1. The null hypothesis is $H_0 : C\beta = (0, 1, -1, 0) \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta \end{bmatrix} = 0$. We want power=0.9 of rejecting this null hypothesis when $H_a : C\beta = 0.5\sigma$. The power of the test is:

$$power = Pr(F_{1,10n-3}(\delta^2) > F_{1,10n-3,.05})$$

where $\delta^2 = \frac{1}{\sigma^2} 0.5\sigma[C(X^TX)^- C^T]^{-1} 0.5\sigma$. In this case,

$$X^TX = \begin{bmatrix} 10n & 5n & 5n & 500n \\ 5n & 5n & 0 & 250n \\ 5n & 0 & 5n & 250n \\ 500n & 250n & 250n & 500n \end{bmatrix}$$

$C(X^TX)^- C^T = 0.4/n$ and $\delta^2 = 0.5 * (1/0.4) * 0.5 = 5n/8$. Choose n such that

$$.90 = power = Pr(F_{1,10n-3}(5n/8) > F_{1,10n-3,.05})$$

The solution is $n = 18$. S-Plus code for computing this sample size is shown below. This code is dangerous in that it may take a long time to execute if the required sample size is very large.

```
> n<-1
> powerfun<-function(n){
  1-pf(qf(.95,1,10*n-3),1,10*n-3,5*n/8) }
> while (powerfun(n)<.90){
     n<-n+1
     powerfun(n)
  }
> n
    [,1]
[1,] 18
```

Additional S-Plus code is in the file posted as power.ssc. This code requires you to specify an upper bound for your serch for the required sample size. It also does not need an explicit formula for the non-centrality parameter. It adds rows to the model matrix as the sample size is increases to compute the value of the noncentrality parameter.

2. (a) $\mathbf{Y} \sim N(\mathbf{X}\underline{\beta}, \sigma^2\mathbf{I})$ where $\mathbf{X}$ is the model matrix, or, $Y_{ij} \sim N(\mu + \alpha_i, \sigma^2)$, for i = 1, ..., 6 and j+1,2, and each Y_{ij} is independent of the others.

   (b) i. This hypothesis is equivalent to $H_0 : \alpha_1 - \alpha_2 = 0$ and $\alpha_2 - \alpha_3 = 0$; or

$$H_0 : C\underline{\beta} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix} \underline{\beta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   Since both $\alpha_1 - \alpha_2$ and $\alpha_2 - \alpha_3$ are estimable and rank(C) = 2 = the number of rows of C, this hypothesis is **testable**.

   ii. This hypothesis is equivalent to $H_0 : \alpha_1 = 0,\ \alpha_2 = 0$ and $\alpha_3 = 0$, but neither α_1 nor α_2 nor α_3 is estimable. To show that α_1 is not estimable, let $\underline{d} = [1, -1, -1, -1, -1, -1, -1]^T$ and $\underline{c} = [0, 1, 0, 0, 0, 0, 0]^T$. Then, $\underline{c}\beta = \alpha_1$ and $\mathbf{X}\underline{d} = 0$, but $\underline{c}^T\underline{d} \neq 0$. Thus, α_1 is not estimable and this hypothesis is **not testable**.

   iii. This hypothesis is equivalent to

$$H_0 : C\underline{\beta} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \underline{\beta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   Since $\mu + \alpha_1$ and $\mu + \alpha_4$ are both estimable and rank(C)=2, this hypothesis is **testable**.

iv. This hypothesis is equivalent to $H_0: \alpha_1-\alpha_2-\alpha_4+\alpha_5=0$ and $\alpha_1-\alpha_3-\alpha_4+\alpha_6=0$. Note that $\alpha_1-\alpha_2-\alpha_4+\alpha_5$ is estimable because $E(\underline{a}\mathbf{Y})=\alpha_1-\alpha_2-\alpha_4+\alpha_5$ for $\underline{a}=[.5,.5,-.5,-.5,0,0,-.5,-.5,.5,.5,0,0]$. Similarly, $\alpha_1-\alpha_3-\alpha_4+\alpha_6$ can also be shown to be estimable. Also note that rank$(C)=2=$ the number of rows of C. Thus, this hypothesis is **testable**. You should recognize the rows of this C matrix as coefficients for interaction contasts.

