Random Variables

Definition 2.1

A random variable is defined as a function $X : S \to \mathcal{X} \subseteq \mathbb{R}$ *mapping the sample space* S *to a subset of the real line.*

The set $\mathcal{X}$ *of possible values of the random variable is called the support of the random variable.*

* We use the notation $(X \in A)$ to denote the event (in S) which is mapped to the set A by the function X.

$$(X \in A) \equiv \{s \in S : X(s) \in A\}$$

* Hence we can define a probability function P_X on $\mathcal{X}$ as

$$P_X(A) = P(X \in A) = P\Big(\{s \in S : X(s) \in A\}\Big)$$

Cumulative Distribution Function

* We talk about the distribution of a random variable to describe the probability that it falls in certain subsets of the real line.

* Of particular interest are right-closed intervals which define the cumulative distribution function.

Definition 2.2

Suppose that X is a random variable defined on a sample space S, then the cumulative distribution function (CDF) of X is defined as

$$F_X(x) = P(X \leqslant x) = P\Big(\{s \in S : X(s) \leqslant x\}\Big)$$

* Two random variables are said to be identically distributed if, and only if, they have the same CDF.

Properties of the CDF

Theorem 2.1

If a function F is a cumulative distribution function then it satisfies the properties

(i) $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to \infty} F(x) = 1$,

(ii) F *is a nondecreasing (monotone increasing) function*

$$x_1 < x_2 \;\Rightarrow\; F(x_1) \leqslant F(x_2),$$

(iii) F *is a right-continuous function*

$$\lim_{x \downarrow x_0} F(x) = F(x_0) \quad \text{for every } x_0 \in \mathbb{R}.$$

∗ It can also be shown that any function F satisfying these conditions is a cumulative distribution function for some random variable.

Continuous and Discrete Random Variables

Definition 2.3

A random variable X is called a continuous random variable if the cumulative distribution function F_X is a continuous function. A continuous random variable has an uncountable support $\mathcal{X}$.

X is said to be a discrete random variable if F_X is a step function. A discrete random variable has finite or countable support $\mathcal{X}$.

Probability Mass and Density Functions

Definition 2.4

If X is a discrete random variable then the probability mass function of X is given by

$$f_X(x) = P(X = x)$$

Definition 2.5

For a continuous random variable X with cumulative distribution function $F_X(x)$, the probability density function of X is the non-negative function f which satisfies

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

Properties of Mass and Density Functions

Theorem 2.2

A function f is a probability density (or mass) function of a random variable if, and only if, it satisfies

(i) $f_X(x) \geqslant 0$ *for every* $x \in \mathbb{R}$.

(ii) $\displaystyle\sum_{x \in \mathcal{X}} f_X(x) = 1$ *(mass function)*

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1 \qquad \text{(density function)}$$

Transformations

* Recall that a random variable X is a function mapping a sample space S to $\mathcal{X} \subseteq \mathbb{R}$.

* Consider a real-valued function g defined on $\mathbb{R}$. Then $Y = g(X)$ is a composition of functions mapping S to $\mathcal{Y} \subseteq \mathbb{R}$ and so is also a random variable.

* For a given random variable X and its associated distribution, we wish to find the distribution of the random variable $Y = g(X)$ for some transformation g.

Defining Probabilities for $Y = g(X)$

* For any set $A \subseteq \mathcal{Y}$ we can define an inverse mapping

$$g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\}$$

* Then we define the event

$$\big(Y \in A\big) = \big(g(X) \in A\big) = \big(X \in g^{-1}(A)\big).$$

* Thus we can define a probability measure

$$P\big(Y \in A\big) = P\big(X \in g^{-1}(A)\big) = P\big(\{s \in S : X(s) \in g^{-1}(A)\}\big)$$

* This satisfies the Axioms of Probability and so is a valid probability measure.

