

# Permutation Tests & Bootstrapping

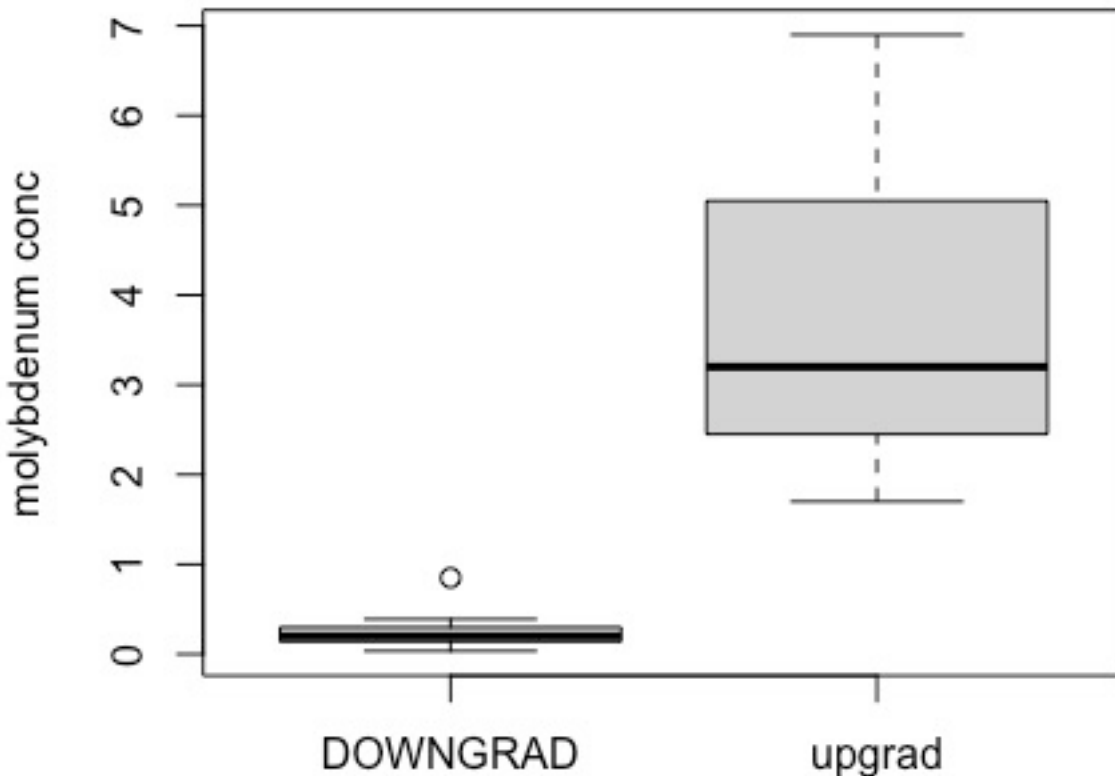
a Practical Stats Online Course

Solutions to Homework

May 17, 2016

## Comparing Two Independent Groups

```
attach(MOLY2)
boxplot(MOLY~LOCAT,ylab="molybdenum conc", col="light gray")
```



Certainly looks like there should be a difference in means. Let's check whether after subtracting off the group means, the residuals have a normal distribution. This is the assumption for a t-test – the difference from group means for both groups together have a normal distribution. We subtract off the means using the `lm` (linear model – regression) command. You could do it by hand but this is quicker.

