

9.2: Central Limit Theorem for Discrete Independent Trials

We have illustrated the Central Limit Theorem in the case of Bernoulli trials, but this theorem applies to a much more general class of chance processes. In particular, it applies to any independent trials process such that the individual trials have finite variance. For such a process, both the normal approximation for individual terms and the Central Limit Theorem are valid.

Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent discrete random variables of an independent trials process with common distribution function $m(x)$ defined on the integers, with mean μ and variance σ^2 . We can prevent this just as we did for Bernoulli trials.

Standardized Sums

Consider the standardized random variable

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}. \quad (9.2.1)$$

This standardizes S_n to have expected value 0 and variance 1. If $S_n = j$, then S_n^* has the value x_j with

$$x_j = \frac{j - n\mu}{\sqrt{n\sigma^2}}. \quad (9.2.2)$$

We can construct a spike graph just as we did for Bernoulli trials. Each spike is centered at some x_j . The distance between successive spikes is

$$b = \frac{1}{\sqrt{n\sigma^2}}, \quad (9.2.3)$$

and the height of the spike is

$$h = \sqrt{n\sigma^2} P(S_n = j). \quad (9.2.4)$$

The case of Bernoulli trials is the special case for which $X_j = 1$ if the j th outcome is a success and 0 otherwise; then $\mu = p$ and $\sigma^2 = \sqrt{pq}$.

We now illustrate this process for two different discrete distributions. The first is the distribution m , given by

$$m = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ .2 & .2 & .2 & .2 & .2 \end{pmatrix}. \quad (9.2.5)$$

In Figure 9.2.1 we show the standardized sums for this distribution for the cases $n = 2$ and $n = 10$. Even for $n = 2$ the approximation is surprisingly good.

For our second discrete distribution, we choose

$$m = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ .4 & .3 & .1 & .1 & .1 \end{pmatrix}. \quad (9.2.6)$$

This distribution is quite asymmetric and the approximation is not very good for $n = 3$, but by $n = 10$ we again have an excellent approximation (see Figure 9.1.5).

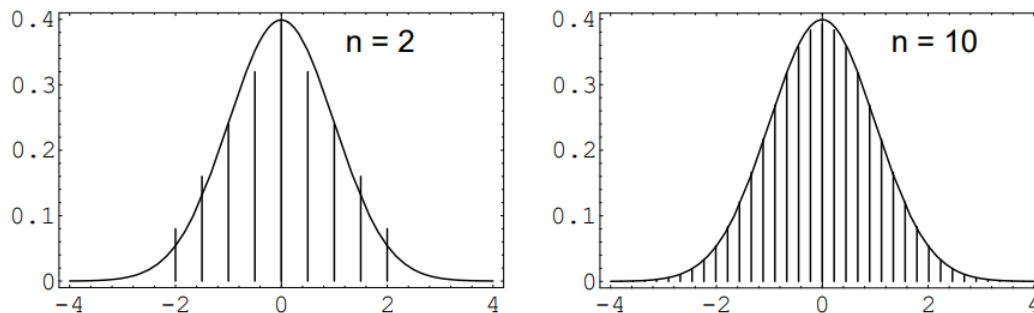


Figure 9.7: Distribution of standardized sums.

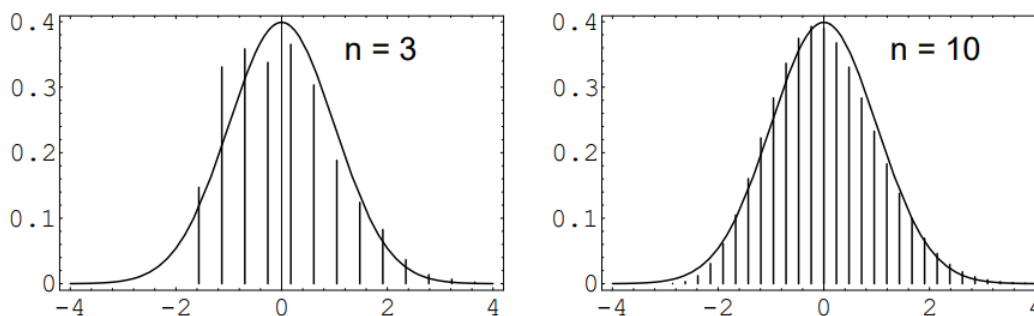


Figure 9.2.1: Distribution of standardized sums.

Approximation Theorem

As in the case of Bernoulli trials, these graphs suggest the following approximation theorem for the individual probabilities.

Theorem 9.2.1

Let X_1, X_2, \dots, X_n be an independent trials process and let $S_n = X_1 + X_2 + \dots + X_n$. Assume that the greatest common divisor of the differences of all the values that the X_j can take on is 1. Let $E(X_j) = \mu$ and $V(X_j) = \sigma^2$. Then for n large,

$$P(S_n = j) \sim \frac{\phi(x_j)}{\sqrt{n\sigma^2}}, \quad (9.2.7)$$

where $x_j = (j - n\mu)/\sqrt{n\sigma^2}$, and $\phi(x)$ is the standard normal density.

