Introduction to Statistics Theory (Fall 2018)

Final Exam

Name: ________________

Results you may use directly:

- $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(x_i - \bar{x})x_i$
- $S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})y_i$
- $var(X+Y) = var(X) + 2cov(X,Y) + var(Y)$
- In simple linear model,

$$var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}, \quad var(\hat{\beta}_0) = \sigma^2(\frac{\bar{x}^2}{S_{xx}} + \frac{1}{n}), \quad cov(\hat{\beta}_1, \hat{\beta}_0) = -\frac{\bar{x}}{S_{xx}}\sigma^2$$

1. Consider a simple linear model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$. The least square estimators are

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}, \qquad \hat{\beta}_0 = \bar{Y} - \bar{x}\hat{\beta}_1.$$

Show the followings

(a) $\mathbb{E}\hat{\beta}_1 = \beta_1$ (10pt).

(b) $\mathbb{E}\hat{\beta}_0 = \beta_0$ (10pt).

(a) $\mathbb{E}\hat{\beta}_1 = \mathbb{E}\frac{\sum_i (x_i - \bar{x})(Y_i - \bar{Y})}{S_{xx}}$

$= \sum_i \frac{(x_i - \bar{x})}{S_{xx}} \cdot \mathbb{E} Y_i$

$= \sum_i \frac{(x_i - \bar{x})}{S_{xx}} (\beta_0 + \beta_1 x_i)$

$= \sum_i \frac{(x_i - \bar{x})}{S_{xx}} (\beta_1 x_i \cancel{+} \beta_1 \bar{x})$

$= \frac{\sum_i (x_i - \bar{x})^2}{S_{xx}} \beta_1 = \beta_1$

(b). $\mathbb{E}\hat{\beta}_0 = \mathbb{E}(\bar{Y} - \bar{x} \cdot \hat{\beta}_1)$

$= \mathbb{E}\frac{\sum Y_i}{n} - \bar{x} \cdot \beta_1$

$= \frac{\sum (x_i \beta_1 + \beta_0)}{n} - \bar{x}\beta_1$

$= \bar{x}\beta_1 + \beta_0 - \bar{x}\beta_1$

$= \beta_0$

2. Based on the conditions in problem 1, consider a summation $\hat{W} = \hat{\beta}_0 + \hat{\beta}_1$. Assuming that we know $\sigma^2 = 1$,

(a) Find the mean of $\hat{W}$ (5pt).

(b) Find the variance of $\hat{W}$ (5pt).

(c) Is $\hat{W}$ normally distributed? (2pt) Why or why not? (3pt)

(d) Construct a two-sided $(1-\alpha)$ confidence interval for $\beta_0 + \beta_1$ (5pt).

(a) $\mathbb{E}\hat{W} = \mathbb{E}\hat{\beta}_0 + \mathbb{E}\hat{\beta}_1 = \beta_0 + \beta_1$

(b)
$$\operatorname{var}(\hat{W}) = \operatorname{var}\hat{\beta}_0 + \operatorname{var}\hat{\beta}_1 + 2\operatorname{cov}(\hat{\beta}_0, \hat{\beta}_1)$$
$$= \left(\frac{\bar{x}^2}{S_{xx}} + \frac{1}{n}\right) + \frac{1}{S_{xx}} + 2\left(-\frac{\bar{x}}{S_{xx}}\right)$$
$$= \frac{\bar{x}^2 + 1 - 2\bar{x}}{S_{xx}} + \frac{1}{n} = \frac{(\bar{x}-1)^2}{S_{xx}} + \frac{1}{n}$$

(c) Yes, because ~~both~~ $\hat{\beta}_0$ and $\hat{\beta}_1$ are jointly normally distribution, $\hat{W} = \hat{\beta}_0 + \hat{\beta}_1$ is a linear combination. hence follows a normal distribution

(d). Due to Normality and known $\sigma^2 = 1$. $(1-\alpha)$ CI is:
$$\left(\hat{W} - Z_{\alpha/2} \cdot \lambda, \ \hat{W} + Z_{\alpha/2}\lambda\right)$$

where
$$\lambda = \sqrt{\frac{(\bar{x}-1)^2}{S_{xx}} + \frac{1}{n}}.$$

3. Suppose that Y is a random sample of size 1 from a population with density function

$$f(y \mid \theta) = \theta y^{\theta - 1}, \qquad 0 \le y \le 1$$

where $\theta > 0$.

(a) Find the most powerful test at significance α for $H_0 : \theta = 1$ vs $H_a : \theta = b$, where $b > 1$ (15pt).

(b) Is the derived test *uniformly* most powerful? (5 pt)

(a) $L(\theta; y) = \theta \cdot y^{\theta-1}$.

