MTH 3210 Written Homework 4

Brady Lamson (blamson@msudenver.edu)

October 10, 2021

3.2 - Probability Distributions for Discrete Random Variables: Exercises: 12-13, 15-16, 18, 22-24

Definition. *For a given sample space S of some experiment, a* ***random variable (rv)*** *is any rule that associates a number with each outcome in S. In mathematical language, a random variable is a function whose domain is the sample space and whose range is the set of real numbers.*

Notation. *Random variables are customarily denoted by uppercase letters, such as X and Y, near the end of the alphabet. Example: $X(\omega) = x$ means that x is the value associated with the outcome ω by the rv X.*

Definition. *Any random variable whose only possible values are 0 and 1 is called a* ***Bernoulli random variable.***

Definition. *A* ***discrete*** *random variable is an rv whose possible values either constitute a finite set or else can be listed in an infinite sequence in which there is a first element, a second element, and so on ("countably" infinite). A random variable is* ***continuous*** *if both of the following apply:*

- *Its set of possible values consist either of all numbers in a single interval on the number line (possibly infinite in extent, eg., from $-\infty$ to ∞) or all numbers in a disjoint union of such intervals (e.g., $[0, 10] \cup [20, 30]$).*
- *No possible value of the variable has positive probability, that is, $P(X = c) = 0$ for any possible value c.*

Notation. *The* ***probability distribution*** *of X says how the total probability of 1 is distributed among the various possible X values. The notation for this will be:*
P(0) = the probability of the X value 0 = P(X = 0)
P(1) = the probability of the X value 1 = P(X = 1)

Definition. *The* ***probability distribution*** *or* ***probability mass function (pmf)*** *of a discrete rv is defined for every number x by $p(x) = p(X = x) = P(all\, \omega \in S : X(\omega) = x)$.*

In terms of words: For every possible value x of the random variable, the pmf specifies the probability of observing that value when the experiment is performed. The conditions, $p(x) \geq 0$ and $\sum_{all\ possible\ x} p(x) = 1$ are required of any pmf.

Problem 12.

Airlines sometimes overbook flights. Suppose that for a plane with 50 seats, 55 passengers have tickets. Define the random variable Y as the number of ticketed passengers who actually show up for the flight. The probability mass function of Y appears in the accompanying table.

y	P[Y = y]
45	.05
46	.10
47	.12
48	.14
49	.25
50	.17
51	.06
52	.05
53	.03
54	.02
55	.01

a. What is the probability that the flight will accommodate all ticketed passengers who show up?

For this problem simply add up the first 6 probabilities. This gives us $\boxed{.83}$.

b. What is the probability that not all ticketed passengers who show up can be accommodated?

For this we add up all probabilities associated with amounts greater than the planes cap. So we have $P[Y > 50] = \boxed{.17}$.

c. If you are the first person on the standby list (which means you will be the first one to get on the plane if there are any seats available after all ticketed passengers have been accommodated), what is the probability that you will be able to take the flight? What is this probability if you are the third person on the standby list?

(i) We can think of this situation as us being any passenger between 1 and 50. So our answer is the probability of at max 49 passengers showing up, which is $\boxed{.58}$

(ii) This one is much the same, instead we have a max of 47 passengers. Our result is $\boxed{.32}$.

Problem 13.

A mail-order computer business has six telephone lines. Let X denote the number of lines in use at a specified time. Suppose the pmf of X is as given in the accompanying table.

x	P[X = x]
0	.10
1	.15
2	.20
3	.25
4	.20
5	.06
6	.04

Calculate the following probabilities:

a. {At most three lines are in use.}
Add up $P[X = 0 : 3]$ inclusive. The result is $\boxed{.7}$

b. {Fewer than three lines are in use.}
For this add up the probabilities of 0, 1 and 2. This gives $\boxed{.45}$.

c. {At least three lines are in use.}
Subtract the above probability from 1. The result is $\boxed{.55}$.

d. {Between two and five lines, inclusive, are in use.}
Add up 2, 3, 4 and 5. The result is $\boxed{.71}$.

e. {Between two and four lines, inclusive, are NOT in use.}
We can reword this as 4, 3 and 2 lines being in use. To explain, if four lines are not in use that means two lines are still in use. Follow the logic for the rest. Calculating this probability gives us $\boxed{.65}$

f. {At least four lines are not in use.}
Using the same logic as in part (e) we can add up the probabilities for 0, 1 and 2 lines being in use. We get $\boxed{.45}$.

Problem 15.

Many manufacturers have quality control programs that include inspection of incoming materials for defects. Suppose a computer manufacturer receives circuit boards in batches of five. Two boards are selected from each batch for inspection. We can represent possible outcomes of the selection process by pairs. For example, the pair (1, 2) represents the selection of boards 1 and 2 for inspection.

a. List the ten different possible outcomes.

$$S = \left\{ \begin{array}{ccccc} (1,2), & (1,3), & (1,4), & (1,5), & (4,5), \\ (2,3), & (2,4), & (2,5), & (3,4), & (3,5) \end{array} \right\}$$

b. Suppose that boards 1 and 2 are the only defective boards in a batch. Two boards are to be chosen at random. Define X to be the number of defective boards observed among those inspected. Find the probability distribution of X.

$$P(X=0) = \{(3,4),(3,5),(4,5)\} = \frac{3}{10}$$

$$P(X=1) = \{(1,3),(1,4),(1,5),(2,3),(2,4),(2,5)\} = \frac{6}{10}$$

$$P(X=2) = \{(1,2)\} = \frac{1}{10}$$

c. Let $F(x)$ denote the cdf of X. First determine $F(0) = P(X \leq 0)$, $F(1)$, and $F(2)$; then obtain $F(x)$ for all other x.

$$F(0) = P(X \leq 0) = \boxed{.3}$$

$$F(1) = P(X \leq 1) = .3 + .6 = \boxed{.9}$$

$$F(2) = P(X \leq 2) = \boxed{1}$$

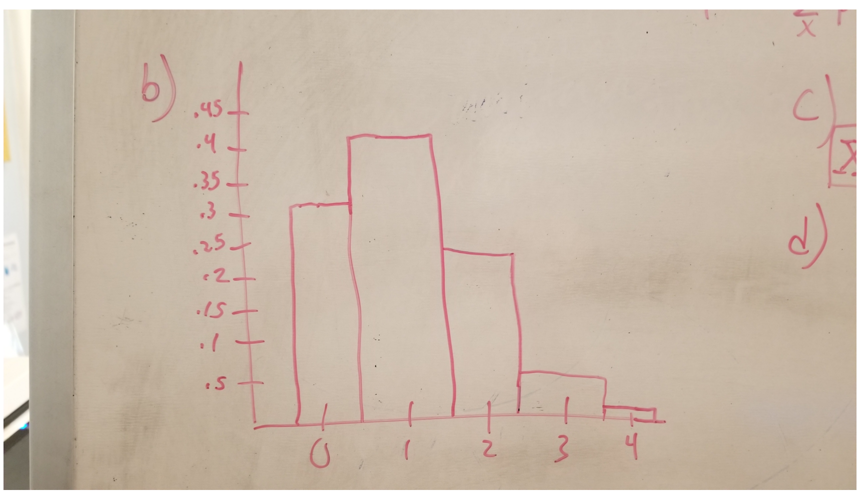
Problem 16.

