Karina Sirabian 9/15/20

Activity 3: Intro to Hypothesis Testing

1.

Let = the average methane flux when lemmings do not graze vegetation and

Let = the average methane flux when lemmings graze vegetation

Let

H0:

Ha:

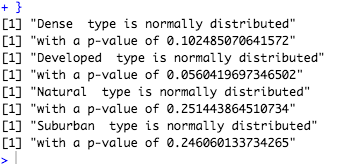
Assuming that both distributions for the average methane flux are normally distributed (based on the results from the shapiro-wilk test in code) and that they both have equal variances (based on the results from the Bartlett test in code), a t-test was run to test if the methane fluxes differ for when lemmings graze vegetation. Since the p-value (0.1399) is greater than , we fail to reject the null hypothesis. We are 95% confident that the methane fluxes do not differ when lemmings graze the vegetation even though > .

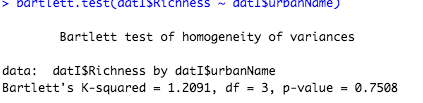
2. To specify you want to run a one-sided t-test, you either set the “alternative” argument to be “less” or “greater” instead of “two.sided” (which is the default value). To specify which value you are testing your mean to be greater than (or less than), set the argument mu to be that value. For example, if you let alternative = “greater” and mu = 1000, then your alternative hypothesis would be testing for if the true mean of your distribution is greater than 1000.

3. H0: There is no difference in insect species richness between types of urban environments

Ha: There is a difference in insect species richness between types urban environments

4. We can assume that the insect data is normally distributed for each urban environment group based on the results from the shapiro-wilk tests (see data from R below) because the p-values are all greater than 0.05. We can also assume that the variance is also relatively equal between both groups based on the result from the Bartlett test (shown below and in R code) because the p-value (0.75) is greater than 0.05. Also, we can assume that Adams et al ensured that these observations are independently sampled from each other and random.





5. I expect the F-ratio in plot A to be higher than the F-ratio in plot B because the F-ratio is used to evaluate differences in data between groups and represents the ratio between the two mean squared values. Thus, a large F ratio means that there is a larger variation between groups and that groups are further away from the global mean. Plot A has more variance in the data than plot B does and the groups in Plot A are all further away from the global mean, so the F-ratio in plot A will probably be higher.

6. Based on the results of the posthoc test (see below), we can assign each urban lancover group a letter for statistical significance. If a p-value between two groups is greater than 0.05, we are 95% confident that there is no difference between those groups, so they would have the same letter. Additionally, if 0 is in the confidence interval then we are confident that there is no difference between groups. On the other hand, if the p-value between 2 groups is less than 0.05, then we are 95% confident that they are different, so the 2 groups would be given different letters.

Developed-Dense: p = 0.88 > 0.05, so they have the same letter (a)

Natural-Dense: p = 0.002 < 0.05, so they have different letters (b, a)

Suburban-Dense: p = 0.14 > 0.05, so they have the same letter (c)

Natural-Developed: p = 0.012 < 0.05, so they have different letters (b, a)

Suburban-Developed: p = 0.621 > 0.05, so they have the same letter (d)

Suburban-Natural: p = 0.065 > 0.05, so they have the same letter (e)

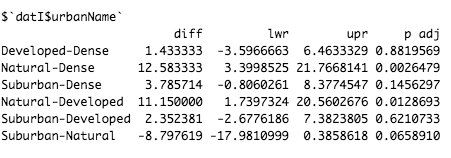
With this information, we can assign letters to each group as follows:

**Dense:**  a, c Since dense, developed, and suburban **Dense:** a

**Developed:** a, d are all similar to each other simplify to **Developed:** a

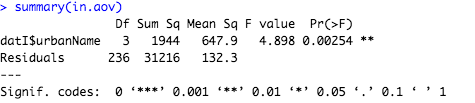
**Suburban:** c, d, e **Suburban:** a, b

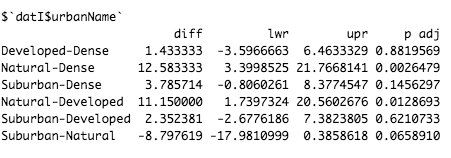
**Natural:** b, e **Natural:** b





7. Based on the results of the ANOVA test (first image below), we are 95% confident that there is a difference in insect species richness between different urban land surfaces because the p-value of the ANOVA test was 0.00254. Since the p-value was less than , we reject the null hypothesis (from number 3) that states that there is no difference in insect species richness between types of urban environments. However, based on the results of Tukey’s Honestly Significant Difference test (second image below), we see that some individual groups do not have a significant difference in insect species richness between them. The only individual pairs of groups that we are 95% confident have a difference in insect species richness are between the groups “Natural-Dense” and “Natural-Developed” because the p-values for those two groups (0.0026, 0.0129) < 0.05. Additionally, 0 is not in their confidence interval indicating that we 95% confident that there is a difference in the groups. All other pairings of groups have a p-value greater than 0.05, so we are 95% confident that there is no difference in insect species richness between these groups (“Developed-Dense”, “Suburban-Dense”, “Suburban-Developed”, and “Suburban-Natural”).



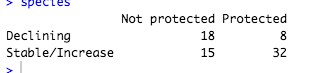


8. Under the null hypothesis that there is no impact of legal protection on species population status, the number of rare plant species with stable/increasing populations should equal the number of rare plant species with decreasing populations within each independent variable. However, this null hypothesis does not affect the number of protected/not protected plants (because this is the independent variable), so the total number of protected plants and total number of not protected plants can be calculated as:

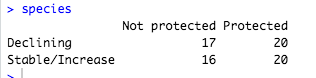
Total number of protected = 32 + 8 = 40

Total number of not protected = 18 + 15 = 33

Because the initial contingency table was



Since there is no impact of legal protection under the null hypothesis, these numbers should be split evenly between the declining and stable/increasing categories. Since 33 is an odd number, it won’t be exactly equal, but we will make it as equal as possible. Thus, under the null hypothesis, the expected values in each category are:



Because 40/2 = 20 and 33/2 = 16.5 -> 17, 16 because we have to use whole numbers

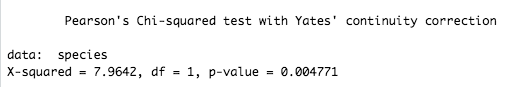
As you can see here, the total number of protected and not protected is still the same, but the number of declining and stable/increasing is now roughly the same. This obeys the null hypothesis because if the relative frequencies of declining and stable/increase are the same, then we can say that there is no impact of legal protection.

9.

H0: There is no relationship between legal protection for rare plant species and plant population status

Ha: There is a relationship between legal protection for rare plant species and plant population status

Let



Since the p-value (0.004) < , we reject the null hypothesis. We are 95% confident that there is a relationship between legal protection and plant population status.

10. <https://github.com/Karina-Sirabian/ENVST_206/blob/master/Activity_3/Activity3.R>