Assignment1 – Probability

Given: Apr 27   Due: May 7

Problem 1.1 (Bayesian Rules) 0 pt

Give the formulas and a one-sentence explanation of the following basic rules in Bayesian inference:

1. Bayes rule
2. Product rule
3. Chain rule
4. Marginalization
5. Normalization

Solution:

1. Bayes rule: $P(A|B) = P(B|A)P(A)/P(B)$
   The conditional probability of A given B multiplied by the probability of B is the same as the conditional probability of B given A multiplied by the probability of A – both are equal to $P(A, B)$. We can use that to compute one conditional probability from the other.

2. Product rule: $P(A, B) = P(A|B)P(B)$, The probability of A and B is the product of the probability of B and the one of A given B. If A and B are independent, this simplifies to $P(A, B) = P(A)P(B)$.

3. Chain rule: $P(A_1, \dots, A_n) = P(A_n|A_{n-1}, \dots, A_1) \cdot P(A_{n-1}|A_{n-2}, \dots, A_1) \cdot \dots$ Iterated application of the product rule.

4. Marginalization of A with respect to Y: $P(A) = \Sigma_{y \in E} P(A, y)$ where E is the set of values of Y. Since the probabilities of the values of Y sum to 1, we can always introduce/remove a sum over all values.

5. Normalization of X with respect to event e: $P(X|e) = \alpha(P(X, e))$ where α is the function that multiplies every element in a vector v (here: the vector $\langle P(X = x_1, e), \dots, P(X = x_n, e) \rangle$ where the x_i are the possible values of X) by $1/\Sigma_i v_i$. The probability of X given e can be obtained by normalizing the joint probability X and e.

Problem 1.2 (Disjunctive Random Variables) 20 pt

We know that given Boolean random variables A and B we have

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

Extend this formula to the case of three random variables $P(A \vee B \vee C)$. Draw a Venn diagram to “prove” your formula.

Solution: $P(A \vee B \vee C) = P(A) + P(B) + P(C) - P(A \wedge B) - P(B \wedge C) - P(C \wedge A) + P(A \wedge B \wedge C)$

Problem 1.3 (Basic Probability) 20 pt

Let A, B, C be Boolean random variables, and let a, b, c denote the atomic events that A, B, C, respectively, are true. Which of the following equalities are always true? Justify each of your answers in one sentence.

1. $P(b) = P(a, b) + P(\neg a, b)$
2. $P(a) = P(a|b) + P(a|\neg b)$
3. $P(a, b) = P(a) \cdot P(b)$
4. $P(a, b|c) \cdot P(c) = P(c, a|b) \cdot P(b)$
5. $P(a \vee b) = P(a) + P(b)$
6. $P(a, \neg b) = (1 - P(b|a)) \cdot P(a)$

Solution:

1. True (marginalization over A)
2. Not true (e.g. $P(a|b) = P(a|\neg b) = 0.6$ would result in $P(a) = 1.2$)
3. Not true (only true if A and B are stochastically independent)
4. True (using product rule, both sides become $P(a, b, c)$)
5. Not true (general form is $P(a \vee b) = P(a) + P(b) - P(a, b)$)
6. True ($1 - P(b|a) = P(\neg b|a)$ and via product rule we get $P(a, \neg b)$)

Problem 1.4 (Chained Production Elements) 20 pt

An apparatus consists of six elements A, B, C, D, E, F. Assume the probabilities $P(b_X)$, that element X breaks down, are all stochastically independent, with $P(b_A) = 5\%$, $P(b_B) = 10\%$, $P(b_C) = 15\%$, $P(b_D) = 20\%$, $P(b_E) = 25\%$, and $P(b_F) = 30\%$.

1. Assume the apparatus works if and only if at least A and B are operational, C and D are operational, or E and F are operational. What is the probability the apparatus works?
2. Consider a different scenario, in which the elements A and C, D and F and B and E are pairwise linked; such that if either of them breaks down, then the linked element is not operational either. What is the probability that the apparatus works now?

Note that we deliberately differentiate between *not being operational* and *being broken*. If an element breaks, it is not operational; if an element is not operational, either it or the linked element broke.

Solution: Let W be a random variable stating that the apparatus works. Let O_X be a random variable indicating that element X is operational.

