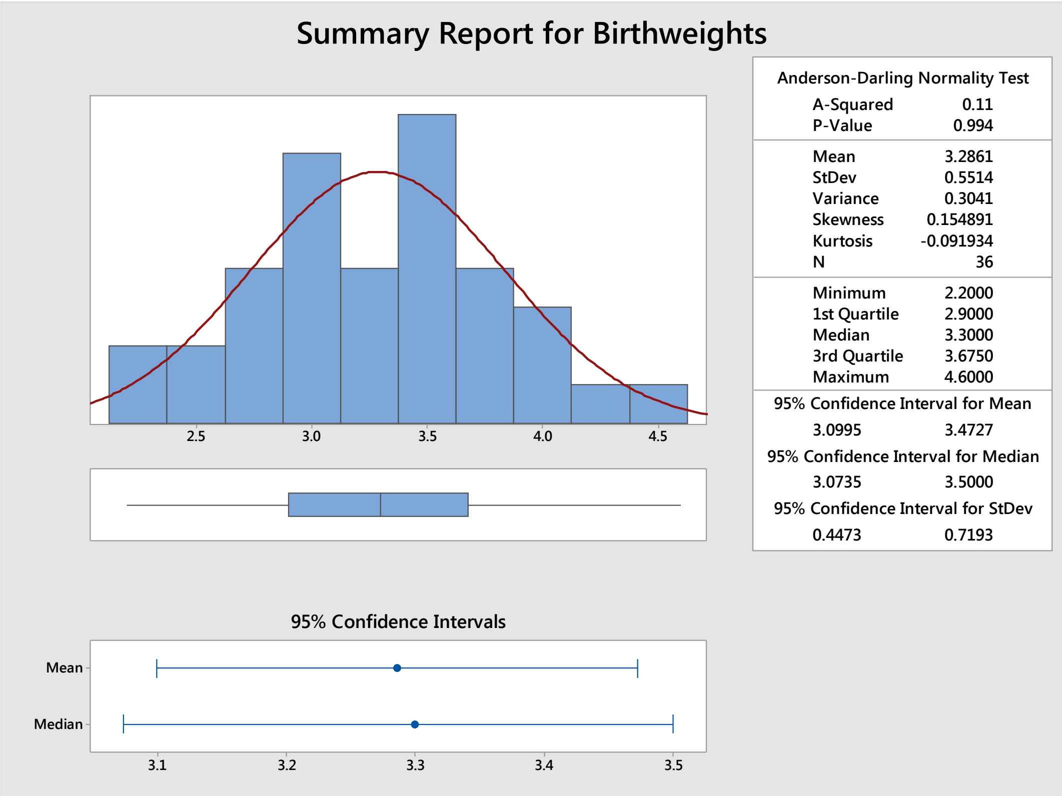
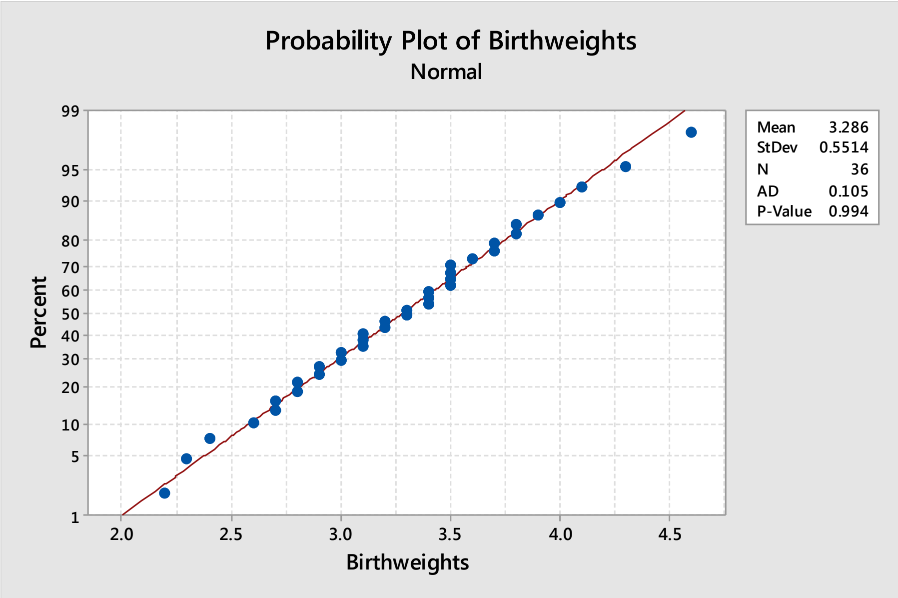
MATU9D2 : Practical Statistics Spring 2017

