Chapter 2 HW

Ahmad M. Osman

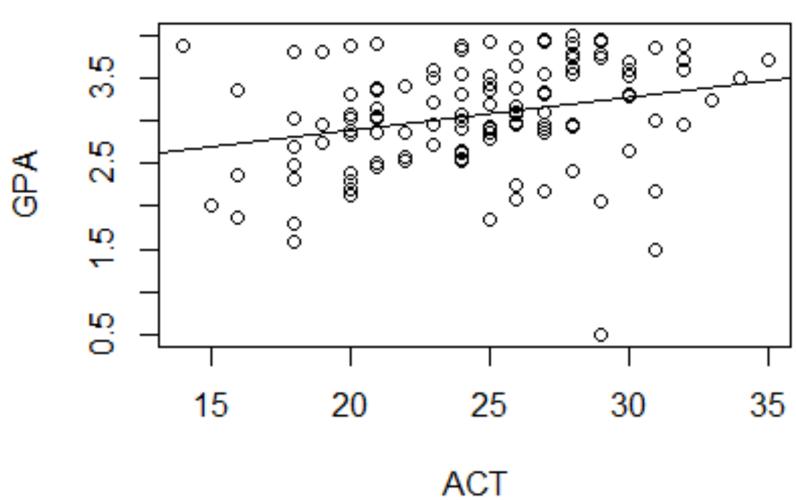
9/18/2017

Chapter 2 - Problems 2.4, 2.13, and 2.23 – GPA Data

Includes answers to:

- 2.4
- 2.13 a, b, and c
- 2.23 a, b, c, d, and e

```
##
## Call:
## lm(formula = GPA ~ ACT)
##
## Coefficients:
## (Intercept)          ACT
##     2.11405      0.03883
```



```
##
## CONFIDENCE LIMITS FOR INTERCEPT AND SLOPE:

##                   0.5 %     99.5 %
## (Intercept) 1.273902675 2.95419590
## ACT         0.005385614 0.07226864

##
## HYPOTHESIS TESTS FOR INTERCEPT AND SLOPE:

##               Estimate Std. Error  t value     Pr(>|t|)
## (Intercept) 2.11404929 0.32089483 6.587982 1.304450e-09
## ACT         0.03882713 0.01277302 3.039777 2.916604e-03

##
## CONFIDENCE INTERVAL FOR A MEAN RESPONSE:

##        fit      lwr      upr
## 1 3.201209 3.084149 3.318269

##
## PREDICTION INTERVAL FOR AN INDIVIDUAL RESPONSE:

##        fit      lwr     upr
## 1 3.201209 2.161538 4.24088

##
## ANOVA TABLE:

## Analysis of Variance Table
##
## Response: GPA
##            Df Sum Sq Mean Sq F value   Pr(>F)
## ACT         1  3.588  3.5878  9.2402 0.002917 **
## Residuals 118 45.818  0.3883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##              DF   Sum Sq
## Corr. Total 119 49.40545

##
## CORRELATION COEFFICIENT:

## [1] 0.2694818
```

Answers Explained:

2.4

a) With a 99% confidence interval, the β1 is between 0.005385614 and 0.07226864. It does not include Zero. Admissions would be interested in that "Zero" because if it exists, then there would be no relationship between the GPA and ACT scores.

b) The T test two alternatives include H_0: $\beta 1 = 0$ and H_a : $\beta 1 \neq 0$. Using the level of significant of 0.01, the confidence interval would be 0.99. According to the decision rule, if $|t^*| \leq t(1 - \alpha/2; n\text{-}2)$, conclude H_0 and if $|t^*| > t(1 - \alpha/2; n\text{-}2)$, conclude H a. In R studio, the results using t(.995, 118) = 2.618137. Since 3.039777 > 2.618137, we can conclude H a, which does not equal 0.

c) From the summary table, the P value = 0.0029 which is less than $\alpha = 0.01$, supporting the H_0.

2.13

a) The 95% confidence interval estimate of the mean freshman GPA for students whose ACT test score is 28 is between 3.061384 and 3.341033. This means that with 95% confidence level, a student with an ACT score of 28 will achieve a GPA between 3.061 and 3.341 in his freshman year. The $\hat{Y}$ estimate is 3.201209.

b) The 95**%** prediction interval estimate of Mary's GPA is between 1.959355 and 4.443063. The fit value remains at 3.201209 which is $\hat{Y}$.

c) Yes, the prediction interval is wider than the confidence interval and it should be as the confidence interval is based on the mean response and the prediction interval is based on a new observation; which is generally because means are less variable than single observations.

