ST 501: Fundamentals Statistical Inference I

Axioms of probability and counting techniques

Minh Tang

Fall 2024

Department of Statistics, North Carolina State University.

Sample space and events

The standard mathematical treatment of probability is via Kolmogorov's axioms of probability.

In this setup there are

1. a sample space Ω
2. a collection of subsets $\mathcal{F}$ of Ω with $\Omega \in \mathcal{F}$.
3. a countably additive function $P\colon \mathcal{F} \mapsto [0,1]$ with $P(\Omega) = 1$.

In this first part of the course, we shall assume that the sample space Ω is discrete, i.e.,

- either $|\Omega|$ is finite
- if not, then Ω is countable

In this first part of the course, we shall assume that the sample space Ω is discrete, i.e.,

- either $|\Omega|$ is finite
- if not, then Ω is countable

We then take $\mathcal{F} = 2^{\Omega}$, i.e., $\mathcal{F}$ is the set of all subsets of Ω.

Definition

Under the above discrete setup, an event A is any $A \subset \Omega$.

- Events A with $|A| = 1$ are simple events.
- Events A with $|A| > 1$ are compound events.

Example If we flip a coin 3 times, we can define

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

where HTT denote that the first coin flip is a head and the second and third coin flips are tails. In addition, $\mathcal{F}$ is the set containing $2^8 = 256$ subsets of Ω.

Example If we measure the number of bacterias in a petri dish, then we can let

$$\Omega = \{0, 1, 2, 3, \dots\} = \mathbb{N}.$$

The sample space Ω is countable. We then let $\mathcal{F}$ be the set of all subsets of $\mathbb{N}$.

Example If we measure the time until failure of an engine (and suppose that we can measure time as precise as we wants), then we can let $\Omega = \{x \in \mathbb{R} \colon x \geq 0\}$, i.e., the sample space is the set of all non-negative real numbers.

Ω is now not countable and we cannot let $\mathcal{F}$ be the set of all subsets of Ω.

Set operations

Let A and B be subsets of some X. Recall the operations

- (Intersection) $A \cap B = \{x \in X : x \in A \text{ and } x \in B\}$.
- (Union) $A \cup B = \{x \in X : x \in A \text{ or } x \in B\}$ (inclusive or).
- (Complement) $\bar{A} = A^c = A' = \{x \in X : x \notin A\}$.
- (Difference) $A \setminus B = A \cap \bar{B} = \{x \in X : x \in A \text{ and } x \notin B\}$.
- (Symmetric difference) $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

$A \triangle B$ is the set of elements x belonging to exactly one of A or B (exclusive or).

Distributive law and De Morgan's law

Proposition

Let A, B, and C be subsets of some X. Then

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (*distributive law*)
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (*distributive law*)
- $\overline{A \cup B} = (A \cup B)^c = \bar{A} \cap \bar{B}$ (*de Morgan's law*)
- $\overline{A \cap B} = (A \cap B)^c = \bar{A} \cup \bar{B}$ (*de Morgan's law*)

The above laws can be generalized to multiple sets, e.g.,

$$\begin{aligned} A \cap (B_1 \cup B_2 \cup B_3) &= (A \cap B_1) \cup (A \cap (B_2 \cup B_3)) \\ &= (A \cap B_1) \cup ((A \cap B_2) \cup (A \cap B_3)) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \end{aligned}$$

The above laws can be generalized to multiple sets, e.g.,

$$\begin{aligned} A \cap (B_1 \cup B_2 \cup B_3) &= (A \cap B_1) \cup (A \cap (B_2 \cup B_3)) \\ &= (A \cap B_1) \cup ((A \cap B_2) \cup (A \cap B_3)) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \end{aligned}$$