Some parts of California are particularly earthquake prone. Suppose that in one metropolitan area, 25% of all homeowners are insured against earthquake damage. Four homeowners are to be selected at random; let X denote the number among the four who have earthquake insurance.

a. Find the probability distribution of X. *Hint:* Let S denote a homeowner who has insurance and F one who does not. Then one possible outcome is SFSS, with probability (.25)(.75)(.25)(.25) and associated X value 3. There are 15 other outcomes.

x	P[X = x]
0	$\left(\frac{3}{4}\right)^4 \approx .316$
1	$\binom{4}{1} \cdot \frac{1}{4} \cdot \left(\frac{3}{4}\right)^3 \approx .422$
2	$\binom{4}{2} \cdot \left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{4}\right)^2 \approx .211$
3	$\binom{4}{3} \cdot \left(\frac{1}{4}\right)^3 \cdot \left(\frac{3}{4}\right) \approx .047$
4	$\left(\frac{1}{4}\right)^4 \approx .0039$
	$\sum_x P[X = x] = .999$

b. Draw the corresponding probability histogram.



c. What is the most likely value for X?

$$\boxed{X = 1}$$

d. What is the probability that at least two of the four selected have earthquake insurance?

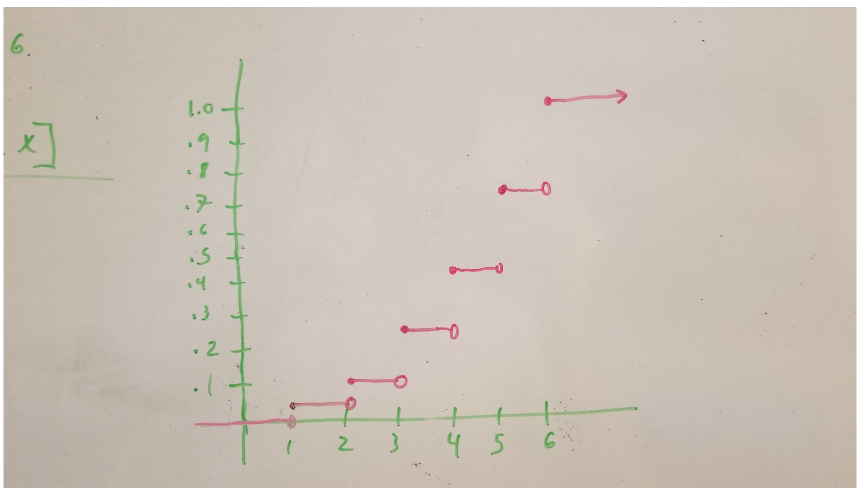
$$P[X \geq 2] \approx \boxed{.2619}$$

Problem 18.

Two fair six-sided dice are tossed independently. Let M = the maximum of the two tosses.

a. What is the pmf of M?
b. Determine the cdf of M and graph it.

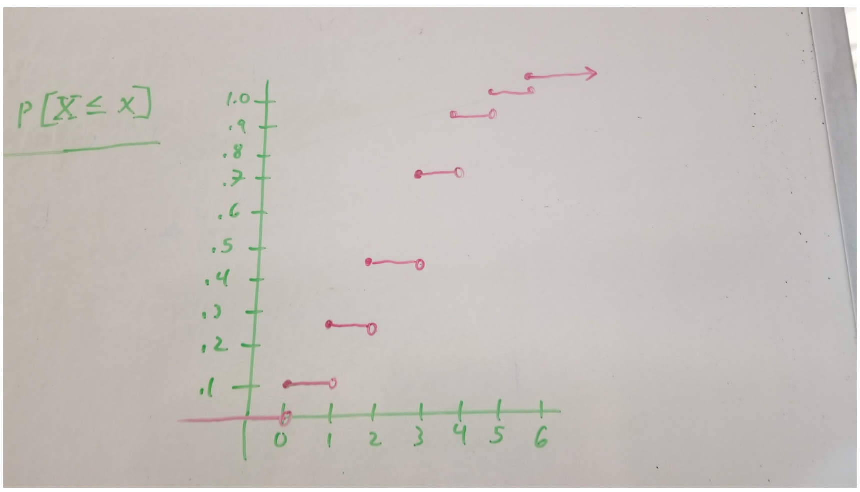
x	$P[M = x]$	$P[M \leq x]$
1	$\left(\frac{1}{36}\right) \approx .028$	$\frac{1}{36} \approx .028$
2	$\left(\frac{3}{36}\right) \approx .083$	$\frac{4}{36} \approx .111$
3	$\left(\frac{5}{36}\right) \approx .139$	$\frac{9}{36} = .25$
4	$\left(\frac{7}{36}\right) \approx .194$	$\frac{16}{36} \approx .444$
5	$\left(\frac{9}{36}\right) = .25$	$\frac{25}{36} \approx .694$
6	$\left(\frac{11}{36}\right) \approx .306$	$\frac{36}{36} = 1$



Problem 22.

Refer to exercise 13, and calculate and graph the cdf. Then use it to calculate the probabilities of the events given in parts (a)-(d) of that problem.

x	$P[X = x]$	$P[X \leq x]$
0	.1	.1
1	.15	.25
2	.2	.45
3	.25	.7
4	.2	.9
5	.06	.96
6	.04	1



a. $P[X \leq 3] = \boxed{.7}$

b. $P[X < 3] = P[X \leq 2] = \boxed{.45}$

c. $P[X \geq 3] = 1 - P[X \leq 3] = \boxed{.55}$

d. $P[2 \leq X \leq 5] = P[X \leq 5] - P[X \leq 1] = \boxed{.71}$

Problem 23.

A branch of a certain bank in New York City has six ATMs. Let X represent the number of machines in use at a particular time of day. The cdf of X is as follows.

$$F(x) = \begin{cases} 0 & X < 0 \\ .06 & 0 \leq X < 1 \\ .19 & 1 \leq X < 2 \\ .39 & 2 \leq X < 3 \\ .67 & 3 \leq X < 4 \\ .92 & 4 \leq X < 5 \\ 1 & 6 \leq X \end{cases}$$

a. $P(X = 2) = .39 - .19 = \boxed{.20}$

b. $P(X > 3) = P(6 \leq X) - P(3 \leq X \leq 4) = \boxed{.33}$

c. $P(2 \leq X \leq 5) = P(5 \leq X < 6) - P(1 \leq X < 2) = \boxed{.78}$

d. $P(2 < X < 5) = P(4 \leq X < 5) - P(2 \leq X < 3) = \boxed{.53}$

Problem 24.

An insurance company offers its policyholders a number of different premium payment options. For a randomly selected policyholder, let X = the number of months between successive payments. The cdf of X is as follows:

$$F(x) = \begin{cases} 0 & X < 1 \\ .3 & 1 \leq X < 3 \\ .4 & 3 \leq X < 4 \\ .45 & 4 \leq X < 6 \\ .6 & 6 \leq X < 12 \\ .1 & 12 \leq X \end{cases}$$

a. What is the pmf of X?

x	P[X = x]
1	.3
3	.1
4	.05
6	.15
12	.4

b. (i) $P(3 \leq X \leq 6) = P(6 \leq X \leq 12) - P(1 \leq X < 3) = \boxed{.3}$

(ii) $P(4 \leq X) = P(12 \leq X) - P(3 \leq X < 4) = \boxed{.6}$

3.3 - Expected Values:
Exercises: 29-32, 36, 44(a)(b)

Prob 29.

