Chapter 4

Continuous random variables

Informally, a continuous random variables describe outcomes in probabilistic situations where the possible values some quantity can take form a continuum. A continuous random variable X is characterized by a sample space $S \subseteq \mathcal{R}$, which denotes the range of possible outcomes, and by a probability density function (pdf) $f(x)$, which gives the relative likelihood of any outcome in a continuum occurring. Similarly, any collection of continuous random variables, $X_1, ...X_k$ is characterized by a collection of sample spaces, $S_1, ..., S_k$ and by a joint probability density function, $f(x_1, ...x_k)$, which gives the relative likelihood of any collection of outcome in a continuum occurring jointly.

4.1. Distribution Functions

Cumulative Distribution Function Let X be a continuous random variable with pdf $f(x)$ defined over the real space $\mathcal{R}$. Then its cumulative distribution function (cdf), $F(x)$ is defined as follows:

$$F(x) = \int_{-\infty}^{x} f(x)dx \quad -\infty \leq x \leq \infty$$

Proposition Let X be a continuous random variable with pdf $f(x)$ defined over the real space $\mathcal{R}$. The probability that X takes value in the interval $[a, b] \subset \mathcal{R}$ is defines as follows:

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

Proof. Notice that:

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \int_{a}^{b} f(x)dx$$

Proposition Let X be a continuous random variable with pdf $f(x)$ defined over the real space $\mathcal{R}$. The probability that X takes value $a \in \mathcal{R}$ is equal to 0.

Proof. Notice that:

$$P(X = a) = \int_a^a f(x)dx = 0$$

Proposition Any density $f(x)$ defined over a sample space S satisfies the completeness axiom, i.e.

$$\int_{x \in S} f(x)dx = 1$$

Joint Cumulative Distribution Function Let X and Y be two continuous random variables, each defined on the real space. Let their joint probability function be $f(x, y)$. Then their joint cumulative distribution function (cdf), $F(x, y)$ is defined as follows:

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x, y)dxdy \quad -\infty \leq x \leq \infty \quad -\infty \leq y \leq \infty$$

Proposition Let X and Y be two continuous random variables, each defined on the real space. Let their joint probability function be $f(x, y)$. The probability that X takes value in the interval $[a, b] \subset \mathcal{R}$ and that Y takes a value in the interval $[c, d]$ is defines as follows:

$$P(a \leq X \leq b, c \leq X \leq d) = \int_c^d \int_a^b f(x, y)dxdy$$

Marginal Density Function Let X and Y be two continuous random variables, each defined on the real space. Let their joint probability function be $f(x, y)$. The marginal density function of X, $f(x)$ assigns probabilities to a range of values of X irrespective of the values of Y can take, and it is defined as follows

$$f(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

Similarly, the marginal density function of Y is equal to

$$f(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

4.2. Moments

Let X be a continuous random variable with pdf $f(x)$ defined over the real space $\mathcal{R}$.

Expected value The expected value of X is defined by

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Variance The variance of X is defined by

$$VAR(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)dx$$

Notice that

$$\begin{aligned} VAR(X) = \int_{-\infty}^{\infty} (x - E[X])^2 f(x)dx = \\ \int_{-\infty}^{\infty} (x^2 - 2xE[X] + E[X]^2) f(x)dx = \\ \int_{-\infty}^{\infty} x^2 f(x)dx - \int_{-\infty}^{\infty} 2xE[X]f(x)dx + \int_{-\infty}^{\infty} E[X]^2 f(x)dx = \\ \int_{-\infty}^{\infty} x^2 f(x)dx - 2E[X]\int_{-\infty}^{\infty} xf(x)dx + E[X]^2 \int_{-\infty}^{\infty} f(x)dx = \\ \int_{-\infty}^{\infty} x^2 f(x)dx - E[X]^2 = E[X^2] - E[X]^2 \end{aligned}$$

where $E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx$.

Let X and Y be two continuous random variables, each defined on the real space. Let their joint probability function be $f(x, y)$.

Covariance The covariance of two continuous random variables X and Y is given by:

$$COV(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - E[X])(y - E[Y])f(x, y)dxdy$$

Notice that

$$\begin{aligned} COV(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - E[X])(y - E[Y])f(x, y)dxdy = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy - xE[Y] - yE[X] + E[X]E[Y])f(x, y)dxdy = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xE[Y]f(x, y)dxdy \\ - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yE[X]f(x, y)dxdy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[X]E[Y]f(x, y)dxdy = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy - \int_{-\infty}^{\infty} xE[Y]\left[\int_{-\infty}^{\infty} f(x, y)dy\right]dx - \int_{-\infty}^{\infty} yE[X]\left[\int_{-\infty}^{\infty} f(x, y)dx\right]dy + E[X]E[Y] = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy - E[Y]\int_{-\infty}^{\infty} xf(x)dx - E[X]\int_{-\infty}^{\infty} yf(y)dy + E[X]E[Y] = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy - 2E[X]E[Y] + E[X]E[Y] = \\ E[XY] - E[X]E[Y] \end{aligned}$$

where $E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy$.

Independence The random variables X and Y are statistically independent if the joint probability density function can be written as the product of the marginal density functions:

$$f(x,y) = f(x)f(y)$$

Proposition If X and Y are statistically independent, then $COV(X,Y) = 0$.

Proof. To see this, notice that

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x)f(y)dxdyz =$$
$$\int_{-\infty}^{\infty} xf(x)\left[\int_{-\infty}^{\infty} yf(y)dy\right]dx = \int_{-\infty}^{\infty} xf(x)E[Y]dx = E[Y]\int_{-\infty}^{\infty} xf(x)dx = E[X]E[Y]$$

which implies that

$$COV(X,Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$$

4.3. Examples of Continuous Random Variables

4.3.1 Uniform distribution

The uniform distribution is a continuous probability distribution that has equal probabilities for all possible outcomes of the random variable. Let X be a uniform random variable distributed over the interval $[a,b] \subset \mathcal{R}$. We write $X \sim U[a,b]$.

