LECTURE NOTES
PROBABILITY | CHAPTER 3

DISCRETE PROBABILITY DISTRIBUTIONS

Random Variable:

Ex: Let ~~an~~ unfair die is tossed 3 times. The sample space is; $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ Let $P(H) = 0.4$ and we are intrested in Number of heads obtained. If X: # of heads obtained, X is a real valued function whose domain is sample space. We have;

$$X(HHH) = 3$$
$$X(HHT) = X(HTH) = X(THH) = 2$$
$$X(HTT) = X(THT) = X(TTH) = 1$$
$$X(TTT) = 0$$

* The probability mass function of a discrete Random Variable X is the function $p(x)$ that gives the probabilities

1. Random variable, Numerical value: $P(X = x) = p(x) \geq 0$

Since the total probability is 1, we have

2. $$\sum_x P(X = x) = \sum_x p(x) = 1$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* we have;

$P(X=3) = 0{,}4 \cdot 0{,}4 \cdot 0{,}4 = 0{,}064$

$P(X=2) = 3 \cdot 0{,}4 \cdot 0{,}4 \cdot 0{,}6 = 0{,}288$

$P(X=1) = 3 \cdot 0{,}4 \cdot 0{,}6 \cdot 0{,}6 = 0{,}432$

$P(X=0) = 0{,}6 \cdot 0{,}6 \cdot 0{,}6 = 0{,}216$

Then;

X	0	1	2	3
$p(x)$	0,216	0,432	0,288	0,064

is the pmf of X.

* A random variable X is discrete if it can take only a finite number, or a countable infinity, of possible values x. In general, X can take Natural Numbers (Count)

Exe X: Number of accidents in E5 highway in a month

X: Number of "3"s obtained among 15 trials of a dice.

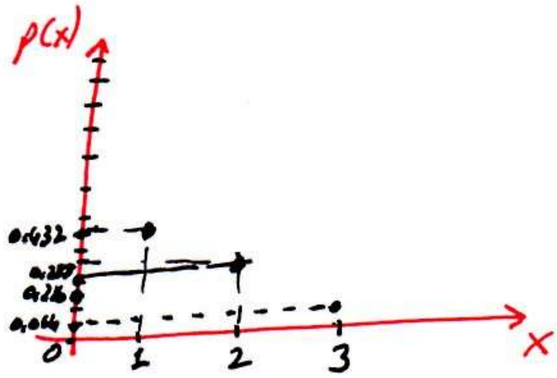
* (Cummulative) distribution function (cdf) of X is

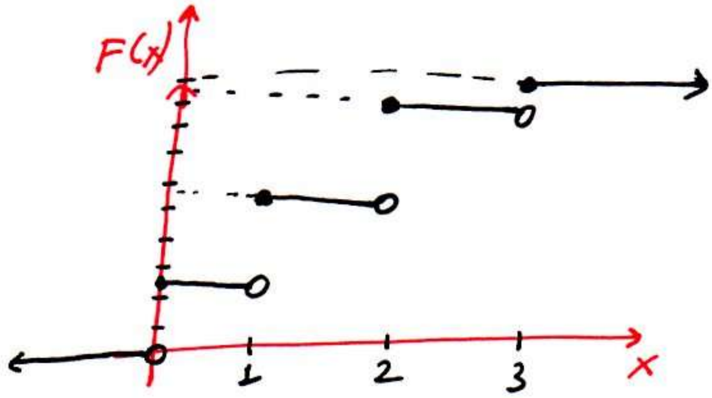
$$\boxed{F(x) = P(X \le x) = \sum_{k=-\infty}^{x} p(k)}$$

For X: # of Heads obtained example;

X	0	1	2	3
$p(x)$	0,216	0,432	0,288	0,064
$F(x)$	0,216	0,648	0,936	1,000

$$F(x) = \begin{cases} 0 & x < 0 \\ 0{,}216 & 0 \le x < 1 \\ 0{,}648 & 1 \le x < 2 \\ 0{,}936 & 2 \le x < 3 \\ 1{,}000 & x \geq 3 \end{cases}$$





Ex: Let $f(x) = \begin{cases} kx, & x \in \{0, 1, 2, 3, 4, 5\} \\ 0 & \text{otherwise} \end{cases}$

is a pmf.

a) Find k

b) Find cdf

c) Find $P(X<2)$ and $P(X \geq 3)$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Ans/k a) $\sum_{x=0}^{5} f(x) = 1$

$$\sum_{x=0}^{5} kx = k \cdot \sum_{x=0}^{5} x = k \cdot \frac{5.6}{2} = 15k = 1$$

$$k = \frac{1}{15}$$

$$f(x) = \begin{cases} \frac{1}{15}x & x = 0, 1, \dots, 5 \\ 0 & \text{o.w.} \end{cases}$$

b) $F(x) = \sum_{b=0}^{x} f(b) = \sum_{b=0}^{x} \frac{1}{15} b = \frac{1}{15} \cdot \frac{x(x+1)}{2} = \frac{x(x+1)}{30}$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x(x+1)}{30} & x = 0, 1, \dots, 5 \\ 1 & x \geq 5 \end{cases}$$ → this is stepwise

c) $P(X < 2) = P(X \leq 1) = F(1) = \frac{1.2}{30} = \frac{1}{15}$

$P(X \geq 3) = 1 - P(X \leq 2) = 1 - F(2) = 1 - \frac{2.3}{30} = \frac{4}{5}$

3.7 Of the people entering a blood bank to donate blood, 1 in 3 have type O^+ blood, and 1 in 15 have type O^- blood. For the next three people entering the blood bank, let X denote the number with O^+ blood and Y the number with O^- blood. Assuming independence among the people with respect to blood type, find the probability distributions for X and Y. Also find the probability distribution for $X + Y$, the number of people with type O blood.