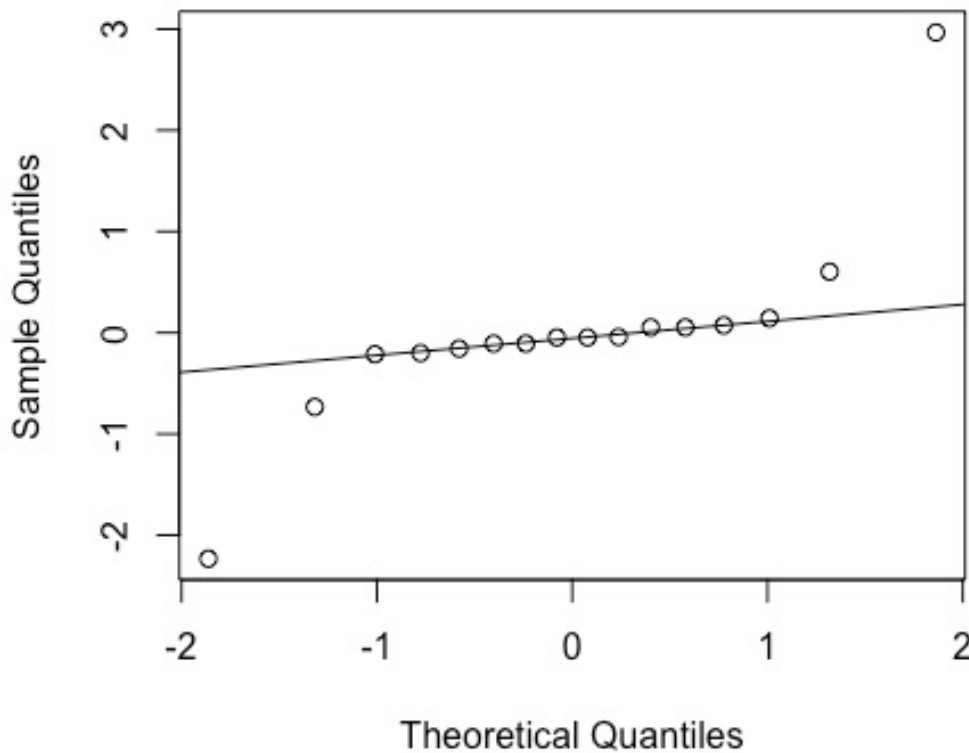
```
shapiro.test(residuals(lm(MOLY~LOCAT)))
qqnorm(residuals(lm(MOLY~LOCAT)))
qqline(residuals(lm(MOLY~LOCAT)))
```

with the results showing that the data do not follow a normal distribution (assumed hypothesis of a normal distribution is rejected – p-value is small):

Shapiro-Wilk normality test

```
data: residuals(lm(MOLY ~ LOCAT))
W = 0.716, p-value = 0.0002546
```

### Normal Q-Q Plot



So this may be a problem for the t-test, but let's just hope all is OK and run it. We're running a 2-sided alternative of difference (one could suspect that a 1-sided test is more appropriate where  $\text{upgrad} > \text{DOWNGRAD}$ , but that's for another day).

```
> t.test(MOLY~LOCAT)
```

Welch Two Sample t-test

```
data: MOLY by LOCAT
t = -2.3836, df = 2.0057, p-value = 0.1396
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -10.321151  2.949254
sample estimates:
mean in group DOWNGRAD    mean in group upgrad
      0.2473846           3.9333333
```

With a p-value of 0.139, the null hypothesis of no difference is NOT rejected – no evidence for a difference in mean concentrations – IF the data followed a normal distribution, which they do not. Is the non-normality affecting the t-test? Very likely. Run a permutation test instead.

```
> choose(16,3)
[1] 560
```

There are only 560 possible rearrangements of this small dataset, so let's compute them all using the "exact.ce" or complete enumeration option of the permTS command:

```
> permTS(MOLY~LOCAT,method='exact.ce',
control=permControl(tsmethod='abs'))
```

Exact Permutation Test (complete enumeration)

```
data:  MOLY by LOCAT
p-value = 0.001786
alternative hypothesis: true mean LOCAT=DOWNGRAD - mean LOCAT=upgrad
is 0
sample estimates:
mean LOCAT=DOWNGRAD - mean LOCAT=upgrad
-3.685949
```

This is a very different result! The p-value is 0.0017, and so we state that the two groups have different means, with strong evidence (a small p-value) for doing so.

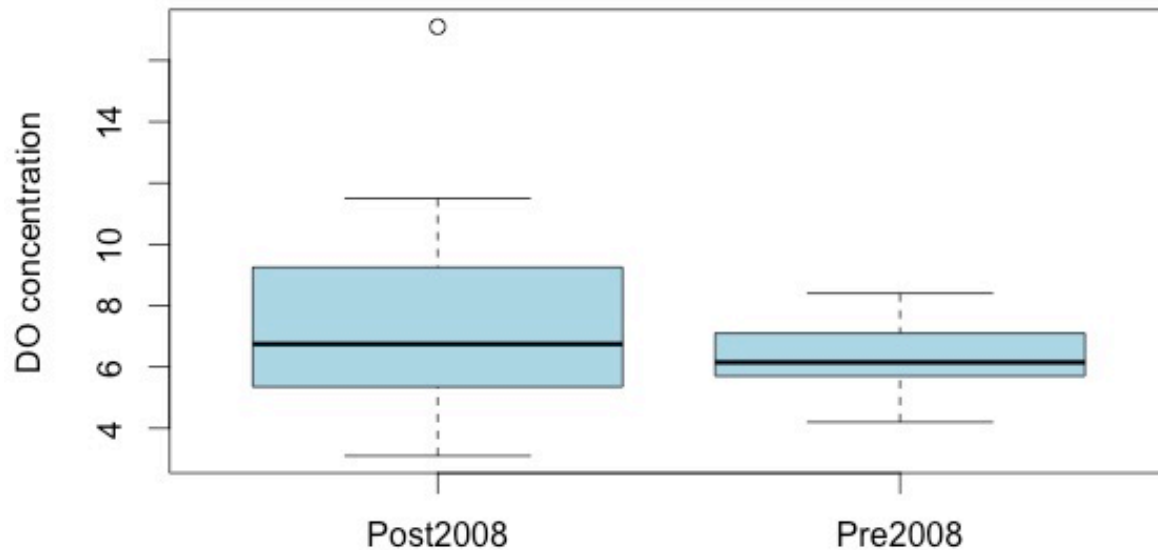
Permutation tests are a huge help for (especially) small, skewed data. That's a very common situation in environmental studies.

