# Partial Observations

by Sven Nilsen, 2020

*In this paper I show that probability distributions can be learned through partial observations.*

For any random binary number of `n` digits, each digit has exactly 50% probability of being `1`:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 00 | 01 | 10 | 11 | | | | |
| 000 | 001 | 010 | 011 | 100 | 101 | 110 | 111 |

However, this only holds if the probability distribution is $1 / 2^n$ per number.
Assume that a non-deterministic function outputs a sample from an arbitrary probability distribution:

$$f : () \to bool^n$$

This probability distribution does not depend on previous output numbers.
In this case, one can simply count up the frequency of each number to learn `f`.

A partial observation of `f` is when a deterministic function `g` removes some of that information:

$$g \cdot f$$

$$g : bool^n \to bool^m \qquad m < n$$

For example, `n = 2` and `m = 1`, where `f` has the following probability distribution:

| 00 | 01 | 10 | 11 |
|---|---|---|---|
| 0.2 | 0.3 | 0.1 | 0.4 |

The bit “sums over” the probabilities in `f`:

| 0 | 1 | | |
|---|---|---|---|
| 0.5 | 0.5 | first bit | |
| 0.3 | 0.7 | second bit | |
| 0.6 | **0.4** | logical AND | learns `f() == 11` |
| **0.2** | 0.8 | logical OR | learns `f() == 00` |

Some functions, such as AND and OR, has the ability to learn parts of the probability distribution.
One can simply count up the frequency of this number and ignore everything else.

|[and] true| == 1 |[or] false| == 1

In general, the probability distribution can be learned by choosing a sub-type for each number:

|[g] b| == 1