Problems on random experiments

Ignacio Cascos, Department of Statistics, Universidad Carlos III de Madrid

2018/19

1. Given three events A, B, and C, express the following events using the union, intersection and complementary event

   1. Out of the three events, only A occurs.

   $$A \cap \overline{B} \cap \overline{C}$$

   2. The three events occur.

   $$A \cap B \cap C$$

   3. A and B occur, but C does not.

   $$A \cap B \cap \overline{C}$$

   4. At least two out of the three event occur.

   $$(A \cap B) \cup (A \cap C) \cup (B \cap C)$$

   5. Exactly two of the three events occur.

   $$(A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C)$$

   6. At most two of the three events occur.

   $$\overline{A \cap B \cap C}$$

   7. At least one of the three events occurs.

   $$A \cup B \cup C$$

   8. One and only one of the three events occurs.

   $$(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$$

   9. None of them occur.

   $$\overline{A \cup B \cup C}$$

   10. If $P(A) = 0.2$, $P(B) = 0.4$, $P(C) = 0.3$, $P(A \cap B) = 0.1$ and $(A \cup B) \cap C = \emptyset$, which is the probability of each of the events described in the 9 previous expressions?

   - $P(A \cap \overline{B} \cap \overline{C}) = 0.1$,
   - $P(A \cap B \cap C) = 0$,
   - $P(A \cap B \cap \overline{C}) = 0.1$,
   - $P\left((A \cap B) \cup (A \cap C) \cup (B \cap C)\right) = 0.1$,
   - $P\left((A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C)\right) = 0.1$,
   - $P(\overline{A \cap B \cap C}) = 1$,
   - $P(A \cup B \cup C) = 0.8$,
   - $P\left((A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)\right) = 0.7$,
   - $P(\overline{A \cup B \cup C}) = 0.2$.

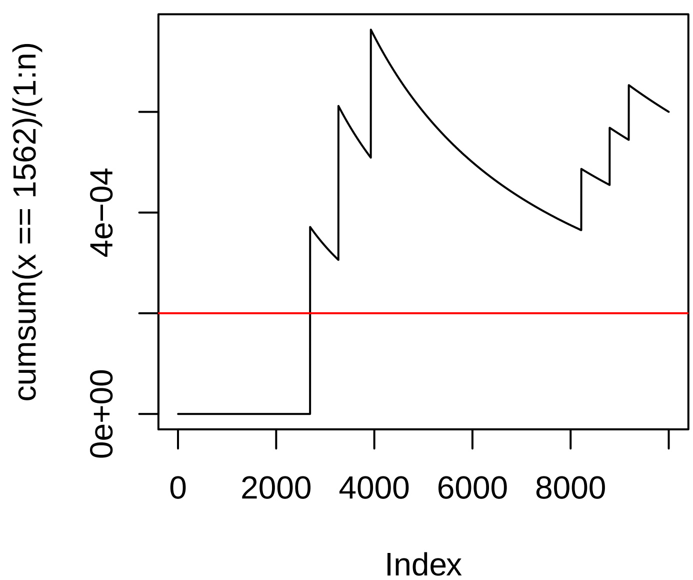
2. Complete the expression $P(A \cup B \cup C) = P(A) + P(B) + P(C) \cdots$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

3. A *flight number*, when combined with the name of the airline and the date, identifies a particular flight. The FAA limits flight numbers to four digits (0001 through 9999). Eastbound and northbound flights are assigned even numbers, while westbound and southbound flights have odd numbers. If all possible flight numbers for an eastbound or northbound flight are equally likely, what is the probability that the flight number of an eastbound flight generated at random is 1562?

$$P(\text{'even number between 0001 and 9999 is 1562'}) = \frac{1}{4999} = 0.0002.$$

```
set.seed(4)
n=10000
x=sample(seq(2,9998,by=2),n,replace=T)
plot(cumsum(x==1562)/(1:n),type="l")
abline(h=1/4999,col="red")
```



4. At a randomized controlled trial (RCT) each study subject is randomly allocated to receive one or other of the alternative treatments under study. Assume five subjects undergo a RCT. Each of them tosses three fair coins, if he obtains exactly one head, he is given a placebo, otherwise, he is treated with the drug under study. What is the probability that four of the subjects are treated with the drug and one with the placebo?

   Denote by D_i the event that the i-th patient is treated with the drug. The complementary to that event, $\overline{D}_i$ represents that the i-th patient is treated with the placebo. Since there are 8 possible outcomes for the toss of the three coins, and three of them contain exactly one head, then $P(\overline{D}_i) = 3/8$ and $P(D_i) = 5/8$. The events D_i and D_j for $i \neq j$ are independent.

$$P(\text{'4 Drug, 1 Placebo'}) = P\left((D_1 \cap D_2 \cap D_3 \cap D_4 \cap \overline{D}_5) \cup \ldots \cup (\overline{D}_1 \cap D_2 \cap D_3 \cap D_4 \cap D_5)\right)$$
$$= P(D_1 \cap D_2 \cap D_3 \cap D_4 \cap \overline{D}_5) + \cdots + P(\overline{D}_1 \cap D_2 \cap D_3 \cap D_4 \cap D_5) = 5\left(\frac{5}{8}\right)^4 \frac{3}{8} = 0.2861.$$

5. Consider a random experiment consisting on tossing twice a biased coin that results in heads 60% of the tosses. Are the events $H_1 \equiv$'head at the first toss' and $E \equiv$'equal result at both tosses' independent? (explain why).