In general, we have

$$A \cap (\bigcup_{i=1}^{n} B_i) = A \cap (B_1 \cup B_2 \cup \cdots \cup B_n) = \bigcup_{i=1}^{n} (A \cap B_i)$$

$$A \cup (\bigcap_{i=1}^{n} B_i) = A \cup (B_1 \cap B_2 \cap \cdots \cap B_n) = \bigcap_{i=1}^{n} (A \cup B_i)$$

$$\overline{\bigcup_{i=1}^{n} B_i} = \overline{B_1 \cup B_2 \cup \cdots \cup B_n} = \bigcap_{i=1}^{n} \bar{B}_i = \bar{B}_1 \cap \bar{B}_2 \cap \cdots \cap B_n$$

$$\overline{\bigcap_{i=1}^{n} B_i} = \overline{B_1 \cap B_2 \cap \cdots \cap B_n} = \bigcup_{i=1}^{n} \bar{B}_i = \bar{B}_1 \cup \bar{B}_2 \cup \cdots \cup B_n$$

Example: Simple Events

1. A coin is tossed three times and the sequence of heads and tails is recorded.

a. List the sample space.

b. List the elements that make up the following events: (1) A = at least two heads, (2) B = the first two tosses are heads, (3) C = the last toss is a tail.

c. List the elements of the following events: (1) A^c, (2) $A \cap B$, (3) $A \cup C$.

10

Axioms of probability

Definition

Let Ω be a discrete sample space and $\mathcal{F} = 2^{\Omega}$. A probability measure P on Ω is a function $P \colon \mathcal{F} \mapsto [0, 1]$ satisfying

1. $P(A) \geq 0$ for all $A \subset \Omega$
2. $P(\Omega) = 1$
3. For any sequence of (pairwise) disjoint events $A_1, A_2, \ldots$

$$P\Big(\bigcup_{i \geq 1} A_i\Big) = \sum_{i \geq 1} P(A_i) \qquad \text{(countable additivity)}$$

A sequence $A_1, A_2, \ldots$ is pairwise disjoint if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.

The above axioms implies the following claims.

Claim 1 $P(\emptyset) = 0$.

12

The above axioms implies the following claims.

Claim 1 $P(\emptyset) = 0$.

Indeed, let $A_1 = \Omega$ and let $A_i = \emptyset$ for all $i \geq 2$. Then

$$
\begin{aligned}
1 = P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \cup \emptyset \cup \dots) &= P(\bigcup_{i \geq 1} A_i) \\
&= \sum_{i \geq 1} P(A_i) = P(\Omega) + \sum_{i \geq 2} P(\emptyset) = 1 + \sum_{i \geq 2} P(\emptyset)
\end{aligned}
$$

and hence $P(\emptyset) = 0$.

Claim 2 For any finite set of disjoint events $A_1, A_2, \ldots, A_n$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \qquad \text{(finite additivity)}$$

Claim 2 For any finite set of disjoint events $A_1, A_2, \ldots, A_n$,

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) \qquad \text{(finite additivity)}$$

Indeed, let $A_j = \emptyset$ for $j \geq n+1$ and write

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) = P(\bigcup_{i\geq 1} A_i) = \sum_{i\geq 1} P(A_i) \\ = \sum_{i=1}^{n} P(A_i) + \sum_{i\geq n+1} P(\emptyset) = \sum_{i=1}^{n} P(A_i) \end{aligned}$$

Q. Why did we assume countable additivity and not just finite additivity ?

14

Q. Why did we assume countable additivity and not just finite additivity ?

A. Suppose $\Omega = \mathbb{N}$ and define P so that

- $P(\Omega) = 1$
- $P(\{i\}) = 0$ for all $i \in \mathbb{N}$.

Then P does not satisfies countable additivity as

$$1 = P(\Omega) = P\Big(\bigcup_{i \geq 1} \{i\}\Big) = \sum_{i \geq 1} P(\{i\}) = \sum_{i \geq 1} 0$$

which is a contradiction.

However, P as defined does satisfies finite additivity.

Claim 3 For any $A \subset \Omega$,

$$P(\bar{A}) = 1 - P(A).$$

15

Claim 3 For any $A \subset \Omega$,

$$P(\bar{A}) = 1 - P(A).$$

Indeed, $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

15

Claim 3 For any $A \subset \Omega$,

$$P(\bar{A}) = 1 - P(A).$$

Indeed, $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

Claim 4a For any events A and B

$$P(A) = P(A \cap B) + P(A \cap \bar{B}).$$

15

Claim 3 For any $A \subset \Omega$,

$$P(\bar{A}) = 1 - P(A).$$

Indeed, $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$.

