Introduction to Probability

Keegan Korthauer
Department of Statistics
UW Madison

Descriptive Statistics

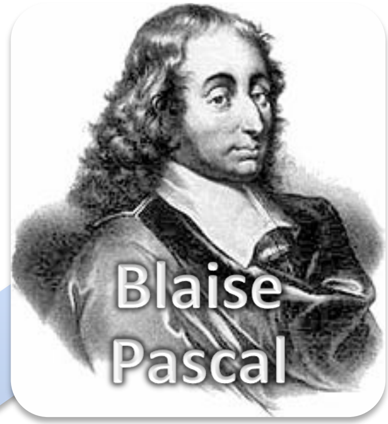
Data Type	Summary Statistics	Graphical Summary
Numerical (Quantitative)	**Location:** Mean, Median, Quartiles, Percentiles	**Univariate:** Histogram, Boxplot, Dotplot
	Spread: Standard Deviation, Variance, Range, IQR	**Bivariate:** Scatterplot
Categorical (Qualitative)	Frequencies (counts), Proportions	Bar Chart, Pie Chart

BASIC TERMINOLOGY

Sample space

Events

History



Early Development

17^{th} Century

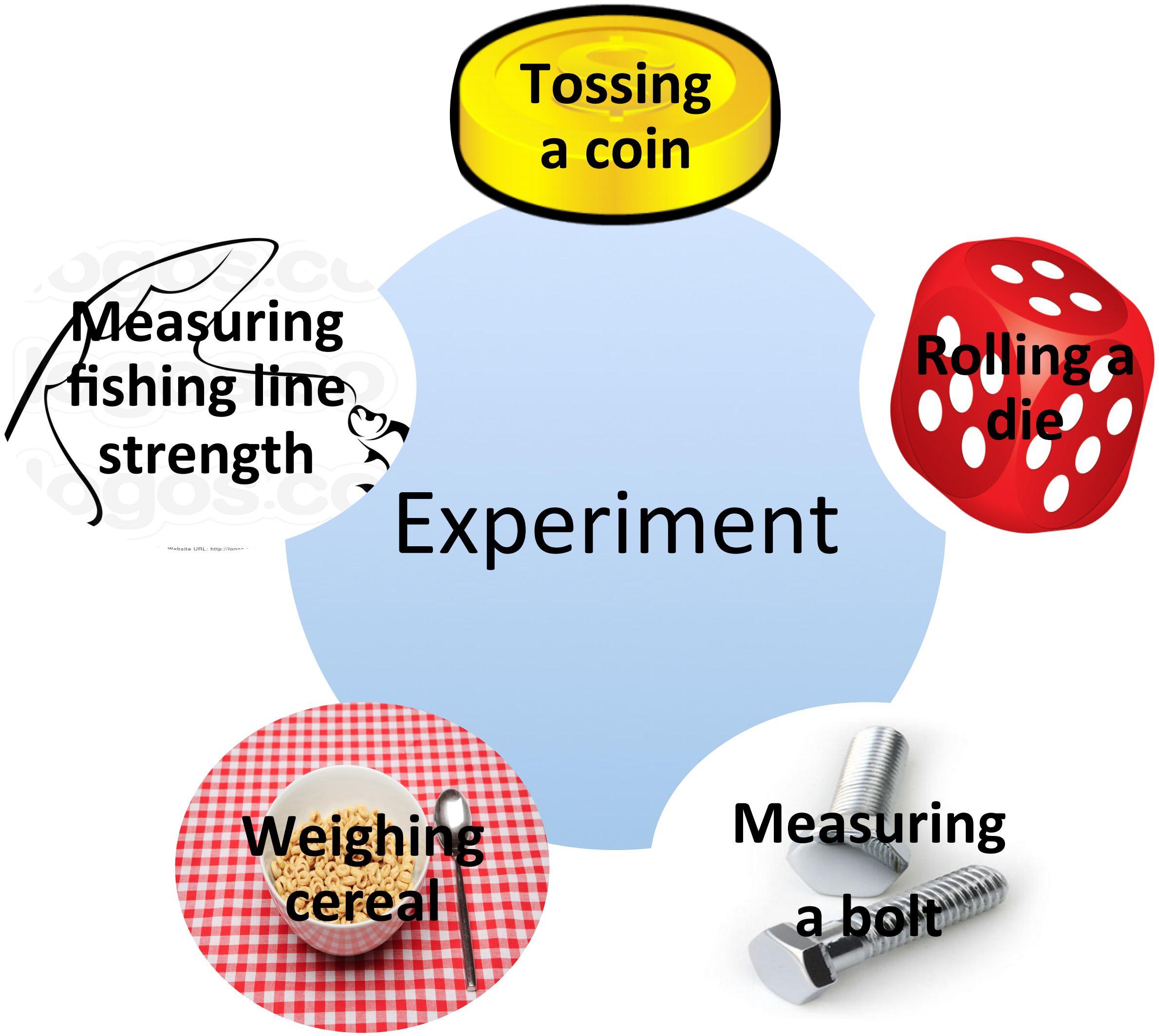
Modern Probability Theory

- Well-established branch of mathematics

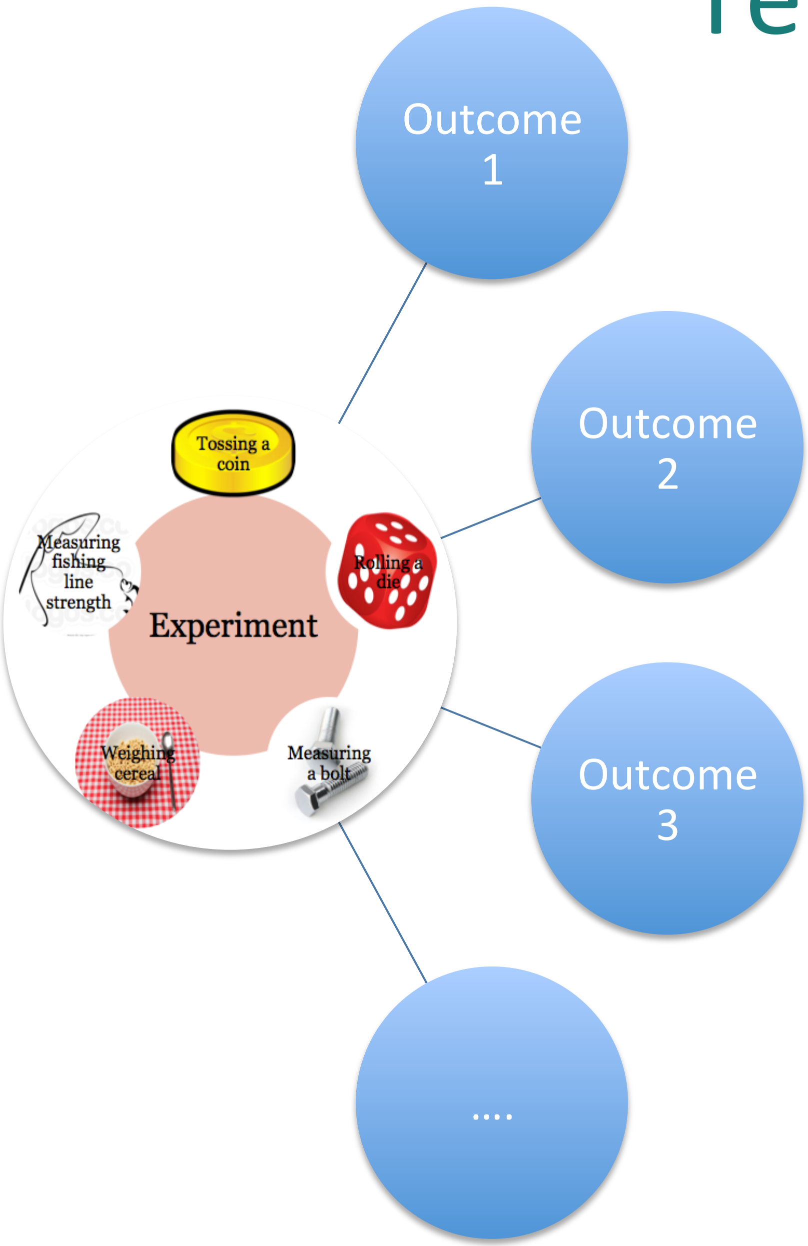
- Wide applications – anything to do with uncertainty and random chance

