

# 01. COMBINATORIAL ANALYSIS

tricky - E18, E20-22, E23, E26

## The Basic Principle of Counting

- combinatorial analysis** → the mathematical theory of counting
- basic principle of counting** → Suppose that two experiments are performed. If exp1 can result in any one of  $m$  possible outcomes and if, for each outcome of exp1, there are  $n$  possible outcomes of exp2, then together there are  $mn$  possible outcomes of the two experiments.
- generalized basic principle of counting** → If  $r$  experiments are performed such that the first one may result in any of  $n_1$  possible outcomes and if for each of these  $n_1$  possible outcomes, and if ..., then there is a total of  $n_1 \cdot n_2 \cdot \dots \cdot n_r$  possible outcomes of  $r$  experiments.

## Permutations

**factorials** -  $1! = 0! = 1$

**N1** - if we know how to count the number of different ways that an event can occur, we will know the probability of the event.

**N2** - there are  $n!$  different arrangements for  $n$  objects.

**N3** - there are  $\frac{n!}{n_1! n_2! \dots n_r!}$  different arrangements of  $n$  objects, of which  $n_1$  are alike,  $n_2$  are alike, ...,  $n_r$  are alike.

## Combinations

**N4** -  $\binom{n}{r} = \frac{n!}{(n-r)! r!}$  represents the number of different groups of size  $r$  that could be selected from a set of  $n$  objects when the order of selection is not considered relevant.

**N4b** -  $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$ ,  $1 \leq r \leq n$

*Proof.* If object 1 is chosen  $\Rightarrow \binom{n-1}{r-1}$  ways of choosing the remaining objects.

If object 1 is not chosen  $\Rightarrow \binom{n-1}{r}$  ways of choosing the remaining objects.

**N5 - The Binomial Theorem** -  $(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$

*Proof.* by mathematical induction:  $n = 1$  is true; expand; sub dummy variable; combine using N4b; combine back to final term

## Multinomial Coefficients

**N6** -  $\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \dots n_r!}$  represents the number of possible divisions of  $n$  distinct objects into  $r$  distinct groups of respective sizes  $n_1, n_2, \dots, n_r$ , where  $n_1 + n_2 + \dots + n_r = n$

*Proof.* using basic counting principle,

$$\begin{aligned} &= \binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \dots \binom{n-n_1-n_2-\dots-n_{r-1}}{n_r} \\ &= \frac{n!}{(n-n_1)! n_1!} \times \frac{(n-n_1)!}{(n-n_1-n_2)! n_2!} \times \dots \times \frac{(n-n_1-n_2-\dots-n_{r-1})!}{0! n_r!} \\ &= \frac{n!}{n_1! n_2! \dots n_r!} \end{aligned}$$

**N7 - The Multinomial Theorem:**  $(x_1 + x_2 + \dots + x_r)^n = \sum_{(n_1, \dots, n_r): n_1 + n_2 + \dots + n_r = n} \frac{n!}{n_1! n_2! \dots n_r!} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}$

## Number of Integer Solutions of Equations

**N8** - there are  $\binom{n-1}{r-1}$  distinct *positive* integer-valued vectors  $(x_1, x_2, \dots, x_r)$  satisfying  $x_1 + x_2 + \dots + x_r = n$ ,  $x_i > 0$ ,  $i = 1, 2, \dots, r$

! cannot be directly applied to N8 as 0 value is not included

**N9** - there are  $\binom{n+r-1}{r-1}$  distinct *non-negative* integer-valued vectors

$(x_1, x_2, \dots, x_r)$  satisfying  $x_1 + x_2 + \dots + x_r = n$

*Proof.* let  $y_k = x_k + 1 \Rightarrow y_1 + y_2 + \dots + y_r = n + r$

# 02. AXIOMS OF PROBABILITY

## Sample Space and Events

- sample space** → The *set* of all outcomes of an experiment (where outcomes are not predictable with certainty)
- event** → Any *subset* of the sample space
- union** of events  $E$  and  $F \rightarrow E \cup F$  is the event that contains all outcomes that are either in  $E$  or  $F$  (or both).
- intersection** of events  $E$  and  $F \rightarrow E \cap F$  or  $EF$  is the event that contains all outcomes that are both in  $E$  and in  $F$ .
- complement** of  $E \rightarrow E^c$  is the event that contains all outcomes that are *not* in  $E$ .
- subset** →  $E \subset F$  if all of the outcomes in  $E$  that are also in  $F$ .
  - $E \subset F \wedge F \subset E \Rightarrow E = F$

