Covariance and Independence:

Recall that joint random variables X and Y are independent if and only if $f(x, y) = g(x) \cdot h(y)$; i.e. their joint distribution is the product of the marginal distributions.

The following theorem is a consequence of this.

Theorem 4.2.13 *If X and Y are independent, then*

$$E(XY) = E(X) \cdot E(Y)$$

and $\sigma_{XY} = 0$.

This theorem is only a one-way implication as seen in the next example.

Example 4.2.14 *Let X and Y be discrete random variables with joint distribution given by,*

$$
\begin{array}{cc|ccc|}
 & & & x & \\
 & & -1 & 0 & 1 \\
\hline
 & -1 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\
y & 0 & 0 & 0 & 0 \\
 & 1 & \frac{1}{6} & 0 & \frac{1}{6} \\
\hline
\end{array}
$$

Find $cov(X, Y)$. Determine whether X and Y are independent.

To find $\text{cov}(X, Y)$ first we need the marginal distributions, $g(x)$ and $h(y)$.

Then we can find μ_X and μ_Y:
$\mu_X = \sum_x x \cdot g(x) = (-1) \cdot \frac{1}{3} + (0) \cdot \frac{1}{3} + (1) \cdot \frac{1}{3} = 0$

$\mu_Y =$

$$
\begin{array}{cc|ccc|c}
 & & & x & & \\
 & & -1 & 0 & 1 & \\
\hline
 & -1 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{2}{3} \\
y & 0 & 0 & 0 & 0 & 0 \\
 & 1 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{3} \\
\hline
 & & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} &
\end{array}
$$

Next we need $E(XY) = \sum_x \sum_y xy \cdot f(x,y)$:

$= ((-1) \cdot (-1)) \cdot \frac{1}{6} + \ldots$

So, $\mathrm{cov}(X,Y) = E(XY) - \mu_X \mu_Y = 0 - (0)(-\frac{1}{3}) = 0$.

However, the random variables are not independent, as can be seen by the example that $f(-1,-1) = \frac{1}{6} \neq g(-1) \cdot h(-1) = \frac{2}{9}$.

$$
\begin{array}{cc|ccc|c}
 & & & x & & \\
 & & 0 & 1 & 2 & \\
\hline
 & 0 & \frac{6}{36} & \frac{12}{36} & \frac{3}{36} & \frac{21}{36} \\
y & 1 & \frac{8}{36} & \frac{6}{36} & & \frac{14}{36} \\
 & 2 & \frac{1}{36} & & & \frac{1}{36} \\
\hline
 & & \frac{15}{36} & \frac{18}{36} & \frac{3}{36} &
\end{array}
$$

Are the random variables X and Y of the caplet example independent?

No. We found that $\text{cov}(X, Y) = -\frac{7}{54} \neq 0$.

Therefore they are not independent.

Example 4.2.15 *Let X and Y be jointly continuous random variables with joint density*

$$f(x) = \begin{cases} \frac{2x^2y}{41} & \text{for } 1 < x < 2, 1 < y < 3 \\ 0 & \text{otherwise} \end{cases}$$

Find the covariance of X and Y and determine whether they are independent.

Conditional Expectations:

Earlier we defined conditional probability $f(x|y)$, for joint random variables X and Y, we can also talk about conditional expectation.

Let X and Y have probability distribution/density $f(x, y)$, and let $u(X)$ be some function of X. The **conditional expected value of** $u(X)$ **given** $Y = y$ is

$$\text{discrete case:} \quad E(X|y) = \sum_x u(x) \cdot f(x|y)$$

$$\text{continuous case:} \quad E(X|y) = \int_x u(x) \cdot f(x|y)\, dx$$

$E(X|y)$ is called the **conditional mean of** X **given** $Y = y$.

Example 4.2.16

$$
y\quad
\begin{array}{c|ccc|c}
 & \multicolumn{3}{c}{x} & \\
 & 0 & 1 & 2 & \\
\hline
0 & \frac{6}{36} & \frac{12}{36} & \frac{3}{36} & \frac{21}{36} \\
1 & \frac{8}{36} & \frac{6}{36} & & \frac{14}{36} \\
2 & \frac{1}{36} & & & \frac{1}{36} \\
\hline
 & \frac{15}{36} & \frac{18}{36} & \frac{3}{36} &
\end{array}
$$

Find the expected value (conditional mean) of X given that $Y = 1$.