## Testing differences between two independent groups.

Using DissOx.RData,

1. *Plot the data first.*

```
> boxplot(DO~Pre.Post,ylab="DO concentration", col="light blue")
```



2. *Run a t-test to determine whether the mean has changed. Don't forget to also test for normality of the residuals.*

```
> shapiro.test(DO[Pre.Post=="Pre2008"])
```

Shapiro-Wilk normality test

```
data: DO[Pre.Post == "Pre2008"]
W = 0.98134, p-value = 0.4431
```

```
> shapiro.test(DO[Pre.Post=="Post2008"])
```

Shapiro-Wilk normality test

```
data: DO[Pre.Post == "Post2008"]
W = 0.91107, p-value = 0.006946
```

```
> t.test(DO~Pre.Post)
```

Welch Two Sample t-test

```
data: DO by Pre.Post
t = 2.0974, df = 39.507, p-value = 0.0424
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.03667192 1.99943919
sample estimates:
mean in group Post2008 mean in group Pre2008
```

7.280556

6.262500

One group doesn't follow a normal distribution. You can't just transform data in one of the time periods, so the perm test may have advantages over the t-test. There appears to be a large enough difference in the group means to overcome this, however, as the t-test results are significant.

*3. Run the 2-sample permutation test to see if results differ from the t-test (it is the main subject of this class, after all).*

Use the PermTS function with method='exact.mc' in the perm package to run the permutation test:

```
> permTS(D0~Pre.Post,method="exact.mc",control=permControl
(nmc=5000,p.conf.level=0.95))
```

Exact Permutation Test Estimated by Monte Carlo

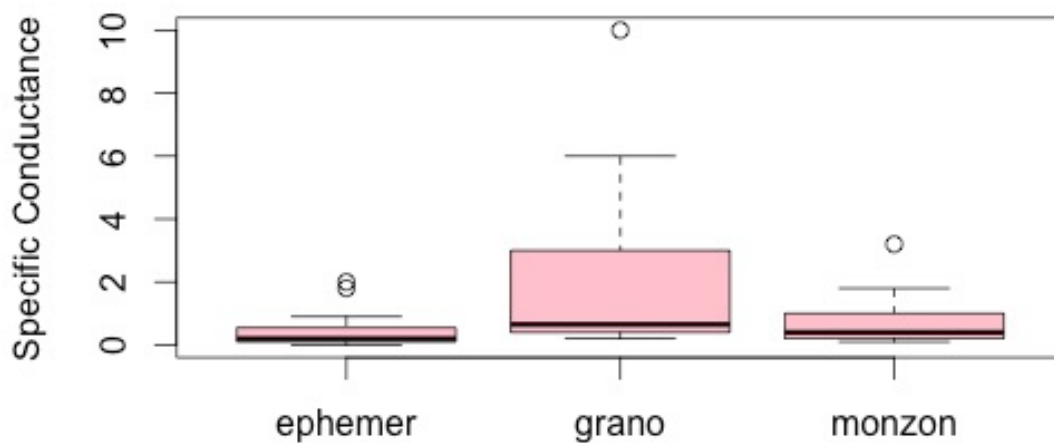
```
data:  D0 by Pre.Post
p-value = 0.009998
alternative hypothesis: true mean Pre.Post=Post2008 - mean Pre.Post=Pre2008 is not
equal to 0
sample estimates:
mean Pre.Post=Post2008 - mean Pre.Post=Pre2008
                        1.018056
p-value estimated from 5000 Monte Carlo replications
95 percent confidence interval on p-value:
 0.006155598 0.014267306
```

The permutation test finds a difference in the means. The non-normality of the Post2008 group, largely caused by the one outlier, doesn't affect the permutation tests as strongly as it did the t-test – the permutation p-values are smaller than the t-test's. The t-test's p-value is outside the 95% CI of the permTS p-value, showing that the permutation test is not giving the same result as the normal-theory t-test.

