

Today's example was motivated from problem 13.4.8 of Walpole, Myers, Myers and Ye, *Probability and Statistics for Engineers and Scientists, 7th ed.*, Prentice Hall 2002. It is an example of an unbalanced design for a one-factor fixed effects model.

The software automatically incorporates the unequal sample sizes. Notice that the simultaneous confidence intervals on the differences of means have different widths. **R**© uses a modification of the Tukey-Kramer bounds given in the text.

The ANOVA table shows that the means are very significantly different. The simultaneous confidence intervals show that the although solvents A and B do not have significantly different sorption rates, solvent C is significantly different than the other two. Note that the diagnostic plots are easily reached by printing the output of the analysis of variance routine. The plot of residuals vs. fitted values shows that the variances for the three treatments are somewhat different, although the variance of treatment C is only twice that of the others which are approximately equal. The QQ-normal plot of the standardized residuals follows the 45° line nicely indicating that there is no evidence that normality assumption is violated.

Data Set Used in this Analysis :

```
# Math 3082          Sorption Data    1-19-2014
#
# Data from Mc Cleve et al, 1997, as quoted in Walpole, Myers & Myers,
# Probability and Statistics for Engineers and Scientists, 6th ed, Prentice
# Hall, Upper Saddle River NJ, 1998, measured sorption rates of three different
# organic chemical solvents used to clean industrial fabricated-metal parts.
# Independent samples were tested from each type and sorption rates were
# recorded as mole percentage.
#
# Solvent  A = Aromatics
#           C = Chloroalkanes
#           E = Esters
#
Solvents Sorption-Rate
A 1.060000000e+000
A 7.900000000e-001
A 8.200000000e-001
A 8.900000000e-001
A 1.050000000e+000
A 9.500000000e-001
A 6.500000000e-001
A 1.150000000e+000
A 1.120000000e+000
C 1.580000000e+000
C 1.450000000e+000
C 5.700000000e-001
C 1.160000000e+000
C 1.120000000e+000
C 9.100000000e-001
C 8.300000000e-001
C 4.300000000e-001
```

E 2.900000000e-001
E 6.000000000e-002
E 4.400000000e-001
E 5.500000000e-001
E 6.100000000e-001
E 4.300000000e-001
E 5.100000000e-001
E 1.000000000e-001
E 5.300000000e-001
E 3.400000000e-001
E 6.000000000e-002
E 9.000000000e-002
E 1.700000000e-001
E 1.700000000e-001
E 6.000000000e-001

R Session:

R version 2.14.0 (2011-10-31)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-apple-darwin9.8.0/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.42 (5933) i386-apple-darwin9.8.0]

[Workspace restored from /home/1004/ma/treibergs/.RData]
[History restored from /home/1004/ma/treibergs/.Rhistory]

```
> tt=read.table("M3082DataSorption.txt", headers=T)
Error in read.table("M3082DataSorption.txt", headers = T) :
  unused argument(s) (headers = T)
> tt=read.table("M3082DataSorption.txt", header=T)
> tt
  Solvents Sorption.Rate
1      A           1.06
2      A           0.79
3      A           0.82
4      A           0.89
```

```

5      A      1.05
6      A      0.95
7      A      0.65
8      A      1.15
9      A      1.12
10     C      1.58
11     C      1.45
12     C      0.57
13     C      1.16
14     C      1.12
15     C      0.91
16     C      0.83
17     C      0.43
18     E      0.29
19     E      0.06
20     E      0.44
21     E      0.55
22     E      0.61
23     E      0.43
24     E      0.51
25     E      0.10
26     E      0.53
27     E      0.34
28     E      0.06
29     E      0.09
30     E      0.17
31     E      0.17
32     E      0.60

