Exercise Sheet 3

The exercise sheet is subdivided into three parts:

- Part A contains warm-up questions you should do in your own time.
- You should solve all questions of Part B and put your solutions in the locker in the undergraduate common room in the Department of Statistical Science (Room number 117) assigned to your tutor group by the deadline stated on the exercise sheet. Don't forget to mark your solutions with your name and student number.

  Reminder: If you miss the deadline for reasons outside your control, for example illness or bereavement, you ***must*** submit a claim for extenuating circumstances, normally within a week of the deadline. Your home department will advise you of the appropriate procedures. For Statistical Science students, the relevant information is on the DOSSSH Moodle page.

  Your tutor will then mark one question selected at random after you handed in your solutions (mark counts towards ICA) and will return your marked solutions to you at the next tutorial at which the solutions to the part B questions will be discussed. Marks available for each part question are indicated in square brackets, e.g. [3]. Very succinct solutions to these questions will usually be made available on Moodle at 9:15am on Thursday mornings, so don't hand in late!
- Solutions to Part C questions will be made available in time for exam revision. These questions will not be discussed at the tutorials.

A: Warm-up questions

1. Let X and Y have joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} x+y, & \text{for } 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

   Determine $\mathrm{corr}(X,Y)$.

2. X_1 and X_2 have joint density $f(x_1,x_2) = 4x_1x_2$ for $0 \le x_1 \le 1, 0 \le x_2 \le 1$. Calculate the marginal and conditional densities of X_1 and X_2, their means and variances and their correlation.

B: Answers to hand in by Thursday, 29th October, 9am

1. Let the joint pdf of the random variables X and Y be given by

$$f_{X,Y}(x,y) = \begin{cases} (\alpha+1)x\exp\left(-x(\alpha y+1)\right) & \text{if } x>0\,,\ y\in(0,1) \\ 0 & \text{otherwise} \end{cases},$$

where $\alpha \in [0,\infty)$ is a parameter.

   (a) Show that the marginal pdf for X is given by

$$f_X(x) = \frac{\alpha+1}{\alpha}\left(1-e^{-\alpha x}\right)e^{-x} \text{ on}, x>0$$

   if $\alpha > 0$. [5]

   (b) Compute the conditional pdf of Y given X, i.e. compute $f_{Y|X}(y|x)$. Note that you need to answer this question for all values $\alpha \in [0,\infty)$. [5]

   (c) Compute the conditional expectation of Y given $X = x$, i.e. compute $\mathbb{E}_{Y|X}[Y|X=x]$ for all values of $\alpha \in [0,\infty)$. [5]

2. Let X_i ($i \in \mathbb{N}$) be independent and identically distributed all following a $\mathrm{Bin}(10,p)$ distribution for some value $p \in [0,1]$. Define $Y_n := \sum_{i=1}^n X_i$.

   (a) What is the expected value of X_i? Use this to find a function $f : \mathbb{N} \to \mathbb{R}$ such that $\mathbb{E}[Y_n - f(n)] = 0$ for all $n \in \mathbb{N}$. [4]

   (b) Compute the covariance $\gamma_{n,m} := \mathrm{Cov}(Y_n, Y_m)$ for $n, m \in \mathbb{N}$ where $n \le m$. [6]

   (c) For which value or values $p \in [0,1]$ is $\gamma_{1,2} = 0$? Justify your answer.
   Show that for all values of $p \in [0,1]$ for which $\gamma_{1,2} = 0$ holds, Y_n and Y_m are in fact independent (regardless of the $n, m \in \mathbb{N}$ chosen). [5]

3. A lecturer performs a random experiment in a lecture by letting a student volunteer pick a spaghetto from a pack at random. Assume that the length L of the spaghetto follows the distribution $N(26\text{cm}, 0.09\text{cm}^2)$. Another volunteer then breaks the spaghetto into exactly two parts at a randomly chosen point. The length of one of the parts is given by $L_1 = UL$ where the distribution of U is assumed to be $U \sim \mathrm{Unif}(0,1)$ and U is assumed to be independent of L.

   (a) Compute the expected length of the spaghetto part L_1. `Hint:` Depending on how you carry out the calculation, thinking of the distribution of U as the conditional distribution, i.e. $U|L = l \sim \mathrm{Unif}(0,1)$ (which does not change the result!) may simplify the calculation. [3]

   (b) Compute the variance of the length of the spaghetto part L_1. `Hint:` same hint as above. [7]

   (c) Approximating the distribution of L_1 as a uniform distribution with the mean and variance computed above, compute the approximate probability of each of the events $B_2 = \{L_1 < 10\text{cm}\}$, $B_3 = \{10\text{cm} \le L_1 \le 20\text{cm}\}$ and $B_1 = \{L_1 > 20\text{cm}\}$ to three significant digits. [5]

C: Exam Practice Questions

1. Let the continuous random variable X follow the distribution given by the following density function:

$$f_X(x) = \begin{cases} a\exp(-x) & \text{if } x > 0 \\ b\exp(2x) & \text{if } x \leq 0 \end{cases}$$

where the numbers $a, b \in \mathbb{R}$ are fixed but unknown.

   (a) Find the values of $(a, b) \in \mathbb{R}^2$ for which f_X is a valid pdf. [15]

   (b) Compute the expectation of X. Express your result in terms of a and b. [10]

   (c) Compute $\mathbb{P}(X > 0)$ in terms of a. [10]

   (d) Here and in the sequel, consider the joint probability density function

$$f_{X,Y}(x, y) = \begin{cases} a\exp(-x) & \text{if } x \geq 0,\ y \in [0, 1) \\ b\exp(2x) & \text{if} x < 0,\ y \in (-1, 0) \\ 0 & \text{otherwise} \end{cases}.$$

   You may use without proof that this is a valid pdf for the same values of a and b as computed in (a) and that f_X is the marginal pdf for X obtained from the joint pdf $f_{X,Y}$. Show that the marginal pdf for Y is given by

$$f_Y(y) = \begin{cases} a & \text{if } y \in [0, 1) \\ \frac{b}{2} & \text{if } y \in (-1, 0) \end{cases}$$

   [10]

   (e) Compute $f_{X|Y}(x|y)$ and name the resulting distribution in the case $y = \frac{1}{2}$. [15]

2. Let $X_1, \ldots, X_n$ be independent random variables with common mean μ and variance σ^2. Define

$$Y = \sum_{i=1}^{n} c^i X_i$$

where c is a positive constant, $0 \leq c \leq 1$. Find the expected value and variance of Y and investigate the special case $c = 1$.
Hint: Recall the sum of a geometric progression: $\sum_{i=0}^{k} a^i = (1 - a^{k+1})/(1 - a)$ for $a \neq 1$.

3. Random variables X and Y have joint probability density function

$$f_{X,Y}(x, y) = xy^2 e^{-xy} e^{-y} \qquad \text{for } x \geq 0 \text{ and } y \geq 0.$$

Find the conditional and marginal densities of X and Y, by using the identities

$$f_{X,Y}(x, y) = f_{X|Y}(x|y) f_Y(y) = f_{Y|X}(y|x) f_X(x)$$

together with the fact that densities must integrate to 1. Name those distributions that you recognise. Hint: The solution requires no integration — but you may need to remind yourself of the form of the Gamma density!