2.23

a) It is done above.

b) $\sigma^2 + \beta_1^2 \sum(X_i - \bar{X})^2$ is estimated by MSR which is equal to 3.5878 , and σ^2 is estimated by MSE, which is equal to 0.3883. When $\beta_1 = 0$, both the MSE and MSR estimate the same quantity.

c) According to the ANOVA table, the F value is equal to 9.2402 and has two alternatives: $H_0: \beta_1 = 0$ and $H_a: \beta_1 \neq 0$. Decision rule is:

- If $F^* \leq F(1-\alpha; df_R - df_F, df_F)$, conclude H_0, and if $F^* > F(1-\alpha; df_R - df_F, df_F)$, conclude H_a.

In R studio, the results using $F(1-\alpha; df_R - df_F, df_F)$, = qf(0.99, 1, 118) = 6.854641. And since F (9.24) > 6.8, we conclude H0.

d) SSR = 3.588 is considered to be the absolute magnitude of the reduction in the variation of Y when C is introduced into the regression model. The relative reduction is 0.0726, found by calculating 3.588/49.406 = 0.0726227584.

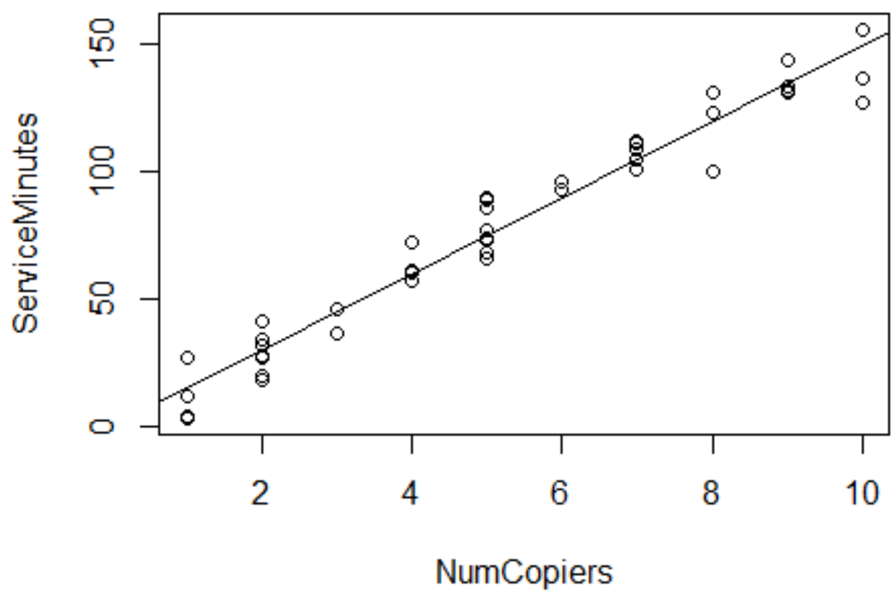
e) r = 0.2694 which is the square root of R^2 found in the previous question. The sign is positive the slope is positive and it is single regression.

Chapter 2 - Problems 2.5 and 2.14 – Copier Data

Includes answers to:

- 2.5 a, d, and e
- 2.14 a and b

```
##
## Call:
## lm(formula = ServiceMinutes ~ NumCopiers)
##
## Coefficients:
## (Intercept)   NumCopiers
##     -0.5802      15.0352
```



```
##
## CONFIDENCE LIMITS FOR INTERCEPT AND SLOPE:

##                   5 %      95 %
## (Intercept) -5.29378  4.133467
## NumCopiers  14.22314 15.847352

## [1] 1.681071

## [1] 0.01891548

##
## HYPOTHESIS TESTS FOR INTERCEPT AND SLOPE:

##                Estimate Std. Error    t value     Pr(>|t|)
## (Intercept) -0.5801567  2.8039411 -0.2069076 8.370587e-01
## NumCopiers  15.0352480  0.4830872 31.1232581 4.009032e-31

##
## CONFIDENCE INTERVAL FOR A MEAN RESPONSE at NumCopiers = 6:

##        fit      lwr     upr
## 1 89.63133 87.28387 91.9788

##
## PREDICTION INTERVAL FOR AN INDIVIDUAL RESPONSE at NumCopiers = 6:

##        fit      lwr      upr
## 1 89.63133 74.46433 104.7983

##
## PREDICTION INTERVAL FOR AN EXPECTED RESPONSE, PER COPIER:

##        fit      lwr     upr
## 1 14.93856 14.54731 15.3298

##
## WORKING-HOTELLING CONFIDENCE BAND (LIMITS) at NumCopiers = 6

##        fit    lower    upper
## 1 89.63133 86.55263 92.71003

##
## ANOVA TABLE:

## Analysis of Variance Table
##
## Response: ServiceMinutes
##            Df Sum Sq Mean Sq F value    Pr(>F)
## NumCopiers  1  76960   76960  968.66 < 2.2e-16 ***
## Residuals  43   3416      79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##               DF  Sum Sq
## Corr. Total 44 80376.8

##
## R SQUARED:

## [1] 0.9574955

##
## CORRELATION COEFFICIENT:

## [1] 0.978517
```

