1. Hand written.

2. Hand written.

3. (a) H_0: $\sigma^2_{batch} = 0$

   H_A: $\sigma^2_{batch} > 0$

   Based on an $F_{4,20}$ statistic of 5.54 and p-value of < 0.0036 there is strong evidence of batch to batch variation.

```
chi.cutL <- pchisq(0.025,20, lower.tail = TRUE)
chi.cutU <- pchisq(0.975,20, lower.tail = TRUE)

sigma2 <- 0.00483
sse <- 0.0876

ll <- sse/chi.cutL
ul <- sse/chi.cutU

ci <- c(sigma2-ll,sigma2+ul)
```

   σ^2 was estimated to be 0.00438 and the approximate 95% confidence interval is between ci[1] and 6.5257365×10^8. That is, the within batch variation calcium content is estimated to be between ci[1] and 6.5257365×10^8 at the 95% confidence level.

   Note that σ^2 must be positive, but the confidence interval given includes negative values.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	4	0.096976	0.024244	5.54	0.0036
Error: MS(Error)	20	0.087600	0.004380		

   (b) $\hat{\sigma^2}_{batch} = MS_{batch} = 0.024244$

   $\hat{\sigma^2} = MS_{error} = 0.00438$

   $\hat{\sigma^2}_t = \frac{MS_{batch} - MS_{error}}{5} = 0.0038812$

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.09697600	0.02424400	5.54	0.0036
Error	20	0.08760000	0.00438000		
Corrected Total	24	0.18457600			

   (c)
```
FOL <- pchisq(0.025,5, lower.tail = TRUE)
FOU <- pchisq(0.975,5,lower.tail = TRUE)
```

```
FL <- qf(0.025,4,20)
FU <- qf(0.975,4,20)


chi.cutL4 <- qchisq(0.025,4, lower.tail = TRUE)
chi.cutU4 <- qchisq(0.975,4, lower.tail = TRUE)

sigma2B <- 0.024244
sst <- 0.09697600

frac.s <- sigma2B/(sigma2B + sigma2)

llB <- (sst*(1-(FU/F0U)))/5*chi.cutL4
ulB <- (sst*(1-(FL/F0L)))/5*chi.cutU4

ciB <- c(frac.s-llB,frac.s+ulB)
```

The percent of batch to batch variability relative to total variability was 0.8338722 with an estimated 95% confidence interval of between 1.7561655 and -4845.290826 percent.

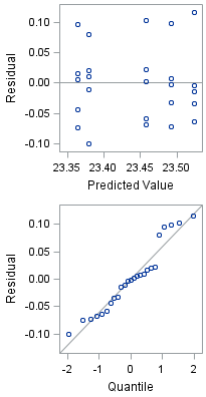
(d) Although the problem said to do an ANOVA, I fit a random effects model.

It appears there is a slight violation of the normality assumption as the quantiles of the residuals show slight deviations than those expected under normality in the tails. This is not a major concern.

It is reasonable to assume the variation is the same within each batch as shown by the fitted vs. residuals plot.

The problem description said that batches were randomly chosen, which makes them independent.

The problem description did not say the "five determinations" were randomly chosen, and so observations within each batch may not be independent of other observations in the same batch.



4. (a) **MODEL**

$y = \mu + \tau_{1hr} + \tau_{2hr} + \tau_{4hr} + \epsilon$

y: BMD loss

μ: The average BMD loss in the SHAM group

τ_{1hr}: The change in average BMD loss between the SHAM group and the 1hr/day PEMF group

τ_{2hr}: The change in average BMD loss between the SHAM group and the 2hr/day PEMF group

τ_{4hr}: The change in average BMD loss between the SHAM group and the 4hr/day PEMF group

ϵ: random error

ANOVA ASSUMPTIONS

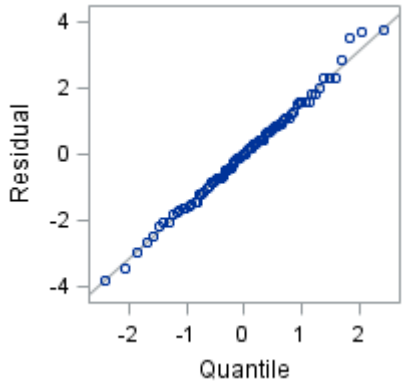
$\epsilon_i \sim iidN(0, \sigma^2)$

The errors are independently and normally distributed with a mean of 0 and with the similar variance of σ^2.

