

**Prepare a document with your answers to each of the following questions. Each question will be worth 10 points. You must include your R code to get any credit for questions that include computation.**

**1)** Seasonal effective disorder can lead to depression during the winter months. A study of mortality in Norway reveals that in a one year period there were 1236 suicides. Of this total 572 occurred between October 10<sup>th</sup> and March 25<sup>th</sup>.

**a)** Use an appropriate test to evaluate the null hypothesis that risk of suicide is unrelated to the season.

**2)** You are teaching a class of 126 students. Of this number 47 are first generation college students. 14 of the first-generation students earned an A in the class. Meanwhile 43 of your students that are not first generation earned an A in the class.

**a)** Evaluate the null hypothesis that being a first-generation college student does not impact the probability of earning an A in your class.

**3)** Two lakes were freshly stocked with fish last season. You go to the field and seine fish from both lakes. The body size of the fish you catch is recorded below. You need to determine whether the fish in the two lakes are the same size.

lake 1	7.69	10.27	9.38	9.89	9.89	8.81	9.82
	9.48	9.78	9.45				
lake 2	9.06	8.93	11.32	10.1	9.8	11.22	8.61
	10.07	8.49	9.02	10.01	8.94	9.02	10.97

**a)** State the null hypothesis of the statistical test you will perform?

**b)** Is the mean fish size in the two lakes significantly different?

**4)** You are performing a study to determine whether fish can learn to distinguish shapes. You have a feeding system that allows fish to press either a circle or square to release food. You record the time required for a fish to release food the first time the feeder is placed in the tank and then later after 10 feeding trials.

	Fish 1	Fish 2	Fish 3	Fish 4	Fish 5	Fish 6	Fish 7	Fish 8	Fish 9	Fish 10
Trial 1	53.98	44.3	52.3	63.6	57.98	51.45	46.79	54.92	43.96	82.75
Trial 10	61.72	52.78	47.16	41.34	64.46	33.09	55.61	48.33	54.63	60.11

**a)** do fish learn to access food more quickly over time.

**5)** Males and females often differ in life history traits like mortality rates across age classes. A long-term study of a species has provided the data below on life span.

females	6	6.2	3.2	5.8	6.8	8.2	4.9	6.2	6	7.4
	5.8	6	5	6	4.8	5.8	5.4	5.8	5.3	7.7
males	5.3	6	5.5	6.9	7.6	5.5	4.7	6.5	6.7	5.5
	6.3	4.3	6.8	4.4	7.7	6.5	8.1	5.9	4	6.9

**a)** Determine whether males or females have more variation in lifespan.

6) You are studying the impact of a soil additive on the germination of seeds. You have collected the proportion of seeds that germinate with four different soil additives plus a control where no soil additive was provided.

	trial 1	trial 2	trial 3	trial 4	trial 5	trial 6	trial 7	trial 8	trial 9	trial 10
control	0.72	0.5	0.54	0.61	0.65	0.62	0.29	0.38	0.47	0.69
treatment 1	0.56	0.56	0.56	0.61	0.68	0.65	0.31	0.34	0.54	0.63
treatment 2	0.44	0.67	0.5	0.63	0.6	0.48	0.77	0.67	0.89	0.67
treatment 3	0.62	0.64	0.6	0.52	0.43	0.65	0.8	0.69	0.62	0.73
treatment 4	0.72	0.76	0.75	0.68	0.8	0.73	0.54	0.76	0.9	0.73

a) Determine if the proportion of seeds that germinate with different soil additives varies. If it does determine which treatments are significantly different.

b) Produce a high-quality plot of this data. Choose a format that illustrates the variability of the data. Insure that you include appropriate axes or legends.

7) You believe that *Thylacinus cynocephalus* (a marsupial wolf analog), though widely believed to have gone extinct in 1933, may still be alive in thinly settled regions of Tasmania. Using museum specimens, you have developed a DNA genotyping test that will accurately identify *T. cynocephalus* DNA 99.9% of the time. You also evaluate your test by running 1000 marsupial samples that are not from the genus *Thylacinus*. You find that 2 of these returned false positives. You go to Tasmania and collect scat, hair samples and bone fragments (all fresh) from a number of locations to test for *Thylacinus* DNA. After testing all 73 of your samples you find that one of them are positive for *Thylacinus* DNA.

a) Use Bayes theorem to determine the probability that you have identified the remains of a *Thylacinus*.

b) Justify your choice of a prior in this problem.

8) In one to two sentences describe what this code is doing.

```
all.data <- read.csv("my.file.csv", as.is = T)
observed.var <- 1.55
rec.result <- vector()
for(i in 1:1000){
  x <- sample(all.data, size = 23, replace = F)
  rec.result[i] <- var(x)
}
sum(rec.result >= observed.var) / length(rec.result)
```

## 9 & 10

Pick 2 questions and write a one paragraph answer for each (max 250 words)

Discuss three causes of the reproducibility crisis.

Discuss the difference in statistical and biological significance.

Discuss the difference in Bayesian and Frequentist statistics