

Basics of Probability — Modules

1. Intro / Examples
2. Set Theory
3. Experiments and Sample Spaces
4. Definition of Probability
5. Finite Sample Spaces
6. Counting Techniques
7. Applications of Counting Techniques
8. Conditional Probability
9. Bayes' Theorem

Mathematical Models for describing observable phenomena:

- Deterministic
- Probabilistic

Deterministic Models

- Ohm's Law ($I = E/R$)
- Drop an object from height h_0 . After t sec, height is $h(t) = h_0 - 16t^2$.
- Deposit \$1000 in a continuously compounding checking 3% account. At time t , it's worth $\$1000e^{.03t}$.

Probabilistic Models — Involve uncertainty

- How much snow will fall tomorrow?
- Will IBM make a profit this year?
- Should I buy a call or put option?
- Can I win in blackjack if I use a certain strategy?

Some Cool Examples

1. Birthday Problem — Assume all 365 days have equal probability of being a person's birthday (ignore Feb 29). Then...

If there are **23** people in the room, the odds are better than 50–50 that there will be a match.

If there are **50** people, the probability is about 97%!

2. Monopoly — In the long run, the property having the highest probability of being landed on is Illinois Ave.

3. Poker — Pick 5 cards from a standard deck. Then

$\Pr(\text{exactly 2 pairs}) \approx 0.0475$

$\Pr(\text{full house}) \approx 0.00144$

$\Pr(\text{flush}) \approx 0.00198$

4. Stock Market — Monkeys randomly selecting stocks could have outperformed most market analysts during the past year.

5. A couple has two kids and at least one is a boy. What's the probability that BOTH are boys?

Possibilities: GG, BG, GB, BB. Eliminate GG since we know that there's at least one boy. Then $\Pr(\text{BB}) = 1/3$.

6. Vietnam Lottery

7. Ask Marilyn. You are a contestant at a game show. Behind one of three doors is a car; behind the other two are goats. You pick door A. Monty Hall opens door B and reveals a goat. Monty offers you a chance to switch to door C. What should you do?

Working Definitions

Probability — Methodology that describes the random variation in systems. (We'll spend about 40% of our time on this.)

Statistics — Uses sample data to draw general conclusions about the population from which the sample was taken. (60% of our time.)

The Joy of Sets

Definition: A **set** is a collection of objects. Members of a set are called **elements**.

Notation:

A, B, C, \dots for sets; a, b, c, \dots for elements

\in for membership, e.g., $x \in A$

\notin for non-membership, e.g., $x \notin A$

U is the universal set (i.e., everything)

\emptyset is the empty set.

Examples:

$$A = \{1, 2, \dots, 10\}. \quad 2 \in A, \quad 49 \notin A.$$

$$B = \{\text{basketball, baseball}\}$$

$$C = \{x \mid 0 \leq x \leq 1\} \quad (\text{"|"} \text{ means "such that"})$$

$$D = \{x \mid x^2 = 9\} = \{\pm 3\} \quad (\text{either is fine})$$

$$E = \{x \mid x \in \mathfrak{R}, x^2 = -1\} = \emptyset \quad (\mathfrak{R} \text{ is the real line})$$

Definition: If every element of set A is an element of set B then A is a **subset** of B , i.e., $A \subseteq B$.

Definition: $A = B$ iff (if and only if) $A \subseteq B$ and $B \subseteq A$.

Properties:

$$\emptyset \subseteq A; A \subseteq U; A \subseteq A$$

$$A \subseteq B \text{ and } B \subseteq C \Rightarrow (\text{implies}) A \subseteq C$$

Remark: Order of element listing is immaterial, e.g,

$$\{a, b, c\} = \{b, c, a\}.$$

Definitions: **Complement** of A with respect to U is $\bar{A} \equiv \{x | x \in U \text{ and } x \notin A\}$.

Intersection of A and B is $A \cap B \equiv \{x | x \in A \text{ and } x \in B\}$.

Union of A and B is $A \cup B \equiv \{x | x \in A \text{ or } x \in B \text{ (or both)}\}$.

(Remember Venn diagrams?)

Example:

Suppose $U = \{\text{letters of the alphabet}\}$, $A = \{\text{vowels}\}$,
and $B = \{a, b, c\}$. Then

$$\bar{A} = \{\text{consonants}\}$$

$$A \cap B = \{a\}$$

$$A \cup B = \{a, b, c, e, i, o, u\}$$

If $A \cap B = \emptyset$, then A and B are **disjoint** (or **mutually exclusive**).

Definitions:

Minus: $A - B \equiv A \cap \bar{B}$

Symmetric difference or XOR:

$$A \Delta B \equiv (A - B) \cup (B - A) = (A \cup B) - (A \cap B)$$

The **cardinality** of A , $|A|$, is the number of elements in A . A is **finite** if $|A| < \infty$.

Examples:

$A = \{3, 4\}$ is finite since $|A| = 2$.

$B = \{1, 2, 3, \dots\}$ is **countably infinite**.

$C = \{x | x \in [0, 1]\}$ is **uncountably infinite**.