Claim 4a For any events A and B

$$P(A) = P(A \cap B) + P(A \cap \bar{B}).$$

Indeed,

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap (B \cup \bar{B})) \\ &= P((A \cap B) \cup (A \cap \bar{B})) = P(A \cap B) + P(A \cap \bar{B}) \end{aligned}$$

as $A \cap B$ and $A \cap \bar{B}$ are disjoint.

15

Claim 4b(Partitioning Rule) Let B be an event and let $A_1, A_2, \ldots$ be a countable sequence of (pairwise) disjoint sets such that $\cup_i A_i = \Omega$. Then

$$P(B) = P(B \cap \Omega) = P(B \cap (\bigcup_i A_i)) = P(\bigcup_i (A_i \cap B)) = \sum_i P(A_i \cap B)$$

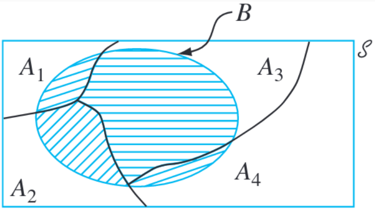


Figure 2.11 Partition of *B* by mutually exclusive and exhaustive A_i's

Claim 5 For any events A and B

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

17

Claim 5 For any events A and B

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

Indeed, $A \cup B = A \cup (B \cap \bar{A})$. Now $A \cap (B \cap \bar{A}) = \emptyset$ and hence, by Claim 4

$$P(A \cup B) = P(A) + P(B \cap \bar{A}) = P(A) + P(B) - P(B \cap A).$$

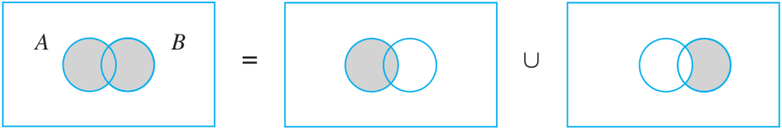


Figure 2.4 Representing A ∪ B as a union of disjoint events

Claim 5 can also be extended to three or more events.

Example For any three events A, B, and C

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

See problem 1.6 in your textbook. This is a special case of the inclusion-exclusion principle.

Example

A news magazine publishes three columns entitled "Art" (A), "Books" (B), and "Cinema" (C). Reading habits of a randomly selected reader with respect to these columns are

Read regularly	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
Probability	.14	.23	.37	.08	.09	.13	.05

Figure 2.9 illustrates relevant probabilities.

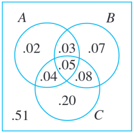


Figure 2.9 Venn diagram for Example 2.26

We have

$$P(A \cup B \cup C) = .14 + .23 + .37 - .08 - .09 - .13 + .05 = .49$$

$$P(A \cap B \cap \bar{C}) = P(A \cap B) - P(A \cap B \cap C) = .08 - .05 = .03$$

$$P(A \cap \bar{B} \cap C) = P(A \cap C) - P(A \cap B \cap C) = .09 - .05 = .04$$

Using the partition rule we have

$$P(A) = P(A \cap B \cap C) + P(A \cap B \cap \bar{C}) + P(A \cap \bar{B} \cap C) \\ + P(A \cap \bar{B} \cap \bar{C})$$

Using the above numbers we have

$$P(A \cap \bar{B} \cap \bar{C}) = .14 - .05 - .03 - .04 = .02$$

20

Two inequalities

From Claim 5 we have the following useful inequality.

Claim 6 (Boole's inequality) For any events A and B

$$P(A \cup B) \leq P(A) + P(B) \qquad \text{(sub-additivity)}$$

21

Two inequalities

From Claim 5 we have the following useful inequality.

