COMP2610/6261: INFORMATION THEORY

Week 2 Tutorial: Probability and Bayesian Interface

Australian National University

SOME KEY CONCEPTS:

Conditional Probability [P(A|B)]: Measure of the probability of an event (A) occurring, given that another event (B) has already occurred. Mathematically, P(A|B) = P(A, B) | P(B).

- For independent events, the joint probability simplifies to P(A, B) = P(A)*P(B). Therefore, P(A|B) = P(A)
- Sum Rule (Law of Total Probability): P(A) = Σ P(A | Bi) * P(Bi), for all i, where the events Bi form a partition of the sample space - meaning every outcome is covered by one of the Bi and the Bi don't overlap.

Chain Rule (Product Rule) of Probability:

- For 2 events: P(A, B) = P(A) * P(B | A) = P(B) * P(A | B)
- For 3 events: P(A, B, C) = P(A) * P(B | A) * P(C | A, B)

Bayesian Interface: Method of statistical interface based on Bayes' theorem. In machine learning, Bayesian Interface is often used to learn probability distributions of the data and make probabilistic predictions and estimations.

- Bayes' theorem: **P(A|B)** = [P(B|A) * P(A)] / P(B), and **P(B|A)** = [P(A|B) * P(B)] / P(A).

Question 1: (From Bishop, 2006) Suppose that we have three colored boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes. Box b contains 1 apple, 1 orange, and 0 limes. Box g contains 3apples, 3 oranges, and 4 limes. A box is chosen at random with probabilities $p(\mathrm{r}) = 0.2$, $p(\mathrm{b}) = 0.2$, $p(\mathrm{g}) =0.6$ and a piece of fruit is removed from the box (with probability of selecting any of the items in the box):

a) What is the probability of selecting an apple?

b) If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

Solution: We wish to calculate the probability of selecting an apple. What we have are the probabilities of selecting each of the boxes. We can also calculate from the information we are given the chance of choosing a particular fruit from a particular box. That is, we have conditional probabilities for fruit given box as well as marginal probabilities of selecting each of the boxes.

a) We marginalize over boxes to express P(apple) in terms of joint probabilities, and then use the chain rule:

$$P(\text{apple}) = P(\text{apple, red}) + P(\text{apple, green}) + P(\text{apple, blue})$$

$$= P(\text{apple} \mid \text{red})P(\text{red}) + P(\text{apple} \mid \text{green})P(\text{green}) + P(\text{apple} \mid \text{blue})P(\text{blue})$$

$$= 0.3 \times 0.2 + 0.3 \times 0.6 + 0.5 \times 0.2 = \mathbf{0.34}$$

b) We want to compute P (green | orange). We know from the definition of conditional probability that,

$$\text{P (green | orange)} = \frac{P\,(green,\;\; orange)}{P\,(orange)}$$

To use this formula, we need P(green, orange) and P(orange). Now

P(orange) = P(orange | red) * P(red) + P(orange | green) * P(green) + P(orange | blue) * P(blue)

= 0.4 × 0.2 + 0.3 × 0.6 + 0.5 × 0.2 = 0.36.

Also, by chain-rule P(orange, green) = P(orange | green) * P(green) = 0.3 × 0.6 = 0.18

Putting all this together, we have P (green | orange) = $\frac{P\,(green,\;\; orange)}{P\,(orange)} = \frac{0.18}{0.36} =$ **0.5**

Question 2: A scientist conducts 1000 trials of an experiment involves variables X,Y , each with possible values {0,1}.He records the outcomes of the trials in the following table.

Counts		X = 0	X = 1
Y	0	100	250
	1	150	500

Compute each of the following, showing all your working.

(a) $p(X = 1, Y = 1)$.

(b) $p(X = 1)$.

(c) $E[X]$.

(d) $p(Y = 1 | X = 1)$.

(e) $p(Y = 1 | X = 0)$.

(f) Let Z be a noisy version of the XOR of X and Y, with

$$p(Z = 1 | X = x, Y = y) = \begin{cases} 0.9 & \text{if } (x, y) = (0, 1) \text{ or } (x, y) = (1, 0) \\ 0.1 & \text{if } (x, y) = (0, 0) \text{ or } (x, y) = (1, 1) \end{cases}.$$

Compute $p(X = 1, Y = 1 | Z = 1)$.

Solution:

(a) By definition, $p(X=1,Y=1) = \frac{500}{1000} = \frac{1}{2}$.

(b) By the sum rule, $p(X=1) = p(X=1,Y=1) + p(X=1,Y=0) = \frac{750}{1000} = \frac{3}{4}$.

(c) By definition of expectation, $E[X] = p(X=1)\cdot 1 + p(X=0)\cdot 0 = p(X=1) = \frac{3}{4}$.

(d) By definition of conditional probability, $p(Y=1|X=1) = \frac{p(X=1,Y=1)}{p(X=1)} = \frac{1}{2}\cdot\frac{4}{3} = \frac{2}{3}$.

(e) By definition of conditional probability, $p(Y=1|X=0) = \frac{p(X=0,Y=1)}{p(X=0)} = \frac{150}{250} = \frac{3}{5}$.

