1. Consider the following model and assumptions. Assume that β_1 and β_2 are fixed unknown constants, and that $\{x_1, \ldots, x_n\}$ are fixed known constants with $x_i \neq x_j$ for at least some $i \neq j$.

$$Y_i = \beta_1 + \beta_2 \exp(x_i) + \epsilon_i, \quad i = 1, \cdots, n, \quad \text{where} \quad \epsilon_i \overset{iid}{\sim} N(0, \sigma^2).$$

You may assume and use any results you know from PSTAT 127 or pre-requisite courses to answer this question.

(a) Write down the distribution of Y_i.
Your answer must include the distribution name, and the mean and variance.

(b) Are $Y_1, \ldots, Y_n$ independently and identically distributed (i.e., are they iid)? (Answer Yes or No.)

(c) Is this a linear model in the sense of PSTAT 126? Why or why not?

(d) Based on your answer to part (1a), write out the likelihood of β_0 and β_1 based on observations $\{y_1, \ldots, y_n\}$. Your answer will be a function of β_0, β_1, $y_1, \ldots, y_n$ and $x_1, \ldots, x_n$.

---

a) for $y_i = \beta_1 + \beta_2 \exp(x_i) + \varepsilon_i$, we are given $\varepsilon_i \sim N(0, \sigma^2)$
under this assumption $X_i^T \beta$ is an unknown constant.

$$X^T = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \\ \vdots & \vdots \\ X_{n1} & X_{2n} \end{bmatrix} = \begin{bmatrix} 1 & \exp(x_1) \\ 1 & \exp(x_2) \\ \vdots & \vdots \\ 1 & \exp(x_n) \end{bmatrix} = \begin{bmatrix} 1 & \exp(X_i) \end{bmatrix}^T \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad \varepsilon_i = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

so $y_i \sim N([1 \ \exp(x_i)]^T \beta, \sigma^2) = N(\beta_1 + \beta_2 \exp(x_i), \sigma^2)$ w/ mean $= \beta_1 + \beta_2 \exp(x_i)$ and variance σ^2. The variance stays the same because the unknown constants only shift the distribution. Nothing related to spread changes.

b) NO, y_is are not independent and identically distributed because the means are different for every i. However they are independent.

c) In the sense of PSTAT 126, $y_i = \beta_1 + \beta_2 \exp(X_i) + \varepsilon_i$ is a linear model. This is mainly because the predictors are linear in the Betas. In this specific case, since β_2 is not included in exp, it is linear in the Betas.

d) Likelihood $= \prod_{i=1}^{n} f(x)$ by independence

*pdf for normal dist: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$ w/ $E(X) = \mu$, $Var(X) = \sigma^2$

$$L(y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-1}{2\sigma^2}(y_i - (\beta_1 + \beta_2 \exp(x_i))^2\right)$$

$$L(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \prod_{i=1}^{n} \exp\left(\frac{-1}{2\sigma^2}(y_i - \beta_1 - \beta_2 \exp(x_i))^2\right)$$

$$L(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i - \beta_1 - \beta_2 \exp\left(\sum_{i=1}^{n} x_i\right)\right)^2\right]$$

2. Review of estimation concepts from 120A/B: Mean square error of a point estimator for a parameter θ.

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from some pdf/pmf $f_Y(y;\theta)$. Let W be a point estimator $W = h(Y_1, Y_2, \ldots, Y_n)$ for θ. The bias of W as a point estimator for θ is defined as

$$\text{Bias}_\theta(W) \quad = \quad E(W) - \theta.$$

The mean square error of a point estimator W for θ is defined as

$$MSE_\theta(W) \quad = \quad E\left[(\ W - \theta\ )^2\right]$$

Using properties of expected values, and the definition of variance from PSTAT 120A/B, show that

$$MSE_\theta(W) \quad = \quad \text{Var}(W) + [\ \text{Bias}_\theta(W)\ ]^2\,.$$

* note $Var(W) = E(W^2) - [E[W]]^2$ (page 4 of Cov Corr Summary)