By definition,

$$E(X|1) = \sum_x x f(x|1).$$

Recall that $f(x|y) = \frac{f(x,y)}{h(y)}$, and so

$$f(0|1) = \frac{8}{14}, \quad f(1|1) = \frac{6}{14}, \quad f(2|1) = 0.$$

Therefore we have

$$E(X|1) = 0 \cdot \frac{8}{14} + 1 \cdot \frac{6}{14} + 2 \cdot 0 = \frac{6}{14} \approx 0.4286.$$

Chapter 5

Special Probability Distributions

This chapter presents some commonly used probability distributions for discrete random variables. Having a pre-determined probability distribution to model a chance experiment prevents from having to re-derive its properties each time (e.g. mean and variance).

The models presented depend on **parameters**; input values which tailor the probability distribution to the particular example.

In some cases the values of the distribution for a range of parameters are recorded in a table which can be used to evaluate probabilities, rather than computing the sums directly, (or integrating in the case of continuous random variables).

This is not only a convenience, in some cases it may be impractical to compute such values on the spot, or impossible if, for example, no exact expression exists for an integral.

5.1 Discrete Uniform Distribution

Suppose a random variable X has a finite range of k values, $\{x_1, x_2, \ldots, x_k\}$. Then X has **discrete uniform distribution** if

$$f(x) = \frac{1}{k}$$

for $x \in \{x_1, x_2, \ldots, x_k\}$. In other words each outcome is equally likely.

Our only parameter in this case is k. For the discrete uniform distribution:

$$\mu = \sum_{i=1}^{k} x_i f(x_i) = \frac{\sum_{i=1}^{k} x_i}{k}.$$

$$\sigma^2 = \sum_{i=1}^{k} (x_i - \mu)^2 f(x_i) = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k} = \frac{\sum_{i=1}^{k} x_i^2}{k} - \left(\frac{\sum_{i=1}^{k} x_i}{k} \right)^2$$

5.2 The Bernoulli Distribution

Consider an experiment with two possible outcomes, either success or failure. (For example, a single coin toss.)

Assign random variable X the value 1 for success and 0 for failure.

If the probability of success is θ, then the probability of a failure is $1 - \theta$.

In this case X is called a **Bernoulli random variable** and has **Bernoulli distribution** given by

$$f(x;\theta) = \theta^x (1-\theta)^{1-x} \quad \text{for } x = 0, 1.$$

Exercise: Show that the Bernoulli distribution has

$$\mu = \theta, \quad \sigma^2 = \theta(1-\theta).$$

5.3 Binomial Distribution

Now consider an experiment with repeated trials, in which the outcome of each trial is either a success or failure.

Random variable X will denote the number of successes, the probability of success is known to be θ, and n is the given number of trials in the experiment.

Then X has **binomial distribution** which is given by

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad \text{for } x = 0, 1, \ldots, n.$$

Random variable X is called a **binomial random variable** if and only if it has this distribution.

The Bernoulli distribution is the special case of the binomial distribution when $n = 1$; a single trial experiment.

Example 5.3.1 *Some examples of binomial random variables:*

- *Number of heads in 35 flips of a coin with 0.63 probability of heads and 0.37 probability of tails.*

  $P(17 \text{ heads}) = b(17; 35, 0.63)$

- *There is a %6.6 chance that a person has O- blood type. In a selection of 20 people what is the probability that 5 of them will have O- blood.*

  $P(5 \text{ people have O-}) = b(5; 20, 0.066)$

Values for $b(x; n, \theta)$ can be found in tables (see the textbook for example). These tables are usually computed for $n = 1, 2, \ldots, 20$ and $\theta = 0.5, 0.10, 0.15, \ldots, 0.50$.

