

3.2.3 Binomial Distribution

The binomial distribution is based on the idea of a *Bernoulli trial*. A Bernoulli trial is an experiment with two, and only two, possible outcomes. A random variable X has a Bernoulli(p) distribution if

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p, \end{cases}$$

where $0 \leq p \leq 1$. The value $X = 1$ is often termed a “success” and $X = 0$ is termed a “failure”. The mean and variance of a Bernoulli(p) random variable are easily seen to be

$$EX = (1)(p) + (0)(1 - p) = p$$

and

$$\text{Var}X = (1 - p)^2p + (0 - p)^2(1 - p) = p(1 - p).$$

In a sequence of n identical, independent Bernoulli trials, each with success probability p , define the random variables X_1, \dots, X_n by

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p. \end{cases}$$

The random variable

$$Y = \sum_{i=1}^n X_i$$

has the binomial distribution and it the number of successes among n independent trials. The probability mass function of Y is

$$\mathbb{P}(Y = y) = \binom{n}{y} p^y (1 - p)^{n-y}.$$

For this distribution,

$$EX = np, \quad \text{Var}(X) = np(1 - p), \quad M_X(t) = [pe^t + (1 - p)]^n.$$

Theorem 3.2.2 (Binomial theorem)

For any real numbers x and y and integer $n \geq 0$,

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}.$$

If we take $x = p$ and $y = 1 - p$, we get

$$1 = (p + (1 - p))^n = \sum_{i=0}^n \binom{n}{i} p^i (1 - p)^{n-i}.$$

Example 3.2.2 (Dice probabilities)

Suppose we are interested in finding the probability of obtaining at least one 6 in four rolls of a fair die. This experiment can be modeled as a sequence of four Bernoulli trials with success probability $p = \frac{1}{6}$. Define

X = total number of 6s in four rolls.

Then $X \sim \text{Bin}(4, \frac{1}{6})$ and

$$P(X > 0) = 1 - P(X = 0) = 1 - \binom{4}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 = 0.518.$$

Now we consider another game; throw a pair of dice 24 times and ask for the probability of at least one double 6. This, again, can be modeled by the binomial distribution with success probability p , where

$$p = P(\text{roll a double 6}) = \frac{1}{36}.$$

Let Y = number of double 6s in 24 rolls, then $Y \sim \text{binomial}(24, \frac{1}{36})$ and

$$P(Y > 0) = 1 - P(Y = 0) = 1 - \binom{24}{0} \left(\frac{1}{36}\right)^0 \left(\frac{35}{36}\right)^{24} = 0.491.$$

This is the calculation originally done in the 18th century by Pascal at the request of the gambler de Mere, who thought both events had the same probability. He began to believe he was wrong when he started losing money on the second bet.

Suppose $X \sim \text{Bin}(n, p)$.

- **Poisson Approximation to Binomial**

If $n \rightarrow \infty$ and $np_n \rightarrow \mu > 0$, then

$$\mathbf{P}(X = k) = \binom{n}{x} p_n^x (1 - p_n)^{n-x} \approx \frac{\mu^x e^{-\mu}}{x!}$$

Example 3.A Suppose that over the long run a manufacturing process produces 1% defective items. What is the chance of getting two or more defective items in a sample of 200 items?

- **Binomial Approximation to Normal**

If $n \rightarrow \infty$ and $np_n \rightarrow \infty$, then

$$\frac{X - np}{\sqrt{np(1-p)}} \approx \text{Normal}(0, 1)$$

In other words,

$$\mathbf{P}(a \leq X \leq b) \approx \Phi\left(\frac{b + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a + 0.5 - np}{\sqrt{np(1-p)}}\right).$$

$X \sim \text{Bin}(100, 1/2)$. Then, what is the probability $\mathbf{P}(45 \leq X \leq 55)$?