ASSIGNMENT

Piyush Soni [SBU ID # 113273375]

Abhishek Bhattad [SBUID# 113277131]

Aditya Nandan Bhide [SBU ID# 113216565]

Niket Shah [SBU ID# 113274949]

A-6

1) $x_1, x_2, \ldots x_n \sim Nor(\theta, \sigma^2)$

$\Rightarrow$ prior $(\theta) \sim Nor(a, b^2)$

$$\Rightarrow f(\theta) = \frac{1}{\sqrt{2\pi}\, b} \exp\left(-\frac{(\theta-a)^2}{2b^2}\right)$$

$\Rightarrow$ posterior $(\theta) \propto \mathcal{L}(p) \cdot f(\theta)$

$$\Rightarrow f(\theta|D) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x_i-\theta)^2}{2\sigma^2}\right) \times \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{(\theta-a)^2}{2b^2}\right)$$

$$\Rightarrow f(\theta|D) \propto \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left(\frac{-1}{2\sigma^2} \Sigma (x_i-\theta)^2\right) \times \frac{1}{(\sqrt{2\pi}\, b)} \exp\left(-\frac{(\theta-a)^2}{2b^2}\right)$$

Since, posterior (θ), i.e. $f(\theta|D)$ is proportional to above expression, if we remove the constants from the above eqn, it will still be proportional to above expression, since all constants are positive so proportionality would not change.

$$\Rightarrow f(\theta|D) \propto \exp\left(\frac{-1}{2\sigma^2} \Sigma (x_i-\theta)^2\right) . \exp\left(-\frac{(\theta-a)^2}{2b^2}\right)$$

$$\propto \exp\left(-\frac{1}{2\sigma^2} \Sigma (x_i-\theta)^2 - \frac{(\theta-a)^2}{2b^2}\right)$$

$$\Rightarrow f(\theta|D) \propto \exp\left(-\frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2 - \frac{(\theta-a)^2}{2b^2}\right)$$

$$\propto \exp\left(\frac{-1}{2\sigma^2}\Sigma(x_i^2+\theta^2-2x_i\theta) - \frac{(\theta^2+a^2-2\theta a)}{2b^2}\right)$$

$$\propto \exp\left[\frac{-1}{2\sigma^2}\left(\Sigma x^2 + n\theta^2 - 2\theta\Sigma x_i\right) - \left[\frac{\theta^2}{2b^2} + \frac{a^2}{2b^2} - \frac{2\theta a}{2b^2}\right]\right]$$

$$\propto \exp\left[\theta^2\left[\frac{-n}{2\sigma^2} - \frac{1}{2b^2}\right] + \theta\left[\frac{2\Sigma x_i}{2\sigma^2} + \frac{2a}{2b^2}\right] - \frac{\Sigma x_i^2}{2\sigma^2} - \frac{a^2}{2b^2}\right]$$

$$\Rightarrow f(\theta|D) \propto \exp\left[\theta^2\left[\frac{-n}{2\sigma^2} - \frac{1}{2b^2}\right] + \theta\left[\frac{n\bar{x}}{\sigma^2} + \frac{a}{b^2}\right] + \text{const}\right] \quad \text{—(1)}$$

for $f(\theta|D)$ to be Normal, $f(\theta|D) \sim Nor(x, y^2)$

i.e. $f(\theta|D) \propto \exp\left[-\frac{(\theta-x)^2}{2y^2}\right]$

$$\Rightarrow f(\theta|D) \propto \exp\left[\theta^2\left[\frac{-1}{2y^2}\right] + \theta\left[\frac{x}{y^2}\right] - \frac{x^2}{2y^2}\right] \quad \text{—(2)}$$

Comparing eq^n ① and ②, cofficients of θ^2 and θ should be same -

$$\Rightarrow \frac{-n}{2\sigma^2} - \frac{1}{2b^2} = \frac{-1}{2y^2} \qquad \{\text{cofficients of } \theta^2\}$$

$$\Rightarrow \frac{1}{y^2} = \frac{n}{\sigma^2} + \frac{1}{b^2} \qquad \Bigg| \quad \underline{\underline{se^2 = \frac{\sigma^2}{n}}}$$

$$\Rightarrow \frac{1}{y^2} = \frac{1}{se^2} + \frac{1}{b^2}$$

$$\Rightarrow \frac{1}{y^2} = \frac{b^2+se^2}{b^2se^2}$$

$$\Rightarrow \boxed{y^2 = \frac{b^2se^2}{b^2+se^2}}$$

Comparing coff of 'θ' -

$$\Rightarrow \frac{n\bar{x}}{\sigma^2} + \frac{a}{b^2} = \frac{x}{y^2}$$

$$\Rightarrow \frac{\bar{x}}{se^2} + \frac{a}{b^2} = x\left(\frac{1}{se^2} + \frac{1}{b^2}\right)$$

$$\Rightarrow \frac{\bar{x}b^2 + a\,se^2}{\cancel{b^2\,se^2}} = x\left(\frac{b^2+se^2}{\cancel{b^2se^2}}\right)$$

$$\Rightarrow \boxed{x = \frac{b^2\bar{x} + se^2 a}{b^2+se^2}}$$

(b)

$$\text{Posterior}(\theta) = f(\theta \mid D) \sim \text{Nor}(x, y^2)$$

$$z_\theta = a\theta + b$$
$$= a\,\text{Nor}(x, y^2) + b$$
$$= \text{Nor}(ax+b,\ a^2y^2)$$

