STATISTICS 411/511 MIDTERM EXAM 1 Name KEY

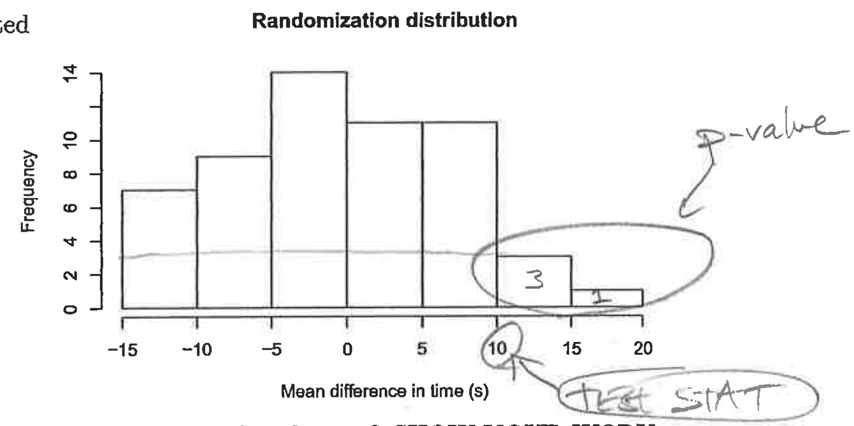
Spring 2018, 200 pts

1. A small study was performed to assess whether a new type of ski decreased skiers' times to complete a course at Bridger Bowl. Eight skiers were randomly selected from Bridger Bowl's ski patrollers to participate. Three skiers were randomly assigned to use the new skis (because there were only 3 sets of the new prototypes available) and the other 5 used conventional skis. The skiers on the new skis had a mean time of 1 minute and 5 seconds. The skiers on the conventional skis had a mean time of 1 minute 15 seconds. Below is the exact randomization distribution for the conventional ski mean time minus the new ski mean time. There are 56 ways to assign 8 units into two groups of 3 and 5.

   (a) (10 pts) State the hypotheses being tested using proper notation.

$H_0: \mu_{conv} = \mu_{new}$

$H_a: \mu_{conv} > \mu_{new}$



   (b) (10 pts) What is the exact *p*-value from the randomization distribution? SHOW YOUR WORK.

Test statistic $= 1'15'' - 1'5'' = 10''$

p-value $= p(\text{Test stat} \geq 10'') = \frac{4}{56} \approx 0.07$

   (c) (10 pts) What is the definition of the above *p*-value?

The probability of observing data (ie test statistic) as extreme or more extreme than actually observed given $H_0: \mu_{conv} = \mu_{new}$ is true.

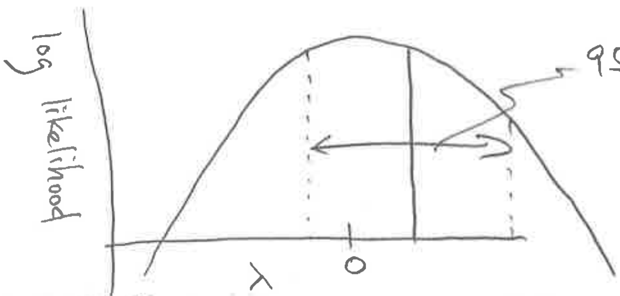
   (d) (30 pts) Give a *Scope of Inference* for this problem.

Because skiers were randomly selected from Bridger ski patrollers, then we can infer these study results to the larger pop. of Bridger ski patrollers. Because of random assignment, we can say that the new skis caused or failed to cause the decrease in ski times.

2. An ANOVA was fit to $\log_{10}$-transformed weights of insects at 4 different Locations, with the following results:

```
Response: log(Weight)
            Df     Sum Sq  Mean Sq    F value    Pr(>F)
Group       --      78.29    26.10     3.2522    0.0243
Residuals  116    -------    -----
```

(a) (10 pts) If the $\log_{10}$-transform was suggested by a Box-Cox transform, draw a picture of the graph from `boxcox()` that would suggest the use of a log-transform.



$$Y = \begin{cases} \log X & \text{if } \lambda = 0 \\ X^{\lambda} & \text{if } \lambda \neq 0 \end{cases}$$

(b) (10 pts) Calculate the missing `Mean Sq` value from the ANOVA table. Show your work! What parameter does it estimate?

$$F = \frac{MSG}{MSE} \Rightarrow 3.2522 = \frac{26.10}{MSE} \Rightarrow MSE = \frac{26.1}{3.2522} = 8.03$$

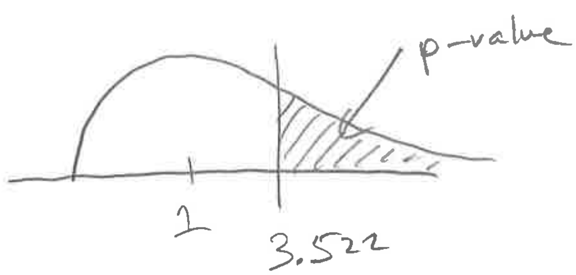
$$\therefore \sigma^2 = \text{constant variance} \doteq 8.03 = MSE$$

(c) (10 pts) Interpret the value of the test statistic in terms of the problem!

F = 3.2522 ⇒ the variability among group means is 3.25 times larger than variability within a group.

