Problem Set 5 - Solutions

Statistics - NYU, Summer 2016
Ercan Karadas

Section 1

[1]

a) - First, determine the possible values that this random variable can assume:

$$X \in \{0, 1, 2\}$$

- Second, for each possible value find the corresponding probability:
$P(X = 0) = P(WW) = \frac{4}{6}\frac{4}{6} = (\frac{4}{6})^2$
$P(X = 1) = P(WB \text{ or } BW) = \frac{4}{6}\frac{2}{6} + \frac{2}{6}\frac{4}{6} = 2(\frac{2}{6})(\frac{4}{6})$
$P(X = 2) = P(BB) = (\frac{2}{6})(\frac{2}{6}) = (\frac{2}{6})^2$

- Third, by combining possible values and the corresponding probabilities we obtain the probability distribution function:

X	0	1	2
$P(X)$	16/36	16/36	4/36

Alternatively, we can present the probability distribution function in a table and we can call this probability distribution table:

X	$P(X)$
0	16/36
1	16/36
2	4/36

These two contain exactly the same information, so you can use either one. In the rest of the problem set I will use the table format.

Note: In part (b) and (c), I will follow the same three steps, but in order not to repeat same thing again and again I will not repeat them. In general, whenever you need to find a probability distribution of given random variable, follow these three steps:
- Step 1. If the random variable is given explicitly (as in this problem), determine the possible values that it can assume. If the random variable is not given explicitly, then first you need to define an appropriate random variable. For example, in problem 3, first you need to determine the random variable.

- Step 2. For each possible value find the corresponding probability. For example, $X = 1$ means we are interested in the outcomes that contain only one black ball. Recall that the sample space for this experiment was $\{WW, WB, BW, BB\}$. Therefore, $X = 1$ possible only if this experiment produces either WB or BW, so we have $P(X = 1) = P(WB \text{ or } BW)$. And since these two outcomes are mutually exclusive we finally obtain $P(X = 1) = P(WB \text{ or } BW) = P(WB) + P(BW)$.

- Step 3. Represent the results in a convenient format, for example as a table.

As we will see when we need to compute the expected value or the variance for a random variable, the key step is again to figure out the probability distribution table.

b) $Y \in \{0, 1, 2\}$
$P(Y = 0) = P(WW) = \frac{4}{6}\frac{3}{5} = \frac{12}{30}$
$P(Y = 1) = P(WB \text{ or } BW) = \frac{4}{6}\frac{2}{5} + \frac{2}{6}\frac{4}{5} = \frac{16}{30}$
$P(Y = 2) = P(BB) = (\frac{2}{6})(\frac{1}{5}) = \frac{2}{30}$
Then,

Y	$P(Y)$
0	12/30
1	16/30
2	2/30

c) $Z \in \{0, 1, 2, \dots 20\}$
Instead of finding $P(Z = 0)$, $P(Z = 1)$, ... $P(Z = 20)$, just observe that actually this is just the Binomial model with $n = 20$ and $p = 1/3$, i.e. $B(20, 1/3)$. Therefore,

$$P(Z = z) = \binom{20}{z}(\frac{1}{3})^z(\frac{2}{3})^{20-z}$$

Also note that the experiment in part (b) is not Binomial because probability of drawing B in the second draw depends on the outcome of the first draw.

[2] a)

$$\begin{aligned}\mu = E(X) &= \sum_{x\in\{0,1,2\}} x \cdot P(X = x) \\ &= 0 \times (\frac{16}{36}) + 1 \times (\frac{16}{36}) + 2 \times (\frac{4}{36}) \\ &= \frac{2}{3}\end{aligned}$$

$$\begin{aligned}\sigma^2 = Var(X) &= \sum_{x\in\{0,1,2\}} (x-\mu)^2 \cdot P(X=x) \\ &= (0-\frac{2}{3})^2 \times (\frac{16}{36}) + (1-\frac{2}{3})^2 \times (\frac{16}{36}) + (2-\frac{2}{3})^2 \times (\frac{4}{36}) \\ &= \frac{4}{9}\end{aligned}$$

b)

