STA237H1F (LEC0201)

Week 2: First principles, conditional probability and independence

(Dekking et al., 2005)

Conditional Probability

- *Probability* (informally): Measure of uncertainty
- Conditional probability: Another measure of uncertainty
- *Condition on what*?: On more information, usually we know that a specific event occured

- *Definition*. Let A and B two events such that $P(B) > 0$. Then the conditional probability of A given B (our new knowledge) is

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

- $P(A|B) = P_B(A)$

Cont…

- *e.g.* Ɛ= birthday month of a person (equally months)
  - S =Short month= $\{Feb, Apr, Jun, Sep, Nov\}$
  - R = Letter R in full name = $\{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec\}$
  - $R \cap S = \{Feb, Apr, Sep, Nov\}$

  - We do not know anything: $P(S) = \frac{5}{12}$
  - We know that the birthday month has a letter R:

    $$P(S|R) = \frac{P(S \cap R)}{P(R)} = \frac{4/12}{8/12} = \frac{4}{8} = \frac{1}{2}$$

  - R is the new sample space:

    $\#R = 8$ and $\#(S \cap R) = 4$ then $P_R(S) = \frac{4}{8}$
  - and $P(R|R) = 1 = P_R(R)$

Cont…

- $P(\cdot\,|B) = P_B(\cdot)$ is a new probability function

- *Therefore it satisfies*
  - $P(\emptyset|B) = 0$
  - $P(A^c|B) = 1 - P(A|B)$
  - $P(A|B) = P(A \cap C|B) + P(A \cap C^c|B)$
  - $P(A \cup C|B) = P(A|B) + P(C|B) - P(A \cap C|B)$
  - If $A \subset C$ then $P(A|B) \leq P(C|B)$

- e.g. (cont…)
  - $P(L|R) = P(S^c|R) = 1 - P(S|R) = 1 - \frac{1}{2} = \frac{1}{2}$

Multiplication rule

- If $P(A) > 0$ and $P(B) > 0$ then
$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$
- e.g. Ɛ= select two balls without replacement from an urn with 4 white and 3 black
  - Let $A =$ both balls are white $P(A) = \frac{{}_4C_2}{{}_7C_2} = \frac{6}{21}$
  - Let $W_i =$ white ball in extraction $i = 1,2$

$P(W_1) = \frac{4}{7}$ and $P(W_2|W_1) = \frac{3}{6}$ then $P(A) = P(W_1 \cap W_2) = \left(\frac{3}{6}\right)\left(\frac{4}{7}\right) = \frac{12}{42}$

  - Can we compute$P(W_1|W_2)$=?
- In general, for $A_1, A_2, \dots, A_n$ such that $P(A_i) > 0$
$$P(A_1 \cap A_2 \cap \cdots \cap A_n)$$
$$= P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Law of total probability

- Let $P(B) > 0$ then

$$P(A) = P(A \cap B) + P(A \cap B^c)$$
$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

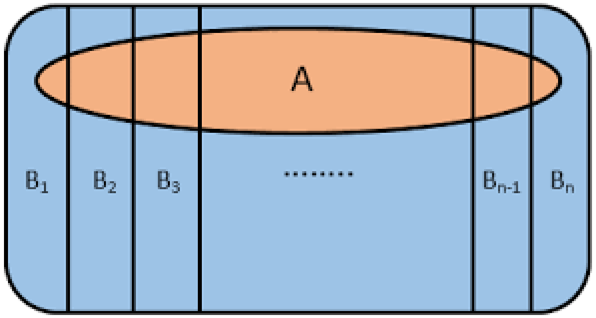
- e.g. $\mathcal{E}$= Testing for HIV
  - Let TP = test is positive, and TN = test is negative
  - VP = person is HIV positive, and VN = person is HIV negative
  - Test: $P(TP|VP) = 0.95$ and $P(TP|VN) = 0.01$
  - Prevalence of HIV in Canada $P(VP) = 0.0017$
    - $P(TP) = P(TP|VP)P(VP) + P(TP|VN)P(VN)$
    - $P(TP) = (0.95)(0.0017) + (0.01)(0.9983) = 0.01159$
  - Prevalence of HIV in South Africa $P(VP) = 0.14$
    - $P(TP) = P(TP|VP)P(VP) + P(TP|VN)P(VN)$
    - $P(TP) = (0.95)(0.14) + (0.01)(0.86) = 0.1416$

Cont…

- *In general*: Let $B_1, B_2, \dots, B_n$ be a finite partition of Ω (i.e. $\bigcup_{i=1}^{n} B_i = \Omega$ and $B_i \cap B_j = \emptyset \ \ \forall i \neq j$). Then

