PTS – Chapter 1

1. Introduction to Probability Theory

P.T. provides
- concepts to speak about uncertainty
  (games of chance, recurring events with some pattern of variation, e.g., measurements, balls in a Galton board, etc.)
- methods to quantify uncertainty

Example: Rolling a die D

What is $P(D=6)$? $\frac{1}{6}$ $\quad P(D \geq 4)$? $\frac{3}{6} = \frac{1}{2}$ $\quad P(D<4)$? $\frac{1}{2}$

Meaning:

1) In the long run, 1/6 of the throws results in 6.
2) There is a chance of 1 in 6 that this throw results in 6.

frequentist

subjective (Bayesian)

view

No consequences for mathematical theory

1.1 Events

Experiments:

S sample space (= set of possible outcomes)

S is known, but not a specific outcome

Rolling a die: $S = \{1, 2, 3, \ldots, 6\}$ $\#S = 6$

Rolling two dice: $S = \{(1,1), (1,2), \ldots, (1,6), \ldots \ldots (3,4) \ldots (6,1), \ldots \ldots (6,6)\}$ $\#S = 6 \cdot 6 = 36$

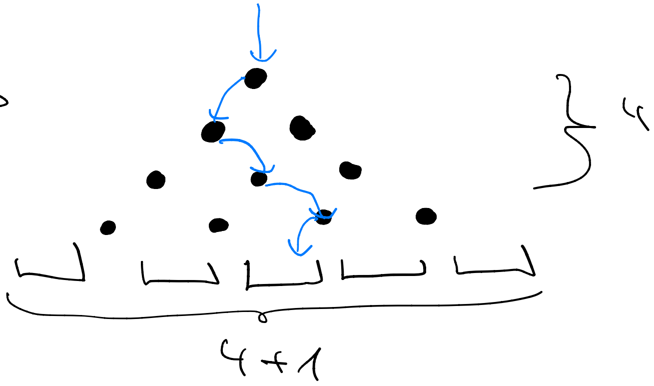
Galton board:

with n levels

has outcomes

$\{1, \ldots, n+1\}$

$\{0, \ldots, n\}$



S can also be infinite!

- measuring distances: any number $m > 0$ is poss. outcome

  — " — $m \in [0.5, 3.00]$

  infinitely many reals in the interval

- How many times do we have to throw a die until a 6 appears

  $\Rightarrow$ consider arbitrarily long sequences of throws, even infinite one

Events:

throw an even number $\mathcal{E}_{even} = \{2,4,6\}$

throw two equal numbers $\mathcal{E}_{equal} = \{(1,1),(2,2),\ldots,(6,6)\}$

In general: $\mathcal{E} \subseteq \mathcal{S}$ ← events are sets

$\mathcal{E}$ occurs if $\underbrace{\text{outcome}}_{}$ $\in \mathcal{E}$

← the outcome of our experiment

in an experiment

Set operations on events

$\mathcal{E} \cup \mathcal{F}$ union ("or")

$\mathcal{E} \cap \mathcal{F}$, $\mathcal{E}\mathcal{F}$ intersection ("and")

$\overline{\mathcal{E}}$ complement ("not"), $\mathcal{S} \setminus \mathcal{E} = \overline{\mathcal{E}}$

$\mathcal{E} \setminus \mathcal{F}$ $\mathcal{E}$ and not $\mathcal{F}$

$= \mathcal{E} \cap \overline{\mathcal{F}}$

Example: $\mathcal{E}_{prime} := \{2, 3.5\}$

$\mathcal{E}_{even} \cup \mathcal{E}_{prime} = \{2, 3, 4, 5, 6\} = \overline{\{1\}}$

$\mathcal{E}_{even} \cap \mathcal{E}_{prime} = \{2\}$

$\overline{\mathcal{E}}_{prime} = \{1, 4, 6\}$

Special case $\emptyset \subseteq \mathcal{S}$, impossible event

Terminology

$\mathcal{E} \cap \mathcal{F} = \emptyset$ disjoint

$\overline{\mathcal{S}} = \emptyset$, $\overline{\emptyset} = \mathcal{S}$

$\overline{\overline{\mathcal{E}}} = \mathcal{E}$

Pronunciation

$\overline{\mathcal{E} \cup \mathcal{F}} = \overline{\mathcal{E}} \cap \overline{\mathcal{F}}$, $\overline{\mathcal{E} \cap \mathcal{F}} = \overline{\mathcal{E}} \cup \overline{\mathcal{F}}$ (De Morgan's Rule)

$\mathcal{E} \equiv \mathcal{F} \iff \mathcal{E} \subseteq \mathcal{F}$ and $\mathcal{F} \subseteq \mathcal{E}$

$\mathcal{E}, \mathcal{F}$ are <u>equivalent</u>

1.2 Axioms of Probability

$P(\mathcal{E})$ probability of $\mathcal{E}$, is a real number

A1: $0 \le P(\mathcal{E}) \le 1$

A2: $P(\mathcal{S}) = 1$

pairwise any two are disjoint

A3: If $\mathcal{E}_1, \mathcal{E}_2, \dots$ are mutually disjoint (i.e., $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset$ for $i \neq j$)

then $P(\mathcal{E}_1 \cup \dots \cup \mathcal{E}_n) = P(\mathcal{E}_1) + \dots + P(\mathcal{E}_n) = \sum_{i=1}^{n} \mathcal{E}_i$, finite

(also holds for infinite series)

What is $P(\emptyset) = 0$ because

$\mathcal{S}, \emptyset$ are disjoint

$$\underline{P(\mathcal{S})} = P(\mathcal{S} \cup \emptyset) = \underline{P(\mathcal{S})} + \underline{P(\emptyset)}$$

$$\Rightarrow 0 = P(\emptyset)$$

$$P(\mathcal{E}_{even} \cup \{1\}) = P(\mathcal{E}_{even}) + P(\{1\})$$

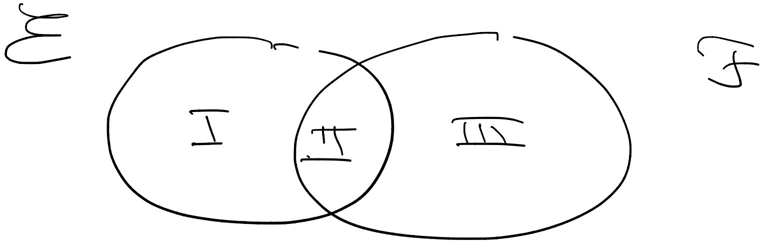
Proposition 1 : $P(\bar{E}) = 1 - P(E)$

Proof : $E \cap \bar{E} = \emptyset$, $E \cup \bar{E} = S$

$$1 \underset{Ax2}{=} P(S) = P(E \cup \bar{E}) \underset{Ax3}{=} P(E) + P(\bar{E})$$

$$\Rightarrow \quad 1 - P(E) = P(\bar{E})$$

What about $P(E \cup F)$ in general?



