Example p. 21 / Beispiel S. 21

FK

automatic

Working directory

```
> setwd("D:/kronthafranz/Documents/01Lehre/06Quantitative Forschungsmethoden
dt en")
```

Load data

```
> load("D:/kronthafranz/Documents/01Lehre/06Quantitative Forschungsmethoden
dt en/05ANOVA/production.RData")
```

Define factor

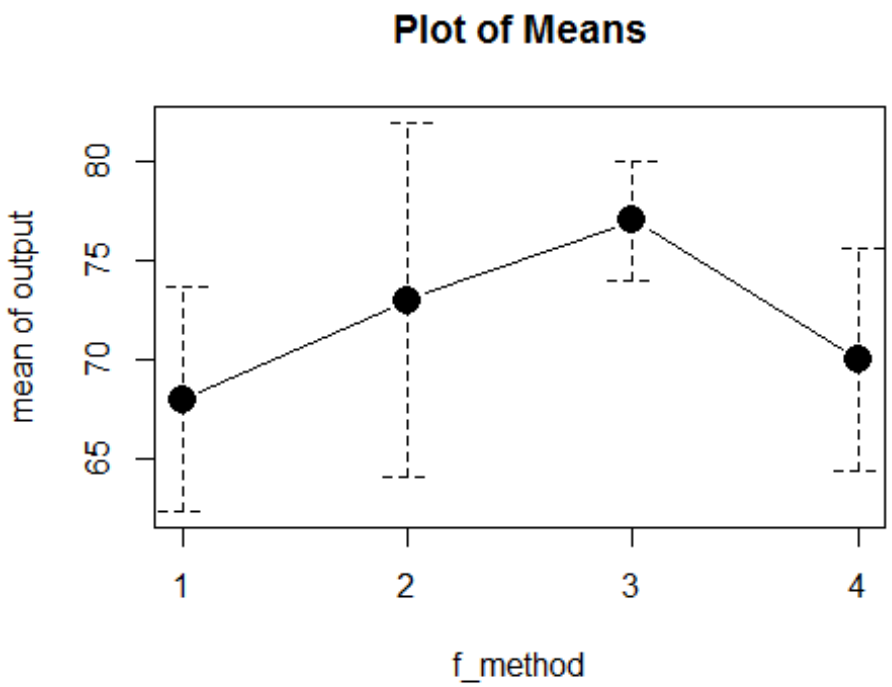
```
> production <- within(production, {
+   f_method <- as.factor(method)
+ })
```

Descriptive statistics

```
> library(abind, pos=15)

> library(e1071, pos=16)

> numSummary(production[,"output", drop=FALSE], groups=production$f_method,
+   statistics=c("mean", "sd", "quantiles"), quantiles=c(0,.25,.5,.75,1))

  mean       sd 0%   25%  50%   75% 100% output:n
1   68 5.366563 61 65.50 67.5 69.50   77        6
2   73 8.532292 62 66.00 75.0 78.75   83        6
3   77 2.898275 73 75.25 77.0 78.75   81        6
4   70 5.329165 62 67.25 70.0 74.25   76        6
```

Plot means

```
> with(production, plotMeans(output, f_method, error.bars="conf.int",
level=0.95,
+   connect=TRUE))
```



