Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 5

Basic Probability – Part I

“Odd Question” # 7

‘Imitate” a coin. That is, write down a sequence of 100 H (for heads) and T (for tails) without tossing a coin–but a sequence that you think will fool everyone into thinking it is the reporting of tossing a fair coin.

Which of these is a real sequence of coin flips?

Exhibit A

H H H H T H H T T T H H T T T H T T T H T T H H T

T T T H T T T H T H T T H H H H T T T T T H H H H

H T T H T T H H T T H H H H H T H H T H T H T H T

T H H T H H T T T H T T T T T T T T T H H T T T T

Exhibit B

H H T H T T T H H T H H H T H T T T H T H H T T T

T H H T T T H H H T H T T T H T T H H T H H T H T

T T H H H T H T T H T H H T T H H H T H T T H H H

T T H H H H T H T T H H T T T H H T H H H T T H T

How could we tell which are the real coin flips?

Hacking (2001, p. 31)

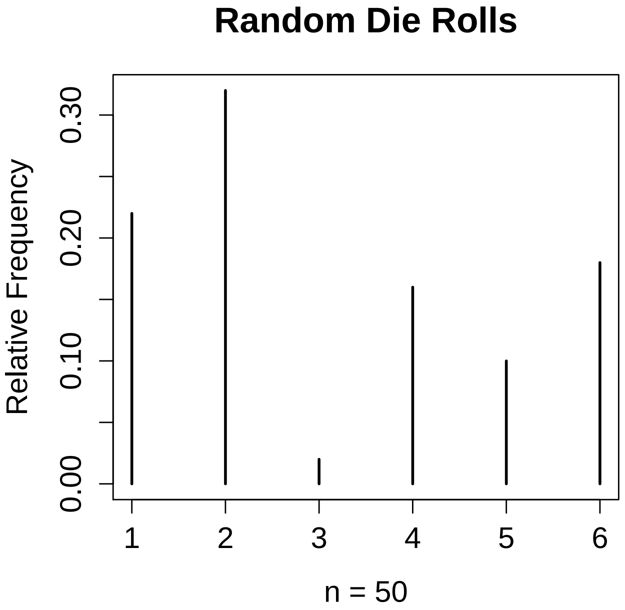
> *Hardly anyone making up a sequence of 10 tosses puts in a run of 7 heads in a row. It is true that the chance of getting 7 heads in a row with a fair coin is only 1/64. But in tossing a coin 100 times, you have at least 93 chances to start tossing 7 heads in a row, because each of the first 93 tosses could begin a run of 7. It is more probable than not, in 100 tosses, that you will get 7 heads in a row. It is certainly more probable than not, that you will get at least 6 heads in a row. Yet almost no one writes down a pretend sequence, in which there are even 6 heads in a row.*

Our Definition of Probability for this Course

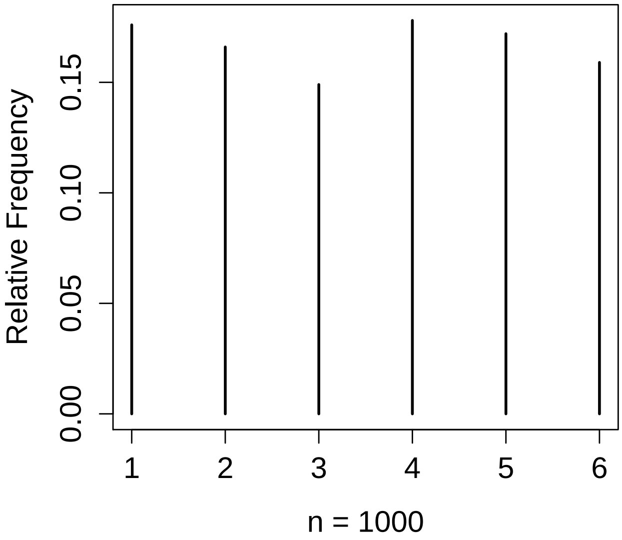
$$\text{Probability} = \text{Long-run Relative Frequency}$$

