Effects of Oil Spill on Plant Growth

Derek Elliott

University of West Florida
STA 5176 Statistics Modeling

Abstract

To test the effectiveness of their cleanup effort after a large oil spill, a study was performed on 80 equal tracts of land, 40 of which were randomly selected from uncontaminated areas, and 40 from the affected areas. The density of Distichlis spicata was taken at each site and recorded. The mean density and variation were found for both groups and it was found that the densities differed by an amount between 5.5 and 17.6 with 95% confidence.

I. Introduction

On January 7, 1992 an oil pipeline burst in San Patricio County, Texas and contaminated a marsh near the Chiltipin Creek. During the process of cleaning the spill, the marsh was burnt and replanted with local fauna. A year after the burning and replanting, the oil company wanted to evaluate the effectiveness of their effort. In order to do this the company randomly chose 40 tracts of land that were affected by the spill, and 40 that were not. The tracts were as similar as possible. The density of *Distichlis spicata* was to be evaluated to measure the effectiveness.

Figure 1: *Distichlis spicata*

II. Data and Methodology

Based on the boxplot and normal probability graphs, it can be shown that the contaminated sites can be considered normally distributed, but the control sites can not. Because of this, we can use the two population t test under the hypothesis:

$$H_0 : \mu_{con} = \mu_{oil} \quad vs. \quad H_I : \mu_{con} > \mu_{oil}$$

.

Table 1: *Summary Statistics*

Control Tracts		Contaminated Tracts	
Mean:	38.48	Mean:	26.93
Median:	41.5	Median	26.00
St. Dev:	16.37	St. Dev	9.88
n:	40	n:	40

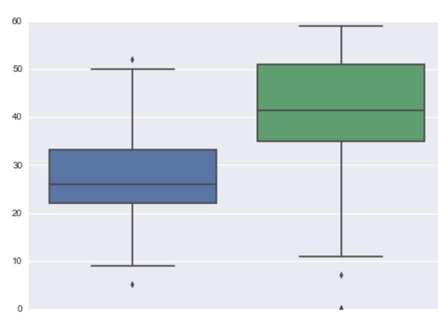

Figure 2: *Box Plot of Contaminated and Control Sites*

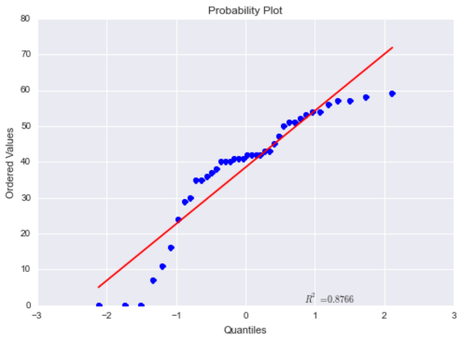

Figure 3: *Normal Probability Plot of Control Sites.*

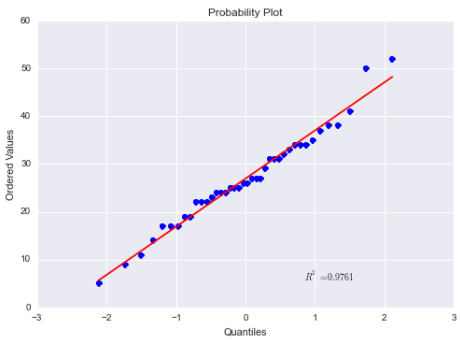

Figure 4: *Normal Probability Plot of Contaminated Sites*

III. Analysis of Data

Because we are using the two population t test, the test statistic is:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_{con}}{n_1} + \frac{S_{oil}}{n_2}}}$$

Where $\bar{x}$, $\bar{y}$, S^2_{con}, and S^2_{oil} are given in Table 1. This gives $t = 3.82$. Given the hypothesis above, in order to reject H_0, it must be shown that $t > t_{\alpha,dof}$. To do this, the degrees of freedom must be computed using the following formula:

$$dof = \frac{(n_{con} - 1)(n_{oil} - 1)}{(1 - c)^2(n_{con} - 1) + c^2(n_{oil} - 1)}$$

Where c is given by

$$c = \frac{\frac{S^2_{con}}{n_{con}}}{\sqrt{\frac{S^2_{con}}{n_{con}} + \frac{S^2_{oil}}{n_{oil}}}}$$

This gives $dof = 64$. Then, from tables, $t_{0.05,64} = 1.699$. Thus, $t = 3.82 > t_{0.05,64} = 1.699$ and H_0 can be rejected. Then, the 95% confidence interval for $\mu_{con} - \mu_{oil}$ is given by:

$$(\mu_{con} - \mu_{oil}) = t_{\alpha/2,dof}\sqrt{\frac{S^2_{con}}{n_{con}} + \frac{S^2_{oil}}{n_{oil}}}$$

Thus, the 95% confidence interval for $\mu_{con} - \mu_{oil}$ is 11.55 ± 6.05.

IV. Concluding Remarks

Based on the results in the previous section, it is known that the density of *Distichlis spicata* is statistically higher in uncontaminated tracts than it is in contaminated tracts, by an amount between 5.05 and 17.6. Further, these results would need to be given back to the environmental scientists who would need to evaluate them and come to a conclusion as to whether the areas that were affected by the oil spill have satisfactorily returned to pre-spill conditions.

References

[Ott and Longnecker, 2010] Ott, R. Lyman. and Longnecker, Michael. (2010). *An Introduction to Statistical Methods and Data Analysis*, 325–330.