The program **CLTIndTrialsLocal** implements this approximation. When we run this program for 6 rolls of a die, and ask for the probability that the sum of the rolls equals 21, we obtain an actual value of .09285, and a normal approximation value of .09537. If we run this program for 24 rolls of a die, and ask for the probability that the sum of the rolls is 72, we obtain an actual value of .01724 and a normal approximation value of .01705. These results show that the normal approximations are quite good.

Central Limit Theorem for a Discrete Independent Trials Process

The Central Limit Theorem for a discrete independent trials process is as follows.

Theorem (Central Limit Theorem) 9.2.1

Let $S_n = X_1 + X_2 + \dots + X_n$ be the sum of n discrete independent random variables with common distribution having expected value μ and variance σ^2 . Then, for $a < b$,

$$\lim_{n \rightarrow \infty} P \left(a < \frac{S_n - n\mu}{\sqrt{n\sigma^2}} < b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx . \quad (9.2.8)$$

Here we consider several examples.

Examples

A die is rolled 420 times. What is the probability that the sum of the rolls lies between 1400 and 1550?

The sum is a random variable

$$S_{420} = X_1 + X_2 + \cdots + X_{420} , \quad (9.2.9)$$

where each X_j has distribution

$$m_X = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix} \quad (9.2.10)$$

We have seen that $\mu = E(X) = 7/2$ and $\sigma^2 = V(X) = 35/12$. Thus, $E(S_{420}) = 420 \cdot 7/2 = 1470$, $\sigma^2(S_{420}) = 420 \cdot 35/12 = 1225$, and $\sigma(S_{420}) = 35$. Therefore,

$$\begin{aligned} P(1400 \leq S_{420} \leq 1550) &\approx P \left(\frac{1399.5 - 1470}{35} \leq S_{420}^* \leq \frac{1550.5 - 1470}{35} \right) \\ &= P(-2.01 \leq S_{420}^* \leq 2.30) \\ &\approx \text{NA}(-2.01, 2.30) = .9670 . \end{aligned}$$

We note that the program **CLTIndTrialsGlobal** could be used to calculate these probabilities.

✓ Example 9.2.6:

A student's grade point average is the average of his grades in 30 courses. The grades are based on 100 possible points and are recorded as integers. Assume that, in each course, the instructor makes an error in grading of k with probability $|p/k|$, where $k = \pm 1, \pm 2, \pm 3, \pm 4, \pm 5$. The probability of no error is then $1 - (137/30)p$. (The parameter p represents the inaccuracy of the instructor's grading.) Thus, in each course, there are two grades for the student, namely the "correct" grade and the recorded grade. So there are two average grades for the student, namely the average of the correct grades and the average of the recorded grades.

We wish to estimate the probability that these two average grades differ by less than .05 for a given student. We now assume that $p = 1/20$. We also assume that the total error is the sum S_{30} of 30 independent random variables each with distribution

$$m_X : \left\{ \begin{array}{cccccccccccc} -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 \\ \frac{1}{100} & \frac{1}{80} & \frac{1}{60} & \frac{1}{40} & \frac{1}{20} & \frac{463}{600} & \frac{1}{20} & \frac{1}{40} & \frac{1}{60} & \frac{1}{80} & \frac{1}{100} \end{array} \right\} . \quad (9.2.11)$$

One can easily calculate that $E(X) = 0$ and $\sigma^2(X) = 1.5$. Then we have

$$\begin{aligned} P \left(-.05 \leq \frac{S_{30}}{30} \leq .05 \right) &= P(-1.5 \leq S_{30} \leq 1.5) \\ &= P \left(\frac{-1.5}{\sqrt{30 \cdot 1.5}} \leq S_{30}^* \leq \frac{1.5}{\sqrt{30 \cdot 1.5}} \right) \\ &= P(-.224 \leq S_{30}^* \leq .224) \\ &\approx \text{NA}(-.224, .224) = .1772 . \end{aligned} \quad (9.2.12)$$

This means that there is only a 17.7% chance that a given student's grade point average is accurate to within .05. (Thus, for example, if two candidates for valedictorian have recorded averages of 97.1 and 97.2, there is an appreciable probability that their correct averages are in the reverse order.) For a further discussion of this example, see the article by R. M. Kozelka.⁵

A More General Central Limit Theorem

In Theorem 9.2.1, the discrete random variables that were being summed were assumed to be independent and identically distributed. It turns out that the assumption of identical distributions can be substantially weakened. Much work has been done in this area, with an important contribution being made by J. W. Lindeberg. Lindeberg found a condition on the sequence $\{X_n\}$ which guarantees that the distribution of the sum S_n is asymptotically normally distributed. Feller showed that Lindeberg's condition is necessary as well, in the sense that if the condition does not hold, then the sum S_n is not asymptotically normally distributed. For a precise statement of Lindeberg's Theorem, we refer the reader to Feller.⁶ A sufficient condition that is stronger (but easier to state) than Lindeberg's condition, and is weaker than the condition in Theorem 9.2.1, is given in the following theorem.

