INF 1004 Mathematics 2
Tutorial #3 Solutions

Woon Jun Wei
2200624

January 23, 2023

Question 1

A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used. Whereas the phrase "free money" is used in only 1% of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?

Sample Solutions

Let S be the event that an email is spam
Let F be the event that an email has the "free money" phrase
$P(S) = 0.8$
$P(S^c) = 0.2$
$P(F \mid S) = 0.1$ (Free money given spam)
$P(F \mid S^c) = 0.01$ (Free money given non-spam)
$P(S \mid F)$ Probability ofspam given free money phrase?

Using bayes rule

$$\begin{aligned} P(S \mid F) &= \frac{P(F \mid S) \cdot P(S)}{P(F)} \\ &= \frac{0.1 \cdot 0.8}{0.8 \cdot 0.1 + 0.2 \cdot 0.01} \\ &= \frac{40}{41} \\ &\approx 0.9756097561 \end{aligned}$$

Question 2

A sample of people, who commute regularly from a town in surrey into London, were asked for an estimate of the time taken on their most recent journey. The replies are summarized in the given table.

Time in minutes	Frequency
35 to 45	12
45 to 55	54
55 to 65	68
65 to 85	41
85 to 105	23

a) Calculate the estimate of the mean and the standard deviation of these times.

b) What is the median?

c) What is the mode?

Sample Solutions

Time in minutes	f	Midpoint = x	$x_i f_i$	$x_i^2 f_i$	cf
35 to 45	12	40	480	19200	12
45 to 55	54	50	2700	135000	66
55 to 65	68	60	4080	244800	134
65 to 85	41	75	3075	230625	175
85 to 105	23	95	2185	207575	198
Sum:	198		12520	837200	

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{12520}{198} = \boxed{63.23}$$

$$\begin{aligned} s^2 &= \frac{1}{n-1}\left[\sum x_i^2 f_i - \frac{(\sum x_i f_i)^2}{n}\right] \\ &= \frac{1}{198-1}\left[837200 - \frac{12520^2}{198}\right] = 231.1234 \end{aligned}$$

$$s = \sqrt{231.1234} = \boxed{15.202}$$

$$\begin{aligned} Q &= L_m + \left(\frac{\frac{n}{2} - cf_{m-1}}{f_m}\right) w_m \\ &= 55 + \frac{\frac{198}{2} - 66}{68} \cdot (65 - 55) \\ &= 59.85294 \\ &\approx \boxed{59.85 \text{ minutes}} \end{aligned}$$

b)
$\frac{198}{2} = 99$
Median class: $\boxed{55 - 65}$
c)
Mode class: $\boxed{55 - 65}$

Question 3

Maurice works at home. At 2 pm he decides to take a break to buy a copy of the chronicle newspaper. There are three nearby newsagents: Arif, Bob and Carol. However, by 2 pm they may have sold all their chronicles and so have none available. The independent probabilities that they have a chronicle available at 2 pm are: Arif 0.4, Bob 0.7, Carol 0.25

(a) Find the probability that none of the three newsagents have the chronicles at 2 pm.?

(b) Maurice decides to visit the newsagent in turn until he obtains a chronicle or until he has visited all three. He tosses a coin. If it lands heads he will visit the three newsagents in the order Bob, Arif, Carol. If it lands tail, he will visit them in order Carol, Arif, Bob. Find the probability that he will obtain a chronicle from Arif.

Sample Solutions

Let the event of Ariv having the chronicles at 2pm be A
Let the event of Bob having the chronicles at 2pm be B
Let the event of Carol having the chronicles at 2pm be C
$P(A) = 0.4, P(B) = 0.7, P(C) = 0.25$
$P(A^c) = 0.6, P(B^c) = 0.3, P(C^c) = 0.75$

a)
None of the three news agents have the chronicles at 2pm

$$P(A^c \cap B^c \cap C^c) = P(A^c) \cdot P(B^c) \cdot P(C^c) = 0.6 \cdot 0.3 \cdot 0.75 = \frac{27}{200} = \boxed{0.135}$$

b)
Find the probability that he will obtain a chronicle from Arif

$$0.5 \cdot 0.3 \cdot 0.4 + 0.5 \cdot 0.75 \cdot 0.4 = \boxed{0.21}$$

Question 4

The following table summarizes the number of late arrivals of all students who attend a certain school on 5 mornings of particular week. (Show your steps for all of the following.)

