Analysis of Means: Examples with Package ANOM

Philip Pallmann

November 18, 2013

Contents

We load

access data stored in `multcomp` and draw graphics with `ggplot2`.

```
library(nlme)
data(ergoStools)
```

1 Durability of detergents

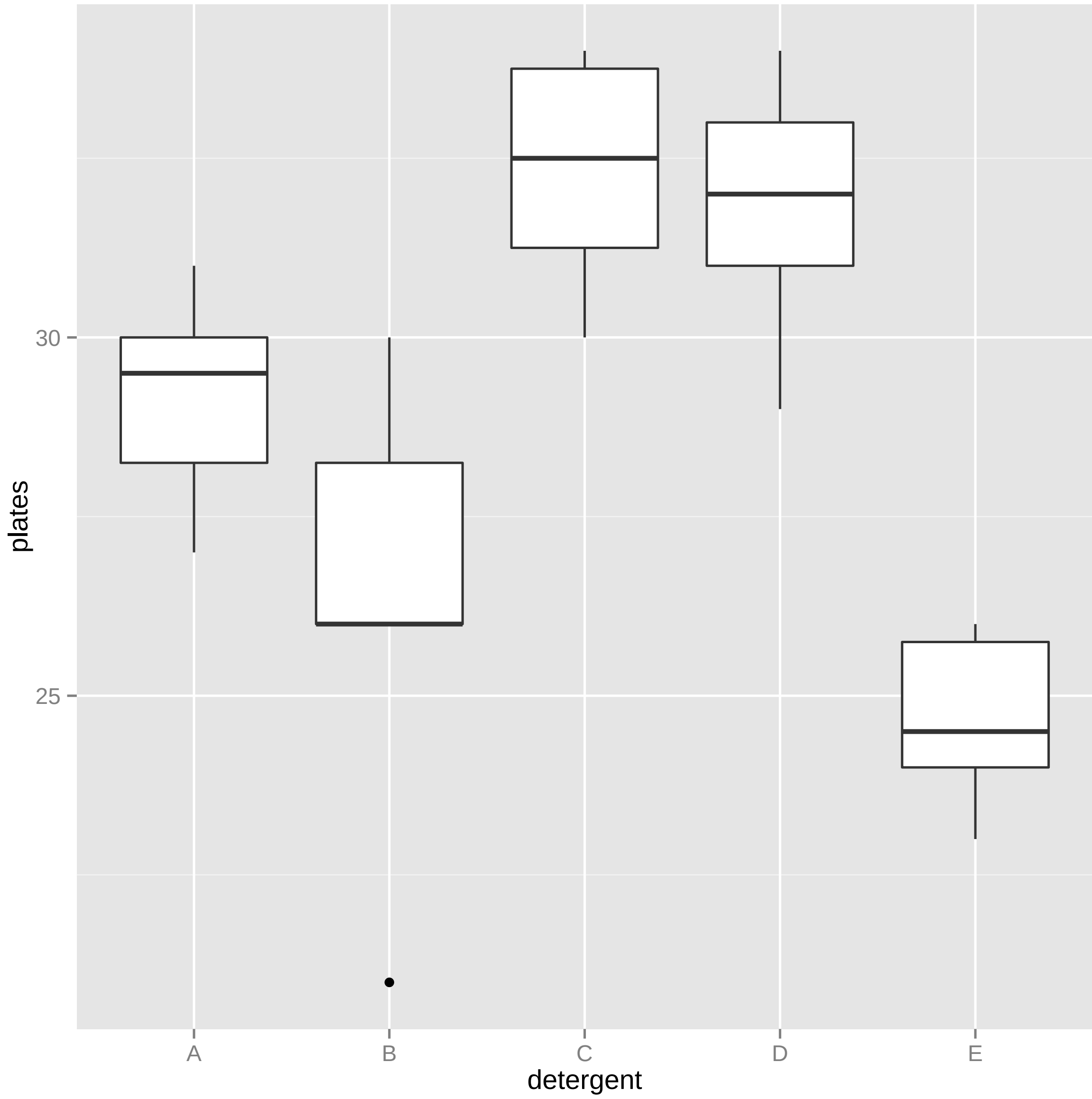
Scheffe (1959, p. 189)

```
detergent

##    detergent block plates
## 1          A   B_1     27
## 2          A   B_2     28
## 3          A   B_3     30
## 4          A   B_4     31
## 5          A   B_5     29
## 6          A   B_6     30
## 7          A   B_7     NA
## 8          A   B_8     NA
## 9          A   B_9     NA
## 10         A  B_10     NA
## 11         B   B_1     26
## 12         B   B_2     26
## 13         B   B_3     29
## 14         B   B_4     NA
## 15         B   B_5     NA
## 16         B   B_6     NA
## 17         B   B_7     30
```

```
## 18          B    B_8      21
## 19          B    B_9      26
## 20          B   B_10      NA
## 21          C    B_1      30
## 22          C    B_2      NA
## 23          C    B_3      NA
## 24          C    B_4      34
## 25          C    B_5      32
## 26          C    B_6      NA
## 27          C    B_7      34
## 28          C    B_8      31
## 29          C    B_9      NA
## 30          C   B_10      33
## 31          D    B_1      NA
## 32          D    B_2      29
## 33          D    B_3      NA
## 34          D    B_4      33
## 35          D    B_5      NA
## 36          D    B_6      34
## 37          D    B_7      31
## 38          D    B_8      NA
## 39          D    B_9      33
## 40          D   B_10      31
## 41          E    B_1      NA
## 42          E    B_2      NA
## 43          E    B_3      26
## 44          E    B_4      NA
## 45          E    B_5      24
## 46          E    B_6      25
## 47          E    B_7      NA
## 48          E    B_8      23
## 49          E    B_9      24
## 50          E   B_10      26
```

```
ggplot(detergent, aes(x = detergent, y = plates)) + geom_boxplot()
## Warning:  Removed 20 rows containing non-finite values (stat_boxplot).
```



