Our new Zoom room for all remaining lectures

• [https://illinois.zoom.us/j/95028894623?pwd=K3lPc0srV2Z6TllhUnVMaDB0UXZFUT09](https://illinois.zoom.us/j/95028894623?pwd=K3lPc0srV2Z6TllhUnVMaDB0UXZFUT09)

• Meeting ID: 950 2889 4623
  Password: 822899
Inductive probability relies on combinatorics or the art of counting combinations.
Counting – Multiplication Rule

• Multiplication rule:
  – Let an operation consist of k steps and
    • \( n_1 \) ways of completing the step 1,
    • \( n_2 \) ways of completing the step 2, ... and
    .......
    • \( n_k \) ways of completing the step k.
  – Then, the total number of ways of carrying the entire operation is:
    • \( n_1 \times n_2 \times \ldots \times n_k \)

Example: DNA 2-mer

\[ n_1 = n_2 = 4 \]
• \( S = \{A, \ C, \ G, \ T\} \) the set of 4 DNA bases
  – Number of k-mers is \( 4^k=4\times4\times4\times\ldots\times4 \) (k –times)
  – There are \( 4^3=64 \) triplets in the genetic code
  – There are only 20 amino acids (AA)+1 stop codon
  – There is redundancy: same AA coded by 1-3 codons
  – Evidence of natural selection: “silent” changes of bases are more common than AA changing ones

• A protein-coding part of the gene is typically 1000 bases long
  – There are \( 4^{1000}=2^{2000}\sim10^{600} \) possible sequences of just one gene
  – Or \( (10^{600})^{25,000}=10^{15,000,000} \) of 25,000 human genes.
  – For comparison, the Universe has between \( 10^{78} \) and \( 10^{80} \) atoms and is \( 4\times10^{17} \) seconds old.
Counting – Permutation Rule

• A permutation is a unique sequence of distinct items.
• If $S = \{a, b, c\}$, then there are 6 permutations
  – Namely: $abc$, $acb$, $bac$, $bca$, $cab$, $cba$ (order matters)
• # of permutations for a set of $n$ items is $n!$
• $n!$ (factorial function) = $n*(n-1)*(n-2)*...*2*1$
• $7! = 7*6*5*4*3*2*1 = 5,040$
• By definition: $0! = 1$
Multiplication and permutation rules are two examples of a general problem, where a set of size \( k \) is drawn from a population of \( n \) distinct objects.
How many ways to choose a sample of \( k \) objects out of a population of \( n \) objects

<table>
<thead>
<tr>
<th>Order matters</th>
<th>Order does not matter</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Replace</strong></td>
<td>( kn \times n \times n \times \ldots \times n = n^k )</td>
</tr>
<tr>
<td><strong>Do not replace</strong></td>
<td>( \frac{n!}{(n-k-1)!} = \frac{n!}{(n-k)!} )</td>
</tr>
</tbody>
</table>
How to solve the problem of \( k \) out of \( n \) with replacement but where order does not matter?

Let's solve the \( n=2 \) problem first:

\[
\begin{align*}
4 \text{ possibilities} & \\
\begin{array}{c}
\text{Object 1} \quad \text{Object 2} \\
1 & 1 \\
2 & 2 \\
3 & 3 \\
4 & 4
\end{array}
\end{align*}
\]

\( n=4, \ k=7 \)

\[
\begin{align*}
\text{\( k \) dots, \( n-1 \) for boundaries} & \\
\binom{k+n-1}{k} = \frac{(k+n-1)!}{k! \ (n-1)!} \quad \text{ways to distribute}
\end{align*}
\]
# Sampling table

How many ways to choose a *sample of k objects* out of *population of n objects*?

<table>
<thead>
<tr>
<th></th>
<th>Order matters</th>
<th>Order does not matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement</td>
<td>((n)^k)</td>
<td>Difficult: (\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)!k!})</td>
</tr>
<tr>
<td>No replacement</td>
<td>(n(n-1)(n-2)..(n-k+1) = \frac{n!}{(n-k)!} = \frac{n!}{(n-k)!k!})</td>
<td>(\binom{n}{k} = \frac{n!}{(n-k)!k!})</td>
</tr>
</tbody>
</table>
A class has $n$ students. What is the smallest $n$ so that there is 100% probability that there is a pair of people with the same birthday, e.g. May 1 (in any year)

A. 366
B. 367
C. 730
D. 32

Get your i-clickers
A class has \( n \) students.
What is the smallest \( n \) so that there is 50% probability that there is a pair people with the same birthday e.g. May 1 (in any year)

A. 734
B. 184
C. 5
D. 23

Get your i-clickers
Probability that $n$ people have different birthdays is:

$$\frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \ldots \times \frac{365-n+1}{365}$$

