Solutions to STAT280 Exam 2

given Mar. 20, 1998

1. [25 points] True–False. (5 pts. each)

T F : If A and B are independent events, then $P(A \cup B) = P(A) + P(B)$.

Solution: FALSE. See p. 291 and p. 295.

T F : The addition rule of means $\mu_{X+Y} = \mu_X + \mu_Y$ holds whether the RV's X and Y are independent or not.

Solution: TRUE. See p. 329 and the bottom of p. 334 to the top of 335.

T F : When a test of hypotheses is done and the null hypothesis is not rejected, then we have strong evidence that it is true.

Solution: FALSE. This was discussed in lecture. Rejecting H_0 means there is strong evidence against it, but not rejecting it just means there wasn't strong evidence against it, not that there was strong evidence for it.

T F : The normal approximation to the binomial distribution is an example of the Central Limit Theorem.

Solution: TRUE. See p. 291 and p. 295.

T F : A researcher computes a 95% confidence interval for the height of adult men as 70.5 in. ± 1.2 in. This means 95% of the adult men have heights between 69.3 in. and 71.7 in.

Solution: FALSE. It means that 95% of the confidence intervals will contain the true mean height of adult men. See exercise 6.3 and the solution in the text.

2. [30 points] *True Fact: About 60% of Rice students ranked in the top 5% of their high school class.* Suppose a SRS of n Rice students is selected.

(a) If $n = 10$, what is the probability that 7 or more in the sample were in the top 5% of their high school class?

Solution: Let X be the number of students in the sample who were in the top 5%.

Then X is a $B(10, .6)$ RV, and we want the probability $P[X \geq 7]$ Now Table C does not cover the case $p = .6$. The key is to change from counting "successes" to counting "failures." An example was given in class. See also Example 5.5, p. 376. So let Y be the number in the sample who were not in the top 5%. Then Y is a $B(10, .4)$ RV and the event $[X \geq 7]$ is the same as the event $[Y \leq 3]$. We can look up these probabilities in Table C.

$$\begin{aligned} P[Y \leq 3] &= P[Y = 0] + P[Y = 1] + P[Y = 2] + P[Y = 3] \\ &= .0060 + .0403 + .1209 + .2150 \\ &= .3822. \end{aligned}$$

Although the Normal Approximation doesn't apply here by our rule of thumb ($\min\{np, n(1-p)\} = \min\{4, 6\} = 4 < 5$), it doesn't work too bad if one uses the continuity correction. Note that $\mu = 6$ and $\sigma = \sqrt{6 * .4} = 1.55$, and

$$\begin{aligned} P[X \geq 7] &= P[X \geq 6.5] && (1) \\ &= P\left[\frac{X - \mu}{\sigma} \geq \frac{6.5 - 6}{1.55}\right] && (2) \\ &\doteq P[Z \geq .32], \quad Z\ N(0,1) && (3) \\ &= 1 - .6255 && (4) \\ &= .3755. \end{aligned}$$

Here, (1) is the continuity correction, (2) is the standardization, (3) is the normal approximation, and (4) is just the table lookup.

(b) Suppose $n = 100$. What is the probability 70 or more in the sample were in the top 5% of their high school class?

Solution: Here, we have to use the normal approximation. Now X is a $B(100, .6)$ RV, and $\mu = 60 > 5$ and $n - \mu = 40 > 5$, so the rule of thumb is easily satisfied.

Applying the continuity correction and all as above, note that $\sigma = \sqrt{60 * .4} = 4.90$,

$$\begin{aligned} P[X \geq 70] &= P[X \geq 69.5] \\ &= P\left[\frac{X-\mu}{\sigma} \geq \frac{69.5-60}{4.9}\right] \\ &\doteq P[Z \geq 1.94], \quad Z\ N(0,1) \\ &= 1 - .9738 \\ &= .0262. \end{aligned}$$

In fact, the exact answer is .0248.

3. [25 points] In a large class of 100 students, 45 are women. In addition, 75 of the students in the class live off campus, and 30 of the women live off campus. A student is selected at random.

(a) What is the probability the random student is a woman?

Solution:

$$\frac{\#\ \textit{Women}}{\#\ \textit{All Students}} = \frac{45}{100} = .45.$$

(b) What is the probability the random student lives off campus?

Solution:

$$\frac{\#\ \textit{Off Campus Students}}{\#\ \textit{All Students}} = \frac{75}{100} = .75.$$

(c) Given that the student is a woman, what is the conditional probablity that she lives off campus?

Solution:

$$\frac{\#\ \textit{Off-Campus Women}}{\#\ \textit{All Women}} = \frac{30}{45} = .667.$$

To give a more formal answer using the definition of conditional probability, let A *be the event that the random student is a woman and* B *the event that the random* *student lives off campus. Then* $P(B) = .75$ *by part (b). Also*

$$P(A \cap B) \;=\; \frac{\#\ \textit{Students Off-Campus AND a Woman}}{\#\ \textit{All Students}} \;=\; \frac{30}{100} \;=\; .30.$$

Then

$$P(B|A) \;=\; \frac{P(A \cap B)}{P(A)} \;=\; \frac{.30}{.45} \;=\; .667.$$

(d) Given that the student lives off campus, what is the condional probility that the student is a woman?

Solution:

$$\frac{\#\ \textit{Off-Campus Women}}{\#\ \textit{All Off-Campus Students}} \;=\; \frac{30}{75} \;=\; .4.$$

The more "formal" argument is

$$P(A|B) \;=\; \frac{.30}{.75} \;=\; .4.$$

(e) Are the event that the random student is woman and the event that the random student lives off campus independent events? Why or why not?

Solution: NOT INDEPENDENT. There are several ways to argue it: $P(A) = .45$ *by part (a) and this is not equal to* $P(A|B) = .4$ *by part (d). Also,* $P(A \cap B) = .3$ $\neq P(A)P(B) = .45 * .75 = .3375$. *Many people thought from an "intuitive" point of view they should be independent, but of course the numerical results give the final answer.*

4. [20 points] A new weight loss program advertises that the participants have lost 22 pounds in five weeks on the average with a standard deviation of $\sigma =$ 10.2 lbs. Suppose in a sample of 100 people we find that the average weight loss is only 19.5 lbs.

(a) Does this provide strong evidence that the advertisers of the weight loss program are overstating the average weight loss? Use the $\alpha = .05$ significance level.

Solution: Let μ be the true mean weight loss. We want to test the hypotheses H_0 $\mu = 22$ and H_A $\mu < 22$. Note that if we reject the null hypothesis we will have strong evidence against it, which is what we are looking for. Then the z statistic is

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{19.5 - 22}{10.2/sqrt100} = 2.45.$$

The area to the right of this value under the $N(0,1)$ density is the same as the area to the left of -2.45, which is .0071. As the P-value (.0071) is less than the significance level (.05), we reject H_0 and conclude there is strong evidence against the claim.

(b) In a civil court case that the FDA brings agains this weight loss program, the judge decides he will find against the company only if he can obtain evidence against their advertised claim at the $\alpha = .001$ significance level. What is the judge's decision?

Solution: As the P-value (.0071) is NOT less than the judge's significance level, we do NOT reject at his level, so the judge finds in favor of the weight loss company. Presumably the FDA's lawyers will appeal.