```

```

> attach(tt)
> solv=ordered(Solvents)
> #
> ##### SUMMARIZE AND BOX PLOT OF DATA #####
> tapply(Sorption.Rate, solv, summary)
$A
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.6500 0.8200 0.9500 0.9422 1.0600 1.1500

$C
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.430 0.765 1.015 1.006 1.232 1.580

$E
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.060 0.135 0.340 0.330 0.520 0.610

>
> #
> plot(Sorption.Rate~solv)
> #

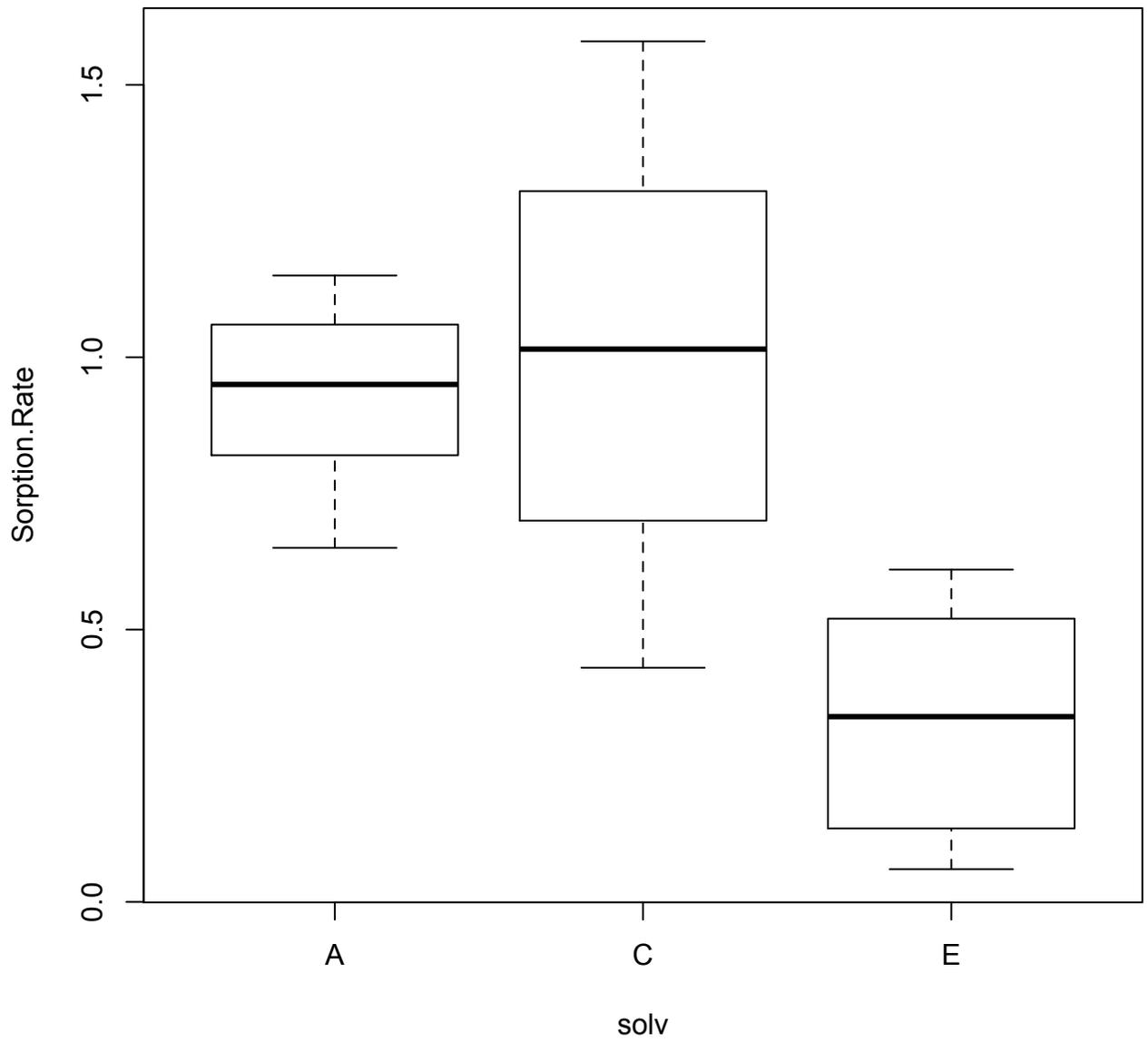
```

```

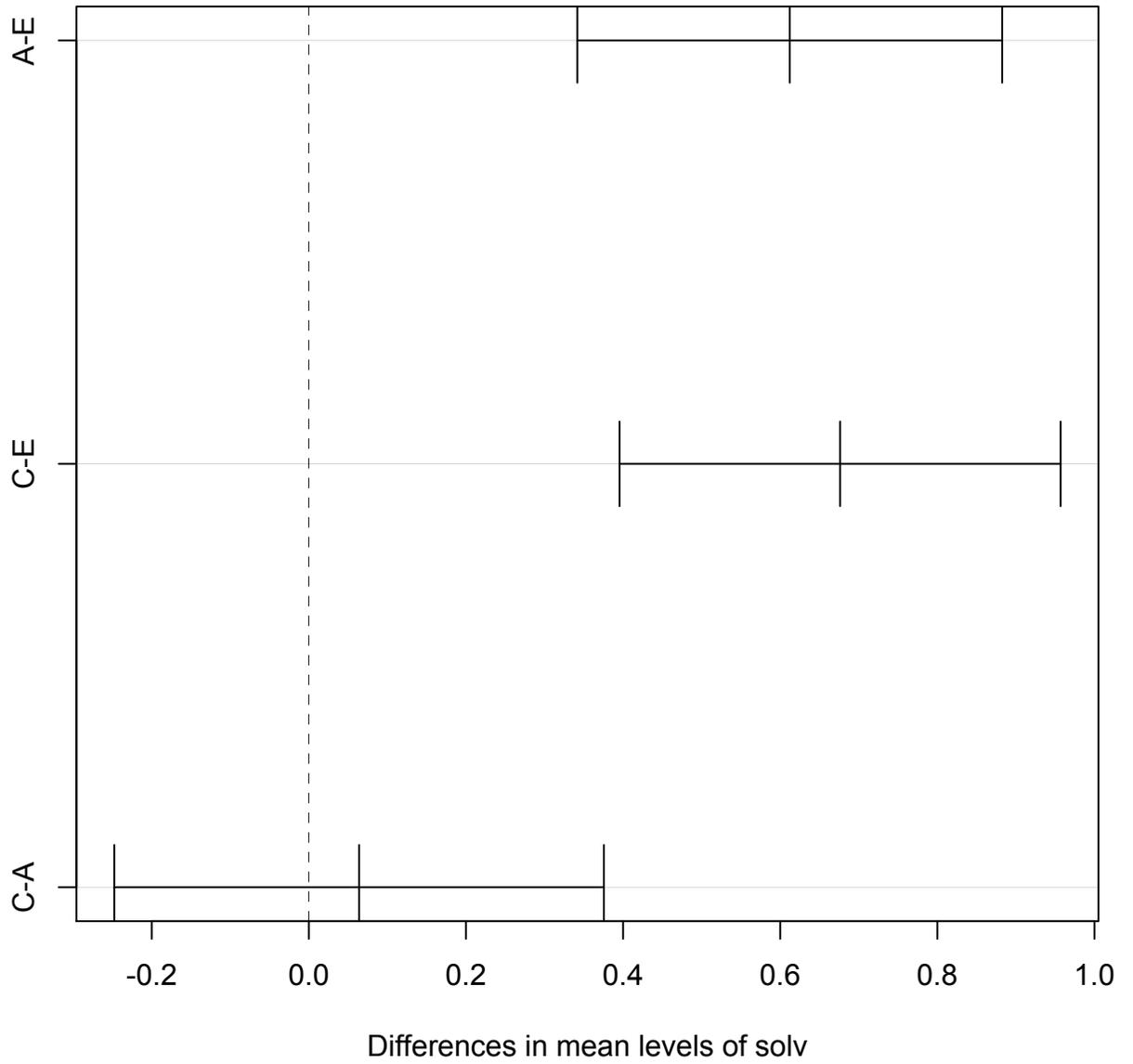
> ##### RUN ANOVA #####
> t1=aov(Sorption.Rate~solv)
> summary(t1)
      Df Sum Sq Mean Sq F value    Pr(>F)
solv     2  3.305   1.6527   24.51 5.86e-07 ***
Residuals 29  1.955   0.0674
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> #
> ##### SIMULTANEOUS CI ON DIFFERENCES OF MEANS #####
> t2=TukeyHSD(t1)
> t2=TukeyHSD(t1,ordered=T)
> plot(t2)
> #
> ##### MODEL CHECKING PLOTS #####
> opar <- par(mfrow = c(2, 2), oma = c(0, 0, 1.1, 0),
+           mar = c(4.1, 4.1, 2.1, 1.1))
> plot(t1)
> par(opar)
>

```



95% family-wise confidence level



aov(Sorption.Rate ~ solv)