{where, z_θ is a std. Normal transformatⁿ of posterior(θ)}

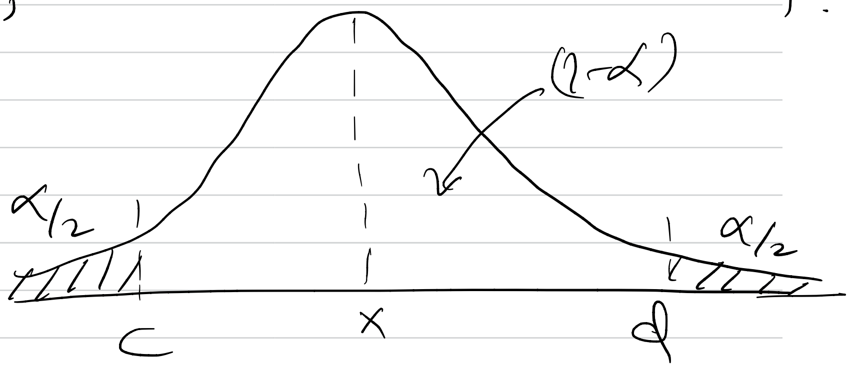
$ax+b=0 \Rightarrow$ $\boxed{b = \frac{-x}{y}}$

$a^2y^2 = 1 \Rightarrow$ $\boxed{a = \frac{1}{y}}$

$$\boxed{z_\theta = \frac{\theta - x}{y}} \quad \text{—— (1)}$$

→ z_θ is std Normal $\sim \text{Nor}(0,1)$

let (c, d) be the CI of posterior(θ), {from part (a)}



then, by defⁿ of CI —

$$\Pr(\theta \le c) = \alpha/2$$

and, $\Pr(\theta \ge d) = \alpha/2$

also,

$$\Pr(c \le \theta \le d) = 1-\alpha$$

$$\Rightarrow \Pr\left(\frac{c-x}{y} \le \frac{\theta - x}{y} \le \frac{d-x}{y}\right) = 1-\alpha$$

$$\Rightarrow \Pr\left(\frac{c-x}{y} \le z_\theta \le \frac{d-x}{y}\right) = 1-\alpha \quad \text{—— (2)}$$

—— {from eqⁿ (1)}

where $z_\theta \sim \text{Nor}(0,1)$

Now, by def^n of $(1-\alpha)$ CI —

$$\Rightarrow \Pr\left(-z_{\alpha/2} \leq z_\theta \leq z_{\alpha/2}\right) = 1-\alpha \quad \text{— (3)}$$

Now, comparing eq^n (2) and (3) —

$$\Rightarrow \frac{c-x}{y} = -z_{\alpha/2}, \quad \frac{d-x}{y} = z_{\alpha/2}$$

$$\Rightarrow c = x - y \cdot z_{\alpha/2}, \quad d = x + y\, z_{\alpha/2}$$

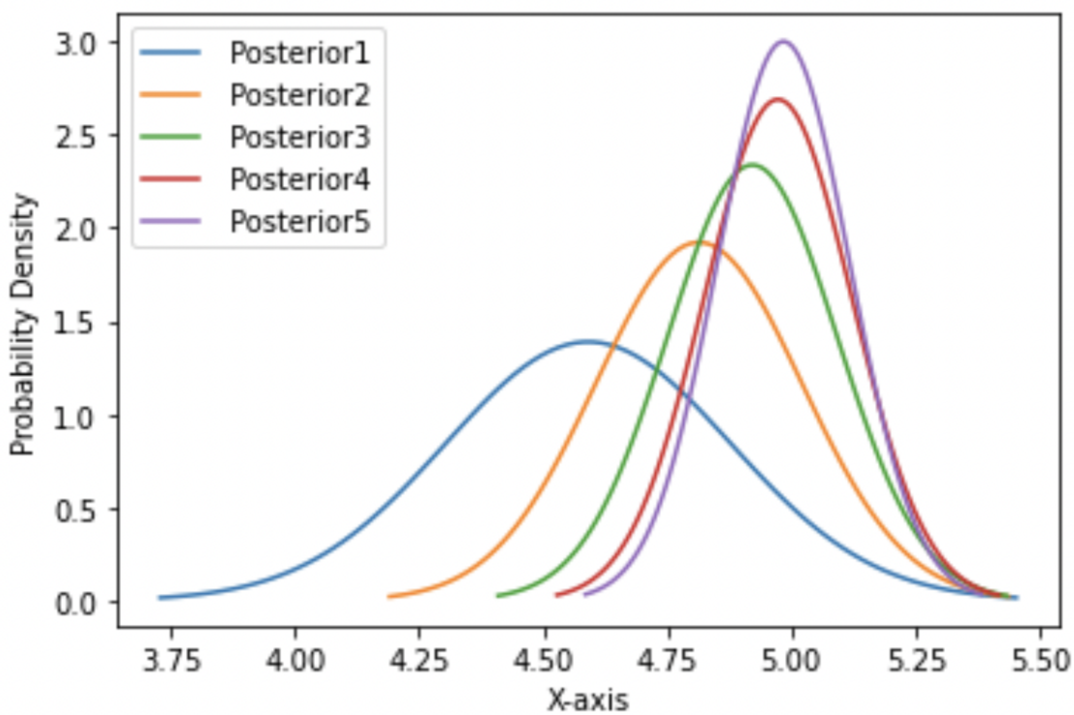
$\Rightarrow$ Thus, $(1-\alpha)$ CI for posterior θ is given by —

$$\Rightarrow \boxed{\left(x - y \cdot z_{\alpha/2},\ x + y \cdot z_{\alpha/2}\right)}$$

Question 2

a]

