Biometry mid-term exam Spring 2021

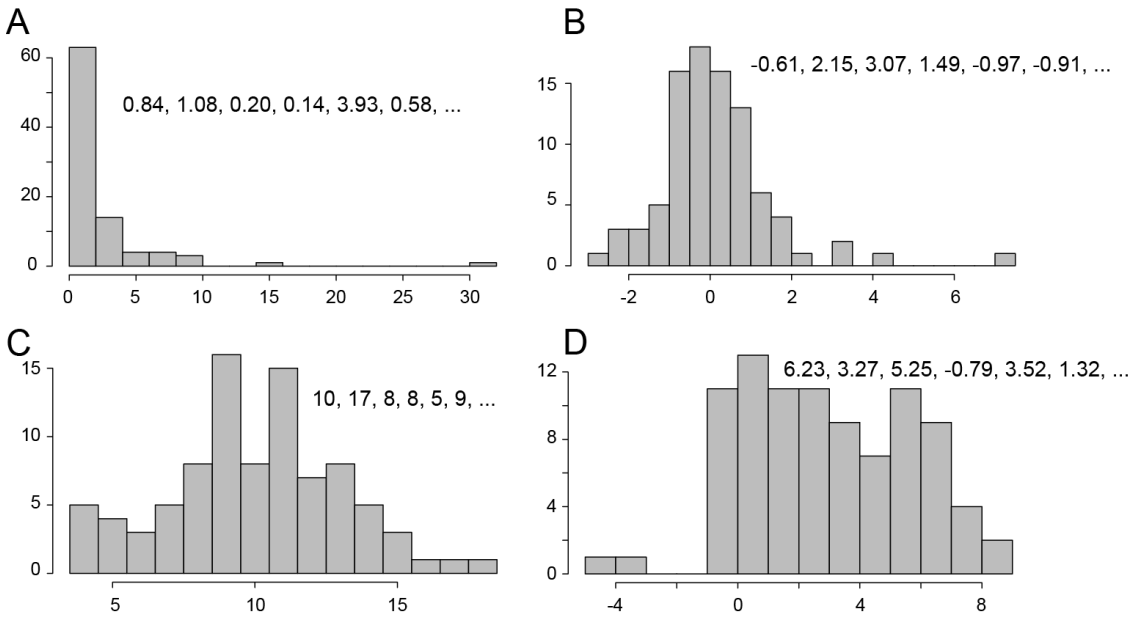
(235 pts total)

Section 1 – Short answer

(I would spend no more than 30 minutes maximum on the short answer section in order to leave yourself enough time for the long answer problems.)

1. (16 pts) Below are four histograms of data, along with a sample of six values from that distribution. For each panel, list **all** the possible distributions from the table below from which these data may have originated. (This is different from asking which distributions might be used to describe the data in an analysis. Here I am asking which distributions, strictly speaking, could have actually generated the data in each plot.)

Normal	Standard normal	Log-normal
Poisson	Binomial	Gamma
Chi-squared	F	t
Beta	Multinomial	Uniform



A)

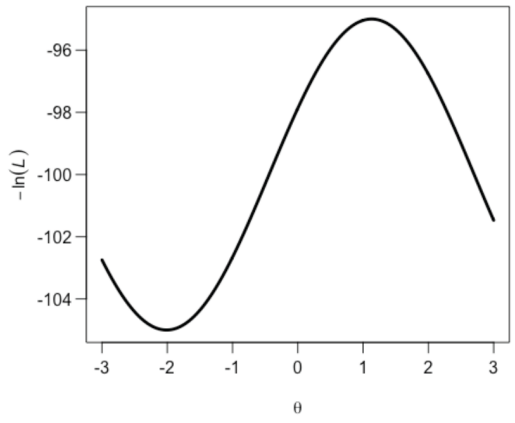
B)

C)

D)

Biometry mid-term exam Spring 2021

2. (5 pts) Based on the figure below, what is the (approximate) MLE for the parameter θ.



3. (5 pts) What is the probability of making a Type I error when making k independent comparisons, assuming α_c is your per-comparison cut-off for significance?

4. (9 pts) Let's say you have a model with a parameter α. Describe in words the interpretation of a 95th percentile confidence interval on a parameter α.

5. (10 pts) Name (in order) the six steps of null hypothesis significance testing.

Biometry mid-term exam Spring 2021

6. (15 pts) Assume two sets of independent random variables

$$X \sim N(\mu, \sigma^2)$$

and

$$Y \sim Pois(\lambda)$$

and a composite variable

$$Z = a + bX + cY$$

where a,b,c are constants, and X and Y are the random variables described above.

a) (7.5 pts) What is E[Z] in terms of μ, σ^2, λ? (Answer need not involve all three parameters.)

b) (7.5 pts) What is Var[Z] in terms of μ, σ^2, and λ? (Answer need not involve all three parameters.)