X	0	1	2	3
$p_X(x)$	$\left(\frac{2}{3}\right)^3$	$3 \cdot \left(\frac{1}{3}\right) \cdot \left(\frac{2}{3}\right)^2$	$3 \cdot \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)$	$\left(\frac{1}{3}\right)^3$

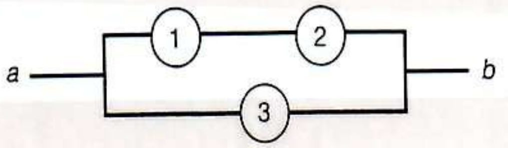
y	0	1	2	3
$p_y(y)$	$\left(\frac{14}{15}\right)^3$	$3 \cdot \left(\frac{1}{15}\right)\left(\frac{14}{15}\right)^2$	$3 \cdot \left(\frac{1}{15}\right)^2 \left(\frac{14}{15}\right)$	$\left(\frac{14}{15}\right)^3$

let $W = X + Y$ (Y, other next three people)

W	0	1	2
$p(w)$	$p_x(0)\cdot p_y(0)$	$p_x(1)\cdot p_y(0) +$ $p_x(0)\cdot p_y(1)$	$p_x(0)\cdot p_y(2) +$ $p_x(1)\cdot p_y(1) +$ $p_x(2)\cdot p_y(0)$

W	3	4	5	6
$p(w)$	$p_x(0)\cdot p_y(3) +$ $p_x(1)\cdot p_y(2) +$ $p_x(2)\cdot p_y(1) +$ $p_x(3)\cdot p_y(0)$	$p_x(1)\cdot p_y(3) +$ $p_x(2)\cdot p_y(2) +$ $p_x(3)\cdot p_y(1)$	$p_x(2)\cdot p_y(3) +$ $p_x(3)\cdot p_y(2)$	$p_x(3)\cdot p_y(3)$

3.10 When turned on, each of the three switches in the diagram below works properly with probability 0.9. If a switch is working properly, current can flow through it when it is turned on. Find the probability distribution for Y, the number of closed paths from a to b, when all three switches are turned on.



$$P(y=0) = P(\bar{3})\cdot[1 - P(1)\cdot P(2)] = 0{,}1\cdot[1 - 0{,}9\cdot 0{,}9] = 0{,}019$$

$$P(y=1) = P(3)\cdot[1 - P(1)\cdot P(2)] + P(\bar{3})\cdot[P(1)\cdot P(2)]$$
$$= 0{,}9\cdot[1 - 0{,}9\cdot 0{,}9] + 0{,}1\cdot[0{,}9\cdot 0{,}9] = 0{,}252$$

$$P(y=2) = P(3)\cdot[P(1)\cdot P(2)] = 0{,}9\cdot 0{,}9\cdot 0{,}9 = 0{,}729$$

y	0	1	2
$p(y)$	$0{,}019$	$0{,}252$	$0{,}729$

Expected Value & Variance

$$\mu = E(X) = \sum_x x \cdot p(x)$$

$$E[g(X)] = \sum_x g(x) \cdot p(x)$$

μ: Expected Value : Mean : Average

Then, $E(X^2) = \sum_x X^2 \cdot p(x)$

$$\sigma^2 = Var(X) = E[(X-\mu)^2] = \sum_x (x-\mu)^2 \cdot p(x)$$

σ^2 is variance of random variable X
σ is standard deviation of random variable X.

* What do mean and variance mean?

μ is location statistics and σ^2 is dispertion statistics. If we generate numbers from a random variable X, on the average, their mean gets closer to μ. Namely;

$$\lim_{n \to \infty} \frac{X}{n} = \mu.$$

On the other hand, consider two random variables X and Y, whose means are the same : μ. Let, X has a larger variance σ^2. This means, X has values away from μ than Y.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Ex4 Let X is the outcome of a fair dice and y is the outcome of a ten faced dice whose values are $\{-1, 0, 1, 2, 3, 4, 5, 6, 7, 8\}$

We have, $f_X(x) = \frac{1}{6}, \; x = 1, 2, \ldots, 6$

$$f_Y(y) = \frac{1}{10}, \; y = -1, 0, 1, \ldots, 8$$

Then; $E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3{,}5 = \mu$

$$E(Y) = -1 \cdot \frac{1}{10} + 0 \cdot \frac{1}{10} + \cdots + 8 \cdot \frac{1}{10} = 3{,}5 = \mu$$

Their variances are;

$$\sigma_X^2 = (1 - 3{,}5)^2 \cdot \frac{1}{6} + \cdots + (6 - 3{,}5)^2 \cdot \frac{1}{6} = 2{,}917$$

$$\sigma_Y^2 = (-1 - 3{,}5)^2 \cdot \frac{1}{10} + \cdots + (8 - 3{,}5)^2 \cdot \frac{1}{10} = 8{,}25$$

Because y's values are on a wider range y has a larger variance. (Note that wider range is NOT the only condition to have a larger variance. This example just gives an idea about variance.)

3.11 You are to pay \$1 to play a game consisting of drawing one ticket at random from a box of numbered tickets. You win the amount (in dollars) of the number on the ticket you draw. Two boxes are available with numbered tickets as shown below:

I $\boxed{0,\ 1,\ 2}$ II $\boxed{0,\ 0,\ 0,\ 1,\ 4}$

a Find the expected value and variance of your *net* gain per play with box I.

b Repeat part (a) for box II.

c Given that you have decided to play, which box would you choose and why?