Solutions to Practical 7 : Minitab

Question 1

The data is quantitative, one sample, small sample (n<30) and the question asks about the standard deviation. Formal technique to answers the question would be a Chisquared Interval for the standard deviation. This assumes that the data is Normally distributed.





Normal Probability Plot looks linear so we can assume that the data follows a Normal distribution.

Test and CI for One Variance: Weights

Method

Null hypothesis	σ = 0.6
Alternative hypothesis	σ ≠ 0.6

The chi-square method is only for the normal distribution.
The Bonett method is for any continuous distribution.

Statistics

Variable	N	StDev	Variance
Weights	36	0.551	0.304

95% Confidence Intervals

Variable	Method	CI for StDev	CI for Variance
Weights	**Chi-Square**	**(0.447, 0.719)**	(0.200, 0.517)
	Bonett	(0.447, 0.720)	(0.200, 0.518)

Tests

Variable	Method	Test Statistic	DF	P-Value
Weights	Chi-Square	29.56	35	0.545
	Bonett	–	–	0.520

H_0: σ = 0.6 H_1 σ ≠ 0.6
95% certain that standard deviation of birthweights lies between 0.45 and 0.72kgs.
So we cannot reject H_0 in favour of H_1 at 5% level i.e. insufficient evidence that the standard deviation of birthweights is significantly different to 0.6kgs.

Data is consistent, with 95% confidence that standard deviation equals 0.6kgs.

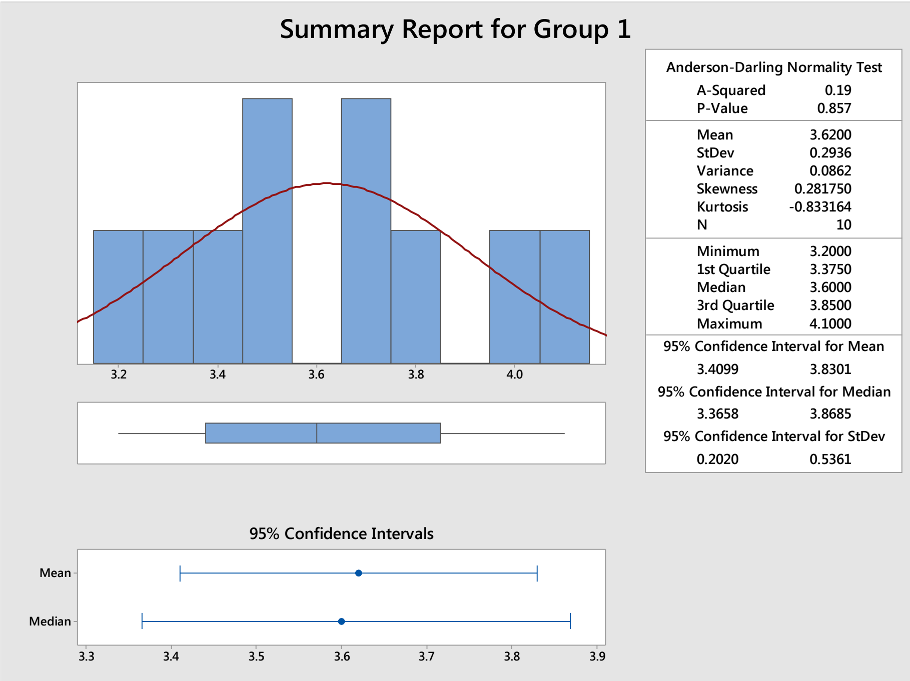
Note that this interval is also presented in the graphical summary above.

