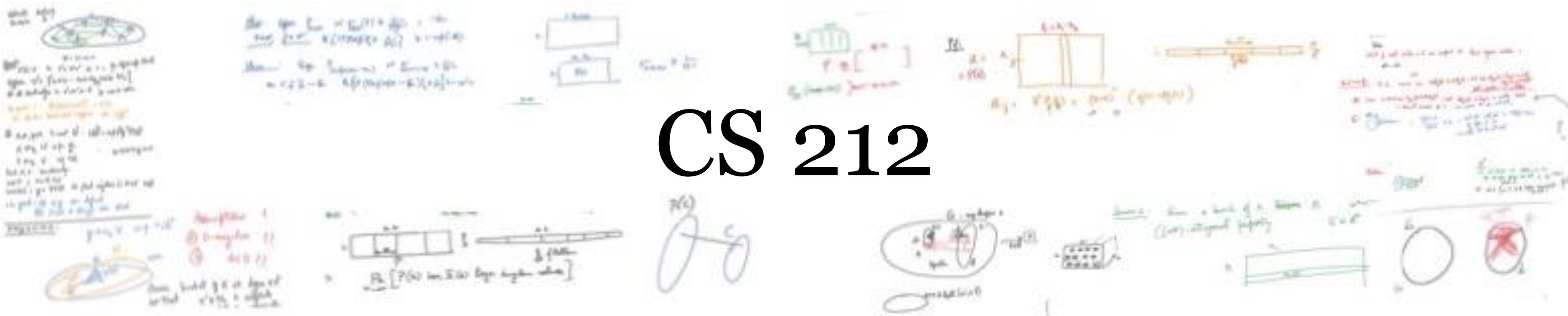


CS 212

Mathematical Foundations of Computer Science

Lecture 11: Probability Basics

Recap: Binomial Theorem

$x = b/a$

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$$

$$= a^n + \binom{n}{1} a^{n-1} b + \cdots + \binom{n}{i} a^{n-i} b^i + \cdots + \binom{n}{1} a b^{n-1} + b^n$$

Different forms:

polynomial

$P(x)$

$$\boxed{(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i}$$

sub. $-x$

$$(1-x)^n = \sum_{i=0}^{n} \binom{n}{i} (-1)^i x^i$$

Recap: Identities Using Binomial Theorem

$\sum_i i\binom{n}{i} = n2^{n-1}$

$\binom{n}{k}$= co-efficient of x^{k} in $(1+x)^n$

Fact: $1+\binom{n}{1}+\cdots+\binom{n}{k}+\cdots+\binom{n}{n-1}+\binom{n}{n} = 2^n$

$\binom{n}{0} + \binom{n}{1}x + \cdots + \binom{n}{k}x^k + \cdots + \binom{n}{n}x^n = (1+x)^n$

Substitute $x=1$

$$\binom{n}{0}+\binom{n}{2}+\cdots\, even\ terms + \cdots = \binom{n}{1}+\binom{n}{3}+\cdots\, odd\ terms + \cdots$$

Substituting $x=-1$

Substitute $x = 1$ in $(1-x)^n = \sum_{i=0}^{n}\binom{n}{i}(-1)^i x^i$

$$0 = (1-x)^n = \sum_{i=0}^{n}\binom{n}{i}(-1)^i$$

Coeff of (x^3) in $x^2(x^3-100x^2+10x-5) = x^5-100x^4+10x^3-5x^2$

$=$ Coeff of (x) in $x^3-100x^2+10x-5$

Counting using Binomial Theorem

Prove that $\binom{n-2}{k}+2\binom{n-2}{k-1}+\binom{n-2}{k-2}=\binom{n}{k}, \forall k\ s.t.\ 2\le k\le n-2$

