MAST20004 Probability

Tutorial Set 8

1. If X and Y are independent random variables with pmf given by

$$p(k) = \begin{cases} 0.4 & \text{if } k = 0 \\ 0.3 & \text{if } k = 1 \\ 0.2 & \text{if } k = 2 \\ 0.1 & \text{if } k = 3 \\ 0 & \text{otherwise,} \end{cases}$$

derive the pmf of $Z = X + Y$.

Solution: $\mathbb{P}(Z = z) = \sum_{x=0}^{z} p(x)p(z-x)$. So, the probability mass function of Z is

$$p_Z(z) = \begin{cases} 0.16 & \text{if } z = 0 \\ 0.24 & \text{if } z = 1 \\ 0.25 & \text{if } z = 2 \\ 0.20 & \text{if } z = 3 \\ 0.10 & \text{if } z = 4 \\ 0.04 & \text{if } z = 5 \\ 0.01 & \text{if } z = 6 \\ 0 & \text{otherwise.} \end{cases}$$

2. If U and V are independent random variables such that $U \stackrel{d}{=} \text{Bi}(m, p)$ and $V \stackrel{d}{=} \text{Bi}(n, p)$, find the distribution of the random variable $W = U + V$

   (a) by thinking about the physical meaning of the random variables, and

   (b) by using the result on page 400 of the lecture slides.

Solution:

(a) U is the number of successes in m trials and V is the number of successes in n trials, all with probability p of success. Therefore $U+V$ is the number of successes in $m+n$ trials with probability p and so $W \stackrel{d}{=} \text{Bi}(m+n, p)$.

(b) For $0 \le w \le m+n$,

$$\begin{aligned} p_W(w) &= \sum_{u=0}^{m} p_U(u)p_V(w-u) \\ &= \sum_{u=0}^{m} \binom{m}{u} p^u (1-p)^{m-u} \binom{n}{w-u} p^{w-u}(1-p)^{n-w+u} I(0 \le w-u \le n) \\ &= \sum_{u=\max(0,w-n)}^{\min(w,m)} \binom{m}{u}\binom{n}{w-u} p^w (1-p)^{m+n-w} \\ &= \binom{m+n}{w} p^w (1-p)^{m+n-w}, \end{aligned}$$

where the last equation follows from the results on Slides 209–212 from lectures (replace D with m, x with u, n with w, and N with $m+n$).

3. Let X denote the amount (in litres) of petrol stocked by a service station at the beginning of a week and suppose that X has a uniform distribution over the interval $[10000, 20000]$. Suppose the amount Y of petrol sold during a week has a uniform distribution over the interval $[10000, X]$.

   (a) Find the joint density function of X and Y.

   (b) If the service station fills up to $15,000$ litres at the beginning of a week, what is the probability that the amount of petrol sold in that week is greater than $12,000$ litres?

   (c) If it is known that the service station sold $12,500$ litres, what is the probability that the amount stocked was greater than $15,000$ litres?

   Solution:

   (a) For $x \in (10000, 20000]$ and $y \in [10000, x]$,

$$\begin{aligned} f_{X,Y}(x, y) &= f_{Y|X}(y|x) f_X(x) \\ &= \frac{1}{10000(x - 10000)}. \end{aligned}$$

   (b)

$$\begin{aligned} \mathbb{P}(Y \geq 12000 | X = 15000) &= \int_{12000}^{15000} \frac{1}{5000} dy \\ &= 3/5. \end{aligned}$$

   (c) For $y \in (10000, 20000]$,

$$\begin{aligned} f_Y(y) &= \int_y^{20000} \frac{1}{10000(x - 10000)} dx \\ &= \frac{\log(10000) - \log(y - 10000)}{10000}. \end{aligned}$$

   Therefore

$$f_{X|Y}(x|y) = \frac{1}{(x - 10000)\left(\log(10000) - \log(y - 10000)\right)}$$

   and

$$\begin{aligned} \mathbb{P}(X \geq 15000 | Y = 12500) &= \int_{15000}^{20000} \frac{1}{(x - 10000)\log(4)} dx \\ &= \frac{1}{\log(4)} \left[\log(x - 10000)\right]_{15000}^{20000} \\ &= \frac{1}{\log(4)} \left(\log(10000) - \log(5000)\right) \\ &= 0.5. \end{aligned}$$

4. Suppose the correlation coefficient between the heights of fathers and sons is $\rho = 0.7$. Also suppose the height X of a father has a mean $\mu_X = 174$cm and a standard deviation $\sigma_X = 5$cm, and the height Y of a son has a mean $\mu_Y = 175$cm and a standard deviation $\sigma_Y = 5$cm. Assume that (X, Y) jointly has a bivariate normal distribution.
   (a) Specify the distribution of Y given that $X = 180$.
   (b) Give the values of $E(Y)$ and $E(Y|X = 180)$.
   (c) Find $\mathbb{P}(Y > 180)$ and $\mathbb{P}(Y > 180|X = 180)$.

