CHAPTER 1. TWO FACTS ABOUT THE PHYSICAL WORLD.

Consider the following situation in the physical world. We have an experiment that we can carry out. The experiment is of a kind such that: (i) each time it is done, the outcome is uncertain; and (ii) it can be repeated over and over again under the same general experimental conditions. Here are some examples of such experiments.

(1) Stop a person on the street whom you have not met before, and ask the month of his or her birth. Here: (i) you are not sure which of the twelve possible outcomes of the experiment will occur; (ii) you can repeat the experiment many times (if you are in a large enough city).

(2) Toss a coin into the air and see which side is facing up after it falls. (In what follows we shall refer to the two faces of a coin as heads and tails.) There are two outcomes, and we can repeat the experiment as often as we like.

(3) Toss a thumbtack into the air so that it falls on a flat surface, and see whether it lands on its back with the point straight up, or on its side with the point touching the surface.

(4) Roll a die and see which of the numbers 1,2,3,4,5,6 appears on the top of the die.

We are now going to describe some facts about what happens when we repeat such an experiment many times. Each time we repeat an experiment, we call it a trial of the experiment. Let us first define some special words that we shall use. If we have an experiment (such as asking about birth-months), and if there is an outcome that we are interested in (say the outcome July), and if we repeat the experiment n times and look at the results, and if the outcome we are interested in occurs x times in those n trials of the experiment, then we say that the outcome has occurred with relative frequency $\frac{x}{n}$ in those n trials. In other words, the relative frequency of an outcome in a series of trials is the proportion of those trials in which that outcome has been observed to occur. For example, if we repeat the birthday experiment 10 times and get the following answers: March, November, July, April, October, July, August, May, March, February; then the outcome July has occurred with relative frequency $\frac{2}{10} = 0.2$ in those 10 trials, and the outcome January has occurred with relative frequency 0.

Sometimes we are interested in a set of possible outcomes instead of a single outcome. For example, in the birthday experiment, we might be interested in the set of outcomes {June, July, August}. Such a set of possible outcomes is called an event. We get the relative frequency of an event in a given series of trials by looking at all the outcomes that are in the event. In the results given above, we see that the event {June, July, August} has occurred with relative frequency $\frac{3}{10} = 0.3$ in those 10 trials.

What happens to relative frequencies when we repeat an experiment many times? Here are some results from an experiment. We toss a thumbtack and see whether or not it lands on its side.

First we toss it 10 times and find that it lands on its side in 7 of those 10 trials. Thus we have a relative frequency of 0.7

Next, we continue tossing until we have a total of 100 trials. We observe that it falls on its side in 61 of those trials for a relative frequency of 0.61.

Next, we continue on to 1000 trials, and observe that it falls on its side in 582 trials for a relative frequency of 0.582.

Next we continue on to 10,000 trials and observe 5924 for a relative frequency of about 0.592.

Finally, we continue on to 100,000 trials and get 58995 for a relative frequency of about 0.590.

You will note that the figures for the relative frequency, 0.7, 0.61, 0.582, 0.592, 0.590, cluster more and more closely together as we carry out more and more trials. This clustering of relative frequency figures is an observed fact, quite apart from any mathematical or physical theory.

You will nearly always observe this clustering when an experiment is repeated many times. As a general rule, for a fixed outcome (or event) in an experiment, you will find that after about 1000 trials, the values of the relative frequency remain clustered in an interval of length 0.06 or less; that after 10,000 trials, they remain clustered in an interval of length 0.02 or less; and that after 100,000 trials they remain clustered in an interval of length 0.006 or less.

Evidence from many such experiments leads to the following general statement and approximate formula over the range of normally observable values of n ($n < 10^{10}$). Let A be an event for an experiment, and consider repeated trials of the experiment. In nearly every sequence of trials, one will find that for every n, all of the values of the relative frequency (of A) beyond the n^{th} trial are clustered in an interval of length $2.7/\sqrt{n}$ or less. (In fact, for smaller values of n such as $n < 10^5$, the clustering is stronger than indicated by the formula $2.7/\sqrt{n}$, and a formula such as $2/\sqrt{n}$ would appear to be satisfactory. On the other hand, the formula $1/\sqrt{n}$ gives too small an interval.) A more accurate formula, for all n, appears to be $\frac{3}{2}\sqrt{\log\log n}/\sqrt{n}$. In later examples we shall use $2/\sqrt{n}$ when $n \leq 10^5$ and $2.7/\sqrt{n}$ when $10^5 < n \leq 10^{10}$.

This fact about the physical world has been observed over and over again by scientists. It is called the strong stability of relative frequencies. It is also sometimes called the strong empirical law of large numbers or the strong square-root-of-n law.

Here is another example. A coin is tossed 100,000 times. The observed relative frequencies of the outcome heads are as follows:

at	10 trials,	0.5 ,
at	100 trials,	0.54 ,
at	1000 trials,	0.501 ,
at	10,000 trials,	0.498 ,
at	100,000 trials,	0.499 .

The strong stability of relative frequencies is a fact which surprises some people. Whether it surprises one or not, it is a fact about the physical world which one can verify for oneself by choosing some experiment (like rolling a die) and repeating it sufficiently often. In the next section we shall describe the mathematics of probability theory, and we shall see that much of the practical usefulness of this mathematics is based upon the strong stability of relative frequencies.

Let us look in more detail at the thumbtack trials reported above. We can break these figures down in a way which shows another, somewhat different, fact about the stability of relative frequencies.

