Total points: 28

1. Despite what it says on the bottle, the amount of beer in a bottle actually varies a little (in line with a normal distribution). I picked a 12 pack beer and measured the amount of beer in each bottle: 12.54, 13.40, 10.41, 12.36, 12.06, 11.42, 12.76, 11.91, 13.19, 11.73, 10.70, 12.21.

a. (6pts) Find a 95% confidence interval for σ^2.

ANSWER: Recall that the MLE for σ^2 of a normal distribution is $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ 1pt

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

Recall that S^2 is related to the χ^2 distribution such that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$ 1pt

Since we have a 12 pack of beer, $n = 12$ and

$$11\frac{S^2}{\sigma^2} \sim \chi^2(11)$$

To construct a 95% confidence interval, let X be any random variable $\sim \chi^2(11)$ then

$P(X < \underline{3.82}) = 0.025$ and 1pt

$P(X < \underline{21.92}) = 0.975$ and ~~1pt~~

$P(3.82 < X < 21.92) = 0.95$

Here we see that $(3.82, 21.92)$ would make a 95% confidence interval for any r.v. $\sim \chi^2(11)$ like $11\frac{S^2}{\sigma^2}$, but we are only interested in σ^2, so we have to transform our interval.

$$P(3.82 < x < 21.92) = 0.95$$

$$P\left(3.82 < 11\frac{S^2}{\sigma^2} < 21.92\right) = 0.95$$

$$P\left(0.347 < \frac{S^2}{\sigma^2} < 1.99\right) = 0.95$$

$$P\left(\frac{1}{0.347} > \frac{\sigma^2}{S^2} > \frac{1}{1.99}\right) = 0.95$$

$$P\left(2.88 > \frac{\sigma^2}{S^2} > 0.502\right) = 0.95$$

We can calculate S^2 from the data. There are two common ways to calculate S^2, both use $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and either is acceptable for this HW:

$$\bar{X} = \frac{1}{12}(12.54 + 13.40 + \ldots + 12.24) = 12.06$$

$$\text{(biased) } S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{12}\left[(12.54-12.06)^2 + \ldots + (12.24-12.06)^2\right] = \underline{0.752}$$

or

$$\text{(unbiased) } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{11}\left[(12.54-12.06)^2 + \ldots\right] = \underline{0.821}$$

1pt for correctly calculating ~~$\bar{X}$ and~~ S^2

1pt $\bar{X}$

Continuing from above...

$$\therefore P\left(2.88 > \frac{\sigma^2}{0.821} > 0.502\right) = 0.95$$

$$P(2.36 > \sigma^2 > 0.412) \approx 0.95$$

So a 95% CI for σ^2 is approximately

$$\underline{(0.412, 2.36)}$$

or

$$\underline{(0.378, 2.17)}$$

1pt for plugging everything in correctly and simplifying to the right answer

1.b. (2pts) Find another 95% Confidence Interval.

ANSWER: How about the 95% CI for σ ?

If $P(0.412 < \sigma^2 < 2.36..) = 0.95$

then $P(\sqrt{0.412} < \sigma < \sqrt{2.36}) = 0.95$

$P(0.642 < \sigma < 1.53) \approx 0.95$

So a 95% CI for σ is $(0.642, 1.53)$ or $(0.614, 1.47)$ 2pts

Since we asked for any 95% CI, you might also have worked out the 95% CI for μ, which would be $(11.48, 12.64)$

or

$(11.51, 12.61)$

Alternatively students may have just created a new 95% CI region for σ^2. e.g. $P(0 < \frac{11 s^2}{\sigma^2} < 19.68) \approx .95$. Award full credit for this too.

1c. (6pts) Find a 90% CI for μ.

ANSWER: Recall that the mle for $\mu = \bar{x} = \frac{1}{n}\sum x_i = 12.06$ 1pt

Recall that $\frac{\bar{x} - E(\bar{x})}{sd(\bar{x})} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim Normal(0,1)$... 0 pts

but also recall that this doesn't apply here since we don't know the true $sd(\bar{x})$. We do know S^2 though, so we can use the other relationship between $\bar{x}$ and a distribution

$$\frac{\bar{x} - E(\bar{x})}{\hat{sd}(\bar{x})} = \frac{\bar{x} - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$ 1pt

Let $Y \sim t_{11}$ then

$P(Y < -1.796) = 0.05$ and 1pt

$P(Y < 1.796) = 0.95$ so 1pt

$P(-1.796 < Y < 1.796) \approx 0.90$

Replacing Y with $\frac{\ldots}{S/\sqrt{n}}$ (which is also [illegible])
we have

$$P\left(-1.796 < \frac{\bar{x}-\mu}{S/\sqrt{n}} < 1.796\right) \approx 0.95$$

using our values of $\bar{x}$ and S^2 from part a, we have

$$P\left(-1.796 < \frac{12.06-\mu}{\sqrt{\frac{0.821}{12}}} < 1.796\right) \approx 0.95$$

$$P(-0.470 < 12.06-\mu < 0.470) \approx 0.95$$

$$P(-12.53 < -\mu < -11.59) \approx 0.95$$

$$P(11.59 < \mu < 12.53) \approx 0.95$$

So a 95% CI for μ = $(11.59, 12.53)$

or, with biased S^2,

$(11.61, 12.5)$

2pts for correctly substituting and simplifying

2. Following is the number of beer bike titles won by Will Rice College each year from 2000 to 2012: 2, 1, 1, 1, 2, 0, 1, 1, 2, 3, 2, 2, 2

a. (8pts) Assuming a binomial distribution, find a 99% CI for p, the probability that Will Rice claims one title. Comment on the quality of the approximation.

