Part Three: Choosing Appropriate Analyses; Answer Key

Identify the statistical analysis appropriate to address each of the following questions.

Answers in arial font.

1 Water quality in Lake Whatcom: What is the largest source of Phosphorous entering Lake?
Potential sources: Fertilizer runoff from residential properties
Fertilizer from golf courses and parks
Debris flows from logging
Sediment runoff from logged slopes
Failing septic systems

Model selection using AIC, followed by multimodel inference.

2 Does leaf litter from alders increase growth rates of young Douglas fir trees?
Douglas firs grown in plots with and without alder leaf litter.

Two-sample hypothesis test. Two-sample t-test if parametric assumptions met; otherwise Mann-Whitney two-sample test.

3 Do whale watching operations increase dive times of orcas?
Orca dive times were recorded (1) in the absence of boats, (2) in the vicinity of many stationary whale watching boats, and (3) in response to pursuing whale watching boats.

Single factor analysis of variance, followed by multiple comparisons test (e.g., Tukey test).

4 Do residual concentrations of mercury remaining in marine sediments differ among various remediation methods? Mercury concentrations in sediments were sampled after the following treatments: (1) no action, (2) capping with clean sand, (3) hydraulic dredging, (4) clamshell dredging.

Single factor analysis of variance.

5 How different are residual mercury concentrations in marine sediments after the above treatments?

Compare means and confidence intervals. Multiple comparisons test to determine whether differences are statistically significant.

6 Travel mode preferences were surveyed among people commuting to WWU. Do mode preferences differ among students, faculty, and staff ?

Chi-squared test for independence (two factors: travel mode, role at WWU).

7 Surveys described in (6) were repeated five years later, after implementation of an extensive transportation demand management program. Have mode preferences of students, faculty, and staff changed during those five years?

Chi-squared test for independence (three factors: travel mode, role at WWU, survey year).

8 Do native butterflies (pollinators) prefer native flowers over introduced flowers?

Chi-squared test for goodness of fit, k=2.

9 Do controlled burns in forested areas increase densities of nesting woodpeckers? Woodpecker densities were measured in 20 plots prior to burning and in the same plots five years after burning.

Paired sample test. Paired sample t-test if parametric assumptions met; otherwise Wilcoxon pared sample test.

10 Does placing engineered large woody debris (LWD) in streams increase densities of aquatic invertebrates immediately downstream? Aquatic invertebrate densities were sampled in 20 locations: ten in areas without LWD and ten immediately downstream of engineered LWD.

Two-sample hypothesis test. Two-sample t-test if parametric assumptions met; otherwise Mann-Whitney two-sample test.

11 Given large spatial variability in aquatic invertebrate densities, how could you address the question in (10) with greater power, but without increasing sample sizes?

Two approaches: (1) Sample ten sites before and after adding LWD, and use paired-sample test. Using same sample locations controls for spatial variability.
(2) Sample the ten sites immediately downstream of the LWD, and ten sites immediately upstream of the LWD – or as far upstream as necessary to be largely beyond the direct influence of the LWD.

12 Bird species richness in habitats around Whatcom County? Which factor(s) is(are) most important?
Potential factors: Foliage height diversity
Snags (# / hectare)
Habitat type (coniferous forest, deciduous forest, shrub, grassland, wetland, …)
Plant density (stems / hectare)
Habitat area
Canopy cover and snag density

Model selection using AIC, followed by multimodel inference.

13 Research question: What is the oral reference dose (RfD) for methylmercury in children? [RfD assumes there is a threshold exposure for toxic effects. The RfD is an estimate of daily exposure in humans that is likely to have no appreciable risk of deleterious effects during a lifetime (EPA definition).] Performance on a finger-tapping test was measured in two hundred children from each of five methylmercury exposure levels (1000 children total) as assayed from blood samples: (1) 0 to 0.01 μg/kg/day, (2) 0.5 μg/kg/day, (3) 1.0 μg/kg/day, (4) 1.5 μg/kg/day, and 2.0 μg/kg/day.

(a) State two kinds of analyses that could be applied to these data to determine methylmercury RfD in children, where "deleterious effect" is defined as a decrease in finger tapping test performance.

Method 1: Single factor analysis of variance followed by multiple comparisons test (e.g., Tukey test), with k=5 groups. Determine greatest dose for which test performance does not differ significantly from performance at zero exposure (0 to 0.01 mg/kg/day).

Method 2: Simple (linear) regression, with n=1000. Determine greatest dose where performance lies within standard error of the intercept (s_a).

(b) Which kind of analysis identified in (a) would yield the most accurate estimate of the real RfD? Why?

Method 2, because it can be used to interpolate between dosages used in the sample. With analysis of variance and multiple comparisons test (Tukey test), conclusions are restricted to the dosages used in the sample. With interpolation, the estimated RfD is likely to be closer to the real RfD than the sampled dosages. (Note: this answer is sensitive to violation of the assumption of a linear relationship. If the data show that test performance depends on methylmercury exposure in a nonlinear manner, then data transformations or nonlinear regression should be used.) For more information on methylmercury RfD studies, see:

http://www.epa.gov/iris/subst/0073.htm.

