Probabilities: a *very* short course

CHANCE EVENT

— the outcome of an experiment
which may have various realisations
Ω — sample space (event space): e_i or A_i — an event.

Ω — sample space is the set of all possible results (outcomes) of a given *statistical experiment, or sampling*.

example — tossing a die

$$\left.\begin{array}{c}
(e_1) \bigcup (e_2) \bigcup (e_3) \bigcup (e_4) \bigcup (e_5) \bigcup (e_6) \\
(e_1 \bigcup e_2),\ (e_1 \bigcup e_3), \ldots,\ (e_5 \bigcup e_6) \\
(e_1 \bigcup e_2 \bigcup e_3),\ (e_1 \bigcup e_2 \bigcup e_4), \ldots,\ (e_4 \bigcup e_5 \bigcup e_6) \\
(e_1 \bigcup e_2 \bigcup e_3 \bigcup e_4) \equiv (\bar{e}_5 \bigcap \bar{e}_6), \ldots \\
(e_1 \bigcup e_2 \bigcup e_3 \bigcup e_4 \bigcup e_5) \equiv (\bar{e}_6), \ldots \\
+ \text{ an event which must happen; e.g. } (e_1 \bigcup e_2 \bigcup e_3 \bigcup e_4 \bigcup e_5 \bigcup e_6) \\
+ \text{ an event which cannot happen: } (\bar{e}_1 \bigcap \bar{e}_2 \bigcap \bar{e}_3 \bigcap \bar{e}_4 \bigcap \bar{e}_5 \bigcap \bar{e}_6)
\end{array}\right\} = \Omega$$

where: $\bigcup$ – means ,,or" (the sum or union of events);
and $\bigcap$ – means ,,and" (the product or intersection of events)

definition of probability (of an event A)

1. Laplace (beg. of 19th C.):

$$\mathcal{P}(A) = \frac{n(A)}{N(total)}.$$

2. von Mises (end of 19th C.):

$$\mathcal{P}(A) = \lim_{n\to\infty} \frac{k_n(A)}{n} \qquad \text{where} \quad k_n(A)$$

is the number of A events in n experiments (or frequency).

3. Kolmogorov (beg. of 20th C.) – (3 axioms):
   - $\mathcal{P}(\text{ of an event }) \in [0, 1]$
   - $\mathcal{P}(\Omega) = 1.$ – AN event from the event space must occur
   - for *mutually exclusive events* A_i; $i = 1, \ldots, n$

$$\mathcal{P}(A_1 \bigcup A_2 \ldots \bigcup A_n) = \sum_{i=1}^{n} \mathcal{P}(A_i).$$

VENN diagrams

THE COMPLEMENT OF AN EVENT A:

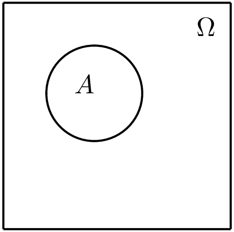
$$\bar{A} = \Omega - A \qquad (a)$$

THE UNION OF (3) EVENTS :

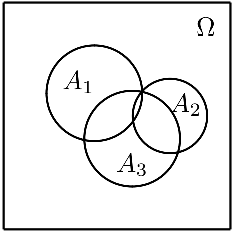
$$A = A_1 \bigcup A_2 \bigcup A_3 \ldots = A_1 + A_2 + A_3 \ldots = \sum_i A_i \quad (b)$$

THE INTERSECTION OF (3) EVENTS :

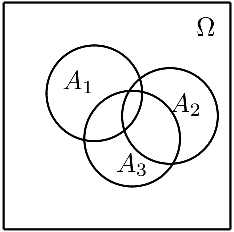
$$A = A_1 \bigcap A_2 \bigcap A_3 \ldots = A_1 A_2 A_3 \ldots = \prod_i A_i \quad (c)$$



(a)



(b)

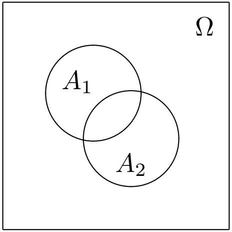


(c)

VENN diagrams, cntd.

