Family Name:________________________

Given Name:________________________

Student Number:________________________

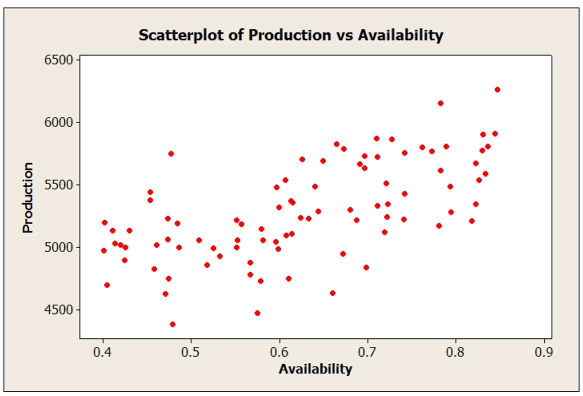
MIE237S Term Test
Examination Type B; Calculator Type 2 Permitted
March 19, 2013, 11:10 A.M.
50 minutes; 20 Marks Available

You should have two booklets. This booklet has 5 pages and consists of the questions and space for you to answer each question. You should have ample space to answer. You may use the backs of pages for extra rough work space. Please do not detach any pages from this booklet.

The second booklet consists of 2 pages printed on both sides. The first side is the aid sheet and the other sides consist of tables of probabilities for the standard normal and t distributions. This booklet is yours to keep.

1.**(16 marks total)** One way to measure the performance of a fleet of haul trucks in an open pit mine is by "availability". The availability of a fleet is (roughly speaking) the proportion of the amount of time the trucks are not being repaired and are actually able to operate.

A reliability engineer at a mine is going to analyze the relationship between haul truck availability and overall mine production in tonnes of ore produced. Here is a plot of $n = 90$ days' worth of data:



The usual simple linear regression model is fit to the data. In case you need it, $S_{xx} = 1.545$. Here is the Minitab output with many entries missing:

```
Regression Analysis: Production versus Availability

The regression equation is
Production = **** + **** Availability


Predictor       Coef  SE Coef      T      P
Constant      4119.2    154.7  26.63  0.000
Availability  ******    241.9   ****  0.000


S = 300.665   R-Sq = 40.4%


Analysis of Variance

Source          DF       SS       MS      F      P
Regression      **  *******    *****  *****  0.000
Residual Error  **  *******    *****
Total           **  *******

Unusual Observations

Obs  Availability  Production     Fit  SE Fit  Residual  St Resid
  4         0.477      5748.0  5008.7    48.0     739.3     2.49R
  7         0.660      4633.6  5351.2    32.8    -717.6    -2.40R
 75         0.575      4469.8  5191.6    34.0    -721.8    -2.42R
 89         0.478      4380.6  5011.5    47.8    -630.9    -2.13R

R denotes an observation with a large standardized residual.
```

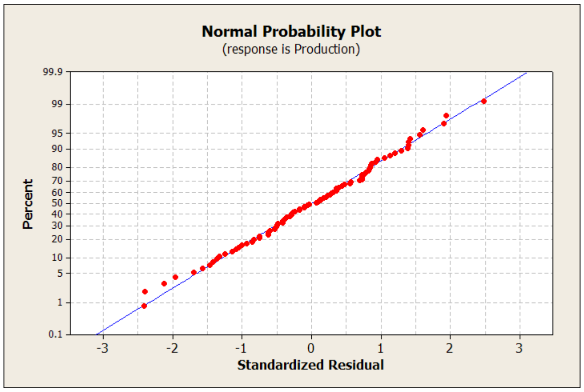
(a) **(4 marks)** Provide the missing entries for the Analysis of Variance part of the Minitab output.

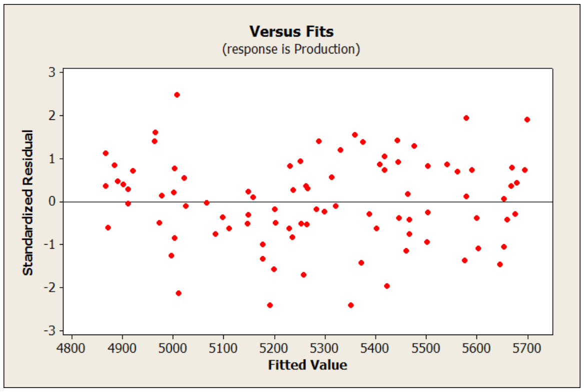
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	___	___________	___________	_______	0.000
Residual Error	___	___________	___________		
Total	___	___________			

(b) **(3 marks)** Provide a 99% confidence interval for the slope of the regression line.

(c) **(3 marks)** Here is a normal probability plot of the standardized residuals and a plot of the standardized residuals versus the fitted values. Comment on whether the usual assumptions of the simple regression model have been satisfied or not.





(d) **(2 marks)** Comment on the possible existence of any outliers or influential points in this dataset.

(e) **(2 marks)** The sample average availability over the 90 days was $\overline{x} = 0.626$. Produce a 95% confidence interval for the mean response at 0.626.

(f) **(2 marks)** The sample average availability over the 90 days was $\overline{x} = 0.626$. The mining company is going to spend thirty million dollars to upgrade the maintenance facility to try to tincrease the average availability to 0.75. Provide a range of plausible values for daily production of ore in tonnes at this new average availability of 0.75, with 95% confidence.

2.**(4 marks total)** Some data $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ is analyzed using the usual simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ with slope estimator $\hat{\beta}_1 = S_{xy}/S_{xx}$.

Suppose a new variable w_i is introduced that is just a linear transformation of the y_i variable. In other words, $w_i = c + dy_i$ for each i (assume $d \neq 0$.) Consider the new simple linear regression model $w_i = \beta_0^{(w)} + \beta_1^{(w)} x_i + \varepsilon_i$.

(a) **(2 marks)** Show that the new slope estimator $\hat{\beta}_1^{(w)}$ is equal to d multiplied by the old slope estimator $\hat{\beta}_1$.

(b) **(2 marks)** Show why the p-value obtained for the hypothesis test $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 \neq 0$ will be exactly the same as the p-value for the hypothesis test $H_0 : \beta_1^{(w)} = 0$ versus $H_1 : \beta_1^{(w)} \neq 0$.