Bayes rule:

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

$$P(EF) = P(E/F)\,P(F) \quad \text{— (1)}$$

$$P(EF) = P(F/E)\,P(E) \quad \text{Also — (2)}$$

Equating (1) and (2)

$$P(E/F)\;P(F) = P(F/E)\,P(E)$$

$$\boxed{P(E/F{=}a) = \frac{P(F/E)\,P(E)}{P(F)}}$$ Bayes rule

Finding probabability of event E given some meesnement F

F can be noisy

Hence we should use another probability meesure which quantifies the uncertainity in F

Chances of labeling C_i for a particular measurement F

$$\underline{P(C_i/F)} = \frac{P(C_i F)}{P(F)} = \frac{\overbrace{\underline{P(F/C_i)}}^{\text{likelihood}}\ \overbrace{P(C_i)}^{\text{prior}}}{P(F)}$$

$S = C_1 \cup C_2$ (Also expressed as $\underline{S = \bigcup_i C_i}$)

union of ME events (classes)

$$F = \bigcup_i F C_i$$

$$P(F) = \sum P(F C_i)$$

$$P(F) = \sum_i P(C_i)\, P(F/C_i)$$

Total probability

Distribution of F (eg. measurements) over the sample space S

---

likelihood : $P(F/C_i)$: Assuming $F \to C_i$ what are chances of this corresponding

Prior: chances of that class actually being present.

Independence

$$\boxed{P(A/B) = P(A)} \rightarrow ①$$

A is not dependent on B

$$\frac{P(AB)}{P(B)} = P(A)$$

$$\Rightarrow \boxed{P(AB) = P(A)\,P(B)} \Rightarrow ②$$

↳ Decomposition of a joint distribution in term of 1D conditional distribution

Practical example in DS : $P(F/C_i)$ → likelihood of one attribute conditioned over a class

A case of multiple attributes : $P(F_1, F_2 \ldots F_n / C)$ → liklihood of many attributes conditioned over a class

If $F_1, F_2 \ldots F_n$ are independent then

$$P(F_1, F_2 \ldots F_n/C) = P(F_1/C)\,P(F_2/C) \ldots P(F_n/C)$$

↳ Decomposition a high-dimensional likelihood in low dim - likelihoods

Conditional Independence

Independence of complement

If A and B independent then A and B^c are also independent

$$\Updownarrow$$

$$P(AB^c) = P(A)P(B^c)$$

Proof: $A = AB \cup AB^c$

$$P(A) = P(AB) + P(AB^c)$$

$$P(A) = P(A)P(B) + P(AB^c)$$

$$P(AB^c) = P(A)(1-P(B))$$

$$P(AB^c) = P(A)P(B^c)$$

Independence of multiple events

Example of 3 events: A, B, C

$$P(ABC) = P(A)P(B)P(C)$$

$$P(AB) = P(A)P(B)$$

$$P(BC) = P(B)P(C)$$

$$P(AC) = P(A)P(C)$$

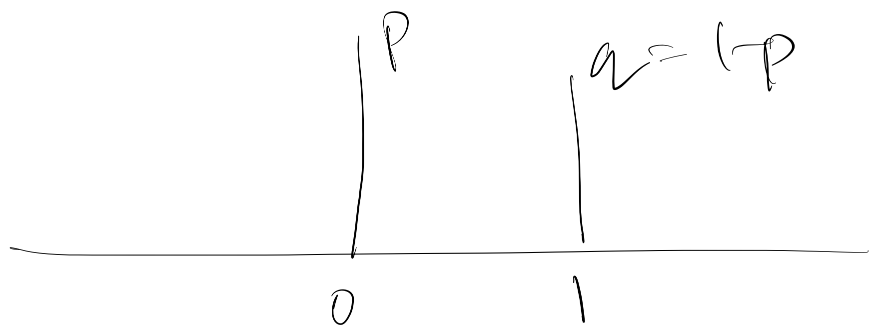
Independence of n events $A_1, A_2, \ldots, A_n$

$\Rightarrow$ independence of all the subsets of $\{A_1, \ldots, A_n\}$

Distributions (simple examples)

Independence $\longrightarrow$ a distribution of two independent events (yes/no)

$$\left.\begin{aligned} P(A=0) &= p \\ P(A=1) &= q = 1-p \end{aligned}\right] \text{ Bernoulli's distribution}$$



$\Downarrow$

Repeat the experiment (which has two outcomes) multiple times

One way to consider probabilities in such a scenario is given by Binomial distribution

Binomial distribution:

If an exp with two outcomes with probability p and q (yes/no), is repeated k times, the probability of k successes in n exps is given by the Binomial distribution

To define this, we need some background in counting

$\Downarrow$

Simple counting $\rightarrow$ complex ways of counting involve relationships, arrangements, selections of different variables.

$\Downarrow$

Two important counting principles

permutation $\longrightarrow$ combinations

Permutations: Arrangements

No. of ways in which k objects can be arranged among n positions is

$$P_{k,n} = n(n-1)(n-2)\cdots(n-k+1) \quad -(1)$$

For $k=n$, $P_{k,n} = n(n-1)\cdots 1 = n! \quad -(2)$

$$\boxed{P_{k,n} = \frac{n!}{(n-k)!}}$$

From ① and ②

Famous example of Permutations: Birthday pairing

(I) No of possible states for each person: 365

No. of persons: k

Total no. of states: 365^k ⟹ denominator in computing probability

⇕

Total possible no. of outcomes ⟶

(II) To compute the no. of cases where atleast 2 share a birthday ⟹

Compute the no. of cases where no 2 people share the birthday

Assumption: $k \leq 365$

$P_{k,365}$ is no. of cases where no two people share birthdays ⟹ probability $= \frac{P_{k,365}}{365^k}$

$\Downarrow$

probability of <u>atleast 2</u> people sharing a birthday = $\boxed{1 - \frac{P_{k,365}}{365^k}}$ $\quad \underline{\underline{k < 365}}$

Combination : <u>selection</u>

How many ways you can <u>select</u> k objects from n objects

eg. For two objects a, b
(a, b) or (b, a) is the same selection
But " " " is a diff arrangement

$\Downarrow$

Permutations has an aspect of <u>arrangements</u> in addition to <u>selection</u>

$\Downarrow$

Permutation = (selection of k) (arrangements of k)

$P_{k,n} = (C_{k,n})\, k!$

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

Bernoulli distribution ⟶ Binomial distribution

⇓

If the yes/no trials are repeated n times then we can have k yes and n−k no

Binomial distribution:
probability of having k successes (yes) in n trials, where the probability of a single success (yes) is p (comes from Bernoulli distribution)

Assumption: Each trial is independent

eg. out of 10 trials, probability of one instance could be as follows

$$P(n,k,p) = \underbrace{(p)(p)(p)(p)}_{4}\underbrace{(1-p)(1-p)\cdots(1-p)}_{6}$$

Other scenarios: different ways of selecting k out of n

e.g. $p(1-p)p(1-p)\ldots$

$pp(1-p)p(1-p)\ldots$

$\Downarrow$

$C_{k,n}$ $\underbrace{\binom{n}{k}}_{\text{another notation of combination}}$

Overall probability considering all possible combinations

$$\boxed{\underbrace{P(p,n,\underline{k})}_{\Downarrow} = \underbrace{C_{k,n}}_{\Downarrow} \underbrace{p^k (1-p)^{n-k}}_{\Downarrow}}$$

$\underline{P(k)_{p,n}}$

$\Downarrow$

probability is a function

no. of possible combinations of yes/no

probability for one instance of combination of yes/no