v. This hypothesis is equivalent to $H_0: \alpha_1-\alpha_2-\alpha_4+\alpha_5=0$ and $\alpha_1-\alpha_3-\alpha_4+\alpha_6=0$ and $\mu+\alpha_6=38$. Note that $\alpha_1-\alpha_2-\alpha_4+\alpha_5$, $\alpha_1-\alpha_3-\alpha_4+\alpha_6$ and $\mu+\alpha_6$ are all estimable and rank$(C)=3=$ the number of rows of C. Thus this hypothesis is **testable**.

vi. Since rank$(C)=2$ is smaller than the number of rows of C (note the first row + the third row = the second row), this hypothesis is **not testable**.

(c) Note that the hypothesis $H_0: C\underline{\beta}=\underline{d}$ is tested by the F statistic given by

$$F=\frac{SS_{H_0}/df_{H_0}}{SSE/df_{resid}}$$

where

$$SS_{H_0}=(C^T\underline{b}-\underline{d})^T(C(X^TX)^-C^T)^{-1}(C^T\underline{b}-\underline{d}),$$

and df_{H_0} and df_{resid} are the degrees of freedom of the null hypothesis and the residuals, respectively. (Here, $df_{resid}=6$ and df_{H_0} =number of rows in C). Also note that if H_0 is true, then $F\sim F_{df_{H_0},df_{resid}}$.

i. The hypothesis to be tested is

$$H_0: (\alpha_1+\alpha_4)/2=(\alpha_2+\alpha_5)/2=(\alpha_3+\alpha_6)/2$$

or

$$H_0: C\underline{\beta}=\begin{bmatrix} 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 1 & -0 & -1 \end{bmatrix}\underline{\beta}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We have $F=117.381$ with df = (2,6) (p-value¡0.0001). The null hypothesis is rejected, i.e., after averaging over the two methods of making cheese the average moisture content, is not the same for all three types of cheese.