We analyze the first 1000 trials in groups of 100, the first 10,000 trials in groups of 1000, and all 100,000 trials in groups of 10,000. The figures are as follows:

on side groups of 100 in first 1000	on side groups of 1000 in first 10,000	on side groups of 10,000 in all 100,000
61	582	5824
67	574	6002
62	584	5846
59	571	5870
66	573	5914
58	590	5965
57	584	5851
63	575	5906
56	614	5934
56	577	5883

Note that in the ten different sequences of 100 trials, we get various relative frequencies lying between 0.560 and 0.670; that in the ten different sequences of 1000 trials, we get relative frequencies lying between 0.571 and 0.614; and that in the ten different sequences of 10,000 trials, we get relative frequencies lying between 0.582 and 0.601. These results again clearly show how, for a given experiment, observed relative frequencies tend to cluster together.

For another example, we have the following data for 10,000 tosses of a coin.

heads groups of 100 in first 1000	heads groups of 1000 in first 10,000
54	501
46	485
55	509
53	536
46	485
54	488
41	500
48	497
51	494
53	484

Here, in ten different sequences of 100 trials, the relative frequencies lie between 0.41 and 0.55 for an interval of length 0.14, and in ten different sequences of 1000 trials, the relative frequencies lie between 0.484 and 0.536 for an interval of length 0.052.

This new and different fact about the stability of relative frequencies can be stated as follows. If repeated blocks (of trials) of length n are observed and if the relative frequency for each block is obtained, then nearly all of these relative frequency values will fall in an interval of length $2/\sqrt{n}$. Note that the formula $2/\sqrt{n}$ is again used. [Footnote. In this case (unlike the previous case of a single, long, increasing sequence of trials), empirical evidence shows that the formula $2/\sqrt{n}$ applies equally well for all values of n.] In the above example, where we observed the interval 0.14 for ten groups of trials of length 100, the formula $2/\sqrt{n} = 2/\sqrt{100}$ gives the enclosing interval 0.2; and where we observed 0.052 for ten groups of trials of length 1000, the formula gives 0.063.

This second fact about the physical world has been observed over and over again by scientists. It is called the weak stability of relative frequencies. It is also sometimes called the weak empirical law of large numbers or the weak square-root-of-n law. We see that the appearance of similar formulas ($2/\sqrt{n}$, $2.7/\sqrt{n}$) in the statements above of both strong and weak empirical stability is, to some extent, a convenient accident.

Remark. Probability theory (to be introduced in Chapter 2) has other important uses besides the study of experiments that can be repeated many times. For example, probability can be used to describe how strongly a person holds a certain belief. We shall later (Chapter 19) consider such other uses, but, for now, we limit ourselves to the kind of repeatable experiment described above. As we shall see, most of the other uses are closely related to the uses which we are considering now.

Predicting variation. The formula $2/\sqrt{n}$ is often useful in predicting the amount of variation that can be expected to occur in a final quantity that is a sum of smaller quantities in a probabilistic experiment. Consider the following two examples.

First example. Every week night, 200 trucks leave the downtown area of a city to travel to other cities. There are only two routes for trucks out of the city. One route, quick and easy, is across a toll bridge which charges $10 per truck. The other route, across a free bridge, is considerably slower and less convenient. It is the custom for each truck driver, as he sets out, to toss a coin. If the coin comes up heads, he takes the toll bridge. If the coin comes up tails, he takes the free bridge. The average total of truck tolls per week night is, as one would expect, 100($10) = $1000. What amount of variation would one expect to see in the totals of truck tolls received each night? Here, the weak square-root-of-n law tells us that nearly all the relative frequencies of choosing the toll bridge should lie in an interval of length $2/\sqrt{200}$. Hence the numbers of trucks crossing the toll bridge per week night should nearly always lie in an interval of length $\frac{2}{\sqrt{200}} \cdot 200 = 29$ around the average value of 100. Hence the totals of money received from trucks each week night should nearly always lie in an interval of length $290 around $1000.

Second example. Every week night, 20,000 cars leave the downtown area of the same city to travel to the suburbs. The only routes available to cars are the same two routes as for trucks: a free bridge and a toll bridge. The toll bridge charges 10¢ per car. Like the truck drivers, the car drivers choose their routes by tossing a coin. The average total of car tolls, per week night is, as one would expect, 10,000($0.10) = $1000. What amount of variation would one expect to see in the totals of car tolls received each night? As before, we use the weak square-root-of-n law. Relative frequencies of choosing the toll bridge should lie in an interval of length $2/\sqrt{20,000}$. Hence the number of cars crossing the toll bridge per week night should nearly always lie in an interval of length $\frac{2}{\sqrt{20,000}} \cdot 20,000 = 283$ around the average value of 10,000. Hence the totals of money received from cars each week night should nearly always lie in an interval of length $28.30 around $1000. We see that the total of car tolls is more stable than the total of truck tolls, because it depends upon a larger number of trials of an individual experiment.

The above examples, with decisions by coin tossing, are somewhat artificial. We shall look at more realistic versions of these examples in Chapter 7.

Further analysis of stability. In the discussion above, we have stated two facts about the stability of relative frequencies. Both facts use the formula $2/\sqrt{n}$ (for $n \leq 10^5$). We now look more carefully at the statements of these two facts, and we introduce some further terminology.

The first fact can be given as follows. When an experiment is repeated N times under the same general conditions (in which the outcome of each trial does not appear to be influenced by the specific outcomes of previous trials--we sometimes speak of this as "doing N successive <u>independent trials</u> of the experiment"), when a particular event A is chosen ahead of time, and when, for each $m \leq N$, we calculate the relative frequency f_m of A in the first m trials, then we nearly always find that for each $n \leq N$, the set of values f_m, $n \leq m \leq N$, all lie in some interval of length $2.7/\sqrt{n}$ or less. Except for the words "nearly always", this is a precise statement of the strong stability law. We shall usually think of N as taken to be quite large.

