = \mathbb{P}(|W| > |w|)$$

$$\approx \mathbb{P}(|Z| > |w|) = 2\Phi(-|w|)$$

$$= 2\Phi(-1.7134) = 0.0866$$

95% confidence interval is

$(\hat{\theta} - z_{\alpha/2}\hat{se}, \hat{\theta} + z_{\alpha/2}\hat{se})$ with $\alpha = 0.05$

i.e. $(0.497, 0.542)$

#5 (a) p: proportion of three letter words found by Twain's.

q: " Snodgrass's.

$\delta = p - q$.

$H_0 : \delta = 0$, $H_1 : \delta \neq 0$.

$n = 8, m = 10$

Let $X_1, \cdots, X_n$ be the data for Twain

$Y_1, \cdots, Y_m$ be the data for Snodgrass

$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = 0.2319$, $\hat{q} = \frac{1}{m}\sum_{i=1}^{m} Y_i = 0.2097$ (Plug-in est.)

$\hat{\delta} = \hat{p} - \hat{q} = 0.0222$ $\quad \hat{se} = \sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}} = 0.006$, S_1^2, S_2^2 : sample var.

$$w = \frac{\hat{\delta} - 0}{\hat{se}} = 3.704$$

p-value $\approx 2\Phi(-|w|) = 2.17 \times 10^{-4}$

95% confidence interval : Theorem 10.10.

$(\hat{\delta} - 1.96\hat{se}, \hat{\delta} + 1.96\hat{se}) = (0.0104, 0.0339)$

(b) $18!$ cases are too many to calculate directly.

Use $B = 10^5$ to approximat p-value for permutation test.

Below gives part of MATLAB code for the calculation.

```
% (b)
Z = [X Y];
l = 18
B = 100000
count = 0
for i = 1:B
    p = randperm(l);
    for j = 1:n
        X(j) = Z(p(j));
    end
    for j = n+1:l
        Y(j) = Z(p(j));
    end
    p = mean(X);
    q = mean(Y);
    d = p-q;
    se = sqrt(var(X)/n + var(Y)/m);
    w = (d-0)/se;
    if (w > w_0)
        count = count + 1;
    end
end
count
count/B
```

This gives count $= 50$,

count$/B = 5\times10^{-4}$.

Hence, approximated p-value $= 5\times10^{-4}$.

#6 $R = \{x^n \mid \frac{\hat{\theta}-\theta_0}{\hat{se}} > z_{\alpha/2}\}$.

$$\beta(\theta_1) = \mathbb{P}_{\theta_1}(|X| \in R)$$
$$= \mathbb{P}_{\theta_1}\left(\left|\frac{\hat{\theta}-\theta_0}{\hat{se}}\right| > z_{\alpha/2}\right)$$
$$\geq \mathbb{P}_{\theta_1}\left(\frac{\hat{\theta}-\theta_0}{\hat{se}} > z_{\alpha/2}\right)$$
$$= \mathbb{P}_{\theta_1}\left(\frac{\hat{\theta}-\theta_1}{\hat{se}} > z_{\alpha/2} - \frac{\theta_1-\theta_0}{\hat{se}}\right).$$

$$\longrightarrow \mathbb{P}(Z \in \mathbb{R}) = 1 \quad \cdots (*)$$

$$\begin{cases} \hat{se} \to 0 \text{ as } n \to \infty \\ \frac{\hat{\theta}-\theta_1}{\hat{se}} \to Z \text{ as } n \to \infty \end{cases}$$ (MLE is asymptotically normal).

Thus, (*) holds. To be specific, theorem 1.8 (continuity of probability) in the textbook is used.

Thus, $\beta(\theta_1) \longrightarrow 1$ as $n \to \infty$.

#7 ~~Likelihood Ratio Test~~

$$\mathcal{L}_n(\mu) = \prod_{i=1}^{n} f(X_i;\mu) = \prod_{i=1}^{n}\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(X_i-\mu)^2}\right)$$

MLE $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$ by Example 9.11

$$\lambda = 2\log\left(\frac{\mathcal{L}(\hat{\mu})}{\mathcal{L}(\hat{\mu}_0)}\right) = 2\log\left(\frac{\mathcal{L}(\hat{\mu})}{\mathcal{L}(\mu_0)}\right)$$

$$= \frac{1}{\sigma^2}\left(\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\hat{\mu})^2\right)$$

$\theta = (\mu, \sigma)$,

$\Theta_0 = \{\theta \mid \mu = \mu_0\}$. Then, $\lambda(x^n) \rightsquigarrow \chi^2_{1,\alpha}$

p-value $= \mathbb{P}(\chi_1^2 > \lambda(x^n)) = \mathbb{P}(|Z| > \sqrt{\lambda(x^n)})$

<u>Wald Test</u> $W = \frac{\hat{\mu}-\mu_0}{\hat{se}}$. p-value $\approx 2\Phi(-|w|)$.

$= \mathbb{P}(|Z| > |w|)$

p-value of likelihood ratio test depends on σ directly, while wald test has info about σ on $\hat{se}$.