Probability Theory Notes

January 21, 2021

1 Introduction to Probability Models (Ross)

1.1 Sample Space and Events

The set of all possible outcomes of an experiment is known as the *sample space* Ω (or denoted S) e.g. when flipping two coins (each once)

$$S = \{(H, H), (H, T), (T, H), (T, T)\} \tag{1}$$

Any subset E of the sample space is known as an *event* e.g. when the event is flipping a heads on the first coin

$$E = \{(H, H), (H, T)\} \tag{2}$$

We use the symbol $\cup$ (*union*)to denote "or", for example, if $E = \{1, 3, 5\}$ and $F = \{1, 2, 3\}$, then:

$$E \cup F = \{1, 2, 3, 5\} \tag{3}$$

We use the symbol $\cap$ (*intersection*) to denote "and", for example, if $E = \{1, 3, 5\}$ and $F = \{1, 2, 3\}$, then:

$$E \cap F = \{1, 3\} \tag{4}$$

Note that $E \cap F$ is more commonly written as EF

If E amd F are *mutually exclusive* e.g. $E = \{H\}$ and $F = \{T\}$, then null event (*the empty set*) is given by $\emptyset$

Can define unions and intersections of more than two events using:

$$\bigcap_{n=1}^{\infty} E_n \text{ and } \bigcup_{n=1}^{\infty} E_n \tag{5}$$

We use the symbol E^c to denote the *compliment* i.e. all outcomes in set that are not in E. Therefore $S^c = \emptyset$

1.2 Probabilities Defined on Events

For events E in S we assume:

$$0 \leq P(E) \leq 1 \tag{6}$$

$$P(S) = 1 \tag{7}$$

$$E_n E_m = \emptyset \text{ when } n \neq m \text{, then:} \tag{8}$$

$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} P(E_n) \tag{9}$$

For two events E and F, we must avoid double counting:

$$P(E \cup F) = P(E) + P(F) - P(EF) \tag{10}$$

Note that when $F = E^c$, $P(EF) = \emptyset$, so $P(E \cup F) = P(E) + P(F)$. This can be extended to multiple events, such as:

$$\begin{aligned}
P(E \cup F \cup G) &= (P((E \cup F) \cup G)) \\
&= P(E \cup F) + P(G) - P((E \cup F)G) \\
&= ... \\
&= P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)
\end{aligned}$$

More generally:

$$\begin{aligned}
P(E_1 \cup E_2 \cup \cdots \cup E_n) = &\sum_i P(E_i) - \sum_{i<j} P(E_i E_j) \\
&+ \sum_{i<j<k} P(E_i E_j E_k) - \cdots + (-1)^{n+1} P(E_1 E_2 \cdots E_n)
\end{aligned}$$

This is the *inclusion-exclusion identity* (when extended to n events), stating that the probability of the union of n events is the sum of the probabilities taken one at a time, minus the sum of the probabilities taken two at a time plus the sum of the probabilities taken three at a time ...

1.3 Conditional Probabilities