CHAPTER 3

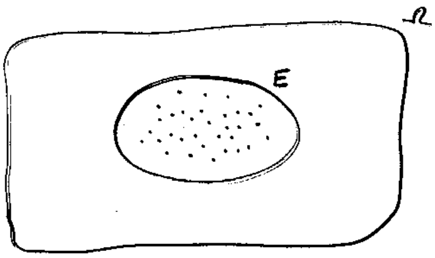
3. CONDITIONAL PROBABILITY

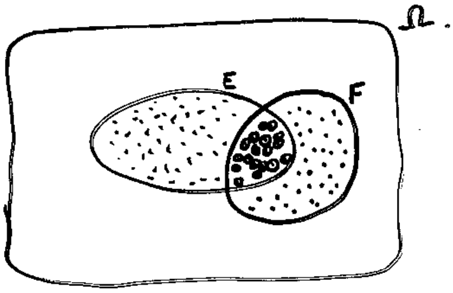
DEFINITION

For two events $E, F \subseteq \Omega$ with $P(F) > 0$, the conditional probability of E given F is denoted $P(E|F)$, and is defined by

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

Visual aid ?





⊙ Points in $E \cap F$

Interpretation :

$P(E|F)$ is the probability that E occurs given the "information" that F occurs

i.e the probability that $\omega_A \in E$ given that $\omega_A \in F$

(recall : ω_A is the actual outcome)

[so, if necessary could write

$$P(E) \equiv P(E|\Omega)$$]

"Classical Probability" interpretation

Ω - n_Ω equally likely outcome

E - n_E

F - n_F $\qquad n_F = n_{E\cap F} + n_{E'\cap F}$

$E \cap F$ - $n_{E\cap F}$

Then prob. $\omega_A \in E$ given that $\omega_A \in F$ is

$$\frac{n_{E\cap F}}{n_F} = \frac{n_{E\cap F} / n_\Omega}{n_F / n_\Omega} = \frac{P(E\cap F)}{P(F)}$$

[the proportion of points in F that are also in E].

- also follows naturally in the FREQUENTIST / SUBJECTIVIST frameworks.

NOTES

(i) $P(\cdot\,|\,\cdot)$ is a set function that takes two arguments

(ii) $P(\cdot\,|\,\cdot)$ obeys the three axioms of probability.

(I) $P(E|F) = \dfrac{P(E \cap F)}{P(F)}$

But $E \cap F \subseteq F \Rightarrow P(E \cap F) \leq P(F)$

$\Rightarrow P(E|F) \leq 1$

$P(E \cap F) \geq 0 \Rightarrow P(E|F) \geq 0.$

(II) $P(\Omega|F) = \dfrac{P(\Omega \cap F)}{P(F)} = \dfrac{P(F)}{P(F)}$

$= 1.$

(III) $E_1, E_2 \subseteq \Omega$ with $E_1 \cap E_2 = \emptyset$

$$P(E_1 \cup E_2 \,|\, F) = \frac{P((E_1 \cup E_2) \cap F)}{P(F)}$$

But

$$(E_1 \cup E_2) \cap F = (E_1 \cap F) \cup (E_2 \cap F)$$

and $(E_1 \cap F)$ and $(E_2 \cap F)$ are mutually exclusive

$$\overset{\text{III}}{\Rightarrow} P((E_1 \cup E_2) \cap F) = P(E_1 \cap F) + P(E_2 \cap F)$$

$$\Rightarrow P(E_1 \cup E_2 \,|\, F) = \frac{P(E_1 \cap F) + P(E_2 \cap F)}{P(F)}$$

$$= P(E_1|F) + P(E_2|F)$$

SIMPLE EXAMPLES

Experiment : Roll of single fair die

Σ - all scores equally likely

$\Omega = \{1,2,3,4,5,6\}$

Let $A \equiv$ "score is odd"

$B \equiv$ "score is >3"

Then $P(A) = \frac{3}{6}$ $P(B) = \frac{3}{6}$ $P(A \cap B) = \frac{1}{6}$

$\Rightarrow P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{1}{3}$

so $P(A|B) < P(A)$

EXPERIMENT Measure failure time of component in electronic system

Suppose $A \equiv$ "fails later than 10 hours"

