Construction of disjoint events from general events

Suppose that $A_1, A_2, ..., A_k, ...$ is a (countable) collection of general events in sample space Ω. Then the collection of events $E_1, E_2, ..., E_k, ...$ defined for $k \geq 1$ by

$$E_k = A_k \bigcap \left(\bigcup_{i=1}^{k-1} A_i \right)'$$

are pairwise mutually exclusive, but

$$\bigcup_{i=1}^{k} E_i = \bigcup_{i=1}^{k} A_i.$$

E_k is the set of sample outcomes in Ω that are elements of A_k but not elements of any of $A_1, ..., A_{k-1}$.

STRUCTURED COLLECTIONS OF EVENTS

DEFINITION

A (countable) collection of subsets, $\mathcal{A}$, of sample space Ω, say

$$\mathcal{A} = \{A_1, A_2, ...\},$$

is an **algebra** if

(I) $\Omega \in \mathcal{A}$

(II) $A_1, A_2 \in \mathcal{A} \implies A_1 \cup A_2 \in \mathcal{A}$

(III) $A \in \mathcal{A} \implies A' \in \mathcal{A}$

Interpretation: An algebra is a set of sets (events) with certain properties, in particular it is *closed* under a **finite** number of union operations (II), that is if $A_1, ...A_k \in \mathcal{A}$, then

$$\bigcup_{i=1}^{k} A_i \in \mathcal{A}.$$

NOTE: If $\mathcal{A}$ is an algebra of subsets of Ω, then

(i) $\emptyset \in \mathcal{A}$

(ii) If $A_1, A_2 \in \mathcal{A}$, then

$$A_1', A_2' \in \mathcal{A} \implies A_1' \cup A_2' \in \mathcal{A} \implies \left(A_1' \cup A_2'\right)' \in \mathcal{A} \implies A_1 \cap A_2 \in \mathcal{A}$$

so $\mathcal{A}$ is also *closed under intersection.*

EXAMPLE Consider $\Omega = [0, 1] = \{\omega : \omega \in \mathbb{R}, 0 \leq \omega \leq 1\}$, and the set of subsets of $[0, 1]$, $\mathcal{A}$, where the elements of $\mathcal{A}$ are countable unions of sub-intervals of $[0, 1]$. Thus, the elements of $\mathcal{A}$ are of the form

$$\bigcup_{i=1}^{k} A_i \quad \text{where} \quad A_i = (a_i, b_i] \quad \text{i=1,2,...,k}$$

for integer $k \geq 0$ real numbers a_i, b_i such that $0 \leq a_i \leq b_i \leq 1$. Then $\mathcal{A}$ is an algebra.

DEFINITION

A (countable) collection of subsets, $\mathcal{A}$, of sample space Ω, say

$$\mathcal{A} = \{A_1, A_2, ...\},$$

is a **σ-algebra** if

(I) $\Omega \in \mathcal{A}$

(II) $A \in \mathcal{A} \implies A' \in \mathcal{A}$

(III) $A_1, A_2, ... \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

Interpretation: A σ-algebra is an algebra that is *closed* under a **countable** (rather than merely finite) number of union operations.

SEQUENCES OF EVENTS

DEFINITION

A **sequence** of events $A_1, A_2, ..., A_k, ...$, denoted $\{A_k\}$, is **non-increasing** if

$$A_1 \supseteq A_2 \supseteq ... \supseteq A_k \supseteq ...$$

and is **non-decreasing** if

$$A_1 \subseteq A_2 \subseteq ... \subseteq A_k \subseteq ...$$

A sequence that is either non-increasing or non-decreasing is termed **monotone**.

NOTES:

(i) If $\{A_k\}$ is non-increasing, then the sequence $\left\{A_k'\right\}$ is non-decreasing, as for $k \geq 1$,

$$A_k \supseteq A_{k+1} \Longrightarrow A_k = A_{k+1} \cup D_k \text{ for some event} D_k$$

$$\Longrightarrow A_k' = (A_{k+1} \cup D_k)' = A_{k+1}' \cap D_k' \subseteq A_{k+1}'$$

Hence, for $k \geq 1$, $A_k' \subseteq A_{k+1}'$.