Claim 6 (Boole's inequality) For any events A and B

$$P(A \cup B) \leq P(A) + P(B) \qquad \text{(sub-additivity)}$$

Claim 6 implies (see problem 1.8 in your textbook)

$$P(\bigcup_i A_i) \leq \sum_i P(A_i)$$

for any countable sequence of events $A_1, A_2, \ldots$

Two inequalities

From Claim 5 we have the following useful inequality.

Claim 6 (Boole's inequality) For any events A and B

$$P(A \cup B) \leq P(A) + P(B) \qquad \text{(sub-additivity)}$$

Claim 6 implies (see problem 1.8 in your textbook)

$$P(\bigcup_i A_i) \leq \sum_i P(A_i)$$

for any countable sequence of events $A_1, A_2, \ldots$

Claim 6, Claim 3, and de Morgan's law together implies

Claim 7 For any countable sequence of events $A_1, A_2, \ldots$

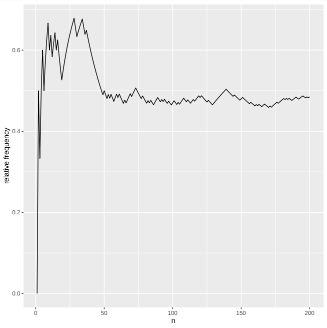
$$P(\bigcap_i A_i) = 1 - P(\overline{\bigcap_i A_i}) = 1 - P(\bigcup_i \bar{A}_i) \geq 1 - \sum_i P(\bar{A}_i).$$

Interpretations of Probability

Some of the most common interpretations are

- Long-term frequency (or objective probability).
- Subjective probability.

Example We throw a *normal* coin n times and let $\hat{p}$ be the *proportion* of time head show up among these n throws.



The main principle behind objective probability is that the probability of any event can be derived (or computed) exactly from some simple set of conditions/assumptions.

Subjective probability is based on opinions or ”gut feelings”.

Examples of these type of statements are

- You are likely to pass the exam.
- The housing market is expected to be even more competitive in 2022.
- Cloudy, with a chance of meatballs.

23

Example: Sum of m die

Suppose we throw $m = 2$ die.

The sample space is $\Omega = \{11, 12, 13, \ldots, 65, 66\}$ and $|\Omega| = 36$.

The number of possible events is then $2^{36} \approx 6.8 \times 10^{10}$.

Suppose we are interested only in the events E_k that the sum of the two faces that appear is equal to k for $k = 2, 3, \ldots, 12$.

Doing a naive count, we have

$$E_2 = \{11\},\ E_3 = \{12, 21\},\ E_4 = \{13, 22, 31\},$$
$$E_5 = \{14, 23, 32, 41\},\ E_6 = \{15, 24, 33, 42, 51\},$$
$$E_7 = \{16, 25, 34, 43, 52, 61\},\ E_8 = \{26, 35, 44, 53, 62\}$$
$$E_9 = \{36, 45, 54, 63\},\ E_{10} = \{46, 55, 64\},\ E_{11} = \{56, 65\},\ E_{12} = \{66\}$$

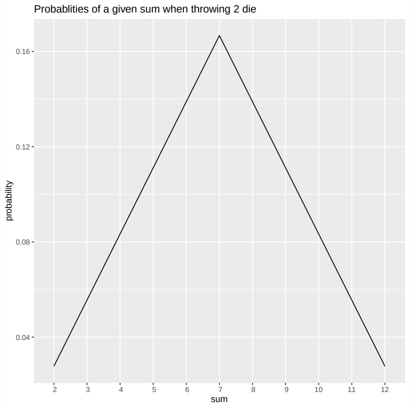
Assuming the sample points are equiprobable, we have

$$P(E_2) = P(E_{12}) = 1/36, P(E_3) = P(E_{11}) = 2/36,$$
$$P(E_4) = P(E_{10}) = 3/36, P(E_5) = P(E_9) = 4/36,$$
$$P(E_6) = P(E_8) = 5/36, P(E_7) = 6/36$$

