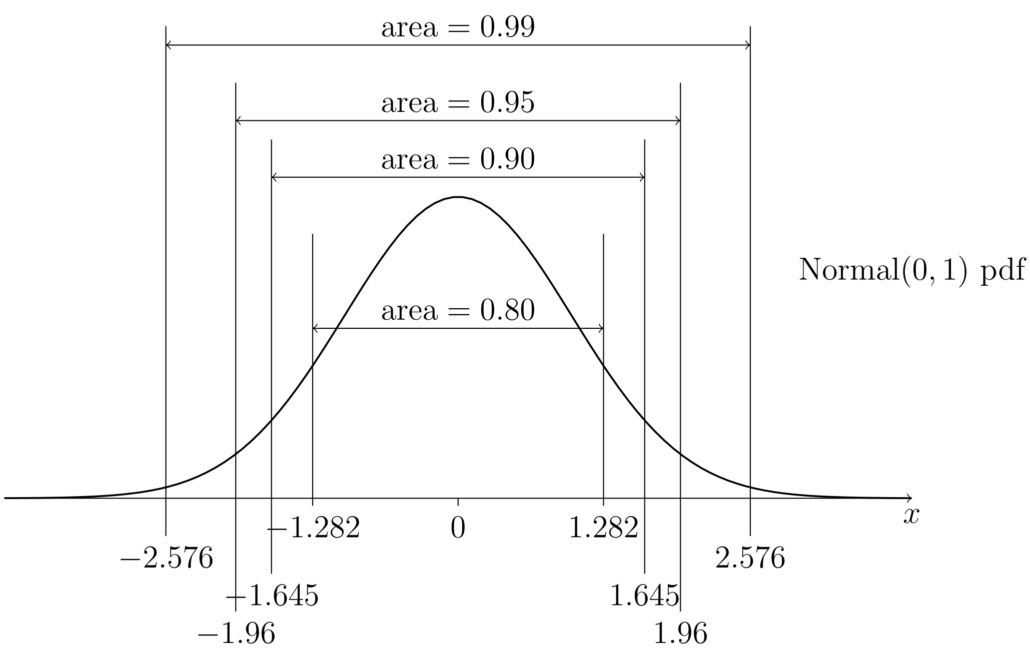
STAT 515 fa 2021 Exam II

Karl B. Gregory

- Do not open this exam until told to do so.
- You may have two handwritten sheet of notes out during the exam.
- You have 75 minutes to work on this exam.
- You may use a simple calculator.
- If you are unsure of what a question is asking for, do not hesitate to ask me for clarification.
- *Good luck, and may the odds be ever in your favor!*



$$\hat{p}_n \pm z_{\alpha/2} \cdot \sqrt{\hat{p}_n(1-\hat{p}_n)/n} \qquad \bar{X}_n \pm t_{n-1,\alpha/2} \cdot S_n/\sqrt{n}$$

$$Z_{\text{test}} = \frac{\hat{p}_n - p_0}{\sqrt{p_0(1-p_0)/n}} \qquad T_{\text{test}} = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}$$

1. Contestants at a new carnival game win a prize if they toss a ring so that it lands around the neck of any of several rubber ducks floating in a whirlpool. The owner has set the cost of playing the game assuming no more than 3% of contestants will succeed. She will increase the cost if she finds that contestants succeed more often than this. In the first days of the carnival, she observes 729 contestants, of which 29 win. Then she comes to you for advice...

   (a) State the null and alternate hypotheses of interest to the owner of the game.

   (b) Give the value of the test statistic for testing the hypotheses in part (a).

   (c) Do you reject the null hypothesis at the $\alpha = 0.05$ significance level?

   (d) Interpret your statistical inference by giving a recommendation to the owner of the game.

   (e) Give an interval in which the p-value lies. Be as precise as you can.

   (f) Construct a 95% confidence interval for the proportion of contestants who win the game.