THE DIFFERENCE OF (2) EVENTS

$$A_1 - A_2 \quad (d)$$



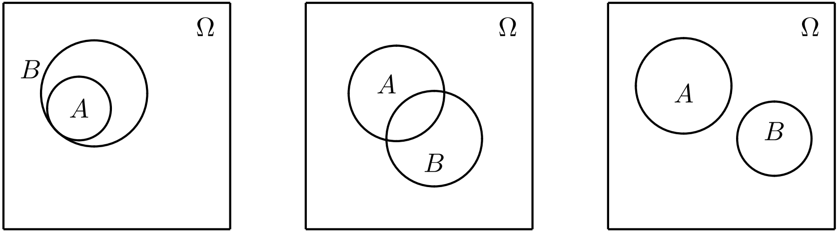
(d)

CONDITIONAL PROBABILITY

$\mathcal{P}(A|B)$ — PROBABILITY OF (an event) A *given that* B occurs
Note: $AB \equiv A \bigcap B$

(1) $$\mathcal{P}(AB) = \mathcal{P}(B)\mathcal{P}(A|B) = \mathcal{P}(A)\mathcal{P}(B|A)$$

(2) $$\mathcal{P}(A|B) = \frac{\mathcal{P}(AB)}{\mathcal{P}(B)} \qquad \mathcal{P}(B|A) = \frac{\mathcal{P}(AB)}{\mathcal{P}(A)}$$



$\mathcal{P}(A|B) = 1$ $\quad 0 < \mathcal{P}(A|B) < 1$ $\quad \mathcal{P}(A|B) = 0$

$A,\ B$:

$$(3) \qquad \mathcal{P}(A|B) = \mathcal{P}(A) \qquad \mathcal{P}(B|A) = \mathcal{P}(B)$$

$$\boxed{\mathcal{P}(AB) \stackrel{(1)}{=} \mathcal{P}(B)\mathcal{P}(A|B) \stackrel{(3)}{=} \mathcal{P}(B)\mathcal{P}(A)}$$

The above formula may be regarded as the fundamental definition of **independent events**

$\mathrm{P}(\text{UNION A} \bigcup \text{B}) \stackrel{?}{=}$

$$\begin{aligned} \mathcal{P}(A \bigcup B) &= \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(AB) \\ \ldots &\quad \ldots \\ \mathcal{P}\left(\sum_{k=1}^{n} A_k\right) &= \sum_{k=1}^{n} \mathcal{P}(A_k) - \sum_{k_1 < k_2} \mathcal{P}(A_{k_1} A_{k_2}) \\ &+ \sum_{k_1 < k_2 < k_3} \mathcal{P}(A_{k_1} A_{k_2} A_{k_3}) + \ldots + (-1)^n \mathcal{P}(A_1 A_2 \ldots A_n) \end{aligned}$$

Let the events $A_1, A_2, \ldots, A_n$ be *a partition* of Ω (sample space) and let B denote an event. We have

$$\mathcal{P}(B) = \mathcal{P}(A_1)\mathcal{P}(B|A_1) + \mathcal{P}(A_2)\mathcal{P}(B|A_2) + \ldots = \sum_{k=1}^{n} \mathcal{P}(A_k)\mathcal{P}(B|A_k)$$

Example: Suppose: 60% of students pass successfully the written exam; 95% (of those who passed the written – to be allowed to enter the oral exam a student *must* obtain a positive grade from the written part) – oral one. What is the probability of a fully successful exam?

Let: E – fully successful exam; $\bar{E}$ – fail; similarly W – successful written exam; $\bar{W}$ – fail (written); O – successful oral exam; $\bar{O}$ – fail (oral);

$\mathcal{P}(\bar{E}) = \mathcal{P}(\bar{E}|W) \cdot \mathcal{P}(W) + \mathcal{P}(\bar{E}|\bar{W}) \cdot \mathcal{P}(\bar{W}) \qquad \mathcal{P}(E) = 1 - \mathcal{P}(\bar{E})$

$\mathcal{P}(\bar{E}) = 0.05 \cdot 0.6 + 1.0 \cdot 0.4 = 0.43 \quad \mathcal{P}(E) = 0.57 = \mathcal{P}(W) \cdot \mathcal{P}(O).$

CONDITIONAL PROBABILITY AND BAYES' THEOREM:

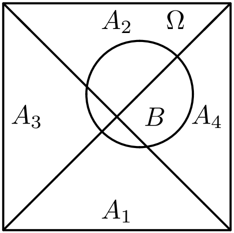
LET US CONSIDER n EVENTS A_i THAT:

1° are mutually exclusive, i.e. $\mathcal{P}(A_l A_m) = 0$ for $l \neq m; l, m = 1, \ldots n$

2° constitute a complete partition of sample space Ω, i.e. $\Omega = \sum_{k=1}^{n} A_k$

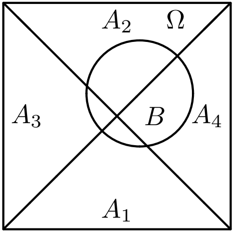
$$B = A_1 B + A_2 B + \ldots + A_n B$$

$$\mathcal{P}(B) = \mathcal{P}(A_1)\mathcal{P}(B|A_1) + \mathcal{P}(A_2)\mathcal{P}(B|A_2) + \ldots = \sum_{k=1}^{n} \mathcal{P}(A_k)\mathcal{P}(B|A_k)$$



$$\mathcal{P}(B) = \sum \mathcal{P}(A_k)\mathcal{P}(B|A_k)$$

BAYES' RULE, CNTD.



$$\mathcal{P}(B) = \sum \mathcal{P}(A_k)\mathcal{P}(B|A_k)$$

$$\mathcal{P}(A_i B) = \mathcal{P}(B)\mathcal{P}(A_i|B) = \mathcal{P}(A_i)\mathcal{P}(B|A_i)$$

(3) $$\mathcal{P}(A_i|B) = \frac{\mathcal{P}(A_i)\mathcal{P}(B|A_i)}{\mathcal{P}(B)} = \frac{\mathcal{P}(A_i)\mathcal{P}(B|A_i)}{\sum_{k=1}^{n}\mathcal{P}(A_k)\mathcal{P}(B|A_k)}$$

$\mathcal{P}(A_i|B)$ — is called the *a posteriori* probability

BAYES' RULE, A PRACTICAL(!) EXAMPLE.

The probability of a disease is one in thousand persons. A routine screening test is positive in 100% of "true" cases and gives an erroneous positive result in 5% of healthy persons. A randomly chosen person is tested and the result is positive What is the probability that the person is really sick?
Denote: S — sick; $\bar{S}$ — healthy; we have

$$\mathcal{P}(+|S) = 1.0 \qquad \mathcal{P}(+|\bar{S}) = 0.05 \qquad \mathcal{P}(-|S) = 0.0 \qquad \mathcal{P}(-|\bar{S}) = 0.95$$

also:

$$\mathcal{P}(+) = \mathcal{P}(S) \cdot \mathcal{P}(+|S) + \mathcal{P}(\bar{S}) \cdot P(+|\bar{S}) = 0.001 \cdot 1 + 0.999 \cdot 0.05 \approx 0.051$$

$$\mathcal{P}(+S) = \mathcal{P}(S) \cdot \mathcal{P}(+|S) = \mathcal{P}(+) \cdot \mathcal{P}(S|+)$$

hence

$$\mathcal{P}(S|+) = \frac{\mathcal{P}(S) \cdot \mathcal{P}(+|S)}{\mathcal{P}(+)} \approx 0.02$$

Question: is it a good screening test?

Adendum – some algebraic beasts

Permutation of n objects taken k at time:

$$P_{n,k} = n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}; \quad P_{n,n} = n!$$

Example: (Feller's problem) Suppose we have 23 persons in the soccer field. What is the probability $\mathcal{P}(A)$ that at least two persons have the same birthday?

$$\Omega = 365^{23}; \quad \mathcal{P}(\bar{A}) = \frac{P_{365,23}}{365^{23}} \approx 0,493 \quad \mathcal{P}(A) = 0.507$$

Combination is the number of distinct subsets of size k taken from n distinct objects (the order within the subset has no importance):

$$C_{n,k} = \frac{P_{n,k}}{k!} \equiv \binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Example: what is the chance of having 'six' out of 49 numbers in the Lotto lottery? — the number of various outcomes is

$$C_{49,6} = \frac{49!}{(43!)!6!} \approx 14 \text{ million}$$