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 $(X_1, Y_1), \ldots, (X_n, Y_n)$. 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

   $$\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} E[Y] \\ E[XY] \end{pmatrix}. \qquad (1)$$

   method of moments estimator

   (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.

2019 Theory I

1. Let $X \sim N(0,1)$ & let Y be an outcome variable
(X,Y) has finite 2nd moment & $E[X^2Y^2] < \infty$
Assume we observe n iid copies (X,Y): $(X_1,Y_1) \cdots (X_n,Y_n)$

(a) consider a linear function for Y, $\alpha + \beta X$, to minimize the following squared loss: $E[\{Y-(\alpha+\beta X)\}^2]$ show that the optimal (α,β) are $\begin{pmatrix}\alpha^*\\ \beta^*\end{pmatrix} = \begin{pmatrix}E[Y]\\ E[XY]\end{pmatrix}$

$$\nabla^2(\cdot) = \begin{pmatrix}1&0\\0&1\end{pmatrix}^*$$

positive definite $\Rightarrow$ minimum.

$$E[\{Y-(\alpha+\beta X)\}^2] = E[Y^2 - 2Y(\alpha+\beta X) + (\alpha+\beta X)^2]$$
$$= E[Y^2 - 2\alpha Y - 2\beta XY + \alpha^2 + 2\alpha\beta X + \beta^2 X^2]$$
$$f(X,Y) = E[Y^2] - 2\alpha E[Y] - 2\beta E[XY] + \alpha^2 + 2\alpha\beta E[X] + \beta^2 E[X^2]$$

$$\frac{\partial f}{\partial \alpha} = -2E[Y] + 2\alpha + 2\beta E[X] \overset{set}{=} 0$$
$$\alpha + \beta E[X] = E[Y] \overset{E[X]=0}{\Longrightarrow} \alpha^* = E[Y]$$

$$\frac{\partial^2 f}{\partial \alpha^2} = 2 > 0 \therefore \text{minim.}$$

$$\frac{\partial f}{\partial \beta} = -2E[XY] + 2\alpha E[X] + 2\beta E[X^2] \overset{set}{=} 0$$
$$\beta E[X^2] + \alpha E[X] = E[XY]$$

$$\frac{\partial^2 f}{\partial \beta^2} = 2 > 0 \therefore \text{min.}$$

we need to solve for α, β in these two equations:

① $\alpha + \beta E[X] = E[Y]$ ② $\beta E[X^2] + \alpha E[X] = E[XY]$

$\alpha = E[Y] - \beta E[X] = E[Y] - \beta \cdot 0 = E[Y]$

NOTE: $E[X] = 0$
$E[X^2] = 1$
$var(X) = 1$

~~plug ① into ②:~~ $\cancel{\beta E[X^2] + E[Y] - \beta E[X]^2 = E[XY]}$

$\cancel{\beta(E[X^2] - E[X]^2) + E[Y] = E[XY]}$

$\cancel{\beta\, var(X) + E[Y] = E[XY] \Rightarrow \beta = \frac{E[XY]-E[Y]}{var(X)} = E[XY]-E[Y]}$ ($var(X) = 1$)

$\cancel{\alpha = E[Y] - (E[XY]-E[Y])E[X] = E[Y]}$

$$\beta E[X^2] + \alpha E[X] = E[XY] \Rightarrow \beta = E[XY]$$

$\therefore \alpha^* = E[Y]$
$\beta^* = E[XY]$ ☺

1.(b) From (a) we estimate (α^*, β^*) as $\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} X_i Y_i$

Give the asymptotic distribution of the obtained estimator after proper normalization.

With finite second moments, for both X·Y and Y, we can apply the CLT to find asymptotic distributions of $\hat{\alpha}$ & $\hat{\beta}$.

$$\sqrt{n}\left(\hat{\alpha} - \alpha^*\right) \xrightarrow[d]{} N(0, \mathrm{var}(Y))$$

and $$\sqrt{n}(\hat{\beta} - \beta^*) \xrightarrow[d]{} N(0, \mathrm{var}(XY))$$

how far do we need to go w/ these variances?