Assuming the sample points are equiprobable, we have

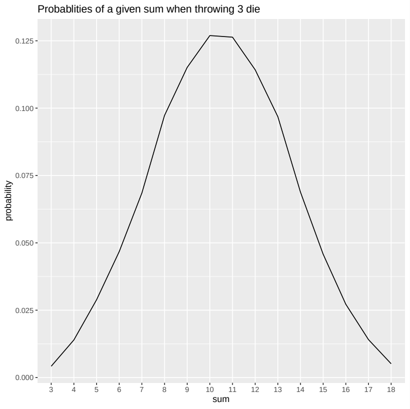
$$P(E_2) = P(E_{12}) = 1/36, P(E_3) = P(E_{11}) = 2/36,$$
$$P(E_4) = P(E_{10}) = 3/36, P(E_5) = P(E_9) = 4/36,$$
$$P(E_6) = P(E_8) = 5/36, P(E_7) = 6/36$$

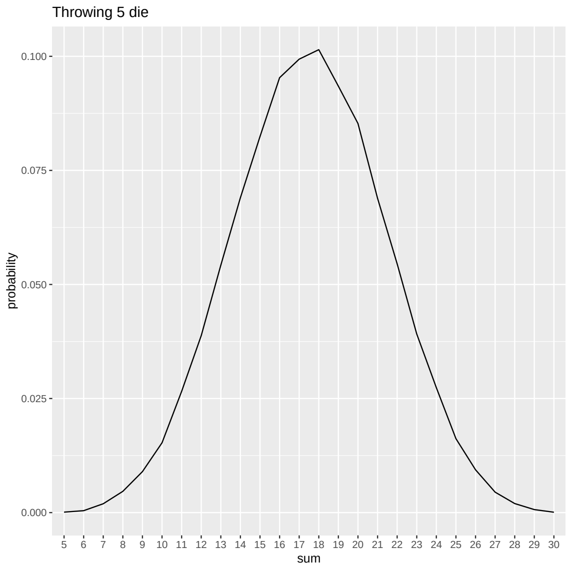
A plot of these probabilities is given below.

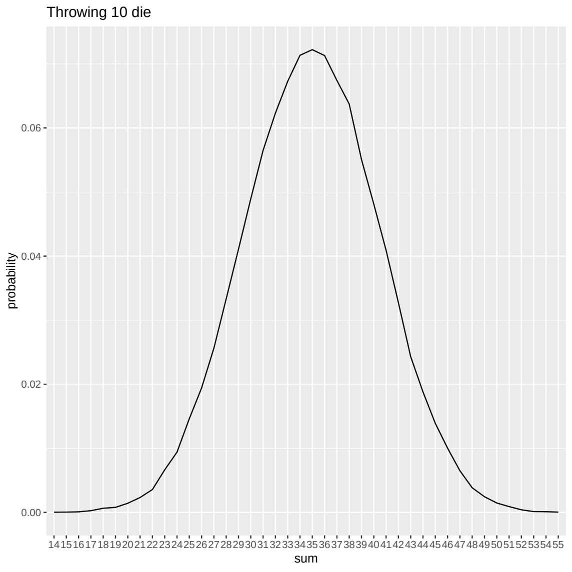


25

Suppose we are interested in throwing m die and counting the sum of the faces. Then the number of sample points is 6^m, the number of events of interest is $5m+1$. A naive count is now extremely tedious, even for $m=3$. A close form formula for the probabilities is available, but quite messy. Simulations yield




Throwing 5 die
probability
0.100
0.075
0.050
0.025
0.000
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
sum


Throwing 10 die
probability
0.06
0.04
0.02
0.00
14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
sum

Counting techniques

Many elementary probability questions assume a discrete sample space Ω where the simple events are equiprobable

The probability of any event $A \subset \Omega$ is then simply

$$P(A) = \frac{|A|}{|\Omega|}.$$

We now describe three simple & important rules for computing $|A|$ without needing to enumerate the elements in A.

Proposition (Product rule)

Suppose A consists of ordered collections of k-elements s.t.

- *there are n_1 choices for the first element.*
- *there are n_2 choices for the second element.*
- *$\ldots$ there are n_k choices for the k-th element.*

Then there are $n_1 \times n_2 \times \cdots \times n_k$ elements in A.

