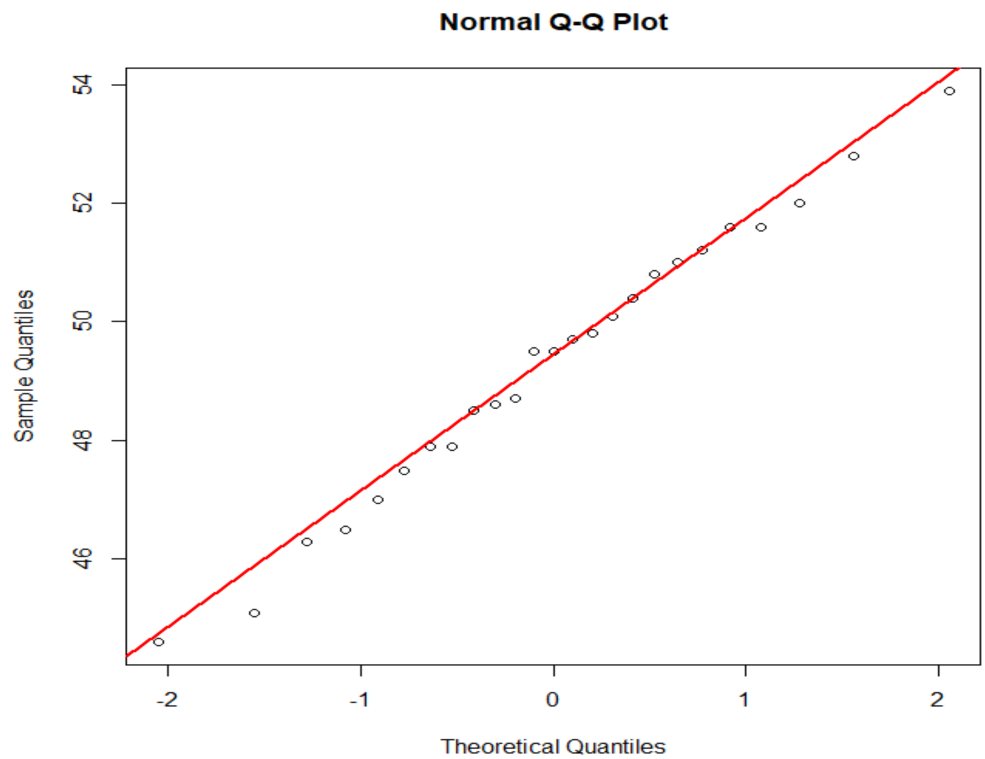
# STAT 231 Tutorial Solutions – Friday, June 28

**1.** A hospital lab has just purchased a new instrument for measuring levels of dioxin (in parts per billion). In order to calibrate the instrument, 25 samples of a "standard" water solution, known to contain 50 parts per billion are measured by this new instrument. The observed data, $y_1, y_2, \ldots, y_{25}$ are given below:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 45.1 | 51.0 | 51.6 | 46.3 | 49.8 | 52.8 | 49.5 | 50.1 | 47.9 | 49.5 | 47.5 | 46.5 | 44.6 |
| 47.0 | 50.8 | 53.9 | 51.6 | 47.9 | 52.0 | 48.7 | 50.4 | 49.7 | 51.2 | 48.5 | | |

Summary information: $\sum_{i=1}^{25} y_i = 1232.5$

It is assumed that the measurements are Normally distributed. Here is a QQ Plot of the sample data:




a. Comment on the assumption of Normality.

**Looking at the above QQ plot, the assumption of Normality looks valid.**

Let $Y_i$ represent the $i^{th}$ measurement by this device.

It seems reasonable to assume $Y_i \sim G(\mu, \sigma)$, i = 1,2,…, 25; independent. In this question, we will assume that $\sigma$ is known to be equal to 2.

b. Using the given information, construct a 90% confidence interval for $\mu$.