Answers Explained:

2.5

a) t(.95;43). The 0.95 is a result of $\alpha = 0.10$ and $\alpha/2 = 0.05$, so for the t test, $1- \alpha/2 = 0.95$. The 43 is a result of subtracting the n= 45-2. This put together qt(.95, 43) which is equal to 1.681071. The standard error for the number of copiers = 0.483 and with this value, we can compute the confidence interval 15.0352 ± 1.6811(.4831), then we can say $14.2231 = \beta_1 = 15.8473$.

d) The test that was conducted to decide whether this standard is being satisfied by Tri-City or not is t-test. The case here is $H_0: \beta_1 \leq 14$ and $H_a: \beta_1 > 14$. *= (15.0352 – 14) / 0.4831 = 2.1428. and since* $t^* = 2.14$, *we conclude* H_a. The error percentage of a type 1 error is at 0.05, to conclude P value, which is equal to 0.0189.

e) Since $\beta_0 = -0.5802$ which is a negative number, we can say it does not relate to the startup time of calls because of that negative sign. T-test on the hypothesis $H_0 = 0$ shows insignificant evidence to reject that $\beta_0 = 0$. Thus the actual β_0 does not significantly vary from $\beta_0 = 0$.

2. 14

a) Using the predict function, number of copiers set to 6, and the interval set to confidence = 89.63133, 87.28387, and 91.9788. which means that $\hat{Y}_6 = 89.63$ and the 90% confidence interval being $87.28387 \leq \hat{Y}_6 \geq 91.9788$ in minutes.

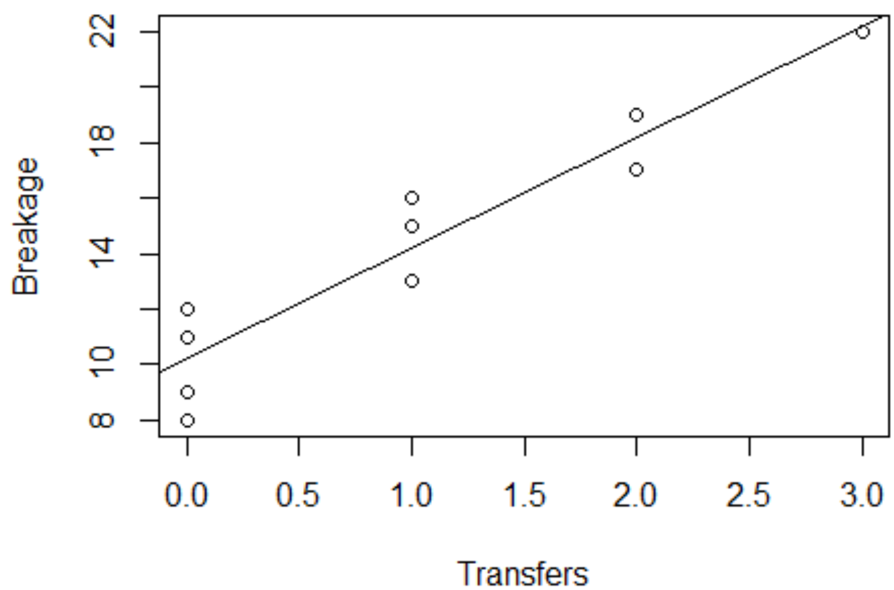
b) The prediction for the next call with 6 copiers needing service is predicted to take 89.63133 minutes, the same as the confidence estimate in part (a). The 90% prediction interval is $74.46433 \leq \hat{Y}_{6\,(new)} \geq 104.7983$ minutes, which is wider than the confidence interval. That is because it is based on a mean value, which is usually more accurate than estimating a new data point.