Solution:

(a) $(Y|X = 180) \stackrel{d}{=} N(175 + 0.7 \times 5 \times \frac{180-174}{5}, 5^2(1 - 0.7^2)) = N(179.2, 12.75)$.

(b) $E(Y) = 175$, $E(Y|X = 180) = 179.2$.

(c)

$$\begin{aligned}\mathbb{P}(Y > 180) &= \mathbb{P}\left(Z > \frac{180 - 175}{5}\right) \\ &= \mathbb{P}(Z > 1) \\ &= 0.1587.\end{aligned}$$

$$\begin{aligned}\mathbb{P}(Y > 180|X = 180) &= \mathbb{P}\left(Z > \frac{180 - 179.2}{\sqrt{12.75}}\right) \\ &= \mathbb{P}(Z > 0.224) \\ &= 0.4114.\end{aligned}$$

5. When a car is stopped by a police patrol, each tyre is checked for wear, and each headlight is checked to see whether it is properly aimed. Let X denote the number of headlights that need adjustment and let Y denote the number of defective tyres.
   (a) If X and Y are independent with $p_X(0) = 0.5$, $p_X(1) = 0.3$, $p_X(2) = 0.2$, and $p_Y(0) = 0.6$, $p_Y(1) = 0.1$, $p_Y(2) = p_Y(3) = 0.05$, $p_Y(4) = 0.2$, display the joint pmf of (X, Y) in a joint probability table.
   (b) Compute $\mathbb{P}(X \leq 1 \text{ and } Y \leq 1)$ from the joint probability table and verify that it equals the product $\mathbb{P}(X \leq 1) \cdot \mathbb{P}(Y \leq 1)$.
   (c) What is $\mathbb{P}(X + Y = 0)$ (the probability of no violations)?
   (d) Compute $\mathbb{P}(X + Y \leq 1)$.

Solution:

(a)

	0	1	2	3	4	$p_X(x)$
0	0.3	0.05	0.025	0.025	0.1	0.5
1	0.18	0.03	0.015	0.015	0.06	0.3
2	0.12	0.02	0.001	0.001	0.04	0.2
$p_Y(y)$	0.6	0.1	0.05	0.05	0.2	

(b) $\mathbb{P}(X \leq 1 \text{ and } Y \leq 1) = 0.56 = \mathbb{P}(X \leq 1) \cdot \mathbb{P}(Y \leq 1)$.

(c) $\mathbb{P}(X + Y = 0) = 0.3$

(d) $\mathbb{P}(X + Y \leq 1) = 0.53$.

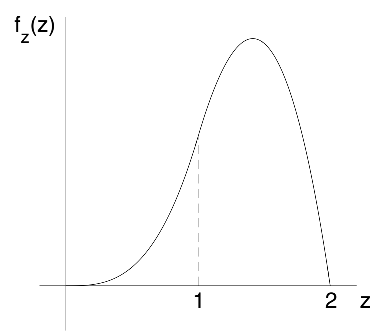
6. If X and Y are independent random variables, each having pdf given by

$$f(x) = 2x, \quad 0 < x < 1,$$

find the pdf of $Z = X + Y$ and sketch its graph.

Solution:

$$\begin{aligned} f_Z(z) &= \int_{S_X} f_X(x) f_Y(z-x)dx \\ &= \int_0^1 4x(z-x)I(0 < z-x < 1)dx \\ &= \begin{cases} \int_0^z 4x(z-x)dx = 2z^3/3, & 0 \le z < 1 \\ \int_{z-1}^1 4x(z-x)dx = -2z^3/3 + 4z - 8/3, & 1 \le z < 2 \\ 0, & \text{elsewhere.} \end{cases} \end{aligned}$$



7. Let X and Y be independent random variables with $\mathbb{E}(X) = \mathbb{E}(Y) = 5$, $V(X) = 1$, and $V(Y) = \sigma^2 > 1$. Let $Z = aX + (1-a)Y$, $0 \le a \le 1$. Find

(a) the value of a that minimizes $V(Z)$ and that minimum value, and

(b) the value of a that maximizes $V(Z)$ and that maximum value.

An investor's interpretation of the problem: Suppose that X and Y are (independent and uncertain!) long-term returns (per \$1) on two different stocks. Then Z is the return on a portfolio where an investor puts $100a$% of her capital in the first stock and the remaining $100(1-a)$% in the second stock. A question of interest is 'Which portfolio should the investor form if she wants to minimise her exposure to risk? [The variance of return is one of the standard measures of risk.] What if she is a risk-seeking person?

