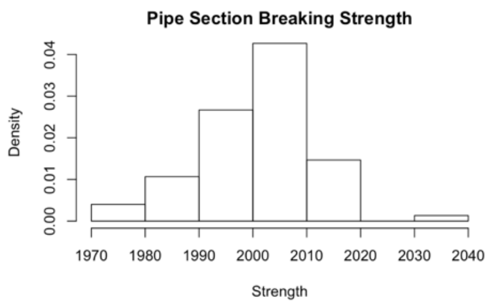
Discussion 8 Soln

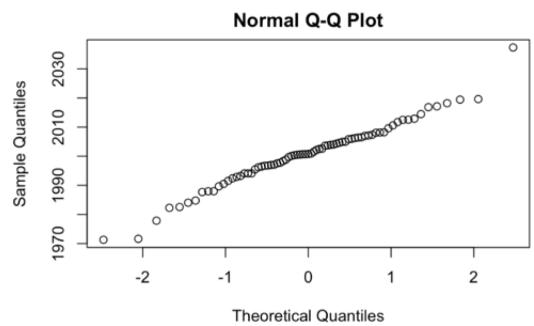
1. Specifications for a water pipe call for a mean breaking strength μ of more than 2000 lb per linear foot. Engineers will perform a test to decide whether or not to use a certain kind of pipe. A random sample of 1 ft sections of pipe is selected and their breaking strengths are measured. The pipe will not be used unless the engineers can conclude (statistically, not with certainty) that the mean breaking strength is greater than 2000.

   (a) Specify appropriate null and alternative hypotheses for this situation.

   ANSWER: $H_o : \mu = 2000$, $H_a : \mu > 2000$

   (b) Based on last week's analsysis, the engineers chose to obtain a sample of 75 random 1 foot pipe sections. The qqplot of the sample data and summary statistics from the sample are given below. Perform a one sample t-hypothesis test at the 5% level after checking that the assumptions for testing are well met and interpret the results in context.





   Pipe Strength Sample Mean: 2001.98; Pipe Strength Sample Standard Deviation: 11.281
   ANSWER: We are assuming these 75 samples are randomly and independently chosen (ie, we aren't only testing from 1 long pipe 75 times). From the qqplot and histogram of the sample data, we do not have strong evidence that the sample came from a non-normal population. Addiitionally, with such a large sample size, we are confident the CLT has kicked in and the distribution of sample means would be normal, even if the population data is not. T test: $T = \frac{2001.98-2000}{11.281/\sqrt{75}} = 1.520016$ *and* $0.05 < P(T_{74} > 1.52) < .10$ *from the table or* $P(T_{74} > 1.52) = 0.06638621$ *from R. Our p value gives us weak evidence against the null and insufficient evidence at the* 5% *level to reject the null. We have insufficient evidence to suggest the mean pipe strength in the population is over 2000 lbs.*

   (c) Another scientist in the lab suggests instead of a t test, a z test could be performed. Explain why either a t or z test will give nearly equivalent p values and conclusions in this case.

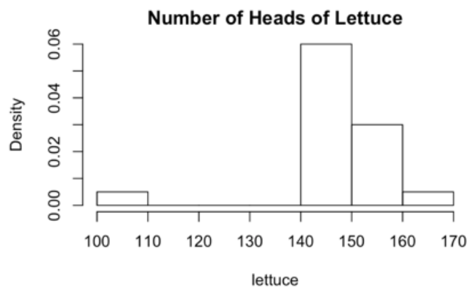
   ANSWER: Since our sample size is so large (75), we are confident that the sample standard deviation is a good approximation for the population standard deviation. Additionally, if we find $P(Z > 1.52) = 0.06425549$

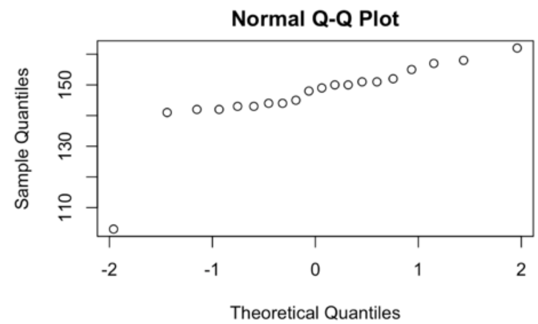
2. A crop scientist evaluating lettuce yields plants 20 plots, treats them with a new fertilizer, lets the lettuce grow, and then measures yield in numbers of heads per plot, with these results:

   145, 142, 144, 141, 142, 155, 143, 157, 152, 143, 103, 151, 150, 148, 150, 162, 149, 158, 144, 151

   The old fertilizer led to an average yield of 145 heads per plot. Test whether the new fertilizer leads to an improved yield via the following steps.

