Essential Mathematics & Statistics for Science: Second Edition by Graham Currell & Antony Dowman (John Wiley & Sons)

Answers to 'Q' Questions

11 ANOVA – Analysis of Variance

To navigate, use the Bookmarks in the PDF file.

Q11.1

$\sigma_{TOT}^2 = \sigma_E^2 + \sigma_F^2$ from [11.1]

$\sigma_{TOT}^2 = 0.05^2 + 0.08^2 = 0.0089$

$\sigma_{TOT} = \sqrt{0.0089} = 0.094$

Q11.2

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-value*	*F crit*
Between Groups	67.7072	4	16.9268	3.22		3.22
Within Groups	236.4800	45	5.255			
Total	304.1872	49				

i) [11.5] Numerator $\boldsymbol{df_B = k - 1}$ and Denominator, $\boldsymbol{df_W = k(n - 1)}$

$k = 5$, $n = 10$ so $df_B = 5 - 1 = 4$ and $df_W = 5(10 - 1) = 45$. The total degrees of freedom will be 4 + 45 = 49 (one less than 50, the total number of values)

$MS_B = SS_B / df_B = 67.7072/4 = 16.9268$

$MS_W = SS_W / df_W = 236.4800/45 = 5.255$

From [11.4] $F = \dfrac{MS_B}{MS_W} = 16.9268/5.255 = 3.22$

F_{CRIT} can be obtained from tables, 1-tailed for $df_B = 4$ (numerator) and $df_W = 45$ (denominator) = 2.58. If tables are not available or do not have the correct degrees of freedom you can use Excel = FINV(0.05,4,45)

ii) As $F > F_{CRIT}$, then $p < 0.05$ but we cannot calculate it without software

Q11.3

The Two-Way ANOVA in EXCEL is performed in Excel file Ch11E.xls sheet Q11.3, with an extract of the results given below:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows (Media)	72	3	24	5.052632	0.044239	4.757055
Columns (Days)	160.1667	2	80.08333	16.85965	0.003447	5.143249
Error	28.5	6	4.75			
Total	260.6667	11				

The results indicate that the Alternative Hypothesis should be accepted at a significance level of 0.05, for both the effects of
Media (Row effect) with $p = 0.044239 < 0.05$, and
Day of collection (Column effect) with $p = 0.003447 < 0.05$.

Q11.4

The Two-Way ANOVA with Replication is performed in Excel file Ch11E.xls sheet Q11.4, with an extract of the results given below:

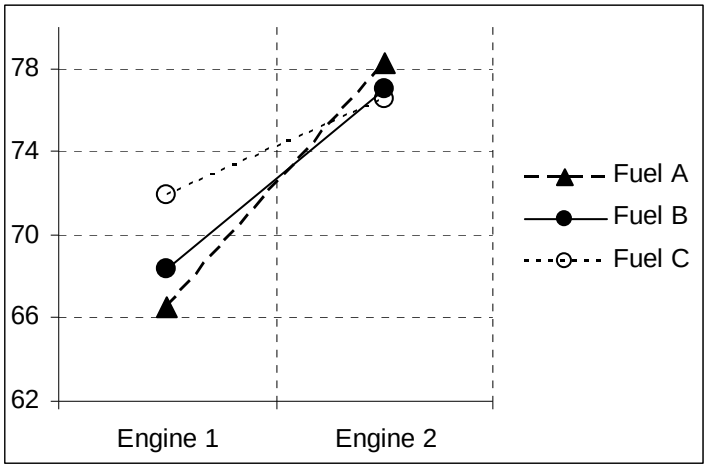
ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample (Temperature)	4.5	1	4.5	1.446429	0.252295	4.747221
Columns (Pressures)	238.7778	2	119.3889	38.375	6.11E-06	3.88529
Interaction	61	2	30.5	9.803571	0.002995	3.88529
Within	37.33333	12	3.111111			
Total	341.6111	17				

The results indicate that the Pressure of the reaction (Column Effect) has a very significant effect on the efficiency of the process:
$p = 6.11E\text{-}06 << 0.05$
and that the Interaction between Pressure and Temperature also has a significant effect:
$p = 0.002995 < 0.05$
However, once the effects of the pressure and its interaction with temperature are taken into account the ANOVA finds that the Temperature, on its own, (Sample effect) no longer has a significant effect on the efficiency:
$p = 0.252295 > 0.05$

Q11.5

The Interaction Plot is performed in Excel file Ch11E.xls sheet Q11.5, with the results given below: (The plots are more easily obtained using MiniTab if the original data is available – see Ch11M.MTP)



Q11.6

These are results from the General Linear Model used in MiniTab. Please see the Video for explanation.