Dishwashing detergent durability in an incomplete two-way design.

How many plates washed before foam in the basin disappears?

Dinner plates soiled in a standard way

5 detergents

10 blocks of 3 basins (and 3 dishwashers rotating after each)

n=6 per detergent

balanced incomplete blocks

which ones are better than average?

two-way ANOVA model without interactions

```
detmod1 <- aov(plates ~ detergent + block, detergent)
summary(detmod1)

##             Df Sum Sq Mean Sq F value  Pr(>F)
## detergent    4  269.8    67.4   35.81 8.5e-08 ***
## block        9   67.5     7.5    3.98  0.0079 **
## Residuals   16   30.1     1.9
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 20 observations deleted due to missingness
```

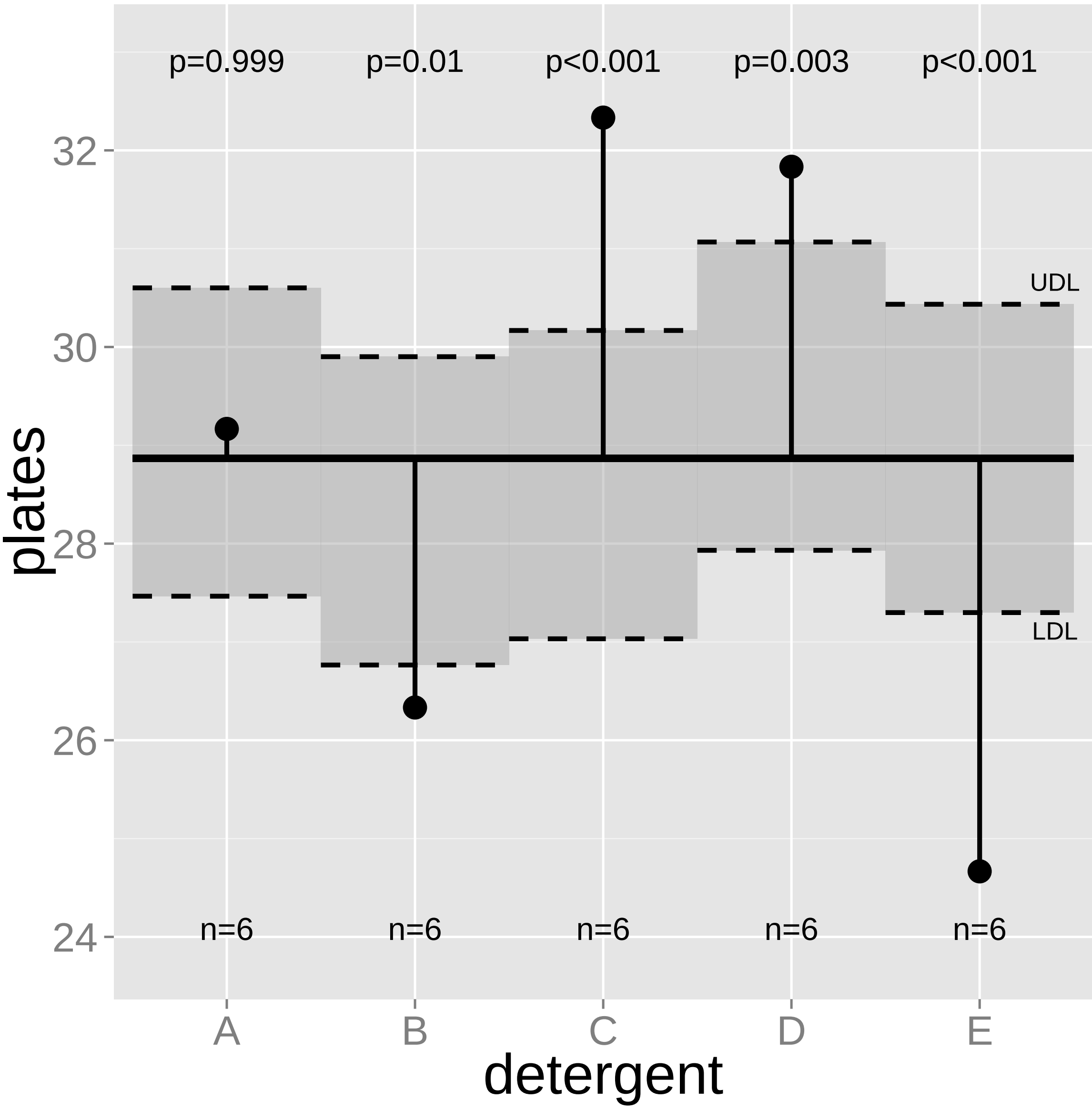
```
det1 <- glht(detmod1, mcp(detergent = "GrandMean"))
summary(det1)
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: GrandMean Contrasts
##
##
## Fit: aov(formula = plates ~ detergent + block, data = detergent)
##
## Linear Hypotheses:
##          Estimate Std. Error t value Pr(>|t|)
## C 1 == 0    0.133      0.549    0.24   0.9987
## C 2 == 0   -2.000      0.549   -3.64   0.0101 *
## C 3 == 0    3.733      0.549    6.80   <0.001 ***
## C 4 == 0    2.333      0.549    4.25   0.0031 **
## C 5 == 0   -4.200      0.549   -7.65   <0.001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

```
confint(det1)
```

```
##
##   Simultaneous Confidence Intervals
##
## Multiple Comparisons of Means: GrandMean Contrasts
##
##
## Fit: aov(formula = plates ~ detergent + block, data = detergent)
##
## Quantile = 2.859
## 95% family-wise confidence level
##
##
## Linear Hypotheses:
##          Estimate lwr    upr
## C 1 == 0  0.133   -1.436  1.703
## C 2 == 0 -2.000   -3.569 -0.431
## C 3 == 0  3.733    2.164  5.303
## C 4 == 0  2.333    0.764  3.903
## C 5 == 0 -4.200   -5.769 -2.631
```