   (a) A histogram and QQ plot of the data is given below. Is the assumption of Normal population appropriate? Will the CLT save us? Why or Why not? Which test should we use?





   ANSWER: 20 plots is pretty big for the CLT, but we are also nervous about the large outlier - qqplot is pretty non-normal

   (b) Run a bootstrap test with set.seed(1).

      i. What assumptions are you making? *Independent observations*

      ii. Which hypotheses are you testing? $H_o : \mu = 145, H_a : \mu > 145$

      iii. What is your observed test statistic? $t_{obs} = \frac{146.5-145}{11.83438/\sqrt{20}} = 0.5668403$

      iv. Find a p-value with the help of 6000 bootstrap replicates. *p=2173/6000=0.3621667*

      v. Draw a conclusion at significance level $\alpha = 0.05$ *Fail to reject the null since our p value is greater than 0.05*

   (c) Perform a t test after stating the assumptions you are using. Compare the p-value and conclusions you draw from the t test to those with the bootstrap.

   Answer: We have to assume that sample size is large enough that CLT tells us $\bar{X} \sim N(\mu_0 = 145, \sigma/\sqrt{20})$ *since QQ plot gives some evidence population is non normal.* $t_{obs} = 0.5668403$ *from above.* $P(T_19 > 0.5668403) = 0.2887292$ *in R or* $p - value > 0.25$ *from table. Also have weak evidence against null. T test is giving stronger evidence against the null.*

   (d) Test whether the lettuce data are compatible with a population median of 145, or rather are strong evidence of a median greater than 145 with a sign test. What assumptions are we making

   Answer: A sign test: 145, 142, 144, 141, 142, 155, 143, 157, 152, 143, 103, 151, 150, 148, 150, 162, 149, 158, 144, 151, corresponds to 0, -, -, -, -, +, -, +, +, -, -, +, +, +, +, +, +, +, -, +, which is 1:0, 8:-, 11:+. To reject $H_o : M = 145$ *for* $H_a : M > 145$*, we'd expect a high number of +. So let B=number of postives (values above mean), drop the zero.* $B \sim Bin(19, .5)$ *and* $B_{obs} = 11$ *so* $pvalue = P(B \geq 11) = 0.3238029$*. No evidence against the null; insufficient evidence at* $\alpha = 0.05$ *Weak evidence that median is above 145. Assumption is just that observations are independent - seems reasonable. Note that the sign test only applies to the population median, not the mean, because we can't say that observations are equally likely to fall above or below the mean without assuming observations come from a symmetric distribution.*

3. A state's Division of Motor Vehicles claims that 60% of teens pass their driving test on the first attempt. An investigative reporter examines an SRS of the DMV records for 125 teens; 86 of them passed on their first try.

   (a) Is there convincing evidence at the $\alpha = 0.05$ significance level that the DMV's claim is incorrect? Do an appropriate hypothesis test.

   ANswer: Let π *be the true proportion of teens who pass their driver's test on the first attempt.* $H_o : \pi = .60, H_A : \pi \neq .60$. *Check normality approximation is ok:* $.60 * 125 > 5$ *and* $.4 * 125 > 5$. *Under* H_o *then,* $\hat{p} \approx N(.60, \sqrt{\frac{.60*.40}{125}}) = N(.60, 0.0438178)$. *The observed test statistic:* $\hat{p} = \frac{86}{125} = 0.688$*So pvalue=*$2 * P(Z > \frac{.688-.60}{0.0438178}) = 2 * P(Z > 2.008316) = 2 * 0.02230486 = 0.0445$. *We have moderate evidence against the null and sufficient at* $\alpha = 0.05$ *to reject the null. Evidence sugests pass rate may be different from advertised* 60%.

   (b) Construct and interpret a 95% confidence interval for the proportion of all teens in the state who passed their driving test on the first attempt. Explain what the interval tells you about the DMV's claim.

   Answer: $\hat{p} \pm z_.025\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .688 \pm 1.96\sqrt{\frac{.688(1-.688)}{125}} = .688 \pm 1.96 * 0.04143969 = (0.607, 0.769)$. *Interval would have also given us evidence to reject since 60% does not fall into the interval.*

   (c) Suppose instead, the reporter gets their information by interviewing a random sample of teens who had taken their driver's test that year. How, if at all, might the test and confidence interval differ from that calculated in a and b?

   Answer: If students inflated their pass rate, interval may have higher center as fewer teens may offer the information that they didn't pass on the first time. Also, the higher the observed $\hat{p}$*, farther away from .6, the p value will be smaller, and we will have stronger evidence against the null.*