Chapter 17

Thinking about Chance

Lecture Slides

Case Study: Thinking about Chance

On February 29, 2012, a woman in Provo, Utah, gave birth on a third consecutive Leap Day, tying a record set in the 1960s.

If birth dates are random and independent, a statistician can show that the chance that three children, selected at random, are all born on Leap Day is about 1 in 3 billion.

By the end of this chapter, you will be able to assess coincidences such as having three children born on Leap Day. Are these events as surprising as they seem?

The Idea of Probability 1

Start by thinking about what would happen if we did this many times.

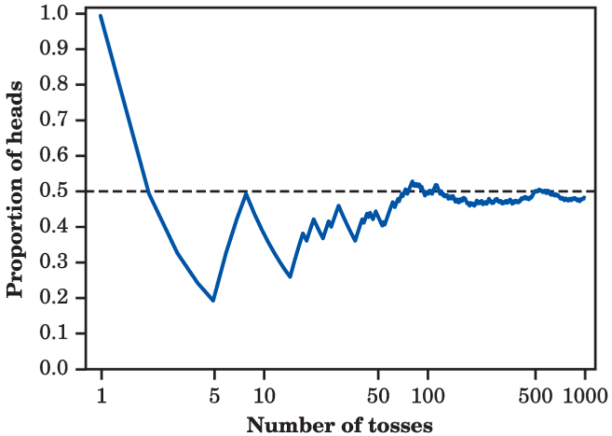
A big fact emerges when we watch coin tosses or the results of random samples closely:

Chance behavior is unpredictable in the short run but has a regular and predictable pattern in the long run.

Example: Coin Tossing

Figure 17.1 shows the proportion of “heads” when tossing a coin 1000 times.

The proportion of tosses that produce “heads” is quite variable at first, but it settles down as we make more and more tosses.



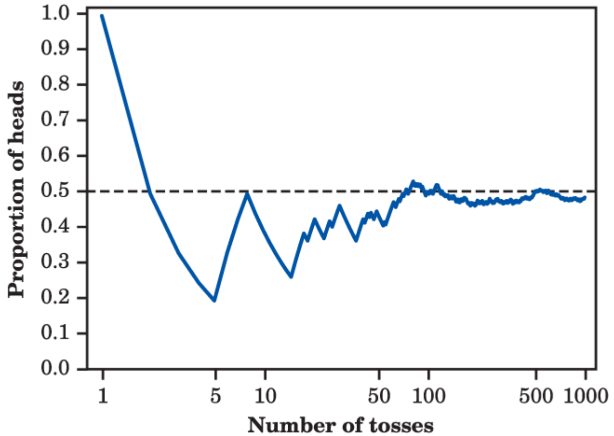
Moore/Notz, *Statistics: Concepts and Controversies*, 10e,

Example: Coin Tossing (continued)

Eventually, this proportion gets close to 0.5 and stays there.

We say that 0.5 is the probability of a head.

The probability 0.5 appears as a horizontal line on the graph.



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020 W. H. Freeman and Company

The Idea of Probability 2

We call a phenomenon **random** if individual outcomes are uncertain, but there is, nonetheless, a regular distribution of outcomes in a large number of repetitions.

The **probability** of any outcome of a random phenomenon is a number between 0 and 1 that describes the proportion of times the outcome would occur in a very long series of repetitions.

The Idea of Probability 3

An outcome with probability 0 never occurs.

An outcome with probability 1 happens on every repetition.

An outcome with probability 1/2 happens half the time in a very long series of trials.

Probability gives us a language to describe the long-term regularity of random behavior.

Example: The Probability of Dying

In 2013, the National Center for Health Statistics reported that the proportion of men aged 20 to 24 years who die in any one year is 0.0012 and the proportion for women is 0.0004.