<u>ANSWER:</u> There are a number of ways to approximate this, they all involve an approximation to an approximation or more, so students should express doubt when commenting on the quality of the approximation

2pts for expressing doubt or reservation that we will use for beer bike titles. This is

$$\# \text{ of titles won} \sim \text{binomial}(3, p)$$

1pt

where $n = 3$ (there are 3 titles up for grabs each year) and p is the unknown probability of winning a title that we wish to estimate.

There is a binomial approximation to the normal distribution that the book uses to construct a CI for p on page 303. According to this a 95% CI for p will be

$$\left(\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

5

where $\hat{p} = \frac{\bar{X}}{n}$.

However, the binomial approximation to the normal only works when n is large ($np \geq 10$). Here $n=3$. The approximation is horrible, and hence doesn't deserve full credit. Nevertheless,

$$\hat{p} = \frac{\bar{X}}{n} = \frac{1.538}{3} = 0.512$$

1pt

$$99\%\ CI = \left(0.512 - 1.96\sqrt{\frac{0.512(1-0.512)}{3}},\ 0.512 + 1.96\sqrt{\frac{0.512(1-0.512)}{3}}\right)$$

$$= (0.053,\ 1.078)$$

1pt

0.038

Alternatively, for full credit students can rely on the Central limit theorem which states that

$$\frac{\bar{X} - E(\bar{X})}{Sd(\bar{X})} \sim N(0,1)$$

First students must notice that

$$E(\bar{x}) = E(x) = np = 3p$$ — 1pt

Next students must calculate $sd(\bar{x})$. We know that for the binomial distribution $var(x) = np(1-p)$, so $var(\bar{x})$ will be $\frac{np(1-p)}{n} = p(1-p)$ and $sd(\bar{x}) = \sqrt{p(1-p)}$ but we cannot calculate this since we do not know p. We can proceed in one of three ways. All require $\bar{x} = \frac{1}{n}\sum X_i = 1.538$

1pt $\bar{x} = 1.538$

1. $$\frac{\bar{x} - 3p}{\sqrt{p(1-p)}} \sim N(0,1)$$ 1pt setup

$$P\left(-2.58 < \frac{\bar{x} - 3p}{\sqrt{p(1-p)}} < 2.58\right) \approx .95$$

$$P\left(-2.58 < \frac{1.538 - 3p}{\sqrt{p(1-p)}} < 2.58\right)$$ 1pt bounds

which does not have a unique solution

no points results

7

2. $$\frac{\bar{x} - 3p}{s/\sqrt{13}} \sim N(0,1) \text{ or } t_{12}$$

$s = 0.745$ $s^2 = 0.603$

1pt $s = 0.745$

(t_{12} isn't appropriate here since the original data are not normal, but it is very close to $N(0,1)$ so don't take off points.)

$$P\left(-2.58 < \frac{\bar{x} - 3p}{s/\sqrt{13}} < 2.58\right)$$ 1pt bands

$$P\left(-2.58 < \frac{1.538 - 3p}{0.745/\sqrt{13}} < 2.58\right)$$

$$P(0.143 < p < 0.883)$$

95% CI = (0.143, 0.883)

or

(0.075, 0.95)

(+ dist)

1pt results

6

2.b. (2pts - bonus) What is the probability of Will Rice Sweeping next beer bike? What assumptions do you make?

ANSWER: Method of moments and mle give the same result here, here's MOM

$$E(X) = np = 3p$$

$$p = \frac{1}{3}E(x)$$

$$\hat{p} = \frac{1}{3}\bar{X} \quad \text{by method of moments}$$

$$\hat{p} = \frac{1}{3} \cdot 1.538 \approx 0.513$$

$$P(\text{Will Rice Sweeps}) = P(X=3) \mid X \sim \text{binomial}(3,p)$$

$$\hat{P}(X=3) = \binom{3}{3}\hat{p}^3(1-\hat{p})^0 = 0.135$$

1pt

We must assume that the probability of Will Rice winning a beer bike title did not change between 2000-2012 and will not change before next year.

We are also assuming that the distribution is in fact binomial. This wouldn't be the case if Will Rice had different probabilities of winning a men's title vs. a woman's title vs. an alumni title.

1pt for reasonable list of assumptions

3. Assume the waiting time for the Rice bus is exponentially distributed. Here are 10 waiting times (in minutes) I observed when waiting for the bus: 1.5, 8.52, 0.5, 0.73, 2.36, 17.64, 8.01, 2.54, 17.78, 20.76

a. (6pts) Find a 95% Confidence Interval for θ.

Answer: Note that $E(\bar{x}) = E(x) = \theta$ 1pt

So we can connect θ to a distribution with

$$\frac{\bar{x} - E(\bar{x})}{\hat{sd}(\bar{x})} = \frac{\bar{x} - \theta}{s/\sqrt{n}} \sim t_{n-1} = t_9 \quad \sim N(0,1)$$ 1pt

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = 8.034$$ 1pt

$$s = \sqrt{\frac{1}{10}\sum_{i=1}^{10}(x_i - 8.034)^2} = 7.514$$ 1pt

Let $Y \sim t_9$

$P(Y < -2.26) = 0.025$ $\frac{1}{2}$pt

$P(Y < 2.26) = 0.975$ $\frac{1}{2}$pt

$P(-2.26 < Y < 2.26) \approx 0.95$

$$P\left(-2.26 < \frac{\hat{\theta} - \theta}{s/\sqrt{n}} < 2.26\right) \approx 0.95$$

$$P\left(-2.26 < \frac{8.034 - \theta}{7.514/\sqrt{10}} < 2.26\right) \approx 0.95$$

$$P\left(2.66 < \theta < 13.41\right) \approx 0.95$$

$\therefore$ 95% CI for θ = $\underline{(2.66, 13.41)}$

1pt for substituting and simplifying correctly