Midterm Exam, STAT 252, AY 22-23, SS
Course instructor: Michael Daniel Lucagbo
Due date: April 30, 2023 (Sunday), at 10am

Instructions

- You are allowed to consult textbooks, notes, or any of our class materials. However, do not consult anyone except for your class Instructor. Work on the exam *individually*. Please email me immediately if you have any questions.
- Your exam answers must be submitted with a careful and concise write-up of the results. Any necessary codes should also be included in the file/s that you submit. However, a solution to a problem that consists only of software code and output will receive no credit.
- Email me your answers as portable document format (PDF) files. Make sure no one else is copied when you email me your answers.
- For the parts involving software coding and output, you are encouraged (but not required) to use R Markdown. For an introduction to R Markdown, you may watch this video:
https://www.youtube.com/watch?v=DNS7i2m4sB0

---

1. **Microbiological method comparison**: A laboratory wants to determine if two different methods (A and B) give similar results for quantifying a particular bacterial species in a particular medium. The bacterial counts are shown below. Under each method, the counts form a random sample. Assume that the counts follow a Poisson distribution, since this distribution a typical model for such data. Let μ_A and μ_B represent the population mean counts for Methods A and B, respectively. Let $\theta = \mu_A - \mu_B$.

	Replicate							
Method	1	2	3	4	5	6	7	8
A	176	125	152	180	159	168	160	151
B	164	121	137	169	144	145	156	139

   (a) Use the bootstrap to estimate the standard error of $\hat{\theta} = \hat{\mu}_A - \hat{\mu}_B$, where $\hat{\mu}_A$ and $\hat{\mu}_B$ are sample means of counts for methods A and B, respectively. Use $B = 2000$ bootstrap samples. Assume the replicates are *unpaired*.

   (b) Provide the histogram for the bootstrap distribution of $\hat{\theta}^*$s from (a) and comment on the features of the histogram.

   (c) Compute a 95% bootstrap percentile confidence interval for θ and interpret the resulting interval. Use $B = 2000$ bootstrap samples.

2. The dataset **fishery.csv** contains 40 annual counts of the numbers of recruits (R) and spawners (S) in a salmon population. The units are in thousands of fish. Recruits are fish that enter the catchable population. Spawners are fish that are laying eggs. Spawners die after laying eggs. The classic model for the relationship between spawners and recruits is

$$R = \frac{1}{\beta_0 + \beta_1/S}, \quad \beta_0 \geq 0 \text{ and } \beta_1 \geq 0. \tag{1}$$

We can fit such a model by using a linear regression given by

$$\frac{1}{R_i} = \beta_0 + \beta_1 \frac{1}{S_i} + \epsilon_i,\ i = 1, \ldots, 40. \tag{2}$$

The S variable can be considered as fixed. Suppose that ϵ_is are iid with mean 0 and finite variance, but their distribution is unknown. The total population abundance will only stabilize if $R = S$. Thus, the stable population level is the point where the line $R = S$ intersects the curve relating R and S. The total population will decline if fewer recruits are produced than the number of spawners who died producing them. If too many recruits are produced, the population will also decline eventually because there is not enough food for them all. Thus, only some middle level of recruits can be sustained indefinitely in a stable population. Let S_0 be the value of S at the stable population level.

(a) Fit the regression model in (2) and estimate S_0. Call this estimate $\hat{S}_0$.

(b) Use the bootstrap to obtain an estimate of the standard error of $\hat{S}_0$. Use $B = 2000$ bootstrap samples.

(c) Provide the histogram for the bootstrap distribution and comment on the features of the histogram.

(d) Compute a 95% bootstrap percentile confidence interval for S_0 and interpret the resulting interval..