4.1 Approximations

Although many approximations of one distribution by another exist, we will give only three here. Others will be given along with the central-limit theorem

Binomial by Poisson We defined the binomial discrete density function, with parameters n and p, as

$$\binom{n}{x} p^x(1-p)^{n-x} \qquad \text{for } x = 0, 1, \ldots, n.$$

If the parameter n approaches infinity and p approaches 0 in such a way that np remains constant, say equal to λ, then

$$\binom{n}{x} p^x(1-p)^{n-x} \to \frac{e^{-\lambda}\lambda^x}{x!} \tag{47}$$

for fixed integer x. The above follows immediately from the following consideration:

$$\begin{aligned}\binom{n}{x} p^x(1-p)^{n-x} &= \frac{(n)_x}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\\ &= \frac{\lambda^x}{x!}\frac{(n)_x}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x} \to \frac{e^{-\lambda}\lambda^x}{x!},\end{aligned}$$

since

$$\frac{(n)_x}{n^x} \to 1, \qquad \left(1-\frac{\lambda}{n}\right)^{-x} \to 1, \qquad \text{and} \qquad \left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda} \qquad \text{as} \quad n \to \infty.$$

Thus, for large n and small p the binomial probability $\binom{n}{x}p^x(1-p)^{n-x}$ can be approximated by the Poisson probability $e^{-np}(np)^x/x!$. The utility of this approximation is evident if one notes that the binomial probability involves two parameters and the Poisson only one.

Binomial and Poisson by normal

Theorem 20 Let random variable X have a Poisson distribution with parameter λ; then for fixed $a < b$

$$P\left[a < \frac{X-\lambda}{\sqrt{\lambda}} < b\right]$$
$$= P[\lambda + a\sqrt{\lambda} < X < \lambda + b\sqrt{\lambda}] \rightarrow \Phi(b) - \Phi(a) \quad \text{as} \quad \lambda \rightarrow \infty. \tag{48}$$

PROOF Omitted. [Eq. (48) can be proved using Stirling's formula, which is given in Appendix A. It also follows from the central-limit theorem.] ////

Theorem 21 De Moivre–Laplace limit theorem Let a random variable X have a binomial distribution with parameters n and p; then for fixed $a < b$

$$P\left[a \leq \frac{X-np}{\sqrt{npq}} \leq b\right] = P[np + a\sqrt{npq} \leq X \leq np + b\sqrt{npq}] \rightarrow$$

$$\Phi(b) - \Phi(a) \qquad \text{as } n \rightarrow \infty. \tag{49}$$

PROOF Omitted. (This is a special case of the central-limit theorem, given in Chaps. V and VI.) ////

Remark We approximated the binomial distribution with a Poisson distribution in Eq. (47) for large n and small p. Theorem 21 gives a normal approximation of the binomial distribution for large n. ////

The usefulness of Theorems 20 and 21 rests in the approximations that they give. For instance, Eq. (49) states that $P[np + a\sqrt{npq} \leq X \leq np + b\sqrt{npq}]$ is approximately equal to $\Phi(b) - \Phi(a)$ for large n. Or if $c = np + a\sqrt{npq}$ and $d = np + b\sqrt{npq}$, then Eq. (49) gives that $P[c \leq X \leq d]$ is approximately equal to

$$\Phi\left(\frac{d-np}{\sqrt{npq}}\right) - \Phi\left(\frac{c-np}{\sqrt{npq}}\right)$$

for large n, and, so, an approximate value for the probability that a binomial random variable falls in an interval can be obtained from the standard normal distribution. Note that the binomial distribution is discrete and the approximating normal distribution is continuous.

EXAMPLE 15 Suppose that two fair dice are tossed 600 times. Let X denote the number of times a total of 7 occurs. Then X has a binomial distribution with parameters $n = 600$ and $p = \frac{1}{6}$. $\mathscr{E}[X] = 100$. Find $P[90 \le X \le 110]$.

$$P[90 \le X \le 110] = \sum_{j=90}^{110} \binom{600}{j} \left(\frac{1}{6}\right)^j \left(\frac{5}{6}\right)^{600-j},$$

a sum that is tedious to evaluate. Using the approximation given by Eq. (49), we have

$$P[90 \le X \le 110] \approx \Phi\left(\frac{110-100}{\sqrt{\frac{500}{6}}}\right) - \Phi\left(\frac{90-100}{\sqrt{\frac{500}{6}}}\right)$$

$$= \Phi(\sqrt{\tfrac{6}{5}}) - \Phi(-\sqrt{\tfrac{6}{5}}) \approx \Phi(1.095) - \Phi(-1.095) \approx .726.$$

////

4.2 Poisson and Exponential Relationship

When the Poisson distribution was introduced in Subsec. 2.4, an experiment consisting of the counting of the number of happenings of a certain phenomenon in time was given special consideration. We argued that under certain conditions the count of the number of happenings in a fixed time interval was Poisson distributed with parameter, the mean, proportional to the length of the interval. Suppose now that one of these happenings has just occurred; what then is the distribution of the length of time, say X, that one will have to wait until the next happening? $P[X > t] = P$[no happenings in time interval of length t] $= e^{-\nu t}$, where ν is the *mean occurrence rate*; so

$$F_X(t) = P[X \le t] = 1 - P[X > t] = 1 - e^{-\nu t} \qquad \text{for } t > 0;$$

that is, X has an exponential distribution. On the other hand, it can be proved, under an independence assumption, that if the happenings are occurring in time in such a way that the distribution of the lengths of time between successive happenings is exponential, then the distribution of the number of happenings in a fixed time interval is Poisson distributed. Thus the exponential and Poisson distributions are related.