GLIB DOLOTOV (U14734371) MA213 SPR 2019

Problem Set 4

Assigned on 02/25. Due on 03/04 (Monday)

1. Textbook problem 4.32, "Solar energy cells", page 194.
2. Textbook problem 4.52, "Reliability of a manufacturing network", page 200.
3. Textbook problem 4.82, "Assigning a passing grade", page 214.
4. Textbook problem 4.104, "Flaws in plastic-coated wire", page 219.
5. Textbook problem 4.126, "Gender discrimination suit", page 223.

4.32 Solar energy cells. According to the Earth Policy Institute (July 2013), 60% of the world's solar energy cells are manufactured in China. Consider a random sample of 5 solar energy cells, and let x represent the number in the sample that are manufactured in China. In Section 4.4, we show that the probability distribution for x is given by the formula

$$p(x) = \frac{(5!)(.6)^x(.4)^{5-x}}{(x!)(5-x)!},$$

where $n! = (n)(n-1)(n-2)\dots(2)(1)$

a. Explain why x is a discrete random variable.
b. Find $p(x)$ for $x = 0, 1, 2, 3, 4,$ and 5.
c. Show that the properties for a discrete probability distribution are satisfied.
d. Find the probability that at least 4 of the 5 solar energy cells in the sample are manufactured in China.

a) x is a discrete random variable because we can specify all the values it can assume (0, 1, 2, 3, 4, 5) and the prob. associated w/ each value (given by the pdf in the problem.

b)
$$p(x) = \frac{(5!)(0.6)^x(0.4)^{5-x}}{(x!)(5-x)!} = \left(\frac{5!}{x!(5-x)!}\right)0.6^x(0.4)^{5-x} = \binom{5}{x}0.6^x(0.4)^{5-x}$$

← this is a binomial distribution $Bin(\underset{n}{5}, \underset{p}{0.6})$

```
> x <- seq(0,5,1)
> x
[1] 0 1 2 3 4 5
> dbinom(x, 5, 0.6)
[1] 0.01024 0.07680 0.23040 0.34560 0.25920 0.07776
> factorial(5) * 0.6^x * 0.4^(5-x) / (factorial(x) * factorial(5-x))
[1] 0.01024 0.07680 0.23040 0.34560 0.25920 0.07776
```

$$\boxed{\begin{aligned} p(0) &= 0.01024\\ p(1) &= 0.07680\\ p(2) &= 0.23040\\ p(3) &= 0.34560\\ p(4) &= 0.25920\\ p(5) &= 0.07776 \end{aligned}}$$

c)
```
> sum( dbinom(x, 5, 0.6) )
[1] 1
```

Trivially $p(x) \geq 0 \quad \forall x \in \{0,1,2,3,4,5\}$

d) P(at least 4 of 5 manufactured in China)

$= P(X = 4 \text{ OR } X = 5)$

$\overset{\text{II}}{=} P(X=4) + P(X=5) = p(4) + p(5) = 0.25920 + 0.07776$

$= \boxed{0.33696}$

4.52 Reliability of a manufacturing network. Refer to the *Journal of Systems Sciences & Systems Engineering* (Mar. 2013) study of the reliability of a manufacturing system that involves multiple production lines, Exercise 4.36 (p. 195). Consider, again, the network for producing integrated circuit (IC) cards with two production lines set up in sequence. The probability distribution of the maximum capacity level of each line is given in Exercise 4.36.

a. Find the mean maximum capacity for each line. Interpret the results practically.

b. Find the standard deviation of the maximum capacity for each line. Interpret the results practically.

a) $\mu = \sum_x x \cdot p(x)$

LINE 1: $\mu = 0 + 0.24 + 0.48 + 34.20 = \boxed{34.92}$

LINE 2: $\mu = 0 + 0.07 + 69.72 = \boxed{69.79}$

Line	Maximum Capacity, x	x^2	$p(x)$	$x \cdot p(x)$	$x^2 \cdot p(x)$
1	0	0	.01	$0.01 \cdot 0 = 0$	$0.01 \cdot 0 = 0$
	12	144	.02	$0.02 \cdot 12 = 0.24$	$0.02 \cdot 144 = 2.88$
	24	576	.02	$0.02 \cdot 24 = 0.48$	$0.02 \cdot 576 = 11.52$
	36	1296	.95	$0.95 \cdot 36 = 34.20$	$0.95 \cdot 1296 = 1231.2$
2	0	0	.002	$0.002 \cdot 0 = 0$	$0.002 \cdot 0 = 0$
	35	1225	.002	$0.002 \cdot 35 = 0.07$	$0.002 \cdot 1225 = 2.45$
	70	4900	.996	$0.996 \cdot 70 = 69.72$	$0.996 \cdot 4900 = 4880.4$

b) $Var = E[X^2] - (E[X])^2$

$= \sum_x x^2 p(x) - \mu^2$

LINE 1 $= (0 + 2.88 + 11.52 + 1231.2) - (34.92)^2 = 26.1936 = \sigma^2 \Rightarrow$ $\boxed{\sigma_1 = 5.118}$