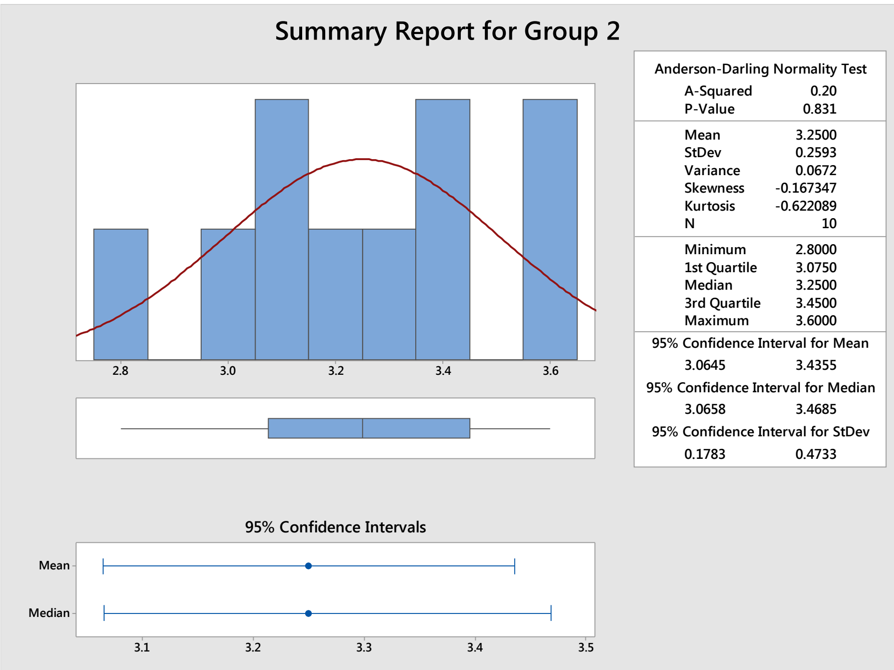
Question 2

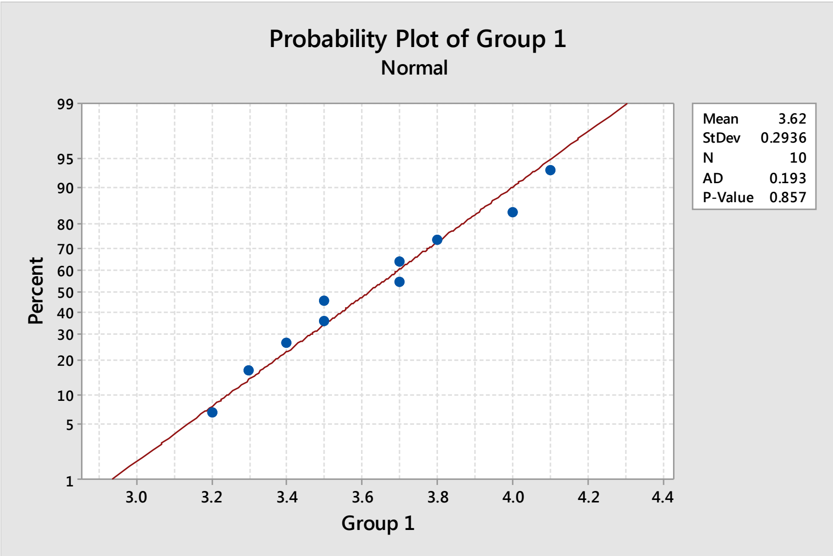
The data is quantitative, two independent samples and the question asks about comparing the variances. An F test to compare variances should be performed.