ii. The hypothesis to be tested is

$$H_0: \alpha_1=\alpha_4,\ \alpha_2=\alpha_5,\ \alpha_3=\alpha_6$$

or

$$H_0: C\underline{\beta}=\begin{bmatrix} 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}\underline{\beta}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We have $F=1.259$ with df = (3,6) (p-value=0.369). The null hypothesis is not rejected, i.e., the average moisture content is not affected by the method of making cheese.

```
> #          #
> # Y vector #
> #          #
> Y <- c( 39.02, 38.79, 35.74, 35.41, 37.02, 36,
+         38.96, 39.01, 35.58, 35.52, 35.7, 36.04)
> n <- length(Y)
>
> #          #
> # X matrix #
> #          #
> X <- matrix( 0, ncol=7, nrow=n )
> X[,1] <- 1
> X[1:2, 2] <- 1; X[3:4, 3] <- 1; X[5:6,4] <- 1
> X[7:8,5] <- 1;  X[9:10,6] <- 1; X[11:12,7] <- 1
> X
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
 [1,]    1    1    0    0    0    0    0
 [2,]    1    1    0    0    0    0    0
 [3,]    1    0    1    0    0    0    0
 [4,]    1    0    1    0    0    0    0
 [5,]    1    0    0    1    0    0    0
 [6,]    1    0    0    1    0    0    0
```

```
  [7,]    1    0    0    0    1    0    0
  [8,]    1    0    0    0    1    0    0
  [9,]    1    0    0    0    0    1    0
 [10,]    1    0    0    0    0    1    0
 [11,]    1    0    0    0    0    0    1
 [12,]    1    0    0    0    0    0    1
>
> #                    #
> # estimation of beta #
> #                    #
>
> betahat <- ginverse(t(X) %*% X) %*% t(X) %*% Y
> Px <- X %*% ginverse(t(X) %*% X) %*% t(X)
> In <- diag( rep(1,n) )
>
> SSE <- t(Y) %*% ( In- Px) %*% Y
> SSE
        [,1]
[1,] 0.66195
> df.sse <- n - qr(X)$rank
> df.sse
[1] 6

>
> #             #
> # 2 (c) (i)   #
> #             #
>
> c1 <- matrix( c( 0, 1, -1, 0, 1, -1, 0,
+                  0, 1, 0, -1, 1, 0, -1 ), ncol=7, byrow=T )
> cbhat1 <- c1 %*% betahat
>
> SS.H0.c1 <-  t(cbhat1) %*% solve(c1 %*% ginverse(t(X) %*% X)
+   %*% t(c1))%*% cbhat1
> SS.H0.c1
         [,1]
[1,] 25.90012
>
> F1 <- (SS.H0.c1/2) / (SSE/(df.sse))
> F1
        [,1]
[1,] 117.381

> pvalue <- 1 - pf( q=F1, df1=2, df2=df.sse )
> pvalue
 [1] 0.00001547711

>
> #             #
> # 2 (c) (ii)  #
> #             #
>
> c2 <- matrix( c( 0, 1, 0, 0, -1, 0, 0,
 +                 0, 0, 1, 0, 0, -1, 0,
 +                 0, 0, 0, 1, 0, 0, -1 ), ncol=7, byrow=T )
> c2
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
 [1,]    0    1    0    0   -1    0    0
 [2,]    0    0    1    0    0   -1    0
 [3,]    0    0    0    1    0    0   -1
> cbhat2 <- c2 %*% betahat
>
> SS.H0.c2 <-  t(cbhat2) %*% solve(c2 %*% ginverse(t(X) %*% X) %*% t(c2))
%*% cbhat2
> SS.H0.c2
         [,1]
[1,] 0.416625
>
> F2 <- (SS.H0.c2/3) / (SSE/(df.sse))
> F2
          [,1]
[1,] 1.258781
> pvalue <- 1 - pf( q=F2, df1=3, df2=df.sse )
> pvalue
[1] 0.0.3693709
```

3. (a) From the S-plus code provided, we have the parameter estimates and their standard errors as below. The associated ANOVA table is also shown.

$$\hat{Y}_i = 15.4324 + 2.8374(time)$$

$$(2.8374) \quad (0.0231)$$

```
Analysis of Variance Table
Response: length
          Df Sum of Sq  Mean Sq  F Value         Pr(F)
     time  1  348.9206 348.9206 31.61394 0.00006289022
Residuals 14  154.5169  11.0369
```

(b) Note $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$ and let $A_1 = P_1$, $A_2 = P_X - P_1$ and $A_3 = I - P_X$. Then,

- A_1, A_2 and A_3 are all $n \times n$ symmetric matrices,
- $I = A_1 + A_2 + A_3$, and
- rank(A_1) + rank(A_2) + rank(A_3) = 1 + (2-1) + (n-2) = n.

Now apply Cochran's Theorem to show

- $\frac{1}{\sigma^2}\underline{Y}^T A_i \underline{Y} \sim \chi^2_{r_i}\left(\frac{1}{2\sigma^2}(X\underline{\beta})^T A_i X\underline{\beta}\right)$, where $r_i = \text{rank}(A_i)$, for each $i \in \{1, 2, 3\}$, and
- $\underline{Y}^T A_1 \underline{Y}$, $\underline{Y}^T A_2 \underline{Y}$, $\underline{Y}^T A_3 \underline{Y}$ are distributed independently.

