1. (a) Whole-plot experimental units are coolers. Split-plot experimental units are cuts of beef.

   (b) The *Source* and *Degress* of *Freedom* columns of the ANOVA table are as follows:

Source	DF
Temp	2
Cooler(Temp)	9
Pres.	1
Temp×Pres.	2
Error=Pres.×Cooler(Temp)	9
C. Total	23

   (c) Cooler(Temp).

   (d) Error or, equivalently, Pres×Cooler(Temp).

2. (a) The true mean responses and corresponding levels for genotype and fertilizer are shown below:

```
> block=factor(rep(1:4,each=12))
> geno=factor(rep(rep(1:3,each=4),4))
> x=rep(seq(0,150,by=50),12)
> fert=factor(x)
> X=model.matrix(~geno+x+I(x^2)+geno:x)
> beta=c(125,15,-10,.4,-0.0015,0,.2)
> d=data.frame(fert = x, geno, mean = X %*% beta)
> d2=d[1:12,]
> mu=matrix(d2[,3],3,4,byrow=T)
> rownames(mu)=c('Geno 1','Geno 2','Geno 3')
> colnames(mu)=c('Fert 0','Fert 50','Fert 100','Fert 150')
> mu
       Fert 0 Fert 50 Fert 100 Fert 150
Geno 1    125  141.25      150   151.25
Geno 2    140  156.25      165   166.25
Geno 3    115  141.25      160   171.25
```

   (b) No, the null hypothesis of no genotype main effects is not true since $\bar{\mu}_{i\cdot}$ is not the same for all i:

```
> rowMeans(mu)
      1       2       3
141.875 156.875 146.875
```

   (c) No, the null hypothesis of no fertilizer main effects is not true since $\bar{\mu}_{\cdot j}$ is not the same for all j:

```
> colMeans(mu)
       0        50       100       150
126.6667 146.2500 158.3333 162.9167
```

(d) No, the null hypothesis of no genotype × fertilizer interactions is not true, since

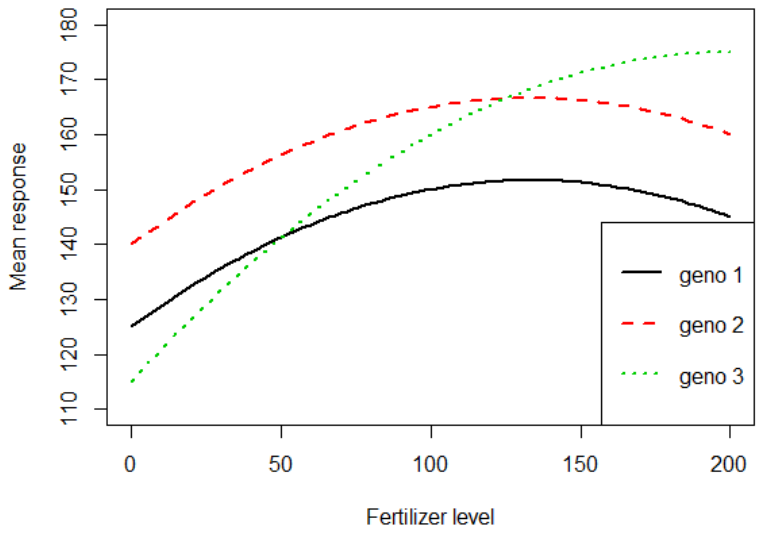
$$(\mu_{11} - \mu_{13}) - (\mu_{31} - \mu_{33}) \neq 0 \quad \text{and} \quad (\mu_{22} - \mu_{23}) - (\mu_{32} - \mu_{33}) \neq 0.$$

```
> mu[1,1] - mu[1,3] - mu[3,1] + mu[3,3]
[1] 20
> mu[2,2] - mu[2,3] - mu[3,2] + mu[3,3]
[1] 10
```

(e) The quadratic equations for three genotypes are

Genotype 1: $f(x) = 125 + 0.4x - 0.0015x^2$
Genotype 2: $f(x) = 125 + 15 + 0.4x - 0.0015x^2 + 0x = 140 + 0.4x - 0.0015x^2$
Genotype 3: $f(x) = 125 - 10 + 0.4x - 0.0015x^2 + 0.2x = 115 + 0.6x - 0.0015x^2$

The plot below was produced by the `R` code that follows:




```
> g1=function(x) 125 + 0.4*x - 0.0015*x^2
> g2=function(x) 140 + 0.4*x - 0.0015*x^2
> g3=function(x) 115 + 0.6*x - 0.0015*x^2
> curve(g1, 0, 200, ylim = c(110,180), type = "l",
+       xlab = "Fertilizer level", ylab = "Mean response", lwd=2)
> curve(g2,col=2, lty=2,lwd=2,add=T)
```

```
> curve(g3,col=3, lty=3,lwd=2,add=T)
> legend("bottomright",paste("geno",1:3),col = c(1:3), lty = 1:3,lwd=2)
```