(ii) If $\{A_k\}$ is non-decreasing, then $\left\{A_k'\right\}$ is non-increasing. Also, for $k \geq 1$, define event E_k by

$$E_k = A_{k+1} \cap A_k'.$$

Then $\{E_k\}$ is a collection of **disjoint** events, and

$$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} A_i$$

DEFINITION

Consider an infinite monotone sequence of events $\{A_k\}$. Then the **limit event** of the sequence, A, is written

$$A = \lim_{k \longrightarrow \infty} A_k$$

and is defined by

$$A = \begin{cases} \displaystyle\bigcap_{k=1}^{\infty} A_k & \text{if } \{A_k\} \text{ is NON-INCREASING} \\ \displaystyle\bigcup_{k=1}^{\infty} A_k & \text{if } \{A_k\} \text{ is NON-DECREASING} \end{cases}$$

Interpretation: If $\{A_k\}$ is non-increasing, then the limit event A corresponds to the collection of sample outcomes that are elements of **all** of the events in the sequence. Thus A occurs if and only if all of $A_1, A_2, ..., A_k, ...$ occur.

If $\{A_k\}$ is non-decreasing, then the limit event A corresponds to the collection of sample outcomes that are elements of **at least one** of the events in the sequence. Thus A occurs if and only if at least one of $A_1, A_2, ..., A_k, ...$ occurs.

NOTE: These limit definitions are only valid for monotone sequences of events.

CONSTRUCTING MONOTONE SEQUENCES

Consider an infinite *general* sequence of events $\{E_k\}$, and the infinite sequences of events $\{A_k\}$ and $\{B_k\}$ uniquely defined by

$$A_k = \bigcup_{j=k}^{\infty} E_j \qquad B_k = \bigcap_{j=k}^{\infty} E_j$$

Then $\{A_k\}$ is a non-increasing sequence and $\{B_k\}$ is a non-decreasing sequence.

DEFINITION

Consider an infinite general sequence of events $\{E_k\}$. Then the **limit superior** (or **limsup**) event, denoted $E^{(S)}$, is defined by

$$E^{(S)} = \limsup_{k \longrightarrow \infty} E_k = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} E_j$$

and the **limit inferior** (or **liminf**) event, denoted $E^{(I)}$, is defined by

$$E^{(I)} = \liminf_{k \longrightarrow \infty} E_k = \bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} E_j$$

NOTE: $E^{(S)}$ is the limit of the non-increasing sequence $\{A_k\}$ of the previous definition, whereas $E^{(I)}$ is the limit of the non-decreasing sequence $\{B_k\}$.

DEFINITION

An infinite general sequence of events $\{E_k\}$ is **convergent** if

$$E^{(S)} = E^{(I)}.$$

Interpretation: $E^{(S)}$ is the collection of sample outcomes that are in *infinitely many* of the events $\{E_k\}$, and $E^{(S)}$ occurs if and only if infinitely many of the E_k events occur.

$E^{(I)}$ is the collection of sample outcomes that are in *all except a finite number* of the events $\{E_k\}$, and $E^{(I)}$ occurs if and only if all except a finite number of the E_k events occur.

Hint: always read "$\bigcup$" as "There exists" and "$\bigcap$" as "For all". Hence to say the event

$$E^{(S)} = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} E_j$$

occurs is equivalent to saying

"**For all** $k \geq 1$, **there exists** a $j \geq k$ such that E_j occurs".

Similarly, to say the event

$$E^{(I)} = \bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} E_j$$

occurs is equivalent to saying

"**There exists** a $k \geq 1$, such that, **for all** $j \geq k$, E_j occurs".

NOTE

Let $k, l \geq 1$. Then, if $i \geq \max(k, l)$, it follows that

$$\bigcap_{j=k}^{\infty} E_j \subseteq E_i \subseteq \bigcup_{j=l}^{\infty} E_j$$

and taking union over k, and then intersection over l, it follows that

$$\bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} E_j \subseteq \bigcap_{l=1}^{\infty} \bigcup_{j=l}^{\infty} E_j$$

so

$$E^{(I)} \subseteq E^{(S)}.$$

Clearly, if $\omega \in E^{(I)}$ then, ω is in *all except a finite number* of E_ks, and therefore ω is in *infinitely many* of them, so $\omega \in E^{(S)}$.

EXAMPLE

Suppose that $\{E_k\}$ is an infinite sequence with $E_i \cap E_j = \emptyset$ for $i \neq j$. Then, for sample outcome ω,

$$\omega \in E^{(S)} \Longrightarrow \omega \in E_k$$

for infinitely many of the E_k. But the E_k events are disjoint, and no such ω exists. Hence

$$E^{(S)} = \emptyset$$

and therefore, as $E^{(I)} \subseteq E^{(S)}$,

$$E^{(I)} = \emptyset$$

and the sequence is convergent.

1.3. THE PROBABILITY FUNCTION

For an event $E \subseteq \Omega$, the **probability that** E **occurs** is written $\mathrm{P}(E)$.