Posterior	Mean	Variance
1	4.59076	0.0825688
2	4.81352	0.0430622
3	4.92126	0.0291262
4	4.97284	0.0220049
5	4.98397	0.0176817

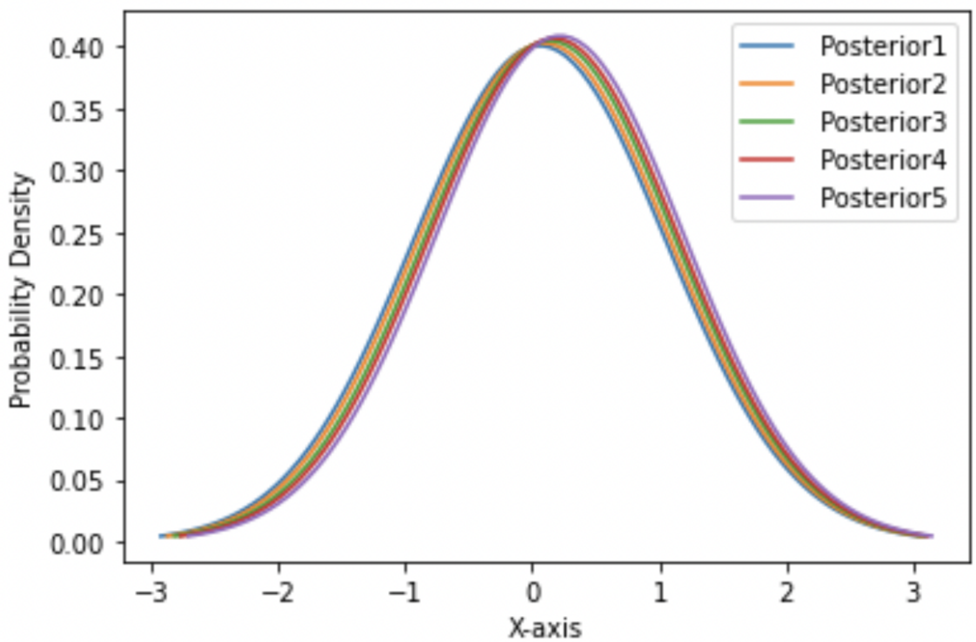


Observations:

1) As we add more data the graph is shifting forward or the mean of the distribution is increasing and it is tending towards the true mean.
2) The variance is reducing drastically initially but the rate of change of variance is slowly reducing.

b]

Posterior	Mean	Variance
1	0.0587162	0.990099
2	0.0950087	0.980392
3	0.138226	0.970874
4	0.171219	0.961538
5	0.218918	0.952381



Observations:

1) As we add more data the mean is increasing slowly and the change in variance is lower compared to the example in part a.

c]
Observation:

1) Posterior tends to move away from the prior at low variance.
2) Posterior and prior are closer to each other at high variance.

3.

$\{(y_1, x_1), \cdots (y_n, x_n)\}$

$\Rightarrow Y = \beta_0 + \beta_1 x + \epsilon_i \quad ; \quad E[\epsilon_i] = 0$

(a) from estimates derived in class —

$$\hat{\beta}_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n(\bar{x})^2}$$

TP

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i y_i - x_i \bar{y} - y_i \bar{x} + \bar{x}\bar{y})}{\sum (x_i^2 + \bar{x}^2 - 2x_i\bar{x})}$$

$$= \frac{\sum x_i y_i - \bar{y}\sum x_i - \bar{x}\sum y_i + n\bar{x}\bar{y}}{\sum x_i^2 + n\bar{x}^2 - 2\bar{x}\sum x_i}$$

$$= \frac{\sum x_i y_i - \frac{\sum y_i}{n}\sum x_i - \frac{\sum x_i}{n}\sum y_i + n\bar{x}\bar{y}}{\sum x_i^2 + \cancel{n}\frac{(\sum x)^2}{n^{\cancel{2}}} - 2\frac{\sum x_i \sum x_i}{n}}$$

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \frac{2\sum x_i \sum y_i}{n} + \frac{\sum x_i}{n} \cdot \frac{\sum y_i}{\cancel{n}} \times \cancel{n}}{\sum x_i^2 + \frac{(\sum x_i)^2}{n} - \frac{2(\sum x_i)^2}{n}}$$

$$= \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{\sum x_i y_i - n \frac{\sum x_i}{n} \frac{\sum y_i}{n}}{\sum x_i^2 - n \frac{(\sum x_i)^2}{n^2}}$$

$$\Rightarrow \boxed{\hat{\beta}_1 = \frac{\sum x_i y_i - n \bar{x} \bar{Y}}{\sum x_i^2 - n(\bar{x})^2}} \Rightarrow$$ which is same as derived in class.

$$\Rightarrow \bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$$

$$\Rightarrow \boxed{\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}}$$

3. (b)

$$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i - \bar{x}) y_i - \bar{y} \sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i - \bar{x}) y_i - \bar{y} \left(\sum x_i - n\bar{x} \right)}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i - \bar{x}) y_i - \bar{y} \left(\cancel{n\bar{x}} - \cancel{n\bar{x}} \right)}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$$

Now,

$$E[\hat{\beta}_1] = E\left[\frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} \right]$$

$$= \frac{\sum (x_i - \bar{x}) E[y_i]}{\sum (x_i - \bar{x})^2} \quad \text{— (1)}$$