3.11) a)

X	0	1	2
$P_X(x)$	$1/3$	$1/3$	$1/3$

$$E(X) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} = 1$$

$$Var(X) = (0-1)^2 \cdot \frac{1}{3} + (1-1)^2 \cdot \frac{1}{3} + (2-1)^2 \cdot \frac{1}{3} = \frac{2}{3} = 0,67$$

b)

y	0	1	4
$P_y(y)$	$\frac{3}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

$$E(y) = 0 \cdot \frac{3}{5} + 1 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} = 1$$

$$Var(y) = (0-1)^2 \cdot \frac{3}{5} + (1-1)^2 \cdot \frac{1}{5} + (4-1)^2 \cdot \frac{1}{5} = 2,4$$

c) Although their means are the same, Variance of game II is higher which means that it is more risky. If you like fun, game two may be more enjoyable. Both games are fair.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* For X is random and a, b constants, we have;

$$\boxed{E(aX+b) = a.E(X)+b}$$ and

$$\boxed{Var(aX+b) = Var(aX) = a^2.Var(X)}$$

Ex// # of parts a repairman repairs is a random variable with mean 5 and variance 3. Repairman gets a fixed money 40TL and 10TL for each part he repairs. Find mean and variance of the money he gets.

Ans// X: # of parts repaired

$E(X)=5$; $Var(X)=3$

y: Money repairman gets.

$y = 40 + 10X$

$$E(y) = E(40+10X) = 40 + 10E(X) = 40 + 10.5 = 90$$

$$Var(y) = Var(40+10X) = 10^2.Var(X) = 100.3 = 300$$

Ex// Show that $Var(X) = E(X^2) - \mu^2$

Ans// $Var(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2]$

$$= E(X^2) - 2\mu \underbrace{E(X)}_{=\mu} + \mu^2 = E(X^2) - 2\mu^2 + \mu^2$$

$$\boxed{Var(X) = E(X^2) - \mu^2}$$

Tchebysheff's Theorem

Let X be a random variable with mean μ and variance σ^2. Then, for $k > 0$,

$$P(|X-\mu| < k.\sigma) \geq 1 - \frac{1}{k^2}$$

or, equivalently,

$$P(\mu - k.\sigma < X < \mu + k.\sigma) \geq 1 - \frac{1}{k^2}.$$

Let the interval $(\mu - k.\sigma \,;\, \mu + k.\sigma)$ is k times σ interval around mean μ. This means that, at least $\left(1-\frac{1}{k^2}\right).100\%$ proportion of X must be in the interval $(\mu - k.\sigma \,;\, \mu + k.\sigma)$. Note that this is a lower bound.

3.17 The number of breakdowns for a university computer system is closely monitored by the director of the computing center, since it is critical to the efficient operation of the center. The number averages 4 per week, with a standard deviation of 0.8 per week.

a Find an interval that must include at least 90% of the weekly figures on number of breakdowns.

b The center director promises that the number of breakdowns will rarely exceed 8 in a one-week period. Is the director safe in making this claim? Why or why not?

3.17) a) $1 - \frac{1}{k^2} = 0{,}90$

$\frac{1}{k^2} = 0{,}10$

$k = 3{,}16$

Since $\mu = 4$ and $\sigma = 0{,}8$;

$4 \pm 3{,}16 . 0{,}8$

$(1{,}472 \,;\, 6{,}528)$

b) $\mu + k.\sigma = 8$

$4 + 0{,}8k = 8$

$k = \frac{8-4}{0{,}8} = 5$

$1 - \frac{1}{5^2} = 0{,}96$

Then, exceeding 8 can occur only 4% of the time (at most). He is safe.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.20 Costs of equipment maintenance are an important part of a firm's budget. Each visit by a field representative to check out a malfunction in a word processing system costs \$40. The word processing system is expected to malfunction approximately five times per month, and the standard deviation of the number of malfunctions is 2.

a Find the expected value and standard deviation of the monthly cost of visits by the field representative.

b How much should the firm budget per month to ensure that the cost of these visits will be covered at least 75% of the time?

a) X: # of malfunctions per month

$\mu_x = E(X) = 5 \; ; \quad \sigma_x = 2$

y: Cost of malfunctions

$\mu_y = E(y) = E(40X) = 40\,E(X) = 40 \cdot 5 = 200$

$\sigma_y^2 = Var(y) = Var(40X) = 40^2\,Var(X) = 40^2 \cdot 2^2 = 80^2$

$\sigma_y = \sqrt{80^2} = 80$

b) $1 - \frac{1}{k^2} = 0{,}75 \Rightarrow k = 2$

$\mu \pm k.\sigma : \; 200 \pm 2.80$

$(40 \; ; \; 360)$

To ensure the cost, 360\$ should be budgeted.

Binomial Distribution

$X \sim \text{Binomial}(n; p)$

$f(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$

$\mu = E(X) = n \cdot p \qquad \sigma^2 = Var(X) = n \cdot p \cdot (1-p)$

n: # of independent trials

p: Fixed probability of success

X: # of success.

3.24 In testing the lethal concentration of a chemical found in polluted water, it is found that a certain concentration will kill 20% of the fish that are subjected to it for 24 hours. If 20 fish are placed in a tank containing this concentration of chemical, find the probability that after 24 hours

- **a** exactly 14 survive.
- **b** at least 10 survive.
- **c** at most 16 survive.

