Name ______________________________

Fall 2021 **STAT 587-2** **Exam II**

(60 points)

Instructions:

- You are allowed to use any resource except aid from another individual.
- Aid from another individual, will automatically earn you a 0.

(intentionally left blank)

1. Let $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$ and we construct an equal-tail $100(1-a)\%$ credible interval for μ. We are interested in the width of the credible interval and how it changes in response to changing the following. For each proposed change (assuming everything else stays the same), indicate whether the credible interval gets wider, narrower, stays the same, or cannot determine based on the available information. (2 pts each)

   Answer: Recall that the formula for this credible interval is

   $$\overline{y} \pm t_{n-1,1-a/2} s/\sqrt{n}$$

   where $\overline{y}$ is the sample mean, s is the sample standard deviation, and n is the sample size.

   (a) a gets larger

   Answer: narrower since $1-a$ gets smaller and thus

   (b) sample mean gets larger

   Answer: stays the same as this only shifts the interval

   (c) sample variance gets larger

   Answer: wider

   (d) sample size increases

   Answer: narrower

   (e) units of Y_i are changed from kilometers to meters

   Answer: wider since the sample standard deviation will be larger (or stays the same if you consider back converting the units)

2. Let $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$. The following are two-sided 95% confidence intervals for μ. What can be said about the p-value for the test $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$? (2 pts each)

   (a) (-1,10)
   Answer: p-value is greater than 0.05

   (b) (1,2)
   Answer: p-value is less than 0.05

   (c) (0,2)
   Answer: p-value is equal to 0.05

3. Let $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$. The following are p-values for the test $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$. What can be said about a two-sided 95% confidence interval for μ? (2 pts each)

   (a) $p = 0.01$
   Answer: the interval does not contain 0

   (b) $p = 0.1$
   Answer: the interval contains 0

4. To detect injury in pigs, researchers conduct a study using a random set of already injured pigs. For each pig, the researcher uses a scale under each foot while a pig is eating. The force from each foot is measured and the researcher calculates the difference (injured foot minus non-injured foot). The researcher hypothesizes that there will be less weight on the injured foot as the pig compensates for the injury.

   Let Y_i be the difference for pig i and assume $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$. The following statistics are observed

n	sample mean	sample standard deviation
10	0.8	1.5

   Answer:

```
n <- 10
sample_mean <- 0.8
sample_sd <- 1.5
```

   (a) Provide an interpretation for μ, i.e. what is μ (in words)? (2 pts)
   Answer: the population mean difference (non-injured minus injured force) amongst the already injured pigs

   (b) What is $\hat{\mu}_{MLE}$? (2 pts)
   Answer: 0.8

   (c) Compute a two-sided 80% confidence interval for μ. (6 pts)
   Answer:

```
a <- 0.2
sample_mean + c(-1,1)*qt(1-a/2, df = n-1)*sample_sd/sqrt(n)

## [1] 0.1439719 1.4560281
```

(d) Compute a p-value for testing $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$. (6 pts)

Answer:

```
2*(1-pt(abs(sample_mean-0)/(sample_sd/sqrt(n)), df = n-1))

## [1] 0.1259668
```

(e) State a conclusion based on the calculated p-value in the previous question? (2 pts)

Answer: there is insufficient evidence to indicate the data are incompatible with the null model, namely $Y_i \stackrel{ind}{\sim} N(0, \sigma^2)$.

(f) What is a more appropriate alternative hypothesis based on the researchers hypothesis? (2 pts)

Answer: $\mu < 0$ since the researcher hypothesizes there will be less force on the injured foot

5. A manufacturer produces nylon with an intended modulus of elasticity of 2.7 GPa. A random sample of nylon from the factory results in 3 of 100 samples with a modulus of elasticity less than 2.7 GPa. Let Y be the number of samples less than 2.7 GPa and assume $Y \sim Bin(n, \theta)$ where θ is the population proportion of samples that have modulus of elasticity less than 2.7 GPa.

Answer:

```
y = 3
n = 100
```

(a) Determine the Bayes estimator for θ, i.e. the posterior expectation. (2 pts)

Answer:

```
(1+y)/(2+n)

## [1] 0.03921569
```

(b) Determine a 70% equal-tail credible interval. (2 pts)

Answer:

```
a <- 0.3
qbeta(c(a/2, 1-a/2), 1+y, 1+n-y)

## [1] 0.02028668 0.05865536
```

(c) Determine the probability that θ is less than 0.05. (2 pts)

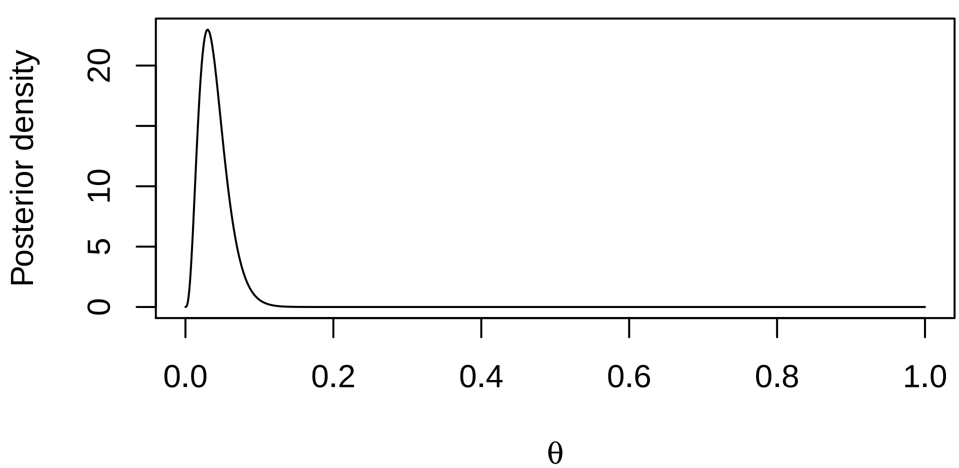
Answer:

```
pbeta(0.05, 1+y, 1+n-y)

## [1] 0.7491401
```

(d) Draw a graph of the posterior for θ. (4 pts)

Answer:



6. The file `data.csv` contains data that are assumed to be normally distributed. Answer the following questions based on these data.

Answer:

```
d <- read.csv("data.csv")
y <- d$y
```

(a) How many observations are there? (2 pts)

Answer:

```
length(y)

## [1] 15
```

(b) What is the sample mean of the data? (2 pts)

Answer:

```
mean(y)

## [1] 28.61893
```

(c) What is the sample variance of the data? (2 pts)

Answer:

```
var(y)

## [1] 116.1606

sd(y) # half points

## [1] 10.77778
```

(d) Construct a two-sided 98% confidence interval for the population mean. (2 pts)

Answer:

```
t.test(y, conf.level = 0.98)$conf.int

## [1] 21.31546 35.92240
## attr(,"conf.level")
## [1] 0.98
```

(e) Compute a p-value for the test with alternative hypothesis that the population mean is greater than 34. (2 pts)

Answer:

```
t.test(y, alternative = "greater", mu = 34)$p.value

## [1] 0.9631846
```