(f) By slide 38 of set 15, an approximate 95% confidence interval for $\mu_{11} - \mu_{12}$ is

$$\bar{y}_{11\cdot} - \bar{y}_{12\cdot} \pm t_{d,0.975}\sqrt{\frac{2}{b}MS_{Error}},$$

where $t_{d,0.975}$ denotes the 0.975 quantile of a t distribution with d degrees of freedom $w(s-1)(b-1) = 3(4-1)(4-1) = 27$

Using the `R` code below,

$$\bar{y}_{11\cdot} - \bar{y}_{12\cdot} = -13.75, \quad \frac{2}{4}MS_{Error} = 19.85,$$

and an approximate 95% confidence interval for $\mu_{11} - \mu_{12}$ is

$$(-22.89, -4.61).$$

```
> Z1=model.matrix(~0+block)
> Z2=model.matrix(~0+geno:block)
> Z=cbind(Z1,Z2)
> set.seed(532)
> u=c(rnorm(4,0,6),rnorm(12,0,7))
> e=rnorm(48,0,6)
> y=round(X%*%beta+Z%*%u+e,1)
> dat=data.frame(block,geno,fert,y)
> est=mean(subset(dat, geno == '1' & fert == '0')$y)-
  mean(subset(dat, geno == '1' & fert == '50')$y)
> est
[1] -13.75
> o=lm(y~block+geno+block:geno+fert+geno:fert, data = dat)
> MS=anova(o)$'Mean Sq'
> df=anova(o)$Df
> var=2 * MS[6] / 4
> var
[1] 19.85307
> est + c(-1,1) * qt(0.975, 27) * sqrt(var)
[1] -22.892296  -4.607704
```

(g) The true value is -16.25, which is contained within the interval computed in part (f).

```
> mu[1,1] - mu[1,2]
[1] -16.25
```

(h) By slide 41 of set 15, an approximate 95% confidence interval for $\mu_{11} - \mu_{21}$ is

$$\bar{y}_{11\cdot} - \bar{y}_{21\cdot} \pm t_{d,0.975}\sqrt{\widehat{\mathrm{Var}}(\bar{y}_{11\cdot} - \bar{y}_{21\cdot})},$$

where $t_{d,0.975}$ denotes the 0.975 quantile of a t distribution with d degrees of freedom computed by Cochran-Satterthwaite and

$$\widehat{\text{Var}}(\bar{y}_{11\cdot} - \bar{y}_{21\cdot}) = \frac{2}{4 \cdot 4} MS_{\text{Blk}\times\text{Geno}} + \frac{2(4-1)}{4 \cdot 4} MS_{\text{Error}} = \frac{1}{8} MS_{\text{Blk}\times\text{Geno}} + \frac{3}{8} MS_{\text{Error}}$$

Using the `R` code below,

$$\bar{y}_{11\cdot} - \bar{y}_{21\cdot} = -22.5, \quad \widehat{\text{Var}}(\bar{y}_{11\cdot} - \bar{y}_{21\cdot}) = 53.50, \quad d = 11.15,$$

and an approximate 95% confidence interval for $\mu_{11} - \mu_{21}$ is

$$(-38.57, -6.43).$$

This agrees with the interval computed by SAS on page 8 of slide set 17 (titled 'geno 1 - geno 2 with no fertilizer').

```
> est=mean(subset(dat, geno == '1' & fert == '0')$y) -
  mean(subset(dat, geno == '2' & fert == '0')$y)
> est
[1] -22.5
> var=MS[4] / 8 + 3 * MS[6] / 8
> var
[1] 53.50212
> d=var^2 / ( (MS[4]/8)^2/df[4]  + (3 * MS[6]/8)^2/df[6] )
> d
[1] 11.15121
> est + c(-1,1) * qt(0.975, d) * sqrt(var)
[1] -38.572543  -6.427457
```

(i) The true value is -15, which is contained within the interval computed in part (h).

```
> mu[1,1] - mu[2,1]
[1] -15
```

(j) Determine an appropriate standard error for the intercept estimate and find its degrees of freedom.

Intercept is the cell mean μ_{11}. By slide 43 of set 15, the standard error for μ_{11} is

$$\begin{aligned}\sqrt{\widehat{\text{Var}}(\bar{y}_{11\cdot})} &= \frac{1}{3 \cdot 4 \cdot 4} \left[MS_{\text{Blk}} + (3-1) MS_{\text{Blk}\times\text{Geno}} + 3(4-1) MS_{\text{Error}} \right] \\ &= \frac{1}{48} \left(MS_{\text{Blk}} + 2 MS_{\text{Blk}\times\text{Geno}} + 9 MS_{\text{Error}} \right)\end{aligned}$$

The degree of freedom can be computed by Cochran-Satterthwaite using the following code.

$$SE(\mu_{11}) = 7.58, \quad df = 6.74.$$

```
> y11=mean(subset(dat, geno == '1' & fert == '0')$y)
> y11
[1] 126.025
> se=sqrt((MS[1]+MS[4]*2+MS[6]*9)/48)
> se
[1] 7.580552
> df2=se^4 /((MS[1]/48)^2/df[1]+(2*MS[4]/48)^2/df[4]+(9*MS[6]/48)^2/df[6])
> df2
[1] 6.743576
```

