Functions of a random variable

1. *Linear function*: $Y = aX + b, a \neq 0$

Suppose X is a continuous random variable and it has cdf $F_X(x)$, find $F_Y(y)$;

$\{Y \leq y\}$ occurs when $A = \{aX + b \leq y\}$ occurs.

(i) $a > 0, A = \left\{X \leq \frac{y-b}{a}\right\}$

$$F_Y(y) = P\left[X \leq \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right),$$

(ii) $a < 0, A = \left\{X \geq \frac{y-b}{a}\right\}$

$$F_Y(y) = P\left[X \geq \frac{y-b}{a}\right] = 1 - F_X\left(\frac{y-b}{a}\right).$$

Using chain rule, $\dfrac{dF}{dy} = \dfrac{dF}{du}\,\dfrac{du}{dy}$

$$f_Y(y) = \begin{cases} \frac{1}{a} f_X\left(\frac{y-b}{a}\right) & a > 0 \\ -\frac{1}{a} f_X\left(\frac{y-b}{a}\right) & a < 0 \end{cases} = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

e.g. $f_X(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2}, \quad -\infty < x < \infty$

then $f_Y(y) = \dfrac{1}{\sqrt{2\pi}|a\sigma|} e^{-(y-b-am)^2/2(a\sigma)^2}.$

Y has mean $b + am$ and standard derivation $|a|\sigma$, and Y remains to be Gaussian.

Example

Suppose we take $a = \dfrac{1}{\sigma}, b = -am = \dfrac{-m}{\sigma}$, that is, $Y = \dfrac{X-m}{\sigma}$, then $f_Y(y) = \dfrac{1}{\sqrt{2\pi}}\, e^{-y^2/2}$. Now, Y is the standard Gaussian random variable with zero mean and unit standard deviation.

2. $Y = X^2$, X is a continuous random variable

$\{Y \le y\}$ occurs when $\{X^2 \le y\}$ or $\{-\sqrt{y} \le X \le \sqrt{y}\}, y \ge 0$. The event is null when y is negative.

$$F_Y(y) = \begin{cases} 0 & y \le 0 \\ F_X(\sqrt{y}) - F_X(-\sqrt{y}) & y > 0 \end{cases}$$

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} - \frac{f_X(-\sqrt{y})}{-2\sqrt{y}}, \quad y > 0$$

$$= \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}.$$

e.g. Let X be Gaussian with mean $m = 0$ and $\sigma = 1$ where $f_X(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$, then

$$f_Y(y) = \frac{e^{-(\sqrt{y})^2/2}}{\sqrt{2\pi}(2\sqrt{y})} + \frac{e^{-(-\sqrt{y})^2/2}}{\sqrt{2\pi}(2\sqrt{y})} = \frac{e^{-y/2}}{\sqrt{2y\pi}}, \quad y > 0.$$

General case

Suppose $g(x) = y$ has n solutions, then

$$f_Y(y) = \sum_{k=1}^{n} f_X(x_k) \left|\frac{dx}{dy}\right| \Big|_{x=x_k}$$

where $x_1, \cdots, x_n$ are the solutions.

e.g. $g(X) = X^2$; for $y > 0, y = x^2$ has two solutions: $\sqrt{y}$ and $-\sqrt{y}$.

Since $\dfrac{dy}{dx} = 2x$, so $f_Y(y) = f_X(\sqrt{y}) \left|\dfrac{1}{2x}\right|_{x=\sqrt{y}} + f_X(-\sqrt{y}) \left|\dfrac{1}{2x}\right|_{x=-\sqrt{y}}$

$$= \frac{f_X(\sqrt{y})}{2\sqrt{y}} + \frac{f_X(-\sqrt{y})}{2\sqrt{y}}.$$

Illustration of the result

Consider the event

$C_y = \{y < Y < y + dy\}$ and B_y be its equivalent event

$$B_y = \{x_1 < X < x_1 + dx_1\} \cup \{x_2 + dx_2 < X < x_2\} \cup \{x_3 < X < x_3 + dx_3\}$$