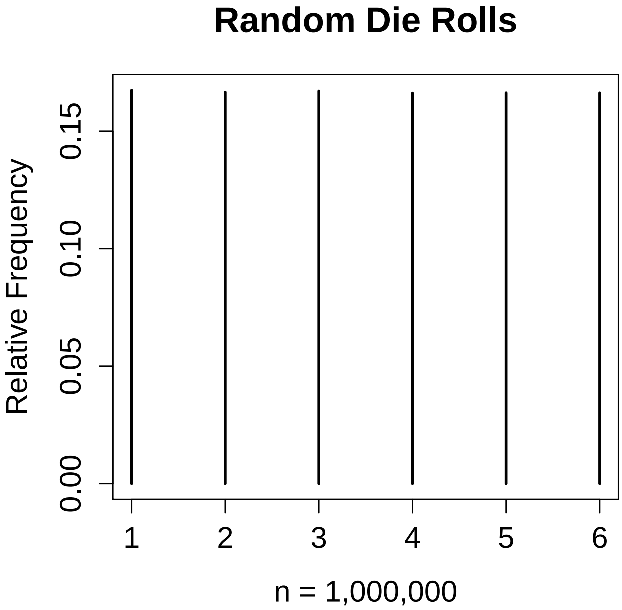
That is, relative frequencies settle down to probabilities if we carry out an experiment over, and over, and over...


Random Die Rolls
Relative Frequency
0.00
0.10
0.20
0.30
1
2
4
5
6
n = 10


Random Die Rolls
Relative Frequency
0.00
0.10
0.20
0.30
1
2
3
4
5
6
n = 50

Random Die Rolls




Random Die Rolls
Relative Frequency
0.00
0.05
0.10
0.15
1
2
3
4
5
6
n = 1,000,000

What do you think of this argument?

- The probability of flipping heads is $1/2$: if we flip a coin many times, about half of the time it will come up heads.
- The last ten throws in a row the coin has come up heads.
- The coin is bound to come up tails next time – it would be very rare to get 11 heads in a row.

(a) Agree

(b) Disagree

The Gambler's Fallacy

Relative frequencies settle down to probabilities, but this does not mean that the trials are dependent.

Dependent = "Memory" of Prev. Trials

Independent = No "Memory" of Prev. Trials

Another Argument

Lucie visits Albert. As she enters, he rolls four dice and shouts "Hooray!" for he has just rolled four sixes. Lucie: "I bet you've been rolling the dice for a long time to get that result!" Now, Lucie may have many reasons for saying this – perhaps Albert is a lunatic dice-roller. But simply on the evidence that he has just rolled four sixes, is her conclusion reasonable?

(a) Yes

(b) No

The *Inverse* Gambler's Fallacy

This is true:

Albert is more likely to get four sixes if he rolls many times than if he rolls only once.

However:

Regardless of how long Albert has been rolling, the probability that he gets four sixes on the particular roll that Lucie observes is unchanged.

The outcome of that roll doesn't tell us anything about whether he has rolled the dice before, just like six heads in a row doesn't mean we're "due" for a tails.

Terminology

Random Experiment

An experiment whose outcomes are random.

Basic Outcomes

Possible outcomes (mutually exclusive) of random experiment.

Sample Space: S

Set of all basic outcomes of a random experiment.

Event: E

A subset of the Sample Space (i.e. a collection of basic outcomes).

In set notation we write $E \subseteq S$.

Example

Random Experiment

Tossing a pair of dice.

Basic Outcome

An ordered pair (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$, e.g. $(2, 5)$

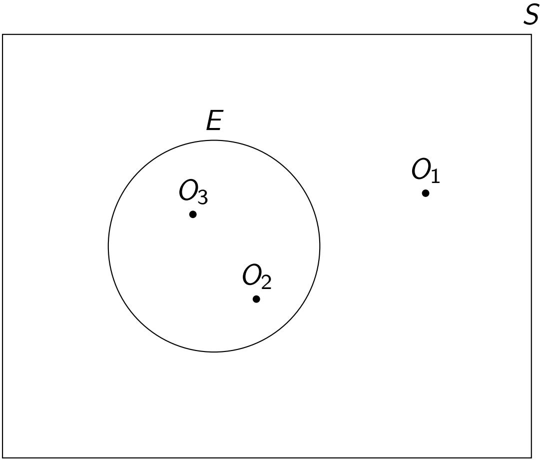
Sample Space: S

All ordered pairs (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$

Event: $E = \{\text{Sum of two dice is less than 4}\}$

$\{(1, 1), (1, 2), (2, 1)\}$

Visual Representation



Probability is Defined on *Sets*, and Events are Sets

Complement of an Event: $A^c = \text{not } A$

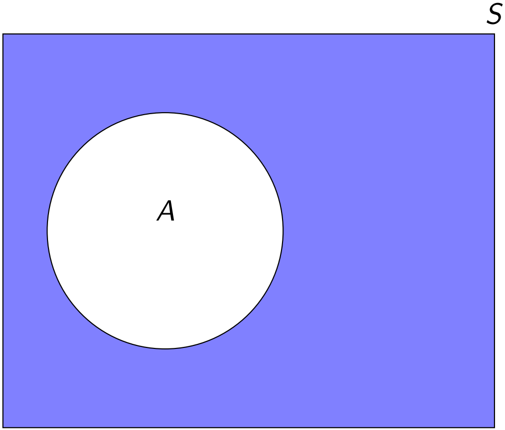


Figure: The complement A^c of an event $A \subseteq S$ is the collection of all basic outcomes from S not contained in A.