30

Example Phone numbers in Raleigh begin with area code 919. What is the probability that the last four digits of a randomly chosen phone number in Raleigh are all even ?

31

Example Phone numbers in Raleigh begin with area code 919. What is the probability that the last four digits of a randomly chosen phone number in Raleigh are all even ?

A. There are 10^7 possible different phone numbers starting with area code 919.

Among these 10^7 numbers, there are $10^3 \times 5^4$ phone numbers whose last four digits are all even.

The probability is therefore $2^{-4} = 1/16$.

31

Permutations and Combinations

We now note two important special cases of the product rule.

Proposition (Permutations)

Let S *be a set with* $|S| = n$*. Let* $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)$ *be an ordered tuple such that* $a_i \in S$ *for all* i *and* $a_i \neq a_j$ *whenever* $i \neq j$*. Then there are* $P_{k,n} = n(n-1)(n-2)\cdots(n-k+1)$ *different possible* $\boldsymbol{a}$*.*

Proposition (Combinations)

Let S be a set with $|S| = n$. Let $\boldsymbol{a} = \{a_1, a_2, \ldots, a_k\} \subset S$ be a subset of k elements from S. Then there are $C_{k,n} = P_{k,n}/k!$ different possible $\boldsymbol{a}$.

In summary we have, for an integer $0 \leq k \leq n$, that

$$P_{k,n} = \frac{n!}{(n-k)!}, \quad C_{k,n} = \frac{n!}{(n-k)! \times k!} = \binom{n}{k} = \binom{n}{n-k}.$$

Convention Note that $\binom{n}{0} = \binom{n}{n} = 1$. We also define $\binom{n}{k} = 0$ whenever $k < 0$ and $k > n$.

34

Example Sam, Ham and Spam are at party with n other people. Six people are chosen to dance on stage. What is the probability that all three of Sam, Ham, and Spam are chosen ?

Example Sam, Ham and Spam are at party with n other people. Six people are chosen to dance on stage. What is the probability that all three of Sam, Ham, and Spam are chosen ?

A. There are $n+3$ people in total. The number of ways of chosing 6 people from among these $n+3$ are $\binom{n+3}{6}$. If Sam, Ham, and Spam are all chosen then there are $\binom{n}{3}$ choices left for the remaining 3 people.

The probability is thus

$$\frac{\binom{n}{3}}{\binom{n+3}{6}} = \frac{120}{(n+1)(n+2)(n+3)}.$$

35

Example Sam, Ham, and Spam are at a party with n other people. Everyone lined up (in a single row) to take a photo. What is the probability that Sam, Ham, and Spam stand next to each other in the photo ?

36

Example Sam, Ham, and Spam are at a party with n other people. Everyone lined up (in a single row) to take a photo. What is the probability that Sam, Ham, and Spam stand next to each other in the photo ?

A. There are $(n+3)!$ different arrangements of n people.

Among these arrangements, there are $3! \times (n+1) \times n!$ arrangements in which Sam, Ham and Spam stands together.

The probability is thus

$$\frac{3! \times (n+1) \times n!}{(n+3)!} = \frac{6}{(n+2)(n+3)}.$$

Example Sam, Ham, and Spam are at a party with n other people. Ten people are chosen to lined up (in a single row) to take a photo. What is the probability that Sam, Ham, and Spam are chosen and stand next to each other in the photo ?

37

Example Sam, Ham, and Spam are at a party with n other people. Ten people are chosen to lined up (in a single row) to take a photo. What is the probability that Sam, Ham, and Spam are chosen and stand next to each other in the photo ?

A. We first determine where Sam, Ham and Spam stands. As they stand next to each other, there are 8 different locations. There are also $3! = 6$ different ordering of Sam, Ham and Spam.

There are $n!/(n-7)!$ different (ordered) selection of the remaining 7 people.

Finally, there are $P_{10,n+3} = (n+3)!/(n-7)!$ different ordered selection of 10 people from $n+3$ total people at the party.