$$\begin{aligned}\mu = E(Y) &= \sum_{y\in\{0,1,2\}} y \cdot P(Y=y) \\ &= 0 \times (\frac{12}{30}) + 1 \times (\frac{16}{30}) + 2 \times (\frac{2}{30}) \\ &= \frac{2}{3} \\ \sigma^2 = Var(Y) &= \sum_{y\in\{0,1,2\}} (y-\mu)^2 \cdot P(Y=y) \\ &= (0-\frac{2}{3})^2 \times (\frac{12}{30}) + (1-\frac{2}{3})^2 \times (\frac{16}{30}) + (2-\frac{2}{3})^2 \times (\frac{2}{30}) \\ &= \frac{16}{45}\end{aligned}$$

c) Since Z is Binomial $E(Z) = np = 20(1/3) = 20/3$, and $V(Z) = np(1-p) = 20(1/3)(2/3) = 40/9$

[3] Let us define a random variable X as the number of days to complete the project.

a) $P(X < 3) = P(X = 1) + P(X = 2) = 0.05 + 0.20 = 0.25$

b)

$$\begin{aligned}\mu_X = E(X) &= \sum_{x\in\{1,2,3,4,5\}} xP(X=x) \\ &= 1.(0.05) + 2.(0.20) + 3.(0.35) + 4.(0.30) + 5.(0.10) \\ &= 3.2\end{aligned}$$

c)

$$\begin{aligned}\sigma^2 = Var(X) &= \sum_{x\in\{1,2,3,4,5\}} (x-\mu)^2 P(X=x) \\ &= (1-3.2)^2.(0.05) + (2-3.2)^2.(0.20) + (3-3.2)^2.(0.35) \\ &+ (4-3.2)^2.(0.30) + (5-3.2)^2.(0.10) \\ &= 1.05987\end{aligned}$$

Therefore, the standard deviation is $\sigma = \sqrt{1.05987} = 1.029$.

d) Let us define another random variable C to denote the total cost of the project as $C = 20,000 + 2,000X$. Then
$E(C) = 20,000 + 2,000.E(X) = 20,000 + (2,000(3.2) = \$26,400$
Since $\sigma^2 = Var(C) = (2,000)^2 Var(X)$, standard deviation of C is equal to $\sigma = \sqrt{Var(C)} = (2,000)\sqrt{Var(X)} = (2000)(1.029) = \$2,059.1$

e) Let us define a random variable X as

$$X = \# \text{ of projects taking at least 4 days to complete.}$$

Then this is a binomial r.v. since we are given that individual projects are independent and there are two possibilities for each project; either it takes at least 4 days or strictly less than 4 days.

Parameters of the Binomial model:
Success probability p=.30 + .10 = .4.
Number of trials n = 3.
We are interested in $P(X \geq 2) =?$
$P(X \geq 2) = P(X = 2) + P(X = 3) = \binom{3}{2}(.4)^2(.6) + \binom{3}{3}(.4)^3(.6)^0 = .352$

[4]

a) $P(X = 3) = \binom{5}{3}(.55)^3(.45)^2 = .3369$
Binomial model, because we can think as if a randomly chosen freshmen will graduate in four years with .55 probability and probabilities are independent for randomly chosen two students.

b) Since sample has 5 elements majority means 3 or more.
$P(X \geq 3) = P(3)+P(4)+P(5) = .3369+(5)(.55)4(.45)+(1)(.55)5(1) = .5931$

c) On average $\mu = np = (80)(.55) = 44$ will graduate in 4 years, and the standard deviation is $\sigma = \sqrt{80(.55)(.45)} = 4.449$.

[5]

a) Let us X be a r.v. denoting the number of one-year-olds who did NOT recognize their mothers' voice. Then X has a Binomial distribution with $n = 20$ and $p = 0.1$. Therefore,
$P(X \geq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) = 0.323.$

b) $P(X = 0) = .122$

c) $E(X) = n.p = 20 \times 0.1 = 2$

d) $V(X) = np(1 - p) = 20 \times 0.1 \times 0.9 = 1.8$

e) $P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 0.957.$

[6]

a) Binomial model with $n = 10$, and $p = 0.9$
Answer: $\binom{10}{3}(0.9)^3(0.10)^7$

b) $E(X) = np = 200(0.9) = 180$
where r.v. X denotes the number of missile attacks that are successfully detected.

