Repeated Trials

Lectures 6

Spring 2002

Bernoulli Trials

Repeated independent trials in which there can be only two outcomes are called Bernoulli trials in honor of James Bernoulli (1654-1705).

- Bernoulli trials lead to the binomial distribution.
- If a the number of trials is large, then the probability of k successes in n trials can be approximated by the Poisson distribution.
- The binomial distribution and the Poisson distribution are closely approximated by the normal (Gaussian) distribution.
- These three distributions are the foundation of much of the analysis of physical systems for detection, communication and storage of information.

Bernoulli Trials

Consider an experiment $\mathbf{E}$ that has two outcomes, say a and b, with probability p and $q = 1 - p$, respectively.

Let $\mathbf{E}_n$ be the experiment that consists of n independent repetitions of $\mathbf{E}$.

The outcomes $\mathbf{E}_n$ are n-sequences with the components a and b.

The outcomes of $\mathbf{E}_2$ are $\{aa\}, \{ab\}, \{ba\}, \{bb\}$, with probabilities p^2, pq, pq, and q^2, respectively.

Bernoulli Trials

Theorem: The outcomes of $\mathbf{E}_n$ are the 2^n sequences of length n. The number of outcomes of $\mathbf{E}_n$ that contain a exactly k times is given by the binomial coefficient. $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Proof: Assume that each of the terms in the expansion of $(a+b)^n$ represents one of the possible outcomes of the experiment $\mathbf{E}_n$.

Multiplying by $(a+b)$ to form $(a+b)^{n+1}$ produces an expression in which each term in $(a+b)^n$ appears twice–once with a appended and once with b appended.

If the assumption is true, then this constructs all possible distinct arrangements of $n+1$ terms.

The assumption is clearly true for $n = 2$.

Bernoulli Trials

Theorem: The probability that the outcome of an experiment that consists of n Bernoulli trials has k successes and $n-k$ failures is given by the *binomial distribution*

$$b(n,k,p) = \binom{n}{k} p^k (1-p)^{n-k}$$

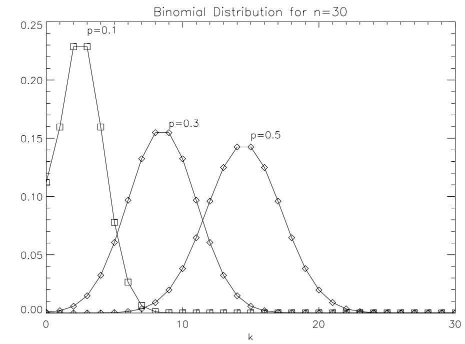
where the probability of success on an individual trial is given by p.

The peak value is near $k = np$, as was established in a homework problem.

In an experiment with n trials one can expect about np successes and $n(1-p$ failures.

Effect of changes in p

A graph of the binomial distribution for $n = 30$ and $p = 0.1$, 0.3 and 0.5 is shown below.



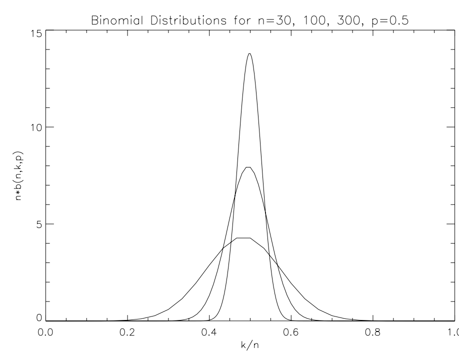
Effect of changes in n

- The mean value of a binomial distribution is np
- The variance of a binomial distribution is $np(1-p)$,
- The standard deviation is $\sqrt{np(1-p)}$
- The standard deviation is a measure of the spread of a distribution about its mean value.
- Both the mean value and the standard deviation increase with the number of trials, but the mean value increases faster.
- The ratio σ/μ is a measure of the spread relative to the mean value.

$$\frac{\sigma}{\mu} = \frac{\sqrt{np(1-p)}}{np} = \frac{1}{\sqrt{n}}\sqrt{\frac{1-p}{p}}$$

Effect of changes in n

A graph of the binomial distribution as a function of the fraction k/n is shown below for $n = 30$, 100 and 300. The vertical values are $nb(n,k,p)$, which compensates for the increase in density of points and gives the plots equal areas.



