Stat 403 - Solution to Assignment 8

Turned in Thursday, November 30, 2000

1. Below are the scores for eleven women golfers for a tournament that consisted of two rounds of golf.

Golfer	11	6	8	5	3	9	1	2	10	4	7
Round 1, X	79	81	83	86	87	88	89	90	91	95	115
Round 2, Y	80	76	87	81	89	91	94	85	88	89	89

(a) Compute the number of concordant and discordant pairs and any ties.

# Concordant	9	9	6	7	2	1	0	3	2	0	39
# Discordant	1	0	2	0	2	4	4	0	0	0	13
Ties	0	0	0	0	2	0	0	0	0	1	3

Kendall's T: $T = \frac{39-13}{55} = \frac{26}{55} = 0.473$

corrected: $T = \frac{39-13}{\sqrt{(55)(55-3)}} = \frac{26}{53.5} = 0.486$

(b) The one-sided P-value is between 0.025 and 0.050. The two-sided P-value is between 0.050 and 0.100. Since the two-sided P-value is not small, there is not a statistically significant association between the scores for the two rounds.

(c) Compute the number of concordant and discordant pairs and any ties.

# Concordant	8	8	5	6	2	1	0	2	1	33
# Discordant	1	0	2	0	2	3	3	0	0	11
Ties	0	0	0	0	1	0	0	0	0	1

Kendall's T: $T = \frac{33-11}{45} = \frac{22}{45} = 0.489$

corrected: $T = \frac{33-11}{\sqrt{(55)(55-1)}} = \frac{22}{44.5} = 0.494$

(d) The one-sided P-value is between 0.023 and 0.036. The two-sided P-value is between 0.046 and 0.072. Using the two-sided P-value we cannot tell for sure without interpolation. However, it is likely that the two-sided P-value is not smaller than 0.05. Therefore, there is not a statistically significant association between the scores for the two rounds.

(e) $\hat{Y} = 63.6 + 0.254(115) = 63.6 + 29.2 = 92.8$ Since golf scores are in whole numbers, the predicted score is 93 with a residual for golfer #7 of 89 - 93 = -4

(f) Summaries: $r_S = 0.606 \quad s_{R(X)} = 3.317 \quad s_{R(Y)} = 3.286$

slope estimate: $r_S \frac{s_{R(Y)}}{s_{R(X)}} = 0.606 \frac{3.286}{3.317} = 0.600$

intercept estimate: $\overline{R(Y)} - slope * \overline{R(X)} = 6 - 0.600 * 6 = 6 - 3.6 = 2.4$

equation: $R(\hat{Y}) = 2.4 + 0.6R(X)$

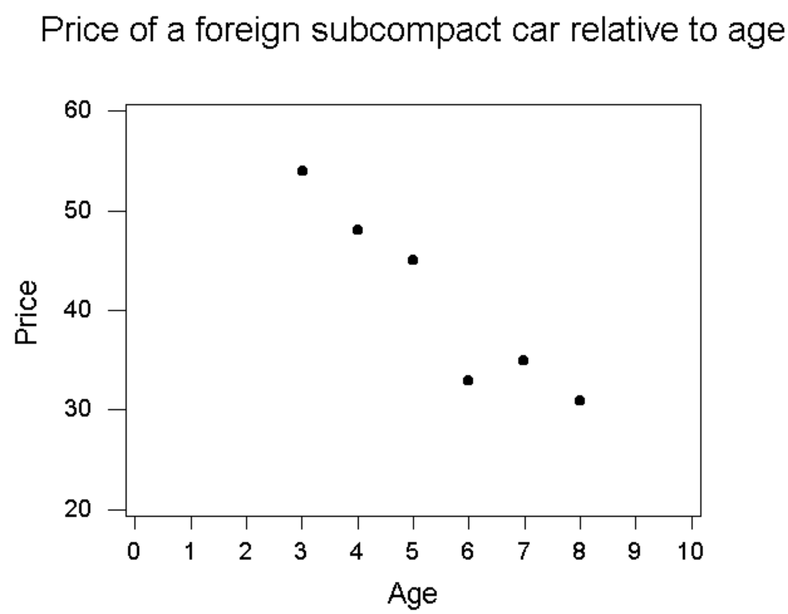
(g) For R(X) = 11, $R(\hat{Y}) = 2.4 + 0.6(11) = 9$, the 9^{th} Y value in order is 89 so the predicted score is 89 and the residual is 89 - 89 = 0.

(h) The rank regression does a better job of predicting the second round score given an unusual first round score. Whether the first round score was a 115 or a 96 we would get the same prediction. This just illustrates that rank regression is not affected as much by unusual or influential points.

2. In an attempt to relate the age (in years), X, of a foreign subcompact car and that car's asking price (in $100), Y, six used cars of different ages are taken and information on asking price is obtained. The data are given below.

Age (years)	3	4	5	6	7	8
Price ($100)	54	48	45	33	35	31

(a) Plot of the car price in $100 versus Age in years.



(b) There is a very strong negative linear relationship between Age and Price. As the Age of the car goes up the Price tends to come down in a linear fashion.

(c) Compute the number of concordant and discordant pairs and any ties.

# Concordant	0	0	0	1	0	1
# Discordant	5	4	3	1	1	14

Kendall's T: $T = \frac{1-14}{15} = \frac{-13}{15} = -0.867$

The one-sided P-value is 0.008. There is a very strong monotone decreasing relationship between Age and Price. As the Age of the car goes up the Price tends to come down.

(d) For a 7 year old car, the predicted price is:

$\hat{Y} = 67.1 - 4.74(7) = 67.1 - 33.2 = 33.9$ or \$3,390

The residual is 35 - 33.9 = 1.1 or \$110.

(e) Compute the 15 pairwise slopes.

	i=1	i=2	i=3	i=4	i=5
j=2	−6.00				
j=3	−4.50	−0.15			
j=4	−7.00	−4.33	−12.00		
j=5	−4.75	−4.33	−5.00	2.00	
j=6	−4.60	−4.25	−4.67	−1.00	−4.00

Median slope $= -4.60$

intercept = Median Y - slope*Median X $= 40 - (-4.6)(5.5) = 40 + 25.3 = 65.3$

equation: $\tilde{Y} = 65.3 - 4.6X$

(f) For a 7 year old car, the predicted price is:

$\tilde{Y} = 65.3 - 4.60(7) = 65.3 - 32.2 = 33.1$ or \$3,310

The residual is 35 - 33.1 = 1.9 or \$190.

(g) The prediction using the least squares line has a smaller residual and so is better. It is interesting to note that the Theil line falls below the least squares line. The only place where the Theil line gives a better prediction is for the 6 year old car.