

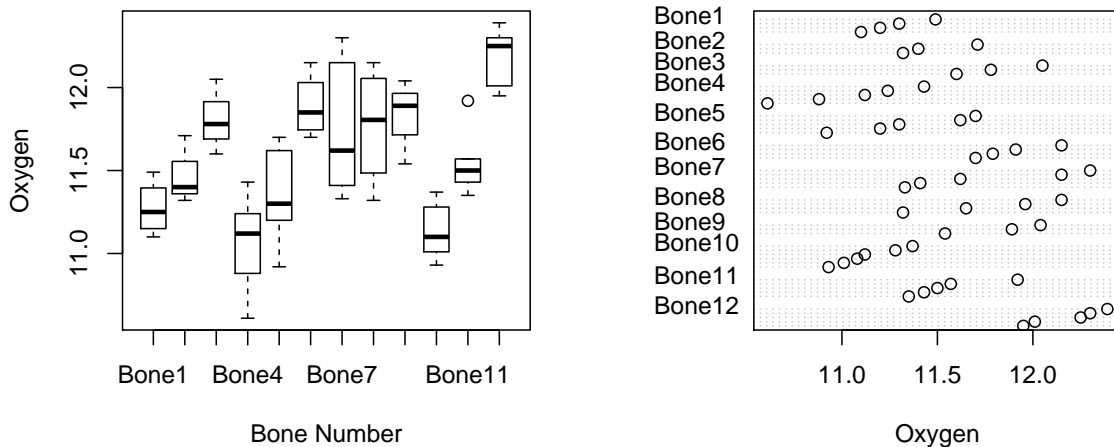
Stat 505 Assignment 2 Solutions

1. Source	df	SS	MS	F	p-value
Between groups	2	18.828	9.414	15.28	1.17e-06
Within groups	124	76.384	0.616		
Total	126	95.212			

2. Barrick and Showers (*Science* v 265:222-224) collected data on oxygen isotopic composition for 12 vertebrae bones (each measured three or more times) from a single *Tyrannosaurus rex* specimen. If compositions are all similar, it would provide evidence that T.rex was warm blooded. If there are strong differences, we would think the dinosaur was cold blooded. We will examine the hypothesis that the dinosaur was warm blooded and that oxygen composition is similar for all vertebrae.

Model: $y_{ij} = \mu + \tau_i + \epsilon_{ij}$ $i = 1, \dots, 12$ $j = 1, 2, \dots, n_i$

Hypotheses: $H_0 : \tau_1 = \tau_2 = \dots = \tau_{12}$, versus $H_A : \text{not all means are equal}$.

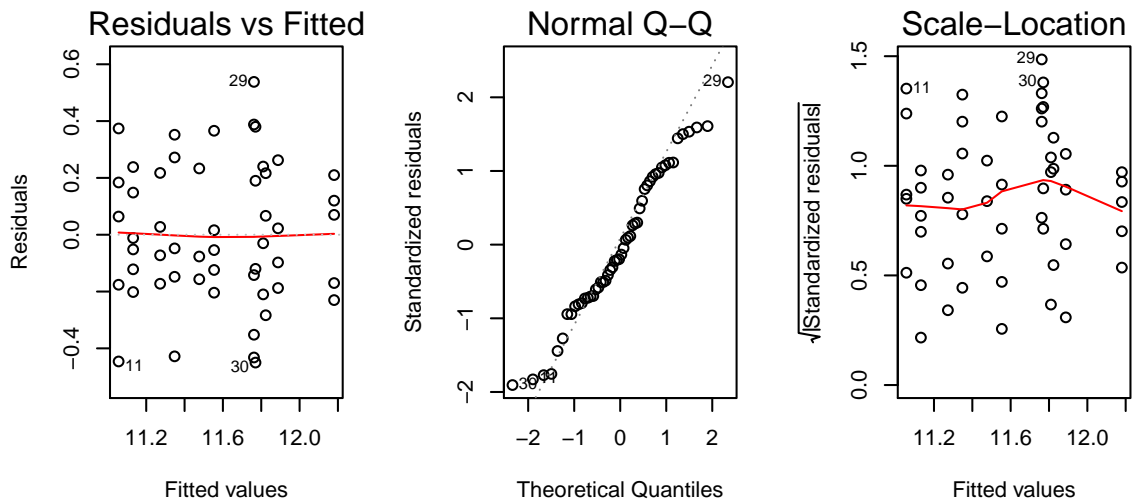


I begin with a box plot and a dot chart to see how much variation there is from bone-to-bone versus within the same bone. I see substantial variation from bone to bone. To test the hypothesis, we need to fit the anova model and do the F test. We get an F stat of 7.42 on 11 and 40 df, with a p-value of 9.7×10^{-7} . This is very strong evidence that mean O_2 composition varies from bone to bone.

I did check the diagnostic plots, and found little reason to worry. Variance seems fairly constant (no trend in Spread-location plot) and residuals are short-tailed in the normal quantile comparison.