$$\mathrm{var}(Y) = E[Y^2] - E[Y]^2$$

$$\mathrm{var}(XY) = E[X^2Y^2] - E[XY]^2$$

similarly,

$$\sqrt{n}\begin{pmatrix} \hat{\alpha} - \alpha^* \\ \hat{\beta} - \beta^* \end{pmatrix} \xrightarrow[d]{} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma\right)$$

where $$\Sigma = \begin{pmatrix} \mathrm{var}(y_i) & \mathrm{cov}(y_i, x_i y_i) \\ & \mathrm{var}(x_i y_i) \end{pmatrix} = \begin{pmatrix} E[y_i^2] - \alpha^{*2} & E[x_i y_i^2] - \alpha^* \beta^* \\ & E[x_i^2 y_i^2] - \beta^{*2} \end{pmatrix}$$

$$\left.\begin{aligned} \mathrm{var}(y_i) &= E[y_i^2] - E[y_i]^2 = E[y_i^2] - \alpha^{*2} \\ \mathrm{Var}(x_i y_i) &= E[x_i^2 y_i^2] - E[x_i y_i]^2 = E[x_i^2 y_i^2] - \beta^{*2} \\ \mathrm{Cov}(y_i, x_i y_i) &= E[x_i y_i^2] - E[y_i]E[x_i y_i] = E[x_i y_i^2] - \alpha^* \beta^* \end{aligned}\right\} \text{plug in above}$$

1\. con't Now suppose that we know the distribution of

2019 Theory I (c)

$Y|X \sim$ lognormal s.t. $\log Y = \gamma X + N(0,\sigma^2) \Rightarrow \log Y|X \sim N(\gamma X, \sigma^2 + \gamma^2)$

let $Z = \log Y \Rightarrow Z|X \sim N(\gamma X, \sigma^2)$

$\Rightarrow Y = e^Z \quad \frac{dz}{dy} = \frac{1}{y} \Rightarrow f_{Y|X}(y|x) = f_{Z|X}(z|x)\frac{1}{y} = \frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{y}\exp\left\{-\frac{1}{2\sigma^2}(\log y - \gamma x)^2\right\}$

To find the joint distribution of (X,Y)

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = \frac{1}{y\sigma}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}(\log y - \gamma x)^2\right\}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{x^2}{2}\right\}$$

$$= \frac{1}{2\pi y\sigma}\exp\left\{-\frac{1}{2\sigma^2}(\log y - \gamma x)^2 - \frac{x^2}{2}\right\}$$

$\alpha^* = E_{(X,Y)}[Y] \quad \beta^* = E_{(X,Y)}[XY]$

all given x.

Note that $E[e^{Zt}] = E[e^{\log Y t}] = E[e^{\log(Y^t)}] = E[Y^t]$

$\therefore E[Y|X] = M_{Z|X}(t=1) = \exp\{\gamma X t + \frac{\sigma^2}{2}t^2\} = \exp\{\gamma X + \frac{\sigma^2}{2}\}$ $\quad E[Y^2|X] = \exp\{2\gamma X + 2\sigma^2\}$

$$E_{(X,Y)}[Y] = E_X[E_{Y|X}[Y|X]] = E_X[\exp\{\gamma X + \tfrac{\sigma^2}{2}\}] = E_X[\exp\{\gamma X\}\exp\{\tfrac{\sigma^2}{2}\}]$$

$$= \exp\{\tfrac{\sigma^2}{2}\}E_X[e^{\gamma X}] = \exp\{\tfrac{\sigma^2}{2}\}M_X(\gamma) = \exp\{\tfrac{\sigma^2}{2}\}\exp\{\tfrac{\gamma^2}{2}\}$$

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

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

$$\beta^* = E_{(X,Y)}[XY] = E_X[E_{Y|X}[XY|X]] = E_X[X E[Y|X]] = E_X[X\exp\{\gamma X + \tfrac{\sigma^2}{2}\}] = E[Xe^{\gamma X}]e^{\sigma^2/2}$$

See trick pg from Kosorok session

$$= \int_{-\infty}^{\infty} x\exp\{\gamma x + \tfrac{\sigma^2}{2}\}\frac{1}{\sqrt{2\pi}}\exp\{-\tfrac{x^2}{2}\}dx = \exp\{\tfrac{\sigma^2}{2}\}\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}}\exp\left\{-\frac{x^2-2\gamma x}{2}\right\}dx$$

$$= \exp\{\tfrac{\sigma^2}{2}\}\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}}\exp\left\{-\frac{x^2-2\gamma x+\gamma^2-\gamma^2}{2}\right\}dx = \exp\{\tfrac{\sigma^2}{2}\}\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\gamma)^2}{2} + \frac{\gamma^2}{2}\right\}dx$$

$$= \exp\left\{\frac{\sigma^2+\gamma^2}{2}\right\}\underbrace{\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\gamma)^2}{2}\right\}dx}_{E[N(\gamma,1)]} = \exp\left\{\frac{\sigma^2+\gamma^2}{2}\right\}\gamma$$

$\therefore \beta^* = \gamma\exp\left\{\frac{\sigma^2+\gamma^2}{2}\right\}$

$\therefore$ we know (α^*, β^*) are functions of (σ, γ), thus to find the MLE of (α^*, γ^*) we can find the MLE's of σ and γ.

