Week #11

- What does the regression line minimise? Draw a picture
- How many regression lines are possible for a given data set?
- Give an example of restriction of range
- Explain the difference between a univariate outlier and a regression outlier, and draw a picture
- Draw an example of how an influential outlier might affect the regression line

Regression

$\bar{Y} = \Sigma(Y_i)/n$
$df_1 = 1$
$df_2 = n - df_1 - 1$
$SP = \Sigma[(X_i - \bar{X})(Y_i - \bar{Y})]$
$SS_X = \Sigma[(X_i - \bar{X})^2]$
$\beta_1 = SP/SS_X$
$\beta_0 = \bar{Y} - \beta_1 \times \bar{X}$

$\hat{Y}_i = \beta_0 + X_i \times \beta_1$
$SS_{\text{tot}} = \Sigma[(Y_i - \bar{Y})^2]$
$SS_{\text{reg}} = \Sigma[(\hat{Y} - \bar{Y})^2]$
$SS_{\text{res}} = SS_{\text{tot}} - SS_{\text{reg}}$
$MS_{\text{reg}} = SS_{\text{reg}}/df_1$
$MS_{\text{res}} = SS_{\text{res}}/df_2$
$F = MS_{\text{reg}}/MS_{\text{res}}$

Critical *F* values

		df_1		
df_2	α	1	2	3
1	0.05	161.4	199.5	215.71
	0.01	4052	4999	5404
2	0.05	18.51	19	19.16
	0.01	98.94	99	99.17
3	0.05	7.71	6.94	6.59
	0.01	34.12	30.82	29.46
4	0.05	7.71	6.94	6.59
	0.01	21.2	18	16.69

Question #1

Test the model fit at an α of 0.05.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
2	4	1	1	1	4.57	-0.43	0.18
1	7	4	4	-4	4.14	-0.86	0.73
6	8	9	9	9	6.29	1.29	1.65
3	1	0	16	0	5.00	0.00	0.00

$SS_X = 14$
$SP = 6$
$\beta_1 = 6/14 = 0.43$
$\bar{Y} = 5$
$\bar{X} = 3$
$\beta_0 = 5 - (0.43 \times 3) = 3.71$
$\hat{Y}_i = 3.71 + (0.43 \times X_i)$
$SS_{\text{tot}} = 30$
$SS_{\text{reg}} = 2.57$

$SS_{\text{res}} = 30 - 2.57 = 27.43$
$df_1 = 1$
$df_2 = 4 - 1 - 1 = 2$
$MS_{\text{reg}} = 2.57/1 = 2.57$
$MS_{\text{res}} = 27.43/2 = 13.71$
$F = 2.57/13.71 = 0.19$
$F_{\text{crit}} = 18.51$
Fail to reject because $0.19 < 18.51$

Question #2

Test the model fit at an α of 0.01.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
4	5	1	0	0	5.07	0.07	0.01
0	6	9	1	-3	4.79	-0.21	0.05
5	7	4	4	4	5.14	0.14	0.02
3	2	0	9	0	5.00	0.00	0.00

$SS_X = 14$
$SP = 1$
$\beta_1 = 1/14 = 0.07$
$\bar{Y} = 5$
$\bar{X} = 3$
$\beta_0 = 5 - (0.07 \times 3) = 4.79$
$\hat{Y}_i = 4.79 + (0.07 \times X_i)$
$SS_{\text{tot}} = 14$
$SS_{\text{reg}} = 0.07$

$SS_{\text{res}} = 14 - 0.07 = 13.93$
$df_1 = 1$
$df_2 = 4 - 1 - 1 = 2$
$MS_{\text{reg}} = 0.07/1 = 0.07$
$MS_{\text{res}} = 13.93/2 = 6.96$
$F = 0.07/6.96 = 0.01$
$F_{\text{crit}} = 98.5$
Fail to reject because $0.01 < 98.5$

Question #3

Test the model fit at an α of 0.01.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
1	9	9	9	-9	9.90	3.90	15.21
2	10	4	16	-8	8.60	2.60	6.76
5	4	1	4	-2	4.70	-1.30	1.69
8	1	16	25	-20	0.80	-5.20	27.04

