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

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

   (b) From (1), we estimate (α^*, β^*) as

   $$\widehat{\alpha} = n^{-1} \sum_{i=1}^{n} Y_i, \quad \widehat{\beta} = n^{-1} \sum_{i=1}^{n} X_i Y_i.$$

   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.

1) $X \sim N(0,1)$ $\quad E[X^2Y^2] < \infty$

Y is outcome $\quad$ Observe $(X_1, Y_1), \ldots, (X_n, Y_n)$

<u>Goal</u>: Obtain best prediction of Y given X for a future subject

(a) Consider linear function, $\alpha + \beta X$, to minimize squared error loss: $E\left[\{Y - (\alpha + \beta X)\}^2\right]$, where the expectation is WRT joint dist. of (Y, X)

Show optimal solution is $\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} E(Y) \\ E(XY) \end{pmatrix}$

$$E_{(Y,X)}\left[Y^2 - 2Y(\alpha + \beta X) + \alpha^2 + 2\alpha\beta X + \beta^2 X^2\right] = E_{(Y|X)}\left[Y^2 - 2Y\alpha - 2\beta XY + \alpha^2 + 2\alpha\beta X + \beta^2 X^2\right]$$

Now, to find optimal α and β, we take derivatives WRT these variables and set $= 0$ to solve:

$$\frac{d E_{(Y,X)}}{d\alpha} = E_{(Y,X)}\left[-2Y + 2\alpha + 2\beta X\right] \overset{set}{=} 0$$

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

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

$\hookrightarrow$ Note: $E_{Y,X}[X] = E_X\left[E_{Y|X}[X|X]\right] = E_X[X] = 0$, since $X \sim N(0,1)$

$$\Leftrightarrow \quad \alpha^* = E[Y]$$

$$\frac{d E_{(Y,X)}}{d\beta} = E_{(Y,X)}\left[-2XY + 2\alpha X + 2\beta X^2\right] \overset{set}{=} 0$$

$$\Leftrightarrow \quad E_{(Y,X)}[XY] = \alpha \underbrace{E_{(Y,X)}[X]}_{0} + \beta^* E_{(Y,X)}[X^2]$$

$\hookrightarrow E_{Y,X}[X^2] = E_X[X^2]$, and $X^2 \sim \chi^2_1$, so $E[X^2] = 1$

$$\Leftrightarrow \quad \beta^* = \frac{E[XY] - 0}{1} = E[XY]$$

Thus, $\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} E(Y) \\ E(XY) \end{pmatrix}$

(b) From (1), 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 estimator after proper normalization.

By CLT, since we have finite second moments, we know

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - E(Y_i)\right) \xrightarrow[d]{} N\left(0, Var(Y_i)\right)$$

From (a), we have $E(Y) = \alpha^*$

$Var(Y) = E(Y^2) - E(Y)^2 = E(Y^2) - \alpha^{*2}$

So $$\sqrt{n}\left(\hat{\alpha} - \alpha^*\right) \xrightarrow[d]{} N\left(0, E(Y^2) - \alpha^{*2}\right)$$

Similarly,

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i Y_i - E(X_i Y_i)\right) \xrightarrow[d]{} N\left(0, Var(X_i Y_i)\right)$$

From (a), $E(XY) = \beta^*$

$Var(XY) = E(X^2Y^2) - E(XY)^2 = E(X^2Y^2) - \beta^{*2} < \infty$ ✓

So, by properties of multivariate Normal,

we have:

$$\sqrt{n}\begin{pmatrix} \hat{\alpha} - \alpha^* \\ \hat{\beta} - \beta^* \end{pmatrix} \xrightarrow[d]{} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} E(Y^2) - \alpha^{*2} & Cov(Y, XY) \\ Cov(Y, XY) & E(X^2Y^2) - \beta^{*2} \end{pmatrix}\right)$$

Note that $Cov(Y, XY) = E[XY^2] - E[Y]E[XY] = E[XY^2] - \alpha^*\beta^*$

Thus,

$$\sqrt{n}\begin{pmatrix} \hat{\alpha} - \alpha^* \\ \hat{\beta} - \beta^* \end{pmatrix} \xrightarrow[d]{} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} E(Y^2) - \alpha^{*2} & E(XY^2) - \alpha^*\beta^* \\ E(XY^2) - \alpha^*\beta^* & E(X^2Y^2) - \beta^{*2} \end{pmatrix}\right)$$

