This homework runs from Thursday 16 November 2023 until 12 noon on Thursday 23 November 2023. Submission is to Gradescope Homework 6.

Question 1

A small commercial passenger aircraft has fuel efficiency that varies under different conditions, such as weather or loading weight. Over time it is observed that in regular use the fuel efficiency X, measured in *miles per UK gallon* (mpg) is distributed with the following PDF and CDF.

$$\begin{aligned} f_X(x) &= 4(x - x^3) \quad 0 \le x \le 1 \qquad & f_X(x) &= 0 \text{ otherwise} \\ F_X(x) &= 2x^2 - x^4 \quad 0 \le x \le 1 \qquad & F_X(x) &= 0 \text{ if } x < 0 \text{ and } 1 \text{ if } x \ge 1 \end{aligned}$$

(a) Calculate the expected value of X.

(b) Calculate the probability that fuel efficiency is between 0.5 and 0.75 miles per gallon.

Random variable Y is an approximate measure of fuel efficiency using the alternate metric unit of *litres of fuel per kilometre.*

$$Y = \frac{2\sqrt{2}}{X}$$

(c) Random variable X always takes values between 0 and 1. What is the range of possible values for Y?

(d) Calculate $P(Y > 5)$.

(e) Calculate the PDF for Y.

Include your working for each part.

[*7 marks*]

Question 2

Two continuous random variables X and Y range between 0 and 1 with the following joint probability distribution.

$$f(x,y) = \begin{cases} \frac{4x}{3} + y^2 & 0 \le x \le 1 \text{ and } 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Calculate the marginal probability distribution functions $f_X(x)$ and $f_Y(y)$.

(b) Calculate $f_X(1)$, $f_Y(1)$, and $f(1,1)$. Use these to show that X and Y are not independent.

(c) Use the following information about X and Y to estimate the expected value and variance of $(X - Y)$ to three decimal places.

$$\mathrm{E}(X) = 0.611 \quad \mathrm{Var}(X) = 0.071 \quad \mathrm{E}(Y) = 0.583 \quad \mathrm{Var}(Y) = 0.082 \quad \mathrm{Cov}(X,Y) = -0.009$$

Include your working for each part.

[*3 marks*]

Solution 1

(a) The expected value is $8/15 \approx 0.533$, calculated using integration over the non-zero range of the PDF.

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx = \int_0^1 4x(x - x^3)\, dx \\ &= 4\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_0^1 = 4\left(\frac{1}{3} - \frac{1}{5}\right) = \frac{8}{15} \approx 0.533 \end{aligned}$$

(b) The probability that X lies in this range can be found through integration of the PDF $f_X(x)$ but is more easily calculated with the CDF provided.

$$\begin{aligned} P(0.5 < X < 0.75) &= F_X(0.75) - F_X(0.5) \\ &= (2 \times 0.75^2 - 0.75^4) - (2 \times 0.5^2 - 0.5^4) = \frac{95}{256} \approx 0.371 \end{aligned}$$

(c) Since $0 \leq X \leq 1$ we can deduce Y may take any value above $2\sqrt{2} \approx 2.83$

(d) Probabilities for Y can be calculated by converting to equivalent statements about X.

$$P(Y > 5) = P(X < 2\sqrt{2}/5) = F_X(2\sqrt{2}/5) = \frac{336}{625} \approx 0.538$$

This uses the fact that the function transforming between X and Y is continuous and monotonically decreasing over the range where it is applied.

(e) The PDF for Y is calculated by working with the function $h(y)$ that is the inverse of the transformation function $g(x)$ which takes X to Y. Function $h(y)$ is both well-defined and differentiable on the relevant interval $Y > 2\sqrt{2}$.

$$\begin{aligned} \text{Transformation function } g(x) &= \frac{2\sqrt{2}}{x} \quad \text{since } Y = \frac{2\sqrt{2}}{X} \\ \text{Inverse function } h(y) &= \frac{2\sqrt{2}}{y} \quad \text{since } X = \frac{2\sqrt{2}}{Y} \\ \text{Derivative } h'(y) &= -\frac{2\sqrt{2}}{y^2} \\ \text{PDF } f_Y(y) &= f_X(h(y)) \cdot |h'(y)| = 4(h(y) - h(y)^3) \cdot |h'(y)| \\ &= 4\left(\left(\frac{2\sqrt{2}}{y}\right) - \left(\frac{2\sqrt{2}}{y}\right)^3\right) \cdot \left|-\frac{2\sqrt{2}}{y^2}\right| = \frac{32}{y^3} - \frac{256}{y^5} \\ &= \frac{32(y^2 - 8)}{y^5} \quad \text{if } y > 2\sqrt{2} \text{ and } f_Y(y) = 0 \text{ otherwise} \end{aligned}$$

There is an alternate method through computing the CDF $F_Y(y)$ in terms of $F_X(x)$ as a generalization of part (d), and then differentiating to get $f_Y(y)$. It is a little longer, though, and requires care to account for the inverse relationship between X and Y.

Solution 2

(a) Marginal probability distributions are calculated by integrating over the full range of the other variable.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{4x}{3} + y^2\,dy = \left[\frac{4xy}{3} + \frac{y^3}{3}\right]_0^1 = \frac{4x+1}{3} \qquad 0 < x < 1$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{4x}{3} + y^2\,dx = \left[\frac{2x^2}{3} + xy^2\right]_0^1 = \frac{2}{3} + y^2 \qquad 0 < y < 1$$

(b) Variables X and Y are independent if their joint probability distribution is equal to the product of their individual marginal distributions: that $f(x,y) = f_X(x)f_Y(y)$ for all x, y. When $x = 1$ and $y = 1$ we have the following.

$$f_X(1) = \frac{4+1}{3} = \frac{5}{3} \qquad f_Y(1) = \frac{2}{3} + 1 = \frac{5}{3} \qquad f(1,1) = \frac{4}{3} + 1 = \frac{7}{3}$$

Thus $f_X(1)f_Y(1) = 25/9 \neq f(1,1)$ and we can conclude that X and Y are dependent.

(c) By linearity of expectations we can calculate the expected value of $(X - Y)$ from the individual expected values of X and of Y. This is valid even though they are not independent. To calculate variance we need to include the covariance.

$$\mathrm{E}(X-Y) = \mathrm{E}(X) - \mathrm{E}(Y) = 0.611 - 0.583 = 0.028$$

$$\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\,\mathrm{Cov}(X,Y) = 0.071 + 0.082 + 2 \times 0.009 = 0.171$$

This formula for variance appeared in the lecture slides and the textbook. Another approach is to calculate $\mathrm{E}(X^2)$, $\mathrm{E}(Y^2)$, and $\mathrm{E}(XY)$ from the values of $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$ and $\mathrm{Cov}(X,Y)$ provided and then recombine these to compute $\mathrm{E}((X-Y)^2)$ and hence $\mathrm{Var}(X-Y)$. This is more long-winded, but reaches the same correct answer.