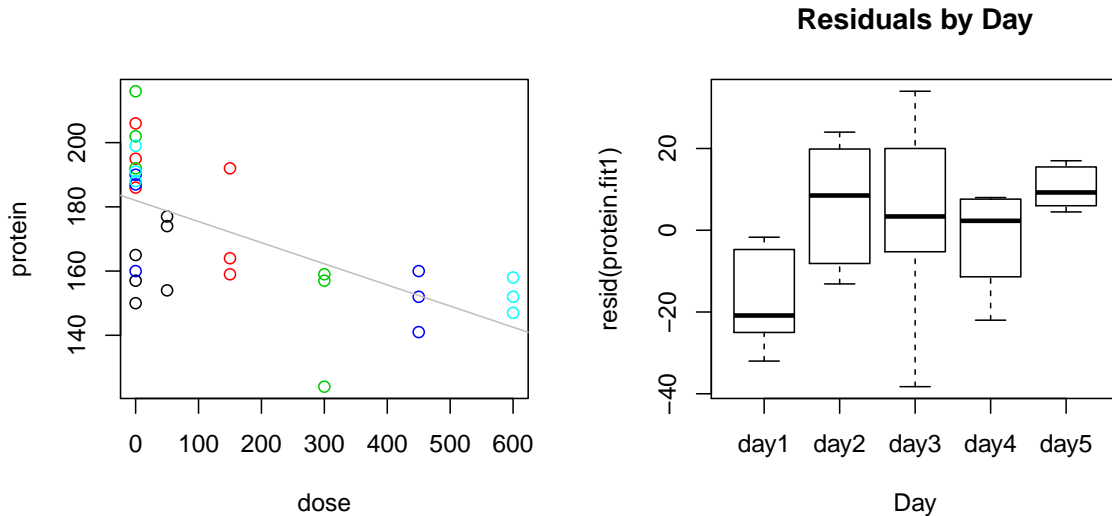
Conclusion: With such a small p-value, we can reject the null hypothesis and conclude that mean O_2 composition varies from bone to bone. Does it vary enough to conclude

the *T.Rex* was warm-blooded? It is a case of significant statistical results which might or might not be scientifically conclusive. Turning to my friend Google, I've learned that in the cold-blooded Komodo dragon, temperature estimates based on isotope analysis show a 9°C difference between bones of the body core and those of extremities (tail, hands, feet) whereas in a warm-blooded cow, the difference in mean temperature estimates is only 1 degree. Source: <http://www.nytimes.com/1992/08/11/science/bone-clues-suggest-warmblooded-dinosaur.html> Unfortunately we don't know how to convert the isotope data into temperatures. I did go back to my source and identified the bones that were core from those in tail, feet and hands. A test of the core versus extremity effect returns $F_{1,10} = 0.724$ with p -value 0.41 (I used bone as a random effect, which you can't do yet) so we would conclude that core mean O_2 composition is not different in the two groups of bones. That suggests that this *T.rex* was warm-blooded. Since within a species, we must have all the same temperature type, one can make inference to the back to the population of *T.rex*'s, but only based on biology, not on the statistics. Other scientists have strengthened the argument by comparing bone structures (which look similar to warm-blooded bird bones).



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
bones	11	6.07	0.55	7.43	0.0000
Residuals	40	2.97	0.07		

3. An experiment reported in Ramsey and Schafer (2002) *The Statistical Sleuth* involves six levels of fatty acid CPFA assigned at random to lab rats. The response is the level of a liver protein.



We first look for any difference in means, testing:

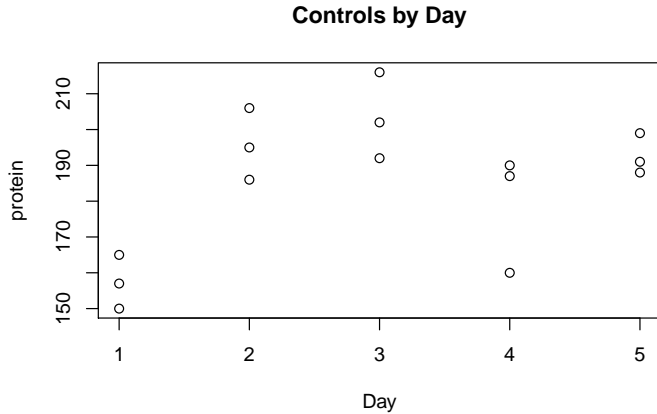
$H_0 : \tau_1 = \tau_2 = \dots = \tau_6$, versus H_A : not all means are equal.

The plot shows a decline in protein as cpfa is increased. Fitting an anova model with a separate mean for each dose gives $F = 4.883$ on 5 and 24 df for a p -value: 0.003, so we have strong evidence that the means are not all the same for each treatment. The picture suggests that much of the change in protein level occurs as a linear function of dose. So I fit protein to a regression on dose with just 1 df. In the regression model, we have only two parameters, slope β_1 and intercept β_0 . The means are defined by slope and intercept, as the control mean is β_0 , and at dose x_i , the mean is $\beta_0 + \beta_1 x_i$. To see if

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
dose	1	5680.31	5680.31	18.40	0.0002
Residuals	28	8642.49	308.66		

the simple linear model was about as good as the full ANOVA model, I compared the two with the `anova` function in R. It provides an F of 1.30 on 4 and 24 df for a rather large p -value of 0.297. We fail to reject the null hypothesis that the extra mean parameters are all zero, and for simplicity, we could use the linear model.