$B \equiv$ "fails later than 5 hours"

rd $P(A) = 0.25 \qquad P(B) = 0.4$

Here $A \cap B \equiv A$

So $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}$$

$$= \frac{0.25}{0.4} = 0.625$$

EXAMPLE A family comprises two children

$\mathcal{E}$ - "Female/male equally likely, successive births identical and mutually unaffected"

$$\Rightarrow \Omega \equiv \{FF, FM, MF, MM\}$$

FF - first child female, second child female.

$$P(\{\omega\}) = \frac{1}{4} \qquad \omega \in \Omega.$$

Two questions

(a) If one child is a boy, what is the probability that the other is also a boy?

(b) If the eldest child is a boy, what is the probability that the other is also a boy?

Let $A \equiv$ "both male" $\quad A \equiv \{MM\}$

$B \equiv$ "at least one male" $\quad B \equiv \{FM, MF, MM\}$

$C \equiv$ "eldest is male" $\quad C \equiv \{MF, MM\}$

(a) Want

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}$$

(b) Want

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{1/4}{2/4} = \frac{1}{2}.$$

Not intuitive?

EXAMPLE THREE CARDS

CARD 1 RED / RED

CARD 2 RED / BLACK

CARD 3 BLACK / BLACK.

A card is selected and one side is exposed; the exposed side is RED

What is the probability that the other side is RED?

Let

$E \equiv$ " RED/RED card selected "

$F \equiv$ " RED side exposed "

$\Rightarrow E \equiv \{R_1, R_2\} \Rightarrow P(E) = \frac{2}{6}$

$F \equiv \{R_1, R_2, R\} \Rightarrow P(F) = \frac{3}{6}$

$E \cap F \equiv \{R_1, R_2\} \Rightarrow P(E \cap F) = \frac{2}{6}$

$\Rightarrow P(E|F) = \frac{P(E\cap F)}{P(F)} = \frac{2/6}{3/6} = \frac{2}{3}$

$\therefore$ IMPORTANT TO LIST CORRECTLY THE EQUALLY LIKELY OUTCOMES.

$\mathcal{E}$ - "each card equally likely to be selected, each side equally likely to be exposed."

Let Ω comprise the possible exposed card sides

i.e. $\Omega \equiv \{\underbrace{R_1, R_2}_{R/R}, \underbrace{R, B}_{R/B}, \underbrace{B_1, B_2}_{B/B}\}$

$P(\{\omega\}) = \frac{1}{6} \qquad \omega \in \Omega.$

EXAMPLE Mass disease screening

100,000 people tested

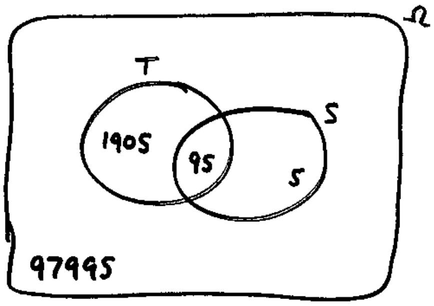
100 actual disease sufferers

Test Results

2000 +ve test results

95 sufferers test +ve.

(+ve test indicates disease)



T - "test is positive"

S - "individual is a sufferer"

so $n_{T \cap S} = 95$ etc.

We have

$$P(S) = \frac{100}{100\,000} = 0 \cdot 001$$

$$P(T \cap S') = \frac{1905}{100\,000} = 0 \cdot 01905$$

$$P(T \cap S) = \frac{95}{100\,000} = 0 \cdot 00095$$

$$P(T' \cap S) = \frac{5}{100\,000} = 0 \cdot 00005$$

So if M = "test gives <u>incorrect</u> indication"

then

$$M = (T \cap S') \cup (T' \cap S)$$

"false positives" "false negatives", or

$$\therefore P(M) = P(T \cap S') + P(T' \cap S)$$

$$= \frac{1910}{100000}$$

To assess the test itself, could look at

$$P(T \mid S') = \frac{P(T \cap S')}{P(S')} = \frac{1905}{99900}$$

$$P(T' \mid S) = \frac{P(T' \cap S)}{P(S)} = \frac{5}{100}$$

[so $P(T \mid S) = \frac{95}{100}$ - QUITE HIGH?

i.e.

$$P(T|S) = 0.95$$

$$P(S|T) = 0.0475$$

- recall that, in practice, the disease status of an individual being tested will not be observed; we merely observe the test result.

