Probability Theory for EOR 2021/2022

—

Assignment 1

Deadline: Friday December 3rd, 19h00.
Hand-in (handwritten okay, but prefer the typed) PDF report via Nestor.

Instructions

- Make this assignment in groups of two from the same tutorial group.
- Explain your answers. If you just answer $P(A) = 0.0030$ you get 0 points, even if this is the correct answer.
- It is better to define and use clear notation. E.g., prefer not to write:

  $$P(\text{we have two pairs in the first four cards}) = \dots$$

  Instead, define the event A: we have two pairs in the first four cards, and then write

  $$P(A) = \dots$$

- Make sure that it is clear what your final answer is.
- Use *first* a formula and *then* fill in the numbers. So suppose you have defined N_A and N_T and have calculated that $N_A = 10$ and $N_T = 20$. The answer could look something like: "<u>Since the outcomes are equally likely, by the naive definition of probability</u>, $P(A) = N_A/N_T = 10/20 = 1/2$." Note that the **underlined part is a crucial piece** of information. **Because it it important to motivate your steps.**
- During the exam, it is okay if you want to use **definitions/theorems/propositions** from the textbook and you do not need to invoke the exact numbers of the **definitions/theorems/propositions** but only need to rephrase the contents via your own words, same rule applies here.
- Good luck!

1. Suppose now A and B are conditional independent events given C, prove that A and B^c are also conditional independent given C and what can you say about (conditional) independence of A and B^c if $C = S$ (S denotes the sample space of all possible outcomes)?

Solution

Remark: This exercise targets at three folds: (1) understand the conditional probability is also a probability function, and it also satisfies the general definition of probability (thus it has all the structures of a probability function); and these two concepts are closely linked and so are the other relevant concepts; (2) be familiar with the (conditional) independence definition: they are simply functions satisfying special identities; (3) understand when we are dealing with probability theory, they are nothing special but just functions (of course, use the right notations to define these functions).

I. **First possible proof:** From the conditional independence definition we know

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

Since conditional probability is also a probability, the second axiom implies that $P(B|C) = (1 - P(B^c|C))$. The above two equations imply that

$$P(A \cap B|C) = P(A|C)(1 - P(B^c|C)) = P(A|C) - P(A|C)P(B^c|C),$$

which means

$$P(A \cap B|C) = P(A|C)(1 - P(B^c|C)) = P(A|C) - P(A|C)P(B^c|C).$$

From $P(A \cap B|C) = P(A|C) - P(A|C)P(B^c|C)$ and the fact (second axiom of probability of the conditional probability version) $P(A|C) - P(A \cap B|C) = P(A \cap B^c|C)$, we know $P(A \cap B^c|C) = P(A|C)P(B^c|C)$ and thus by the definition of the conditional independence we know A and B^c are also independent given C.

Second possible proof (shorter and smarter one): Since conditional probability is also a probability and now if we regard the conditional probability function $P(\cdot|B)$ as a new probability function $P^*(\cdot)$, then A and B are independent with respect to probability $P^*(\cdot)$. Therefore, according to the proposition (Proposition 2.5.3 from the textbook) such that if A and B are independent then A and B^c are independent, we know A and B^c are also independent with respect to probability $P^*(\cdot)$. Then we know A and B^c are also independent given C.

Remark: The second solution is a bit abstract, but I trust you can understand it.

II. **When** $C = S$, we would have that A and B are independent since $P(D|S) = P(D)$ (knowing S occurred does not provide any extra information). This shows that once we conditional on the whole sample space, we are operating with the same probability, and conditional independence is simply independence. Then the result we prove above would imply that if A and B are independent then A and B^c are also independent.

Remark: When we work with conditionings, we are using the control variates method by narrowing down the sample space to smaller(could be the same as shown here) specific events. Here when we choose a specific event (sample space) to conditional on, we could derive properties of independence from properties of conditional independence.

2. **(I)**. Consider drawing 6 (six!) cards from a well-shuffled standard deck (52 (13×4) cards without jokers). We provide R codes for you to verify the results in 2.(I).(a) and 2.(I).(b). Just like flipping a fair coin many times, the average value will be closer to 1/2 (get 1 for head, 0 for bottom), you can check if the simulated results are closer to your theoretically derived ones.

