1. ANNOUNCEMENTS

First midterm: Jan 31, Tuesday

Covers up to the Thursday before the test day. Does not cover lecture material on day of test.

2. LAW OF TOTAL PROBABILITY

Recall conditional probability:

\[ P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \]

What is the relationship between \( P(A|B) \) and \( P(B|A) \)?

\[ \frac{P(A|B)}{P(B|A)} = \frac{P(A)}{P(B)} \]

Consider \( \mathcal{R}_i \) for \( i = 1, \ldots, n \). That is, \( \Omega \) (the universe) is made up of \( n \) regions. Then, \( \mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n \) partition \( \Omega \) if 1) \( \mathcal{R}_i \cap \mathcal{R}_j = \emptyset, i \neq j \) (mutually disjoint) 2) \( \mathcal{R}_1 \cup \mathcal{R}_2 \cup \ldots \cup \mathcal{R}_n = \Omega \) (collectively exhaustive).

Assume \( P(\mathcal{R}_i) \geq 0 \) for all \( i = 1, \ldots, n \).

Let \( A \) belong to \( \Omega \). Set \( B_i = A \cap \mathcal{R}_i \). Then \( B_i \cap B_j = \emptyset \) if \( i \neq j, i = 1, \ldots, n \) and \( A = B_1 \cup B_2 \cup \cdots \cup B_n \), i.e., the \( B_i \)'s partition \( A \).

\[ P(A) = P(B_1) + P(B_2) + \cdots + P(B_n), \text{ with } P(B_i) = P(\mathcal{R}_i)P(A|\mathcal{R}_i) = P(A \cap \mathcal{R}_i). \]

Therefore, the law of total probability states that

\[ P(A) = P(\mathcal{R}_1)P(A|\mathcal{R}_1) + P(\mathcal{R}_2)P(A|\mathcal{R}_2) + \cdots + P(\mathcal{R}_n)P(A|\mathcal{R}_n). \]

2.1. Example. Imagine there are five channels: Channel 1 sells 20% of the tickets with 0.4 profitable tickets; Channel 2 sells 30% of the tickets with 0.6 profitable tickets; Channel 3 sells 10% of the tickets with 0.2% profitable tickets; Channel 4 sells 15% of the tickets with 0.8 profitable tickets and Channel 5 sells 25% of the tickets with 0.9 profitable tickets.

The percentages define a partition of the sample space. The decimals define the conditional probabilities.

\( A = \) event that ticket is profitable

\( \mathcal{R}_i = \) event sold through Channel \( i \)

\[ P(A) = P(\mathcal{R}_1)P(A|\mathcal{R}_1) + \cdots + P(\mathcal{R}_5)P(A|\mathcal{R}_5) \]

\[ = 0.2 \times 0.4 + 0.3 \times 0.6 + \cdots + 0.25 \times 0.9. \]
3. Independent events

Two events, $A$ and $B$, are independent IF $P\{A \cap B\} = P\{A\}P\{B\}$.

$P\{A \cap B\} = P\{A|B\}P\{B\} = P\{A\}P\{B\}$ if $A$ and $B$ are independent.

Compare to mutually exclusive (which means strongly dependent).

Independent does not imply transitive, i.e. $A$, $B$ independent and $B$, $C$ independent does not imply $A$, $C$ are independent. However, it does imply that $\bar{A}$, $B$ are independent AND $A$, $\bar{B}$ are independent AND $\bar{A}$, $\bar{B}$ are independent.

Notes: $A \cap B$ and $\bar{A} \cap B$ are mutually exclusive AND $B = (A \cap B) \cup (\bar{A} \cap B)$, therefore

$P\{B\} = P\{A \cap B\} + P\{\bar{A} \cap B\} = P\{A\}P\{B\} + P\{\bar{A} \cap B\}$. We assume $A$ and $B$ are independent. So by definition these two events must be independent.

To show 1000 events are mutually independent, we must show any intersection of any number of events must be independent.