c)

$$\begin{aligned} P(\text{At least one of them will detect}) &= 1 - P(\text{none of them will detect}) \\ &= 1 - (0.1)(0.1) = 0.99 \end{aligned}$$

d) $\binom{10}{8}(0.99)^8(0.01)^2 + \binom{10}{9}(0.99)^9(0.01)^1 + \binom{10}{10}(0.99)^{10}(0.01)^0$

Section 2

[7] Let X be the number of vehicles available for use on a given day. Then

a)

$$\begin{aligned} P(X = 0) &= (0.05)(0.1)(0.1)(0.2) = 0.0001 \\ P(X = 1) &= (0.95)(0.1)(0.1)(0.2) + 2(0.05)(0.9)(0.1)(0.2) \\ &\quad + (0.05)(0.1)(0.1)(0.8) = 0.0041 \\ P(X = 2) &= 2(0.95)(0.9)(0.1)(0.2) + 2(0.05)(0.9)(0.1)(0.8) \\ &\quad + (0.05)(0.9)(0.9)(0.2) + (0.95)(0.1)(0.1)(0.8) = 0.0571 \\ P(X = 3) &= (0.05)(0.9)(0.9)(0.8) + 2(0.95)(0.1)(0.9)(0.8) \\ &\quad + (0.95)(0.9)(0.9)(0.2) = 0.3231 \\ P(X = 4) &= (0.95)(0.9)(0.9)(0.8) = 0.6156 \end{aligned}$$

Then the probability distribution table is given by:

Y	$P(Y)$
0	0.0001
1	0.0041
2	0.0571
3	0.3231
4	0.6156

b) $E(X) = 3.55$

c) $V(X) = 0.3875$ and therefore $\sigma_x = \sqrt{Var(X)} = 0.6225$.

[8] a) Team A can win the series by winning 4 out of either 4, 5, 6 or 7 games.
P(A wins in the first 4 games) $= (0.6)^4 = 0.1296$
P(A wins in the first 5 games) $= \left(C_3^4(0.6)^3(0.4)\right) \times (0.6) = 0.20736$ (To win

in the 5th game, it has to win 3 games in the first 4 and then the fifth game)
P(A wins in the first 6 games) $= \left(C_3^5 (0.6)^3 (0.4)^2\right) \times (0.6) = 0.20736$
P(A wins in the first 7 games) $= \left(C_3^6 (0.6)^3 (0.4)^3\right) \times (0.6) = 0.165888$
Thus P(A wins) $= 0.1296 + 2(0.20736) + 0.165888 = 0.71021$.

b)

$$\begin{aligned} P(7^{th} \text{ game needed}) &= P(\text{ A wins 3 and B wins 3 of first 6 games }) \\ &= C36(0.6)3(0.4)3 \\ &= 0.27648 \end{aligned}$$

[9] Let the number of returns be denoted by the random variable X. We can use the binomial model with n = 50, and p = 0.15.

a)

$$\begin{aligned} P(X \leq 3) &= \binom{50}{0}(0.15)^0(0.85)^{50} + \binom{50}{1}(0.15)^1(0.85)^{49} \\ &+ \binom{50}{2}(0.15)^2(0.85)^{48} + \binom{50}{3}(0.15)^3(0.85)^{47} \\ &= 0.046 \end{aligned}$$

b) $E(X) = 50(0.15)$ and $V(X) = 50(0.15)(0.85)$

c) Total Cost (TC)$= 100X$, so $E(TC) = 100[50(0.15)]$ and $V(TC) = 100^2[50(0.15)(0.85)]$.

[10]

a)

X	$P(X)$
0	16/49
1	24/49
2	9/49

b)

Y	$P(Y)$
0	12/42
1	24/42
2	6/42

c)

Z	$P(Z)$
0	4/81
1	28/81
2	49/81

d)

T	$P(T)$
0	30/72
1	36/72
2	6/72

[11] a) $E(X) = 42/49$, $V(X) = .4897$

b) $E(Y) = 36/42$, $V(Y) = .4081$

c) $E(Z) = 126/81$, $V(Z) = .34567$