<u>Discrete Random Variables</u>

Let X be a *discrete random variable* with outcomes, $x_1, x_2, \ldots, x_n$. The probability that the outcome of experiment X is x_i is $P(X = x_i)$ or $p_X(x_i)$:

- $\forall_i \;\; p_X(x_i) \geq 0$
- $\sum_{i=1}^{n} p_X(x_i) = 1$

p_X is termed the *probability mass function*.

Joint Discrete Random Variables

Let Y be a discrete random variable with outcomes, $y_1, y_2, \ldots, y_m$. The probability that the outcome of experiment X is x_i and the outcome of experiment Y is y_j is the *joint probability*, $P(X = x_i, Y = y_j)$ or $p_{XY}(x_i, y_j)$:

- $\forall_{i,j} \;\; p_{XY}(x_i, y_j) \geq 0$
- $\sum_{i=1}^{n} \sum_{j=1}^{m} p_{XY}(x_i, y_j) = 1$

p_{XY} is termed the *joint probability mass function*.

Marginal Probabilities

It is possible to recover the *marginal* p.m.f., p_X (or p_Y), from the joint p.m.f., p_{XY}, by summing across its rows (or columns):

$$
\begin{aligned}
p_X(x_i) &= \sum_{j=1}^{m} p_{XY}(x_i, y_j) \\
p_Y(y_j) &= \sum_{i=1}^{n} p_{XY}(x_i, y_j).
\end{aligned}
$$

PXY

X \ Y	red	green	PX
ripe	0.36	0.06	0.42
unripe	0.04	0.54	0.58
PY	0.4	0.6	

PX|Y

X \ Y	red	green
ripe	0.9	0.1
unripe	0.1	0.9

PY|X

Y \ X	ripe	unripe
red	0.857	0.07
green	0.143	0.93

Figure 1: Joint and conditional distributions (dependent random variables).

Conditional Probabilities

$$
\begin{aligned}
p_{X|Y}(x_i \,|\, y_j) &= \frac{p_{XY}(x_i, y_j)}{p_Y(y_j)} \\
&= \frac{p_{XY}(x_i, y_j)}{\sum_{k=1}^{n} p_{XY}(x_k, y_j)}
\end{aligned}
$$

$$
p_{XY}(x_i, y_j) = p_{X|Y}(x_i \,|\, y_j) p_Y(y_j)
$$

$$
\begin{aligned}
p_{Y|X}(y_j \,|\, x_i) &= \frac{p_{XY}(x_i, y_j)}{p_X(x_i)} \\
&= \frac{p_{XY}(x_i, y_j)}{\sum_{k=1}^{m} p_{XY}(x_i, y_k)}
\end{aligned}
$$

$$
p_{XY}(x_i, y_j) = p_{Y|X}(y_j \,|\, x_i) p_X(x_i)
$$

Bayes' Rule

Sometimes we know $p_{Y|X}(y_j, x_i)$ and want to compute $p_{X|Y}(x_i, y_j)$. *Bayes' rule* allows us to do this:

$$\begin{aligned} p_{X|Y}(x_i|y_j) &= \frac{p_{XY}(x_i, y_j)}{p_Y(y_j)} \\ &= \frac{p_{Y|X}(y_j|x_i)p_X(x_i)}{p_Y(y_j)}. \end{aligned}$$

PXY

Y

X		heads	tails		PX
	ripe	0.2	0.2		0.4
	unripe	0.3	0.3		0.6

PY	0.5	0.5

PX|Y

Y

X		heads	tails
	ripe	0.4	0.4
	unripe	0.6	0.6

PY|X

X

Y		ripe	unripe
	heads	0.5	0.5
	tails	0.5	0.5

Figure 2: Joint and conditional distributions (independent random variables).

Statistical Independence

When knowledge of the outcome of Y gives no information about the outcome of X then

$$p_{X|Y}(x_i \,|\, y_j) = p_X(x_i).$$

Since

$$p_{XY}(x_i, y_j) = p_{X|Y}(x_i \,|\, y_j) p_Y(y_j)$$

it follows that

$$p_{XY}(x_i, y_j) = p_X(x_i) p_Y(y_j).$$

Statistical Independence (contd.)

Furthermore, given that

$$p_{XY}(x_i, y_j) = p_X(x_i)p_Y(y_j)$$

it follows that knowledge of the outcome of X gives no information about the outcome of Y:

$$\begin{aligned} p_Y(y_j) &= \frac{p_{XY}(x_i, y_j)}{p_X(x_i)} \\ &= \frac{p_{Y|X}(y_j|x_i)p_X(x_i)}{p_X(x_i)} \\ &= p_{Y|X}(y_j|x_i). \end{aligned}$$

X and Y are said to be *statistically independent*.