Law of Large Numbers

Define S_n to be the number of successes on n trials. Then

$$P\left[S_n = k\right] = b(n,k,p)$$

We would like to know the behavior of the function in the central region and on the tails. We can examine the tail on the right with the ratio (The tail on the left will be symmetric.)

$$\frac{b(n,r,p)}{b(n,r-1,p)} = \frac{(n-r+1)p}{rq} = 1 + \frac{(n+1)p - r}{rq} < 1 - \frac{r-np}{rq}$$

Example: $n = 120$, $p = 0.01$, $r = 15$. Then $r - np = 15 - 12 = 3$ while $r - (n+1)p = 15 - 13 = 2$. Hence,

$$1 - \frac{r-(n+1)p}{rq} < 1 - \frac{r-np}{rq}$$

Law of Large Numbers

The ratio of successive terms is a number that is decreasing. Therefore, the sum is smaller than the sum over a geometric series, in which the ratio of terms is a constant. A bound on the probability $P[S_n \geq r]$ is therefore given by the geometric sum with ratio ρ if ρ is the ratio for the first pair of terms.

$$P[S_n \geq r] \leq \sum_{k=0}^{\infty} b(n,r,p)\rho^k = b(n,r,p)\frac{1}{1-\rho}$$

Substitution of $\rho = 1 - \frac{r-np}{rq}$ now leads to the upper bound

$$P[S_n \geq r] \leq b(n,r,p)\frac{rq}{r-np}$$

Law of Large Numbers

We now need to replace $b(n,r,p)$ with an upper bound that is easy to work with. We do this by noting that all of the terms between the center, m, and r are greater than $b(n,r,p)$ and that the total of those terms must be less than 1. The number of such terms is no more than $r-np$, so $(r-np)b(n,r,p) < 1$ so that $b(n,r,p) < 1/(r-np)$. Putting this into the above equation yields the simple upper bound

$$\begin{aligned} P[S_n \geq r] &\leq b(n,r,p)\frac{rq}{r-np} \\ &\leq \left(\frac{1}{r-np}\right)\left(\frac{rq}{r-np}\right) \\ &\leq \frac{rq}{(r-np)^2} \quad \text{if } r > np \end{aligned}$$

Law of Large Numbers

A similar analysis could be performed on the left tail. However, this can be avoided by observing that saying that there are at most r successes is the same as saying there are at least $(n-r)$ failures. Exchanging $n-r$ for r and p for q on the right side above then yields, after simplification,

$$P[S_n \leq r] \leq \frac{(n-r)p}{(np-r)^2} \quad \text{if } r < np$$

Law of Large Numbers

Let us now look at the probability that the number of successes is much different from np. We can address this by using the above results. Let $r = n(p+\varepsilon)$. Then

$$P[S_n \geq n(p+\varepsilon)] \leq \frac{n(p+\varepsilon)q}{(n(p+\varepsilon)-np)^2} = \frac{n(p+\varepsilon)q}{(n\varepsilon)^2} \to 0$$

because the denominator grows as n^2 while the numerator grows as n. In the same way, the probability on the left tail also decreases with n, so that $P[S_n \leq n(p-\varepsilon)] \to 0$.

Almost all the probability is in the central region, which is of width $n\varepsilon$. Since the location of the center is $m = np$, the ratio of the width to the center point is ε/p, which can be as small as one wishes.

Law of Large Numbers

Theorem: The probability that the ratio S_n/n differs from p by less than ε in a set of n Bernoulli trials approaches unity as n increases.

$$P\left[\left|\frac{S_n}{n} - p\right| < \varepsilon\right] \to 1$$

As n increases, the probability that the average number of successes differs from p by more than ε tends to zero.*

We find application of the law of large numbers in many areas of science and engineering. One prominent example is in Shannon's development of the noisy channel coding theorem.

*For further discussion of the law of large numbers see William Feller, *An Introduction to Probability Theory and its Applications*, Vol I, page 152..