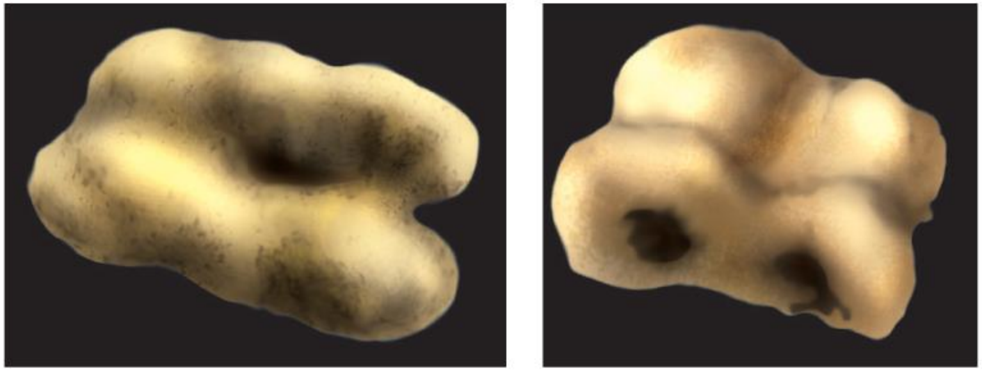
An insurance company knows that it will have to pay out next year on about 0.12% of the policies sold on men's lives and on about 0.04% of the policies sold on women's lives.

It will charge more to insure a man because the probability of having to pay is higher.

The Ancient History of Chance 1

The most common method of randomization in ancient times was “rolling the bones”—that is, tossing several astragali.

The astragalus (Figure 17.2) is a six-sided animal heel bone that, when thrown, will come to rest on one of four sides (the other two sides are rounded).



Sheep Dog

Moore/Notz, *Statistics: Concepts and Controversies*, 10e,

The Ancient History of Chance 2

Modern dice are made so they are all similar and should show the same pattern of outcomes when rolled.

Professional gamblers noticed the regular pattern of outcomes of dice or cards and tried to adjust their bets to the odds of success.

“How should I bet?” is the question that launched mathematical probability.

The Ancient History of Chance 3

The systematic study of randomness began when seventeenth-century French gamblers asked French mathematicians for help in figuring out the “fair value” of bets on games of chance.

Probability theory, the mathematical study of randomness, originated with Pierre de Fermat and Blaise Pascal in the seventeenth century and was well developed by the time statisticians took it over in the twentieth century.

Myths about Chance Behavior

The idea of probability seems straightforward. It answers the question, "What would happen if we did this many times?"

We meet chance behavior constantly, and psychologists tell us that we deal with it poorly.

The myth of short-run regularity: The idea of probability is that randomness is regular in the long run. Unfortunately, our intuition about randomness tries to tell us that random phenomena should also be regular in the short run. When they aren't, we look for some explanation other than chance variation.

Example: What looks random?

Toss a fair coin six times and record heads (H) or tails (T) on each toss. Which of these outcomes is most probable?

HTHTTH HHHTTT TTTTTT

Almost everyone says that HTHTTH is more probable, because TTTTTT and HHHTTT do not “look random.”

All three are equally probable since coins have no memory and having half heads and half tails is a pattern that emerges in the long run, not the short term.

The Myth of the Surprising Coincidence 1

On November 18, 2006, Ohio State beat Michigan in football by a score of 42 to 39.

Later that day, the winning numbers in the Pick 4 Ohio lottery were 4239.

What an amazing coincidence!

Well, maybe not . . .

The Myth of the Surprising Coincidence 2

There are 32 NFL teams, 235 NCAA Division I teams, 150 NCAA Division II teams, 231 NCAA Division III teams, and over 25,000 high school teams.

All play a number of games during the season.

There are 38 states with a Pick 3 or Pick 4 lottery game, with winning numbers often drawn multiple times per week.

That’s a lot of opportunities to match a Pick 3 or Pick 4 lottery number that has digits that could conceivably be a football score such as 217 or 4239.

Example: Winning the Lottery Twice

In 1986, Evelyn Marie Adams won the New Jersey State lottery for the second time, adding $1.5 million to her previous $3.9 million jackpot.

The New York Times (February 14, 1986) claimed that the odds of one person winning the big prize twice were about 1 in 17 trillion.

Two statistics professors disagreed in a letter that appeared in the *Times* 2 weeks later.

Example: Winning the Lottery Twice (continued)

The chance of winning twice is tiny, but it is almost certain that someone among the millions of regular lottery players in the United States would win two jackpot prizes.