1.(c) con't

MLEs of (σ, γ):

$$L(\sigma,\gamma) = \prod_{i=1}^{n} \frac{1}{2\pi y_i \sigma} \exp\left\{-\frac{1}{2\sigma^2}(\log y_i - \gamma x_i)^2 - \frac{x_i^2}{2}\right\} = \frac{1}{2\pi\sigma}\left(\prod_{i=1}^{n}\frac{1}{y_i}\right)\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y_i - \gamma x_i)^2 - \frac{\sum x_i^2}{2}\right\}$$

$$l(\sigma,\gamma) \propto \sum_{i=1}^{n} -\log(y_i) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y_i - \gamma x_i)^2 - \frac{1}{2}\sum_{i=1}^{n} x_i^2 - \frac{1}{2}\log(2\pi\sigma^2)$$

$$\propto -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y_i - \gamma x_i)^2 - \frac{n}{2}\log(2\pi\sigma^2)$$

$$\frac{\partial l}{\partial \gamma} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y_i - \gamma x_i)(-2x_i) \overset{set}{=} 0$$

$$\sum_{i=1}^{n}(\log y_i)x_i - \gamma\sum_{i=1}^{n} x_i^2 \overset{set}{=} 0$$

$$\sum_{i=1}^{n}\log(y_i)x_i = \gamma\sum_{i=1}^{n} x_i^2 \Rightarrow \hat{\gamma} = \frac{\sum_{i=1}^{n} x_i \log(y_i)}{\sum_{i=1}^{n} x_i^2}$$

$$\frac{\partial l}{\partial \sigma^2} = \sum_{i=1}^{n}\frac{(\log(y_i) - \gamma x_i)^2}{2\sigma^4} - \frac{1}{2\sigma^2} \overset{set}{=} 0$$

$$\frac{1}{2\sigma^4}\sum_{i=1}^{n}(\log(y_i) - \gamma x_i)^2 = \frac{n}{2\sigma^2} \Rightarrow \frac{1}{n}\sum_{i=1}^{n}(\log(y_i) - \hat{\gamma} x_i)^2 = \hat{\sigma}^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log(y_i) - \frac{x_i \sum x_i \log(y_i)}{\sum x_i^2}\right)^2$$

$\therefore$ By invariance of MLE:

$$\hat{\alpha}^* = \exp\left\{\frac{1}{2}(\hat{\sigma}^2 + \hat{\gamma}^2)\right\} = \exp\left\{\frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\log(y_i) - x_i\frac{\sum_{i=1}^{n} x_i \log(y_i)}{\sum_{i=1}^{n} x_i^2}\right) + \frac{\sum_{i=1}^{n} x_i \log(y_i)}{\sum_{i=1}^{n} x_i^2}\right)\right\}$$

Assuming regularity conditions hold,

$$\sqrt{n}\begin{pmatrix}\hat{\gamma} - \gamma \\ \hat{\sigma}^2 - \sigma\end{pmatrix} \xrightarrow{d} N(0, I(\theta)^{-1}) \quad \text{where } I(\theta) = \lim_{n\to\infty} -\frac{1}{n}E\left[\begin{pmatrix}\frac{\partial^2 l}{\partial\gamma^2} & \frac{\partial^2 l}{\partial\gamma\partial\sigma} \\ \frac{\partial^2 l}{\partial\gamma\partial\sigma} & \frac{\partial^2 l}{\partial\sigma^2}\end{pmatrix}\right]$$

$$\frac{\partial^2 l}{\partial\gamma^2} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2x_i^2 = -\sum_{i=1}^{n}\frac{x_i^2}{\sigma^2} \qquad \frac{\partial^2 l}{\partial\sigma^2{}^2} = -\left(\sum_{i=1}^{n}\frac{(\log(y_i) - \gamma x_i)^2}{(\sigma^2)^3} - \frac{1}{2(\sigma^2)^2}\right)$$

$$\frac{\partial^2 l}{\partial\gamma\partial\sigma} = -\frac{1}{(\sigma^2)^2}\sum_{i=1}^{n} x_i(\log y_i - \gamma x_i)$$

$$E\left[-\frac{\partial^2 l}{\partial\gamma^2}\right] = E\left[\sum\frac{x_i^2}{\sigma^2}\right] = \sum_{i=1}^{n}\frac{E[x_i^2]}{\sigma^2} = \frac{nE[x_i^2]}{\sigma^2} = \frac{n}{\sigma^2}$$

$$E[(z_i - \gamma x_i)^2] \;\; \| \qquad z - \gamma x_i \mid x_i \sim N(0, \sigma^2)$$