	T 1	T 2	T 3	T 4
	67.2	66.4	62.3	63.2
	62.4	63.2	59.1	62.4
	66.4	64	60.7	62.4
Means =	65.33	64.53	60.70	62.67

General Linear Model: Yield_1 versus Temp_1

```
Factor   Type   Levels  Values

Temp_1   fixed       4  1, 2, 3, 4


Analysis of Variance for Yield_1, using Adjusted SS for Tests


Source   DF  Seq SS  Adj SS  Adj MS     F      P

Temp_1    3  38.449  38.449  12.816  4.22  0.046

Error     8  24.320  24.320   3.040

Total    11  62.769
```

$p < \alpha = 0.05$

S = 1.74356 R-Sq = 61.25% R-Sq(adj) = 46.73%

Unusual Observations for Yield_1

Obs	Yield_1	Fit	SE Fit	Residual	St Resid
2	62.4000	65.3333	1.0066	-2.9333	-2.06 R

R denotes an observation with a large standardized residual.

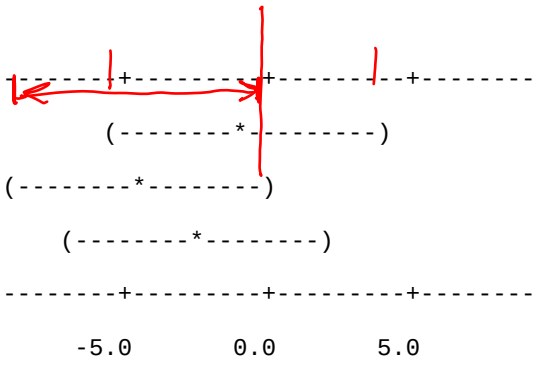
Tukey 95.0% Simultaneous Confidence Intervals

Response Variable Yield_1

All Pairwise Comparisons among Levels of Temp_1

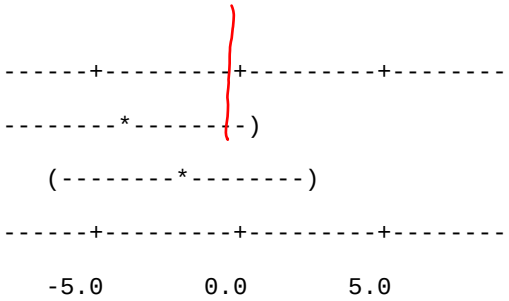
Temp_1 = 1 subtracted from:

Temp_1	Lower	Center	Upper
2	-5.360	-0.800	3.76010
3	-9.193	-4.633	-0.07323
4	-7.227	-2.667	1.89343



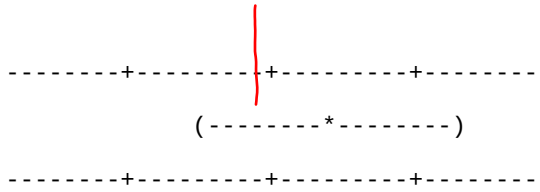
Temp_1 = 2 subtracted from:

Temp_1	Lower	Center	Upper
3	-8.393	-3.833	0.7268
4	-6.427	-1.867	2.6934



Temp_1 = 3 subtracted from:

Temp_1	Lower	Center	Upper
4	-2.593	1.967	6.527



-5.0 0.0 5.0

Tukey Simultaneous Tests

Response Variable Yield_1

All Pairwise Comparisons among Levels of Temp_1

Temp_1 = 1 subtracted from:

Temp_1	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
2	-0.800	1.424	-0.562	0.9406
3	-4.633	1.424	-3.255	0.0465
4	-2.667	1.424	-1.873	0.3103

$p < \alpha = 0.05$

Temp_1 = 2 subtracted from:

Temp_1	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-3.833	1.424	-2.693	0.1026
4	-1.867	1.424	-1.311	0.5815

Temp_1 = 3 subtracted from:

Temp_1	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	1.967	1.424	1.381	0.5429

The only significant difference is between T1 and T3 where the p- value is 0.046