```
ANOM(det1)
```



2 fattyacid

3 penicillin

SCHEFFE 1959

4 ANOM with linear mixed-effects models

Pinheiro and Bates (2000) describe data from an investigation of ergonomic stools performed by Wretenberg et al. (1993). They had nine people assess the physical effort to rise from each of four stool types: a stool of ordinary height, a low ordinary stool, a one-legged stool, and a pneumatic stool. Each subject tested each stool type once and rated the perceived exertion on the so-called Borg scale, which takes values from 6 to 20. We wish to detect types of stools that are significantly more or less difficult to rise from compared to average.

This is a complete block design; each of the four stools under study is evaluated nine times. One challenge is that the scoring is clearly subjective: a person who has general difficulties to rise is more likely to assign lower ratings than a healthy person. To account for this variability between raters, a linear mixed-effects model is our method of choice.

The model for the ith individual has the general form

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i$$

where $\mathbf{y}_i$ are the outcomes, $\boldsymbol{\beta}$ the fixed and $\mathbf{b}_i$ the random effects, $\mathbf{X}_i$ and $\mathbf{Z}_i$ the design matrices for the fixed and random effects, and $\boldsymbol{\epsilon}_i$ the residuals. We assume that $\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$ and $\mathbf{e}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$, and additionally $\mathrm{cov}(\mathbf{b}_i, \mathbf{e}_i) = \mathbf{0}$.

In our ergonomic stools example, we model fixed stool effects and treat the test persons as random. We hereby make the assumption that the testers are representative of all people who could sit on these stools and that their scorings are normally distributed with between-rater variance σ_b.

There are several concurrent and (in part) complementary frameworks for mixed-effects modeling which coexist in the R world. Here we illustrate ANOM using two of the most popular and best-established packages, `nlme` (Pinheiro et al. 2013) and `lme4` (Bates et al. 2013).

4.1 Using `nlme`

We model the effort to rise (on the Borg scale) as a function of the type of stool and include random subject effects to acknowledge between-rater variability:

```
library(nlme)
esmodel1 <- lme(effort ~ Type, random = ~1 | Subject, data = ergoStool)
es1 <- glht(esmodel1, mcp(Type = "GrandMean"), alternative = "two.sided")
ANOM(es1, xlabel = "Stool Type", ylabel = "Exertion (Borg Scale)")
```

4.2 Using `lme4`

An identical graphic can be generated with the functionality in `lme4`:

```
library(lme4)
esmodel2 <- lmer(effort ~ Type + (1 | Subject), data = ergoStool)
es2 <- glht(esmodel2, mcp(Type = "GrandMean"), alternative = "two.sided")
ANOM(es2, xlabel = "Stool Type", ylabel = "Exertion (Borg Scale)")
```

Note that there are instances when the two mixed-effects model packages cannot be used interchangeably. Only `nlme` supports imposing a pattern like compound symmetry or AR(1) on the covariance matrix of the random effects and/or residuals whereas `lme4`'s strong points are sophisticated random effects hierarchies and nested structures, and it also fits generalized linear mixed-effects models.

4.3 Ignoring the repeated structure

One may wonder what would happen if we did not account for the fact that each chair was tested by several individuals. We fit a simple linear model and observe that the decision limits drift apart so that stool type 4 is no longer significant. The reason is quite simple: when fitting a mixed-effects model, we partition the total variance of the data into a component explained by the variability among the nine raters and a residual component. If we ignore the clustered structure of the data and fit a standard linear model, all variability is summed up in the residual variance, thus making it harder to detect significant stool effects.

