MATH1180
Homework 6
Due Friday, March 17, 2017

Please (re)read Chapter 4 of the Baglivo textbook.

Please submit solutions to the following problems. Use 4 decimal places of accuracy for probabilities unless a problem says otherwise.

When submitting homework, please remember the following:

- Show all work leading to each solution.
- *Staple* all sheets together. A paper clip is not acceptable.
- Do not submit crossed-out or sloppy work.
- Do not submit ripped or torn pages.
- Be sure to submit your own work.

PROBLEM 1. A standard deck of 52 playing cards has 12 face cards and 40 other cards. A poker hand is a subset of 5 cards from the standard deck. Let X be the number of face cards in a poker hand. Assume that each choice of poker hand is equally likely.

(*a*) Write the formula you would use to find $P(X = x)$. Your answer will be an expression in x.

(*b*) Find the probability that the poker hand contains exactly 2 face cards.

(*c*) Find the mean and standard deviation of X. Interpret the mean in the context of this problem.

PROBLEM 2. A committee of 10 politicians consists of 7 Democrats and 3 Republicans. A subcommittee is chosen at random from these 10 politicians (meaning that the subcommittee is chosen in such a way that each possible choice is equally likely).

(*a*) Suppose that the subcommittee consists of 3 politicians. What is the probability that the subcommittee has at least 1 Republican and at least 1 Democrat?

(*b*) Suppose that the subcommittee consists of 5 politicians. What is the probability that the subcommittee has at least 1 Republican and at least 2 Democrats?

PROBLEM 3 (BM). Each child born to a couple with a particular set of genetic characteristics has probability 0.35 of having blood type O. Suppose that a couple with the given characteristics has 5 children (with no multiple births), and let X be the number of children with blood type O.

(*a*) Write the formula you would use to find $P(X = x)$. Your answer will be an expression in x.

(*b*) Complete the following probability distribution table; use 4 decimal places of accuracy.

x	0	1	2	3	4	5
$P(X = x)$	???	???	???	???	???	???

Which event is the most likely event? That is, for which value of x is $P(X = x)$ the largest?

(c) Find $P(X = 1 \mid X \leqslant 2)$, and interpret your answer in the context of this problem.

(d) Find $E(X)$ and $SD(X)$. Interpret the mean in the context of this problem.

PROBLEM 4 (AF). For the following random variables, explain why at least one condition needed for a binomial distribution or for a binomial approximation is unlikely to be satisfied.

(a) Let X be the number of people in a family of size 4 who go to church on a given Sunday, when any one of them goes 50% of the time in the long run. NOTE: If all conditions held, then a binomial distribution with $n = 4$ and $p = 0.50$ would be used.

(b) Let X be the number of people voting for the Democratic candidate out of the 100 votes in the first precinct that reports results, when 60% of the population voted for the Democrat. NOTE: If all conditions held, then a binomial distribution with $n = 100$ and $p = 0.60$ would be used.

(c) Let X be the number of females in a simple random sample of size 4 students from a class of size 20, when half the class is female. NOTE: If all conditions held, then a binomial distribution with $n = 4$ and $p = 0.50$ would be used.

PROBLEM 5. A group of 35 college students is working on an important community project. The group has 10 members who are left-handed, and the rest are right-handed. A radio station wishes to interview 5 of the students, and the faculty advisor chooses the 5 students randomly. Let X be the number of left-handed students chosen.

(a) Write the formula you would use to find $P(X = x)$. Your answer will be an expression in x.

(b) Complete the following probability distribution table; use 4 decimal places of accuracy.

x	0	1	2	3	4	5
$P(X = x)$	???	???	???	???	???	???

(c) Find $E(X)$ and $SD(X)$.

(d) For which value of x is $P(X = x)$ largest?

PROBLEM 6 (PG). According to the Behavioral Risk Factor Surveillance System, 65.0% of low income adult Americans adhere to a sedentary lifestyle as compared to 48.5% of high income adult Americans.

(a) Let X be the number of individuals who adhere to a sedentary lifestyle in a simple random sample of 6 adults chosen from the subpopulation of *low income* Americans.

   (i) Find $E(X)$ and $SD(X)$. Interpret the mean in the context of this problem.

   (ii) Find the probability that at least 3 of the 6 individuals adheres to a sedentary lifestyle.

(*b*) Repeat part (*a*) where X is the number of individuals who adhere to a sedentary lifestyle in a simple random sample of 6 adults chosen from the subpopulation of *high income* Americans.

