Statistical Data Analysis
Problem sheet 2
Due Monday, 24 October 2022

Exercise 1 [10 marks]: Consider the joint pdf for the continuous random variables x and y

$$f(x, y) = 6x(1-x)\,\frac{1}{\xi}e^{-(y-x)/\xi}$$

for $0 \leq x \leq 1$, $x \leq y < \infty$ and positive parameter ξ.

(a) [**1 mark**] Are x and y independent? Justify your answer.

(b) [**5 marks**] Define the new variables $u = y - x$ and $v = x$. Find the joint pdf of u and v. Are u and v independent?

(c) [**4 marks**] Show that the marginal pdfs for u and v are

$$g_u(u) = \frac{1}{\xi}e^{-u/\xi}$$

with $u \geq 0$ and

$$g_v(v) = 6v(1-v)$$

for $0 \leq v \leq 1$.

Exercise 2 [5 marks]: Consider n random variables $\vec{x} = (x_1, \ldots, x_n)$ that follow a joint pdf $f(\vec{x})$ and constants $c_0, c_1, \ldots, c_n$.

(a) [**1 mark**] Starting from the definition of the expectation value for continuous random variables, show that

$$E\left[c_0 + \sum_{i=1}^{n} c_i x_i\right] = c_0 + \sum_{i=1}^{n} c_i E[x_i]\ .$$

(b) [**4 marks**] Using the result from (a), show that the variance is

$$V\left[c_0 + \sum_{i=1}^{n} c_i x_i\right] = \sum_{i,j=1}^{n} c_i c_j \mathrm{cov}[x_i, x_j]\ .$$

For the variance above, find what this reduces to in the case where the variables $x_1, \ldots, x_n$ are uncorrelated.

Exercise 3 [5 marks]: Consider two random variables x and y and a constant α. From the previous exercise we have (no need to rederive)

$$V[\alpha x + y] = \alpha^2 V[x] + V[y] + 2\alpha\mathrm{cov}[x, y] = \alpha^2\sigma_x^2 + \sigma_y^2 + 2\alpha\rho\sigma_x\sigma_y\ ,$$

where $\sigma_x^2 = V[x]$, $\sigma_y^2 = V[y]$, and the correlation coefficient is $\rho = \mathrm{cov}[x, y]/\sigma_x\sigma_y$. Using this result, show that the correlation coefficient always lies in the range $-1 \leq \rho \leq 1$. (Use the fact that the variance $V[\alpha x + y]$ is always greater than or equal to zero and consider the cases $\alpha = \pm\sigma_y/\sigma_x$.)