1. (a) We might use subscripts U and L to write $\bar{x}_U = 17.5$ $\bar{x}_L = 13.67$, as these means are *sample statistics*.

   (b) The research question asks if species counts are reduced, so it is a 1-sided alternative hypothesis:

$$\mathbf{H}_0: \ \mu_U - \mu_L = 0 \qquad \text{vs.} \qquad \mathbf{H}_a: \ \mu_U - \mu_L > 0.$$

   (c) *randomization*; we use bootstrapping for confidence intervals

   (d) If this were a randomization distribution it would be centered on the null value 0. So it must be a bootstrap distribution.

   (e) We might write the 12 numbers for unlogged plots on individual slips of paper and place them in one bag. Slips with the 9 numbers for logged plots would be placed in a separate bag. We would sample 12 times with replacement from the unlogged bag and compute a mean, sample 9 times from the logged bag with replacement and compute the mean of those numbers, then subtract the second mean from the first.

   (f) In a matched pairs approach, each case/row gives us a number for both group samples. We might do this by focusing only on plots that will be (or have been) logged, but for which we have species counts right before logging begins. So each plot would contribute a difference in species count, the number prior to logging minus the number 8 years after logging.

2. (iv)

3. A Type II error is more likely when $\alpha = 0.01$ than when $\alpha = 0.05$.

4. (a) $\mathbf{H}_0: \ \mu = 7$ vs. $\mathbf{H}_a: \ \mu \neq 7$.

   (b) The larger n is, the smaller the standard error, so choose $n = 150$.

   (c) In this context, a Type I error means we concluded the average pH-levels for spring water in natural habitats is different than 7 despite that, in truth, it is 7.

   (d) We fail to reject the null hypotheses—that the average pH-level for springs is 7—at the 1% level, having not found significant evidence to refute the null hypothesis.

   (e) (iii)

5. (a) Our sample statistic/point estimate is $\hat{p} = 903/1714 \doteq 0.5268$. . To that we will add and subtract 1.96 times the standard error:

$$0.5268 \pm (1.96)(0.012), \qquad \text{or} \qquad [0.5033, 0.5503].$$

   (b) Since 0.5 is not inside the 95% CI of part (a), $[.5033, 0.5503]$, it follows that, using the same data set, a test of hypotheses would result in a P-value smaller than 0.05.

   (c) It would be less wide.

   (d) (iii)