$$P[C_y] = f_Y(y)|dy|$$

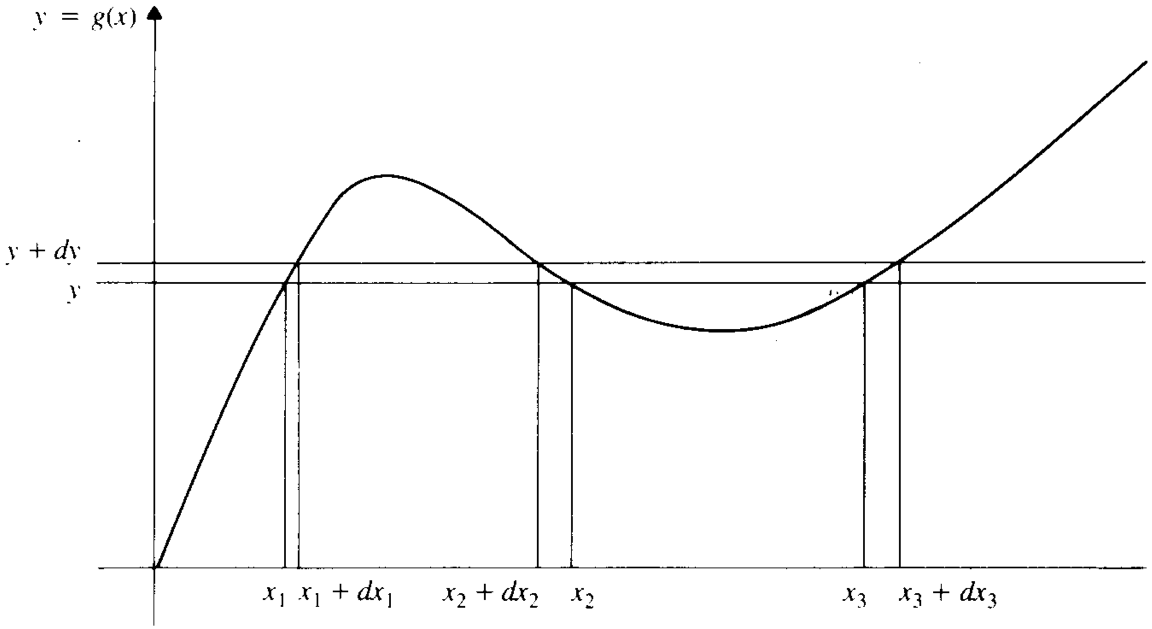
and

$$P[B_y] = f_X(x_1)|dx_1| + f_X(x_2)|dx_2| + f_X(x_3)|dx_3|;$$

so

$$f_Y(y) = \sum_k \frac{f_X(x_k)}{|dy/dx|_{x=x_k}} = \sum_k f_X(x_k) \left|\frac{dx}{dy}\right|_{x=x_k}.$$

Note that each $f_X(x_k)$ is multiplied by the scaling factor $|dx/dy|_{x=x_k}$.



Example

Consider $Y = \cos X$, where X is uniformly distributed in $[0, 2\pi]$. Since X is uniformly distributed, $f_X(x) = \begin{cases} c & \text{for} \quad x \in [0, 2\pi] \\ 0 & \text{for} \quad x \notin [0, 2\pi] \end{cases}$. By observing $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$, we obtain $c = \dfrac{1}{2\pi}$.

For $-1 < y < 1, y = \cos x$ has two solutions: $x_1 = \cos^{-1} y$ and $x_2 = 2\pi - \cos^{-1} y$.

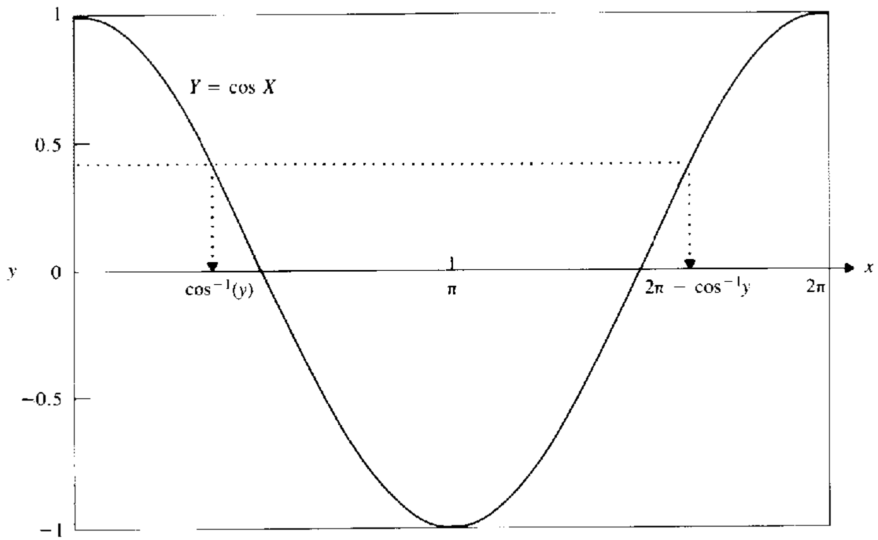
$$\left.\frac{dy}{dx}\right|_{x_1} = -\sin x_1 = -\sin(\cos^{-1} y) = -\sqrt{1 - y^2}, \quad 0 < x_1 < \pi.$$