$SS_X = 30$
$SP = -39$
$\beta_1 = -39/30 = -1.3$
$\bar{Y} = 6$
$\bar{X} = 4$
$\beta_0 = 6 - (-1.3 \times 4) = 11.2$
$\hat{Y}_i = 11.2 + (-1.3 \times X_i)$
$SS_{\text{tot}} = 54$
$SS_{\text{reg}} = 50.7$

$SS_{\text{res}} = 54 - 50.7 = 3.3$
$df_1 = 1$
$df_2 = 4 - 1 - 1 = 2$
$MS_{\text{reg}} = 50.7/1 = 50.7$
$MS_{\text{res}} = 3.3/2 = 1.65$
$F = 50.7/1.65 = 30.73$
$F_{\text{crit}} = 98.5$
Fail to reject because $30.73 < 98.5$

Question #4

Test the model fit at an α of 0.05.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
2	7	9	0	0	7.75	0.75	0.56
6	10	1	9	3	6.75	-0.25	0.06
8	4	9	9	-9	6.25	-0.75	0.56
9	6	16	1	-4	6.00	-1.00	1.00
0	8	25	1	-5	8.25	1.25	1.56

$SS_X = 60$
$SP = -15$
$\beta_1 = -15/60 = -0.25$
$\bar{Y} = 7$
$\bar{X} = 5$
$\beta_0 = 7 - (-0.25 \times 5) = 8.25$
$\hat{Y}_i = 8.25 + (-0.25 \times X_i)$
$SS_{\text{tot}} = 20$
$SS_{\text{reg}} = 3.75$

$SS_{\text{res}} = 20 - 3.75 = 16.25$
$df_1 = 1$
$df_2 = 5 - 1 - 1 = 3$
$MS_{\text{reg}} = 3.75/1 = 3.75$
$MS_{\text{res}} = 16.25/3 = 5.42$
$F = 3.75/5.42 = 0.69$
$F_{\text{crit}} = 10.13$
Fail to reject because $0.69 < 10.13$

Question #5

Test the model fit at an α of 0.01.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
6	2	1	4	2	3.60	-0.40	0.16
8	10	1	36	6	4.40	0.40	0.16
9	1	4	9	-6	4.80	0.80	0.64
5	3	4	1	2	3.20	-0.80	0.64

$SS_X = 10$
$SP = 4$
$\beta_1 = 4/10 = 0.4$
$\bar{Y} = 4$
$\bar{X} = 7$
$\beta_0 = 4 - (0.4 \times 7) = 1.2$
$\hat{Y}_i = 1.2 + (0.4 \times X_i)$
$SS_{\text{tot}} = 50$
$SS_{\text{reg}} = 1.6$

$SS_{\text{res}} = 50 - 1.6 = 48.4$
$df_1 = 1$
$df_2 = 4 - 1 - 1 = 2$
$MS_{\text{reg}} = 1.6/1 = 1.6$
$MS_{\text{res}} = 48.4/2 = 24.2$
$F = 1.6/24.2 = 0.07$
$F_{\text{crit}} = 98.5$
Fail to reject because $0.07 < 98.5$

Question #6

Test the model fit at an α of 0.01.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
1	5	16	4	8	6.57	-0.43	0.19
2	8	9	1	-3	6.67	-0.33	0.11
6	9	1	4	2	7.11	0.11	0.01
9	7	16	0	0	7.43	0.43	0.19
7	6	4	1	-2	7.22	0.22	0.05

$SS_X = 46$
$SP = 5$
$\beta_1 = 5/46 = 0.11$
$\bar{Y} = 7$
$\bar{X} = 5$
$\beta_0 = 7 - (0.11 \times 5) = 6.46$
$\hat{Y}_i = 6.46 + (0.11 \times X_i)$
$SS_{\text{tot}} = 10$
$SS_{\text{reg}} = 0.54$