3.25 Refer to Exercise 3.24.

- **a** Find the number expected to survive out of 20.
- **b** Find the variance of the number of survivors out of 20.

3.24) X: # of fish that are survived

$X \sim \text{Binomial}(n = 20;\ p = \overbrace{0{,}8}^{=1-0{,}2})$

$P(X = x) = f(x) = \binom{20}{x} 0{,}8^x (1 - 0{,}8)^{20-x}$

a) $P(X = 14) = \binom{20}{14} \cdot 0{,}8^{14} \cdot 0{,}2^6 = 0{,}1091$

b) $P(X \leq 16) = 1 - P(X \geq 17) = 1 - [f(17) + f(18) + f(19) + f(20)]$

c) $P(X \geq 10) = \sum_{x=10}^{20} f(x)$

3.25) $E(X) = 20 \cdot 0{,}8 = 16$

$Var(X) = 20 \cdot 0{,}8 \cdot 0{,}2 = 3{,}2$

3.31 A missile protection system consists of n radar sets operating independently, each with probability 0.9 of detecting an aircraft entering a specified zone. (All radar sets cover the same zone.) If an airplane enters the zone, find the probability that it will be detected if

a $n = 2$ **b** $n = 4$

3.32 Refer to Exercise 3.31. How large must n be if it is desired to have probability $1-10^{+6}$ ~~0.99~~ of detecting an aircraft entering the zone?

3.31) X: # of radars detecting an aircraft.

$X \sim Binomial(n; p = 0{,}9)$

$$f(x) = \binom{n}{x} \cdot 0{,}9^x \cdot 0{,}1^{n-x}$$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - f(0) = ?$$

a) $n = 2 \Rightarrow P(X \geq 1) = 1 - \binom{2}{0} \cdot 0{,}9^0 \cdot 0{,}1^2 = 0{,}99$

b) $n = 4 \Rightarrow P(X \geq 1) = 1 - \binom{4}{0} \cdot 0{,}9^0 \cdot 0{,}1^4 = 0{,}9999$

3.32) $P(X \geq 1) = 1 - 10^{+6}$

$$1 - \binom{n}{0} \cdot 0{,}9^0 \cdot 0{,}1^{+n} = 1 - 10^{+6}$$

$$1 - 0{,}1^n = 1 - 10^6$$

$n = 6$ ~~[illegible]~~ radars should be used.

3.34 An oil exploration firm is to drill ten wells, with each well having probability 0.1 of successfully producing oil. It costs the firm \$10,000 to drill each well. A successful well will bring in oil worth \$500,000.

a Find the firm's expected gain from the ten wells.

b Find the standard deviation of the firm's gain.

3.34) X: # of wells producing oil.

$X \sim \text{Binomial}(n=10;\ p=0{,}1)$

$y = -10 \cdot 10000 + 500000 \cdot X$

$y = 5X - 1$ (in 100 thousand dollars)

a) $E(X) = n \cdot p = 10 \cdot 0{,}1 = 1$

$E(y) = E(5X-1) = 5E(X) - 1 = 5 - 1 = 4$

$E(y) = 400000$ dollars

b) $Var(X) = n \cdot p \cdot (1-p) = 10 \cdot 0{,}1 \cdot 0{,}9 = 0{,}9$

$Var(y) = Var(5X-1) = 25\, Var(X) = 25 \cdot 0{,}9 = 22{,}5$

$\sigma_y = \sqrt{22{,}5} = 4{,}74342$

$\sigma_y = 474342$ dollars.

(Note that, y can be Negative. in fact; ~~$y \in \{5k-1, k \in \{0,$~~
$y \in \{5k-1 \mid k \in \{0, 1, \ldots, 10\}\}$
in 100 thousand dollars.)

3.37 Ten motors are packaged for sale in a certain warehouse. The motors sell for \$100 each, but a "double-your-money-back" guarantee is in effect for any defectives the purchaser might receive. Find the expected net gain for the seller if the probability of any one motor being defective is 0.08. (Assume that the quality of any one motor is independent of the quality of the others.)

3.37) X: # of motors defective.

$X \sim \text{Binomial}(n=10;\ p=0{,}08)$

$E(X) = 10 \cdot 0{,}08 = 0{,}8$

Let y: Net gain. Then;

$$y = 100.(10-X) - 200.X = 1000 - 100X - 200X$$
$$= 10000 - 300X$$

$$E(y) = E(10000 - 300X) = 10000 - 300.E(X)$$
$$= 10000 - 300.0{,}8 = 9760$$

3.35 A firm sells four items randomly selected from a large lot known to contain 10% defectives. Let Y denote the number of defectives among the four sold. The purchaser of the item will return the defectives for repair, and the repair cost is given by

$$C = 3Y^2 + Y + 2$$

Find the expected repair cost.

3.35) y: # of defectives

$y \sim Binomial(n=4;\ p=0{,}1)$

$\mu = E(y) = 4.0{,}1 = 0{,}4$

$Var(y) = 4.0{,}1.0{,}9 = 0{,}36$

Note that; $Var(y) = E(y^2) - \mu^2$

Then; $E(y^2) = Var(y) + \mu^2 = 0{,}36 + 0{,}4^2 = 0{,}52$

$$C = 3y^2 + y + 2$$

$$E(y) = E(3y^2 + y + 2) = 3E(y^2) + E(y) + 2$$
$$= 3.0{,}52 + 0{,}04 + 2 = 3{,}6$$

Geometric Distribution

$X \sim Geometric(p)$

p: Fixed probability of success.