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

- Graphically



Bayes rule

- If $P(A) > 0$ and $P(B) > 0$ then

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

- Rule of inverse probabilities

- Learning rule:

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

  - $P(B)$ = prior probability
  - $P(B|A)$ = posterior probability

Cont…

- e.g. Ɛ= Testing for HIV
  - Let TP = test is positive, and TN = test is negative
  - VP = person is HIV positive, and VN = person is HIV negative
  - Test: $P(TP|VP) = 0.95$ and $P(TP|VN) = 0.01$

  - Prevalence of HIV in Canada $P(VP) = 0.0017 \Rightarrow P(TP) = 0.01159$
    - $P(VP|TP) = \frac{P(TP|VP)P(VP)}{P(TP)} = \frac{(0.95)(0.0017)}{0.01159} = 0.1392$

  - Prevalence of HIV in South Africa $P(VP) = 0.14 \Rightarrow P(TP) = 0.1416$
    - $P(VP|TP) = \frac{P(TP|VP)P(VP)}{P(TP)} = \frac{(0.95)(0.14)}{0.1416} = 0.9392$

Cont…

- *In general*: Let $B_1, B_2, \dots, B_n$ be a finite partition of Ω (i.e. $\bigcup_{i=1}^{n} B_i = \Omega$ and $B_i \cap B_j = \emptyset \ \forall i \neq j$). Then

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)}$$

  - for j $= 1, \dots, 5$ and $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$

- e.g. Five urns numbered 1 to 5 each with 10 balls, i white and $10 - i$ black, $i = 1, \dots, 5$. $\mathcal{E} =$ Select one urn at random and then take one ball from the urn selected
  - Let $W =$ white ball, and $U_i =$ select urn $i = 1, \dots, 5$
  - $P(W) = \sum_{i=1}^{5} P(W|U_i)P(U_i) = \sum_{i=1}^{5} \left(\frac{i}{10}\right)\left(\frac{1}{5}\right) = \frac{15}{50} = \frac{3}{10}$
  - $P(U_1|W) < \cdots < P(U_5|W) = 1/3$

Independence

- *Definition*: events A and B are independent if one of the following is satisfied
  - (a) $P(A|B) = P(A)$
  - (b) $P(B|A) = P(B)$
  - (c) $P(A \cap B) = P(A)P(B)$

- E.g. Ɛ= birthday month
  - $P(S) = \frac{5}{12}$ and $P(S|R) = \frac{1}{2}$ then S and R are not independent
  - $S^c = L$, therefore $S \cap L = \emptyset \Leftrightarrow S, L$ are disjoint, but $P(S|L) = 0$ so S, L are not independent
- E.g. Ɛ= select a card from a deck of 52
  - $F = number\ 4$, $R = red$, $P(F) = \frac{4}{52}$, $P(F|R) = \frac{2}{26}$, thus F, R are indep but $F \cap R \neq \emptyset$, then F, R are not disjoint

Exercise

- Each of 2 balls is painted either black or gold and then placed in an urn. Suppose that each ball is colored with equal probability
  - (a) Compute the probability that both balls are painted gold
  - (b) Suppose that you obtain information that the gold paint has been used (and thus at least one of the balls is painted gold). Compute the conditional probability that both balls are painted gold
  - (c) Suppose now that the urn tips over and 1 ball falls out. It is painted gold. What is the probability that both balls are gold in this case?
    - $\Omega = \{(G, G), (G, B), (B, G), (B, B)\}$
    - $A =$ two gold balls, $\Rightarrow P(A) = 1/4$
    - $B =$ gold paint has been used $= \{(G, G), (G, B), (B, G)\} \Rightarrow P(A|B) = 1/3$
    - $C =$ 1st ball is golden $= \{(G, G), (G, B)\} \Rightarrow P(A|C) = 1/2$
    - $\therefore A$ and B are not independent, A and C are not independent

Rstudio exercises

- Suggested commands: $<-$, matrix, sample, $\mathrm{function}$
- Simulate a conditional probability in R:
  - $\mathcal{E}$=draw two dice of six numbers
    - trial <- sample(1:6,2,replace=TRUE)
  - A = first dice is a two
  - B = sum of the two dice is seven
  - Draw multiple trials to determine
    - $P(A)$ = relative frequency of occurrences when the first dice is a two from total draws
    - $P(A|B)$ = relative frequency of occurrencies when the first dice is a two from those where the sum is a seven
  - Fix repeatitions at $N = 10000$
  - Use $0 - 1$ identification vectors to indicate if a specific trial satisfies a condition and use command mean