? Theorem Central Limit Theorem 9.2.2

(Let $X_1, X_2, \dots, X_n, \dots$ be a sequence of independent discrete random variables, and let $S_n = X_1 + X_2 + \dots + X_n$. For each n , denote the mean and variance of X_n by μ_n and σ_n^2 , respectively. Define the mean and variance of S_n to be m_n and s_n^2 , respectively, and assume that $s_n \rightarrow \infty$. If there exists a constant A , such that $|X_n| \leq A$ for all n , then for $a < b$,

$$\lim_{n \rightarrow \infty} P\left(a < \frac{S_n - m_n}{s_n} < b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx. \quad (9.2.13)$$

The condition that $|X_n| \leq A$ for all n is sometimes described by saying that the sequence $\{X_n\}$ is uniformly bounded. The condition that $s_n \rightarrow \infty$ is necessary (see Exercise 9.2.15).

We illustrate this theorem by generating a sequence of n random distributions on the interval $[a, b]$. We then convolute these distributions to find the distribution of the sum of n independent experiments governed by these distributions. Finally, we standardize the distribution for the sum to have mean 0 and standard deviation 1 and compare it with the normal density. The program **CLTGeneral** carries out this procedure.

In Figure 9.2.2 we show the result of running this program for $[a, b] = [-2, 4]$, and $n = 1, 4$, and 10. We see that our first random distribution is quite asymmetric. By the time we choose the sum of ten such experiments we have a very good fit to the normal curve.

The above theorem essentially says that anything that can be thought of as being made up as the sum of many small independent pieces is approximately normally distributed. This brings us to one of the most important questions that was asked about genetics in the 1800's.

The Normal Distribution and Genetics

When one looks at the distribution of heights of adults of one sex in a given population, one cannot help but notice that this distribution looks like the normal distribution. An example of this is shown in Figure 9.2.3. This figure shows the distribution of heights of 9593 women between the ages of 21 and 74. These data come from the Health and Nutrition Examination Survey I (HANES I). For this survey, a sample of the U.S. civilian population was chosen. The survey was carried out between 1971 and 1974.

A natural question to ask is "How does this come about?". Francis Galton, an English scientist in the 19th century, studied this question, and other related questions, and constructed probability models that were of great importance in explaining the genetic effects on such attributes as height. In fact, one of the most important ideas in statistics, the idea of regression to the mean, was invented by Galton in his attempts to understand these genetic effects.

Galton was faced with an apparent contradiction. On the one hand, he knew that the normal distribution arises in situations in which many small independent effects are being summed. On the other hand, he also knew that many quantitative attributes, such as height, are strongly influenced by genetic factors: tall parents tend to have tall offspring. Thus in this case, there seem to be two large effects, namely the parents. Galton was certainly aware of the fact that non-genetic factors played a role in determining the height of an individual. Nevertheless, unless these non-genetic factors overwhelm the genetic ones, thereby refuting the hypothesis that heredity is important in determining height, it did not seem possible for sets of parents of given heights to have offspring whose heights were normally distributed.

