MAST20004 Probability

Tutorial Set 5

1. A Poisson distribution has $\mathbb{P}(X = 1) = \mathbb{P}(X = 2)$. Find $\mathbb{P}(X = 0)$.

   Solution: For a Poisson random variable with parameter λ,

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

   So $\mathbb{P}(X = 1) = \mathbb{P}(X = 2)$ implies that

$$\lambda e^{-\lambda} = \frac{\lambda^2 e^{-\lambda}}{2},$$

   and so $2\lambda = \lambda^2$ which, in turn implies that $\lambda = 0$ or 2. The physical solution is $\lambda = 2$, in which case $\mathbb{P}(X = 0) = e^{-2}$.

2. Let X be the number of zeros in $n = 50$ independent random decimal digits (each digit from $0, 1, \ldots, 9$ has equal probability of occurring). Find
   (a) the exact probability $\mathbb{P}(X = 2)$;
   (b) the Poisson approximation to this probability.

   Solution:

   (a) X has a binomial distribution with parameters $n = 50$ and $p = 1/10$. Thus the probability that $X = 2$ is given by

$$P_X(2) = \binom{50}{2}(1/10)^2(9/10)^{48} = 0.0779.$$

   (b) The Poisson approximation assumes that the distribution of X is approximated by a Poisson distribution with parameter $\lambda = np = 5$. Thus,

$$\mathbb{P}(X = 2) \approx 5^2 e^{-5}/2 = 0.0842.$$

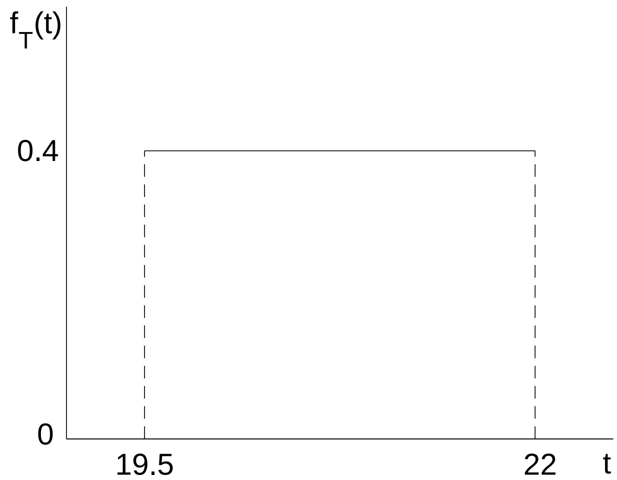
3. The temperature T in a restaurant is maintained within the range 19.5°C to 22°C. The digital thermometer on the control panel of the plant shows the value of T rounded to the nearest integer. Assuming that T is uniformly (continuously) distributed in the indicated range,
   (a) write down and plot the probability density function and distribution function of T,
   (b) write down and plot the probability mass function and distribution function of the thermometer reading U, and
   (c) write down the expected value and variance of T and U. Comment on your answers.

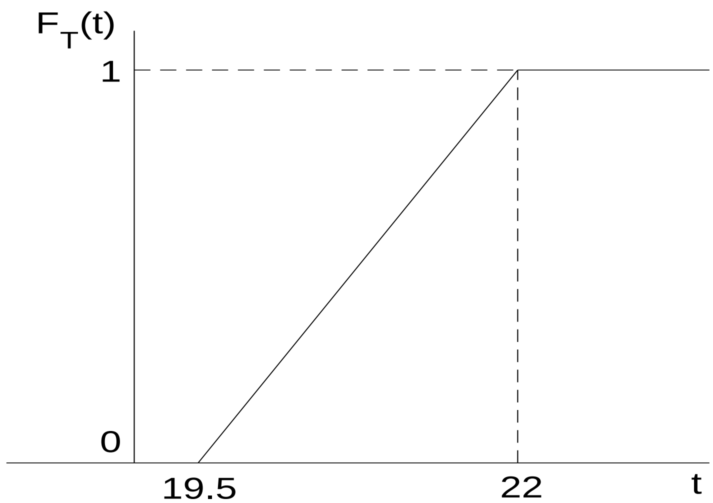
   Solution:

   (a)

$$f_T(t) = \begin{cases} 2/5 & \text{if } t \in [19.5, 22] \\ 0 & \text{otherwise,} \end{cases}$$

$$F_T(t) = \begin{cases} 0 & \text{if } t < 19.5 \\ \dfrac{t - 19.5}{2.5} & \text{if } t \in [19.5, 22] \\ 1 & \text{if } t > 22. \end{cases}$$

