7.12 Function of one random variable - derived distribution

X is a RV. Let a quantity Y be functionally dependent on X:

$$Y = g(X) \tag{7.50}$$

Then, Y too is a random variable, if the function g holds certain properties (P&P p. 123):

1. Its domain must include the range of X.
2. It must be a Borel function, i.e., for every y, the set R_y such that $g(x) \le y$ must consist of the union and intersection of a countable no. of intervals. Only then is $Y \le y$ an event.
3. The events $g(X) = \pm\infty$ must each have zero probability.

What are the probabilistic characteristics of Y? We are interested to find the distribution of Y and its statistics such as mean and s.d. from the distribution of X. As we will see, the event $\{Y \le y\}$ implies $\{X \in R_y\}$, where R_y is some range depending on y, so that, the CDF of Y is:

$$F_Y(y) = P[X \in R_y] \tag{7.51}$$ aaa

Two cases will be looked at for the function of one RV, one when g is a monotonic function, and two, when it is not. F_X is known in every case.

7.12.1 g(X) is one to one

7.12.1.1 X is discrete

Since g is one to one, $Y= g(X)$ too is discrete.

$$\begin{aligned} p_Y(y_i) &= P[Y = y_i] = P[X = g^{-1}(y_i)] = p_X(g^{-1}(y_i)) \\ F_Y(y) &= P[Y \le y] = P[g(X) \le y] = P[X \le g^{-1}(y)] = \sum_{\text{all } x_i \le g^{-1}(y)} p_X(x_i) \end{aligned} \tag{7.52}$$

7.12.1.2 X is continuous

Since g is one to one, $Y = g(X)$ too is continuous.

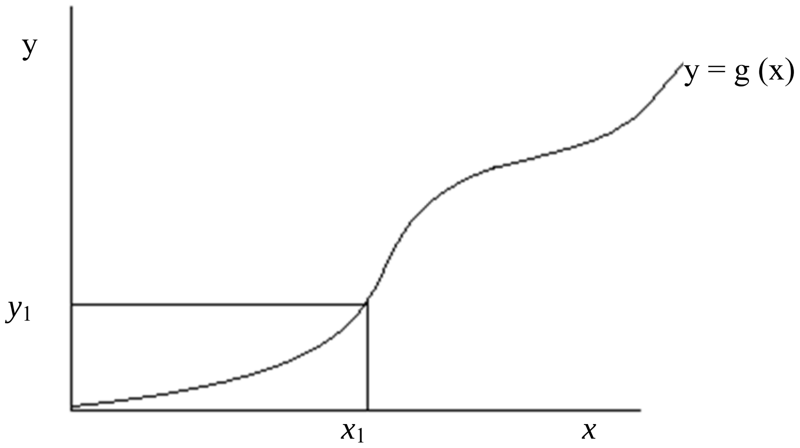


Figure 7-1: monotonic tranformation from x to y

It is clear from Figure 7-2 that $\{Y \le y_1\} \Leftrightarrow \{X \le x_1\}$ for any (x_1, y_1) pair. Hence, the CDF of *Y* can be written as:

$$F_Y(y) = P[Y \le y] = P[g(X) \le y] = P[X \le g^{-1}(y)] \tag{7.53}$$

which, by definition, is the integral:

$$F_Y(y) = P[X \le g^{-1}(y)] = \int_{-\infty}^{g^{-1}(y)} f_X(x)dx \tag{7.54}$$

Using Leibniz rule[7] for differentiation under the integral sign, we obtain the PDF of Y as:

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(g^{-1}(y)) \\ &= \frac{d}{dg^{-1}}F_X(g^{-1}(y))\frac{dg^{-1}(y)}{dy} \\ &= f_X(g^{-1}(y))\,|\frac{dg^{-1}(y)}{dy}| \end{aligned} \tag{7.55}$$

The absolute value is imposed to account for cases when *g*(*X*) is a decreasing function.