In summary the probability is

$$\frac{8 \times 6 \times n!/(n-7)!}{(n+3)!/(n-7)!} = \frac{48}{(n+1)(n+2)(n+3)}.$$

More on combinations

We list some simple but important identities for $\binom{n}{k}$.

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \qquad \text{(Pascal's triangle)},$$

$$k\binom{n}{k} = n\binom{n-1}{k-1},$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n,$$

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a+b)^n, \qquad \text{(binomial theorem)}$$

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = 2^{n-1},$$

$$\sum_{k=0}^{m} \binom{r}{k}\binom{s}{m-k} = \binom{r+s}{m}, \qquad \text{(Vandermonde's identity)}$$

38

Birthday problem (Example E in Chapter 1)

Q. You are at a party with $n-1$ other classmates. How likely is it that two of you share a common birthday ?

39

Birthday problem (Example E in Chapter 1)

Q. You are at a party with $n-1$ other classmates. How likely is it that two of you share a common birthday ?

A. Suppose there are n people in a party. Suppose that the birthday of a person is equiprobably distributed in $\{1, 2, \ldots, 365\}$ (the number of days in a non-leap year).

Sample space Ω is the collection of sequences $(b_1, b_2, \ldots, b_n)$ with $b_i \in \{1, 2, \ldots, 365\}$.

Thus $|\Omega| = 365^n$ and the sample points in Ω are equiprobable

We are interested $P(A_n)$ where

$$A_n = \{s = (b_1, b_2, \ldots, b_n) \in \Omega \colon b_i = b_j \text{ for some } i \neq j\}$$

Now $\bar{A}_n$ is the event that every person has a distinct birthday.

For a sample point $s = (b_1, b_2, \ldots, b_n) \in \bar{A}_n$, we have

$$b_1 \in \{1, 2, \ldots, 365\}, b_2 \in \{1, 2, \ldots, 365\} \setminus \{b_1\}, \ldots,$$
$$b_{k+1} \in \{1, 2, \ldots, 365\} \setminus \{b_1, b_2, \ldots, b_k\}.$$

Now $\bar{A}_n$ is the event that every person has a distinct birthday.

For a sample point $s = (b_1, b_2, \ldots, b_n) \in \bar{A}_n$, we have

$$b_1 \in \{1, 2, \ldots, 365\}, b_2 \in \{1, 2, \ldots, 365\} \setminus \{b_1\}, \ldots,$$
$$b_{k+1} \in \{1, 2, \ldots, 365\} \setminus \{b_1, b_2, \ldots, b_k\}.$$

We thus have $|\bar{A}_n| = 365 \times 364 \times \cdots \times (365 - n + 1)$ and hence

$$\begin{aligned} P(\bar{A}_n) &= \frac{365 \times 364 \times 363 \times \cdots \times (365 - n + 1)}{365^n} \\ &= \prod_{i=1}^{n-1} \left(1 - \frac{i}{365}\right) \end{aligned}$$

In summary

$$P(A_n) = 1 - P(\bar{A}_n) = 1 - \prod_{i=1}^{n-1}\Big(1 - \frac{i}{365}\Big).$$

We can now evaluate

$$P(A_{20}) = 0.41; \quad P(A_{30}) = 0.7; \quad P(A_{40}) = 0.89; \quad P(A_{50}) = 0.97$$

Multinomial coefficients

PROPOSITION **C**

The number of ways that n objects can be grouped into r classes with n_i in the ith class, $i = 1, \ldots, r$, and $\sum_{i=1}^{r} n_i = n$ is

$$\binom{n}{n_1 n_2 \cdots n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

Proof

This can be seen by using Proposition B and the multiplication principle. (Note that Proposition B is the special case for which $r = 2$.) There are $\binom{n}{n_1}$ ways to choose the objects for the first class. Having done that, there are $\binom{n-n_1}{n_2}$ ways of choosing the objects for the second class. Continuing in this manner, there are

$$\frac{n!}{n_1!(n-n_1)!} \frac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \cdots \frac{(n-n_1-n_2-\cdots-n_{r-1})!}{0!n_r!}$$

choices in all. After cancellation, this yields the desired result. ■