Laws of Operation:

1. Complement Law: $A \cup \bar{A} = U, A \cap \bar{A} = \emptyset, \bar{\bar{A}} = A$

2. Commutative: $A \cup B = B \cup A, A \cap B = B \cap A$

3. DeMorgan's: $\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$

4. Associative: $A \cup (B \cup C) = (A \cup B) \cup C,$

$$A \cap (B \cap C) = (A \cap B) \cap C$$

5. Distributive: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Proofs: Easy. Could use Venn diagrams or many other ways.

Experiments and Sample Spaces

Consider a “random” experiment:

E_1 : Toss a die and observe the outcome.

Definition: A **sample space** associated with an experiment E is the set of *all* possible outcomes of E . It's usually denoted by S or Ω .

Examples:

E_1 has sample space $S_1 = \{1, 2, 3, 4, 5, 6\}$.

Another sample space for E_1 is $S'_1 = \{\text{odd, even}\}$.

So a sample space doesn't have to be unique!

E_2 : Toss a coin 3 times and observe the sequence of H 's and T 's.

$$S_2 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

E_3 : A new light bulb is tested to see how long it lasts.

$$S_3 = \{t | t \geq 0\}.$$

Definition: An **event** is a set of possible outcomes. Thus, any subset of S is an event.

Example (for E_1): If A_1 is the event “an even number occurs,” then $A_1 = \{2, 4, 6\}$, i.e., when the die is tossed, we get 2 or 4 or 6.

Remark: \emptyset is an event of S (“nothing happens”)
 S is an event of S (“something happens”)

Remark: If A is an event, then \bar{A} is the complementary (opposite) event.

Example (for E_1):

$A_1 = \{2, 4, 6\} \Rightarrow \bar{A}_1 = \{1, 3, 5\}$ (i.e., “an odd number occurs”)

Remark: If A and B are events, then $A \cup B$ and $A \cap B$ are events.

Example (for E_2): Let

$$\begin{aligned} A_2 &= \text{“exactly one } T \text{ was observed”} \\ &= \{HHT, HTH, THH\} \end{aligned}$$

$$B_2 = \text{“no } T\text{'s observed”} = \{HHH\}$$

$$\begin{aligned} C_2 &= \text{“first coin is } H\text{”} \\ &= \{HHH, HHT, HTH, HTT\} \end{aligned}$$

Then

$$\begin{aligned} A_2 \cup B_2 &= \text{“at most one } T \text{ observed”} \\ &= \{HHT, HTH, THH, HHH\} \end{aligned}$$

$$A_2 \cap C_2 = \{HHT, HTH\}$$

Probability Basics

Suppose A is some event for a sample space S . What's the prob that A will occur, i.e., $\Pr(A)$?

Example: Toss a fair coin. $S = \{H, T\}$. What's the prob that H will come up?

$$\Pr(\{H\}) = \Pr(H) = 1/2.$$

What does this mean?

Frequentist view: If the experiment were repeated n times, where n is very large, we'd expect about $1/2$ of the tosses to be H 's.

$$\frac{\text{Total \# of } H\text{'s out of } n \text{ tosses}}{n} \approx 1/2.$$

Example: Toss a fair die. $S = \{1, 2, 3, 4, 5, 6\}$, where each individual outcome has prob $1/6$. Then $\Pr(1, 2) = 1/3$.

Definition: With each event $A \subseteq S$, we associate a number $\Pr(A)$, called “the **probability** of A ,” satisfying the following *axioms*:

(1) $0 \leq \Pr(A) \leq 1$ (prob's are *always* betw. 0 and 1).

(2) $\Pr(S) = 1$ (prob of *some* outcome is 1). Example:

Die. $\Pr(S) = \Pr(1, 2, 3, 4, 5, 6) = 1$.

(3) If $A \cap B = \emptyset$, then $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

Example: $\Pr(1 \text{ or } 2) = \Pr(1) + \Pr(2) = 1/6 + 1/6 = 1/3$.

(4) Suppose A_1, A_2, \dots is a sequence of disjoint events (i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$). Then

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

Example: Toss a coin until the first H appears.

$$S = \{H, TH, TTH, TTTH, \dots\}.$$

Define the *disjoint* events

$$A_1 = \{H\}, A_2 = \{TH\}, A_3 = \{TTH\}, \dots$$

Then

$$1 = \Pr(S) = \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

More Nifty Properties

Theorem 1: $\Pr(\emptyset) = 0$.

Proof: Since $A \cap \emptyset = \emptyset$, we have that A and \emptyset are disjoint. So Axiom (3) implies

$$\Pr(A) = \Pr(A \cup \emptyset) = \Pr(A) + \Pr(\emptyset).$$

Note: Converse is *false*: $\Pr(A) = 0$ does *not* imply $A = \emptyset$. Example: Pick a random number betw. 0 and 1.

Theorem 2: $\Pr(\bar{A}) = 1 - \Pr(A)$.

Proof:

$$\begin{aligned} 1 &= \Pr(S) \text{ (by Axiom (2))} \\ &= \Pr(A \cup \bar{A}) \\ &= \Pr(A) + \Pr(\bar{A}) \text{ (} A \cap \bar{A} = \emptyset \text{; Axiom (3)).} \end{aligned}$$

Example: The probability that it'll rain tomorrow is one minus the probability that it won't rain.