Number of late arrivals	Number of Pupils
0	275
1	111
2	33
3	12
4	13
5	16

(a) Let random variable X=number of late arrivals. Is the random variable discrete or continuous?

(b) Draw the probability distribution for X. Calculate the mean and the standard deviation of the data in the table.

(c) Calculate the expected value.

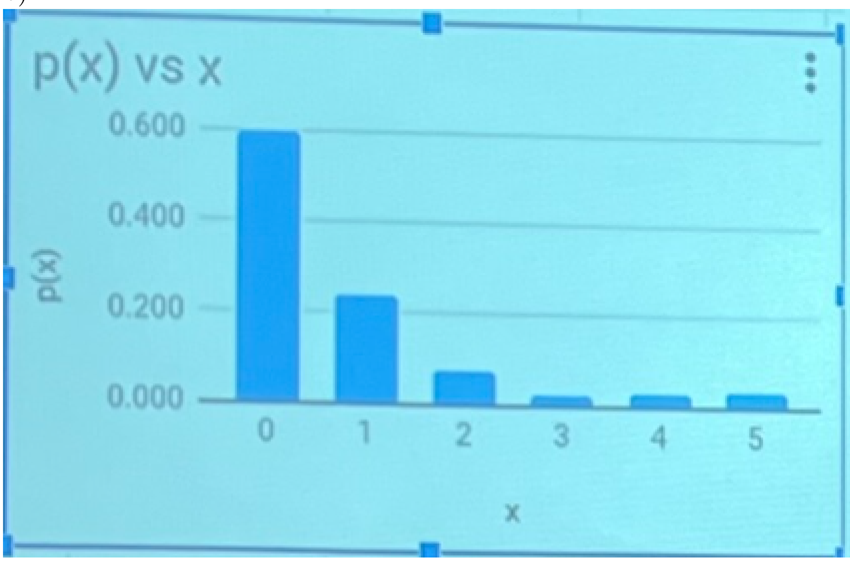
Sample Solutions

Part 1

a)

The random variable is $\boxed{\textit{discrete}}$.

b)



x	f	$P(x) = \frac{f}{460}$	$P(x) \cdot x$	(x-μ)^2	(x-μ)$^2 \cdot P(x)$
0	275	0.598	0	0.5625	0.336
1	111	0.241	0.241	0.0625	0.01508
2	33	0.072	0.143	1.5625	0.11209
3	12	0.026	0.0782	5.0625	0.132065
4	13	0.028	0.1130	10.5625	0.29850
5	16	0.035	0.173913	18.0625	0.62826
sum	460	1	0.75		1.5222

$$E(X) = \mu = \sum xP(X = x) = \boxed{0.75}$$

$$Var(X) = \sigma^2 = \sum (x - \mu)^2 P(X = x) = \boxed{1.52228}$$
$$\sigma = \sqrt{1.52228} = \boxed{1.23}$$

c)

$$E(X) = \boxed{0.75}$$

Question 5

A Terrence walks to school every morning. The probability that he arrives late is 0.15 and independent of whether he arrives late on any other morning. For a week, in which he decides to walk to school on five mornings, find:

(a) Probability he arrives late on two or fewer mornings.

(b) Probability he arrives late on three or more mornings.

(c) Find the mean and standard deviation of the number of mornings on which he arrives late.?

Sample Solutions

Let X be the number of days he arrives late.
n = 5
p(X) = 0.15

$$X \sim B(5, 0.15)$$

Part 1

$$\begin{aligned} P(X \leq 2) &= P(X=0) + P(X=1) + P(X=2) \\ &= \binom{5}{2}(o.15)^2(1-0.15)^{5-2} + \binom{5}{1}(0.15)(1-0.15)^{5-1} + \binom{5}{0}(1-0.15)^5 \\ &= 0.973388125 \\ &\approx \boxed{0.9734} \end{aligned}$$

Part 2

$$\begin{aligned} P(X \geq 3) &= P(X=3) + P(X=4) + P(X=5) \\ &= \binom{5}{3}(0.15)^3(1-0.15)^{5-3} + \binom{5}{4}(0.15)^4(1-0.15)^{5-4} + \binom{5}{5}(0.15)^5 \\ &= 0.026611875 \\ &\approx \boxed{0.0266} \end{aligned}$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) \\ &= 1 - 0.9734 \\ &= 0.026611875 \\ &\approx \boxed{0.0266} \end{aligned}$$

Part 3

$$\begin{aligned} E(x) &= 5 \cdot 0.15 = \boxed{0.75} \\ Var(X) &= 0.75(1-0.15) = 0.6375 \\ s &= \sqrt{0.6375} \approx \boxed{0.799} \text{ (answer given was 0.7984)} \end{aligned}$$

Question 6

Bob is a student who travels to and from University by bus. He believes that the probability of having to wait more than six minutes to catch a bus is 0.4 and is independent of the time of day and direction of travel.