---

[7] Leibniz rule for differentiation under the integral sign:

$$\frac{d}{d\alpha}\int_{\phi_1(\alpha)}^{\phi_2(\alpha)} F(x,\alpha)dx == \int_{\phi_1(\alpha)}^{\phi_2(\alpha)} \frac{\partial F}{\partial \alpha}dx + F(\phi_2,\alpha)\frac{d\phi}{d\alpha} - F(\phi,\alpha)\frac{d\phi_1}{d\alpha}$$

7.12.1.3 Example: Normal to lognormal transformation

$$X \sim N(\mu_X, \sigma_X) \qquad f_X(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_X} e^{-\frac{1}{2}(\frac{x-\mu X}{\sigma_X})^2}$$

$$Y = e^X$$

$$f_Y(y) = f_X(\ln y)\left|\frac{d \ln y}{dy}\right| = \frac{1}{y} f_X(\ln y)$$

$$x = g^{-1}(y) = l\,\mathrm{n}\, y \qquad = \frac{1}{y}\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_X} e^{-\frac{1}{2}(\frac{l\,\mathrm{n}\, y - \mu X}{\sigma X})^2}$$

7.12.1.4 Example: laser directed at a random angle

In Section 6.1.2 we looked at the Cauchy distribution as the tangent of a Uniform random variable. The problem was posed as a laser gun being directed on a wall at a random angle: the distribution of the angle was known and the distribution of the projection was required. Here we look at the problem differently: the distribution of the location is known, that of the angle is required.

7.12.2 g(X) is many to one

7.12.2.1 *X* and *Y* are discrete

$$y = g(x) \quad \text{and} \quad g^{-1}(y) = \{x_1, x_2 ...... x_k\}$$

$$\therefore \{Y = y\} = \bigcup_{i=1}^{k} (X = x_i) = \bigcup_{i=1}^{k(y)} \{X = x_i(y)\}$$

$$\therefore p_Y(y) = \sum_{i(y)=1}^{k(y)} p_X(x_i(y))$$

7.12.2.2 X and Y are continuous

If *X* and *Y* are continuous, then the PDF of *Y*,

$$f_Y(y) = \sum_{i=1}^{k} f_X(x_i)\left|\frac{dx_i}{dy}\right|, \quad where \ g^{-1}(y) = \{x_1, x_2, ..., x_k\} \tag{7.56}$$

and the CDF of Y is given by Eq aaa above.

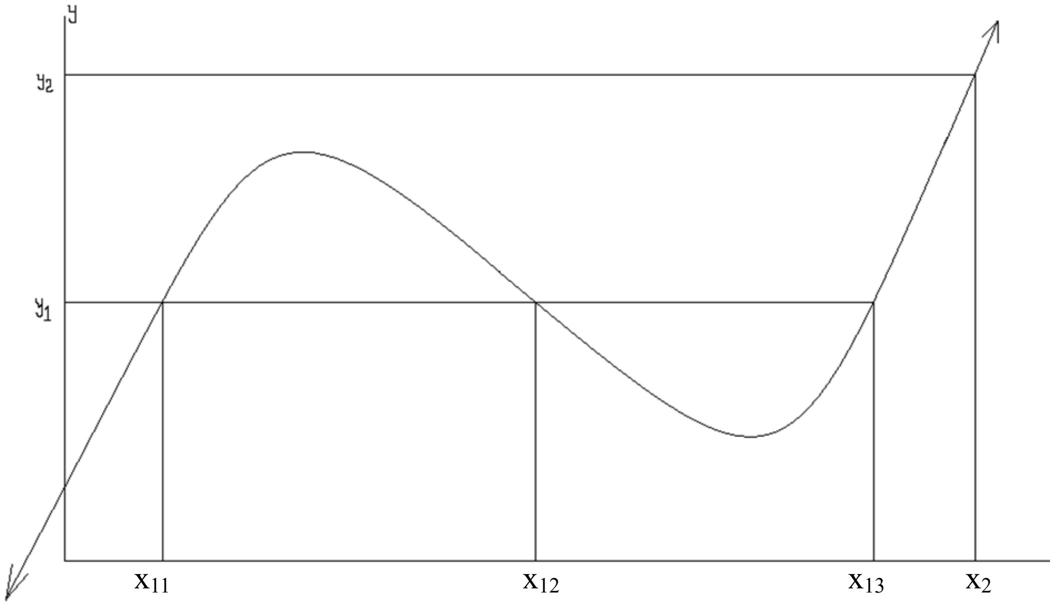


Figure 7-2: non-monotonic tranformation from x to y

For example, in Figure 7-2, y_2 has only one inverse but y_1 has three. Consequently, $\{Y \le y_2\} = \{X \le x_2\}$ and $\{Y \le y_1\} = \{X \le x_{11} \bigcup x_{12} \le X \le x_{13}\}$

