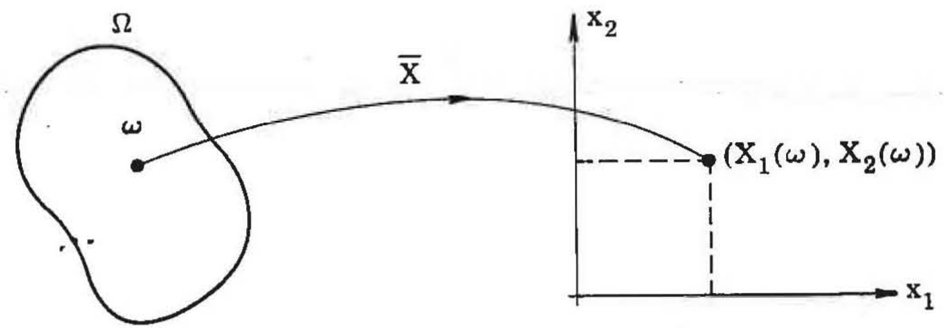
RANDOM VECTORS

P. Thoft-Christensen, M. J. Baker
Structural reliability theory and its applications, 1982

The concept of a random variable is basically used in a one-dimensional sense.

A random variable is a real-valued function $X : \Omega \rightarrow R$ mapping the sample space Ω into the real line R.



It can easily be extended to a vector-valued random variable $\bar{X}:\Omega \to R^n$ called a *random vector (random n-tuple)*,where $R^n = R \times R \times ... \times R$

An *n*-dimensional random vector $\bar{X}:\Omega \to R^n$ is an ordered set $\bar{X} = (X_1, X_2, ..., X_n)$ of one-dimensional random variables $X_i:\Omega \to R, i = 1, ..., n$.

All $X_1, X_2, ..., X_n$ are defined on the same sample space Ω.

Let X_1 and X_2 be two random variables. The range of the random vector $\bar{X} = (X_1, X_2)$ is then a subset of R^2 as shown in figure.

The range of an *n*-dimensional random vector is a subset of R^n.

Consider again two random variables X_1 and X_2 and their corresponding distribution functions F_{X_1} and F_{X_2}.

The latter give no information on the joint behaviour of X_1 and X_2.

Thus *joint probability distribution function*

$F_{X_1,X_2}: R^2 \to R$ is defined:

$$F_{X_1,X_2}(x_1,x_2) = P\left(\left(X_1 \le x_1\right) \cap \left(X_2 \le x_2\right)\right)$$

we use $F_{\bar{X}}$ for F_{X_1,X_2}, where $\bar{X} = (X_1, X_2)$

The definition can be generalized to the *n*-dimensional case

$$F_{\bar{X}}(\bar{x}) = P\left(\bigcap_{i=1}^{n} \left(X_i \le x_i\right)\right)$$

where $\bar{X} = (X_1,\ldots,X_n)$ and $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_n)$

Discrete or continuous random vectors exist, the latter of our concern only.

Our analysis is restricted to two-dimensional random vectors only, to be generalized easily.

The *joint probability density function* for the random vector $\bar{X} = (X_1, X_2)$ is given

$$f_{\bar{X}}(\bar{x}) = \frac{\partial^2}{\partial x_1 \partial x_2} F_{\bar{X}}(\bar{x})$$

The inverse formula is

$$F_{\bar{X}}(\bar{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_1} f_{\bar{X}}(x_1', x_2') \, dx_1' dx_2'$$

The following functions exist

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{\bar{X}}(x_1, x_2) \, dx_2$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{\bar{X}}(x_1, x_2) \, dx_1$$

They are *marginal density functions* of a random vector X – one-dimensional functions.

Example

...