Pf. RHS= $\binom{n}{k}$ = coefficient of x^k in polynomial P$(x)=(1+x)^n$

Note: $\binom{n-2}{k}$ = coeff. of x^k in $(1+x)^{n-2}$

$P(x) = (1+x)^n = (1+x)^{n-2}\times(1+x)^2 = (1+x)^{n-2}(1+2x+x^2)$

$P(x) = (1+x)^{n-2} + 2x(1+x)^{n-2} + x^2(1+x)^{n-2}$

Coeff of x^k in $P(x)$ = Coeff (x^k) in $(1+x)^{n-2}$ + Coeff (x^k) in $2x(1+x)^{n-2}$ + $\boxed{\text{Coeff}(x^k)}$ in $\boxed{x^2}(1+x)^{n-2}$

$= \binom{n-2}{k} + 2\cdot$ coeff (x^{k-1}) in $(1+x)^{n-2}$ + $\boxed{\text{Coeff}(x^{k-2})}$ in $(1+x)^{n-2}$

$=$

$\binom{n}{k}$ = RHS

$= \binom{n-2}{k} + 2\cdot\binom{n-2}{k-1} + \binom{n-2}{k-2}$ = LHS ∎

Probability Theory

Questions that we'll answer

- Estimating success/ failure of algorithms.
- Measure of Information
- Statistical Significance of Hypothesis
- What is the chance that a given system fails?

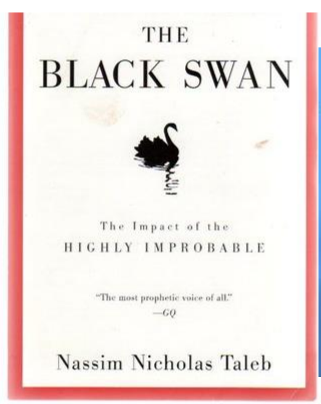
SHANNON ENTROPY

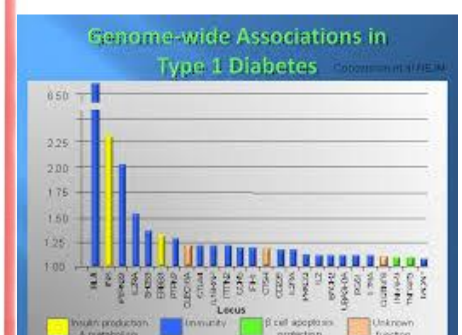
$$H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

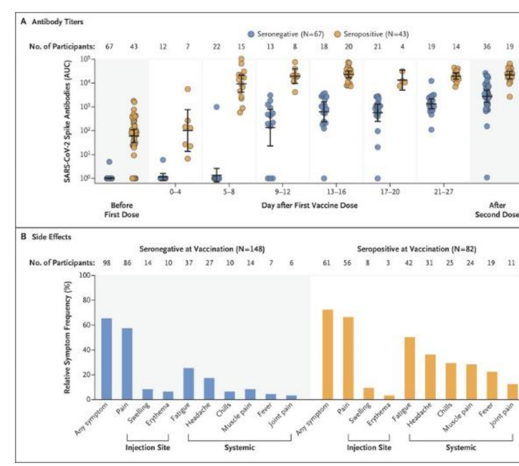
Weissman SCORE ™

$$W = \alpha \frac{r}{\overline{r}} \frac{\log \overline{T}}{\log T}$$

NOTE: r and T refer to the compression ratio and time-to-compress for the target algorithm, $\overline{r}$ and $\overline{T}$ refer to the same quantities for a standard universal compressor (e.g. gzip or FLAC), and α is a scaling constant. By normalizing by the performance of a standard compressor, we take away variation in compressive performance between types of data.