$$E[(z_i - \gamma x_i)^2] = E_x\left[E_{z|x}[(z_i - \gamma x_i)^2]\right] = E_x[\sigma^2 + 0^2] = \sigma^2$$

$$E\left[-\frac{\partial^2 l}{\partial(\sigma^2)^2}\right] = E\left[\sum\frac{(\log(y_i) - \gamma x_i)^2}{(\sigma^2)^3} - \frac{1}{2(\sigma^2)^2}\right] = -\frac{n}{2(\sigma^2)^2} + \frac{1}{\sigma^6}\sum_{i=1}^{n}E[(\log(y_i) - \gamma x_i)^2]$$

$$= -\frac{n}{2\sigma^4} + \frac{n\sigma^2}{\sigma^6} = -\frac{n}{2\sigma^4} + \frac{n}{2\sigma^4} = \frac{n}{2\sigma^4}$$

$$E\left[-\frac{\partial^2 l}{\partial\gamma\partial\sigma^2}\right] = E\left[\frac{1}{(\sigma^2)^2}\sum_{i=1}^{n} x_i(\log y_i - \gamma x_i)\right] = \sum_{i=1}^{n}\frac{E[x_i(\log y_i - \gamma x_i)]}{\sigma^4} = \sum_{i=1}^{n}\frac{E_x E_{z|x}[x_i(z_i - \gamma x_i) \mid x_i]}{\sigma^4} = 0$$

$$\therefore I_1(\theta) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/2\sigma^4 \end{pmatrix} \Rightarrow I_1(\theta)^{-1} = \frac{1}{1/\sigma^6} \begin{pmatrix} 1/\sigma^4 & 0 \\ 0 & 1/2\sigma^4 \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}$$

$$\therefore \sqrt{n} \begin{pmatrix} \hat{\gamma} - \gamma \\ \hat{\sigma}^2 - \sigma^2 \end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix} \right)$$

let $g(a,b) = \begin{pmatrix} \exp\{\frac{1}{2}(b+a^2)\} \\ a\exp\{\frac{b+a^2}{2}\} \end{pmatrix} \Rightarrow \nabla g(a,b) = \begin{pmatrix} a\exp\{\frac{1}{2}(b+a^2)\} & \frac{1}{2}\exp\{\frac{1}{2}(b+a^2)\} \\ (1+a^2)\exp\{\frac{1}{2}(b+a^2)\} & \frac{1}{2}a\exp\{\frac{1}{2}(b+a^2)\} \end{pmatrix}$

$\therefore$ By Delta-method,

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}^* - \alpha^* \\ \hat{\beta}^* - \beta^* \end{pmatrix} \xrightarrow{d} \nabla g(\gamma, \sigma^2)\, N\left(0, \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}\right)$$

$$\nabla g(\gamma, \sigma^2) = \exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\} \begin{pmatrix} \gamma & 1/2 \\ 1+\gamma^2 & \gamma/2 \end{pmatrix}$$

$\Sigma =$

$$\nabla g(\gamma,\sigma^2) \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix} \nabla g'(\gamma,\sigma^2) = \exp\{\gamma^2+\sigma^2\} \begin{pmatrix} \gamma & 1/2 \\ 1+\gamma^2 & \gamma/2 \end{pmatrix} \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix} \begin{pmatrix} \gamma & 1+\gamma^2 \\ 1/2 & \gamma/2 \end{pmatrix}$$

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

$$\therefore \sqrt{n} \begin{pmatrix} \hat{\alpha}^* - \alpha^* \\ \hat{\beta}^* - \beta^* \end{pmatrix} \xrightarrow{d} N(0, \Sigma)$$

where Σ found above.

1.(d) Calculate the ARE between the MLE for β^* and $\hat{\beta}$ given in (b) | 2019 Theory I

Based on (b) the asymptotic variance of $\hat{\beta}$ is $var(X_iY_i)$

$var(X_iY_i) = E[(X_iY_i)^2] - E[X_iY_i]^2$ we found $E[X_iY_i]$ in (c), $E[X_iY_i] = \gamma \exp\{\frac{\sigma^2+\gamma^2}{2}\}$

$$E[X_i^2Y_i^2] = E_X[E_{Y|X}[X_i^2Y_i^2|X_i]] = E_X[X_i^2 E[Y_i^2|X_i]] = E_X[X_i^2 M_{Y|X}(2)]$$
$$= E_X[X_i^2 \exp\{2X_i\gamma + 2\sigma^2\}] = \exp(2\sigma^2)E_X[X_i^2\exp(2X_i\gamma)]$$

