On average value and standard deviation of number of successes in Bernoulli trials

Expected value of number of successes in Bernoulli trials

Probability $P_N(X = k)$ of k successes in N Bernoulli trials, i.e. that the random variable X assumes value k is equal to

$$P_N(X = k) = \binom{N}{k} p^k q^{N-k} \tag{1}$$

where p is probablity of success and q is probability of failure. We always have that $p + q = 1$. The expected value $E[X]$ of a discrete random variable X assuming the values $x_1, x_2, \ldots, x_N$ with probability $P(X = x_i)$ is defined as

$$E[X] = \sum_{i=1}^{N} x_i P(X = x_i) \tag{2}$$

If $P_N(X = k)$ is the probability that in N Bernoulli trials the random variable X denoting the number of successes assumes the value k then the expected value $E[X]$, or in other words the average number of successes in N Bernoulli trials, is equal to

$$E[X] = \sum_{k=1}^{N} k P_N(X = k) = \sum_{k=1}^{N} k \binom{N}{k} p^k q^{N-k} \tag{3}$$

To compute the expected value $E[X]$ of number of successes in N Bernoulli trials we use the identity

$$(p + q)^N = \sum_{k=1}^{N} \binom{N}{k} p^k q^{N-k} \tag{4}$$

We may write the following function of variable x

$$(px+q)^N = \sum_{k=1}^{N} \binom{N}{k} p^k x^k q^{N-k} \tag{5}$$

and compute the derivatives of the left and right side of the equation (4) with respect to x

$$\frac{d(px+q)^N}{dx} = N(px+q)^{N-1}p \tag{6}$$

$$= \frac{d}{dx} \sum_{k=1}^{N} \binom{N}{k} p^k x^k q^{N-k} = \sum_{k=1}^{N} \binom{N}{k} p^k k x^{k-1} q^{N-k}$$

When in equation

$$N(px+q)^{N-1}p = \sum_{k=1}^{N} \binom{N}{k} p^k k x^{k-1} q^{N-k} \tag{7}$$

we substitute $x = 1$ and notice that $p + q = 1$, we obtain

$$Np = \sum_{k=1}^{N} k \binom{N}{k} p^k q^{N-k} \tag{8}$$

and based on equation (3) we see that the expected value $\mu = E[X]$ of number of successes in N Bernoulli trials equals to Np

$$\mu = Np \tag{9}$$

Standard deviation of number of successes in Bernoulli trials

In order to compute standard deviation $\sigma = \sqrt{E[(X - E[X])^2]}$ of random variable X one can express $E[(X - E[X])^2]$ in terms of $E[X^2]$ and $E[X]$

$$\begin{aligned} E[(X - E[X])^2] &= E[X^2 - 2XE[X] + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - 2(E[X])^2 + (E[X])^2 \\ &= E[X^2] - (E[X])^2 \end{aligned} \tag{10}$$

We differentite the expressions on the left and right side of the equation (7) with respect to x obtaining

$$\frac{dN(px+q)^{N-1}p}{dx} = N(N-1)(px+q)^{(N-2)}p^2 \tag{11}$$

$$= \frac{d}{dx}\sum_{k=1}^{N}\binom{N}{k}p^k k x^{k-1} q^{N-k} = \sum_{k=1}^{N}\binom{N}{k}p^k k(k-1)x^{k-2}q^{N-k}$$

$$= \sum_{k=1}^{N}\binom{N}{k}p^k k^2 x^{k-2} q^{N-k} - \sum_{k=1}^{N}\binom{N}{k}p^k k x^{k-2} q^{N-k}$$

In the above equation we substitute $x = 1$, $p + q = 1$ and receive

$$N(N-1)p^2 = \sum_{k=1}^{N} k^2 \binom{N}{k} p^k q^{N-k} - \sum_{k=1}^{N} k \binom{N}{k} p^k q^{N-k} \tag{12}$$

We notice that the last expression in the equation (12) based on equation (8) equals Np. Therefore we may write

$$E[X^2] = \sum_{k=1}^{N} k^2 \binom{N}{k} p^k q^{N-k} = N(N-1)p^2 + Np \tag{13}$$

To calculate $E[X^2] - (E[X])^2$ we write

$$E[X^2] - (E[X])^2 = N(N-1)p^2 + Np - N^2p^2 \tag{14}$$

$$= N^2p^2 - Np^2 + Np - N^2p^2 = Np - Np^2$$

$$= Np(1-p) = Npq$$

Therefrom we obtain the standard deviation of number of successes in Bernoulli trials as

$$\sigma = \sqrt{Npq} \tag{15}$$

References

[1] Kenneth S. Miller (1956) *Engineering Mathematics* Dover Publications, Inc., New York

Pawel Jan Piskorz (paweljs@gmail.com)