STEP 3 – Combinations and Functions of Random Variables (Sheet 1)

Independent random variables

Understand and use the idea of independent random variables.

Algebra of expectation

Know, understand and use the algebra of expectation:

$\mathrm{E}(aX + bY + c) = a\mathrm{E}(X) + b\mathrm{E}(Y) + c$

$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$

$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$

and for independent random variables:

$\mathrm{Var}(aX \pm bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$

Use cumulative distribution functions to calculate the probability density function of a related random variable; for example, X^2 from X.

Knowledge of generating functions will not be required.

Q1, (STEP III, 2006, Q13)

Two points are chosen independently at random on the perimeter (including the diameter) of a semicircle of unit radius. The area of the triangle whose vertices are these two points and the midpoint of the diameter is denoted by the random variable A. Show that the expected value of A is $(2+\pi)^{-1}$.

Q2, (STEP III, 2102, Q12)

(i) A point P lies in an equilateral triangle ABC of height 1. The perpendicular distances from P to the sides AB, BC and CA are x_1, x_2 and x_3, respectively. By considering the areas of triangles with one vertex at P, show that $x_1 + x_2 + x_3 = 1$.

Suppose now that P is placed at random in the equilateral triangle (so that the probability of it lying in any given region of the triangle is proportional to the area of that region). The perpendicular distances from P to the sides AB, BC and CA are random variables X_1, X_2 and X_3, respectively. In the case $X_1 = \min(X_1, X_2, X_3)$, give a sketch showing the region of the triangle in which P lies.

Let $X = \min(X_1, X_2, X_3)$. Show that the probability density function for X is given by

$$\mathrm{f}(x) = \begin{cases} 6(1-3x) & 0 \leqslant x \leqslant \frac{1}{3}, \\ 0 & \text{otherwise}. \end{cases}$$

Find the expected value of X.

(ii) A point is chosen at random in a regular tetrahedron of height 1. Find the expected value of the distance from the point to the closest face.
[The volume of a tetrahedron is $\frac{1}{3} \times$ area of base $\times$ height and its centroid is a distance $\frac{1}{4} \times$ height from the base.]

Q3, (STEP III, 2013, Q13)

(a) The continuous random variable X satisfies $0 \leqslant X \leqslant 1$, and has probability density function $\mathrm{f}(x)$ and cumulative distribution function $\mathrm{F}(x)$. The greatest value of $\mathrm{f}(x)$ is M, so that $0 \leqslant \mathrm{f}(x) \leqslant M$.

(i) Show that $0 \leqslant \mathrm{F}(x) \leqslant Mx$ for $0 \leqslant x \leqslant 1$.

(ii) For any function $\mathrm{g}(x)$, show that

$$\int_0^1 2\mathrm{g}(x)\mathrm{F}(x)\mathrm{f}(x)\mathrm{d}x = \mathrm{g}(1) - \int_0^1 \mathrm{g}'(x)\big(\mathrm{F}(x)\big)^2\mathrm{d}x\,.$$

(b) The continuous random variable Y satisfies $0 \leqslant Y \leqslant 1$, and has probability density function $k\mathrm{F}(y)\mathrm{f}(y)$, where f and F are as above.

(i) Determine the value of the constant k.

(ii) Show that

$$1 + \frac{nM}{n+1}\mu_{n+1} - \frac{nM}{n+1} \leqslant \mathrm{E}(Y^n) \leqslant 2M\mu_{n+1}\,,$$

where $\mu_{n+1} = \mathrm{E}(X^{n+1})$ and $n \geqslant 0$.

(iii) Hence show that, for $n \geqslant 1$,

$$\mu_n \geqslant \frac{n}{(n+1)M} - \frac{n-1}{n+1}\,.$$

Q4, (STEP III, 2015, Q13)

Each of the two independent random variables X and Y is uniformly distributed on the interval $[0, 1]$.

(i) By considering the lines $x + y =$ constant in the x-y plane, find the cumulative distribution function of $X + Y$.

Hence show that the probability density function f of $(X+Y)^{-1}$ is given by

$$\mathrm{f}(t) = \begin{cases} 2t^{-2} - t^{-3} & \text{for } \tfrac{1}{2} \leqslant t \leqslant 1 \\ t^{-3} & \text{for } 1 \leqslant t < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Evaluate $\mathrm{E}\Big(\dfrac{1}{X+Y}\Big)$.

(ii) Find the cumulative distribution function of Y/X and use this result to find the probability density function of $\dfrac{X}{X+Y}$.

Write down $\mathrm{E}\Big(\dfrac{X}{X+Y}\Big)$ and verify your result by integration.

Q5, (SEP III, 2014, Q12)

The random variable X has probability density function $f(x)$ (which you may assume is differentiable) and cumulative distribution function $F(x)$ where $-\infty < x < \infty$. The random variable Y is defined by $Y = e^X$. You may assume throughout this question that X and Y have unique modes.

(i) Find the median value y_m of Y in terms of the median value x_m of X.

(ii) Show that the probability density function of Y is $f(\ln y)/y$, and deduce that the mode λ of Y satisfies $f'(\ln \lambda) = f(\ln \lambda)$.

(iii) Suppose now that $X \sim N(\mu, \sigma^2)$, so that

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}.$$

Explain why

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-\mu-\sigma^2)^2/(2\sigma^2)}\, dx = 1$$

and hence show that $E(Y) = e^{\mu+\frac{1}{2}\sigma^2}$.

(iv) Show that, when $X \sim N(\mu, \sigma^2)$,

$$\lambda < y_m < E(Y).$$

Q6, (STEP III, 2016, Q13)

Given a random variable X with mean μ and standard deviation σ, we define the *kurtosis*, κ, of X by

$$\kappa = \frac{E\big((X-\mu)^4\big)}{\sigma^4} - 3.$$

Show that the random variable $X - a$, where a is a constant, has the same kurtosis as X.

(i) Show by integration that a random variable which is Normally distributed with mean 0 has kurtosis 0.

(ii) Let $Y_1, Y_2, \ldots, Y_n$ be n independent, identically distributed, random variables with mean 0, and let $T = \sum_{r=1}^{n} Y_r$. Show that

$$E(T^4) = \sum_{r=1}^{n} E(Y_r^4) + 6\sum_{r=1}^{n-1}\sum_{s=r+1}^{n} E(Y_s^2)E(Y_r^2).$$

(iii) Let $X_1, X_2, \ldots, X_n$ be n independent, identically distributed, random variables each with kurtosis κ. Show that the kurtosis of their sum is $\dfrac{\kappa}{n}$.

Q7, (STEP III, 2017, Q13)

The random variable X has mean μ and variance σ^2, and the function V is defined, for $-\infty < x < \infty$, by

$$\mathrm{V}(x) = \mathrm{E}\big((X - x)^2\big).$$

Express $\mathrm{V}(x)$ in terms of x, μ and σ.

The random variable Y is defined by $Y = \mathrm{V}(X)$. Show that

$$\mathrm{E}(Y) = 2\sigma^2. \qquad (*)$$

Now suppose that X is uniformly distributed on the interval $0 \leqslant x \leqslant 1$. Find $\mathrm{V}(x)$. Find also the probability density function of Y and use it to verify that $(*)$ holds in this case.