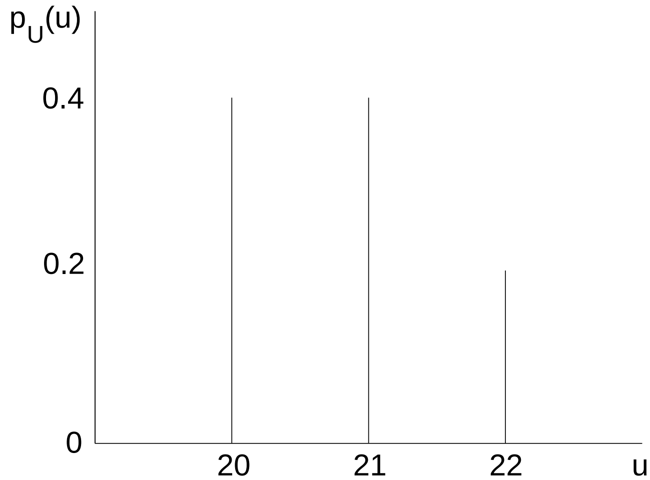


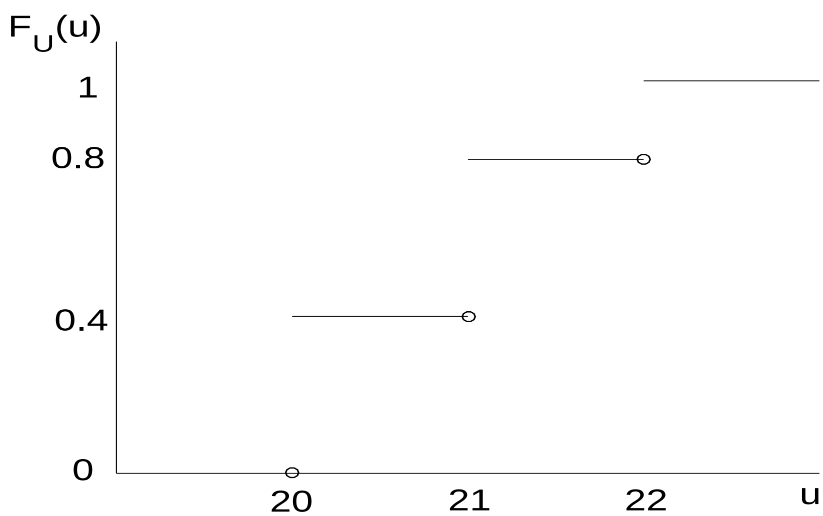


(b)

$$p_U(u) = \begin{cases} 2/5 & \text{if } u = 20 \\ 2/5 & \text{if } u = 21 \\ 1/5 & \text{if } u = 22, \end{cases}$$

$$F_U(u) = \begin{cases} 0 & \text{if } u < 20 \\ 2/5 & \text{if } u \in [20, 21) \\ 4/5 & \text{if } u \in [21, 22) \\ 1 & \text{if } u \geq 22. \end{cases}$$





(c) From lectures, $\mathbb{E}[T] = (22 + 19.5)/2 = 20.75$ and $V(T) = (22 - 19.5)^2/12 = 0.5208$.

$\mathbb{E}[U] = 20(2/5) + 21(2/5) + 22(1/5) = 104/5 = 20.8$. $\mathbb{E}[U^2] = 400(2/5) + 441(2/5) + 484(1/5) = 2166/5 = 433.2$ and so $V(U) = 433.2 - (20.8)^2 = 0.56$.

The mean of the rounded-off variable U is a reasonable approximation for the mean temperature. The variance of U slightly overestimates the variance of T.

4. A call centre employs five operators who receive calls independently of one another, each according to a Poisson distribution with parameter $\lambda = 0.5t$ for any given t-minute period.

(a) What is the probability that during a given 2-minute period, the first operator receives no requests?

(b) What is the probability that during a given 2-minute period, exactly four of the five operators receive no requests?

(c) Write an expression for the probability that during a given 2-minute period, all of the operators receive exactly the same number of requests.

Solution:

(a) Let X be the number of calls the first operator receives in a 2-minute period. Then $X \stackrel{d}{=} \text{Pn}(1)$, and $\mathbb{P}(X = 0) = e^{-1} = 0.3679$.

(b) $\mathbb{P}(\text{exactly 4 of the 5 operators receive no requests}) = \binom{5}{4} e^{-4} \times (1 - e^{-1}) = 0.0579$.

(c) $\mathbb{P}(\text{all operators receive the same number of requests}) = e^{-5} \sum_{k=0}^{\infty} (k!)^{-5} = 0.0137$.

5. Consider a sequence of Bernoulli trials where the probability of success is p. Let X be the number of trials until the rth success.
   (a) Which values can X take?
   (b) Derive the probability mass function p_X.
   (c) Write down expressions for $\mathbb{E}[X]$ and $V(X)$.
   (d) Let Y be the number of successes in x trials. Show that $\mathbb{P}(X \leq x) = \mathbb{P}(Y \geq r)$.

   Solution:

   (a) $S_X = \{r, r+1, r+2, \ldots\}$.

   (b) The probability mass function $p_X(x)$ can be derived in two ways.