The statisticians estimated even odds (a probability of 1/2) of another double winner within 7 years.

Sure enough, Robert Humphries won his second Pennsylvania lottery jackpot ($6.8 million total) in May 1988.

The Myth of the Law of Averages 1

Roaming the gambling floors in Las Vegas, watching money disappear into the drop boxes under the tables, is revealing.

You can see some interesting human behavior in a casino.

When the shooter in the dice game craps rolls several winners in a row, some gamblers think she has a “hot hand” and bet that she will keep on winning.

Others say that “the law of averages” means that she must now lose so that wins and losses will balance out.

The Myth of the Law of Averages 2

Believers in the law of averages think that if you toss a coin six times and get TTTTTT, the next toss must be more likely to give a head.

It’s true that, in the long run, heads must appear half the time. What is myth is that future outcomes must make up for an imbalance such as six straight tails.

After 10,000 tosses, the results of the first six tosses don’t matter.

They are overwhelmed by the results of the next 9994 tosses, not compensated for.

What Is the Law of Averages? 1

Is there a “law of averages”? There is, although it is sometimes referred to as the “law of large numbers.”

The law of large numbers states that in a large number of “independent” repetitions of a random phenomenon (such as coin tossing), averages or proportions are likely to become more stable as the number of trials increases, whereas sums or counts are likely to become more variable.

What Is the Law of Averages? 2

next slide

In Figure 17.1, we see that the proportion of heads gradually becomes closer and closer to 0.5 as the number of tosses increases. This illustrates the law of large numbers.

next slide

In Figure 17.3, we see that for these same tosses, the total number of heads differs from exactly half of the tosses being heads. We see how this difference varies more and more as the number of tosses increases.

The law of large numbers does not apply to sums or counts.

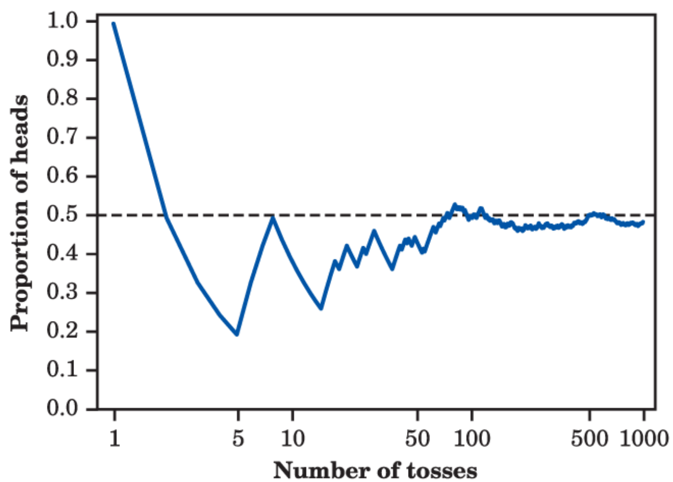
What Is the Law of Averages? 3

$$\underbrace{\frac{\#(H)}{n}}\to\frac{1}{2}$$

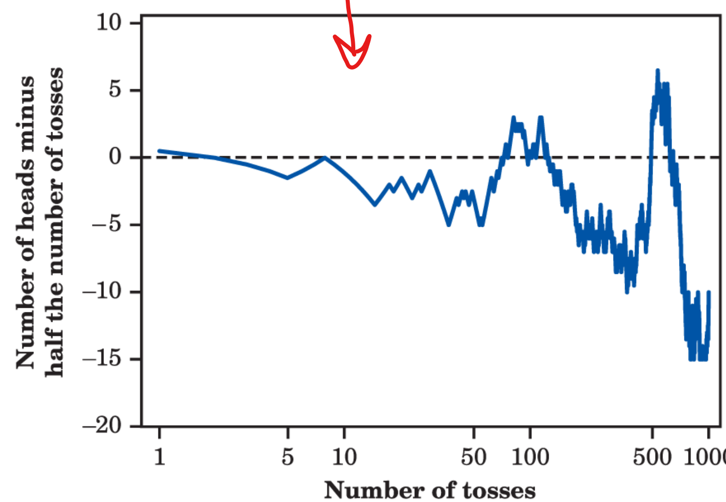
The same as

$$\frac{\#(H)-\frac{1}{2}n}{n}\to 0$$

$$\underbrace{\#(H)-\frac{1}{2}n}$$



Moore/Notz, *Statistics: Concepts and Controversies*, 10e,
© 2020 W. H. Freeman and Company



Moore/Notz, *Statistics: Concepts and Controversies*, 10e,
© 2020 W. H. Freeman and Company

Personal Probabilities 1

Joe sits staring into his beer as his favorite baseball team, the Chicago Cubs, loses another game.

