Name:______________________________________ **Date:**____________________

PSTAT 5A: Discussion Worksheet 06

Spring 2023, with *Ethan P. Marzban*

1. (a) If $X \sim \mathcal{N}(-3,\ 1.5)$, what is $\mathbb{P}(X \leq -2)$?

Solution: We must first standardize:

$$\mathbb{P}(X \leq -2) = \mathbb{P}\left(\frac{X - (-3)}{1.5} \leq \frac{-2 - (-3)}{1.5} \right) = \mathbb{P}(Z \leq 0.67)$$

From a normal table, we see that the desired probability is $0.7486 = 74.86\%$

(b) Find $\pi_{0.77}$, the value such that $\mathbb{P}(Z \leq \pi_{0.77}) = 0.77$ where $Z \sim \mathcal{N}(0,\ 1)$.

Solution: Note that this problem is effectively asking for the 77$^{\text{th}}$ percentile of the standard normal distribution. From a normal table we see that $\mathbb{P}(Z \leq 0.74) = 0.7704$, meaning we can approximately take $\pi_{0.77} \approx 0.74$.

(c) Find the 21.77$^{\text{th}}$ percentile of the standard normal distribution.

Solution: Note that this problem is effectively asking for the 77$^{\text{th}}$ percentile of the standard normal distribution. From a normal table we see that $\mathbb{P}(Z \leq -0.78) = 0.2177$, meaning the desired result is approximately -0.78.

(d) Find the 10$^{\text{th}}$ percentile of the t_{21} distribution.

Solution: We consult the $t-$table. The 10$^{\text{th}}$ percentile will have a left-tail area of 0.100, meaning we look in the column with a one-tailed value of 0.100. To find the appropriate row, we note that t_{21} denotes a $t-$distribution with 21 degrees of freedom; hence, we find the desired value to be 1.32.

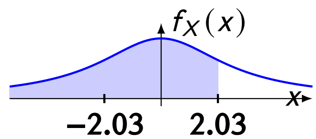
(e) If $T \sim t_{37}$, what is $\mathbb{P}(T \leq -2.03)$?

Name:________________________________ Date:________________

Solution: We again consult the t−table. Note that the value of 2.03 in the row with df $=$ 37 corresponds to a one-tailed probability of 0.025. This means that $\mathbb{P}(T \leq -2.03) =$ 0.025 = 2.5%.

(f) If $T \sim t_{37}$, what is $\mathbb{P}(T \leq 2.03)$? **Hint:** Draw a picture!

Solution: Following the hint, we draw a picture. This is the area we seek:



The unshaded area, by the symmetry of the t distribution as well as our answer to part (e) above, is 0.025; therefore, by the complement rule, the desired probability is $1 - 0.025 =$ 0.975 = 97.5%.

2. Mark is interested in performing inference on the true proportion of UCSB students that use *Venmo*. As such, he takes a representative sample of 92 UCSB students and finds that 57% of these students use *Venmo*.

(a) Identify the population.

Solution: The population is the set of all UCSB students.

(b) Identify the sample.

Solution: The sample is the set of 92 UCSB students included in the representative sample.

(c) Define the parameter of interest.

Solution: The parameter of interest is p = the true proportion of UCSB students that use *Venmo*.

(d) Define the random variable of interest.

Name:____________________________________ **Date:**__________________

Solution: The random variable of interest is $\widehat{P}$ = the proportion of UCSB students, in a representative sample of 92, that use *Venmo*.

(e) Construct a 95% confidence interval for the true proportion of UCSB students that use *Venmo*. Be sure to check any/all relevant conditions, and interpret your interval.

Solution: We first check the success-failure conditions. Since we don't have access to the value of p, we use the substitution approximation to check:

1) $n\widehat{p} = (92)(0.57) = 52.544 \geq 10 \checkmark$

2) $n(1 - \widehat{p}) = (92)(1 - 0.57) = 39.56 \geq 10 \checkmark$

Since the substitution approximation to the success-failure conditions are met, we can invoke the Central Limit Theorem for Proportions to say

$$\widehat{P} \sim \mathcal{N}\left(p, \ \sqrt{\frac{p(1-p)}{n}}\right)$$

As such, our confidence interval will take the form

$$\widehat{p} \pm z_\alpha \cdot \sqrt{\frac{p(1-p)}{n}}$$

But, since we don't have access to p, we will instead use

$$\widehat{p} \pm z_\alpha \cdot \sqrt{\frac{\widehat{p} \cdot (1-\widehat{p})}{n}}$$

and since we are using a 95% confidence level we take $z_\alpha = 1.96$ to compute

$$(0.57) \pm 1.96\sqrt{\frac{(0.57)\cdot(1-0.57)}{92}} = (0.57) \pm 0.101 = [0.469 \, , \, 0.671]$$

The interpretation of this interval is: we are 95% confident that the true proportion of UCSB students that use *Venmo* is between 46.9% and 67.1%.

(f) Construct a 77% confidence interval for the true proportion of UCSB students that use *Venmo*.

Solution: We again use

$$\widehat{p} \pm z_\alpha \cdot \sqrt{\frac{\widehat{p} \cdot (1 - \widehat{p})}{n}}$$

however now, since we are using a 77% confidence level, we first compute

$$\frac{1 - 0.77}{2} = 0.115$$

meaning we need to find the 11.5^{th} percentile of the standard normal distribution. Consulting a normal table, we see that the 11.5^{th} percentile of the standard normal distribution is around -1.20, meaning we take $z_\alpha = 1.2$ and

$$(0.57) \pm 1.20\sqrt{\frac{(0.57) \cdot (1 - 0.57)}{92}} = (0.57) \pm 0.062 = [0.508\ ,\ 0.632]$$

3. A quality-control checker takes a representative sample of 35 *GauchoSnip*-brand scissors and finds that the sampled scissors have an average weight of 6 oz and a standard deviation of 1.4 oz.

(a) Define the parameter of interest.

Solution: The parameter of interest is μ = the true average weight of *GauchoSnip*-brand batteries.

(b) Define the random variable of interest.

Solution: The random variable of interest is $\overline{X}$ = the average weight of 35 randomly-selected *GauchoSnip*-brand scissors.

(c) Construct a 95% confidence interval for the true average weight of *GauchoSnip*-brand scissors, and interpret your interval.

Solution: We first need to answer a set of questions:

- Can we assume the population (i.e. the distribution of weights of <u>all</u> *GauchoSnip*-brand scissors) is normally distributed? **No**.
- Is our sample size large enough? **Yes**, since $n = 35 \geq 30$.

- Do we have access to σ, or s? **We have access to s.**

Based on the answers to these questions, we know that we need to use the **t–distribution**. Specifically, our confidence interval will be of the form

$$\overline{x} \pm c \cdot \frac{s}{\sqrt{n}}$$

where c is the 2.5$^{\text{th}}$ percentile of the $t_{n-1} = t_{35-1} = t_{34}$ distribution, scaled by -1. From the t–table we find $c = 2.03$, meaning our confidence interval becomes

$$6 \pm (2.03) \cdot \frac{1.4}{\sqrt{35}} = 6 \pm 0.48 = [5.52\,,\ 6.48]$$

The interpretation of this interval is: we are 95% confident that the true average weight of a *GauchoSnip*-brand pair of scissors is between 5.52 oz and 6.48 oz.