   (i) The probability of a sequence of r successes and $x-r$ failures with the rth success occurring on the xth trial is $p^r(1-p)^{x-r}$. Such a sequence can occur in $\binom{x-1}{r-1}$ different ways. Therefore, for $x = r, r+1, r+2, \ldots$,

$$p_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}.$$

   (ii) Let $Z \stackrel{d}{=} \mathrm{Nb}(r,p)$. That is, Z counts the number of failures before the rth success (see Slides 184–185). Then $X = Z + r$ (see Slide 192). Then, for $x = r, r+1, r+2, \ldots$,

$$\begin{aligned} p_X(x) &= p_Z(x-r) \\ &= \binom{x-1}{r-1} p^r (1-p)^{x-r}. \end{aligned}$$

   (c) From Slide 193 we have

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X+r] \\ &= \mathbb{E}[X] + r \\ &= \frac{r(1-p)}{p} + r \\ &= \frac{r}{p}, \end{aligned}$$

   and

$$\begin{aligned} V(X) &= V(Z+r) \\ &= V(Z) \\ &= \frac{r(1-p)}{p^2}. \end{aligned}$$

   (d) We have $Y \stackrel{d}{=} \mathrm{Bi}(x,p)$. Then, for $x = r, r+1, r+2, \ldots$,

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(\text{at most } x \text{ trials until the } r\text{th success}) \\ &= \mathbb{P}(\text{at least } r \text{ successes in } x \text{ trials}) \\ &= \mathbb{P}(Y \geq r). \end{aligned}$$

   See Slide 196.

6. (A harder question) The number of emissions from a radioactive source that occur in the time period $[0, t]$ follows a Poisson distribution with parameter λ. Each of these emissions is detected by a Geiger counter with probability p or missed with probability $1 - p$. What is the distribution of the number of detected emissions in $[0, t]$?

   Solution: Let X be the number of emissions detected in $[0, t]$ and N be the number of emissions that actually occurred. Then, conditional on $N = n$, X is binomially-distributed with parameters n and p. So, for $x \leq n$,

$$\mathbb{P}(X = x | N = n) = \binom{n}{x} p^x (1-p)^{n-x}$$

   Now we use the Law of Total Probability to write $\mathbb{P}(X = x) = \sum_{n=x}^{\infty} \mathbb{P}(X = x | N = n)\mathbb{P}(N = n)$ and so

$$\begin{aligned} \mathbb{P}(X = x) &= \sum_{n=x}^{\infty} \binom{n}{x} p^x (1-p)^{n-x} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= e^{-\lambda} \frac{(\lambda p)^x}{x!} \sum_{n=x}^{\infty} \frac{(\lambda(1-p))^{n-x}}{(n-x)!} \\ &= e^{-\lambda} \frac{(\lambda p)^x}{x!} e^{\lambda(1-p)} \\ &= \frac{e^{-\lambda p}(\lambda p)^x}{x!}, \end{aligned}$$

   and so X has a Poisson distribution with parameter λp.

MAST20004 Probability

Computer Lab 5

In this lab you

- investigate the shape of the pmf for a negative binomial random variable for different values of its parameters.
- study the spatial distribution of various facilities in the city of Coventry in England.

Exercise A - Negative binomial distribution pmf

1. The m-file **Lab5ExA.m** calculates and plots the pmf for a negative binomial random variable. When you run the program in the command window you are prompted to enter the values for the parameters r and p. This program handles the generalised form of the negative binomial so it works for any $r > 0$, not just integer r (see lecture slides 185–187).

   (a) Copy the program **Lab5ExA.m** from the server and study it. Using lecture slide 193, check that the formulae used for the mean and standard deviation are accurate. Using lecture slide 185, make sure you understand how the code calculating the negative binomial coefficient $\binom{-r}{z}$ works.

   (b) Run the program for $r = 3$ and $p = 0.5$. Verify using the pmf directly that the probabilities plotted at 1 and 2 are actually identical. Referring to $r(z)$ as defined on lecture slide 194, would you predict each of the values $r(1)$, $r(2)$, and $r(3)$ to be greater than, equal to, or less than 1? Check your predictions by hand calculations.

   (c) Run the program for $r = 1.5$ and $p = 0.5$. Verify the value of the probability mass function at $z = 3$ directly using the pmf on slide 188.

   (d) Run the program for a variety of r and p values including non-integer r. Summarise the effect of the parameters on the symmetry (or lack of it) of the distribution. For what values does the distribution appear similar to a 'bell-shaped' curve?

Exercise B - Coventry spatial data

1. As introduced in lectures, the Poisson distribution counts the number of discrete events occurring in continuous time. The Poisson distribution can also be used to model the positions of points in two-dimensional space where those points are 'scattered uniformly'. In this exercise we look at the distribution of various facilities, specifically fish and chip shops, churches and post offices - across the city of Coventry in England - to see if they fit the Poisson model.

   Copy the Excel spreadsheet Coventry.xls from the MAST20004 folder and look at the Coventry map (try zooming the view to make the map larger - this will help you later). The basic idea is to divide the total map area into a number of squares and count the number of 'events' (here facilities) in each square. This distribution should be Poisson. Also we need to estimate the Poisson parameter first before comparing the fit.

   Work through all the spreadsheet questions. The spreadsheet instructions are self explanatory and answers are available once you have completed the exercises so you can check your results. This spreadsheet (as well as many other interesting exercises) came from the website maintained by Neville Hunt and Sidney Tyrrell of Coventry University.