{Treating x's as const}

$$\Rightarrow \quad Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

$$\Rightarrow \quad E[Y_i] \overset{\text{LOE}}{=} E[\beta_0] + E[\beta_1 X_i] + E[\epsilon_i]$$

$$= \beta_0 + \beta_1 X_i + 0$$

$$= \beta_0 + \beta_1 X_i \quad \text{--- (2)} \qquad \left\{ \because E[\epsilon_i] = 0 \text{ Assumpt}^n \text{ (1)} \right\}$$

$\Rightarrow$ Substituting eq^n (2) in (1) –

$$\Rightarrow \quad E[\hat{\beta}_1] = \frac{\sum (X_i - \overline{X})(\beta_0 + \beta_1 X_i)}{\sum (X_i - \overline{X})^2}$$

$$= \beta_0 \frac{\sum (X_i - \overline{X})}{\sum (X_i - \overline{X})^2} + \beta_1 \frac{\sum (X_i - \overline{X}) X_i}{\sum (X_i - \overline{X})^2}$$

$$= \beta_0 \frac{(\sum X_i - n\overline{X})}{\sum (X_i - \overline{X})^2} + \beta_1 \frac{\sum (X_i - \overline{X}) X_i}{\sum (X_i - \overline{X})^2}$$

$$= \beta_0 \frac{(\cancel{n\overline{X}} - \cancel{n\overline{X}})}{\sum (X_i - \overline{X})^2} + \beta_1 \frac{\sum (X_i - \overline{X}) X_i}{\sum (X_i - \overline{X})^2}$$

$$= \beta_1 \frac{\sum (X_i - \overline{X}) X_i}{\sum (X_i - \overline{X})^2}$$

$$\Rightarrow E[\hat{\beta}_1] = \beta_1 \frac{\sum(x_i - \bar{x})x_i}{\sum(x_i - \bar{x})^2} \qquad \left\{\bar{x} = \frac{\sum x_i}{n}\right\}$$

$$= \beta_1 \frac{\sum(x_i^2 - \bar{x}x_i)}{\sum(x_i^2 + \bar{x}^2 - 2x_i\bar{x})}$$

$$= \beta_1 \frac{(\sum x_i^2 - \bar{x}\sum x_i)}{(\sum x_i^2 + \sum \bar{x}^2 - 2\bar{x}\sum x_i)}$$

$$= \beta_1 \frac{\sum x_i^2 - n\bar{x}^2}{\sum x_i^2 + n\bar{x}^2 - 2\bar{x}(n\bar{x})}$$

$$= \beta_1 \frac{\cancel{(\sum x_i^2 - n\bar{x}^2)}}{\cancel{(\sum x_i^2 - n\bar{x}^2)}}$$

$$\Rightarrow \boxed{E[\hat{\beta}_1] = \beta_1}$$ Thus, $\hat{\beta}_1$ is unbiased.

Now, $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$

$$\Rightarrow E[\hat{\beta}_0] = E[\bar{Y} - \hat{\beta}_1\bar{X}]$$

$$\overset{\text{LoE}}{=} E[\bar{Y}] - E[\hat{\beta}_1\bar{X}]$$

$$= \frac{\sum E[Y_i]}{n} - \bar{X}E[\hat{\beta}_1] \qquad \left\{\begin{array}{l}\text{Assuming } X\text{'s as const} \\ \text{and } \bar{Y} = \frac{\sum Y_i}{n}\end{array}\right\}$$

$$= \frac{\sum(\beta_0 + \beta_1 X_i)}{n} - \bar{X}\beta_1 \qquad \{\text{from eq}^n \text{ 2}\}$$

$$\Rightarrow E[\hat{\beta}_0] = \frac{\Sigma(\beta_0 + \beta_1 x_i)}{n} - \beta_1 \frac{\Sigma x_i}{n}$$

$$= \frac{\Sigma \beta_0}{n} + \cancel{\beta_1 \frac{\Sigma x_i}{n}} - \cancel{\beta_1 \frac{\Sigma x_i}{n}}$$

$$= \frac{n\beta_0}{n}$$

$$\boxed{E[\hat{\beta}_0] = \beta_0}$$

Thus, $\hat{\beta}_0$ is also unbiased.

Q4.

