Answers (Problem set 9)

Online S520

1. *In a May 2016 Gallup poll, 61% of a sample of 1025 U.S. adults supported same-sex marriage.*

   (a) *Treating the data as a simple random sample, find a 95% confidence interval for the percentage of all U.S. adults who support same-sex marriage.*

   ```
   > .61 - qnorm(.975) * sqrt(.61*.39) / sqrt(1025)
   [1] 0.5801404
   > .61 + qnorm(.975) * sqrt(.61*.39) / sqrt(1025)
   [1] 0.6398596
   ```

   A 95% confidence interval for the percentage of all U.S. adults who supported same-sex marriage at the time of the survey is 58% to 64%.

   (b) *Suppose we wanted to have a 95% confidence interval for the percentage of all U.S. adults who support same-sex marriage with total length 2% (i.e. 0.02.) How large a simple random sample would we need?*

   We set the total length of the confidence interval to 0.02:

   $$2q\sqrt{\frac{p(1-p)}{n}} = 0.02$$

   Rearranging to make n the subject gives

   $$n = p(1-p)\left(\frac{2q}{0.02}\right)^2$$

   q is `qnorm(0.975)` ≈ 1.96.

   For p, we can be safe and use $p = 0.5$ or we can be risky and use $p = 0.61$.

   The risky choice gives $n = 9139$ and the safe choice gives $n = 9604$, so there's not too much difference in the scheme of things.

   The required sample size is over 9000, which is almost certainly too big for a phone survey but might be achievable with an online sample (but questions abound as to whether an online sample really represents the population.)

2. *Trosset chapter 10 Problem Set A, question 1.*

   (a) That we have s but not σ points us toward the t-test, although the sample size is large enough that the standard normal test won't be much different.

   $$t = \frac{3.194887 - 0}{\sqrt{104.0118/400}} = 6.26.$$

   (b) Either (iv) or (iii): (iv) is best, but (iii) gives the same answer to three sig figs.

   (c) The P-value is less than α, so true.

3. Glycemic index

   (a) *What is the experimental unit? What measurements are taken on the experimental units? Is this a problem with one or two independent samples?*

   Experimental unit: Subject with diabetes. Measurements: Glycemic index for dates without coffee and glycemic index for dates with coffee.

   (b) *Give null and alternative hypotheses for an appropriate two-tailed t-test, and calculate the t-statistic.*

   Let μ be the average difference between GI of dates without coffee minus GI with coffee. $H_0 : \mu = 0, H_1 : \mu \neq 0$.

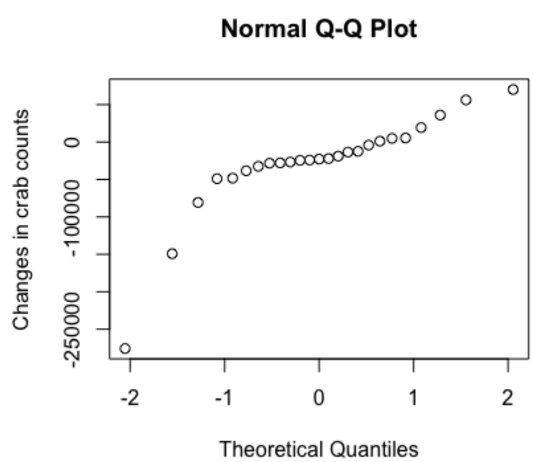
$$T = \frac{11.5}{\frac{21}{\sqrt{10}}} = 1.73.$$

   (c) *The P-value (significance probability) was calculated to be 0.12, so the null hypothesis was not rejected. From this and the other information given, is it correct to conclude that we are sure that on average, dates have the same glycemic index with or without coffee? Explain.*

   Nope. The sample size is much too small to conclude there's no or negligible difference. In fact, the observed mean difference was 11.5: if that's accurate, that seems like a pretty big difference.

4. Counting crustaceans

```
crabs = read.table("crab-counts.txt", header = TRUE)
qqnorm(crabs$CHANGE, ylab="Changes in crab counts")
```



The major problem is with the normality assumption. From the normal quantile plots (or just about any other plot,) the year-to-year changes are not normally distributed. The t-test may thus be misleading with a relatively small sample — especially with the presence of near-outliers (e.g. Big Stone).

5. *Trosset chapter 10 Problem Set D, questions 1 & 2*

(a)
```
ratios = c(.693, .662, .690, .606, .570, .749, .672, .628, .609, .844,
.654, .615, .668, .601, .576, .670, .606, .611, .553, .933)
qqnorm(ratios)
qqnorm(log(ratios))
```

Frankly, neither QQ plot looks like a straight line. Which scale to use is a matter of judgment — the logged data gives a plot slightly closer to a straight line, but transformations hinder interpretation. If I have to choose one from these two, I would go with the logged data, but reasonable statisticians might differ.

(b) Suppose we're working on the log scale. Let $\mu_0 = \log(2/(1+\sqrt{5})) = 0.618034$. We test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_1$.

```
golden.ratio = 2 / (1 + sqrt(5))
t.stat = (mean(log(ratios)) - log(golden.ratio)) / (sd(log(ratios))/sqrt(length(ratios
2 * (1 - pt(abs(t.stat), df = length(ratios) - 1))
```

The P-value is 0.058.

If we use a strict significance level of 0.05, then we do not reject the null hypothesis. (If you did the test on the original scale, you get a P-value of 0.054, and the same conclusion.) More holistically, we could say that there's some evidence of inconsistency with the golden ratio hypothesis — more data would help resolve the issue.