To evaluate $b(x; n, \theta)$ from these tables for when $\theta > 0.50$ we can use the following property:

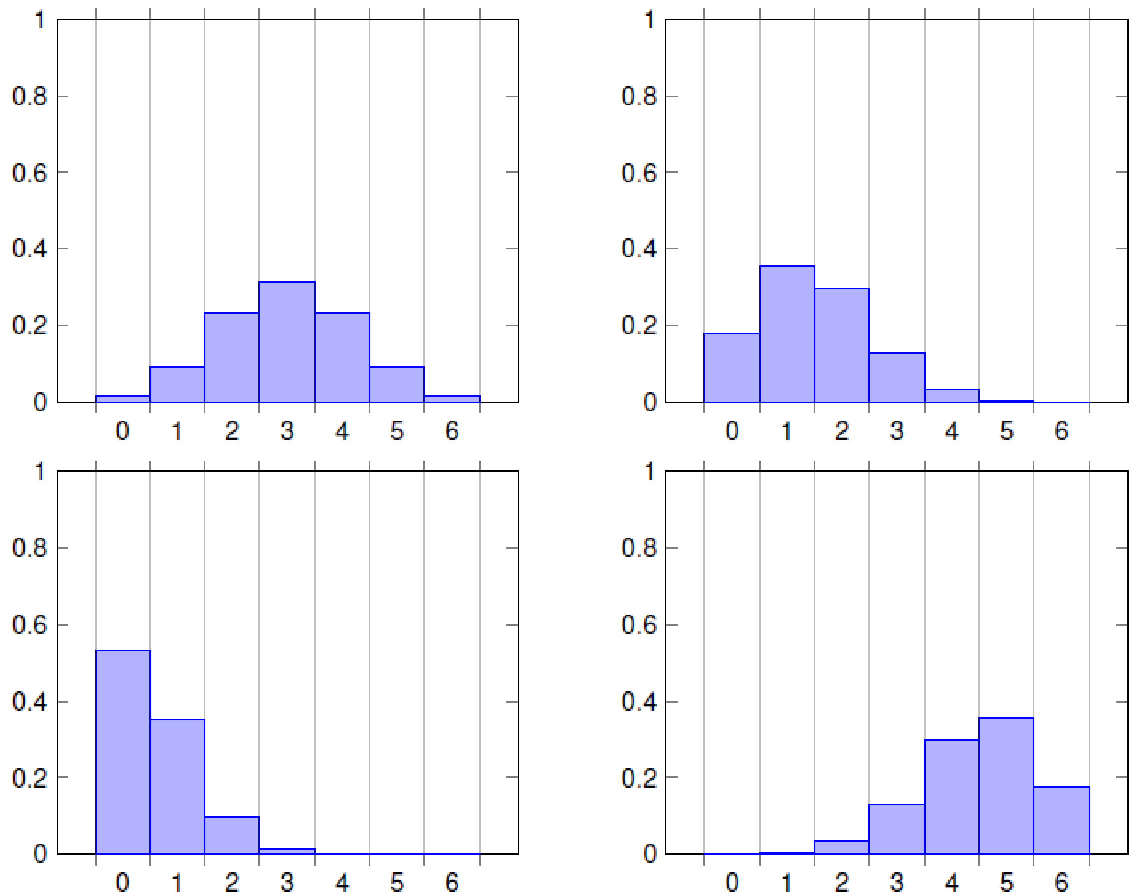
Theorem 5.3.2

$$b(x; n, \theta) = b(n - x; n, 1 - \theta)$$

For example,

$$b(7; 11, 0.75) = b(4; 11, 0.25) = 0.1721$$

Exercise: Show that the theorem holds.



For each graph of $b(x; n, \theta)$ we have $n = 6$. Determine which of these has $\theta = 0.1, 0.25, 0.5$, and 0.75

Moments of the Binomial Distribution:

Theorem 5.3.3 *Moment generating function of the binomial distribution is*

$$M_X(t) = (1 + \theta(e^t - 1))^n.$$

Proof 5.3.4

$$\begin{aligned}
M_X(t) &= E(e^{tX}) \\
&= \sum_{x=0}^{n} (e^{tx}) \cdot \binom{n}{x} \theta^x (1-\theta)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (e^t\theta)^x (1-\theta)^{n-x} \\
&= (e^t\theta + (1-\theta))^n \quad \textit{(by the binomial theorem)} \\
&= (1 + \theta(e^t - 1))^n.
\end{aligned}$$

From the moment generating function, we can find the mean.

Finding the **mean** from the moment generating function:

$$\mu = \frac{d}{dt} M_X(t) \Big|_{t=0}$$

$$= \frac{d}{dt} (1 + \theta(e^t - 1))^n \Big|_{t=0}$$

$$= n(1 + \theta(e^t - 1))^{n-1} \cdot (\theta e^t) \big|_{t=0}$$

$$= n(1 + \theta(e^0 - 1))^{n-1} \cdot (\theta e^0)$$

$$= n\theta.$$