$$E_X[X_i^2\exp(2X_i\gamma)] = \int_{-\infty}^{\infty} X_i^2\exp(2X_i\gamma)\frac{1}{\sqrt{2\pi}}\exp\{-\tfrac{1}{2}X_i^2\}dx$$
$$= \int_{-\infty}^{\infty}\frac{X_i^2}{\sqrt{2\pi}}\exp\{-\tfrac{1}{2}(X_i^2-4X_i\gamma-4\gamma^2)+2\gamma^2\} = \exp\{2\gamma^2\}\underbrace{\int_{-\infty}^{\infty}\frac{X_i^2}{\sqrt{2\pi}}\exp\{-\tfrac{1}{2}(X_i-2\gamma)^2\}}_{E[X_i^2] \text{ where } X_i\sim N(2\gamma,1)}$$
$$= \exp\{2\gamma^2\}(1+4\gamma^2)$$

$\therefore E[X_i^2Y_i^2] = \exp\{2(\sigma^2+\gamma^2)\}(1+4\gamma^2)$

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

$$\therefore \sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, (1-4\gamma^2)\exp\{2(\sigma^2+\gamma^2)\} - \gamma^2\exp\{(\sigma^2+\gamma^2)\})$$

we found in (c)

$$\sqrt{n}(\beta^*-\beta) \xrightarrow{d} N(0, \sigma^2(1+\gamma^2)^2 + \frac{\gamma^2\sigma^4}{4})$$

Thus the ARE of β^* and $\hat{\beta}$ is

$$ARE(\beta^*, \hat{\beta}) = \frac{\sqrt{n}[(1-4\gamma^2)\exp\{2(\sigma^2+\gamma^2)\} - \gamma^2\exp\{\sigma^2+\gamma^2\}]}{\sqrt{n}[\sigma^2(1+\gamma^2)^2 + \frac{\gamma^2\sigma^4}{2}]\exp\{\sigma^2+\gamma^2\}}$$
$$= \frac{(1-4\gamma^2)\exp\{2(\sigma^2+\gamma^2)\} - \gamma^2\exp\{\sigma^2+\gamma^2\}}{(\sigma^2(1+\gamma^2)^2 + \frac{\gamma^2\sigma^4}{2})\exp\{\sigma^2+\gamma^2\}}$$

Show strictly $> 1 \Rightarrow \hat{\beta}$ more efficient than β^*

1.(e) We want to find the best function to minimize $E[\{Y-g(x)\}^2]$, what is the optimal $g(x)$ in terms of (γ, σ^2)? Hint: consider minimization conditional on X.

$$E[\{Y-g(x)\}^2 | X] = E[Y^2 - 2g(X)Y + g^2(X) | X] = E[Y^2 | X] - 2g(X)E[Y|X] + g^2(X)$$

$$\frac{\partial}{\partial g(x)} = -2E[Y|X] + 2g(x) \overset{set}{=} 0$$

$$\Rightarrow g(x) = E[Y|X] = \exp\{\gamma x + \frac{\sigma^2}{2}\}$$ as found in (c).

$\therefore$ The best function to minimize $E[\{Y-g(x)\}^2]$ is $g(x) = \exp\{\gamma x + \frac{\sigma^2}{2}\}$

look @ Kosorok session page
trick w/ minimizing ~~fom~~ arbitrary functions.

1.(c) side-note

$E[xe^{\gamma x}]e^{\sigma^2/2}$

this trick works w/ mgfs & cfs.

trick pg.

2019 Theory 1

Kosorok session 7/16/20

$E[e^{tx}] = e^{t^2/2}$

$\frac{d}{dt}E[e^{tx}] = \frac{d}{dt}e^{t^2/2}$

$E[xe^{tx}] = te^{t^2/2}$

$\frac{d}{dt}E[xe^{tx}] = \frac{d}{dt}(te^{t^2/2})$

$E[x^2e^{tx}] = (1+t^2)e^{t^2/2}$

Thus $E[xe^{\gamma x}]e^{\sigma^2/2} = \gamma e^{\gamma^2/2}e^{\sigma^2/2}$

$= \gamma \exp\left\{\frac{\gamma^2+\sigma^2}{2}\right\}$

by Fubini, because "things exist" and are positive
all the necessary moments exist
- when integrals don't converge, the order change can't happen.

jointly measureable

1(e). $E[(Y-g(x)-t\,h(x))^2]$ minimize WRT t.