Interpretation : $\mathrm{P}(.)$ is a *set-function* that assigns "weight" to collections of possible outcomes of an experiment. There are many ways to think about precisely how this assignment is achieved;

CLASSICAL : "Consider equally likely sample outcomes ..."

FREQUENTIST : "Consider long-run *relative frequencies* ..."

SUBJECTIVE : "Consider personal degree of belief ..."

or merely think of $\mathrm{P}(.)$ as a set-function.

1.4. PROPERTIES OF P(.): THE AXIOMS OF PROBABILITY

AXIOM 1 Let $E \subseteq \Omega$. Then $0 \leq \mathrm{P}(E) \leq 1$.

AXIOM 2 $\mathrm{P}(\Omega) = 1$.

AXIOM 3 If $E, F \subseteq \Omega$, with $E \cap F = \emptyset$, then $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F)$.

Corollaries : For events $E, F \subseteq \Omega$

$\mathrm{P}(E') = 1 - \mathrm{P}(E)$.
$\mathrm{P}(\emptyset) = 0$.
If $E \subseteq F$, then $\mathrm{P}(E) \leq \mathrm{P}(F)$.
In general, $\mathrm{P}(E \cup F) = \mathrm{P}(E) + \mathrm{P}(F) - \mathrm{P}(E \cap F)$.

Extension : If events $E_1, ..., E_k$ are **pairwise** mutually exclusive, then

$$\mathrm{P}\left(\bigcup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} \mathrm{P}(E_i)$$

Example: Consider an experiment with possible outcomes $s_1, ..., s_k$, and let E_i correspond to the event "outcome is s_i". Suppose $p_i = \mathrm{P}(E_i)$. Then the axioms require that, for $i = 1, ..., k$,

(a) $0 \leq p_i \leq 1$
(b) $p_1 + ... + p_k = 1$.

1.5. CONDITIONAL PROBABILITY

For events $E, F \subseteq \Omega$ the **conditional probability** that F occurs **given** that E has occurred is written $\mathrm{P}(F|E)$, and is defined by

$$\mathrm{P}(F|E) = \frac{\mathrm{P}(E \cap F)}{\mathrm{P}(E)}$$

if $\mathrm{P}(E) > 0$.

NOTE: $\mathrm{P}(E \cap F) = \mathrm{P}(E)\mathrm{P}(F|E)$, and in general, for events $E_1, ..., E_k$,

$$\mathrm{P}\left(\bigcap_{i=1}^{k} E_i\right) = \mathrm{P}(E_1)\mathrm{P}(E_2|E_1)\mathrm{P}(E_2|E_1 \cap E_2)...\mathrm{P}(E_k|E_1 \cap E_2 \cap ... \cap E_{k-1}).$$

This result is known as the CHAIN or MULTIPLICATION RULE.

DEFINITION

Events E and F are **independent** if

$$\mathrm{P}(E|F) = \mathrm{P}(E)$$

so that

$$\mathrm{P}(E \cap F) = \mathrm{P}(E)\mathrm{P}(F)$$

Extension : Events $E_1, ..., E_k$ are independent if, for **every** subset of events of size $l \leq k$, indexed by $\{i_1, ..., i_l\}$, say,

$$\mathrm{P}\left(\bigcap_{j=1}^{l} E_{i_j}\right) = \prod_{j=1}^{l} \mathrm{P}(E_{i_j}).$$

1.6. THE THEOREM OF TOTAL PROBABILITY

THEOREM

Let $E_1, ..., E_k$ be a partition of Ω, and let $F \subseteq \Omega$. Then

$$\mathrm{P}(F) = \sum_{i=1}^{k} \mathrm{P}(F|E_i)\mathrm{P}(E_i)$$

Proof

$E_1, ..., E_k$ form a partition of Ω, and $F \subseteq \Omega$, so

$$\begin{aligned} F &= (F \cap E_1) \cup ... \cup (F \cap E_k) \\ \Rightarrow \quad \mathrm{P}(F) &= \sum_{i=1}^{k} \mathrm{P}(F \cap E_i) \qquad \text{(by AXIOM 3, as } E_i \cap E_j = \emptyset) \\ &= \sum_{i=1}^{k} \mathrm{P}(F|E_i)\mathrm{P}(E_i) \end{aligned}$$

Example : See Simpson's Paradox.