$I = E \setminus F$

$II = E \cap F$

$III = F \setminus E$

$$P(E \cup F) = P(I \cup II \cup III)$$
$$= P(I) \cup P(II) \cup P(III) \quad (Ax\,3)$$

$$P(E) = P(I) + P(II)$$

$$P(F) = P(II) + P(III)$$

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

Proposition 2 : $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Example 3 :

Drinkers Quiz

B drinks beer, W drinks wine

$$P(B \cup W) = P(B) + P(W) - P(BW)$$

$$= .68 + .49 - .35 = 1.17 - .035 = .82$$

$$P(\overline{B \cup W}) = 1 - .82 = .18$$

Definition 4: The odds of $\mathcal{E}$ is

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

Says how much more likely $\mathcal{E}$ is than $\overline{\mathcal{E}}$

Quiz: Odds of throwing a 4

$$\mathcal{E} = \{4\}, \quad P(\mathcal{E}) = \frac{1}{6}$$

$$\text{odds}(\mathcal{E}) = \frac{P(\mathcal{E})}{1 - P(\mathcal{E})} = \frac{\frac{1}{6}}{\frac{5}{6}} = \frac{1}{5}$$

Trump example: $P(\underbrace{\text{"T will be el."}}_{\mathcal{E}}) = .4$

$$\text{odds}(\mathcal{E}) = \frac{.4}{.6} = \frac{2}{3}$$

1.3 Uniformity

Often: all outcomes are equally likely (only possible if $\mathcal{S}$ finite)
and have prob. > 0

$\boxed{\# \mathcal{S} = n}$, say $\mathcal{S} = \{1, \dots n\}$

$$\Rightarrow P(\{1\}) = P(\{2\}) = \dots = P(\{n\}) = p$$

$$\Rightarrow 1 \underset{A2}{=} P(\mathcal{S}) = P(\{1, \dots, n\}) \underset{A3}{=} P(\{1\}) + \dots + P(\{n\}) = n \cdot p$$

$$\Rightarrow 1 = n \cdot p \Rightarrow p = \frac{1}{n} = P(\{i\}), \quad 1 \leq i \leq n$$

Generalize: $\boxed{\mathcal{E} \subseteq \mathcal{S}} \Rightarrow P(\mathcal{E}) = \boxed{\frac{\# \mathcal{E}}{n}}$

$\# \mathcal{E}$ is the card. of $\mathcal{E}$

Counting Principle

\# outcomes : throw 2 or 3 dice

first throw a die, determines a stack of cards

then pick a card (of 32)

Combinations of experiments, sequential executions

E_1 has m outcomes, $E2$ has n outcomes

\# outcomes of "$E1$ then $E2$" $= m \cdot n$

Outcomes correspond to matrix

$$\begin{matrix} (1,1), & \dots & (1,n) \\ (2,1) & & \\ \vdots & & \\ (m,1) & \dots & (m,n) \end{matrix}$$

In general : E_i has n_i outcomes

$\Rightarrow$ E_1 then E_2 ... then E_r has $n_1 \cdot \dots \cdot n_r$ outcomes

Quiz 3: Socks in a Box

$$P(\underbrace{\text{2 socks have diff. colour}}_{\mathcal{E}}) = \frac{\# \mathcal{E}}{\# \mathcal{S}}$$

8 black, 7 white

$\mathcal{S}$ = Pairs of socks picked, first, then second $\Rightarrow \# \mathcal{S} = 15 \times 14$

$\mathcal{E}$ = Pairs with $(\text{1st} = w \wedge \text{2nd} = b) \cup (\text{1st} = b \wedge \text{2nd} = w)$

$\# \mathcal{E} \qquad 7 \times 8 + 8 \times 7$

$$\frac{\# \mathcal{E}}{\# \mathcal{S}} = \frac{7 \cdot 8 + 8 \cdot 7}{15 \times 14} = \frac{\cancel{2} \cdot \cancel{7} \cdot 8}{15 \cdot \cancel{14}} = \frac{8}{15}$$

Quiz 4: How many Words?

Words = # possibilities 1st bit

× # — " — 2nd — " —

. . .

× # — " — 32nd bit

2
×
2
×
⋮
×
2

$$= 2^{32} \quad // \quad 2^{10} = 1024 \approx 1000 = 10^3$$

$$= 2^{30} \cdot 2^2 = (2^{10})^3 \cdot 2^2$$

$$\approx (10^3)^3 \cdot 2^2 = 4{,}000{,}000{,}000$$

Quiz 5 : People on an Elevator

$$P(\underbrace{\text{"all off at same floor"}}_{\mathcal{E}})$$

$$\mathcal{E}_i = \text{"all off at floor } i\text{"} \qquad (i = 1, \dots, 4)$$

$$\mathcal{S} = \{ (f_1, f_2, f_3, f_4) \mid f_i \in \{1, 2, 3, 4\},\ i = 1, \dots, 4 \}$$

$$= \{ (1,1,1,1), (1,1,1,2), \dots \}$$

$$\boxed{\# \mathcal{S}} = 4^4 \qquad // \text{ In general: } f = \# \text{floors},\ p = \# \text{persons}$$

$$\boxed{f^p}$$

$$\mathcal{E}_i = \{ (i, i, i, i) \} \Rightarrow \# \mathcal{E} = 1$$

$$\mathcal{E} = \biguplus_{i=1}^{f} \mathcal{E}_i \Rightarrow \boxed{\# \mathcal{E}} = \boxed{f}$$

$$\frac{\# \mathcal{E}}{\# \mathcal{S}} = \frac{\cancel{f}^{1}\ 1}{f^{\cancel{p}\ p-1}} = \frac{1}{f^{p-1}}. \qquad \text{Here: } \frac{\# \mathcal{E}}{\# \mathcal{S}} = \frac{1}{4^3} = \frac{1}{2^6}$$

$$= \frac{1}{64}$$

disjoint union,
i.e., the $\mathcal{E}_i$ are mut. disjoint

Example 6 10 books: 4 CS, 3 Math, 2 stat, 1 Hist

Organize so that book of same subject are together:

E.g., SS. CCCC. H. MMM

How many possibilities?

1) # permutations of subjects $4 \cdot 3 \cdot 2 \cdot 1 = 4!$
2) permutation within subjects, e.g.

CS : 4!
M : 3!
S : 2!
H : 1!

$$\# \text{Arrangements} = 4! \cdot 4! \cdot 3! \, 2! \, 1!$$

$$= (2^3 \cdot 3) \cdot (2^3 \cdot 3) \cdot (2^1 \cdot 3) \cdot 2^1$$

$$= 2^8 \cdot 3^3 = 256 \cdot 27$$

$$\approx 250 \cdot 28$$

$$= 250 \cdot 4 \cdot 7 = 7 \cdot 1000 = 7000$$

$$4! = 4 \cdot 3 \cdot 2$$
$$= 2^2 \cdot 3 \cdot 2$$
$$= 2^3 \cdot 3$$

\# Arrangement, w/ book mixed up arbitrarily

$$10 \cdot 9 \cdot 8 \cdots 1 = 10!$$

Random arrang. of book leading to subjects together

$$\approx \frac{7000}{10!}$$

Example 7 Course with 5 male, 3 female students.

