Some Discrete Distribution Families

Many families of discrete distributions have been studied; we shall discuss the ones that are most commonly found in applications.

In each family, we need a formula for the probability mass function (pmf), and knowing the population expected value and standard deviation is also important.
Discrete Uniform Distribution

If a random variable $X$ takes only a finite number of values $x_1, x_2, \ldots, x_n$, and they have equal probabilities $1/n$, it has a discrete uniform distribution.

Expected value:

$$
\mu_X = \frac{1}{n} \sum_{i=1}^{n} x_i
$$

Variance:

$$
\sigma_X^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_X)^2
$$
Most often, the values are equally spaced integers:

\[ x_1 = a \]
\[ x_2 = a + 1, \]
\[ \ldots \]
\[ x_n = a + n - 1 = b \]

for some minimum and maximum integer values \( a \) and \( b \).

The expected value and variance are then

\[ \mu_X = \frac{a + b}{2} \]
\[ \sigma_X^2 = \frac{(b - a + 1)^2 - 1}{12} = \frac{n^2 - 1}{12}. \]
Example: Dice

Discrete uniform distributions have relatively few applications; the simplest is the number of spots shown when a fair die is rolled:

\[ a = 1 \]
\[ b = 6 \]
\[ n = b - a + 1 = 6 \]
\[ \mu_X = 3.5 \]
\[ \sigma^2_X = 2.916667. \]
Example: Sampling

One way to draw a single item from a finite population of size $N$ is:

- label the items from 1 to $N$;
- sample $X$ from the discrete uniform distribution on 1, 2, \ldots, $N$;
- choose the item labeled $X$. 
Binomial Distribution

Suppose that we carry out a sequence of independent Bernoulli trials: simple experiments with only two outcomes, “success” and “failure”, each with the same probability $p$ of success.

Then $X$, the number of successes after $n$ trials, has a binomial distribution.
Probability mass function

<table>
<thead>
<tr>
<th>$X = 0$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 0$</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$1 - p$</td>
<td>$p$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$(1 - p)^2$</td>
<td>$2p(1 - p)$</td>
<td>$p^2$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$(1 - p)^3$</td>
<td>$3p(1 - p)^2$</td>
<td>$3p^2(1 - p)$</td>
<td>$p^3$</td>
</tr>
<tr>
<td>4</td>
<td>$(1 - p)^4$</td>
<td>$4p(1 - p)^3$</td>
<td>$6p^2(1 - p)^2$</td>
<td>$4p^3(1 - p)$</td>
</tr>
</tbody>
</table>

Every entry is $1 - p$ times the entry directly above plus $p$ times the entry above and to the left.

In the $n^{th}$ row, the probabilities are the terms in the binomial expansion

$$[(1 - p) + p]^n = (1 - p)^n + np(1 - p)^{n-1} + \cdots + p^n.$$
In general,

\[ f(x; n, p) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \ldots, n, \]

where the binomial coefficient \( \binom{n}{x} \) is

\[ \binom{n}{x} = \frac{n!}{x!(n-x)!}. \]
Heuristic

Consider any one way of getting $x$ successes in $n$ trials: $SFF \ldots FS$, with $x$ $S$s and $n - x$ $F$s;

By independence, the probability of this sequence is

$$p \times (1 - p) \times (1 - p) \times \cdots \times (1 - p) \times p = p^x(1 - p)^{(n-x)};$$

There are $\binom{n}{x}$ ways of arranging $x$ $S$s and $n - x$ $F$s, so the probability is

$$\binom{n}{x} \times p^x(1 - p)^{(n-x)}. $$
Expected value:

\[ \mu_X = np \]

Variance:

\[ \sigma^2_X = np(1 - p) \]

Example: Inspection

Suppose the probability that a cell phone camera flash unit fails to conform to specifications is \( p \).

When the production process is in statistical control, units are probabilistically independent.

The number of non-compliant units in a sample of size \( n \) has a binomial distribution.
Geometric Distribution

Consider the same context of a sequence of independent Bernoulli trials. We wait until the first success, and $X$ is the number of that trial.