7.12.2.3 Example: Normal to Chi squared transformation

$$X \sim N(0,1)$$

$$Y = X^2$$

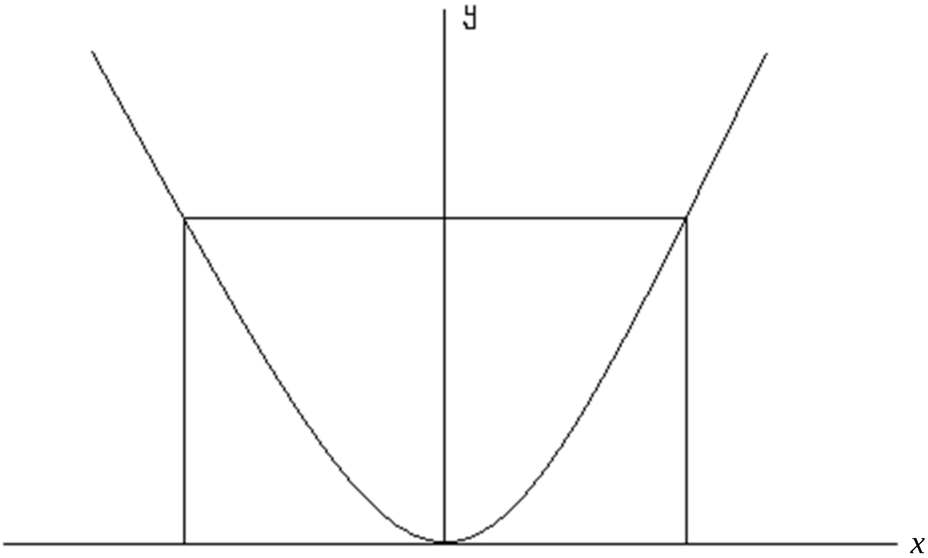
$$P[Y \le y] = P[-\sqrt{y} \le X \le \sqrt{y}]$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$f_Y(y) = f_X(\sqrt{y})\frac{d}{dy}\sqrt{y} - f_X(-\sqrt{y})\frac{d(-\sqrt{y})}{dy}$$

$$= f_X(\sqrt{y})\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\frac{1}{2\sqrt{y}}$$

$$= \frac{1}{2\sqrt{y}}[\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y} + \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y}] = \frac{1}{\sqrt{2\pi y}}e^{-\frac{1}{2}y}$$



Example:

$U = CP^2, \qquad P \sim N(\mu_\rho, \sigma_\rho)$

$f_U = ?$

Set $U = u$, then $P = \pm\sqrt{\frac{u}{c}}$, so that $p_1 = -\sqrt{u/c}$ and $p_2 = +\sqrt{u/c}$

Now,

$$\left|\frac{dp_1}{du}\right| = \frac{1}{2}\frac{1}{\sqrt{uc}}$$

$$\left|\frac{dp_2}{du}\right| = \frac{1}{2}\frac{1}{\sqrt{uc}}$$

Then, by eqn (7.56) the pdf of U is:

$$f_U(u) = \sum_{i=1}^{2} f_p(pi)\left|\frac{dpi}{du}\right|$$

$$= \frac{1}{\sqrt{2\pi}\sigma_P}[e^{-\frac{1}{2}(\frac{-\sqrt{u/c}-\mu P}{\sigma p})^2} + e^{-\frac{1}{2}(\frac{\sqrt{u/c}-\mu p}{\sigma p})^2}]\frac{1}{2}\frac{1}{\sqrt{uc}}$$

Alternately, the problem can be solved from first principles:

