Brief Notes #5
Functions of Random Variables and Vectors

(a) Functions of One Random Variable

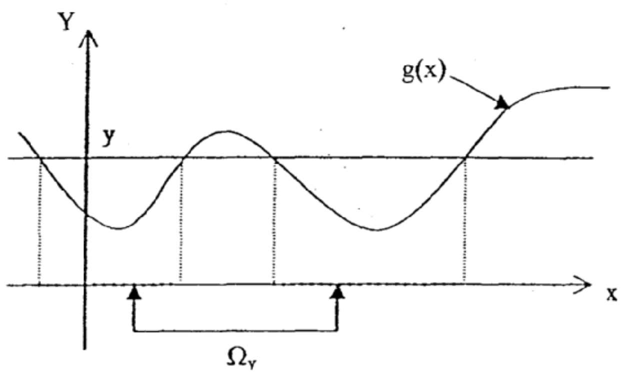
- **Problem:**

Given the CDF of the random variable X, $F_X(x)$, and a deterministic function $Y = g(x)$, find the (derived) distribution of the random variable Y.

- **General solution:**

Let $\Omega_y = \{x: g(x) \le y\}$. Then:

$$F_Y(y) = P[Y \le y] = P[x \in \Omega_y] = \int_{\Omega_y} f_X(x)dx$$



- **Special cases:**
  - Linear functions:

$Y = g(x) = a + bx$

If $b > 0$:

$$X(y) = \frac{y-a}{b};\ \Omega_y = \{x : a + bx \le y\} = \left(-\infty, \frac{y-a}{b}\right]$$

$$F_Y(y) = P[x \in \Omega_y] = F_X\left(\frac{y-a}{b}\right)$$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y-a}{b}\right) = \frac{1}{b}f_X\left(\frac{y-a}{b}\right)$$

If b < 0:

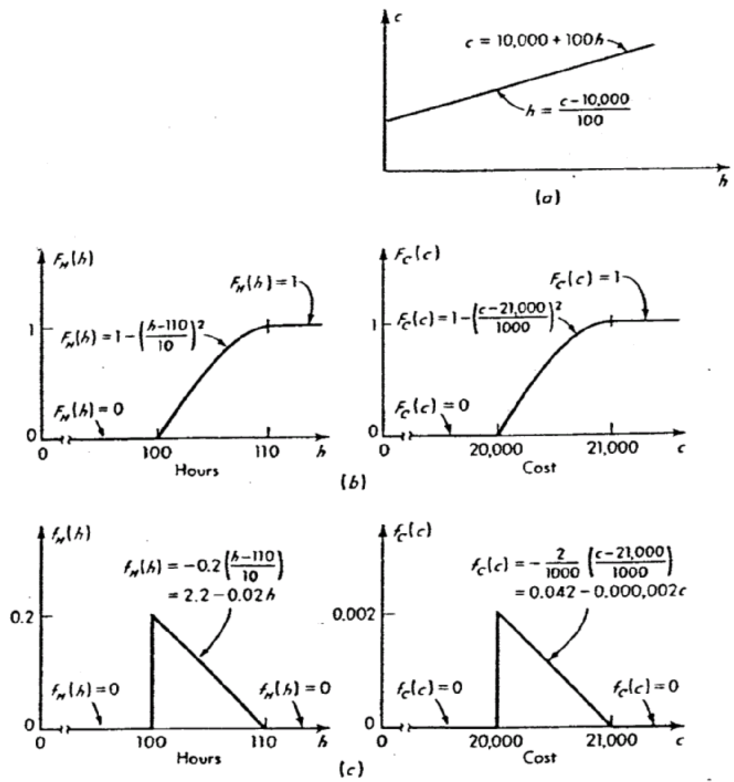
$$\Omega_y = \left[\frac{y-a}{b}, \infty\right)$$

$$F_Y(y) = 1 - F_X\left(\frac{y-a}{b}\right)$$

$$f_Y(y) = -\frac{1}{b} f_X\left(\frac{y-a}{b}\right) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$$

For any b ≠ 0:

$$f_Y(y) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$$



Example of linear transformation (b > 0): derived distributions, construction-cost illustration, *C* = 10,000 + 100*H*. (a) Functional relationship between cost and time; (b) cumulative distribution function of *H*, given, and *C*, derived; (c) probability density function of *H*, given, and *C*, derived.