Binomial Coefficient

The *binomial coefficient* is the number of subsets of size k drawn from a set of size n:

$$\binom{n}{k}.$$

The number of sequences of length k drawn from a set of size n is:

$$n(n-1)...(n-k+1).$$

There are $k!$ different orderings for each of these sequences. It follows that:

$$n(n-1)...(n-k+1) = \binom{n}{k}k!.$$

Binomial Coefficient (contd.)

Consequently,

$$\binom{n}{k} = \frac{n(n-1)...(n-k+1)}{k!}.$$

Multiplying numerator and denominator by $(n-k)!/(n-k)! = 1$ yields the familiar formula:

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)...(n-k+1)(n-k)!}{k!(n-k)!} \\ &= \frac{n!}{(n-k)!k!}.\end{aligned}$$

Binomial Distribution

A probabilistic experiment, X, has two outcomes, x_1 and x_2, which occur with probabilities, $p_X(x_1) = \theta$ and $p_X(x_2) = 1 - \theta$. Let k be the number of times the outcome of X is x_1 in n repeated trials. The distribution of the random variable, K, is given by

$$p_K(k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}.$$

This is called the *binomial distribution*.

Example

An anthropologist knows from records kept at a dig site that the probability that a recovered fossil human skull is female is 0.6. The probability that out of six skulls, exactly four will be female is:

$$p_K(4) = \binom{6}{4}(0.6)^4(0.4)^2 = 0.311.$$

Expected Value[1]

Let X be a discrete random variable with numerical outcomes, $\{x_1, \ldots, x_n\}$. The *expected value* of X, is defined as follows:

$$\langle X \rangle = \sum_{i=1}^{n} p_X(x_i) x_i.$$

Variance

The *variance* of X is defined as the expected value of the squared difference of X and $\langle X \rangle$:

$$\left\langle [X - \langle X \rangle]^2 \right\rangle = \sum_{i=1}^{n} p_X(x_i) \left[x_i - \langle X \rangle \right]^2.$$

[1] "God is or He is not...Let us weight the gain and the loss in choosing...'God is.' If you gain, you gain all, if you lose, you lose nothing. Wager, then, unhesitatingly, that He is." – Blaise Pascal

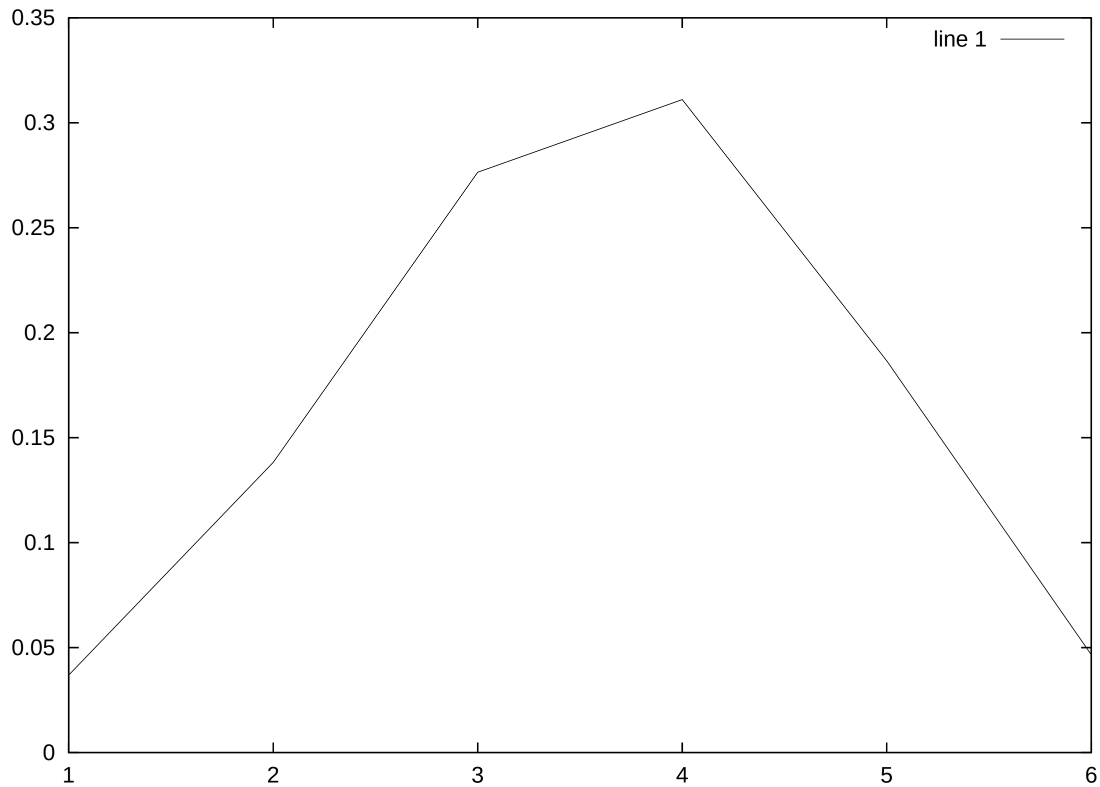

Figure 3: Binomial distribution for $n = 6$ and $\theta = 0.6$.