- Statistics, physics, finance, meteorology, medicine

Terminology



Terminology



Sample Space:

The set of **all possible outcomes** of an experiment

An Event:

A subset of a sample space

Sample Space

- Coin tossing - {Heads, Tails} (finite)
- Die rolling – {1, 2, 3, 4, 5, 6}
- Bolt measuring – {x cm | $0.4 < x < 0.45$} (infinite)
- Cereal weighing – {x gram | $10 < x < 12$}
- Fishing line strength measuring – {x lb | $3.5 < x < 4.9$}

Example 2.1

Two Boxes of Resistors

4 resistors labeled 10 ohms

4 resistors Labeled 20 ohms

Actual Resistances

9 10 11 12

18 19 20 21

The engineer **chooses one resistor from each box** and **determines the resistance of each**

Event A – 1st resistor has a resistance > 10
Event B – 2nd resistor has a resistance < 19
Event C – the sum of the resistances = 28

Example 2.1 Continued

- Find a sample space for this experiment

S = {(9,18), (9,19), (9,20), (9,21),
(10,18), (10,19), (10,20), (10,21),
(11,18), (11,19), (11,20), (11,21),
(12,18), (12,19), (12,20), (12,21)}

Each pair here represents one **outcome** of the experiment

- Specify the subsets corresponding to the events A, B, and C.

Event A = {(11,18), (11,19), (11,20), (11,21), (12,18), (12,19), (12,20), (12,21)}