The pmf of the amount of memory X (GB) in a purchased flash drive was given in Example 3.13 as:

x	P[X = x]
1	.3
3	.1
4	.05
6	.15
12	.4

x	P[X = x]	$x \cdot P[X = x]$	$(x - \mu_X)^2 \cdot P[X = x]$
1	.05	.05	$(1 - 6.45)^2 \cdot .05 \approx 1.485$
2	.10	.20	$(2 - 6.45)^2 \cdot .1 \approx 1.98$
4	.35	1.4	$(4 - 6.45)^2 \cdot .35 \approx 2.1$
8	.40	3.2	$(8 - 6.45)^2 \cdot .4 \approx 0.961$
16	.10	1.6	$(16 - 6.45)^2 \cdot .1 \approx 9.12$
		$\mu_X = 6.45$	V(X) = 15.646

Compute the following:
a. $E(X) = \mu_X = \boxed{6.45}$

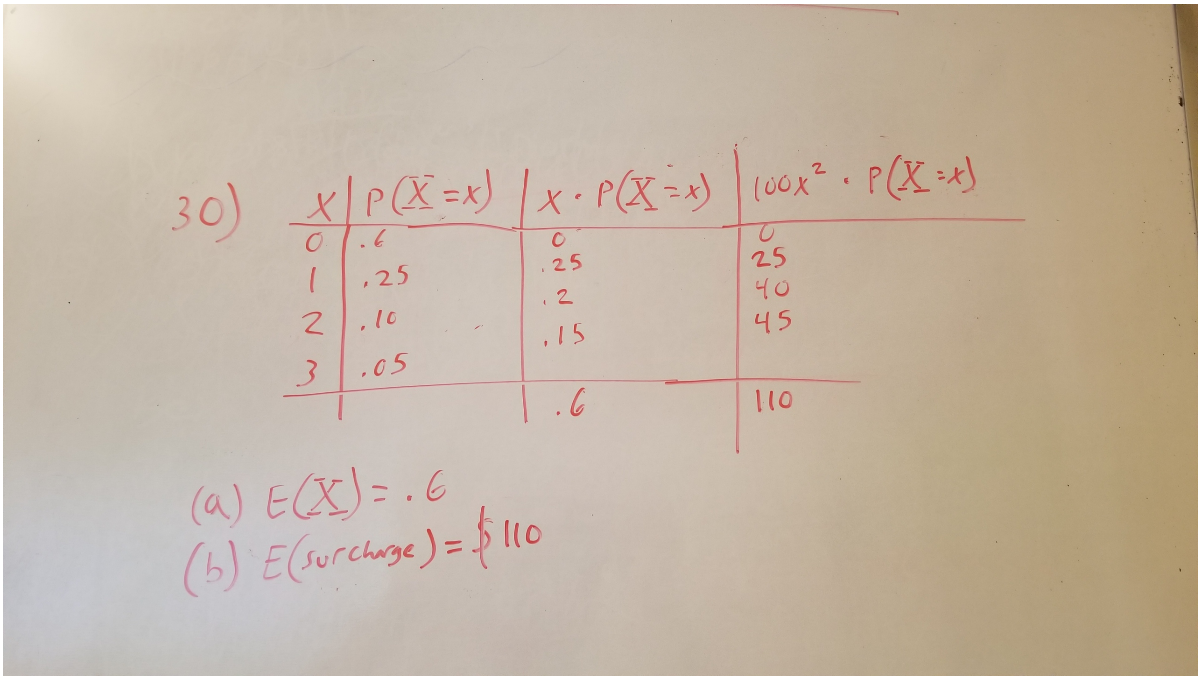
b. V(X) directly from the definition
$V(X) = \sum_x (x - \mu_X)^2 \cdot P[X = x] = \boxed{15.646}$

c. The standard deviation of X
$\sigma = \sqrt{V(X)} = \sqrt{15.646} \approx \boxed{3.96}$

d. V(X) using the shortcut formula
$V(X) = E(X^2) - [E(X)]^2 = 57.25 - 6.45^2 = \boxed{15.648}$

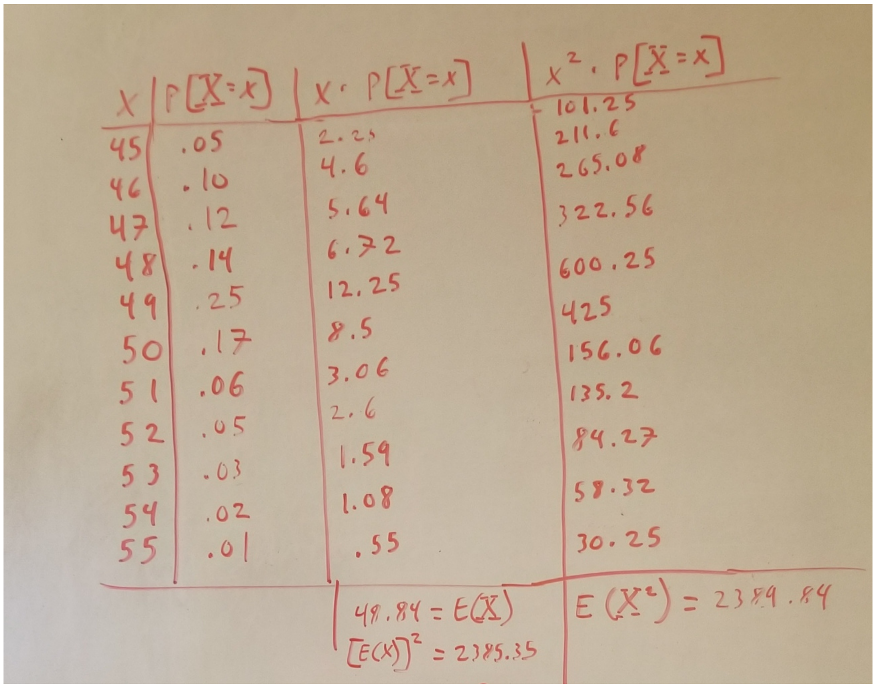
Problem 30.

An individual who has automobile insurance from a certain company is randomly selected. Let Y be the number of moving violations for which the individual was cited during the last 3 years.

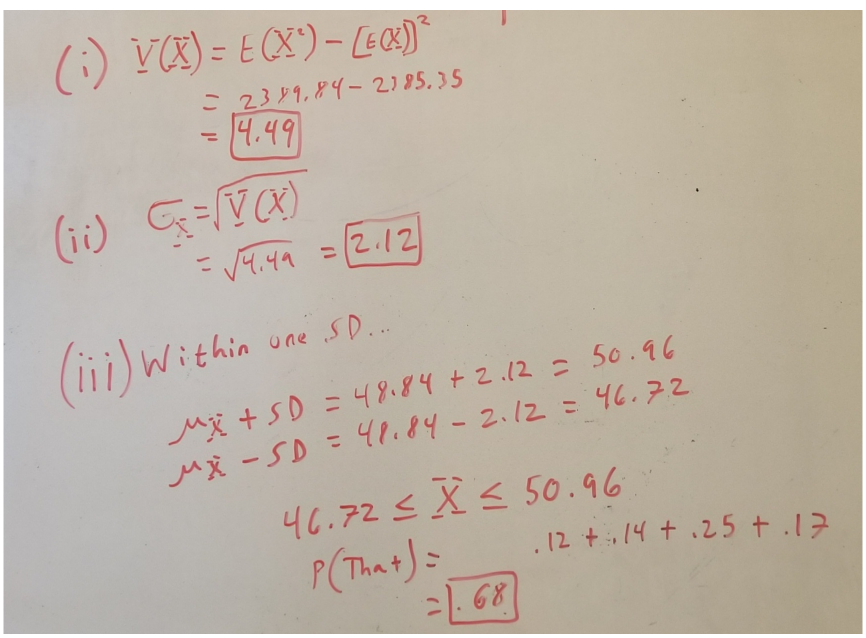


Problem 31.

Refer to Exercise 12 and calculate V(Y) and σ_Y. Then determine the probability that Y is within 1 standard deviation of its mean value.



x	$P[X=x]$	$x \cdot P[X=x]$	$x^2 \cdot P[X=x]$
45	.05	2.25	101.25
46	.10	4.6	211.6
47	.12	5.64	265.08
48	.14	6.72	322.56
49	.25	12.25	600.25
50	.17	8.5	425
51	.06	3.06	156.06
52	.05	2.6	135.2
53	.03	1.59	84.27
54	.02	1.08	58.32
55	.01	.55	30.25
		$48.84 = E(X)$ $[E(X)]^2 = 2385.35$	$E(X^2) = 2389.84$



(i) $V(X) = E(X^2) - [E(X)]^2$

$= 2389.84 - 2385.35$

$= \boxed{4.49}$

(ii) $\sigma_X = \sqrt{V(X)}$

$= \sqrt{4.49} = \boxed{2.12}$

(iii) Within one SD...

