Introductory Statistics Tutorial
Chapter 5 – Discrete Random Variables

Section A: Discrete Random Variables

1. Random variable Y has the following probability function:

y	5	6	7	8	9	10	11	12
$\text{pr}(Y = y)$	0.06	0.12	0.16	0.21	0.25	0.14	0.05	0.01

(a) Fill in the gaps in the cumulative probability function.

y	5	6	7	8	9	10	11	12
$\text{pr}(Y \leq y)$	0.06	0.18				0.94		

(b) Find the probability that:

(i) Y is no more than 7

(ii) Y is equal to 10

(iii) Y is at least 6

(iv) Y is at least 6 and at most 10

(v) Y is greater than 8

(vi) Y is more than 7 but less than 12.

2. Random variable X has the following probability function:

x	3	4	5
$\text{pr}(X = x)$	0.1	0.6	0.3

Calculate $E(X)$ and $\text{sd}(X)$.

Section B: Binomial Distribution

1. The manufacturer of disk drives for a well-known brand of computers expects 5% of the drives to malfunction during the computer's warranty period. Let X be the number of disk drives, in a batch of 10 randomly selected disk drives, which malfunction during this period. X has a Binomial distribution.

(a) Identify n and p, the parameters of the Binomial random variable.

(b) In the context of this exercise, state the assumptions required for X to have a Binomial distribution.

(c) Are the assumptions satisfied here?

(d) Using the MINITAB output below, calculate the probability that:

(i) no disk drive will malfunction during the warranty period.

(ii) at least two disk drives will malfunction during the warranty period.

(iii) between 1 and 4 (inclusive) disk drives will malfunction during the warranty period.

(e) Calculate the mean and standard deviation of X.

MINITAB Output

```
Binomial with n = 10 and p = 0.0500000

       x      P( X = x )                x     P( X <= x )
    0.00        0.5987               0.00        0.5987
    1.00        0.3151               1.00        0.9139
    2.00        0.0746               2.00        0.9885
    3.00        0.0105               3.00        0.9990
    4.00        0.0010               4.00        0.9999
    5.00        0.0001               5.00        1.0000
```

Section C: Poisson Distribution

1. What three conditions are necessary for X to have a Poisson distribution?

(a)

(b)

(c)

2. Use the MINITAB output below to calculate the following probabilities.

(a) If $X \sim$ Poisson ($\lambda = 4.8$) then pr($X \leq 5$) =

(b) If $X \sim$ Poisson ($\lambda = 4.8$) then pr($X = 3$) =

(c) If $X \sim$ Poisson ($\lambda = 4.8$) then pr($X \geq 2$) =

(d) If $X \sim$ Poisson ($\lambda = 4.8$) then pr($2 \leq X \leq 6$) =

MINITAB Output

```
Poisson with mu = 4.80000

        x     P( X = x )              x     P( X <= x )
     0.00        0.0082            0.00          0.0082
     1.00        0.0395            1.00          0.0477
     2.00        0.0948            2.00          0.1425
     3.00        0.1517            3.00          0.2942
     4.00        0.1820            4.00          0.4763
     5.00        0.1747            5.00          0.6510
     6.00        0.1398            6.00          0.7908
```


Questions 3 to 5 are based on questions from the Test in Semester 2 1999

Questions 3 to **5** refer to the following information.

Suppose that a hospital has, on average, 2.2 births each hour. Let X be the number of births in an hour at this hospital.

3. Which of the following statements, if true, would cause concern about using the Poisson distribution to model the distribution of the random variable X?

I: Mothers are more likely to give birth late at night or in the early hours of the morning.
II: It is not possible for 2.2 births to occur in a particular hour.
III: Multiple births, such as twins or triplets, can occur.
IV: Two pregnant women, who both intend having their babies delivered at the hospital, have the same estimated due date for the birth of their baby.

(1) II and III only.
(2) I, II and IV only.
(3) I and III only.
(4) I, II and III only.
(5) III only.

Questions 4 and **5** refer to the following additional information.

Suppose a Poisson ($\lambda = 2.2$) distribution is used as a model for the distribution of the random variable X.

4. Use the MINITAB output below to find the probability that fewer than 3 babies are born at this hospital in a particular hour is approximately:

(1) 0.623
(2) 0.819
(3) 0.377
(4) 0.181
(5) 0.197

5. Over a period of 5 consecutive hours the probability that exactly 15 babies are born at this hospital can be calculated by:

(1) pr($X = 3$) where $X \sim$ Poisson ($\lambda = 2.2$)
(2) 5 x pr($X = 3$) where $X \sim$ Poisson ($\lambda = 2.2$)
(3) $(\text{pr}(X = 3))^5$ where $X \sim$ Poisson ($\lambda = 2.2$)
(4) pr($Y = 15$) where $Y \sim$ Poisson ($\lambda = 11.0$)
(5) 2 x pr($Y = 7.5$) where $Y \sim$ Poisson ($\lambda = 5.5$)

MINITAB Output

```
Poisson with mu = 2.20000

        x     P( X = x )              x     P( X <= x )
     0.00        0.1108            0.00          0.1108
     1.00        0.2438            1.00          0.3546
     2.00        0.2681            2.00          0.6227
     3.00        0.1966            3.00          0.8194
     4.00        0.1082            4.00          0.9275
```

Section D: Choosing an Appropriate Probability Model

In questions 1 to 5, state an appropriate model for the distribution of random variable X. Choose from: Binomial, Poisson, or neither Binomial nor Poisson. State the parameters of your model and briefly discuss the assumptions underlying the model.

1. A fire occurred at a warehouse in Wellington. 200 boxes out of 1200 were water damaged when the fire was extinguished. These boxes were mixed up with all of the other boxes while the warehouse was being repaired. The company decided to sell all boxes at a reduced price (informing prospective buyers that some of the boxes were damaged). The purchasing company randomly selects 25 boxes and delivers them to their retail outlet. Let X be the number of these 25 boxes that have been damaged.

2. A shuttle bus service between the Tamaki Campus and City Campus has a capacity of 30 passengers. Information collected over a long period of time shows that, on average, 40% of passengers are female. For a Tamaki Campus to City Campus journey, let X be the number of male passengers who get on the shuttle before the first female boards the shuttle.

3. The number of deaths due to strokes in the Auckland region each year varies randomly with 550 deaths per year, on average. Let X be the number of deaths due to strokes in a given 6-month period.

4. An Auckland car sales company is offering customers an incentive when they purchase a new car. At the end of the month each customer purchasing a car will be offered an opportunity to win either a major prize or a minor prize determined by the spinning of a wheel. There is a one-in-ten chance of winning a major prize (a \$2000 refund). If they don't win a major prize, they win a minor prize (a \$500 refund). At the end of a month 45 cars had been sold. Let X be the number of customers who receive a major prize.

5. In a particular semester 1400 students are enrolled in the Stage I Statistics papers. Of these students, 280 had not done any mathematics in the previous two years. One of the streams had 200 students randomly allocated to it. Let X be the number of students in that lecture stream who had not done any mathematics in the previous two years.