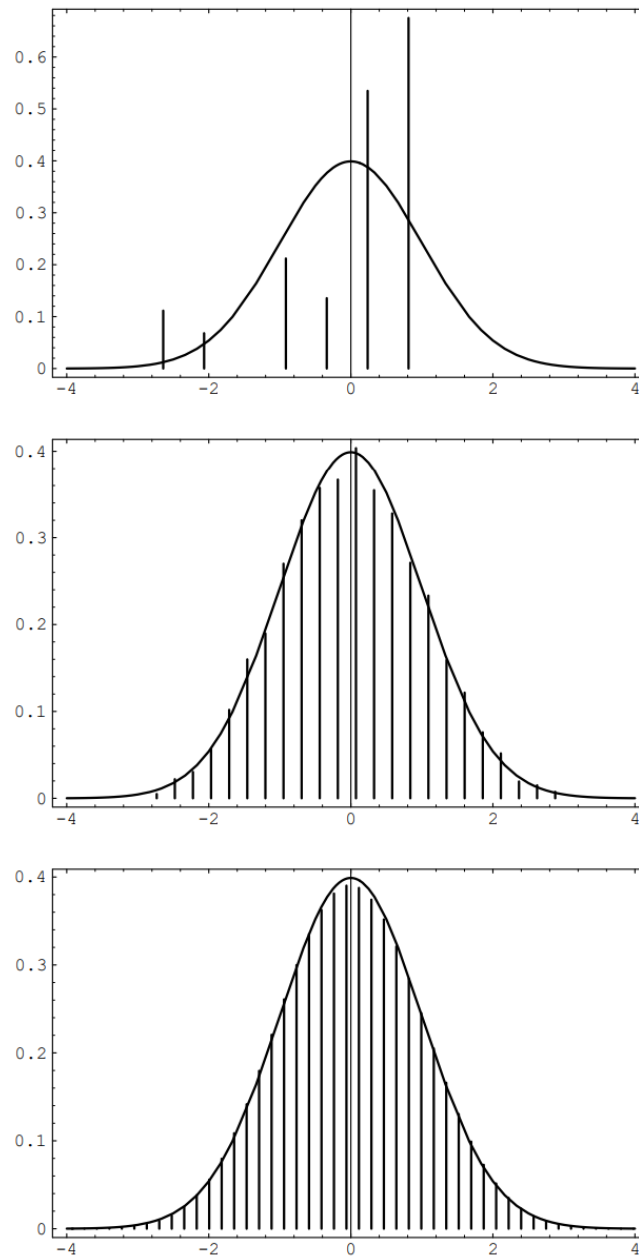


Figure 9.2.2: Sums of randomly chosen random variables.

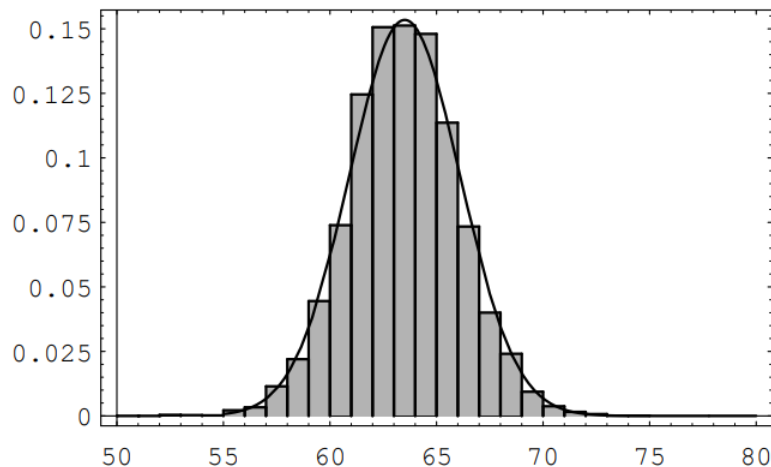


Figure 9.2.3: Distribution of heights of adult women.

One can express the above problem symbolically as follows. Suppose that we choose two specific positive real numbers x and y , and then find all pairs of parents one of whom is x units tall and the other of whom is y units tall. We then look at all of the offspring of these pairs of parents. One can postulate the existence of a function $f(x, y)$ which denotes the genetic effect of the parents' heights on the heights of the offspring. One can then let W denote the effects of the non-genetic factors on the heights of the offspring. Then, for a given set of heights $\{x, y\}$, the random variable which represents the heights of the offspring is given by

$$H = f(x, y) + W, \quad (9.2.14)$$

where f is a deterministic function, i.e., it gives one output for a pair of inputs $\{x, y\}$. If we assume that the effect of f is large in comparison with the effect of W , then the variance of W is small. But since f is deterministic, the variance of H equals the variance of W , so the variance of H is small. However, Galton observed from his data that the variance of the heights of the offspring of a given pair of parent heights is not small. This would seem to imply that inheritance plays a small role in the determination of the height of an individual. Later in this section, we will describe the way in which Galton got around this problem.

We will now consider the modern explanation of why certain traits, such as heights, are approximately normally distributed. In order to do so, we need to introduce some terminology from the field of genetics. The cells in a living organism that are not directly involved in the transmission of genetic material to offspring are called somatic cells, and the remaining cells are called germ cells. Organisms of a given species have their genetic information encoded in sets of physical entities, called chromosomes. The chromosomes are paired in each somatic cell. For example, human beings have 23 pairs of chromosomes in each somatic cell. The sex cells contain one chromosome from each pair. In sexual reproduction, two sex cells, one from each parent, contribute their chromosomes to create the set of chromosomes for the offspring.

Chromosomes contain many subunits, called genes. Genes consist of molecules of DNA, and one gene has, encoded in its DNA, information that leads to the regulation of proteins. In the present context, we will consider those genes containing information that has an effect on some physical trait, such as height, of the organism. The pairing of the chromosomes gives rise to a pairing of the genes on the chromosomes.

In a given species, each gene can be any one of several forms. These various forms are called alleles. One should think of the different alleles as potentially producing different effects on the physical trait in question. Of the two alleles that are found in a given gene pair in an organism, one of the alleles came from one parent and the other allele came from the other parent. The possible types of pairs of alleles (without regard to order) are called genotypes.

