Solution to the test

1.(a)
$\hat{\beta} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_x-\bar{x})^2} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{13469-100\times10\times11.8}{12820-100\times10^2} = 0.5918$

$\hat{\alpha} = \bar{y} - \bar{x}\hat{\beta} = 11.8 - 0.5918 \times 10 = 5.882$

An unbiased estimator for σ^2 is $s^2 = MSE = \frac{SSE}{n-2}$
And $SSE = \sum(y_i - \hat{y}_i)^2 = SS_{YY} - SSB, (VV = TV - EV)$
$TV = SS_{YY} = \sum(y_i - \bar{y})^2 = \sum y_i^2) - n\bar{y}^2 = 25543 - 100 \times (11.8)^2 = 11619$
$EV = SSB = n\hat{\beta}^2 S_X^2 = \hat{\beta}^2 SS_{XX} = \hat{\beta} \cdot SS_{XY} = 0.5918 \times 1669 = 987.7$
Thus, $VV = SSE = SS_{YY} - SSB = 11619 - 987.7 = 10631.3$
$MSE = \frac{10631.3}{98} = 108.48$
$R^2 = \frac{EV}{TV} = 1 - \frac{VV}{TV} = 1 - \frac{10631.3}{11619} = 0.085$ Only 8.5% of the total variability is explained by the straight line model.

(b)
A 95% CI for α is:
$\hat{\alpha} \pm t_{n-2,\alpha/2} \cdot s(\hat{\alpha})$
$= 5.882 \pm t_{98,0.025} \cdot \sqrt{s^2(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i-\bar{x})^2})}$
$= 5.882 \pm 1.99\sqrt{108.48 \times (\frac{1}{100} + \frac{10^2}{2820})}$
$= 5.882 \pm 1.99 \times 2.22$
$= 5.882 \pm 4.4178$
$= [1.4642, 10.2998]$

A 95% CI for α is:
$\hat{\beta} \pm t_{n-2,\alpha/2} \cdot s(\hat{\beta})$
$= 0.5918 \pm t_{98,0.025} \cdot \sqrt{\frac{s^2}{\sum(x_i-\bar{x})^2}}, t_{98,0.025} is approx 1.99$
$= 0.5918 \pm 1.99 \times \sqrt{\frac{108.48}{2820}}$
$= 0.5918 \pm 1.99 \times 0.196$
$= 0.5918 \pm 0.39$
$= [0.2018, 0.9818]$

(c)

For testing

$$H_0 : \beta = 0$$

$$H_a : \beta \neq 0$$

$t = \frac{\hat{\beta}}{s(\hat{\beta})} = \frac{0.5918}{0.196} = 3.019$, since$|t| > t_{98,0.025} = 1.99$, we reject $H_0 : \beta = 0$.

For testing

$$H_0 : \alpha = 0$$

$$H_a : \alpha \neq 0$$

$t = \frac{\hat{\alpha}}{s(\hat{\alpha})} = \frac{5.882}{2.22} = 2.65$, since$|t| > t_{98,0.025} = 1.99$, we reject $H_0 : \alpha = 0$.

2.
(a)
$\min_{\beta} \sum (y_i - \beta x_i)^2$
$\frac{\partial}{\partial \beta} = 2 \sum (y_i - \beta x_i) \cdot x_i = 0 \Rightarrow \hat{\beta} = \frac{\sum y_i x_i}{\sum x_i^2}$

(b)
$\hat{\beta} = \frac{\sum y_i x_i}{\sum x_i^2} = \sum \frac{x_i}{\sum x_i^2} \cdot y_i$ is a linear combination of y_i and y_i are normal, thus $\hat{\beta}$ is normal.

$E[\hat{\beta}] = E[\sum \frac{x_i}{\sum x_i^2} \cdot y_i] = \sum \frac{x_i}{\sum x_i^2} \cdot E[y_i] = \sum \frac{x_i \beta x_i}{\sum x_i^2} = \beta \frac{\sum x_i^2}{\sum x_i^2} = \beta$,unbiased.

$Var[\hat{\beta}] = Var[\sum \frac{x_i}{\sum x_i^2} \cdot y_i] = \sum (\frac{x_i}{\sum x_i^2})^2 \sigma^2 = \frac{\sigma^2}{\sum x_i^2}$, since y_i are independent, and $Var[y_i] = \sigma^2$.

3.
(a)
$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{(y_i - \bar{y})^2}} = 0 \Rightarrow \sum (x_i - \bar{x})(y_i - \bar{y}) = 0$ Then $\hat{\gamma} = 0$ and $\hat{\delta} = \bar{y} - \hat{\gamma}\bar{x} = \bar{y}$. The regression line is $\hat{y} = \bar{y}$.

(b)The LSE and the MLE agree under normality,
Since $l(y_1, ...y_n) = \frac{1}{(2\pi\sigma^2)^(n/2)} \exp\{-\frac{1}{2} \sum (y_i - \delta - \gamma x_i)^2\}$.

