

第六次作业 2017 年 11 月 30 日

注：本次作业必须在 2017 年 12 月 7 日上课前交。

1. 为研究屠宰场消除牛尸上的大肠杆菌的抗菌处理功效，华盛顿大学的研究人员调查了一些屠宰场。他们比较了 3 种处理技术：6 天干燥老化制冷处理 (Dry_aging_6)，醋酸喷雾处理 (Acetic_acid) 和低压热水处理 (Hot_water)。每个屠宰场仅应用一种处理技术。对每一种处理技术，他们调查了 4 至 5 个屠宰场，一共有 14 个屠宰场。在每个屠宰场里抽样 4 至 16 个牛尸，总共抽取 123 个牛尸。记录在处理前和处理后每个牛尸上的大肠杆菌数量(CFU)。这样得到处理功效(用 logCFU_diff 表示)为它们之间的对数差，即用处理后的 log-CFU 减去处理前的 log-CFU。下面是两个主要模型(M1 和 M2)的输出，以及模型 M1 的残差图。涉及到分布的一些 p -值结果也已列出。

```
> M1 = lmer(logCFU_diff ~ trt + (1|facility)) > M2 = lm(logCFU_diff ~ trt)

M1: Linear mixed model fit by REML
Formula: logCFU_diff ~ trt + (1 | facility)
Data: dat
AIC BIC logLik deviance REMLdev
324.6 338.7 -157.3 313.7 314.6
Random effects:
Groups Name Variance Std.Dev.
facility (Intercept) 0.46338 0.68072
Residual 0.61403 0.78360
Number of obs: 123, groups: facility, 14

Fixed effects:
Estimate Std. Error t value
(Intercept) 0.818744 0.329637 2.484
trtAcetic_acid 0.615125 0.498149 1.235
trtHot_water -0.009394 0.463872 -0.020

> anova(M1)
Analysis of Variance Table
Df Sum Sq Mean Sq F value
trt 2 1.2205 0.61025 0.9938

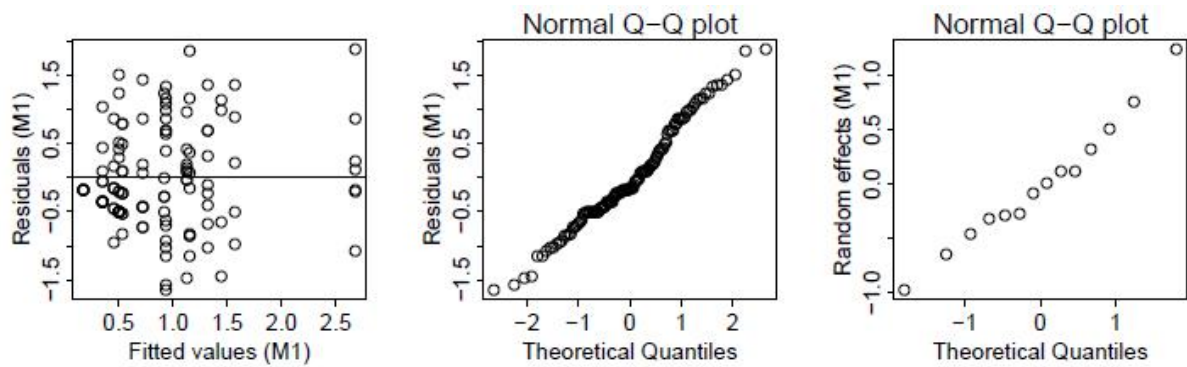
> M3 = update(M1, .~. - trt)
> anova(M3, M1)
Df AIC BIC logLik Chisq Df Pr(>Chisq)
M3 3 321.6 330.1 -157.8
M1 5 323.7 337.7 -156.8 2.016 2 0.364

M2: formula: logCFU_diff ~ trt
Coefficients:
Estimate Std.Err t val Pr(>|t|)
(Intercept) 0.754 0.144 5.251 6.62e-07 ***
trtAcetic_acid 0.623 0.232 2.684 0.0083 **
trtHot_water 0.045 0.198 0.226 0.8215

Residual std. error: 0.964 on 120 DF
Multiple R-squared: 0.065
Adjusted R-squared: 0.050
F-statistic:4.20 on 2 and 120 DF, p-val:0.0172

> logLik(M1, REML=T)
'log Lik.' -157.304 (df=5)
> logLik(M2, REML=T)
'log Lik.' -171.4011 (df=4)
> logLik(M1, REML=F)
'log Lik.' -156.8279 (df=5)
> logLik(M2, REML=F)
'log Lik.' -168.504 (df=4)

> pchisq(2*14.0971, df=1, lower.tail=F)
[1] 1.10e-07
> pchisq(2*11.6761, df=1, lower.tail=F)
[1] 1.35e-06
> pf(0.9938, df1=2, df2=11, lower.tail=F)
[1] 0.4010983
> pf(0.9938, df1=2, df2=109, lower.tail=F)
[1] 0.3734961
```



- 1). 写出模型 M1 拟合的线性混合效应方程，包括所有参数的估计值，及随机效应及误差的分布。(提示：处理技术为属性变量，在 M1 中用 `trt` 表示。从 M1 结果中可见没有出现 `Dry_aging_6`，是因为在实际计算中将其作为比较基准)
- 2). 在模型 M1 中，将屠宰场的影响看成固定效应而不是随机效应恰当吗？请简要解释。(提示：注意到模型 M1 和 M2 的差异)
- 3). 列出用 REML 方法关于模型 M1 的至少 3 个假设，并根据 M1 的残差图来评估这些假定是否正确。
- 4). 请利用上面的输出结果，检验使用相同处理技术的屠宰场之间所产生的平均处理效应是否存在差异。指出检验的名称，检验统计量，在原假设下的分布及 p -值。
- 5). 有证据表明不同处理技术的功效存在差异吗？请用 F 检验回答此问题，指出检验统计量，在原假设下的分布及 p -值。
- 6). 未来的类似实验将关注某一单项处理技术，期望获得功效的更精细刻画。假设资源充足，有两种可供选择的抽样方案：
 - 方案一：从 15 个屠宰场中各取 5 个牛尸；
 - 方案二：从 8 个屠宰场中各取 20 个牛尸；
 假设方差与当前实验估计得到的相同，哪种方案更优？请给出解释。(提示：从极小化平均处理功效的标准差来考虑)