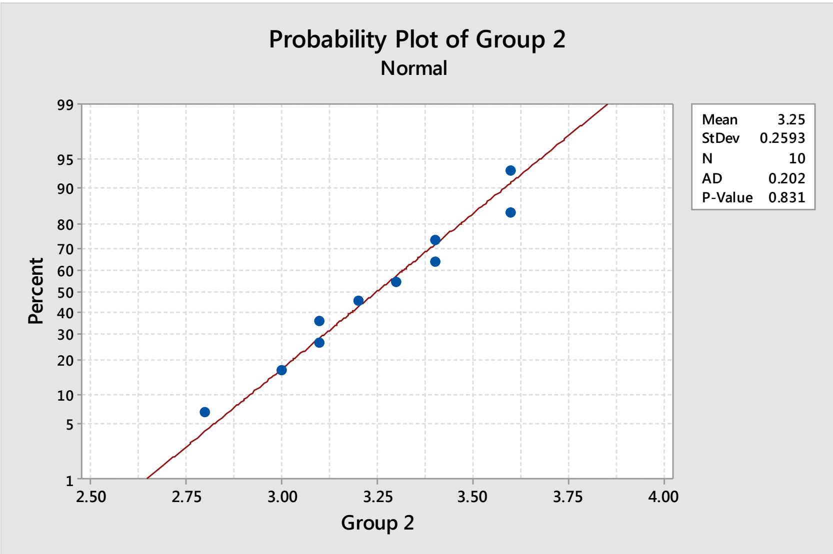
This assumes that the data in both samples is Normally distributed.
The assumption of Normality is valid if the data follows a Normal distribution.

Examining the Normal probability plots below, we can assume that the data from both groups follow Normal distribution since the plots are approximately linear.







