STOR 151 MIDTERM 2
PRACTICE EXAM, MARCH 2024

This is a practice exam — open book, 75 minutes for the whole exam; there are 6 questions here but a typical exam would ask you to do 3 questions in that time limit.

SHOW ALL WORKING — even correct answers will not get full credit if it's not clear how they were obtained. Incorrect answers will gain substantial credit if the method of working is substantially correct.

1. (From Midterm 2, Fall 2018.) Sam is a player on the UNC basketball team. In the long run, she succeeds (scores a basket) on about 76% of her free-throw attempts. During the ACC tournament, she takes a total of 41 free throws. What is the probability that she scores (a) at least 27 baskets, (b) more than 27 baskets? (c) Explain the distinction between (a) and (b) and show how you take that into account in your calculation. (Hint: use the normal approximation with a continuity correction.)

2. (From Final, Fall 2018.) An irregular eight-sided die has the following probabilities:

Number on face of the die	1	2	3	4	5	6	7	8
Probability	0.1	0.1	0.1	0.1	0.1	0.1	x	x

   (a) Find x.

   (b) Let A be the event "the die shows an odd number", Let B be the event "the die shows one of 1, 2, 3, 4". Find $P(A)$, $P(B)$, $P(A \text{ and } B)$ and $P(A \text{ or } B)$. Are the events A and B independent?

   (c) Suppose the die is tossed three times. What is the probability that the total is greater than 22?

   (d) If the same die is thrown a large number of times, the mean score is 5.1 and the standard deviation is 2.39 (you can assume these numbers without proof). If the die is thrown 100 times, what is the approximate probability that the *average* score is at least 5.5?

3. (From Final, Fall 2018.) A large survey of 17-year-old boys in the USA included the question "Do you work out with weights at least twice a week?" 71% of the respondents said yes. Assume that this proportion applies to the full population of 17-year-old boys in the USA.

   (a) Among a sample of 10 boys, what is the probability that at least 8 of them work out with weights at least twice a week?

   (b) Among a sample of 31 boys, what is the probability that exactly 22 (71% of 31) of them work out with weights at least twice a week?

   (c) Among a sample of 121 boys, what is the approximate probability that at least 91 of them work out with weights at least twice a week?

   (d) Now suppose I tell you that the sample in part (c) actually came from Germany, i.e. not part of the original survey, and that the sample reported exactly 91 boys (of the 121) who do work out with weights at least twice a week. Does this sample provide evidence that 17-year-old boys in Germany have different exercise habits from those in the USA? State carefully what assumptions you are making and how you reached your conclusion.

4. (From Midterm 2, Fall 2019.) In the final stages of international soccer competitions such as the World Cup, if the game is tied after 120 minutes of play, it has to be decided by a "penalty shoot-out" in which each team gets five penalty shots (if the game is still tied, there are additional penalties until it is settled). According to an analysis some years ago by the polling company FiveThirtyEight, in penalty shoot-outs in World Cup competitions, the probability of success on a single penalty is 0.715.

   (a) What distribution would apply if we wanted to know the number of penalties scored by a particular team in their five attempts? What assumptions does this involve?

   (b) In a game between England and Colombia in the 2018 Men's World Cup, the game was decided on penalties, and England won 4-3. Under the assumptions you stated in (a), what is the probability of *exactly* this score?

   (c) In an effort to improve his team's penalty scoring performance, the England manager arranges some training sessions where the team has to take 50 penalties (total). If the probability of success on a single penalty does not change, what is the probaility that the team scores on at least 40 of the 50 penalty kicks? State any additional assumptions you make.

   (d) Some commentators have noted that when penalties occur in regulation play (when one team commits a foul in the penalty area) the percentage of successful penalties is higher than that in a penalty shoot-out. Why do you think this happens and in what respect might this information make you reconsider your answers to parts (a)–(c)?

5. (From Final, Fall 2019.) In a game played at the NC State Fair, players are invited to throw a baseball at a set of three blocks on a table; if they knock all three blocks of the table, they win a prize. Past results suggest that the probability of winning a prize on a single play of the game is about 0.12.

   (a) Jenny plays the game six times. What is the probability that she wins at least twice?

   (b) Peter decides to keep playing until he wins once. What are the mean and standard deviation of the number of throws he requires?

   (c) Over the course of a two-hour session, the game is played 150 times. What is the probability that at least 25 of those games result in winning a prize?

6. (From Final, Fall 2019.) A woman claims to be able to tell the difference between Coke and Pepsi in a blind tasting. To test this, 60 unlabelled cups are set out, 30 containing Coke and the other 30 containing Pepsi. Each one, she takes a sip and then states which of Coke or Pepsi she thinks it is. If she truly cannot tell the difference, she should get the right answer 50% of the time. In fact, she is successful in 37 guesses (61.7%).

