STAT 515 fa 2020 Exam II

Karl B. Gregory

This is a take-home test due to COVID-19. Do not communicate with classmates about the exam until after its due date/time. You may

- *Use your notes and the lecture notes.*
- *Use books.*
- *NOT work together with others.*

Write all answers on blank sheets of paper; then take pictures and merge to a PDF. Upload a single PDF to Blackboard.

1. Copy down this sentence on your answer sheet and put your signature underneath: *I have not collaborated with any other student on this exam. The work I have presented is my own.*

2. Suppose the number of birds that visit your bird feeder while you are eating breakfast each day follows a Poisson($\lambda = 8$) distribution.

   (a) Find the probabilities of the following events:

   i. Exactly 8 birds visit the feeder during breakfast.

   Letting X be the number of birds that visit the feeder during breakfast, we have

   $$P(X = 8) = \frac{e^{-8}8^8}{8!} = \texttt{dpois(8,8)} = 0.1395865.$$

   ii. At least one bird visits the feeder during breakfast.

   We have

   $$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-8} = 0.9996645.$$

   iii. Fewer than 12 birds visit the feeder during breakfast.

   We have

   $$P(X \leq 11) = \sum_{y=0}^{11} \frac{e^{-8}8^y}{y!} = \texttt{ppois(11,8)} = 0.888076.$$

   (b) Let Y be the number of birds that visit your feeder while you are eating breakfast over the course of a given week. Find

   i. $\mathbb{E}Y$.

   We have $Y \sim \text{Poisson}(\lambda = 56)$, so $\mathbb{E}Y = 56$.

   ii. $P(Y = 60)$.

   We have

   $$P(Y = 60) = \frac{e^{-56}56^{60}}{60!} = \texttt{dpois(60,56)} = 0.04473174.$$

   iii. $P(Y > 50)$.

   We have

   $$P(Y > 50) = 1 - P(Y \leq 50) = \texttt{1 - ppois(50,56)} = 0.7657168.$$

3. Suppose the weights of peaches of a certain variety sold at a farmers market are Normally distributed with mean 147 grams and standard deviation 4 grams.

   (a) Let X be the weight of a randomly selected peach.

      i. Find $P(X < 140)$.

      We have

      $$P(X < 140) = P(Z < (140 - 147)/4) = P(Z < -1.75) = 0.04005916.$$

      ii. Find $P(|X - 147| < 3)$.

      We have

      $$P(|X - 147| < 3) = P(-3/4 < Z < 3/4) = 0.5467453.$$

      iii. $P(X = 147)$.

      This is 0 because X is a continuous random variable.

      iv. Find the weight exceeded by exactly 1% of the peaches of this variety.

      The 0.99 quantile is

      $$147 + 4 \cdot z_{0.01} = 147 + 4 \cdot 2.326348 = 156.3054.$$

   (b) Let $\bar{X}_n$ be the mean weight of $n = 5$ randomly selected peaches.

      i. Find $P(\bar{X}_n < 144)$.

      We have

      $$P(\bar{X}_n < 144) = P(Z < \sqrt{5}(144 - 147)/4) = P(Z < -1.677051) = 0.04676625.$$

      ii. Find $P(|\bar{X}_n - 147| < 3)$.

      We have

      $$P(|\bar{X}_n - 147| < 3) = P(-\sqrt{5} \cdot 3/4 < Z < \sqrt{5} \cdot 3/4) = 0.9064675.$$

      iii. Find $P(\bar{X}_n > 148)$.

      We have

      $$P(\bar{X}_n > 148) = P(Z > \sqrt{5}(148 - 147)/4) = P(Z > 0.559017) = 0.2880751.$$

iv. Find $P(\bar{X}_n = 147)$.

This is equal to zero since $\bar{X}_n$ is a continuous random variable.

v. Find the probability that the total weight of a bag of 5 peaches exceeds 740 grams.

The total weight exceeds 740 if and only if $\bar{X}_n$ exceeds 148. This was computed earlier as 0.2880751.

4. A new variety of peaches has come to the farmers market. Suppose you buy a basket of 20 peaches having a total weight of 3,000 grams and that the standard deviation computed on the weights of the 20 peaches is $S_n = 3.5$. Assume that the peach weights are Normally distributed.

(a) Build a 95% confidence interval for the mean weight of peaches of this variety.

We have $\bar{X}_n = 3000/20 = 150$. A 95% CI for μ is given by

$$150 \pm \underbrace{t_{20-1,0.05/2}}_{2.093024} \frac{3.5}{\sqrt{20}} = (148.3619, 151.6381).$$

(b) Build a 99% confidence interval for the mean weight of peaches of this variety.

A 95% CI for μ is given by

$$150 \pm \underbrace{t_{20-1,0.001/2}}_{2.860935} \frac{3.5}{\sqrt{20}} = (147.761, 152.239).$$

(c) Would a 90% confidence interval be wider or narrower than the 95% confidence interval?

