Lecture 2

David Armstrong

STATS 67 - UCI

Probability

- Let the events $A_1, A_2, ..., A_M$ form a partition of the sample space S. Thus $A_1 \cup A_2 \cup ... \cup A_M = S$.

- The *law of total probability* states that if $A_1, A_2, ..., A_M$ form a partition then $P(B) = P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_M) = \sum_{i=1}^{M} P(B \cap A_i)$.

Example:

- Let B be the event that you flip a coin and it lands on head.

- Let A_1 denote the event you roll a die and it lands on 1, A_2 be the event it lands on 2,..., and A_6 be the event it lands on 6.

- What is the sample space if you roll one die and flip a coin at the same time?

- What is the probability if each possible event?

- Therefore:
  $P(B) = P(A_1 \cap B) + P(A_2 \cap B) + ... + P(A_6 \cap B) = \frac{1}{12} + ... + \frac{1}{12} = \frac{6}{12} = 0.5$

Probability

Example:

	D_A	D_B	D_C	D_D	D_E	D_F	Total
G_M	0.182	0.124	0.072	0.092	0.042	0.082	0.595
G_F	0.024	0.006	0.131	0.083	0.087	0.075	0.405
Total	0.206	0.129	0.203	0.175	0.129	0.158	1.000

- This is a *contingency* table of gender and department applied to (by proportions).
- Let G denote the gender of an applicant (G_M is male and G_F is female).
- Let D denote the department applied to (D_A is department A, D_B is department B,..., and D_F is department F.
- What is the probability of being a male and applying for department A?
- What is the probability of being a male?
- What is the probability of applying to department F?

Conditional Probability

- Now we come to what is known as a *conditional* probability.

- This is the probability of an event occurring given that another event is already known to have occurred.

- Notationally this is stated as $A|B$, which is meant to define the event A occurring conditional (or given) on B having occurred.

- The event B has occurred and been observed, what remains random and yet to be realized is event A.

- $P(A|B)$ signifies the probability of event A occurring, conditional on the fact (or given) that event B has occurred.

Example

- If A is the event someone drinks coffee regularly and B is the event someone stays up late regularly, then $P(A|B)$ will define the probability someone drinks coffee regularly given that they stay up late.

- $P(B|A)$ will define the probability that someone stays up late regularly given that they drink coffee regularly.

  - What does $P(A|B^c)$ represent?

- $P(A)$ is the probability that someone drinks coffee regularly. Remember this is the unconditional probability of an event.

  - It is the probability of A occurring marginalized over the outcomes of event B.

Conditional Probability

To compute conditional probabilities, a theorem knowns as Baye's theorem is used.

- $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ and $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$

  Using the formula, one can see that:

$$P(A \cap B) = P(A|B)P(B) \qquad \text{and} \qquad P(A \cap B) = P(B|A)P(A).$$

- Remember that if A has two outcomes, then $P(B) = P(A \cap B) + P(A^c \cap B)$.
  - In general $P(B) = \sum_{i=1}^{M} P(B \cap A_i)$.

Conditional Probability: Independent Events

If the knowledge that B occurred has no bearing on the likelihood that A will occur (similarly if knowing A has no bearing on B occurring), this means that A and B are *independent*.

- Notationally this is $P(A|B) = P(A)$ (and also that $P(B|A) = P(B)$).
- As a result, $P(A \cap B) = P(A)P(B)$
  - The probability of A and B occurring is equal to the probability of A occurring times the probability of B occurring.

Example - Say you flip a coin twice. Let A be the event the first flip is heads and B be the event the second flip is heads.

- What is the probability of the second flip being a heads, given the first flip is a heads?
- Notationally this is $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(A \cap B)}{P(A \cap B) + P(A \cap B^c)}$.
  - Calculate $P(B|A)$
- What is the probability the second flip ends in a heads?
- The events A and B are independent. Knowing that A has occurred does not change the likelihood that B will occur.

Conditional Probability

Example:

	D_A	D_B	D_C	D_D	D_E	D_F	Total
G_M	0.182	0.124	0.072	0.092	0.042	0.082	0.595
G_F	0.024	0.006	0.131	0.083	0.087	0.075	0.405
Total	0.206	0.129	0.203	0.175	0.129	0.158	1.000

- What is the probability that someone applied to department C given that they were female?

- What is the probability that someone is male given they applied to department F?

- What is the probability that an applicant was male?

- Compare the last two probabilities. Are gender and department applied to independent events?.

- If they were independent, then $P(G_M) = P(G_M|D_F)$ (in fact for all departments $D_A, ..., D_F$).

- Also $P(D_C) = 0.203$ but $P(D_C|G_F) = 0.323$.

Conditional Probability: Independent Events

Extending the concept of independent events to the scenario of events $A_1, A_2, ..., A_M$.

- If events $A_1, A_2, ..., A_M$ are independent, then $P(A_1 \cap A_2 \cap ... \cap A_M) = P(A_1) * P(A_2) * ... * P(A_M) = \prod_{i=1}^{M} P(A_i)$

- Example, say you flip a coin 3 times and A_i is the event that a heads shows up on the i^{th} flip ($i = 1, 2, 3$).

- Note that the sample space is: $S = \{HHH, HTH, HTT, ..., TTH, TTT\}$.

  - What is the probability of flipping a head, then a tail, then a head?

Conditional Probability

Example.

Say there is a bioassay type exam that is a test for the presence of a certain type of disease. The exam will correctly return a positive test, $+$, when the person actually has the disease, D, with probability 0.95.

The exam will incorrectly return a positive test when the person actually does not have the disease, D^c, with probability 0.02.

The prevalence of the disease is 0.01 (this is $P(D) = 0.01$).

Use a tree diagram to calculate the probability of getting a positive test on the medical exam (not conditional on whether someone has the disease or not.)

Conditional Probability

Example: Continued from previous page.

- What is the probability of having the disease given a positive medical result.

- What is the probability of having the disease?

- Are having the disease and getting a positive result independent events?

- If they were independent, then $P(D) = P(D|+)$.