   Denote by H_2 the event head at the second toss, while T_1 and T_2 are respectively tail at the first and second toss. We know $P(H_1) = P(H_2) = 0.6$ and H_1 is independent of H_2. Clearly

   $$P(E) = P((H_1 \cap H_2) \cup (T_1 \cap T_2)) = P(H_1)P(H_2) + P(T_1)P(T_2) = 0.36 + 0.16 = 0.52$$

   and

   $$P(E \cap H_1) = P(H_1 \cap H_2) = P(H_1)P(H_2) = 0.36 \neq 0.312 = 0.52 \times 0.6 = P(E)P(H_1).$$

   Use R package `prob` to simulate 1000 times the previous experiment (set your NIU, DNI, or NIE as seed and type your code and results). What fraction of experiments resulted in head at the first toss?, what fraction of experiments resulted in identical outcomes at both tosses? If we restrict to those experiments with identical outcomes at both tosses, what fraction of them resulted in head at the first toss? Does the empirical evidence about the independence of H_1 and E coincide with your answer to the first question?

```r
library(prob)
S=probspace(tosscoin(2),probs=c(.36,.24,.24,.16))
set.seed(1)
empirical(sim(S,ntrials=1000))
```

```
##   toss1 toss2 probs
## 1     H     H 0.358
## 2     T     H 0.231
## 3     H     T 0.248
## 4     T     T 0.163
```

The proportion of pairs of tosses (out of the 1000) with equal outcomes is 0.521, the proportion of pairs of tosses with head at the first toss is 0.606, while the proportion of pairs of tosses with equal oucomes and head at the first toss (so two heads) is 0.358. Since this last number is clearly different from the product of the first two, the events are not independent.

6. An insurance company has clients classified as high, medium, and low risk. These clients have a probability of claiming equal to 0.02, 0.01, and 0.0025 respectively. If the probability of being a client of high risk is 0.1, 0.2 of medium risk, and 0.7 of low risk, what is the probability that a claim selected at random comes from a high risk client?

Denote the events

$$H \equiv \text{“High risk client”}$$
$$M \equiv \text{“Medium risk client”}$$
$$L \equiv \text{“Low risk client”}$$
$$C \equiv \text{“the client Claims”}$$

with probabilities $P(H) = 0.1$, $P(M) = 0.2$, $P(L) = 0.7$, and $P(C|H) = 0.02$, $P(C|M) = 0.01$, and $P(C|L) = 0.0025$.

$$P(H|C) = \frac{P(C|H)P(H)}{P(C)} = \frac{0.02 \times 0.1}{P(C|H)P(H) + P(C|M)P(M) + P(C|L)P(L)} = \frac{0.002}{0.00575} = 0.3475.$$

7. A company aims to predict the likelihood of a flight arrival delay up to six hours before airlines notify passengers by crunching data on weather, a flight's prior inbound airplane's status, FAA updates, historical data and other information. Assume that 20% of all flights suffer some kind of delay and that the company predicts as delayed 50% of the flights that are actually delayed, while it (wrongly) predicts as delayed 5% of the flights landing on time.

   1. If you are taking three flights on three different days this summer, what is the probability that at least one of them is delayed? In case you needed to make any assumption, explain which.
   2. What percentage of flights are announced to be delayed by the company?
   3. If the company announces a delay at a given flight, what is the probability that it is actually delayed?

- $D \equiv$'Flight is delayed', $P(D) = 0.2$;
- $AD \equiv$'Company announces that the flight is delayed', $P(AD|D) = 0.5$, $P(AD|\overline{D}) = 0.05$;
- $D_i \equiv$'i-th flight is delayed', $P(D_i) = 0.2$, D_is independent if flights are on different days.

   1. Probability that at least one out of three flights (on different days) is delayed: 0.488

$$\begin{aligned} P(D_1 \cup D_2 \cup D_3) &= 1 - P(\overline{D_1 \cup D_2 \cup D_3}) = 1 - P(\overline{D_1} \cap \overline{D_2} \cap \overline{D_3}) \\ &= 1 - P(\overline{D_1})P(\overline{D_2})P(\overline{D_3}) = 1 - (1 - P(D))^3 = 1 - 0.8^3 = 0.488\,. \end{aligned}$$

   2. Percentage of flights announced to be delated: 14%

$$P(AD) = P(AD|D)P(D) + P(AD|\overline{D})P(\overline{D}) = 0.5 \times 0.2 + 0.05 \times 0.8 = 0.14\,.$$

   3. Probability that a flight announced to be delated is actually delayed: 0.714

$$P(D|AD) = \frac{P(AD|D)P(D)}{P(AD)} = \frac{0.5 \times 0.2}{0.14} = \frac{5}{7} = 0.714\,.$$