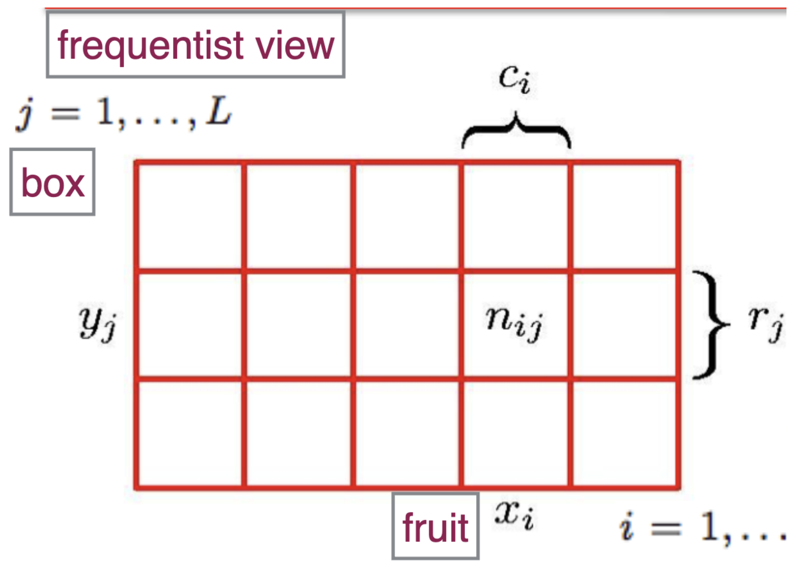
Bayesian Probabilities

Simulate Monty Hall Problem: three binary variables, run the simulations and get the fractions, which converges to the final probability.

Case study: determine the precise year the change in regulation was introduced
Method: Bayesian principle + MC chain

Probability Theory
Frequentist view: repeat the experiments infinite number of times, and the result obtained is the probability, regardless of whether the event is composite.

A trial is conducted n times, and the results fall into the following grid.



Sum Rule : $p(X) = \sum_Y p(X, Y)$
Product Rule : $p(X, Y) = p(Y|X)p(X)$
Bayes' Theorem : $p(Y|X) = \dfrac{p(X|Y)p(Y)}{p(X)}$

Marginal probability: p($X = x_i$)=$\dfrac{c_i}{N}$
Conditional probability: p($Y = y_j|X = x_i$)=$\dfrac{n_{ij}}{c_i}$

Prior probability : the probability that the knight captures a troll
Posterior probability
Apply Bayes' Theorem on Trolls under the bridge: calculate $P(H_1|T)$

Kolmogorov probability

sample space S, event A as an area in the sample space
(Ω, F, P) probability space : Ω -- sample space; F -- set of events;

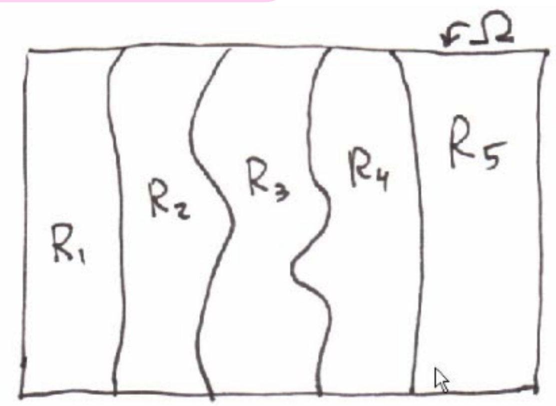
- each event is a subset of Ω cotaining zero or more outcomes

Axioms: (satisfied by frequentist definition of probability)

1. P(A) $\geq$ 0 for an event A
2. P(Ω) = 1
3. if A ∩ B = 0, then P(A ∪ B)=P(A)+P(B)

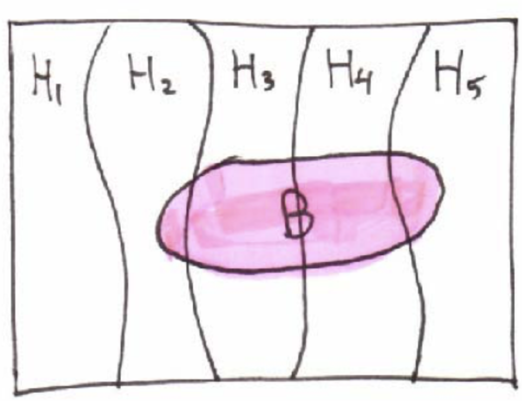
Venn Diagrams: think of probability as area