3. This is a split-plot experiment, where block = GH, whole-plot factor = WL, and split-plot factor = GENO. We can separate the ANOVA table into whole- and split-plot parts, which has the skeleton

Source	DF
GH	3
WL	2
WP Error (= GH:WL)	6
GENO	1
WL:GENO	2
SP Error (= GH:GENO + GH:WL:GENO)	3+6=9
c. total	(4)(3)(2) - 1 = 23

   (a) The numerator should be based on WL, which is the whole-plot factor. Hence, the denominator should be based on the whole-plot error, GH:WL. Therefore,

$$F = \frac{SS_{\text{WL}}/df_{\text{WL}}}{SS_{\text{GH:WL}}/df_{\text{GH:WL}}} = \frac{321.8/2}{116.4/6} = 8.29.$$

   (b) The numerator should be based on GENO, which is the split-plot factor. Hence, the denominator should be based on the split-plot error, GH:GENO + GH:WL:GENO. Therefore,

$$\begin{aligned} F &= \frac{SS_{\text{GENO}}/df_{\text{GENO}}}{(SS_{\text{GH:GENO}} + SS_{\text{GH:WL:GENO}})/(df_{\text{GH:GENO}} + df_{\text{GH:WL:GENO}})} \\ &= \frac{2.5/1}{(11.7 + 14.5)/(3 + 6)} \\ &= 0.859. \end{aligned}$$

   (c) The numerator should be based on WL:GENO, which is falls under the split-plot part of the ANOVA table. Hence, the denominator should be based on the split-plot error, GH:GENO + GH:WL:GENO. Therefore,

$$\begin{aligned} F &= \frac{SS_{\text{WL:GENO}}/df_{\text{WL:GENO}}}{(SS_{\text{GH:GENO}} + SS_{\text{GH:WL:GENO}})/(df_{\text{GH:GENO}} + df_{\text{GH:WL:GENO}})} \\ &= \frac{75.1/2}{(11.7 + 14.5)/(3 + 6)} \\ &= 12.90. \end{aligned}$$

4. (a) Let

$$\begin{aligned} a_{ik} &= \frac{y_{i1k} + y_{i2k}}{2} = \bar{\mu}_{i.} + p_k + \bar{e}_{i.k} \\ &= \bar{\mu}_{i.} + \varepsilon_{ik}, \end{aligned}$$

where $\varepsilon_{ik} = p_k + \bar{e}_{i.k}$. Note that the ε_{ik} terms are $iid\ N(0, \sigma^2)$, where $\sigma^2 = \sigma_p^2 + \frac{\sigma_e^2}{2}$. Thus, a two sample t-test can be used to test $H_0 : \bar{\mu}_{1.} = \bar{\mu}_{2.}$. From the R output of the analysis of averages, we have

$$t = \frac{84.892 - 80.454}{\sqrt{2.169^2 + 1.534^2}}.$$

(b) Let

$$\begin{aligned} d_{ik} &= y_{i1k} - y_{i2k} = \mu_{i1} - \mu_{i2} + e_{i1k} - e_{i2k} \\ &= \delta_i + \eta_{ik}, \end{aligned}$$

where $\delta_i = \mu_{i1} - \mu_{i2}$ and $\eta_{ik} = e_{i1k} - e_{i2k}$. Note that the η_{ik} terms are $iid\ N(0, \sigma_\eta^2)$, where $\sigma_\eta^2 = 2\sigma_e^2$. The test of infection main effect is a test of $H_0 : \frac{\mu_{11}+\mu_{21}}{2} = \frac{\mu_{12}+\mu_{22}}{2} \iff H_0 : \mu_{11} - \mu_{21} + \mu_{21} - \mu_{22} = 0 \iff \delta_1 + \delta_2 = 0$. From the last analysis of the differences in R, we can test $H_0 : \delta_1 + \delta_2 = 0$ with

$$t = \frac{8.250 + 1.492}{\sqrt{2.439^2 + 1.724^2}}.$$

(c) It is straightforward to see that a test for interaction is a test of $H_0 : \delta_1 = \delta_2 \iff H_0 : \delta_1 - \delta_2 = 0$. Thus,

$$t = \frac{8.250 - 1.492}{\sqrt{2.439^2 + 1.724^2}}$$

is the relevant test statistic.

(d) $\hat{\sigma}_\eta^2 = 2\hat{\sigma}_e^2 = 5.974^2 \implies \hat{\sigma}_e^2 = \frac{5.974^2}{2}$.

(e) $\hat{\sigma}^2 = \hat{\sigma}_p^2 + \frac{\hat{\sigma}_e^2}{2} = 5.313^2 \implies \hat{\sigma}_p^2 = 5.313^2 - \frac{5.974^2}{4}$.

The answers to parts a) through e) above match tests and estimates obtained by fitting the full linear mixed effects model $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$.