LINE 2 $= (0 + 2.45 + 4880.4) - (69.79)^2 = 12.2059 = \sigma^2 \Rightarrow$ $\boxed{\sigma_2 = 3.494}$

```r
# Problem 2)
library(distr)
library(distrEx)

# cap <- c(0,12,24,36)
# p <- c(0.01, 0.02, 0.02, 0.95)

line1 <- DiscreteDistribution(supp = c(0,12,24,36), prob = c(0.01, 0.02, 0.02, 0.95))
line2 <- DiscreteDistribution(supp = c(0,35,70), prob = c(0.002, 0.002, 0.996))

E(line1)
E(line2)

var(line1)
var(line2)

sd(line1)
sd(line2)
```

```
> E(line1)
[1] 34.92
> E(line2)
[1] 69.79
> var(line1)
[1] 26.1936
> var(line2)
[1] 12.2059
> sd(line1)
[1] 5.117968
> sd(line2)
[1] 3.493694
```

$\sigma_2 < \sigma_1 \Rightarrow$ maximum capacity of line 1 is more likely to fall farther from its mean (34.9) than line 2 [from its respective mean (69.8)]

4.82 Assigning a passing grade. A literature professor decides to give a 20-question true–false quiz to determine who has read an assigned novel. She wants to choose the passing grade such that the probability of passing a student who guesses on every question is less than .05. What score should she set as the lowest passing grade?

$X \sim Bin(\underset{n}{\uparrow}{20}, \underset{p}{\uparrow}{0.5})$ is probability dist'n of correct answers given by guessing randomly

$P(\text{pass}) < 0.05$ $\quad P(X > \text{Pass-grade}) < 0.05$

$\iff 1 - P(X \leq \text{Pass-grade}) < 0.05$

$\iff P(X \leq \text{Pass-grade}) > 0.95$

$\iff F(\text{pass-grade}) > 0.95$

cdf

$\iff \text{Pass-grade} > F^{-1}(0.95; n=20, p=0.5)$

```
> qbinom(0.95, 20, 0.5)
[1] 14
```

$\text{Pass-grade} > 14$

$\boxed{\text{Lowest passing grade: } 15/20 \ (75\%)}$

4.104 Flaws in plastic-coated wire. The British Columbia Institute of Technology provides on its Web site (www.math.bcit.ca) practical applications of statistics to mechanical engineering. The following is a Poisson application. A roll of plastic-coated wire has an average of .8 flaws per 4-meter length of wire. Suppose a quality control engineer will sample a 4-meter length of wire from a roll of wire 220 meters in length. If no flaws are found in the sample, the engineer will accept the entire roll of wire. What is the probability that the roll will be rejected? What assumption did you make to find this probability?

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

$$\lambda = 0.8$$

$$P(\text{rejected}) = P(1 \text{ or no flaws})$$
$$= 1 - P(0 \text{ flaws})$$
$$= 1 - p(0; \lambda = 0.8)$$
$$= 1 - \frac{0.8^0 \cdot e^{-0.8}}{0!} = 1 - e^{-0.8} = \boxed{0.551}$$

Assumptions: # of flaws in 4 meters of wire follows a POISSON DISTRIBUTION

Characteristics of a Poisson Random Variable

1. The experiment consists of counting the number of times a certain event occurs during a given unit of time or in a given area or volume (or weight, distance, or any other unit of measurement).
2. The probability that an event occurs in a given unit of time, area, or volume is the same for all the units.
3. The number of events that occur in one unit of time, area, or volume is independent of the number that occur in other units.
4. The mean (or expected) number of events in each unit is denoted by the Greek letter lambda (λ).

4.126 Gender discrimination suit. The *Journal of Business & Economic Statistics* (July 2000) presented a case in which a charge of gender discrimination was filed against the U.S. Postal Service. At the time, there were 302 U.S. Postal Service employees (229 men and 73 women) who applied for promotion. Of the 72 employees who were awarded promotion, 5 were female. Make an inference about whether or not females at the U.S. Postal Service were promoted fairly.

	men	women	total
prom	67	5	72
not prom	162	68	230
total	229	73	302

$\Rightarrow$

	men	women	total
prom	0.2219	0.0166	0.2384
not prom	0.5364	0.2252	0.7616
total	0.7583	0.2417	302

Assuming binomial dist'n:

X = # of women promoted

$$X \sim Bin(\underset{n}{73}, \underset{\substack{p = \text{prob of} \\ \text{prom (man or woman)}}}{0.2384}) \Rightarrow E[X] = \overset{73 \times 0.2384}{17.4}$$

$$P(X \leq 10) = \underset{\text{cdf } Bin(73,\ 0.2384)}{F(10)} =$$

```
> pbinom(10, 73, 0.2384)
[1] 0.02404401
```

In other words, it is extremely unlikely that promotions were distributed fairly between men and women.