2.7 Sums of Random Variables: The Convolution Theorem

It is often necessary to calculate the probability distribution of sums of random variables. Specifically, if X_1 and X_2 are random variables that have some joint probability distribution, then we may want to compute the probability distribution of a new random variable Y, where

$$Y = X_1 + X_2$$

EXAMPLE Industrial Accidents

Suppose that the times between successive industrial accidents are continuous random variables X_1, X_2 then the time (from time zero) at which the second accident occurs is given by random variable $Y = X_1 + X_2$.

The following theorem provides a way of calculating the distribution of Y.

THEOREM

Suppose that X_1 and X_2 are discrete independent random variables with probability mass functions f_{X_1} and f_{X_2} respectively. If the random variable Y is defined by $Y = X_1 + X_2$, then the probability mass function of Y is given by

$$f_Y(y) = \sum_{x_1=-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1)$$

Proof

$$\begin{aligned}
f_Y(y) = \mathrm{P}[Y = y] &= \sum_{x_1=-\infty}^{\infty} \mathrm{P}[Y = y | X_1 = x_1]\mathrm{P}[X_1 = x_1] \quad \text{by the Thm. of Total Prob.} \\
&= \sum_{x_1=-\infty}^{\infty} \mathrm{P}[X_2 = y - x_1]\mathrm{P}[X_1 = x_1] \qquad X_1, X_2 \text{ independent} \\
&= \sum_{x_1=-\infty}^{\infty} f_{X_2}(y - x_1) f_{X_1}(x_1)
\end{aligned}$$

Note : Some terms in the summation may be zero.

EXAMPLE Suppose that X_1 and X_2 are independent and identically distributed random variables, each taking values on the set $\mathbb{X} = \{0, 1\}$, where the mass function of X_1 and X_2 is given by

$$f_{X_1}(x) = \begin{cases} 1 - \theta & x = 0 \\ \\ \theta & x = 1 \end{cases}$$

Let $Y = X_1 + X_2$. Then Y takes values on range $\mathbb{Y} = \{0, 1, 2\}$, and, for $y \in \mathbb{Y}$ we have, by the Theorem of Total Probability,

$$\begin{aligned}
\mathrm{P}[\ Y = y\ ] &= \mathrm{P}[\ Y = y \mid X_1 = 0\ ]\mathrm{P}[\ X_1 = 0\ ] + \mathrm{P}[\ Y = y \mid X_1 = 1\ ]\mathrm{P}[\ X_1 = 1\ ] \\
&= \mathrm{P}[\ X_2 = y\ ]\mathrm{P}[\ X_1 = 0\ ] + \mathrm{P}[\ Y = y - 1\ ]\mathrm{P}[\ X_1 = 1\ ] \\
&= \sum_{x_1=0}^{1} \mathrm{P}[\ X_1 = x_1\ ]\mathrm{P}[\ X_2 = y - x_1\ ]
\end{aligned}$$

$$\Longrightarrow f_Y(y) = \sum_{x_1=0}^{1} f_{X_1}(x_1) f_{X_2}(y - x_1)$$

Therefore

$$
\begin{aligned}
y=0: &\quad f_Y(0) = f_{X_1}(0)f_{X_2}(0) && = (1-\theta)^2 \\
y=1: &\quad f_Y(1) = f_{X_1}(0)f_{X_2}(1) + f_{X_1}(1)f_{X_2}(0) && = 2\theta(1-\theta) \\
y=2: &\quad f_Y(2) = f_{X_1}(1)f_{X_2}(1) && = \theta^2
\end{aligned}
$$

and $f_Y(y) = 0$ for all other values of y.

Continuous version

Suppose that X_1 and X_2 are continuous independent random variables with probability density functions f_{X_1} and f_{X_2} respectively. If the random variable Y is defined by $Y = X_1 + X_2$, then the probability density function of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1)\ dx_1$$

Note : The integrand may be zero on intervals of $\mathbb{R}$.

EXAMPLE Suppose X_1 and X_2 are independent and **identically** distributed random variables with range $(0,1)$, and p.d.f. that is constant, $f_{X_1}(x) = f_{X_1}(x) = 1$, on that range and zero otherwise. Then, if $Y = X_1 + X_2$, Y takes values on the range $(0,2)$, and

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1)\ dx_1 = \int_0^1 f_{X_2}(y - x_1)\ dx_1$$

The integral is to be evaluated for a **fixed** y value, and the integrand is only non-zero when $0 < y - x_1 < 1$. Hence

$$f_Y(y) = \begin{cases} \displaystyle\int_0^y dx_1 = y & 0 < y \leq 1 \\[2ex] \displaystyle\int_{y-1}^1 dx_1 = 2 - y & 1 < y < 2 \end{cases}$$

Sums of More than Two Random Variables

If $X_i \sim f_{X_i}$, $i = 1, ..., n$, are n independent random variables, and random variable $Y_n = X_1 + X_2 + ... + X_n$, then the probability distribution of Y_n can be calculated by writing

$$Y_n = Y_{n-1} + X_n,$$

which is the sum of **two** random variables, and therefore the simple convolution result can be used. Similarly, the distribution of Y_{n-1} can be calculated by writing

$$Y_{n-1} = Y_{n-2} + X_{n-1},$$

which is the sum of **two** random variables, and again the simple convolution result can be used, etc.