Intersection of Events: $A \cap B = A$ and B

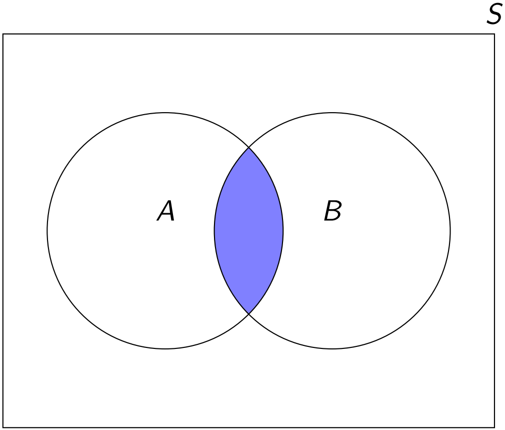


Figure: The intersection $A \cap B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in both A and B

Union of Events: $A \cup B = A$ or B

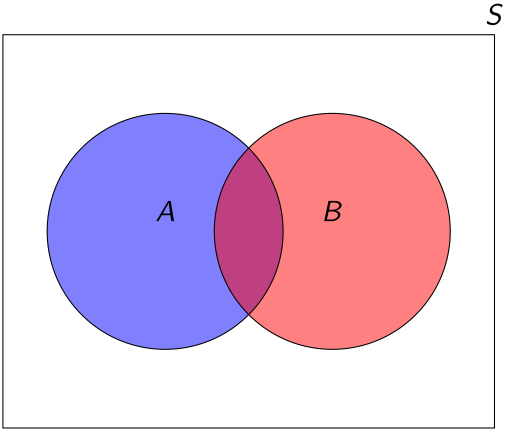


Figure: The union $A \cup B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in A, B or both.

Mutually Exclusive and Collectively Exhaustive

Mutually Exclusive Events

A collection of events $E_1, E_2, E_3, \ldots$ is *mutually exclusive* if the intersection $E_i \cap E_j$ of *any two different events* is empty.

Collectively Exhaustive Events

A collection of events $E_1, E_2, E_3, \ldots$ is *collectively exhaustive* if, taken together, they contain *all of the basic outcomes in S*.
Another way of saying this is that the union $E_1 \cup E_2 \cup E_3 \cup \cdots$ is S.

Implications

Mutually Exclusive Events

If one of the events occurs, then none of the others did.

Collectively Exhaustive Events

One of these events *must* occur.

Mutually Exclusive but *not Collectively Exhaustive*

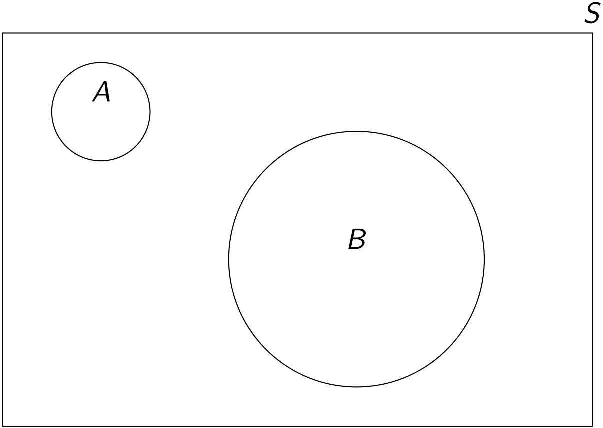


Figure: Although A and B don't overlap, they also don't cover S.

Collectively Exhaustive but *not Mutually Exclusive*

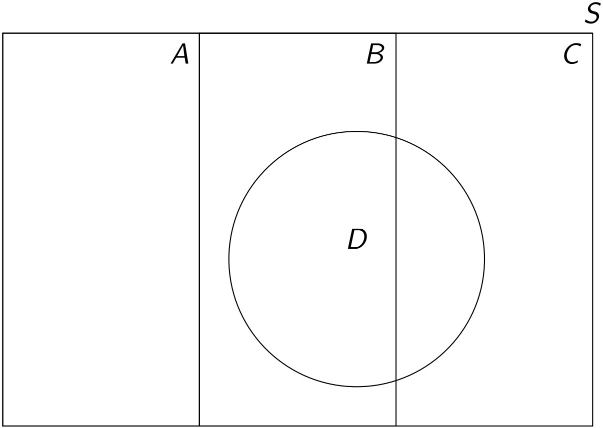


Figure: Together A, B, C and D cover S, but D overlaps with B and C.