$\mu_X + SD = 48.84 + 2.12 = 50.96$

$\mu_X - SD = 48.84 - 2.12 = 46.72$

$46.72 \leq X \leq 50.96$

$P(\text{That}) = .12 + .14 + .25 + .17$

$= \boxed{.68}$

Problem 32.

A certain brand of upright freezer is available in three different rated capacities: $16ft^3, 18ft^3$ and $20ft^3$. Let X = the rated capacity of a freezer of this brand sold at a certain store. Suppose X has the pmf given below.

a. Compute $E(X), E(X^2)$, and $V(X)$

b. If the price of a freezer having capacity X is $70X - 650$, what is the expected price paid by the next customer to buy a freezer?

c. What is the variance of the price paid by the next customer?

d. Supose that although the rated capacity of a freezer is X, the actual capacity is $h(X) = X - .008X^2$. What is the expected actual capacity of the freezer purchased by the next customer?

32)

x	$P[X=x]$	$x \cdot P[X=x]$	$x^2 \cdot P[X=x]$	$(70x-650) \cdot P[X=x]$
16	.2	3.2	51.2	94
18	.5	9	162	305
20	.3	6	120	225
		$E(X) = 18.2$ $[E(X)]^2 = 331.24$	$E(X^2) = 333.2$	$E(\text{That}) = 624$ $[E(\text{That})]^2 = 389,376$

a) $V(X) = 333.2 - 331.24$
$= \boxed{1.96}$

b) $E(70x - 650) = \boxed{\$624}$

x	$(70x - 650)^2 \cdot P[X=x]$		$(x - .008x^2) \cdot P[X=x]$
16	44180		2.79
18	186050		7.7
20	168750		5.04
	$E(\text{This}) = 398980$		15.53

(d) $= \boxed{15.53}$

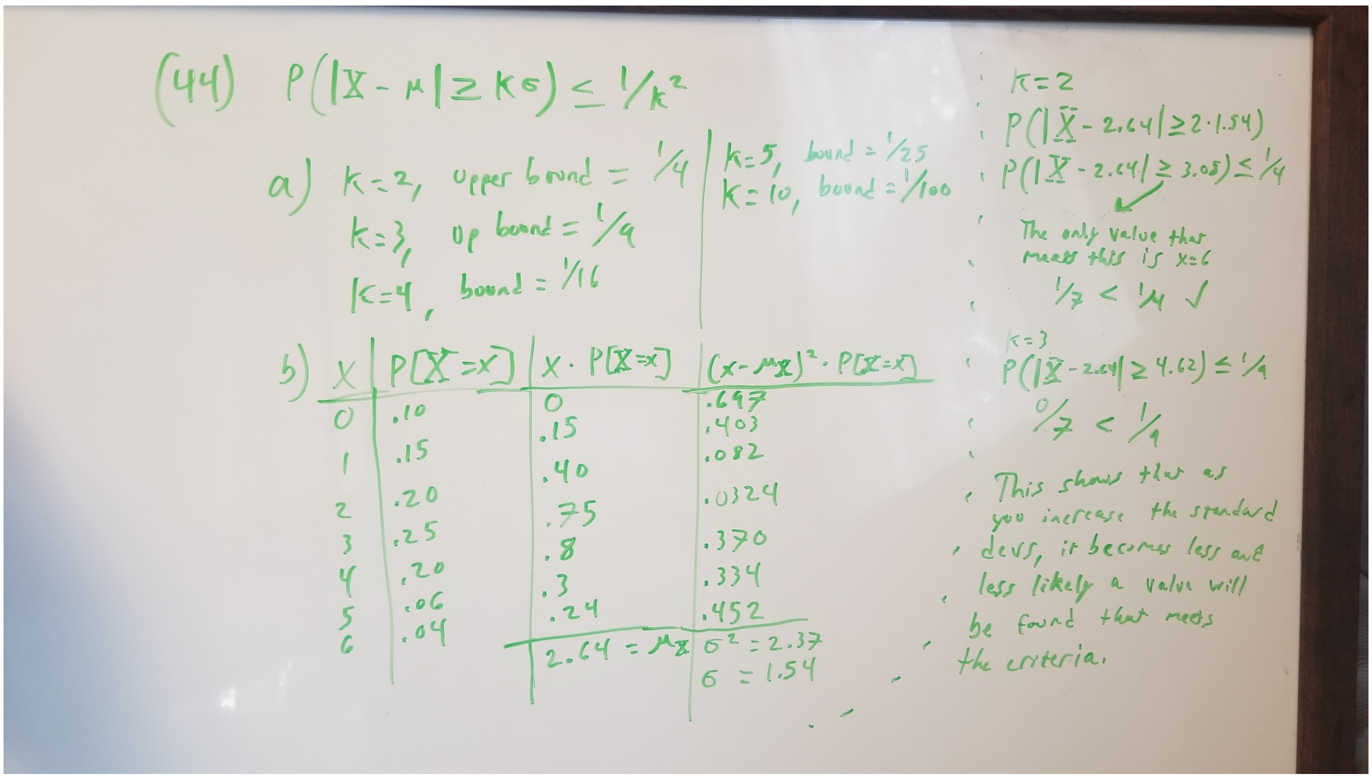
c) $V(\text{price}) = 398980 - 389376$
$= \boxed{9604}$

Problem 44. (a) (b)

A result called **Chebyshev's inequality** states that for any probability distribution of an rv X and any number k that is at least 1, $P(|X i \mu| \geq k\sigma) \leq \frac{1}{k^2}$. In words, the probability that the value of X lies at least k standard deviation from its mean is at most $\frac{1}{k^2}$.

a. What is the value of the upper bound for k = 2, 3, 4, 5 and 10?

b. Compute μ and σ for the distribution of Exercise 13. Then evaluate $P(|X i \mu| \geq k\sigma)$ for the values of k given in part (a). What does this suggest about the upper bound relative to the corresponding probability?



3.4 - The Binomial Probability Distribution:
Exercises: 46-49, 53-54

Problem 46.

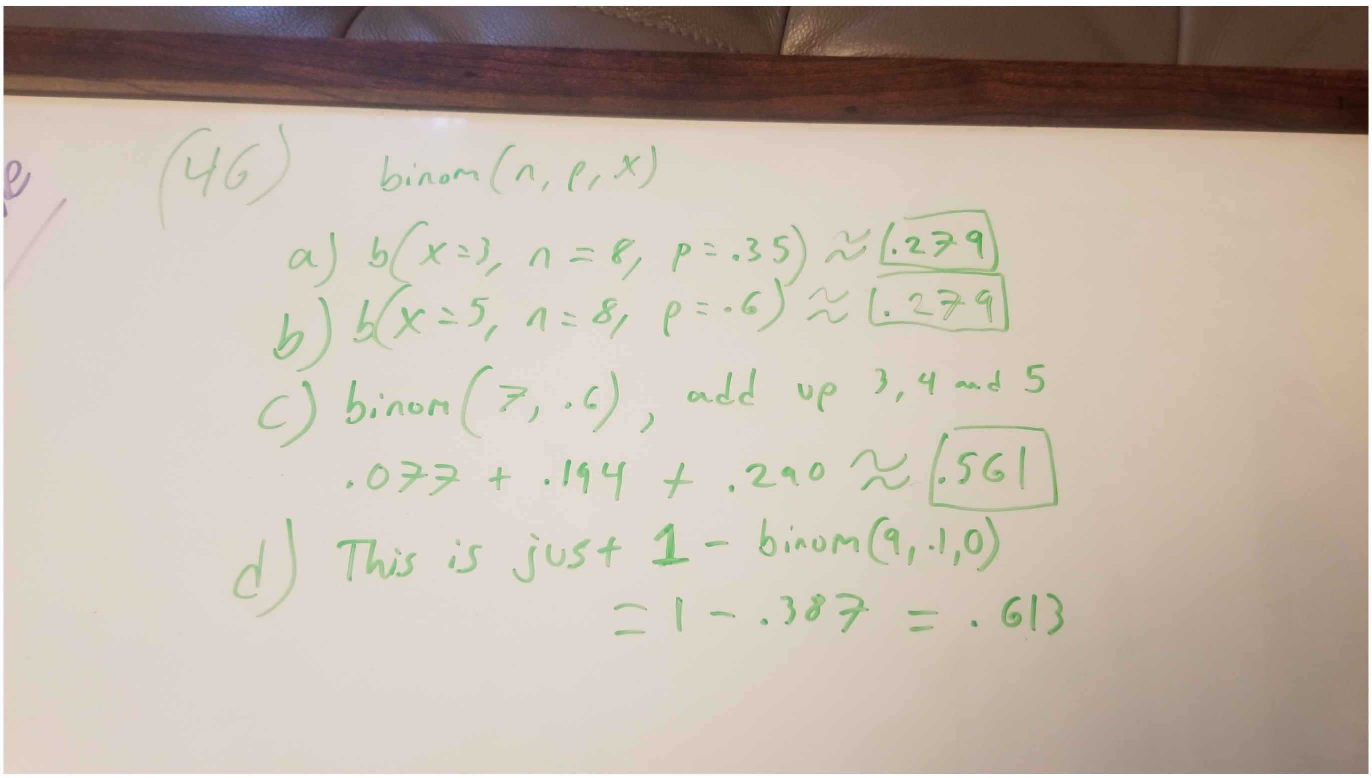
Calculate the following binomial probabilities directly from the formula for b(x; n, p):