(c)
$e_i = y_i - \hat{y}_i = 0, \forall i, \Rightarrow \sum e_i^2 = 0$ i.e $SSE = 0$.
$R^2 = 1 - \frac{SSE}{SS_{YY}} = 1$, perfect fit.

4

(a)

From the computer output, we get $\hat{\alpha} = 0.62228, \hat{\beta} = 0.72646, \hat{y} = 0.62228 + 0.72646x, MSE = \hat{\sigma}^2 = \frac{0.63933}{8-2} = 0.106555$.

(b)

From the computer output, we ge $\hat{y} = 0.19896 + 0.758333x, R^2 = 98\%, \hat{\sigma}^2 = s^2 = 0.0115972$.

6.

(a)

$\bar{y} = \frac{\sum y_i}{n} = \frac{3240}{8} = 405$
$\bar{x} = \frac{\sum x_i}{n} = \frac{44}{8} = 5.5$
$\hat{\beta}_1 = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_x-\bar{x})^2} = \frac{\sum x_iy_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{18060-8\times5.5\times405}{257-8\times(5.5)^2} = \frac{240}{15} = 16$
$\hat{\beta}_0 = \bar{y} - \bar{x}\hat{\beta}_1 = 405 - 5.5 \times 16 = 317$
$\hat{y} = 317 + 16x$

(b)

$s^2 = MSE = \frac{SSE}{n-2}$

And $SSE = \sum(y_i - \hat{y}_i)^2 = SS_{YY} - n\hat{\beta}_1^2 S_X^2 = SS_{YY} - \hat{\beta}_1 \cdot SS_{XY} = 8000 - 3840 = 4160$
$SS_{YY} = \sum(y_i - \bar{y})^2 = \sum y_i^2) - n\bar{y}^2 = 1320200 - 8 \times (405)^2 = 8000$
$EV = SSB = n\hat{\beta}_1^2 S_X^2 = \hat{\beta}_1^2 SS_{XX} = \hat{\beta}_1 \cdot SS_{XY} = 16 \times 240 = 3840$
$MSE = \frac{4160}{6} = 693.3$

(c)

F statistics for regression test:

$$H_0 : \beta_1 = 0$$

$$H_a : \beta_1 \neq 0$$

$F = \frac{EV/1}{MSE} = \frac{3840}{693.3} = 5.5387$
C.R: $\{F > F_{1,6,\alpha}\}$

(d)

A 95% CI for the mean when x=0.613 is:
$\hat{y}_{x_h} \pm t_{n-2,y/2} \cdot s(\hat{y}_{x_h})$
$= \hat{y}_{x_h} \pm t_{6,0.025} \cdot \sqrt{s^2(\frac{1}{n} + \frac{(x_h-\bar{x}^2)^2}{\sum(x_i-\bar{x})^2})}$
$\hat{y}_{x_h} = 317 + 16 \times 0.613 = 326.8$
$s(\hat{y}_{x_h}) = \sqrt{s^2(\frac{1}{n} + \frac{(x_h-\bar{x}^2)^2}{\sum(x_i-\bar{x})^2})}$
$= \sqrt{693.3(\frac{1}{8} + \frac{(0.613-5.5)^2}{15})}$

$= \sqrt{1190.524}$
$= 34.5$
$t_{6,0.025} = 2.447$
Thus, the 95% CI for the mean when x=0.613 is:
$326.8 \pm 2.447 \times 34.5 = 326.8 \pm 84.4215 = [242.3785, 411.2215]$.

7
$\ln y_i = \ln\alpha + \beta x_i + e_i$
$\ln y = \{0, 2.197, 4.499, 6.8, 9.393\}$
$\hat{\beta} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_x-\bar{x})^2} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{13469-100\times10\times11.8}{12820-100\times10^2} \approx 2.3389$

8
A 95% CI for α is:
$\hat{\beta} \pm t_{n-2,\alpha/2} \cdot s(\beta)$
$= 3 \pm t_{8,0.025} \cdot \sqrt{\frac{s^2}{\sum(x_i-\bar{x})^2}}, t_{8,0.025} = 2.306$
Since, $s^2 = \hat{S}_R^2 = 2 = \frac{SSE}{8} \Rightarrow SSE = 16$
$0.92 = R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{16}{SST} \Rightarrow SST = 200$
$SSR = EV = 200 - 16 = 184$
$184 = EV = SSR = \hat{\beta}^2 \sum(x_i - \bar{x})^2 = 3^2 \sum(x_i - \bar{x})^2$
$\Rightarrow \sum(x_i - \bar{x})^2 = \frac{184}{9} = 20.4$
$s(\hat{\beta}) = \sqrt{\frac{2}{20.4}} = \sqrt{0.098} = 0.313$
$CI(\beta) = 3 \pm 2.306 \times 0.313 = 3 \pm 0.72 = [2.28, 3.72] \Rightarrow$ yes, we can reject the null hypothesis,because 0 not in $CI(\beta)$