Theorem 3: For any two events A and B ,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Proof: First observe that $B = (A \cap B) \cup (\bar{A} \cap B)$ where $A \cap B$ and $\bar{A} \cap B$ are disjoint. Thus

$$\Pr(B) = \Pr(A \cap B) + \Pr(\bar{A} \cap B) \quad (*)$$

and so

$$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(\bar{A} \cap B) \quad (A, \bar{A} \cap B \text{ are disjoint}) \\ &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \quad (\text{by } (*)). \end{aligned}$$

Remark: Can also do an easy Venn diagram proof.
(Subtract $\Pr(A \cap B)$ to avoid double-counting.)

Remark: Axiom (3) is a “special case” of this theorem
in which $A \cap B = \emptyset$.

Example: Suppose there's...

40% chance of colder weather

10% chance of rain *and* colder weather

80% chance of rain *or* colder weather.

Find the chance of rain.

$$\begin{aligned}\Pr(R) &= \Pr(R \cup C) - \Pr(C) + \Pr(R \cap C) \\ &= 0.8 - 0.4 + 0.1 = 0.5.\end{aligned}$$

Theorem 4: For any three events A , B , and C ,

$$\begin{aligned} & \Pr(A \cup B \cup C) \\ &= \Pr(A) + \Pr(B) + \Pr(C) \\ &\quad - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) \\ &\quad + \Pr(A \cap B \cap C) \end{aligned}$$

Remark: See any good prob book for a generalization to $\Pr(A_1 \cup A_2 \cup \dots \cup A_n)$. It's called the **principle of inclusion-exclusion**.

Example: 75% of Atlantans jog (J), 20% like ice cream (I), and 40% enjoy music (M). Also, 15% J and I , 30% J and M , 10% I and M , and 5% do all three. Find the probab that a random resident will engage in at least one of the three activities.

$$\begin{aligned} &P(J \cup I \cup M) \\ &= P(J) + P(I) + P(M) \\ &\quad - P(J \cap I) - P(J \cap M) - P(I \cap M) \\ &\quad + P(J \cap I \cap M) \\ &= .75 + .20 + .40 - .15 - .30 - .10 + .05 = .85. \end{aligned}$$

Find the prob of precisely one activity.

$$\begin{aligned} P(J \cap \bar{I} \cap \bar{M}) + P(\bar{J} \cap I \cap \bar{M}) + P(\bar{J} \cap \bar{I} \cap M) \\ = .35 + 0 + .05 = .40. \end{aligned}$$

How'd we get those?? Use Venn diagram, starting from the center and building out.

(Bonus) Theorem 5: $A \subseteq B \Rightarrow \Pr(A) \leq \Pr(B)$.

Proof:

$$\begin{aligned}\Pr(B) &= \Pr(A \cup (\bar{A} \cap B)) \\ &= \Pr(A) + \Pr(\bar{A} \cap B) \\ &\geq \Pr(A).\end{aligned}$$

Remark: $A \subseteq B$ means that B occurs whenever A occurs; so the Theorem makes intuitive sense.

Finite Sample Spaces

Suppose S is finite, say $S = \{a_1, a_2, \dots, a_n\}$.

Let B be an event consisting of r ($\leq n$) outcomes, i.e., $B = \{a_{j_1}, a_{j_2}, \dots, a_{j_r}\}$, where the j_i 's represent r indices from $\{1, 2, \dots, n\}$. Then $\Pr(B) = \sum_{i=1}^r \Pr(a_{j_i})$.

Note: “Choosing an object **at random**” means that each object has the same prob of being chosen.

1.5 Finite Sample Spaces

Example: You have 2 red cards, a blue card, and a yellow. Pick one card at random.

$$S = \{\text{red, blue, yellow}\} = \{a_1, a_2, a_3\}$$

$$\Pr(a_1) = 1/2, \Pr(a_2) = 1/4, \Pr(a_3) = 1/4.$$

$$\Pr(\text{red or yellow}) = \Pr(a_1) + \Pr(a_3) = 3/4.$$

Definition: A **simple sample space** (SSS) is a finite sample space in which all outcomes are *equally likely*.

Remark: In the above example, S is *not* simple since $\Pr(a_1) \neq \Pr(a_2)$.

Example: Toss 2 fair coins.

$S = \{HH, HT, TH, TT\}$ is a SSS (all prob's are $1/4$).

$S' = \{0, 1, 2\}$ (number of H 's) is *not* a SSS. Why?

1.5 Finite Sample Spaces

Theorem: For any event A in a SSS S ,

$$\Pr(A) = \frac{|A|}{|S|} = \frac{\# \text{ elements in } A}{\# \text{ elements in } S}.$$

Example: Die. $A = \{1, 2, 4, 6\}$ (each with prob $1/6$).

$$\Pr(A) = 4/6.$$

1.5 Finite Sample Spaces

Example: Roll a pair of dice. Possible results (each w.p. $1/36$):

$$\begin{array}{cccc} 1,1 & 1,2 & \dots & 1,6 \\ 2,1 & 2,2 & \dots & 2,6 \\ & \vdots & & \\ 6,1 & 6,2 & \dots & 6,6 \end{array}$$

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Counting Techniques — count the elements in events from a SSS.

Multiplication Rule

Addition Rule

Permutations

Combinations

Multiplication Rule

Two operations are performed one after the other:

(a) The first operation can be done in n_1 ways.

(b) Regardless of the way in which the first operation was performed, the second can be performed in n_2 ways.