Given, $U = cP^2$, we can write:

$$P[U \le u] = P[cP^2 \le u]$$

$$= P[-\sqrt{\frac{u}{c}} \le P \le \sqrt{\frac{u}{c}}]$$

$$= \phi[\frac{\sqrt{u/c} - \mu p}{\sigma_p}] - \phi[\frac{-\sqrt{u/c} - \mu p}{\sigma_p}]$$

$$f_U(u) = \quad \phi(\frac{\sqrt{\frac{u}{c}} - \mu p}{\sigma_P}]\frac{1}{2\sqrt{uc}}\frac{1}{\sigma p} - \phi(\frac{-\sqrt{u/c} - \mu p}{\sigma_P})\frac{-1}{2\sqrt{uc}\sigma_P}.$$

$$= \frac{1}{2\sqrt{2\pi}\sqrt{uc}\sigma_P}[e^{-\frac{1}{2}(\frac{\sqrt{u/c}-\mu p}{\sigma_P})^2} + e^{-\frac{1}{2}(\frac{-\sqrt{u/c}-\mu p}{\sigma_P})^2}]$$

Example:

The range R of a javelin thrower is given by :

$$R = \frac{V_0^2}{g}\sin 2\phi$$

where V_0 is the initial velocity, g is acceleration due to gravity, ϕ is the angle made with the horizontal at the time of throw.

a) Variabilities in the thrower's performance makes V_0 a random variable, with mean m0 and standard deviation s0. Find the approximate mean and standard deviation of R in terms of m0, s0, ϕ and g.

b) Now assume that V_0 is Logormally distributed, with m0 = 20 m/sec and s0 = 2m/sec. For a throwing angle ϕ = 45 degrees, what is the probability that the thrower's range will be more than 52 meters?

Example: wind induced wave ht, Z = $\frac{f}{14000}V^2$

where f= fetch, d = depth of lake, V = wind speed .

Since V = random, so is Z .

Example:

Wave force on cylinder:

F = $F_{drag} + F_{inertia}$ = $C_d \frac{\rho}{2g} aU|U| + C_m \frac{\rho}{g} v\dot{U}$

Cd=drag coeff

Cm=inertia coeff

a=area of cylinder/length

v= volume of cylinder / length

U= water velocity

U(dot)=water acceleration

7.12.3 X is continuous but Y is discrete

Continuous to discrete transforms are also possible.

say *X* is measurement of damage (continuous) and *Y* is damage class (discrete).

So, $X \in A_1 \Rightarrow Y = y_1$, then:

$$P[Y = y_i] = P[X \in A_i] = \int_{ai}^{bi} f_X(x)dx$$

7.13 Function of several random variables

$$Y = g(X_1, X_2, ...X_n) \tag{7.57}$$

Special case when *g* is a linear combination is discussed in Section 7.15.

PDF of Y?

7.13.1 Example: convolution of two random variables

Wave height (*H*) and wave period (*T*) at a location off Hawaii during a storm are jointly distributed random variables. The joint probability density function is given by:

$$f_{H,T}(h,t) = \begin{cases} k(35 - h - t), & 0 < h < 20,\ 0 < t < 15 \\ 0, & \text{otherwise} \end{cases}$$

where *h* is in feet and *t* is in seconds.

a) Find *k*.

b) It has been found that a certain offshore installation will be safe as long as $H + T < 10$. What is the probability that the installation will be safe?

Answer:

(a) By equating the volume under the joint PDF to one, we obtain k = 1/5250

(b) The probability that the installation is safe is given by:

$$\begin{aligned} P[H+T<10] &= \int_{\text{all } t}\int_{\text{all } h} I(h+t<10)\, f_{H,T}(h,t)\, dh\, dt \\ &= k\int_{t=0}^{10}\int_{h=0}^{10-t} (35-h-t)\, dh\, dt \\ &= k\int_{t=0}^{10} \left(35h - h^2/2 - ht\right)_0^{10-t} dt \\ &= k\int_{t=0}^{10} (t^2/2 - 35t + 300)\, dt \\ &= 0.270 \end{aligned}$$

7.13.2 Example: difference of two exponential random variables

A structure has exponentially distributed capacity with mean μ_C. The load, independent of the capacity, is also exponential with mean μ_D.