Since rank$(P_X - P_1)$=1,

$$\frac{1}{\sigma^2}\underline{Y}^T(P_X - P_1)\underline{Y} \sim \chi^2_1(\delta^2)$$

where

$$\delta^2 = \frac{1}{\sigma^2}(X\underline{\beta})^T(P_X - P_1)X\underline{\beta}$$

(c) As shown in part (b), $\frac{1}{\sigma^2}\underline{Y}^T(P_X - P_1)\underline{Y}$ and $\frac{1}{\sigma^2}\underline{Y}^T(I - P_X)\underline{Y}$ are independent chi-square distributions with 1 df and $n-2$ df, respectively. Then,

$$\frac{\frac{1}{\sigma^2}\underline{Y}^T(P_X - P_1)\underline{Y}/1}{\frac{1}{\sigma^2}\underline{Y}^T(I - P_X)\underline{Y}/(n-2)} = \frac{R(\beta_1|\beta_0)}{MS_{resid}} \sim F_{1,14}(\delta^2)$$

where

$$\delta^2 = \frac{1}{\sigma^2}(X\underline{\beta})^T(P_X - P_1)X\underline{\beta}.$$

(d) We need to examine the non-centrality parameter δ^2 carefully. Note that

$$\begin{aligned}
\delta^2 &= \frac{1}{\sigma^2}(X\underline{\beta})^T(P_X - P_1)X\underline{\beta} \\
&= \frac{1}{\sigma^2}[(X\underline{\beta})^T(I - P_1)X\underline{\beta} - (X\underline{\beta})^T(I - P_X)X\underline{\beta}] \\
&= \frac{1}{\sigma^2}(X\underline{\beta})^T(I - P_1)X\underline{\beta} \qquad ((X\underline{\beta})^T(I - P_X)X\underline{\beta} = 0) \\
&= \frac{1}{\sigma^2}(X\underline{\beta})^T(I - P_1)(I - P_1)X\underline{\beta} \\
&= \frac{1}{\sigma^2}[(I - P_1)X\underline{\beta}]^T(I - P_1)X\underline{\beta} \\
(I - P_1)X &= [(I - P_1)\underline{1} \quad (I - P_1)\underline{X}] \qquad \underline{X} = [X_1\ X_2\ \ldots\ X_n]^T \\
&= [\underline{0} \quad (\underline{X} - \bar{X}\underline{1})] \qquad \bar{X} = \sum_{i=1}^{n} X_i/n \\
(I - P_1)X\underline{\beta} &= \beta_1(\underline{X} - \bar{X}\underline{1}).
\end{aligned}$$

Then,

$$\begin{aligned}
\delta^2 &= \frac{1}{\sigma^2}\beta_1^2(\underline{X} - \bar{X}\underline{1})^T(\underline{X} - \bar{X}\underline{1}) \\
&= \frac{1}{\sigma^2}\beta_1^2\sum_{i=1}^{n}(X_i - \bar{X})^2
\end{aligned}$$

Since $X_1 \neq X_2$, for example, $\delta^2 = 0$ if and only if $\beta_1 = 0$. Therefore, the null hypothesis is $H_0 : \beta_1 = 0$, and the non-centrality parameter δ^2 is zero if and only if this null hypothesis is true.

(e) We have $F = 31.61$ with df=(1,$n-2$)=(1,14) and p-value $= 0.0001$, suggesting that the coefficient β_1 is significantly different from zero. (This result can be easily obtained from the ANOVA table produced by S-Plus or SAS. In the following code, however, this it was computed manually.)

```
> n <- nrow(crystals)
> ones <- rep(1,n)
> X <- cbind( ones, crystals$time )
> Y <- crystals$length
> Px <- X %*% ginverse(t(X) %*% X) %*% t(X)
> P1 <- ones %*% ginverse(t(ones) %*% ones) %*% t(ones)
> In <- diag( rep(1,n) )
> SS.resid <- t(Y) %*% ( In - Px ) %*% Y
> MS.resid <- SS.resid / (n-2)
> c(SS.resid, MS.resid)
[1] 154.51691  11.03692
> R.b1.b0 <- t(Y) %*% (Px - P1) %*% Y
> R.b1.b0
         [,1]
[1,] 348.9206
>
> F.b1 <- R.b1.b0/MS.resid
> pvalue <- 1 - pf( q=F.b1, df1=1, df2=n-2 )
> round( c(F.b1, pvalue), 4)
[1] 31.6139  0.0001
```

(f) The plot of the residuals against time exhibits a quadratic pattern, indicating that the proposed straight line model is inadequate.