The # ways to perform the two operations together is n_1n_2 .

1.6 Counting Techniques

Example: 3 ways to go from City A to B, and 4 ways to go from B to C. Then the you can go from A to C (via B) in 12 ways.

Example: Roll 2 dice. How many outcomes? (Assume $(3, 2) \neq (2, 3)$.) Answer is 36.

Example: Select 2 cards from a deck without replacement and *care about order* (i.e., $(Q\spadesuit, 7\clubsuit) \neq (7\clubsuit, Q\spadesuit)$). How many ways can you do this? Answer: $52 \cdot 51 = 2652$.

Example: Box of 10 sox — 2 red and 8 black. Pick 2 w/o repl.

(a) Let A be the event that both are red.

$$\Pr(A) = \frac{\# \text{ ways to pick 2 reds}}{\# \text{ ways to pick 2 sox}} = \frac{2 \cdot 1}{10 \cdot 9} = \frac{1}{45}.$$

(b) Let B be the event that both are black.

$$\Pr(B) = \frac{8 \cdot 7}{10 \cdot 9} = \frac{28}{45}.$$

(c) Let C be one of each color.

$$\begin{aligned}\Pr(C) &= 1 - \Pr(\bar{C}) \\ &= 1 - \Pr(A \cup B) \\ &= 1 - \Pr(A) - \Pr(B) \quad (A \text{ and } B \text{ disjoint}) \\ &= 16/45.\end{aligned}$$

1.6 Counting Techniques

Remark: The multiplication rule can be extended to more than 2 operations.

Example: Flip 3 coins. $2 \times 2 \times 2 = 8$ possible outcomes.

Example: Toss n dice. 6^n possible outcomes.

Addition Rule

Can use method A in n_A ways.

Can use method B in n_B ways.

If only one method can be used, you have $n_A + n_B$ ways of doing so.

Example: Go to Starbucks and have a muffin (blueberry or oatmeal) or a bagel (sesame, plain, salt), but not both. $2 + 3 = 5$ choices.

Permutations

Definition: An arrangement of n symbols in a *definite order* is a **permutation** of the n symbols.

Example: How many ways to arrange the numbers 1,2,3? Answer: 6 ways — 123, 132, 213, 231, 312, 321.

Example: How many ways to arrange $1, 2, \dots, n$?

(choose first)(choose second) \cdots (choose n th)

$$n(n - 1)(n - 2) \cdots 2 \cdot 1 = n!.$$

Example: Baseball manager has 9 players on his team. Find the # of possible batting orders. Answer: $9! = 362880$.

Definition: The # of r -tuples we can make from n different symbols (each used at most once) is called the # of **permutations of n things taken r -at-a-time**,

$$P_{n,r} \equiv \frac{n!}{(n-r)!} \quad (*).$$

Note that $0! = 1$ and $P_{n,n} = n!$.

Example: How many ways can you take two symbols from a, b, c, d ? Ans: $P_{4,2} = 4!/2! = 12$ — $ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc$.

Proof (of (*)):

$$\begin{aligned} P_{n,r} &= (\text{choose first})(\text{second})\cdots(r\text{th}) \\ &= n(n-1)(n-2)\cdots(n-r+1) \\ &= \frac{n(n-1)\cdots(n-r+1)(n-r)\cdots 2 \cdot 1}{(n-r)\cdots 2 \cdot 1} \\ &= \frac{n!}{(n-r)!}. \end{aligned}$$

1.6 Counting Techniques

Example: How many ways to fill the first 4 positions of a batting order?

$n = 9$ players, $r = 4$ positions.

$$P_{9,4} = 9! / (9 - 4)! = 3024 \text{ ways.}$$

Example: How many of these 3024 ways has Smith batting first?

Method 1: First 4 positions: (Smith,?,?,?). This is equiv to choosing 3 players from the remaining 8.

$$P_{8,3} = 8!/(8 - 3)! = 336 \text{ ways.}$$

Method 2: It's clear that each of the 9 players is equally likely to bat first. Thus, $3024/9 = 336$.

Example: How many license plates of 6 digits can be made from the numbers 1,2,...,9...

(a) with no repetitions? (e.g., 123465) $P_{9,6} = 9!/3! = 60480$.

(b) allowing repetitions? (e.g., 123345 or 123465)
 $9 \times 9 \times \dots \times 9 = 9^6 = 531441$.

(c) containing repetitions? $531441 - 60480 = 470961$.

Combinations

Suppose we only want to count the number of ways to choose r out of n objects *without* regard to order, i.e., count the number of different subsets of these n objects that contain exactly r objects.

Example: How many subsets of $\{1, 2, 3\}$ contain exactly 2 elements? (Order isn't important.)

3 subsets — $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$

Definition: The # of subsets with r elements of a set with n elements is called the number of **combinations of n things taken r -at-a-time**.

Notation: $\binom{n}{r}$ or $C_{n,r}$ (read as “ n choose r ”). These are also called **binomial coefficients**.

Difference between permutations and combinations:

Combinations — not concerned w/order: $(a, b, c) = (b, a, c)$.

Permutations — concerned w/order: $(a, b, c) \neq (b, a, c)$.

The number of permutations of n things taken r -at-a-time is always at least as large as the number of combinations. In fact, . . .