Let a 2-dimensional discrete random vector $\bar{X} = (X_1, X_2)$ be defined on Ω by

P(4, 3) = 0.1

P(4, 4) = 0.1

P(5, 3) = 0.3

P(5, 4) = 0.2

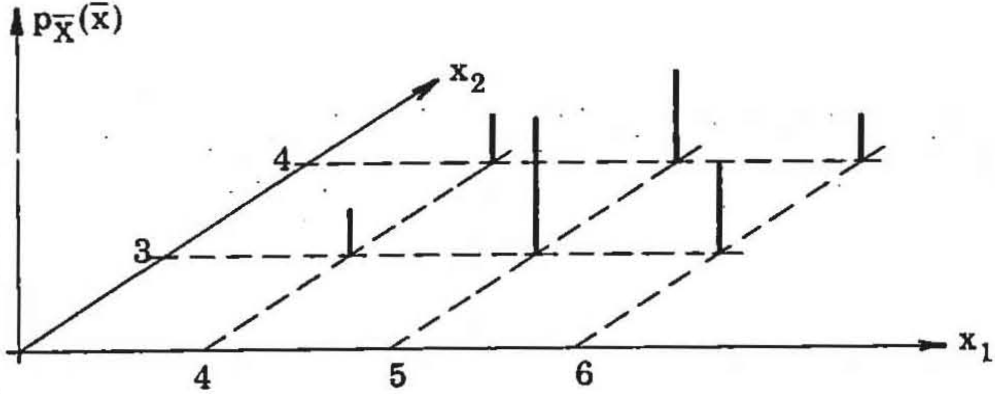
P(6, 3) = 0.2

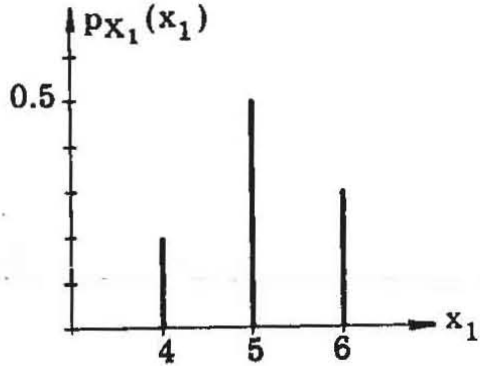
P(6, 4) = 0.1

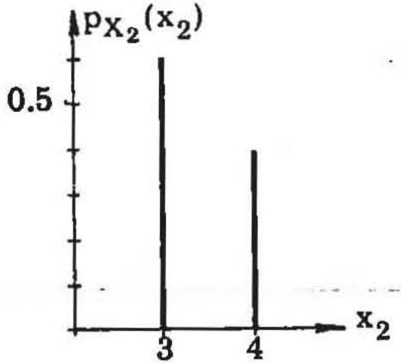
The joint mass function $p_{\bar{X}}$, and the marginal mass functions p_{X_1} and p_{X_2} are illustrated in Figures below.

Note that

$$p_{\bar{X}}(x_1, x_2) \neq p_{X_1}(x_1)\, p_{X_2}(x_2)$$


$p_{\overline{X}}(\overline{x})$
x_2
4
3
x_1
4
5
6


$p_{X_1}(x_1)$
0.5
x_1
4
5
6


$p_{X_2}(x_2)$
0.5
x_2
3
4

CONDITIONAL DISTRIBUTIONS

Probability of occurrence of event E_1 conditional upon the occurrence of event E_2 was defined by

$$P\left(E_1 \middle| E_2\right) = \frac{P\left(E_1 \cap E_2\right)}{P\left(E_2\right)}$$

The *conditional probability mass function* for two jointly distributed discrete random variables X_1 and X_2 is defined

$$p_{X_1|X_2}\left(x_1 \middle| x_2\right) = \frac{p_{X_1,X_2}\left(x_1, x_2\right)}{p_{X_2}\left(x_2\right)} = \frac{\text{joint density}}{\text{marginal density}}$$

Continuous cases define the *conditional probability density function*

$$f_{X_1|X_2}\left(x_1 \middle| x_2\right) = \frac{f_{X_1,X_2}\left(x_1, x_2\right)}{f_{X_2}\left(x_2\right)}$$

where $f_{X_2}\left(x_2\right) > 0$ and where f_{X_2} is a marginal PDF.

...

Mind the discrete │ continuous diversity: $p_{X_1|X_2}$ is a conditional mass function, $f_{X_1|X_2}$ a conditional density function.

Two random variables X_1 and X_2 are *independent* if

$$f_{X_1|X_2}\left(x_1|x_2\right)= f_{X_1}\left(x_1\right)$$

Chich implies

$$f_{X_1|X_2}\left(x_1|x_2\right)= f_{X_1}\left(x_1\right) f_{X_2}\left(x_2\right)$$

Integrating with respect to x_1 gives *conditional distribution function*

$$F_{X_1}\left(x_1\right)=\int_{-\infty}^{\infty} F_{X_1|X_2}\left(x_1|x_2\right) f_{x_2}\left(x_2\right) dx_2$$

similarly the x_2 case.

