BASIC PHD WRITTEN EXAMINATION

THEORY, SECTION 1

(9:00 AM–1:00 PM, July 23, 2019)

INSTRUCTIONS:

(a) This is a **CLOSED-BOOK** examination.

(b) The time limit for this examination is four hours.

(c) Answer both questions that follow.

(d) Put the answers to different questions on separate sets of paper.

(e) Put your exam code, **NOT YOUR NAME**, on each page. The same code is used for Section 1 and Section 2 of the PhD Theory Exam. Please keep the code confidential and do not share this information with any students or faculty. Sharing your code with either students or faculty is viewed as a violation of the UNC honor code.

(f) Return the examination with a signed statement of the UNC honor pledge, separately from your answers. The pledge statement is given on the last page of the exam handout.

(g) In the questions to follow, you are required to answer only what is asked, and not to tell all you know about the topics involved.

1. (25 points) Let X denote a random variable from $N(0,1)$, and let Y be an outcome variable. The joint distribution of (X,Y) has a finite second moment and $E[X^2Y^2] < \infty$. Assume that we observe n i.i.d copies of (X,Y), denoted by $(X_1,Y_1),...,(X_n,Y_n)$. The goal is to obtain the best prediction of Y given X for a future subject.

   (a) One simple prediction is to consider a linear function, $\alpha + \beta X$, to minimize the following squared loss:
   $$E\left[\{Y - (\alpha + \beta X)\}^2\right],$$
   where the expectation is with respect to the joint distribution of (Y,X). Show that the optimal solution for (α,β), denoted by (α^*,β^*), is given by
   $$\begin{pmatrix}\alpha^* \\ \beta^*\end{pmatrix} = \begin{pmatrix} E[Y] \\ E[XY]\end{pmatrix}. \tag{1}$$

   (b) From (1), we estimate (α^*,β^*) as
   $$\widehat{\alpha} = n^{-1}\sum_{i=1}^{n} Y_i, \quad \widehat{\beta} = n^{-1}\sum_{i=1}^{n} X_iY_i.$$
   Give the asymptotic distribution of the obtained estimator after a proper normalization.

   Now suppose that we know the distribution of Y given X is from a log-normal family, i.e.,
   $$\log Y = \gamma X + N(0,\sigma^2).$$

   (c) Obtain the maximum likelihood estimators for α^* and β^* given in (1) and derive their asymptotic distribution.

   (d) Calculate the asymptotic relative efficiency between the maximum likelihood estimator for β^* and $\widehat{\beta}$ given in (b).

   (e) If we allow the prediction function to be arbitrary, that is, we aim to find the best function, $g(X)$, to minimize
   $$E\left[\{Y - g(X)\}^2\right],$$
   what is the optimal $g(X)$ in terms of (γ,σ^2)?
   Hint: consider minimization conditional on X.

   Points: (a) 5; (b) 5; (c) 5; (d) 5; (e) 5.

2. (25 points) Let $X_1, \ldots, X_n$ be i.i.d samples from a distribution with density function

$$f(x) = \theta^{-1} e^{(a-x)/\theta} I(x > a), \text{ where } \theta > 0.$$

(a) When a is known, derive the uniformly most powerful test of size α for testing $H_0 : \theta \geq \theta_0$ versus $\theta < \theta_0$, where θ_0 is a known constant.

(b) When a is known, derive the asymptotic distribution of the maximum likelihood estimator of θ.

In the rest questions, we assume $a = \theta$, i.e. the density is $f(x) = \theta^{-1} e^{(\theta-x)/\theta} I(x > \theta)$.

(c) Prove that both $\bar{X}/\theta$ and $X_{(1)}/\theta$ are pivotal quantities, where $\bar{X}$ is the sample mean and $X_{(1)}$ is the smallest order statistic.

(d) Obtain two confidence intervals with confidence coefficient $1 - \alpha$ for θ, based on two pivotal quantities in (c).

(e) When n is sufficiently large, which of the two confidence intervals has shorter length? Justify your answer.

Points: (a) 5; (b) 5; (c) 5; (d) 5; (e) 5.

Statement of the UNC honor pledge:

"In recognition of and in the spirit of the honor code, I certify that I have neither given nor received aid on this examination and that I will report all Honor Code violations observed by me."

(Signed) ______________________________
NAME

(Printed) ______________________________
NAME

2019 S1 Q1

Q1.

$X \sim N(0,1)$. $\mathbb{E}X^2Y^2 < \infty$

(a) $\mathbb{E}[\{Y-(\alpha+\beta X)\}^2] = \mathbb{E}[\{Y-(\alpha^*+\beta^* X) + (\alpha^*-\alpha+(\beta^*-\beta)X)\}^2]$

$$= \underbrace{\mathbb{E}[\{Y-(\alpha^*+\beta^*X)\}^2]}_{\text{fixed}} + 2\mathbb{E}[\{Y-(\alpha^*+\beta^*X)\}\cdot\{(\alpha^*-\alpha)+(\beta^*-\beta)X\}] + \mathbb{E}[\{(\alpha^*-\alpha)+(\beta^*-\beta)X\}^2]$$

$$= C + 2\{\mathbb{E}Y\cdot(\alpha^*-\alpha) + \mathbb{E}YX\cdot(\beta^*-\beta) - (\alpha^*+\beta^*\mathbb{E}X)\cdot(\alpha^*-\alpha)$$
$$- \alpha^*(\beta^*-\beta)\mathbb{E}X - \beta^*(\beta^*-\beta)\mathbb{E}X^2\}$$
$$+ (\alpha^*-\alpha)^2 + 2(\alpha^*-\alpha)(\beta^*-\beta)\mathbb{E}X + (\beta^*-\beta)\mathbb{E}X^2$$

$$= C + 2\{\alpha^*(\alpha^*-\alpha) + \beta^*(\beta^*-\beta) - \alpha^*(\alpha^*-\alpha) - \beta^*(\beta^*-\beta)\cdot 1\}$$
$$+ (\alpha^*-\alpha)^2 + (\beta^*-\beta)^2$$

$= C + (\alpha^*-\alpha)^2 + (\beta^*-\beta)^2$: minimized when $\alpha=\alpha^*$, $\beta=\beta^*$.