X: # of independent trials to obtain FIRST sucess.

$$f(x) = (1-p)^{x-1} \cdot p$$

$$\mu = E(X) = \frac{1}{p} \quad ; \quad \sigma^2 = Var(X) = \frac{1-p}{p^2}$$

Negative Binomial Distribution

$X \sim Negative\ Binomial(r; p)$

p: Fixed probability of success

r: # of success required

X: # of independent trials to obtain r^{th} success.

$$f(x) = \binom{x-1}{r-1} \cdot (1-p)^{x-r} \cdot p^r$$

$$\mu = E(X) = \frac{r}{p} \qquad \sigma^2 = Var(X) = \frac{r \cdot (1-p)}{p^2}$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.40 Suppose 10% of the engines manufactured on a certain assembly line are defective. If engines are randomly selected one at a time and tested, find the probability that the first nondefective engine is found on the second trial.

3.40) X: # of engines selected to obtain FIRST nondefective one.

$X \sim Geometric(p = \underbrace{0{,}90}_{= 1 - 0{,}10})$

$f(x) = 0{,}10^{x-1} \cdot 0{,}90$

$P(X=2) = f(2) = 0{,}10^{2-1} \cdot 0{,}90 = 0{,}09$

3.41 Refer to Exercise 3.40. Find the probability that the third nondefective engine is found

a on the fifth trial.

b on or before the fifth trial.

3.42 Refer to Exercise 3.40. Given that the first two engines are defective, find the probability that at least two more engines must be tested before the first nondefective is found.

3.43 Refer to Exercise 3.40. Find the mean and variance of the number of the trial on which

a the first nondefective engine is found.

b the third nondefective engine is found.

3.41) a) $P(X=5) = 0{,}10^{5-1} \cdot 0{,}90 = 0.9 \cdot 10^{-5}$

b) $P(X \leq 5) = \sum_{x=1}^{5} f(x)$

3.42) Since trials are independent, $P(X \geq 2) = ?$

$P(X \geq 2) = 1 - [f(1)] = 1 - 0{,}10^{1-1} \cdot 0{,}90 = 0{,}10$

3.43) a) $E(X) = \frac{1}{0{,}90} = 1{,}11$; $Var(X) = \frac{1-0{,}90}{0{,}90^2} = 0{,}123$

b) Let $Y \sim$ Negative Binomial $(r=3; p=0{,}90)$

$E(Y) = \frac{3}{0{,}90} = 3{,}33$; $Var(Y) = \frac{3 \cdot (1-0{,}90)}{0{,}90^2} = 0{,}369$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.44 The employees of a firm that manufactures insulation are being tested for indications of asbestos in their lungs. The firm is requested to send three employees who have positive indications of asbestos to a medical center for further testing. If 40% of the employees have positive indications of asbestos in their lungs, find the probability that ten employees must be tested to find three positives.

3.45 Refer to Exercise 3.44. If each test costs \$20, find the expected value and variance of the total cost of conducting the tests to locate three positives. Do you think it is highly likely that the cost of completing these tests would exceed \$350?

3.44) X: # of employees to test to find 3 positives

$X \sim$ Negative Binomial$(r=3;\ p=0{,}40)$

$$f(x) = \binom{x-1}{3-1} 0{,}60^{x-3} \cdot 0{,}40^{3}$$

$$P(X=10) = \binom{9}{2} 0{,}60^{7} \cdot 0{,}40^{3} = 0{,}0645$$

3.45) $E(X) = \frac{3}{0{,}40} = 7{,}5$; $Var(X) = \frac{3.(1-0{,}40)}{0{,}40^2} = 11{,}25$

y: Cost of the tests $\Rightarrow y = 20X$

$\mu_y = E(y) = E(20X) = 20 \cdot E(X) = 20 \cdot 7{,}5 = 150$

$Var(y) = Var(20X) = 20^2\, Var(X) = 20^2 \cdot 11{,}25 = 4500$

$\sigma_y = \sqrt{4500} = 67{,}08$

$\mu_y + k \cdot \sigma_y = 350$

$150 + k \cdot 67{,}08 = 350$

$k = \frac{350-150}{67{,}08} = 2{,}98$

$1 - \frac{1}{k^2} = 0{,}888$

At least 88,8% of the costs are within (0;350)

At most 12,2% of the costs will exceed \$350.

Not highly likely.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.51 An appliance comes in two colors, white and brown, which are in equal demand. A certain dealer in these appliances has three of each color in stock, although this is not known to the customers. Customers arrive and independently order these appliances. Find the probability that

a the third white is ordered by the fifth customer.

b the third brown is ordered by the fifth customer.

c all of the whites are ordered before any of the browns.

d all of the whites are ordered before all of the browns.

a) X: # of customers to order 3^{rd} white

$$X \sim \text{Negative Binomial}(p = \tfrac{1}{2})$$

$$f(x) = \binom{x-1}{2}\left(\tfrac{1}{2}\right)^{x-3}\left(\tfrac{1}{2}\right)^{3}$$

$$P(X=5) = \binom{4}{2}\left(\tfrac{1}{2}\right)^{2}\left(\tfrac{1}{2}\right)^{3} = 0{,}1875$$

b) 0,1875 because white and Brown are in equal demand

c) $P(X=3) = \binom{2}{2}\left(\tfrac{1}{2}\right)^{0}\left(\tfrac{1}{2}\right)^{3} = \frac{1}{8} = 0{,}125$

d) $P(X<6) = 1 - P(X=6) = 1 - \binom{5}{2}\cdot\left(\tfrac{1}{2}\right)^{6-3}\cdot\left(\tfrac{1}{2}\right)^{3} = 0{,}844$

Poisson Distribution

X: # of success / unit time, space, etc.

$$X \sim \text{Poisson}(\lambda)$$

$$f(x) = \frac{e^{-\lambda}\cdot\lambda^{x}}{x!}$$

$$\mu = E(X) = \lambda \; ; \quad \sigma^2 = Var(X) = \lambda$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* Note that we use Poisson distribution if we want to find the # of success and we only know the average.