Let's find $n$ when this is approximately 1:

$$\frac{1}{2} \approx \exp \left( \sum_{k=1}^{n} \log \left( 1 - \frac{k-1}{365} \right) \right)$$
\[
\log \left( 1 - \frac{k-1}{365} \right) \approx - \frac{k-1}{365}
\]

\[
\sum_{k=1}^{n} \log \left( 1 - \frac{k-1}{365} \right) \approx - \frac{(n-1)n}{2 \cdot 365}
\]

We need

\[
\frac{1}{2} = \exp \left( - \frac{n(n-1)}{2 \cdot 365} \right)
\]

or

\[
-\log 2 \approx -\frac{n}{2 \cdot 365}
\]

\[
n \approx \sqrt{2 \cdot 365 \cdot \log 2} \approx 22.5
\]
SOLUTION  Because each person can celebrate his or her birthday on any one of 365 days, there are a total of \((365)^n\) possible outcomes. (We are ignoring the possibility of someone having been born on February 29.) Furthermore, there are \((365)(364)(363) \cdot (365 - n + 1)\) possible outcomes that result in no two of the people having the same birthday. This is so because the first person could have any one of 365 birthdays, the next person any of the remaining 364 days, the next any of the remaining 363, and so on. Hence, assuming that each outcome is equally likely, we see that the desired probability is

\[
\frac{(365)(364)(363) \cdots (365 - n + 1)}{(365)^n}
\]

It is a rather surprising fact that when \(n \geq 23\), this probability is less than \(\frac{1}{2}\). That is, if there are 23 or more people in a room, then the probability that at least two of them have the same birthday exceeds \(\frac{1}{2}\). Many people are initially surprised by this result, since 23 seems so small in relation to 365, the number of days of the year. However, every pair of individuals has probability \(\frac{365}{(365)^2} = \frac{1}{365}\) of having the same birthday, and in a group of 23 people there are \(\binom{23}{2} = 253\) different pairs of individuals. Looked at this way, the result no longer seems so surprising. □
Let’s check it on our class

Please type your month and date of birth in chat
For our Zoom room with 66 people, the probability of all having different birthdays is

\[ \frac{365}{365} \cdot \frac{364}{365} \cdot \ldots \cdot \frac{365 - 66 + 1}{365} \approx \exp \left( - \frac{66 \cdot 65}{2 \cdot 365} \right) \approx 0.0027 \]

Probability to have at least one pair with the same birthday is 99.7%
Out of 55 people who submitted their birthday there were two matches

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17-Dec</td>
<td></td>
<td></td>
<td></td>
<td>Abhishek Bhattacharjee</td>
</tr>
<tr>
<td>17-Dec</td>
<td></td>
<td></td>
<td></td>
<td>Bruno Suarez</td>
</tr>
<tr>
<td>29-Sep</td>
<td></td>
<td></td>
<td></td>
<td>Yoshi</td>
</tr>
<tr>
<td>29-Sep</td>
<td></td>
<td></td>
<td></td>
<td>Daria Wendell</td>
</tr>
</tbody>
</table>
Probability Axioms, Conditional Probability, Statistical (In)dependence, Circuit Problems
Axioms of probability

Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties:

If $S$ is the sample space and $E$ is any event in a random experiment,

(1) $P(S) = 1$

(2) $0 \leq P(E) \leq 1$

(3) For two events $E_1$ and $E_2$ with $E_1 \cap E_2 = \emptyset$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

These axioms imply that:

$$P(\emptyset) = 0$$

$$P(E') = 1 - P(E)$$

if the event $E_1$ is contained in the event $E_2$

$$P(E_1) \leq P(E_2)$$
Addition rules following from the Axiom (3)

If $A$ and $B$ are mutually exclusive events, i.e. $A \cap B = \emptyset$

$P(A \cup B) = P(A) + P(B)$ (2-2)

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (2-1)
\[ P(A \cup B \cup C) = P(A) + P(B) + P(C) - \\
- P(A \cap B) - P(A \cap C) - P(B \cap C) + \\
+ P(A \cap B \cap C). \]
The **conditional probability** of an event $B$ given an event $A$, denoted as $P(B|A)$, is

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

for $P(A) > 0$.

This definition can be understood in a special case in which all outcomes of a random experiment are equally likely. If there are $n$ total outcomes,

$$P(A) = \frac{\text{number of outcomes in } A}{n}$$

Also,

$$P(A \cap B) = \frac{\text{number of outcomes in } A \cap B}{n}$$

Consequently,

$$\frac{P(A \cap B)}{P(A)} = \frac{\text{number of outcomes in } A \cap B}{\text{number of outcomes in } A}$$

Therefore, $P(B|A)$ can be interpreted as the relative frequency of event $B$ among the trials that produce an outcome in event $A$. 