OMCIT 592 Module 07 Self-Paced 01 (instructor Val Tannen)

Reference to this self-paced segment in seg.07.04.

This is a segment that contains material meant to be learned *at your own pace*. We are trying to assist you in this endeavor by organizing the material in a manner similar to the way it is outlined in the recorded segments, however with one additional suggestion.

When you see the following marker:

we suggest that you stop and make sure you thoroughly understood the material presented so far before you proceed further.

Proofs of Probability Properties P0-P7

As you may recall, in the segment "Probability properties" you learned about seven probability properties, accompanied by brief justifications on why they hold. In this segment we will thoroughly prove all these properties, as well as the generalizations of two of them.

For all the propositions bellow, consider a probability space $(\Omega, \Pr)$ and arbitrary events E, A, B in this space.

We will begin with:

$$\textbf{(P0)} \quad \Pr[E] \geq 0$$

Proof. Recall that an event E in the probability space $(\Omega, \Pr)$ is a subset $E \subseteq \Omega$. Moreover, we have that the probability of this event is:

$$\Pr[E] = \sum_{w \in E} \Pr[w]$$

Recall that the definition of the probability space states that it is a pair $(\Omega, \Pr)$ where Ω is a finite non-empty set of outcomes and $\Pr : \Omega \longrightarrow [0, 1]$ is a function that associates with each outcome $w \in \Omega$ its probability $\Pr[w]$ which is a real number between 0 and 1 (inclusive).

Since for any $w \in \Omega$ we have $\Pr[w] \geq 0$ and since any sum of numbers that are each ≥ 0 is itself ≥ 0 we conclude that $\Pr[E] \geq 0$.

The next property also follows directly from our definitions for probability space and event probabilities:

$$\textbf{(P1)} \quad \Pr[\Omega] = 1$$

Proof. Recall that our definition of probability space also stipulates

$$\sum_{w \in \Omega} \Pr[w] = 1$$

Now, by the definition of the probability of an event:

$$\Pr[\Omega] = \sum_{w \in \Omega} = 1$$

(CONTINUED)

Proofs of Probability Properties P0-P7 (continued)

We prove now the **addition rule**:

$$\textbf{(P2)} \quad A \cap B = \emptyset \quad \Rightarrow \quad \Pr[A \cup B] = \Pr[A] + \Pr[B]$$

Proof. To prove (P2) we note the following **general property of sums**. If I and J are two **disjoint** sets then

$$\sum_{k \in I \cup J} a_k = \sum_{i \in I} a_i + \sum_{j \in J} a_j$$

Using this

$$\Pr[A \cup B] = \sum_{w \in A \cup B} \Pr[w] = \sum_{w \in A} \Pr[w] + \sum_{w \in A} \Pr[w] = \Pr[A] + \Pr[B]$$

Recall that in the segment "Probability properties" we also stated that (P2) generalizes for n finite sets as follows [this is the **(generalized) addition rule**]:

(P2Gen) If $\{A_1, \ldots, A_n\}$ are pairwise disjoint, then $\Pr[A_1 \cup \ldots \cup A_n] = \Pr[A_1] + \ldots + \Pr[A_n]$

Proof. We prove (P2Gen) by ordinary induction on n.

Base Case (BS): $n = 1$ is trivial since $\Pr[A_1] = \Pr[A_1]$.

Induction Step (IS): Let $k \geq 1$ arbitrary. Assume (IH) that:

$$\Pr[A_1 \cup \ldots \cup A_k] = \Pr[A_1] + \ldots + \Pr[A_k]$$

Now let $B = A_1 \cup \ldots \cup A_k$. Since A_{k+1} is disjoint from each of $A_1, \ldots, A_k$ it is also disjoint from B. Hence, by (P2), we have that $\Pr[B \cup A_{k+1}] = \Pr[B] + \Pr[A_{k+1}]$. Using also the (IH) we obtain:

$$\Pr[A_1 \cup \ldots \cup A_k \cup A_{k+1}] = \Pr[B \cup A_{k+1}] = \Pr[B] + \Pr[A_{k+1}] = \Pr[A_1] + \ldots + \Pr[A_k] + \Pr[A_{k+1}]$$

(CONTINUED)

Proofs of Probability Properties P0-P7 (continued)

Now we will prove (P3) which is called **monotonicity** and states that:

$$\textbf{(P3)} \quad A \subseteq B \Rightarrow \Pr[A] \leq \Pr[B]$$

Proof. To prove (P3) we observe that if $A \subseteq B$ then $B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \emptyset$, so we can apply (P2):

$$\Pr[B] = \Pr[A] + \Pr[B \setminus A] \geq \Pr[A]$$

because $\Pr[B \setminus A] \geq 0$

Next we will prove (P4) which states that:

$$\textbf{(P4)} \quad \Pr[\overline{E}] = 1 - \Pr[E] \text{ where } \overline{E} = \Omega \setminus E$$

Proof. To prove (P4) we observe that $E \cap \overline{E} = \emptyset$ and $E \cup \overline{E} = \Omega$ so we have:

$$\begin{aligned} \Pr[E \cup \overline{E}] &= \Pr[E] + \Pr[\overline{E}] \quad && \text{by (P2)} \\ \Pr[E \cup \overline{E}] &= \Pr[\Omega] = 1 && \text{by (P1)} \end{aligned}$$

It follows that:

$$\Pr[E] + \Pr[\overline{E}] = 1 \qquad \text{therefore} \qquad \Pr[\overline{E}] = 1 - \Pr[E]$$

(CONTINUED)

Proofs of Probability Properties P0-P7 (continued)

Recall that you proved (P5) in an activity in the segment ”Probability properties”:

$$\textbf{(P5)} \quad \Pr[\emptyset] = 0$$

We repeat the proof here.