• <u>General monotonic (one-to-one) functions</u>

• *Monotonically increasing functions*

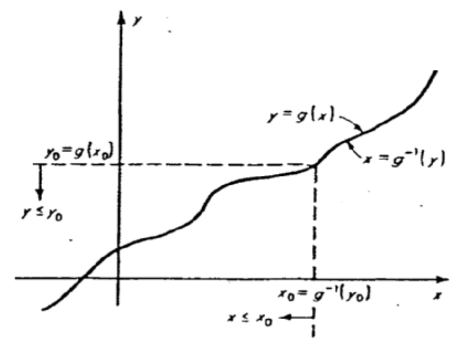
$$F_Y(y) = F_X[x(y)]$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dx(y)}{dy} \cdot f_X[x(y)]$$

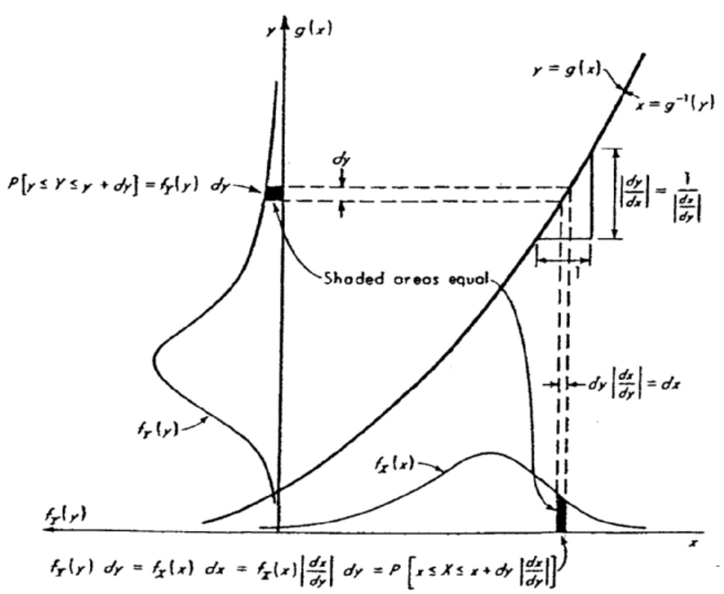
• *Monotonically decreasing functions*

$$F_Y(y) = 1 - F_X[x(y)]$$

$$f_Y(y) = \left|\frac{dx(y)}{dy}\right| \cdot f_X[x(y)]$$



A monotonically increasing one-to-one function relating Y to X.



Graphical interpretation of $f_Y(y) = \frac{dx}{dy} f_X(x)$.

Examples of Monotonic Transformations

Consider an exponential variable $X \sim EX(\lambda)$ with cumulative distribution function $F_X(x) = 1 - e^{-\lambda x}$, $x \geq 0$.

Exponential, Power and Log Functions

Exponential Functions

Suppose $Y = e^X$, $\Rightarrow$ $X = \ln Y$, $y \geq 0$. This is a monotonic increasing function, and $F_Y(y) = F_X(x(y)) = 1 - e^{-\lambda \ln y} = 1 - y^{-\lambda}$. This distribution is known as the (strict) Pareto Distribution.

Power Functions

Suppose $Y = X^{\frac{1}{\alpha}}$, $\alpha > 0$ $\Rightarrow$ $X = \ln Y$, $y \geq 0$. This is a monotonic increasing function, and $F_Y(y) = F_X(x(y)) = 1 - e^{-\lambda y^{\alpha}}$. This distribution is known as the Weibull (Extreme Type III) Distribution.