But for diagnosis purposes,

(do not observe disease status)

might look at

$$P(S|T) = \frac{P(S \cap T)}{P(T)} = \frac{95}{2000} = 0.0475$$

$$P(S|T') = \frac{P(S \cap T')}{P(T')} = \frac{5}{98000}$$

SUMMARY

So, in general, for two events E, F

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad (\text{if } P(F) > 0)$$

NOTE: In general

$$P(E|F) \neq P(E)$$
$$P(E|F) \neq P(F|E)$$

and $P(E \cap F) = P(E|F)P(F)$

DEFINITION

Events E and F are independent if

$$P(E|F) = P(E)$$

("pairwise independent")

or equivalently

$$P(F|E) = P(F)$$
$$P(E \cap F) = P(E)P(F)$$

[Also: straightforward to show

E, F independent $\iff E, F'$ independent]

Interpretation

- information about the occurrence of F (or F') does not affect our assessment of the probability that E (or E') occurs.

i.e

$$P(E_1 | E_2) = P(E_1 | E_3)$$

$$= P(E_2 | E_3) \quad \text{etc}$$

$$= \frac{1}{2}$$

But

$$P(E_1 | E_2 \cap E_3) = 0$$

$\therefore$ E_1, E_2, E_3 not independent mutually.

EXAMPLE ROLL TWO FAIR DICE

$\mathcal{E}$ - "all 36 pairs of scores equally likely"

E_1 - "1^{st} score odd"

E_2 - "2^{nd} score odd"

E_3 - "total score odd"

E_1, E_2 pairwise independent
E_1, E_3 pairwise independent
E_2, E_3 pairwise independent

EXTENSION

Events $E_1, E_2, \ldots, E_k$ are mutually independent if

$$P\left(\bigcap_{i \in Z_k} E_i\right) = \prod_{i \in Z_k} P(E_i)$$

for all subsets Z_k of $\{1, 2, \ldots k\}$

e.g. $Z_k = \{1,2\}, \{1,3\}, \ldots$
$\{1,2,3\}, \ldots$
$\{1,2,3,4\} \ldots$
etc.

If $k=3$; require

$$P(E_1 \cap E_2) = P(E_1)P(E_2)$$
$$P(E_1 \cap E_3) = P(E_1)P(E_3)$$
$$P(E_2 \cap E_3) = P(E_2)P(E_3)$$

and

$$P(E_1 \cap E_2 \cap E_3) = P(E_1)P(E_2)P(E_3)$$

all to hold.

NOTE

PAIRWISE INDEPENDENCE $\not\Rightarrow$ MUTUAL INDEPENDENCE

We will commonly assume that sequences of repeats of experiments lead to outcomes that are mutually independent and have identical probabilistic specifications.

– coins, dice etc.

DEFINITION

For events E_1, E_2 and F with $P(F)>0$, E_1 and E_2 are conditionally independent given F if

$$P(E_1 \mid E_2 \cap F) = P(E_1 \mid F)$$

or equivalently

$$P(E_1 \cap E_2 \mid F) = P(E_1 \mid F)P(E_2 \mid F).$$

THE GENERAL MULTIPLICATION RULE

("CHAIN RULE")

From the Conditional Probability definition, we have for E, F with $P(F)>0$.

$$P(E \cap F) = P(E \mid F)P(F)$$

For events $E_1, E_2, \dots, E_n$ we have

$$P(E_1 \cap E_2 \cap \dots \cap E_n)$$
$$= P(E_1)P(E_2 \cap E_3 \cap \dots \cap E_n \mid E_1)$$
$$= P(E_1)P(E_2 \mid E_1)P(E_3 \cap E_4 \cap \dots \cap E_n \mid E_1 \cap E_2)$$

etc.