$\frac{\partial}{\partial t}E[(Y-g(x)-th(x))^2] = -2E[(Y-g(x)-th(x))h(x)] = 0\ \forall t$, chose a convenient t,

$\frac{\partial}{\partial t}[-2E[(Y-g(x)-t(h(x)))h(x)]] = -2E[h^2(x)] > 0$ if $h(x) > 0$

$\therefore$ for any choice of $h()$, generally speaking moving around the parameter space in a non-trivial way, the 2nd deriv is pos.

set $t=0$ thus solving for $h(x)$ a.s. will be a convex $\Rightarrow$ score soln is the minimum.

$E[(Y-g(x))h(x)] = 0 \quad \forall\ h(x) \neq 0$

$E[(E[Y|x]-g(x))h(x)] = 0$ WTFind $g(x)$ s.t. $\supset = 0\ \forall\ h(x) > 0$

let $h(x) = E[Y|X]-g(x)$

$E[(E[Y|X]-g(x))^2] = 0$ is only true if $g(x) = E[Y|X]$

t-space

if you can solve the score, we will have a universal minimum

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.

2019 Theory I

2. Let $X_1 \cdots X_n \overset{iid}{\sim} f(x) = \frac{1}{\theta} e^{(a-x)/\theta} I(x > a)$ where $\theta > 0$

1a) When a is known, derive the UMP test of size α for testing $H_0: \theta \geq \theta_0$ vs. $H_1: \theta < \theta_0$ where θ_0 is a known constant.

$$f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{(a-x_i)/\theta} I(x_i > a) = \theta^{-n} \exp\left\{\sum_{i=1}^{n} \frac{a - x_i}{\theta}\right\} I(x_{(1)} > a)$$

$$= \theta^{-n} \exp\left\{\frac{1}{\theta}\left(an - \sum_{i=1}^{n} x_i\right)\right\} I(x_{(1)} > a) = \theta^{-n} \exp\left\{\frac{an}{\theta} - \frac{\sum_i x_i}{\theta}\right\} I(x_{(1)} > a)$$

$\therefore$ we have an exponential family where $c(\theta) = \theta^{-n} \exp\left(\frac{an}{\theta}\right)$, $h(x) = I(x_{(1)} > a)$, $Q(\theta) = -\frac{1}{\theta}$, $T(x) = \sum_{i=1}^{n} x_i$

$\therefore$ we know the UMP test is of the form

$$\phi(x) = \begin{cases} 1 & T(x) < c \\ \cancel{\gamma} & \cancel{T(x) = c} \\ 0 & T(x) < c \end{cases} \quad \text{where } E_{\theta_0}[\phi(x)] = \alpha$$

$$\alpha = E_{\theta_0}[\phi(x)] = P(T(x) < c) + \cancel{\gamma P(T(x) = c)}$$

$$= P\left(\sum_i x_i < c\right) + \cancel{\gamma P\left(\sum x_i = c\right)} = P\left(\sum_{i=1}^{n} x_i - an < c - an\right)$$

$$\boxed{x - a \sim \exp(\theta_0) \Rightarrow \sum_{i=1}^{n}(x-a) = \sum_{i=1}^{n} x_i - an \sim \text{Gamma}(n, \theta_0)}$$

where x represents $\sum x_i - an$

$$\alpha = \int_0^{c-an} \frac{1}{\Gamma(n)\theta_0^n} x^{n-1} e^{-x/\theta_0} \, dx$$

$$\therefore \; c = F^{-1}(\alpha) + an$$

2. 1(b) When a is known, derive the asymptotic distribution of the MLE of θ. | 2019 Theory I

$$L(\theta) = \theta^{-n} \exp\left\{ \sum_{i=1}^{n} (a-x_i)/\theta \right\} I\{x_{(1)} > a\}$$

$$\ell(\theta) \propto -n\log\theta + \frac{1}{\theta}\sum(a-x_i)$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = \frac{n}{\theta^2} + \frac{2\sum(a-x_i)}{\theta^3}$$

$$= \frac{n\theta + 2\sum(a-x_i)}{\theta^3} < 0$$ <0 as $x_i > a$ $\therefore$ we have a maximum

$$\frac{\partial \ell}{\partial \theta} = -\frac{n}{\theta} - \frac{\sum(a-x_i)}{\theta^2} \overset{set}{=} 0$$

$$\frac{n}{\theta} = \frac{\sum(x_i-a)}{\theta^2} \Rightarrow \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i-a)$$ $x_i - a \sim \exp(\theta)$

with finite second moment $\sum(x_i-a) \sim \text{Gamma}(n,\theta)$

$\therefore$ by CLT $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i-a)$ converges in distribution s.t.

$$\sqrt{n}\left(\hat{\theta}-\theta\right) \xrightarrow{d} N(0,\theta^2)$$ as $E[\exp(\theta)] = \theta$, $var[\exp(\theta)] = \theta^2$

$$\therefore \frac{\sqrt{n}(\hat{\theta}-\theta)}{\theta} \xrightarrow{d} N(0,1) \equiv Z$$