It would be narrower than both the 95% and the 99% confidence interval.

(d) You would like to build a 95% confidence interval with a margin of error of no more than 1 gram. What sample size should you use?

We should use $n = 48$, since $1.96 \cdot 3.5/\sqrt{n} \leq 1$ for all $n \geq 47.0596$.

(e) Suppose you want to test whether peaches of this new variety are heavier on average than peaches of the variety in question 3. Formulate the relevant hypotheses.

We would be interested in testing

$$H_0\colon\ \mu \leq 147 \text{ versus } H_1\colon\ \mu > 147.$$

(f) Based on your basket of 20 peaches, what is your conclusion about these hypotheses at the $\alpha = 0.05$ significance level.

> The test statistic is
>
> $$T_{\text{test}} = \frac{150 - 147}{3.5/\sqrt{20}} = 3.833259 > t_{20-1,0.05} = 1.729133.$$
>
> Since it is greater than the critical value we reject H_0 at the $\alpha = 0.05$ significance level.

(g) What is the p-value for testing these hypotheses based on your basket of 20 peaches?

> The p-value is the area under the pdf of the t_{20-1} pdf to the right of $T_{\text{test}} = 3.833259$. We can get this from R with
>
> `1 - pt(3.833259,20-1)` = 0.0005606611.

(h) Say whether you would reject the null hypothesis when testing

$$H_0\colon\ \mu = 147 \text{ versus } H_1\colon\ \mu \neq 147.$$

at the $\alpha = 0.01$ significance level based on your basket of 20 peaches.

> We would reject H_0 because our 99% confidence interval for μ does not contain the value 147.

(i) Make a 95% confidence interval for the variance σ^2 of the weights of peaches of this new variety.

> Using
>
> $$\chi^2_{20-1,0.975} = \texttt{qchisq(.025,20-1)} = 8.906516$$
> $$\chi^2_{20-1,0.975} = \texttt{qchisq(.975,20-1)} = 32.85233.$$
>
> we have
>
> $$\left(\frac{(20-1)3.5^2}{\chi^2_{20-1,0.975}}, \frac{(20-1)3.5^2}{\chi^2_{20-1,0.025}}\right) = (7.084734, 26.13255).$$

5. Suppose it is of interest to estimate the proportion p of students who eat a vegetarian diet.

   (a) Suppose you take a sample of 100 students and 5 say they are vegetarian.

   i. Build a 95% confidence interval for p.

> We compute the Agresti-Coull interval since $n\hat{p}_n = 5$. We have $\tilde{p}_n = 7/104 = 0.0673$. The interval is given by
>
> $$0.0673 \pm 1.96\sqrt{0.0673(1 - 0.0673)/104} = (0.019, 0.115).$$

ii. If the true proportion of vegetarians is $p = 0.034$, give the exact value of $P(\hat{p}_n \geq 0.05)$ for $n = 100$.

> Letting $Y \sim \text{Binomial}(100, 0.034)$ be the number of vegetarians in the sample, we have
>
> $$P(\hat{p}_n \geq 0.05) = P(Y \geq 5) = 1 - P(Y \leq 4) = \texttt{1 - pbinom(4,100,0.034)} = 0.2538522.$$

iii. Suppose you want to build a 95% confidence interval for p with a margin of error no greater than 1 percentage point. Using your sample of 100 students as a pilot study, suggest a sample size.

> We choose the smallest whole number n which satisfies
>
> $$1.96\sqrt{0.05(1-0.05)/n} \leq 0.01 \iff \underbrace{(1.96 \cdot \sqrt{0.05(1-0.05)}/0.01)^2}_{1824.76} \leq n.$$
>
> So we take n to be 1825.

(b) A food truck operator who serves vegetarian meals does his own study of the student population: in a random sample of 1000 students, he finds that 53 are vegetarian. He will visit campus with his food truck if he determines that the proportion of students who are vegetarian exceeds 4%.

i. State the hypotheses of interest to the food truck operator.

> He is interested in testing
>
> $$H_0\colon\ p \leq 0.04 \text{ versus } H_1\colon\ p > 0.04.$$

ii. State whether the food truck operator should reject or not reject his null hypothesis at significance level $\alpha = 0.05$ based on the data he collected.

> We have
>
> $$Z_{\text{test}} = \frac{0.053 - 0.04}{\sqrt{0.04(1-0.04)/1000}} = 2.097866.$$
>
> The critical value is $z_{0.05} = \texttt{qnorm(.95)} = 1.645$. Since $2.097866 > 1.645$, he rejects H_0 at the $\alpha = 0.05$ significance level.

iii. Compute the p-value for testing his hypotheses based on the data he collected.

> The p-value is given by
>
> $$P(Z > 2.097866), \quad Z \sim \text{Normal}(0,1),$$
>
> which is 0.01795849.