Ex. Find the probability that a random string from \(a, b, \ldots, z\) of length 10 has no repeated letters.

\[ S = \text{Letters of len. 10 from a, \ldots, z} \]

\[ |S| = 26^{10} \]

\[ E = \text{Strings with no repeats} \]

\[ |E| = 26 \cdot 25 \cdot \ldots \cdot 17 = \frac{26!}{(26-10)!} = \frac{26!}{16!} \]

\[ P(E) = \frac{26!}{16!} \cdot \frac{1}{26^{10}} = 0.1365 \]
In general, what is the probability that a string of length \( n \) from an alphabet of size \( n \) has no repeated letters?

Note: if \( k > n \) then this probability is 0, by the pigeonhole property. Consider the case when \( k < n \).

\[ |\mathcal{S}| = n^k, \quad |\mathcal{E}| = \frac{n^k}{(n-k)!} = \frac{n(n-1) \cdots (n-k+1)}{(n-k)!} \]

\[ x^k = x(x-1)(x-2) \cdots (x-k+1) \]

So, \( |\mathcal{E}| = n^k \)

\[ \mathbb{P}(E) = \frac{n^k}{n^k} = \frac{x(x-1)(x-2) \cdots (x-k+1)}{n \cdot n \cdot \cdots \cdot n} \]

\[ = \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \cdots \cdot \frac{n-k+1}{n} \]
Ex 5 dice are thrown. What is the probability that no number is repeated?

Each throw gives a random string of length 5 from the alphabet \( \{1, 2, 3, 4, 5, 6\} \)

\[ |S| = 6^5 = 7776 \]

\[ |E| = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 = 720 \]

\[ P(E) = \frac{720}{7776} = 0.0925 \]
Ex: Same 5 dice. What is probability of getting 1 pair and 3 distinct numbers

- Pick # in pair: \( \binom{6}{1} = 6 \) ways
- Pick 3 distinct #s: \( \binom{5}{3} = \frac{5!}{2!} = 10 \)
- Pick an arrangement of the 3 distinct #s: \( 3! = 6 \)

Let \( x, y, z \) be the 3 distinct #s. The spaces between \( x, y, z \) are 4 bins that can be filled with either

- 2 members of pair in same bin: \( \binom{4}{1} \)
- or 2 members of pair in different bins: \( \binom{4}{2} \)
x y z

# bins to hold pain

# ways of distributing pain

into bins: \( \binom{4}{1} + \binom{4}{2} = 4 + \frac{4 \cdot 3}{2} = 10 \)

By Product rule:

\[ |E| = 6 \cdot 10 \cdot 6 \cdot 10 = 3600 \]

\[ P(E) = \frac{3600}{7776} = 0.4629 \]
Exercise:

determine probability of

- 1 triple, 2 distinct 1200 \( \frac{\text{Prob}}{.1543} \)
- 2 pairs, 1 distinct 1800 \( \text{Prob} \) = .2314
- 1 pair, 1 triple 300 \( \text{Prob} \) = .0385
- 1 quad, 1 distinct 150 \( \text{Prob} \) = .01929
- 1 quint 6 \( \text{Prob} \) = .00077

\( \checkmark \) all distinct 720 \( \text{Prob} \) = .0925
\( \checkmark \) 1 pair, 3 distinct 3600 \( \text{Prob} \) = .4629

7776 1
7.2 Probability Theory

It may not be true that all outcomes are equally likely.

Define

A probability distribution is a function

\[ P : S \rightarrow \mathbb{R} \]

satisfying

1. \( \forall x \in S : 0 \leq P(x) \leq 1 \)

2. \( \sum_{x \in S} P(x) = 1 \)
we extend \( \mathcal{D} \) to a function

\[ \mathcal{D} : \mathcal{P}(S) \to \mathbb{R} \]

by

\[ \mathcal{D}(E) = \sum_{x \in E} \mathcal{D}(x) \]

Then we have 'identities':

- \( \mathcal{D}(\emptyset) = 1 - \mathcal{D}(E) \)
- \( \mathcal{D}(E \cup F \cup \ldots \cup F) = \mathcal{D}(E) + \mathcal{D}(F) - \mathcal{D}(E \cap F) \)
- if \( E_1, E_2, \ldots, E_k \) are pairwise disjoint events, then

\[ \mathcal{D}(\bigcup_{i=1}^{k} E_i) = \sum_{i=1}^{k} \mathcal{D}(E_i) \]
Exercise: Prove these formulae

Proof of 1st bullet

Note: $E \cup \overline{E} = S$ and $E \cap \overline{E} = \emptyset$.