Event B = {(9,18), (10,18), (11,18), (12,18)}

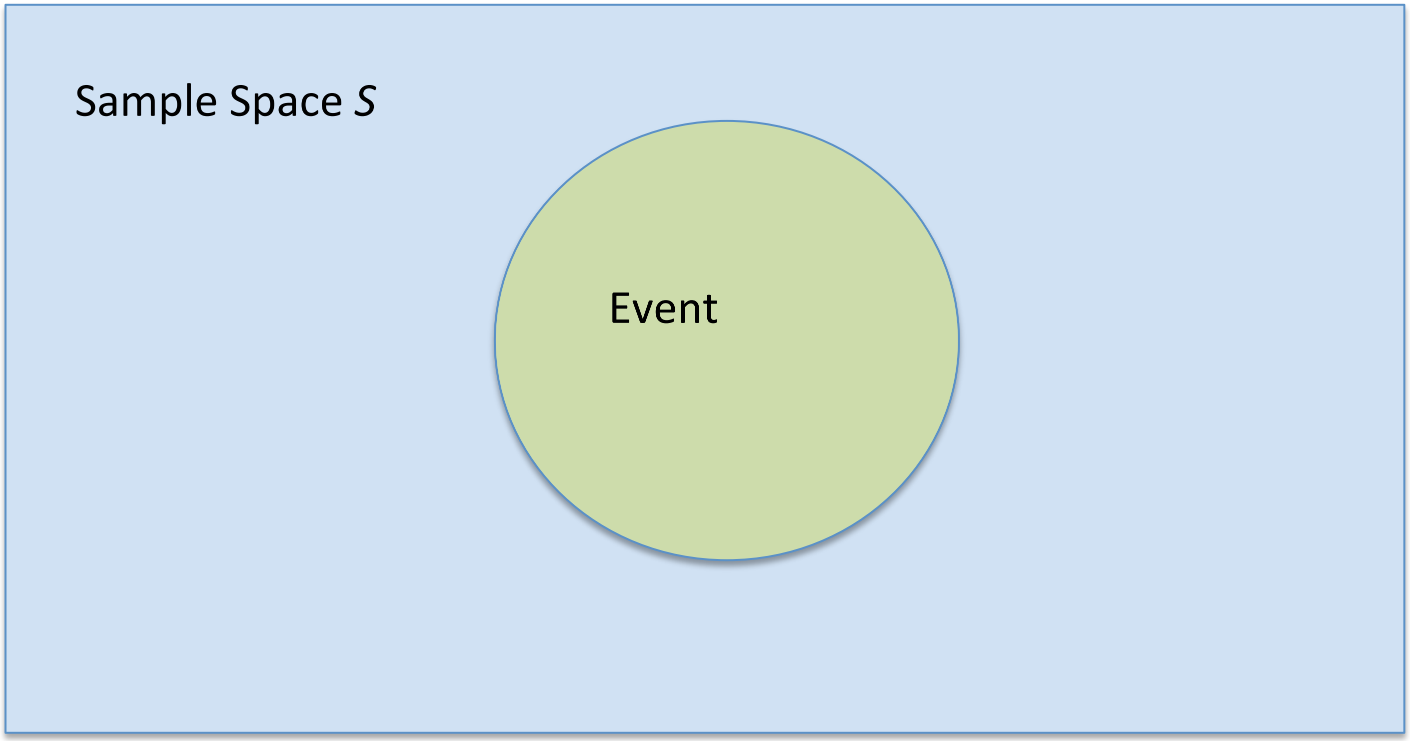
Event C = {(9,19), (10,18)}

Notes

- Empty set $\emptyset$ and the entire sample space S are both events of the sample space

- A given event is said to have occurred if the outcome of the experiment is one of the outcomes in the event, e.g. if a die comes up 2, the events {2,4,6} and {1,2,3} have both occurred

Venn Diagram



Combining Events

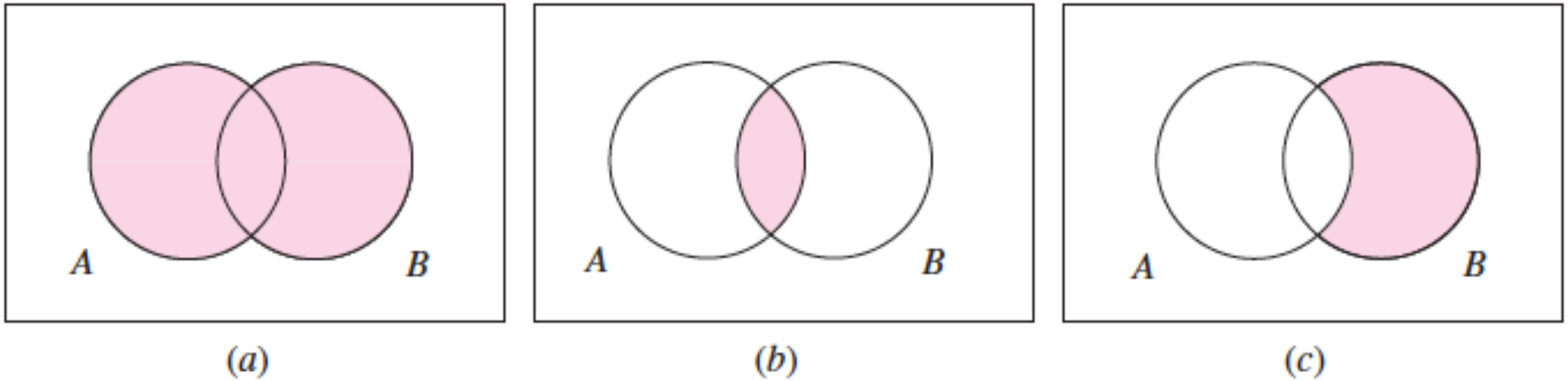


FIGURE 2.1 Venn diagrams illustrating various events: **(a)** $A \cup B$, **(b)** $A \cap B$, **(c)** $B \cap A^c$.