Collectively Exhaustive *and* Mutually Exclusive

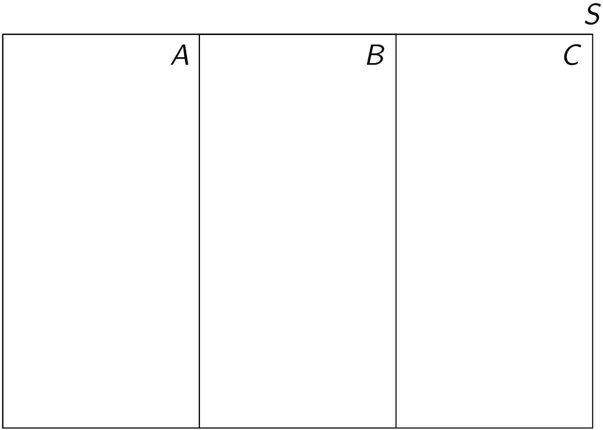


Figure: A, B, and C cover S and don't overlap.

Axioms of Probability

We assign every event A in the sample space S a real number $P(A)$ called the probability of A such that:

Axiom 1 $0 \leq P(A) \leq 1$

Axiom 2 $P(S) = 1$

Axiom 3 If $A_1, A_2, A_3, \ldots$ are mutually exclusive events, then
$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \ldots$$

"Classical" Probability

When all of the basic outcomes are equally likely, calculating the probability of an event is simply a matter of counting – count up all the basic outcomes that make up the event, and divide by the total number of basic outcomes.

Recall from High School Math:

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, $\ldots$, n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Corollary – Number of Possible Orderings

$k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1 = k!$

Permutations – Order n people in k slots

$P_k^n = \frac{n!}{(n-k)!}$ (Order Matters)

Combinations – Choose committee of k from group of n

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$, where $0! = 1$ (Order Doesn't Matter)

Poker – Deal 5 Cards, Order Doesn't Matter

Basic Outcomes

$\binom{52}{5}$ possible hands

How Many Hands have Four Aces?

Poker – Deal 5 Cards, Order Doesn't Matter

Basic Outcomes

$\binom{52}{5}$ possible hands

How Many Hands have Four Aces?

48 (# of ways to choose the single card that is not an ace)

Probability of Getting Four Aces

$48/\binom{52}{5} \approx 0.00002$

Poker – Deal 5 Cards, Order Doesn't Matter

What is the probability of getting 4 of a kind?

- 13 ways to choose *which* card we have four of
- 48 ways to choose the last card in the hand
- $13 \times 48 = 624$

$$624/\binom{52}{5} \approx 0.00024$$

A Fairly Ridiculous Example

Roger Federer and Novak Djokovic have agreed to play in a tennis tournament against six Penn professors. Each player in the tournament is randomly allocated to one of the eight rungs in the ladder (next slide). Federer always beats Djokovic and, naturally, either of the two pros always beats any of the professors. What is the probability that Djokovic gets second place in the tournament?

Rung 1

Quarterfinalist 1

Rung 2

Semifinalist 1

Rung 3

Quarterfinalist 2

Rung 4

Winner

Rung 5

Quarterfinalist 3

Rung 6

Semifinalist 2

Rung 7

Quarterfinalist 4

Rung 8

Solution: Order Matters!

Denominator

8! basic outcomes – ways to arrange players on tournament ladder.

Numerator

Sequence of three decisions:

1. Which rung to put Federer on? (8 possibilities)
2. Which rung to put Djokovic on?
    - For any given rung that Federer is on, only 4 rungs prevent Djokovic from meeting him until the final.
3. How to arrange the professors? (6! ways)

$$\frac{8 \times 4 \times 6!}{8!} = \frac{8 \times 4}{7 \times 8} = 4/7 \approx 0.57$$

Even if the basic outcomes are equally likely, the events of interest may not be...

“Odd Question” # 4

To throw a total of 7 with a pair of dice, you have to get a 1 and a 6, or a 2 and a 5, or a 3 and a 4. To throw a total of 6 with a pair of dice, you have to get a 1 and a 5, or a 2 and a 4, or a 3 and another 3. With two fair dice, you would expect:

(a) To throw 7 more frequently than 6.

(b) To throw six more frequently than 7.

(c) To throw 6 and 7 equally often.

Basic Outcomes Equally Likely, Events of Interest Aren't

		Second Die					
		1	2	3	4	5	6
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
First	3	4	5	6	7	8	9
Die	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Table: There are 36 equally likely basic outcomes, of which 5 correspond to a sum of six and 6 correspond to a sum of seven.

$P(7) = 6/36 = 1/6$

$P(6) = 5/36$