PDF The probability density function of X is equal to

$$f(x) = \frac{1}{b-a} \quad a \leq X \leq b$$

CDF The cumulative distribution function of X is equal to

$$F(z) = P(X \leq z) = \int_{a}^{z} \frac{1}{b-a}dx = \frac{1}{b-a}\int_{a}^{z} dx = \frac{1}{b-a}[x]_a^z = \frac{1}{b-a}(z-a) = \frac{z-a}{b-a}$$

Expected value The expected value of X is equal to

$$E(z) = \int_{a}^{b} xf(x)dx = \int_{a}^{b} x\frac{1}{b-a}dx = \frac{1}{b-a}\int_{a}^{b} xdx =$$
$$\frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{b-a}(\frac{b^2}{2} - \frac{a^2}{2}) = \frac{(b^2-a^2)}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$$

Variance The expected value of X is equal to

$$VAR(z) = E[X^2] - E[X]^2 = \int_a^b x^2 \frac{1}{b-a} dx - \left(\int_a^b x f(x) dx \right)^2 =$$

$$\frac{1}{b-a} \int_a^b x^2 dx - \frac{b+a}{2} = \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b - \frac{b+a}{2} =$$

$$\frac{1}{3(b-a)} \left[b^3 - a^3 \right] - \frac{b+a}{2} = \frac{(b-a)^2}{12}$$

4.3.2 Standard Normal distribution

A standard normal random variable, denoted Z, has a probability density function defined as follows

$$f(z) = \frac{1}{\sqrt[2]{2\pi}} \exp^{-\frac{z^2}{2}}$$

Proposition The density of a standard normal random variable is symmetric about zero, i.e. $f(z) = f(-z)$

Proof. To see this, notice that

$$f(-z) = \frac{1}{\sqrt[2]{2\pi}} \exp^{-\frac{(-z)^2}{2}} = \frac{1}{\sqrt[2]{2\pi}} \exp^{-\frac{z^2}{2}} = f(z)$$

Expected value The expected value of standard normal random variable is zero.

Proof.

$$E[Z] = \int_{-\infty}^{\infty} z f(z) dz = \int_{-\infty}^{0} z f(z) dz + \int_{0}^{\infty} z f(z) dz$$

By symmetry, $\int_{-\infty}^{0} z f(z) dz = - \int_{0}^{\infty} z f(z) dz$, which implies that $E[Z] = 0$.

Variance The variance of standard normal random variable is one.

Proof. $VAR[Z] = E[Z^2] - E[Z]^2 = E[Z^2]$. Since $E[Z^2] = 1$, then $VAR[Z] = 1$.

4.3.3 Normal distribution

A normal variable X is defined as a linear transformation of the standard normal:

$$X = \mu + \sigma Z$$

where $\sigma > 0$. We write $X \sim \mathcal{N}(\mu, \sigma^2)$.

PDF The probability density function of normal random variable is equal to

$$f(x) = \frac{1}{\sigma \sqrt[2]{2\pi}} \exp^{-\frac{\left(\frac{x-\mu}{\sigma}\right)^2}{2}}$$

Expected value The expected value of normal random variable is μ.

Proof. $E[X] = E[\mu + \sigma Z] = E[\mu] + E[\sigma Z] = \mu + \sigma E[Z] = \mu$

Variance The variance of normal random variable is σ^2.

Proof. $VAR[X] = VAR[\mu + \sigma Z] = VAR[\mu] + VAR[\sigma Z] = 0 + \sigma^2 VAR[Z] = \sigma^2$

Proposition The probability that a normal r.v. X falls into the interval $[a, b]$ is equal to

$$P(a \leq X \leq b) = P(a \leq \mu + \sigma Z \leq b) =$$
$$P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

4.4. Exercises

Exercise 1 X is a continuous random variable with probability density function (PDF)

$$f(x) = \frac{1}{9}x^2, \quad 0 \leq x \leq 3$$

- Find the following probabilities:
  - $P(0 \leq X \leq 1)$
  - $P(0 \leq X \leq 2)$
  - $P(1 \leq X \leq 2)$
- Find the expected value and variance of X.

Exercise 2 X is a continuous random variable with the following PDF

$$f(x) = 3x^2, \quad 0 \leq x \leq 1$$

Compute E$[X]$ and VAR$[X]$.

Exercise 3 The random variable Z has a standard normal distribution.

- Find the following probabilities:
  - $P(0 < Z < 1.20)$

- $P(-1.33 < Z < 0)$
- $P(Z > 1.33)$
- $P(-0.77 < Z < 1.68)$

- Find x given that $P(x < Z < 1.68) = 0.2$

Exercise 4 The tread life of a particular brand of tyre has a normal distribution with mean 35000 miles and standard deviation 4000 miles.

- What is the probability that a tyre of this brand will have a tread life between 35000 and 38000 miles?
- What is the probability that a tyre of this brand will have a tread life of less than 32000 miles?

Exercise 5 A repair team is responsible for a stretch of oil pipe 2 miles long. The distance at which any fracture occurs can be represented by a uniformly distributed random variable, with the PDF, $f(x) = 0.5$.

- Find the CDF of X
- Find the probability that any given fracture occurs between 0.5 mile and 1.5 miles along the stretch pipeline.

Exercise 6 A client has an investment portfolio whose mean value is equal to 1000000GBP, with a standard deviation of 30000GBP. Assume that the value of the portfolio follows a Normal distribution. Determine the probability that the portfolio is between 970000GBP and 1060000GBP.