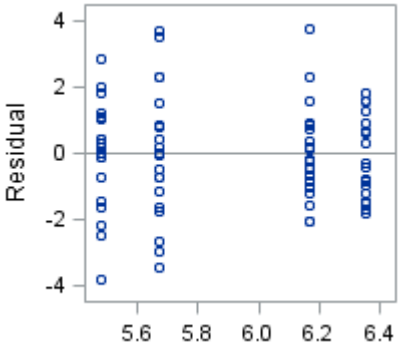
(b) Note that assuming $\Sigma\tau_i = 0$ specifies the means model, which is a different parameterization than the effects model written in (a).

Parameter	Estimate	Standard Error
PEMF 1h/day	0.24887500	0.31100703
PEMF 2h/day	-0.43862500	0.31100703
PEMF 3h/day	0.43387500	0.31100703
Sham	-0.24412500	0.31100703

(c) With the Normal Q-Q plot we are assessing whether it is reasonable to assume the residuals are normally distributed. There is slight deviation from normality in the upper tail of the distribution, but for the most part it is reasonable to assume normality in the residuals.



(d) There is slightly less variation in the residuals from the PEMF1hr/day treatment and the PEMF3hr/day treatment groups, which is particularly noticeable in the PEMF3hr/day treatment group. The variances appear similar enough that the slightly less variable group may be due to random chance and is likely not a concern in terms of the accuracy of ANOVA results interpretations.



(e) Higher bone mineral densities are desirable. We are given responses of percent losses in bone mineral density. Bone mineral density is ideal, and so we would like to minimize the percent loss in bone mineral density.

We would like to do a lower tailed test for whether the percentage of BMD loss has decreased.

Using 90% family-wise confidence intervals, there is no evidence that any of the PEMF treatments reduce the percentage of BMD loss when compared to SHAM.

Comparisons significant at the 0.1 level are indicated by ***.				
Trt Comparison	Difference Between Means	Simultaneous 90% Confidence Limits		
PEMF3 - Sham	0.6780	-Infinity	1.5686	
PEMF1 - Sham	0.4930	-Infinity	1.3836	
PEMF2 - Sham	-0.1945	-Infinity	0.6961	

(f) At the 90% confidence level, it is estimated that PEMF1hr/day caused percentage loss in BMD to be 1.3836% or less on average than those receiving SHAM. The study description did not indicate random sampling, so inference was only made to "those" in the study.

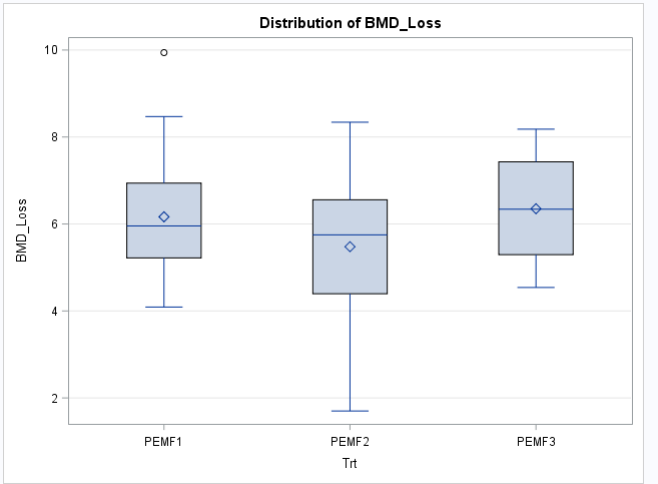
5. (a) $\Gamma_L = -1\tau_{1hr} + 0\tau_{2hr} + 1\tau_{4hr}$