We do have a problem with days on which the data were collected. If we plot just the controls – which should have a single mean – we see that the three controls on day 1 are low, as is one control on day 4. This could happen if the technicians are learning how to make good measurements and improved their technique between day one and day two. We must also worry that their learning also biasing the protein levels of the treatment groups, so it might help to adjust for days of the week.



Adding day seems to improve the regression model, as the F stat is 3.43 on 4 and 24 df with p -value 0.024. With day in the model, dose has a slightly stronger effect with slope estimate -0.0839 and $SE = 0.015$ as compared to an estimate of -0.0658 and $SE = 0.015$ without day. However, seeing the shift in controls from day 1 to days 2-5 along with the fact that each dose was run on only one day tells me that day and treatment are confounded to some degree. The statistics do not take into account possible measurement errors and biases, so I have lost confidence in any model I might fit. Adding days as a factor into the model does not cure the poor study design.

This problem could have been avoided if the researchers had run several doses on each day instead of using this design:

	Their Design						Suggested Design					
	0	50	150	300	450	600	0	50	150	300	450	600
day1	3	3	0	0	0	0	2	0	0	2	0	2
day2	3	0	3	0	0	0	2	2	2	0	0	0
day3	3	0	0	3	0	0	2	0	2	0	0	2
day4	3	0	0	0	3	0	2	2	0	0	2	0
day5	3	0	0	0	0	3	2	0	0	2	2	0

For instance, if we take as a constraint that they can run just 6 per day, 2 of those could be control, and we could choose 2 treatment levels at random to replicate twice under the constraint that each treatment is used in duplicate on each of two days. We would lose some precision about the level of protein for controls, but the comparisons would be much nicer. Also, it appears that they ran the doses in increasing order instead of randomly assigning a dose to each day. Without random allocation of treatments, we cannot make causal inference.

If we did not have the problem of the confounding day variable, we would infer that CPFA causes decrease in the liver protein, but could extend back only to the rats in the lab from which these rats were sampled. Researchers use strains of rats which are highly uniform genetically, so expert opinion might allow one to push the inference back to all lab rats.

R Code

```
#####
SST <- 95.212 - 76.384
MST <- SST/2
MSE <- 76.384/124
Fstat <- MST/MSE
print(rbind(c(2,SST,MST,Fstat, 1-pf( Fstat , 2,124)),
            c(124,76.384,MSE,NA,NA)))

#####
trex <- read.table("http://www.math.montana.edu/~jimrc/classes/stat505/data/Trex02.txt",head=T)
  ## reorder numerically
  bones <- factor( as.character(trex$bonegrp), levels=paste("Bone",1:12,sep=""))
par(mfrow=c(1,2))
boxplot(oxygen ~ bones, data=trex)
with(trex, dotchart(oxygen, groups=bones))

#####
trex.fit <- lm(oxygen ~ bones, trex)
anova(trex.fit)
par(mfrow=c(1,3))
plot(trex.fit, which = c(1,2,3))
  ## adding in core versus extreme bones:
trex$core <- factor(rep(1:0,c(29,23)), labels=c("core","extreme"))
require(nlme)
anova(lme(oxygen ~ core, random = ~1|bonegrp, trex, method="ML"))
#           numDF denDF  F-value p-value
#(Intercept)      1    40 15057.189 <.0001
#core             1    10   0.827 0.3846

#####
cpfa <- read.table("http://www.math.montana.edu/~jimrc/classes/stat505/data/fattyAcid-RS5-18.txt",head=T)
levels(cpfa$treatmnt)[1] <- "cntr0"
dose <- as.numeric(substr(levels(cpfa$treatmnt),5,10))[unclass(cpfa$treatmnt)]
plot(protein ~ dose, data = cpfa, col = as.numeric(day))
protein.fit1 <- lm( protein ~ dose, data = cpfa)
abline( protein.fit1, col="grey")
protein.fullfit <- lm( protein ~ treatmnt, data = cpfa)
anova(protein.fit1)
anova(protein.fit1, protein.fullfit)
plot(resid(protein.fit1)~ cpfa$day, xlab = "Day", main = "Residuals by Day")

#####
plot(protein ~ unclass(day), cpfa , subset = treatmnt=="cntr0", xlab = "Day", main = "Controls by Day")
  anova(lm( protein ~ day +dose+treatmnt, data = cpfa))

#####
asRun <- with(cpfa, table(day, dose))
xtable::xtable( asRun, label="As Run", digits=0)
mydesign <- 0*asRun
mydesign[,1] <- 1 -> diag(mydesign)
mydesign[cbind(1:5,2:6)] <- 1
mydesign[,2:6] <- mydesign[, sample(2:6)]
mydesign <- mydesign[, sample(5),]
dimnames(mydesign) <- dimnames(asRun)
xtable::xtable(mydesign*2, label="Proposed Design", digits=0)
```