Worksheet 7.4

Max's Lecture
MATH 55

July 22, 2019

Exercise A (example in book). Suppose 1 person in 100,000 has a particular rare disease for which their is a fairly accurate diagnostic test. This test is correct 99 percent of the time when given to a person who has the disease, and it is correct 99.5 percent of the time when given to a person at random who does not have the disease.

Given this information, can we find:

1. the probability that a person who tests positive for the disease has the disease?
2. the probability that a person who tests negative for the disease really does not have the disease?

Both of these are found on page 496-497 of the book.

Exercise B. Prove the theorem that expected value is linear, i.e. prove the following two facts if $X, X_1, X_2 \ldots, X_n$ are random variables on the same sample space and a, b are real numbers:

1. $E(X_1 + \cdots + X_n) = E(X_1) + \ldots E(X_n)$.
2. $E(aX + b) = aE(X) + b$.

1. We first show this for 2 random variables:

$$E(X_1+X_2) = \sum_{s\in S} p(s)(X_1(s)+X_2(s))$$

$$= \sum_{s\in S} p(s)\,X_1(s) + \sum_{s\in S} p(s)\,X_2(s)$$

$$= E(X_1)+E(X_2)$$

The case for n random variables follows from induction.

2. This also follows from the definition of expectation

$$E(aX+b) = \sum_{s\in S} p(s)(aX(s)+b)$$

$$= a\sum_{s\in S} p(s)X(s) + b\sum_{s\in S} p(s)$$

$$= a\,E(X)+b \quad \text{since } \sum_{s\in S} p(s)=1$$

Exercise C (example in book). Suppose you roll a pair of dice. Use the theorem in Exercise B to compute the expected value of their sum.

Let X be the sum of the dice, X_1 be the value of the first die, and X_2 be the value of the second die.

We can see that $X = X_1 + X_2$. Using the linearity of expectation, we first compute $E(X_1)$ and $E(X_2)$ and use this to find $E(X)$.

$$E(X_1) = 1\tfrac{1}{6} + 2\tfrac{1}{6} + 3\tfrac{1}{6} + 4\tfrac{1}{6} + 5\tfrac{1}{6} + 6\tfrac{1}{6} = \frac{7}{2}$$

$$E(X_2) = E(X_1) = \frac{7}{2}$$

So $E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = 7$.

Exercise D. 1. Suppose you run n Bernoulli trials, where the probability of success is p. What is the expected value of the number of success? Let X be # of successes

2. Suppose you keep repeating a Bernoulli trial until you get a success, and then stop after the first success where the probability of success is p. What is the expected number of trials you will do?

Note: The first part has you compute the expected value of a binomial distribution, and the second has you compute the expected value of a geometric distribution.

1. We will do this by using the linearity of expectation.

Let $X_i = \begin{cases} 1 & \text{if the } i^{th} \text{ trial is a success} \\ 0 & \text{if the } i^{th} \text{ trial is a failure.} \end{cases}$

We first ~~comp~~ compute $E(X_i) = p \cdot 1 + (1-p)\,0 = p.$

Since we can see that $X = X_1 + X_2 + \cdots + X_n$,

$$E(X) = \underbrace{E(X_1) + E(X_2) + \dots E(X_n)}_{\text{each is equal to } p} = \underbrace{p + p + \cdots + p}_{n \text{ times}} = np.$$

There is an alternate way to do this directly using the definition of expectation.

2. We will do this by using the definition of expectation.

First, $p(X=j) = (1-p)^{j-1}p$ since you must fail the first $j-1$ times and then succeed on the j^{th} time. So by the laws of expectation,

$$E(X) = \sum_{j=1}^{\infty} j \cdot p(X=j) = \sum_{j=1}^{\infty} j(1-p)^{j-1}p = p\underbrace{\sum_{j=1}^{\infty} j(1-p)^{j-1}}_{\substack{\text{In class} \\ \text{I told you} \\ \text{this sum is } 1/p^2.}} = p/p^2 = \frac{1}{p}.$$

Exercise E. 1. What is the variance of the random variable X with $X(t) = 1$ if Bernoulli trial is a sucess and and $X(t) = 0$ is the Bernoulli trial is a failure, where p is the probability of success and q is the probability of failure?

2. Use the above to find the variance of the number of successes in scenario 1 of exercise D.

1. $V(X) = E(X^2) - (E(X))^2$. We do the following computations:

$E(X) = 1 \cdot p + 0q = p$ (where $q = 1-p$)

$E(X^2) = 1^2 \cdot p + 0^2 q = p$

So $V(x) = p - p^2 = p(1-p) = pq$.

2. Since each trial is independent,

$V(Y)$ where Y is number of successes over the n trials, is

$$\underbrace{V(X) + \cdots + V(X)}_{n \text{ times}} = npq.$$

Exercise F. You do 100 bernoulli trials (with the chance of success being 1/2 for each trial). The random variable X records the number of successes. What is the expected value of X? What is an upper bound of the probability that $X(s)$ differs from the expected value by at least 25?

NOTE: The problem was incorrect on the version of the worksheet that I handed out in class. I want an upper bound, not the actual probability.

First: Since this is the binomial distribution we know $E(X) = np = 100 \cdot \frac{1}{2} = 50$

Another way of phrasing the probability it asks for.

$p(|X(s) - E(X)| \geq 25)$, which by Chebyshev inequality is

$$\leq \frac{V(X)}{25^2} = \frac{npq}{25^2} = \frac{100 \cdot \frac{1}{2} \cdot \frac{1}{2}}{25^2} = \frac{25}{25^2} = \frac{1}{25}$$

So $\frac{1}{25}$ is an upper bound on the probability