$$\Rightarrow P(E_1 \cap E_2 \cap \dots \cap E_n)$$
$$= P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2) \times \dots \times P(E_n|E_1 \cap E_2 \cap \dots \cap E_{n-1})$$

product of n terms

This is the General Multiplication rule.

or Chain Rule

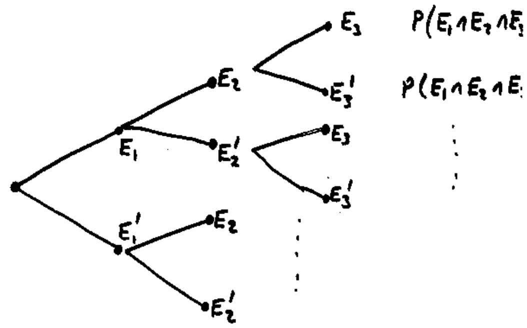
SPECIAL CASE: $E_1, E_2, \dots, E_n$ MUTUALLY INDEPENDENT

$$P\left(\bigcap_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} P(E_i)$$

"SEQUENTIAL PROBABILITY"

Visual Representation?

- PROBABILITY TREE



eg BRANCH 1 BRANCH 2 BRANCH 3

$P(E_1)$ $P(E_2|E_1)$ $P(E_3|E_1 \cap E_2)$

→ MULTIPLY ALONG BRANCHES TO GET

ie $P(E_1 \cap E_2 \cap E_3)$

EXAMPLE A bag contains 10 balls

6 WHITE
4 RED

4 balls selected "without replacement"

$\mathcal{E}$ - "on each selection, all balls equally likely to be chosen from those remaining"

Let $A \equiv$ "all selected are WHITE"

$E_i \equiv$ "WHITE selected on i^{th} draw" $i = 1, \dots, 4$

so that

$$A \equiv E_1 \cap E_2 \cap E_3 \cap E_4$$

Then

$$P(A) = P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2)P(E_4|E_1 \cap E_2 \cap$$

$$= \frac{6}{10} \times \frac{5}{9} \times \frac{4}{8} \times \frac{3}{7}$$

If $\mathcal{E}$ - "balls selected with replacement"

Then

$$P(E_2|E_1) \equiv P(E_2) \equiv P(E_1)$$
$$P(E_3|E_1 \cap E_2) \equiv P(E_3) \equiv P(E_1)$$

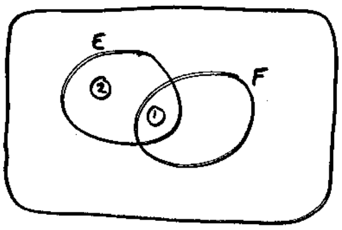
... etc

So $P(A) = \left(\frac{6}{10}\right)^4$

3.1 THE THEOREM OF TOTAL PROBABILITY

Recall the earlier expression for $E, F \subseteq \Omega$:

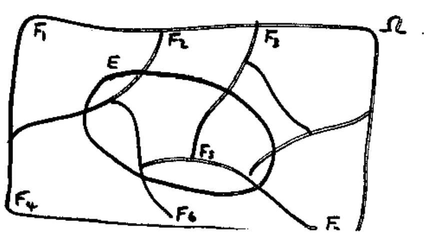
$$E = (E \cap F) \cup (E \cap F')$$
$$\quad ① \qquad\quad ②$$



This result uses the fact that F and F' form a partition of Ω.

For a more general result: consider events $F_1, F_2, \dots, F_n$ that form a partition of Ω. Then

$$E \equiv \bigcup_{i=1}^{n} (E \cap F_i)$$



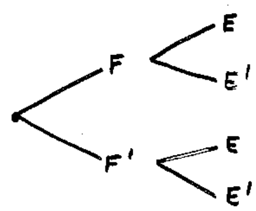
Axiom III

$$\Rightarrow P(E) = P(E \cap F) + P(E \cap F')$$

Conditional Prob. Def^N

$$\Rightarrow P(E) = P(E|F)P(F) + P(E|F')P(F')$$

i.e.



