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:

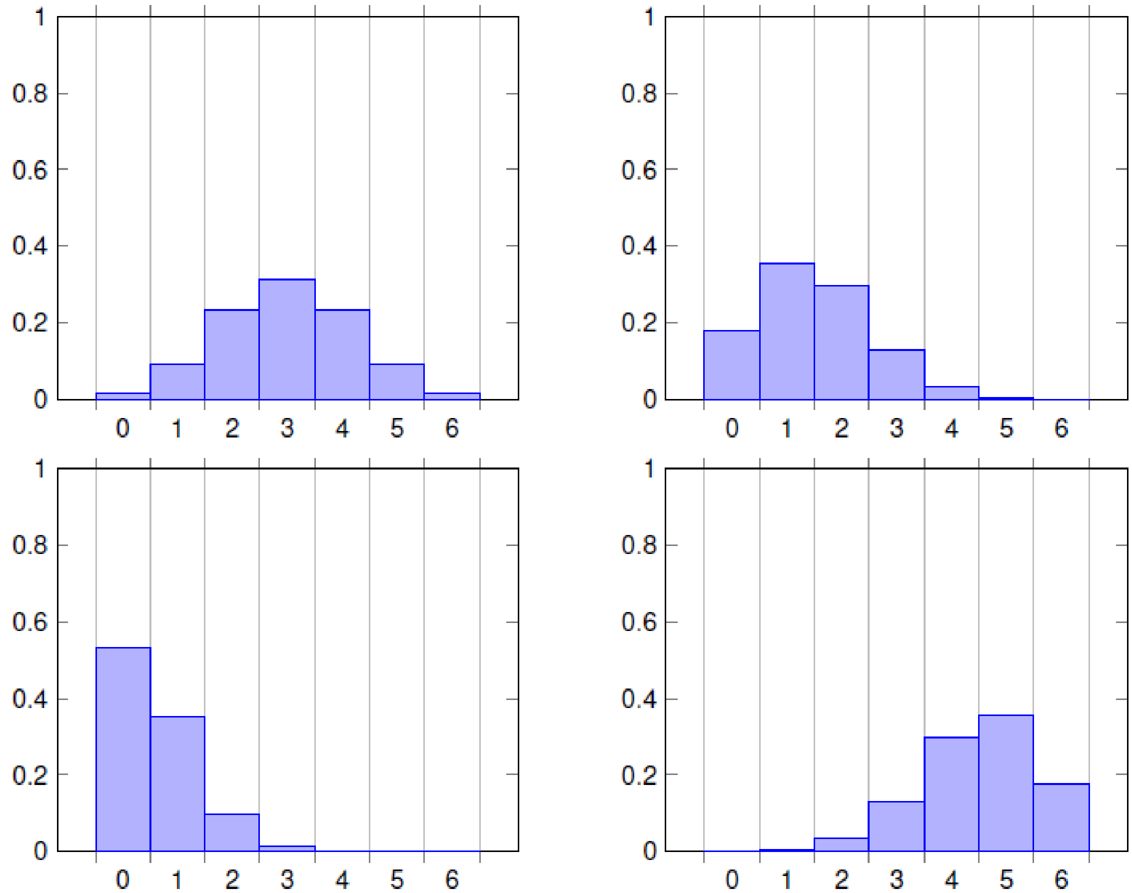
Theorem 5.3.2

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

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:

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

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

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

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

$$= n\theta.$$

Next we want to find the **variance**.
First, we need the second moment about the origin:

$$\begin{aligned}
E(X^2) &= \frac{d^2}{dt^2} M_X(t)\bigg|_{t=0} \\
&= \frac{d^2}{dt^2}(1+\theta(e^t-1))^n\bigg|_{t=0} \\
&= \frac{d}{dt} n\theta e^t (1+\theta(e^t-1))^{n-1}\bigg|_{t=0} \\
&= n\theta e^t (1+\theta(e^t-1))^{n-1} \\
&\quad + n(n-1)\theta e^t (1+\theta(e^t-1))^{n-2} \cdot (\theta e^t)\big|_{t=0} \\
&= n\theta + n(n-1)\theta^2
\end{aligned}$$

Finally, we can use the formula for the variance:

$$\sigma^2 = E(X^2) - \mu^2 = n\theta + n(n-1)\theta^2 - (n\theta)^2 = n\theta - n\theta^2 = n\theta(1-\theta).$$

We obtained the following theorem:

Theorem 5.3.5 *The mean and variance of the binomial distribution:*

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

Consider a random variable Y that denotes the proportion of successes in n trials.

So, $Y = \frac{X}{n}$, where X is the binomial random variable.
Then, the following holds.

Theorem 5.3.6 *Let X be a binomial random variable and let $Y = \frac{X}{n}$. Then*

$$E(Y) = \theta, \quad \sigma_Y^2 = \frac{\theta(1-\theta)}{n}.$$

By Chebyshev's Theorem, with $C = k\sigma$ ($k = \frac{C}{\sigma}$) we have

$$P(|Y - \theta| < C) \geq 1 - \frac{1}{k^2} = 1 - \frac{1}{(\frac{C^2}{\sigma^2})} = 1 - \frac{\theta(1-\theta)}{C^2 n}$$

Thus for any value of $C > 0$ we have

$$P\left(\left|\frac{X}{n} - \theta\right| < C\right) \geq 1 - \frac{\theta(1-\theta)}{C^2 n}.$$

When n is large, the fraction on the right side gets small, and so

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - \theta\right| < C\right) = 1.$$

This holds for any $C > 0$, no matter how small.

Explanation: The more trials we perform, the more likely it is that the proportion of successes will be close to the probability of a success θ.

Example: In repeatedly flipping a balanced coin, the more flips we perform (n), the more likely that the proportion of heads obtained ($\frac{X}{n}$) will be 0.5 (θ).