Transformation of Random Variables

Problem 1

Given the probability density function $f_X(x)$ of a random variable X. Determine the probability density function $f_Y(y)$ of Y for the linear transformation $Y = aX + b$, where $a \neq 0$.

Problem 2

Consider the statistically independent, Gaussian distributed random variables $X_1, \ldots, X_n$ with zero mean and variance $\sigma^2 = 1$.

2.1 Determine the probability density function $f_Y(y)$ of the random variable

$$Y = X_1^2 + \ldots + X_n^2 .$$

2.2 Now assume $n = 2$. Sketch $f_Y(y)$ and determine the expectation value $\mathrm{E}\{Y\}$.

Please note:
In some literatures, you will find different notations for the probability density function and cumulative probability distribution function.

According to the present script, in the following, we will use for the probability density function and the cumulative probability distribution function of a random variable X the denotations $f_X(x)$ and $F_X(x)$, respectively!

Remarks on Transformation of Random Variables

A problem that arises frequently in practical applications of probability is the following.

Given a random variable X, which is characterized by its probability density function $f_X(x)$, determine the probability density function of the random variable $Y = g(X)$, where $g(X)$ is some given function of X.

When the mapping g from X to Y is one-to-one, the determination of $f_Y(y)$ is relatively straightforward. However, when the mapping is not one-to-one, as is the case, e.g. $Y = X^2$, we must be very careful in our derivation of $f_Y(y)$.

Example 1: Consider the random variable Y defined as

$$Y = aX + b\,, \quad a \neq 0\,, \tag{1}$$

where a and b are constants. Determine the probability density function $f_Y(y)$.

See the solution of problem 1 in our exercise.

Example 2: Consider the random variable Y defined as

$$Y = aX^3 + b\,, \quad a \neq 0\,. \tag{2}$$

Determine the probability density function $f_Y(y)$.

Analogous to the solution of problem 1, the mapping between X and Y is one-to-one. Hence, if $a > 0$ we have

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(aX^3 + b \leq y\right) \\ &= \Pr\left[X \leq \left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right] = F_X\left[\left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right]\,. \end{aligned} \tag{3}$$

If $a < 0$ we can get similarly

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(aX^3 + b \leq y\right) \\ &= \Pr\left[X \geq \left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right] = 1 - \Pr\left[X \leq \left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right] \\ &= 1 - F_X\left[\left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right]\,. \end{aligned} \tag{4}$$

Differentiation of (3) and (4) with respect to y yields the desired relationship between the probability density functions $f_X(x)$ and $f_Y(y)$ as

$$\begin{aligned} f_Y(y) &= \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} \\ &= \frac{1}{3\,|a|\left(\frac{y-b}{a}\right)^{\frac{2}{3}}} \cdot f_X\left[\left(\frac{y-b}{a}\right)^{\frac{1}{3}}\right]\,, \quad \text{for} \quad a \neq 0\,. \end{aligned} \tag{5}$$

Example 3: The random variable Y is defined as

$$Y \;=\; aX^2 + b\,, \quad a > 0\,. \tag{6}$$

Determine the probability density function $f_Y(y)$.

In contrast to the above examples, the mapping between X and Y is not one-to-one. To determine the cumulative probability distribution function of Y, we observe that

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) \;=\; \Pr\left(aX^2 + b \le y\right) \\ &= \Pr\left(|X| \le \sqrt{\frac{y-b}{a}}\right)\,. \end{aligned} \tag{7}$$

Hence,

$$F_Y(y) \;=\; F_X\left(\sqrt{\frac{y-b}{a}}\right) + 1 - F_X\left(-\sqrt{\frac{y-b}{a}}\right)\,. \tag{8}$$

Differentiating (8) with respect to y, we obtain $f_Y(y)$ in terms of the probability density function of X in the form

$$f_Y(y) \;=\; \frac{f_X\left(\sqrt{\frac{y-b}{a}}\right)}{2a\sqrt{\frac{y-b}{a}}} + \frac{f_X\left(-\sqrt{\frac{y-b}{a}}\right)}{2a\sqrt{\frac{y-b}{a}}}\,. \tag{9}$$

In this example, we observe that the equation $g(x) \;=\; ax^2 + b \;=\; y$ has two real roots,

$$x_1 \;=\; \sqrt{\frac{y-b}{a}} \quad \text{and} \quad x_2 \;=\; -\sqrt{\frac{y-b}{a}}$$

and that $f_Y(y)$ consists of two terms corresponding to these two solutions. That is

$$f_Y(y) \;=\; \frac{f_X\left(x_1 = \sqrt{\frac{y-b}{a}}\right)}{\left|g'\left(x_1 = \sqrt{\frac{y-b}{a}}\right)\right|} + \frac{f_X\left(x_2 = -\sqrt{\frac{y-b}{a}}\right)}{\left|g'\left(x_2 = -\sqrt{\frac{y-b}{a}}\right)\right|}\,, \tag{10}$$

where $g'(x_i) \;=\; \frac{\mathrm{d}}{\mathrm{d}x} g(x)\,|_{x=x_i}$ denotes the first derivative of $g(x)$.