The second fact about stability can be stated as follows. For each n, if we repeatedly carry out blocks of n independent trials of the experiment and observe the relative frequency of A in each block, then nearly all of the observed relative frequencies lie in an interval of length $2/\sqrt{n}$. Again, except for the words "nearly all", this is a precise statement of the weak stability law. The following empirical connection between weak and strong stability can also be observed and can be stated as part of the weak law. If λ is the limiting value to which the relative frequencies converge (under the strong law), then the particular interval $(\lambda - 1/\sqrt{n}, \lambda + 1/\sqrt{n})$ may be used (for each n) as the interval of length $2/\sqrt{n}$ for the weak law.

In the discussion above about tossing a thumbtack and tossing a coin, we used the observed data to illustrate both strong and weak stability. (In the case of strong stability, we had $N = 100{,}000$.) A somewhat stronger version of the weak stability law will be given in Chapter 6.

What more precise meanings can be given to the words "nearly always" and "nearly all" in these laws? If we fix n, and make extensive observations of blocks of n independent trials for the weak law, we find empirically that "nearly all" means "at least 95%", provided that $n \geq 9$. If we make extensive observations for the strong law by looking at many different individual long sequences of independent trials, we find empirically that "nearly always" means "for more than 99% of those observed long sequences" when the sequences have $N \approx 10^5$ total trials, and "for more than 95% of those observed long sequences" when the sequences have $N \approx 10^{10}$ total trials.

Theoretical note. The laws of weak and strong stability are empirical laws. That is to say, they are facts about nature that one can observe directly, quite apart from any mathematical theory. Once we have set up a mathematical theory of probability, however, we can attempt to deduce these laws from the assumptions of our theory. We then find that it is possible to deduce the weak stability law. It becomes a special case of a more general theorem known as the weak law of large numbers. (The special case will be deduced in Chapter 6, the general theorem in Chapter 16.)

With regard to the strong stability law, the reader with a knowledge of limits will note that the strong law implies that, if we imagine a series of trials to be carried out without end, then, nearly always, the observed sequence of relative frequencies will approach a limiting value in the strict mathematical sense of limit. [Footnote. A proof can be given which is subtle but correct. It uses the Cauchy criterion for convergence of a sequence.] The statement that this infinite sequence of observed values almost always approaches a limit can also be deduced from the assumptions of probability theory. It is a special case of a more general theorem known as the <u>strong law of large numbers</u>. The specific empirical statement of the strong stability law, given above, using the approximate formula $2.7/\sqrt{n}$ (for $n \leq 10^{10}$), however, has not yet been deduced from theory.

To get a theoretical statement that has been proved for <u>all</u> n, the formula $2.7/\sqrt{n}$ must be modified to $4\sqrt{\log \log n}/\sqrt{n}$ where <u>log</u> means natural logarithm.

Note that the factor $\sqrt{\log (\log n)}$ grows extremely slowly. It only reaches the value 2 at about $n = 10^{10}$.

[Footnote. A reader already familiar with advanced probability theory will note that the statements usually given

and for the strong law of large numbers, while elegant, are non-constructive in that they refer to the completed totality of an infinite sequence of trials. An example of such a statement is the assertion that, almost always, $\lim_{n\to\infty} f_n = 1/2$ where f_n is the relative frequency of heads at the n^{th} toss in a single, infinite sequence of tosses of a coin. This statement only has meaning if we conceive of the entire sequence as completed. Such statements cannot be tested against observation. It is only by examining and analyzing the constructive content of the proofs of these statements, that one finds (as above) versions of the theorems that can be tested against observation. Analysis of proofs for their constructive content is an important concern of mathematical logic. The usual non-constructive statements of these laws have the advantages of intellectual economy and of clarity and simplicity for further theoretical use. They have the disadvantages of having no direct operational meaning and of omitting information implicit in their proofs.]

The reasons for using the words "strong" and "weak" in connection with the stability laws are indicated by the following remark. The empirical fact of strong stability with the formula $2/\sqrt{n}$ logically implies (as is easily verified) the empirical fact of weak stability, provided that, in the statement of weak stability, we alter the formula $2/\sqrt{n}$ to $4/\sqrt{n}$. On the other hand, there is no simple argument to show that empirical weak stability directly implies some form of empirical strong stability. (Indeed, as we have already noted above,

strong stability with the formula $2/\sqrt{n}$ becomes incorrect for $n > 10^5$, while weak stability remains true.)

<u>Examples</u>. Figure 1.1 is a record of $N = 100$ tosses of a coin. For each n, the irregular curve gives f_n, the relative frequency of <u>heads</u> in the first n trials. The value of $2/\sqrt{n}$ is given, for each n, by the distance between the two smooth curves. The two places where the relative frequency curve falls below the lower smooth curve are not violations of strong stability, since strong stability only requires that all later frequencies cluster in <u>some</u> interval of length $2/\sqrt{n}$ (not necessarily the specific interval symmetrical about 0.5 given by the two smooth curves.)

Figure 1.2 is a record of 5 blocks of 35 tosses each of a coin. The same smooth curves are drawn as before. If we look at some particular value of n (for example $n = 35$), this figure illustrates weak stability. If we were to record more blocks of 35 tosses, we would eventually find relative frequencies falling outside (both above and below) the smooth curves at various values of n, but, for each value of $n \geq 9$, more than 95% of the blocks (like all five of the curves at $n = 35$ in the figure), would fall between the two smooth curves.