If we assume that the height of a human being is largely controlled by a specific gene, then we are faced with the same difficulty that Galton was. We are assuming that each parent has a pair of alleles which largely controls their heights. Since each parent contributes one allele of this gene pair to each of its offspring, there are four possible allele pairs for the offspring at this gene location. The assumption is that these pairs of alleles largely control the height of the offspring, and we are also assuming that genetic factors outweigh non-genetic factors. It follows that among the offspring we should see several modes in the height

distribution of the offspring, one mode corresponding to each possible pair of alleles. This distribution does not correspond to the observed distribution of heights.

An alternative hypothesis, which does explain the observation of normally distributed heights in offspring of a given sex, is the multiple-gene hypothesis. Under this hypothesis, we assume that there are many genes that affect the height of an individual. These genes may differ in the amount of their effects. Thus, we can represent each gene pair by a random variable X_i , where the value of the random variable is the allele pair's effect on the height of the individual. Thus, for example, if each parent has two different alleles in the gene pair under consideration, then the offspring has one of four possible pairs of alleles at this gene location. Now the height of the offspring is a random variable, which can be expressed as

$$H = X_1 + X_2 + \cdots + X_n + W, \quad (9.2.15)$$

if there are n genes that affect height. (Here, as before, the random variable W denotes non-genetic effects.) Although n is fixed, if it is fairly large, then Theorem 9.2.2 implies that the sum $X_1 + X_2 + \cdots + X_n$ is approximately normally distributed. Now, if we assume that the X_i 's have a significantly larger cumulative effect than W does, then H is approximately normally distributed.

Another observed feature of the distribution of heights of adults of one sex in a population is that the variance does not seem to increase or decrease from one generation to the next. This was known at the time of Galton, and his attempts to explain this led him to the idea of regression to the mean. This idea will be discussed further in the historical remarks at the end of the section. (The reason that we only consider one sex is that human heights are clearly sex-linked, and in general, if we have two populations that are each normally distributed, then their union need not be normally distributed.)

Using the multiple-gene hypothesis, it is easy to explain why the variance should be constant from generation to generation. We begin by assuming that for a specific gene location, there are k alleles, which we will denote by A_1, A_2, \dots, A_k . We assume that the offspring are produced by random mating. By this we mean that given any offspring, it is equally likely that it came from any pair of parents in the preceding generation. There is another way to look at random mating that makes the calculations easier. We consider the set S of all of the alleles (at the given gene location) in all of the germ cells of all of the individuals in the parent generation. In terms of the set S , by random mating we mean that each pair of alleles in S is equally likely to reside in any particular offspring. (The reader might object to this way of thinking about random mating, as it allows two alleles from the same parent to end up in an offspring; but if the number of individuals in the parent population is large, then whether or not we allow this event does not affect the probabilities very much.)

For $1 \leq i \leq k$, we let p_i denote the proportion of alleles in the parent population that are of type A_i . It is clear that this is the same as the proportion of alleles in the germ cells of the parent population, assuming that each parent produces roughly the same number of germ cells. Consider the distribution of alleles in the offspring. Since each germ cell is equally likely to be chosen for any particular offspring, the distribution of alleles in the offspring is the same as in the parents.

We next consider the distribution of genotypes in the two generations. We will prove the following fact: the distribution of genotypes in the offspring generation depends only upon the distribution of alleles in the parent generation (in particular, it does not depend upon the distribution of genotypes in the parent generation). Consider the possible genotypes; there are $k(k+1)/2$ of them. Under our assumptions, the genotype $A_i A_i$ will occur with frequency p_i^2 , and the genotype $A_i A_j$, with $i \neq j$, will occur with frequency $2p_i p_j$. Thus, the frequencies of the genotypes depend only upon the allele frequencies in the parent generation, as claimed.

This means that if we start with a certain generation, and a certain distribution of alleles, then in all generations after the one we started with, both the allele distribution and the genotype distribution will be fixed. This last statement is known as the Hardy-Weinberg Law.

We can describe the consequences of this law for the distribution of heights among adults of one sex in a population. We recall that the height of an offspring was given by a random variable H , where

$$H = X_1 + X_2 + \cdots + X_n + W, \quad (9.2.16)$$

with the X_i 's corresponding to the genes that affect height, and the random variable W denoting non-genetic effects. The Hardy-Weinberg Law states that for each X_i , the distribution in the offspring generation is the same as the distribution in the parent generation. Thus, if we assume that the distribution of W is roughly the same from generation to generation (or if we assume that its effects are small), then the distribution of H is the same from generation to generation. (In fact, dietary effects are part of W , and it is clear that in many human populations, diets have changed quite a bit from one generation to the next in recent times. This

change is thought to be one of the reasons that humans, on the average, are getting taller. It is also the case that the effects of W are thought to be small relative to the genetic effects of the parents.)

