UWO CS2214

Tutorial #10

Problem 1 (Basic probability calculations) In a roulette, a wheel with 38 numbers is spun. Of these, 18 are red, and 18 are black. The other two numbers, which are neither black nor red, are 0 and 00. The probability that when the wheel is spun it lands on any particular number is 1/38.

1. What is the probability that the wheel lands on a red number?
2. What is the probability that the wheel lands on a black number twice in a row?
3. What is the probability that the wheel lands on 0 or 00?
4. What is the probability that in five spins the wheel never lands on either 0 or 00?

Provide detailed justifications of your answers.

Solution 1

1. There are 18 favorable outcomes for a total of $18 + 18 + 2$ possible outcomes. Hence the probability is $18/(18 + 18 + 2)$ is $9/19 \approx 0.473$.
2. There are 18×18 favorable outcomes for a total of $(18 + 18 + 2)^2$ possible outcomes. Hence the probability is $18^2/(18 + 18 + 2)^2$ is $81/361 \approx 0.2243$.
3. There are 2 favorable outcomes for a total of $18 + 18 + 2$ possible outcomes. Hence the probability is $2/(18 + 18 + 2)$ is $1/19 \approx 0.0526$.
4. We use the Bernoulli trials formula. We see "wheel lands on 0 or 00" as a success. The number of trials is $n = 5$, the probability of a success is $p = 2/19$ and the desired number of success is $k = 0$. Thus the requested probability is

$$\binom{n}{k} p^k (1-p)^{n-k} = 1(1/19)^0(18/19)^5 \approx 0.763.$$

Problem 2 (Bernoulli Trials) A multiple-choice exam consists of 20 questions. Each question offers 4 choices, one and only one of them being correct. The exam is passed if at least 16 questions are answered correctly. What is the probability that someone answering randomly passes the exam.

Solution 2

Let $P(k)$ the probability of answering correctly exactly k questions out of 16. We use the Bernoulli trials formula $b(k : n, p)$ from the lectures, with $n = 20$ and $p = 1/2^4 = 1/15$. Hence, we have $P(k) = b(k : 20, 1/16)$, that is:

$$P(k) = \binom{20}{k} (1/16)^k (15/16)^{20-k}. \tag{1}$$

We are asked to compute $P(16) + P(17) + P(18) + P(19) + P(20)$

In Maple, we obtain the following

```
  > sum(binomial(20,k) * (1/16)^k * (15/16)^(20-k), k=16..20);
                                                62292169
                                        ------------------------
                                        302231454903657293676544

> evalf(%);
                                                          -15
                                           0.2061074980 10
```

Problem 3 (Law of total probability) Let S be a sample set. Thus S is a finite set. Let $B_1, \ldots, B_n$ be events of S. that is, subsets of S. We assume that $\{B_1, \ldots, B_n\}$ forms a *partition* of S that is,

(1) each of $B_1, \ldots, B_n$ is non-empty and is a proper subset of S,

(2) the sets $B_1, \ldots, B_n$ are pairwise disjoint, that is, $B_i \cap B_j = \emptyset$ for all $1 \leq i < j \leq n$,

(3) the union of all $B_1, \ldots, B_n$ equals S.

Then, for every event A of S, we have:

$$P(A) = \sum_{i=1}^{i=m} P(A \cap B_i).$$

Solution 3 From the hypothesis, we have

$$\begin{array}{rcll} A & = & A \cap S & \text{since } A \subseteq S \\ & = & A \cap (B_1 \cup \cdots \cup B_n) & \text{from (3)} \\ & = & (A \cap B_1) \cup \cdots \cup (A \cap B_n) & \text{from distributiity} \end{array}$$

From (1), the sets $(A \cap B_i)$ are pairwise disjoint, therefore, using a theorem of the lecture slides, we deduce:

$$P(A) = \sum_{i=1}^{i=m} P(A \cap B_i).$$

Problem 4 (Conditional probabilty) Prove that, for any three events A, B, C, from some sample space, each event having positive probability, we have

1. $P(A|B) = 1 - P(\overline{A}|B)$,
2. $P(A \cap B \cap C) = P(A)\, P(B|A)\, P(C|(A \cap B)$.