**With Normal observations, and $\sigma$ known, a 90% confidence interval for $\mu$ is given by:**

$\boldsymbol{\bar{y} \pm z_{0.95} * \frac{\sigma}{\sqrt{n}}}$ **=>** $\boldsymbol{\frac{1232.5}{25} \pm 1.645 * \frac{2}{\sqrt{25}}}$ **=> 49.3 $\pm$0.658**

**The 90% CI for $\mu$ is [48.642, 49.958]**

**So, with 90% confidence, we believe that the true value of $\mu$ is between 48.642 and 49.958.**

c. Now, suppose we want to test $H_o$: $\mu = 50$ (versus a two-sided alternative),

$H_A$: $\mu \neq 50$). We start by determining the appropriate discrepancy measure / test statistic, D, and state the distribution of D, assuming that $H_o$ is true.

**We have a random sample of independent observations drawn from a Normal population with a known standard deviation. The discrepancy measure / test statistic is:**

$\boldsymbol{D = \frac{|\bar{Y}-\mu_o|}{\sigma/\sqrt{n}}}$**. In this case, D ~ G(0, 1), assuming Ho is true.**

d. Now calculate the observed value of the test statistic, d, and use it to calculate the appropriate p-value.

**We calculated the sample mean already, when we constructed the 90% CI.**

**The sample mean is,** $\boldsymbol{\bar{y} = \frac{1232.5}{25} = 49.3}$**. Plugging in the other values, the observed value of our discrepancy measure / test statistic is:**

$$\boldsymbol{d = \frac{|\bar{y}-\mu_o|}{\sigma/\sqrt{n}} \Rightarrow d = \frac{|49.3-50|}{2/\sqrt{25}} = 1.75}$$

**We use the observed value of the test statistic to determine the p-value. In this case, we need a two-sided p-value, as this is a two-sided test (i.e. $H_A$ is an inequality).**

**So, the appropriate p-value = P(|D|≥ 1.75) (where D ~ G(0, 1), assuming $H_o$ is true)**

**The two-sided p-value = 2*(1 – P(Z ≤ 1.75)) = 2*(1 – 0.95994) = 0.08012**

e. Use the Table 5.1 guidelines to assess the amount of evidence against $H_o$.

**With a p-value = 0.08012, according to Table 5.1 guidelines, we have weak evidence against $H_o$.**

f. Using a threshold value, $\alpha$ = 10%, and the p-value from part e., decide whether to reject or not reject $H_o$. Justify your answer.

**Since the p-value is less than $\alpha$ = 10%, we must reject $H_o$.**

g. Now, let's revisit the 90% CI for $\mu$: [48.642, 49.958]. Based on this CI, do your answers make sense in parts e. and f.?

**Yes. Everything makes sense. It turns out that a CI is another method to test hypotheses. With a 90% confidence level, there is an implied level of significance of 10% (the complement of our confidence level).**

**Remember that we rejected $H_o$ in part f. Our 90% CI does not include the null value of 50. This makes sense. The data doesn't support $H_o$ at the 10% level of significance. We see this with the p-value and the CI.**

h. Suppose that we had conducted our hypothesis test at a 5% level of significance. Would our decision in part f. change?

**Yes! At the 5% level of significance, we CANNOT REJECT Ho, as our p-value (0.08012) is now greater than 5%. So, you see, if you change the threshold / level of significance, your decision could change.**

i. **Apply the Concepts:** Would you expect the corresponding 95% CI for $\mu$ to include the null value of 50? Explain.

**Yes! You can go ahead and construct the 95% CI to verify, if you like. We know that a 95% CI would be wider, so it COULD include 50. However, we also know the p-value, so we can answer this question.**

**With a p-value greater than 5%, this implies that the null value of 50 would lie inside the corresponding 95% CI. In case you aren't convinced.....**

**Check: 95% CI for $\mu$ is: $\bar{y} \pm z_{0.975} * \frac{\sigma}{\sqrt{n}}$ => $\frac{1232.5}{25} \pm 1.96 * \frac{2}{\sqrt{25}}$ => 49.3 $\pm$0.784**

**The 95% CI for $\mu$ is [48.516, 50.084]. The 95% CI DOES include 50, as expected.**