## DeMorgan's Laws

$$\left( \bigcup_{i=1}^n E_i \right)^c = \bigcap_{i=1}^n E_i^c$$

*Proof.* to show  $LHS \subset RHS$ : let  $x \in \left( \bigcup_{i=1}^n E_i \right)^c$   
 $\Rightarrow x \notin \bigcup_{i=1}^n E_i \Rightarrow x \notin E_1$  and  $x \notin E_2 \dots$  and  $x \notin E_n$   
 $\Rightarrow x \in E_1^c$  and  $x \in E_2^c \dots$  and  $x \in E_n^c$   
 $\Rightarrow x \in \bigcap_{i=1}^n E_i^c$   
 to show  $RHS \subset LHS$ : let  $x \in \bigcap_{i=1}^n E_i^c$

$$\left( \bigcap_{i=1}^n E_i \right)^c = \bigcup_{i=1}^n E_i^c$$

*Proof.* using the first law of DeMorgan, negate LHS to get RHS

## Axioms of Probability

### definition 1: relative frequency

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}$$

problems with this definition:

- $\frac{n(E)}{n}$  may not converge when  $n \rightarrow \infty$
- $\frac{n(E)}{n}$  may not converge to the same value if the experiment is repeated

### definition 2: Axioms

Consider an experiment with sample space  $S$ . For each event  $E$  of the sample space  $S$ , we assume that a number  $P(E)$  is defined and satisfies the following 3 axioms:

- $0 \leq P(E) \leq 1$
- $P(S) = 1$
- For any sequence of mutually exclusive events  $E_1, E_2, \dots$  (i.e., events for which  $E_i E_j = \emptyset$  when  $i \neq j$ ),

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

$P(E)$  is the probability of event  $E$ .

## Simple Propositions

**N1** -  $P(\emptyset) = 0$

**N2** -  $P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i)$  (aka axiom 3 for a finite  $n$ )

**N3 - strong law of large numbers** - if an experiment is repeated over and over again, then with probability 1, the proportion of time during which any specific event  $E$  occurs will be equal to  $P(E)$ .

**N6** - the definitions of probability are mathematical definitions. They tell us which set functions can be called **probability functions**. They do not tell us what value a probability function  $P(\cdot)$  assigns to a given event  $E$ .

probability function  $\iff$  it satisfies the 3 axioms.

**N7** -  $P(E^c) = 1 - P(E)$

**N8** - if  $E \subset F$ , then  $P(E) \leq P(F)$

**N9** -  $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

**N10** - Inclusion-Exclusion identity where  $n = 3$

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

**N11 - Inclusion-Exclusion identity** -

$$\begin{aligned} P(E_1 \cup E_2 \cup \dots \cup E_n) &= \sum_{i=1}^n P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} E_{i_2}) + \dots \\ &\quad + (-1)^{r+1} \sum_{i_1 < i_2 < \dots < i_r} P(E_{i_1} E_{i_2} \dots E_{i_r}) + \dots \\ &\quad + (-1)^{n+1} P(E_1 E_2 \dots E_n) \end{aligned}$$

*Proof.* Suppose an outcome with probability  $\omega$  is in exactly  $m$  of the events  $E_i$ , where  $m > 0$ . Then

**LHS:** the outcome is in  $E_1 \cup E_2 \cup \dots \cup E_n$  and  $\omega$  will be counted once in  $P(E_1 \cup E_2 \cup \dots \cup E_n)$

**RHS:**

- the outcome is in exactly  $m$  of the events  $E_i$  and  $\omega$  will be counted exactly  $\binom{m}{1}$  times in  $\sum_{i=1}^n P(E_i)$

- the outcome is contained in  $\binom{m}{2}$  subsets of the type  $E_{i_1} E_{i_2}$  and  $\omega$  will be counted  $\binom{m}{2}$  times in  $\sum_{i_1 < i_2} P(E_{i_1} E_{i_2})$

- ... and so on

hence  $RHS = \binom{m}{1} \omega - \binom{m}{2} \omega + \binom{m}{3} \omega - \dots \pm \binom{m}{m} \omega$

$$\begin{aligned} &= \omega \sum_{i=0}^m \binom{m}{i} (-1)^i = \text{binomial theorem where } x = -1, y = 1 \\ &= 0 = LHS \end{aligned}$$

e.g. For an outcome with probability  $\omega$  and  $n = 3$

- Case 1.**  $w = P(E_1 E_2)$   
 LHS =  $\omega$   
 RHS =  $(\omega + \omega + 0) - (\omega + 0 + 0) + 0 = \omega$
- Case 2.**  $\omega = P(E_1 \cap E_2 \cap E_3)$   
 LHS =  $\omega$   
 RHS =  $(\omega + \omega + \omega) - (\omega + \omega + \omega) + \omega = \omega$

**N12** -

$$(i) \quad P\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n P(E_i)$$

$$(ii) \quad P\left(\bigcup_{i=1}^n E_i\right) \geq \sum_{i=1}^n P(E_i) - \sum_{j < i} P(E_i E_j)$$

$$(iii) \quad P\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n P(E_i) - \sum_{j < i} P(E_i E_j) + \sum_{k < j < i} P(E_i E_j E_k)$$

(iv) and so on.