Solution 4

1. By definition of a conditional probability, we have:

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

Since $\{\overline{A}, A\}$ is a partition of S, from the law of total probability, we have:

$$P(B) = P(\overline{A} \cap B) + P(A \cap B).$$

Therefore, we have:

$$P(\overline{A}|B) = \frac{P(B) - P(A \cap B)}{P(B)} = 1 - P(A|B).$$

2. By definition of a conditional probability, we have:

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

which gives:

$$P(A \cap B \cap C) = P(A \cap B)\, P(C|(A \cap B)).$$

By definition again of a conditional probability, we have:

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

that is,

$$\frac{1}{P(A \cap B)} = P(B|A)P(A).$$

From where we derive:

$$P(A \cap B \cap C) = P(B|A)P(A)\, P(C|(A \cap B)).$$

Problem 5 (Bayes theorem) Suppose that 8% of all bicycle racers use steroids, that a bicyclist who uses steroids tests positive for steroids 96% of the time, and that a bicyclist who does not use steroids tests positive for steroids 9% of the time (this is a "false positive" test result).

1. What is the probability that a randomly selected bicyclist who tests positive for steroids actually uses steroids?

2. What is the probability that a randomly selected bicyclist who tests negative for steroids did not use steroids?

Solution 5 Let S be the event that the "a bicyclist uses steroids" Let E be the event "a test is positive". What we know from the problem statement is:

$$P(S) = \frac{8}{100}, P(\overline{S}) = \frac{92}{100}, P(E|S) = \frac{96}{100}, P(E|\overline{S}) = \frac{9}{100}.$$

From there, we deduce:

$$P(\overline{E}|\overline{S}) = 1 - P(E|\overline{S}) = \frac{91}{100}, P(\overline{E}|S) = 1 - P(E|S) = \frac{4}{100}.$$

1. We want $P(S|E)$. Hence, we use Bayes' formula:

$$\begin{aligned} P(S|E) &= \frac{P(E|S)p(S)}{p(E|S)p(S)+p(E|\overline{S})p(\overline{S})} \\ &= \frac{96/100\times 8/100}{96/100\times 8/100+9/100\times 92/100} \\ &= \frac{64}{133} \approx 0.481. \end{aligned}$$

2. We use Bayes' formula:

$$\begin{aligned} P(\overline{S}|\overline{E}) &= \frac{P(\overline{E}|\overline{S})p(\overline{S})}{p(\overline{E}|\overline{S})p(\overline{S})+p(\overline{E}|S)p(S)} \\ &= \frac{91/100\times 92/100}{91/100\times 92/100+4/100\times 8/100} \\ &= \frac{2093}{2101} \approx 0.9961 \end{aligned}$$

Problem 6 (Generalized Bayes theorem) Suppose that E is an event from a sample space S and that $F_1, F_2, \ldots, F_n$ are events forming a partition of S. Assume that $p(E) \neq 0$ for $i = 1, 2, \ldots, n$. Then, prove that we have:

$$p(F_j \mid E) = \frac{p(E \mid F_j)p(F_j)}{\sum_{i=1}^{n} p(E \mid F_i)p(F_i)}$$

Solution 6 From the simple version, or first version, of Bayes theorem, we have:

$$p(F_j \mid E) = \frac{p(E \mid F_j)p(F_j)}{P(E)}$$

From the total law of probability, we have:

$$p(E) = \sum_{i=1}^{n} p(E \mid F_i)p(F_i).$$

The conclusion follows.

Problem 7 (Bayesian spam filter) Suppose that a Bayesian spam filter is trained on a set of 500 spam messages and 200 messages that are not spam. The word "exciting" appears in 40 spam messages and in 25 messages that are not spam. Would an incoming message be rejected as spam if it contains the word "exciting" and the threshold for rejecting spam is 0.9? (Assume, for simplicity, that the message is equally likely to be spam as it is not to be spam.)

Solution 7 Let S be the event that the message is spam. Let E be the event that the message contains the word "exciting". We have:

$$\begin{aligned} P(S|E) &= \frac{P(E|S)p(S)}{p(E|S)p(S)+p(E|\overline{S})p(\overline{S})} \\ &= \frac{40/500\times 1/2}{40/500\times 1/2+25/200\times 1/2} \\ &= \frac{16}{41} \approx 0.390 \end{aligned}$$

Problem 8 (Bayesian spam filter) Let S be the event that a message is spam. let E_1 and E_2 denote the events that the message contains the words w_1 and w_2 respectively. We assume that the events E_1 and E_2 are independent given S, that is, we have:

$$p(E_1 \cap E_2|S) = p(E_1|S)p(E_2|S).$$

We also assume $p(S) = \frac{1}{2}$. Prove that we have:

$$p(S \mid E_1 \cap E_2) = \frac{p(E_1 \mid S)p(E_2 \mid S)}{p(E_1 \mid S)p(E_2 \mid S) + p(E_1 \mid \overline{S})p(E_2 \mid \overline{S})}.$$

Solution 8 The proof was given during the lecture video recording.