Chapter 2 - Problems 2.6, 2.15, and 2.25 – Airfreight Data

Includes answers to:

- 2.6 a, c, and d
- 2.15 a and b
- 2.25 a, b, and d

```
##
## Call:
## lm(formula = Breakage ~ Transfers)
##
## Coefficients:
## (Intercept)    Transfers
##        10.2          4.0
```




```
## [1] 2.306004

## [1] 0.05394531

##
## CONFIDENCE LIMITS FOR INTERCEPT AND SLOPE:
```

```
##                 2.5 %    97.5 %
## (Intercept) 8.670370 11.729630
## Transfers   2.918388  5.081612

##
## HYPOTHESIS TESTS FOR INTERCEPT AND SLOPE:

##             Estimate Std. Error   t value     Pr(>|t|)
## (Intercept)     10.2  0.6633250 15.377079 3.178273e-07
## Transfers        4.0  0.4690416  8.528029 2.748669e-05

##
## CONFIDENCE LIMITS FOR INTERCEPT AND SLOPE:

##                1.25 %   98.75 %
## (Intercept) 8.374846 12.025154
## Transfers   2.709421  5.290579

##
## CONFIDENCE INTERVAL FOR A MEAN RESPONSE at Transfers = 2 or 4:

##              fit      lwr      upr
## Transfers=2 18.2 15.97429 20.42571
## Transfers=4 26.2 21.22316 31.17684

##
## PREDICTION INTERVAL FOR AN INDIVIDUAL RESPONSE at Transfers = 2:

##    fit      lwr      upr
## 1 18.2 12.74814 23.65186

##
## PREDICTION LIMITS FOR THE MEAN OF 3 NEW OBSERVATIONS at Transfers=2:

##    fit    lower    upper
## 1 18.2 15.15908 21.24092

##
## PREDICTION LIMITS FOR THE TOTAL OF 3 NEW OBSERVATIONS at Transfers=2:

##    fit    lower    upper
## 1 54.6 45.47724 63.72276

##
## WORKING-HOTELLING CONFIDENCE BAND (LIMITS) at Transfers = 2 or 4

##    fit    lower    upper
## 1 18.2 15.44116 20.95884
## 2 26.2 20.03104 32.36896

##
## ANOVA TABLE:
```

```
## Analysis of Variance Table
##
## Response: Breakage
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Transfers  1  160.0   160.0  72.727 2.749e-05 ***
## Residuals  8   17.6     2.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##             DF Sum Sq
## Corr. Total  9  177.6

## [1] 5.317655

##
## R SQUARED:

## [1] 0.9009009

##
## CORRELATION COEFFICIENT:

## [1] 0.949158

## [1] 0.978517
```

Answers Explained:

2.6

a) n = 10, α = 0.05 and α/2 = 0.025, so (1- α/2) = 0.975 and n-2 = 8. Therefor qt(0.975;8) is equal to 2.306004. The confidence interval becomes 4.0 ± 2.306(.469). With 95% confidence, the estimate is 4.0 ampules per transfer.

c) Using the confint(fit,level=.95) the 95% confident interval for β_0 is equal to 8.670370 ampules ≤ β_0 ≥ 11.729630 ampules. The estimation is 10.2 ampules which is within the 95% confident interval. There is about 10 ampules that are broken even when no transfers are made.

d) An appropriate test for that would be a t test with alternatives H_0: $\beta_0 \leq 9$, and H_a: $\beta_0 > 9$. *The t value from this function is 2.306004, leading us to conclude* H_0. The P value is 0.05394531.

2.15

a) When X=2, with 99% confidence interval, at transfers 2 or 4. For those two transfers, the fit breakage is 18.2 ampules and the confidence interval is between 15.97 and 20.42. When X=4 the fit value is 26.2 and the confidence interval is between 21.22 and 31.17 ampules. The number of ampules broken between 2 and 4 transfers lie within the stated intervals.

b) When the next shipment has 2 transfers and 99% prediction interval. The broken ampules for this shipment is between 12.748 and 23.65. The fit value remains 18.2. But since this is just a prediction, usually the prediction interval is wider than the two transfers confidence interval in the previous part.

2.25

a) The ANOVA table is done above. The additive elements are the columns Sum Sq and df.

b) With $\alpha = 0.05$, the alternatives are: $H_0: \beta_1 = 0$, and $H_a: \beta_1 = 0$. The qf(.95, 1, 8) is equal to 5.317655. The decision rule is: If $F^* \leq 5.317655$, conclude H_0, and if $F^* > 5.317655$, then conclude H_a. F is 72.727 (found from the ANOVA table) and that value is larger than 5.31 so we conclude H_a.

d) The $R^2 = 0.9009009$ and $r = 0.949158$. The proportion of the variation in Y is accounted for by introducing X into the regression model is about 90%.