The Cubbies have some good young players, so let's ask Joe, "What's the chance that the Cubs will go to the World Series next year?"

Joe brightens up. "Oh, about 10%," he says.

Does Joe assign probability 0.10 to the Cubs' appearing in the World Series?

Personal Probabilities 2

The outcome of next year's pennant race is certainly unpredictable, but we can't reasonably ask what would happen in many repetitions.

If probability measures "what would happen if we did this many times," Joe's 0.10 is not a probability.

Probability is based on data about many repetitions of the same random phenomenon.

Joe is giving us something else: his personal judgment.

Personal Probabilities 3

Yet we often use the term *probability* in a way that includes personal judgments of how likely it is that some event will happen.

We make decisions based on these judgments: we take the bus downtown because we think the probability of finding a parking spot is low.

A **personal probability** of an outcome is a number between 0 and 1 that expresses an individual's judgment of how likely the outcome is.

Probability and Risk 1

Once we understand that “personal judgment of how likely” and “what happens in many repetitions” are different ideas, we have a good start toward understanding why the public and the experts disagree so strongly about what is risky and what isn’t.

Experts use probabilities from data.

Individuals and society seem to ignore data.

We worry about some risks that almost never occur while ignoring others that are much more probable.

Example: Asbestos in the Schools

High exposure to asbestos is dangerous. Low exposure is not very risky.

The probability that a teacher who works for 30 years in a school with typical asbestos levels will get cancer from the asbestos is around 15/1,000,000.

The risk of dying in a car accident during a lifetime of driving is about 15,000/1,000,000.

Example: Asbestos in the Schools (continued)

Driving regularly is about 1000 times more risky than teaching in a school where asbestos is present!

Risk does not stop us from driving. Yet the much smaller risk from asbestos launched massive cleanup campaigns and a federal requirement that every school inspect for asbestos and make the findings public.

Probability and Risk 2

Why do we worry about very unlikely threats such as tornadoes and terrorists more than we worry about heart attacks?

We feel safer when a risk seems under our control than when we cannot control it. We are in control (or so we imagine) when we are driving, but we can't control the risk from asbestos or tornadoes or terrorists.

Probability and Risk 3

It is hard to comprehend very small probabilities. Probabilities of 15 per million and 15,000 per million are both so small that our intuition cannot distinguish between them.

Psychologists have shown that we generally overestimate very small risks and underestimate higher risks. Perhaps this is part of the general weakness of our intuition about how probability operates.

Probability and Risk 4

The probabilities for risks such as asbestos in the schools are not as certain as probabilities for tossing coins. They must be estimated from complicated statistical studies by experts. Perhaps it is safest to suspect that the experts may have underestimated the level of risk.

Statistics in Summary 1

- Some things in the world, both natural and of human design, are **random**. That is, their outcomes have a clear pattern in very many repetitions even though the outcome of any one trial is unpredictable.

- **Probability** describes the long-term regularity of random phenomena. The probability of an outcome is the proportion of very many repetitions on which that outcome occurs. A probability is a number between 0 (the outcome never occurs) and 1 (always occurs). We emphasize this kind of probability because it is based on data.

Statistics in Summary 2

- Probabilities describe only what happens in the long run. Short runs of random phenomena such as tossing coins or shooting a basketball often don't look random to us because they do not show the regularity that, in fact, emerges only in very many repetitions.

- **Personal probabilities** express an individual's personal judgment of how likely outcomes are. Personal probabilities are also numbers between 0 and 1. Different people can have different personal probabilities, and a personal probability need not agree with a proportion based on data about similar cases.