Basic Statistics

Lecture notes of Alexander Wood
CSCI 360 Cryptography and Cryptanalysis
awood@jjay.cuny.edu

John Jay College of Criminal Justice

Statistical Profiles

Statistical profiles can mean many things. For instance, last class we created a statistical profile of single letter distributions in a text document.

Probability Distributions

When we build statistical profiles we will be creating **probability distributions**. A probability distribution is a set of possible **outcomes**, $\Omega = \{x_1, x_2, \ldots, x_n\}$, together with probabilities $P(X = x_1), P(X = x_2), \cdots, P(X = x_n)$ which represent the **likelihoods** of the respective outcomes occurring.

The symbol X above denotes the (unknown) actual outcome, and $P(X = x_i)$ denotes the probability that the **outcome** is x_i.

Example 1: D=Coin Flip

Lets build a probability distribution for a single coin flip. The possible outcomes are heads or tails, denoted $\Omega = \{H, T\}$. The probability of each is $1/2$.

$$P(X = H) = P(X = T) = 1/2$$

Example 2: Single Roll of 6-Sided Die

There are six outcomes: $\Omega = \{1, 2, 3, 4, 5, 6\}$, each with equal likelihood.

$$P(X = 1) = P(X = 2) = P(X = 3)$$

$$= P(X = 4) = P(X = 5) = P(X = 6) = 1/6$$

Example 3: Two Rolls of a Fair-Sided Die

Now lets look at an example where each outcome does not have the same likelihood. The sum of two rolls of a fair-sided die can take values $\Omega = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Example 3 Continued

Below is the probability P that the roll of two fair six sided dice will add up to Ω.

Ω	P
2	3%
3	6%
4	8%
5	11%
6	14%
7	17%
8	14%
9	11%
10	8%
11	6%
12	3%

References

- The random module: `https://docs.python.org/3/library/random.html`