(d) (10 pts) How did R calculate the p-value? Indicate the distribution and which tail(s) of the distribution were used.

$p\text{-value} = P(F > 3.522) = 0.0243$ using upper tail from $F(3, 116)$:



(e) (15 pts) The data analyst found that the mean $\log_{10}$-transformed weight for insects in a warmer location were heavier than the mean $\log_{10}$-transformed weight for insects in a colder location (p-value $= 0.009$, t=2.5) with a 95% CI for the mean difference equal to [0.035, 0.921] $\log_{10}(mg)$. Write a proper summary of the results with respect to the original scale of the data. Include in your summary the appropriate p-value, test statistic and CI, and provide an interpretation of the CI that explicitly indicates which location has larger weights.

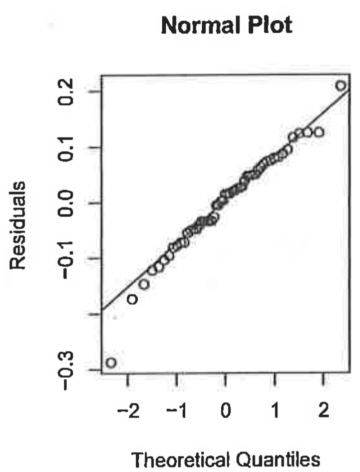
We are 95% confident that insects in warmer locations have a median weight that is 8% to 734% ($10^{0.035}$, $10^{0.921}$) larger than median weight of insects in a colder location ($p = 0.009$, $t = 2.5$)

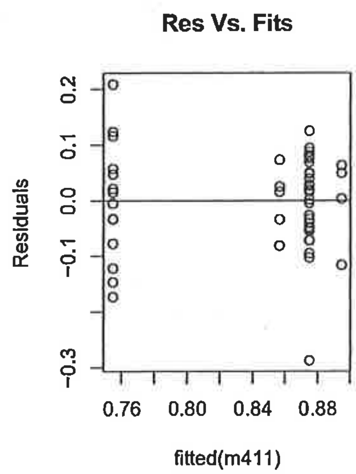
(f) (15 pts) If there are large outliers in the data set, what other kind of ANOVA should you try? Describe how this other ANOVA deals with outliers.

Kruskal Wallis deals with outliers by first assigning ranks to data then applying ANOVA. The magnitude of an outlier does not affect the rank, so Kruskal Wallis is resistant to outliers.

3. A study was conducted at a single hospital in Cleveland to determine the number of healthcare workers (HCW) who enter patient rooms. All 50 rooms in the hospital were monitored for a day. Patient rooms were classified according to the type of hospital unit: intensive care (ICU) with true mean number of HCWs μ_1, surgical (Surg) with true mean μ_2, Oncology with true mean μ_3 or Pediatric with true mean μ_4.

```
m411 = aov(HCW.prop~ UnitType)
summary(m411)
##             Df Sum Sq Mean Sq F value   Pr(>F)
## UnitType     3 1514.9  504.96  6.2801 0.001082 **
## Residuals 49 3939.9   80.41
```





INDEPENDENCE BETWEEN GROUPS: No! Patients and HCWS move between rooms in different units.

```
TukeyHSD(m411)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
##                           diff        lwr          upr      p adj
## Surg-ICU              1.808097  -9.711158   13.3273534  0.9752382
## Oncology-ICU          3.785622 -12.211515   19.7827604  0.9221135
## Pediatric-ICU       -10.218724 -22.642791    2.2053442  0.1410754
## Oncology-Surg         1.977525 -10.716077   14.6711268  0.9757651
## Pediatric-Surg      -1.2026821 -19.745407   -4.3082355  0.0007555
## Pediatric-Oncology  -1.4004346 -27.524395   -0.4842969  0.0397167

estimable(aov(HCW.prop~ UnitType-1),c(0.333,0.333,0.333,-1))
##                           Estimate Std. Error  t value DF     Pr(>|t|)
## c(0.333,0.333,0.333,-1)   11.99576    3.17062 3.783411 __ 0.0004217171
```

(a) (40 pts) What assumptions must be met for the results of the ANOVA to be valid? Check each of these assumptions and state how you are checking the assumption.

DATA must be normal for these small sample sizes in each group. This assumption does not appear to be violated due to normal prob. plot, points close to line.

CONSTANT VARIANCE of data in each group appears not to be violated due to roughly equal spread of residuals in residual vs. fits plot.

Independence within a GROUP: rooms within each unit are not independent! Same HCWS in the units!

(b) (20 pts) The researchers' planned primary research question was whether Pediatric units and ICUs have a different number of HCWs entering and exiting patient rooms on the average. They are disappointed with Tukey's results for this comparison. Is there a "better" test? If so, state explicitly what the better test is and why it can be applied. What will be better about the p-value and CI that you get from this other test?

Because it's planned, we can use a follow-up t-test w/ DFE = 49 and an individual α-level = 0.05 & individual 95% CI. This will result in smaller p-value and more narrow CI compared to Tukeys $p = 0.141$ and 95% CI = [-22.6, 2.2] that maintains a family-wise error rate & confidence level.

(c) (10 pts) The researcher's planned secondary research question was that Pediatric units had a LOWER mean number of HCWs compared to the overall mean of the 3 other units. From the R output above, test this hypothesis. Simply provide the p-value and the degrees of freedom for the test.

$p = 0.0004/2$, DFE = 49.