① $E \cap F$

② $E \cap F'$

"to end up in E, choose either F or F' route, then choose E."

Assuming that $P(F_i) > 0 \ \forall i$, we have

$$P(E) = P\left(\bigcup_{i=1}^{n} (E \cap F_i)\right)$$
$$= \sum_{i=1}^{n} P(E \cap F_i) \qquad \text{Ax III}$$
$$= \sum_{i=1}^{n} P(E|F_i)P(F_i) \qquad \text{C.P.}$$

- consider a tree diagram with n initial branches (corresponding to $F_1, \dots, F_n$), with each branch having two second stage branches (E, E')
- sum the probs. for branches that end at E

This result is the THEOREM OF TOTAL PROBABILITY

Key elements :

(i) partition of Ω $F_1, F_2, .., F_n$

(ii) formula

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

(iii) easy form for $P(E|F_i)$?

[extension to n infinite is also valid]

Clearly we have

$$A \equiv (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3)$$

$\therefore$ theorem of total prob. gives

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) && = \tfrac{4}{8} \times \tfrac{1}{3} \\ &+ P(A|B_2)P(B_2) && + \tfrac{1}{11} \times \tfrac{1}{3} \\ &+ P(A|B_3)P(B_3) && + \tfrac{7}{18} \times \tfrac{1}{3} \end{aligned}$$

$$= \frac{97}{297}$$

[NOTE: $\neq \dfrac{4+1+7}{8+11+18} = \dfrac{12}{37}$

which we may have guessed ?]

EXAMPLE Three bags containing

BAG 1	4 RED ,	4 WHITE	balls
BAG 2	1 RED ,	10 WHITE	
BAG 3	7 RED ,	11 WHITE	

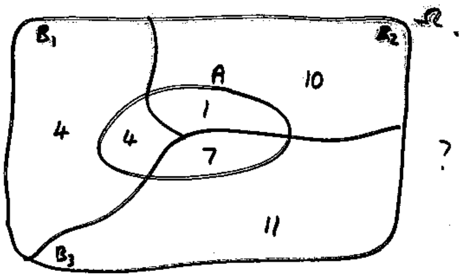
A bag is selected, and a single ball drawn.

$\mathcal{E}$ - "All bags / balls in a bag equally likely to be selected"

Let A - "ball selected is RED"

B_i - "bag i is selected"

$\Omega \equiv$ "all possible selections of single balls".



Why can we not use this diagram to say $P(A) = \dfrac{12}{37}$?

– because the 37 balls are not equally likely to be selected.

i.e $\mathcal{E}$ is not defined in this way

EXAMPLE Games between two players

– sequence of repeated games; for each game

Player A – wins with prob p.

Player B – wins with prob q

– Game is drawn with prob. r.

(Axioms $\Rightarrow r = 1-p-q$)

Match ends on the first occasion that a player wins a game.

What is the prob. that player A wins the match?

$\mathcal{E}$ – "games identical; results are mutually independent"

Let $A \equiv$ "A wins match"

$A_n \equiv$ "A wins match on n^{th} game"

$H_k \equiv$ "k^{th} game is drawn"

$W_k \equiv$ "A wins game k"

Then $A = \bigcup_{n=1}^{\infty} A_n$ (PARTITION)

and Axiom III $\Rightarrow P(A) = \sum_{n=1}^{\infty} P(A_n)$

But

$$P(A_n) = P(H_1 \cap H_2 \cap \ldots \cap H_{n-1} \cap W_n)$$

$$= \left\{ \prod_{k=1}^{n-1} P(H_k) \right\} P(W_n)$$

CHAIN RULE + INDEPENDENCE.

$$= r^{n-1} p$$

as $P(H_k) = r \quad \forall k$

$P(W_k) = p$

$$\Rightarrow P(A) = \sum_{n=1}^{\infty} r^{n-1} p.$$

$$= \frac{p}{1-r} \qquad \text{(sum of g.p.)}$$

$$= \frac{p}{p+q}. \qquad r = 1-p-q.$$

Alternative solution:

Partition A according to result of first game.