(g) If the owner is interested in the hypotheses

$$H_0\colon\ p = 0.03 \text{ versus } H_1\colon\ p \neq 0.03,$$

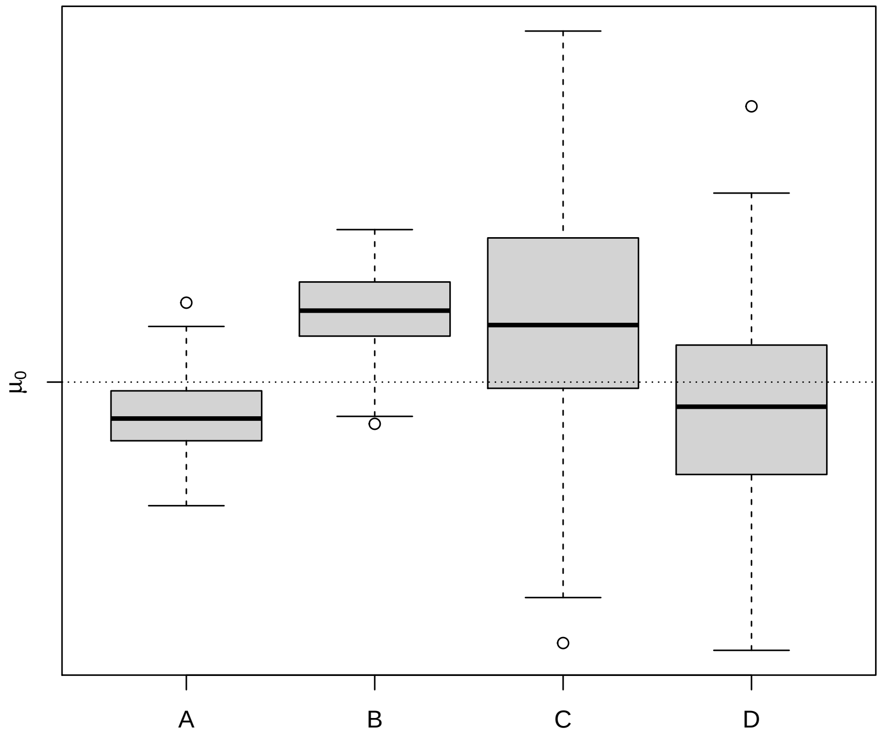
does she reject H_0 at $\alpha = 0.05$?

(h) If the owner wishes to estimate the true proportion of contestants who win the game with a margin of error no greater than 1%, with 95% confidence, how many contestants should she observe? Use the data collected in the first days of the carnival as a best guess for p.

2. Each of three scientists is interested in a set of hypotheses about a mean μ:

    1. Alain is interested in testing H_0: $\mu \leq \mu_0$ versus H_1: $\mu > \mu_0$.
    2. Muriel is interested in testing H_0: $\mu \geq \mu_0$ versus H_1: $\mu < \mu_0$.
    3. Étienne is interested in testing H_0: $\mu = \mu_0$ versus H_1: $\mu \neq \mu_0$.

    Consider these boxplots of data sets A, B, C, and D that the scientists might observe:



    Fill in each blank with A, B, C, or D.

    (a) For Alain, data set ____________ would give the largest p-value, while data set ____________ would give the smallest p-value.

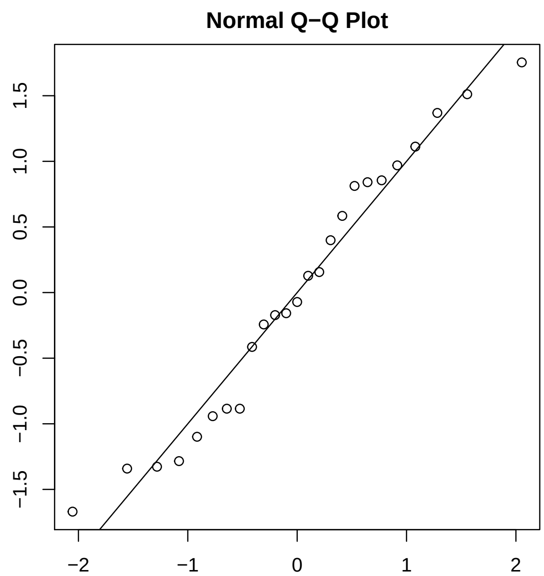
    (b) For Muriel, data set ____________ would give the largest p-value, while data set ____________ would give the smallest p-value.

    (c) For Étienne, data set ____________ would give the largest p-value, while data set ____________ would give the smallest p-value.

3. Each of a large number of barbunya beans was photographed by a computer and its size measured in the number of pixels covered by the bean [1]. Here is a sample of 25 of these measurements, in thousands of pixels, rounded to the nearest decimal place:

71.5	69.5	76.5	67.7	70.1	73.4	76.3	64.4	71.7	62.9	58.9	80.2	
81.2	78.4	68.9	76.6	61.3	82.9	69.4	64.4	61.2	64.0	61.6	74.7	77.4

Suppose it is of interest to test whether the mean pixel area of barbunya beans is less than sixty-seven thousand pixels based on these 25 measurements. The sample mean and standard deviation for these data are $\bar{X}_n = 70.60$ and $S_n = 7.01$. A Normal quantile-quantile plot of the standardized values is here:



(a) State the null and alternate hypotheses of interest.

(b) Explain why it is important to look at the Normal quantile-quantile plot.

(c) Give the value of the test statistic for testing the hypotheses in part (a).

(d) Do you reject the null hypothesis at significance level $\alpha = 0.05$?

(e) Give an interval in which the p-value lies. Be as precise as you can.

(f) Construct a 95% confidence interval for the mean pixel area of barbunya beans.

(g) Based on these data, would you reject the null hypothesis when testing

$$H_0\colon\ \mu = 67 \text{ versus } H_1\colon\ \mu \neq 67$$

at the $\alpha = 0.05$ significance level?

References

[1] Koklu M. and Ozkan I.A. Multiclass classification of dry beans using computer vision and machine learning techniques. *Computers and Electronics in Agriculture*, 174(105507), 2020.