4. We can follow the same argument as in Problem 3(b) using Cochran's theorem but here note that now we have $\underline{Y} \sim N(Z\underline{\lambda}, \omega^2 I)$.

(a)

$$\frac{1}{\omega^2}\underline{Y}^T(I - P_X)\underline{Y} \sim \chi^2_{n-2}(\delta^2) \qquad \text{with} \qquad \delta^2 = \frac{1}{\omega^2}(Z\underline{\lambda})^T(I - P_X)Z\underline{\lambda}$$

(b)

$$\frac{1}{\omega^2}\underline{Y}^T(P_X - P_1)\underline{Y} \sim \chi^2_1(\delta^2) \qquad \text{with} \qquad \delta^2 = \frac{1}{\omega^2}(Z\underline{\lambda})^T(P_X - P_1)Z\underline{\lambda}$$

(c) No. For a ratio of chi-square distributions to have an F distribution, the denominator must have a central chi-square distribution. Here, the non-centrality parameter for $\frac{1}{\sigma^2}\underline{Y}^T(P_X - P_1)\underline{Y}$ is not zero because $Z^T(I - P_X)Z$ is not a matrix of zeros.

5. Now we have $\underline{Y} \sim N(X\underline{\beta}, \sigma^2 I)$ again.

(a)

$$\frac{1}{\sigma^2}\mathbf{Y}^T(I - P_Z)\mathbf{Y} \sim \chi^2_{n-3}(\delta^2)$$

$$\delta^2 = \frac{1}{\sigma^2}(X\underline{\beta})^T(I - P_Z)X\underline{\beta}$$

(b) Yes. Here the non-centrality parameter for $\frac{1}{\sigma^2}\mathbf{Y}^T(I-P_Z)\mathbf{Y}$ is zero since

$$\begin{aligned}
\delta^2 &= \frac{1}{\sigma^2}(X\underline{\beta})^T(I-P_Z)X\underline{\beta} \\
&= \frac{1}{\sigma^2}(X\underline{\beta})^T(I-P_Z)P_X X\underline{\beta} && (X = P_X X) \\
&= \frac{1}{\sigma^2}(X\underline{\beta})^T(P_X-P_Z P_X)X\underline{\beta} \\
&= \frac{1}{\sigma^2}(X\underline{\beta})^T(P_X-P_X)X\underline{\beta} && (P_Z P_X = P_X) \\
&= 0
\end{aligned}$$

6. Note that $\underline{Y} \sim N(W\underline{\gamma}, \tau^2 I)$.

(a) Following the same argument used in Problem 3, it is easy to show that

$$\begin{aligned}
\frac{1}{\sigma^2}\underline{Y}^T(I-P_W)\underline{Y} &\sim \chi^2_{n-7} \\
\frac{1}{\sigma^2}\underline{Y}^T(P_W-P_Z)\underline{Y} &\sim \chi^2_4\left(\frac{1}{\tau^2}(W\underline{\gamma})^T(P_W-P_Z)W\underline{\gamma}\right).
\end{aligned}$$

and that these quadratic forms are independent random variables. Consequently, $F = \frac{\underline{Y}^T(P_W-P_Z)\underline{Y}/(4)}{\underline{Y}^T(I-P_W)\underline{Y}}$ has a non-central F-distribution with (4,9) degrees of freedom. Look at the non-centrality parameter carefully.