- **Union** ($A \cup B$) – outcomes belonging either to A **or** B, **or** both
- **Intersection** ($A \cap B$) – outcomes belonging **both** to A and to B
- **Complement** (A^c) – outcomes **not** belonging to A
- **Difference** (B-A) – outcomes belonging to B **but not** to A

Mutually Exclusive Events

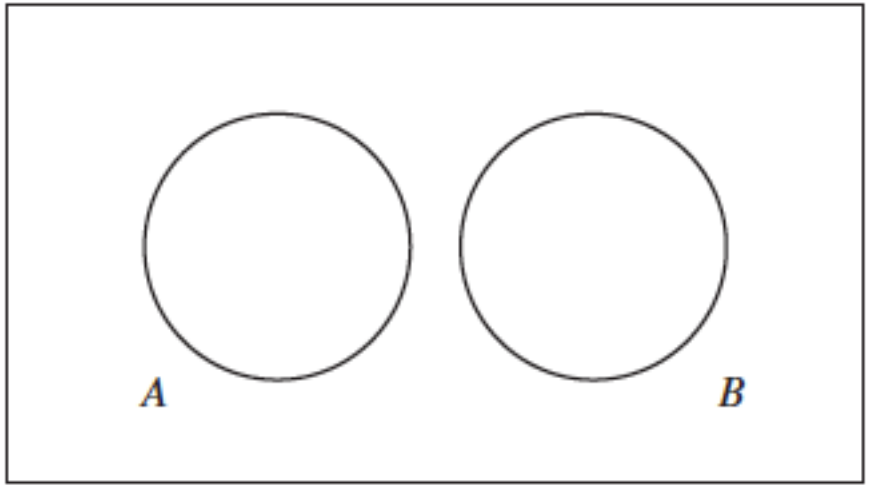


FIGURE 2.2 The events A and B are mutually exclusive.

Intersection between the events is empty.

Example 2.3

- Refer to the events of Example 2.1

$A = \{(11, 18), (11, 19), (11, 20), (11, 21), (12, 18), (12, 19), (12, 20), (12, 21)\}$
$B = \{(9, 18), (10, 18), (11, 18), (12, 18)\}$
$C = \{(9, 19), (10, 18)\}$

- If the experiment is performed, is it possible for events A and B both to occur?
- How about B and C?
- A and C? Which pair of events is mutually exclusive?

Answers:

- A and B both occur with outcomes (11,18) and (12,18)
- B and C both occur with outcome (10,18)
- A and C are mutually exclusive because they have no common outcomes

AXIOMS OF PROBABILITY

Defining probability

Axioms of probability

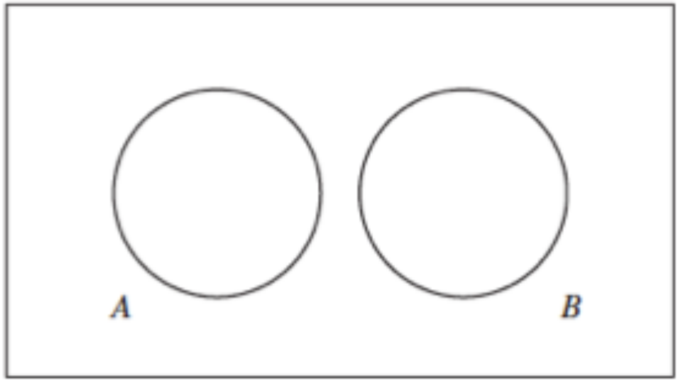
Definition

- **Probability**: A quantitative measure of how likely the event is to occur

- The probability of event A occurring is denoted as P(A)

The Three Axioms of Probability

1. P(S) = 1, where S denotes the sample space



2. $0 \leq P(A) \leq 1$ for any event A

3. If A and B are mutually exclusive events, then
   $P(A \cup B) = P(A) + P(B)$

   If A_1, A_2, … are mutually exclusive events, then
   $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$

