Name ____

Fall 2020 **STAT 587-2** **Exam II**

(45 points)

Instructions:

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

1. Let $Y_i \overset{ind}{\sim} N(\mu, 1)$ and use $\hat{\mu} = \overline{Y} + 2$ as an estimator for μ.

   (a) Is $\hat{\mu}$ unbiased for μ?
   Answer: No, because $E[\hat{\mu}] = \mu + 2$.

   (b) What is the variance for $\hat{\mu}$?
   Answer:

   $$Var[\hat{\mu}] = Var[\overline{Y} + 2] = Var[\overline{Y}] = 1/n$$

   (c) Is $\hat{\mu}$ consistent for μ?
   Answer: No, we know $\overline{Y}$ is consistent for μ, but $\hat{\mu} = \overline{y} + 2$ is always off by 2. Thus $\hat{\mu}$ is not consistent for μ.

   (d) If we collect 9 observations with $\overline{y} = -1$, what is the value for $\hat{\mu}$?
   Answer:

   $$\hat{\mu} = \overline{y} + 2 = -1 + 2 = 1$$

2. For the following questions determine, with everything else held constant, whether the confidence interval width for a normal mean will increase, decrease, stay the same, or the answer cannot be determined.

   (a) Sample size increase
   Answer: Width will decrease

   (b) Sample standard deviation decreases
   Answer: Width will decrease

   (c) Confidence level increases
   Answer: Width will increase

   (d) Sample average increases
   Answer: Width will stay the same

   (e) t-critical value increases
   Answer: Width will increase

3. During the 2020 general election, an exit poll worker randomly selects (through a Bernoulli random number generator app) voters in Ames Precinct 2-4 and asks them if they voted for anybody in the Iowa 4th district race. The poll worker found that 17 of 20 people indicated that they had voted in that race and nobody declined to answer. In the following, **probability** refers to the probability that a voter in this precinct voted in this race.

   (a) Is this a random sample of voters in Ames Precinct 2-4?

   Answer: Yes, selection was done through a random number generator app.

   (b) Which model is the most appropriate to analyze these data?

   Answer: binomial

   (c) Determine the value of the maximum likelihood estimate for the probability.

   Answer:

```
y = 17
n = 20
y/n

## [1] 0.85
```

   (d) Determine the value of the posterior mean when using a Unif(0,1) prior on the probability.

   Answer:

```
a = b = 1
(a+y)/(a+b+n)

## [1] 0.8181818
```

   (e) Using `binom.test`, construct a 90% equal-tail confidence interval for the probability.

   Answer:

```
binom.test(y, n, conf.level = 0.9)$conf.int

## [1] 0.6563362 0.9578306
## attr(,"conf.level")
## [1] 0.9
```

   (f) Construct a 99% equal-tail credible interval for the probability.

   Answer:

```
alpha = 1-.99
qbeta(c(alpha/2, 1-alpha/2), a+y, b+n-y)

## [1] 0.5678352 0.9660529
```

4. In four test fires of Falcon 9 Full Thrust first-stage boosters, SpaceX measured the average sea-level thrust to be 7,607 kN with a standard deviation of 1,034 kN.

   (a) Is this a random sample of Falcon 9 Full Thrust first-stage boosters?

   Answer: No, most likely they built a booster. Tested it. Then built another booster. Thus, no random process was used to select these boosters.

   (b) Which model is the most appropriate to analyze these data?

   Answer: normal

   (c) Assuming the default prior, what is the posterior mean for the average thrust?

   Answer:

```
n = 4
sample_mean = 7607
sample_sd   = 1034
```

   7607 kN

   (d) Calculate a 95% confidence interval for the average thrust.

   Answer:

```
a = .05
sample_mean + c(-1,1)*qt(1-a/2, df = n-1) * sample_sd /sqrt(n)

## [1] 5961.675 9252.325
```

   (e) Using the default prior, calculate the posterior probability that the average thrust is greater than 7,300 kN.

   Answer:

```
mu0 = 7300
t = (mu0-sample_mean)/(sample_sd/sqrt(n))
```

   Let μ be the average thrust and $\mu|y \sim t_{n-1}(\overline{y}, s^2/n)$. Thus

$$P(\mu > 7300|y) = P\left(\frac{\mu - \overline{y}}{s/\sqrt{n}} > \frac{7300 - 7607}{1034/\sqrt{4}}\right) = P(T > -0.5938104) = 1 - P(T \leq -0.5938104).$$

```
1-pt(t, df = n-1)

## [1] 0.7027818
```

5. The file `concrete.csv` contains data on an experiment to test the tensile strength (psi) of random selected concrete beams. Answer the following questions based on these data.

   Answer:

```
d = readr::read_csv("concrete.csv")

## Parsed with column specification:
## cols(
##   tensile_strength = col_double()
## )

y = d$tensile_strength
```

   (a) How many beams were tested?

   Answer:

```
length(y)

## [1] 15
```

   (b) What is the sample average tensile strength?

   Answer:

```
mean(y)

## [1] 50.18654
```

   (c) What is the sample standard deviation of tensile strength?

   Answer:

```
sd(y)

## [1] 1.819735
```

   (d) Construct a 95% credible interval for the population mean tensile strength.

   Answer:

```
t.test(y)$conf.int

## [1] 49.17880 51.19427
## attr(,"conf.level")
## [1] 0.95
```

   (e) Compute a p-value for the test that the population mean tensile strength is equal to 50 psi versus the alternative that the strength is greater than 50 psi.

   Answer:

```
t.test(y, mu = 50, alternative = "greater")$p.value

## [1] 0.3486733
```

6. Let $Y_i \stackrel{ind}{\sim} Exp(\lambda)$ where $E[Y_i] = \frac{1}{\lambda}$. Using the prior $\lambda \sim Ga(a, b)$ (rate parameter b), the posterior is $\lambda|y \sim Ga(a+n, b+n\overline{y})$. Answer the following questions based on observing $n = 10$, $\overline{y} = 3.3$, and assuming the prior $\lambda \sim Ga(1, 1)$.

   Answer:

```
n = 10
ybar = 3.3
a = b = 1
```

   (a) What is the posterior expectation?

   Answer:

```
(a+n)/(b+n*ybar)

## [1] 0.3235294
```

   (b) What is the posterior standard deviation?

   Answer:

```
sqrt(a+n)/(b+n*ybar)

## [1] 0.09754779
```

   (c) What is the posterior median?

   Answer:

```
qgamma(.5, a+n, b+n*ybar)

## [1] 0.3137801
```

   (d) Compute an 80% equal-tail credible interval.

   Answer:

```
qgamma(c(.1,.9), a+n, b+n*ybar)

## [1] 0.2064925 0.4531365
```

   (e) Compute the posterior probability that the exponential mean $(1/\lambda)$ is greater than 3.

   Answer:

$$P(1/\lambda > 3) = P(\lambda \leq 1/3)$$

```
pgamma(1/3, a+n, b+n*ybar)

## [1] 0.5792485
```