a. $b(3; 8, .35)$

b. $b(5; 8, .6)$

c. $P(3 \leq X \leq 5)$ when n = 7 and p = .6

d. $P(1 \leq X)$ when n = 9 and p = .1



Problem 47

The article **"Should you report that fender-bender?"** (***Consumer Reports***), **Sept. 2013:15** reported that 7 in 10 auto accidents involve a single vehicle. Suppose 15 accidents are randomly selected. Use Appendix Table A.1 to answer each of the following questions.

a. What is the probability that at most 4 involve a single vehicle?

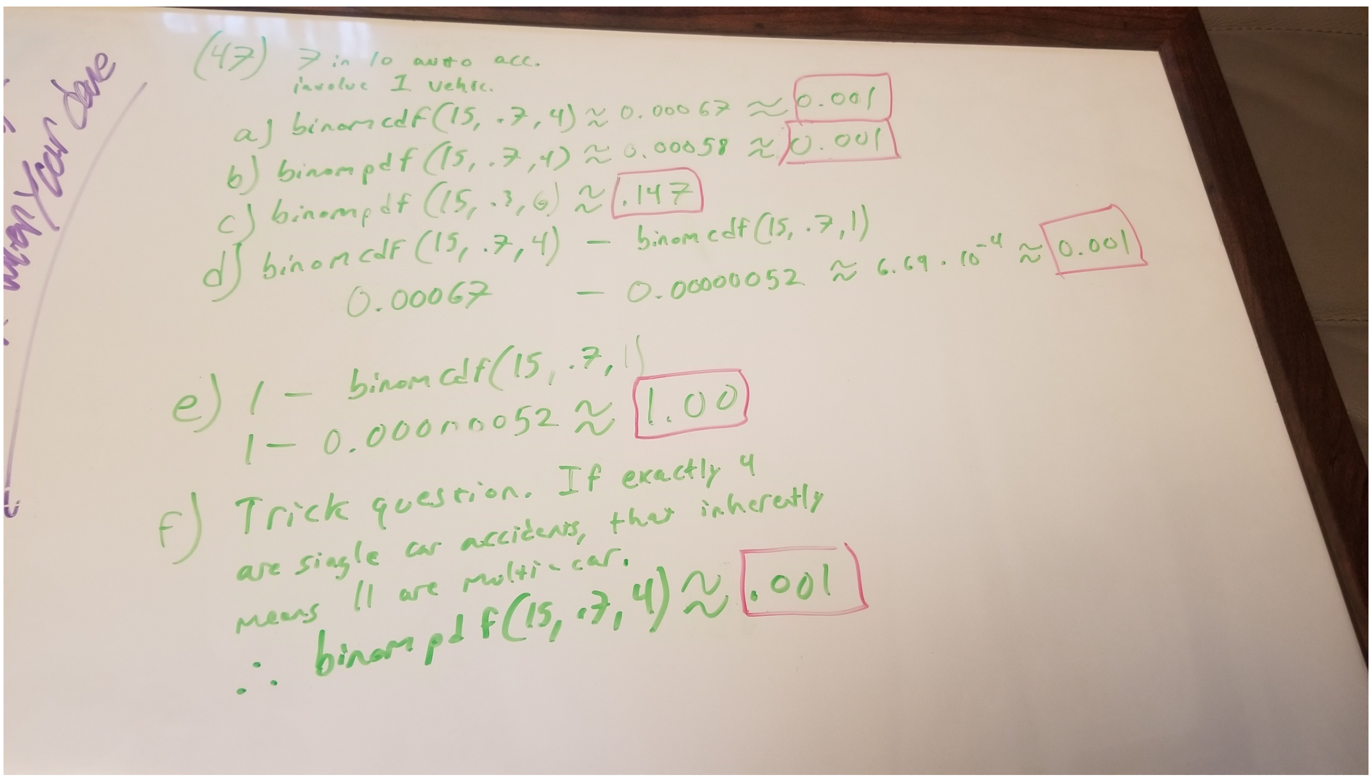
b. What is the probability that exactly 4 involve a single vehicle?

c. What is the probability that exactly 6 involve multiple vehicles?

d. What is the probability that between 2 and 4, inclusive, involve a single vehicle?

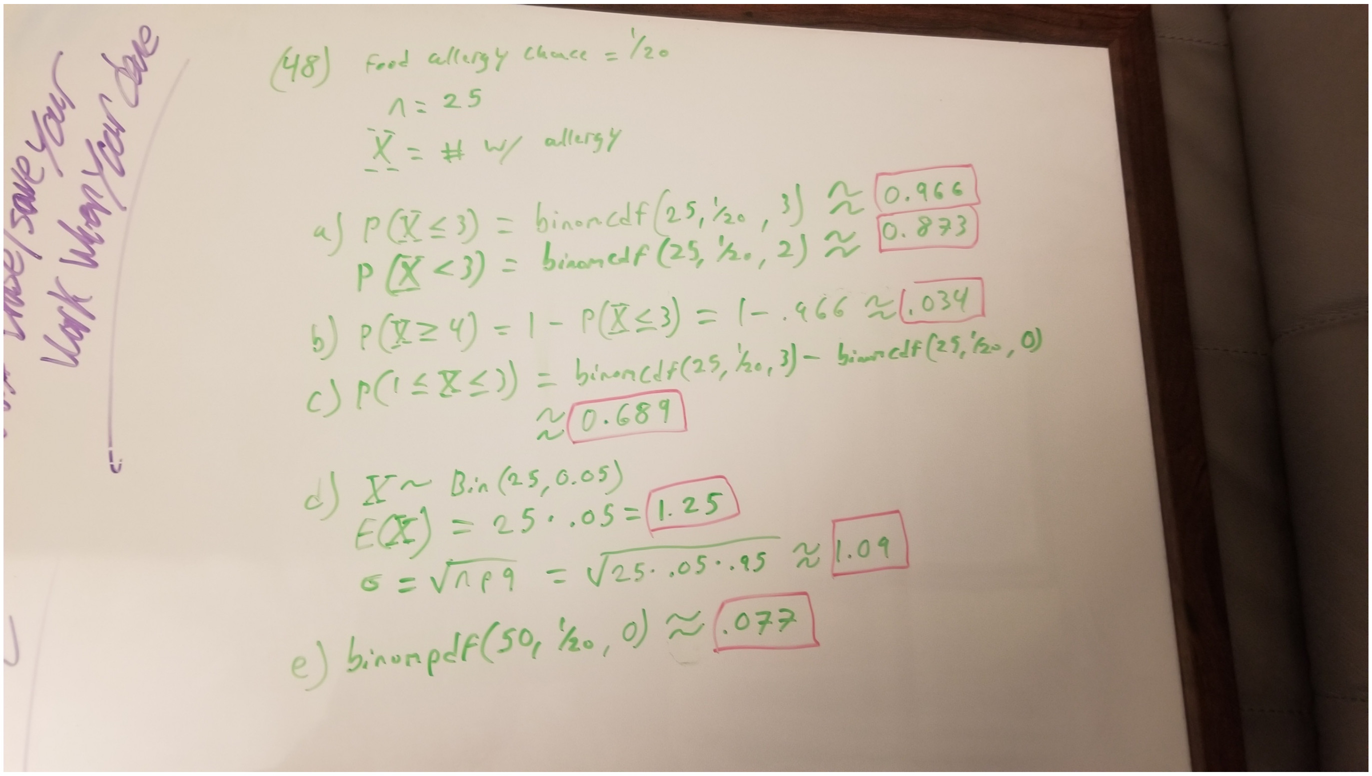
e. What is the probability that at least 2 involve a single vehicle?