For the following research questions, determine the following:

(a) the kind of statistical analysis appropriate to address the question
(b) null and alternative hypotheses
(c) assumptions necessary in using the appropriate statistical analysis
(d) criteria for rejection of the null hypothesis or hypotheses
(e) formula for calculating the test statistic

14 Do WWU students favor left vs. right hands in the same proportion as the global human population? Use this class as a sample of WWU students.

a) Chi-squared test for goodness of fit, k=2.

b) p(right) = p(right, human pop); p(left) = p(left, human pop)

c) This class is a random sample of WWU students, relative to hand preference.

d) $\nu = k-1 = 1$. $\chi^2_{calc} \geq \chi^2_{crit}$ If $\alpha = 0.05$, $\chi^2_{0.05,1} = 3.841$

e) Use continuity correction because k=2. $$\chi_c^{\ 2} = \sum_{i=1}^{2} \frac{\left(\left|f_i - \hat{f}_i\right| - 0.5\right)^2}{\hat{f}_i}$$

15 (In memory of the late Linus Pauling) Does daily consumption of vitamin C (ascorbic acid) tablets reduce cold virus infection rate? An experimental trial was conducted using 1000 human subjects, 250 each given daily doses of one of the following: (1) placebo, (2) 250 mg, (3) 500 mg, (4) 1000 mg.

Two approaches appropriate: (1) Single factor analysis of variance; (2) simple linear regression, if relationship is linear or can be made linear by transformation of dosages (X).

Approach (1): Single factor analysis of variance

a) Single factor analysis of variance.

b) H_0: $\mu_{placebo} = \mu_{250} = \mu_{500} = \mu_{1000}$

c) Residuals (differences between measurements and means) are normally distributed.
Sample variances are equal.

d) groups DF = k-1 = 4-1 =3; error DF = N-k = 1000 – 4 = 996
Criterion: $F_{calc} \geq F_{crit}$ If $\alpha = 0.05$, $F_{crit} = F_{0.05(1),3,996} = 2.62$ (use error DF = 500)
Tukey test: $q_{calc} \geq q_{crit}$ If $\alpha = 0.05$, $q_{crit} = q_{0.05(1),996,4} = 4.781$ (use error DF = 120)

e) ANOVA: $F_{calc} = \dfrac{groups\,MS}{error MS}$ Tukey test: $q = \dfrac{\overline{X}_B - \overline{X}_A}{SE}$ $\text{SE} = \sqrt{\dfrac{\text{error MS}}{n}}$

Approach (2): Simple linear regression (next page)

15 (continued) Approach (2): Simple linear regression

a) Simple linear regression, if relationship linear or made linear w/ transformation of dosage (*X*).

b) H_0: $\beta \geq 0$

c) Residuals are normally distributed (for each value of *X*, normal distribution of *Y*).
Variance in *Y* equal throughout range of *X*. (Variance of ε equal throughout range of *X*.)
Relationship between *Y* and *X* is linear.

d) Must use *t*-test for one-tailed hypothesis. $t_{crit} = t_{\alpha(1),n-2}$ if α=0.05, $t_{0.05(1),1000-2} = t_{0.05(1),998} \approx 1.647$
(use d.f. = 900)

e) $t_{calc} = b/s_b$ where $s_b = \sqrt{\frac{s^2_{Y\cdot X}}{\sum x^2}}$ were $s^2_{Y\cdot X}$ = residual MS

16 Does driver use of cell phones increase risk of motor vehicle accidents? The accident rate (number of accidents per 1000 driving hours) was recorded for 100 people in each of the three categories (300 people total): (1) no cell phone use, (2) use hands-free cell phones while driving, (3) use hand-held cell phones while driving.

a) Single factor analysis of variance, followed by Tukey multiple comparison test.

b) H_0: $\mu_{none} = \mu_{hands\text{-}free} = \mu_{hand\text{-}held}$ (pairwise hypotheses for Tukey test)

c) Residuals (differences between measurements and means) are normally distributed.
Sample variances are equal.

d) groups DF = *k*-1 = 3-1 =2; error DF = *N*-*k* = 300 – 3 = 297
Criterion: $F_{calc} \geq F_{crit}$ If α = 0.05, $F_{crit} = F_{0.05(1),2,297} = 3.04$ (use error DF = 200)

e) $F_{calc} = \frac{groups\,MS}{error MS}$

Note: This question was studied for real, and published in:
Redelmeier DA, Tibshirani RJ. 1997. *New England J. Medicine* 336(7): 453-458. (13 Feb. 1997)
The authors found that driver cell phone use increased vehicle accident four-fold, for both hand-held and hands-free cell phones (*P* < 0.001). Increased risk was similar regardless of driver characteristics, including age and driving experience. Ironically, “Thirty-nine percent of the drivers called emergency services after [their] collision, suggesting that having a cellular telephone may have had advantages in the aftermath of an event.”