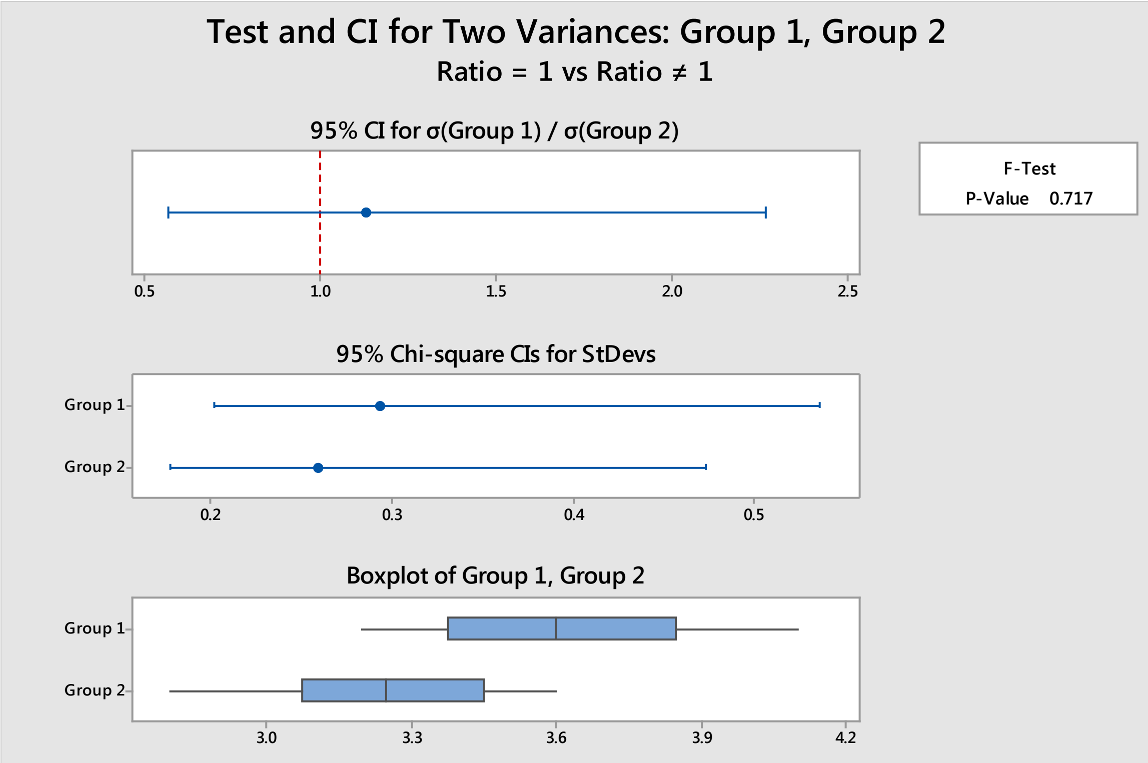


$H_o: \quad \sigma_1^2 = \sigma_2^2 \quad H_1 \quad \sigma_1^2 \neq \sigma_2^2$

See the F test results on the next page

p = 0.717 >0.05 so we cannot reject H_o in favour of H_1 at 5% level so insufficient evidence that the variance in drug levels are significantly different.

So we can assume that the variances are equal and the unpaired t test completed is appropriate.



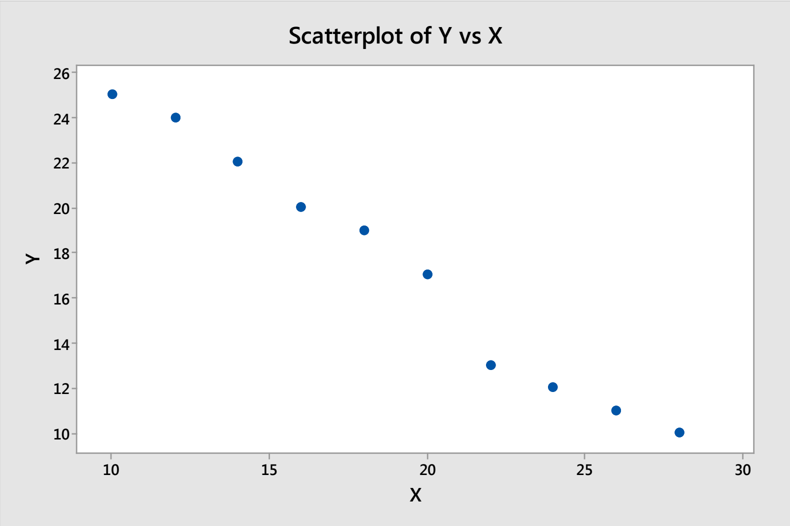
Test and CI for Two Variances: Group 1, Group 2

```
Method

Null hypothesis          σ(Group 1) / σ(Group 2) = 1
Alternative hypothesis   σ(Group 1) / σ(Group 2) ≠ 1
Significance level       α = 0.05

F method was used. This method is accurate for normal data only.

Statistics
                                      95% CI for
Variable   N  StDev  Variance          StDevs
Group 1   10  0.294     0.086  (0.202, 0.536)
Group 2   10  0.259     0.067  (0.178, 0.473)

Ratio of standard deviations = 1.133
Ratio of variances = 1.283

95% Confidence Intervals
                                CI for
          CI for StDev        Variance
Method        Ratio             Ratio
F        (0.564, 2.272)  (0.319, 5.164)


Tests

                         Test
Method  DF1  DF2    Statistic  P-Value
F         9    9         1.28    0.717
```

Repeats the result of the F test given in the graphic above (p=0.717) but also includes the Observed Test Statistic (1.28)

Question 3

(a)