(b) Let $Z_i := X_iY_i$. $\mathbb{E}Z_i^2 < \infty$

Let $T_i = \begin{bmatrix} Y_i \\ Z_i \end{bmatrix}$. Then, $\mathbb{E}T_i = \begin{bmatrix} \alpha^* \\ \beta^* \end{bmatrix}$, $Var(T_i) = \begin{bmatrix} Var\,Y_i & Cov(Y_i,Z_i) \\ Cov(Z_i,Y_i) & Var\,Z_i \end{bmatrix}$

$$= \begin{bmatrix} Var\,Y_i & \mathbb{E}X_iY_i^2 - \mathbb{E}Y_i\cdot\mathbb{E}X_iY_i \\ \text{"} & \mathbb{E}X_i^2Y_i^2 - (\mathbb{E}X_iY_i)^2 \end{bmatrix}$$: all components $<\infty$

Note that $Cov(Y_i,Z_i)^2 \le Var(Y_i)Var(Z_i) < \infty$

Thus, by CLT,

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n T_i - \mathbb{E}T_1\right) \to_d N\left(\begin{bmatrix}0\\0\end{bmatrix}, Var(T_1)\right)$$

$$\Rightarrow \sqrt{n}\left(\begin{bmatrix}\hat\alpha\\\hat\beta\end{bmatrix} - \begin{bmatrix}\alpha^*\\\beta^*\end{bmatrix}\right) \to_d N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix} Var\,Y_1 & \mathbb{E}X_1Y_1^2 - \alpha^*\beta^* \\ \mathbb{E}X_1Y_1^2 - \alpha^*\beta^* & \mathbb{E}X_1^2Y_1^2 - \beta^{*2}\end{bmatrix}\right)$$

(c) First,

$\alpha^* = \mathbb{E}Y = \mathbb{E}\exp(\gamma X+\varepsilon)$, where $\varepsilon \sim N(0,\sigma^2)$

$= \mathbb{E}\exp(\gamma X)\cdot\mathbb{E}\exp(\varepsilon)$ $(\because)$ $X \perp\!\!\!\perp \varepsilon$

$= \exp(\frac{1}{2}\gamma^2)\cdot\exp(\frac{1}{2}\sigma^2\cdot 1^2)$ $(\because)$ $N(\mu,\rho^2)$ mgf $= \exp(\mu t + \frac{1}{2}\rho^2t^2)$

$= \exp(\frac{1}{2}\gamma^2 + \frac{1}{2}\sigma^2)$

$\beta^* = \mathbb{E}XY = \mathbb{E}\,X\cdot\exp(\gamma X+\varepsilon)$

$= \mathbb{E}X\cdot\exp(\gamma X)\cdot\mathbb{E}\exp(\varepsilon)$

$= \frac{d}{d\gamma}(\mathbb{E}\cdot\exp(\gamma x))\cdot\mathbb{E}\exp(\varepsilon)$

$= \frac{d}{d\gamma}(\exp(\frac{1}{2}\gamma^2))\cdot\exp(\frac{1}{2}\sigma^2)$

$= \gamma\cdot\exp(\frac{1}{2}\gamma^2 + \frac{1}{2}\sigma^2)$

Note that $(\gamma,\sigma^2) \to (\alpha^*,\beta^*)$ is 1-1.

Thus, if $(\hat{\gamma}, \hat{\sigma}^2)$ is MLE of (γ, σ^2), then $\hat{\alpha}^* = \exp(\frac{1}{2}\hat{\gamma}^2 + \frac{1}{2}\hat{\sigma}^2)$, $\hat{\beta}^* = \hat{\gamma} \cdot \exp(\frac{1}{2}\hat{\gamma}^2 + \frac{1}{2}\hat{\sigma}^2)$ are MLE of α^*, β^*.

Likelihood

$$L(\gamma, \sigma^2) = \prod_{i=1}^n f_{X_i, Y_i}(x_i, y_i ; \gamma, \sigma^2)$$
$$= \prod_{i=1}^n f_{Y_i|X_i}(y_i | x_i ; \gamma, \sigma^2) \cdot f_{X_i}(x_i)$$
$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log y_i - \gamma x_i)^2\right) \cdot \frac{1}{\sigma y_i} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right)$$

Here $f_{Y|X}(y|x ; \gamma, \sigma^2) = \frac{d}{dy} P(Y \le y | X = x) = \frac{d}{dy} P\left(\exp(\gamma X + \varepsilon) \le y \middle| X = x\right)$

$$= \frac{d}{dy} P\left(\exp(\gamma x + \varepsilon) \le y \middle| X = x\right)$$
$$= \frac{d}{dy} P\left(\exp(\gamma x + \varepsilon) \le y\right) \quad (\because)\ X \perp\!\!\!\perp \varepsilon$$
$$= \frac{d}{dy} P\left(\varepsilon \le \log y - \gamma x\right)$$
$$= \frac{d}{dy} P\left(\frac{\varepsilon}{\sigma} \le \frac{\log y - \gamma x}{\sigma}\right)$$
$$= \frac{d}{dy} \Phi\left(\frac{\log y - \gamma x}{\sigma}\right)$$
$$= \phi\left(\frac{\log y - \gamma x}{\sigma}\right) \cdot \frac{1}{\sigma} \cdot \frac{1}{y}$$
$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log y - \gamma x)^2\right) \cdot \frac{1}{\sigma y}$$

$$L(\gamma, \sigma^2) = (2\pi\sigma)^{-n} \cdot \prod_{i=1}^n \exp\left(-\frac{1}{2\sigma^2}(\log y_i)^2 + \frac{\gamma}{\sigma^2} \log y_i \cdot x_i - \frac{\gamma^2}{2\sigma^2} x_i^2\right) \cdot \frac{1}{y_i} \cdot \exp\left(-\frac{1}{2}x_i^2\right)$$
$$= (2\pi\sigma)^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_i (\log y_i)^2 + \frac{\gamma}{\sigma^2} \cdot \sum_i \log y_i \cdot x_i - \frac{\gamma^2}{2\sigma^2} \cdot \sum_i x_i^2\right) \cdot \prod_i \frac{1}{y_i} \exp\left(-\frac{1}{2}x_i^2\right)$$
$$l(\gamma, \sigma^2) = -\frac{n}{2} \log 4\pi^2\sigma^2 - \quad '' \quad + \quad '' \quad - \quad '' \quad) + \sum_i -\log y_i - \frac{1}{2}x_i^2$$