   (a) Calculate the probability of drawing exactly two pair. Does it make sense to you that this number is higher/lower than the probability of two pair when drawing 5 cards?

Solution Define the event of drawing two pairs in j cards as D_j. There are $N_r = \binom{13}{2}$ ways to draw the ranks of the pairs. Then for each rank, there are $N_p = \binom{4}{2}$ ways to draw the cards with this particular rank. We also need to select the remaining two cards. These need to have a different rank from the others and from each other, so we draw two ranks from the remaining 11 ranks, i.e. $N_d = \binom{11}{2}$. For each of these, we have $N_o = \binom{4}{1}$ options to draw a card. In total, there are $N_t = \binom{52}{6}$ possible hands. Using the naive definition of probability, we conclude that the probability of drawing exactly two pair equals

$$P(D_6) = \frac{N_r N_p^2 N_d N_o^2}{N_t} = \frac{381}{3139} = 0.1214$$

It is not clear cut whether drawing an extra card is favorable to drawing two pair. It can work against you: you might have two pair after five cards and then the extra card can make it three pair. Or it helps: you might not have two pair after five cards, but get two pair on the sixth.

Intuitively though, it feels like drawing an extra card should help. Suppose that you have not drawn two pair with the first five cards. To be able to end up with two pair, the only relevant case is where you have drawn one pair and three cards of different ranks. This one pair hand is quite common, it happens with probability $352/833 \approx 0.426$. Given this hand, $3 \cdot 3 = 9$ out of the remaining 47 cards are favorable to you, still about a 0.192 probability. So drawing one pair in 5 and making that into a two pair on the sixth has probability $0.426 \cdot 0.192 = 0.0809$. This is already higher than the probability of two pair in 5 which we calculated in the exercises, so definitely the extra card is going to help. More precisely, denote by S_j the event of drawing a single pair in j cards, then

$$\begin{aligned} P(D_6) &= P(D_6|D_5)P(D_5) + P(D_6|S_5)P(S_5) \\ &= \frac{40}{47}\frac{198}{4165} + \frac{9}{47}\frac{352}{833} \\ &= 0.1214 \end{aligned}$$

Remark: so to show whether an etra card helps or not, we discuss the case where it does not help (we have two pairs already D_5) and where it does help (turn one pair into two pairs), and given the second case outweight the probability value of event D_5 we know it is benefitial for sure.

(b) Calculate the probability of drawing three pair.

Solution Define the event of drawing three pairs as T. There are $N_r = \binom{13}{3}$ ways to draw the three ranks. Then for each rank, there are $N_p = \binom{4}{2}$ ways to draw the cards. In total, there are $N_t = \binom{52}{6}$ possible hands. Using the naive definition of probability and the multiplication rule, we conclude that

$$P(T) = \frac{N_r N_p^3}{N_t} = 0.0030$$

(c) What is the probability of not drawing two pair with the first four cards and ending up with three pair? You can use that the probability of drawing two pair with the first four cards occurs with probability $\frac{216}{20825} = 0.0104$, and the conditional probability of drawing three pairs given first four cards forming two pairs is $\frac{\binom{11}{1}\binom{4}{2}}{\binom{48}{2}} = 0.0585$.

Solution Denote by D_4^C the complement of D_4. By the law of total probability

$$P(T) = P(T|D_4)P(D_4) + P(T \cap D_4^C)$$
$$\Rightarrow$$
$$P(T \cap D_4^C) = P(T) - P(T|D_f)P(D_4)$$

We now have all the information to find the probability of not drawing a two pair with the first four cards and ending up with three pair,

$$P(T \cap D_4^C) = P(T) - P(T|D_4)P(D_4) = 0.0030 - 0.0585 \cdot 0.0104 = 0.0024.$$

Remark: here is how we calculate $P(T|D_4)$ and thanks Hans Ligtenberg for pointing this solution out:

$$P(T|D_4) = \frac{\binom{11}{1}\binom{4}{2}}{\binom{48}{2}}.$$

The idea behind is that by symmetry, if we know event M occurred that the first two pairs are 7's and 2's should not change the above conditional probability, and thus we can use $P(T|D_4, M)$ to calculate $P(T|D_4)$.