## Comparing 3+ Groups

Chloride concentrations were measured by Feth et al. (1964) -- *Sources of mineral constituents in water from granitic rocks, Sierra Nevada, California and Nevada; USGS Water Supply Paper 1535-1* – in shallow ephemeral springs and waters from two granitic rock types in the Sierra Nevada mountains. This was a classic paper in geochemistry, not to mention that field work must have been done in awesome scenery. Test whether the mean chloride concentration differs among the groups of springs using anova and the permutation test. Test whether median concentrations differ using the nonparametric analog. Explain your findings.

```
> attach(feth)
> boxplot(Chloride~rocktype,ylab="Specific Conductance",col="pink")
```



### Parametric method: ANOVA

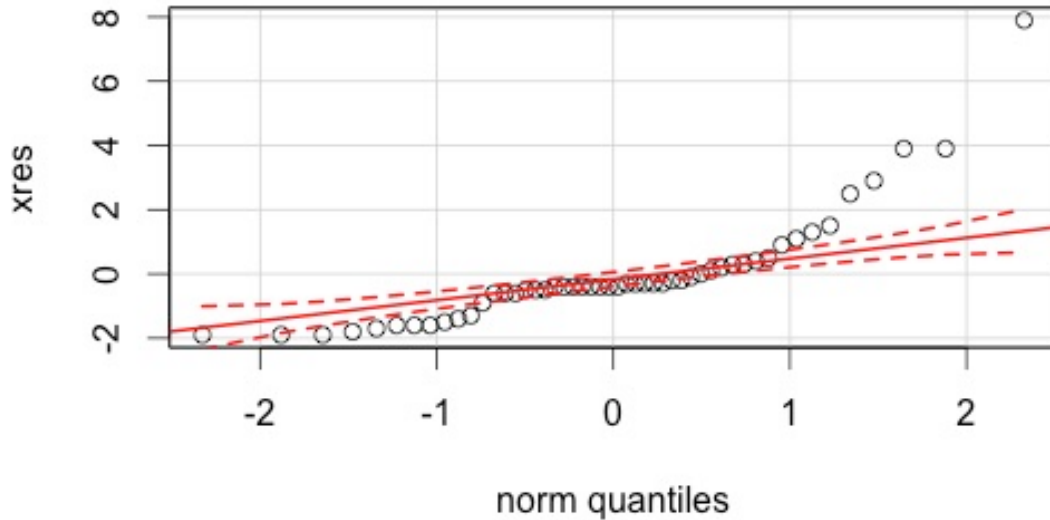
Check residuals for a normal distribution and compute the ANOVA:

```
> xres=residuals(aov(Chloride~rocktype))
> shapiro.test(xres)
```

Shapiro-Wilk normality test

```
data: xres
W = 0.76254, p-value = 1.33e-07
```

```
> qqPlot(xres)    #qqPlot is in the car library
```



```
> summary(aov(Chloride~rocktype))
              Df Sum Sq Mean Sq F value Pr(>F)
rocktype      2  25.93  12.967    4.088 0.0231 *
Residuals    47 149.09   3.172
```

There is a significant difference between the means of the three groups that can be seen even with significant non-normality.

### Permutation method

Use the permKS function in the perm package:

```
> permKS(Chloride,rocktype,method="exact.mc", control=
permControl(nmc=5000,p.conf.level=0.95))
```

K-Sample Exact Permutation Test Estimated by Monte Carlo

```
data: Chloride and rocktype
p-value = 0.0132
```

```
p-value estimated from 5000 Monte Carlo replications
95 percent confidence interval on p-value:
0.01004707 0.01653994
```

The two permutation test results agree (the p-value from perm1way is within the confidence interval for the p-value from permKS). The chloride concentrations are significantly different.