We had an exam: all students got different mark

$$P(\text{"}\underbrace{\text{all female students got the top marks}}_{\mathcal{E}}\text{"}) =$$

$\mathcal{S}$ = all possible rankings, $\#\mathcal{S} = 8!$

Females Top Rankings:

$$\underbrace{\boxed{\quad F \quad}}_{3!}\ \underbrace{\boxed{\quad M \quad}}_{5!}$$

$$P(\mathcal{E}) = \frac{\#\mathcal{E}}{\#\mathcal{S}} = \frac{3!\,5!}{8!} = \frac{3!\,\cancel{5!}}{8\cdot 7\cdot 6\cdot \cancel{5!}} = \frac{\cancel{3!}}{8\cdot 7\cdot \cancel{6}}$$

$$= \frac{1}{8\cdot 7} = \frac{1}{56}$$

Ex 2 Revisited: The top 3 students are a randomly chosen set of 3 out of 8.

We asked: What is the probability that a specific set is chosen?

How many choices of 3 out of 8 are possible?

1.) Choose 3 out of 8 in sequence: How many poss?

$$8 \cdot 7 \cdot 6$$

2.) A set of 3 can be obtained in $3!$ ways

a, b, c   b, c, a
a, c, b   c, a, b
b, a, c   c, b, a

In total: $\dfrac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = \dfrac{(8 \cdot 7 \cdot 6) \cdot (5 \cdot 4 \cdot \ldots \cdot 1)}{(3 \cdot 2 \cdot 1) \cdot (5 \cdot 4 \cdot \ldots \cdot 1)}$

$$= \frac{8!}{3!\,5!} = \binom{8}{3} = \binom{8}{5}$$

Generalize: Choose r out of n

$$\frac{n!}{r!\,(n-r)!} = \binom{n}{r}$$

"n choose r"

$$= \binom{n}{n-r}$$

"binomial coefficients"

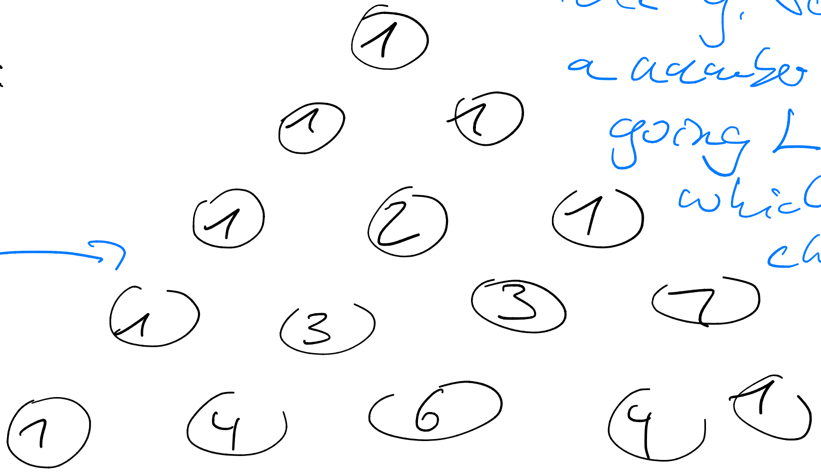
1.) Select r elements out of n:

$n(n-1)\cdots(n-r+1)$

2.) We get the same set r! times

$$(x+y)^n = \binom{n}{0}x^n y^0 + \binom{n}{1}x^{n-1}y^1 + \binom{n}{2}x^{n-2}y^2 + \ldots$$

Pascal's triangle:



only one path leading to a pin on the sides

We reach a node in the G. board by making a number of choices of going L or R, which correspond to choices of x and y

each node has the # of paths leading to it in the G. board

Example 9: Given 5 men, 8 women. Randomly select 5 persons.

P("2 men, 3 women are select") =

$\mathcal{S}$ = all possible choices of 5 out of 13

$\# \mathcal{S} = \binom{13}{5}$

$\mathcal{E}$ = selection 2 men ot of 5,
3 women out of 8

$$\# \mathcal{E} = \binom{5}{2} \cdot \binom{8}{3}$$

$$P(\mathcal{E}) = \frac{\binom{5}{2} \cdot \binom{8}{3}}{\binom{13}{5}} = \frac{560}{1287}$$

on handw.
notes

__Example 10:__ Given objects $1, \dots, n$. Select subset of size k.

$$P(\underbrace{\text{"1 is in the selection"}}_{E}) =$$

S = all subsets of size k of $\{1, \dots, n\}$

$$\# S = \binom{n}{k}$$

E = all subsets of size k __containing 1__

$$\# E = \binom{n-1}{k-1}$$

i.e., only choose remaining $k-1$ out of $n-1$

$$P(E) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\cancel{(n-1)!}}{\cancel{(k-1)!}\,\underbrace{\cancel{(n-1-(k-1))!}}_{\cancel{(n-k)!}}} \cdot \frac{k\cancel{!}\,\cancel{(n-k)!}}{n\cancel{!}} = \frac{k}{n}$$

1.4 Conditional Probabilities

Throw two dice

$$P(D_1 + D_2 = 8) = \frac{\#\{(2,6),(3,5),(4,4),(5,3),(6,2)\}}{\#\mathcal{S}}$$

$$= \frac{5}{36}$$

(What if $D_1 = 1$? $\Rightarrow P(\cdot) = 0$)

Suppose, we know $D_1 = 3$.

$$P(D_1 + D_2 = 8 \mid D_1 = 3) = ?$$

condition

New sample space $\mathcal{S}' = \{(3,1),(3,2) \ldots, (3,6)\}$

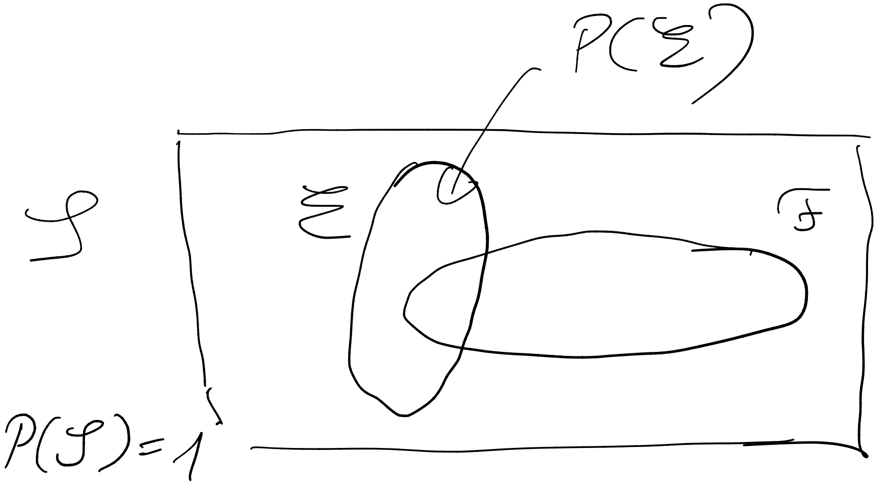
$$\#\mathcal{S}' = 6$$

$$\mathcal{E}' = \{(3,5)\} \Rightarrow \#\mathcal{E}' = 1$$

$$P(D_1 + D_2 = 8 \mid D_1 = 3) = \frac{\#\mathcal{E}'}{\#\mathcal{S}'} = \frac{1}{6}$$