   (a) Construct a 99% confidence interval for p, the true probability of success in a single trial. What assumptions are involved here?

   (b) Carry out a formal test of the null hypothesis $H_0: \ p = 0.5$ against the alternative $H_A: \ p > 0.5$ What is the p-value in this case? What conclusion do you draw?

(c) Suppose we require that the confidence interval in part (a) has margin of error less than 0.1. What sample size would be required for this?

7. (Additional problem, added March 19, 2024.) An urn contains 12 white balls and 6 black balls. A sample of 7 balls is drawn at random (without replacement). What is the probability that the number of white balls in the sample is (a) 0, (b) 1, (c) 2, (d) 3, (e) 4, (f) 5, (g) 6, (h) 7? In all cases, round off your answer to 4 decimal places.

SAMPLE SOLUTIONS

Note: As with any exam questions of mine, there is no unique "right answer". This is intended as a guideline: it doesn't mean that other answers are necessRarily wrong.

1. If X is the number of baskets scored, then the distribution of X is binomial with $n = 41,\ p = 0.76$. So by the formulas for mean and SD, $\mu = 41 \times 0.76 = 31.16,\ \sigma = \sqrt{41 \times 0.76 \times 0.24} = 2.735$.

   Part (a): calculate $P(X < 27)$ and subtract from 1. With $x = 26.5$ (continuity correction), we have $z = \frac{26.5-31.16}{2.735} = -1.70$, the corresponding normal probability is 0.0446. Therefore, the answer is $1 - 0.0446 = 0.9554$.

   Part (b): calculate $P(X \leq 27)$ and subtract from 1. With $x = 27.5$ (continuity correction), we have $z = \frac{27.5-31.16}{2.735} = -1.34$ the corresponding normal probability is 0.0901. Therefore, the answer is $1 - 0.0901 = 0.9099$.

   Part (c): the difference between the answers to (a) and (b), which is 0.0455, accounts for the probability that X exactly equals 27. [Although it wasn't part of the required answer that you calculate this, you could also work it out as $\frac{41!}{27! \times 14!} \times 0.76^{27} \times 0.24^{14} = 0.0449$, not exactly the same but close enough to validate use of the normal approximation.]

2. (a) The probabilities must add to 1, so $x = 0.2$.

   (b) 0.5, 0.4, 0.2, 0.7. We have $P(A \text{ and B}) = 0.2 = 0.5 \times 0.4 = \text{P(A)} \times \text{P(B)}$, so A and B *are* independent.

   Added after grading: Some students just *assumed* A and B were independent, which isn't correct. The event (A and B) occurs when the die shows 1 or 3; the probability of that is $0.1 + 0.1 = 0.2$. Apart from that, the most common error was to assume the eight sides were all equally likely, which leads to a quite different set of answers.

   (c) Possible combinations are three 8s (total 24), or two 8s and one 7 (total 23). The latter could be arranged one of three ways ((8,8,7) or (8.7.8) or (7,8,8)). The corresponding probabilities are 0.2^3 and 3×0.2^3 so the answer is $4 \times 0.2^3 = 4 \times 0.008 = 0.032$.

   (d) For $n = 100$, the standard deviation of $\bar{x}$ is $\frac{2.39}{\sqrt{100}} = \frac{2.39}{10} = 0.0239$, so for a normal probability calculation, $z = \frac{5.5-5.1}{0.239} = 1.67$. The left tail probability is 0.9525 and therefore the right tail probability is $1 - 0.9525 = 0.0475$. This relies on the Central Limit Theorem which is valid in this case because $n > 30$. (Quite a few students forgot to divide by 10 to get the standard deviation of $\bar{x}$, the generic formula for which is $\frac{\sigma}{\sqrt{n}}$.)

3. (a) $\binom{10}{8}(0.71)^8(0.29)^2 + \binom{10}{9}(0.71)^9(0.29) + (0.71)^{10} = 0.2444 + 0.1330 + 0.0326 = 0.410$. Here, as with all these questions, small deviations from the stated answer due to rounding error will not be penalized. Note that $\binom{10}{8} = \frac{10!}{8!2!} = \frac{10\times9}{2} = 45$, while $\binom{10}{9} = 10$.

   (b) $\binom{31}{22}(0.71)^{22}(0.29)^9 = 20,160,075 \times (0.71)^{22} \times (0.29)^9 = 0.1562$. *Side note.* $\binom{31}{22} = \frac{31!}{22!9!}$. You should be able to evaluate the factorials on your calculator but the alternative way to do it is $\frac{31\times30\times29\times28\times27\times26\times25\times24\times23}{9\times8\times7\times6\times5\times4\times3\times2\times1} = \frac{31\times30\times29\times28\times27\times26\times25\times24\times23}{362,880} = 20,160,075$.

