**Stat 468/568 Biostatistics**

**HW 2 (Chapter 4), Total points 15**

**Due 02/01/2022**

**Last Name: First Name:**

**Student class:** Graduate/Undergrad (select one as applies)

***Special Instruction: No extension of the due date is allowed to any particular student. However, if an extension is required, whole class will be allowed an extension. Start doing any homework early to allow significant amount of time to work on it depending on your need.***

(1) [Graduate Students Only]. Let the random variable follow a Binomial distribution with parameters and , i.e., . Prove that

(a) the expected value of X, .

(b) the variance of , where

(**Hints**: There are many ways to do it. The easiest one is to consider that Binomial distribution is the sum of the i.i.d Bernoulli random variables)

Proof:

1) Let be i.i.d. Bernoulli with success probability p.

Then, and

Note that if , then X=. Then,

(a)

(b)

Alternative proof:

(a) By definition, we have

, since , discussed in the class

Because by the binomial theory . Can you verify, how?

(b) To prove result (b), we use the fact that

and (both were discussed in the class)

**Now,**

Since by comparing with the binomial theory , with and. Can you verify, how?

So,

Then,

(2) Let be a discrete random variable that has the probability function for . Show that

where is called the cumulative distribution function (CDF) of the random variable . (**Hints**: You can just start from the definition of and then expand it, and try to re-express terms as a CDF).

Proof:

Method 1: Starting on left hand side (LHS),

+

+

+

Method 2: Starting on the RHS,

(3) High blood pressure costs the nation $47.5 billion annually in direct medical expenses and $3.5 billion each year in lost productivity.A recent research shows that 30% American adults have high blood pressure. In other words, a randomly selected adult has probability of 0.3 to have high blood pressure. We randomly select 20 American adults. Let denote the number of adults with high blood pressure among the 20 randomly selected American adults.

1. Write out the probability density function of , i.e., specify .
2. What is the probability that 8 American adults will have a high blood pressure?
3. What is the probability that at most 8 American adults will have a high blood pressure?
4. Find the expected value of the number of people with high blood pressure.
5. Find the variance of the number of people with high blood pressure.

Solution (See the SAS output for b and c at the end).

We have X~B(n=20, p=0.3)

(a) p(x)= ; x=0,1,2,…,20

(b) p(8)= =pdf("binomial",**8**,**0.3**,**20**)=0.1144;

(c) = cdf("binomial",**8**,**0.3**,**20**)=0.88667;

(d) E(X)=np=20\*0.3=6

(e) V(X)=npq=20\*0.3\*0.7=4.2

SAS Code:

**data** hw2Chapt4SASol;

/\*pdf or cdf("binomial",x,p,n);\*/

bp8=pdf("binomial",**8**,**0.3**,**20**);

cF8=cdf("binomial",**8**,**0.3**,**20**);

**run**;

**proc** **print**;

title "Output for 3 (b): p(8) and (c): F(8)";

**run**;

|  |
| --- |
| Output for 3 (b): p(8) and (c): F(8) |

| **Obs** | **bp8** | **cF8** |
| --- | --- | --- |
| **1** | 0.11440 | 0.88667 |