In the rest of the questions we assume $a = \theta$
the density is $f_1(x) = \theta^{-1} e^{(\theta - x)/\theta} I(x > \theta)$

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

$$f(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\theta} e^{(\theta - x_i)/\theta} I(x_i > \theta) = \frac{1}{\theta^n} \exp\left\{\sum_{i=1}^{n} \frac{\theta - x_i}{\theta}\right\} I(x_i > \theta) = \theta^{-n} \exp\{n - \Sigma x_i/\theta\} I(X_{(1)} > \theta)$$

$f_1(x)$ is the density for a shifted exponential, a member of the location-scale family

Thus $X - \theta \sim \exp(\theta) \Rightarrow \frac{X-\theta}{\theta} \sim \exp(1)$

$$\sum_{i=1}^{n} \frac{X-\theta}{\theta} \sim \text{Gamma}(n, 1) \qquad \frac{1}{n}\sum_{i=1}^{n} \frac{X-\theta}{\theta} \sim \text{Gamma}(n, 1/n)$$

$$\therefore \frac{1}{n}\sum_{i=1}^{n} \frac{X-\theta}{\theta} = \frac{1}{n\theta}(\Sigma X_i - n\theta) = \frac{\bar{X}}{\theta} - 1 \sim \text{Gamma}(n, 1/n)$$

$\therefore \frac{\bar{X}}{\theta} \sim \text{Gamma}(n, 1/n, -1)$ a shifted exponential whose distribution is $\perp\!\!\!\perp \theta$ $\therefore \frac{\bar{X}}{\theta}$ is a pivotal quantity.

To work w/ $X_{(1)}$ we want the CDF of X.

$$F_X(x) = \int_\theta^x \frac{1}{\theta} e^{-(t-\theta)/\theta} dt = \frac{1}{\theta}(-\theta) e^{-(t-\theta)/\theta}\Big|_{t=\theta}^{x} = -e^{1 - t/\theta}\Big|_{t=\theta}^{x} = -e^{\frac{\theta - x}{\theta}} + 1$$
$$= 1 - e^{(\theta - x)/\theta}$$

$$P(X_{(1)} < z) = 1 - P(X_{(1)} > z) = 1 - P(X_1 \cdots X_n > z) = 1 - P(X_1 > z)^n = 1 - (1 - P(X_1 < z))^n$$
$$= 1 - (1 - 1 + e^{\theta - z/\theta})^n = 1 - e^{\frac{n}{\theta}(\theta - z)}$$

$$\therefore X_{(1)} - \theta \sim \exp(\theta/n) \Rightarrow \frac{X_{(1)} - \theta}{\theta} \sim \exp(1/n) \Rightarrow \frac{X_{(1)}}{\theta} - 1 \sim \exp(1/n)$$

$\Rightarrow \frac{X_{(1)}}{\theta} \sim$ shifted exponential $(1/n, -1)$, whose distribution is $\perp\!\!\!\perp \theta$

$\therefore \frac{X_{(1)}}{\theta}$ is a pivotal quantity.

21(d). Obtain two confidence intervals w/ confidence coefficient $1-\alpha$ for θ based on the two pivotal quantities found in (c). 2019 Theory I

① Using $\frac{\bar{X}}{\theta}$: we know $\frac{\bar{X}}{\theta}-1 \sim \text{Gamma}(n, 1/n)$

let $\Gamma_{\alpha/2}(n, 1/n)$, $\Gamma_{1-\alpha/2}(n, 1/n)$ be the $\alpha/2^{th}$ & $1-\alpha/2^{th}$ percentile of the Gamma$(n, 1/n)$ distribution, respectively. Then

$$1-\alpha = P\left(\Gamma_{\alpha/2}(n, 1/n) < \frac{\bar{X}}{\theta}-1 < \Gamma_{1-\alpha/2}(n, 1/n)\right)$$

$$= P\left(\Gamma_{\alpha/2}(\,)+1 < \frac{\bar{X}}{\theta} < \Gamma_{1-\alpha/2}(\,)+1\right)$$

$$= P\left(\frac{\bar{X}}{\Gamma_{1-\frac{\alpha}{2}}(\,)+1} < \theta < \frac{\bar{X}}{\Gamma_{\alpha/2}(\,)+1}\right)$$

$\therefore$ a $1-\alpha(100)\%$ CI for θ is $\left\{\theta : \frac{\bar{X}}{\Gamma_{1-\frac{\alpha}{2}}(n, 1/n)+1} < \theta < \frac{\bar{X}}{\Gamma_{\alpha/2}(n, 1/n)+1}\right\}$