Subjective Impression

Negative linear relationship

(b) & (c) Correlation: X, Y

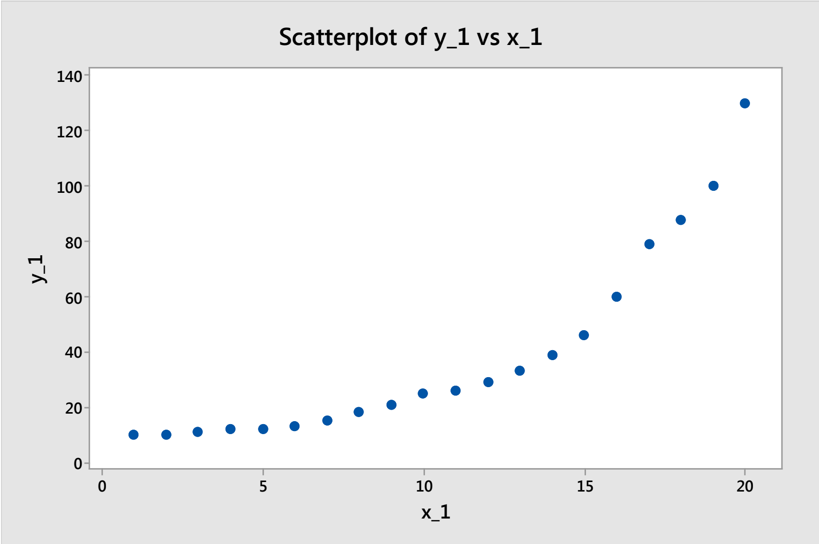
```
Pearson correlation of X and Y = -0.991
P-Value = 0.000
```

r = -0.991 (as the data is almost all on a straight line in the negative direction close to -1)

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

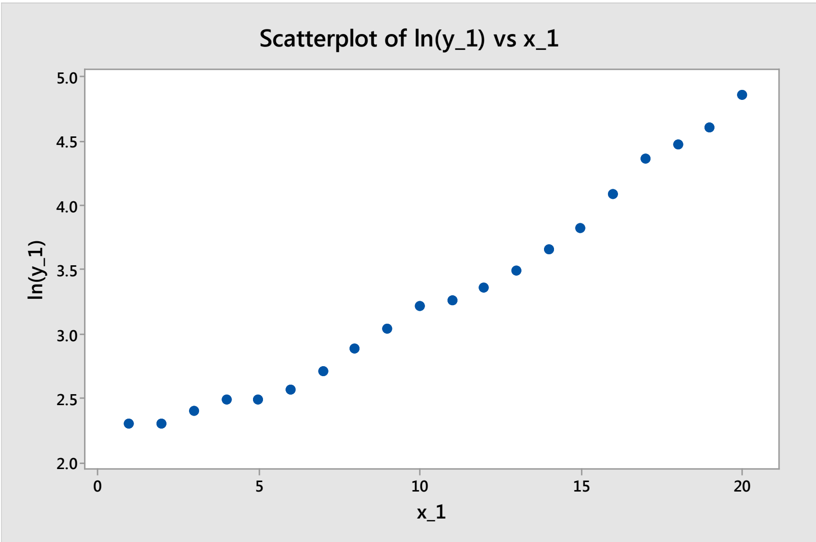
$p < 0.001$ so can reject H_o in favour of H_1 at 1% level and conclude that the correlation is significantly different to zero. i.e. significant relationship

Question 4



Subjective Impression

Positive non- linear relationship



Subjective Impression

Transformation has linearised the relationship i.e. linear relationship between ln(y) and x

Correlation: x_1, y_1, ln(y_1)

```
              x_1      y_1
y_1         0.887
            0.000

ln(y_1)     0.985    0.945
            0.000    0.000


Cell Contents: Pearson correlation
               P-Value
```

y vs x

r = 0.887 & p<0.001 so if we had not looked at the graph we would have said that there is a significant relationship.

ln(y) vs x

r = 0.985 & p<0.001
r has increased since closer to a straight line and since p<0.001 can reject Ho in favour of H1 and conclude correlation significantly different to zero.