The distribution of the sum of n independent random variables can therefore be calculated using this iterative approach.

2.8 Expectation and Variance

DEFINITION

For a **discrete** random variable X taking values in set $\mathbb{X}$ with mass function f_X, the **expectation** of X is defined by

$$\mathrm{E}_{f_X}[\,X\,] = \sum_{x\in\mathbb{X}} x f_X(x) \equiv \sum_{x=-\infty}^{\infty} x f_X(x)$$

as $f_X(x) \equiv 0$ for $x \notin \mathbb{X}$.

For a **continuous** random variable X taking values in interval $\mathbb{X}$ with pdf f_X, the expectation of X is defined by

$$\mathrm{E}_{f_X}[X] = \int_{\mathbb{X}} x f_X(x)\;dx \equiv \int_{-\infty}^{\infty} x f_X(x)\;dx$$

as $f_X(x) \equiv 0$ for $x \notin \mathbb{X}$.

DEFINITION

The **variance** of X is defined by

$$\mathrm{Var}_{f_X}[X] = \mathrm{E}_{f_X}[(X - \mathrm{E}_{f_X}[X])^2] = \mathrm{E}_{f_X}[X^2] - \{\mathrm{E}_{f_X}[X]\}^2.$$

Interpretation : The expectation and variance of a probability distribution can be used to aid description, or to characterize the distribution;

The EXPECTATION is a measure of *location*
The VARIANCE is a measure of *scale* or *spread*

of the distribution.

NOTES

(i) Take care when carrying out summation/integration to only include non-zero terms/integrand.

(ii) Sum/integral may be infinite ("divergent").

EXAMPLE Suppose that X is a discrete random variable taking values on $\mathbb{X} = \{0, 1, 2, ...\}$ with pdf

$$f_X(x) = \frac{\lambda^x}{x!}e^{-\lambda} \qquad x = 0, 1, 2, ...$$

and zero otherwise.

Then

$$\begin{aligned}
\mathrm{E}_{f_X}[\,X\,] &= \sum_{x=-\infty}^{\infty} x f_X(x)\;dx = \sum_{0}^{\infty} x\frac{\lambda^x}{x!}e^{-\lambda} \\
&= \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!} \\
&= \lambda e^{-\lambda}e^{\lambda} = \lambda
\end{aligned}$$

using the power series expansion for the exponential function

$$e^t = \sum_{x=0}^{\infty}\frac{t^x}{x!} = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + ...$$

EXAMPLE Suppose that X is a continuous random variable taking values on $\mathbb{X} = \mathbb{R}^+$ with pdf

$$f_X(x) = \frac{2}{(1+x)^3} \qquad x > 0.$$

Then

$$\mathrm{E}_{f_X}[\ X\ ] = \int_{-\infty}^{\infty} x f_X(x)\ dx = \int_0^{\infty} \frac{2x}{(1+x)^3}\ dx = 1$$

after integrating by parts.

RESULTS INVOLVING EXPECTATIONS

Suppose that X_1 and X_2 are independent random variables, and a_1 and a_2 are constants. Then if $Y = a_1X_1 + a_2X_2$,

$$\begin{aligned}\mathrm{E}_{f_Y}[Y] &= a_1\mathrm{E}_{f_{X_1}}[X_1] + a_2\mathrm{E}_{f_{X_2}}[X_2] \\ \mathrm{Var}_{f_Y}[Y] &= a_1^2\mathrm{Var}_{f_{X_1}}[X_1] + a_2^2\mathrm{Var}_{f_{X_2}}[X_2]\end{aligned}$$

so that, in particular (when $a_1 = a_2 = 1$) we have

$$\begin{aligned}\mathrm{E}_{f_Y}[Y] &= \mathrm{E}_{f_{X_1}}[X_1] + \mathrm{E}_{f_{X_2}}[X_2] \\ \mathrm{Var}_{f_Y}[Y] &= \mathrm{Var}_{f_{X_1}}[X_1] + \mathrm{Var}_{f_{X_2}}[X_2]\end{aligned}$$

so we have a simple additive property for expectations and variances. Note also that if $a_1 = 1, a_2 = -1$, then

$$\begin{aligned}\mathrm{E}_{f_Y}[Y] &= \mathrm{E}_{f_{X_1}}[X_1] - \mathrm{E}_{f_{X_2}}[X_2] \\ \mathrm{Var}_{f_Y}[Y] &= \mathrm{Var}_{f_{X_1}}[X_1] + \mathrm{Var}_{f_{X_2}}[X_2]\end{aligned}$$

Functions of a random variable

Suppose that X is a random variable, and $g(.)$ is some function. Then we can define the expectation of $g(X)$ (that is, the expectation of a function of a random variable) by

$$\mathrm{E}_{f_X}[\ g(X)\ ] = \begin{cases} \displaystyle\sum_{x=-\infty}^{\infty} g(x) f_X(x) & \text{DISCRETE} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f_X(x)\ dx & \text{CONTINUOUS} \end{cases}$$

Note that $Y = g(X)$ is also a random variable whose probability distribution we can calculate from the probability distribution of X.