7. PROBABILITY THEORY

Probability shows you the likelihood, or chances, for each of the various future outcomes, based on a set of assumptions about how the world works.

• Allows you to handle randomness (uncertainty) in a consistent, rational manner.

• Forms the foundation for statistical inference (drawing conclusions from data), sampling, linear regression, forecasting, risk management.

The Link Between Probability and Statistics

With **Statistics**, you go from observed data to generalizations about how the world works.

For example, if we observe that the seven hottest years on record occurred in the most recent decade, we may conclude (perhaps without justification) that there is global warming.

With **probability**, you start from an assumption about how the world works, and then figure out what kinds of data you are likely to see.

In the above example, we could assume that there is no global warming and ask how likely we would be to get such high temperatures as we have been observing recently.

So probability provides the justification for statistics!

Indeed, probability is the only scientific basis for decision-making in the face of uncertainty.

Questions:

If you toss a coin, what is the probability of getting a head?

Explain your answer in two different ways.

What did you mean by "probability"?

If you toss a coin twice, what is the probability of getting exactly one Head? Suggest a practical way to verify your answer.

(Toss a coin twice, to test your conjecture!)

If you toss a coin 10 times and count the total number of Heads, do you think   Prob (0 Heads) = Prob (5 Heads)?
Do you think   Prob (4 Heads) = Prob (6 Heads)?

(Again, let's try it.)

Random Experiment: A process or course of action that results in one of a number of possible outcomes. The outcome that occurs cannot be predicted with certainty.

Sample Space: The set of all possible outcomes of the experiment.

Eg: If the experiment is "Toss a coin twice", the sample space is

$$S = \{HH, HT, TH, TT\}.$$

Note: HT means "Heads on first toss, tails on second toss".

The individual outcomes of the sample space are called **simple events**. So S is composed of the simple events HH, HT, TH, TT.

Do you think that HT and TH are really different?
(Consider the evidence from our experiments).

Event: Any subset of the sample space.

Equivalently, an event is any collection of simple events.

Eg: In coin-tossing, let A = {Exactly One Head}. Then A is an event. Specifically, A = {HT,TH}.

Attempts to Define Probability

The sad truth: "Probability" has no precise definition!!

All attempts to define probability must ultimately rely on circular reasoning.

Roughly speaking, the probability of an event is the "chance" or "likelihood" that the event will occur.

To each event A, we want to attach a number P(A), called the **probability** of A, which represents the likelihood that A will occur.

There are various ways to define P(A), but in order to make sense, any definition must satisfy

- P(A) is between zero and 1.
  (Zero represents "impossibility" and 1 represents "certainty".)

- $P(E_1) + P(E_2) + \cdots = 1$,
    where E_1, E_2, $\cdots$ are the simple events in the sample space.

The three most useful approaches to obtaining a definition of probability are:

> The **classical approach,** the **relative frequency approach**, and the **subjective approach**.

The Classical Approach

Assume that all simple events are equally likely. Define the classical probability that an event A will occur as

$$P(A) = \frac{\#\text{Simple Events in A}}{\#\text{Simple Events in S}}$$

So P(A) is the number of ways in which A can occur, divided by the number of possible individual outcomes, assuming all are equally likely.

Eg: In tossing a coin twice, if we take

$$S = \{HH, HT, TH, TT\},$$

then the classical approach assigns probability 1/4 to each simple event. If

$$A = \{\text{Exactly One Head}\} = \{HT, TH\}, \text{ then}$$

$$P(A) = 2/4 = 1/2 .$$

Question: Does this tell you how often A would occur if we repeated the experiment ("toss a coin twice") many times?

The relative frequency approach

The probability of an event is the long run frequency of occurrence.

To estimate P(A) using the frequency approach, repeat the experiment *n* times (with *n* large) and compute *x/n*, where *x* = # Times A occurred in the *n* trials.

The larger we make *n*, the closer *x/n* gets to P(A).

- $\dfrac{x}{n} \to P(A)$.

Eg: If there have been 126 launches of the Space Shuttle, and two of these resulted in a catastrophic failure, we can estimate the probability that the next launch will fail to be
$$2/126 = 0.016.$$

The frequency approach allows us to determine the probability empirically from actual data. It is more widely applicable than the Classical approach, since it doesn't require us to specify a sample space consisting of equally likely simple events.