$$MSE(W) = \underbrace{E[W^2] - E[W]^2}_{Var(W)} + \underbrace{[E(W) - \theta]^2}_{Bias^2}$$

$$= E[W^2] - \cancel{E[W]^2} + (\cancel{E[W]^2} - 2\theta E[W] + \theta^2]$$

$$= E[W^2] - 2\theta E[W] + \theta^2$$

$$= E[W^2] - E[2\theta W] + E[\theta^2]$$

$$= E[W^2 - 2\theta W + \theta^2]$$

$$= E[(W - \theta)^2] = MSE(W)$$

3. Suppose that $Y_1, \ldots, Y_n \overset{iid}{\sim} f_Y(y;\theta)$ where

$$f_Y(y;\theta) = \begin{cases} \left(\dfrac{2y}{\theta^2}\right), & \text{if } 0 < y < \theta \\ 0, & \text{otherwise} \end{cases}$$

(a) Find $E(Y_i)$ and $\mathrm{Var}(Y_i)$ for each $i \in \{1, \ldots, n\}$, using any approach you have studied.

(b) Consider the estimator $\hat{\theta} = \overline{Y}$.

- Show that $\hat{\theta}$ is a biased estimator of θ, and find the bias.
- Find the variance of $\hat{\theta}$.
- Find the mean square error of $\hat{\theta}$, when estimating θ, i.e., find $MSE_\theta(\hat{\theta})$.

---

a) $E[y_i] = \int_0^\theta y \frac{2y}{\theta^2}\,dy = \frac{1}{\theta^2}\int_0^\theta 2y^2\,dy = \frac{1}{\theta^2}\left[\frac{2}{3}y^3\Big|_0^\theta\right] = \frac{1}{\theta^2}\left[\frac{2}{3}\theta^3 - \frac{2}{3}(0)^3\right] = \frac{1}{\theta^2}\left[\frac{2}{3}\theta^3\right] = \frac{2\theta}{3}$

$\therefore E[y_i] = \frac{2\theta}{3}$

$var[y_i] = E[y_i^2] - E[y_i]^2$

$E[y_i^2] = \int_0^\theta y^2 \frac{2y}{\theta^2}\,dy = \frac{1}{\theta^2}\int_0^\theta 2y^3\,dy = \frac{1}{\theta^2}\left[\frac{2y^4}{4}\Big|_0^\theta\right] = \frac{1}{\theta^2}\left[\frac{\theta^4}{2} - \frac{0^4}{2}\right] = \frac{1}{\theta^2}\left(\frac{\theta^4}{2}\right) = \frac{\theta^2}{2}$

$var[y_i] = \frac{\theta^2}{2} - \left(\frac{2\theta}{3}\right)^2 = \frac{\theta^2}{2} - \frac{4\theta^2}{9} = \frac{9\theta^2}{18} - \frac{8\theta^2}{18} = \frac{\theta^2}{18}$

$\therefore var(y_i) = \frac{\theta^2}{18}$

b) i) find bias

$Bias = E[\hat{\theta}] - \theta = E[\bar{y}] - \theta = E\left[\frac{\sum y_i}{n}\right] - \theta$

$= \frac{1}{n}\sum E[y_i] - \theta = \frac{1}{n}\left[n\,\frac{2\theta}{3}\right] - \theta = -\frac{\theta}{3}$

$\therefore \hat{\theta}$ is biased

ii) find variance $(\hat{\theta})$

$var(\hat{\theta}) = var(\bar{y}) = var\left(\frac{1}{n}\sum y_i\right) = \left(\frac{1}{n}\right)^2 var\left(\sum y_i\right) = \left(\frac{1}{n}\right)^2 \sum var(y_i)$