(c) Now $Y|X$ is logNormal: $\log Y = \gamma X + N(0,\sigma^2)$

→ Obtain MLEs of $\alpha^* = E(X)$ and $\beta^* = E(XY)$ and derive their asymptotic distribution

- Let $Z \doteq \log Y$; so $Z|X = \log Y|X \sim N(\gamma X, \sigma^2)$

  ↳ since $Z = \log Y$; $\frac{dz}{dy} = \frac{1}{y}$, so $f_{Y|X}(y) = f_{Z|X}(\log y) \cdot \frac{1}{y}$

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

- We know $X \sim N(0,1)$

- To get $f_{Y,X}(y,x)$, the joint distribution, we multiply $f_{Y|X} \cdot f_X$:

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

Now, we want to find $\alpha^* = E_{(Y,X)}[Y]$

→ Option 1: Brute force it using joint distribution [this proved to be very difficult for me!]

→ Option 2: Note $E_{(Y,X)}[Y] = E_X\left[E_{Y|X}[Y|X]\right]$

Above, we set $Z = \log(Y)$, so $Z|X \sim N(\gamma X, \sigma^2)$

We know the MGF of Z is: $E[e^{Zt}] = \exp\left\{\gamma X t + \frac{1}{2}\sigma^2 t^2\right\}$

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

So, given X, we can find the moments of Y by:

$$E[Y^t|X] = \exp\left\{\gamma X t + \frac{1}{2}\sigma^2 t^2\right\}$$

$$\Rightarrow E[Y|X] = \exp\left\{\gamma X + \frac{1}{2}\sigma^2\right\}$$

Now, we need $E_X\left[\exp\{\gamma X + \frac{1}{2}\sigma^2\}\right]$ ⟶

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

$$= \exp\{\tfrac{1}{2}\sigma^2+\underline{\underline{\tfrac{1}{2}\gamma^2}}\}\underbrace{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\{-\tfrac{1}{2}\underbrace{(x^2-2\gamma x+\underline{\underline{\gamma^2}})}_{(x-\gamma)^2}\}\,dx}_{\text{PDF of } N(\gamma,1)\ \to 1}$$

So, $\alpha^* = E[Y] = \exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\}$

Next, to find $\beta^* = E_{(Y,X)}[XY]$

$$E_{(Y,X)}[XY] = E_X\left[E_{Y|X}[XY|X]\right] = E_X\left[X\,E[Y|X]\right] = E_X\left[X\exp\{\gamma x+\tfrac{1}{2}\sigma^2\}\right]$$

$$= \exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\}\underbrace{\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}}\exp\{-\tfrac{1}{2}(x-\gamma)^2\}\,dx}_{\text{This is just the expectation of a } N(\gamma,1) \text{ distributed variable } =\gamma}$$

So, $\beta^* = E[XY] = \gamma\exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\}$

Now that we have $\alpha^* = \exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\}$ and $\beta^* = \gamma\exp\{\tfrac{1}{2}(\sigma^2+\gamma^2)\}$, we can find the MLEs of σ^2 and γ, and use invariance of MLEs to find the MLEs of α^* and β^* (call them $\hat{\alpha}^*$ and $\hat{\beta}^*$).

Our joint Likelihood is: $L = \prod_{i=1}^{n}\frac{1}{y_i}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(\log y_i-\gamma x_i)^2\}\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}x_i^2\}$

$\Rightarrow$ log-Likelihood is: $\ell\text{-}L \propto \sum_{i=1}^{n} -\log(y_i) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\log y_i-\gamma x_i)^2 - \frac{1}{2}x_i^2$

Then, $\frac{d\ell\text{-}L}{d\gamma} = \sum_{i=1}^{n} -\frac{1}{\sigma^2}(\log y_i-\gamma x_i)(-x_i) \overset{\text{set}}{=} 0$

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

$\to$

and $\frac{\partial \ell\text{-}L}{\partial \sigma^2} = \sum_{i=1}^{n} -\frac{1}{2\sigma^2} + \frac{(\log Y_i - \gamma X_i)^2}{2\sigma^4} \overset{set}{=} 0$