Definition 11: Events $\mathcal{E}, \mathcal{F}$, $P(\mathcal{F}) > 0$

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



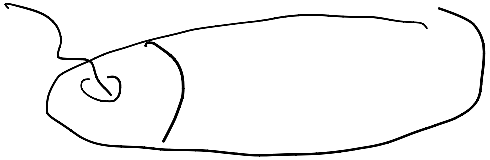
Note: $P(\cdot \mid \mathcal{F})$ is again a probability on $\mathcal{S}$

E.g.: $P(\mathcal{E}_1 + \dots + \mathcal{E}_k \mid \mathcal{F})$

$= P(\mathcal{E}_1 \mid \mathcal{F}) + \dots + P(\mathcal{E}_k \mid \mathcal{F})$ if $\mathcal{F}_i$ are mutually disjoint

Now, only outcomes in $\mathcal{F}$ are considered possible

$\mathcal{E}' = \mathcal{E} \cap \mathcal{F}$



$\mathcal{S}' = \mathcal{F}$

Normalize P to $P' = P(\cdot \mid \mathcal{F})$, such that $P'(\mathcal{S}') = 1$

$$P'(\mathcal{E}) = P(\mathcal{E} \mid \mathcal{F}) = \frac{P(\mathcal{E} \cap \mathcal{F})}{P(\mathcal{F})}$$

Subjective (= Bayesian) view of probabilities

We have seen the subj. view of cond. prob.

Frequentist view: n experiments

$\mathcal{F} \simeq n \cdot P(\mathcal{F})$ many times

$\mathcal{E}\mathcal{F} \simeq n \cdot P(\mathcal{E}\mathcal{F})$ many times

Ignore outcomes not $\mathcal{F}$.

Among the $n \cdot P(\mathcal{F})$ many $\mathcal{F}$-outcomes, there are $n\, P(\mathcal{E}\mathcal{F})$ many $\mathcal{E}\mathcal{F}$-outcomes

$$\Rightarrow P(\mathcal{E} \mid \mathcal{F}) = \frac{n \cdot P(\mathcal{E}\mathcal{F})}{n \cdot P(\mathcal{F})} = \frac{P(\mathcal{E}\mathcal{F})}{P(\mathcal{F})}$$

Example 12 Box with 32 transistors:

20 working, 8 partly working, 4 deficient

Exp: Choose 1 transistor,
Suppose it does not fail. What is the prob. that it is working?

Three events: W, P, D (pick a working, ... transistor)

$$P(W \mid \bar{D}) = \frac{P(W\bar{D})}{P(\bar{D})} = \frac{P(W)}{P(W \cup P)} = \frac{\frac{20}{32}}{\frac{28}{32}} = \frac{20}{28} = \frac{5}{7}$$

$$P(E \mid F) = \frac{P(EF)}{P(F)}$$

Quiz 7: Tossing Coins

$$\mathcal{E}$$

$$\mathcal{S} = \{ \underbrace{\underbrace{(H,H)}_{\mathcal{E}}, (H,T), (T,H)}_{\mathcal{F}}, (T,T) \}$$

$$P(\underbrace{2 \text{ heads}}_{\mathcal{E}} \mid \underbrace{\geq 1 \text{ head}}_{\mathcal{F}}) =$$

$$= \frac{P(\mathcal{E}\mathcal{F})}{P(\mathcal{F})} = \frac{P(2 \text{ heads} \wedge \geq 1 \text{ head})}{P(\geq 1 \text{ head})}$$

$$= \frac{P(\{(H,H)\})}{P(\{(H,H),(T,H),(H,T)\})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Quiz 8: Champions

$$P(\text{Champions}) = P(\overset{\text{reaching the final}}{F}\ \overset{\text{winning the final}}{W}) = P(W|F)\,P(F)$$

$$= .5 \times .2 = \underline{\underline{0.1}}$$

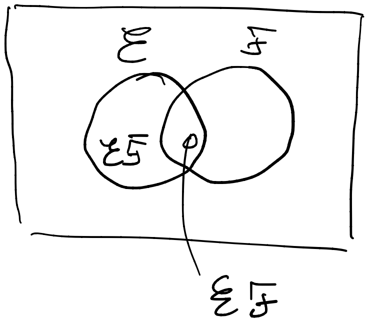
$$P(F) = .2$$

$$P(W|F) = .5$$

$$P(W|F) = \frac{P(WF)}{P(F)}$$

$$\Rightarrow P(W|F)\,P(F) = P(WF)$$

1.5 Bayes' Formula



$\mathcal{E}, \mathcal{F}$ events $\Rightarrow$ $\mathcal{E} = \mathcal{E}\mathcal{F} \,\dot{\cup}\, \mathcal{E}\bar{\mathcal{F}}$

$$\Rightarrow P(\mathcal{E}) = P(\mathcal{E}\mathcal{F}) + P(\mathcal{E}\bar{\mathcal{F}})$$

$$= P(\mathcal{E}|\mathcal{F})P(\mathcal{F}) + P(\mathcal{E}|\bar{\mathcal{F}})P(\bar{\mathcal{F}})$$

Here: $P(\mathcal{E})$ can be computed by conditioning on some $\mathcal{F}$

special case of the law of total probability (LOTP)

Example 15 Insurance company: People are $\overset{\mathcal{R}}{\text{risk}}$ taker (30%) or not. Every year, 40% of risk takers have an $\underset{\mathcal{A}}{\text{accident}}$, only 20% of non-risk takers. What is $P(\mathcal{A})$?

$$P(\mathcal{A}) = P(\mathcal{A}|\mathcal{R})P(\mathcal{R}) + P(\mathcal{A}|\bar{\mathcal{R}})P(\bar{\mathcal{R}})$$

$$= .4 \times .3 + .2 \times .7 = .12 + .14 = 0.26$$

updating beliefs in the presence of new information

Example 16 : Suppose, a client has an accident.