```
esmodel3 <- lm(effort ~ Type, ergoStool)
es3 <- glht(esmodel3, mcp(Type = "GrandMean"), alternative = "two.sided")
ANOM(es3, xlabel = "Stool Type", ylabel = "Exertion (Borg Scale)")
```

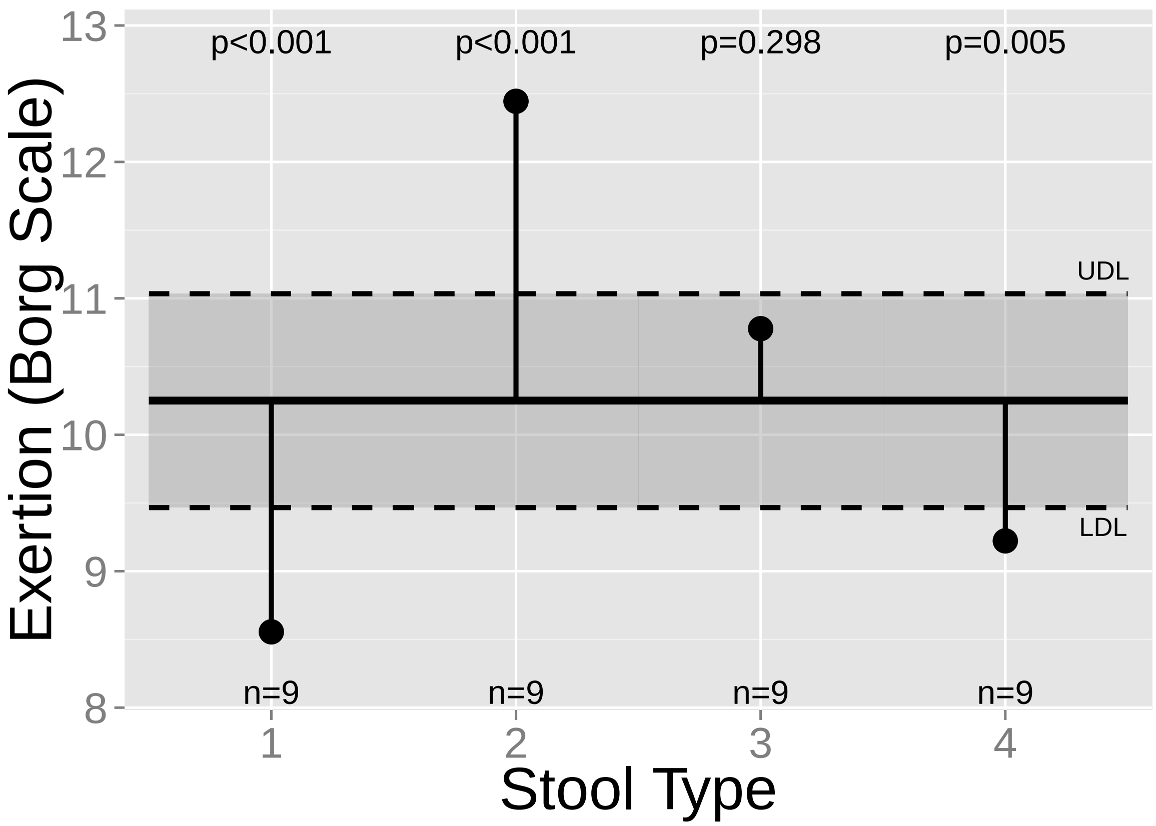


Figure 1: ANOM decision chart for the ergonomic stool data based on a linear mixed-effects model.

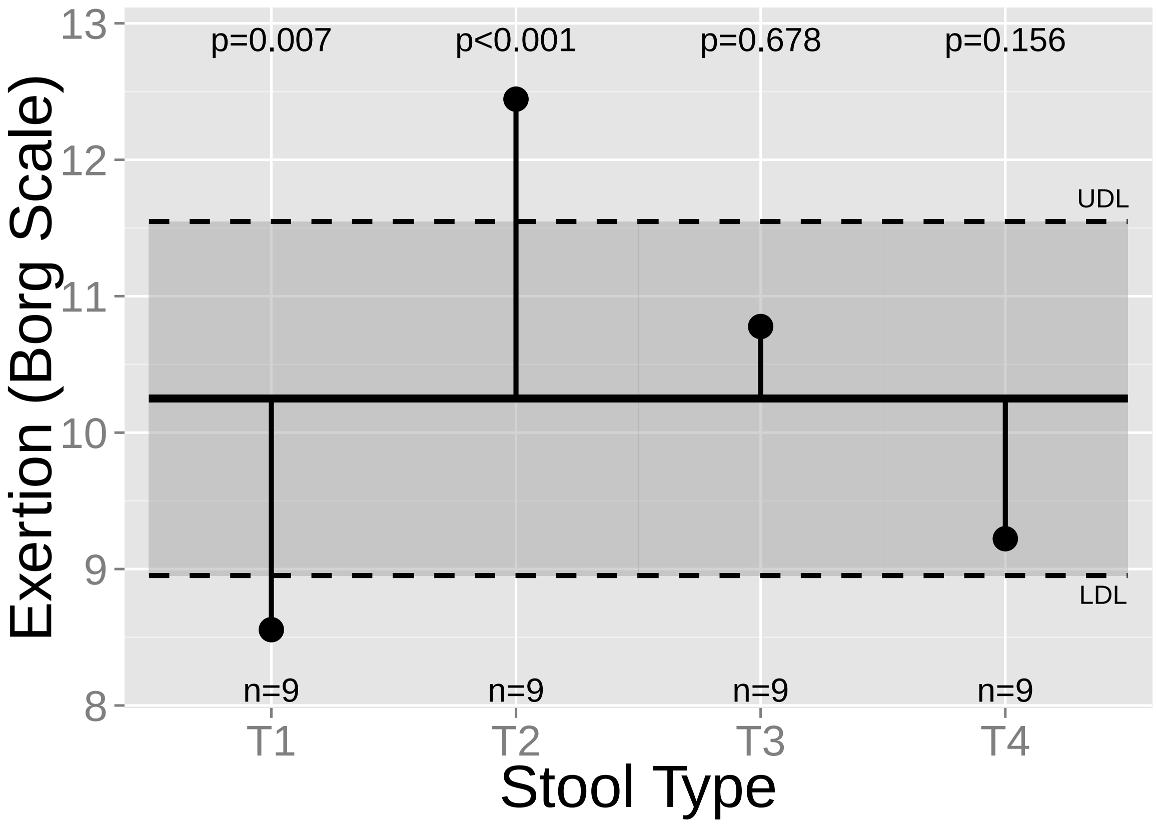


Figure 2: ANOM decision chart for the ergonomic stool data based on a standard linear model.

5 ANOM in a two-way layout

Nelson et al. (2005) describe a trial on 30 male cancer patients in a balanced complete two-way layout. Each patient was treated with one therapy (chemo or radiation) and one out of three drugs (named 1, 2, and 3), and the level of hemoglobin was measured afterwards. The six drug-therapy combinations may be compared in a simplified pseudo-one-way layout where the new factor's levels are the combinations of therapy and drug:

```
hemoglobin$the <- as.factor(abbreviate(hemoglobin$therapy))
hemoglobin$td <- with(hemoglobin, the:drug)
hemodel <- lm(level ~ td, hemoglobin)
he <- glht(hemodel, mcp(td = "GrandMean"), alternative = "two.sided")
ANOM(he, xlabel = "Treatment", ylabel = "Hemoglobin Level")
```

We find that chemotherapy together with drug 3 leads to hemoglobin levels that are significantly above the grand mean. On the contrary, if radiation therapy is combined with either drug 1 or 2, the levels of hemoglobin are close to the lower decision limit.

Acting on the assumption that the two treatment factors "therapy" and "drug" do not interact, we may also investigate the marginal drug effects across both therapies:

```
hemodel2 <- lm(level ~ drug + therapy, hemoglobin)
he2 <- glht(hemodel2, mcp(drug = "GrandMean"), alternative = "two.sided")
ANOM(he2, xlabel = "Drug", ylabel = "Hemoglobin Level")
```

Figure **??** shows that drug 3 raises hemoglobin levels significantly compared to the grand mean. Notice that this type of pooled analysis would be inept in the presence of a therapy-drug interaction; here one should rather perform separate drug comparisons for radiation and for chemotherapy.

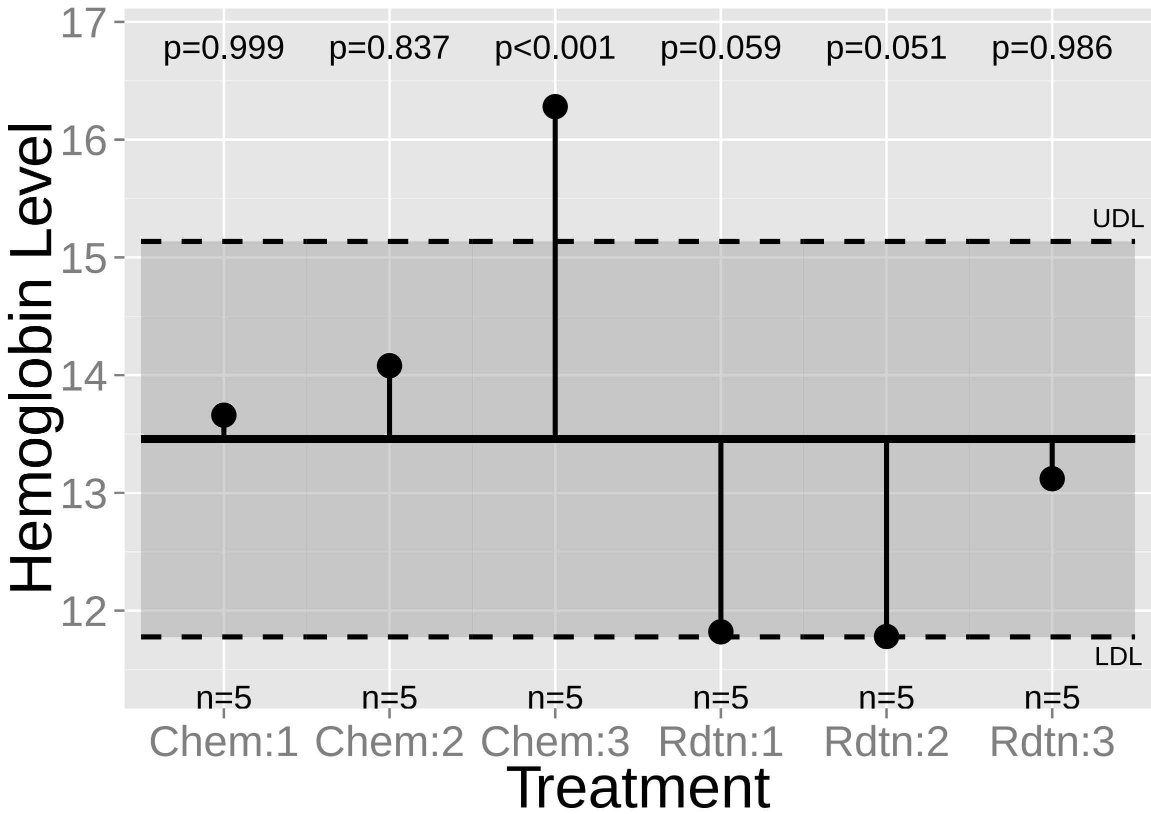