② Using $\frac{X_{(1)}}{\theta}$: we know $\frac{X_{(1)}}{\theta}-1 \sim \exp(1/n)$

let $E_{\alpha/2}(1/n)$, $E_{1-\alpha/2}(1/n)$ be the $\frac{\alpha}{2}^{th}$ & $1-\frac{\alpha}{2}^{th}$ percentile of the $\exp(1/n)$ distribution, respectively. Then

$$1-\alpha = P\left(E_{\alpha/2}(1/n) < \frac{X_{(1)}}{\theta}-1 < E_{1-\frac{\alpha}{2}}(1/n)\right)$$

$$= P\left(\frac{X_{(1)}}{E_{1-\frac{\alpha}{2}}(\cdot)+1} < \theta < \frac{X_{(1)}}{E_{\alpha/2}(\cdot)+1}\right)$$

$\therefore$ a $1-\alpha(100)\%$ CI for θ is $\left\{\theta : \frac{X_{(1)}}{E_{1-\frac{\alpha}{2}}(\frac{1}{n})+1} < \theta < \frac{X_{(1)}}{E_{\alpha/2}(\frac{1}{n})+1}\right\}$

2(e) When n is sufficiently large, which of the two CIs have shorter length? Justify your answer. 2019 Theory I

length ① $\dfrac{\bar{X}}{1+\Gamma_{\frac{\alpha}{2}}(n,\frac{1}{n})} - \dfrac{\bar{X}}{1+\Gamma_{1-\frac{\alpha}{2}}(n,\frac{1}{n})} = \bar{X}\left(\dfrac{(1+\Gamma_{1-\frac{\alpha}{2}}(n,1/n)) - (1+\Gamma_{\frac{\alpha}{2}}(n,1/n))}{(1+\Gamma_{1-\frac{\alpha}{2}}(n,1/n))(1+\Gamma_{\frac{\alpha}{2}}(n,1/n))}\right)$

$$= \bar{X}\underbrace{\left[\dfrac{\Gamma_{1-\frac{\alpha}{2}}(n,1/n) - \Gamma_{\frac{\alpha}{2}}(n,1/n)}{(1+\Gamma_{1-\frac{\alpha}{2}}(n,1/n))(1+\Gamma_{\frac{\alpha}{2}}(n,1/n))}\right]}_{c} \xrightarrow{d} N(\overset{\neq 0}{\wedge}, \overset{0}{\wedge}) \quad \text{as } n\to\infty$$

as c is a constant & $\bar{X} \xrightarrow{d} N(\cdot,\cdot)$ by LLN.

length ② $\dfrac{X_{(1)}}{1+E_{\frac{\alpha}{2}}(1/n)} - \dfrac{X_{(1)}}{1+E_{1-\frac{\alpha}{2}}(\frac{1}{n})}$

$L = E_{\frac{\alpha}{2}}(1/n) \Rightarrow \frac{\alpha}{2} = 1-e^{-nL}$

$1-\frac{\alpha}{2} = e^{-nL}$

$\log(1-\frac{\alpha}{2}) = -nL$

$L = -\dfrac{\log(1-\frac{\alpha}{2})}{n}$

similarly $1-\frac{\alpha}{2} = 1-e^{-nU}$

$\log(\alpha/2) = -nU$

$-\frac{1}{n}\log(\alpha/2) = U = E_{1-\alpha/2}(1/n)$

$\therefore$ ② $= \dfrac{X_{(1)}}{1-\frac{1}{n}\log(1-\frac{\alpha}{2})} - \dfrac{X_{(1)}}{1-\frac{1}{n}\log(\alpha/2)} = X_{(1)}\left(\dfrac{1-1/n\log(\alpha/2) - 1 + \frac{1}{n}\log(1-\alpha/2)}{(1-\frac{1}{n}\log(1-\frac{\alpha}{2}))(1-\frac{1}{n}\log(\frac{\alpha}{2}))}\right)$

$$= X_{(1)}\left(\dfrac{\overset{n\to\infty}{\overbrace{\tfrac{1}{n}}^{\to 0}}(\log(1-\frac{\alpha}{2}) - \log(\frac{\alpha}{2}))}{\underbrace{(1-\frac{1}{n}\log(1-\frac{\alpha}{2}))}_{\xrightarrow{n\to\infty} 1}\underbrace{(1-\frac{1}{n}\log(\frac{\alpha}{2}))}_{\xrightarrow{n\to\infty} 1}}\right) \to 0$$

$\therefore$ for sufficiently large n, the length of the CI based in the minimum, $X_{(1)}$ will go towards 0 & result in a smaller CI than the CI based in $\bar{X}$.