(II). Repeat the exact game described in (I) three times (we draw three 6-card hands with replacement). We use the following notations:

- X_i is the random variable for the 6 cards you draw from the ith game ($i = 1, 2, 3$);

- We say $X_i = X_j$ if they are the same 6 cards.

  Order does not matter, so if you get 777722 (first 2 with the spades, and second 2 with the heart suit) from the first draw, and 727772 from the second draw (first 2 with the heart suit, and second 2 with the spade), then $X_2 = X_1 = \{$*get the following 6-card hand: 227777 (2's with the heart and the spades)*$\}$.

(a) What is $P(X_1 = X_2)$?

Solution There are $\binom{n}{k}$ $(n = 52, k = 6)$ total possible outcomes in S for the ith game $(i = 1, 2, 3)$, denote $\{X_i = a\}$ as the event that ith get the 6-card hand a. Then by LOTP:

$$\begin{aligned}
P(X_1 = X_2) &= \sum_{a \in S} P(X_1 = X_2 | X_1 = a) P(X_1 = a) \\
&= \sum_{a \in S} \frac{1}{\binom{n}{k}} \frac{1}{\binom{n}{k}} \\
&= \binom{n}{k} \frac{1}{\binom{n}{k}} \frac{1}{\binom{n}{k}} = \frac{1}{\binom{n}{k}} \approx 4.91194841e^{-8}.
\end{aligned}$$

Alternatively, you can also consider this as drawing 6-card hands with replacement for two times, and count there are $\binom{n}{k}$ outcomes among the total equally likely $\binom{n}{k}^2$ outcomes, and the result remains the same.

(b) What is the probability that you have at least two X_i's of $N = 3$ drawings that are equal to each other?

Solution

$$\begin{aligned}
P\left(\cup_{i<j}\{X_i = X_j\}\right) &= P\left(\{X_1 = X_2\} \cup \{X_1 = X_3\} \cup \{X_2 = X_3\}\right) \\
&= \sum_m P(A_m) - \sum_{m<n} P(A_m \cap A_n) + P(A_1 \cap A_2 \cap A_3) \\
&=_{\textit{by symmetry}} 3P(A_1) - 3P(A_1 \cap A_2) + P(A_1 \cap A_2 \cap A_3) \\
&= 3P(A_1) - 3P(X_1 = X_2 = X_3) + P(X_1 = X_2 = X_3) \\
&= 3\frac{1}{\binom{n}{k}} - 2\frac{1}{\binom{n}{k}^2} = \frac{3\binom{n}{k} - 2}{\binom{n}{k}^2} \approx 1.47358448e^{-7}.
\end{aligned}$$

where we denote $A_1 = \{X_1 = X_2\}, A_2 = \{X_1 = X_3\}, A_3 = \{X_2 = X_3\}$.

For nearly any problem you'll encounter in business, consultancy, you'll need the computer to do the computations and simulations. The earlier you become comfortable with using the computer, the better.

Here we provide a short discussion in `R`. Later on in other courses, e.g., *Probability distribution* you may need to learn to read these as well. Both `R` and Python are used widely!

R help!

We can use simulation to verify answers 2.(I).(a) and 2.(I).(b). You need to draw N different hands (with replacement) from a deck. Provide the simulated probability for $N = \{10, 100, 1000, 10000\}$ (using 1 decimal precision for $R = 10$, 2 decimals for $N = 100$, 3 decimals for $N = 1000$ and 4 decimals for $N = 10000$. Intuitively, for a given value of N, would you trust your simulation results more for the answer to a or b?

Solution About 90% of the groups will get something between the following numbers for (a)

R	Lower	Upper
10	0	0.3
100	0.07	0.18
1000	0.10	0.14
10000	0.11	0.13

And for (b)

R	Lower	Upper
10	0	0
100	0	0.01
1000	0.001	0.006
10000	0.0021	0.0039

Intuitively, there are two (opposing) answers: (a) can be simulated more precisely for smaller values of R since the event occurs more frequently. (b) seems to be simulated more precisely (the range is much smaller), since for such a rare event 0 is very close to the correct answer. It turns out that the second answer is more precise if we think of precision in terms of the variance (which we will define later). The variance is an absolute measure of deviation from the true values calculated in (a) and (b).