Expected Value of Binomial r.v.

The expected value of a binomial random variable with parameters n and θ:

$$
\begin{aligned}
\langle K \rangle &= \sum_{k=1}^{n} k p_K(k) \\
&= \sum_{k=1}^{n} k \binom{n}{k} \theta^k (1-\theta)^{n-k} \\
&= \sum_{k=1}^{n} \frac{kn!}{(n-k)!k!} \theta^k (1-\theta)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!} \theta^k (1-\theta)^{n-k} \\
&= n\theta \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} \theta^{k-1} (1-\theta)^{n-k}
\end{aligned}
$$

Expected Value of Binomial r.v. (contd.)

Letting $\ell = k - 1$:

$$\begin{aligned}
\langle K \rangle &= n\theta \sum_{\ell=0}^{n-1} \frac{(n-1)!}{(n-1-\ell)!\ell!} \theta^{\ell} (1-\theta)^{n-1-\ell} \\
&= n\theta \underbrace{\sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \theta^{\ell} (1-\theta)^{n-1-\ell}}_{1} \\
&= n\theta
\end{aligned}$$

A Useful Equality

$$\lim_{n\to\infty} \frac{\frac{n!}{(n-k)!}}{n^k}$$

$$= \lim_{n\to\infty} \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-(k-1))}{n^k}$$

$$= \lim_{n\to\infty} \frac{\overbrace{n \cdot n \cdot \cdots \cdot n}^{k} + \cdots}{n^k}$$

$$= 1$$

Another Useful Equality

Applying the *binomial formula*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

to $\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n$ yields:

$$\begin{aligned}
\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n &= \lim_{n\to\infty}\sum_{k=0}^{n}\binom{n}{k} 1^{n-k}\frac{(-\mu)^k}{n^k} \\
&= \lim_{n\to\infty}\sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}\frac{(-\mu)^k}{n^k} \\
&= \lim_{n\to\infty}\sum_{k=0}^{n}\underbrace{\frac{\frac{n!}{(n-k)!}}{n^k}}_{1}\;\frac{(-\mu)^k}{k!} \\
&= \sum_{k=0}^{\infty}\frac{(-\mu)^k}{k!} \\
&= e^{-\mu}.
\end{aligned}$$

Poisson Distribution

Consider a cube of uranium. On average, μ atoms in the cube transmute into lead per unit time. The actual number of atoms which transmute per unit time, k, is a *Poisson* random variable. The *Poisson distribution* is given by:

$$p_K(k) = \frac{\mu^k e^{-\mu}}{k!}.$$

The Poisson distribution can be derived from the Binomial distribution:

$$p_K(k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}.$$

Poisson Distribution (contd.)

Let n be the number of atoms and let θ be the probability that any single atom transmutes into lead in a unit of time. The average number atoms which decay per unit time, μ, is then:

$$\mu = n\theta.$$

1) Substituting μ/n for θ; 2) Expanding the binomial coefficient; and 3) Taking the limit, $n \to \infty$, yields:

$$\begin{aligned} p_K(k) &= \lim_{n\to\infty} \frac{n!}{(n-k)!k!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k} \\ &= \frac{\mu^k}{k!} \lim_{n\to\infty} \left[\frac{\frac{n!}{(n-k)!}}{n^k} \cdot \frac{\left(1 - \frac{\mu}{n}\right)^n}{\left(1 - \frac{\mu}{n}\right)^k} \right]. \end{aligned}$$

Poisson Distribution (contd.)

$$
\begin{aligned}
p_K(k) &= \frac{\mu^k}{k!} \lim_{n\to\infty} \left[\frac{\frac{n!}{(n-k)!}}{n^k} \cdot \frac{\left(1-\frac{\mu}{n}\right)^n}{\left(1-\frac{\mu}{n}\right)^k} \right] \\
&= \frac{\mu^k}{k!} \cdot \left[1 \cdot \frac{e^{-\mu}}{(1-0)^k} \right] \\
&= \frac{\mu^k e^{-\mu}}{k!}
\end{aligned}
$$

Random Processes and Random Variables

- Bernoulli trial
  - defined by parameter θ
- Repeated Bernoulli trials
  - defined by parameters θ and n
  - binomial random variable

    $p_K(k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$

    $\langle K \rangle = n\theta$
- Poisson process
  - defined by parameter μ
  - Poisson random variable

    $p_K(k) = \frac{\mu^k e^{-\mu}}{k!}$

    $\langle K \rangle = \mu$