(a) Find the reliability of the structure.

(b) A proof load test is performed on the structure as follows. A known load, c_0, is placed on the structure, and the structure survives without any damage. With this new information, find the updated reliability of the structure.

Answer:

(a) Taking advantage of the indepdence between C and D we can write the reliability = P$[D < C]$ as: $\int_{\text{all}\,c,d} I(d < c) f_{C,D}(c,d)\,dc\,dd = \int_{\text{all}\,c} F_D(c) f_C(c)\,dc$, which upon solving the integration yields P$[D < C]$ = $\mu_C / (\mu_C + \mu_D)$.

7.14 Expected value of a function of random variable(s)

7.14.1 Function of one random variable

Regardless of whether g is one to one or not, the expectation is defined as:

$$E(Y) = E(g(X)] = \begin{cases} \int_{-\infty}^{\infty} g(x)\, f_X(x) dx \text{ for continuous } X \\ \sum_{all\,x} g(x)\, p_X(x) \text{ for discrete } X \end{cases} \tag{7.58}$$

To obtain an approximate mean of Y, expand $g(X)$ in Taylor series around μ_X.

$$E(Y) \approx g(\mu_X) + \frac{1}{2} g''(\mu_X)\sigma_X^2$$

7.14.2 Function of several random variables

The expected value of $Y = g(X_1, X_2, ...X_n)$ is:

$$E(Y) = E(g(X)] = \begin{cases} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\underline{x})\, f_{\underline{X}}(\underline{x}) d\underline{x} \text{ for continuous } \underline{X} \\ \sum_{all\,x_1} \cdots \sum_{all\,x_n} g(\underline{x})\, p_{\underline{X}}(\underline{x}) \text{ for discrete } \underline{X} \end{cases} \tag{7.59}$$

To obtain an approximate mean of Y expand g($\underline{X}$) in Taylor series around $\mu_{\underline{X}}$.

$$E(Y) \approx g(\underline{\mu}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \left.\frac{\partial^2 g}{\partial x_i \partial x_j}\right|_{\underline{\mu}} \sigma_{ij}$$ where $\underline{\mu}$ is the mean vector and σ_{ij} is the covariance.

7.15 Sum of several RVs

$\underline{X} = [X_1, X_2, ..., X_n]^T$ is a vector of *n* jointly distributed random variables with mean vector:

$$\underline{\mu}_X = \begin{Bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{Bmatrix} \tag{7.60}$$

covariance matrix:

$$V_X = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{n1} & \cdots & \cdots & \sigma_{nn} \end{bmatrix} \tag{7.61}$$

and joint density function $f_{\underline{X}}(\underline{x})$.

Let the random variable *Y* be a linear combination of $\underline{X}$.

$$Y = a_0 + \underline{a}\underline{X} \qquad \text{(7.62)sum7aaa}$$

where a_0 is a scalar and $\underline{a} = \{a_1, a_2, ..., a_n\}$ is a row vector multiplying the column $\underline{X}$. We wish to find the statistics of *Y*.

7.15.1 Mean and variance of the sum

Regardless of the distribution of $\underline{X}$, the mean of *Y* is,

$$\mu_Y = a_0 + \underline{a}\underline{\mu}_X \tag{7.63}$$

and the variance of *Y* is:

$$\sigma_Y^2 = \underline{a}\underline{V}_X\underline{a}^T \tag{7.64}$$

As we shall see in Section 7.17 below, if $\underline{X}$ is jointly normal , then the linear combination *Y* too is Normal.

7.15.2 Example: sum of two IID geometric RVs

The PMF of the Pascal RV can be derived by convolution using Eq (5.10). For r = 2, we have:

$$
\begin{aligned}
P[X_2 = n] &= P[G_1 + G_2 = n], \qquad P(G_i = m) = q^{m-1} p \\
&= \sum_{m=1}^{n-1} P[G_2 = n-m \mid G_1 = m] P[G_1 = m] \\
&= \sum_{m=1}^{n-1} P[G_2 = n-m] P[G_1 = m] \quad \text{since } G_2 \text{ and } G_1 \text{ are independent} \\
&= \sum_{m=1}^{n-1} q^{n-m-1} p \, q^{m-1} p \\
&= \sum_{m=1}^{n-1} q^{n-2} p^2 \\
&= (n-1) q^{n-2} p^2
\end{aligned}
$$