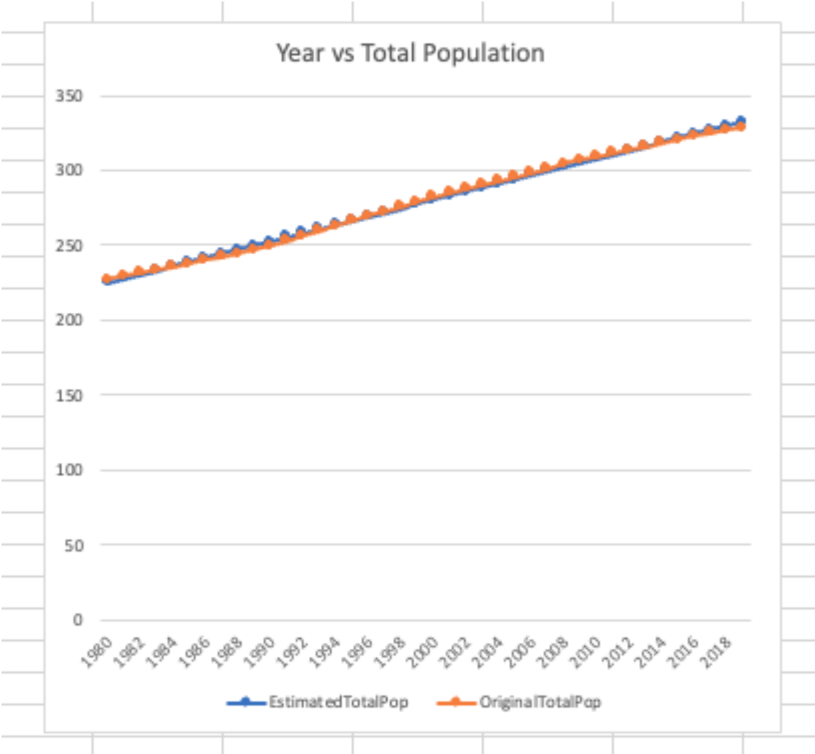
a) 1. Year vs Total US Population

beta0(hat) = -5236.6206

beta1(hat) = 2.7584

Yi(hat) = -5236.6206 + 2.7584*Xi

SSE = 123.8535



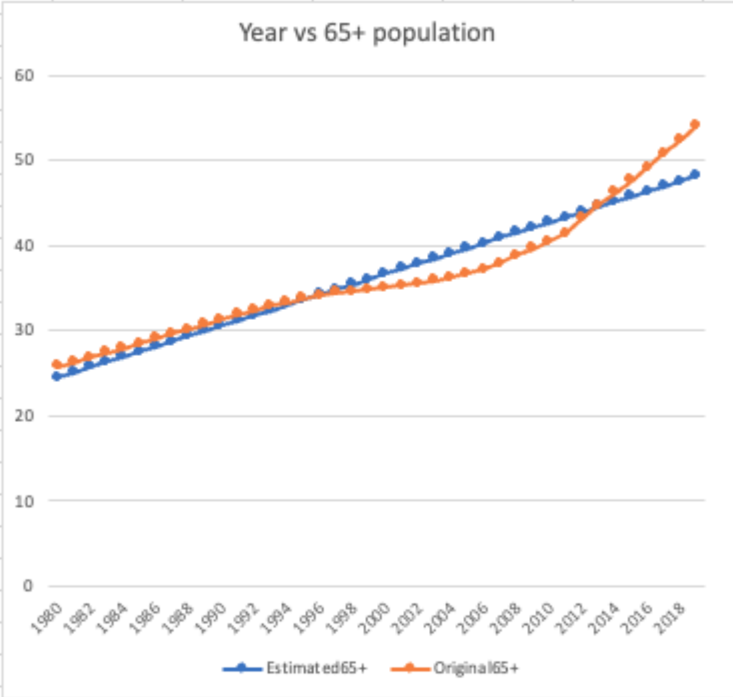
2. Year vs 65+ US Population

beta0(hat) = -1177.01269

beta1(hat) = 0.60682

Yi(hat) = -1177.01269 + 0.60682*Xi

SSE = 176.0385



The SSE of “Year vs Total US Population” is less than “Year vs 65+ US Population”, so Year vs Total US Population is more suitable. Also this can be seen from the graph. There are few data points, from 1998 to 2010 and 2014 to 2019 that don't really fit the linear regression line in “Year vs 65+ US Population” whereas the “Year vs Total US Population” regression line perfectly fits the data except for a few points towards the end.

b) 1. From 1980 to 2018
beta0(hat) = -1131.6868
beta1(hat) = 0.58407

Yi(hat) = -1131.6868 + 0.58407*Xi

SSE = 137.6930652

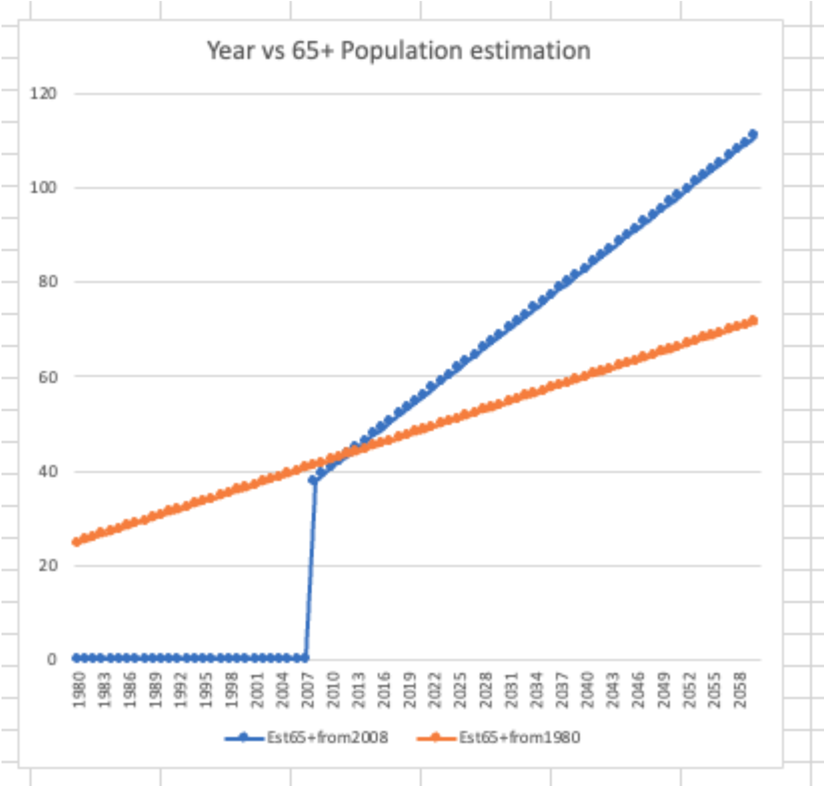
Prediction in 2060 = 71.4974 million

2. From 2008 to 2018
beta0(hat) = -2778.111068
beta1(hat) = 1.4024

Yi(hat) = -2778.111068 + 1.4024*Xi

SSE = 2.008763174

Prediction in 2060 = 110.832932 million



We have a different number of years to compare SSE values. So, we will take MSSE values.
MSE (method 1) = **137.6930652/39 = 3.530591415**
MSE (method 2) = **2.008763174/11 = 0.182614834**

From method 2 i.e. 2008 to 2018, we get low MSE value. So we should trust the data with low MSE value. Taking into consideration the trusted data, the media was correct with predicting the population getting doubled and as predicted it's more than doubled (110.83 million) in 2060.

