Statistical Data Analysis
Problem sheet 2
Due Monday, 21 October 2024

Exercise 1 [10 marks]: Consider (as in Problem Sheet 1) the joint pdf for the continuous random variables x and y

$$f(x,y) = \begin{cases} \frac{1}{\pi R^2} & x^2 + y^2 \leq R^2 \, , \\ 0 & \text{otherwise.} \end{cases}$$

Define the new variables

$$\begin{aligned} u &= \sqrt{x^2 + y^2} \, , \\ v &= \tan^{-1}(y/x) \, . \end{aligned}$$

That is, u corresponds to the radius and v to the azimuthal angle in plane polar coordinates, with $u \geq 0$ and $0 \leq v < 2\pi$.

(a) [**5**] Find the joint pdf of u and v. (Use the inverse of the transformation $x = u \cos v$, $y = u \sin v$.) Are u and v independent? Justify your answer.

(b) [**5**] Find the marginal pdfs for u and v.

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

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

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)

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

where $\sigma_x^2 = V[x]$, $\sigma_y^2 = V[y]$, and the correlation coefficient is $\rho = \text{cov}[x,y]/\sigma_x\sigma_y$. 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 $\alpha = \pm\sigma_y/\sigma_x$.)