* When X is in fact Binomial but n is too high and p is too small, we can approximate Binomial Probabilities by setting $\lambda = n.p$. Aproximation is practically good when $np < 7$.

* If X is # of ~~units~~ events / unit time and # of units in t units of time is asked, we have; $X \sim Poisson(\lambda) \Leftrightarrow Y \sim Poisson(\lambda t)$

where Y is # of events / t units time.

* λ is sometimes called "rate"

Ex/ Among 200 telephone bills, ~~...~~ probability that are is incorrect is independently 0,01. WPT we have at least 2 incorrect bills?

Ans/ $X \sim Binomial(n=200; p=0,01)$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \left[\binom{200}{0} 0{,}01^0 . 0{,}99^{200} + \binom{200}{1} 0{,}01^1 0{,}99^{199}\right]$$
$$= 0{,}59535$$

setting $\lambda = n.p = 2$

$$P(X \geq 2) = 1 - \left[\frac{e^{-2} . 2^0}{0!} + \frac{e^{-2} . 2^1}{1!}\right] = 0{,}59399$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.62 The number of imperfections in the weave of a certain textile has a Poisson distribution with a mean of four per square yard.

a Find the probability that a 1-square-yard sample will contain at least one imperfection.

b Find the probability that a 3-square-yard sample will contain at least one imperfection.

3.63 Refer to Exercise 3.62. The cost of repairing the imperfections in the weave is \$10 per imperfection. Find the mean and standard deviation of the repair costs for an 8-square-yard bolt of the textile in question.

3.64 The number of bacteria colonies of a certain type in samples of polluted water has a Poisson distribution with a mean of two per cubic centimeter.

a If four 1-cubic-centimeter samples are independently selected from this water, find the probability that at least one sample will contain one or more bacteria colonies.

b How many 1-cubic-centimeter samples should be selected to have a probability of approximately 0.95 of seeing at least one bacteria colony?

3.62) X: # of imperfections / sq yard

$X \sim \text{Poisson}(\lambda = 4 / \text{sq yard})$

a) $$f(x) = \frac{e^{-4} 4^x}{x!}$$

$$P(X \geq 1) = 1 - P(X=0) = 1 - \frac{e^{-4} 4^0}{0!} = 0{,}9817$$

b) y: # of imperfections / 3-sq yard

$E(y) = \lambda_y = 3.4 = 12$

$$f_y(y) = \frac{e^{-12} \cdot 12^y}{y!}$$

$$P(y \geq 1) = 1 - P(y=0) = 1 - \frac{e^{-12} 12^0}{0!} \cong 1$$

3.63) N: # of imperfections / 8-sq yard

$E(N) = 8.4 = 32$ $\quad$ $Var(N) = 32$

$E(C) = E(10N) = 320$ $\quad$ $Var(C) = Var(10N) = 3200$

3.64) X: # of bacteria / cm^3

$$f(x) = \frac{e^{-2} \cdot 2^x}{x!}$$

a) $p = P(X \geq 1) = 1 - f(0) = 1 - \frac{e^{-2} \cdot 2^0}{0!} = 0{,}865$

Y: # of samples contain at least one bacteria

$Y \sim$ Binomial ($n = 4$; $p = 0{,}865$)

$$f_Y(y) = \binom{4}{y} 0{,}865^y \cdot 0{,}135^{4-y}$$

$$P(Y \geq 1) = 1 - P(Y=0) = 1 - \binom{4}{0} 0{,}865^0 \cdot 0{,}135^4 = 0{,}9997$$

b) $1 - \binom{n}{0} 0{,}865^0 \, 0{,}135^n = 0{,}95$

$$0{,}05 = 0{,}135^n$$

$$\ln 0{,}05 = n \cdot \ln 0{,}135$$

$$n = \frac{\ln 0{,}05}{\ln 0{,}135} = 1{,}5$$

$n = 2$ samples ~~have~~ have prob. higher than 0,95

3.66 A food manufacturer uses an extruder (a machine that produces bite-size foods such as cookies and many snack foods) that produces revenue for the firm at the rate of \$200 per hour when in operation. However, the extruder breaks down an average of two times for every 10 hours of operation. If Y denotes the number of breakdowns during the time of operation, the revenue generated by the machine is given by

$$R = 200t - 50Y^2$$

where t denotes hours of operation. The extruder is shut down for routine maintenance on a regular schedule and operates like a new machine after this maintenance. Find the optimal maintenance interval t_0 to maximize the expected revenue between shutdowns.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

3.66) $Y \sim Poisson(0{,}2t)$

$2/10$ hours $\Rightarrow 0{,}2/\text{hour} \Rightarrow 0{,}2t / t$ hours

$$E(Y) = 0{,}2t \qquad Var(Y) = 0{,}2t$$

$$E(Y^2) = 0{,}2t + (0{,}2t)^2 = 0{,}2t + 0{,}04t^2$$

$$E(B) = 200t - 50(0{,}2t + 0{,}04t^2)$$

$$E(B) = 190t + 2t^2$$

$$\frac{dE(B)}{dt} = 190 + 4t = 0 \Rightarrow t = 47{,}5 \text{ hours.}$$

Hypergeometric Distribution

k: Type I items

$N-k$: Non-Type I items

N: Total items

select n items randomly

X: # of type I items

$$X \sim Hypergeometric\,(N; k; n)$$

$$f(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$\mu = E(X) = n \cdot \frac{k}{N} \; ; \quad \sigma^2 = Var(X) = n\frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)$$

* For N is large and k is small, Hypergeometric distribution is very close to Binomial with $p=\frac{k}{N}$.
Practically, if $\frac{n}{N}=5\%$, approximation is applicable.
Notice the similarity with $E(X)=np$ and $Var(X)=np(1-p)$. The term $\frac{N-n}{N-1}$ is called Finite Population Correction factor, which is close to 1 for $\frac{n}{N}<5\%$.