Discussion

Generally speaking, the Central Limit Theorem contains more information than the Law of Large Numbers, because it gives us detailed information about the of the distribution of S_n^* ; for large n the shape is approximately the same as the shape of the standard normal density. More specifically, the Central Limit Theorem says that if we standardize and height-correct the distribution of S_n , then the normal density function is a very good approximation to this distribution when n is large. Thus, we have a computable approximation for the distribution for S_n , which provides us with a powerful technique for generating answers for all sorts of questions about sums of independent random variables, even if the individual random variables have different distributions.

Historical Remarks

In the mid-1800's, the Belgian mathematician Quetelet⁷ had shown empirically that the normal distribution occurred in real data, and had also given a method for fitting the normal curve to a given data set. Laplace⁸ had shown much earlier that the sum of many independent identically distributed random variables is approximately normal. Galton knew that certain physical traits in a population appeared to be approximately normally distributed, but he did not consider Laplace's result to be a good explanation of how this distribution comes about. We give a quote from Galton that appears in the fascinating book by S. Stigler⁹ on the history of statistics:

First, let me point out a fact which Quetelet and all writers who have followed in his paths have unaccountably overlooked, and which has an intimate bearing on our work to-night. It is that, although characteristics of plants and animals conform to the law, the reason of their doing so is as yet totally unexplained. The essence of the law is that differences should be wholly due to the collective actions of a host of independent influences in various combinations...Now the processes of heredity...are not petty influences, but very important ones...The conclusion is...that the processes of heredity must work harmoniously with the law of deviation, and be themselves in some sense conformable to it.

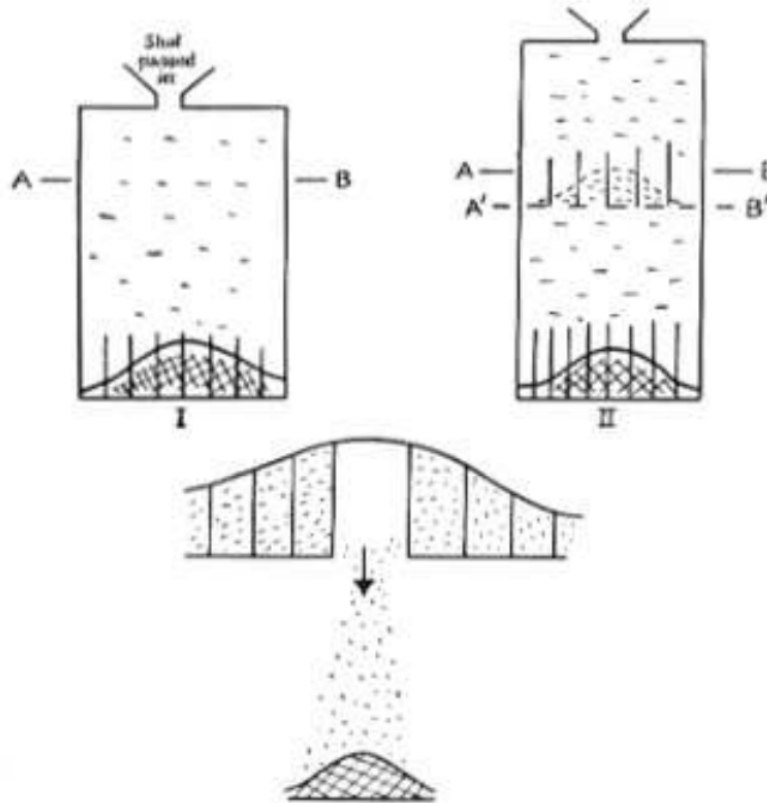


Figure 9.2.4: Two-stage version of the quincunx

Galton invented a device known as a quincunx (now commonly called a Galton board), which we used in Example 3.2.1 to show how to physically obtain a binomial distribution. Of course, the Central Limit Theorem says that for large values of the parameter n , the binomial distribution is approximately normal. Galton used the quincunx to explain how inheritance affects the distribution of a trait among offspring.

We consider, as Galton did, what happens if we interrupt, at some intermediate height, the progress of the shot that is falling in the quincunx. The reader is referred to Figure [fig 9.62]. This figure is a drawing of Karl Pearson,¹⁰ based upon Galton's notes. In this figure, the shot is being temporarily segregated into compartments at the line AB. (The line A'B' forms a platform on which the shot can rest.) If the line AB is not too close to the top of the quincunx, then the shot will be approximately normally distributed at this line. Now suppose that one compartment is opened, as shown in the figure. The shot from that compartment will fall, forming a normal distribution at the bottom of the quincunx. If now all of the compartments are opened, all of the shot will fall, producing the same distribution as would occur if the shot were not temporarily stopped at the line AB. But the action of stopping the shot at the line AB, and then releasing the compartments one at a time, is just the same as convoluting two normal distributions. The normal distributions at the bottom, corresponding to each compartment at the line AB, are being mixed, with their weights being the number of shot in each compartment. On the other hand, it is already known that if the shot are unimpeded, the final distribution is approximately normal. Thus, this device shows that the convolution of two normal distributions is again normal.

