Repeated Trials (Bernoulli Trials)

Consider a series of independent experiments. Suppose that the probability of event a in each experiment is

$$P(a) = p\,. \tag{1}$$

Let also

$$P(\bar{a}) = q \tag{2}$$

with

$$p + q = 1\,. \tag{3}$$

The probability that event a occurs k times in a specific order in n trials is $p^k q^{n-k}$. The number of ways that a can occur k times in n trials is equal to $\binom{n}{k}$. Therefore, the probability that a occurs k times in n trials in any order is:

$$P_n(k) = \binom{n}{k} p^k q^{n-k}\,. \tag{4}$$

Figure 1 shows the variation of the probability as given by Equation (4).

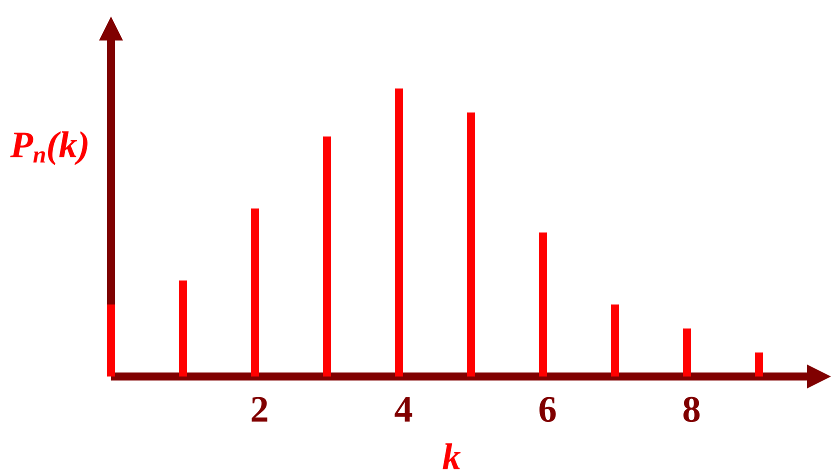

Figure 1. Variation of repeated trial probability.

Most Likely Number of Success

The value of k for which $P_n(k)$ is maximum is referred to as the most likely number of success. It may be shown that

$$k_{max} = \begin{cases} k_1, & k_1 = Greatest\ Integer \le (n+1)P \quad if\ (n+1)P \ne Integer \\ k_1\ and\ k_1 - 1, & k_1 = (n+1)P \quad if\ (n+1)P = Integer \end{cases}. \tag{5}$$

Probability that event a occurs k times with $k_1 \le k \le k_2$ is given as

$$P(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{n-k} . \tag{6}$$

Asymptotic Theorems

When the number of trials is very large, approximate asymptotic expressions for the probability may be used.

DeMoivre-Laplace Theorem

Let n be a large number and $npq >> 1$. Then for values of k in the $\sqrt{npq}$ neighborhood of its most likely value np (i.e. $|k - np|$ of the order of $\sqrt{npq}$) it may be shown that

$$P_n(k) = \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}} . \tag{7}$$

This approximation is known as the DeMoivre-Laplace theorem. The proof is based on Stirling formula

$$n! \approx n^n e^{-n} \sqrt{2\pi n} \qquad \text{as} \qquad n \to \infty . \tag{8}$$

Approximate Evaluation of P_n $(k_1 \le k \le k_2)$

In a similar limiting case, it follows that

$$P_n(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{2\pi npq}} \sum_{k=k_1}^{k_2} e^{-\frac{(k-np)^2}{2npq}} , \tag{9}$$

or

$$P_n(k_1 \le k \le k_2) \approx \frac{1}{\sqrt{2\pi npq}} \int_{k_1}^{k_2} e^{-\frac{(x-np)^2}{2npq}} dx \,. \tag{10}$$

Introducing the error function

$$erfx = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{y^2}{2}} dy \,, \qquad erf(\infty) = \frac{1}{2}, \quad -erf(-x) = erf(x), \tag{11}$$

we find

$$P_n(k_1 \le k \le k_2) \approx erf\,\frac{k_2 - np}{\sqrt{npq}} - erf\,\frac{k_1 - np}{\sqrt{npq}} \,. \tag{12}$$

Gaussian Functions

The Gaussian or normal function is defined as

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} , \qquad g(-x) = g(x). \tag{13}$$