$\Gamma_Q = 1\tau_{1hr} + -2\tau_{2hr} + 1\tau_{4hr}$

(b) There was no evidence of the linear trend and some evidence of the quadratic trend. The boxplots are consistent with this as the median is slightly above 6 in the PEMF1 and PEMF3 (corresponds to PEMF4) treatments and slightly below 6 in the PEMF2 treatment.

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Linear	1	0.34225000	0.34225000	0.16	0.6881

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Quadratic	1.56000000	0.79420901	1.96	0.0544

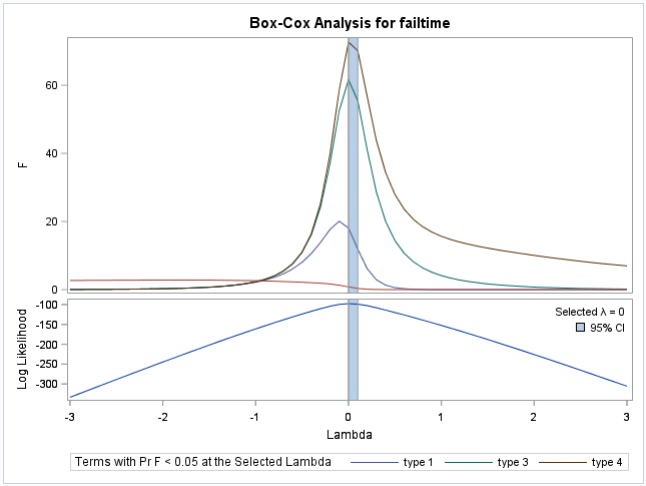


(c) If I had used all treatment groups in the analysis, $SS_L + SS_Q \neq SS_{trt}$ because there are four treatment groups, we would need to add SS_{cubic} to the linear and quadratic SS as that will completely partition the SS_{trt}. However, I omitted the SHAM group from the analysis so the mathematical statement $SS_L + SS_Q = SS_{trt}$ should be true as there are 3 treatment groups and therefore we can fit up to an order of 2 polynomial model.

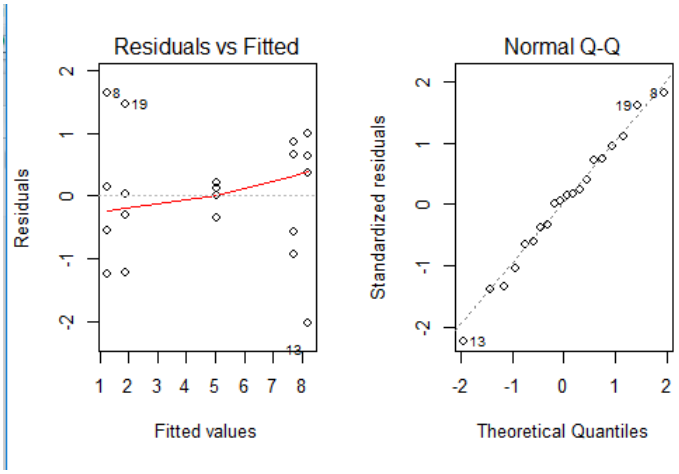
6. Yes, there is a problem. Using the coefficients for the contrasts that were given in the notes we have to have equally spaced increments of treatments if the treatments are on the continuous scale and we are performing tests of linear and quadratic orthoganol contrasts. In this problem, the treatment increments are on the continuous scale and are not equally spaced.

   A side question: Does the order of categorical treatments matter for these coefficients? In problem (4) we treated going from PEMF1hr/day to PEMF2hr/day as the same increment as PEMF2hr/day to PEMF4hr/day when doing the orthoganol trend contrasts, which may not be correct in that case unless that assumption is reasonable. Is that correct?