What is the prob. this was a risk taker?

$$\boxed{P(R \mid A)} = \frac{P(RA)}{P(A)} = \boxed{\frac{P(A \mid R)\, P(R)}{P(A)}}$$

Bayes' Law

Remember:

$$P(A \mid R) = \frac{P(AR)}{P(R)} = \frac{P(RA)}{P(R)}$$

$$\Rightarrow P(RA) = P(A \mid R)\, P(R)$$

$$= \frac{.4 \times .3}{0.26}$$

$$= \frac{.12}{.26} = \frac{12}{26} = \frac{6}{13}$$

$$= 0.46$$

Bayes' Formula

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

Quiz 9: Testing for a Disease

$\mathcal{D}$ person has disease $\qquad$ $\mathcal{T}$ test is positive

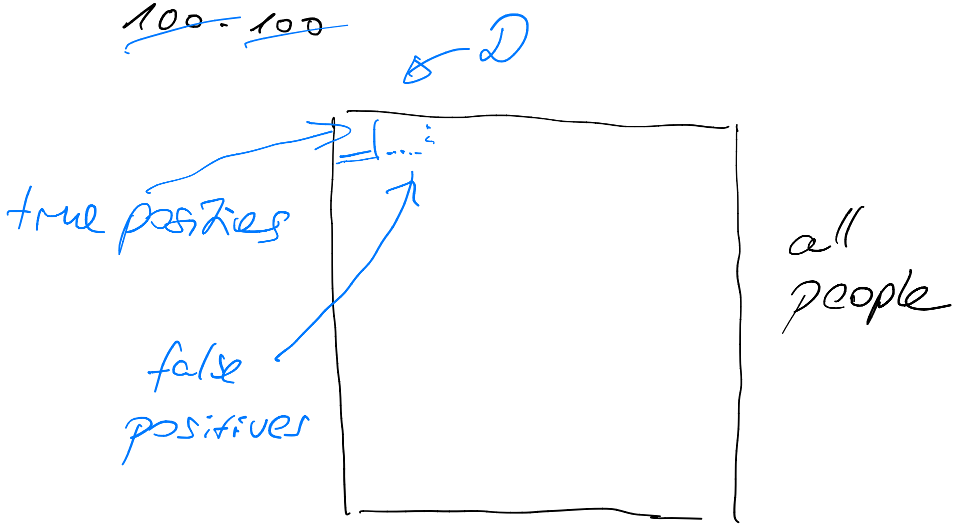
$$P(\mathcal{T} \mid \mathcal{D}) = \frac{99}{100} \qquad P(\mathcal{T} \mid \bar{\mathcal{D}}) = \frac{1}{100} \qquad P(\mathcal{D}) = \frac{1}{100}$$

$$P(\mathcal{D} \mid \mathcal{T}) = \frac{P(\mathcal{T} \mid \mathcal{D})\, P(\mathcal{D})}{P(\mathcal{T})} = \frac{\dfrac{\cancel{99}}{\cancel{100}} \cdot \dfrac{1}{\cancel{100}}}{\dfrac{2 \cdot \cancel{99}}{\cancel{100} \cdot \cancel{100}}} = \frac{1}{2}$$

$$P(\mathcal{T}) = P(\mathcal{T} \mid \mathcal{D})\, P(\mathcal{D}) + P(\mathcal{T} \mid \bar{\mathcal{D}})\, P(\bar{\mathcal{D}})$$

$$= \frac{99}{100} \cdot \frac{1}{100} + \frac{1}{100} \cdot \frac{99}{100}$$

LOTP



General Law of Total Probability (LOTP)

Let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be a partition of $\mathcal{S}$, i.e.

- $\mathcal{F}_i \cap \mathcal{F}_j = \emptyset, \quad i \neq j$
- $\bigcup_{i=1}^{n} \mathcal{F}_i = \mathcal{F}_1 \cup \dots \cup \mathcal{F}_n = \mathcal{S}$

Note:

$$\sum_{i=1}^{n} P(\mathcal{F}_i) = 1$$

$\mathcal{E} \subseteq \mathcal{S}$ event $\Rightarrow \mathcal{E} = \biguplus_{i=1}^{n} \mathcal{E}\mathcal{F}_i$

$$\Rightarrow P(\mathcal{E}) = \sum_{i=1}^{n} P(\mathcal{E}\mathcal{F}_i) = \sum_{i=1}^{n} P(\mathcal{E} \mid \mathcal{F}_i)\, P(\mathcal{F}_i)$$

Quiz 11: Marbles

LOTP: What are $\mathcal{E}$ and the $\mathcal{F}_i$?

$\mathcal{E}$ = Marble drawn is red

$\mathcal{F}_i$ = The i-th bag is chosen

$$P(\mathcal{E}) = P(\mathcal{E} | \mathcal{F}_1) P(\mathcal{F}_1) + P(\mathcal{E} | \mathcal{F}_2) P(\mathcal{F}_2) + P(\mathcal{E} | \mathcal{F}_3) P(\mathcal{F}_3)$$

$$P(\mathcal{F}_1) = P(\mathcal{F}_2) = P(\mathcal{F}_3) = \frac{1}{3}$$

$$P(\mathcal{E} | \mathcal{F}_1) = \frac{80}{100} \qquad P(\mathcal{E} | \mathcal{F}_2) = \frac{55}{100} \qquad P(\mathcal{E} | \mathcal{F}_3) = \frac{45}{100}$$

$$\Rightarrow \quad P(\mathcal{E}) = \frac{1}{3} \frac{80 + 55 + 45}{100} = \frac{1}{\underset{1}{\cancel{3}}} \frac{\overset{60}{\cancel{180}}}{100} = \frac{60}{100}$$

$(F_i)_{i=1}^{n}$ partition of S, E event

$$P(F_i \mid E) =$$

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

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

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

Generalized Bayes' Formula

Quiz 10: Bivid-20 Diagnosis

B person has Bivid

T test is positive

Interesting: $P(B|T) = ?$

$$P(T|B) = \frac{99}{100}$$

$$P(\bar{T}|\bar{B}) = \frac{99}{100} \Rightarrow P(T|\bar{B}) = \frac{1}{100}$$

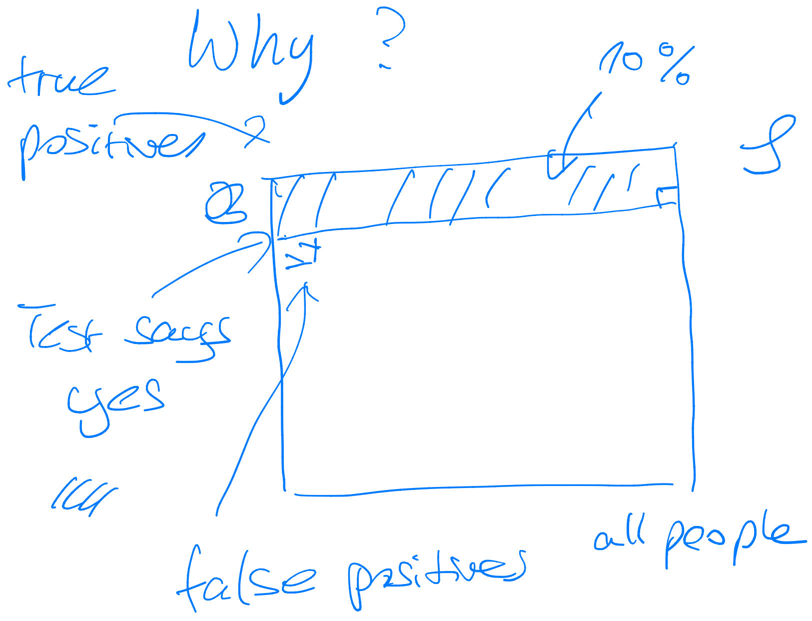
$$P(B) = \frac{10}{100}$$

$$P(B|T) = \frac{P(T|B)\,P(B)}{P(T)} = \frac{\frac{99}{\cancel{100}}\cdot\frac{\cancel{10}}{\cancel{100}}}{\frac{99}{\cancel{100}}\cdot\frac{\cancel{10}}{\cancel{100}} + \frac{1}{\cancel{100}}\cdot\frac{\overset{?}{\cancel{90}}}{\cancel{100}}}$$

$$P(T) = P(T|B)\,P(B) + P(T|\bar{B})\cdot P(\bar{B})$$

$$= \frac{99}{99+9} = \frac{99}{108} = 91.67$$

$$\approx 92$$



1.6 Independent Events

Example : Consider a deck of French cards

$\mathcal{E}$ = draw a red card

$\mathcal{F}$ = draw an ace

$$P(\mathcal{F}) = \frac{4}{52} = \frac{1}{13}$$

red aces

$$P(\mathcal{F} \mid \mathcal{E}) = \frac{2}{26}$$

\# red cards

$$P(\mathcal{E}) = \frac{1}{2}, \qquad P(\mathcal{E} \mid \mathcal{F}) = \frac{1}{2}$$