(c) Use the normal approximation to the binomial distribution (preferably, with the continuity correction). The mean and standard deviation are $\mu = 121 \times 0.71 = 85.91$ and $\sigma = \sqrt{121 \times 0.71 \times 0.29} = 4.99$ so for a normal probability calculation (for "greater than or equal to 91", which we also interpret as "greater than 90.5") we have $z = \frac{90.5-85.91}{4.99} = 0.92$. The left tail probability is 0.8212 and therefore the right tail probability (which we want) is $1 - 0.8212 = 0.1788$. *Alternative solutions:* Some students misinterpreted the problem as *exactly 91* rather than *at least 91* — if you work that out using the binomial formula, the answer is 0.04897694. I have partial credit for that though it wasn't the answer wanted. At least one student tried to use the binomial formula for each of 91, 92,...,121 and add them all up to get the answer to the question asked, but that's a lot of work to do on a pocket calculator and in fact the student concerned did not get the right answer. However, if anyone wants to know, the exact answer (computed in R — pbinom(90.5,121,.71,lower.tail=F)) is 0.1794476.

(d) This is a hypothesis testing problem. Suppose p is the proportion of all 17-year-old boys in Germany who work out with weights at least twice a week. We want to test $H_0:\ p = 0.71$ against $H_A:\ p \neq 0.71$. The null hypothesis corresponds to the proportion in Germany being exactly the same as in the USA. Here, 91 is the number of boys in the sample who say yes to the question asked (whether they work out with weights at least twice a week). Under H_0, the mean of that quantity is 85.91. Since $91 > 85.91$, the one-sided p-value is the probability that the number of positive responses is 91 or greater when the null hypothesis is true. By the answer to (c), that probability is approximately 0.1788. Since it's a two-sided test, the two-sided p-value is twice that, or 0.3576. This is clearly not a statistically significant conclusion so we conclude there is no evidence that 17-year-old boys in Germany have different exercise habits from 17-year-old boys in the USA.

4. (a) Binomial distribution assumes (i) number of penalty kicks is fixed (5), (ii) they are independent, (iii) the probability of success is the same. Part (iii) might be more questionable because different players have different skill levels for penalty kicks.

(b) Note that this is the combination of two probabilities — 4/5 for England and 3/5 for Colombia. The combined probability is $\binom{5}{4} \cdot \binom{5}{3} \cdot (0.715)^7 \cdot (0.285)^3 = 0.1106$.

Some students interpreted the question that there were a total of seven penalty kicks and that four of them were "won" by England (this is not the way penalty kicks work — the two teams take turns to shoot at the goal.) If you did it this way, the answer would be $\binom{7}{4} \times 0.715^4 \times 0.285^3 = 35 \times 0.2614 \times 0.0231 = 0.211$. This answer got partial credit.

(c) Use normal approximation with continuity correction: $\mu = 50 \times 0.715 = 35.75, \sigma = \sqrt{50 \times 0.715 \times 0.285} = 3.192, x = 39.5$ (note the question says "at least 40", not "more than 40", so we subtract 0.5). So $z = \frac{39.5-35.75}{3.192} = 1.17$ to two decimal places. From the normal table, the left-tail probability is 0.879 so the right-tail probability is $1 - 0.879 = 0.121$.

Alternative solutions: (i) The exact probability using the binomial distribution is 0.1179; (ii) Without continuity correction: $z = \frac{40-35.75}{3.192} = 1.33$ — would lead to a probability 0.0918. Quite a few students interpreted the question as "exactly 40" instead of "at least 40"; in that case the answer would be $\binom{50}{40} \times 0.715^{40} \times 0.285^{10} = 0.054$. I gave partial credit for that solution.

(d) This question was (intentionally) more varied in the range of acceptable answers. The most common one by far was "effect of stress" or some equivalent wording, implying that players are more likely to miss a penalty in a high-stakes penalty shootout than in regular play. I consider that an acceptable answer but I'm not personally convinced — it sounds to me rather like the "hot hand" theory in basketball, a theory that many people subscribe to but for which direct evidence is lacking. A couple of students mentioned the following scenario (you have to know a bit more about the rules of soccer to think of this); in a penalty shootout, if the penalty is missed or saved the play ends immediately (the shooter doesn't get a second chance) whereas if a penalty is missed or saved in regular play, and it bounces back on to the field, the ball is still in play and could be picked up by one of the offensive team for a goal. I don't actually think that's the correct explanation, but it's interesting. The third possible explanation is that in regulation play, the team tends to choose their best penalty kicker to take the kick, resulting in higher success rate. Nobody mentioned the possibility of reporting bias (i.e. there isn't really any difference in the probability of success), but that could also be an issue. Effects on your other answers: except for the last one, each of these explanations implies the probability of success is not constant, violating one of the assumptions of the binomial distribution.

