Assignment1 – Probability

Given: Apr 25   Due: May 2

Problem 1.1 (Simple Sample Spaces)

In many important situations including all problems treated in this course, the sample space, probability measure, domains, and random variables can be given in a simplified form, namely by:

- a list of random variable declarations $X_1, \ldots, X_n$, each consisting of
  - a name such as X
  - a finite domain such as $D_X = \{0, 1, 2, 3\}$
- a probability function $\Omega \to [0; 1]$ where $\Omega = D_{X_1} \times \ldots \times D_{X_n}$ such that $\Sigma_{e \in \Omega} P(e) = 1$

Define the corresponding probability space $\langle \Omega, Q \rangle$ and show that it satisfies the Kolmogorov axioms.

Define the random variables $Y_1, \ldots, Y_n$ induced by the respective X_i.

Solution: $Q : \mathcal{P}(\Omega) \to [0; 1]$ is defined by $Q(A) = \Sigma_{e \in A} P(e)$. (Note this is a finite sum because all the D_X and thus Ω and A are finite.)

To show the Kolmogorov axioms:

- $Q(\Omega) = \Sigma_{e \in \Omega} P(e) = 1$
- $Q(\bigcup_i A_i) = \Sigma_{e \in \bigcup_i A_i} P(e) =$ (because the A_i are pairwise disjoint.) $\Sigma_i \Sigma_{e \in A_i} P(e) = \Sigma_i Q(A_i)$.

For each X_i with domain D_{X_i}, we define a random variable $Y_i : \Omega \to D_{X_i}$ by $Y_i(x_1, \ldots, x_n) = x_i$ for $(x_1, \ldots, x_n) \in \Omega$.

Problem 1.2 (Bayesian Rules)

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

1. Bayes rule

   Solution: $P(A \mid B) = P(B \mid A)P(A)/P(B)$

   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

   Solution: $P(A, B) = P(A \mid B)P(B)$, 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

Solution: $P(A_1, \ldots, A_n) = P(A_n \mid A_{n-1}, \ldots, A_1) \cdot P(A_{n-1} \mid A_{n-2}, \ldots, A_1) \cdot \ldots$ Iterated application of the product rule.

4. Marginalization

Solution: 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

Solution: Normalization of X with respect to event e: $P(X \mid e) = \alpha(P(X, e))$ where α is the function that multiplies every element in a vector v (here: the vector $\langle P(X = x_1, e), \ldots, P(X = x_n, e)\rangle$ 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 1.3 (Basic Probability)

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. $P(b) = P(a, b) + P(\neg a, b)$

Solution: True (marginalization over A)

2. $P(a) = P(a \mid b) + P(a \mid \neg b)$

Solution: Not true (e.g. $P(a \mid b) = P(a \mid \neg b) = 0.6$ would result in $P(a) = 1.2$)

3. $P(a, b) = P(a) \cdot P(b)$

   Solution: Not true (only true if A and B are stochastically independent)

4. $P(a, b \mid c) \cdot P(c) = P(c, a \mid b) \cdot P(b)$

   Solution: True (using product rule, both sides become $P(a, b, c)$)

5. $P(a \vee b) = P(a) + P(b)$

   Solution: Not true (general form is $P(a \vee b) = P(a) + P(b) - P(a, b)$)

6. $P(a, \neg b) = (1 - P(b \mid a)) \cdot P(a)$

   Solution: True ($1 - P(b \mid a) = P(\neg b \mid a)$ and via product rule we get $P(a, \neg b)$)

Problem 1.4 (Chained Production Elements)

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 $P(b_F) = 30\%$.

Note: 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.

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?

   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.

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. 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?

Solution: 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

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

Problem 1.5 (Probabilities in Python)

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/`