3.72 A foreman has ten employees from whom he must select four to perform a certain undesirable task. Among the ten employees, three belong to a minority ethnic group. The foreman selected all three minority employees (plus one other) to perform the undesirable task. The members of the minority group then protested to the union steward that they had been discriminated against by the foreman.

The foreman claimed that the selection was completely at random. What do you think?

3.72)

$\boxed{\begin{array}{l}3 \text{ ethnic} \\ 7 \text{ non ethnic}\end{array}}$ → Select $n=3$ employees

$N=10$ employees

X: # of employees that are from ethnic group.

$X \sim \text{Hypergeometric}(N=10; k=3; n=3)$

$$f(x)=\frac{\binom{3}{x}\binom{7}{3-x}}{\binom{10}{3}}$$

$$P(X=3)=\frac{\binom{3}{3}\binom{7}{0}}{\binom{10}{3}}=0{,}0083 \rightarrow \text{Low probability.}$$

I do NOT think so because Random selection has low probability of this outcome.

3.74 Used photocopying machines are returned to the supplier, cleaned, and then sent back out on lease agreements. Major repairs are not made, and, as a result, some customers receive malfunctioning machines. Among eight used photocopiers in supply today, three are malfunctioning. A customer wants to lease four of these machines immediately. Hence four machines are quickly selected and sent out, with no further checking. Find the probability that the customer receives

a no malfunctioning machines.

b at least one malfunctioning machine.

c three malfunctioning machines.

3.74)

3 malfunction
5 Not malfunc.

→ Select $n=4$ machines

X: # of machines that are malfunctioning.

$N=8$ machines

$X \sim \text{Hypergeometric}\,(N=8;\ k=3;\ n=4)$

$$P(X=0) = \frac{\binom{3}{0}\binom{5}{4}}{\binom{8}{4}} = 0{,}071$$

$$P(X \geq 1) = 1 - P(X=0) = 1 - 0{,}929$$

$$P(X=3) = \frac{\binom{3}{3}\binom{5}{0}}{\binom{8}{4}} = 0{,}014$$

3.79 Lot acceptance sampling procedures for an electronics manufacturing firm call for sampling n items from a lot of N items and accepting the lot if $Y \leq c$, where Y is the number of nonconforming items in the sample. From an incoming lot of 20 printer covers, 5 are to be sampled. Find the probability of accepting the lot if $c = 1$ and the actual number of nonconforming covers in the lot is

a 0 **b** 1 **c** 2 **d** 3 **e** 4

3.79)

k
N-k

→ Select $n=5$ items

Y: # of defective items

$Y \sim \text{Hypergeometric}\,(N=20;\ k;\ n=5)$

$N=20$ items

$P(\text{Accept}) = P(Y \leq 1)$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

a-b) If $k=0$ OR 1, $P(\text{Accept})=1$

c) $k=2 \Rightarrow P(Y\le 1)=1-P(Y=2)=1-\dfrac{\binom{2}{2}\binom{8}{3}}{\binom{10}{5}}=0{,}777$

d) $k=3 \Rightarrow P(Y\le 1)=\dfrac{\binom{3}{0}\binom{7}{5}+\binom{3}{1}\binom{7}{4}}{\binom{10}{5}}=0{,}5$

e) $k=4 \Rightarrow P(Y\le 1)=\dfrac{\binom{4}{0}\binom{6}{5}+\binom{4}{1}\binom{6}{4}}{\binom{10}{5}}=0{,}262$

3.81 Two assembly lines (I and II) have the same rate of defectives in their production of voltage regulators. Five regulators are sampled from each line and tested. Among the total of ten tested regulators, there were four defectives. Find the probability that exactly two of the defectives came from line I.

→ line I : $X_1 \sim \text{Binomial}(n_1=5;p)$

→ line II : $X_2 \sim \text{Binomial}(n_2=5;p)$

$Y=X_1+X_2 \sim \text{Binomial}(n=10;p)$

$$P(X_1=2 \mid Y=4)=\frac{P(X_1=2,\ X_2=2)}{P(Y=4)}=\frac{P(X_1=2)\cdot P(X_2=2)}{P(Y=4)}$$

$$=\frac{\binom{5}{2}p^2\cdot(1-p)^3\cdot\binom{5}{2}\cdot p^2\cdot(1-p)^3}{\binom{10}{4}p^4\cdot(1-p)^6}=\frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}}=\frac{10}{21}$$

Note that the problem turns into hypergeometric distribution; $X_1 \sim \text{Hypergeometric}(N=10;\ K=5;\ n=4)$

Because they have the same defective rate.

Moment Generating Function

* Remember $E[g(x)] = \sum_x g(x) f(x)$ where $f(x)$ is pmf of X.

Now, consider the function $g(t) = e^{tX}$.