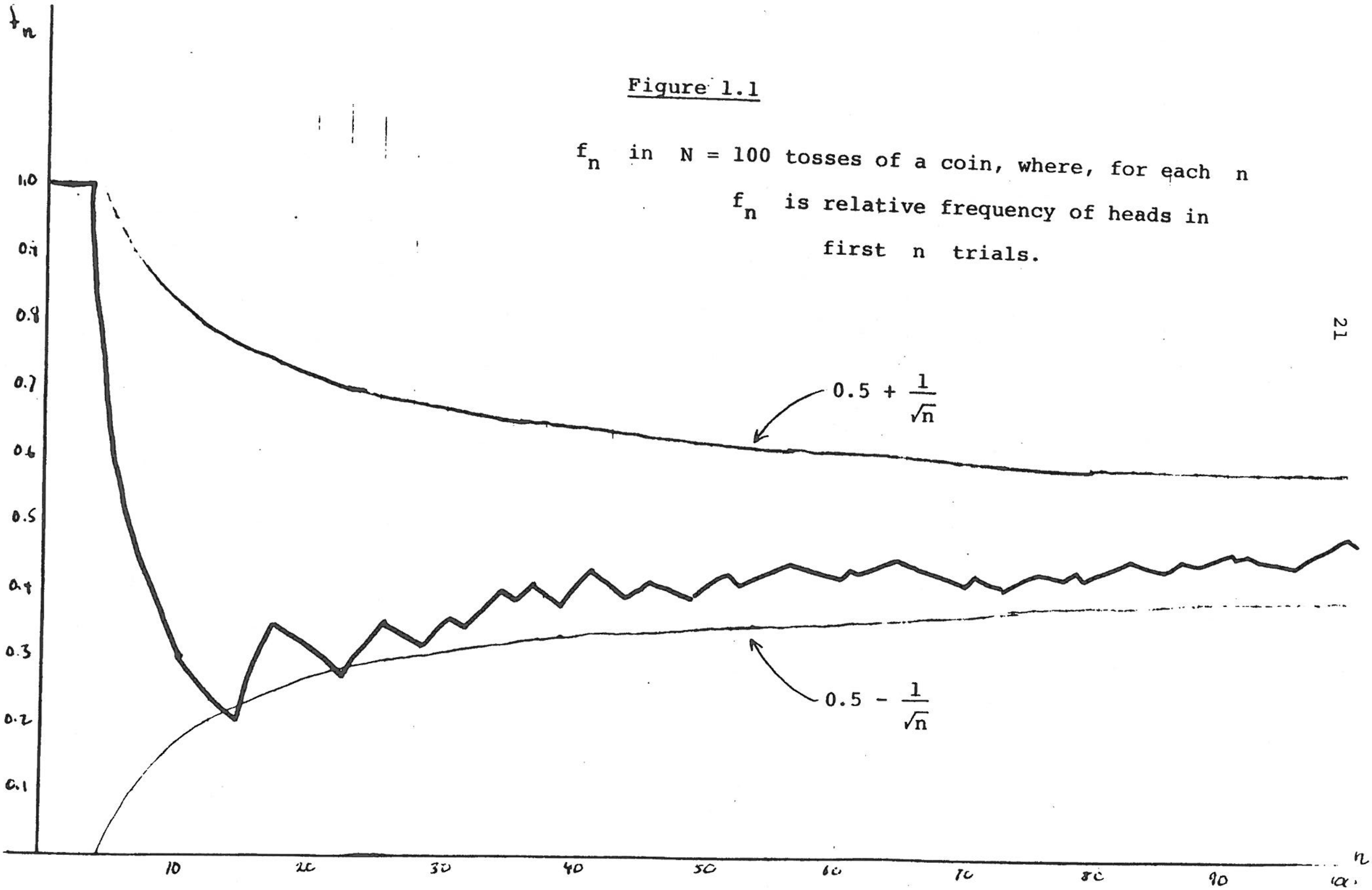

Figure 1.1

f_n in $N = 100$ tosses of a coin, where, for each n f_n is relative frequency of heads in first n trials.