(a) Assuming Bob beliefs are correct. Calculate values for the mean and standard deviation of the number of times he has to wait more than six minutes to catch a bus in a week when he catches 10 buses.

(b) During a thirteen-week period, the number of times (out of 10) he had to wait more than six minutes to catch a bus were as follows:

$$4, 8, 8, 9, 3, 2, 2, 7, 0, 1, 5, 2, 0$$

(i) Calculate the mean and standard deviation of this data.

(ii) Does the answer in (b)i support Bob's beliefs that the probability of having to wait more than six minutes to catch a bus is constant as 0.4 regardless of any factors?

Sample Solutions

a)
Binomial n = 10, p = 0.4
Mean is $np = 10 \cdot 0.4 = \boxed{4}$
Variance is $np(1-p) = 10 \cdot 0.4 \cdot 0.6 = \boxed{2.4}$
Standard deviation is $s = \sqrt{2.4} \approx \boxed{1.549}$

b)

$$4, 8, 8, 9, 3, 2, 2, 7, 0, 1, 5, 2, 0$$

Number waits per week (X)	Frequency	$x_i f_i$	$x_i^2 f_i$
0	2	0	0
1	1	1	1
2	3	6	12
3	1	3	9
4	1	4	16
5	1	5	25
6	0	0	0
7	1	7	49
8	2	16	128
9	1	9	81
total	13	51	341

i)

Mean $= \frac{51}{13} = \boxed{3.923076923}$
Variance $= \frac{1}{13-1} \sum (x_i - 3.6923)^2 = \boxed{10.07692308}$
Standard deviation $= \sqrt{10.07692308} \approx \boxed{3.17441}$
ii)
The mean is similar to binomial's mean, but the standard deviation is significantly higher which both indicate that 'p' isn't constant. This doesn't support Bob's belief.

Question 7

A shop sells two types of bath cubes: relaxing and invigorating. The owner of a hotel buys 50 cubes from this shop. For each cube, there is independently a probability of 0.4 that it will be relaxing.

(a) Find the probability that more than 90% of the 50 cubes will be relaxing cubes.

(b) The 50 cubes included 22 relaxing cubes. Give a reason, whether or not a binomial distribution will provide an appropriate model for the random variable, R, in each of the following cases:

(i) A guest at the hotel randomly selects one of the 50 cubes. If it is not a relaxing cube, he replaces it and again selects a cube at random. He continues this procedure until he obtains a relaxing cube. The random variable R denotes the number of cubes he selects until he obtains a relaxing cube.

(ii) The owner randomly selects 20 of the 50 cubes and places them in a bowl. The random variable R denotes the number of relaxing cubes placed in the bowl; each being evaluated as one trial.

My Solution

$$X \sim B(50, 0.4)$$

n = 50, p = 0.4. Let X be the number of cubes of the 50 which are relaxing cubes
a)
90% of 50 = 45

$$\begin{aligned} P(x > 45) &= P(x = 46) + P(x = 47) + P(x = 48) + P(x = 49) + P(x = 50) \\ &= \binom{50}{46}(0.4)^{46}(0.6)^{50-46} + \binom{50}{47}(0.4)^{47}(0.6)^{50-47} \\ &+ \binom{50}{48}(0.4)^{48}(0.6)^{50-48} + \binom{50}{49}(0.4)^{49}(0.6)^{50-49} \\ &+ \binom{50}{50}(0.4)^{50}(0.6)^{50-50} \\ &= \boxed{1.65639 \cdot 10^{-14}} \end{aligned}$$

b)
i)
R = number of cubes selected until a relaxing cube. $\boxed{\text{Not binomial as n is not fixed}}$
ii)
R = number of relaxing cubes placed in a bowl.
Perhaps not enough information; $\boxed{\text{but not binomial as p is not constant}}$