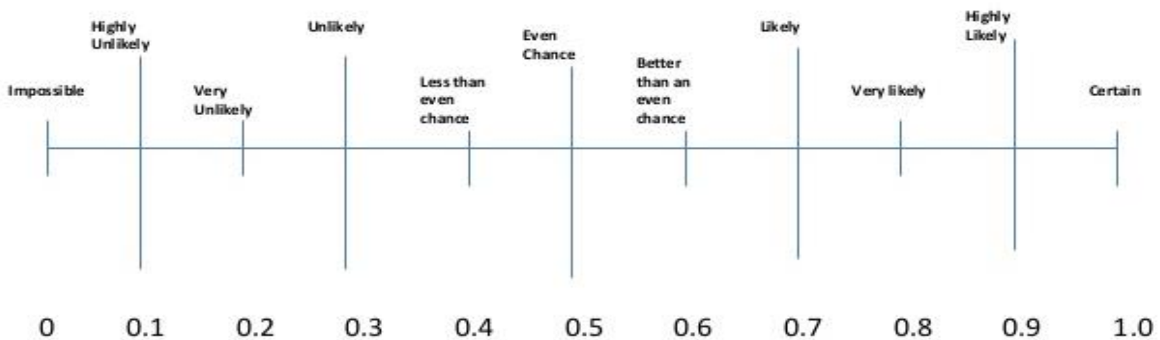




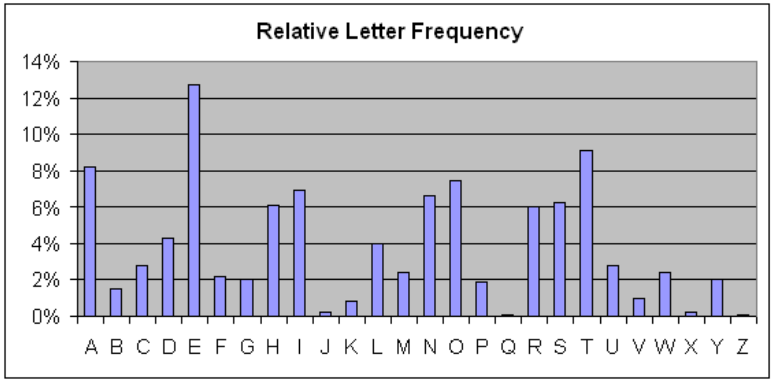
mRNA vaccines

What is Probability?