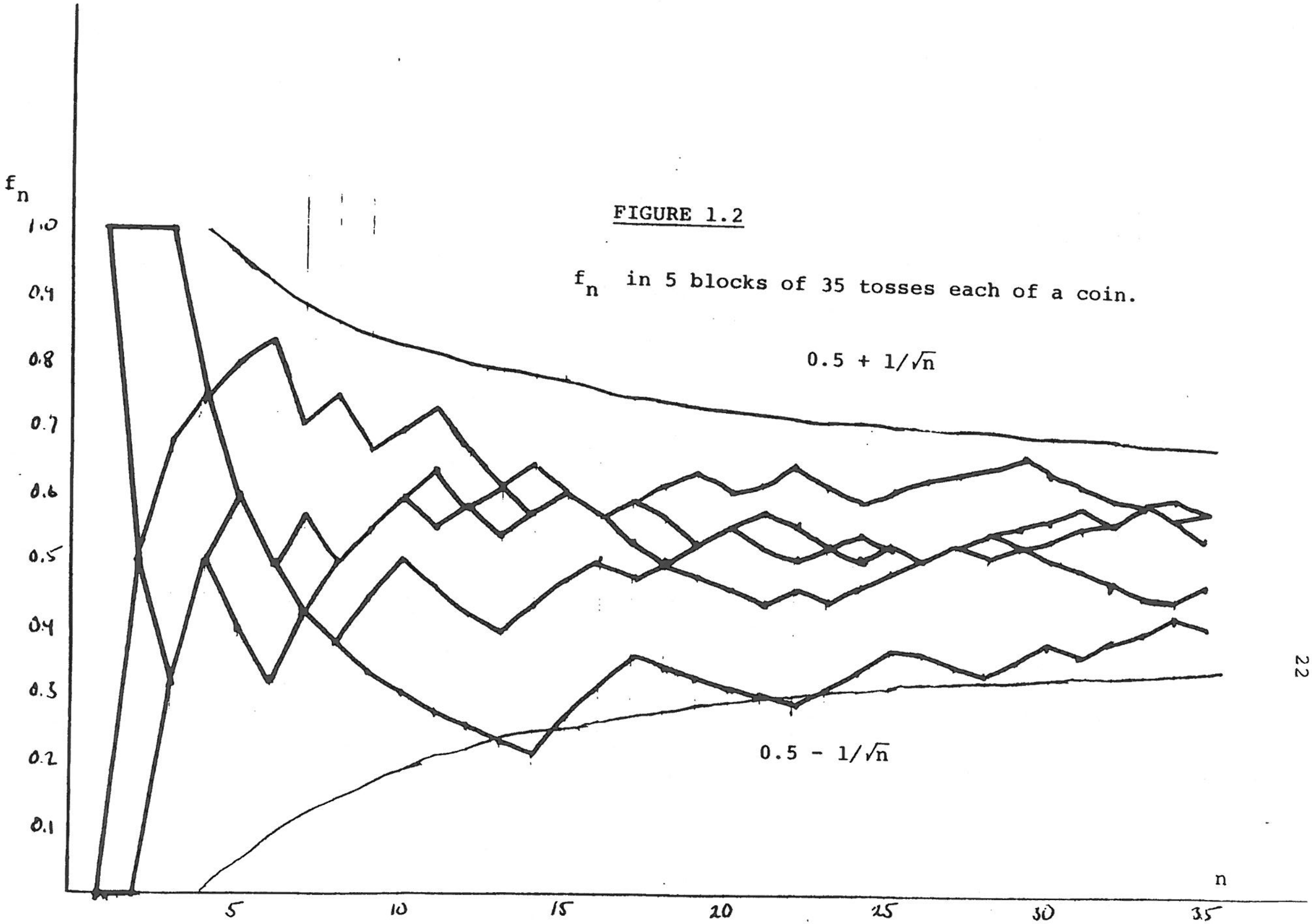

FIGURE 1.2

f_n in 5 blocks of 35 tosses each of a coin.

Sources of data. The results stated earlier in this chapter for 100,000 tosses of a thumbtack do not come from an actual experiment with a thumbtack, but are taken from a different physical experiment which was used to simulate the experiment of tossing a thumbtack. (See Chapter 8.) The same is true for the results given above for tossing a coin. [Footnote. The observed coin tossing data are taken from Chapter I of An Introduction to Probability Theory and its Applications, volume 1, by William Feller (3rd edition, 1968, John Wiley and Sons.) Other aspects of these data are described and discussed in Chapter III of that book, and a graph of the data is given on page 87 of that book. We consider certain further features of such data in Chapter 8 below.]

The law of averages. The profits of gambling casinos are based upon the stability of relative frequencies. Assume that a gambler in a casino is playing a particular game (say roulette [Footnote. The game of roulette is outlined in the exercises below.]) and that on each successive play of the game, the gambler bets the same amount of money, ℓ, on the event A. On any play, if A occurs, the gambler wins w dollars, and if A does not occur, the gambler loses the ℓ dollars that were bet. (In roulette, if A is the occurrence of a certain number and if $\ell = 1$, then $w = 35$.) Let x be the number of times the gambler wins in the first n plays. Then, after n plays, the gambler has won $wx - \ell(n-x)$ dollars. (Here, a negative quantity represents a loss to the gambler.) Thus, after n plays, the gambler's average winnings per play will be

$$\frac{1}{n}(wx-\ell(n-x)) = w\frac{x}{n} - \ell(1 - \frac{x}{n})$$

$$= wf_n - \ell(1 - f_n)$$

where f_n is the relative frequency of wins in the first n plays. The gambler will be ahead after n plays if $wf_n - \ell(1-f_n) > 0$; that is to say, if

$$f_n > \frac{\ell}{w+\ell} .$$

(Thus, in the case of roulette, the gambler will be ahead if $f_n > \frac{1}{35+1} = \frac{1}{36}$.)

Now in honestly-run gambling games, repeated plays occur under the same general conditions. Hence by the strong stability of relative frequencies, we expect the values of f_n to become and remain more and more closely clustered about some limiting value λ as n increases. If $\lambda > \frac{\ell}{w+\ell}$, it follows that (nearly always) the gambler's average winnings per play will eventually become and remain positive. We hence say that the bet is a <u>favorable</u> one for the gambler if $\lambda > \frac{\ell}{w+\ell}$. If $\lambda < \frac{\ell}{w+\ell}$, it follows that (nearly always) the average winnings per play will become and remain negative. We hence say that the bet is <u>unfavorable</u> if $\lambda < \frac{\ell}{w+\ell}$. In the special case where $\lambda = \frac{\ell}{w+\ell}$, we say that the bet is <u>fair</u>. For repeated bets on a single number at roulette, it is a matter of empirical fact (in the game as played in North America) that with a carefully balanced and operated roulette wheel $\lambda \approx 1/38$. (In Europe, roulette wheels have 37 rather than 38 compartments, and $\lambda \approx \frac{1}{37}$.) Thus these roulette bets are <u>unfavorable</u> for our bettor, and we can expect the bettor's average winnings per game eventually to become and remain negative.

How long will it take for losses to occur? The stability laws give us ways to answer this question. We consider, first, strong stability. <u>Permanent</u> loss will occur as soon as f_n becomes and <u>remains</u> below $\frac{\ell}{w+\ell}$. That is to say, permanent loss will occur as soon as f_n becomes and remains within the distance $\frac{\ell}{w+\ell} - \lambda$ of the limiting value λ. In the case of bets on a single number at roulette,

$$\frac{\ell}{w+\ell} - \lambda = \frac{1}{36} - \frac{1}{38} = \frac{2}{36\cdot 38} \approx 0.0015.$$

Now the interval given by the strong stability law must always include λ. Hence, to be sure that f_n becomes and remains within 0.0015 of λ, we must, by the strong stability law, have

$$2.7/\sqrt{n} \leq 0.0015.$$

This gives $n \geq \frac{7.3}{(0.0015)^2} \approx 3.2 \times 10^6$. Hence we expect that permanent loss will be inevitable after 3,200,000 plays of the game. [<u>Footnote</u>. We set $2.7/\sqrt{n} \leq 0.0015$ rather than $1.35/\sqrt{n} \leq 0.0015$, because the strong stability law, as stated above, only guarantees that the relative frequency will remain within <u>some</u> interval of length $2.7/\sqrt{n}$. It does not imply that this interval must be the interval symmetric about λ.