*Proof.*  $\bigcup_{i=1}^n E_i = E_1 \cup E_1^c E_2 \cup E_1^c E_2^c E_3 \cup \dots \cup E_1^c E_2^c \dots E_{n-1}^c E_n$

$$P\left(\bigcup_{i=1}^n E_i\right) = P(E_1) + P(E_1^c E_2) + P(E_1^c E_2^c E_3) + \dots + P(E_1^c E_2^c \dots E_{n-1}^c E_n)$$

## Sample Space having Equally Likely Outcomes

tricky - 14, 15, 16, 18, 19, 20

Consider an experiment with sample space  $S = \{e_1, e_2, \dots, e_n\}$ . Then  $P(\{e_1\}) = P(\{e_2\}) = \dots = P(\{e_n\}) = \frac{1}{n}$  or  $P(\{e_i\}) = \frac{1}{n}$ .

**N1** - for any event  $E$ ,  $P(E) = \frac{\# \text{ of outcomes in } E}{\# \text{ of outcomes in } S} = \frac{\# \text{ of outcomes in } E}{n}$

**increasing sequence** of events  $\{E_n, n \geq 1\} \rightarrow$

$E_1 \subset E_2 \subset \dots \subset E_n \subset E_{n+1} \subset \dots$

$$\lim_{n \rightarrow \infty} E_n = \bigcup_{i=1}^{\infty} E_i$$

**decreasing sequence** of events  $\{E_n, n \geq 1\} \rightarrow$

$$E_1 \supset E_2 \supset \dots \supset E_n \supset E_{n+1} \supset \dots$$

$$\lim_{n \rightarrow \infty} E_n = \bigcap_{i=1}^{\infty} E_i$$

### 03. CONDITIONAL PROBABILITY AND INDEPENDENCE

tricky - E6, urns (p.37)

#### Conditional Probability

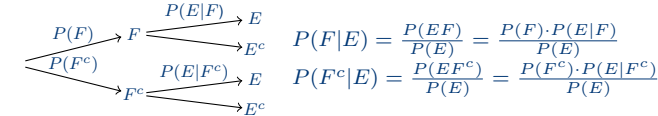
**N1** - if  $P(F) > 0$ . then  $P(E|F) = \frac{P(E \cap F)}{P(F)}$

**N2** - **multiplication rule** -  $P(E_1 E_2 \dots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2) \dots P(E_n|E_1 E_2 \dots E_{n-1})$

#### Total Probability & Bayes' Theorem

**conditioning formula** -  $P(E) = P(E|F)P(F) + P(E|F^c)P(F^c)$

**tree diagram** -



#### Total Probability

**theorem of total probability** - Suppose  $F_1, F_2, \dots, F_n$  are mutually exclusive

events such that  $\bigcup_{i=1}^n F_i = S$ , then  $P(E) = \sum_{i=1}^n P(EF_i) = \sum_{i=1}^n P(F_i)P(E|F_i)$

#### Bayes Theorem

$$\begin{aligned} P(F_j|E) &= \frac{P(EF_j)}{P(E)} \\ &= \frac{P(F_j)P(E|F_j)}{\sum_{i=1}^n P(F_i)P(E|F_i)} \end{aligned}$$

#### application of bayes' theorem

$$P(B_1 \mid A) = \frac{P(A|B_1) \cdot P(B_1)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)}$$

Let  $A$  be the event that the person test positive for a disease.

$B_1$ : the person actually has the disease.

$B_2$ : the person does not have the disease.

true positives: $P(B_1 \mid A)$	false negatives: $P(\bar{A} \mid B_1)$
false positives: $P(A \mid B_2)$	true negatives: $P(\bar{A} \mid B_2)$

#### Independent Events

**N1** -  $E$  and  $F$  are independent  $\iff$

$$P(EF) = P(E) \cdot P(F) \quad \text{or} \quad P(E|F) = P(E)$$

**N2** - if  $E$  and  $F$  are independent, then  $E$  and  $F^c$  are independent.

**N3** - if  $E, F, G$  are independent, then  $E$  will be independent of any event formed from  $F$  and  $G$ . (e.g.  $F \cup G$ )

<b>commutative</b>	$E \cup F = F \cup E$	$E \cap F = F \cap E$
<b>associative</b>	$(E \cup F) \cup G = E \cup (F \cup G)$	$(E \cap F) \cap G = E \cap (F \cap G)$
<b>distributive</b>	$(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$	$(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$
<b>DeMorgan's</b>	$(\bigcup_{i=1}^n E_i)^c = \bigcap_{i=1}^n E_i^c$	$(\bigcap_{i=1}^n E_i)^c = \bigcup_{i=1}^n E_i^c$