1.7. BAYES THEOREM

THEOREM

Suppose $E, F \subseteq \Omega$, with $\mathrm{P}(E), \mathrm{P}(F) > 0$. Then

$$\mathrm{P}(E|F) = \frac{\mathrm{P}(F|E)\mathrm{P}(E)}{\mathrm{P}(F)}$$

Proof

$$\mathrm{P}(E|F)\mathrm{P}(F) = \mathrm{P}(E \cap F) = \mathrm{P}(F|E)\mathrm{P}(E), \text{ so } \mathrm{P}(E|F)\mathrm{P}(F) = \mathrm{P}(F|E)\mathrm{P}(E).$$

NOTE: in general, $\mathrm{P}(E|F) \neq \mathrm{P}(F|E)$ - see Medical Testing example.

1.8. COUNTING TECHNIQUES

Suppose that an experiment has N equally likely sample outcomes. If event E corresponds to a collection of sample outcomes of size $n(E)$, then

$$\mathrm{P}(E) = \frac{n(E)}{N}$$

so it is necessary to be able to evaluate $n(E)$ and N in practice.

Multiplication principle

If operations labelled $1, ..., r$ can be carried out in $n_1, ..., n_r$ ways respectively, then there are $\prod_{i=1}^{r} n_i = n_1 ... n_r$ ways of carrying out the r operations in total.

EXAMPLE : If each of r trials of an experiment has N possible outcomes, then there are N^r possible sequences of outcomes in total.

e.g.(i) If a multiple choice exam has 20 questions, each of which has 5 possible answers, then there are 5^{20} different ways of completing the exam.

(ii) There are 2^m subsets of m elements (as each element is either **in** the subset, or **not in** the subset, which is equivalent to m trials each with two outcomes).

Sampling from a finite population

Consider a collection of N items, and a sequence of operations labelled $1, ..., r$ such that the ith operation involves **selecting** one of the items remaining after the first $i - 1$ operations have been carried out. Let n_i denote the number of ways of carrying out the ith operation, for $i = 1, ..., r$. Then there are two distinct cases;

(a) **Sampling with replacement :** an item is returned to the collection after selection. Then $n_i = N$ for all i, and there are N^r ways of carrying out the r operations.

(b) **Sampling without replacement :** an item is not returned to the collection after selected. Then $n_i = N - i + 1$, and there are $N(N - 1)...(N - r + 1)$ ways of carrying out the r operations.

e.g. Consider selecting 5 cards from 52. Then (a) leads to 52^5 possible selections, whereas (b) leads to 52.51.50.49.48 possible selections

Note: The **order** in which the operations are carried out may be important

e.g. in a raffle with three prizes and 100 tickets, the draw $\{45, 19, 76\}$ is different from $\{19, 76, 45\}$.

Note: the items may be **distinct** (unique in the collection), or **indistinct** (of a unique type in the collection, but not unique individually).

e.g. The numbered balls in the National Lottery, or individual playing cards, are **distinct**. However balls in the lottery are regarded as "WINNING" or "NOT WINNING", or playing cards are regarded in terms of their suit only, are **indistinct**.

PERMUTATIONS AND COMBINATIONS

DEFINITION

A **permutation** is an *ordered* arrangement of a set of items.

A **combination** is an *unordered* arrangement of a set of items.

RESULT 1 The number of permutations of n distinct items is $n! = n(n-1)...1$.

RESULT 2 The number of permutations of r from n distinct items is

$$P_r^n = \frac{n!}{(n-r)!} = n(n-1)...(n-r+1) \qquad \text{(by the Multiplication Principle).}$$

If the **order** in which items are selected is not important, then

RESULT 3 The number of combinations of r from n distinct items is

$$C_r^n = \begin{pmatrix} n \\ r \end{pmatrix} = \frac{n!}{r!(n-r)!} \qquad (\text{as } P_r^n = r!C_r^n).$$

-recall the **Binomial Theorem**, namely

$$(a+b)^n = \sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} a^i b^{n-i}.$$

Then the number of subsets of m items can be calculated as follows; for each $0 \leq j \leq m$, choose a subset of j items from m. Then

$$\text{Total number of subsets} = \sum_{j=0}^{m} \begin{pmatrix} m \\ j \end{pmatrix} = (1+1)^m = 2^m.$$

If the items are **indistinct**, but each is of a unique type, say Type I, ..., Type κ say, (the so-called **Urn Model**) then

RESULT 4 The number of distinguishable permutations of n indistinct objects, comprising n_i items of type i for $i = 1, ..., \kappa$ is $\dfrac{n!}{n_1!n_2!...n_\kappa!}$.

Special Case : if $\kappa = 2$, then the number of distinguishable permutations of the n_1 objects of type I, and $n_2 = n - n_1$ objects of type II is $C_{n_2}^n = \dfrac{n!}{n_1!(n-n_1)!}$.

Also, there are C_r^n ways of partitioning n **distinct** items into two "cells", with r in one cell and $n-r$ in the other.