In the present range of values, the formula $2.7/\sqrt{n}$ gives a valid but crude estimate. In fact, on the basis of observation, permanent losses appear to occur sooner than this.]

The strong stability of relative frequencies, as it applies to such betting situations, is popularly referred to as the law of averages. We have made several special assumptions in our discussion. First, we have assumed that each bet is the same size. Second, we have assumed that each bet is on the same event. The assumption that the size of the bet is constant is important. If the bettor is allowed to vary the size of the bet, without restriction, from play to play, then it is no longer possible to conclude, from the strong stability of relative frequencies, that losses are inevitable. The total available resources of the gambler (and of the casino) may become an important factor in the situation, and we must take account, in our analysis, of the possibility that the gambler (or the casino) can be ruined (that is to say, bankrupted) as a result of a single large bet. We consider this further in Chapter 8 . In order to assure that the law of averages will apply (and hence that an inevitable profit will occur), casinos normally place a bound on the size of individual bets. (It is not hard to show that the analysis for a constant size bet can also be carried through for bets of varying but bounded size.) The placing of a maximum bound rules out, for example, the system sometimes known as "doubling" or "martingale", in which, on each bet, the gambler bets an amount that will, if he or she wins, recoup the total of all losses so far.

The assumption that each bet is on the same event turns out to be less critical. If the gambler systematically changes his or her bet from event to event on successive plays, then, so long as each bet is unfavorable and so long as there is a maximum bound on bet size, eventual losses will be inevitable and will occur at or before the time predicted by the strong stability law. This conclusion (that changing the event will not help the bettor) rests on certain empirical facts about the irregularity of observed results in repeated plays. These facts go somewhat deeper than the simple statement of the strong stability law. We do not discuss these facts here.

Provided that there is no collusion among bettors (to get around the maximum bound on bet size), a casino can view each bet placed as a bet on a separate play of the game against that bettor. Hence, against bettors who are betting the same amount of money on individual numbers at roulette, a casino will be assured of a continuing profit as soon as 3,200,000 bets have been placed. [Footnote. From observation, as noted before, profits appear to be certain after many fewer bets have been made. It should also be noted that certain other standard casino games, such as chuck-a-luck (to be described in Chapter 8), give the casino a bigger advantage than does roulette.]

So far we have considered the implications of strong stability for the bettor and the casino. Weak stability can also be used to give information. We again assume bets of constant size. Weak stability tells us, for any single value of $n \geq 9$, that f_n is almost certain to lie in the interval $(\lambda-1/\sqrt{n},\lambda+1/\sqrt{n})$. Hence, for a given n such that $1/\sqrt{n} < \frac{\ell}{w+\ell} - \lambda$, the gambler is almost certain to have a net loss. This implies

$$n > \frac{1}{(\frac{\ell}{w+\ell} - \lambda)^2} .$$

In the case of roulette, we have

$$n > \frac{1}{(\frac{1}{36} - \frac{1}{38})^2} ,$$

or, approximately, $n > 450,000$. The difference between this weak law analysis and the previous strong law analysis is that the weak law tells us that the bettor is almost certain to be behind for a given $n > 450,000$, while the strong law analysis tells us that the bettor is almost certain to become and remain behind after $n \approx 3,200,000$.

If a bet is especially unfavorable, but the unfavorable nature of the bet is not immediately evident, the bet is sometimes called a sucker bet. An example is the following. The bettor bets even money ($\ell = w$) that at least one six will appear in three rolls of a single die. Observation shows that $\lambda \approx 0.42$ in this case. Since $\frac{\ell}{w+\ell} = 0.5$, weak law analysis tells us that for any given value of $n > \frac{1}{(0.5-0.42)^2} \approx 156$ plays, the better is almost certain to be behind.

The formula

$$n > \frac{1}{(\lambda_0-\lambda)^2} ,$$

where λ is the observed limiting frequency of winning and where $\lambda_0 = \frac{\ell}{w+\ell}$ is the limiting relative frequency for which the bet would be a fair bet, is a useful one in considering unfavorable bets. (In using this formula, we must always take $n \geq 9$, since the weak law requires $n \geq 9$.) For example, in the dice game knows as <u>craps</u>, the bettor rolls the dice and bets even money that an event called <u>passing</u> will occur. In this case, λ is observed to be ≈ 0.493. (The game of craps is described in an exercise.) Hence the formula tells us that the bettor is almost certain to have a loss for any given n such that $n > \frac{1}{(0.5-0.493)^2} \approx 20{,}000$.

In Chapter 6, we shall develop a stronger, but still empirically correct, weak stability formula. The stronger formula says that, for values of n larger than both $9(\frac{1-\lambda}{\lambda})$ and $9(\frac{\lambda}{1-\lambda})$, the bettor is almost certain to be behind for any given value of n such that

$$n > \frac{4(\lambda-\lambda^2)}{(\lambda_0-\lambda)^2} .$$

This improved formula gives $n > 46{,}000$ for bets on a single number at roulette, $n > 152$ for the sucker bet with a die described above, and $n > 20{,}000$, as before, for craps. (For $\lambda = 1/2$, the improved formula reduces to the previous formula.)

