COUNTING AND PROBABILITY

Reading: CLR Sections 6.1 -- 6.3; 6.6.

Counting

Given a set S with n elements, how many ways are there to choose k elements from S? The answer depends on whether (1) order matters and (2) duplicates are allowed.

<table>
<thead>
<tr>
<th></th>
<th>Duplicates not allowed</th>
<th>Duplicates allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(k ≤ n)</td>
<td>(k may be &gt; n)</td>
</tr>
<tr>
<td>Order matters</td>
<td>k-permutation</td>
<td>k-tuple (k-string)</td>
</tr>
<tr>
<td>(sequences)</td>
<td>\frac{n!}{(n - k)!}</td>
<td>(n^k)</td>
</tr>
<tr>
<td>Order doesn't matter</td>
<td>k-combination</td>
<td>k-selection</td>
</tr>
<tr>
<td>(sets)</td>
<td>\frac{n!}{k! (n - k)!}</td>
<td>\frac{(n - 1 + k)!}{k! (n - 1)!}</td>
</tr>
</tbody>
</table>

Example: S = {a, b, c, d}, k = 2

<table>
<thead>
<tr>
<th>Type</th>
<th>Elements</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-permutations of S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-combination of S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-strings of S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-selections of S</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Exercise: Try k = 3
Sample Spaces

A sample space is the set of all possible outcomes of an experiment.

Experiment A: Flip three coins.
Sample space =

Experiment B: Flip a coin until you get heads.
Sample space =

A sequential sample spaces describing a sequence of experiments is naturally viewed as a tree:

Experiment A:

Experiment B:
Events

An event is a subset of a sample space.

Experiment A:

- Event A1: First flip is a head =
- Event A2: Second flip is a tail =
- Event A3: Exactly two tails =
- Event A4: Two consecutive flips the same =

Experiment B:

- Event B1: First flip is a head =
- Event B2: First flip is a tail =
- Event B3: Even number of flips =

If A and B are events in sample space S,
- "A and B" is translated "A ∩ B"
- "A or B" is translated "A ∪ B"
- "not A" is translated "S − A"

Two events A and B are mutually exclusive if A ∩ B = Ø.

Of the four events in Experiment A, which pairs are mutually exclusive?

Of the three events in Experiment B, which pairs are mutually exclusive?
Probability Distributions

A **probability distribution** $\Pr\{}$ on a sample space $S$ is any mapping from events to $[0..1]$ such that the following axioms hold:

1. $\Pr\{A\} \geq 0$ for any event $A$.
2. $\Pr\{S\} = 1$
3. If $A \cap B = \emptyset$ then $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$

Some useful theorems:

- $\Pr\{\emptyset\} = 0$.
- $\Pr\{S - A\} = 1 - \Pr\{A\}$.
- If $A \subseteq B$ then $\Pr\{A\} \leq \Pr\{B\}$
- $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\}$

A sequential sample space gives rise to a **probability tree**. If events at distinct stages of tree are independent (defined below), probability of leaf is product of probabilities on path to leaf.

Experiment A with $H_1 = (1/3)$, $H_2 = (1/4)$, $H_3 = (1/5)$

Experiment B with $H_1 = (1/3)$
Conditional Probability

Probability of A given B = \( \Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} \)

Example 1: In Experiment A with \( H_1 = (1/3) \), \( H_2 = (1/4) \), \( H_3 = (1/5) \), what is probability of \( A_3 \) (exactly two tails) given \( A_4 \) (two consecutive flips the same)?

Example 2: Ann has two children, one of which is a girl. What is the probability that the other child is a boy? (Assume a child is girl or boy with equal probability.)

Independence

Events A and B are independent if \( \Pr\{A \cap B\} = \Pr\{A\}\Pr\{B\} \).
(Equivalently, \( \Pr\{A \mid B\} = \Pr\{A\} \))

Of the four events in Experiment A, which pairs are independent?
Discrete Random Variables

A discrete random variable on a sample space $S$ is a function $f: S \rightarrow \text{reals}$.

E.g., in Experiment A let

$h(x) = \text{number of heads in outcome } x$.

$r(x) = \text{length of longest run in outcome } x$.

\[
\begin{array}{c|c|c}
\text{Outcome} & h & r \\
\hline
s & \uparrow & \uparrow \\
\hline
H_1 H_2 H_3 & 3 & 1 \\
\hline
H_1 H_2 T_3 & 3 & 1 \\
\hline
H_1 T_2 H_3 & 2 & 2 \\
\hline
H_1 T_2 T_3 & 2 & 2 \\
\hline
T_1 H_2 H_3 & 2 & 2 \\
\hline
T_1 H_2 T_3 & 2 & 2 \\
\hline
T_1 T_2 H_3 & 2 & 2 \\
\hline
T_1 T_2 T_3 & 2 & 2 \\
\end{array}
\]
Probability Density Function

The pre-image of a value $v$ under $f = f^{-1}(v) = \{ x \in S \mid f(x) = v \}$
(This differs from CLR's notation "$f = v$", which I find confusing.)

Each pre-image is an event:

$h^{-1}(1) =$

$r^{-1}(3) =$

The probability density function (PDF) of discrete random variable $f$ is another function that maps each target value $v$ of $f$ to its probability. This is determined by summing the probabilities of all sources $x$ that $f$ maps to $v$:

$$[PDF(f)](v) = Pr\{f^{-1}(v) \} = \sum_{\{ x \in S \mid f(x) = v \}} Pr\{x\}$$

What is the function $PDF(h)$?

What is the function $PDF(r)$?

Discrete random variables $f$ and $g$ are independent if for all $w$ and $v$, the events $f^{-1}(w)$ and $g^{-1}(v)$ are independent.

Are $h$ and $r$ independent?
**Expected Value**

The **expected value** $E[f]$ of a discrete random variable $f$ is a weighted average in which target values $v$ of $f$ are weighted by their probability:

$$E[f] = \sum_v v \cdot \Pr\{f^{-1}(v)\}$$

**Examples:**

- Expected number of heads in Experiment A = $E[h] = \ldots$

- Expected length of runs in Experiment A = $E[r] = \ldots$

- Expected number of flips in Experiment B = $\ldots$

Expectations are linear. If $f$ & $g$ are discrete random variables and $c$ is a constant:

- $E[f + g] = E[f] + E[g]$  \{The sum of functions $f$ and $g$ is a function $(f + g)(x) = f(x) + g(x)$.\}

- $E[cf] = c[E[f]]$  \{The scaling by $c$ of a function $f$ is a function $(cf)(x) = cf(x)$.\}

If $f$ and $g$ are independent, then

- $E[fg] = E[f]E[g]$  \{The product of functions $f$ and $g$ is a function $(fg)(x) = f(x)g(x)$.\}