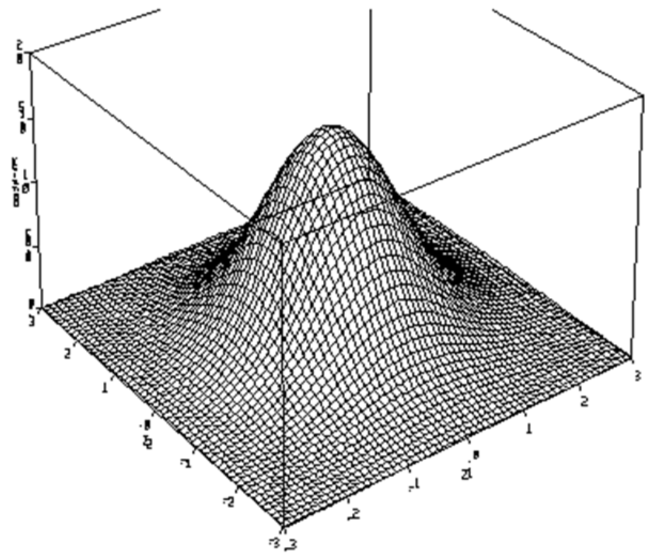
## Joint, Marginal and Conditional Distributions

Often, we want to make probability statements about more than one random variable at a time.

Probability functions defined over a pair of random variables are known as **joint distributions**.



If $X$ and $Y$ are discrete random variables then the probability that $X = x$ and $Y = y$ is given by the **joint distribution** Pr($X$=$x$, $Y$=$y$).

*Rolling a pair of dice:*

Consider rolling a pair of dice. There are 36 possible outcomes: (1,1), (1,2), ..., (6,6). Each outcome is equally likely with probability 1/36. Define two random variables:

$X$ = the number of 3's in a typical outcome
$Y$ = the number of 5's in a typical outcome

Joint distribution of X and Y when a pair of dice is rolled:

| | | $X$ = 0 | $X$ = 1 | $X$ = 2 |
|---|---|---|---|---|
| $Y$ | 0 | 16/36 | 8/36 | 1/36 |
| | 1 | 8/36 | 2/36 | 0 |
| | 2 | 1/36 | 0 | 0 |

Given a joint distribution function, we can obtain the probability distributions of the individual random variables.

If $X$ and $Y$ are discrete random variables then

$$\Pr(X = x) = \sum_{i=1}^{n} \Pr(X = x, Y = y_i)$$

is the **marginal distribution** of $X$.

*Rolling a pair of dice:*

| | | X: 0 | X: 1 | X: 2 | $f(y)$ |
|---|---|---|---|---|---|
| Y | 0 | 16/36 | 8/36 | 1/36 | |
| | 1 | 8/36 | 2/36 | 0 | |
| | 2 | 1/36 | 0 | 0 | |
| | $f(x)$ | | | | |

In addition to knowing the probability of the joint occurrence of the two random variables and their individual probability distributions, we might also be interested in knowing the probability of a particular random variable given that another random variable has already occurred.

For example, we might want to know the probability that the price of a house is $500,000 given that it's 2,000 square feet.

The **conditional distribution** of $Y$ given $X$ is given by

$$\Pr(Y = y|X = x) = \frac{\Pr(X = x, Y = y)}{\Pr(X = x)}$$

*Rolling a pair of dice:*

| | | $Y$ = 0 | $Y$ = 1 | $Y$ = 2 |
|---|---|---|---|---|
| Given $X$ | 0 | 0.64 | 0.32 | 0.04 |
| | 1 | 0.80 | 0.20 | 0 |
| | 2 | 1.00 | 0 | 0 |

## Conditional Expectation and Conditional Variance

The expected value of $Y$ given $X$ is known as the **conditional expectation** of $Y$ given $X$.

It is the mean of $Y$ using the conditional distribution.

If $X$ and $Y$ are discrete random variables then

$$E(Y|X = x) = \sum_{i=1}^{n} y_i \Pr(Y = y_i|X = x).$$

*Rolling a pair of dice:*

The concept of conditional expectation is easily extended to variances to obtain the **conditional variance**.

If $X$ and $Y$ are discrete random variables then

$$Var(Y|X = x) = \sum_{i=1}^{n}[y_i - E(Y|X = x)]^2 \Pr(Y = y_i|X = x).$$

*Rolling a pair of dice:*

In some cases the two random variables are not related, that is, they are **independent** (**statistically independent** or **independently distributed**).

If $X$ and $Y$ are **independently distributed** then

$$\Pr(Y = y, X = x) = \Pr(Y = y)\Pr(X = x),$$

the joint distribution equals the product of the marginal distributions.

Notice, this also means $\Pr(Y = y|X = x) = \Pr(Y = y)$.

A useful property of conditional expectations is known as the **law of iterated expectations** (LIE).

The unconditional expectation of $Y$ if $X$ takes on one of $k$ values is

$$E(Y) = \sum_{i=1}^{k} E(Y|X = x_i)\Pr(X = x_i) = E[E(Y|X)].$$

When we are dealing with two random variables, one of the main items of interest is how closely they are associated. The concepts of **covariance** and **correlation** are two ways to measure closeness of two random variables.

## Covariance and Correlation

The covariance between $X$ and $Y$ is $Cov(X,Y) = \sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$.

If $X$ and $Y$ are discrete random variables taking $k$ and $n$ values respectively,

$$Cov(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{k}\left(x_j-\mu_X\right)(y_i-\mu_Y)\Pr(X=x_j, Y=y_i).$$

It follows $Cov(X,X) = Var(X)$.

*Rolling a pair of dice:*

Although covariance measure is useful in identifying the nature of the association between $X$ and $Y$ its units in which it is measure are difficult to interpret.

To avoid this problem, a normalized covariance measure is used.

The **correlation coefficient** between $X$ and $Y$ is

$$Corr(X,Y) = \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

It can be shown that the correlation coefficient lies between -1 and 1.

Notice if $X$ and $Y$ are independent random variables then they are uncorrelated. The converse is not necessarily true (i.e. zero correlation need not imply independence).

**Some properties for two random variables:**

For any random variables X and Y and constants a, b, and c

- $E(a) = a$
- $E(aX + bY) = a\mu_X + b\mu_Y$
- $Var(X) = E(X^2) - \mu_X^2$
- $Var(aX + bY + c) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}$
- $Var(X + Y) = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$
- $Cov(X, Y) = E(XY) - \mu_X\mu_Y$
- $Cov(a + bX + cV, Y) = b\sigma_{XY} + c\sigma_{VY}$