UCSC AMS 7 **Quiz 2 – Solutions** **Summer Session II, 2016**

1. (5 pts) Suppose that the proportion of HIV infections in the *at-risk* population is 0.08. If three members of the at-risk population are selected at random (and independently of each other) and their blood samples are combined, what is the probability that the combined sample tests positive?

Note: You may assume that the combined blood test returns a positive result if any of the individual samples is infected, and that the sampling is like drawing with replacement.

*The combined sample tests positive if **at least one** of three samples came from an infected person. The complementary event is that **none** of the three samples came from an infected person. So...*

$$\begin{aligned} P(\textit{combined sample tests positive}) &= 1 - P(\textit{none of the individual samples are infected}) \\ &= 1 - \overbrace{(0.92) \times (0.92) \times (0.92)}^{\textit{samples are independent}} \\ &= 1 - 0.778688 \approx 0.2213. \end{aligned}$$

2. Suppose that 4 marbles are selected at random, with replacement from a jar containing 2 red marbles and 8 black marbles.

Observation: *If $X = \#$ of red marbles out of 4 drawn, then $X \sim B(4, 0.2)$, because*

$$P(\textit{red marble}) = 0.2$$

*for one draw from the jar, and the **four** draws are done with replacement.*

(a) (2 pts) What is the probability that ***exactly one*** of the selected marbles is red?

$$P(X = 1) = \frac{4!}{3! \times 1!} \cdot (0.2)^1 \cdot (0.8)^3 = 4 \cdot (0.2) \cdot (0.512) = 0.4096.$$

(b) (2 pts) What is the probability that ***no more than one*** of the selected marbles are red?

"No more than one" *is the same as $X = 0$ or $X = 1$, so*

$$P(\text{no more than one red}) = P(X = 0) + P(X = 1) = (0.8)^4 + 0.4096 = 0.8192.$$

(c) (1 pts) What is the probability that ***at least one*** of the selected marbles is red?

"at least one" *is the opposite of **none**, so*

$$P(\text{at least one red}) = 1 - P(\text{all black}) = 1 - (0.8)^4 = 0.5904.$$

3. Suppose that 100 marbles are selected at random with replacement from the jar in problem 2. Let X = number of red marbles in the sample.

(a) (2 pts) What are the ***mean*** and ***standard deviation*** of X?

$X \sim B(100, 0.2)$, so $\mu_X = 100 \cdot (0.2) = 20$ and $\sigma_X = \sqrt{100 \cdot (0.2) \cdot (0.8)} = 4$.

(b) (3 pts) What is $P(17 \leq X \leq 22)$? Use the normal approximation.

Observe that $np = 20 > 5$ and $nq = 80 > 5$, so use of the ***normal approximation to the binomial*** *is appropriate. Accordingly:*

$$\begin{aligned} P(17 \leq X \leq 22) &\approx P\left(\frac{16.5-20}{4} < Z < \frac{22.5-20}{4}\right) \\ &= P(-0.875 < Z < 0.625) \\ &= T(0.625) - (1 - T(0.875)) = \ldots? \end{aligned}$$

Since there is no entry for $Z = 0.875$ or $Z = 0.625$ in our normal table we have to approximate $T(0.875)$ and $T(0.625)$. We can...

$\Rightarrow$ *Average the table values around 0.875 and 0.625:*

$$T(0.875) \approx \frac{T(0.87) + T(0.88)}{2} = 0.8092$$

and

$$T(0.625) \approx \frac{T(0.62) + T(0.63)}{2} = 0.73405$$

$\Rightarrow$ *Round 0.875 and 0.625 to two decimal places, and use the table entries for the rounded values. In this case, we should round one of them up and the other one down (consider the corresponding picture of the normal curve to see why). I.e., we use*

$$T(0.625) \approx \overbrace{T(0.62)}^{down} = 0.7324 \quad \textit{and} \quad T(0.875) \approx \overbrace{T(0.88)}^{up} = 0.8106$$

or

$$T(0.625) \approx \overbrace{T(0.63)}^{up} = 0.7357 \quad \textit{and} \quad T(0.875) \approx \overbrace{T(0.87)}^{down} = 0.8078$$

Returning to the calculation of $P(17 \leq X \leq 22)$, we have

$$\begin{aligned} \textit{either } P(17 \leq X \leq 22) &\approx 0.8092 + 0.73405 - 1 = 0.54325, \\ \textit{or } P(17 \leq X \leq 22) &\approx 0.8106 + 0.7324 - 1 = 0.543, \\ \textit{or } P(17 \leq X \leq 22) &\approx 0.8078 + 0.7357 - 1 = 0.5435, \end{aligned}$$

and we can conclude that $P(17 \leq X \leq 22) \approx 0.543$.

Comment: *You are not meant to use all three. I used all three to illustrate the fact that you will get very similar results whichever method you choose. The first method* ***does*** *give the most accurate result for $P(-0.875 < Z < 0.625)$, and all of them slightly underestimate $P(17 \leq X \leq 22) \approx 0.5466$ (calculated using an on-line binomial calculator).*

4. (5 pts) The height of women in the U.S. has a normal distribution with mean $\mu = 63.6$ inches and standard deviation $\sigma = 2.5$ inches. What is the probability that the average height in a simple random sample of 100 U.S. women is greater than 64 inches? (You may invoke the central limit theorem.)

Since heights are assumed to be normally distributed and the sample is simple-random we may conclude that

$$\overline{h} \sim N\left(63.6, \frac{2.5}{\sqrt{100}}\right) = N(63.6, 0.25)$$

where $\overline{h}$ is the sample mean (even ***without*** *the central limit theorem). It follows that*

$$P(\overline{h} > 64) \approx P\left(Z > \frac{64 - 63.6}{0.25}\right) = P(Z > 1.6) = 1 - T(1.6) = 0.0548.$$