f. What is the probability that exactly 4 involve a single vehicle and the other 11 involve multiple vehicles?



Problem 48.

NBC News reported on May 2, 2013, that 1 in 20 children in the United States have a food allergy of some sort. Consider selecting a random sample of 25 children and let X be the number in the sample who have a food allergy. Then $X \sim Bin(25, .05)$.



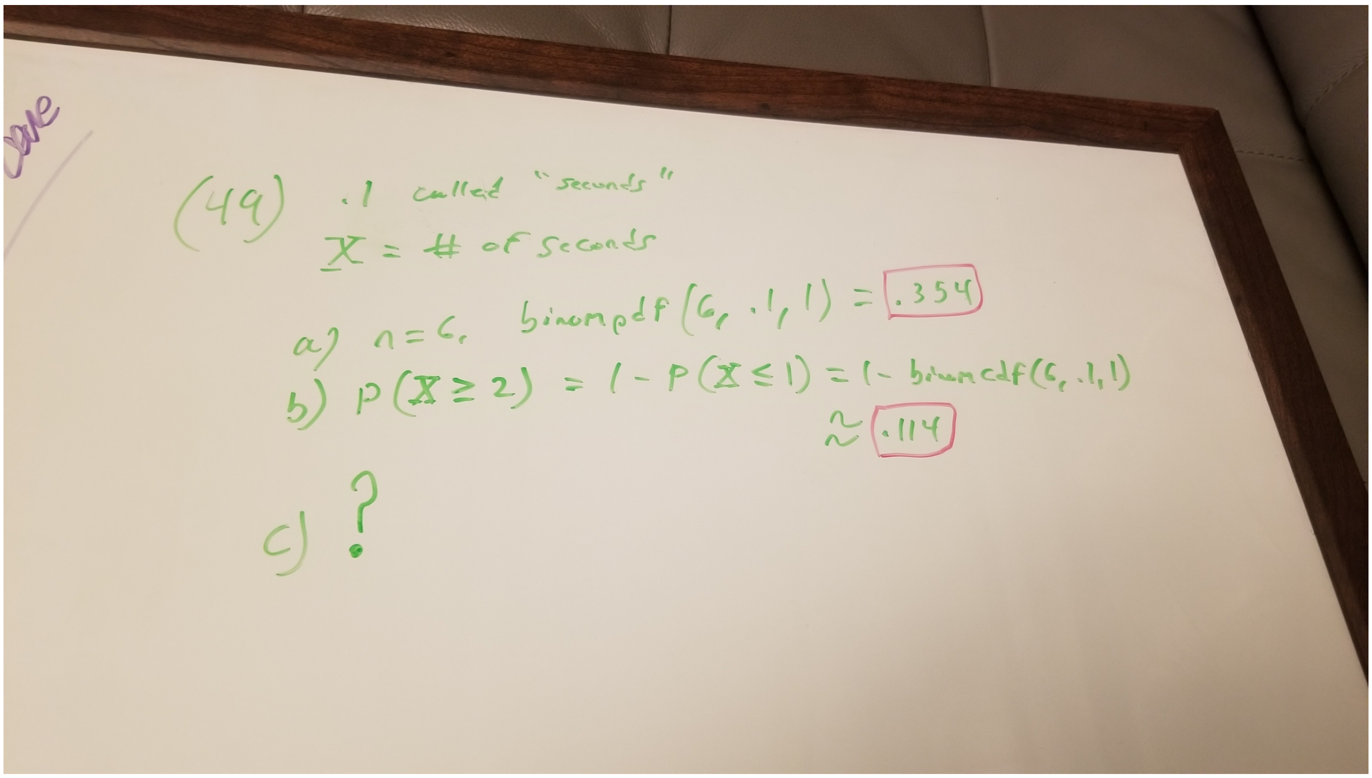
Problem 49.

A company that produces fine crystal knows from experience that 10% of its goblets have cosmetic flaws and must be classified as "seconds."

a. Among six randomly selected goblets, how likely is it that only one is a second?

b. Among six randomly selected goblets, what is the probability that at least two are seconds?

c. If goblets are examined one by one, what is the probability that at most five must be selected to find four that are not seconds?