Table 1—Probabilities of Various Baseball Outcomes Based on 1989 Data

Event	American League	National League
Out	.672	.685
Walk	.092	.091
Single	.171	.159
Double	.040	.040
Triple	.005	.006
Home run	.020	.019
Errors	2% of outs	
Sacrifice fly	12.5% of outs (if applicable)	
Double play	17.5% of outs (if applicable)	

The Subjective Approach

This approach is useful in betting situations and scenarios where one-time decision-making is necessary. In cases such as these, we wouldn't be able to assume all outcomes are equally likely and we may not have any prior data to use in our choice.

The subjective probability of an event reflects our personal opinion about the likelihood of occurrence. Subjective probability may be based on a variety of factors including intuition, educated guesswork, and empirical data.

Eg: In my opinion, there is an 85% probability that Stern will move up in the rankings in the next Business Week survey of the top business schools.

Relationship Between Classical and Frequency Approaches:

If we can find a sample space in which the simple events really are equally likely, then the **Law of Large Numbers** asserts that the classical and frequency approaches will produce the same results.

Eg: For the experiment "Toss a coin once", the sample space is S = {H, T} and the classical probability of Heads is 1/2.

According to the Law of Large Numbers (LLN), if we toss a fair coin repeatedly, then the proportion of Heads will get closer and closer to the Classical probability of 1/2.

For a demonstration of the LLN, see the website at:
http://users.ece.gatech.edu/~gtz/java/cointoss/index.html

Questions:

In Roulette, if you've seen a long run of red, does it make sense to start betting on black?

If you've been losing at the racetrack, do you bet more on the last race? Why?

In craps, if the shooter throws a crap on the come-out, does this improve the chances of getting a 7 or 11 on the next roll? (The gamblers call this an "apology").

If the market has been moving up recently, does this increase the chances of a sudden drop? (Financial analysts call this a "correction").

Do you believe in "hot hands" in gambling? In sports? In business?

If a financial analyst has made all the "right" decisions in the past do you think it is likely that he/she will continue to do so?

If there are thousands of financial analysts, isn't it likely that one could make all the "right" decisions just by chance? Should we make a hero out of that person?


What is the "Law of Averages" that gamblers talk about?

"Law of Averages: the proposition that the occurrence of one extreme will be matched by that of the other extreme so as to maintain the normal average."
(Oxford American Dictionary, 1980).

Most gamblers interpret this to mean that if you've been losing badly your luck will have to improve if you keep playing.

This is *not* true!

In Minitab, I simulated 1000 rounds of a coin-tossing game where at each round, you win $1 for Heads and lose $1 for Tails. I plotted the player's winnings, and the proportion of Heads.

Player's Winnings

Proportion of Heads
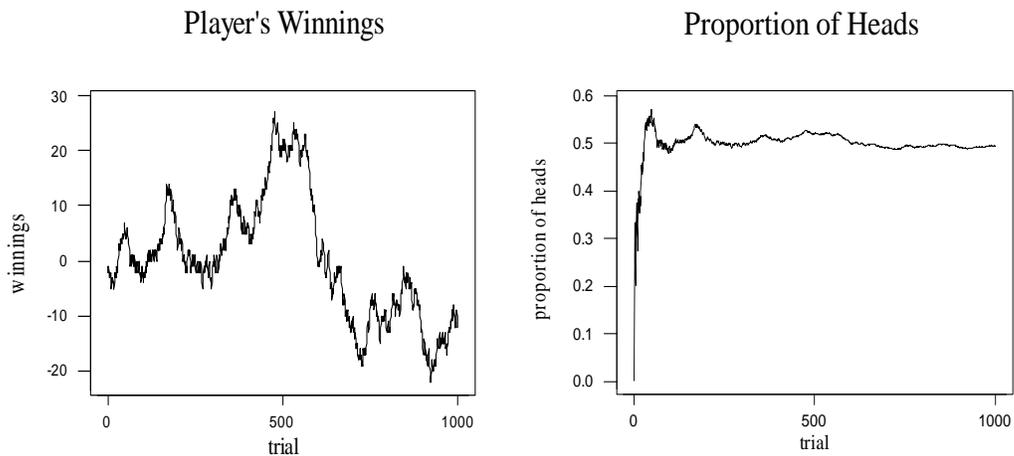
The proportion of Heads does tend to 1/2, as the Law of Large Numbers says it should.