Profile likelihood wrt fixed σ^2.

$$\frac{\partial l}{\partial \gamma} = 0 \iff + \frac{1}{\sigma^2} \sum_i \log y_i \cdot x_i - \frac{2\gamma}{2\sigma^2} \sum_i x_i^2 = 0$$
$$\iff \hat{\gamma} = \left(\sum_i \log y_i \cdot x_i\right) \Big/ \left(\sum_i x_i^2\right)$$

$$g(\sigma^2) = l(\hat{\gamma}, \sigma^2) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2}(\sigma^2)^{-1} \sum_i (\log y_i - \hat{\gamma} x_i)^2 + \text{constant}$$
$$\frac{\partial g}{\partial \sigma^2} = 0 \iff -\frac{n}{2}(\sigma^2)^{-1} + \frac{1}{2}(\sigma^2)^{-2} \cdot \sum_i (\log y_i - \hat{\gamma} x_i)^2 = 0$$
$$\iff \hat{\sigma}^2 = \frac{1}{n} \sum_i (\log y_i - \hat{\gamma} x_i)^2$$

Now, by expressing $\log y_i = \gamma x_i + \epsilon_i$,

$$\hat{\gamma} = \frac{\sum_i \gamma x_i^2 + \epsilon_i x_i}{\sum_i x_i^2} = \gamma + \frac{\sum_j \epsilon_j x_j}{\sum_j x_j^2}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_i \left(\gamma x_i + \epsilon_i - \left(\gamma + \frac{\sum_j \epsilon_j x_j}{\sum_j x_j^2} x_i\right)\right)^2 = \frac{1}{n} \sum_i \left(\epsilon_i - \frac{\sum_j \epsilon_j x_j}{\sum_j x_j^2} x_i\right)^2 = \frac{1}{n}\left\{\sum_i \epsilon_i^2 - \frac{(\sum_j \epsilon_j x_j)^2}{\sum_j x_j^2}\right\}$$

To get their joint dist^n, first assume $(X_1, \cdots, X_n) = (x_1, \cdots, x_n)$ is given.

Then, $\hat{\gamma}|_{X=x} = \gamma + \sum_i \frac{x_i}{\sum_j x_j^2}\cdot\varepsilon_i$

$\hat{\sigma}^2|_{X=x} = \frac{1}{n}\left\{\sum_i \varepsilon_i^2 - \frac{1}{\sum_j x_j^2}\left(\sum_j x_j\varepsilon_j\right)^2\right\}$

Letting $w_i = \frac{x_i}{\sum_j x_j^2}$, where $\sum w_i^2 = \frac{1}{\sum x_j^2}$, $x_i = \frac{w_i}{\sum w_j^2}$

$\hat{\gamma}|_{X=x} = \gamma + \sum_i w_i\varepsilon_i = \gamma + w^T\varepsilon$

$\hat{\sigma}^2|_{X=x} = \frac{1}{n}\left\{\sum_i \varepsilon_i^2 - \frac{1}{\sum_j w_j^2}\left(\sum_j w_j\varepsilon_j\right)^2\right\} = \frac{1}{n}\left\{\varepsilon^T\varepsilon - (w^Tw)^{-1}(w^T\varepsilon)^2\right\}$

$= \frac{1}{n}\left\{\varepsilon^T\varepsilon - \varepsilon^T w(w^Tw)^{-1}w^T\varepsilon\right\}$

$= \frac{1}{n}\varepsilon^T\left(I - w(w^Tw)^{-1}w^T\right)\varepsilon$.

where $w = (w_1,\dots,w_n)^T$, $\varepsilon = (\varepsilon_1,\dots,\varepsilon_n)^T \sim N_n(0, \sigma^2 I_n)$

Then, $\hat{\gamma} \perp\!\!\!\perp \hat{\sigma}^2 | X=x$ because $w^T\varepsilon \perp\!\!\!\perp (I - w(w^Tw)^{-1}w^T)\varepsilon$ from $w^T\cdot Cov(\varepsilon)\cdot(I - w(w^Tw)^{-1}w^T)^T = 0$.

$\bar{X}$ $\quad S = \frac{1}{n-1}\sum(x_i-\bar{x})^2$

$\cancel{\perp\!\!\!\perp}$

Furthermore,

$\hat{\gamma}|_{X=x} \sim N(\gamma, \sigma^2 w^Tw)$ $\leftarrow$ $\frac{1}{\sum x_i^2}$ $\quad (\hat{\gamma}-\gamma)\cdot\sqrt{\sum x_i^2}\,\Big|_{X=x} \sim N(0,\sigma^2) \Rightarrow (\hat{\gamma}-\gamma)\sqrt{\sum X_i^2} \sim N(0,\sigma^2)$

$\hat{\sigma}^2|_{X=x} \sim \frac{\sigma^2}{n}\cdot\chi^2_{n-1}$ $\leftarrow$ NOT depend on "x" $\quad \hat{\sigma}^2 \perp\!\!\!\perp X \quad \hat{\sigma}^2 \sim \frac{\sigma^2}{n}\cdot\chi^2_{n-1}$

(∵) $\frac{1}{\sigma^2}\varepsilon^T H\varepsilon \sim \chi^2_{tr(H)}$, where $H = I - w(w^Tw)^{-1}w^T$

$tr(H) = n-1$

Then, $\sqrt{n}\left(\frac{\hat{\sigma}^2}{\sigma^2} - 1\right)\Big|_{X=x} = \sqrt{n}\left(\frac{1}{n}\chi^2_{n-1} - 1\right) \to_d N(0,2)$

from CLT: $\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n-1} Z_k^2 - EZ_1^2\right) \to_d N(0, Var Z_1^2)$ where $Z_i^2 \overset{iid}{\sim} \chi^2_1$

$\sqrt{n}(\hat{\gamma}-\gamma)|_{X=x} = \frac{\sqrt{n}(\hat{\gamma}-\gamma)}{(w^Tw)^{1/2}}\Big|_{X=x}\cdot(w^Tw)^{1/2} \sim N(0,\sigma^2)\cdot(w^Tw)^{1/2}$