7.15.3 Example: sum of two IID uniforms

$X_1 \sim U(0,1), X_2 \sim U(0,1)$ and X_1, X_2 are independent of each other.
Find the distribution of their sum, $Y = X_1 + X_2$

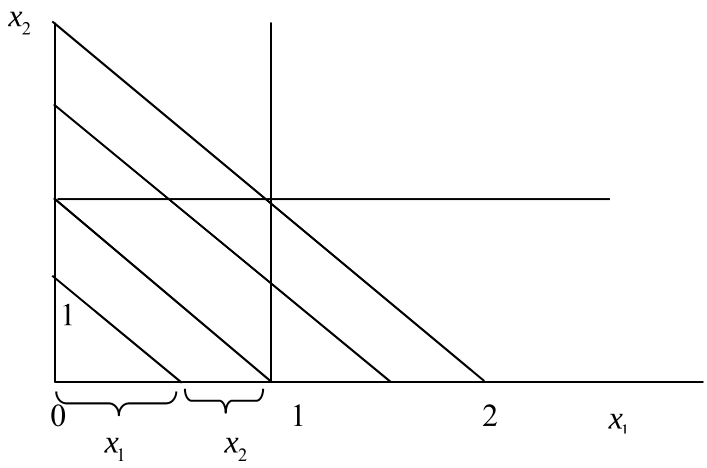


Figure 7-3: sum of two uniforms

The distribution of the sum can be conveniently written with the help of the indicator function as,

$$F_Y(y) = \int_{\text{all } x_1, x_2} I(x_1 + x_2 \le y) f_{X_1}(x_1) f_{X_2}(x_2) \, dx_1 \, dx_2$$

where the independence of X_1 and X_2 has been utilized. Differentiating, we obtain the PDF of Y which involves the delta function as follows:

$$f_Y(y) = \int_{\text{all } x_1, x_2} \delta(x_1 + x_2 = y) f_{X_1}(x_1) f_{X_2}(x_2)\, dx_1\, dx_2$$

$$= \int_{\text{all } x_1} f_{X_1}(x_1) f_{X_2}(y - x_1)\, dx_1$$

Recognizing that the PDFs of X1 and X2 are non-zero only in the interval [0,1] we simplify the PDF of *Y* as:

$$f_Y(y) = \int_{x_1=0}^{1} (1) I\left(0 < y - x_1 < 1\right) dx_1$$

For 0 < y < 1, we need to restrict 0< x1 < y, yielding,

$$f_Y(y) = \int_0^y (1)(1)\, dx_1 = y,\ 0 < y < 1$$

For 1 < y < 2, x_1 does not need any restriction, yielding,

$$f_Y(y) = \int_0^1 (1) I\left(y - 1 < x_1 < y\right) dx_1$$

$$= \int_{x_1 = y-1}^{1} (1)(1)\, dx_1$$

$$= 2 - y, \quad 1 < y < 2$$

Thus $f_y(y) = \begin{cases} y, & 0 < y < 1 \\ 2 - y, & 1 < y < 2 \end{cases}$

which is the triangular density function.

7.16 Several functions of several random variables

Papoulis p. 143, 183.

Let $\underline{X} = \{X_1 ....X_n\}$. Let *k* functions be defined on $\underline{X}$:

$$\begin{aligned} Y_1 &= g_1(\underline{X}) \\ &\vdots \\ Y_k &= g_k(\underline{X}) \end{aligned} \qquad (7.65)$$

f_Y

7.16.1 The joint density

If $k = n$, the joint density function of $\underline{Y}$ is:

$$f_{\underline{Y}}(y_1 .... y_n) = \frac{f_{\underline{X}}(x_1 .... x_n)}{|\, J(x_1 .... x_n)\,|} \qquad (7.66)$$

where the Jacobian of the transformation is:

$$J = \begin{vmatrix} \partial g_1 / \partial x_1 & \cdots & \partial g_1 / \partial x_n \\ \vdots & \ddots & \vdots \\ \partial g_n / \partial x_1 & & \partial g_n / \partial x_n \end{vmatrix}$$