But the player's winnings

(Total number of Heads – Total number of Tails)

does *not* tend to zero! It just wanders all over the place.

•Even though the Law of Large Numbers is true, the Law of Averages is a fallacy!

Lucking Out: Weird Rituals and Strange Beliefs

Psychologists tell me that superstitions are rooted in false correlations between a particular act and a particular result, not unlike rain falling on Indian crops after a rain dance. It need not rain every time the dance is done for a superstitious belief in the dance's power to evolve. For both sports fans and athletes, it often takes but a single victory to convince them that wearing a certain article of clothing, or doing or *not* doing a particular act can influence the result of a game.

As Michael J. Mahoney, a psychologist, sees it, superstitious behavior in sports is a way of assuming control over chance factors that are really uncontrollable.

"Superstitions are primitive but very powerful rituals that acknowledge there are limits to what a person can control," said Mahoney, a member of the department of psychology at the University of North Texas in Denton.

The New York Times, January 27, 1991

Odds

Odds are often used to describe the payoff for a bet.

Consider horseracing, for example.

If the odds against a horse are $a{:}b$, then the bettor must risk b dollars to make a profit of a dollars.

If the true probability of the horse winning is $b/(a+b)$, then this is a fair bet.

In the 1999 Belmont Stakes, the odds against Lemon Drop Kid were 29.75 to 1, so a $2 ticket paid $61.50.

The ticket returns two times the odds, plus the $2 ticket price.

Odds and payoffs for horseracing.

BELMONT STAKES

9th race yesterday at Belmont Park
1½ Miles. Purse: $1,000,000-added 3-year-olds.
131st running.
Stake value of race: $1,000,000. Value to winner $600,000; Second $200,000; Third $100,000; Fourth $60,000;
Fifth $30,000.

Horse	Wgt	PP	¼	½	1M	1¼	Str	Fin	Jockey	To $1
Lemon Drop Kid	126	6	8-3½	8-2	8-2½	4-hd	1-½	1-hd	Santos	29.75
Vision and Verse	126	2	3-hd	4-hd	4-2½	5-2½	3-1½	2-1½	Castillo	54.75
Charismatic	126	4	2-1	2-½	2-½	1-hd	2-hd	3-4¾	Antley	1.60
Best of Luck	126	12	11-5	11-9	10-½	6-½	6-2	4-½	Samyn	13.00
Stephen Got Even	126	11	5-1½	3-½	3-½	2-hd	4-1½	5-3½	Sellers	9.30
Patience Game	126	7	6-½	7-1	9-hd	8-1½	7-5	6-no	Desormeaux	12.50
Silverbulletday	121	3	1-hd	1-hd	1-½	3-1	5-1½	7-7½	Bailey	5.10
Menifee	126	10	4-hd	5-1	5-hd	9-4½	8-6	8-3¼	Day	2.60
Pineaff	126	5	12	12	12	12	10-½	9-8	LeJeune	60.00
Prime Directive	126	9	7-½	6-½	6-hd	11-4	11-1	10-nk	Smith	76.00
Teletable	126	1	9-1	9-2½	11-20	10-hd	9-2	11-2¾	Velazquez	75.75
Adonis	126	8	10-6	10-4½	7-½	7-½	12	12	Chavez	46.25

Time: 23.79; 47.60; 1:12.8; 1:36.57; 2:1.90; 2:27.88.
6 (6) Lemon Drop Kid . . .61.50 26.00 10.60
2 (2) Vision and Verse. 44.40 17.00
4 (4) Charismatic 3.60
Off 5:29. Start: Good. Track: Fast.
Exacta (6-2) paid $1,537.00; Trifecta (6-2-4) paid $5,343.00; Pick 3 (4-3-6) 3 Correct Paid $2,450.00.
Trainer: F. Schulhofer.
Scratched: None; Overweights: None.

The New York Times, June 6, 1999

If a fair coin is tossed once, the odds on Heads are 1 to 1.

If a fair die is tossed once, the odds on a six are 5 to 1.

In the game of Craps, the odds on getting a 6 before a 7 are 6 to 5. (We will show this later).