Similarly, $\left.\frac{dy}{dx}\right|_{x_2} = \sqrt{1-y^2}, \pi < x_2 < 2\pi$. By applying the formula:

$$f_Y(y) = \frac{1}{2\pi\sqrt{1-y^2}} + \frac{1}{2\pi\sqrt{1-y^2}} = \frac{1}{\pi\sqrt{1-y^2}} \quad \text{for} \quad -1 < y < 1.$$

$$\text{cdf of } Y = F_Y(y) = \int_{-\infty}^{y} f_Y(y')\, dy' = \begin{cases} 0 & y < -1 \\ \frac{1}{2} + \frac{\sin^{-1} y}{\pi} & -1 \le y \le 1 \\ 1 & y > 1 \end{cases};$$

Y is called the *arc sine distribution*.



Some properties on expected value

Suppose the pdf is symmetric about a point m, that is,

$$f_X(m-x) = f_X(m+x), \quad \text{for all } x.$$

Assuming $E[X]$ exists and consider

$$\begin{aligned}
m - \int_{-\infty}^{\infty} t f_X(t)\ dt &= \int_{-\infty}^{\infty} (m-t) f_X(t)\ dt \\
&= \int_{-\infty}^{m} (m-t) f_X(t)\ dt + \int_{m}^{\infty} (m-t) f_X(t)\ dt \\
&= \int_{-\infty}^{0} -u f_X(m+u)\ du + \int_{0}^{\infty} -u f_X(m+u)\ du, \quad u = t-m \\
&= \int_{0}^{\infty} x f_X(m-x)\ dx - \int_{0}^{\infty} u f_X(m+u)\ du = 0,
\end{aligned}$$

so that $m = E[X]$.

For example, the pdf of a Gaussian random variable is symmetric about $x = m$, and so $E[X] = m$.

When X is a non-negative random variable

(i) $E[X] = \int_0^\infty [1 - F_X(t)]\, dt, \quad X$ is continuous

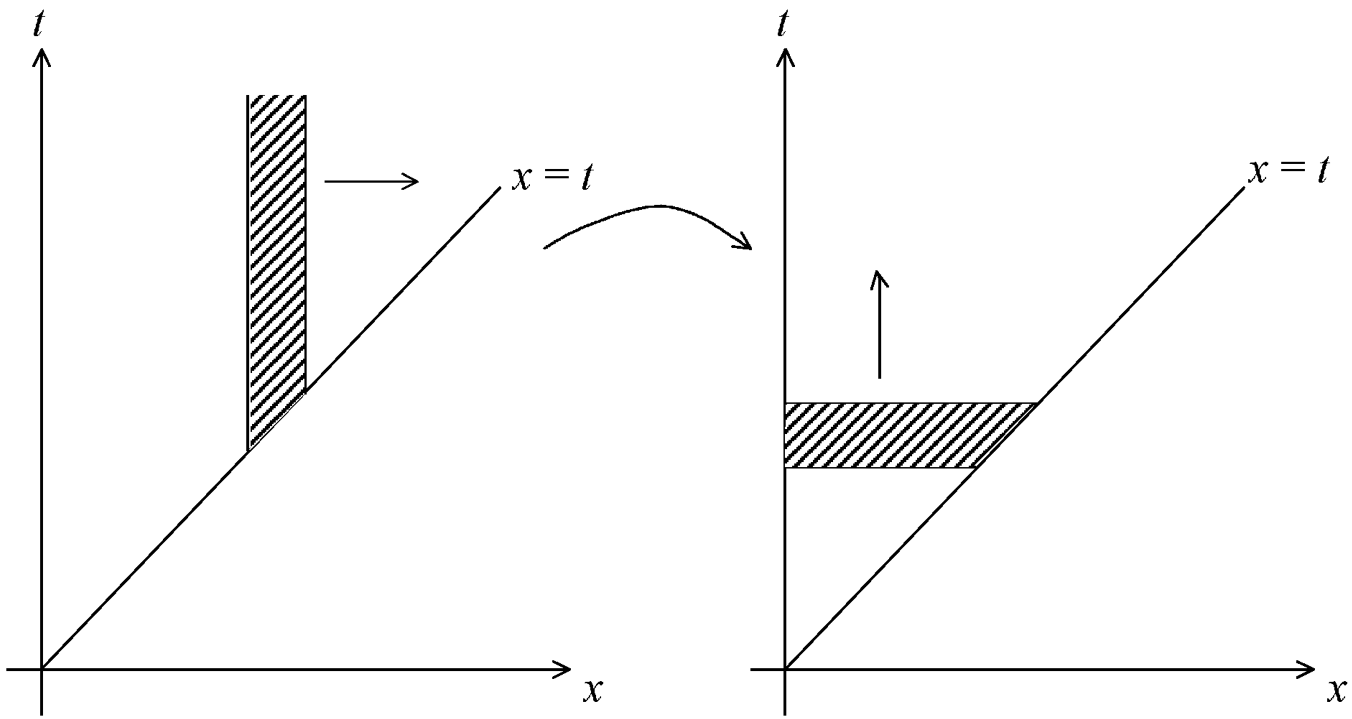
(ii) $E[X] = \sum_{k=0}^{\infty} P[X > k]$, if X is discrete and assumes non-negative integer values.