## Testing differences in two paired groups

Using DP2v3\_TP.Rdata

---

*What characteristic is used to pair data between the groups?*

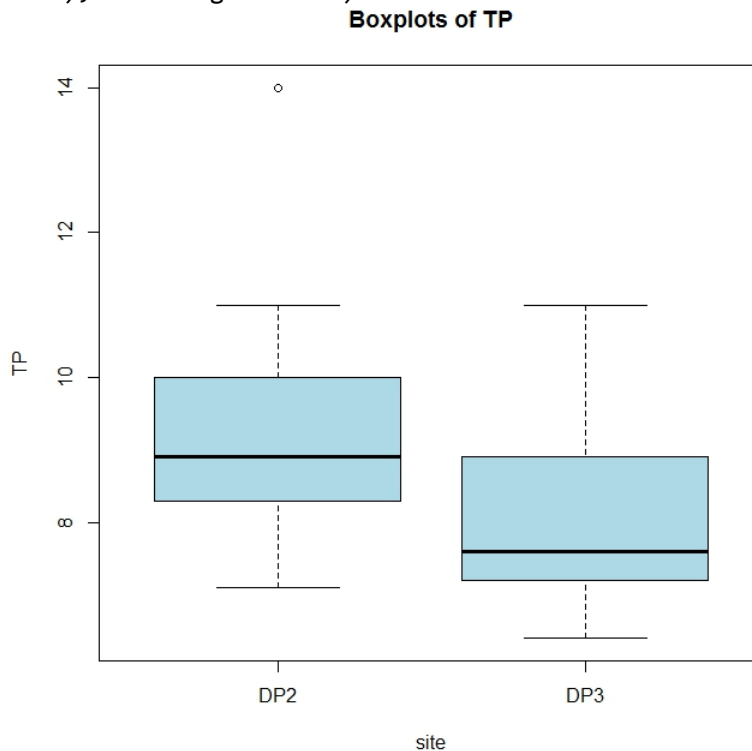
These are paired by date of collection.

*Compute the paired t-test and the paired permutation test and interpret the results.*

```
> attach(DP2v3_TP)
```

Draw a boxplot of data at the two locations. Note the arguments to see how the look of a boxplot can be customized:

```
> boxplot(DP2,DP3, main="Boxplots of TP",ylab="TP", xlab="site",
names=c("DP2","DP3"), col="light blue")
```

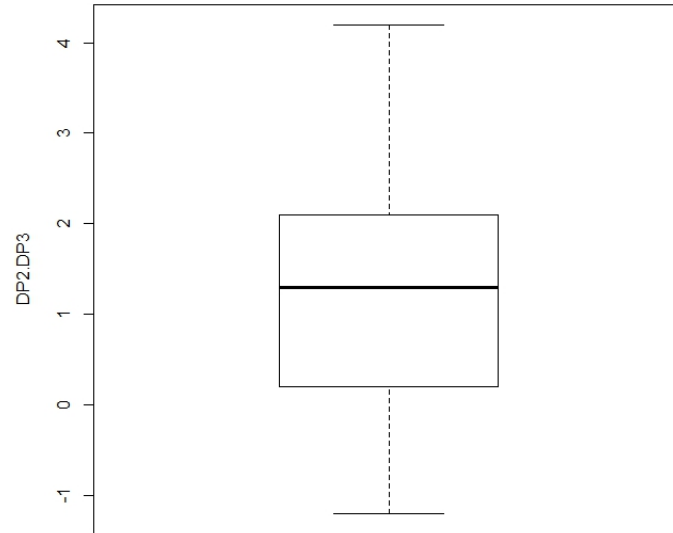


The data appear to be skewed. This might cause problems for a parametric t-test. The two groups do look like their variability is approximately the same, however, so perhaps we will still be able to see differences. The medians appear to differ.

The data are paired by sampling date. By doing so, any differences due to different flows or other influences that change with time of year are factored out. Paired tests actually test whether the mean or median difference DP2-DP3 is zero. A boxplot of the differences is



## Solutions 9



To run the paired t-test:

```
> t.test(DP2, DP3, paired=TRUE)
```

```
Paired t-test
data: DP2 and DP3
t = 5.1391, df = 32, p-value = 1.326e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.7188849 1.6629333
sample estimates:
mean of the differences
      1.190909
```

Note that the 95% confidence interval on MeanDP2-MeanDP3 does not include zero. This states that the p-value for the test will be below  $1-0.95 = 0.05$ , and so we will reject the null hypothesis. Zero is not a plausible 95% estimate for the difference in means. The best estimate is 1.19, with the mean at DP2 higher than at DP3.

The paired permutation test can be computed with the permMatched script. The script also outputs the t-test parametric results:

```
> permMatched(DP2,DP3)
Permutation One-Sample Test
data: DP2 and DP3
  alternative = two.sided
p-value = 0
```

```
The normal-theory one-sample t test is:
t= 5.139149  p= 1.325794e-05
mean = 1.190909  95% CI = ( 0.7188849 1.662933 )
```

Your permutation test result will vary slightly from this because it is randomly selecting only a few thousand of the many thousands of possible results.