If the first success is at trial $x$, it must have been preceded by $x - 1$ failures, so

$$P(X = x) = (1 - p)^{x-1}p, \quad x = 1, 2, \ldots$$

$X$ has a geometric distribution.
Expected value:

\[ \mu_X = \frac{1}{p} \]

Variance:

\[ \sigma^2_X = \frac{1 - p}{p^2} \]

Example: Inspection

Suppose the probability that a cell phone camera flash unit fails to conform to specifications is \( p \).

If the production process is in statistical control, the number of units that have been inspected when the first non-compliant unit is detected has a geometric distribution.
More generally, we could wait until we have seen \( r \) successes. The number \( X \) of trials up to the \( r^{\text{th}} \) success has a negative binomial distribution.

Probability mass function:

\[
P(X = x) = \binom{x - 1}{r - 1} (1 - p)^{x-r} p^r
\]

Expected value:

\[
\mu_X = \frac{r}{p}
\]

Variance:

\[
\sigma_X^2 = \frac{r(1 - p)}{p^2}
\]
Hypergeometric Distribution

Suppose that a random sample of size $n$ is drawn from a finite population of $N$ items, each of which can be classified as a “Success” or a “Failure”. Assume that there are $K$ successes. The number $X$ of successes in the sample has a hypergeometric distribution.

Example: Acceptance sampling

A sample of 10 items is taken from a shipment of 200 items, of which $K$ are non-compliant; $X$ is the number of non-compliant items in the sample.
The population of size $N$ has $\binom{N}{n}$ possible samples of size $n$, and taking a “random sample” means that each of them is equally likely to be chosen.

If the sample contains $x$ successes, they could be taken in $\binom{K}{x}$ ways from the $K$ successes in the population.

Also it contains $n - x$ failures which could be taken in $\binom{N-K}{n-x}$ ways from the $N - K$ failures in the population.

So the probability mass function is

$$P(X = x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}, 0 \leq x \leq K \text{ and } 0 \leq n - x \leq N - K.$$
Write \( p = K/N \), the fraction of successes in the population.

**Expected value:**

\[
\mu_X = np
\]

**Variance:**

\[
\sigma^2_X = np(1 - p) \left( \frac{N - n}{N - 1} \right)
\]

Note that the expected value has the same expression as for the binomial distribution, and the variance differs from the binomial case in the factor \( \frac{N - n}{N - 1} \), known as the *finite population correction factor*. 
Poisson Distribution

Both the binomial distribution and the hypergeometric distribution are used to model counts, but in each case the counts have a definite upper limit of $n$, the sample size.

Some counts, like the number of defects on a silicon wafer, do not have such an upper bound, and the Poisson distribution is often used as a model.
Probability mass function:

\[ P(X = x) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, 2, \ldots \]

Expected value:

\[ \mu_X = \theta \]

Variance:

\[ \sigma^2_X = \theta \]
The items being counted are often distributed over some region, and the parameter $\theta$ is then related to a *rate* of occurrence.

**Example: Workplace accidents**

The number of accidents in a workplace might be modeled as having a Poisson distribution with

$$\theta = \lambda T,$$

where $\lambda$ is the rate of occurrence, say in accidents per week, and $T$ is the number of weeks over which accidents were counted.
Example: Yarn defects

The number of defects in a length of yarn might be modeled as having a Poisson distribution with

\[ \theta = \lambda T, \]

where \( \lambda \) is the rate of occurrence, say in defects per meter, and \( T \) is the length in meters of yarn inspected.
Example: Paper quality

The number of visible specks in a sample of white paper might be modeled as having a Poisson distribution with

$$\theta = \lambda T,$$

where $\lambda$ is the rate of occurrence, say in specks per square meter, and $T$ is the area in square meters of paper inspected.
Approximations

We can show that for large $N$, the hypergeometric pmf

$$P(X = x) = \frac{(K)_x (N-K)_{n-x}}{(N)_n}$$

is approximately the same as the binomial pmf

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

with $p = K/N$.

That is, sampling from a large population is almost the same as sampling from an infinite population.
Also, if $n$ is large and $p = \theta/n$ is small, then the binomial pmf

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

is approximately the same as the Poisson pmf

$$P(X = x) = \frac{\theta^x e^{-\theta}}{x!}$$

Example: Auto insurance

The number of auto insurance policies experiencing claims in a given period is bounded above by the number of policies the insurer has written, but might nevertheless be modeled as having a Poisson distribution.