Module 7: Assessing Site Reclamation

7.1 Assessing Site Reclamation

Assessing Site Reclamation

- The question of interest is:
  - Has sufficient cleanup been performed to consider a contaminated site reclaimed?
- Reclaimed can mean different things to different people:
  - Perfectly pristine
  - Like it was before it was contaminated
  - Like it would be now if hadn't been contaminated

Assessing Site Reclamation

- Reclaimed can mean different things to different people:
  - Like its neighbors (or some other control site)
  - Clean enough to use for some other purpose
  - Clean enough not to pose a risk
  - As clean as we can make it with the $ we have

Assessing Site Reclamation

- Deciding what “reclaimed” means is not a statistical issue, it’s a policy question
- Generally, “how clean is clean enough” is decided on through a process of negotiation between interested parties like the owner, regulator, and members of the community.
- However, once a standard of cleanliness is set, deciding if it has been met does require the use of statistical methods

Assessing Site Reclamation

- Generally, we will try and err on the side of being conservative.
- Conservative will mean that we will assume the site is in the condition that it has most recently been in until sufficient evidence exists to change that condition.

Assessing Site Reclamation

- Another reason to make this assumption is that the power of a test is a function of sample size.
- Any small difference, even one of no practical significance, can be statistically significant given enough data
- Conversely, an important difference may not be detected from a small sample size
- So, the results of a study can be biased by not taking enough data or taking lots and lots of data

Assessing Site Reclamation

- This is the difference between statistical significance and practical significance
- To be of interest, a result should be both statistically significant (a real difference exists) and practically significant (it's large enough to care about)

Assessing Site Reclamation

- Setting up the hypotheses:
  - If a site has been clean but now is suspected of contamination then the null hypothesis will be that it is equal to a control site, i.e. clean
  - The alternative hypothesis is that it is contaminated.
  - The test checks if sufficient evidence exists to reject the null hypothesis and declare the site contaminated

Assessing Site Reclamation

- Setting up the hypotheses:
  - If the site has been known to be contaminated and we're checking if it has been cleaned up, then the null hypothesis is that it is contaminated
  - The alternative hypothesis is that it is clean
  - Sufficient evidence must exist to reject the null and declare the site reclaimed
  - Note: This is not generally done in practice so may be a new concept to stakeholders. This technique may or may not be acceptable in your workplace.

Assessing Site Reclamation

- Once the hypotheses are set up in this way, standard statistical tests can be used
  - t tests
  - ANOVA
  - Other tests discussed in the Impact Assessment Module

Assessing Site Reclamation

- Alternatively, the concept of bioequivalence can be used.
- A remediated site can be declared to be bioequivalent to a control site if certain site characteristics are similar
  - The site characteristics must be defined
  - Similar must be defined

Assessing Site Reclamation

- Examples of site characteristics that may be used:
  - biomass
  - percent vegetated
  - percent of vegetation that consists of desirable species
  - use by local fauna
  - number of flora and fauna observed
  - prevalence of an indicator species

Assessing Site Reclamation

- Examples of how “similar” might be defined:
  - the mean of the remediated site must be at least some set value
  - the mean must be X% of the control mean
  - the mean must lie within a defined range of the control mean
  - the ratio of the means must be at least some set value

Assessing Site Reclamation

- Once the characteristic has been set
  - Example: biomass at the remediated site
- and similar has been defined
  - equal to the biomass at the control site
- hypothesis are set up
  - Null: sites are not equal therefore the mean of the difference is not zero
  - Alternative: sites are equal therefore the mean of the difference is close to zero

Assessing Site Reclamation

- data can be taken
  - sample 5 randomly selected pairs of sites in remediated and control area
- and analyzed
  - calculate the sample mean and standard deviation of the differences between the remediated and control sites (paired test)

Assessing Site Reclamation

- and hypothesis test is performed
  - use two one-sided test (TOST)
- then a conclusion can be drawn
  - remediated site is clean enough or it isn't

Assessing Site Reclamation

- Example:
  - Question: Is a remediated site clean enough to be considered reclaimed?

$$H_o: \mu_d < \mu_{dL} \text{ or } \mu_d > \mu_{dH}$$

$$H_A: \mu_{dL} <= \mu_d <= \mu_{dH}$$

Sample Mean of the Differences=Sample $Mean_d$ (n=5) = 9.5 grams/meter

Sample $Standard\ Deviation_1$ (n=5) = 5

Sample $Standard\ Deviation_2$ (n=5) = 7

Assessing Site Reclamation

Assume that it's decided that the sites would be considered equal if the mean difference was less than or equal to ten grams per meter

$\mu_{dL} = -10$

$\mu_{dH} = 10$

Assessing Site Reclamation

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$SE(\bar{d}) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$t_L = \frac{\bar{d} - \mu_{dL}}{SE(\bar{d})} \qquad t_H = \frac{\bar{d} - \mu_{dH}}{SE(\bar{d})}$$

Assessing Site Reclamation

- $S_p^2 = (4*5^2 + 4*7^2)/(5+5-2) = 296/8 = 37$
- $SE(d) = 6.08*0.63 = 3.85$
- $t_L = (9.5 - (-10))/3.85 = 5.06$
- $t_H = (9.5 - 10)/3.85 = -0.13$
- Compare the two calculated t values to the appropriate critical values from the t table

Assessing Site Reclamation

- Compare t_L to the upper 5% point of the t distribution with 8 df
- Compare t_H to the lower 5% point of the t distribution with 8 df
- If t_L is greater than or equal to the upper 5% point from the table and t_H is less than or equal to the lower 5% point, Reject Ho

Assessing Site Reclamation

- $t_L = 5.06$
- Compare t_L to the upper 5% point of the t distribution with 8 df = 1.86
- $t_L > 1.86$ so $\mu_d < \mu_{dL}$ is not credible
- $t_H = -0.13$
- Compare t_H to the lower 5% point of the t distribution with 8 df = -1.86
- t_H not< -1.86 so $\mu_d > \mu_{dH}$ is still credible
- So, we can reject the part of the null hypothesis $\mu_d < \mu_{dL}$ but we can't reject the part that suggests $\mu_d > \mu_{dH}$

Assessing Site Reclamation

- So, the conclusion from the example is that, even though 9.5 is less than 10, there still isn't strong enough evidence to reject the null hypothesis.
- Why and what could be done next if you are the researcher?