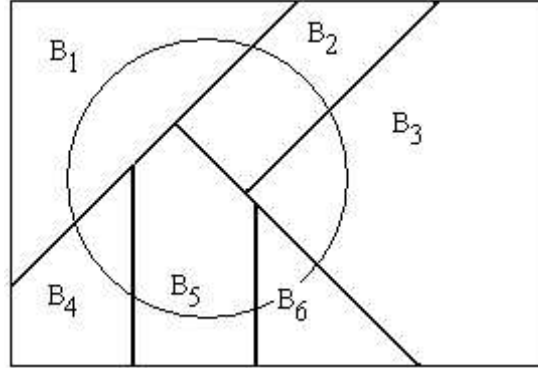


Bayes' Formula.



Partition B_1, \dots, B_k
 $\bigcup_{i=1}^k B_i = S, B_i \cap B_j = \emptyset$ for $i \neq j$
 $\mathbb{P}(A) = \sum_{i=1}^k \mathbb{P}(AB_i) = \sum_{i=1}^k \mathbb{P}(A|B_i)\mathbb{P}(B_i)$ - total probability.

Example: In box 1, there are 60 short bolts and 40 long bolts. In box 2, there are 10 short bolts and 20 long bolts. Take a box at random, and pick a bolt. What is the probability that you chose a short bolt?

B_1 = choose Box 1.
 B_2 = choose Box 2.

$$\mathbb{P}(\text{short}) = \mathbb{P}(\text{short}|B_1)\mathbb{P}(B_1) + \mathbb{P}(\text{short}|B_2)\mathbb{P}(B_2) = \frac{60}{100}(\frac{1}{2}) + \frac{10}{30}(\frac{1}{2})$$

Example:

Partitions: B_1, B_2, \dots, B_k and you know the distribution.

Events: A, A, ..., A and you know the $\mathbb{P}(A)$ for each B_i

If you know that A happened, what is the probability that it came from a particular B_i ?

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\mathbb{P}(A|B_1)\mathbb{P}(B_1) + \dots + \mathbb{P}(A|B_k)\mathbb{P}(B_k)} : \text{Bayes's Formula}$$

Example: Medical detection test, 90% accurate.

Partition - you have the disease (B_1), you don't have the disease (B_2)

The accuracy means, in terms of probability: $\mathbb{P}(\text{positive}|B_1) = 0.9, \mathbb{P}(\text{positive}|B_2) = 0.1$

In the general public, the chance of getting the disease is 1 in 10,000.

In terms of probability: $\mathbb{P}(B_1) = 0.0001, \mathbb{P}(B_2) = 0.9999$

If the result comes up positive, what is the probability that you actually have the disease? $\mathbb{P}(B_1|\text{positive})$?

$$\begin{aligned} \mathbb{P}(B_1|\text{positive}) &= \frac{\mathbb{P}(\text{positive}|B_1)\mathbb{P}(B_1)}{\mathbb{P}(\text{positive}|B_1)\mathbb{P}(B_1) + \mathbb{P}(\text{positive}|B_2)\mathbb{P}(B_2)} \\ &= \frac{(0.9)(0.0001)}{(0.9)(0.0001) + (0.1)(0.9999)} = 0.0009 \end{aligned}$$

The probability is still very small that you actually have the disease.

Example: Identify the source of a defective item.

There are 3 machines: M_1, M_2, M_3 . $\mathbb{P}(\text{defective})$: 0.01, 0.02, 0.03, respectively.

The percent of items made that come from each machine is: 20%, 30%, and 50%, respectively.

Probability that the item comes from a machine: $P(M_1) = 0.2, P(M_2) = 0.3, P(M_3) = 0.5$

Probability that a machine's item is defective: $P(D|M_1) = 0.01, P(D|M_2) = 0.02, P(D|M_3) = 0.03$

Probability that it came from Machine 1:

$$P(M_1|D) = \frac{P(D|M_1)P(M_1)}{P(D|M_1)P(M_1) + P(D|M_2)P(M_2) + P(D|M_3)P(M_3)}$$

$$= \frac{(0.01)(0.2)}{(0.01)(0.2) + (0.02)(0.3) + (0.03)(0.5)} = 0.087$$

Example: A gene has 2 alleles: A, a. The gene exhibits itself through a trait with two versions.

The possible phenotypes are "dominant," with genotypes AA or Aa, and "recessive," with genotype aa.

Alleles travel independently, derived from a parent's genotype.

In a population, the probability of having a particular allele: $\mathbb{P}(A) = 0.5, \mathbb{P}(a) = 0.5$

Therefore, the probabilities of the genotypes are: $\mathbb{P}(AA) = 0.25, \mathbb{P}(Aa) = 0.5, \mathbb{P}(aa) = 0.25$

Partitions: genotypes of parents: (AA, AA), (AA, Aa), (AA, aa), (Aa, Aa), (Aa, aa), (aa, aa).

Assume pairs match regardless of genotype.

Parent genotypes	Probabilities	Probability that child has dominant phenotype
(AA, AA)	$\frac{1}{16}$	1
(AA, Aa)	$2 \times (\frac{1}{4})(\frac{1}{2}) = \frac{1}{4}$	1
(AA, aa)	$2 \times (\frac{1}{4})(\frac{1}{4}) = \frac{1}{8}$	1
(Aa, Aa)	$(\frac{1}{2})(\frac{1}{2}) = \frac{1}{4}$	$\frac{3}{4}$
(Aa, aa)	$2 \times (\frac{1}{2})(\frac{1}{4}) = \frac{1}{4}$	$\frac{1}{2}$
(aa, aa)	$\frac{1}{16}$	0

If you see that a person has dark hair, predict the genotypes of the parents:

$$P((AA, AA)|A) = \frac{1 \times \frac{1}{16}}{\frac{1}{16}(1) + \frac{1}{4}(1) + \frac{1}{8}(1) + \frac{1}{4}(\frac{3}{4}) + \frac{1}{4}(\frac{1}{2}) + \frac{1}{16}(0)} = \frac{1}{12}$$

You can do the same computation to find the probabilities of each type of couple.

Bayes's formula gives a prediction inside the parents that you aren't able to directly see.

Example: You have 1 machine.

In good condition: defective items only produced 1% of the time. $P(\text{in good condition}) = 90\%$

In broken condition: defective items produced 40% of the time. $P(\text{broken}) = 10\%$

Sample 6 items, and find that 2 are defective. Is the machine broken?

This is very similar to the medical example worked earlier in lecture:

$P(\text{good}|2 \text{ out of } 6 \text{ are defective}) =$

$$= \frac{P(2 \text{ of } 6 | \text{good})P(\text{good})}{P(2 \text{ of } 6 | \text{good})P(\text{good}) + P(2 \text{ of } 6 | \text{broken})P(\text{broken})}$$

$$= \frac{\binom{6}{2}(0.01)^2(0.99)^4(0.9)}{\binom{6}{2}(0.01)^2(0.99)^4(0.9) + \binom{6}{2}(0.4)^2(0.6)^4(0.1)} = 0.04$$

** End of Lecture 7