But AXIOM III / THM OF TOTAL PROB

$$\Rightarrow P(T) = P(T \cap S) + P(T \cap S')$$

$$= P(T|S)P(S) + P(T|S')P(S')$$

Now "GOOD TEST" suggests

$P(T|S) \approx 1$ say $P(T|S) = 1-\alpha$

$P(T|S') \approx 0$ $P(T|S') = \beta$

– may be able to quantify

α, β

via lab testing?

So if the proportion of sufferers in the population is p

i.e. $P(S) = p$

then

$$P(T) = (1-\alpha)p + \beta(1-p)$$

May be able to identify p numerically from population studies.

In practice, we do not observe whether a tested person is a sufferer or not; we merely observe T or T'.

$\therefore$ Can quantify $P(T|S)$, $P(T|S')$

but may really be interested in $P(S|T)$ or $P(S|T')$

BEFORE TEST "PRIOR"		AFTER TEST "POSTERIOR"
$P(S)$	$\longrightarrow$	$P(S\|T)$ or $P(S\|T')$

PROBABILITY OF BEING A SUFFERER BEFORE / AFTER TEST IS CARRIED OUT.

The mechanism for modifying

$P(S)$ to $P(S|T)$

is given to us by the conditional prob. definition

i.e.

$$P(S|T) = \frac{P(T \cap S)}{P(T)} = \frac{P(T|S)P(S)}{P(T)}$$

i.e. $$P(S|T) = \left[\frac{P(T|S)}{P(T)}\right] P(S)$$

i.e. we must multiply by

$$\frac{P(T|S)}{P(T)}.$$

3.2 Bayes Theorem

Theorem

For events $E, F \subseteq \Omega$ with $P(E), P(F) > 0$,

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

Proof

By the conditional probability def^n

$P(E \cap F) = P(E|F)P(F)$

and $P(E \cap F) = P(F|E)P(E)$

and the result follows by equating these two equations.

Proof

$$E \equiv \bigcup_{k=1}^{n} (E \cap F_k)$$

$$\Rightarrow P(E) = \sum_{k=1}^{n} P(E \cap F_k)$$

$$= \sum_{k=1}^{n} P(E|F_k)P(F_k)$$

But $$P(F_i|E) = \frac{P(E|F_i)P(F_i)}{P(E)}$$

hence result.

Extension: $F_1, F_2, \ldots$ (countable no. of events partition Ω)

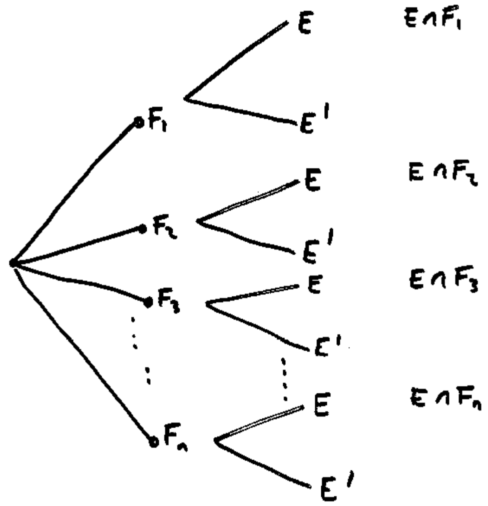
General Version

If $F_1, \ldots, F_n$ partition Ω (wlog assume $P(F_i) > 0 \ \forall i$), then for $E \subseteq \Omega$ with $P(E) > 0$

$$P(F_i|E) = \frac{P(E|F_i)P(F_i)}{P(E)}$$

$$= \frac{P(E|F_i)P(F_i)}{\sum_{k=1}^{n} P(E|F_k)P(F_k)}$$

Interpretation ? Prob. Tree.



STAGE 1

n branches $F_1, F_2, \dots, F_n$

STAGE 2

2 sub-branches E, E'

$\therefore$ $2n$ end points, n of which end within E.

BAYES TH^m : calculates the conditional probability that "we took branch F_i, given that we ended up in E."

EXAMPLE Recall

BAG 1	4 RED / 4 WHITE
BAG 2	1 RED / 10 WHITE
BAG 3	7 RED / 11 WHITE

A - "selected ball is RED"

B_i - "ball came from bag i"

If the ball is RED, what is the probability that it came from BAG 2?