Remark: Choosing a permutation is the same as first choosing a combination *and* then putting the elements in order, i.e.,

$$\frac{n!}{(n-r)!} = \binom{n}{r} r!$$

So

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

$$\binom{n}{r} = \binom{n}{n-r}, \quad \binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n.$$

Binomial Theorem:

$$(x + y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$$

This is where Pascal's \triangle comes from!

Corollary: Surprising fact:

$$\sum_{i=0}^n \binom{n}{i} = 2^n.$$

Proof:

$$2^n = (1 + 1)^n = \sum_{i=0}^n \binom{n}{i} 1^i 1^{n-i}.$$

Example: An NBA team has 12 players. How many ways can the coach choose the starting 5?

$$\binom{12}{5} = \frac{12!}{5!7!} = 792.$$

Example: Smith is one of the players on the team. How many of the 792 starting line-ups include him?

$$\binom{11}{4} = \frac{11!}{4!7!} = 330.$$

(Smith gets one of the five positions for free; there are now 4 left to be filled by the remaining 11 players.)

1.6 Counting Techniques

Example: 7 red shoes, 5 blues. Find the number of arrangements.

R B R R B B R R R B R B

I.e., how many ways to put 7 reds in 12 slots?

Answer: $\binom{12}{7}$.

Some applications of counting techniques.

Hypergeometric problems

Permutations vs. Combinations

Birthday problem

Poker probabilities

Multinomial coefficients

Hypergeometric Distribution

You have a objects of type 1 and b objects of type 2.

Select n objects w/o replacement from the $a + b$.

$$\begin{aligned}
 & \text{Pr}(k \text{ type 1's were picked}) \\
 &= \frac{(\# \text{ ways to choose } k \text{ 1's})(\text{choose } n - k \text{ 2's})}{\# \text{ ways to choose } n \text{ out of } a + b} \\
 &= \frac{\binom{a}{k} \binom{b}{n - k}}{\binom{a + b}{n}} \text{ (the **hypergeometric distr'n**).}
 \end{aligned}$$

1.7 Counting Applications

Example: 25 sox in a box. 15 red, 10 blue. Pick 7 w/o replacement.

$$\Pr(\text{exactly 3 reds are picked}) = \frac{\binom{15}{3} \binom{10}{4}}{\binom{25}{7}}$$

Permutations vs. Combinations — It's all how you approach the problem!

Example: 4 red marbles, 2 whites. Put them in a row in random order. Find...

- (a) $\Pr(2 \text{ end marbles are } W)$
- (b) $\Pr(2 \text{ end marbles aren't both } W)$
- (c) $\Pr(2 \text{ } W\text{'s aren't side by side})$

Method 1 (using permutations): Let the sample space

$$S = \{\text{every random ordering of the 6 marbles}\}.$$

(a) A : 2 end marbles are W — $WRRRRW$.

$$|A| = 2!4! = 48 \Rightarrow \Pr(A) = \frac{|A|}{|S|} = \frac{48}{720} = \frac{1}{15}.$$

(b) $\Pr(\bar{A}) = 1 - \Pr(A) = 14/15$.

(c) B : 2 W's side by side — WWRRRR or RWWRRR
or ... or RRRRWW

$$\begin{aligned}
 |B| &= (\# \text{ ways to select pair of slots for 2 W's}) \\
 &\quad \times (\# \text{ ways to insert W's into pair of slots}) \\
 &\quad \times (\# \text{ ways to insert R's into remaining slots}) \\
 &= 5 \times 2! \times 4! = 240.
 \end{aligned}$$

$$\Pr(B) = \frac{|B|}{|S|} = \frac{240}{720} = \frac{1}{3}.$$

But — The above method took too much time! Here's an easier way...

Method 2 (using combinations): Which 2 positions do the W 's occupy? Now let

$S = \{\text{possible pairs of slots that the } W\text{'s occupy}\}.$

Clearly, $|S| = \binom{6}{2} = 15.$

(a) Since the W 's must occupy the end slots in order for A to occur, $|A| = 1 \Rightarrow \Pr(A) = |A|/|S| = 1/15.$

(b) $\Pr(\bar{A}) = 14/15.$

(c) $|B| = 5 \Rightarrow \Pr(B) = 5/15 = 1/3.$

Birthday Problem

n people in a room. Find the prob that at least two have the same birthday. (Ignore Feb. 29, and assume that all 365 days have equal prob.)

A : All birthdays are different.

$S = \{(x_1, \dots, x_n) : x_i = 1, 2, \dots, 365\}$ (x_i is person i 's birthday), and note that $|S| = (365)^n$.

$$|A| = P_{365,n} = (365)(364) \cdots (365 - n + 1)$$

$$\begin{aligned} \Pr(A) &= \frac{(365)(364) \cdots (365 - n + 1)}{(365)^n} \\ &= 1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365} \end{aligned}$$

We want

$$\Pr(\bar{A}) = 1 - \left(1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365} \right)$$

1.7 Counting Applications

Notes: When $n = 366$, $\Pr(\bar{A}) = 1$.

For $\Pr(\bar{A})$ to be $> 1/2$, n must be ≥ 23 . (surprising)

When $n = 50$, $\Pr(\bar{A}) = 0.97$.