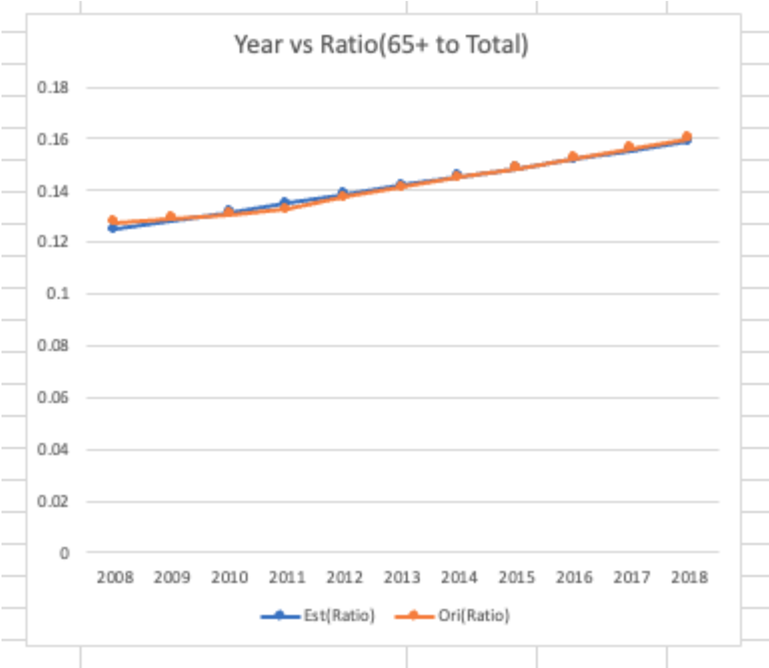
c) 1. I have taken the ratio of Total US Population to 65+ population for the year 2008 to 2018.

beta0(hat) = -6.719157552
beta1(hat) = 0.00340838

Yi(hat) = -6.719157552 + 0.00340838*Xi

Prediction in 2019 = 0.16236237

SSE: 1.79804*10^-5



2. 2008 to 2018

From part b,
Predicted population of 65+ in 2019 = 53.334532 million

For Total population from 2008 to 2018,
beta0(hat) = -4280.563141
beta1(hat) = 2.283433396

Yi(hat) = -4280.563141 + 2.283433396*Xi
Predicted total population in 2019 = 329.68886

Ratio = 53.334532/329.688886 = 0.1617723207

From original data, ratio for 2019 = 54.058263/328.329953 = **0.1646461509**

Our method 1 answer matches closer to the original ratio. Therefore, method 1 (taking the ratio first and then predicting) is more suitable than method 2 (predicting individually and then taking ratios). But the difference between the ratios for both the approaches is not significant.

In method 1, we observe from the graph that the regression line almost fits the data points. In method 2, our 65+ population regression line is similar to method 1 ratio regression line. Assuming both carry almost similar errors, we are left with errors for the total population variable in method 2. If we account for errors in method 1 and 2, we have one error for method 1 and we add up two errors of 65+ population and total population. So, our method 1 is more suitable to method 2 but as evident from the calculations, the difference is not significant.

Question 5

a]
Equation for part a:
-0.002910668651756 3407 x_0 +0.00323153935083614 x_1
+0.019909621463865344 x_2 +0.0005760918070320722 x_3
+0.02319266851656047 x_4 +0.1308982007428861 3 x_5 +0.05682042842085549 5
x_6

SSE: 0.3164114091068155

b]
Equation for part b:
+0.0038865467616105006 x_0 +0.04187384651642297 x_1
+0.04825699298089394 x_2

SSE: 0.6403887599866868

c]
Equation for part c:
-0.004106308666515591 x_0 +0.23571158786308255 x_1

SSE: 0.4638050678629464

d]
1) If we use all features SSE reduces so we can say that getting more data we can more accurately predict if a person will get the admission or not.
2) As we are reducing the amount of data SSE will increase.
3) If we select 'TOEFL Score', 'SOP', 'LOR' our SSE is higher compared to ‘GRE Score’, ‘GPA’. So we can conclude that according to data ‘GRE Score’ and ‘GPA’ matters more while deciding whether a person will get admission.
4) In part ‘a’ we can also observe that the coefficient of features like ‘GPA’ is much higher so we can say that it is contributing more towards the final decision of admission.