$= \left(\frac{1}{n}\right)^2 (n)\left(\frac{\theta^2}{18}\right) = \frac{\theta^2}{18n}$ $\quad \therefore var(\hat{\theta}) = \frac{\theta^2}{18n}$

iii) find $MSE(\hat{\theta})$

$MSE(\hat{\theta}) = var(\hat{\theta}) + Bias(\hat{\theta})^2$

$= \frac{\theta^2}{18n} + \left(-\frac{\theta}{3}\right)^2$

$= \frac{\theta^2}{18n} + \frac{\theta^2}{9} = \frac{\theta^2 + 2n\theta^2}{18n}$

$= \frac{\theta^2(1+2n)}{18n}$

$\therefore MSE(\hat{\theta}) = \frac{\theta^2}{18n} + \frac{\theta^2}{9} = \frac{\theta^2(1+2n)}{18n}$

4. This exercise requires you to review the definitions of certain distributions seen in 120A/B, that also were used in 126:

Suppose $X_1, \ldots, X_6 \stackrel{iid}{\sim} N(3, 16)$, and that $Y_1, Y_2 \stackrel{iid}{\sim} N(0,1)$, and that $X_1, \ldots, X_6$ are independent of Y_1 and Y_2.

Clearly write down the distribution (*including distribution name and the numeric values of any parameters or degrees of freedom if appropriate*) for each of the following random variables.

Provide clear reasoning in each case by stating the results and definitions you are using from pre-requisite classes.

$* \; U = \sum_{i=1}^{n} (a_i X_i + b_i) \sim \text{Normal}\left(\sum_{i=1}^{n} (a_i \mu_i + b_i), \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$

if $Z \sim N(0,1)$, then $Z^2 \sim \chi_1^2$ $*$

if $Z_1, \ldots, Z_n \stackrel{iid}{\sim} N(0,1)$, then $\sum Z_i^2 \sim \chi_n^2$ $*$

(a) $T = \sum_{i=1}^{6} X_i$ $\quad T \sim \text{Normal}(18, 96)$

Using $*$, $T = \sum_{i=1}^{n=6} ((1) X_i + 0)$ $\quad a_i = 1$, $b_i = 0$

so $\mu = \sum_{i=1}^{n} (a_i \mu_i + b_i) = \sum_{i=1}^{6} (1)(3) = (6)(3) = 18$

$\sigma^2 = \sum_{i=1}^{n} (a_i^2 \sigma_i^2) = \sum_{i=1}^{6} (1)^2 (16) = (6)(16) = 96$

(b) $W = \dfrac{\left(\sum_{i=1}^{6} X_i\right)}{4}$ $\quad W \sim \text{Normal}(9/2, 6)$

$W = \frac{1}{4}\left(\sum_{i=1}^{6} X_i\right) = \sum_{i=1}^{6} \frac{1}{4} X_i + 0$ $\quad a_i = 1/4$, $b_i = 0$ $\quad (*)$

$\mu = \sum_{i=1}^{n} \frac{1}{4}\mu_i = \frac{1}{4}(3) + (\frac{1}{4})(3) + \cdots + (\frac{1}{4})(3) = (\frac{1}{4})(3)(6) = 9/2$

$\sigma^2 = \sum_{i=1}^{n} (\frac{1}{4})^2 \sigma_i^2 = (\frac{1}{16})(16) + (\frac{1}{16})(16) + \ldots + (\frac{1}{16})(16)(6) = 6$

(c) $V = 3Y_1 - 2X_3 + 7$ $\quad V \sim \text{Normal}(1, 73)$

$E[V] = E[3Y_1] - E[2X_3] + E[7]$
$= 3E[Y_1] - 2E[X_3] + 7$
$= 3[0] - 2[3] + 7 = 7 - 6 = 1$

$X_1 \perp\!\!\!\perp X_2 \quad X_1 \sim N(\mu_1, \sigma_1^2)$ & $X_2 \sim N(\mu_2, \sigma_2^2)$

$X_1 + X_2 \sim N(\mu_1 + \mu_2, var(X_1) + var(X_2))$ $*$

$var[V] = var[3Y_1] + var[-2X_3] + var(7)$
$= 9 var(Y_1) + 4 var[X_3] + 0$
$= (9)(1) + (4)(16) = 9 + 64 = 73$