Poker Problems

Draw 5 cards at random from a standard deck.

of possible hands is $|S| = \binom{52}{5} = 2,598,960$.

Terminology:

rank = 2, 3, ..., Q, K, A,

suit = , , , 

(a) 2 pairs — e.g., $A\heartsuit, A\clubsuit, 3\heartsuit, 3\diamondsuit, 10\spadesuit$

Select 2 ranks (e.g., $A, 3$). Can do this $\binom{13}{2}$ ways.

Select 2 suits for first pair (e.g., \heartsuit, \clubsuit). $\binom{4}{2}$ ways.

Select 2 suits for second pair (e.g., \heartsuit, \diamondsuit). $\binom{4}{2}$ ways.

Select remaining card to complete the hand. 44 ways.

$$|2 \text{ pairs}| = \binom{13}{2} \binom{4}{2} \binom{4}{2} 44 = 123,552$$

$$\Pr(2 \text{ pairs}) = \frac{123,552}{2,598,960} \approx 0.0475.$$

(b) Full house (1 pair, 3-of-a-kind) —
 e.g., $A\heartsuit, A\clubsuit, 3\heartsuit, 3\diamonds, 3\spadesuit$

Select 2 *ordered* ranks (e.g., $A, 3$). $P_{13,2}$ ways.

Select 2 suits for pair (e.g., \heartsuit, \clubsuit). $\binom{4}{2}$ ways.

Select 3 suits for 3-of-a-kind (e.g., $\heartsuit, \diamonds, \spadesuit$). $\binom{4}{3}$ ways.

$$|\text{full house}| = 13 \cdot 12 \binom{4}{2} \binom{4}{3} = 3744$$

$$\Pr(\text{full house}) = \frac{3744}{2,598,960} \approx 0.00144.$$

(c) Flush (all 5 cards from same suit)

Select a suit. $\binom{4}{1}$ ways.

Select 5 cards from that suit. $\binom{13}{5}$ ways.

$$\Pr(\text{flush}) = \frac{5148}{2,598,960} \approx 0.00198.$$

(d) Straight (5 ranks in a row)

Select a starting point for the straight ($A, 2, 3, \dots, 10$).

$\binom{10}{1}$ ways.

Select a suit for each card in the straight. 4^5 ways.

$$\Pr(\text{straight}) = \frac{10 \cdot 4^5}{2,598,960} \approx 0.00394.$$

(e) Straight flush

Select a starting point for the straight. 10 ways.

Select a suit. 4 ways.

$$\Pr(\text{straight flush}) = \frac{40}{2,598,960} \approx 0.0000154.$$

Multinomial Coefficients

Example: n_1 blue sox, n_2 reds. # of assortments is $\binom{n_1 + n_2}{n_1}$ (binomial coefficients).

Generalization (for k types of objects): $n = \sum_{i=1}^k n_i$

of arrangements is $n!/(n_1!n_2!\cdots n_k!)$.

1.7 Counting Applications

Example: How many ways can “Mississippi” be arranged?

$$\frac{\# \text{ perm's of 11 letters}}{(\# m's)! (\# p's)! (\# i's)! (\# s's)!} = \frac{11!}{1!2!4!4!} = 34,650.$$

Conditional Probability

Definition

Properties

Independence

Conditional Probability

Example: Die. $A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4, 5\}$. So $\Pr(A) = 1/2$, $\Pr(B) = 5/6$.

Suppose we *know* that B occurs. Then the prob of A “given” B is

$$\Pr(A|B) = \frac{2}{5} = \frac{|A \cap B|}{|B|}$$

So the prob of A depends on the info that you have! The info that B occurs allows us to regard B as a new, restricted sample space. And...

$$\Pr(A|B) = \frac{|A \cap B|}{|B|} = \frac{|A \cap B|/|S|}{|B|/|S|} = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

Definition: If $\Pr(B) > 0$, the **conditional prob of A given B** is $\Pr(A|B) \equiv \Pr(A \cap B)/\Pr(B)$.

Remarks: If A and B are disjoint, then $\Pr(A|B) = 0$. (If B occurs, there's no chance that A can also occur.)

What happens if $\Pr(B) = 0$? Don't worry! In this case, makes no sense to consider $\Pr(A|B)$.

Example: Toss 2 dice and take the sum.

A : odd toss = $\{3, 5, 7, 9, 11\}$

B : $\{2, 3\}$

$$\Pr(A) = \Pr(3) + \dots + \Pr(11) = \frac{2}{36} + \frac{4}{36} + \dots + \frac{2}{36} = \frac{1}{2}.$$

$$\Pr(B) = \frac{1}{36} + \frac{2}{36} = \frac{1}{12}.$$

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(3)}{\Pr(B)} = \frac{2/36}{1/12} = 2/3.$$

Example: 4 white sox, 8 red. Select 2 w/o repl.

A : 1st sock W; B : 2nd W; C : Both W ($= A \cap B$).

$$\Pr(C) = \Pr(A \cap B) = \Pr(A)\Pr(B|A) = \frac{4}{12} \cdot \frac{3}{11} = \frac{1}{11}.$$

$$\begin{aligned} \Pr(B) &= \Pr(A \cap B) + \Pr(\bar{A} \cap B) \\ &= \Pr(A)\Pr(B|A) + \Pr(\bar{A})\Pr(B|\bar{A}) \\ &= \frac{4}{12} \cdot \frac{3}{11} + \frac{8}{12} \cdot \frac{4}{11} = \frac{1}{3}. \end{aligned}$$

Could you have gotten this result w/o thinking?

