CPSC 351 — Mathematics Review
Part Four: Discrete Probability Theory

Ideally, everything in the document should be a ***review*** of material that you learned about in a prerequisite course. It will be assumed that you understand and can use all of it in CPSC 351.

The primary reference that was used, when preparing text, is Stirzaker's text *Probability and Random Variables: A Beginner's Guide* [1].

Experiments, Outcomes, and Events

Definition 1. An ***experiment*** is a procedure (or process) that yields one of a given set of possible ***outcomes***. The set of possible outcomes of the experiment — which we will often name Ω — is called the ***sample space***.

It is possible (but not guaranteed) that experiments whose sample spaces are *finite sets* received most (or all) attention in a prerequisite course where discrete probability theory was introduced. Experiments with infinite sample spaces will sometimes be considered in this course too.

Definition 2. For any set Ω, let $\mathcal{P}(\Omega)$ denote the set of all ***subsets*** of Ω.

Definition 3. An ***event*** is a subset of the experiment's sample space Ω.

Thus every *event* is an element of the set $\mathcal{P}(\Omega)$. Indeed, $\mathcal{P}(\Omega)$ is set of all events (for this experiment) — so that if Ω is finite then $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$.

Definition 4. An ***elementary event*** is an event with size one — that is, it is an event that only includes a single outcome.

Probability Distributions

Definition 5. Consider an experiment with sample space Ω. A ***probability distribution*** is a (total) function

$$\mathsf{P} : \Omega \to \mathbb{R}$$

such that $0 \leq \mathsf{P}(x) \leq 1$ for every outcome $x \in \Omega$, and such that

$$\sum_{x \in \Omega} \mathsf{P}(x) = 1.$$

Definition 6. If Ω is a finite set then the ***uniform probability distribution*** (for Ω) defines the probability of every outcome to be the same: This is the function $\mathsf{P} : \Omega \to \mathbb{R}$ such that

$$\mathsf{P}(x) = \frac{1}{|\Omega|}$$

for every outcome $x \in \Omega$.

A probability distribution P (on an experiment with a countable sample space) is "extended" to get a function

$$\mathsf{P} : \mathcal{P}(\Omega) \to \mathbb{R}$$

by setting

$$\mathsf{P}(A) = \sum_{x \in A} \mathsf{P}(x)$$

for every event $A \subseteq \Omega$ (that is, for all $A \in \mathcal{P}(\Omega)$).

Suppose that $A \subseteq \Omega$ is an event. Then, if P is the ***uniform distribution*** for Ω, then

$$\mathsf{P}(A) = \sum_{x \in A} \mathsf{P}(x) = \frac{|A|}{|\Omega|}.$$

Combinations

If Ω is a sample space for an experiment and $A \subseteq \Omega$ is an event, then the ***complement***[1] of the event A, $\overline{A}$, is the set of outcomes that *are not* in A. That is, $\overline{A} = \{x \in \Omega \mid x \notin A\}$.

Theorem 7. *Let Ω be a sample space with probability distribution $P : \Omega \to \mathbb{R}$, and let $A \subseteq \Omega$. Then the probability of the complement, $\overline{A}$, of the event A is*

$$\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A).$$

[1]It is also OK if you use A^C to represent the complement of A, as we did for *languages*. These are both commonly used as the name for this set.

Theorem 8. *Let Ω be a sample space with probability distribution P : $\Omega \to \mathbb{R}$. Then, for any events $A, B \subseteq \Omega$,*

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B).$$

Since $\mathsf{P}(A \cap B) \geq 0$ for all events $A, B \subseteq \Omega$, Theorem 8 implies the following.

Corollary 9. *Let Ω be a sample space with probability distribution P : $\Omega \to \mathbb{R}$. Then, for any events $A, B \subseteq \Omega$,*

$$\mathsf{P}(A \cup B) \leq \mathsf{P}(A) + \mathsf{P}(B).$$

Random Variables

Definition 10. Let Ω be a sample space. A ***random variable over* $\boldsymbol{\Omega}$** is a (total) function $X : \Omega \to \mathbb{R}$.

Suppose X is a random variable and $x \in \mathbb{R}$. Then "$X = x$" is the event including all outcomes for which the value of the random variable X is the real number x, that is, "$X = x$" is the event

$$\{\sigma \in \Omega \mid X(\sigma) = x\}.$$

Independence

Once again, let Ω be a sample space with a probability distribution P : $\Omega \to \mathbb{R}$. Then, for any events $A, B \subseteq \Omega$,

Definition 11. Let $A, B \subseteq \Omega$. Then the events A and B are ***independent*** (with respect to the probability distribution P) if

$$\mathsf{P}(A \cap B) = \mathsf{P}(A) \times \mathsf{P}(B).$$

Definition 12. Let $X, Y : \Omega \to \mathbb{R}$. Then the random variables X and Y are ***independent*** if

$$\mathsf{P}(X = x \cap Y = y) = \mathsf{P}(X = x) \times \mathsf{P}(Y = y)$$

for all $x, y \in \mathbb{R}$.

Thus the ***events*** X and Y are independent if and only if the ***events*** "$X = x$" and "$Y = y$" are independent for all $x, y \in \mathbb{R}$.

References

[1] Davvid Stirzaker. *Probability and Random Variables: A Beginner's Guide*. Cambridge University Press, 1999.