Proof

(i) Write

$$\int_0^\infty P[X > x]\ dx = \int_0^\infty \int_x^\infty f_X(t)\ dt\ dx.$$

Interchange the order of integration



$$\int_0^\infty P[X > x]\ dx = \int_0^\infty \left(\int_0^t dx \right) f_X(t)\ dt = \int_0^\infty t f_X(t)\ dt = E[X].$$

(ii) Consider $\sum_{k=0}^{N} P[X > k] = \sum_{k=0}^{N} \sum_{\ell=k+1}^{N} P_X(\ell) = \sum_{k=0}^{N} kP_X(k)$.

$$
\begin{array}{ll}
k = 0: & P_X(1) + P_X(2) + \cdots + P_X(N) \\
k = 1: & \qquad\quad + P_X(2) + \cdots + P_X(N) \\
 & \qquad\qquad \vdots \qquad\qquad \vdots \\
k = N-1: & \qquad\qquad\qquad\qquad P_X(N) \\
\hline
 & P_X(1) + 2P_X(2) + \cdots + NP_X(N)
\end{array}
$$

Taking the limit $N \to \infty$, we have

$$\sum_{k=0}^{\infty} P[X > k] = \sum_{k=0}^{\infty} kP_X(k) = E[X].$$

Expected value of $Y = g(X)$

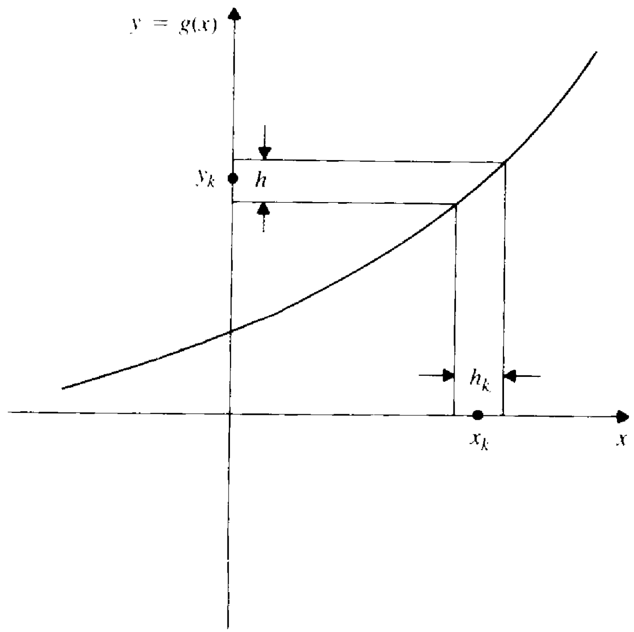
Direct approach: we first find the pdf of Y

$$E[Y] \approx \sum_k y_k f_Y(y_k) h.$$

Suppose $g(x)$ is *strictly increasing*

$$\begin{aligned} f_Y(y_k)h &= f_X(x_k)h_k \\ E[Y] &\approx \sum_k g(x_k) f_X(x_k) h_k. \end{aligned}$$

Taking $h \to 0$, $E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$. The result is valid even if $g(x)$ is not strictly increasing.



Example Indicator function

$$Y = g(X) = I_C(X) = \begin{cases} 0 & X \text{ not in } C \\ 1 & X \text{ in } C \end{cases}$$

$$E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x)\ dx = \int_C f_X(x)\ dx = P[X \text{ in } C].$$

The integration over C refers to the integration over the interval of x that corresponds to the occurrence of the event C.