* The support $\mathcal{Y}$ of Y is given by

$$\mathcal{Y} = \{y : y = g(x) \text{ for some } x \in \mathcal{X}\}$$

Transformations of Discrete Random Variables

* For a discrete random variable we can find the probability mass function of Y from that for X.

$$f_Y(y) = \sum_{\{x \in \mathcal{X} : g(x) = y\}} f_X(x) = \sum_{x \in g^{-1}(y)} f_X(x) \quad \text{for } y \in \mathcal{Y}$$

* The cumulative distribution function for Y is found by summing its probability mass function

$$F_Y(y) = \sum_{t \leqslant y} f_Y(t) = \sum_{\{x \in \mathcal{X} : g(x) \leqslant y\}} f_X(x)$$

Transformations of Continuous Random Variables

* For a continuous random variable, it is generally easiest to get the cdf first.

$$F_Y(y) = \int_{\{x \in \mathcal{X} : g(x) \leqslant y\}} f_X(x)\, dx$$

* We can then find the probability density function using the relation

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

Monotone Transformations

* g is said to be monotone if $u > v \;\Rightarrow\; g(u) > g(v)$(increasing) or $u > v \;\Rightarrow\; g(u) < g(v)$ (decreasing).

* A monotone g is one-to-one and so g^{-1} is also single-valued and monotone.

Theorem 2.3

Suppose that X has cdf F_X on support $\mathcal{X}$ and let $Y = g(X)$ be defined on $\mathcal{Y} = g(\mathcal{X})$.

(i) *If g is an increasing function then $F_Y(y) = F_X\left(g^{-1}(y)\right)$ for any $y \in \mathcal{Y}$.*

(ii) *If g is a decreasing function and X is a continuous random variable then $F_Y(y) = 1 - F_X\left(g^{-1}(y)\right)$ for any $y \in \mathcal{Y}$.*

Monotone Transformations of Continuous Random Variables

Theorem 2.4

Let X be a continuous random variable with continuous pdf f_X on a support $\mathcal{X}$ and let $Y = g(X)$ where g is a monotone function on $\mathcal{X}$. Let $\mathcal{Y} = g(\mathcal{X})$ and suppose that g^{-1} has a continuous derivative on $\mathcal{Y}$. Then the pdf of Y is

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right)\left|\dfrac{dg^{-1}(y)}{dy}\right| & \text{for } y \in \mathcal{Y} \\ 0 & \text{otherwise.} \end{cases}$$

Extension for Piecewise Monotone Transformations

Theorem 2.5

Let X be a continuous random variable with pdf f_X on the support $\mathcal{X}$ and let $Y = g(X)$. Let $A_0, A_1, \ldots, A_k$ be a partition of $\mathcal{X}$ such that $P(X \in A_0) = 0$ and f_X is continuous on each A_i. If there exist functions $g_1, \ldots, g_k$ defined on $A_1, \ldots, A_k$ such that

(i) $g(x) = g_i(x)$ *for every* $x \in A_i$*;*

(ii) g_i *is monotone on* A_i *for each* $i = 1, \ldots, k$*,*

(iii) *the set* $\mathcal{Y} = \{y : y = g_i(x)$ *for some* $x \in A_i\}$ *is the same for each* $i = 1, \ldots, k$*,*

(iv) g_i^{-1} *has continuous derivative on* $\mathcal{Y}$ *for each* $i = 1, \ldots, k$*,*

then the pdf of Y is

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right) \left| \frac{dg_i^{-1}(y)}{dy} \right| & \text{for } y \in \mathcal{Y} \\ 0 & \text{otherwise.} \end{cases}$$

Probability Integral Transform

Theorem 2.6

Let X have continuous cdf F_X and define the random variable $Y = F_X(X)$. Then Y is distributed as a uniform random variable on the interval $(0, 1)$.

That is the pdf of Y is

$$f_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Expectations

Definition 2.6

If X is a discrete random variable with probability mass function f_{sssX} on support $\mathcal{X}$ then the expected value or mean of $g(X)$ for any real-valued function g is

$$\mathsf{E}\left(g(X)\right) = \sum_{x \in \mathcal{X}} g(x) f_X(x)$$

provided that $\sum |g(x)| f_X(x) < \infty$, otherwise we say that the mean does not exist.

If X is a continuous random variable with probability density function $f_X(x)$ the expected value of $g(X)$ is

$$\mathsf{E}\left(g(X)\right) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

provided that $\int |g(x)| f_X(x)\, dx < \infty$, otherwise we say that the mean does not exist.

