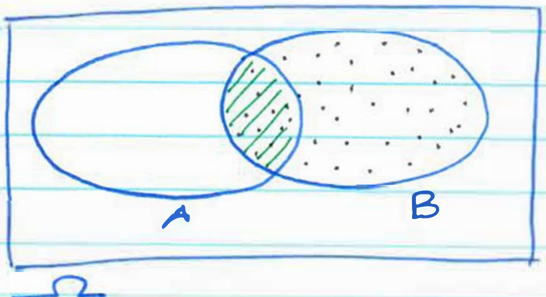
Conditional Probability

Conditional probability of an event A assuming another event M is

$$P(A \mid M) \triangleq \frac{P(A \cap M)}{P(M)} \qquad \left[\text{assuming that } P(M) \neq 0\right]$$



$$P(A \mid B) = \frac{\text{"green"}}{\text{"dots"}}$$

$\boxed{\text{Eg 1}}$: Toss a fair die, i.e probability of outcomes = 1/6.

What is the probability of observing a 6 given that we will observe an even number?

$$A = \{6\} \qquad M = \{2, 4, 6\}$$

$$P(A \mid M) = \frac{P(A \cap M)}{P(M)} = \frac{P(\{6\})}{P(\{2,4,6\})}$$

$$= \frac{1/6}{1/6 + 1/6 + 1/6} = \frac{1}{3}$$

Eg 2 We toss two fair dice in succession.

(a) What is the probability that the total exceeds 6?

(b) If the first dice shows 3, what is the conditional probability that the total exceeds 6?

Sample Space

$$\Omega = \left\{ \begin{matrix} (1,1), & (1,2) & \cdots, & (1,6) \\ (2,1), & (2,2) & \cdots, & (2,6) \\ \vdots & \vdots & & \vdots \\ (6,1), & (6,2) & \cdots, & (6,6) \end{matrix} \right\}$$

Total # of outcomes $= 36$

A = Event that the total exceeds 6

$$= \left\{ \begin{matrix} (1,6) \\ (2,5), (2,6) \\ (3,4), (3,5), (3,6) \\ \vdots \\ (6,1), (6,2), \cdots, (6,6) \end{matrix} \right\}$$

21 possible outcomes lead to event A

For part (a), $P(A) = \dfrac{21 \times (\cancel{1/36})}{36 \times (\cancel{1/36})} = \dfrac{21}{36} = \dfrac{7}{12}$

B = Event that the first dice shows 3

$$= \{(3,1), (3,2), \cdots, (3,6)\}$$

For part (b), we are interested in

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

$A \cap B = \{(3,4), (3,5), (3,6)\}$

$$\Rightarrow P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3 \times 1/36}{6 \times 1/36} = \frac{1}{2}$$

TOTAL PROBABILITY and BAYES' THEOREM

Total Probability Theorem:

If $U = [A_1, A_2, \dots, A_n]$ is a partition of $\underbrace{S}_{\text{sample space}}$,

then for any event B,

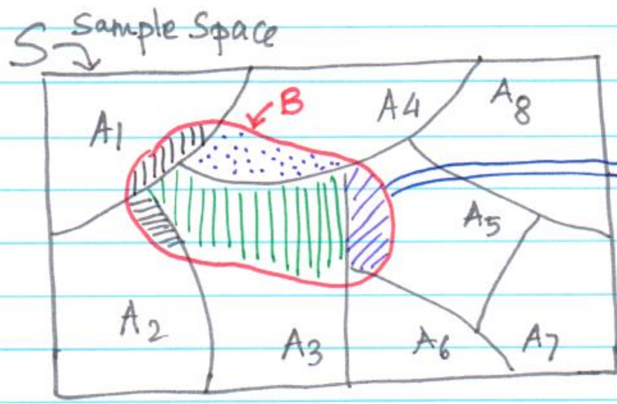
$$P(B) = P(B|A_1)P(A_1) + \dots + P(B|A_n)P(A_n)$$

Proof:

$$\begin{aligned} B &= B \cap S \\ &= B \cap (A_1 \cup A_2 \dots \cup A_n) \\ &= (B \cap A_1) \cup (B \cap A_2) \dots \cup (B \cap A_n) \end{aligned}$$

these are mutually exclusive

$$\Rightarrow P(B) = P(B \cap A_1) + \dots + P(B \cap A_n)$$
$$= P(B|A_1)P(A_1) + \dots + P(B|A_n)P(A_n)$$