$\frac{(\hat{\gamma}-\gamma)}{(w^Tw)^{1/2}}\Big|_{X=x} \sim N(0,\sigma^2) \Rightarrow \left(\sum x_i^2\right)^{1/2}\cdot(\hat{\gamma}-\gamma) \sim N(0,\sigma^2)$

$\Rightarrow \sqrt{n}(\hat{\gamma}-\gamma)|_{X=x} \to_d N(0,\sigma^2)$ (?) $\quad \not\Rightarrow \sqrt{n}(\hat{\gamma}-\gamma) \to_d N(0,\sigma^2)$ $\quad ∵ \frac{(\sum x_i^2)^{1/2}}{\sqrt{n}} \to_p 1$

$\sqrt{n}(\hat{\gamma}-\gamma)$

$\hat{\gamma}-\gamma|_{X=x} = \sum_i w_i\varepsilon_i$

Assume $\sqrt{n}(\hat{\gamma}-\gamma)|_{X=x} \to_d N(0,\sigma^2)$

$$∴ \sqrt{n}\left(\begin{bmatrix}\hat{\gamma}\\ \hat{\sigma}^2\end{bmatrix} - \begin{bmatrix}\gamma\\ \sigma^2\end{bmatrix}\right) \to_d N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma^2 & 0\\ 0 & 2\sigma^4\end{bmatrix}\right)$$

$$\sqrt{n}(\hat{\gamma}-\gamma)\big|_{X=x} = \frac{\sqrt{n}(\hat{\gamma}-\gamma)}{\sqrt{n}(\omega^T\omega)^{1/2}}\Big|_{X=x}\cdot(\sqrt{n}\,\omega^T\omega)^{1/2} \to_d \underline{N(0,\sigma^2)}\cdot 1$$

$\hat{\sigma}^2|_{X=x}$

$\frac{1}{n}\chi^2_{n-1} \to_d N$ $\quad Z_i^2 \sim \chi^2_1$

$$\sqrt{n}\Big(\frac{1}{n}\sum_{k=1}^{n-1} Z_k^2 - E Z^2\Big) \to_d N(0, Var(Z^2))$$

$$\sqrt{n}\Big(\frac{1}{n}\chi^2_{n-1}-1\Big) \to_d N(0,2)$$

$$\sqrt{n}\Big(\frac{\hat{\sigma}^2}{\sigma^2}-1\Big) \to_d N(0,2)$$

$$\sqrt{n}(\hat{\sigma}^2-\sigma^2)\big|_{X=x} \to_d N(0,2\sigma^4)$$

From this,

$$\frac{\hat{\gamma}-\gamma}{\sigma(\omega^T\omega)^{1/2}}\Big|_{X=x} \sim N(0,1)$$

$$\frac{n}{\sigma^2}\hat{\sigma}^2\Big|_{X=x} \sim \chi^2_{n-1}$$

Since thes two cond'l dist^n does not depend on $X=x$, we can get

$$A = \frac{\hat{\gamma}-\gamma}{\sigma}\cdot\Big(\sum_i x_i^2\Big)^{1/2} \sim N(0,1) \qquad \& \quad A \perp\!\!\!\perp B \;..$$

$$B = \frac{n}{\sigma^2}\hat{\sigma}^2 \sim \chi^2_{n-1}$$

$$(\because)\; P(A\in\mathcal{A}, B\in\mathcal{B}) = \mathbb{E}[I(A\in\mathcal{A}, B\in\mathcal{B})] = \mathbb{E}[\mathbb{E}[I(\cdots)\,|\,X]]$$
$$= \mathbb{E}[P(A\in\mathcal{A}, B\in\mathcal{B}\,|\,X)]$$
$$= \mathbb{E}[\underbrace{P(A\in\mathcal{A}|X)}_{\text{does not depend on } X}\cdot\underbrace{P(B\in\mathcal{B}|X)}_{\text{does not depend on } X}]$$
$$= \mathbb{E}[P(A\in\mathcal{A})\cdot P(B\in\mathcal{B})] = P(A\in\mathcal{A})\cdot P(B\in\mathcal{B})$$

Finally, $(\alpha^*, \beta^*) = g(\gamma,\sigma^2) = \big(\exp(\tfrac{1}{2}\gamma^2+\tfrac{1}{2}\sigma^2),\; \gamma\cdot\exp(\tfrac{1}{2}\gamma^2+\tfrac{1}{2}\sigma^2)\big)$

$$\Rightarrow \frac{\partial g}{\partial(\gamma,\sigma^2)^T} = \begin{bmatrix} \gamma\exp(\frac{1}{2}\gamma^2+\frac{1}{2}\sigma^2) & \frac{1}{2}\exp(\frac{1}{2}\gamma^2+\frac{1}{2}\sigma^2) \\ \exp(\frac{1}{2}\gamma^2+\frac{1}{2}\sigma^2) + \gamma^2\exp(\frac{1}{2}\gamma^2+\frac{1}{2}\sigma^2) & \gamma\exp(\frac{1}{2}\gamma^2+\frac{1}{2}\sigma^2)\cdot\frac{1}{2} \end{bmatrix} = \exp(\tfrac{1}{2}\gamma^2+\tfrac{1}{2}\sigma^2)\cdot\begin{bmatrix} \gamma & \frac{1}{2} \\ 1+\gamma^2 & \frac{1}{2}\gamma \end{bmatrix}$$

