1

a

$$X_{ij} \sim n(\mu_i, \sigma^2)$$
$$X_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

where

$$\epsilon_{ij} \sim n(0, \sigma^2)$$

b

Use an F test to test whether $\mu_i = \mu_j \forall i, j$.

```
bonedensity=read.table(
    "http://www.uio.no/studier/emner/matnat/math/STK2120/v16/bonedensity.txt",
    header=T)
aov.fit=aov(density~factor(treatment),data=bonedensity)
print(summary(aov.fit))

#                      Df   Sum Sq   Mean Sq F value  Pr(>F)
#factor(treatment)    2 0.003186 0.0015928   7.718  0.0014 **
#Residuals           42 0.008668 0.0002064

# Significant at 0.01

tukey.fit = TukeyHSD(aov.fit ,ordered=T)
png("oblig1.png")
plot(tukey.fit)
dev.off()
```

Isoflavons are significant $\alpha = 0.01$

2

a

We could use single factor ANOVA, but block is better since the person could be a significant factor

b

$$X_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

where

$$\epsilon_{ij} \sim n(0, \sigma^2)$$

c

```
rullestol=
read.table("http://www.uio.no/studier/emner/matnat/math/STK2120/v16/rullestol.txt",
           header=T)
aov.fit = aov(trykk~factor(pute)+factor(person), data=rullestol)
summary(aov.fit)


#                Df Sum Sq Mean Sq F value    Pr(>F)
#factor(pute)     4   2131   532.7   6.554 0.000454 ***
#factor(person)   9   7455   828.3  10.191 1.36e-07 ***
```

```
#Residuals       36   2926    81.3
#

tukey.fit = TukeyHSD(aov.fit,ordered=T)

#        diff        lwr        upr     p adj
#BD-JP    5.8  -5.775221  17.37522  0.6076304
#SC-JP    9.1  -2.475221  20.67522  0.1826418
#RF-JP   13.1   1.524779  24.67522  0.0199029
#SF-JP   19.3   7.724779  30.87522  0.0002649
#SC-BD    3.3  -8.275221  14.87522  0.9232171
#RF-BD    7.3  -4.275221  18.87522  0.3834630
#SF-BD   13.5   1.924779  25.07522  0.0154371
#RF-SC    4.0  -7.575221  15.57522  0.8571574
#SF-SC   10.2  -1.375221  21.77522  0.1063365
#SF-RF    6.2  -5.375221  17.77522  0.5456224
#
plot(tukey.fit)
```

Both pute and person are significant factors.

d

We use the studentized range distribution to get confidence intervals for each $\mu_i - \mu_j$ such that the simultaneous confidece level is $1-\alpha$. Then we say that each pair i, j where 0 is not in the CI are significantly different.

We find that RF-JP, SF-JP and SF-BD are significantly different.

3

a

$$\begin{aligned}
\sum_i \sum_j (X_{ij} - \mu_{ij})^2 &= \sum_i \sum_j (X_{ij} - (\mu + \alpha_i))^2 \\
&= \sum_i \sum_j ((X_{ij} - \bar{X}_{i\cdot}) + (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot} - \alpha_i) + (\bar{X}_{\cdot\cdot} - \mu))^2 \\
&= \sum_i \sum_j (a_{ij} + b_i + c)^2 \\
&= \sum_i \sum_j (a_{ij}^2 + b_i^2 + c^2 + 2a_{ij}b_i + 2b_i c + 2a_{ij}c) \\
&= \sum_i \sum_j (a_{ij}^2 + b_i^2 + c^2) + 2\sum_i b_i \sum_j a_{ij} + 2Jc\sum_i b_i + 2c\sum_i \sum_j a_{ij} \\
&= \sum_i \sum_j (a_{ij}^2 + b_i^2 + c^2) + 2\sum_i b_i 0 + 2Jc0 + 2c\sum_i 0 \\
&= \sum_i \sum_j (a_{ij}^2 + b_i^2 + c^2) \\
&= \sum_i \sum_j a_{ij}^2 + J\sum_i b_i^2 + IJc^2
\end{aligned}$$

c

$$\begin{aligned}
\hat{\mu} &= \bar{X}_{..} \\
&= \frac{1}{IJ}\sum_i \sum_j X_{ij} \\
&= \frac{1}{IJ}\sum_i \sum_j \mu + \alpha_i + \epsilon_{ij} \\
&= \frac{1}{IJ}\sum_i \sum_j (\mu + \alpha_i + \epsilon_{ij}) \\
&= \mu + \frac{1}{IJ}\sum_i \sum_j (\epsilon_{ij})
\end{aligned}$$

$$\begin{aligned}
E(\hat{\mu}) &= \mu \\
Var(\hat{\mu}) &= 1/(IJ)^2 \sum_{ij} Var(\epsilon_{ij}) \\
&= \sigma^2 IJ/(IJ)^2 = \sigma^2/(IJ)
\end{aligned}$$