By Neyman-Pearson

$$\frac{L(\theta_0; y)}{L(b; y)} = \frac{1}{b} \cdot y^{1-1} / y^{b-1}$$

$$= \frac{1}{b} \cdot y^{1-b} \le k$$

since $1-b<0 \Rightarrow y \ge k^*$

we need $P(Y \ge k^*) = \alpha$.

$$\int_{k^*}^{1} \theta \cdot y^{\theta-1} \cdot dy = y^\theta \Big|_{k^*}^{1} = 1 - (k^*)^\theta = \alpha.$$

under H_0, $\theta = 1 \Rightarrow k^* = 1 - \alpha$.

$\Rightarrow$ The rejection region is $\{ y \ge 1 - \alpha \}$

(b) since the rejection region does not vary with b, therefore the derived test is the uniformly most powerful test.

4. Assume $Y_1, \dots Y_n \overset{iid}{\sim} \mathrm{N}(\mu, \sigma^2)$, $\bar{Y} = \sum_{i=1}^n Y_i/n$.

(a) Show that $\bar{Y}$ is an unbiased estimator for μ (5pt).

(b) Show that $\bar{Y}$ is a *minimum sufficient* statistic for μ (10pt).

(c) Show that $\bar{Y}$ is a minimum variance unbiased estimator for μ (5pt).

(a) $$\mathbb{E}\,\bar{Y} = \sum \frac{\mathbb{E} Y_i}{n} = \frac{n\mu}{n} = \mu.$$

(b) $$L(y_1 \dots y_n) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left[- \frac{\sum (y_i - \mu)^2}{2\sigma^2} \right]$$

$$= \frac{1}{(2\pi\sigma^2)^n} \cdot \exp\left(- \frac{\sum y_i^2 - 2\mu \sum y_i + n\mu^2}{2\sigma^2} \right)$$

$$= \underbrace{\frac{1}{(2\pi\sigma^2)^n} \exp\left(- \frac{\sum y_i^2}{2\sigma^2} \right)}_{h(y_1 \dots y_n)} \cdot \underbrace{\exp\left(- \frac{-2\mu \sum y_i + n\mu^2}{2\sigma^2} \right)}_{g(\mu, y_1 \dots y_n)}$$

$\Rightarrow$ $\sum Y_i$ is sufficient for μ.

And. $$\frac{L(y_1 \dots y_n)}{L(x_1 \dots x_n)} = \exp\left(- \frac{(\sum y_i^2 - \sum x_i^2)}{2\sigma^2} \right) \cdot \exp\left(- \frac{-2\mu(\sum y_i - \sum x_i)}{2\sigma^2} \right)$$

is free from μ if and only if $\sum Y_i = \sum X_i$

$\Rightarrow$ $\sum Y_i$ is minimum sufficient for μ.

$\Rightarrow$ $\bar{Y} = \frac{\sum Y_i}{n}$ is minimum sufficient for μ.

(c) By Rao-Blackwell Theorm, $\bar{Y}$ is minimum variance unbiased estimator for μ.

5. Assume that $Y_1, \ldots, Y_n$ are iid with density function

$$f(y) = \frac{1}{\theta}, \qquad 0 < y < \theta$$

Consider the estimator $Y_{max} = \max\{Y_1, \ldots, Y_n\}$ for the parameter θ.

(a) Find out $P(Y_{max} \le k)$, where $0 \le k \le \theta$ (5pt).

(b) Setting $k = \theta - \epsilon$, prove or disprove that Y_{max} is a consistent estimator for θ (10pt).

(c) Prove or disprove that $\frac{n}{(n+1)^2} Y_{max}$ is a consistent estimator for θ (5pt).

(a) $P(Y_{max} \le k) = P(Y_1 \le k, \ldots, Y_n \le k)$
$= \left(\frac{k}{\theta}\right)^n$

(b) $P(Y_{max} \le (\theta - \varepsilon)) = \left(\frac{\theta - \varepsilon}{\theta}\right)^n$.

also $P(Y_{max} \geq \theta + \varepsilon) = 0$

$\Rightarrow P(|Y_{max} - \theta| \geq \varepsilon) = \left(\frac{\theta - \varepsilon}{\theta}\right)^n \rightarrow 0$

as $n \to \infty$.

$\Rightarrow$ Y_{max} is a consistent estimator for θ.

(c) By (b). $Y_{max} \xrightarrow{P} \theta$

$\frac{n}{(n+1)^2} \longrightarrow 0$ as $n \to \infty$

$\Rightarrow \frac{n}{(n+1)^2} \cdot Y_{max} \xrightarrow{P} 0$.

but $\theta > 0$, therefore $\frac{n}{(n+1)^2} Y_{max}$ is NOT a consistent estimator for θ.