Its derivatives with respect to t are;

$$g^{(1)}(t) = \frac{d\,g(t)}{d(t)} = \frac{d\,e^{tX}}{dt} = X.\,e^{tX}$$

$$g^{(2)}(t) = \frac{d^2 g(t)}{dt^2} = \frac{d}{dt} X.e^{tX} = X^2.\,e^{tX}$$

and in general, $g^{(k)}(t) = X^k.\,e^{tX}$.

* The expectation $E(X^k)$ is called "k^{th} moment" of X.

Let $M^{(k)}(t) = E[g^{(k)}(t)] = E[X^k.\,e^{tX}]$

Then, $M(t)$ is called "Moment generating function" because;

$$\boxed{M^{(k)}(0) = E(X^k)}$$

* Note that; $\mu = E(X) = M^{(1)}(0)$ and

$$\sigma^2 = E(X^2) - \mu^2 = M^{(2)}(0) - [M^{(1)}(0)]^2$$

3.82) Bernoulli random Variable is Binomial with $n=1$. Therefore, n independent trials are called Bernoulli trials. We have; $X \sim \text{Bernoulli}(p)$

$$f(x) = p^{x}(1-p)^{1-x}, \; x \in \{0,1\}$$

$$M(t) = E(e^{tX}) = \sum_{x=0}^{1} e^{tx} f(x) = e^{t.0} f(0) + e^{t.1} f(1)$$

$$M(t) = (1-p) + e^{t} . p = pe^{t} + (1-p)$$

3.83) Let X_i are iid Bernoulli trials, $i = 1, 2, \ldots, n$
independent and identically distributed

Then; $Y = X_1 + X_2 + \cdots + X_n \sim \text{Binomial}(n; p)$

$$M_Y(t) = M_1(t) . M_2(t) . \ldots . M_n(t) = [pe^{t} + (1-p)]^{n}$$

$$M_Y^{(1)}(t) = n . p . e^{t} [p . e^{t} + (1-p)]^{n-1}$$

$$\mu = M_Y^{(1)}(0) = E(Y) = n . p \, [p + 1 - p]^{n-1} = np$$

$$M_Y^{(2)}(t) = npe^{t} (n-1) . p [p . e^{t} + (1-p)]^{n-2} + npe^{t} . [pe^{t} + (1-p)]^{n-1}$$

$$M_Y^{(2)}(0) = E(Y^2) = np^2 (n-1) + np$$

$$\sigma^2 = E(Y^2) - \mu^2 = n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np(1-p)$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

* The following properties of Moment Generating Function is useful to find Mean and Variance.

(i) Let $Y = X_1 + X_2 + \dots + X_n$, X_i are independent

Then; $\boxed{M_Y(t) = M_1(t) \cdot M_2(t) \cdot \dots \cdot M_n(t)}$ because;

$$M_Y(t) = E(e^{Yt}) = E\left(e^{(X_1+X_2+\dots+X_n)t}\right) = E\left(e^{X_1 t} \cdot \dots \cdot e^{X_n t}\right)$$

$$= E(e^{X_1 t}) \cdot E(e^{X_2 t}) \cdot \dots \cdot E(e^{X_n t}) = M_1(t) \cdot \dots \cdot M_n(t)$$

by independence

(ii) If X has mgf $M(t)$ and $Y = aX+b$ then

$\boxed{M_Y(t) = e^{tb} M(at)}$ because

$$M_Y(t) = E(e^{Yt}) = E(e^{(aX+b)t}) = E(e^{aXt} \cdot e^{bt})$$

$$= e^{bt} \cdot E(e^{at \cdot X}) = e^{bt} \cdot M(at)$$

3.82 Find the moment-generating function for the Bernoulli random variable.

3.83 Show that the moment-generating function for the binomial random variable is given by

$$M(t) = [pe^t + (1-p)]^n$$

Use this result to derive the mean and variance for the binomial distribution.

3.84 Show that the moment-generating function for the Poisson random variable with mean λ is given by

$$M(t) = e^{\lambda(e^t - 1)}$$

Use this result to derive the mean and variance for the Poisson distribution.

*Note that the relationship between Geometric and Negative Binomial Distributions is the same.

$X_i \sim \text{Geometric}(p)$

$\Rightarrow Y = X_1 + X_2 + \dots + X_r \sim \text{Negative Binomial}(r; p)$

So, same procedure is applied to find mean and variance of Negative Binomial Distribution.

*Also we can show that,

$$\left.\begin{array}{l} X_1 \sim \text{Binomial}(n_1; p) \\ X_2 \sim \text{Binomial}(n_2; p) \end{array}\right\} \Rightarrow Y = X_1 + X_2 \sim \text{Binomial}(n_1 + n_2; p)$$

3.84) $X \sim \text{Poisson}(\lambda) \Rightarrow f(x) = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} = e^{-\lambda} \underbrace{\sum_{x=0}^{\infty} \frac{[e^t \cdot \lambda]^x}{x!}}_{= e^{\lambda \cdot e^t}}$$

$$= e^{-\lambda} \cdot e^{\lambda \cdot e^t} = e^{\lambda[e^t - 1]}$$

$$M^{(1)}(t) = \lambda \cdot e^t e^{\lambda \cdot [e^t - 1]} \; ; \; M^{(2)}(t) = \lambda^2 e^{2t} e^{\lambda[e^t - 1]} + \lambda \cdot e^t \cdot e^{\lambda[e^t - 1]}$$

$\mu = E(X) = M^{(1)}(0) = \lambda$; $E(X^2) = M^{(2)}(0) = \lambda^2 + \lambda$

$\sigma^2 = Var(X) = (\lambda^2 + \lambda) - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com