Finding event probabilities

- If A is an event containing outcomes $O_1,\ldots,O_n$ (so $A=\{O_1,\ldots,O_n\}$), then

$$P(A) = P(O_1) + P(O_2) + \ldots + P(O_n)$$

- Since **outcomes** are mutually exclusive, this follows from axiom three
- Be careful not to confuse **events** with **outcomes**

Lottery Revisited

- a.k.a. Sample Spaces with Equally Likely Outcomes

- A population from which an item is sampled at random can be thought of as a sample space with equally likely outcomes

- S - a sample space with N equally likely outcomes

  A - an event containing k outcomes

P(A) = (# of ways A can happen) / (total # possible outcomes)

= k / N

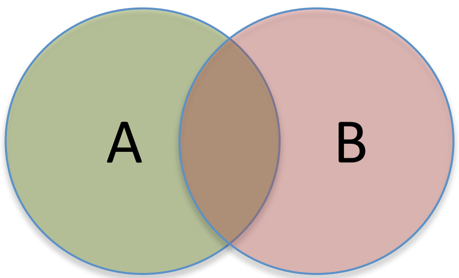
Properties

(a) If A, B $\in$ *S* and A $\subset$ B, then P(A) $\leq$ P(B)

(b) P(A) = 1 – P(A^c)

(c) If A and B are mutually exclusive, then P(A $\cap$ B) = 0

(d) P(A $\cup$ B) = P(A) + P(B) - P(A $\cap$ B)



Example 2.8

In aluminum can manufacturing, the probability that a can has a flaw

- on its side is 0.02
- on the top is 0.03
- on both the side and the top is 0.01

(a) What is the probability that a randomly chosen can has a flaw?

(b) What is the probability that it has no flaw?

Example 2.8 (a)

P(Flaw on side) = 0.02

P(Flaw on top) = 0.03

P(Flaw on side **AND** Flaw on top) = 0.01

What property/formula to use? Property (d)

P(A randomly chosen can has a flaw)

= P(Flaw on side **OR** top)

= P(Flaw on side) + P(Flaw on top)

– P(Flaw on side **AND** top)

= 0.02 + 0.03 – 0.01 = 0.04

Example 2.8 (b)

What property/formula to use? Property (b)

P(A randomly chosen can has no flaw)

= 1 – P(A randomly chosen can has a flaw)

= 1 – 0.04

= 0.96

Bonferroni Inequality

If A and B are two events in the same sample space, then
$P(A \cap B) \geq P(A) + P(B) - 1$.

Proof:

We have property (d) which says $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

We also know that $P(A \cup B) \leq 1$ by axiom 2.

So by rearranging the terms, $P(A \cap B) \geq P(A) + P(B) - 1$.

Additional Property

If events A_1, A_2, and A_3 are in the same sample space, then
$P(A_1 \cup A_2 \cup A_3) \leq P(A_1) + P(A_2) + P(A_3)$.

Proof:

$P(A_1 \cup B) = P(A_1) + P(B) - P(A_1 \cap B) \leq P(A_1) + P(B)$

Now let $B = A_2 \cup A_3$. Similar argument shows
$P(B) \leq P(A_2) + P(A_3)$.

Then we get
$P(A_1 \cup A_2 \cup A_3) \leq P(A_1) + P(B) \leq P(A_1) + P(A_2) + P(A_3)$.

COUNTING METHODS

Fundamental principle of counting

Permutations

combinations

Fundamental Principle of Counting

The Fundamental Principle of Counting

Assume that k operations are to be performed. If there are n_1 ways to perform the first operation, and if for each of these ways there are n_2 ways to perform the second operation, and if for each choice of ways to perform the first two operations there are n_3 ways to perform the third operation, and so on, then the total number of ways to perform the sequence of k operations is $n_1 n_2 \cdots n_k$.

Example 2.11

- When ordering a certain type of computer, there are 3 choices of hard drive, 4 choices for the amount of memory, 2 choices of video card, and 3 choices of monitor.

- In how many ways can a computer be ordered?
- Answer: 3*4*2*3 = 72.

Next

- Permutations and combinations (2.2)
- Conditional probability (2.3)