$SS_{\text{res}} = 10 - 0.54 = 9.46$
$df_1 = 1$
$df_2 = 5 - 1 - 1 = 3$
$MS_{\text{reg}} = 0.54/1 = 0.54$
$MS_{\text{res}} = 9.46/3 = 3.15$
$F = 0.54/3.15 = 0.17$
$F_{\text{crit}} = 34.12$
Fail to reject because $0.17 < 34.12$

Question #7

Test the model fit at an α of 0.05.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
6	2	1	4	-2	4.15	0.15	0.02
4	1	1	9	3	3.85	-0.15	0.02
8	3	9	1	-3	4.46	0.46	0.21
2	6	9	4	-6	3.54	-0.46	0.21
9	8	16	16	16	4.62	0.62	0.38
1	4	16	0	0	3.38	-0.62	0.38

$SS_X = 52$
$SP = 8$
$\beta_1 = 8/52 = 0.15$
$\bar{Y} = 4$
$\bar{X} = 5$
$\beta_0 = 4 - (0.15 \times 5) = 3.23$
$\hat{Y}_i = 3.23 + (0.15 \times X_i)$
$SS_{\text{tot}} = 34$
$SS_{\text{reg}} = 1.23$

$SS_{\text{res}} = 34 - 1.23 = 32.77$
$df_1 = 1$
$df_2 = 6 - 1 - 1 = 4$
$MS_{\text{reg}} = 1.23/1 = 1.23$
$MS_{\text{res}} = 32.77/4 = 8.19$
$F = 1.23/8.19 = 0.15$
$F_{\text{crit}} = 7.71$
Fail to reject because $0.15 < 7.71$

Question #8

Test the model fit at an α of 0.05.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
10	6	25	4	10	5.72	1.72	2.97
7	2	4	4	-4	4.69	0.69	0.48
5	5	0	1	0	4.00	0.00	0.00
3	7	4	9	-6	3.31	-0.69	0.48
0	0	25	16	20	2.28	-1.72	2.97

$SS_X = 58$
$SP = 20$
$\beta_1 = 20/58 = 0.34$
$\bar{Y} = 4$
$\bar{X} = 5$
$\beta_0 = 4 - (0.34 \times 5) = 2.28$
$\hat{Y}_i = 2.28 + (0.34 \times X_i)$
$SS_{\text{tot}} = 34$
$SS_{\text{reg}} = 6.9$

$SS_{\text{res}} = 34 - 6.9 = 27.1$
$df_1 = 1$
$df_2 = 5 - 1 - 1 = 3$
$MS_{\text{reg}} = 6.9/1 = 6.9$
$MS_{\text{res}} = 27.1/3 = 9.03$
$F = 6.9/9.03 = 0.76$
$F_{\text{crit}} = 10.13$
Fail to reject because $0.76 < 10.13$

Question #9

Test the model fit at an α of 0.05.

X_i	Y_i	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}_i$	$(\hat{Y}_i - \bar{Y})$	$(\hat{Y}_i - \bar{Y})^2$
2	4	4	1	2	5.68	0.68	0.46
10	3	36	4	-12	2.96	-2.04	4.16
3	6	1	1	-1	5.34	0.34	0.12
1	7	9	4	-6	6.02	1.02	1.04

$SS_X = 50$
$SP = -17$
$\beta_1 = -17/50 = -0.34$
$\bar{Y} = 5$
$\bar{X} = 4$
$\beta_0 = 5 - (-0.34 \times 4) = 6.36$
$\hat{Y}_i = 6.36 + (-0.34 \times X_i)$
$SS_{\text{tot}} = 10$
$SS_{\text{reg}} = 5.78$

$SS_{\text{res}} = 10 - 5.78 = 4.22$
$df_1 = 1$
$df_2 = 4 - 1 - 1 = 2$
$MS_{\text{reg}} = 5.78/1 = 5.78$
$MS_{\text{res}} = 4.22/2 = 2.11$
$F = 5.78/2.11 = 2.74$
$F_{\text{crit}} = 18.51$
Fail to reject because $2.74 < 18.51$