Visual / Venn-diag. interpretation of Total Probability Theorem

Bayes' Theorem

$$P(B \cap A_i) = P(A_i|B)\,P(B) = P(B|A_i)\,P(A_i)$$

$$\Rightarrow \quad P(A_i|B) = \frac{P(B|A_i)\cdot P(A_i)}{P(B)}$$

$$\boxed{P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)\,P(A_i)}}$$

we often use the terms

"a priori" for $P(A_i)$ $\Big\} \rightarrow$ without knowing anything, what a priori is the prob of A_i?

and

"a posteriori" for $P(A_i|B)$ $\Big\} \rightarrow$ Given that event B has occured, what is prob of A_i?

Independence of Events

Two events A and B are independent if

$$P(A \cap B) = P(A)\,P(B)$$

or $\quad P(A \cap B) = P(A|B)\,P(B) = P(A)\,P(B)$

$\Rightarrow$ A & B are independent if $\boxed{P(A|B) = P(A)}$

V. Imp

n events $A_1, A_2, \dots, A_n$ are independent if any $k < n$ of them are idependent and

$$P(A_1 \cap A_2 \cdots \cap A_n) = P(A_1)\,P(A_2) \cdots P(A_n)$$

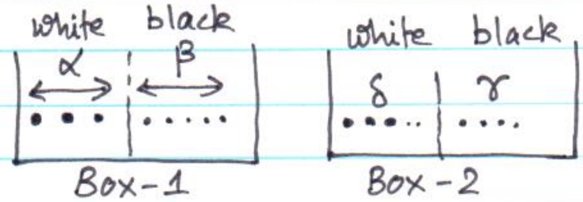
Eg: Events A_1, A_2 and A_3 are independent if

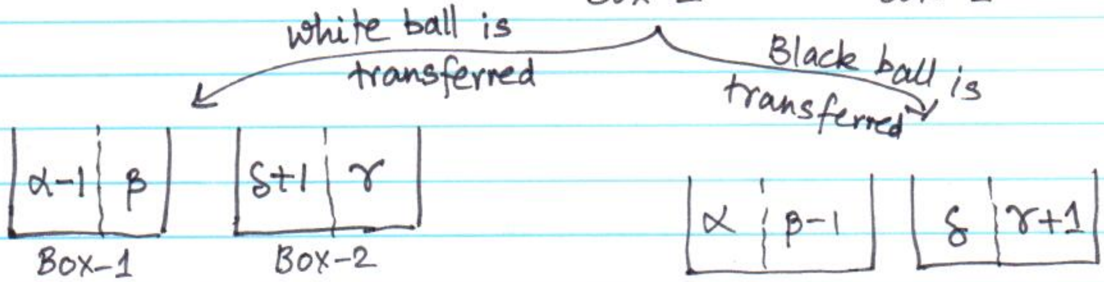
$$\left.\begin{aligned} P(A_1 \cap A_2) &= P(A_1)\,P(A_2) \\ P(A_1 \cap A_3) &= P(A_1)\,P(A_3) \\ P(A_3 \cap A_2) &= P(A_3)\,P(A_2) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1)\,P(A_2)\,P(A_3) \end{aligned}\right\}$$

All 4 conditions must be Valid

Example: (2.14 in "PP") We have two boxes, box 1 has α white balls and β black balls. Box 2 has δ white balls and γ black balls. We transfer one ball from box 1 to box 2. Next, a ball is drawn from box 2. What is the prob. that it will be white?

Solution: Original System





Let $W \rightarrow$ be the event that a white ball is transferred to box 2

$B \rightarrow$ black ball is transferred to Box 2.

$$P(W) = \frac{\alpha}{\alpha+\beta}$$

$$P(B) = \frac{\beta}{\alpha+\beta}$$

Recall that W and B are mutually exclusive events.

Quiz: Are W and B independent events?

Ans: No!! Why? $P(W|B) = \frac{P(W \cap B)}{P(B)} = 0$

However, for independence, we need $P(W|B) = P(A) \Rightarrow W$ and B are NOT independent.