5. (a) Binomial distribution: the probability Jenny wins either 0 times or 1 times are respectively $0.88^6 = 0.4644$ and $6 \times 0.88^5 \times 0.12 = 0.3800$. Therefore, the probability she wins at least twice is $1 - 0.4644 - 0.3800 = 0.1556$.

   (b) This is a geometric distribution with $p = 0.12$; the mean is $\mu = \frac{1}{p} = 8.3333$ and the standard deviation is $\sigma = \sqrt{\frac{1-p}{p^2}} = 7.817$,

   (c) With $n = 150,\ p = 0.12$, we have $np = 18 > 10,\ n(1-p) = 132 > 10$ so the normal approximation to the binomial distribution is applicable. The mean number of success is $n \times p = 18$ and the standard deviation is $\sqrt{n \times p \times (1-p)} = 3.9799$. Using the continuity correction, we calculate $z = \frac{24.5-18}{3.9799} = 1.63$; the left-tail probability associated with that is 0.9484. Therefore, the answer is (approximately) $1 - 0.9484 = 0.0516$. Two comments here. Some students wrongly interpreted the questions to mean the probability of exactly 25 successes. In that case, the answer would be $\binom{150}{25} 0.12^{25} 0.88^{125}$ which comes to 0.02145. Some students did give that exact answer, but since it's answering the wrong question, I only gave partial credit. The other common error was that students used the normal approximation but forgot about the continuity correction. Given the number of times we have mentioned this in the course, I felt it was appropriate to require that for full credit. Partial credit if you used the normal approximation but not the continuity correction.

6. (a) The standard error is $\sqrt{\frac{0.617 \times 0.383}{60}} = 0.0628$ so the 99% confidence interval is $0.617 \pm 2.58 \times 0.0628 = 0.617 \pm 0.162 = (0.455, 0.779)$.

   (b) For a hypothesis test we assume $p = 0.5$, so the standard error is $\sqrt{\frac{0.5 \times 0.5}{60}} = 0.0645$ and the z score for the test statistic is $\frac{0.117}{0.0645} = 1.814$. The one-sided p-value is $1 - 0.9649 = 0.0351$ which is "strong but not overwhelming" evidence that the null hypothesis is false. So the data do, to some extent, support the woman's claim. Here with $n = 60,\ p = 0.5$, both $np = n(1-p) = 30 > 10$, so the normal approximation to the binomial distribution is satisfied.

(c) In this case we require that $\sqrt{\frac{0.5\times 0.5}{n}} \times 2.58 < 0.1$: this leads to an n of at least 167. Note that we assume the default $p = 0.5$ in questions of this nature because we do not know what the true value of p will be in the future experiment.

7. Hypergeometric distribution: the probability of k white balls in the sample is $\frac{\binom{12}{k}\times\binom{6}{7-k}}{\binom{18}{7}}$. For $k = 0$, the answer to (a) is 0 (simplest way to see this: if there are 0 white balls in the sample the there must be 7 black balls in the sample, but there are only 6 black balls in the urn, so this is impossible). For the remaining values of k, first evaluate $\binom{18}{7} = \frac{18!}{7!\times 11!} = \frac{18\times 17\times 16\times 15\times 14\times 13\times 12}{7\times 6\times 5\times 4\times 3\times 2\times 1} = 31824$. Then we have:

(b) (b) $k = 1$: $\frac{\binom{12}{1}\times\binom{6}{6}}{\binom{18}{7}} = \frac{12\times 1}{31824} = 0.0004$,

(c) $k = 2$: $\frac{\binom{12}{2}\times\binom{6}{5}}{\binom{18}{7}} = \frac{66\times 6}{31824} = 0.0124$,

(d) $k = 3$: $\frac{\binom{12}{3}\times\binom{6}{4}}{\binom{18}{7}} = \frac{220\times 15}{31824} = 0.1037$,

(e) $k = 4$: $\frac{\binom{12}{4}\times\binom{6}{3}}{\binom{18}{7}} = \frac{495\times 20}{31824} = 0.3111$,

(f) $k = 5$: $\frac{\binom{12}{5}\times\binom{6}{2}}{\binom{18}{7}} = \frac{792\times 15}{31824} = 0.3733$,

(g) $k = 6$: $\frac{\binom{12}{6}\times\binom{6}{1}}{\binom{18}{7}} = \frac{924\times 6}{31824} = 0.1742$,

(h) $k = 7$: $\frac{\binom{12}{7}\times\binom{6}{0}}{\binom{18}{7}} = \frac{792\times 1}{31824} = 0.0249$.