Check assumptions

Independence of observations

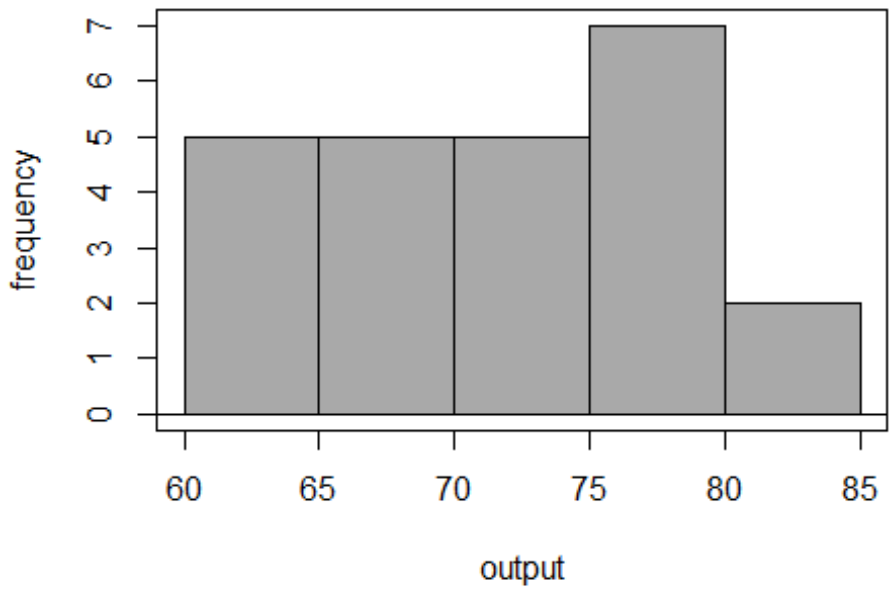
Matter of design of the experiment

Equality of variance

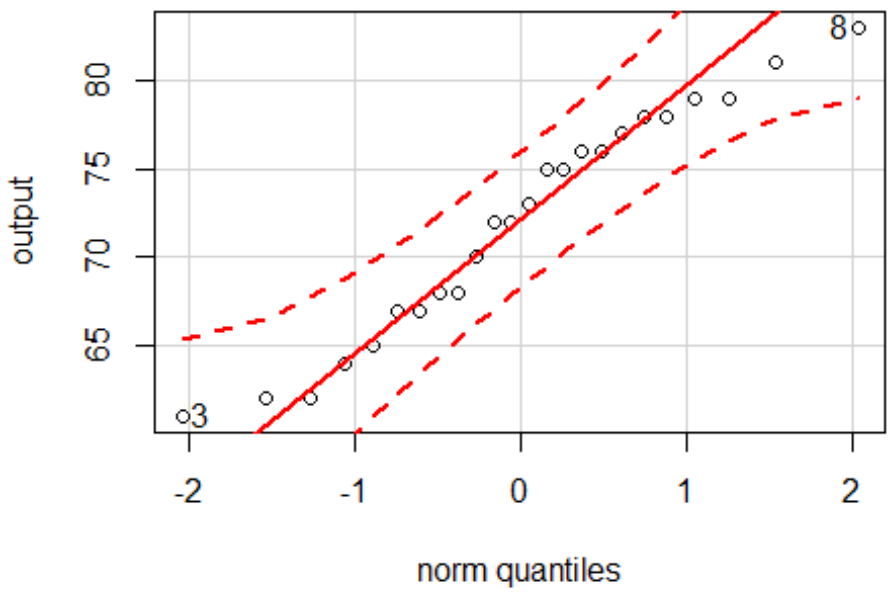
```
> with(production, tapply(output, f_method, var, na.rm=TRUE))

   1    2    3    4
28.8 72.8  8.4 28.4

> with(production, tapply(output, f_method, var, na.rm=TRUE))

   1    2    3    4
28.8 72.8  8.4 28.4

> leveneTest(output ~ f_method, data=production, center="mean")

Levene's Test for Homogeneity of Variance (center = "mean")
      Df F value Pr(>F)
group  3  2.6873  0.074 .
      20
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Normal distribution

```
> with(production, Hist(output, scale="frequency", breaks="Sturges",
col="darkgray"))
```




```
> with(production, qqPlot(output, dist="norm", id.method="y", id.n=2,
+    labels=rownames(production)))
```




```
 3  8
 1 24
```

```
> normalityTest(~output, test="shapiro.test", data=production)


    Shapiro-Wilk normality test

data:  output
W = 0.95365, p-value = 0.3245
```

ANOVA

```
> AnovaModel.1 <- aov(output ~ f_method, data=production)

> summary(AnovaModel.1)

            Df Sum Sq Mean Sq F value Pr(>F)
f_method     3    276    92.0   2.659 0.0761 .
Residuals   20    692    34.6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> with(production, numSummary(output, groups=f_method, statistics=c("mean",
"sd")))

  mean       sd data:n
1   68 5.366563      6
2   73 8.532292      6
```

```
3   77 2.898275      6
4   70 5.329165      6

> TukeyHSD(AnovaModel.1)

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = output ~ f_method, data = production)

$f_method
    diff         lwr     upr     p adj
2-1    5  -4.5054003 14.5054 0.4717578
3-1    9  -0.5054003 18.5054 0.0674888
4-1    2  -7.5054003 11.5054 0.9342653
3-2    4  -5.5054003 13.5054 0.6472435
4-2   -3 -12.5054003  6.5054 0.8134157
4-3   -7 -16.5054003  2.5054 0.1998434
```

Interpretation:

H0 is rejected (significance level of 10%)

There is a difference in the productivity of production methods

Tukey Post Hoc shows that there is a difference between method 1 and method 3