So $\hat{\mu}$ is normally distributed with $E = \mu$ and $Var = \sigma^2/(IJ)$

d

$$\begin{aligned}
E(\hat{\alpha_i}) &= E(\bar{X}_{i\cdot}) - E(\bar{X}_{..}) \\
&= \mu + \alpha_i - \mu \\
&= \alpha_i \\
Var(\hat{\alpha_i}) &= Var(\bar{X}_{i\cdot} - \bar{X}_{..}) \\
&= Var(\bar{X}_{i\cdot}) + Var(\bar{X}_{..})) - Covar(\bar{X}_{i\cdot}, \bar{X}_{..}) \\
&= \sigma^2/J + \sigma^2/(IJ) - 2\sum_k Covar(\bar{X}_{i\cdot}, \bar{X}_{k\cdot})/I \\
&= \sigma^2/J + \sigma^2/(IJ) - 2Var(\bar{X}_{i\cdot})/I \\
&= \sigma^2/J + \sigma^2/(IJ) - 2\sigma/(JI) \\
&= \sigma^2/J - \sigma^2/(IJ) \\
&= \frac{I-1}{IJ}\sigma^2
\end{aligned}$$

e

```
> mean(bonedensity$density)
[1] 0.2232889
> mean(bonedensity$density[bonedensity$treatment==1])
[1] 0.2188667
> mean(bonedensity$density[bonedensity$treatment==2])
[1] 0.2159333
> mean(bonedensity$density[bonedensity$treatment==3])
[1] 0.2350667
```

$$
\begin{aligned}
\hat{\mu} &= \bar{X}.. = 0.2232889 \\
\hat{\alpha_1} &= \bar{X}_1. - \bar{X}.. = 0.2188667 - 0.2232889 = -0.0044222 \\
\hat{\alpha_2} &= \bar{X}_2. - \bar{X}.. = 0.2159333 - 0.2232889 = -0.0073556 \\
\hat{\alpha_2} &= \bar{X}_2. - \bar{X}.. = 0.2350667 - 0.2232889 = 0.0117778
\end{aligned}
$$

f

We have that $E(MSE) = \sigma^2$ so $MSE = 0.0002064$ is an estimate for σ^2. So standard error for $\hat{\mu}$ is estimated by $\sqrt{MSE/IJ} = \sqrt{0.0002064/(3*15)} = 0.002141651$, and the standard error for α_i is estimated by $\sqrt{MSE(I-1)/IJ} = \sqrt{0.0002064*2/(3*15)} = 0.003028751$

g

```
bonedensity=read.table(
    "http://www.uio.no/studier/emner/matnat/math/STK2120/v16/bonedensity.txt",
    header=T)
options(contrasts=c("contr.sum","contr.poly"))
fit.lm=lm(density~factor(treatment),data=bonedensity)
summary(fit.lm)


#Coefficients:
#                      Estimate Std. Error t value Pr(>|t|)
#(Intercept)           0.223289   0.002142 104.267   <2e-16 ***
#factor(treatment)1   -0.004422   0.003029  -1.460   0.1517
#factor(treatment)2   -0.007356   0.003029  -2.429   0.0195 *
#
options(contrasts=c("contr.treatment","contr.poly"))
fit.lm=lm(density~factor(treatment),data=bonedensity)

#Coefficients:
#                      Estimate Std. Error t value Pr(>|t|)
#(Intercept)           0.218867   0.003709  59.007  < 2e-16 ***
#factor(treatment)2   -0.002933   0.005246  -0.559  0.57900
#factor(treatment)3    0.016200   0.005246   3.088  0.00356 **
#
```

We get $SE_{\hat{\mu}} = 0.002142$ and $SE_{\hat{\alpha_i}} = 0.003029$ which is the same as the results from f

h

The estimates can be expressed:

$$
\begin{aligned}
\hat{\mu} &= \bar{X}_1. \\
\hat{\alpha}_i &= \bar{X}_i. - \bar{X}_1.
\end{aligned}
$$