Answers

1. (b) only

2. (a) IV (b) III (c) I (d) II

3. Deduce the correct answers from how yours were marked/scored.

4. (i) <u>0.1841</u>
   (ii) <u>0.2743</u>
   (iii) <u>0.4602</u>
   (iv) <u>0.7257</u>

5. (a) . . . , choose a sample of size at least 30.
   (b) . . . , you should quadruple the sample size (i.e., obtain a sample which is 4 times as large).

6. Only (iii) gives a correct interpretation.

7. (a) One might imagine a consumer advocate is out to detect instances when a company's product doesn't live up to its claims. This is why I choose a one-sided alternative hypothesis:

$$\mathbf{H}_0: \quad \mu = 26, \qquad \mathbf{H}_a: \quad \mu < 26.$$

   (b) One does not prove the truth of null hypotheses. We fail to reject the null hypothesis, for the reason that test statistics as extreme as the one from our sample are not extremely rare in a world where the null hypothesis is true. In other words, there is no inconsistency between getting this test statistic and the truth of the null hypothesis. But there are plenty of other null values for μ which are also consistent with our data.
   (c) Whatever test statistic was produced from a random sample of Ford F-150 trucks, it is not a rare one in a world where $\mu = 26$. In fact, test statistics as extreme or more so have a probability of arising 31% of the time.

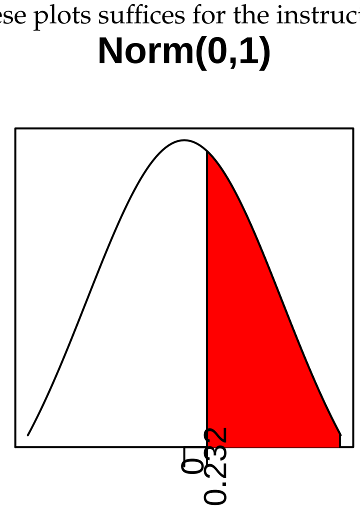
8. (a) For p, the proportion of U.S. adults who would declare themselves to be very interested in environmental issues, we have hypotheses

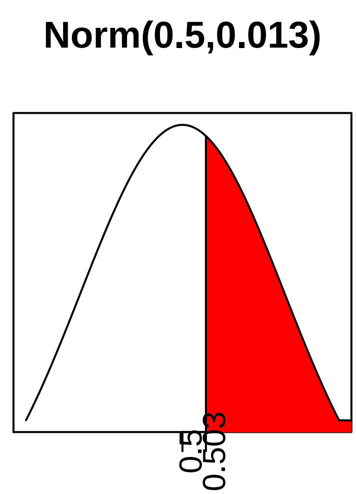
$$\mathbf{H}_0: \ p = 0.5, \qquad \mathbf{H}_a: \ p > 0.5.$$

   (b) Here, $\hat{p} = 751/1493 \doteq 0.503$, and the standard error (the one we would have in a world where the null hypothesis is true) is SE= $\sqrt{(.5)^2/1493} \doteq 0.0129$. Thus, our standardized z-statistic (since we are dealing with proportions, not means) is

$$z = \frac{0.503 - 0.5}{0.129} \doteq 0.232.$$

   (c) Either of these plots suffices for the instructions:





9. (a) Does the new drug cause a greater reduction in blood pressure than the current "best" drug?

   (b) The explanatory variable, *treatment group*, is categorical (values are "old drug" or "new drug"). The response variable, *amount blood pressure is reduced*, is quantitative.

   (c) There is the mean amount that patients see their blood pressure reduced by under the "old" (current best) drug, and the mean amount of reduction under the "new" drug. These can be labeled μ_1 and μ_2, or μ_N and μ_O. Using the latter, the hypotheses are $\mathbf{H}_0: \mu_N - \mu_O = 0$, $\mathbf{H}_a: \mu_N - \mu_O > 0$.

   (d) We might denote by $\overline{x_N}$ and $\overline{x_O}$ the average change in blood pressure in the sample of those taking new and old drug, respectively. The natural point estimate (sample statistic) for our test is the difference of these two: $\overline{x_N} - \overline{x_O}$ (though they may be subtracted the other way, too).

   (e) One way to produce a randomization statistic is to write the response value (difference in blood pressure) of each patient on a slip of paper and place these 46 slips in a bag, mixing them up. Draw out these slips (*without replacement*), treating the first 23 numbers as if they come from recipients of the new drug (the average of these numbers is $\overline{x_N}$), and the last 23 as recipients of the old drug (their mean is $\overline{x_O}$). The difference $\overline{x_N} - \overline{x_O}$ is the randomization statistic.

   (f) In our setting, a Type I error means the two drugs actually produce the same average reduction in blood pressure, but our sample produced data that was *statistically significant* leading us to conclude the new drug is better.

   (g) The statement is **false**. Reducing the significance level α always makes it harder to reject the null hypothesis, thereby decreasing the chances of a Type I error (i.e., less likely we might convict an innocent defendant). At the same time, this increases the chances of a Type II error (that is, it becomes more likely we fail to convice a guilty defendant).

   (h) This *P*-value is not less than the required significance level 0.01, and so we fail to reject the null hypothesis. The researchers do not have compelling evidence the new drug is an improvement over the old one.

10. (a) From the "Descriptive Statistics for One Quantitative Variable" section of StatKey, one can enter the nine differences from the "difference: impaired - unimpaired" row and learn that the sample mean and standard deviation are $\bar{x}_{\text{Diff}} = 0.963$, $s = 0.358$.

   (b) We have hypotheses

$$\mathbf{H}_0: \mu_{\text{Diff}} = 0, \qquad \mathbf{H}_a: \mu_{\text{Diff}} > 0.$$

   The standardized t test statistic is

$$t = \frac{\bar{x}_{\text{Diff}} - 0}{s/\sqrt{9}} = \frac{0.963}{0.358/3} = 8.07.$$

   Using StatKey's t-distribution calculator with df $= 8$, the amount of right-tailed area beyond $t = 8.07$ is small. Specifically, the *P*-value is approximately 0.000021. Thus, we reject the null hypothesis at the $\alpha = 0.05$ level, and conclude there is a positive mean difference in stoppage time wearing the goggles vs. not wearing them.

11. (a) The approximate standard error is

$$\text{SE}_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{6.8^2}{32} + \frac{5.7^2}{45}} \doteq 1.472.$$

   (b) Our point estimate is $\bar{x}_1 - \bar{x}_2 = 1.1$. We will use critical value $t^* = 2.04$, which comes from the t-distribution with 31 degrees of freedom. Thus, our 95% CI is

$$1.1 \pm (2.04)(1.472), \qquad \text{or} \qquad (-1.90, 4.10).$$

   (c) Again employing the t-distribution calculator with 31 degrees of freedom, we have that 99% of values lie between (-2.744) and 2.744. So, the rejection region is $t < -2.744$ (or $(-\infty, -2.744)$) coupled with $t > 2.744$.