(f) By Bayes' rule,

$$p(X=1,Y=1|Z=1) = \frac{p(Z=1|X=1,Y=1)\cdot p(X=1,Y=1)}{p(Z=1)}.$$

By the sum rule (or marginalisation),

$$\begin{aligned}
p(Z=1) &= \sum_{x,y} p(Z=1,X=x,Y=y) \\
&= \sum_{x,y} p(Z=1|X=x,Y=y)\cdot p(X=x,Y=y) \\
&= 0.9\cdot(p(X=1,Y=0)+p(X=0,Y=1))+ \\
&\quad 0.1\cdot(p(X=1,Y=1)+p(X=0,Y=0)) \\
&= 0.9\cdot\left(\frac{1}{4}+\frac{15}{100}\right)+0.1\cdot\left(\frac{1}{2}+\frac{1}{10}\right) \\
&= 0.42.
\end{aligned}$$

Question 3: (From Barber, 2011) Two balls are placed in a box as follows:

- a fair coin is tossed
- a white ball is placed in the box if a head occurs; otherwise, a red ball is placed in the box
- the coin is tossed again
- a red ball is placed in the box if a tail occurs; otherwise, a white ball is placed in the box.

Balls are drawn from the box three times in succession (always with replacing the drawn ball back in the box). It is found that on all three occasions a red ball is drawn. What is the probability that both balls in the box are red?

Solution: According to question, P(first ball is red) = P(second ball is red) = 0.5.

We wish to calculate P(2 red balls in the box| 3 red balls selected from the box). Using Bayes rule we find that,

P(2 red balls in the box | 3 red balls selected from the box) = $\frac{P(2\, red\ balls\ in\ the\ box, 3\, red\ balls\ selected\ from\ the\ box)}{P(\,3\, red\ balls\ selected\ from\ the\ box)}$

Since it is impossible to select two red balls if both balls in the box are white. Hence,

P(3 red balls selected from the box) = P(3 red balls selected from the box|2 red balls in the box) * P(2 reds in the box)

+ P(3 red balls selected from the box|1 red ball in the box) * P(1 reds in the box)

Solution (Cont.):

$$= 1 \times 0.25 + 0.125 \times 0.5 = 0.3125$$

and so, P(2 red balls in the box | 3 red balls selected from the box) = 0.25 / 0.3125 = **0.800**.

Question 4: Several of the results we have seen in lectures generalize in the expected way when we condition on additional random variables.

(a) For random variables X,Y, the conditional probability $p(X|Y)$ may be defined as

$$p(X|Y) = \frac{p(X,Y)}{p(Y)}.$$

Using this definition, show that for random variables X,Y,Z, the following conditional version of Bayes rule holds:

$$p(X|Y,Z) = \frac{p(Y|X,Z)p(X|Z)}{p(Y|Z)}.$$

Question 4 (Cont.):

(b) The sum rule (or marginalisation) says that for random variables $X_1,...,X_n$,

$$p(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) = \sum_x p(X_1, \ldots, X_{i-1}, X_i = x, X_{i+1}, \ldots, X_n).$$

Using this fact, and the definition of conditional probability, show that

$$p(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n|Y) = \sum_x p(X_1, \ldots, X_{i-1}, X_i = x, X_{i+1}, \ldots, X_n|Y)$$

where Y is another random variable.

Solution: **(a)** The definition is nothing but a restatement of the chain rule, **p(X, Y) = p(X|Y) * p(Y).** Similarly, for 3 random variables, **p(X, Y, Z) = p(X | Y, Z) * p(Y, Z) ……….(1).**

The chain rule also implies that, **p(X, Y, Z) = p(Y |X, Z) * p(X, Z) ……….(2).**

Also, p(Y, Z) = p(Y |Z) * p(Z), and p(X, Z) = p(X / Z) * p(Z). By substituting p(Y,Z) and p(X,Z) in eqn. (1) and (2) and equating, we get p(Y | X, Z) * p(X | Z) * p(Z) = p(X | Y, Z) * p(Y |Z) * p(Z), or

p(Y | X , Z) * p(X | Z) = **p(X | Y, Z) * p(Y |Z)**, which is the desired statement.

Solution (Cont.):

(b) By definition of conditional probability,

$$p(X_1, \ldots, X_n | Y) = \frac{p(X_1, \ldots, X_n, Y)}{p(Y)}$$

Thus,

$$\begin{aligned}
p(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n | Y) &= \frac{p(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n, Y)}{p(Y)} \\
&= \frac{\sum_x p(X_1, \ldots, X_{i-1}, X_i = x, X_{i+1}, \ldots, X_n, Y)}{p(Y)} \\
&= \sum_x \frac{p(X_1, \ldots, X_{i-1}, X_i = x, X_{i+1}, \ldots, X_n, Y)}{p(Y)} \\
&= \sum_x p(X_1, \ldots, X_{i-1}, X_i = x, X_{i+1}, \ldots, X_n | Y).
\end{aligned}$$

THANK YOU

Contact Details:

Office: *B137A, Brian Anderson Building (Ground Floor)*

Email: *manish.kumar@anu.edu.au*

Australian National University