Galton also considered the quincunx from another perspective. He segregated into seven groups, by weight, a set of 490 sweet pea seeds. He gave 10 seeds from each of the seven group to each of seven friends, who grew the plants from the seeds. Galton found that each group produced seeds whose weights were normally distributed. (The sweet pea reproduces by self-pollination, so he did not need to consider the possibility of interaction between different groups.) In addition, he found that the variances of the weights of the offspring were the same for each group. This segregation into groups corresponds to the compartments at the line AB in the quincunx. Thus, the sweet peas were acting as though they were being governed by a convolution of normal distributions.

He now was faced with a problem. We have shown in Chapter 7 and Galton knew, that the convolution of two normal distributions produces a normal distribution with a larger variance than either of the original distributions. But his data on the sweet pea seeds

showed that the variance of the offspring population was the same as the variance of the parent population. His answer to this problem was to postulate a mechanism that he called , and is now called . As Stigler puts it:¹¹

The seven groups of progeny were normally distributed, but not about their parents' weight. Rather they were in every case distributed about a value that was closer to the average population weight than was that of the parent. Furthermore, this reversion followed "the simplest possible law," that is, it was linear. The average deviation of the progeny from the population average was in the same direction as that of the parent, but only a third as great. The mean progeny reverted to type, and the increased variation was just sufficient to maintain the population variability.

Galton illustrated reversion with the illustration shown in Figure [fig 9.63].¹² The parent population is shown at the top of the figure, and the slanted lines are meant to correspond to the reversion effect. The offspring population is shown at the bottom of the figure.

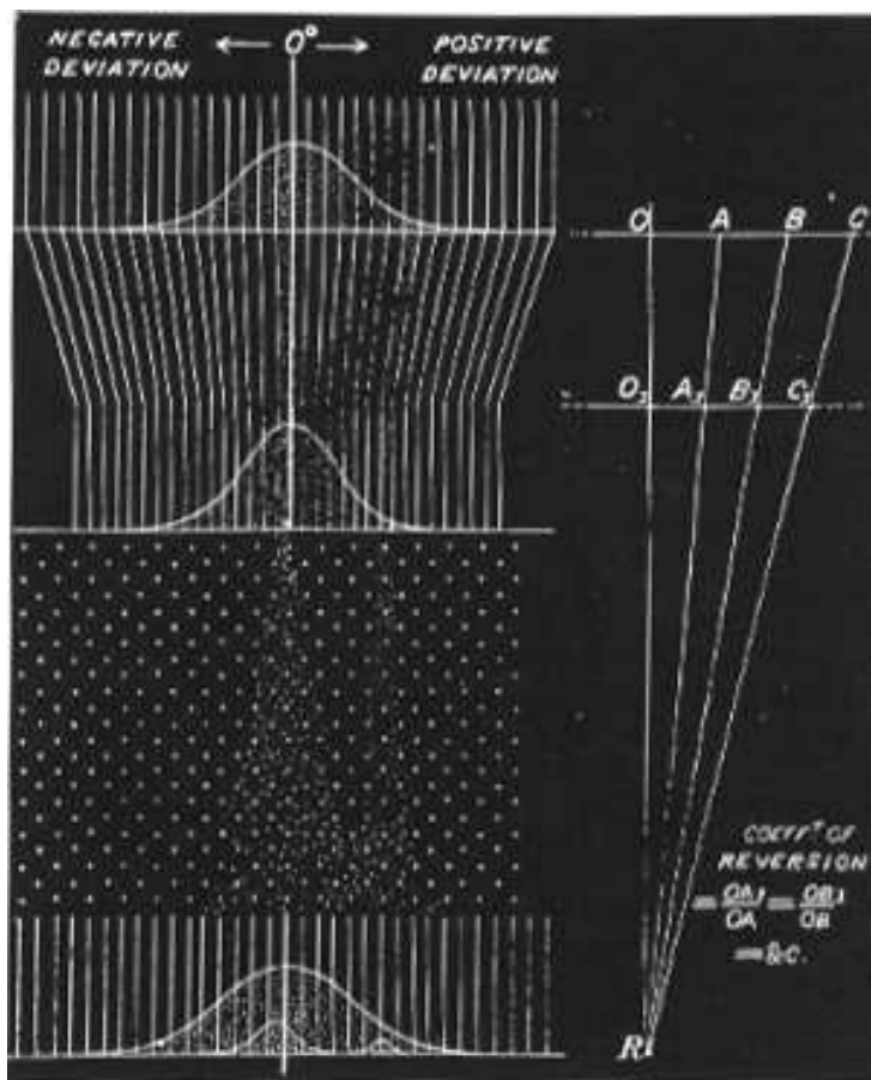


Figure 9.2.4: Galton's explanation of reversion.

