Conditional Probability 2 – Solutions
COR1-GB.1305 – Statistics and Data Analysis

The Birthday Problem

1. A class has 50 students. What is the probability that at least two students have the same birthday? Assume that each person in the class was assigned a random birthday between January 1 and December 31.

Solution: Assume that everyone in the class is randomly assigned a birthday, which corresponds to number between 1 and 365 representing the day of the year. It turns out to be much easier to compute the probability using the complement rule, as

$$\mathrm{P}(\text{at least 2 people have the same birthday}) = 1 - \mathrm{P}(\text{all 50 birthdays are different}).$$

The next trick is to write the event that all 50 birthdays are different in a redundant way:

$$\begin{aligned}\{\text{all 50 birthdays are different}\} &= \{\text{first 2 are different}\} \cap \{\text{first 3 are different}\} \\ &\quad \cap \{\text{first 4 are different}\} \cap \{\text{first 5 are different}\} \cap \cdots \cap \{\text{first 50 are different}\}.\end{aligned}$$

In class we showed how to use the multiplicative rule repeatedly to get:

$$\begin{aligned}\mathrm{P}(\text{all 50 birthdays are different}) &= \mathrm{P}(\text{first 2 diff.}) \cdot \mathrm{P}(\text{first 3 diff.} \mid \text{first 2 diff.}) \\ &\quad \cdot \mathrm{P}(\text{first 4 diff.} \mid \text{first 3 diff.}) \cdot \mathrm{P}(\text{first 5 diff.} \mid \text{first 4 diff.}) \cdot \cdots \cdot \mathrm{P}(\text{first 50 diff.} \mid \text{first 49 diff.}).\end{aligned}$$

Using this expression we can compute

$$\mathrm{P}(\text{all 50 birthdays are different}) = \frac{364}{365} \cdot \frac{363}{365} \cdot \frac{362}{365} \cdot \cdots \cdot \frac{365-49}{365}.$$

We can do a similar calculation for other class sizes. The following table shows the probabilities of having at least two students with the same birthday for various class sizes:

Class Size	P(all diff.)	P(at least 2 same)
10	88%	12%
20	59%	41%
30	29%	71%
40	11%	89%
50	3%	97%
60	0.6%	99.4%
70	0.08%	99.92%

Independence

2. Suppose that you flip two fair coins. Let $A =$"the first coin shows Heads," $B =$ "The second coin shows Heads." Find the probability of getting Heads on both coins, i.e. find $P(A \cap B)$.

Solution: The long way to solve this problem is to write out the elementary outcomes and their probabilities:

Outcome	Probability
HH	$\frac{1}{4}$
HT	$\frac{1}{4}$
TH	$\frac{1}{4}$
TT	$\frac{1}{4}$

Since $A \cap B = \{\text{HH}\}$, it follows that

$$\mathrm{P}(A \cap B) = \tfrac{1}{4}.$$

We can solve this problem much more expediently using the independence of A and B:

$$\begin{aligned} \mathrm{P}(A \cap B) &= \mathrm{P}(A)\,\mathrm{P}(B \mid A) \\ &= \mathrm{P}(A)\,\mathrm{P}(B) \\ &= (\tfrac{1}{2})(\tfrac{1}{2}) \\ &= \tfrac{1}{4}. \end{aligned}$$

3. Suppose that you roll two dice. What is the probability of getting exactly one 6?

Solution: Define the following events:

$$\begin{aligned} A &= \text{“6 on the first roll,”} \\ B &= \text{“6 on the second roll,”} \end{aligned}$$

Using the shorthand $\bar{A} = A^c$ and $A\bar{B} = A \cap \bar{B}$, the event "exactly one 6" can be written as

$$\text{“exactly one 6”} = A\bar{B} \cup \bar{A}B$$

These events are mutually exclusive, so

$$\mathrm{P}(\text{exactly one 6}) = \mathrm{P}(A\bar{B}) + \mathrm{P}(\bar{A}B)$$

Using the independence of events A and B, we get

$$\mathrm{P}(A\bar{B}) = \mathrm{P}(A)\,\mathrm{P}(\bar{B}) = (\tfrac{1}{6})(\tfrac{5}{6})$$
$$\mathrm{P}(\bar{A}B) = \mathrm{P}(\bar{A})\,\mathrm{P}(B) = (\tfrac{5}{6})(\tfrac{1}{6})$$

Note that these two expressions are equal. Thus,

$$\mathrm{P}(\text{exactly one } 6) = 2 \cdot \tfrac{1}{6} \cdot \tfrac{5}{6} \approx 28\%$$

4. Suppose that you sell fire insurance policies to two different buildings in Manhattan, located in different neighborhoods. You estimate that the buildings have the following chances of being damaged by fire in the next 10 years: 5%, and 1%. Assume that fire damages to the two buildings are independent events. Compute the probability that exactly one building gets damaged by fire in the next 10 years.

Solution:

$$(.05)(.99) + (.95)(.01) = .059 = 5.9\%$$

5. Consider the following experiment. A hat contains two coins:

- one coin, the "fair" coin, has 50% chance of heads and 50% chance of tails on every flip;
- the other coin, the "heads" coin, has heads on both sides, so it always lands heads on every flip.

You reach into the hat and pull out a random coin, equally likely to get the fair coin or the heads coin. Then, you flip this coin twice.

Define events A and B as

$$\begin{aligned} A &= \text{the first flip lands heads} \\ B &= \text{the second flip lands heads.} \end{aligned}$$

(a) Based on your intuition, do you think that A and B independent events?