$$\begin{aligned}
\frac{1}{\tau^2}(W\underline{\gamma})^T(P_W-P_Z)W\underline{\gamma} &= \frac{1}{\tau^2}[(W\underline{\gamma})^T(I-P_Z)W\underline{\gamma} - (W\underline{\gamma})^T(I-P_W)W\underline{\gamma}] \\
&= \frac{1}{\tau^2}(W\underline{\gamma})^T(I-P_Z)W\underline{\gamma}. \\
&= \frac{1}{\tau^2}((I-P_Z)W\underline{\gamma})^T(I-P_Z)W\underline{\gamma}.
\end{aligned}$$

This is zero if and only if $(I-P_Z)W\underline{\gamma} = \underline{0}$. Note that

$$\begin{aligned}
(I-P_Z)W\underline{\gamma} &= (I-P_Z)[\ Z \quad \underline{w}_4 \quad \underline{w}_5 \quad \cdots \quad \underline{w}_{10}\ ]\underline{\gamma} \\
&= [\ \mathbf{0} \quad (I-P_Z)\underline{w}_4 \quad (I-P_Z)\underline{w}_5 \quad \cdots \quad (I-P_Z)\underline{w}_{10}\ ]\underline{\gamma} \\
&= \alpha_1(I-P_Z)\underline{w}_4 + \alpha_2(I-P_Z)\underline{w}_5 + \ldots + \alpha_7(I-P_Z)\underline{w}_{10} \\
&= (I-P_Z)(\alpha_1\underline{w}_4 + \alpha_2\underline{w}_5 + \ldots + \alpha_7\underline{w}_{10})
\end{aligned}$$

where $\underline{w}_j$ is the j-th column of W. Since $\underline{w}_4, ..., \underline{w}_{10}$ span a space of dimension 6 and the columns of Z span a space of dimension 3, it follows that $\alpha_1\underline{w}_4 + \alpha_2\underline{w}_5 + \ldots + \alpha_7\underline{w}_{10}$ is not in the space spanned by the columns of Z for all possible values of $\alpha_1, \alpha_2, ..., \alpha_7$. Furthermore, $\alpha_1\underline{w}_4 + \alpha_2\underline{w}_5 + \ldots + \alpha_7\underline{w}_{10}$ is in the space spanned by the columns of Z if and only if it is equal to $Z\underline{c}$ for some vector $\underline{c}$. In that case,

$$(I-P_Z)(\alpha_1\underline{w}_4 + \alpha_2\underline{w}_5 + \ldots + \alpha_7\underline{w}_{10}) = (I-P_Z)Z\underline{c} = \underline{0}$$

and the noncentrality paramter is zero. However, when

$$\alpha_1\underline{w}_4 + \alpha_2\underline{w}_5 + \ldots + \alpha_7\underline{w}_{10} = Z\underline{c}$$

the model becomes $\underline{Y} = Z(\gamma_1\ \gamma_2\ \gamma_3)^T + Z\underline{c} + \underline{\epsilon}$ which is the model from problem 3. Hence, the noncentrality parameter is zero if and only if the model from problem 3 is correct.

(b) It is not appropriate to use $\underline{Y}^T(I-P_Z)\underline{Y}/(n-3)$ here. When the null hypothesis (that the model in problem 3 is correct)is false, $\frac{1}{\sigma^2}\underline{Y}(I-P_Z)\underline{Y}$ has a chi-square distribution with 7 degrees of freedom and non-centrality parameter

$$\delta^2 = \frac{1}{\tau^2}(W\underline{\gamma})^T(I-P_Z)W\underline{\gamma},$$

which is not zero. The noncentrality parameter is zero when the null hypothesis (that the model in problem 3 is correct) is true. In that case, $rank(W) = 3$ and $\underline{Y}^T(I - P_Z)\underline{Y}$ should be divided by $(n-3)$ instead of $(n-7)$.