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

Plugging in $\hat\gamma \Rightarrow \hat\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log Y_i - \left(\frac{\sum_{i=1}^{n} X_i \log Y_i}{\sum_{i=1}^{n} X_i^2}\right)X_i\right)^2$

So, by invariance of MLE, we have:

$$\hat\alpha^* = \exp\left\{\tfrac{1}{2}(\hat\sigma^2 + \hat\gamma^2)\right\} \text{ and } \hat\beta^* = \hat\gamma \exp\left\{\tfrac{1}{2}(\hat\sigma^2+\hat\gamma^2)\right\}$$

where $\hat\gamma = \frac{\sum_{i=1}^{n} X_i \log Y_i}{\sum_{i=1}^{n} X_i^2}$ and $\hat\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\log Y_i - \hat\gamma X_i)^2$

Now, assuming regularity conditions hold, then by properties of MLE, we know

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

Find components of $I(\theta)$:

$$\frac{\partial^2 \ell\text{-}L}{\partial\gamma^2} = -\sum_{i=1}^{n}\frac{X_i^2}{\sigma^2} \Longrightarrow -\frac{1}{n}E\left[-\sum_{i=1}^{n} X_i^2/\sigma^2\right] = \frac{nE[X_i^2]}{n\sigma^2} = \frac{1}{\sigma^2}$$

$$\frac{\partial^2 \ell\text{-}L}{\partial\hat\sigma^{2^2}} = \sum_{i=1}^{n}\frac{1}{2\sigma^4} - \frac{4\sigma^2(\log Y_i - \gamma X_i)^2}{4\sigma^8} = \sum_{i=1}^{n}\frac{1}{2\sigma^4} - \frac{(\log Y_i - \gamma X_i)^2}{\sigma^6}$$

$$\Rightarrow -\frac{1}{n}E\left[\sum_{i=1}^{n}\frac{1}{2\sigma^4} - \frac{(\log Y_i - \gamma X_i)^2}{\sigma^6}\right] = \frac{-1}{2\sigma^4} + \frac{1}{n}\frac{1}{\sigma^6}\sum_{i=1}^{n}\underbrace{E[(\log Y_i - \gamma X_i)^2]}_{E[(Z_i - \gamma X_i)^2]}$$

$$= -\frac{1}{2\sigma^4} + \frac{1}{n}\frac{1}{\sigma^6}\cdot n\sigma^2$$

$Z - \gamma X \sim N(0,\sigma^2)$ so
$E[(Z-\gamma X)^2] = Var(Z-\gamma X) + E(Z-\gamma X)^2$
$= \sigma^2 + 0$

$$= -\frac{1}{2\sigma^4} + \frac{1}{\sigma^4} = \frac{1}{\sigma^4}$$

$$\frac{\partial^2 \ell\text{-}L}{\partial\gamma\partial\sigma} = \sum_{i=1}^{n}\frac{-X_i(\log Y_i - \gamma X_i)}{\sigma^4}\ ; \ E\left[X_i(\log Y_i - \gamma X_i)\right] = E\left[X_i \underbrace{E(\log Y_i - \gamma X_i | X_i)}_{0}\right]$$

b/c $Z - \gamma X_i \sim N(0,1)$

So, $-\frac{1}{n}E\left[\frac{\partial^2 \ell}{\partial\gamma\partial\sigma}\right] = 0$

$\longrightarrow$ 5

So, we have: $\sqrt{n}\begin{pmatrix}\hat{\gamma}-\gamma\\ \hat{\sigma}^2-\sigma\end{pmatrix} \xrightarrow[d]{} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^4\end{pmatrix}\right)$

Since $I(\theta)^{-1} = \begin{pmatrix}1/\sigma^2 & 0\\ 0 & 1/\sigma^4\end{pmatrix}^{-1} = \begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^4\end{pmatrix}$

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

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

Then, by Delta-Method:

$$\sqrt{n}\begin{pmatrix}\hat{\alpha}^* - \alpha^*\\ \hat{\beta}^* - \beta^*\end{pmatrix} \xrightarrow[d]{} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \underbrace{\nabla g\begin{pmatrix}\gamma\\ \sigma^2\end{pmatrix} I(\theta)^{-1}\, \nabla g\begin{pmatrix}\gamma\\ \sigma^2\end{pmatrix}'}_{\text{call this } \Sigma}\right)$$

where

$$\Sigma = \exp\left\{\tfrac{1}{2}(\gamma^2+\sigma^2)\right\}^2 \begin{pmatrix}\gamma & 1/2\\ \gamma^2+\gamma & \gamma/2\end{pmatrix}\begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^4\end{pmatrix}\begin{pmatrix}\gamma & \gamma^2+\gamma\\ 1/2 & \gamma/2\end{pmatrix}$$