Example

Consider two jointly distributed discrete random variables X_1 and X_2 again.

Note that

$$p_{X_1,X_2}(5,3) = p_{X_1}(5)\,p_{X_2}(3)$$

but for example

$$p_{X_1,X_2}(6,4) \neq p_{X_1}(6)\,p_{X_2}(4)$$

Therefore, X_1 and X_2 are dependent.

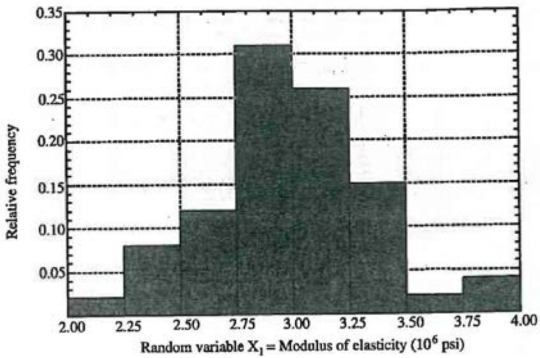
EXAMPLE

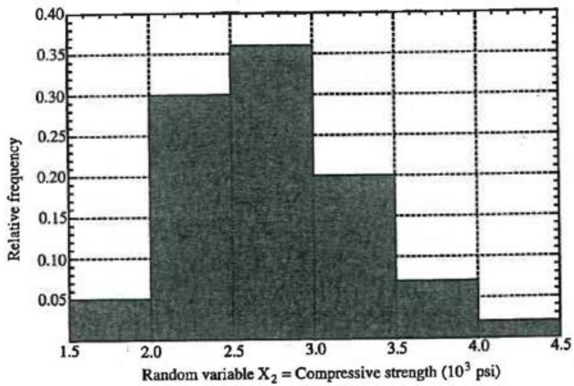
Consider a set of tests in which two quantities are measured: modulus of elasticity, X_1, and compressive strength, X_2.
Since the values of these variables vary from test to test, as seen in Table, it is appropriate to treat them as random variables.

TABLE
Values of modulus of elasticity and compressive strength

Sample number	f'_c, psi	E, psi	Sample number	f'_c, psi	E, psi	Sample number	f'_c, psi	E, psi
1	3,059	3,335,000	34	2,266	2,797,000	67	2,634	3,216,000
2	3,397	3,280,000	35	3,414	3,087,000	68	2,309	2,725,000
3	2,575	3,117,000	36	2,973	3,122,000	69	3,062	3,317,000
4	3,803	3,271,000	37	2,397	2,933,000	70	2,336	2,995,000
5	2,887	3,201,000	38	3,629	3,522,000	71	2,325	2,512,000
6	2,774	3,067,000	39	2,797	3,042,000	72	2,600	2,840,000
7	3,187	3,252,000	40	3,164	2,890,000	73	2,197	2,636,000
8	2,804	2,814,000	41	2,063	2,421,000	74	3,635	3,304,000
9	1,563	2,354,000	42	2,521	3,074,000	75	1,938	2,483,000
10	2,258	2,606,000	43	2,643	2,962,000	76	2,557	2,618,000
11	2,753	3,233,000	44	4,072	3,814,000	77	3,566	3,990,000
12	2,156	2,854,000	45	2,249	2,920,000	78	2,432	3,112,000
13	2,752	3,020,000	46	3,107	3,485,000	79	2,903	3,408,000
14	2,933	3,080,000	47	3,009	2,942,000	80	2,776	2,963,000
15	2,821	3,455,000	48	2,452	2,901,000	81	3,239	3,497,000
16	2,209	2,464,000	49	2,361	2,917,000	82	2,393	2,960,000
17	2,774	2,853,000	50	2,780	3,010,000	83	3,459	3,545,000
18	2,391	2,685,000	51	3,113	3,454,000	84	2,423	3,097,000
19	3,251	2,931,000	52	3,071	3,182,000	85	2,330	2,697,000
20	2,933	2,841,000	53	2,577	2,962,000	86	3,199	3,318,000
21	3,049	3,034,000	54	2,421	2,803,000	87	3,101	3,188,000
22	2,079	2,473,000	55	1,878	2,534,000	88	2,509	2,516,000
23	3,615	3,895,000	56	3,470	3,377,000	89	3,306	2,823,000
24	2,724	2,937,000	57	2,977	3,342,000	90	2,402	2,935,000
25	2,690	2,999,000	58	2,140	2,635,000	91	2,524	2,856,000
26	2,722	2,880,000	59	2,087	2,208,000	92	2,318	2,214,000
27	2,170	2,985,000	60	2,551	2,810,000	93	2,884	3,089,000
28	2,509	2,790,000	61	4,025	3,977,000	94	2,803	3,014,000
29	2,172	2,663,000	62	2,303	2,362,000	95	2,983	3,308,000
30	3,450	3,236,000	63	1,650	2,335,000	96	2,877	2,965,000
31	3,729	3,201,000	64	2,683	2,823,000	97	2,192	2,553,000
32	1,807	2,344,000	65	3,280	3,214,000	98	2,631	3,179,000
33	2,438	3,144,000	66	3,801	3,060,000	99	2,456	2,904,000
						100	2,725	3,150,000