(c) P(4 are all good) + P(first of the four is 'second') $\cdot$ .9

binompdf$(4, .9, 4)$ + binompdf(4,.9,3) $\cdot$ .9 $\approx$ $\boxed{.918}$

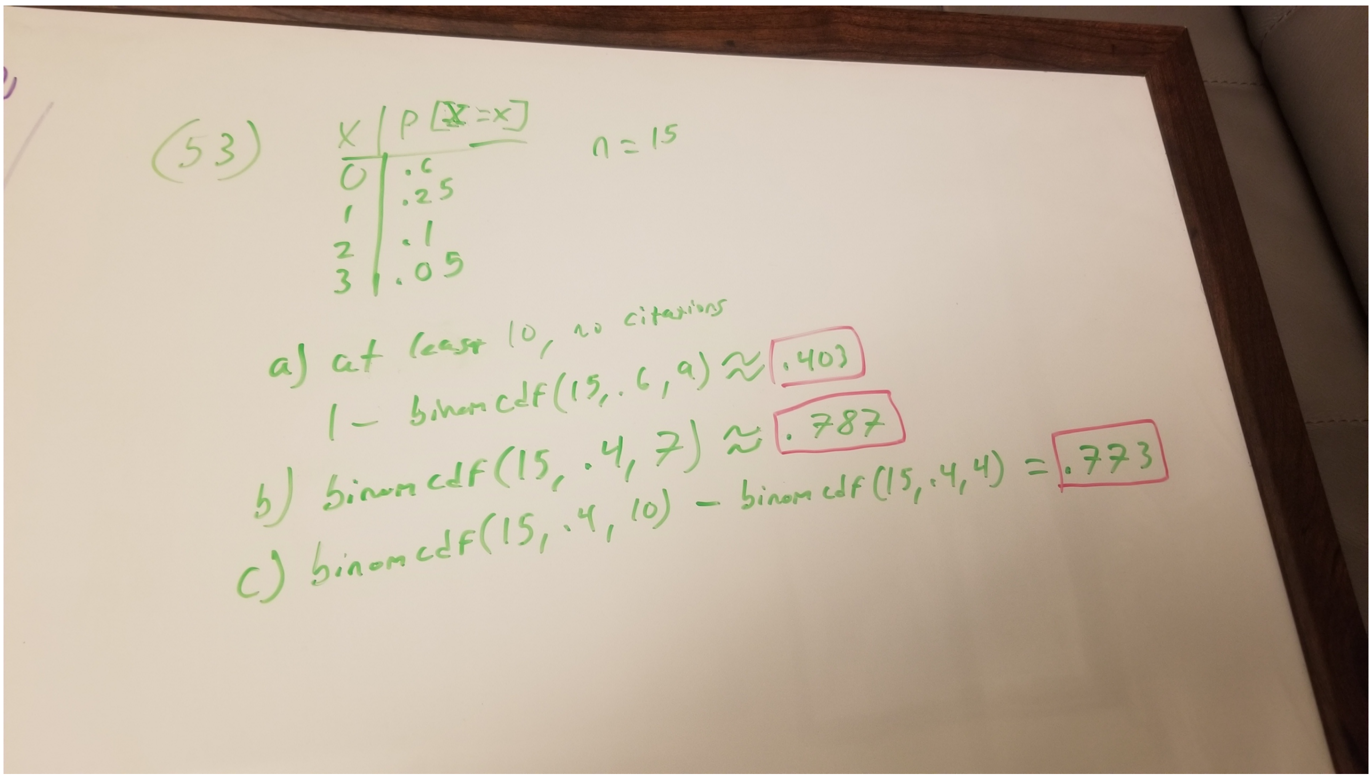
Problem 53

Exercise 30 (Section 3.3) gave the pmf of Y, the number of traffic citations for a randomly selected individual insured by a particular company. What is the probability that among 15 randomly chosen such individuals

a. At least 10 have no citations?

b. Fewer than half have at least one citation?

c. The number that have at least one citation is between 5 and 10, inclusive?



Problem 54

A particular type of tennis racket comes in a midsize version and an oversize version. Sixty percent of all customers at a certain store want the oversize version.

a. Among ten randomly selected customers who want this type of racket, what is the probability that at least six want the oversize version?

b. Among ten randomly selected customers, what is the probability that the number who want the oversize version is within 1 standard deviation of the mean value?

c. The store currently has seven rackets of each version. What is the probability that all of the next ten customers who want this racket can get the version they want from current stock?

(54) sizes = mid and over

$6/10$ want over

a) $X \sim B(10, .6)$

$P(X \geq 6)$

$1 - \text{binomcdf}(10, .6, 5) = \boxed{.633}$

b) $\mu_X = n \cdot p$

$= 10 \cdot .6$

$= 6$

$\sigma = \sqrt{n \cdot p \cdot (1-p)}$

$= \sqrt{10 \cdot .6 \cdot (1-.6)}$

$= \sqrt{6 \cdot .4}$

≈ 1.55

within 1 SD = $4.45 \leq x \leq 7.55$

floor the decimals.

$4 \leq x \leq 7$

$\text{binomcdf}(10, .6, 7) - \text{binomcdf}(10, .6, 3)$

$\approx \boxed{.780}$

3.6 - The Poisson Probability Distribution:
Exercises: 79-82

Problem 79.

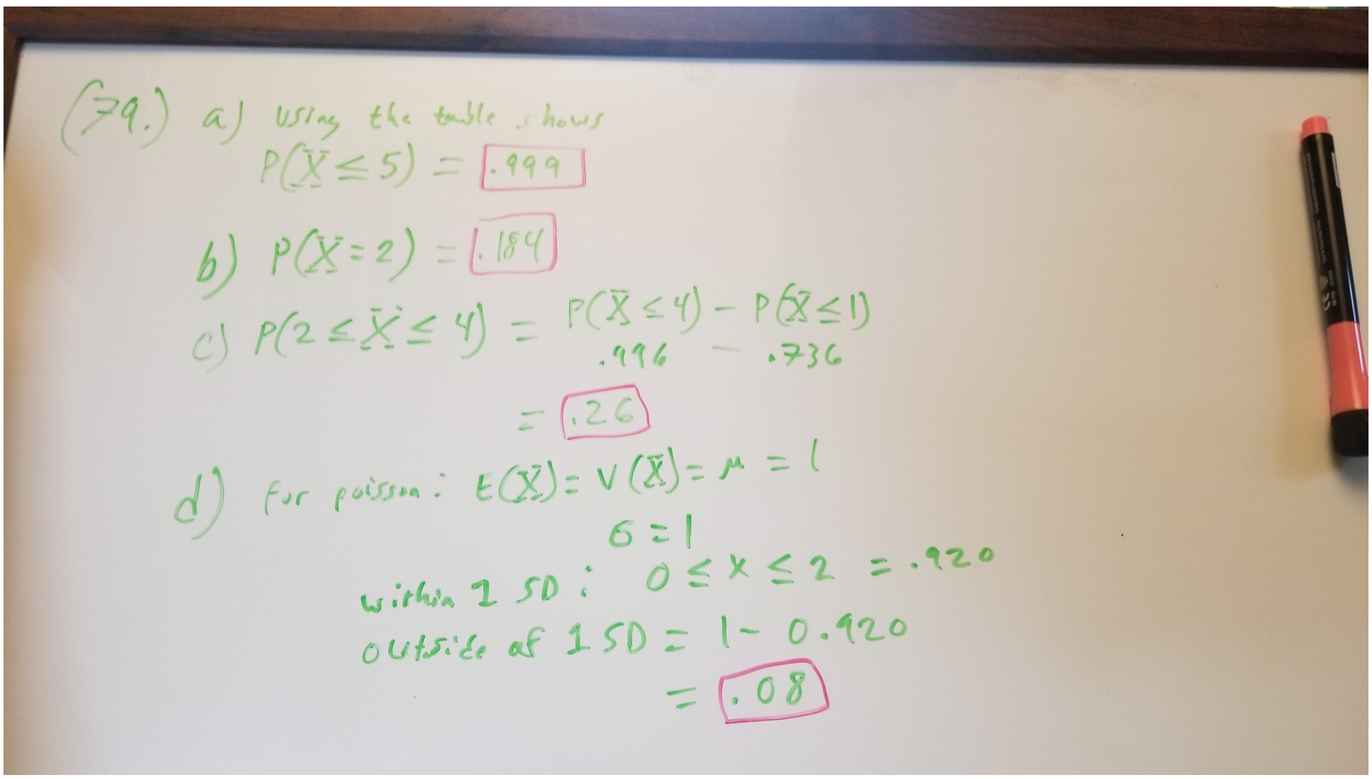
The article "Expectation Analysis of the Probability of Failure for Water Supply Pipes" (J. of Pipeline Systems Engr. and Practice, May 2012: 36–46) proposed using the Poisson distribution to model the number of failures in pipelines of various types. Suppose that for cast-iron pipe of a particular length, the expected number of failures is 1 (very close to one of the cases considered in the article). Then X, the number of failures, has a Poisson distribution with $\mu = 1$.

a. Obtain $P(X \leq 5)$ by using Appendix Table A.2.

b. Determine $P(X = 2)$ first from the pmf formula and then from Appendix Table A.2.

c. Determine $P(2 \leq X \leq 4)$.

d. What is the probability that X exceeds its mean value by more than one standard deviation?



Problem 80.