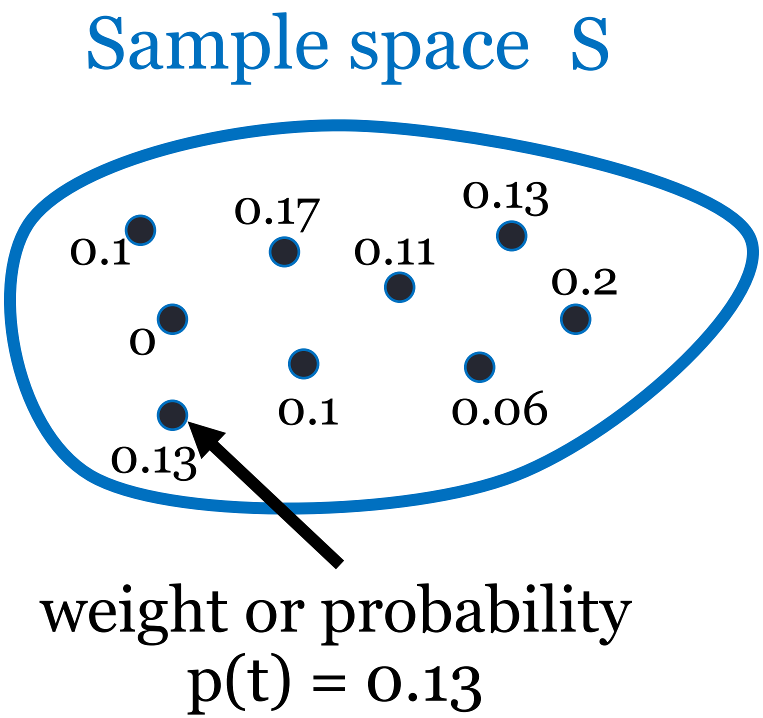
- Measuring uncertainty or chance



- Frequency of events



Probability Distribution



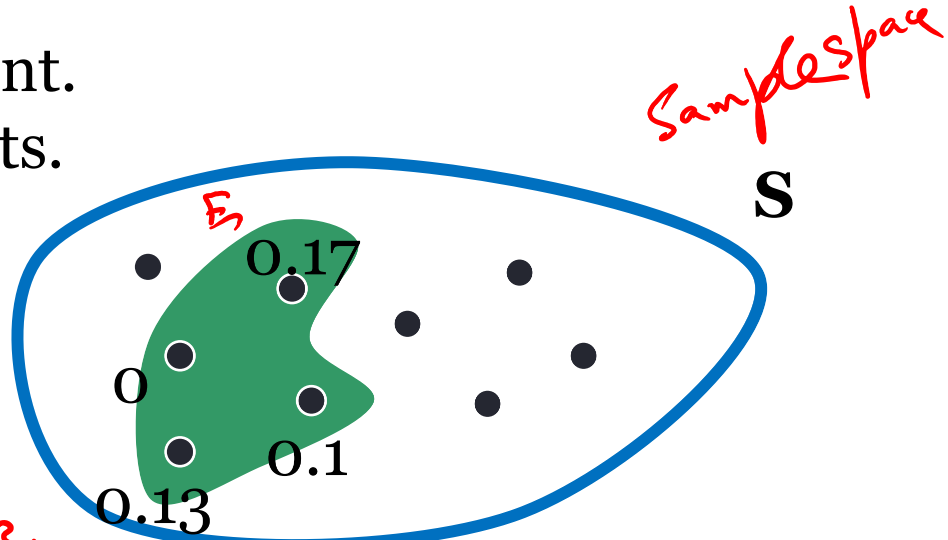
A (finite) probability distribution D is a finite set S of elements, where each element t in S has a non-negative real weight, proportion, or probability $p(t)$

Weights must satisfy: $\boxed{\sum_{t \in S} p(t) = 1}$

S is sample space , elements t ∈ S are called samples/atoms.

Events

- Any set $E \subseteq S$ is called an event.
- Elements called Atomic events.



$$\Pr_D[E] = \sum_{t \in E} p(t)$$

probability distribution

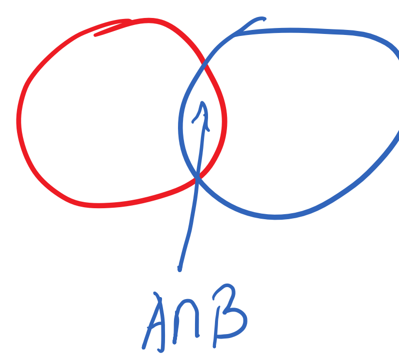
$\Pr_D[E] = 0.4$

$= 0 + 0.17 + 0.13 + 0.1$

Union of two events:

If A and B are events, then

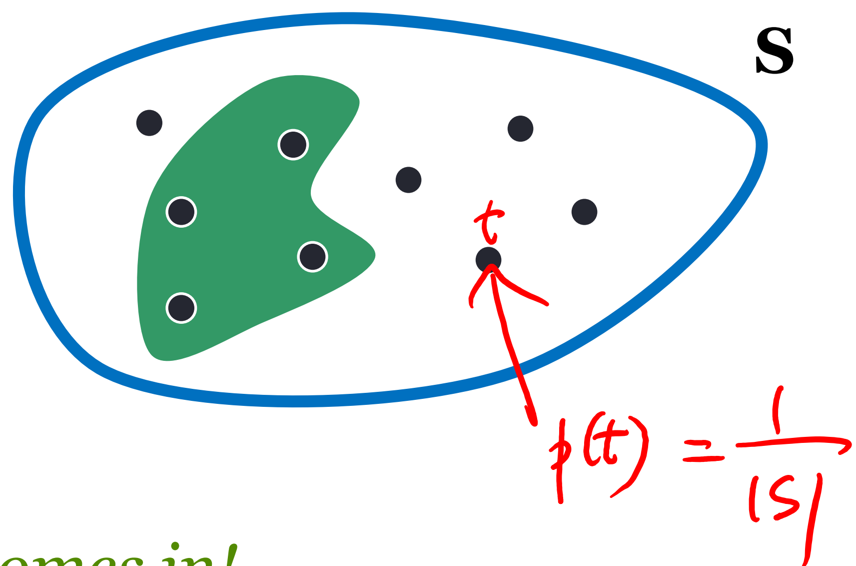
$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$$



Uniform Distribution

If each element (atomic event) has equal probability, the distribution is said to be uniform

$$\Pr_D[E] = \sum_{t \in E} p(t) = \frac{|E|}{|S|}$$



This is where Counting comes in!