Given that

$$P(A|B_1) = \frac{4}{8} \quad P(A|B_2) = \frac{1}{11} \quad P(A|B_3) = \frac{7}{18}$$

BAYES TH^m B_1, B_2, B_3 partition Ω.

$$\therefore P(B_2|A) = \frac{P(A|B_2)P(B_2)}{P(A)}$$

$$= \frac{P(A|B_2)P(B_2)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)}$$

$$= \frac{\frac{1}{11} \times \frac{1}{3}}{\left(\frac{4}{8} \times \frac{1}{3}\right) + \left(\frac{1}{11} \times \frac{1}{3}\right) + \left(\frac{7}{18} \times \frac{1}{3}\right)}$$

$$= \frac{9}{97}$$

Bayes Theorem is deduced immediately from the Conditional Probability definition

- it is a formula for relating conditional probs.

- allows us to compute

P ["UNOBSERVED EVENT" | "OBSERVED"]

from

P ["OBSERVED" | "UNOBSERVED"]

– "REVERSE" or "INVERSE" PROB.

i.e.

$$A \equiv (A \cap W_1) \cup (A \cap H_1) \cup (A \cap (W_1' \cap H_1'))$$

THEOREM OF TOTAL PROB:

$\Rightarrow$

$$P(A) = P(A|W_1)P(W_1) + P(A|H_1)P(H_1) + P(A|W_1' \cap H_1')\,P(W_1' \cap H_1')$$

But

$$P(A|W_1) = 1$$

$$P(A|H_1) = P(A) \quad \text{INDEPENDENCE.}$$

$$P(A|W_1' \cap H_1') = 0$$

$$\Rightarrow P(A) = 1.p + P(A).r + 0.q$$

$$\Rightarrow P(A) = \frac{p}{1-r} = \frac{p}{p+q}.$$

as before.

EXAMPLE Medical Screening

Diagnostic test for disease

- not 100% accurate

- occasionally, we observe that

NON-SUFFERERS TEST POSITIVE

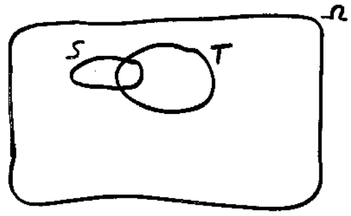
SUFFERERS TEST NEGATIVE

Let Ω - "the population of individuals who may be tested."

Let

$S \equiv$ "individual is a sufferer"

$T \equiv$ "individual tests positive"



S and S' partition Ω.

$$\therefore T \equiv \underbrace{(T \cap S)}_{\text{"CORRECT +ve RESULTS"}} \cup \underbrace{(T \cap S')}_{\text{"INCORRECT +ve RESULTS"}}$$

"FALSE POSITIVES"

NOTE For events E, F

In general

$$P(E|F) \neq P(F|E)$$

eg
$$P[\text{"MEASLES"} \mid \text{"SPOTS"}] \neq P[\text{"SPOTS"} \mid \text{"MEASLES"}]$$

e.g. $$P[\text{"GUILT"} \mid \text{"EVIDENCE"}] \neq P[\text{"EVIDENCE"} \mid \text{"GUILT"}]$$

"THE PROSECUTOR'S FALLACY".

EXAMPLE Recall MEDICAL SCREENING.

S - "sufferer"
T - "test is positive"

Given

$$P(T|S) = 1-\alpha \qquad \alpha, \beta \to 0 \,?$$
$$P(T|S') = \beta$$
$$P(S) = p.$$

THEOREM OF TOTAL PROB.

$$P(T) = P(T|S)P(S) + P(T|S')P(S') = (1-\alpha)p + \beta(1-p)$$

Interested in $P(S|T)$, $P(S|T')$

BAYES TH^m

$$\Rightarrow P(S|T) = \frac{P(T|S)P(S)}{P(T)} = \frac{(1-\alpha)p}{(1-\alpha)p + \beta(1-p)}$$

$$P(S|T') = \frac{P(T'|S)P(S)}{P(T'|S)P(S) + P(T'|S')P(S')} = \frac{\alpha p}{\alpha p + (1-\beta)(1-p)}$$

Numerically suppose

$\alpha = 0.05$ FALSE NEGATIVES
$\beta = 0.10$ FALSE POSITIVES
$p = 0.1$

$$\Rightarrow P(S|T) = 0.514.$$

If $p = 0.001$

$$P(S|T) = 0.009$$

-very different in magnitude from

$$P(T|S) = 1-\alpha = 0.95.$$

Is this result counter intuitive?