Intuition: Knowing $\mathcal{F}$ doesn't tell us anything about $\mathcal{E}$

$\mathcal{E}, \mathcal{F}$ are independent

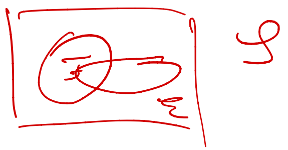
In general, $P(\mathcal{E} \mid \mathcal{F}) \neq P(\mathcal{E})$

prior probability of $\mathcal{E}$

posterior prob. of $\mathcal{E}$

Definition : $\mathcal{E}, \mathcal{F}$ independent $\qquad P(\mathcal{E} \mid \mathcal{F}) = P(\mathcal{E})$

Note: the definition is symmetric



$$P(\mathcal{E} \mid \mathcal{F}) = P(\mathcal{E}) \iff P(\mathcal{E}) = \frac{P(\mathcal{E}\mathcal{F})}{P(\mathcal{F})}$$

$$\iff \boxed{P(\mathcal{E})P(\mathcal{F}) = P(\mathcal{E}\mathcal{F})}$$

if $P(\mathcal{F}) \neq 0$,
$P(\mathcal{E}) \neq 0$

The first definition assumes that $P(\mathcal{F}) \neq 0$

Alternative:

$\mathcal{E}, \mathcal{F}$ indep. iff $P(\mathcal{E}\mathcal{F}) = P(\mathcal{E}) \cdot P(\mathcal{F})$

$$P(\mathcal{E})P(\mathcal{F}) = P(\mathcal{E}\mathcal{F}) \Rightarrow P(\mathcal{F}) = \frac{P(\mathcal{E}\mathcal{F})}{P(\mathcal{E})}$$

$$= \frac{P(\mathcal{F}\mathcal{E})}{P(\mathcal{E})} = P(\mathcal{F} \mid \mathcal{E})$$

Quiz: Dice and independence

$\mathcal{E}$ = "$D_1 + D_2 = 7$"

$\mathcal{F}$ = "$D_1 + D_2 = 8$"

$\mathcal{G}$ = "$D_1 = 3$"

$\mathcal{F}: \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$

$\Rightarrow \# \mathcal{F} = 5$

$\mathcal{G} = \{(3,1), \dots, (3,6)\}$

$\Rightarrow \# \mathcal{G} = 6$

$\mathcal{E}$ ind $\mathcal{F}$: $P(\mathcal{E}\mathcal{F}) = 0 \neq \frac{1}{6} \cdot \frac{5}{36} = P(\mathcal{E})P(\mathcal{F})$ No!

$\mathcal{E}$ ind $\mathcal{G}$: $P(\mathcal{E}\mathcal{G}) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(\mathcal{E})P(\mathcal{G})$ Yes!

$\mathcal{F}$ ind $\mathcal{G}$: $P(\mathcal{F}\mathcal{G}) = \frac{1}{36} \neq \frac{5}{36} \cdot \frac{1}{6} = P(\mathcal{F})P(\mathcal{G})$ No!

$$P(\mathcal{E}) = \frac{6}{36} = \frac{1}{6}$$

$$P(\mathcal{F}) = \frac{5}{36}$$

$$P(\mathcal{G}) = \frac{6}{36} = \frac{1}{6}$$

$$P(\mathcal{E}\mathcal{F}) = 0$$

$$P(\mathcal{E}\mathcal{G}) = \frac{1}{36}$$

$$P(\mathcal{F}\mathcal{G}) = \frac{1}{36}$$

Example: Consider a deck of French cards

$\mathcal{E}$ = draw a red card

$\mathcal{F}$ = draw an ace

$$P(\mathcal{E}) = \frac{1}{2}, \qquad P(\mathcal{E} / \mathcal{F}) = \frac{1}{2}$$

$$P(\mathcal{F}) = \frac{1}{13}, \qquad P(\mathcal{F} / \mathcal{E}) = \frac{1}{13}$$

We observe as well:

$$P(\bar{\mathcal{E}} / \mathcal{F}) = \frac{1}{2}$$

$$P(\bar{\mathcal{F}} / \mathcal{E}) = \frac{12}{13}$$

$$P(\bar{\mathcal{E}}) = \frac{1}{2}$$

$$P(\bar{\mathcal{F}}) = \frac{12}{13}$$

Symmetry:

$$\boxed{P(\mathcal{E}\mathcal{F}) = P(\mathcal{E})\,P(\mathcal{F})}$$

Definition of independence of events

$$\Longleftrightarrow P(\mathcal{E} / \mathcal{F}) = P(\mathcal{E})$$

$$\Longleftrightarrow P(\mathcal{F} / \mathcal{E}) = P(\mathcal{F})$$

if $P(\mathcal{E}) \neq 0 \neq P(\mathcal{F})$

Proposition 20 $\mathcal{E}, \mathcal{F}$ independent $\Rightarrow$ $\mathcal{E}, \bar{\mathcal{F}}$ independent

Proof: Show $P(\mathcal{E}\bar{\mathcal{F}}) = P(\mathcal{E})P(\bar{\mathcal{F}})$

$$P(\mathcal{E}) = P(\mathcal{E}\mathcal{F} \cup \mathcal{E}\bar{\mathcal{F}}) = P(\mathcal{E}\mathcal{F}) + P(\mathcal{E}\bar{\mathcal{F}})$$

$$= P(\mathcal{E}) \cdot P(\mathcal{F}) + P(\mathcal{E}\bar{\mathcal{F}})$$

$$\Rightarrow \underset{\displaystyle (1 - P(\mathcal{F}))P(\mathcal{E})}{\underset{\parallel}{P(\mathcal{E})}} - P(\mathcal{E}) \cdot P(\mathcal{F}) = \underline{P(\mathcal{E}\bar{\mathcal{F}})}$$

$$(1 - P(\mathcal{F}))P(\mathcal{E}) = \underline{P(\bar{\mathcal{F}}) \cdot P(\mathcal{E})}$$

$\mathcal{E}, \mathcal{F}$ ind. $\Rightarrow$ $\bar{\mathcal{E}}, \bar{\mathcal{F}}$ ind.