In the general case, suppose that $x_1,\; x_2,\; \ldots,\; x_n$ are the real roots of the equation $g(x) \;=\; y$. Then the probability density function of the random variable $Y = g(X)$ may be expressed as

$$f_Y(y) \;=\; \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|}\,, \tag{11}$$

where $g'(x_i) = \frac{\mathrm{d}}{\mathrm{d}x} g(x)\,|_{x=x_i}$ denotes the first derivative of $g(x)$ and the roots x_i $(i = 1,\; 2,\; \ldots,\; n)$ are functions of Y.

Solution of Problem 1

First, let us calculate the cumulative probability distribution function $F_Y(y)$ of the random variable $Y = aX + b$ with $a \neq 0$.

$$F_Y(y) = \Pr(Y \leq y) = \Pr(aX + b \leq y) = \Pr(aX \leq y - b)$$

$$= \begin{cases} \Pr\left(X \leq \dfrac{y-b}{a}\right) & , \quad a > 0 \\ \Pr\left(X \geq \dfrac{y-b}{a}\right) & , \quad a < 0 \end{cases}$$

$$= \begin{cases} \Pr\left(X \leq \dfrac{y-b}{a}\right) & , \quad a > 0 \\ 1 - \Pr\left(X \leq \dfrac{y-b}{a}\right) & , \quad a < 0 \end{cases}$$

$$= \begin{cases} F_X\left(\dfrac{y-b}{a}\right) & , \quad a > 0 \\ 1 - F_X\left(\dfrac{y-b}{a}\right) & , \quad a < 0 . \end{cases} \tag{12}$$

From the equation $y = ax + b$ with $a \neq 0$, man can get

$$\mathrm{d}y = a \cdot \mathrm{d}x . \tag{13}$$

Differentiation of (12) with respect to y and inserting (13) yields

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

$$= \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}y} F_X\left(\dfrac{y-b}{a}\right) & , \quad a > 0 \\ -\dfrac{\mathrm{d}}{\mathrm{d}y} F_X\left(\dfrac{y-b}{a}\right) & , \quad a < 0 \end{cases}$$

$$= \begin{cases} \dfrac{1}{a} \cdot \dfrac{\mathrm{d}}{\mathrm{d}x} F_X\left(\dfrac{y-b}{a}\right) & , \quad a > 0 \\ -\dfrac{1}{a} \cdot \dfrac{\mathrm{d}}{\mathrm{d}x} F_X\left(\dfrac{y-b}{a}\right) & , \quad a < 0 \end{cases}$$

$$= \begin{cases} \dfrac{1}{a} \cdot f_X\left(\dfrac{y-b}{a}\right) & , \quad a > 0 \\ -\dfrac{1}{a} \cdot f_X\left(\dfrac{y-b}{a}\right) & , \quad a < 0 . \end{cases} \tag{14}$$

The probability density function $f_Y(y)$ is then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) , \quad \forall\, a \neq 0 . \tag{15}$$

A linear transformation of a random variable yields only a shift and scaling of its probability density function.

Solution of Problem 2

2.1 The probability density function of a Gaussian distributed random variable X is defined as

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] , \tag{16}$$

where μ is the expectation value or the mean value of X and σ^2 is the variance of X.

If the Gaussian random variable X_i has zero mean and unitary variance, then $f_{X_i}(x_i)$ can be written as

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i^2}{2}\right) . \tag{17}$$

When $y < 0$,

$$\begin{aligned} \text{since} \quad Y \geq 0 \quad &\Rightarrow \quad \Pr\{Y \leq 0\} = F_Y(y) = 0 , \quad \forall\, y < 0 \\ &\Rightarrow \quad f_Y(y) = 0 , \quad \forall\, y < 0 . \end{aligned} \tag{18}$$

When $y \geq 0$, we start with

$$\begin{aligned} \Pr\{y \leq Y \leq y + \mathrm{d}y\} &= f_Y(y)\mathrm{d}y \\ \Leftrightarrow \quad \Pr\left\{y \leq \sum_{i=1}^{n} X_i^2 \leq y + \mathrm{d}y\right\} &= f_Y(y)\mathrm{d}y . \end{aligned} \tag{19}$$