If $k \neq n,$

When $k < n,$ choose $\begin{matrix} Y_{k+1} = X_{k+1} \\ Y_n = X_n \end{matrix}$

If $k > n.$

Express $Y_{n+1}.........Y_k$ in term of $Y_1 \ldots\ldots\ldots Y_n$

If there are several solutions to the problem, i.e., several $\underline{x}$ vectors $\{\underline{x}^{(1)}, \underline{x}^{(2)}, ...\}$ give rise to the same $\underline{y}$ vector:

Use the sum over all such solutions:

$$f_{\underline{Y}}(y_1........y_n) = \sum_i \frac{f_{\underline{X}}(x_1^{(i)},...,x_n^{(i)})}{|J(x_1^{(i)},...,x_n^{(i)})|} \qquad (7.67)$$

7.16.2 Linear combination

We genralize Eq sum7aaa to obtain *m* linear combinations $\underline{Y} = [Y_1, Y_2, ..., Y_m]^T$ of $\underline{X}$:

$$\underline{Y} = \begin{Bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{Bmatrix} = \underline{A}_0 + \underline{A}\underline{X} = \begin{Bmatrix} a_{01} \\ a_{02} \\ \vdots \\ a_{0m} \end{Bmatrix} + \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{Bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{Bmatrix} \qquad (7.68)$$

where $\underline{A}_0$ is an *m*-dimensional column vector and $\underline{A}$ is an $m \times n$ coefficient matrix. The mean and covariance matrix of Y are respectively,

$$\underline{\mu}_Y = \underline{A}_0 + \underline{A}\underline{\mu}_X \qquad (7.69)$$

$$\underline{V}_Y = \underline{A}\,\underline{V}_X\,\underline{A}^T \qquad (7.70)$$

We shall see next in Section 7.17 that if $\underline{X}$ is jointly normal , then $\underline{Y}$ too is jointly normal.

Consider the special case $m = n, \underline{\mu}_X = \{\underline{0}\}, \underline{V}_X = [I]$, that is, $\underline{X}$ is an IID standard normal n-vector. Then $\underline{Y}$ is an n-dimensional normal with $\underline{\mu}_Y = \underline{A}_0$, $\underline{V}_Y = \underline{A}\underline{A}^T$. This property is used for simulating correlated normals (Section 8.6.2).

7.17 Joint Normal distribution

7.17.1 Bivariate normal

Recall that X is said to have a normal distribution with mean μ and variance $\sigma^2 > 0$ if its density function is of the form:

$$N(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma^{1/2}} \exp[-\frac{1}{2} Q_1(x;\mu,\sigma^2)] \tag{7.71}$$

where, $Q_1(x;\mu,\sigma^2) = \frac{1}{\sigma^2}(x-\mu)^2 = (x-\mu)\sigma^{-2}(x-\mu)$. It is related to the standard normal form ϕ through: $N(x;\mu,\sigma^2) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$.

In a parallel manner, a two dimensional RV $\underline{X} = (X_1, X_2)^T$ is said to have a non-singular bivariate normal distribution if its density function is of the form:

$$N_2(\underline{x};\underline{\mu},V) = \frac{1}{2\pi|V|^{1/2}} \exp[-\frac{1}{2} Q_2(\underline{x};\underline{\mu},V)] \tag{7.72}$$

where,

$$Q_2(\underline{x};\underline{\mu},V) = (\underline{x}-\underline{\mu})^T V^{-1}(\underline{x}-\underline{\mu})$$

$$\underline{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

The correlation coefficient is $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}, |\rho| < 1$ since $|\sigma_{12}| < \sigma_1\sigma_2$. The bivariate N.D. is also denoted as $N_2(\underline{\mu}, V)$. Eq (7.72) can be expanded to:

$$\begin{aligned} &N_2(x_1,x_2;\mu_1,\mu_2;\sigma_1,\sigma_2;\rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ &\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \end{aligned} \tag{7.73}$$

Its standard form (zero means, unit variances) is: