Assignment 6 - Due Monday 23

1. (2 points) Liam is interested in making a one sample z-interval for the proportion of adults who claim they are good at math. In order to estimate p, the proportion of parents who claim to be 'math-heads', what is the smallest sample size possible to obtain a margin of error of ≤ 0.05?

• Assume 95% confidence.

So $Z^* = \text{invNorm}(0.975) = 1.96$

$$\hat{p} \pm \underbrace{Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{0.05}$$

need

$$Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq 0.05$$

• we are not given a estimated value for p, so we will use p= 0.5 which will maximize p(1-p).

so we have:

$$1.96 \sqrt{\frac{0.5(0.5)}{n}} \leq 0.05$$

$$\frac{(1.96)^2 \cdot (0.5)^2}{n} \leq (0.05)^2$$

$$n \geq \frac{(1.96)^2 (0.5)^2}{(0.05)^2} = 384.16$$

so you should sample 385 people.

2. (2 points) Members of an online gaming league play thousands of games over the course of a year. Suppose that scores of individual games have a known standard deviation of $\sigma = 30$ points. Raunak plans on taking a random sample of n games from this population to make a 95% confidence interval for the mean score. He wants the margin of error to be no more than 10 points. What is the smallest approximate sample size required?

$\sigma = 30$

$Z^* = 1.96$ as $\alpha = 0.05$

$= \text{invNorm}(0.975)$

for a confidence interval for μ with Known σ we have.

$$\bar{x} \pm \underbrace{Z^* \frac{\sigma}{\sqrt{n}}}_{10}$$

$$1.96 \cdot \frac{30}{\sqrt{n}} \leq 10$$

$$\frac{(1.96)^2 \cdot (30)^2}{n} \leq 100$$

$$n \geq \frac{(1.96)^2 \cdot (30)^2}{100} = 34.5744$$

So Raunak will need to sample 35 people.

3. (2 points) Peter works at a toy panda factory and would like to estimate the mean weight in grams of the factory's toy pandas. he'll sample n pandas to build a 90% confidence interval for the mean with a margin of error of no more than 15 g. Preliminary data suggests that $\sigma = 60$ is a reasonable estimate for the standard deviation of these weights.

given σ known $\sigma = 60$ so C.I. is given by:

$$\bar{x} \pm \overbrace{Z^* \frac{\sigma}{\sqrt{n}}}^{15g}$$

$$Z^* = \text{invNorm}(0.95) = 1.645$$

$$Z^* \cdot \frac{\sigma}{\sqrt{n}} \leq 15$$

$$1.645 \cdot \frac{60}{\sqrt{n}} \leq 15$$

$$\frac{\sqrt{n}}{1.645 \cdot 60} \geq \frac{1}{15}$$

$$\sqrt{n} \geq \frac{1.645 \cdot 60}{15}$$

$$n \geq \left(\frac{1.645 \cdot 60}{15}\right)^2 = 43.2964$$

So you must sample $\boxed{n = 44}$ toy Pandas.

4. (2 points) Anika wants to use a one-sample z-interval to estimate what proportion of voters in a country plan on voting for a certain candidate. She wants the margin of error to be no more than $\pm 3\%$ at 99% confidence. What is the smallest sample size required to obtain the estimate?

no estimated proportion was proposed, so we will use a conservative value 0.05 which maximizes $p(1-p)$.

with $\alpha = 0.01$, $Z^* = \text{invNorm}(0.995)$
$= 2.576$

we have:

$$\hat{p} \pm \underbrace{Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{0.03}$$

$$2.576 \sqrt{\frac{(0.5)^2}{n}} \leq 0.03$$

$$\frac{(0.5)}{\sqrt{n}} \leq 0.03$$

$$\frac{\sqrt{n}}{(2.576)(0.5)} \geq \frac{1}{(0.03)}$$

$$\sqrt{n} \geq \frac{(2.576)(0.5)}{(0.03)}$$

$$n \geq \left(\frac{(2.576)(0.5)}{(0.03)}\right)^2 = 1843.27$$

So sample at least $\boxed{1844 \text{ people}}$

5. A simple random sample of 34 legendary Pokemon and 28 non-legendary Pokemon have attack means $\overline{x}_1 = 71.4$, $\overline{x_2} = 109$, with $s_1^2 = 935$, and $s_2^2 = 966$. Is there statistical evidence supporting a significant difference between the true mean of legendary and non-legendary Pokemon?

$n_1 = 34$ $n_2 = 28$

$\bar{x}_1 = 71.4$ $\bar{x}_2 = 109$

$\underbrace{s_1^2 = 935}_{\text{Legendary}}$ $\underbrace{s_2^2 = 966}_{\text{Non-Legendary}}$

Appears I put these backwards as legendary Pokemon appear significantly weaker.

Procedure: 2-sample t-test for $\mu_1 - \mu_2$,

$H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

$H_a: \mu_1 - \mu_2 > 0$ or $\mu_1 > \mu_2$

Conditions: We do have a simple random sample, n_1 and n_2 are clearly $< 10\%$ of N_1 and N_2 in this case. we have $n = 34 > 30$ Legendary Pokemon, but only $n = 28 < 30$ Non-legendary Pokemon. Note 28 is NOT large enough for CLT to apply. We will assume attacks are normally distributed for non-legendary Pokemon.

- For this example it appears $S_1 \approx S_2$ so we are using a pooled t-test.

- Using 2-sample T-test on calculator gives:

$t\text{-calc} = -4.7828$ note $df = 60$ $(n_1 + n_2 - 2)$

$p = 0.999\ldots$

Conclusion: Since we have a very large p-value $0.999 > \alpha = 0.05$, there is not sufficient evidence to reject H_0.

In other words there is not evidence to suggest that legendary Pokemon have higher attack values than non-legendary.