- consider the rate of "FALSE POSITIVE" results. This result is due to the fact that the majority of Positive Test results are FALSE POSITIVES.

$$P(S|T) = \frac{P(T\cap S)}{P(T)} = \frac{n_{T\cap S}}{n_T}.$$

- this ratio is low if $n_{T\cap S}$ is small compared to n_T.

If the suspect is GUILTY

Chance of evidence E_1 is 0.90
Chance of evidence E_2 is 0.99

If the suspect is NOT GUILTY

Chance of evidence E_1 is 0.20
Chance of evidence E_2 is 0.01

i.e.

$$P(E_1|G) = 0.90$$
$$P(E_2|G) = 0.99$$

$$P(E_1|G') = 0.20$$
$$P(E_2|G') = 0.01$$

EXAMPLE

Two pieces of evidence are used in a court case to attempt to assess the probability of a suspect's guilt.

E_1 - "eye-witness"

E_2 - "finger print match".

Suppose that the chance that a piece of evidence is misleading is assessed as follows:

Want

$$P(G|E_1\cap E_2) = \frac{P(E_1\cap E_2|G)P(G)}{P(E_1\cap E_2)} \quad (1)$$

where

$$P(E_1\cap E_2) = P(E_1\cap E_2|G)P(G) + P(E_1\cap E_2|G')P(G') \quad (2)$$

$$= P(E_1|G)P(E_2|G)P(G) + P(E_1|G')P(E_2|G')P(G') \quad (3)$$

(1) - by BAYES THEOREM

(2) - by THM OF TOTAL PROB.

(3) - by CONDITIONAL INDEPENDENCE

From (1) and (3)

$$P(G|E_1 \cap E_2) = \frac{P(E_1|G)P(E_2|G)P(G)}{P(E_1 \cap E_2)}$$

$$= \frac{0.90 \times 0.99 \times p}{(0.90 \times 0.99 \times p) + (0.2 \times 0.01 \times (1-p))}$$

$P(G)$	$P(G\|E_1 \cap E_2)$
0.01	0.818
0.05	0.959
0.1	0.980
0.2	0.991

- depends on $p = P(G)$

THE TIMES

Evidence of theorem recipe for confusion

" While there could be no possible objection to the prosecution presenting DNA evidence based as it was on statistical data, reliance on evidence of the Bayes Theorem in relation to non-scientific evidence was a recipe for confusion, misunderstanding and misjudgment. Accordingly, in such cases, in the absence of special features, Bayesian evidence should not be admitted. "

" . . . . a recipe for confusion, misunderstanding and misjudgment, possibly among counsel, probably among judges and almost certainly among jurors. "

COURT OF APPEAL
16/10/97

DEFINITION

The ODDS on event E are given by

$$\frac{P(E)}{P(E')} = \frac{P(E)}{1-P(E)}$$

In the medical screening example, the odds on being a sufferer are

$$\frac{P(S)}{1-P(S)} \quad \text{"PRIOR ODDS"}$$

before the test is carried out.

After a positive test result, the odds on being a sufferer have become

$$\frac{P(S|T)}{1-P(S|T)} = \frac{P(T|S)}{P(T|S')} \frac{P(S)}{P(S')}$$

"POSTERIOR ODDS"

i.e. odds have been modified by a factor

$$\frac{P(T|S)}{P(T|S')} = \frac{1-\alpha}{\beta}$$

i.e. $\dfrac{\text{POSTERIOR ODDS}}{\text{PRIOR ODDS}} = \dfrac{P(T|S)}{P(T|S')}$

"ODDS RATIO"