Log Functions

Suppose $Y = -\ln X$, $\Rightarrow$ $X = e^{-Y}$, $-\infty \leq y \leq \infty$. This is a monotonic decreasing function, and $F_Y(y) = 1 - F_X(x(y)) = e^{-\lambda e^{-y}}$. This distribution is known as the Gumbel (Extreme Type I) Distribution.

(b) Functions of Two or More Random Variables

- **Problem:**

Given the JCDF of the random vector $\begin{bmatrix} X \\ Y \end{bmatrix}$, $F_{X,Y}(x, y)$, and a deterministic function $Z = g(x, y)$, find the (derived) distribution of the random variable Z.

- **General solution:**

Let $\Omega_z = \{x, y: g(x, y) \le z\}$. Then:

$$F_Z(z) = P[Z \le z] = P[(x, y) \in \Omega_z] = \iint_{\Omega_z} f_{X,Y}(x, y)dxdy$$

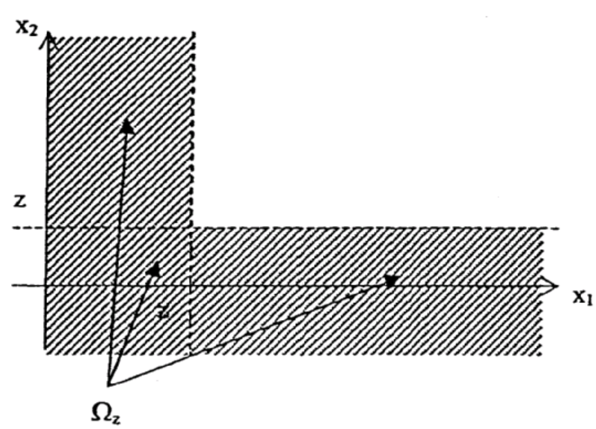
- **Special cases:**

  - Minimum/maximum functions i.e. $Z = Min[X_1, X_2, \ldots, X_n]$ (e.g. minimum strength) or $Z = Max[X_1, X_2, \ldots, X_n]$ (e.g. maximum load)

  - **$Z = Min[X_1, X_2, \ldots, X_n]$.** For n = 2,

$$F_Z(z) = P[Z \le z] = \iint_{\Omega_z} f_{X_1,X_2}(x_1, x_2)dx_1dx_2 \quad \text{with } \Omega_z \text{ shown in the figure}$$

$$= 1 - \int_z^{\infty} dx_1 \int_z^{\infty} f_{X_1,X_2}(x_1, x_2)dx_2$$



If X_1 and X_2 are independent:

$$\int_z^\infty dx_1 \int_z^\infty f_{X_1,X_2}(x_1,x_2)dx_2 = [1 - F_{X_1}(z)][1 - F_{X_2}(z)]$$

Therefore,

$$F_Z(z) = 1 - [1 - F_{X_1}(z)][1 - F_{X_2}(z)]$$

For n iid variables:

$$F_Z(z) = P[Z \le z] = 1 - P[(X_1 > z) \cap \ldots \cap (X_n > z)]$$

$$= 1 - [1 - F_X(z)]^n$$

or, with $G_X(x) = 1 - F_X(x)$,

$$G_Z(z) = P[Z > z] = [G_X(z)]^n$$

$$f_Z(z) = \frac{d}{dz}F_Z(z) = -\frac{d}{dz}G_Z(z) = n[G_X(z)]^{n-1} f_X(z)$$

- **$Z = \mathrm{Max}[X_1, X_2, \ldots, X_n]$**

$$F_Z(z) = P\left[\bigcap_i (X_i \le z)\right] = F_{\underline{X}}\begin{bmatrix} z \\ \vdots \\ z \end{bmatrix}$$

$$= \prod_i F_{X_i}(z) \qquad \text{(if } X_i\text{'s are independent)}$$

$$= [F_X(z)]^n \qquad \text{and} \qquad f_Z(z) = n[F_X(z)]^{n-1} f_X(z) \qquad \text{(if } X_i\text{'s are iid)}$$