PROBABILITY CALCULATIONS

Recall that if an experiment has N equally likely sample outcomes, and event E corresponds to a collection of sample outcomes of size $n(E)$, then

$$\mathrm{P}(E) = \frac{n(E)}{N}$$

EXAMPLE 1. A True/False exam has 20 questions. Let

$$E = \text{“16 answers correct at random”}$$

Then

$$\mathrm{P}(E) = \frac{\text{Number of ways of getting 16 out of 20 correct}}{\text{Total number of ways of answering 20 questions}} = \frac{\binom{20}{16}}{2^{20}} = 0.0046$$

EXAMPLE 2. Sampling without replacement.

Consider an Urn Model with 10 Type I objects and 20 Type II objects, and an experiment involving sampling five objects without replacement. Let

$$E = \text{“precisely 2 Type I objects selected”}$$

We need to calculate N and $n(E)$ in order to calculate $\mathrm{P}(E)$. In this case N is the number of ways of choosing 5 from 30 items, and hence

$$N = \binom{30}{5}.$$

To calculate $n(E)$, we think of E occurring by first choosing 2 Type I objects from 10, and then choosing 3 Type II objects from 20, and hence, by the multiplication rule,

$$n(E) = \binom{10}{2}\binom{20}{3}.$$

Therefore

$$\mathrm{P}(E) = \frac{\binom{10}{2}\binom{20}{3}}{\binom{30}{5}} = 0.360$$

NOTE: This result can be obtained using a conditional probability argument; consider event $F \subseteq E$, where

$$F = \text{“sequence of objects 11222 obtained”}$$

Then $F = \bigcap_{i=1}^{5} F_{ij}$ where

$$F_{ij} = \text{“type } j \text{ object obtained on draw } i\text{”} \quad i = 1, ..., 5, j = 1, 2$$

Then

$$\begin{aligned} \mathrm{P}(F) &= \mathrm{P}(F_{11})\mathrm{P}(F_{21}|F_{11})...\mathrm{P}(F_{52}|F_{11}, F_{21}, F_{32}, F_{42}) \\ &= \frac{10}{30}\frac{9}{29}\frac{20}{28}\frac{19}{27}\frac{18}{26} \end{aligned}$$

Now consider event G where

$$G = \text{"sequence of objects 12122 obtained"}$$

Then

$$\mathrm{P}(G) = \frac{10}{30}\frac{20}{29}\frac{9}{28}\frac{19}{27}\frac{18}{26}$$

i.e. $\mathrm{P}(G) = \mathrm{P}(F)$. In fact, **any** sequence containing two Type I and three Type II objects has this probability, and there are $\binom{5}{2}$ such sequences. Thus, as all such sequences are mutually exclusive,

$$\begin{aligned}\mathrm{P}(E) &= \binom{5}{2}\frac{10}{30}\frac{9}{29}\frac{20}{28}\frac{19}{27}\frac{18}{26} \\ &= \frac{\binom{10}{2}\binom{20}{3}}{\binom{30}{5}}\end{aligned}$$

as before.

EXAMPLE 3. Sampling with replacement.

Consider an Urn Model with 10 Type I objects and 20 Type II objects, and an experiment involving sampling five objects with replacement. Let

$$\mathrm{E} = \text{"precisely 2 Type I objects selected"}$$

Again, we need to calculate N and $n(E)$ in order to calculate $\mathrm{P}(E)$. In this case N is the number of ways of choosing 5 from 30 items with replacement, and hence

$$N = 30^5$$

To calculate $n(E)$, we think of E occurring by first choosing 2 Type I objects from 10, and 3 Type II objects from 20 in any order. Consider such sequences of selection

Sequence	Number of ways
11222	10.10.20.20.20
12122	10.20.10.20.20
.	.

etc., and thus a sequence with 2 Type I objects and 3 Type II objects can be obtained in $10^2 20^3$ ways. As before there are $\binom{5}{2}$ such sequences, and thus

$$\mathrm{P}(E) = \frac{\binom{5}{2}10^2 20^3}{30^5} = 0.329.$$

Again, this result can be obtained using a conditional probability argument; consider event $F \subseteq E$, where

$$F = \text{“sequence of objects 11222 obtained”}$$

Then

$$\mathrm{P}(F) = \left(\frac{10}{30}\right)^2 \left(\frac{20}{30}\right)^3$$

as the results of the draws are **independent**. This result is true for any sequence containing two Type I and three Type II objects, and there are $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ such sequences that are mutually exclusive, so

$$\mathrm{P}(E) = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \left(\frac{10}{30}\right)^2 \left(\frac{20}{30}\right)^3$$