$$1 = \sum_{x \in S} P(x) = \sum_{x \in E \cup \overline{E}} P(x) = \sum_{x \in E} P(x) + \sum_{x \in \overline{E}} P(x)$$

$$\therefore 1 = P(E) + P(\overline{E})$$

$$\therefore P(\overline{E}) = 1 - P(E).$$

Ex. Let $S \neq \emptyset$ be a finite sample space and suppose

$$P(x) = \frac{1}{|S|} \quad \text{for all } x \in S.$$
observe:

1) $0 \leq \mathbf{P}(x) \leq 1$ for all $x \in \mathcal{S}$

2) $\sum_{x \in \mathcal{S}} \mathbf{P}(x) = \sum_{x \in \mathcal{S}} \frac{1}{|\mathcal{S}|} = |\mathcal{S}| \cdot \frac{1}{|\mathcal{S}|} = 1$

so $\mathbf{P}(\cdot)$ is a probability dist. in which all outcomes are equally likely.

Let $E \subseteq \mathcal{S}$. Then

$\mathbf{P}(E) = \sum_{x \in E} \mathbf{P}(x) = \sum_{x \in E} \frac{1}{|\mathcal{S}|} = |E| \cdot \frac{1}{|\mathcal{S}|}$

i.e. $\mathbf{P}(E) = \frac{|E|}{|\mathcal{S}|}$.

We recover the defn of probability from (7.1). We call this the uniform probability distribution.
Data

Let $E, F \subseteq S$ with $P(F) \neq 0$.

The conditional probability of $E$ given $F$ is

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}$$

Interpretation: Given the knowledge that event $F$ occurs, the prob. that event $E$ occurs is $P(E \mid F)$.
Ex. A bit string of length 3 is selected at random (i.e. uniform dist.)
Given that the 1st bit is 0, what is the prob. that there are exactly 2 1's?

\[ S = \{000, 001, 010, 011, 100, 101, 110, 111\} \]
\[ E = \{011, 101, 110\} \text{ (exactly 2 1's)} \]
\[ F = \{000, 001, 010, 011\} \text{ (1st bit 0)} \]

\[ P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4} = 0.25 \]

\[ E \cap F = \{011\} \]
Events $E, F \subseteq \mathcal{S}$ are called **independent** if

$$P(EnF) = P(E) \cdot P(F)$$

**Note:**

- If $P(F) \neq 0$, this is equivalent to

$$P(E) = \frac{P(EnF)}{P(F)} = P(E|F)$$

- If $P(E) \neq 0$, this is equivalent to

$$P(F) = \frac{P(FnE)}{P(E)} = P(F|E)$$
Ex. Previous bit string problem.

are E and F independent?

\[ P(E \cap F) = \frac{1}{8} \]

\[ P(E) \cdot P(F) = \frac{3}{8} \cdot \frac{4}{8} = \frac{3}{16} \neq \frac{1}{8} \]

not independent.

also

\[ P(E|F) = \frac{1}{4} \neq \frac{3}{8} = P(E) \]

\[ P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3} \neq \frac{4}{8} = P(F) \]
Ex.

A permutation of \{1, 2, 3\} is selected at random. In the event that 1 precedes 2, independent of the event that 3 is not first.

\[ S = \{123, 132, 213, 231, 312, 321\} \]

\[ E = \{123, 132, 312\} \quad (1 \text{ before } 2) \]

\[ F = \{123, 132, 213, 231\} \quad (3 \text{ not first}) \]

\[ E \cap F = \{123, 132\} \]

\[ P(E) = \frac{3}{6} = \frac{1}{2}, \quad P(F) = \frac{4}{6} = \frac{2}{3} \]

\[ P(E \cap F) = \frac{2}{6} = \frac{1}{3} = \frac{1}{2} \cdot \frac{2}{3} = P(E) \cdot P(F) \]

So \( E \) and \( F \) are independent.
\[ \text{Ex. } S = \{ 1, 2, 3, 4, 5 \} \]

<table>
<thead>
<tr>
<th>$x$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>0.1</td>
</tr>
</tbody>
</table>

\[ E = \{ 1, 2, 3 \}, \quad F = \{ 3, 4, 5 \} \]

Find: $P(E)$, $P(F)$, $P(E \cap F)$, $P(E|F)$, $P(F|E)$

$P(E) = 0.1 + 0.2 + 0.2 = 0.5$

$P(F) = 0.2 + 0.4 + 0.1 = 0.7$

$P(E \cap F) = 0.2$

$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{0.2}{0.7} = \frac{2}{7} \approx 0.5$%

$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{0.2}{0.5} = \frac{2}{5} \approx 0.7$

Observe $E$ and $F$ are not independent.
Example 5 \[ P(x) \].

Let \( C = \{1, 5\} \).

Are \( E \) and \( C \) independent?

\[ E \cap C = \{1\} \]

\[ P(C) = 0.1 + 0.1 = 0.2 \]

\[ P(E \cap C) = 0.1 = 0.5 \times 0.2 = P(E) \cdot P(C) \]

So yes, \( E \perp C \) are independent.