Joint Probability Distributions

In many experiments, two or more random variables have values that are determined by the outcome of the experiment.

For example, the binomial experiment is a sequence of trials, each of which results in success or failure.

If

\[ X_i = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ trial is a success} \\ 0 & \text{otherwise,} \end{cases} \]

then \( X_1, X_2, \ldots, X_n \) are all random variables defined on the whole experiment.
To calculate probabilities involving two random variables $X$ and $Y$ such as

$$P(X > 0 \text{ and } Y \leq 0),$$

we need the *joint* distribution of $X$ and $Y$.

The way we represent the joint distribution depends on whether the random variables are discrete or continuous.
Two Discrete Random Variables

If $X$ and $Y$ are discrete, with ranges $R_X$ and $R_Y$, respectively, the joint probability mass function is

$$p(x, y) = P(X = x \text{ and } Y = y), x \in R_X, y \in R_Y.$$ 

Then a probability like $P(X > 0 \text{ and } Y \leq 0)$ is just

$$\sum_{x \in R_X: x > 0} \sum_{y \in R_Y: y \leq 0} p(x, y).$$
Marginal Distribution

To find the probability of an event defined only by $X$, we need the *marginal* pmf of $X$:

$$p_X(x) = P(X = x) = \sum_{y \in R_Y} p(x, y), x \in R_X.$$  

Similarly the marginal pmf of $Y$ is

$$p_Y(y) = P(Y = y) = \sum_{x \in R_X} p(x, y), y \in R_Y.$$
Two Continuous Random Variables

If $X$ and $Y$ are continuous, the joint probability density function is a function $f(x, y)$ that produces probabilities:

$$P[(X, Y) \in A] = \int \int_A f(x, y) \, dy \, dx.$$ 

Then a probability like $P(X > 0 \text{ and } Y \leq 0)$ is just

$$\int_0^\infty \int_{-\infty}^0 f(x, y) \, dy \, dx.$$
Marginal Distribution

To find the probability of an event defined only by $X$, we need the *marginal* pdf of $X$:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, dy, \quad -\infty < x < \infty.$$ 

Similarly the marginal pdf of $Y$ is

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx, \quad -\infty < y < \infty.$$
Independent Random Variables

Independent Events

Recall that events $A$ and $B$ are independent if

$$P(A \text{ and } B) = P(A)P(B).$$

Also, events may be defined by random variables, such as $A = \{X \geq 0\} = \{s \in S : X(s) \geq 0\}$.

We say that random variables $X$ and $Y$ are independent if any event defined by $X$ is independent of every event defined by $Y$. 
Independent Discrete Random Variables

Two \textit{discrete} random variables are independent if their joint pmf satisfies

\[ p(x, y) = p_X(x)p_Y(y), \quad x \in R_X, \quad y \in R_Y. \]

Independent Continuous Random Variables

Two \textit{continuous} random variables are independent if their joint pdf satisfies

\[ f(x, y) = f_X(x)f_Y(y), \quad -\infty < x < \infty, \quad -\infty < y < \infty. \]

Random variables that are not independent are said to be \textit{dependent}. 
More Than Two Random Variables

Suppose that random variables $X_1, X_2, \ldots, X_n$ are defined for some experiment.

If they are all discrete, they have a *joint pmf*:

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n).$$

If they are all continuous, they have a *joint pdf*:

$$P(a_1 < X_1 \leq b_1, \ldots, a_n < X_n \leq b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \, dx_n \ldots \, dx_1.$$
Full Independence

The random variables $X_1, X_2, \ldots, X_n$ are independent if their joint pmf or pdf is the product of the marginal pmfs or pdfs.

Pairwise Independence

Note that if $X_1, X_2, \ldots, X_n$ are independent, then every pair $X_i$ and $X_j$ are also independent.

The converse is not true: pairwise independence does not, in general, imply full independence.
Conditional Distribution
If $X$ and $Y$ are discrete random variables, then

$$P(Y = y | X = x) = \frac{P(X = x \text{ and } Y = y)}{P(X = x)} = \frac{p(x, y)}{p_X(x)}.$$  

We write this as $p_{Y|X}(y|x)$.

If $X$ and $Y$ are continuous random variables, we still need to define the distribution of $Y$ given $X = x$.

But $P(X = x) = 0$, so the definition is not obvious; however,

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

may be shown to have the appropriate properties.
Expected Value, Covariance, and Correlation

As you might expect, for a function of several discrete random variables:

$$E[h(x_1, \ldots, x_n)] = \sum_{x_1} \cdots \sum_{x_n} h(x_1, \ldots, x_n) p(x_1, \ldots, x_n)$$

For a function of several continuous random variables:

$$E[h(x_1, \ldots, x_n)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \ldots, x_n) f(x_1, \ldots, x_n) \, dx_n \cdots dx_1.$$
Covariance

Recall the variance of $X$:

$$V(X) = E \left[ (X - \mu_X)^2 \right].$$

The covariance of two random variables $X$ and $Y$ is

$$Cov(X, Y) = E \left[ (X - \mu_X)(Y - \mu_Y) \right].$$

Note: the covariance of $X$ with itself is its variance.
Correlation

Just as the units of $V(X)$ are the square of the units of $X$, so the units of $\text{Cov}(X, Y)$ are the product of the units of $X$ and $Y$.

The corresponding dimensionless quantity is the correlation:

$$\text{Corr}(X, Y) = \rho_{X, Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X) V(Y)}}.$$
Properties of Correlation

If \( ac > 0 \),
\[
\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y).
\]

For any \( X \) and \( Y \),
\[
-1 \leq \text{Corr}(X, Y) \leq 1.
\]

If \( X \) and \( Y \) are independent, then \( \text{Corr}(X, Y) = 0 \), but not conversely. That is, \( \text{Corr}(X, Y) = 0 \) does not in general mean that \( X \) and \( Y \) are independent.

If \( \text{Corr}(X, Y) = \pm 1 \), then \( X \) and \( Y \) are exactly linearly related:
\[
Y = aX + b \text{ for some } a \neq 0.
\]