PROBLEM 7 (BM). A rapid test for the presence in the blood of antibodies to HIV, the virus that causes AIDS, gives a positive result with probability 0.0011 when a person who is free of HIV antibodies is tested. Suppose that a clinic tests 2000 people who are free of HIV antibodies.

(*a*) Find the probability there are at most two "false positives" among these 2000 individuals.

(*b*) Find the probability that there are at least four "false positives" among these 2000 individuals.

PROBLEM 8 (BR). The National Center for Health Statistics reports that 4 in 1000 elderly drivers nationwide are involved in motor vehicle accidents in a given year.

(*a*) In a certain community, 6 of 500 elderly drivers who take diabetes medications regularly were involved in motor vehicle accidents within a one year period. Find the probability of seeing 6 or more accidents among seniors taking diabetes medications if the national rate holds for this subpopulation.

(*b*) In a certain community, 6 of 850 elderly drivers who take arthritis medication regularly were involved in motor vehicle accidents within a one year period. Find the probability of seeing 6 or more accidents among seniors taking arthritis medication if the national rate holds for this subpopulation.

(*c*) Epidemiologists say that there is an *excess risk* in a subpopulation if the probability of seeing at least that many events (in this case, motor vehicle accidents in a given year) is less than 0.10. Using this definition, would epidemiologists assign excess risk of motor vehicle accidents to the subpopulation of seniors who take medications regularly for either diabetes or arthritis?

PROBLEM 9 (BM). Typhoid fever is a life-threatening bacterial illness caused by *Salmonella typhi* and transmitted by ingestion of contaminated food and drink. On average, the Centers for Disease Control (CDC) receives reports of 1.15 cases of typhoid fever per day from all over the United States, although most cases were acquired while traveling internationally.

(*a*) Find the probability that more than 3 cases are reported to the CDC in a single day.

(*b*) Find the probability that between 4 and 6 cases are reported to the CDC in a 5-day period.

PROBLEM 10 (NW). According to a recent report from the Centers for Disease Control and Prevention, the infant mortality rate in the United States is 6.1 deaths per 1000 live births. (Thus, the probability that an infant dies is $\frac{6.1}{1000} = 0.0061$.)

Let X be the number of infant deaths in 500 randomly selected live births in the United States.

(a) Use a binomial model with $n = 500$ and $p = 0.0061$ to complete the following table:

	$P(X = 0)$	$P(X = 1)$	$P(X = 2)$	$P(X = 3)$	$P(X \leqslant 3)$
Binomial Probabilities:	???	???	???	???	???

Use 6 decimal places of accuracy throughout.

(b) Use a Poisson approximate model with $\lambda = np$ to complete the following table:

	$P(X = 0)$	$P(X = 1)$	$P(X = 2)$	$P(X = 3)$	$P(X \leqslant 3)$
Poisson Probabilities:	???	???	???	???	???

Use 6 decimal places of accuracy throughout.

(c) Are your answers to parts (a) and (b) substantially different?

NOTE: *The Washington Post* called the rate reported here a "national embarrassment" in a September 2014 blog post because the U.S. is ranked 27th among the wealthy nations of the world. Finland is ranked first, with the smallest infant mortality rate: 2.3 deaths per 1000 live births.

PROBLEM 11 (SK). According to recent Federal Bureau of Investigation (FBI) statistics, there are 3.27 bank robberies per day, on average, in the Southern Region of the United States.

(a) Find the probability that 4 or fewer bank robberies occur on a given day.

(b) Find the probability that no bank robberies occur in a two-day period.

(c) Assume that more than 3 bank robberies occur on a given day. What is the chance that exactly 4 occur on that day?

PROBLEM 12 (RG). In the imaginary city of Artinia, the average daily high temperature this past winter was $22.3°$, with a standard deviation of $4.5°$. The publicity director of Artinia claimed that during the 90 winter days, there were at most 60 days in which the temperature never went above freezing ($32°$). Why do you know that the publicity director is lying?

PROBLEM 13 (RG). Suppose that the prevalence of a disease, $P(D)$, is 0.23. Suppose that the predictive value of a positive test, $P(D|Pos)$, is 0.87, and the predictive value of a negative test, $P(D^c|Neg)$, is 0.96.

(a) What is $P(D|Neg)$?

(b) You now have enough information to compute $P(Pos)$ and $P(Neg)$. What are those values? HINT: It is helpful to use the equation $P(Pos) + P(Neg) = 1$.

(c) What is are the sensitivity $P(Pos|D)$ and specificity $P(Neg|D^c)$?