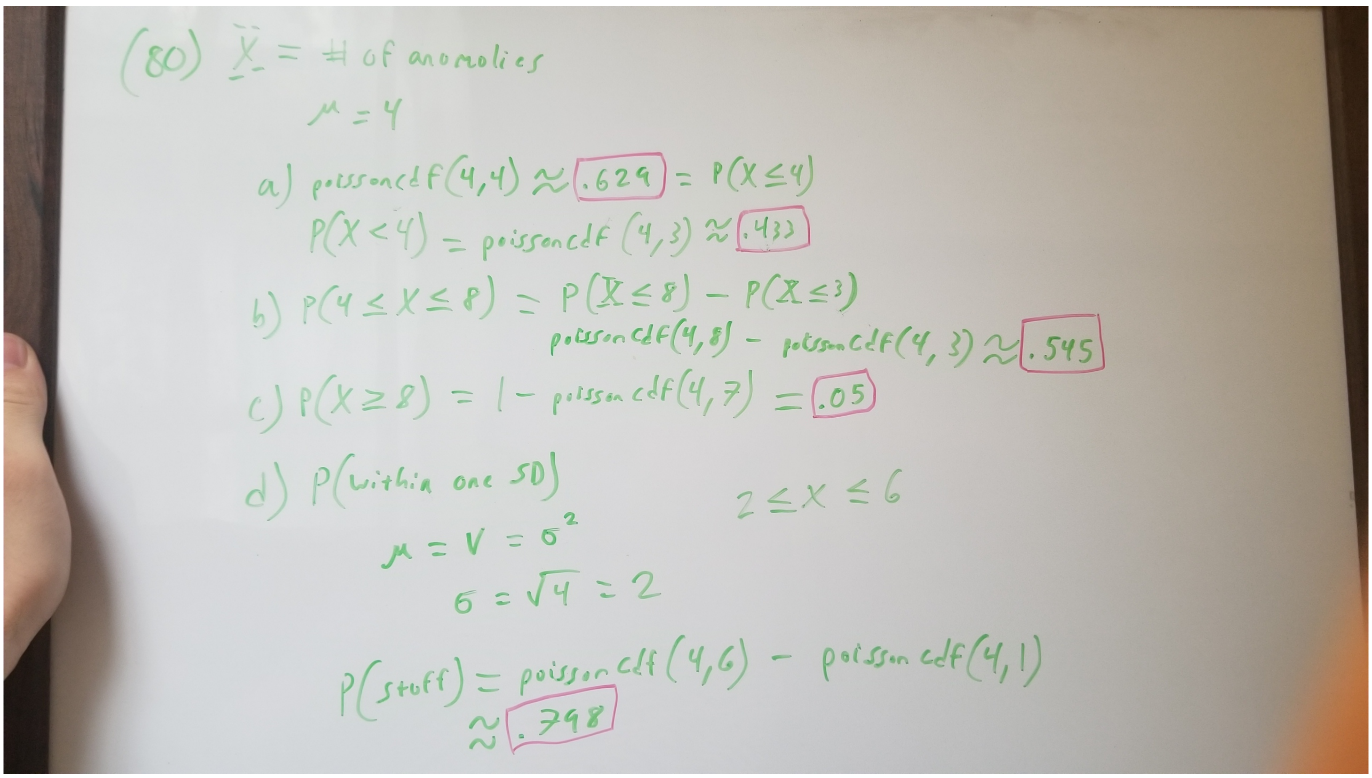
Let X be the number of material anomalies occurring in a particular region of an aircraft gas-turbine disk. The article "Methodology for Probabilistic Life Prediction of Multiple-Anomaly Materials" (Amer. Inst. of Aeronautics and Astronautics J., 2006: 787–793) proposes a Poisson distribution for X. Suppose that $\mu = 4$.

a. Compute both $P(X \leq 4)$ and $P(X < 4)$.

b. Compute $P(4 \leq X \leq 8)$.

c. Compute $P(8 \leq X)$.

d. What is the probability that the number of anomalies exceeds its mean value by no more than one standard deviation?



Problem 81.

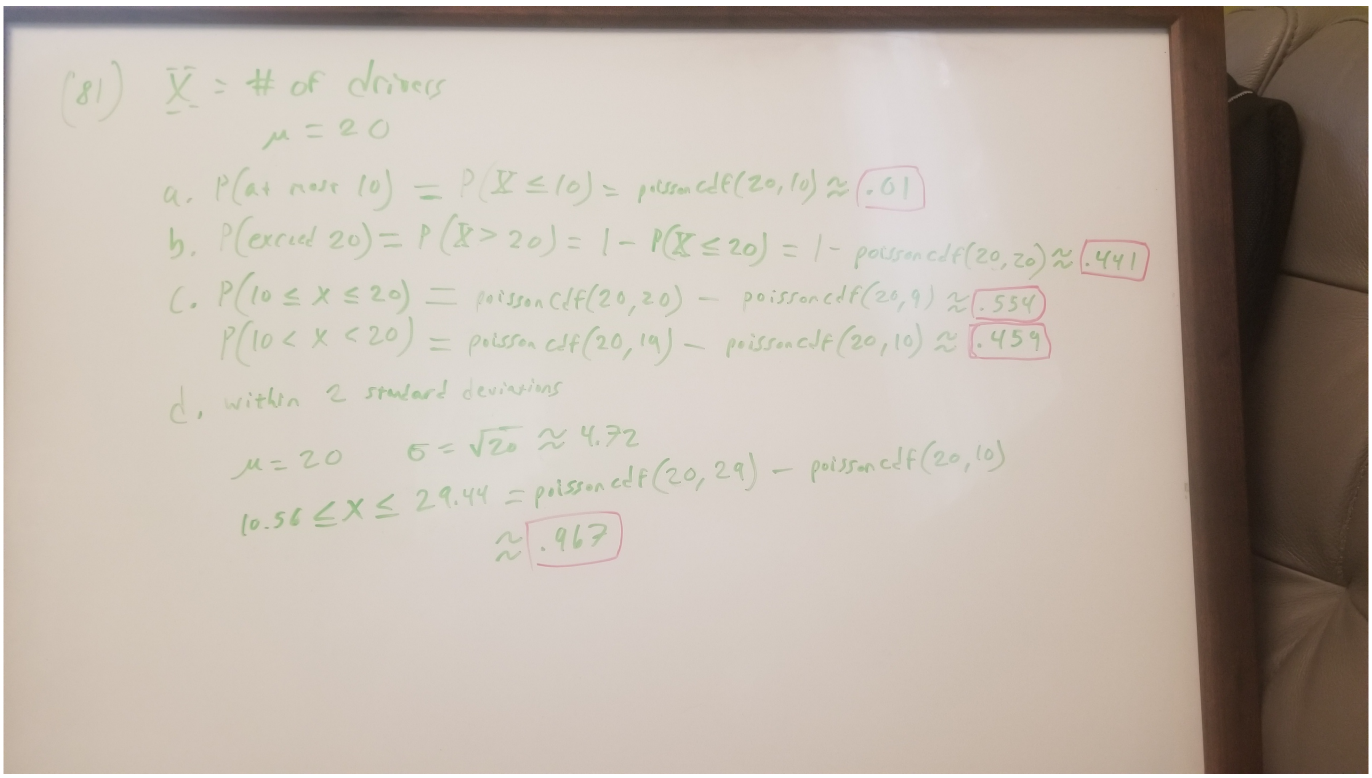
Suppose that the number of drivers who travel between a particular origin and destination during a designated time period has a Poisson distribution with parameter $\mu = 20$ (suggested in the article "Dynamic Ride Sharing: Theory and Practice," J. of Transp. Engr., 1997: 308–312). What is the probability that the number of drivers will

a. Be at most 10?

b. Exceed 20?

c. Be between 10 and 20, inclusive? Be strictly between 10 and 20?

d. Be within 2 standard deviations of the mean value?



Problem 82.

Consider writing onto a computer disk and then sending it through a certifier that counts the number of missing pulses. Suppose this number X has a Poisson distribution with parameter $\mu = .2$. (Suggested in "Average Sample Number for Semi-Curtailed Sampling Using the Poisson Distribution," J. Quality Technology, 1983: 126–129.)

a. What is the probability that a disk has exactly one missing pulse?

X = Number of missing pulses.

$$P(X = 1) = \text{poissonpdf}(.2, 1) \approx \boxed{.164}$$

b. What is the probability that a disk has at least two missing pulses?

$$P(X \leq 2) = 1 - P(X \leq 1) \approx \boxed{.018}$$

c. If two disks are independently selected, what is the probability that neither contains a missing pulse?

$$P(X = 0)^2 = (\text{poissonpdf}(.2, 0))^2 \approx \boxed{.67}$$