$A_1, A_2, \ldots, A_n$ are mutually independent $\iff$ for each set of $k$ distinct numbers (with $k \in [2, n]$), then

$P\{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\} = P\{A_{i_1}\}P\{A_{i_2}\} \cdots P\{A_{i_k}\}$.

Mutually exclusive $P\{\text{unions}\} = \sum P_i$ (sums)

Mutually independent $P\{\text{intersections}\} = \Pi P_i$ (products)

3.1. Example. $\Omega = \{(i, j)|i, j \in [1, 6]\}$ Each sample point has probability 1/36 because the dice are fair.

$A = \{ \text{die 1 results in a } 1, 2, 3 \}$.

$B = \{ \text{die 1 results in } 3, 4, 5 \}$.

$C = \{ \text{sum of two faces is } 9 \} = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$.

Recall: In order for multiple events to be independent, all possibilities must be independent.

$P\{C\} = \frac{1}{9} = \frac{4}{36}$. $P\{A\} = \frac{1}{2}$. $P\{B\} = \frac{1}{2}$.

$A \cap B = \{(3, 1), (3, 2), \ldots, (3, 6)\}$.

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

$P\{A \cap B \cap C\} = \frac{1}{36}$. $P\{A\}P\{B\}P\{C\} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{9} = \frac{1}{36}$.

$P\{B \cap C\} = \frac{3}{36} = \frac{1}{12} \neq P\{B\}P\{C\} = \frac{1}{2} \cdot \frac{1}{9}$.

Thus $A$, $B$, $C$ are not independent.
4. Bayes Theorem

$A_1, A_2, \ldots, A_n$ mutually exclusive and mutually exhaustive (i.e. They partition $\Omega$).

Bayes Theorem is that

$$P\{A_i|A\} = \frac{P\{A_i\} P\{A|A_i\}}{\sum_{j=1}^{n} P\{A_j\} P\{A|A_j\}} = \frac{P\{A_i \cap A\}}{P\{A\}}.$$

$P\{A_i\}$ “A priori”

$P\{A_i|A\}$ “A posteriori”

4.1. Example. 80% of programs written in $C$.

20% of programs written in $J$.

20% of $C$ programs compile OK on 1st try = $P\{E|C\}$.

60% of $J$ programs compile OK on 1st try = $P\{E|J\}$.

Event $E$ = \{ program compiled OK \}.

Event $C$ = \{ program written in $C$ \}.

Event $J$ = \{ program written in $J$ \}.

$P\{C|E\} = \frac{P\{C \cap E\}}{P\{E\}} = \frac{P\{C\} P\{E|C\}}{\pi}$, with

$E = (C \cap E) \cup (J \cap E)$ thus $(\ast) = P\{C\} P\{E|C\} + P\{J\} P\{E|J\}$.

5. Performance criteria

$D$, member of a class

$S$, test indicates that individual is a member

$\eta = P\{S|D\}$ “sensitivity”

$\theta = P\{\bar{S}|\bar{D}\}$ “specificity”

Want $\eta$ and $\theta$ to be close to 1 for a good test.

$\pi = P\{D\}$: When this number is small, you get lots of false positives.

$P\{S|\bar{D}\} = 1 - P\{\bar{S}|\bar{D}\} = 1 - \theta$.

“Predictive value of a positive test” = $P\{D|S\}$. Want it to be 1.
\[
P\{D|S\} = \frac{P\{D\}P\{S|D\}}{P\{D\}P\{S|D\} + P\{D\}P\{S|D\}} \\
= \frac{\pi \eta}{\pi \eta + (1 - \pi)(1 - \theta)}.
\]

5.1. **Example.** Let \( \pi = 0.0001 \).

\( \eta = 0.977 \) will correctly identify defect that is present

\( \theta = 0.926 \) if there is a defect, with probability 0.926, then it will say so.

100,000 chips

Expect 10 defectives

9.77 correctly diagnosed

Expect 7399 false positives.