Coming back to the Example, we are interested in the event

$$A = \{\text{white ball is drawn from second box}\}$$

Since W and B are mutually exclusive events and form a partition of the sample space, using Baye's theorem, we can write

$$P(A) = P(W)P(A|W) + P(B)P(A|B)$$

Easy to check that

$$P(A|W) = \frac{\delta+1}{\delta+\gamma+1} \quad \text{and} \quad P(A|B) = \frac{\delta}{\delta+\gamma+1}$$

$$\Rightarrow \quad P(A) = \left(\frac{\alpha}{\alpha+\beta}\right) \times \left(\frac{\delta+1}{\delta+\gamma+1}\right) + \left(\frac{\beta}{\alpha+\beta}\right)\left(\frac{\delta}{\delta+\gamma+1}\right).$$

~~**Example:** A cancer diagnostic device is 95% accurate. A person takes the test and the result is positive. It is also known that the apriori probability of cancer in the person's town is 0.02. What is the probability that he/she has cancer?~~

~~Let H → denote the set of healthy patients~~
~~C → " " cancer "~~

~~T = Correct be the event that device is accurate~~
~~= Incorr " " " inaccurate.~~

~~or $T = \begin{cases} H & \text{if device detects healthy} \\ C & \text{" " cancer} \end{cases}$~~

~~We are given $P(T=H|H) = 0.95$~~
~~$P(T=C|H) = 0.95$~~

Example : A cancer diagnostic device is 95 % accurate. A person takes the test & the result is positive. It is given that apriori probability of cancer in the person's town is 0.02. What is the prob that the person has cancer?

Solution: Let T denote the outcome of the diagnostic device.

$$T = \begin{cases} H & \text{(healthy)} \\ C & \text{(cancer)} \end{cases}$$

We are given that
$$\left.\begin{aligned} P(T=H \mid H) &= 0.95 \\ \text{and} \quad P(T=C \mid C) &= 0.95 \end{aligned}\right\} \text{95 \% accurate}$$

The device detects cancer, i.e $T=C$.

$\boxed{\text{We want to find} \quad P(C \mid T=C)}$

Using Baye's rule:
$$P(C \mid T=C) = \frac{P(C \cap T=C)}{P(T=C)}$$

$$= \frac{P(C) \cdot P(T=C \mid C)}{P(T=C)}$$

$$= \frac{(0.02) \times (0.95)}{P(T=C)}$$

$$= \frac{(0.02) \times 0.95}{P(C)\, P(T=C \mid C) + P(H)\, P(T=C \mid H)}$$

$$= \frac{(0.02) \times 0.95}{(0.02) \times 0.95 + (0.98) \times 0.05}$$

$$P(C \mid T=C) = \boxed{0.278}$$

Lecture 2-continued

Evaluation metrics

* Consider an experiment in which the outcomes are labeled either as positive (p) or negative (n).

* Example : E-mails in your Inbox

Spam (p) — Not-Spam (n)

Group of Patients

Have a disease (p) — Do NOT have a disease (n).

* For the Spam Example, your browser has a "Spam-Classifier" which classifies each email as either Spam (p) or No-Spam (n).
It could use some "features" of the e-mail, such as content, origin etc...

There can be 4 possible scenarios for a "Spam Classifier"

TRUE POSITIVE (TP) →	Classifier predicts Spam (P)	and	Email is a Spam (P)
FALSE POSITIVE (FP) →	Classifier predicts Spam (P)	and	Email is NOT a Spam (N)
TRUE NEGATIVE (TN) →	Classifier predicts Not Spam (N)	and	Email is NOT a Spam (N)
FALSE NEGATIVE (FN) →	Classifier predicts NOT Spam (N)	and	Email is a Spam (P)