Independence is inherited by complements

Consider $\mathcal{E}$ ind of $\mathcal{F}$, $\quad \not\Longrightarrow \quad \mathcal{E}$ ind $\mathcal{F} \cap \mathcal{G}$
$\mathcal{E}$ ind of $\mathcal{G}$

<u>Example:</u> Throw two dice

$\mathcal{E} =$ "$D_1 + D_2 = 7$" $\qquad \mathcal{F} =$ "$D_1 = 1$" $\qquad \mathcal{G} =$ "$D_2 = 6$"

$$P(\mathcal{E}) = \frac{1}{6} \qquad P(\mathcal{F}) = \frac{1}{6} \qquad P(\mathcal{G}) = \frac{1}{6}$$

$$P(\mathcal{E}\mathcal{F}) = \frac{1}{36} \qquad P(\mathcal{E}\mathcal{G}) = \frac{1}{36} \qquad P(\mathcal{F}\mathcal{G}) = \frac{1}{36}$$

$\Longrightarrow$ $\mathcal{E}, \mathcal{F}$ ind. $\qquad \mathcal{E}, \mathcal{G}$ ind

What about ind. of $\mathcal{E}$ and $\mathcal{F} \cap \mathcal{G} = \mathcal{F}\mathcal{G}$?

$$P(\mathcal{E})P(\mathcal{F}\mathcal{G}) = \frac{1}{6}\frac{1}{36} \neq \frac{1}{36} = P(\mathcal{E}\mathcal{F}\mathcal{G}) = P(\mathcal{E}(\mathcal{F}\mathcal{G}))$$

$\Longrightarrow$ $\mathcal{E}$ and $\mathcal{F}\mathcal{G}$ are <u>not</u> ind.

Let $\mathcal{E}$ ind of $\mathcal{F}$, $\mathcal{E}$ ind of $\mathcal{G}$. Can we conclude that $\mathcal{E}$ ind of $\mathcal{F}\mathcal{G}$ $(= \mathcal{F} \cap \mathcal{G})$?

No!

Example: Throw two dice!

$\mathcal{E} =$ "$D_1 + D_2 = 7$" $\quad \mathcal{F} =$ "$D_1 = 1$" $\quad \mathcal{G} =$ "$D_2 = 6$"

$P(\mathcal{E}) = \frac{1}{6} \quad P(\mathcal{F}) = \frac{1}{6} \quad P(\mathcal{G}) = \frac{1}{6}$

$P(\mathcal{E}\mathcal{F}) = \frac{1}{36} \quad P(\mathcal{E}\mathcal{G}) = \frac{1}{36} \quad P(\mathcal{F}\mathcal{G}) = \frac{1}{36}$

$\Rightarrow \mathcal{E}, \mathcal{F}$ ind. $\quad \Rightarrow \mathcal{E}, \mathcal{G}$ ind.

What about the independence of $\mathcal{E}$ and $\mathcal{F}\mathcal{G}$?

Compare $P(\mathcal{E}(\mathcal{F}\mathcal{G}))$ and $P(\mathcal{E}) \cdot P(\mathcal{F}\mathcal{G})$!

$$P(\mathcal{E}(\mathcal{F}\mathcal{G})) = P(\mathcal{E}\mathcal{F}\mathcal{G}) = \frac{1}{36}$$

$$P(\mathcal{E}) \cdot P(\mathcal{F}\mathcal{G}) = \frac{1}{6} \cdot \frac{1}{36}$$

Since $\frac{1}{36} \neq \frac{1}{6} \cdot \frac{1}{36}$, $\mathcal{E}$ and $\mathcal{F}.\mathcal{G}$ are not independent

Definition 22 : $\mathcal{E}, \mathcal{F}, \mathcal{G}$ are *independent* if

- $\mathcal{E}, \mathcal{F}$ and $\mathcal{E}, \mathcal{G}$ and $\mathcal{F}, \mathcal{G}$ are ind.
  (pairw. ind.)
- $P(\mathcal{E}\mathcal{F}\mathcal{G}) = P(\mathcal{E})P(\mathcal{F})P(\mathcal{G})$

Now, $\mathcal{E}$ and $\mathcal{F}\mathcal{G}$ are ind.

$$P(\mathcal{E}\cdot(\mathcal{F}\mathcal{G})) = P(\mathcal{E}\mathcal{F}\mathcal{G}) = P(\mathcal{E})\,\underbrace{P(\mathcal{F})\;P(\mathcal{G})}$$

$$= P(\mathcal{E})\;P(\mathcal{F}\mathcal{G}) \leftarrow \text{ind. of } \mathcal{F}, \mathcal{G}$$

Remark: $\mathcal{E}, \mathcal{F}, \mathcal{G}$ ind $\Longrightarrow$ $\mathcal{E}$ and $(\mathcal{F} \cup \mathcal{G})$ ind.

$$
\begin{aligned}
P(\mathcal{E}(\mathcal{F} \cup \mathcal{G})) &= P(\mathcal{E}\mathcal{F} \cup \mathcal{E}\mathcal{G}) \\
&= P(\mathcal{E}\mathcal{F}) + P(\mathcal{E}\mathcal{G}) - P(\mathcal{E}\mathcal{F}\mathcal{G}) \\
&= P(\mathcal{E})P(\mathcal{F}) + P(\mathcal{E})P(\mathcal{G}) - P(\mathcal{E})\overbrace{P(\mathcal{F})P(\mathcal{G})}^{=P(\mathcal{F}\mathcal{G})} \\
&= P(\mathcal{E})\left(P(\mathcal{F}) + P(\mathcal{G}) - P(\mathcal{F}\mathcal{G})\right) \\
&= P(\mathcal{E}) \; P(\mathcal{F} \cup \mathcal{G})
\end{aligned}
$$

Generalized Definition

$E_1, \dots, E_n$ are independent iff

for every subset of $F_1, \dots, F_m$:

$$P(F_1 \cdots F_m) = P(F_1) \cdots P(F_m)$$

Examples:

- Sequences of experiments; E_i refers to i-th execution of experiment

  E.g.: Rolling die

- Sequences of disease tests:

  Intuition for one test being positive for people w/ and w/o disease:

$$P(T \mid D) = .99 \qquad P(\bar{T} \mid \bar{D}) = .99$$

$$\underline{P(T \mid \bar{D})} = .01$$

prob. of false positives

- Sequences of disease tests:

Intuition for one test being positive for people w/ and w/o disease:

$$P(T|D) = .99 \qquad P(\bar{T}|\bar{D}) = .99$$

sensitivity

$$\underline{P(T|\bar{D})} = .01$$

prob. of false positives

- Let T_1, T_2 be two applications of test to same person w/ pos. outcome

$$P(T_1 T_2 | D) \qquad \frac{99}{100} \cdot \frac{99}{100} \approx .98$$

sensitivity does not go much down

cond. independence

$$\leadsto \overset{?}{=} P(T_1|D) \cdot P(T_2|D)$$

$$P(T_1 T_2 | \bar{D}) \; \left(= P((T_1 \cap T_2) | \bar{D}) \right)$$

$$\overset{?!}{=} P(T_1|\bar{D}) \cdot P(T_2|\bar{D}) = \frac{1}{100} \cdot \frac{1}{100} = \frac{1}{10^4}$$

probability for false positives is small

- Sequences of disease tests

Example probabilities for a test being positive for people w/ and w/o disease

specificity

$$P(T \mid D) = .99 \qquad P(\bar{T} \mid \bar{D}) = .99$$

sensitivity

$$\Leftrightarrow P(T \mid D) = .01$$

If $P(D)$ is low (e.g., 1%), then we cannot rely on the test because there are as many true positive as false positive test results.

