Chapter 2

2.13 Refer to Grade point average. Page 98

a- Calculate R^2. What proportion of the variation in Y is accounted for by introducing X into the regression model? From page 98

$\bar{X} = 24.725, \sum_{i=1}^{n=120}(X_i - \bar{X})^2 = 2379.925$

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	1	3.5878	SSR=3.5878	3.58785	9.24024	0.002917
Xi	1	3.5878	3.5878	3.58785	9.24024	0.002917
Error	118	45.8176	SSE=45.8176	MSE=0.38828		
Lack-of-Fit	19	6.4857	6.4857	0.34135	0.85919	0.632449
Pure Error	99	39.3319	39.3319	0.39729		
Total	119	SST=49.4055				

Summary of Model

S = 0.623125	R-Sq = 7.26%	R-Sq(adj) = 6.48%
PRESS = 47.6103	R-Sq(pred) = 3.63%	

$$R^2 = \frac{SSR}{SSTo} = \frac{3.588}{49.405} = 0.0726$$

$$R^2 = 1 - \frac{SSE}{SSTo} = 1 - \frac{45.818}{49.405} = 1 - 0.9274 = 0.0726$$

This means that the 7.26% of change in the mean freshman OPA for students is by ACT test score.

b. Obtain a 95 percent interval estimate of the mean freshman OPA for students whose ACT test score is 28. Interpret your confidence interval. From page 76- to 79

$$E(Y_h) = \hat{Y}_h = b_0 + b_1 X_h$$

$$s^2(\hat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)$$

Confidence Interval for $E(Y_h)$ is: $\hat{Y}_h \pm t_{\left(1-\frac{\alpha}{2};n-2\right)} s\left(\hat{Y}_h\right)$

$$\alpha = 0.05, \qquad \frac{\alpha}{2} = 0.025$$

At $X_h = 28$

$$\hat{Y}_h = 2.114 + 0.0388\,(28) = 3.2012$$

$$s^2(\widehat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)$$

$$s^2(\widehat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = \mathbf{0.3883}\left(\frac{1}{120} + \frac{(28 - 24.725)^2}{2379.925}\right) = 0.004986$$

$$s(\widehat{Y}_h) = \sqrt{0.007776} = 0.0706$$

$$t_{\left(1-\frac{\alpha}{2};n-2\right)} = t_{(0.975;118)} = 1.9807$$

$$3.22012 \pm 1.9807(0.0706)$$

$$3.0614 < E(Y_h) < 3.3410$$

c. Mary Jones obtained a score of 28 on the entrance test. Predict her freshman OPA-using a 95 percent prediction interval. Interpret your prediction interval.

$$s^2(\widehat{Y_{new}}) = MSE\left(1 + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)$$

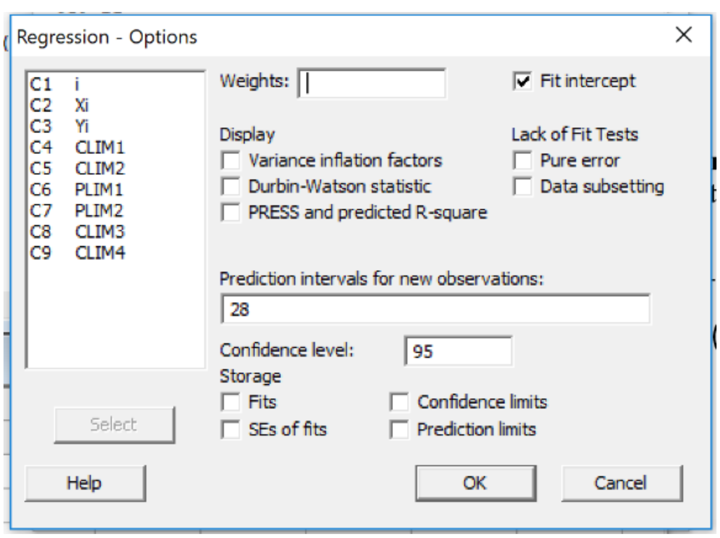
$$\widehat{Y}_h \pm t_{\left(1-\frac{\alpha}{2};n-2\right)} s(\widehat{Y_{new}})$$

$$s^2(\widehat{Y_{new}}) = MSE\left(1 + \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = \mathbf{0.3883}\left(1 + \frac{1}{120} + \frac{(28 - 24.725)^2}{2379.925}\right) = 0.39328$$

$$s(\widehat{Y_{new}}) = 0.6271$$

$$3.22012 \pm 1.9807(0.6271)$$

$$1.9594 < Y_{h(new)} < 4.4430$$



d. Is the prediction interval in part (c) wider than the confidence interval in part (b)? Should it be?