The set of points $(X_1, \ldots, X_n)$ lie within a spherical layer S_y with the radius $\sqrt{y}$ and the thickness $d\sqrt{y}$. Hence, we look for the probability $\Pr\{S_y\}$

$$f_Y(y)\mathrm{d}y = \Pr\{S_y\} = \int_{S_y} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n . \tag{20}$$

Since $X_1, \ldots, X_n$ are statistically independent, we have

$$\begin{aligned} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) , \end{aligned} \tag{21}$$

and consequently

$$f_Y(y)\mathrm{d}y = \int_{S_y} \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n . \tag{22}$$

Within the spherical layer S_y, the integrand

$$\frac{1}{(2\pi)^{\frac{n}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right) \;=\; (2\pi)^{-\frac{n}{2}}\,\mathrm{e}^{-\frac{y}{2}}$$

can be regarded as constant. The value of the integral in (22) is therefore the product of this constant value $(2\pi)^{-\frac{n}{2}}\,\mathrm{e}^{-\frac{y}{2}}$ and the volume of S_y. The volume can be calculated by the product of its surface $c(n)(\sqrt{y})^{n-1}$ and its thickness $\mathrm{d}\sqrt{y} \;=\; \dfrac{\mathrm{d}y}{2\sqrt{y}}$ as follows

$$f_Y(y)\mathrm{d}y \;=\; (2\pi)^{-\frac{n}{2}}\frac{c(n)}{2}\mathrm{e}^{-\frac{y}{2}}y^{\frac{n}{2}-1}\mathrm{d}y\,, \quad y \geq 0\,. \tag{23}$$

With the normalization $\displaystyle\int_{-\infty}^{\infty} f_Y(y)\mathrm{d}y \;=\; 1$, we can determine $c(n)$

$$\begin{aligned}\int_{-\infty}^{\infty} f_Y(y)\mathrm{d}y \;&=\; \frac{c(n)}{2}(2\pi)^{-\frac{n}{2}}\int_{0}^{\infty} y^{\frac{n}{2}-1}\mathrm{e}^{-\frac{y}{2}}\mathrm{d}y \\ &=\; \frac{c(n)}{2}(2\pi)^{-\frac{n}{2}}2^{\frac{n}{2}}\int_{0}^{\infty} t^{\frac{n}{2}-1}\mathrm{e}^{-t}\mathrm{d}t \\ &=\; \frac{c(n)}{2}(2\pi)^{-\frac{n}{2}}2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) \\ &=\; 1\,,\end{aligned} \tag{24}$$

where $t \;=\; \dfrac{y}{2}$ and $\Gamma(k)$ denotes the Gamma-function

$$\Gamma(k) \;=\; \int_{0}^{\infty} t^{k-1}\mathrm{e}^{-t}\mathrm{d}t\,.$$

Finally, inserting the result from (24) into (23) and comparing left and right side yields

$$f_Y(y) \;=\; \begin{cases} 0 & , \quad y<0 \\ 2^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)^{-1} y^{\frac{n}{2}-1}\mathrm{e}^{-\frac{y}{2}} & , \quad y \geq 0\,. \end{cases} \tag{25}$$

2.2 For $n=2$ holds

$$f_Y(y) \;=\; \begin{cases} 0 & , \quad y<0 \\ \frac{1}{2}\mathrm{e}^{-\frac{y}{2}} & , \quad y \geq 0\,, \end{cases} \tag{26}$$

where we have used $\Gamma(1) \;=\; \displaystyle\int_{0}^{\infty}\mathrm{e}^{-t}\mathrm{d}t \;=\; 1$.

The corresponding sketch is given as below.

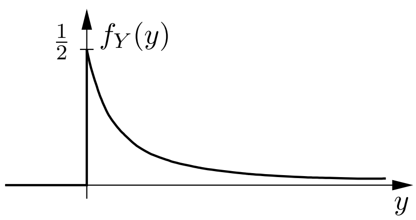

Figure 1: Sketch of $f_Y(y)$ for $n = 2$.

The expectation value $\mathrm{E}\{Y\}$ can be calculated by

$$\begin{aligned} \mathrm{E}\{Y\} &= \int_{-\infty}^{\infty} y f_Y(y)\mathrm{d}y \;=\; \frac{1}{2}\int_0^{\infty} y \mathrm{e}^{-\frac{y}{2}}\mathrm{d}y \\ &= 2\int_0^{\infty} t\mathrm{e}^{-t}\mathrm{d}t \;=\; 2\left(\left[t(-\mathrm{e}^{-t})\right]_0^{\infty} - \int_0^{\infty}\left(-\mathrm{e}^{-t}\right)\mathrm{d}t\right) \\ &= 2\int_0^{\infty} \mathrm{e}^{-t}\mathrm{d}t \\ &= 2\,. \end{aligned} \tag{27}$$