A couple has two kids and at least one is a boy.
What's the prob that BOTH are boys?

$S = \{GG, GB, BG, BB\}$, (' BG ' means 'boy then girl')

C : Both are boys = $\{BB\}$.

D : At least 1 is a boy. = $\{GB, BG, BB\}$.

$$\Pr(C|D) = \frac{\Pr(C \cap D)}{\Pr(D)} = \frac{\Pr(C)}{\Pr(D)} = 1/3.$$

(My intuition was $1/2$ — the *wrong* answer! The problem was that we didn't know whether D meant the first or second kid.)

Properties — analogous to Axioms of probability.

$$(1') \quad 0 \leq \Pr(A|B) \leq 1.$$

$$(2') \quad \Pr(S|B) = 1.$$

$$(3') \quad A_1 \cap A_2 = \emptyset \Rightarrow \Pr(A_1 \cup A_2|B) = \Pr(A_1|B) + \Pr(A_2|B).$$

(4') If A_1, A_2, \dots are all disjoint, then

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i \mid B\right) = \sum_{i=1}^{\infty} \Pr(A_i|B).$$

Independence Day — Any unrelated events are independent.

A : It rains on Mars tomorrow.

B : Coin lands on H .

Definition: A and B are **independent** iff $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

Example: If $\Pr(\text{rains on Mars}) = 0.2$ and $\Pr(H) = 0.5$, then $\Pr(\text{rains and } H) = 0.1$.

Note: If $\Pr(A) = 0$, then A is indep of any other event.

Remark: Events don't have to be physically unrelated to be indep.

Example: Die. $A = \{2, 4, 6\}$, $B = \{1, 2, 3, 4\}$, $A \cap B = \{2, 4\}$, so $\Pr(A) = 1/2$, $\Pr(B) = 2/3$, $\Pr(A \cap B) = 1/3$.

$\Pr(A)\Pr(B) = 1/3 = \Pr(A \cap B) \Rightarrow A, B$ indep.

More natural interpretation of independence...

Theorem: Suppose $\Pr(B) > 0$. Then A and B are indep iff $\Pr(A|B) = \Pr(A)$.

Proof: A, B indep $\Leftrightarrow \Pr(A \cap B) = \Pr(A)\Pr(B) \Leftrightarrow \Pr(A \cap B)/\Pr(B) = \Pr(A)$.

Remark: So if A and B are indep, the prob of A doesn't depend on whether or not B occurs.

(Bonus) Theorem: A, B indep $\Rightarrow A, \bar{B}$ indep.

Proof: $\Pr(A) = \Pr(A \cap \bar{B}) + \Pr(A \cap B)$, so that

$$\begin{aligned}\Pr(A \cap \bar{B}) &= \Pr(A) - \Pr(A \cap B) \\ &= \Pr(A) - \Pr(A)\Pr(B) \quad (A, B \text{ indep}) \\ &= \Pr(A)[1 - \Pr(B)] = \Pr(A)\Pr(\bar{B}).\end{aligned}$$

Don't confuse independence with disjointness!

Theorem: If $\Pr(A) > 0$ and $\Pr(B) > 0$, then A and B can't be indep and disjt at the same time.

Proof: A, B disjt ($A \cap B = \emptyset$) $\Rightarrow \Pr(A \cap B) = 0 < \Pr(A)\Pr(B)$. Thus A, B not indep. Similarly, indep implies not disjt.

Extension to more than two events.

Definition: A, B, C are indep iff

$$(a) \Pr(A \cap B \cap C) = \Pr(A)\Pr(B)\Pr(C)$$

and

(b) All *pairs* must be indep:

$$\Pr(A \cap B) = \Pr(A)\Pr(B)$$

$$\Pr(A \cap C) = \Pr(A)\Pr(C)$$

$$\Pr(B \cap C) = \Pr(B)\Pr(C)$$

Note that condition (a) by itself isn't enough.

Example: $S = \{1, 2, \dots, 8\}$ (each element w.p. $1/8$).

$A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6, 7\}$, $C = \{1, 2, 3, 8\}$.

(a) $A \cap B \cap C = \{1\}$. $\Pr(A \cap B \cap C) = \Pr(A)\Pr(B)\Pr(C) = 1/8$, so (a) is satisfied. However, (b) is *not*...

(b) $A \cap B = \{1\}$. $\Pr(A \cap B) = 1/8 \neq 1/4 = \Pr(A)\Pr(B)$.

Also note that (b) by itself isn't enough.

Example: $S = \{1, 2, 3, 4\}$ (each element w.p. $1/4$).

$A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$. $A \cap B = A \cap C = B \cap C = A \cap B \cap C = \{1\}$.

(b) $\Pr(A \cap B) = 1/4 = \Pr(A)\Pr(B)$. Same deal with A, C and B, C . So (b) is OK. But (a) *isn't*...

(a) $\Pr(A \cap B \cap C) = 1/4 \neq 1/8 = \Pr(A)\Pr(B)\Pr(C)$.