(d) $S = (Y_1)^2$ $\quad S \sim \chi_1^2$

By $*$, since $Y_1 \sim \text{Normal}(0,1)$, then

$S = (Y_1)^2 \sim \chi_1^2$

(e) $U = \left(\frac{X_2 - 3}{4}\right)$ $U \sim Normal(0,1)$

$U = \frac{1}{4}X_2 - \frac{3}{4} = \sum_{i=2}^{n=2}\left(\frac{1}{4}X_i + \frac{-3}{4}\right)$ $a_i = \frac{1}{4}$ $b_i = -3/4$ (*)

$\mu = \sum_{i=2}^{2}\left(\left(\frac{1}{4}\right)(\mu_i) + {}^{-3}/4\right) = \left(\frac{1}{4}\right)(3) - 3/4 = \frac{3}{4} - \frac{3}{4} = 0$

$\sigma^2 = \sum_{i=2}^{2}\left(\frac{1}{4}\right)^2(\sigma_i^2) = \left(\frac{1}{16}\right)(16) = 1$

(f) $Q = \sum_{i=1}^{6}\left[\left(\frac{X_i - 3}{4}\right)^2\right]$ $Q \sim \chi_6^2$

Since $U = \frac{X_i - 3}{4}$ and $U \sim Normal(0,1)$, then

$$\sum_{i=1}^{6}[U^2] = \sum_{i=1}^{6}\left(\frac{X_i-3}{4}\right)^2 \sim \chi_6^2 \quad (*)$$

(g) $R = \frac{Y_2}{\sqrt{(Y_1^2)}}$ $R \sim t_1$

Since $Y_2 \sim Normal(0,1)$ and $(Y_1^2) \sim \chi_1^2$ (*)

then

$$R = \frac{Y_2}{\sqrt{(Y_1^2)}} \sim t_1$$

proof

$W \sim N(0,1)$

$V \sim \chi_r^2$

$U = \frac{W}{\sqrt{V/r}} \sim t_r$

5. Suppose a random variable X has probability density function

$$f_X(x) = \begin{cases} e^x, & x < 0 \\ 0 & elsewhere \end{cases}$$

(a) Show that the moment generating function (mgf) for X is $M_X(t) = \dfrac{1}{(t+1)}$.

(b) Using this moment generating function, find the mean of X.

(c) Using this moment generating function, find the variance of X.

Note: if your class did not study mgf's when you took PSTAT 120AB, please email Prof Meiring as soon as possible.

---

a) $M_x(t) = E[e^{tx}] = \int_{-\infty}^{0} e^{tx} e^x \, dx = \int_{-\infty}^{0} e^{x(t+1)} \, dx = \frac{1}{t+1} e^{x(t+1)} \Big|_{-\infty}^{0} = \frac{1}{t+1} e^0 - \frac{1}{t+1} e^{-\infty} = \frac{1}{t+1} - 0 = \frac{1}{t+1}$

$\therefore M_x(t) = \frac{1}{t+1}$

b) Since $E[X] = \frac{d}{dt} M_x(t) \Big|_{t=0} = M_x'(0)$,

$M_x(t) = (t+1)^{-1}$

$M_x'(t) = -(t+1)^{-2} = \dfrac{-1}{(t+1)^2}$

$M_x'(0) = \dfrac{-1}{(0+1)^2} = -1$

$\therefore E[X] = -1$

c) Since $E[X^2] = \frac{d}{dt} M_x'(t) \Big|_{t=0} = M_x''(0)$

$M_x''(t) = \frac{d}{dt} -(t+1)^{-2} = 2(t+1)^{-3} = \dfrac{2}{(t+1)^3}$

$M_x''(0) = \dfrac{2}{(0+1)^3} = 2 = E[X^2]$

Since $Var(X) = E[X^2] - [E[X]]^2$

$Var(X) = 2 - [-1]^2 = 2 - 1 = 1$

$\therefore Var(X) = 1$