Section 2 – Long answer

7. (35 pts) Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with probability density function

$$f(x|\sigma) = \frac{1}{2\sigma} e^{-\frac{|x|}{\sigma}}$$

A) (5 pts) Find the E[X].

B) (20 pts) Find the maximum likelihood estimator for the parameter σ.

C) (10 pts) How would you use bootstrap to find the standard error of the maximum likelihood estimator?

Biometry mid-term exam Spring 2021

8. (40 pts) A researcher is studying Chytridiomycosis (an infection disease caused by the fungus *Batrachochytrium dendrobatidis*) in Costa Rican variable harlequin toads, and surveys a series of locations for their infection status. The researcher returns from sampling and, following some analysis, reports that the number of sites with Chytridiomycosis infection (Y) can be modelled as

$$Y \sim N(\hat{\mu} = 100, \widehat{\sigma^2} = 25)$$

a) (20 pts) A colleague scratches her head and says: "You should have used a Binomial distribution $Binom(n, p)$ instead of a Normal distribution."

Using the information that you have, estimate n and p. Since you do not have a calculator, you can leave your answer in the form of an expression (that is, the mathematical expression you would type into a calculator to get the answer if you had one).

b) (20 pts) Another colleague scratches his head and says: "Your sample size is too small, because we need the 95th percentile confidence intervals on $\hat{p}$ to be no bigger than 0.01". How many sites should the researcher sample in order to get 95th percentile confidence intervals that are no wider than 0.01. Since you do not have a calculator, you can leave your answer in the form of an expression.

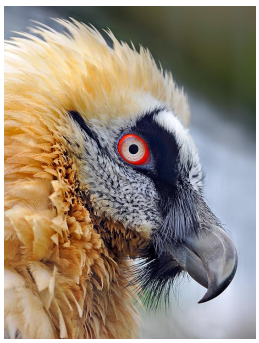
Biometry mid-term exam Spring 2021

9. (30 pts) Consider the recent article from the *New York Times* provided to you with the exam.

a) (15 points) The authors make some valid points in this article. Name two of them.

b) (15 pts) The authors make at least one major error in this story. What is it? Why are the authors incorrect?

Biometry mid-term exam Spring 2021



10. (65 points) Bearded Vulture (*Gypaetus barbatus*) is one of the most charismatic species of birds of prey; as Doug Futuyma says "if you are going to have a favorite vulture, this is it!". Unfortunately, they nest in secluded cliffs; so their nests are hard to find, and even harder to monitor. You have a research team that is responsible with monitoring a small Bearded Vulture population in Beypazari, Turkey. The population consists of only 10 pairs. They are territorial birds and tend to occupy the same sites across years (high site fidelity). However, for various reasons a pair may not breed in any given breeding season and that site will remain unoccupied.

a) (15 points) Let p_i be the probability that the i^{th} site is occupied ($i = (1,2,3 \dots 10)$), and J_i be the number of juveniles produced that season by a pair in site i. Assume that the probability of occupancy and the number of juveniles produced by each occupied nest is independent between nests, and that occupancy and the number of juveniles can be surveyed with no error. Write the joint likelihood for the dataset you have on occupancy status and number of juveniles. (Note: You have to decide on the distribution for J_i.) (Hint: Implicit in the wording is that fact that you have two datasets here. One dataset on occupancy and another on count conditional on occupancy.)

b) (15 points) In order to detect occupancy, your research team goes to previously known nesting sites to monitor the area using telescopes and binoculars. Let π be the probability of detecting vultures for a site that is occupied in a single survey visit.

$$P(detection|occupied) = \pi$$

and that this probability is the same across all sites. If all 10 sites are occupied in a given breeding season, how many visits does your team have to make to each site in order detect (on average) occupancy on 5 sites of those sites? (Hint: Your result will be in terms of π.)

Biometry mid-term exam Spring 2021

Nearest Neighbor Distance (NND) is one of the indicators used to measure the territoriality of a species. It is the distance of a nest site to the nearest nest site of the same species. Higher distances indicate stronger territoriality. In this regard, it can be considered as measure of intraspecific competition for nest sites. NND can also be calculated between species, as in the distance of a nest site to the nearest nest site of a competitor species. This is a measure of interspecific competition.

The Egyptian Vulture (*Neophron percnopterus*) is a closely related species to the Bearded Vulture. They both nest in cliffs and generally in the same area.

c) (15 points) Using NND as a measure of competition, suggest a parametric method to test whether intraspecific competition has the same strength as interspecific competition for the Bearded Vulture.

d) (15 points) Using NND as a measure of competition, suggest a permutation-based method to test whether intraspecific competition has the same strength as interspecific competition for the Bearded Vulture.

e) (5 points) Why might the permutation-based method be preferred in this case?