$\boxed{\text{Note}}$ $\quad g: \mathbb{R}^n \to \mathbb{R}^m$

$$\underset{m\times 1}{g(X)} = \underset{m\times 1}{g(\mu)} + \underset{m\times n}{\frac{\partial g}{\partial\mu^T}}\underset{n\times 1}{(X-\mu)} + \cdots$$

$$\Rightarrow \frac{\partial g}{\partial(\gamma,\sigma^2)^T}\begin{bmatrix}\sigma^2 & 0 \\ 0 & 2\sigma^4\end{bmatrix}\Big(\frac{\partial g}{\partial(\gamma,\sigma^2)^T}\Big)^T = \exp(\gamma^2+\sigma^2)\cdot\begin{bmatrix}\gamma & \frac{1}{2} \\ 1+\gamma^2 & \frac{1}{2}\gamma\end{bmatrix}\begin{bmatrix}\sigma^2\gamma & \sigma^2+\sigma^2\gamma^2 \\ \sigma^4 & \sigma^4\gamma\end{bmatrix}$$

$$= \exp(\gamma^2+\sigma^2)\begin{bmatrix}\sigma^2\gamma^2+\frac{1}{2}\sigma^4 & \sigma^2\gamma+\sigma^2\gamma^3+\frac{1}{2}\sigma^4\gamma \\ \sigma^2\gamma+\sigma^2\gamma^3+\frac{1}{2}\sigma^4\gamma & \sigma^2(1+\gamma^2)^2+\frac{1}{2}\sigma^4\gamma^2\end{bmatrix}$$

$$\sqrt{n}\left(\begin{bmatrix}\hat{\gamma} \\ \hat{\sigma}^2\end{bmatrix} - \begin{bmatrix}\gamma \\ \sigma^2\end{bmatrix}\right) \to_d N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}\sigma^2 & 0 \\ 0 & 2\sigma^4\end{bmatrix}\right)$$

Δ - method :

$$\sqrt{n}\left(\begin{bmatrix}\hat{\alpha}^* \\ \hat{\beta}^*\end{bmatrix} - \begin{bmatrix}\alpha^* \\ \beta^*\end{bmatrix}\right) \to_d N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \exp(\gamma^2+\sigma^2)\cdot\sigma^2\begin{bmatrix}\gamma^2+\frac{1}{2}\sigma^2 & \gamma+\gamma^3+\frac{1}{2}\sigma^2\gamma \\ \gamma+\gamma^3+\frac{1}{2}\sigma^2\gamma & (1+\gamma^2)^2+\frac{1}{2}\sigma^2\gamma^2\end{bmatrix}\right)$$

(d)

From (b): $\sqrt{n}\left(\hat{\beta}^*_{(b)} - \beta^*\right) \to_d N\left(0, \mathbb{E}X^2Y^2 - \beta^{*2}\right)$

$$\mathbb{E}X^2Y^2 = \mathbb{E}\, X^2 \exp(2\gamma X + 2\varepsilon) \qquad (\because)\ \log Y = \gamma X + \varepsilon \Rightarrow Y = \exp(\gamma X + \varepsilon)$$

$$= \mathbb{E}\, X^2 \cdot \exp(2\gamma X) \cdot \mathbb{E}\exp(2\varepsilon)$$

$$\left(\mathbb{E}\exp(tX) = \exp\left(\tfrac{1}{2}t^2\right) \Rightarrow \mathbb{E}\, X \cdot \exp(tX) = \exp\left(\tfrac{1}{2}t^2\right)\cdot t \right.$$
$$\left. \Rightarrow \mathbb{E}\, X^2 \cdot \exp(tX) = \exp\left(\tfrac{1}{2}t^2\right)\cdot t^2 + \exp\left(\tfrac{1}{2}t^2\right) \right)$$

$$= \left\{ \exp\left(\tfrac{1}{2}(2\gamma)^2\right)\cdot(2\gamma)^2 + \exp\left(\tfrac{1}{2}(2\gamma)^2\right) \right\} \cdot \exp\left(\tfrac{1}{2}\sigma^2 \cdot 2^2\right)$$

$$= \exp(2\gamma^2 + 2\sigma^2)\{4\gamma^2 + 1\}$$

$$\beta^{*2} = \gamma^2 \exp(\gamma^2 + \sigma^2)$$

$$\therefore \sigma^2_{(b)} = \exp(2\gamma^2 + 2\sigma^2)\{4\gamma^2+1\} - \gamma^2 \exp(\gamma^2 + \sigma^2)$$

$$\mathbb{E}XY$$
$$= \mathbb{E}\, X \exp(\gamma X + \varepsilon)$$
$$= \underline{\mathbb{E}\, X \exp(\gamma X)} \cdot \mathbb{E}\exp(\varepsilon)$$
$$= \gamma \exp\left(\tfrac{1}{2}\gamma^2\right) \cdot \exp\left(\tfrac{1}{2}\sigma^2 \cdot 1^2\right)$$

From (c): $\sqrt{n}\left(\hat{\beta}^*_{(c)} - \beta^*\right) \to_d N\left(0, \underbrace{\exp(\gamma^2+\sigma^2)\cdot\sigma^2\cdot\left\{(1+\gamma^2)^2 + \tfrac{1}{2}\sigma^2\gamma^2\right\}}_{=\sigma^2_{(c)}}\right)$

$$\mathrm{ARE}\left(\hat{\beta}^*_{(b)}, \hat{\beta}^*_{(c)}\right) = \frac{\sigma^2_{(c)}}{\sigma^2_{(b)}} = \frac{\sigma^2\left\{(1+\gamma^2)^2 + \frac{1}{2}\sigma^2\gamma^2\right\}}{\exp(\gamma^2+\sigma^2)\cdot\{4\gamma^2+1\} - \gamma^2}$$

(e) $\mathbb{E}\left[\{Y - g(X)\}^2\right] = \mathbb{E}\left[\mathbb{E}\left[\{Y - g(X)\}^2 \mid X\right]\right]$

$$\mathbb{E}\left[\{Y-g(X)\}^2 \mid X\right] = \mathbb{E}\left[\left\{(Y - \mathbb{E}[Y|X]) + (\mathbb{E}[Y|X] - g(X))\right\}^2 \mid X\right]$$

$$= \mathbb{E}\left[(Y - \mathbb{E}[Y|X])^2 \mid X\right] + 2\,\mathbb{E}\left[(Y - \mathbb{E}[Y|X])\cdot \underbrace{(\mathbb{E}[Y|X] - g(X))}_{ft^n \text{ of } X} \mid X\right] + \mathbb{E}\left[\underbrace{(\mathbb{E}[Y|X] - g(X))^2}_{ft^n \text{ of } X} \mid X\right]$$

$$= \mathrm{Var}(Y|X) + 2\left(\mathbb{E}[Y|X] - g(X)\right)\cdot \underbrace{\mathbb{E}\left[Y - \mathbb{E}[Y|X] \mid X\right]}_{\overset{\text{"}}{0}} + \left(\mathbb{E}[Y|X] - g(X)\right)^2$$

$$= \mathrm{Var}(Y|X) + \left(\mathbb{E}[Y|X] - g(X)\right)^2$$

$$\geq \mathrm{Var}(Y|X). \qquad \text{"="} \text{ iff } g(X) = \mathbb{E}[Y|X]$$

