Unit 3 Question Bank

Instructions

Problem set completion is defined as completing one problem from each group.

1. Confidence Intervals for Sample Means
2. Confidence Intervals for Proportions
3. Null Hypothesis Significance Testing
4. Fisherian Hypothesis Testing
5. Neyman-Pearson Hypothesis Testing

Confidence Intervals for Sample Means

1. A biologist is studying Total Nitrogen concentrations (mg/L) in groundwater across agricultural fields. She collects a sample of 40 wells and measures the nitrate concentration in each. She calculates the mean Total Nitrogen at 27.028 mg/L and finds the standard deviation to be 12.885 mg/L. Her boss has suggested she produce confidence intervals on the estimated mean to help affirm the results, but didn't provide guidance on what confidence **level**.

   (a) Calculate the margin of error for a 90%, 95%, and 99% interval on mean Total Nitrogen

   (b) State which interval she should use if the **true** mean and standard deviation are $\mu = 29$ and $\sigma = 9$.

2. A researcher studying forage quality in mixed-grass prairie wants to understand how much dry biomass individual tillers typically produce under mid-season grazing. After clipping a random set of 35 tillers, the researcher dries them and records their masses. The sample averages 87.23 grams, and the variability across samples corresponds to a standard deviation of 21.96 grams. Construct a 99% confidence interval on the mean dry biomass.

3. An agricultural engineering group is testing the reliability of a new seed-metering mechanism designed for precision planters. To evaluate the uniformity of seed delivery, the team runs a controlled bench test and records the spacing distance (in millimeters) between consecutive seeds released by the device. After conducting the trial, the sample of 30 observed spacings has an average distance of 17.45 mm, with a sample standard deviation of 3.1 mm.

   (a) Calculate the point estimate for spacing distance.

   (b) Calculate the t-critical value for a 90% interval.

   (c) Which critical value will be larger, a t value at 95% or a z value at 95%? Justify your answer.

4. A developmental biologist is studying variability in larval body mass within a population of ground beetles raised under identical environmental conditions. After the larvae reach a particular instar, the researcher randomly selects 40 individuals and records each larva's dry mass after a standardized desiccation period. The results are entered properly but the data files are corrupted during analysis, leaving some of it missing.

   Given a sample mean dry mass of 77.76 mg and a 95% interval of $(73.108, 82.411)$, determine the original sample standard deviation.

Confidence Intervals for Proportions

5. A unique striping pattern occurs in about 28.1% of fence lizard populations. A field ecologists takes a survey of local fence lizards and finds that out of 608 adults, 153 were found to have that striping pattern.

   Calculate a 95% confidence interval on the proportion of fence lizards from the survey found to have the striping pattern and compare it with the assumed population proportion.

6. During a germination study on prairie dropseed, researchers track whether each seed successfully produces a viable seedling within a controlled growth chamber. In one trial, 460 seeds were sown, and 189 produced viable seedlings. Construct a 90% confidence interval for the proportion of seeds in this trial that successfully germinated into viable seedlings.

7. An experiment on colony establishment in a soil microbe saw that out of 100 plated cells that were individually tracked under a microscope, 96 succeeded in forming visible colonies.

   (a) Calculate the point estimate for colony formation rate.

   (b) Explain why a confidence interval can't be constructed for this point estimate.

8. Raise the sample size in the previous question to 110 without changing the number of successful colony formations.

   (a) Calculate the point estimate and margin of error.

   (b) Assume that the true proportion is 27%. What is a possible reason why the point estimate is so different from the truth?

Null Hypothesis Significance Testing

9. Using the data from question 4, test at the $\alpha = 5\%$ level whether or not the larva being studied have a different mean dry mass from the population mean of $\mu = 81$ mg. State the relevant hypotheses, p-value, and decision.

10. Using the data from question 5, test at the $\alpha = 0.05$ level whether the lizards sampled by the ecologist have a lower prevalence of striping pattern than the population. State the relevant hypotheses, p-value, and decision.

11. Two independent groups of patients are being evaluated for their response to a new therapeutic regimen designed to reduce a specific inflammatory biomarker. Each patient provides a blood sample after four weeks of treatment, and the change in biomarker level is recorded.

    In the first treatment group, 22 patients show an average reduction of 9.843 units with a standard deviation of 2.764 units.

    In the second treatment group, 10 patients show an average reduction of 11.971 units with a standard deviation of 1.488 units.

    Test whether the second treatment group demonstrates a significantly greater mean reduction in the biomarker at the $\alpha = 0.05$ level. Assume both samples are normally distributed. State the relevant hypotheses, p-value, and decision.

12. A clinical trial is investigating whether a new medication reduces symptom severity in patients with a chronic inflammatory condition. Each patient is evaluated before beginning the medication and again after four weeks of treatment.

    For the 10 patients enrolled in this phase of the trial, symptom-severity scores are recorded as follows:

Before trt:	4	3	8	9	10	8	7	10	7	3
After trt:	10	9	9	7	9	6	6	1	2	8

    Determine whether symptom severity significantly changed following treatment at the $\alpha = 0.05$ level. State the relevant hypotheses, p-value, and decision.

Fisherian Hypothesis Testing

13. A workplace health program is comparing two drug-testing procedures to determine whether they yield different rates of positive results. During a routine evaluation, one group of employees is screened using a new rapid test, while another group is screened using the current standard test. Follow-up laboratory confirmation classifies each result as either positive or negative.

    Among employees tested with the rapid procedure, 45 were confirmed positive and 63 were confirmed negative.

    Among employees tested with the standard procedure, 15 were confirmed positive and 91 were confirmed negative.

    The laboratory that develops the tests claims that there's no long run difference between the two procedures.

    (a) State the relevant hypotheses for checking the laboratory's claims.

    (b) Develop the contingency table for the two testing procedures, include labels.

14. Write out the formula, using the contingency table in question 13, to compute an exact test on the claim that the two drug tests are not different.

15. Interpret an OR with a 95% confidence interval of $(2.13, 9.07)$ in the context of question 13.

16. A study conducted on COVID hospitalizations between vaccinated and non-vaccinated groups chose to perform an exact test on the count data. When they found an OR that crossed 1 and saw a p-value of 0.00001, they chose to reject the null hypothesis. The lab manager in charge of the team conducting the study cited that these results were invalid due to mis-interpreted statistics. State at least two of the errors that the lab manager found in their interpretation of the results.

Neyman-Pearson Hypothesis Testing

A parasitology research team is testing whether a new antihelmintic protocol reduces gastrointestinal parasite loads in sheep more effectively than the current deworming regimen. In a preliminary trial, animals were randomly assigned to receive either the standard treatment or the new protocol. After 28 days, each sheep's fecal egg count (FEC) is measured as an index of parasite burden.

The prelim results are:

Sheep receiving the standard treatment have an average FEC of 812 eggs per gram (EPG) with a standard deviation of 265 EPG.

Sheep receiving the new protocol have an average FEC of 611 EPG with a standard deviation of 231 EPG.

Each group includes 12 sheep.

17. State the main and alternative hypotheses for the helminth study.
18. Estimate the effect size from the prelim trial.
19. NP 3 Calculate the power of the prelim trial.
20. Estimate the minimum sample size needed for a subsequent trial to have a power of 80%.