Class Exercise - Solutions

1. **Rebel Magpies**

In a 2022 pilot study, scientists fitted 5 Australian magpies with tiny GPS trackers that could only be removed using a magnet or scissors. Within 10 minutes, a dominant female had successfully removed the tracker from a younger bird, and within 3 days all of the devices had been removed.

In a larger follow-up study, a wildlife ecologist fits a random sample of 40 magpies with the same type of tracker and records how long (in hours) each device stays on before being removed by the bird or a flock-mate. The sample yields an average attachment time of $\bar{x} = 20.0$ hours. From previous deployments, the population standard deviation of attachment times is known to be 4.0 hours.

(a) Construct and interpret a 95% confidence interval for the true mean attachment time of the trackers on Australian magpies.

Solution

We have: $n = 40, \quad \overline{x} = 20, \quad \sigma = 4.$

Want 95% CI for $\mu \quad \because \sigma$ is known $\Rightarrow \quad Z - table \quad ; \quad Z_{0.025} = 1.96$

$$\overline{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \qquad \Rightarrow \qquad 20 \pm 1.96 \cdot \frac{4}{\sqrt{40}} \qquad \Rightarrow \qquad 20 \pm 1.2396$$

$$18.7604 < \mu < 21.2396$$

With repeated sampling, we are 95% confident that the average attachment time of the trackers is between 18.7604 and 21.2396 hours.

(b) A technology company that manufactures the trackers claims that the devices stay attached for *at least* 22 hours on average when used with magpies. Based on the 95% confidence interval you constructed, can this claim be validated? Justify your answer.

Solution

No. The company cannot make this claim, since 22 is not in the interval, and we are 95% confident that the average attachment time of the trackers is between 18.7604 and 21.2396 hours.

(c) If the ecologist wants the margin of error for a 95% confidence interval to be no more than 0.75 hours, what is the minimum sample size needed?

Solution

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.96 \cdot 4}{0.75}\right)^2 = 109.2722 \qquad \Rightarrow \qquad n = 110\, magpies$$

2. **Dragonfly Evasion**

Female dragonflies fake sudden death to avoid unwanted male suitors.

To estimate how common this behaviour is in the wild, a researcher examines 180 randomly selected female dragonflies during the peak mating season and finds that 117 of them display death-feigning behaviour when approached by a male.

(a) Construct and interpret a 90% confidence interval for the true proportion of female dragonflies that use death-feigning behaviour to avoid unwanted mates.

Solution

Let x = the number of female dragonflies in the sample that use death-feigning behaviour

We have: $n = 180, \quad x = 117, \quad \hat{p} = \frac{x}{n} = \frac{117}{180} = 0.65$
Want 90% CI for $p \quad \Rightarrow \quad Z - table \quad ; \quad Z_{0.05} = 1.645$

$$\hat{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{\sqrt{n}}} \quad \Rightarrow \quad 0.65 \pm 1.645 \cdot \sqrt{\frac{(0.65) \cdot (1 - 0.65)}{\sqrt{180}}} \quad \Rightarrow \quad 0.65 \pm 0.0585$$

$$0.5915 < p < 0.7085$$

With repeated sampling, we are 90% confident that the actual percentage of female dragon flies that use death-feigning behaviours is between 59.15% and 70.85%

(b) The researcher wants to estimate this true proportion with a 90% confidence level and a margin of error of no more than 4%. A preliminary field study suggests that about 68% of females use death-feigning as an avoidance strategy. What is the minimum sample size that should be used?

Solution

$$n = p \cdot (1 - p) \left(\frac{Z_{\alpha/2}}{E}\right)^2 = (0.68) \cdot (1 - 0.68) \left(\frac{1.645}{0.04}\right)^2 = 368.0194 \quad \Rightarrow \quad n = 369$$

(c) Construct and interpret a 99% lower confidence bound for the true proportion of female dragonflies that engage in death-feigning to avoid male suitors.

Solution

Want 99% LCB for $p \quad \Rightarrow \quad Z-table \quad ; \quad Z_{0.01} = 2.33$

$$\hat{p} - Z_\alpha \cdot \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{\sqrt{n}}} \quad \Rightarrow \quad 0.65 - 2.33 \cdot \sqrt{\frac{(0.65)\cdot(1-0.65)}{\sqrt{180}}} \quad \Rightarrow \quad 0.65 - 0.0828$$

$$0.5672 \leq p$$

With repeated sampling, we are 99% confident that the actual percentage of female dragon flies that use death-feigning behaviours is at least 56.72%

3. **ChatGPT Energy Use**

A single question to ChatGPT uses ten times as much electricity as a Google search and consumes half a litre of water.

At a clean-tech research lab, an engineer collects data on the electricity required (in watt-hours) for each of 12 randomly selected ChatGPT queries under controlled conditions. The sample of 12 queries yields a mean electricity consumption of 8.2 watt-hours per query, with a sample standard deviation of 2.5 watt-hours.

(a) Construct and interpret a 95% confidence interval for the true mean electricity consumption of a ChatGPT query.

Solution

We have: $n = 12, \quad \overline{x} = 8.2, \quad s = 2.5.$
Want 95% CI for $\mu \quad \because \sigma$ is unknown $\quad \Rightarrow \quad t-table \quad ; \quad t_{0.025,11} = 2.201$

$$\overline{x} \pm t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}} \quad \Rightarrow \quad 8.2 \pm 2.201 \cdot \frac{2.5}{\sqrt{12}} \quad \Rightarrow \quad 8.2 \pm 1.5884$$

$$6.6116 < \mu < 9.7884$$

With repeated sampling, we are 95% confident that the true average electricity consumption of a ChatGPT query is between 6.6116 and 9.7884 watt-hours.

(b) A "typical" Google search uses about 0.8 watt-hours of electricity. Using your 95% confidence interval, does the data provide evidence that a ChatGPT query uses more energy on average than a Google search? Explain.

Solution

Yes. Since the entire 95% confidence interval for ChatGPT's average energy use (6.6116 to 9.7884 watt-hours) is far above 0.8 watt-hours, the data provides clear statistical evidence that a ChatGPT query uses more energy on average than a Google search.

(c) Construct and interpret a 95% upper confidence interval for the true mean electricity consumption of a ChatGPT query.

Solution

Want 95% UCB for μ $\quad \because \sigma$ is unknown $\quad \Rightarrow \quad t-table \quad ; \quad t_{0.05,11} = 1.796$

$$\overline{x} + t_{\alpha,n-1} \cdot \frac{s}{\sqrt{n}} \quad \Rightarrow \quad 8.2 + 1.796 \cdot \frac{2.5}{\sqrt{12}} \quad \Rightarrow \quad 8.2 + 1.2962$$

$$\mu \leq 9.4962$$

With repeated sampling, we are 95% confident that the true average electricity consumption of a ChatGPT query is at most 9.4962 watt-hours.