7. (a) The selected transformation was $\lambda = 0$. Therefore, we should log the response to stabilize the variance.



(b) Normality is reasonable, however, of major concern is the violation of homogeneity of variance as one group has much less variable residuals than the others. The p-values from the ANOVA will be too small and we will find significant differences more often than we would expect if the HOV assumption was met.




```
mat <- read.csv("material.csv")
mat$failtime.ln <- log(mat$failtime)

mat.mean <- tapply(mat$failtime.ln, as.factor(mat$type),mean)
mat.s <- data.frame((tapply(mat$failtime.ln, as.factor(mat$type),var)))
colnames(mat.s) <- "Sample Var"
rownames(mat.s) <- c("1", "2", "3", "4", "5")

print(xtable(mat.s))
```

	Sample Var
1	0.06
2	1.53
3	0.79
4	1.86
5	1.23

```
lm1 <- lm(failtime.ln~as.factor(type), data = mat)

par(mfrow=c(2,2))
#plot(lm1)
```

(c) Based on an $F_{4,15} = 37.657$ and a p-value of < 0.001 there is strong evidence that one of the material types has a different mean failure time than the others.

```
summary(lm1)
```

```
Call:
lm(formula = failtime.ln ~ as.factor(type), data = mat)

Residuals:
     Min       1Q   Median       3Q      Max
-2.00945 -0.55716  0.08608  0.64969  1.64792

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)        5.0516     0.5234   9.651 7.95e-08
as.factor(type)2  -3.8091     0.7402  -5.146 0.000119
```

```
as.factor(type)3   2.6648     0.7402   3.600 0.002625
as.factor(type)4   3.1624     0.7402   4.272 0.000668
as.factor(type)5  -3.1476     0.7402  -4.252 0.000695

Residual standard error: 1.047 on 15 degrees of freedom
Multiple R-squared:  0.9094,Adjusted R-squared:  0.8853
F-statistic: 37.66 on 4 and 15 DF,  p-value: 1.176e-07
```

```
anova(lm1)
```

```
Analysis of Variance Table

Response: failtime.ln
                Df  Sum Sq Mean Sq F value    Pr(>F)
as.factor(type)  4 165.056  41.264  37.657 1.176e-07
Residuals       15  16.437   1.096
```

8. Below we see in the table that the means increase from a to c, however, the 95% bonferroni adjusted CI for a-b is completely below zero and the 95% bonferroni adjusted CI for a-c contains zero. These would correspond to rejecting the null of no difference and failing to reject the null of no difference respectively.

```
set.seed(7)
a <- rnorm(100,10,5)
b <- rnorm(1000, 14, 10)
c <- rnorm(5,12,5)

dat <- data.frame(matrix(0,nrow=length(c(a,b,c)), ncol = 2))
dat$norm <- c(a,b,c)
dat$group <- c(rep("a",length(a)), rep("b", length(b)), rep("c",length(c)))
dat <- data.frame(dat)

lm.dat <- lm(norm~as.factor(group), data = dat)
#summary(lm.dat)

write.csv(dat, file = "prob7.csv")

tb <- c(mean(a), mean(b), mean(c))
tb <- data.frame(tb)

rownames(tb) <- c("a", "b", "c")
colnames(tb) <- c("means")

print(xtable(tb, align = "||l|l||"))
```

	means
a	10.69
b	13.77
c	20.61

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	1102
Error Mean Square	90.16932
Critical Value of t	2.39764

Comparisons significant at the 0.05 level are indicated by ***.

group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
c - b	6.8393	-3.3680	17.0467	
c - a	9.9122	-0.5212	20.3455	
b - c	-6.8393	-17.0467	3.3680	
b - a	3.0728	0.6850	5.4607	***
a - c	-9.9122	-20.3455	0.5212	
a - b	-3.0728	-5.4607	-0.6850	***