هل فترة الثقة للتنبؤ في الجزء (س) أوسع من فترة الثقة في الجزء (ب)؟ هل يجب أن تكون؟

Yes, Yes

2.15. Refer to Airfreight breakage Problem 1.21.

$$\bar{X} = 1, \quad \sum_{i=1}^{10} (X_i - \bar{X})^2 = 10$$

ANOVA TABLE

Source of Variation	d.f	SS	MS	F	p-value
Regression	1	SSR=160	$MSR = 160$	72.72	0.00
Error	8	SSE=17.6	$MSE = 2.2$		
Total	9	SSTo= 177.6			

a. Because of changes in airline routes, shipments may have to be transferred more frequently than in the past. Estimate the mean breakage for the following numbers of transfers: X = 2, 4. Use separate 99 percent confidence intervals. Interpret your results.

At $X_h = 2$

$\hat{Y}_h = 10.2 + 4\,(2) = 18.2$

$$s^2(\hat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = \mathbf{2.2}\left(\frac{1}{10} + \frac{(2-1)^2}{10}\right) = 0.44$$

$s(\hat{Y}_h) = \sqrt{0.44} = 0.6633$

$t\left(1 - \frac{\alpha}{2}; n-2\right) = t(0.995; 8) = 3.355$

$18.2 \pm 3.355(0.6633)$

$15.976 < E(Y_h) < 20.424$

At $X_h = 4$

$\hat{Y}_h = 10.2 + 4\,(4) = 26.2$

$$s^2(\hat{Y}_h) = MSE\left(\frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right) = \mathbf{2.2}\left(\frac{1}{10} + \frac{(4-1)^2}{10}\right) = 2.2$$

$s(\hat{Y}_h) = \sqrt{2.2} = 1.483$

$t\left(1-\frac{\alpha}{2}; n-2\right) = t(0.995; 8) = 3.355$

$26.2 \pm 3.355(1.483)$

$12.748 < E(Y_h) < 23.652$

We conclude that the mean number of ampules found to be broken upon arrival when 2 transfers from one aircraft to another over the shipment route of 2 are produced is somewhere between 15.976 and 20.424 ampules

أن متوسط عدد الأمبولات التي وجدت مكسوره عند وصولها،وذلك بعد انتقالها مرتين من طائرة إلى آخرى خلال مسار الشحنة , بين 15.976 و 20.424 أمبوله.

We conclude that the mean number of ampules found to be broken upon arrival when 4 transfers from one aircraft to another over the shipment route are produced is somewhere between 12.748 and 23.652 ampules.

أن متوسط عدد أمبولات وجدت منكسره عند وصولهم عندما تم نقله عبر 4 مرات من طائرة واحدة إلى آخر عبر مسار الشحنة , بين 12.748 و 23.652 أمبوله.

b. The next shipment will entail two transfers. Obtain a 99 percent prediction interval for the number of broken ampules for this shipment. Interpret your prediction interval.

$s^2(\widehat{Y_{new}}) = MSE\left(1+\frac{1}{n}+\frac{(X_h-\bar{X})^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\right) = \mathbf{2.2}\left(1+\frac{1}{10}+\frac{(2-1)^2}{10}\right) = 2.64$

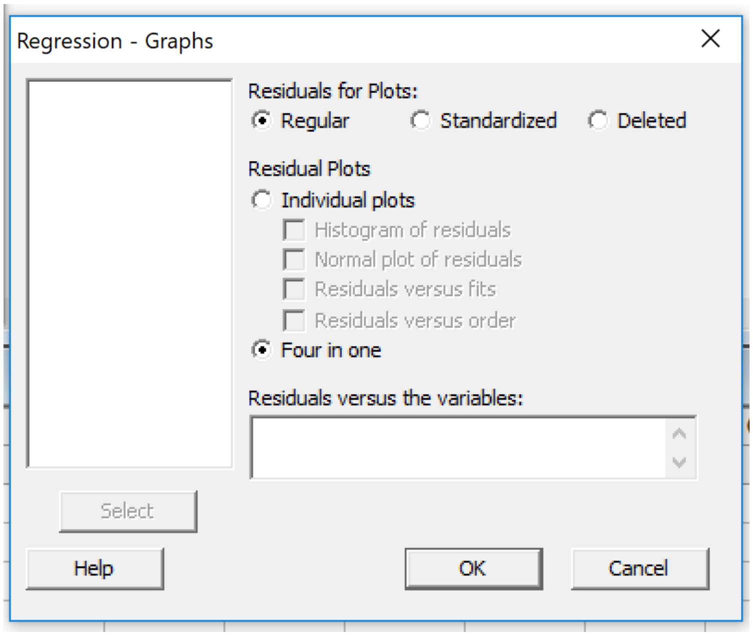
$s(\widehat{Y_h}) = \sqrt{2.64} = 1.6248$