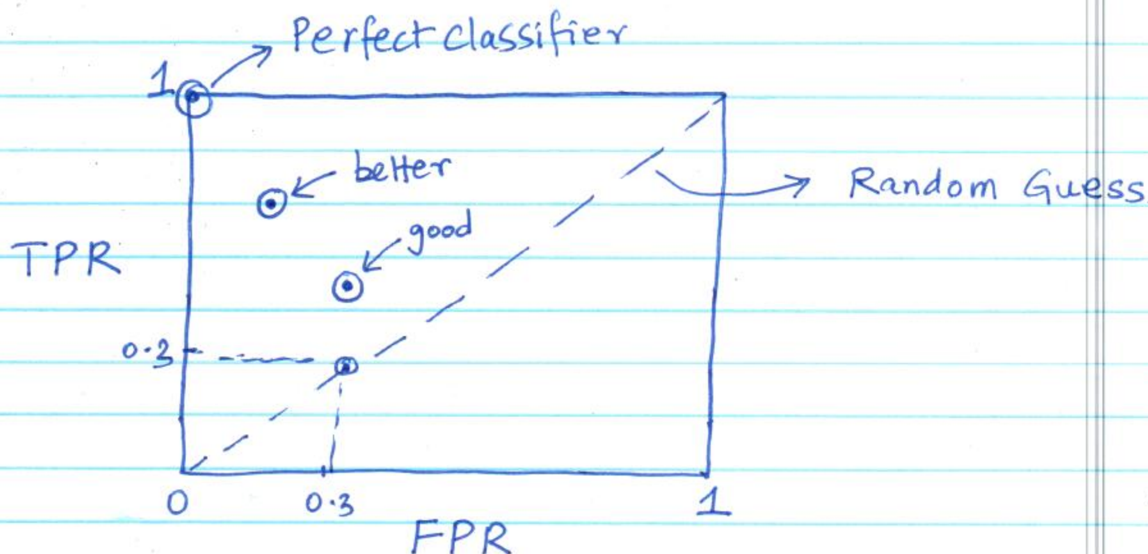
False Positive Rate = Prob(Decision = P | Truth = N)
(FPR)

True Positive Rate = Prob (Decision = P | Truth = P)
(TPR)

Ideally, we want TPR → High (close to 1)

FPR → Small (close to 0)

Receiver Operating characteristic (or ROC)



For a "random Guess", i.e a Decision which is taken independently from the Truth;

$$P(\text{Decision} = P \mid \text{Truth} = N) = P(\text{Decision} = P)$$

and $P(\text{Decision} = P \mid \text{Truth} = P) = P(\text{Decision} = P)$

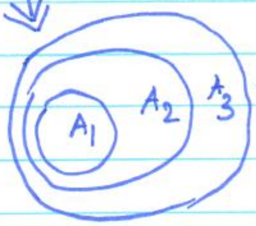
$$\Rightarrow \boxed{TPR = FPR = P(\text{Decision} = P)}$$ → for a random Guess

Generalized Additive Laws

* If $A_1 \subset A_2 \subset A_3 \ldots,$ then

$$P\left(\bigcup_k A_k\right) = \lim_{n \to \infty} P(A_n)$$

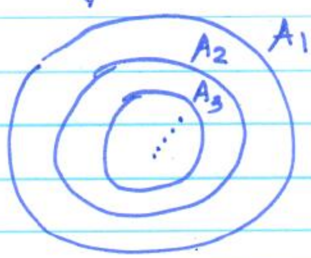
*



$\Rightarrow$ A_i's are "increasing" set of events

* If $A_1 \supset A_2 \supset A_3 \ldots,$ then

$$P\left(\bigcap_k A_k\right) = \lim_{n \to \infty} P(A_n)$$



$\Rightarrow$ A_i's are "decreasing" set of events

$\boxed{\text{UNION BOUND}}$

For any sequence of events $A_1, A_2, \ldots,$

$$P\left(\bigcup_k A_k\right) \le \sum_k P(A_k)$$

Let us prove union bound for two events A_1 and A_2:

$$P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$$

* Recall : $A_1 \cup A_2 = A_1 \cup (A_2 \cap \bar{A_1})$.

* Now note that A_1 and $A_2 \cap \bar{A_1}$ are mutually exclusive events.

$$\Rightarrow P(A_1 \cup A_2) = P(A_1 \cup (A_2 \cap \bar{A_1}))$$

$$= P(A_1) + P(A_2 \cap \bar{A_1})$$

$$\leq P(A_1) + P(A_2)$$

why? because $A_2 \cap \bar{A_1} \subseteq A_2$

$$\Rightarrow \boxed{P(A_1 \cup A_2) \leq P(A_1) + P(A_2)}$$