6. $H \sim \{0, 1\}$ → bad soil; ↓ good soil

prior ⇓

priors –

H_0 : good soil $\quad P(H=0) = P$

H_1 : bad soil $\quad P(H=1) = 1-P$

$$\boxed{f_H(h) = P^{1-h}(1-P)^{h}}$$

water content $(W) \rightarrow f_W(w/H=0) = Nor(w; -\mu, \sigma^2)$

(a)

$$f_W(w/H=1) = Nor(w; \mu, \sigma^2)$$

sample, $w = \{w_1 --- w_n\}$ cond^n I

Posterior: $Pr(H=0 \mid \{w_1 -- w_n\}) \overset{\text{Bayes theorem}}{=} \dfrac{Pr(H=0 \cap \{w_1 -- w_n\}) \, Pr(H=0)}{Pr(\{w_1 -- w_n\})}$

$$\overset{\text{chain rule}}{=} \frac{Pr(\{w_1 -- w_n\} \mid H=0) \, Pr(H=0)}{Pr(\{w_1 -- w_n\})}$$

→ since, $Pr(\{w_1 -- w_n\})$ is independent of 'H', thus replaced by proportionality –

$$\Rightarrow Pr(H=0/\{w_1 - w_n\}) \propto Pr(\{w_1 -- w_n\}/H=0) \, P(H=0)$$

$$\propto \prod_{i}^{n} pr(w_i/H=0) \; P(H=0)$$

$$\propto \left(\prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(w_i-(-\mu))^2}{2\sigma^2}\right)\right)\right) \cdot P$$

$\left\{ \because Pr(H=0) = P \text{ and, } P(w/H=0) \sim Nor(-\mu, \sigma^2) \right\}$

$$\propto \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\Sigma(w_i+\mu)^2}{2\sigma^2}\right) \cdot P$$

Replacing all constants and proportionality with a single const 'c' -

$$\Rightarrow Pr\left(H=0 / \{w_1 -- w_n\}\right) = c \cdot p \cdot \exp\left(-\frac{\Sigma(w_i + \mu)^2}{2\sigma^2}\right) \quad —(1)$$

Similarily,

$$\Rightarrow Pr\left(H=1 / \{w_1 -- w_n\}\right) = c \cdot (1-p) \exp\left(-\frac{\Sigma(w_i - \mu)^2}{2\sigma^2}\right) \quad —(2)$$

since, $Pr(H=1) = 1-p$

and, $f(w/H=1) \sim Nor(\mu, \sigma^2)$

Now, for choosing $C=0$ i.e. good soil, —

$$\Rightarrow Pr\left(H=0 / \{w_1 -- w_n\}\right) \geq Pr\left(H=1 / \{w_1 -- w_n\}\right)$$

{Given}

Using eq^n (1) and (2)

$$\Rightarrow \cancel{c} \cdot p \cdot \exp\left(-\frac{\Sigma(w_i + \mu)^2}{2\sigma^2}\right) \geq \cancel{c} \cdot (1-p) \cdot \exp\left(\frac{-\Sigma(w_i - \mu)^2}{2\sigma^2}\right)$$

$$\Rightarrow \frac{p}{1-p} \geq \exp\left(\frac{\Sigma(w_i + \mu)^2}{2\sigma^2} - \frac{\Sigma(w_i - \mu)^2}{2\sigma^2}\right)$$

$$\Rightarrow \frac{p}{1-p} \geq \exp\left(\frac{\Sigma\left((w_i + \mu)^2 - (w_i - \mu)^2\right)}{2\sigma^2}\right)$$

$$\Rightarrow \frac{p}{1-p} \geq \exp\left(\frac{\sum \left((w_i+\mu)^2 - (w_i-\mu)^2 \right)}{2\sigma^2} \right)$$

$$\Rightarrow \frac{p}{1-p} \geq \exp\left(\frac{\sum (2w_i)(2\mu)}{2\sigma^2} \right)$$

$$\Rightarrow \frac{p}{1-p} \geq \exp\left(\frac{2\mu}{\sigma^2} \sum w_i \right)$$

$$\Rightarrow \ln\left(\frac{p}{1-p}\right) \geq \frac{2\mu}{\sigma^2} \sum w_i$$

$$\Rightarrow \boxed{\sum w_i \leq \frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)}$$

condition for $\underline{\underline{C=0}}$

Question 6

b]

```
For P(H_0) = 0.1, the hypotheses selected are :  0 1 0 0 1 0 1 1 0 1
For P(H_0) = 0.3, the hypotheses selected are :  0 1 0 0 1 0 1 1 0 1
For P(H_0) = 0.5, the hypotheses selected are :  0 1 0 0 1 0 1 1 0 1
For P(H_0) = 0.8, the hypotheses selected are :  0 1 0 0 1 0 1 1 0 1
```

6. $w = \{w_1 \, --- \, w_n\}$

(c)

$\rightarrow AEP = P(C=0 \mid H=1)P(H=1) + P(C=1 \mid H=0)P(H=0)$

from part (a) —

$C=0$; if following condⁿ holds —

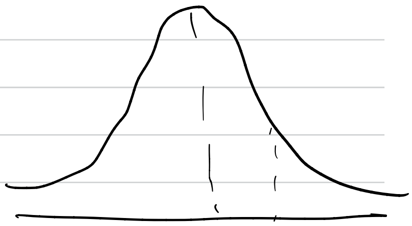
$$\sum w_i \le \frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)$$

So, $P(C=0 \mid H=1) = Pr\left(\underbrace{\sum w_i}_{Y} \le \underbrace{\frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)}_{d} \,\middle|\, H=1\right)$

$\Rightarrow w_i \sim Nor(\mu, \sigma^2)$ {if $H=1$}

$\therefore \underbrace{\sum_{i=1}^{n} w_i}_{Y} \sim Nor(n\mu, n\sigma^2)$ {Sum of independent Normals}

$\Rightarrow Pr(C=0 \mid H=1) = Pr(Y \le d)$ — (1)



Let, $Z = aY + b$
(St. Nor.)