Using the concept of histograms, we can get an idea of the general shape of the probability density function (PDF) for each individual variable and the joint probability density function and joint probability distribution function.
For each individual variable, we define appropriate intervals of values and then count the number of observations within each interval.
The resulting relative frequency histogram for each variable is shown in Figure





To consider the joint histogram, we need to define "two-dimensional intervals”.

For example, one "interval" would be for values of $X_1(E)$ between 3.0×10^6 psi and 3.25×10^6 psi and values of $X_2(f_c')$ between 2.5×10^6 psi and 3.0×10^3 psi.

Looking at Table, we see that there are 15 samples that satisfy both requirements simultaneously; these samples are highlighted in the table.

Therefore, we have 15 observations in this interval out of 100 total observations, and the relative frequency value is 15/100 = 0.15.

This value is indicated as the shaded block in Figure, the relative frequency histogram.

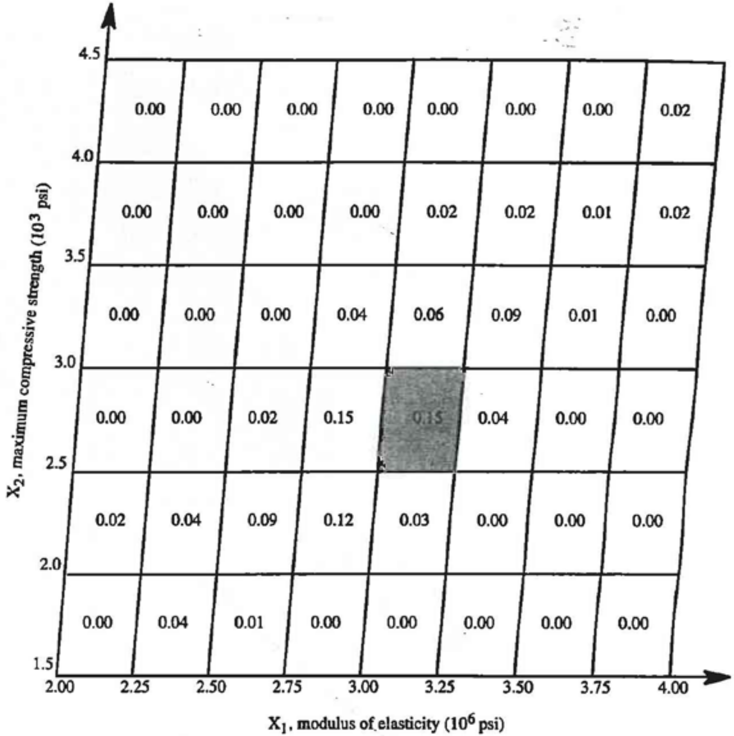
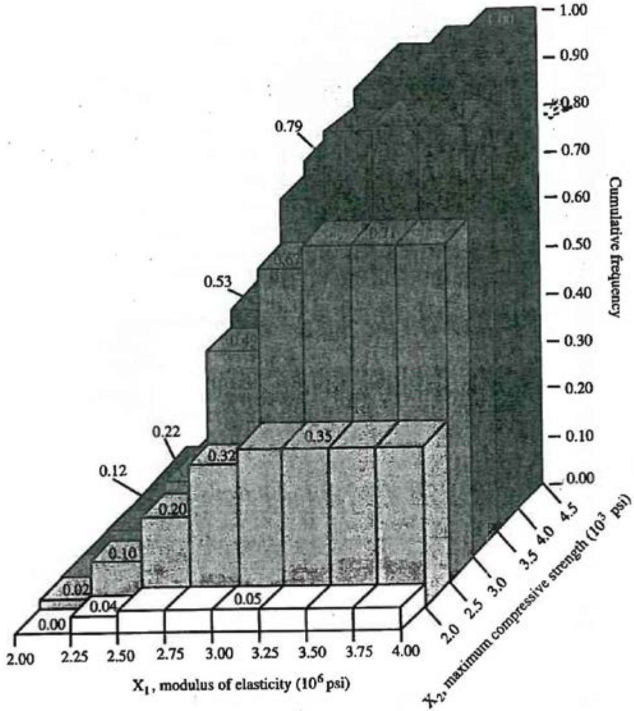