Figure 3: ANOM decision chart for the hemoglobin data based one a pseudo-one-way analysis.

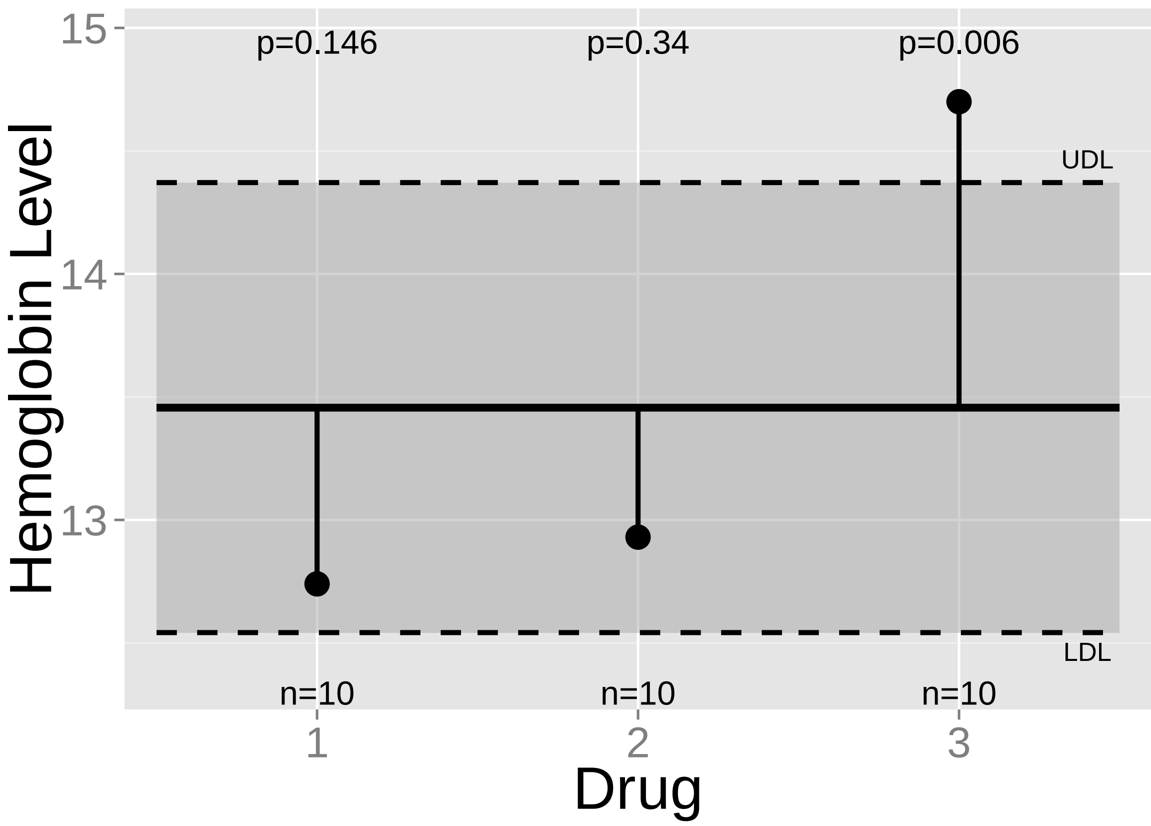


Figure 4: ANOM decision chart for the hemoglobin data based on a two-way analysis.

6 STUFF

Fragestellung

Figures

References

Douglas Bates, Martin Maechler, Ben Bolker, and Steven Walker. *lme4: Linear mixed-effects models using Eigen and S4*, 2013. URL `http://CRAN.R-project.org/package=lme4`. R package version 1.0-5.

Peter R. Nelson, Peter S. Wludyka, and Karen A. F. Copeland. *The Analysis of Means: A Graphical Method for Comparing Means, Rates, and Proportions.* ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia, PA, and ASA, Alexandria, VA, 2005. ISBN 0-89871-592-X.

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. *nlme: Linear and nonlinear mixed effects models*, 2013. URL `http://CRAN.R-project.org/package=nlme`. R package version 3.1-111.

José C. Pinheiro and Douglas M. Bates. *Mixed-Effects Models in S and S-PLUS.* Springer, New York, NY, 2000. ISBN 0-387-98957-9.

Per Wretenberg, Ulf P. Arborelius, and Fredrik Lindberg. The effects of a pneumatic stool and a one-legged stool on lower limb joint load and muscular activity during sitting and rising. *Ergonomics*, 36 (5):519–535, 1993.