Exercises

Exercise 9.2.1:

A die is rolled 24 times. Use the Central Limit Theorem to estimate the probability that

- the sum is greater than 84.
- the sum is equal to 84.

Exercise 9.2.2:

A random walker starts at 0 on the x -axis and at each time unit moves 1 step to the right or 1 step to the left with probability $1/2$. Estimate the probability that, after 100 steps, the walker is more than 10 steps from the starting position.

Exercise 9.2.3:

A piece of rope is made up of 100 strands. Assume that the breaking strength of the rope is the sum of the breaking strengths of the individual strands. Assume further that this sum may be considered to be the sum of an independent trials process with 100 experiments each having expected value of 10 pounds and standard deviation 1. Find the approximate probability that the rope will support a weight

- of 1000 pounds.
- of 970 pounds.

Exercise 9.2.4:

Write a program to find the average of 1000 random digits 0, 1, 2, 3, 4, 5, 6, 7, 8, or 9. Have the program test to see if the average lies within three standard deviations of the expected value of 4.5. Modify the program so that it repeats this simulation 1000 times and keeps track of the number of times the test is passed. Does your outcome agree with the Central Limit Theorem?

Exercise 9.2.5:

A die is thrown until the first time the total sum of the face values of the die is 700 or greater. Estimate the probability that, for this to happen,

- more than 210 tosses are required.
- less than 190 tosses are required.
- between 180 and 210 tosses, inclusive, are required.

Exercise 9.2.6:

A bank accepts rolls of pennies and gives 50 cents credit to a customer without counting the contents. Assume that a roll contains 49 pennies 30 percent of the time, 50 pennies 60 percent of the time, and 51 pennies 10 percent of the time.

- Find the expected value and the variance for the amount that the bank loses on a typical roll.
- Estimate the probability that the bank will lose more than 25 cents in 100 rolls.
- Estimate the probability that the bank will lose exactly 25 cents in 100 rolls.
- Estimate the probability that the bank will lose any money in 100 rolls.
- How many rolls does the bank need to collect to have a 99 percent chance of a net loss?

Exercise 9.2.7:

A surveying instrument makes an error of -2 , -1 , 0 , 1 , or 2 feet with equal probabilities when measuring the height of a 200-foot tower.

- Find the expected value and the variance for the height obtained using this instrument once.
- Estimate the probability that in 18 independent measurements of this tower, the average of the measurements is between 199 and 201, inclusive.

Exercise 9.2.8:

For Example 9.2.9 estimate $P(S_{30} = 0)$. That is, estimate the probability that the errors cancel out and the student's grade point average is correct.

Exercise 9.2.9:

Prove the Law of Large Numbers using the Central Limit Theorem.

Exercise 9.2.10:

Peter and Paul match pennies 10,000 times. Describe briefly what each of the following theorems tells you about Peter's fortune.

- The Law of Large Numbers.
- The Central Limit Theorem.

Exercise 9.2.11:

A tourist in Las Vegas was attracted by a certain gambling game in which the customer stakes 1 dollar on each play; a win then pays the customer 2 dollars plus the return of her stake, although a loss costs her only her stake. Las Vegas insiders, and alert students of probability theory, know that the probability of winning at this game is $1/4$. When driven from the tables by hunger, the tourist had played this game 240 times. Assuming that no near miracles happened, about how much poorer was the tourist upon leaving the casino? What is the probability that she lost no money?

Exercise 9.2.12:

We have seen that, in playing roulette at Monte Carlo (Example [exam 6.7]), betting 1 dollar on red or 1 dollar on 17 amounts to choosing between the distributions

$$m_X = \begin{pmatrix} -1 & -1/2 & 1 \\ 18/37 & 1/37 & 18/37 \end{pmatrix} \quad (9.2.17)$$

or

$$m_X = \begin{pmatrix} -1 & 35 \\ 36/37 & 1/37 \end{pmatrix} \quad (9.2.18)$$

You plan to choose one of these methods and use it to make 100 1-dollar bets using the method chosen. Using the Central Limit Theorem, estimate the probability of winning any money for each of the two games. Compare your estimates with the actual probabilities, which can be shown, from exact calculations, to equal .437 and .509 to three decimal places.

Exercise 9.2.13:

In Example 9.2.9 find the largest value of p that gives probability .954 that the first decimal place is correct.

Exercise 9.2.14:

It has been suggested that Example 9.2.9 is unrealistic, in the sense that the probabilities of errors are too low. Make up your own (reasonable) estimate for the distribution $m(x)$, and determine the probability that a student's grade point average is accurate to within .05. Also determine the probability that it is accurate to within .5.

Exercise 9.2.15:

Find a sequence of uniformly bounded discrete independent random variables $\{X_n\}$ such that the variance of their sum does not tend to ∞ as $n \rightarrow \infty$, and such that their sum is not asymptotically normally distributed.

This page titled [9.2: Central Limit Theorem for Discrete Independent Trials](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.