For example, C is the event that "5" appears in a tossing of a fair dice. We have $E[I_C(X)] = 1/6$, where X is the discrete random variable representing the number shown on the dice.

Linearity of expectation operator

Suppose $Y = \sum_{k=1}^{n} g_k(X)$, we have

$$\begin{aligned} E[Y] &= E\left[\sum_{k=1}^{n} g_k(X)\right] = \int_{-\infty}^{\infty} \sum_{k=1}^{n} g_k(x) f_X(x)\ dx \\ &= \sum_{k=1}^{n} \int_{-\infty}^{\infty} g_k(x) f_X(x)\ dx = \sum_{k=1}^{n} E[g_k(X)]. \end{aligned}$$

Example

$$Y = g(X) = a_0 + a_1 X + \cdots + a_n X^n$$

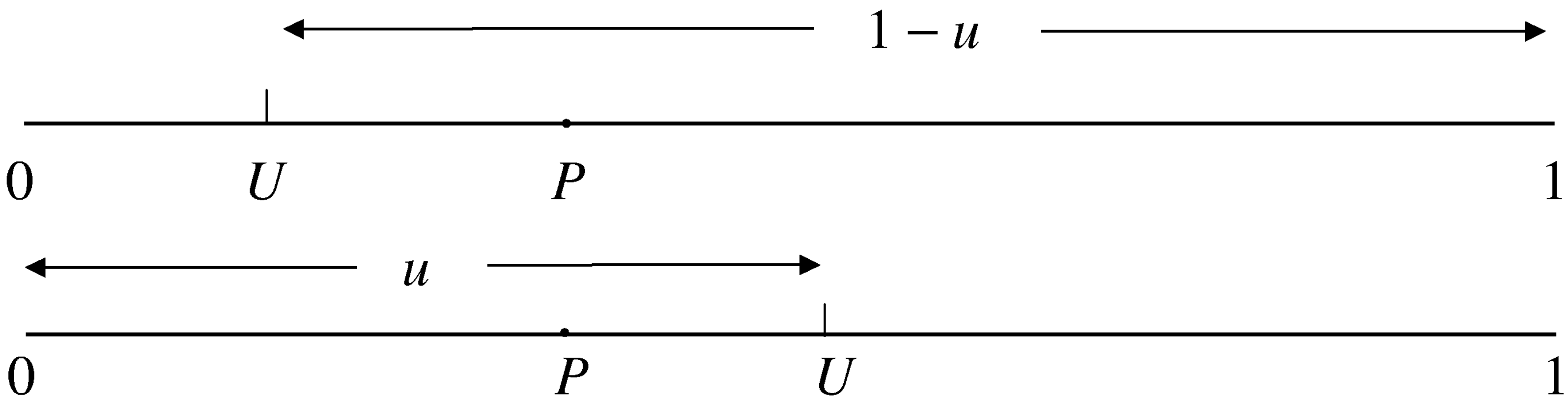
$$\begin{aligned} E[Y] &= E[a_0] + E[a_1 x] + \cdots + E[a_n X^n] \\ &= a_0 + a_1 E[X] + \cdots + a_n E[X^n]. \end{aligned}$$

Remark

$$n\text{th moment of } X = E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\ dx.$$

$$\begin{aligned} \text{VAR}[X] &= E[X^2] - E[X]^2 \\ &= 2^{\text{nd}}\text{moment} - (1^{\text{st}}\text{moment})^2. \end{aligned}$$

Example A stick of unit length is split at a point U that is *uniformly distributed* over $(0, 1)$. Determine the expected length of the piece that contains the particular point P. Here, let p be the distance of P from the left end "0", $0 \leq p \leq 1$.



Solution Let u be the distance of U from the end "0" and $L_p(u)$ denote the length of the substick that contains the point P, and note that

$$L_p(u) = \begin{cases} 1 - u & u < p \\ u & u > p \end{cases} .$$

Note that $f_U(u) = 1$ for $0 \le u \le 1$ and

$$\begin{aligned} E[L_p(u)] &= \int_0^1 L_p(u) f_U(u)\ du \\ &= \int_0^p (1-u)\ du + \int_p^1 u\ du \\ &= \left(p - \frac{p^2}{2}\right) + \left(\frac{1}{2} - \frac{p^2}{2}\right) \\ &= \frac{1}{2} + p(1-p). \end{aligned}$$

Since $p(1-p)$ is maximized when $p = \dfrac{1}{2}$, it is interesting to note that the expected length of the substick containing the point P is maximized when P is the midpoint of the original stick. When $p = \dfrac{1}{2}, E[L_p(u)] = \dfrac{3}{4}$.