(b) Compute $P(A)$.

Solution: There are two possibilities with equal chances: either we pick the fair coin, or we pick the heads coin. We know that

$$P(\text{fair coin}) = P(\text{heads coin}) = 0.5$$

Given the coin, it is easy to compute the probabilities of heads:

$$\begin{aligned} P(\text{heads on first flip} \mid \text{fair coin}) &= 0.5 \\ P(\text{heads on first flip} \mid \text{heads coin}) &= 1.0 \end{aligned}$$

Finally,

$$\begin{aligned} P(A) &= P(\text{fair coin})P(\text{heads on first} \mid \text{fair coin}) \\ &\quad + P(\text{heads coin})P(\text{heads on first} \mid \text{heads coin}) \\ &= (0.5)(0.5) + (0.5)(1.0) \\ &= 0.75. \end{aligned}$$

(c) Compute $P(A \cap B)$.

Solution: Given the coin, the first and the second flips are independent:

$$\begin{aligned} P(\text{heads on both flips} \mid \text{fair coin}) &= P(\text{heads on first flip} \mid \text{fair coin}) \\ &\quad \cdot P(\text{heads on second flip} \mid \text{fair coin}) \\ &= (0.5)(0.5) \\ &= 0.25. \end{aligned}$$

Similarly,

$$P(\text{heads on both flips} \mid \text{heads coin}) = 1.0.$$

Now,

$$\begin{aligned}P(A) &= P(\text{fair coin})P(\text{heads on both} \mid \text{fair coin}) \\ &\quad + P(\text{heads coin})P(\text{heads on both} \mid \text{heads coin}) \\ &= (0.5)(0.25) + (0.5)(1.0) \\ &= 0.625.\end{aligned}$$

(d) Use your answers to parts (b) and (c) to either prove or disprove that A and B are independent.

Solution: To check for independence, we compare the product $P(A)P(B)$ with $P(A \cap B)$. Noting that $P(A) = P(B)$, we have

$$\begin{aligned}P(A)P(B) &= (0.75)(0.75) \\ &= 0.5625.\end{aligned}$$

Clearly, $P(A)P(B) \neq P(A \cap B)$. Thus, the events are not independent.
To get some more intuition for what is happening here, note that

$$\begin{aligned}P(B) &= 0.75, \\ P(B \mid A) &= P(B \cap A)/P(A) \\ &= 0.625/0.75 \\ &= 0.833.\end{aligned}$$

That is, before performing the experiment, we have a 75% chance of getting a heads on the second flip. In the middle of the experiment, if we see that the first flip is heads, then we have an 83.3% chance of getting heads on the next flip. Why is this? After we see event A, we gain some information relevant to event B, namely that it is more likely we have selected the heads coin.

Bayes' Rule

6. Every year in March there is a standardized exam for people who want to be licensed sheep herders. It happens that, with probability 0.45, a person will pass this exam. In the process of screening people, it turns out that among those who passed the exam, 60% had taken college courses in biology. It happens also that 30% of all those who take the exam had college courses in biology. Find the probability that a person with college courses in biology will pass the exam.

Solution: The information in the problem is

$$\begin{aligned} \mathrm{P}(\text{Pass}) &= .45 \\ \mathrm{P}(\text{Bio}) &= .30 \\ \mathrm{P}(\text{Bio} \mid \text{Pass}) &= .60 \end{aligned}$$

The problem is asking us to compute the quantity P(Pass | Bio). Using Bayes' rule,

$$\begin{aligned} \mathrm{P}(\text{Pass} \mid \text{Bio}) &= \mathrm{P}(\text{Bio} \mid \text{Pass}) \cdot \frac{\mathrm{P}(\text{Pass})}{\mathrm{P}(\text{Bio})} \\ &= (.60) \cdot \frac{(.45)}{(.30)} \\ &= .90. \end{aligned}$$

That is, there is a 90% chance that a person with college courses in biology will pass the exam.

7. Amazon.com maintains a list of all registered customers, along with their email addresses. During July, they sent coupons to 20% of their customers. They recorded that 5% of their customers made purchases in July, and 40% of all purchases were made with coupons. In this problem we will compute the proportion of customers sent a coupon in July who made a purchase in that month. For simplicity, we will assume that customers either make 0 or 1 purchases in July.

   (a) Consider a random customer, and define two events:

$$\begin{aligned} \text{Coupon} &= \text{the customer received a coupon in July}, \\ \text{Purchase} &= \text{the customer made a purchase in July}. \end{aligned}$$

   Express all percentages given in the problem statement as probabilities or conditional probabilities of these two events. Example: P(Coupon) = 0.20.

Solution: The information in the problem is

$$\begin{aligned} \text{P(Coupon)} &= .20 \\ \text{P(Purchase)} &= .05 \\ \text{P(Coupon} \mid \text{Purchase)} &= .40 \end{aligned}$$

(b) Use Bayes' rule to compute the proportion of custoers sent a coupon in July who made a purchase that month.

Solution: The problem is asking us to compute the quantity P(Purchase | Coupon). Using Bayes' rule,

$$\begin{aligned} \text{P(Purchase} \mid \text{Coupon)} &= \text{P(Coupon} \mid \text{Purchase)} \cdot \frac{\text{P(Purchase)}}{\text{P(Coupon)}} \\ &= (.40) \cdot \frac{(.05)}{(.20)} \\ &= .10. \end{aligned}$$