The bettor's eventual average loss per play at craps is

$$\ell(1-\lambda) - w\lambda = 1(0.507) - 1(0.493) = 0.014,$$

while the eventual average loss per play at roulette is

$$1\left(\frac{37}{38}\right) - 35\left(\frac{1}{38}\right) = \frac{2}{38} = 0.053.$$

The reader will note an apparent paradox. Although the disadvantage at roulette, as measured by this average loss, is almost four times as great as at craps, the losses at craps will occur twice as soon as the losses at roulette, as measured by the values of n calculated above from the weak law. In Chapter 6, we shall further explore this interesting feature of the law of averages, namely that in some cases a game with smaller average loss may lead to earlier certain loss for the bettor. We shall there verify, for example, that if we consider a large number of bettors who have each made 10,000 successive bets (of the same amount on a single number) at roulette and a large number of bettors who have each made 10,000 successive bets (of the same amount on passing) at craps, we can expect about 16% of the roulette bettors to have come out ahead, but only about 8% of the craps bettors to have come out ahead. The explanation of this puzzling result is that although there are fewer losers at roulette, those who _have_ lost will have lost much more, on the average, than the losers at craps.

Comment. In Chapters 2, 3, and 4 on probability spaces and combinatorial methods, we shall learn, among other things, how to calculate the limits of certain relative frequencies on the basis of certain other empirical facts. For example, we shall calculate that the limiting observed relative frequency for winning the sucker bet above is 0.42, assuming that the limiting observed relative frequency for each number on a die is 1/6. Extensive compilations of empirically observed relative frequencies for a variety of experiments and events also exist. One such is The World Book of Odds, by Neft, Cohen, and Deutsch (Grosset and Dunlap, New York, 1978.) [Footnote. In this reference, relative frequencies are given in the form of true odds. The true odds against an event are the ratio that w should have to ℓ in order for a bet on that event to be fair. Thus true odds of m to n express a relative frequency of $\frac{n}{m+n}$.] In Chapter 8, we shall return to the subject of bets. We shall consider more complex forms of bet and learn to calculate expected winnings or losses in a less crude and more informative way. In Chapter 19, we shall consider bets in non-repeatable experimental situations.

The weak stability law in physics. The weak stability law plays an important role in physics, where it is used to provide a connection between apparently random microscopic events such as molecular positions, speeds, and collisions, and such stable macroscopic quantities as density, temperature, and pressure. For this purpose, a special form of the weak square-root-of-n law will be developed and described in Chapter 7.

Final note. Some readers may feel that our convention of using "nearly all" to mean "more than 95%" and of using "nearly always" to mean "more than 95% of the time" is unsatisfactory. They would argue that, in practical terms, 5% is a far from negligible fraction. The criticism is not a deep one, however.

If we use $3/\sqrt{n}$ in place of $2/\sqrt{n}$, observation shows that "nearly all", in the statement of weak stability, means "more than 99.5%". If we use $4/\sqrt{n}$ in place of $2/\sqrt{n}$, "nearly all" means "more than 99.995%". Similar changes occur for the statement of strong stability. In each case, the discussion and analysis in the text (for example, the analysis of casino profit) can be modified to agree with the new meanings for "nearly all" and "nearly always".

EXERCISES FOR CHAPTER 1.

1-1. Consider the following record of 100 rolls of a single die:

46363	21262	51531	32555	13534	31333	54644
34612	21436	31631	34113	42613	45126	66516
56246	41344	35523	25645	63561	55533.	

(a) Find the relative frequency, in these 100 trials, of each of the following events:

{1}, {2}, {3}, {4}, {5}, {6}, {5,6}, {4,5,6}.

(b) Let f_n be the relative frequency of an event after n trials. For the given trials, find the values of $f_5, f_{10}, f_{15}, \ldots, f_{100}$ for the event {4,5,6} and make a plot similar to Figure 1.1. Does this plot agree with strong stability?

(c) Do the same as (b) for the event {5,6}.

(d) Take the above data as 10 blocks of 10 trials each. For each block, find the relative frequency for the event {4,5,6}, and see if these results agree with weak stability.

(e) Take the above data as 5 blocks of 20 trials each. Do the same as in (d).

1-2. Consider the toll bridge and the free bridge as described in the text under "predicting variation". Each week night, 200,000 pedestrians walk home across these bridges. Each pedestrian tosses a coin to decide which bridge to cross. The toll bridge charges each pedestrian one cent. What average total of pedestrian tolls per week night would one expect, and what amount of variation would one expect to see in the total of pedestrian tolls received each night?

1-3. Assume the strong stability law with the formula $2/\sqrt{n}$. Show that weak stability must logically follow, provided that we are allowed to use the formula $4/\sqrt{n}$ (in place of the formula $2/\sqrt{n}$) in the statement of weak stability.

1-4. In North America, the game of roulette is played as follows. A metal disk (the wheel) is mounted on a vertical axis. On the circumference of the wheel are 38 open box-shaped compartments. The compartments are labeled with the numbers 0,00,1,2,3,...,36. The compartments 0 and 00 are colored green. Half of the remaining compartments are red and half are black. Immediately adjacent to the circumference of the wheel is a fixed, inward-sloping annular surface. On each play of the game, the wheel is set in rotation and a small metal ball is set in motion in the opposite direction on the annular surface. Eventually, the ball comes to rest in one of the compartments on the turning wheel. Bettors can bet on various different events for each play of the game by placing money (or chips representing money) in an appropriate position on a large diagram on which the possible bets are indicated. (This diagram is marked on the surface of the roulette table.) Bets may be placed on individual numbers, on certain pairs of numbers (with $(\ell = 1,\ w = 17)$, on certain sets of three numbers $(\ell = 1,\ w = 11)$, of four numbers $(\ell = 1,\ w = 8)$, of six numbers ($\ell = 1$, $w = 5$), of twelve numbers ($\ell = 1$, $w = 2$), and of eighteen numbers ($\ell = 1$, $w = 1$). A popular form of bet is on the set of eighteen red numbers or on the set of eighteen black numbers.