Therefore,

$\mathbb{E}\left[\{Y - g(X)\}^2\right] \geq \mathbb{E}\left[\{Y - \mathbb{E}[Y|X]\}^2\right]$ proved. $\therefore g(X)_{min} = \mathbb{E}[Y|X]$

$$= \mathbb{E}\left[\exp(\gamma X + \varepsilon) \mid X\right]$$
$$= \exp(\gamma X)\cdot \mathbb{E}\varepsilon$$
$$= \exp\left(\gamma X + \tfrac{1}{2}\sigma^2\right)$$

2019 S1 Q2

$X_1, \dots, X_n \overset{iid}{\sim} f(\cdot)$. $f(x) = \frac{1}{\theta} e^{(a-x)/\theta} I(x>a)$, $\theta > 0$

(a) $H_0: \theta \ge \theta_0$ vs $H_1: \theta < \theta_0$. UMP size α test ?

First consider $H_0(\theta_0): \theta = \theta_0$ vs $H_1: \theta = \theta_1$, where $\theta_1 < \theta_0$ known.

NP lemma : MP size α test of $H_0(\theta_0)$ vs H_1 is given by

(i) $\phi(x) = \begin{cases} 1 & L(x;\theta_1)/L(x;\theta_0) > K \\ \gamma & = \\ 0 & < \end{cases}$ (ii) $E_{\theta_0}\phi(X) = \alpha$

$$\frac{L(x;\theta_1)}{L(x;\theta_0)} = \frac{\prod_{i=1}^n \frac{1}{\theta_1} e^{(a-x_i)/\theta_1} I(x_i>a)}{\prod_{i=1}^n \frac{1}{\theta_0} e^{(a-x_i)/\theta_0} I(x_i>a)} = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left(\sum_{i=1}^n \frac{a-x_i}{\theta_1} - \sum_{i=1}^n \frac{a-x_i}{\theta_0}\right)$$

$$= \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left(\left(\frac{1}{\theta_1} - \frac{1}{\theta_0}\right)\left(na - \sum_i x_i\right)\right)$$

Since $X_1 \sim X_n$ are from continuous dist^n, we can omit $\frac{L(x;\theta_1)}{L(x;\theta_0)} = K$ case (zero prob.)

$\frac{L(x;\theta_1)}{L(x;\theta_0)} > K \iff \sum_i x_i < C$, $\exists c$

$\phi(x) = I\left(\sum_i x_i < c\right)$, where $P_{\theta_0}\left(\sum_i X_i < c\right) = \alpha$

To get exact value of c, observe that $Y_i := X_i - a \Rightarrow f_{Y_i}(y) = \frac{1}{\theta} e^{-y/\theta} I(y>0)$

$\Rightarrow Y_i \sim Ga(1, \theta)$

Then, $\sum_i Y_i = \sum_i (X_i - a) \sim Ga(n, \theta) \Rightarrow \frac{2}{\theta}\sum_i Y_i \sim Ga(n, 2) \overset{d}{=} \chi^2_{2n}$

$\therefore \alpha = P_{\theta_0}\left(\sum_i X_i < c\right) = P_{\theta_0}\left(\sum_i (X_i - a) < c - na\right) = P_{\theta_0}\left(\sum_i Y_i < c - na\right)$

$= P_{\theta_0}\left(\frac{2}{\theta_0}\sum_i Y_i < \frac{2}{\theta_0}(c - na)\right)$

$\therefore \frac{2}{\theta_0}(c - na) = F^{-1}(\alpha)$, where F: cdf of χ^2_{2n}

$\therefore C = na + \frac{\theta_0}{2} F^{-1}(\alpha)$: NOT depend on θ_1

$\Rightarrow$ We can extend this test as the UMP size α test of

$H_0(\theta_0): \theta = \theta_0$ vs $H_1: \theta < \theta_0$

Finally, compute power at θ

$\beta(\theta) = E_\theta \phi(X) = P_\theta\left(\sum_i X_i < na + \frac{\theta_0}{2} F^{-1}(\alpha)\right) = P_\theta\left(\frac{2}{\theta}\sum_i (X_i - a) < \frac{\theta_0}{\theta} F^{-1}(\alpha)\right)$

$= F\left(\frac{\theta_0}{\theta} F^{-1}(\alpha)\right)$ ($\because$) $\frac{2}{\theta}\sum_i (X_i - a) \sim \chi^2_{2n}$

: decreasing ft^n of θ.

Thus, $\sup_{\theta \ge \theta_0} E_\theta \phi(X) = E_{\theta_0}\phi(X)$ holds.

Now, Consider an arbitrary size α test of $H_0: \theta \ge \theta_0$ vs $H_1: \theta < \theta_0$, denoted by ϕ^*

Then, $\sup_{\theta \ge \theta_0} \mathbb{E}\, \phi^*(X) = \alpha$.

$$\Rightarrow \quad \mathbb{E}_{\theta_0} \phi^*(X) \le \alpha$$

$$\Rightarrow \quad \text{size } \alpha \text{ test of } H_0(\theta_0): \theta = \theta_0 \text{ vs } H_1: \theta < \theta_0$$

Since ϕ: UMP of the above hypothesis,

$$\mathbb{E}_{\theta_1} \phi(X) \ge \mathbb{E}_{\theta_1} \phi^*(X), \ \forall\, \theta_1 < \theta_0$$

Therefore, ϕ achieves maximal power on $\theta < \theta_0$ among all size α test of $H_0: \theta \ge \theta_0$.