{Given $H=1$, $\sum w_i = Y \sim Nor(n\mu, n\sigma^2)$}

$\Rightarrow Nor(0,1) \equiv a \times Nor(n\mu, n\sigma^2) + b$

$\Rightarrow an\mu + b = 0$

$\Rightarrow a^2 n\sigma^2 = 1 \Rightarrow \boxed{a = \frac{1}{\sigma\sqrt{n}}}$

$\boxed{b = -\frac{\mu\sqrt{n}}{\sigma}}$

$\Rightarrow Z = \frac{Y}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}$

From eq^n 1

$$\Rightarrow Pr(C=0|H=1) = Pr(Y \le d) = F_Y(d)$$

also, $$Z = \frac{Y}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}$$

$$\Rightarrow \boxed{d_z = \frac{d}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}}$$

Thus, $$F_Y(d) = \Phi\left(\frac{d}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}\right)$$

$$\therefore Pr(C=0|H=1) = \Phi\left(\frac{d}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}\right)$$

$$= \Phi\left(\frac{\frac{\sigma^2}{2\mu}\ln\left(\frac{P}{1-P}\right)}{\sigma\sqrt{n}} - \frac{\mu\sqrt{n}}{\sigma}\right)$$

$$\boxed{Pr(C=0|H=1) = \Phi\left(\frac{\sigma}{2\mu\sqrt{n}}\ln\left(\frac{P}{1-P}\right) - \frac{\mu\sqrt{n}}{\sigma}\right)}$$

— 2

Now, for $Pr(C=1 \mid H=0)$

for $C=1$,

$$Pr(H=0 \mid \{w_1 \dots w_n\}) < Pr(H=1 \mid \{w_1 \dots w_n\})$$

from part (a) —

for $C=1 \Rightarrow$

$$\Rightarrow \sum w_i > \frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)$$

$$\therefore Pr(C=1 \mid H=0) = Pr\left(\sum w_i > \frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right) \;\middle|\; H=0\right)$$

$$= 1 - Pr\Big(\underbrace{\sum w_i}_{X} \leq \underbrace{\frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)}_{C} \;\Big|\; H=0\Big)$$

$$= 1 - Pr(X \leq C \mid H=0)$$

Now, for given $H=0$,

$$w_i \sim Nor(-\mu, \sigma^2)$$

$$\therefore \sum w_i \sim \sum Nor(-\mu, \sigma^2)$$

for sum of independent Normals —

$$\Rightarrow \sum w_i \sim Nor(-n\mu, n\sigma^2)$$

$$\Rightarrow X \sim Nor(-n\mu, n\sigma^2)$$

Thus, for $H=0$

and, $X \sim Nor(-n\mu, n\sigma^2)$

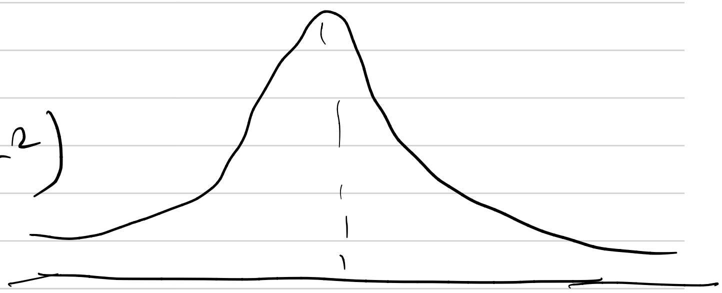
$\Rightarrow Pr(C=1 \mid H=0) = 1 - F_X(c)$

where, $c = \frac{\sigma^2}{2\mu} \ln\left(\frac{p}{1-p}\right)$

Now,

$Z = aX + b$

$\Rightarrow Nor(0,1) = a\,Nor(-n\mu, n\sigma^2) + b$



$\therefore -na\mu + b = 0$

$b = n\mu a$

$\Rightarrow a^2 n\sigma^2 = 1$

$\Rightarrow \boxed{a = \frac{1}{\sigma\sqrt{n}}}$, $\boxed{b = \frac{\mu\sqrt{n}}{\sigma}}$

$\therefore Z = \frac{X}{\sigma\sqrt{n}} + \frac{\mu\sqrt{n}}{\sigma}$ | $F(c) = \Phi(Z_c)$

$\Rightarrow Z_c = \frac{c}{\sigma\sqrt{n}} + \frac{\mu\sqrt{n}}{\sigma}$

$\Rightarrow Pr(C=1 \mid H=0) = 1 - \Phi(Z_c)$

$$\Rightarrow \Pr(C=1 \mid H=0) = 1 - \Phi\left(\frac{c}{\sigma\sqrt{n}} + \frac{\mu\sqrt{n}}{\sigma}\right)$$

$$= 1 - \Phi\left(\frac{\frac{\sigma^2}{2\mu}\ln\left(\frac{p}{1-p}\right)}{\sigma\sqrt{n}} + \frac{\mu\sqrt{n}}{\sigma}\right)$$

$$\boxed{\Pr(C=1 \mid H=0) = 1 - \Phi\left(\frac{\sigma}{2\mu\sqrt{n}}\ln\left(\frac{p}{1-p}\right) + \frac{\mu\sqrt{n}}{\sigma}\right)}$$

from eq^n 2 and eq^n 3 — — 3

$$\Rightarrow AEP = \Pr(C=0 \mid H=1)\Pr(H=1) + \Pr(C=1 \mid H=0)\Pr(H=0)$$

$$\boxed{AEP = (1-p)\left[\Phi\left(\frac{\sigma}{2\mu\sqrt{n}}\ln\left(\frac{p}{1-p}\right) - \frac{\mu\sqrt{n}}{\sigma}\right)\right] + p_0\left(1 - \Phi\left(\frac{\sigma}{2\mu\sqrt{n}}\ln\left(\frac{p}{1-p}\right) + \frac{\mu\sqrt{n}}{\sigma}\right)\right)}$$