FIGURE. Relative frequency histogram for both X_1 and X_2

A cumulative frequency histogram can also be constructed as shown in Figure.



For example, to find the cumulative value of the number of times that X_1 is less than or equal to 3.0×10^6 psi and X_2 is less than or equal to 2.35×10^3 psi, we add all the relative frequency values in

Figure that satisfy this requirement.

The result would be

0 + 0.04 + 0.01 + 0 + 0.02 + 0.04 + 0.09 + 0.12 = 0.32.

Functions of random variables

A continuous random variable Y which is a function $f\left(X\right)$ of a continuous random variable X is defined, the density function f_Y may determined given the density function f_X as follows

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

where $x=f^{-1}\left(y\right)$

Expanding the problem we have a random vector $\bar{Y}=\left(Y_1,Y_1,\ldots,Y_n\right)$ - function $\bar{f}=\left(f_1,f_1,\ldots,f_n\right)$ of a random vector $\bar{X}=\left(X_1,X_2,\ldots,X_n\right)$, that is $Y_i=f_i\left(X_1,\ldots,X_n\right)$, where i = 1, 2, ... ,n.

Each function f_i i = 1, 2, ... , n is a one-to-one mapping, so inverse relations exist:

$$X_i=g_i\left(Y_1,\ldots,Y_n\right)$$

It can then be shown that $f_{\overline{Y}}\left(\overline{y}\right)=f_{\overline{X}}\left(\overline{x}\right)\left|J\right|$

where $\overline{x}=\left(x_1,x_2,\ldots,x_n\right)$ and $\overline{y}=\left(y_1,y_2,\ldots,y_n\right)$

$$J=\begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \ldots & \dfrac{\partial x_1}{\partial y_n} \\ \ldots & \ldots & \ldots \\ \dfrac{\partial x_n}{\partial y_1} & \ldots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}$$

is the Jacobian determinant.

Let the random variable, Y be a function f of the random vector $\overline{X}=\left(X_1,X_2,\ldots,X_n\right)$

It can be shown that

$$E\left(Y\right)=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}f\left(\overline{x}\right)f_{\overline{X}}\left(\overline{x}\right)dx_1\ldots dx_n$$

where $\overline{x}=\left(x_1,x_2,\ldots,x_n\right)$ and $f_{\overline{X}}\left(\overline{x}\right)$ is the probability density function for the random vector $\overline{X}$.

CORRELATION

Basic Definitions

Let X_1 and X_2 be two random variables with means μ_{X_1} and μ_{X_2} and standard deviations σ_{X_1} and σ_{X_2}.

The *covariance* of X_1 and X_2 is defined as

$$\mathrm{Cov}\left[X_1, X_2\right] = E\left[\left(X_1 - \mu_{X_1}\right)\left(X_2 - \mu_{X_2}\right)\right] =$$

$$= E\left[X_1 X_2 - X_1\mu_{X_1} - X_2\mu_{X_2} + \mu_{X_1}\mu_{X_2}\right]$$

where $E[\;]$ denotes expected value.

Note that $\mathrm{Cov}\left[X_1, X_2\right] = \mathrm{Cov}\left[X_2, X_1\right]$.