$\therefore$ ϕ: UMP size α test of $H_0: \theta \ge \theta_0$ vs $H_1: \theta < \theta_0$

(b) $L(x;\theta) = \theta^{-n} \exp\left(\frac{1}{\theta}\left(na - \sum_i x_i\right)\right) I\left(x_{(1)} > a\right)$: exp'l family.

$$l(x;\theta) = -n\log\theta + \theta^{-1}\left(na - \sum_i x_i\right)$$

$$\frac{\partial l}{\partial \theta} = 0 \Leftrightarrow -n\theta^{-1} - \theta^{-2}\left(na - \sum_i x_i\right) = 0 \Leftrightarrow \hat{\theta} = \bar{x} - a$$

<u>Asymp dist^n theory of MLE of exp'l family</u> : $\sqrt{n}(\hat{\theta} - \theta) \to_d N(0, I(\theta)^{-1})$

$$\frac{\partial^2 l}{\partial \theta^2} = n\theta^{-2} + 2\theta^{-3}\left(na - \sum_i x_i\right)$$

$$\Rightarrow I(\theta) = \mathbb{E}\left[-\frac{\partial^2 l}{\partial \theta^2}\right] = -n\theta^{-2} - 2\theta^{-3} na + 2\theta^{-3}\, \mathbb{E}\left[\sum_i X_i\right]$$

$$= -n\theta^{-2} - 2na\theta^{-3} + 2\theta^{-3}(n\theta + na) = 2n\theta^{-2} - n\theta^{-2} = n\theta^{-2}$$

$$\left(Y_i = X_i - a \sim Ga(1,\theta) \Rightarrow \mathbb{E}\left[\sum_i X_i\right] = \mathbb{E}\left[\sum_i Y_i\right] + na = n\theta + na\right)$$

$$\therefore \sqrt{n}(\hat{\theta} - \theta) \to_d N\left(0, \frac{\theta^2}{n}\right)$$

(c) Let $Y_i = X_i - \theta$. Then, $f_{Y_i}(y) = \theta^{-1} e^{-y/\theta} I(y > 0) \Rightarrow Y_i \sim Ga(1,\theta)$

$$\Rightarrow \sum_i Y_i \sim Ga(n,\theta) \Rightarrow \frac{2}{\theta}\sum_i Y_i \sim Ga(n,2) \ : \ \underline{\text{NOT}} \text{ depend on } \theta$$

$$\Rightarrow \frac{\bar{X}}{\theta} = \frac{1}{n}\frac{\sum X_i}{\theta} = \frac{1}{n}\frac{\sum Y_i}{\theta} + 1 \sim \frac{1}{2n}\cdot Ga(n,2) + 1 \ : \ \underline{\text{NOT}} \text{ depend on } \theta \ \therefore \text{ pivotal}$$

Compute $\frac{X_{(1)}}{\theta}$'s dist^n.

$$P\left(\frac{X_{(1)}}{\theta} \ge k\right) = P\left(X_{(1)} \ge \theta k\right) = P(\forall X_i \ge \theta k) = \prod_{i=1}^n P(X_i \ge \theta k) = \prod_{i=1}^n \int_{\theta k}^\infty \frac{1}{\theta} e^{-(x-\theta)/\theta}\, dx$$

$(k \ge 1)$

$$= \left(\left[-e^{-(x-\theta)/\theta}\right]_{\theta k}^\infty\right)^n = \left(e^{(\theta k - \theta)/\theta}\right)^n = \left(e^{k-1}\right)^n = e^{n(k-1)} \ : \ \underline{\text{NOT}} \text{ depend on } \theta$$

$\therefore$ pivotal

(d)

$f(x) = \theta^{-1} e^{(\theta-x)/\theta} I(x>\theta)$

$Z_i := \frac{1}{\theta}(X_i-\theta)$.

$P(Z_i \le k) = P(X_i \le \theta + k\theta) = \int_\theta^{\theta+k\theta} \theta^{-1} e^{(\theta-x)/\theta}\, dx = \int_0^k \theta^{-1} e^{-z}\cdot\theta\, dz \quad \left(z=\frac{1}{\theta}(x-\theta)\right)$

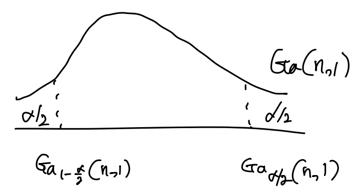
$= \left[-e^{-z}\right]_0^k = 1-e^{-k}$: cdf of $Exp(1) = Ga(1,1)$

∴ Z_i iid $Exp(1)$.

① Based on $\frac{\bar X}{\theta} = \frac{1}{\theta}\frac{1}{n}\sum_{i=1}^n(\theta Z_i+\theta) = \frac{1}{n}\sum_{i=1}^n Z_i + 1$

Note that $\sum_{i=1}^n Z_i \sim Ga(n,1)$

$\Rightarrow P\left(a_1 \le \frac{\bar X}{\theta} \le a_2\right) = P\left(\underbrace{n(a_1-1)}_{= Ga_{1-\frac{\alpha}{2}}(n,1)} \le \sum_{i=1}^n Z_i \le \underbrace{n(a_2-1)}_{= Ga_{\alpha/2}(n,1)}\right)$



⇒ $(1-\alpha)$ CI for θ is given by

$1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(n,1) \le \frac{\bar X}{\theta} \le 1+\frac{1}{n}Ga_{\alpha/2}(n,1) \Rightarrow \theta \in \left[\frac{\bar X}{1+\frac{1}{n}Ga_{\alpha/2}(n,1)}, \frac{\bar X}{1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(n,1)}\right]$

② Based on $\frac{X_{(1)}}{\theta} = \frac{1}{\theta}(\theta Z_{(1)}+\theta) = Z_{(1)}+1$

Note that $f_{Z_{(1)}}(z_1) = \frac{n!}{0!\cdot 1\cdot (n-1)!}\left[F(z_1)\right]^0 \cdot f(z_1)\cdot\left[1-F(z_1)\right]^{n-1} = n e^{-z_1}(e^{-z_1})^{n-1} I(z_1>0)$

$= n e^{-nz_1} I(z_1>0)$

$\Rightarrow Z_{(1)} \sim Exp\left(\frac{1}{n}\right) = Ga\left(1,\frac{1}{n}\right) \Rightarrow nZ_{(1)} \sim Ga(1,1)$

$\Rightarrow P\left(b_1 \le \frac{X_{(1)}}{\theta} \le b_2\right) = P\left(\underbrace{n(b_1-1)}_{= Ga_{1-\frac{\alpha}{2}}(1,1)} \le nZ_{(1)} \le \underbrace{n(b_2-1)}_{= Ga_{\frac{\alpha}{2}}(1,1)}\right)$

⇒ $(1-\alpha)$ CI for θ is given by

$1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(1,1) \le \frac{X_{(1)}}{\theta} \le 1+\frac{1}{n}Ga_{\alpha/2}(1,1) \Rightarrow \theta \in \left[\frac{X_{(1)}}{1+\frac{1}{n}Ga_{\alpha/2}(1,1)}, \frac{X_{(1)}}{1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(1,1)}\right]$

$X_i = \theta Z_i + a$. $E X_i = \theta\cdot 1 + a = \theta + a$ $\quad Var X_i = \theta^2\cdot 1 = \theta^2$

$P(n(\bar x - a) \le k) = P(n(\theta\bar z + a - a)\le k) = P(\theta\cdot\sum z_i \le k) = P\left(\sum z_i \le \frac{k}{\theta}\right) = F_{Ga(n,\theta)}(k)$

$n(\bar x - a) \to_d Ga(n,\theta)$

$P(n(x_{(1)}-a)\le k) = P(n(\theta z_{(1)} + a - a)\le k) = P\left(z_{(1)} \le \frac{k}{n\theta}\right) = 1 - P\left(\forall z_i > \frac{k}{n\theta}\right) = 1-\left(1-\left(1-e^{-\frac{k}{n\theta}}\right)\right)^n = 1-e^{-\frac{k}{\theta}}$

$\Rightarrow n(x_{(1)}-a) \to_d Exp(\theta) = Ga(1,\theta)$

CI length based on $\bar{x}$:

$$L_1 = \bar{x}\left(\frac{1}{1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(n,1)} - \frac{1}{1+\frac{1}{n}Ga_{\frac{\alpha}{2}}(n,1)}\right)$$

$$= \bar{x}\,\frac{1}{(1+\quad)(1+\quad)}\cdot\frac{1}{n}\left(Ga_{\frac{\alpha}{2}}(n,1) - Ga_{1-\frac{\alpha}{2}}(n,1)\right)$$

$$= \frac{\bar{X}}{n}\cdot\left(Ga_{\frac{\alpha}{2}}(n,1) - Ga_{1-\frac{\alpha}{2}}(n,1)\right)\cdot C_{1n} \qquad \left(C_{1n} \xrightarrow[n]{} C_1 : \text{constant}\right)$$

CI length based on $X_{(1)}$:

$$L_2 = X_{(1)}\left(\frac{1}{1+\frac{1}{n}Ga_{1-\frac{\alpha}{2}}(1,1)} - \frac{1}{1+\frac{1}{n}Ga_{\frac{\alpha}{2}}(1,1)}\right)$$

$$= X_{(1)}\,\frac{1}{(1+\quad)(1+\quad)}\cdot\frac{1}{n}\left(Ga_{\frac{\alpha}{2}}(1,1) - Ga_{1-\frac{\alpha}{2}}(1,1)\right)$$

$$= \frac{X_{(1)}}{n}\cdot C_{2n} \qquad \left(C_{2n} \xrightarrow[n]{} C_2 : \text{constant}\right)$$

Compare the convergence rates of these two.

First, use the fact that $Ga_{\frac{\alpha}{2}}(n,1) \simeq n + \sqrt{n}\, Z_{\frac{\alpha}{2}}$

$(\because)$ F_n : cdf of $Ga(n,1)$. $V_n \sim Ga(n,1) \overset{d}{=} Ga(1,1) + \cdots + Ga(1,1) \Rightarrow \sqrt{n}\left(\frac{1}{n}V_n - 1\right) \rightarrow_d N(0,1)$

$$1-\frac{\alpha}{2} = P\left(V_n \le Ga_{\frac{\alpha}{2}}(n,1)\right) = P\left(\sqrt{n}\left(\frac{1}{n}V_n - 1\right) \le \sqrt{n}\left(\frac{1}{n}Ga_{\frac{\alpha}{2}}(n,1) - 1\right)\right)$$

$$\underset{n\to\infty}{\simeq} P\left(Z \le \sqrt{n}\left(\frac{1}{n}Ga_{\frac{\alpha}{2}}(n,1) - 1\right)\right)$$

$$\Rightarrow \sqrt{n}\left(\frac{1}{n}Ga_{\frac{\alpha}{2}}(n,1) - 1\right) \simeq Z_{\frac{\alpha}{2}}$$

$$\Rightarrow Ga_{\frac{\alpha}{2}}(n,1) \simeq n + \sqrt{n}\cdot Z_{\frac{\alpha}{2}}$$

Then, $L_1 \simeq \frac{\bar{X}}{n}\left(n + \sqrt{n}\, Z_{\frac{\alpha}{2}} - n - \sqrt{n}\, Z_{1-\frac{\alpha}{2}}\right)\cdot C_{1n}$

$$\simeq \frac{\bar{X}}{n}\sqrt{n}\left(Z_{\frac{\alpha}{2}} - Z_{1-\frac{\alpha}{2}}\right)\cdot C_{1n} \quad : \text{rate of } \left(\frac{\bar{X}}{\sqrt{n}}\right)$$

$$L_2 = \frac{X_{(1)}}{n} \quad : \text{rate of } \left(\frac{X_{(1)}}{n}\right)$$

Comparison of L_1 vs L_2 $\Leftrightarrow$ $\frac{\bar{X}}{\sqrt{n}}$ vs $\frac{X_{(1)}}{n}$ $\Leftrightarrow$ $\bar{X}$ vs $\frac{X_{(1)}}{\sqrt{n}}$

Since $\bar{X} \rightarrow_p EX_1 = 2\theta$

$X_{(1)} \rightarrow_p \theta \Rightarrow \frac{X_{(1)}}{\sqrt{n}} \rightarrow_p 0$], $\frac{\bar{X}}{\sqrt{n}} = O_p(1)$ while $\frac{X_{(1)}}{n} = o_p\left(\frac{1}{\sqrt{n}}\right)$

$\therefore$ L_2 is shorter when suff. large n.