Solution:

$$V(Z) = a^2 V(X) + (1-a)^2 V(Y) = a^2 + (1-a)^2\sigma^2.$$

(a) We want to minimise

$$f(a) = (\sigma^2+1)a^2 - 2\sigma^2 a + \sigma^2$$

with respect to $a \in [0, 1]$.

$$f'(a) = 2(\sigma^2+1)a - 2\sigma^2$$

which is equal to zero when

$$a = a^* = \frac{\sigma^2}{\sigma^2+1},$$

which is clearly in $[0, 1]$.

$$f''(a) = 2(\sigma^2+1),$$

which is positive, so this point minimises $f(a)$. The minimum value of the variance is given by

$$f(a^*) = \frac{\sigma^4}{\sigma^2+1} - \frac{2\sigma^4}{\sigma^2+1} + \frac{\sigma^4+\sigma^2}{\sigma^2+1} = \frac{\sigma^2}{\sigma^2+1}.$$

(b) The maximum value of the variance occurs either at $a = 0$ when $V(Z) = \sigma^2$, or $a = 1$ when $V(Z) = 1$. Since $\sigma^2 > 1$, the maximum occurs at $a = 0$ when $V(Z) = \sigma^2$.

MAST20004 Probability

Computer Lab 8

In this lab you

- investigate how to generate observations on a standard bivariate normal distribution with correlation coefficient 0 by using the distributions for the polar coordinates derived in lecture slides 384–386.
- use a similar argument to derive the joint and marginal distributions for the polar coordinates of a uniform random point inside the unit circle and use these results to simulate the distribution.
- simulate the distribution from Tutorial 8, Question 6 and use your program to cross-check your answer for the pdf of Z.

Exercise A - Standard bivariate normal pdf with $\rho = 0$

Start this exercise by reviewing lecture slides 384–386. These slides show that you can simulate a random point with a standard bivariate normal distribution with $\rho = 0$ by generating its polar coordinates with the independent distributions for R and Θ given on slide 386.

The **incomplete** Matlab m-file **Lab8ExA.m** will simulate `npts` observations from this distribution, once you have typed in some additional code that is required. The program plots the observations and also plots empirical marginal pdfs for both X and Y. You can change the value of `npts` in the program itself.

The additional code required generates the required observations on R using the 'inverse transformation method'. This method uses the fact that if $U \stackrel{d}{=} R(0,1)$ then $F_X^{-1}(U)$ has distribution function F_X, as explained in lectures.

1. Find the distribution function F_R for R.
2. Hence find the inverse distribution function F_R^{-1} .
3. Type in the additional code to generate observations on R (only two lines are required!) and study the program to make sure you understand how it works (don't worry about the detail of the plotting commands).
4. Run the modified program and think about how to visually check the outcome. What marginal distributions do you expect for X and Y?
5. As a check on the accuracy of the marginal distributions, uncomment the lines calculating the proportion of observations between -1 and 1 for both X and Y and check the answer against the appropriate probability tables.

Exercise B - Uniformly distributed point inside the unit circle

In this exercise you need to make some minor changes to your Matlab m-file **Lab8ExA.m** so start by saving a copy of it as **Lab8ExB.m**. You want **Lab8ExB.m** to simulate a random point P uniformly distributed inside the unit circle by generating its polar coordinates.

Note: Let (X, Y) be the coordinates of P. Then the joint pdf of (X, Y) is

$$f_{(X,Y)}(x, y) = \begin{cases} 1/\pi & x^2 + y^2 \leq 1 \\ 0 & \text{elsewhere.} \end{cases}$$

1. Using a very similar argument to that used on lecture slide 384–386, derive the joint and marginal densities for the polar coordinates (R, Θ) of P . Check that your marginal densities integrate to 1 over the appropriate range.
2. Derive the distribution function F_R for R.
3. Hence find the inverse distribution function F_R^{-1} .
4. Modify your program to generate observations on R and also to modify the plot title appropriately.
5. Run the modified program and think about how to visually check the outcome. What shape do you expect to see for the marginal distributions X and Y?
6. Extension task: As one check on the accuracy of the marginal distributions modify the lines calculating the proportion of observations in a given range (choose your own) and check the simulated answer.

Exercise C - Tutorial 8, Question 6

Suitably modifying the **incomplete** program **Lab8ExC.m** will generate observations on the bivariate distribution considered in Tutorial 8, Question 6.

You will need to type in the commands to generate the observations on X and Y .

1. Add the required commands and run the program. Visually check the output and the expected marginals.
2. Think of dividing the unit square into four quadrants. Using the graphical output as a guide put the quadrants in increasing order ranked by the probabilities that the random point hits each one of them (no calculations are required).
3. Uncomment the lines which calculate the proportion of times that the sum $Z = X + Y$ exceeds 1 and use the simulation result to check your answer to Tutorial 8, question 6.