General Definition: A_1, \dots, A_k are indep iff
 $\Pr(A_1 \cap \dots \cap A_k) = \Pr(A_1) \cdots \Pr(A_k)$ *and*
all subsets of $\{A_1, \dots, A_k\}$ are indep.

Independent Trials: Perform n trials of an experiment such that the outcome of one trial is indep of the outcomes of the other trials.

Example: Flip 3 coins indep'ly.

(a) $\Pr(\text{1st coin is } H) = 1/2$. Don't worry about the other two coins since they're indep of the 1st.

(b) $\Pr(\text{1st coin } H, \text{ 3rd } T) = \Pr(\text{1st coin } H)\Pr(\text{3rd } T) = 1/4$.

Remark: For indep trials, you just multiply the individual probs.

Example: Flip a coin infinitely many times (each flip is indep of the others).

$$\begin{aligned} p_n &\equiv \Pr(\text{1st } H \text{ on } n\text{th trial}) \\ &= \Pr(\underbrace{TT \cdots T}_{n-1} H) \\ &= \underbrace{\Pr(T)\Pr(T) \cdots \Pr(T)}_{n-1} \Pr(H) = 1/2^n. \end{aligned}$$

$$\Pr(H \text{ eventually}) = \sum_{n=1}^{\infty} p_n = 1.$$

Bayes' Theorem

Partitions

Bayes' Theorem

Examples

Partition of a sample space — split the sample space into disjoint, yet all-encompassing subsets.

Definition: The events A_1, A_2, \dots, A_n form a **partition** of the sample space S if

- (1) A_1, A_2, \dots, A_n are disjoint.
- (2) $\bigcup_{i=1}^n A_i = S$.
- (3) $\Pr(A_i) > 0$ for all i .

Remark: When an experiment is performed, *exactly one* of the A_i 's occurs.

Example: A and \bar{A} form a partition.

Example: “vowels” and “consonants” form a partition of the letters.

Example: Suppose A_1, A_2, \dots, A_n form a partition of S , and B is some arbitrary event. Then

$$B = \bigcup_{i=1}^n (A_i \cap B).$$

So if A_1, A_2, \dots, A_n is a partition,

$$\begin{aligned}\Pr(B) &= \Pr\left(\bigcup_{i=1}^n (A_i \cap B)\right) \\ &= \sum_{i=1}^n \Pr(A_i \cap B) \quad (A_1, A_2, \dots, A_n \text{ are disjoint}) \\ &= \sum_{i=1}^n \Pr(A_i) \Pr(B|A_i) \quad (\text{by defn of cond'l prob}).\end{aligned}$$

This is the **Law of Total Probability**.

Example: $\Pr(B) = \Pr(A)\Pr(B|A) + \Pr(\bar{A})\Pr(B|\bar{A})$,
which we saw in the last module.

Bayes Theorem: If A_1, A_2, \dots, A_n form a partition of S and B is any event, then

$$\Pr(A_i|B) = \frac{\Pr(A_i \cap B)}{\Pr(B)} = \frac{\Pr(A_i)\Pr(B|A_i)}{\sum_{j=1}^n \Pr(A_j)\Pr(B|A_j)}.$$

The $\Pr(A_i)$'s are **prior** probabilities (“before B ”).

The $\Pr(A_i|B)$'s are **posterior** probabilities (“after B ”).

The $\Pr(A_i|B)$ add up to 1. That's why the funny-looking denominator.

Example: In a certain city with good police,
 $\Pr(\text{Any defendent brought to trial is guilty}) = 0.99.$

In any trial,

$\Pr(\text{Jury acquits if defendent is innocent}) = 0.95.$

$\Pr(\text{Jury convicts if defendent is guilty}) = 0.95.$

Find $\Pr(\text{Defendent is innocent} | \text{Jury acquits}).$

Events: I = “innocent”, G = “guilty” = \bar{I} , A = “acquittal”. Since the partition is $\{I, G\}$, Bayes' \Rightarrow

$$\begin{aligned}\Pr(I|A) &= \frac{\Pr(I)\Pr(A|I)}{\Pr(I)\Pr(A|I) + \Pr(G)\Pr(A|G)} \\ &= \frac{(0.01)(0.95)}{(0.01)(0.95) + (0.99)(0.05)} \\ &= 0.161.\end{aligned}$$

Notice how the posterior prob's depend strongly on the prior prob's.

Example: A store gets $1/2$ of its items from Factory 1, $1/4$ from Factory 2, and $1/4$ from Factory 3.

2% of Factory 1's items are defective.

2% of Factory 2's items are defective.

4% of Factory 3's items are defective.

An item from the store is found to be bad. Find the prob it comes from Factory 1. [Answer should be $< 1/2$ since bad items favor Factory 3.]

Events: $F_i =$ "Factory i ", $D =$ "defective item". Partition is $\{F_1, F_2, F_3\}$.

$$\begin{aligned}\Pr(F_1|D) &= \frac{\Pr(F_1)\Pr(D|F_1)}{\sum_{j=1}^3 \Pr(F_j)\Pr(D|F_j)} \\ &= \frac{(0.5)(0.02)}{(0.5)(0.02) + (0.25)(0.02) + (0.25)(0.04)} \\ &= 0.4.\end{aligned}$$

It turns out that $\Pr(F_2|D) = 0.2$ and $\Pr(F_3|D) = 0.4$.