1. In this problem, o_X is equivalent to $\neg b_X$ for all elements X.

$$\begin{aligned}
P(w) &= P((o_A \wedge o_B) \vee (o_C \wedge o_D) \vee (o_E \wedge o_F)) \\
&= 1 - P(\underbrace{\neg(o_A \wedge o_B) \wedge \neg(o_C \wedge o_D) \wedge \neg(o_E \wedge o_F)}_{\text{all events are independent}}) \\
&= 1 - P(\neg(o_A \wedge o_B)) \cdot P(\neg(o_C \wedge o_D)) \cdot P(\neg(o_E \wedge o_F)) \\
&= 1 - P(\neg o_A \vee \neg o_B) \cdot P(\neg o_C \vee \neg o_D) \cdot P(\neg o_E \vee \neg o_F) \\
&= 1 - P(b_A \vee b_B) \cdot P(b_C \vee b_D) \cdot P(b_E \vee b_F) \\
&= 1 - (P(b_A) + P(b_B) - P(b_A) \cdot P(b_B)) \cdot (P(b_C) + P(b_D) - P(b_C) \cdot P(b_D)) \\
&\quad \cdot (P(b_E) + P(b_F) - P(b_E) \cdot P(b_F)) \\
&= 1 - (0.05 + 0.1 - (0.05 \cdot 0.1)) \cdot (0.15 + 0.2 - (0.15 \cdot 0.2)) \cdot (0.25 + 0.3 - (0.25 \cdot 0.3))
\end{aligned}$$

2. Using the exclusion principle:

$$\begin{aligned}
P(w) &= P((o_A \wedge o_B) \vee (o_C \wedge o_D) \vee (o_E \wedge o_F)) \\
&= P(o_A, o_B) + P(o_C, o_D) + P(o_E, o_F) - P(o_A, o_B, o_C, o_D) - P(o_A, o_B, o_E, o_F) - P(o_C, o_D, o_E, o_F) \\
&\quad + P(o_A, o_B, o_C, o_D, o_E, o_F)
\end{aligned}$$

Due to the links, o_X is equivalent to $\neg b_X \wedge \neg b_{\widetilde{X}}$ where $\widetilde{X}$ is the element that X is linked with. So, for example, o_A is equivalent to $\neg b_A \wedge \neg b_C$. This gives us

$$\begin{aligned}
P(w) = {} & \underbrace{P(\neg b_A, \neg b_C, \neg b_B, \neg b_E)}_{=P(o_A, o_B)} + P(\neg b_A, \neg b_C, \neg b_D, \neg b_F) + P(\neg b_B, \neg b_E, \neg b_D, \neg b_F) \\
&- P(\neg b_A, \neg b_B, \neg b_C, \neg b_D, \neg b_E, \neg b_F) - P(\neg b_A, \neg b_B, \neg b_C, \neg b_D, \neg b_E, \neg b_F) \\
&- P(\neg b_A, \neg b_B, \neg b_C, \neg b_D, \neg b_E, \neg b_F) + P(\neg b_A, \neg b_B, \neg b_C, \neg b_D, \neg b_E, \neg b_F) \\
= {} & P(\neg b_A, \neg b_C, \neg b_B, \neg b_E) + P(\neg b_A, \neg b_C, \neg b_D, \neg b_F) + P(\neg b_B, \neg b_E, \neg b_D, \neg b_F) \\
&- 2P(\neg b_A, \neg b_B, \neg b_C, \neg b_D, \neg b_E, \neg b_F) \\
= {} & P(\neg b_A)P(\neg b_B)P(\neg b_C)P(\neg b_E) + P(\neg b_A)P(\neg b_C)P(\neg b_D)P(\neg b_F) + P(\neg b_B)P(\neg b_D)P(\neg b_E)P(\neg b_F) \\
&- 2P(\neg b_A)P(\neg b_B)P(\neg b_C)P(\neg b_D)P(\neg b_E)P(\neg b_F) \\
= {} & 0.5450625 + 0.4522 + 0.378 - 2 \cdot 0.305235 \approx 76\%
\end{aligned}$$

Problem 1.5 (Probabilities in Python) 40 pt

Complete the partial implementation of probabilities at `https://kwarc.info/teaching/AI/resources/AI2/probabilities/`

Solution: See `https://kwarc.info/teaching/AI/resources/AI2/probabilities/`