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

and $\hat{\alpha}^* = \exp\{\frac{1}{2}(\hat{\sigma}^2+\hat{\gamma}^2)\}$, $\hat{\beta}^* = \hat{\gamma}\exp\{\frac{1}{2}(\hat{\sigma}^2+\hat{\gamma}^2)\}$ (as shown on previous page).

(d) Calculate asymptotic relative efficiency between β^* and $\hat{\beta}$

$$\beta^* = \hat{\gamma} \exp\left\{ \tfrac{1}{2} (\hat{\sigma}^2 + \hat{\gamma}^2) \right\} \quad ; \quad \hat{\beta} = \tfrac{1}{n} \sum_{i=1}^{n} X_i Y_i$$

We know $\sqrt{n}\,(\beta^* - \beta) \xrightarrow{d} N\left(0, \left(\sigma^2(\gamma^2+\gamma)^2 + \frac{\sigma^4\gamma^2}{4}\right)\exp(\gamma^2+\sigma^2)\right)$ from (c)

From (b), we know $\sqrt{n}\,(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, E[X^2Y^2] - E[XY]^2\right)$

Now that we have a distribution for Y, we can find $E[X^2Y^2]$ & $E[XY]$:

→ We actually found $E[XY]$ in part (c);

$$E[XY] = \gamma \exp\left\{ \tfrac{1}{2}(\sigma^2 + \gamma^2) \right\}$$

So $E[XY]^2 = \gamma^2 \exp(\sigma^2 + \gamma^2)$

→ $E[X^2Y^2] = E_X\left[X^2 E[Y^2|X]\right]$

We know $E[Y^t|X] = \exp\left\{ \gamma X t + \tfrac{1}{2}\sigma^2 t^2 \right\}$

So $E[Y^2|X] = \exp\left\{ 2\gamma X + 2\sigma^2 \right\}$

Then $E_X\left[X^2 \exp(2\gamma X + 2\sigma^2)\right] = \exp(2\sigma^2)\, E_X\left[X^2 e^{X(2\gamma)}\right]$

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

$$= \exp(2\sigma^2 + 2\gamma^2) \underbrace{\int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} \exp\left\{ -\tfrac{1}{2}(x - 2\gamma)^2 \right\} dx}_{\text{Second moment of variable } N(2\gamma,1)}$$

$\Rightarrow 1 + (2\gamma)^2 = 1 + 4\gamma^2$

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

So, putting this together, $\sqrt{n}\,(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \gamma^2\exp(\sigma^2+\gamma^2) - (1+4\gamma^2)\exp(2\sigma^2+2\gamma^2)\right)$

Then, finally:

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

(Not sure how far to go...)

(e) Minimize $E[\{Y-g(X)\}^2]$, what is optimal $g(X)$ in terms of (γ, σ^2)?

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

To minimize WRT $g(X)$, take derivative WRT $g(X)$:

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

$$E[Y|X] = E[g(X)|X]$$

We know from (c) that $E[Y|X] = \exp\{\gamma X + \frac{1}{2}\sigma^2\}$.

So, the optimal $g(X)$ in terms of (γ, σ^2) is $\exp\{\gamma X + \frac{1}{2}\sigma^2\}$

$$E[(Y-g(X)-th(X))^2]$$

$$\frac{d}{dt} = -2E[(Y-g(X)-th(X))h(X)]$$

$$\frac{d^2}{dt^2} = 2E[h^2(X)] > 0 \Rightarrow$$ so we are at a minimum if we set $\frac{d}{dt} = 0$

↳ Now, setting $t=0$ gives score: $E[(Y-g(X))h(X)] = 0 \quad \forall h \neq 0$

- Now condition on X:

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

Since true $\forall h(X)$, let $h(X) = E[Y|X] - g(X)$

$$\Rightarrow E[(E(Y|X) - g(X))^2] = 0$$

Which is only true if $g(X) = E(Y|X)$