Some Properties of Expectations

Theorem 2.7

Let X be a random variable with support $\mathcal{X}$. Then for any real-valued functions g_1 and g_2 whose expectations exist and any real constants a, b and c

(i) $\mathsf{E}\left(ag_1(X) + bg_2(X) + c\right) \;=\; a\,\mathsf{E}\left(g_1(X)\right) + b\,\mathsf{E}\left(g_2(X)\right) + c.$

(ii) *If* $g_1(x) \geqslant 0$ *for all* $x \in \mathcal{X}$ *then* $\mathsf{E}\left(g_1(X)\right) \geqslant 0$

(iii) *If* $g_1(x) \geqslant g_2(x)$ *for all* $x \in \mathcal{X}$ *then* $\mathsf{E}\left(g_1(X)\right) \geqslant \mathsf{E}\left(g_2(X)\right)$

(iv) *If* $\leqslant g_1(x) \leqslant b$ *for all* $x \in \mathcal{X}$ *then* $a \leqslant \mathsf{E}\left(g_1(X)\right) \leqslant b$

Moments

Definition 2.7

Suppose that X is a random variable with cdf F_X. For any positive integer r, the r^{th} moment of X (more accurately of F_X) is $\mu_r' = \mathsf{E}\left(X^r\right)$

The r^{th} central moment of X is $\mu_r = \mathsf{E}\left((X-\mu)^r\right)$ where $\mu = \mu_1' = \mathsf{E}(X)$.

Definition 2.8

The second central moment is called the variance of X

$$\mathrm{Var}(X) = \mathsf{E}\left((X-\mu)^2\right) = \mathsf{E}\left(X^2\right) - \left(\mathsf{E}(X)\right)^2$$

The positive square root of the variance is called the standard deviation of X.

Moment Generating Functions

Definition 2.9

Let X be a random variable with cdf F_X. The moment generating function of X (or of F_X) is

$$M_X(t) \;=\; \mathsf{E}\left(\mathsf{e}^{tX}\right)$$

provided this expectation exists for t in some neighbourhood of 0.

Theorem 2.8

If X has moment generating function $M_X(t)$ then for any integer r

$$\mu'_r = \mathsf{E}\left(X^r\right) = \frac{d^r}{dt^r} M_X(t)\bigg|_{t=0}.$$

Properties of the Moment Generating Function

Theorem 2.9

Let X be a random variable with moment generating function $M_X(t)$ which exists in a neighbourhood of 0 and let a and b be two real constants. The the moment generating function of the random variable $aX + b$ is

$$M_{aX+b}(t) = \mathrm{e}^{bt} M_X(at).$$

* If the moment generating function of a random variable exists then all moments of that random variable exist.
* The reverse is not true (e.g. the log-normal distribution).

Equality of Distributions

Theorem 2.10
Let X and Y be two random variables with cdfs F_X and F_Y respectively, all of whose moments exist.

(i) *If X and Y have bounded support then $F_X(u) = F_Y(u)$ for all u if, and only if $\mathsf{E}(X^r) = \mathsf{E}(Y^r)$ for all positive integers r.*

(ii) *If the moment generating functions $M_X(t)$ and $M_Y(t)$ exist in a neighbourhood of 0 and $M_X(t) = M_Y(t)$ for all t in that neighbourhood then $F_X(u) = F_Y(u)$ for all u.*

Convergence of Distribution and Moment Generating Functions

Theorem 2.11

Let $X_1, X_2, \ldots$ be a sequence of random variables each with moment generating function $M_{X_n}(t)$, and cumulative distribution function F_{X_n} for $i = 1, 2, \ldots$. Suppose that

$$\lim_{n\to\infty} M_{X_n}(t) \;\to\; M_X(t) \qquad \text{for all } t \text{ in a neighbourhood of 0}$$

where $M_X(t)$ is a moment generating function for some random variable X. Then there exists a unique cumulative distribution function F_X whose moments are determined by $M_X(t)$ and

$$\lim_{n\to\infty} F_{X_n}(x) \;\to\; F_X(x)$$

at every x where F_X is continuous.