Total Probability Theorem

- **Claim.** If $B \subset A$ then $\Pr(B) \leq \Pr(A)$.
- **Proof.** $A = B \cup (A \setminus B)$, so

$$\Pr(A) = \Pr(B) + \Pr(A \setminus B) \geq \Pr(B).$$

- **Def.** The events $A_1, \ldots, A_n$ form a partition of the sample space Ω if
  1. A_i are mutually exclusive: $A_i \cap A_j = \emptyset$ for $i \neq j$.
  2. $A_1 \cup \ldots \cup A_n = \Omega$.
- **Total Probability Theorem.** Let $A_1, \ldots, A_n$ be a partition of Ω. For any event B,

$$\Pr(B) = \sum_{j=1}^{n} \Pr(A_j) \Pr(B|A_j).$$

- **Proof.** $B = \cup(B \cap A_j)$ (disjoint union), so

$$\Pr(B) = \sum_{j=1}^{n} \Pr(B \cap A_j).$$

  The theorem follows from $\Pr(B \cap A_j) = \Pr(A_j) \Pr(B|A_j)$.
- The latter holds for A_j with $\Pr(A_j) = 0$ if we define $\Pr(A_j) \Pr(B|A_j) := 0$ since then $P(B \cap A_j) = 0$

Example

- In a certain county
  - 60% of registered voters are Republicans
  - 30% are Democrats
  - 10% are Independents.
- When those voters were asked about increasing military spending
  - 40% of Republicans opposed it
  - 65% of the Democrats opposed it
  - 55% of the Independents opposed it.
- What is the probability that a randomly selected voter in this county opposes increased military spending?

- $\Omega = \{$registered voters in the county$\}$
- $R = \{$registered republicans$\}$, $\Pr(R) = 0.6$
- $D = \{$registered democrats$\}$, $\Pr(D) = 0.3$
- $I = \{$registered independents$\}$, $\Pr(I) = 0.1$
- $B = \{$registered voters opposing increased military spending$\}$
- $\Pr(B|R) = 0.4$, $\Pr(B|D) = 0.65$, $\Pr(B|I) = 0.55$.

By the total probability theorem:

$$\begin{aligned}\Pr(B) &= \Pr(B|R)\Pr(R) + \Pr(B|D)\Pr(D) + \Pr(B|I)\Pr(I) \\ &= (0.4 \cdot 0.6) + (0.65 \cdot 0.3) + (0.55 \cdot 0.1) = 0.49.\end{aligned}$$

Bayes' Theorem

- **Bayes Theorem.** Let $A_1, \ldots, A_n$ be a partition of Ω. For any event B

$$\Pr(A_i|B) = \frac{\Pr(A_i)\Pr(B|A_i)}{\Sigma_{j=1}^{n}\Pr(A_j)\Pr(B|A_j)}.$$

- **Proof.**

$$\begin{aligned}\Pr(A_i|B) &= \frac{\Pr(A_i \cap B)}{\Pr(B)} \\ &= \frac{\Pr(A_i)\Pr(B|A_i)}{\Sigma_{j=1}^{n}\Pr(A_j)\Pr(B|A_j)}.\end{aligned}$$

- Example.
  - · A registered voter from our county writes a letter to the local paper, arguing against increased military spending. What is the probability that this voter is a Democrat?
  - · Presumably that is $\Pr(D|B)$, so by Bayes' theorem:

$$\begin{aligned}\Pr(D|B) &= \frac{0.65 \cdot 0.3}{(0.4 \cdot 0.6) + (0.65 \cdot 0.3) + (0.55 \cdot 0.1)} \\ &= \frac{0.195}{0.49} \approx 0.398.\end{aligned}$$

AIDS

- Just for the heck of it Bob decides to take a test for AIDS and it comes back positive.
- The test is 99% effective (1% FP and FN).
- Suppose 0.3% of the population in Bob's "bracket" has AIDS.
- What is the probability that he has AIDS?
- · $\Omega = \{$all the people in Bob's bracket$\}$.

  · $A_1 = \{$people in Ω with AIDS$\}$, $\Pr(A_1) = 0.003$

  · $A_2 = \{$people in Ω without AIDS$\}$, $\Pr(A_2) = 0.997$

  · $B = \{$people in Ω who would test positive$\}$

  · $\Pr(B|A_1) = .99$ and $Pr(B|A_2) = .01$

  · By Bayes' rule

$$\begin{aligned}\Pr(A_1|B) &= \frac{0.99 \cdot 0.003}{(0.99 \cdot 0.003) + (0.01 \cdot 0.997)} \\ &\approx \frac{0.003}{0.003 + 0.01} \\ &\approx 0.23.\end{aligned}$$

Random Variables

- What is the probbaility that:
  - in k out of n flips a coin will land on its head
  - k out of n randomly selected people will test positive to AIDS
  - k out of n rolled die will show an even number
  - a student will correctly guess k out of n multiple choice questions with a fixed number of choices.
- Mathematicians realized that often a problem in probability can be summarized in terms of a "random variable".
- **Definition.** A *random variable* on a sample space Ω is a function from Ω to $\mathbb{R}$.
- **Example.**A coin is flipped 10 times.
  - Ω is the space of all sequences of H and T of length 10.
  - Let X count the number of heads.
  - For example $X(HHTHHTTHTH) = 6$.
- Can you think of an analogous X for the other problems mentioned above?

- **Example.** What is the probability that a randomly chosen person in the class will weigh more than 160 lbs.?
    - A natural random variable in this case is the weight of the selected student.
- What is the difference between the range of values our two random variables can attain?

Probability Distributions

- There is a natural probabilistic structure induced on a random variable X defined on Ω:
  - The set $\{\omega \in \Omega : X(\omega) = c\}$ is an event.
  - So we can ask for

$$\Pr(X = c) = \Pr(\{\omega \in \Omega : X(\omega) = c\}).$$

  Example. A biased coin ($\Pr(H) = 2/3$) is flipped twice.

  - Let X count the number of heads:

  $\Pr(X = 0) = \Pr(\{TT\}) = (1/3)^2 = 1/9.$
  $\Pr(X = 1) = \Pr(\{HT, TH\}) = 2 \cdot 1/3 \cdot 2/3 = 4/9.$
  $\Pr(X = 2) = \Pr(\{HH\}) = (2/3)^2 = 4/9.$

- Similarly we might be interested in:

$$\Pr(X \le c) = \Pr(\{\omega \in \Omega : X(\omega) \le c\},$$

  and more generally, for any $T \subset \mathbb{R}$:

$$\Pr(X \in T) = \Pr(\{\omega \in \Omega : X(\omega) \in T\}.$$

- In our coin example,

  $\Pr(X \le 1) = \Pr(\{TT, HT, TH\}) = 1/9{+}4/9 = 5/9.$