Law of Exhaustion



If R_i is EME (exhaustive and mutually exclusive) $\rightarrow \sum_i P(R_i) = 1$
Inclusion-exclusion principle: $P(E^c) = P(\Omega \backslash E) = 1 - P(E)$

Law of Total Probability/Law of de-Anding

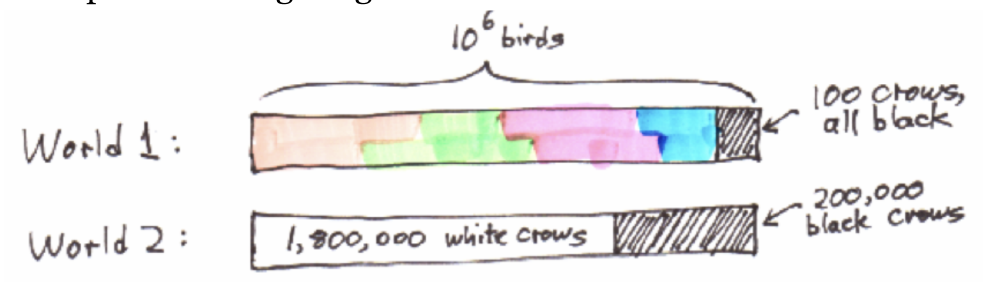


$P(B) = P(BH1) + P(BH2) + \ldots = \sum_i P(BH_i)$
$P(B) = \sum_i P(B|H_i)P(H_i)$

$P(H_i|B) \propto P(B|H_i)P(H_i)$

Normalization factor: $\frac{1}{\sum_j P(B|H_j)P(H_j)}$

Example of Bird sighting



$\frac{H1|D}{H2|D} = \frac{P(D|H1)P(H1)}{P(D|H2)P(H2)}$

Calculate Monty Problem using $P(H_i|O_3) \propto P(O_3|H_i)P(H_i)$

$P(O_3|H_i)$: probability that Monty opens door 3 if the prize is behind door i.

Bayesian Parameter Estimates

Example of releasing prisoners: only two prisoners will be released

$P(A|S_B) = P(AB|S_B) + P(AC|S_B)$=$\frac{1}{1+x}$

$P(S_B|BC) = x$: the probability of BC being released while the jailer says B

Bernoulli Distribution

Bernoulli trials : 2 possible outcomes; independent identical events; single parameter x (probability of one outcome); N & N_B

$P(data|x) = x^{N_B}(1-x)^{N-N_B}$

$P(x|data) \propto x^{N_B}(1-x)^{N-N_B} \times P(x|I)$ calculate x given data

Central limit theorem

$u_i =< x^i >= \int x^i p(x)dx$

$M_i = \int (x- < x >)^i p(x)dx$

$M_2 = Var(x)$

Skew(x)=$\frac{M_3}{M_2^{3/2}}$

Gaussian Distribution / Normal distribution

$p(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}[\frac{x-\mu}{\sigma}]^2)$

Cauchy Distribution

$p(x) = \frac{1}{\pi\sigma}(1 + [\frac{x-\mu}{\sigma}]^2)^{-1}$

Slowest falling tails, area=1 (M_0), divergent M_1 and M_2

Semi-invariants: higher moments that are additive

Characteristic Function of a distribution is its Fourier Transform

$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} p_X(x) dx$

S=X+Y

$p_S(s) = \int p_X(u) p_Y(s-u) du$

$\phi_S(t) = \phi_X(t)\phi_Y(t)$

the simple sum of a larger number of random variables is normally distributed, with variance equal to the sum of the variances

$p \sum X_i(\cdot) \sim Normal(0, \sum \sigma_i^2)$

Maximum a posteriori (MAP): the maximum of the posterior distribution given some data for a uniform prior