- What can we do?

Idea: Apply the test twice! First T_1, then T_2.

But: Need to ensure that probabilities multiply!

That is

$$P(T_1, T_2 | D) = P(T_1 | D)\, P(T_2 | D)$$

numbers from our example ↙

$$= \frac{99}{100} \cdot \frac{99}{100} = \frac{9801}{10\,000} \approx 98\%$$

This property of T_1, T_2, D is spelt out as

"T_1 and T_2 are conditionally independent given D"

It is independence of T_1, T_2 with regard to the probability measure

$$P(\cdot | D)$$

remember that for every probability $P(\cdot)$ on S also $P(\cdot | D)$ is a probability on S.

Assume that T_1, T_2 are also independent given $\bar{D}$.

Then

$$P(T_1, T_2 \mid \bar{D}) = P(T_1 \mid \bar{D}) \cdot P(T_2 \mid \bar{D})$$

numbers
from
↙ our example

$$= \frac{1}{100} \cdot \frac{1}{100} = \frac{1}{10\,000}$$

This shows that the probability of false positives has been sharply reduced:
The relationship between false and true positives now is

numbers
from
↙ our example

$$\frac{1}{9801}$$

A Bayesian analysis of what can be concluded from the other test results ($T_1 \bar{T}_2$, $\bar{T}_1 T_2$ and $\bar{T}_1 \bar{T}_2$) will be part of the assignment.

$$\boxed{\text{What is } P(T_1 \bar{T_2} \mid D)\,?}$$

$$= P(T_1 \mid D)\, P(\bar{T_2} \mid D) = \frac{99}{100} \cdot \frac{1}{100} = \frac{99}{10^4}$$

$$P(T_1 \bar{T_2} \cup \bar{T_1} T_2 \mid D) \approx \frac{200}{10.000} = \frac{2}{100}$$

$$P(D \mid T_1 \bar{T_2} \cup \bar{T_1} T_2)$$

$$= \frac{P(T_1 \bar{T_2} \cup \bar{T_1} T_2 \mid D) \cdot P(D)}{P(T_1 \bar{T_2} \cup \bar{T_1} T_2)} = \frac{2\, P(T_1 \bar{T_2} \mid D) \cdot P(D)}{2\, P(T_1 \bar{T_2})}$$

$$P(T_1 \bar{T_2}) = P(T_1 \bar{T_2} \mid D) \cdot P(D) + P(T_1 \bar{T_2} \mid \bar{D})\, P(\bar{D})$$

$$= \frac{99}{10^4} \cdot \frac{1}{100} + \frac{99}{10^4} \cdot \frac{99}{100}$$

$$= \frac{\cancel{2} \cdot \frac{99}{\cancel{10^4}} \cdot \frac{1}{\cancel{100}}}{\cancel{2} \cdot \frac{99 + 99^2}{\cancel{10^6}}} = \frac{99}{(1+99) \cdot 99} = \boxed{\frac{1}{100}}$$

$$\boxed{\text{What is } P(T_1 \bar{T}_2 \mid \bar{D})\,?}$$

$$= P(T_1 \mid \bar{D})\, P(\bar{T}_2 \mid \bar{D}) = \frac{1}{100} \cdot \frac{99}{100} = \frac{99}{10^4}$$

$$P(T_1 \bar{T}_2 \cup \bar{T}_1 T_2 \mid \bar{D}) \approx \frac{200}{10.000} = \frac{2}{100}$$

$$P(\bar{D} \mid T_1 \bar{T}_2 \cup \bar{T}_1 T_2)$$

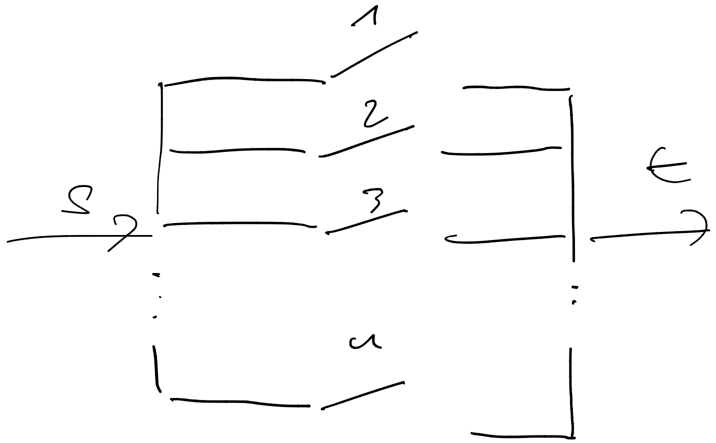
$$= \frac{P(T_1 \bar{T}_2 \cup \bar{T}_1 T_2 \mid \bar{D}) \cdot P(\bar{D})}{P(T_1 \bar{T}_2 \cup \bar{T}_1 T_2)} = \frac{2\, P(T_1 \bar{T}_2 \mid \bar{D}) \cdot P(\bar{D})}{2\, P(T_1 \bar{T}_2)}$$

$$P(T_1 \bar{T}_2) = P(T_1 \bar{T}_2 \mid D) \cdot P(D) + P(T_1 \bar{T}_2 \mid \bar{D})\, P(\bar{D})$$

$$= \frac{99}{10^4} \cdot \frac{1}{100} + \frac{99}{10^4} \cdot \frac{99}{100}$$

$$= \frac{\cancel{2} \cdot \frac{99}{\cancel{10^4}} \cdot \frac{99}{\cancel{100}}}{\cancel{2} \cdot \frac{99 + 99^2}{\cancel{10^6}}} = \frac{99^{\cancel{2}}}{(1+99) \cdot \cancel{99}} = \boxed{\frac{99}{100}}$$

Example 23



components are ind.

work with p_i for comp i

System works if ≥ 1 comp. works

$\mathcal{E}$ = "system works", F_i = "comp i works" $\Rightarrow P(\mathcal{E}) = ?$

$\overline{\mathcal{E}}$ = "system doesn't work" if no comp. works

if $\overline{F}_1$ and $\overline{F}_2$ ... $\overline{F}_n$

Note: (F_i) ind $\Rightarrow$ $(\overline{F}_i)$ ind

$$P(\mathcal{E}) = 1 - P(\overline{\mathcal{E}}) = 1 - P(\overline{F}_1 \cdots \overline{F}_n) = 1 - P(\overline{F}_1) \cdots P(\overline{F}_n)$$

$$= 1 - \prod_{i=1}^{n} P(\overline{F}_i) = 1 - \prod_{i=1}^{n} (1 - P(F_i)) = 1 - \prod_{i=1}^{n} (1 - p_i)$$