If X and Y are continuous random variables then this formula becomes

$$\mathrm{CoV}\left(X_1, X_2\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x_1 - \mu_{X_1}\right)\left(x_2 - \mu_{X_2}\right) f_{XY}\left(x, y\right) dxdy$$

The *coefficient of correlation* (also called the correlation coefficient) between two random variables X_1 and X_2 is defined as

$$\rho_{X_1X_2} = \frac{\mathrm{Cov}\left[X_1, X_2\right]}{\sigma_{X_1}\sigma_{X_2}}$$

It can be proven that the coefficient of correlation is limited to values between –1 and 1 inclusive, that is

$$-1 \le \rho_{X_1X_2} \le 1$$

The value of $\rho_{X_1X_2}$ indicates the degree of *linear* dependence between the two random variables *X* and *Y*.

If $\rho_{X_1X_2}$ is close to 1, then *X* and *Y* are linearly correlated.

If $\rho_{X_1X_2}$ is close to zero, then the two variables are not *linearly* related to each other.

Note the emphasis on the word "linearly."

Two random variables X_1 and X_2 are *uncorrelated* if $\rho_{X_1X_2} = 0$.

The following identity

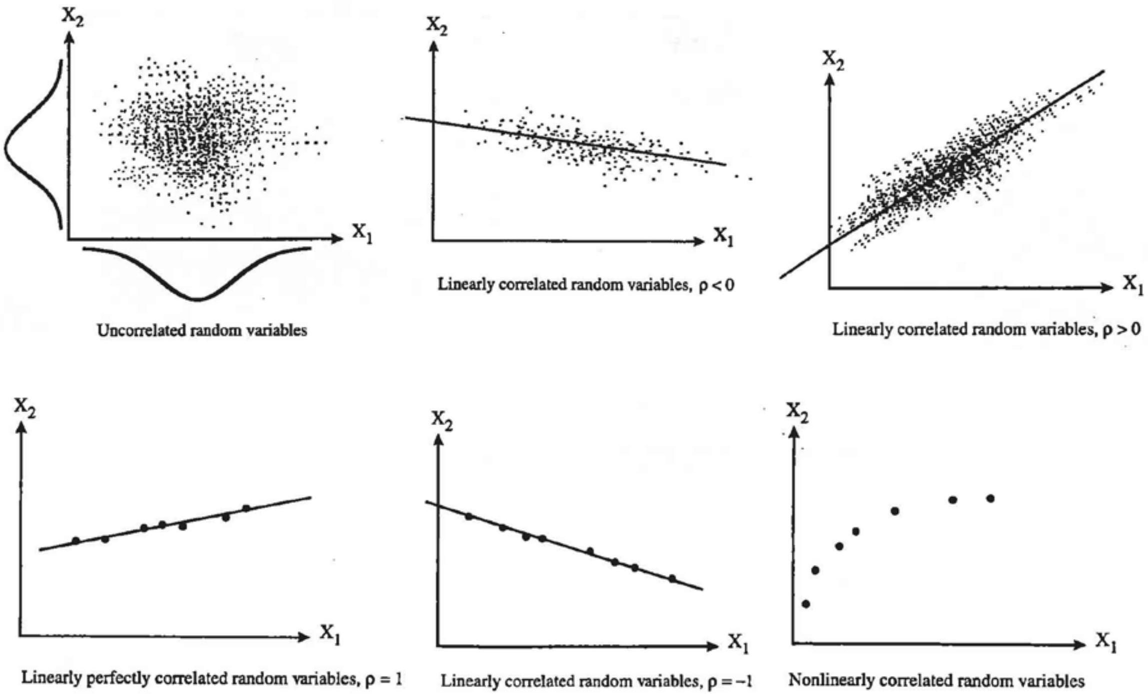
$$\text{Cov}\left[X_1, X_2\right] = E\left[\left(X_1 - \mu_{X_1}\right)\left(X_2 - \mu_{X_2}\right)\right] = E\left[X_1 \cdot X_2\right] - E\left[X_1\right]E\left[X_2\right]$$

is specified in the case of uncorrelated random variables X_1 and X_2

$$E\left[X_1 \cdot X_2\right] = E\left[X_1\right]E\left[X_2\right]$$

When $\rho_{X_1X_2}$ is close to zero, it does not mean that there is no dependence at all; there may be some nonlinear relationship between the two variables.