There are two differences in European roulette. First, there are only 37 compartments (there is no compartment labeled 00). Second, bets on red or black are treated in a special way. If a bettor bets on red and red appears, the bettor wins (with $\ell = 1$, $w = 1$); if black appears, the bettor loses; but if 0 appears, the bettor does not immediately lose. Instead, the bet is said to be "in prison", and the bettor must then wait until red or black appears on a later play. If red appears first, the original wager ($\ell = 1$) is released and returned to the bettor (but with no winnings added). If black appears first, the bettor loses and the original wager is turned over to the casino. (We shall learn, in Chapter 16, how to analyze a bet on red in European roulette.)

There is also a third form of roulette played in Central America. This roulette differs from North American roulette in that there are 39 compartments (the additional compartment is labeled with an eagle.)

There are a variety of popular systems for betting the colors at roulette. All are variants, in some form, of the martingale system, which is to bet, successively, the amounts 1,2,4,8,16,..., until a win occurs, and then to start over. Application of the martingale system is limited by whatever bound the casino may place on the size of a bet (and by the available resources of the bettor). Some of these systems will be considered further in Chapter 16.

(a) Give a strong stability analysis (like that in the text for North American roulette) for bets on a single number at European roulette. Do the same for Central American roulette. (Assume $\lambda \approx 1/37$ for European roulette and $\lambda \approx 1/39$ for Central American roulette.)

(b) Give a weak stability analysis for the same cases (using the improved formula $\frac{4(\lambda-\lambda^2)}{(\lambda_0-\lambda)^2}$ as in the text.)

(c) Consider a bet on the <u>odd</u> numbers ($\ell = 1$, $w = 1$). Compare the eventual average loss per play for bets on <u>odd</u> with the eventual average loss per play for bets on a single number.

(d) For bets on <u>odd</u>, carry out a strong stability analysis for European roulette, for North American roulette, and for Central American roulette. (Assume $\lambda \approx 18/37$, $\lambda \approx 18/38$, and $\lambda \approx 18/39$ respectively.)

(e) Do the same as (d) with a weak stability analysis using the improved formula. [<u>Footnote</u>. In Chapter 6 we shall see that of a large number of bettors who have each made 10,000 successive bets on <u>odd</u>, we would expect about 0.3% of the bettors at European roulette to have come out ahead and none of the bettors at the other two roulettes to have come out ahead.]

1-5. The game of craps is played as follows. An individual known as the shooter makes a series of rolls with two dice. The game ends when either the event pass or the event don't-pass occurs. These events are defined as follows. Numbers refer to the sum appearing on the two dice on a single roll.

(i) 7 or 11 on the first roll. Pass occurs and the game ends.

(ii) 2, 3, or 12 on the first roll. Don't-pass occurs and the game ends.

(iii) One of the six numbers 4, 5, 6, 8, 9, 10 occurs on the first roll. This number is now called the shooter's point The shooter continues to roll until either the point appears again or 7 appears. If the point appears first, pass occurs and the game ends. The shooter is said to have made the point. If 7 appears first, don't-pass occurs and the game ends.

In a casino, a variety of bets against the casino may be made by a gambler as a craps game progresses. The gambler's winnings for each available bet are established by the rules of the casino. (These rules may vary substantially from casino to casino.) For example, if the shooter rolls the point 6 on the first roll of a game, a gambler may then bet on pass. For this bet with $\ell = 1$, some casinos offer $w = 1$, some offer $w = 1.1$, and some offer $w = 1.16$.

(a) For bets on making the point 6, λ is observed to be ≈ 0.455. Carry out a weak stability analysis, using the improved formula, for each of the three casino pay-off rules described above. What is the eventual average loss per bet in each case?

(b) For bets on making the point 4, λ is observed to be 0.333. Casinos usually pay $w = 1.8$ for $\ell = 1$. Carry out a weak law analysis using the improved formula. What is the eventual average loss per bet?

[Footnote. The pay-off rule adopted by a casino on a particular bet is often stated in the form of odds offered by the casino. "Odds of m to n" means that $w = \frac{m}{n}$ when $\ell = 1$. Odds are usually stated in terms of whole numbers. Thus the pay-off $w = 1.8$ for making the point 4 can be described as odds of 9 to 5, and the pay-off $w = 1.16$ for making the point 6 can be described as odds of 29 to 25. Often, in connection with craps, casinos use the word "for" rather than "to" in stating their pay-off rules. A pay-off of "m for n" then means the same as odds of $m - n$ to n. Thus, for example, "3 for 1" is the same as odds of 2 to 1.]

1-6. Consider the following sucker bet. The bettor bets even money that a six will not occur in six rolls of one die. Here, observation shows that $\lambda \approx 0.335$.

(a) Apply a weak stability analysis.

(b) We refer to the bettor's opponent in a sucker bet as the hustler. How many plays of this game (with $\ell = 1$) does the hustler need in order to be sure (in the sense of "almost always") of a continuing and permanent net profit?

(With many sucker bets, an incorrect but somewhat plausible argument can be used to justify the bet. In the present case, one might argue: "From stability of relative frequencies, I expect a six to occur on the average once in every six rolls. Sometimes it will occur more than once in six rolls. Sometimes it will not occur at all. Therefore, I expect it not to occur about half the time.")

SPECIAL PROJECT I

Estimate the percentage of private automobiles in your community area which are white. Your report should include:

1. Your estimate.
2. A description of your procedure and calculation.
3. A statement about the reliability of your estimate.
4. A critique of your procedure.
5. A critique of the formulation of the original problem and a description of the decisions you have made to clarify it and to make it more precise.

(You are not expected to know or to discover a single right way to do this project. Make your best common-sense effort to get what seems to be a good estimate.)