- <u>Linear transformations</u>

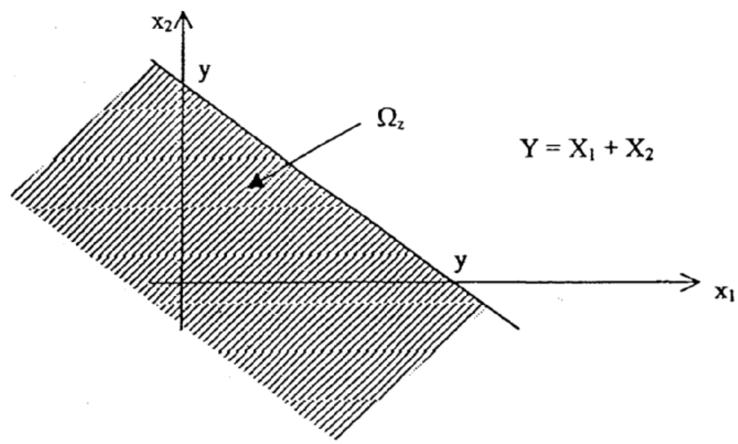
$$Y = \sum_i a_i x_i$$

- **Simplest case: $Y = X_1 + X_2$**

$$F_Y(y) = P[Y \le y] = P[x_1 + x_2 \le y] = \iint\limits_{x_1+x_2\le y} f_{X_1,X_2}(x_1, x_2)dx_1dx_2$$

$$= \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{y-x_2} f_{X_1,X_2}(x_1, x_2)dx_1$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1,X_2}(y - x_2, x_2)dx_2$$



If X_1 and X_2 are independent, then:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(y - x_2)f_{X_2}(x_2)dx_2 \quad \text{(convolution)}$$

- **Example: derivation of Gamma distribution**

Consider $Y = X_1 + X_2$, where X_1 and X_2 are iid exponential, with density:

$$f_{X_i}(x) = \begin{Bmatrix} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{Bmatrix}$$

Then,

$$f_Y(y) = \int_0^\infty f_X(y - x_1) f_X(x_1) dx_1$$

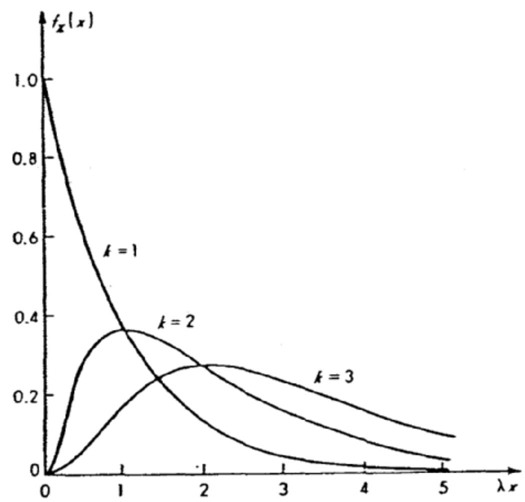
$$= \lambda^2 y e^{-\lambda y} \qquad \text{(Rayleigh or Gamma(2) distribution)}$$

In general, for any n, the probability density of $Y = X_1 + X_2 + \ldots + X_n$, where the X_i are iid exponential, is:

$$f_Y(y) = \frac{\lambda(\lambda y)^{n-1} e^{-\lambda y}}{\Gamma(n)}, \quad y \geq 0, \text{ where } \Gamma(n) = (n-1)!$$

(Gamma(n) distribution)

Note: for n = 1, the Gamma distribution reduces to the exponential distribution.



Gamma distributions $G(k,\lambda)$.