Figure 2.36 illustrates the concept of correlation.



Examples of correlated and uncorrelated random variables.

It is interesting to note what happens when two variables are uncorrelated (i.e., $\rho_{X_1X_2} = 0$).

This implies that the covariance is equal to zero.

When

$$\mathrm{CoV}\left[X_1, X_2\right]=0$$

$$E\left(X_1 X_2\right) = \mu_{X_1}\mu_{X_2}$$

the expected value of the product X_1X_2 is the product of the expected values.

It is important to emphasize that the terms "statistically independent" and "uncorrelated" are not always synonymous.

Statistically independent is a much stronger statement than uncorrelated.

If two variables are statistically independent, then they must also be uncorrelated.

...

However, the converse is not, in general, true.
If two variables are uncorrelated, they are not necessarily statistically independent.
The foregoing comments on correlation pertain to two random variables.

When dealing with a random vector, a *covariance matrix* is used to describe the correlation between all possible pairs of the random variables in the vector.
For a random vector with *n* random variables, the covariance matrix, [C], is defined as

$$[C] = \begin{bmatrix} \mathrm{Cov}[X_1, X_1] & \mathrm{Cov}[X_1, X_2] & \ldots & \mathrm{Cov}[X_1, X_n] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Cov}[X_2, X_2] & \ldots & \mathrm{Cov}[X_2, X_n] \\ \ldots & \ldots & \ldots & \ldots \\ \mathrm{Cov}[X_n, X_1] & \mathrm{Cov}[X_n, X_2] & \ldots & \mathrm{Cov}[X_n, X_n] \end{bmatrix}$$

Note that $\text{Cov}[X_i, X_i] = \text{Var}[X_i]$.

In some cases, it is more convenient to work with a matrix of coefficients of correlation [*p*] defined as

$$[p] = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}$$

Note two things about the matrices [C] and [p].

First, they are symmetric matrices.

Second, the terms on the main diagonal of the [C] matrix can be simplified using the fact that $\text{Cov}(X_i, X_i) = \text{Var}(X_i) = \sigma_{X_i}^2$.

The diagonal terms in [*p*] are equal to 1.

If all *n* random variables are *uncorrelated*, then the off-diagonal terms are equal to zero and the covariance matrix becomes a diagonal matrix of the form

$$[C]=\begin{bmatrix} \sigma_{X_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{X_2}^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_{X_n}^2 \end{bmatrix}$$

The matrix [p] becomes a diagonal matrix with 1's on the diagonal

Statistical Estimate of the Correlation Coefficient

In practice we often do not know the underlying distributions of the variables we are observing, and thus we have to rely on test data and observations to estimate parameters.
When we have observed data for two random variables *X* and *Y*, we can estimate the correlation coefficient as follows.
Assume that there are *n* observations $\{x_1, x_2, \ldots, x_n\}$ of variable *X* and n observations $\{y_1, y_2, \ldots, y_n\}$ of variable *Y*.
The correlation coefficient can be calculated using

$$\hat{\rho} = \frac{1}{n-1} \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{s_X s_Y} = \frac{1}{n-1} \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}}{s_X s_Y}$$

where $\overline{x}$ and $\overline{y}$ are sample means and s_X and s_y sample standard deviations

...

MULTIVARIATE DISTRIBUTIONS

Animportant joint density function of two continuous random variables X_1 and X_2 is the *bivariate normal density function*

$$f_{X_1,X_2}(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 -2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right]$$

where $-\infty\le x_1\le\infty$, $-\infty\le x_2\le\infty$, and μ_1, μ_2 are the means σ_1, σ_2 the standard deviations and ρ the coefficient of X_1, X_2.

The *multivariate normal density function* is defined

$$f_{\bar{X}}(\bar{x})=\frac{1}{(2\pi)^{n/2}}\frac{1}{C^{1/2}}\exp\left[-\frac{1}{2}\sum_{i,j=1}^{n}(x_i-\mu_j)M_{ij}(x_j-\mu_i)\right]$$

$\bar{x}=(x_1,x_2,\dots,x_n)$, $M=C^{-1}$, and where C is the covariance matrix.