Integral of $g(y)$ is given as

$$G(x) = \int_{-\infty}^{x} g(y) dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy \,, \qquad G(\infty) = 1 \,, \tag{14}$$

and

$$erf(x) = G(x) - \frac{1}{2}, \qquad G(-x) = 1 - G(x). \tag{15}$$

Using g and G, it follows that

$$P_n(k) = \binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sqrt{npq}} g\left(\frac{k - np}{\sqrt{npq}}\right), \tag{16}$$

$$P_n(k_1 \le k \le k_2) = G\left(\frac{k_2 - np}{\sqrt{npq}}\right) - G\left(\frac{k_1 - np}{\sqrt{npq}}\right) \,. \tag{17}$$

Generalized Bernoulli Trials

Assume that in the probability experiment $\Im$ the events a_1, a_2, … a_r are mutually exclusive (i.e. $a_i \cap a_j = 0$ for $i \neq j$) and $a_1 \cup a_2 \cup ... \cup a_r = S$. Let $P(a_1) = P_1$, $P(a_2) = P_2$, …, $P(a_r) = P_r$ with $P_1 + P_2 + ... + P_r = 1$. The probability that event a_1 occurs k_1 times, a_2 occurs k_2 times, …, a_r occurs k_r times in n independent trials with $k_1 + k_2 + ... + k_r = n$ is given as

$$P_n(k_1, k_2, ..., k_r) = \frac{n!}{k_1! k_2! ... k_r!} P_1^{k_1} P_2^{k_2} ... P_r^{k_r} . \tag{18}$$

For large n, if k_i is in the $\sqrt{n}$ vicinity of nP_i, the Demoivre-Laplace theorem implies that

$$P_n(k_1, k_2, ..., k_r) \approx \frac{\exp\left\{-\frac{1}{2}\left[\frac{(k_1 - nP_1)^2}{nP_1} + ... + \frac{(k_r - nP_r)^2}{nP_r}\right]\right\}}{\sqrt{(2\pi n)^{r-1} P_1 ... P_r}} . \tag{19}$$

Poisson Theorem

For a Bernoulli trial, suppose that n is very large but P is very small and such that $nP = a$ of the order of one. It then follows that as $n \to \infty$, $nP \to a$,

$$P_n(k) = \binom{n}{k} P^k q^{n-k} \approx e^{-nP} \frac{(nP)^k}{k!} = e^{-a} \frac{a^k}{k!} . \tag{20}$$

Furthermore,

$$P_n(k_1 \le k \le k_2) \approx e^{-nP} \sum_{k=k_1}^{k_2} \frac{(nP)^k}{k!} . \tag{21}$$

Random Poisson Points

We place at random n points in the interval $(0, T)$. Let $t_2 - t_1 = t_a$. The probability of finding k points in t_a is given as

$$P(k \ POINTS \ IN \ t_a) = \binom{n}{k} P^k q^{n-k}, \quad P = \frac{t_a}{T} . \tag{22}$$

Suppose $n >> 1$ and $t_a << T$ such that $n_P = \frac{nt_a}{T}$ is finite. From the Poisson Theorem we find

$$P(k \ Points \ in \ t_a) \approx e^{-\frac{nt_A}{T}} \frac{\left(\frac{nt_a}{T}\right)^k}{k!} \text{ as } n \to \infty, \ T \to \infty. \tag{23}$$

Let $\lambda = \frac{n}{T}$. Then,

$$P(k \ Point \ in \ t_a) \approx e^{-\lambda t_a} \frac{(\lambda t_a)^k}{k!}. \tag{24}$$

Useful Formula

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^n y^{n-k}, \qquad 2^n = \sum_{k=0}^{n} \binom{n}{k}, \quad \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$$

$$(x_1 + x_2 + ... + x_n)^n = \sum_{k_1, k_2, ..., k_r > 0} \binom{n}{k_1, k_2, ..., k_r} x_1^{k_1} x_2^{k_2} ... x_r^{k_r}$$

$$\sum_{k_1, k_2, ..., k_r} \binom{n}{k_1, k_2, ..., k_r} = r^n$$

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \qquad e = \sum_{k=0}^{\infty} \frac{1}{k!}$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 ... = \sum_{k=1}^{\infty} (-1)^{k-1} x^{k-1}$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + ... = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$$

$$\ln \frac{1+x}{1-x} = 2 \sum_{k=1}^{\infty} \frac{1}{2k-1} x^{2k-1}$$

$$\ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$$

$$Arctgx = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \ldots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}$$

$$Arcthx = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} \ldots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}$$