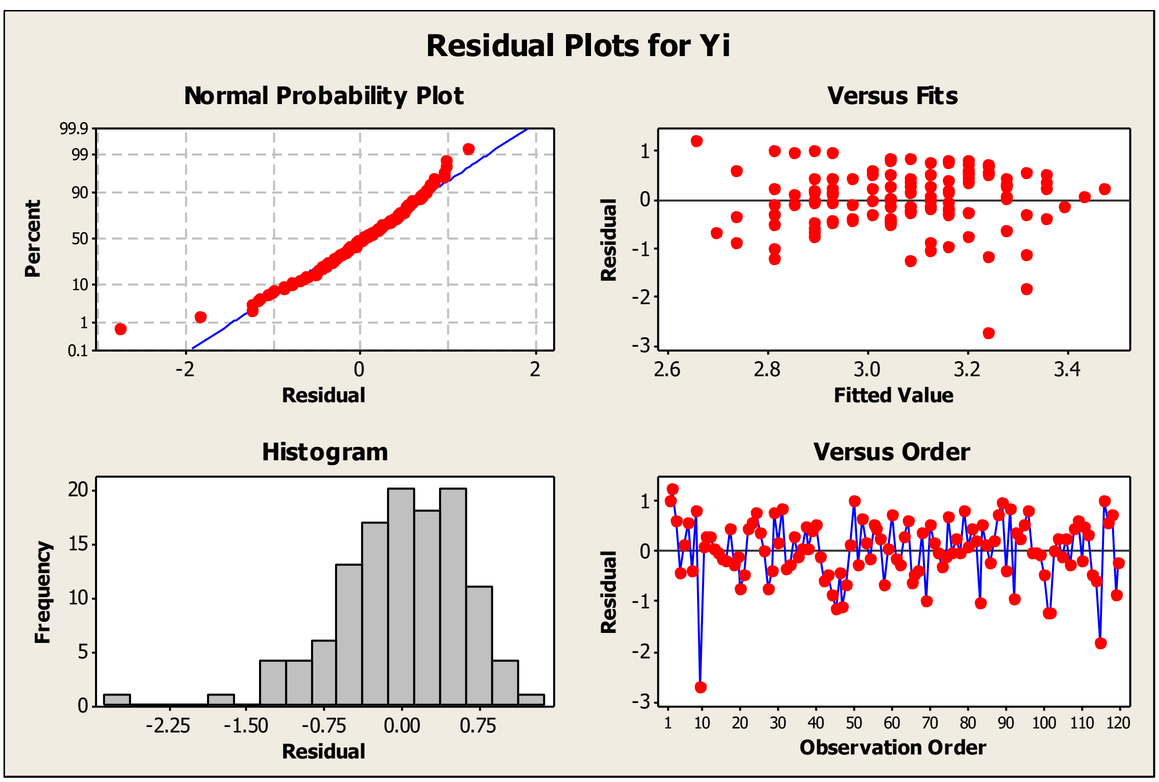
$18.2 \pm 3.355(1.6248)$

$$12.748 < Y_{h(new)} < 23.652$$

With confidence coefficient 0.99, we predict that the mean number of ampules found to be broken upon arrival when 2 transfers from one aircraft to another over the shipment route of 2 are produced is somewhere between 12.748 and 23.652 ampules.

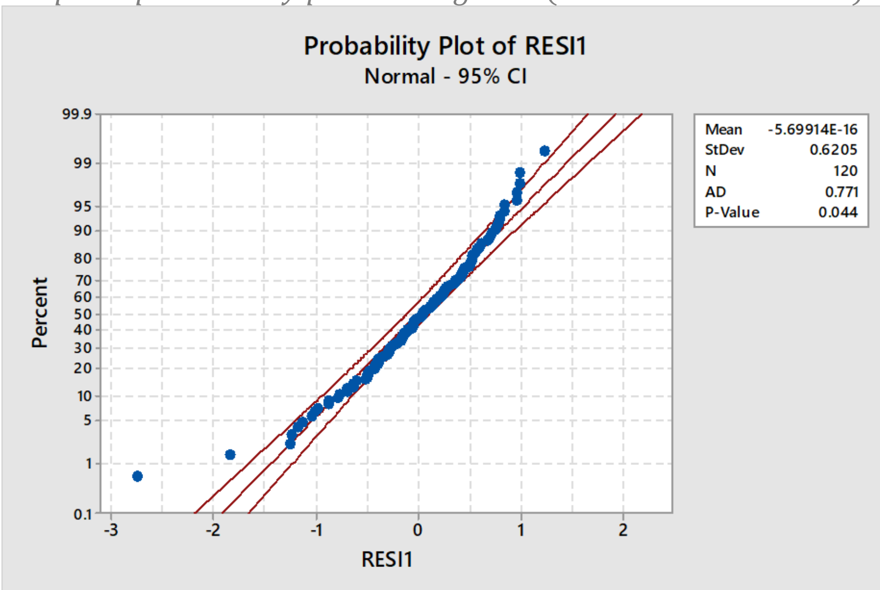
Refer to Plastic hardness.
For test normality of residuals
Stat >> Regression>> Regression >> Graphs





For test normality of residuals

$Graph \rightarrow probability\ plot \rightarrow single \rightarrow (distribution\ Normal) \rightarrow RIS \rightarrow ok$



At 0.01 it is normal but at 0.05 , it is not normal