Proof. To prove (P5) we observe that $\emptyset = \overline{\Omega}$ so we have:

$$\begin{aligned} \Pr[\emptyset] = \Pr[\overline{\Omega}] &= 1 - \Pr[\Omega] && \text{by (P4)} \\ &= 1 - 1 = 0 && \text{by (P1)} \end{aligned}$$

Now we will prove (P6) which as mentioned in the segment ”Probability properties” is called **inclusion-exclusion** and states that:

$$\textbf{(P6)} \quad \Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$$

Proof. To prove (P6) we apply (P2) twice. Since A and $B \setminus A$ are disjoint and $A \cup (B \setminus A) = A \cup B$ we have by (P2):

$$\Pr[A \cup B] = \Pr[A] + \Pr[B \setminus A]$$

And since $B \setminus A$ and $A \cap B$ are disjoint and $(B \setminus A) \cup (A \cap B) = B$ we also have, again by (P2):

$$\Pr[B] = \Pr[B \setminus A] + \Pr[A \cap B]$$

Hence by some algebraic manipulation:

$$\begin{aligned} \Pr[A \cup B] &= \Pr[A] + \Pr[B \setminus A] \\ &= \Pr[A] + \Pr[B] - \Pr[A \cap B] \end{aligned}$$

(CONTINUED)

Proofs of Probability Properties P0-P7 (continued)

Now we will prove (P7) which is called **union bound**, and as mentioned in lecture it plays a major role in the analysis of probabilistic algorithms.

$$\textbf{(P7)} \quad \Pr[A \cup B] \leq \Pr[A] + \Pr[B]$$

Proof. The proof for (P7) follows directly from (P6) by noting that $\Pr[A \cap B] \geq 0$ by (P0).

Recall that in the segment "Probability properties" we also generalized (P7) for n finite sets as follows:

$$\textbf{(P7Gen)} \quad \Pr[A_1 \cup \ldots \cup A_n] \leq \Pr[A_1] + \ldots + \Pr[A_n]$$

Proof. We prove (P7Gen) by ordinary induction on n.

Base Case (BS): $n = 1$ is trivial since $\Pr[A_1] \leq \Pr[A_1]$.

Induction Step (IS): Let $k \geq 1$ arbitrary. Assume (IH) that:

$$\Pr[A_1 \cup \ldots \cup A_k] \leq \Pr[A_1] + \ldots + \Pr[A_k]$$

Now let $B = A_1 \cup \ldots \cup A_k$, and let A_{k+1} be another event in the same probability space. By (P6) we have that:

$$\Pr[B \cup A_{k+1}] = \Pr[B] + \Pr[A_{k+1}] - \Pr[B \cap A_{k+1}]$$

By (IH) we have that:

$$\Pr[B] = \Pr[A_1 \cup \ldots \cup A_k] \leq \Pr[A_1] + \ldots + \Pr[A_k]$$

And by (P0) we have that:

$$\Pr[B \cap A_{k+1}] \geq 0$$

It follows that:

$$\Pr[B \cup A_{k+1}] \leq \Pr[A_1] + \ldots + \Pr[A_k] + \Pr[A_{k+1}]$$

(CONTINUED)

Proofs of Probability Properties P0-P7 (continued)

We end with some observations. We noted in lecture that cardinality (of finite sets) has properties analogous to those above. In particular, it holds that:

$$\begin{array}{ll} \textbf{(P2)} & A \cap B = \emptyset \Rightarrow |A \cup B| = |A| + |B| \\ \textbf{(P3)} & A \subseteq B \Rightarrow |A| \leq |B| \\ \textbf{(P5)} & |\emptyset| = 0 \\ \textbf{(P6)} & |A \cup B| = |A| + |B| - |A \cap B| \\ \textbf{(P7)} & |A \cup B| \leq |A| + |B| \end{array}$$

This analogy is not accidental! Both cardinality and probability are instances of the mathematical concept of "measure". Other examples of measure are area and volume.

Moreover, you must have noticed that we proved only (P0), (P1) and (P2) from the definition of finite probability space. All the other properties were shown as consequence of the first three. This is also not accidental. The first three properties correspond to Kolmogorov's axiomatization of the general theory of probability on spaces of outcomes that can be infinite, but in which only "measurable" events have probability. By comparison, the probability theory on finite sets of outcomes is simpler since all events are finite and hence have probability.