Introductory Statistics Tutorial Answers
Chapter 5 – Discrete Random Variables

Section A: Discrete Random Variables

1. (a)

y	5	6	7	8	9	10	11	12
pr($Y \le y$)	0.06	0.18	**0.34**	**0.55**	**0.80**	0.94	**0.99**	**1.00**

(b) (i) pr($Y \le 7$) = 0.34
(ii) pr($Y = 10$) = 0.14
(iii) pr($Y \ge 6$) = 1 – pr($Y \le 5$) = 1 – 0.06 = 0.94
(iv) pr($6 \le Y \le 10$) = pr($Y \le 10$) – pr($Y \le 5$) = 0.94 – 0.06 = 0.88
(v) pr($Y > 8$) = 1 – pr($Y \le 8$) = 1 – 0.55 = 0.45
(vi) pr($7 < Y < 12$) = pr($Y \le 11$) – pr($Y \le 7$) = 0.99 – 0.34 = 0.65

2. E(X) = 4.2, sd(X) = 0.6

Section B: Binomial Distribution

1. (a) $n = 10$, $p = 0.05$
(b) There is a fixed number of trials, 10. Each disk drive is a trial.
Each trial has 2 outcomes: Disk drive malfunctions or disk drive does not malfunction.
The disk drives are independent.
The probability that a disk drive malfunctions is constant.
(c) The first two assumptions will be satisfied.
The disk drives may not be independent. Disk drives could be made from the same batch of materials or may have the same systematic fault.
The probability of a disk drive malfunctioning will not be constant because it will depend on how a disk drive is used.
(d) (i) pr($X = 0$) = 0.5987
(ii) pr($X \ge 2$) = 1 – pr($X \le 1$) = 1 – 0.9139 = 0.0861
(iii) pr($1 \le X \le 4$) = pr($X \le 4$) – pr($X \le 0$) = 0.9999 – 0.5987 = 0.4012
(e) E(X) = np = 10 x 0.05 = 0.5
sd(X) = $\sqrt{np(1-p)}$ = $\sqrt{10 \times 0.05 \times 0.95}$ = 0.689 (to 3 d.p.)

Section C: Poisson Distribution

1. (a) The events occur at a constant average rate of λ per unit time.
(b) Occurrences are independent of one another.
(c) The probability of 2 or more occurrences in a time interval of length d tends to zero, as d tends to zero.
2. (a) pr($X \le 5$) = 0.6510
(b) pr($X = 3$) = 0.1517
(c) pr($X \ge 2$) = 1 – pr($X \le 1$) = 1 – 0.0477 = 0.9523
(d) pr($2 \le X \le 6$) = pr($X \le 6$) – pr($X \le 1$) = 0.7908 – 0.0477 = 0.7431
3. (3)
4. (1)
5. (4)

Section D: Choosing an Appropriate Probability Model

1. There is a fixed number of trials, 25 boxes. Each trial has 2 outcomes, water damaged or not. But sampling is from a finite population of 1200 boxes so conditions will not be the same for each trial. However $\frac{n}{N} = \frac{25}{1200} \approx 0.02 < 0.1$ so conditions for each trial will be approximately the same.
An appropriate model is $X \sim$ Binomial ($n = 25$, $p = \frac{200}{1200}$)
2. Neither.
3. The average rate may not be exactly constant over the 6-month period because strokes may occur more often at one time of the year.
Deaths due to strokes are likely to be independent.
Deaths by stroke are unlikely to occur at exactly the same time.
An appropriate model is $X \sim$ Poisson ($\lambda = 275$)
4. There are 45 trials (each person who has bought a new car is a trial). There are 2 outcomes for each trial (wins a major prize or not). Each person's result is independent of the other people's result. The probability of a major prize being won is 0.1 for each person.
An appropriate model is $X \sim$ Binomial ($n = 45$, $p = 0.1$)
5. Fixed number of trials, 200 students. Each trial has 2 outcomes, student has not done any mathematics in the previous two years or they have. But sampling is from a finite population of 1400 boxes so conditions will not be the same for each trial. However $\frac{n}{N} = \frac{200}{1400} \approx 0.14 > 0.1$ so conditions for each trial will not be approximately the same.
Neither the Binomial distribution nor the Poisson distribution is appropriate.