The `R`notebook file also on Nestor. Open it with Rstudio, and the rests come easily as shown by our introduction video on Nestor. Exam of our course this time **would not** involve any codes, but you will need to learn some basic coding ideas for the following course, e.g., Probability Distribution; and when you get stuck somewhere, simulations by codes are also a nice way to explore possible directions.

Each line of code can be executed using ctrl + Enter in Rstudio.

Construct a deck using

```
rank   <- 1:13
suit   <- c(1,1,1,1)
deck   <- kronecker(rank,suit)
```

You now have created a vector with four 1's (representing the four aces), four 2s, …, four 13s. You can get a draw of $k = 6$ cards from the deck using

```
k    <- 6
draw <- sample(deck,k,replace=FALSE)
```

We need to get many of these draws (say $N = 100$ of them), for which we use the `replicate()` function, i.e.

```
N     <- 100
draws <- replicate(N,sample(deck,k,replace=FALSE))
```

This will store the hands in a matrix with 6 rows (the number of cards we draw) and N columns.

Now we are going to count the number of 1s, 2s, ..., 13s that are selected in each hand. Here's the command, we won't get into the details.

```
freq <- t(sapply(1:13,function(a) colSums(draws==a)))
```

If you look at `freq[,1]`, R will show you the frequencies of the numbers 1 to 13 in the first hand. You get two pair when there are exactly two 2s in this vector. So you want how many columns contain exactly two 2s. The command is

```
M    <- freq == 2
TwoP <- sum(colSums(M)==2)
```

What is this command doing? `M <- freq==2` generates a matrix M with 13 rows and N columns. In each column, M has a 1 on the ith row if i appears two times in the hand. So we have two pairs if the column sum (`colSums(M)`) is equal to 2. We then count how many times this happens in the N hands that we have drawn by taking the outer `sum` on the second line of the above piece of code.

```r
N        <- 200000 # Number of hands

# Generate a deck to draw from
rank     <- 1:13
suit     <- c(1,1,1,1)
deck     <- kronecker(rank,suit)

# Draw N hands
draw     <- replicate(N,sample(deck,6,replace=FALSE))

# Get the frequencies
freq     <- t(sapply(1:13,function(a) colSums(draw==a)))

# Probability of exactly two pairs
TwoP     <- sum(colSums(freq==2)==2)/N
# Probability of exactly three pairs
ThreeP   <- sum(colSums(freq==2)==3)/N

# Conditional probabilities
# Count hands with two pair in the first four cards
drawF    <- draw[1:4,]
freqF    <- t(sapply(1:13,function(a) colSums(drawF==a)))
TwoPF    <- sum(colSums(freqF==2)==2)/N
# Count hands with one pair in the last two cards
# given that there were two pair in the first four cards
drawL    <- draw[5:6,]
freqL    <- t(sapply(1:13,function(a) colSums(drawL==a)))
TcDf     <- sum((colSums(freqL==2)==1)*(colSums(freqF==2)==2))/(TwoPF*N)

# Count hands with Three Pair given that we did not get
# two pair in the first four cards
TiDf     <- sum((colSums(freq==2)==3)*(colSums(freqF==2)!=2))/N

# Gather and display results
sim                <- c(TwoP,ThreeP,TwoPF,TcDf,TiDf)
theory             <- c(0.1214, 0.0030,0.0104,0.0585,0.0024)
results            <- cbind(sim,theory,sim/theory)
colnames(results) <- c("Simulation","Theory", "Ratio")
rownames(results) <- c("D", "T","Df","T|Df","T\&Dfc")
print(results)
```

Here are also codes in Python

```
import numpy as np
import random
import collections
N = 2000 # Number of hands to draw (with replacement)
k =6
# Generate a deck to draw from
rank =  np.arange(13)+1
suit = np.ones([1,4])
deck =np.kron(rank, suit)
deck = deck[0]
# Draw N hands
draw=[]

# in r we work with k times N matrix, here we works with N times k matrix, each
# row would store one hand
for i in np.arange(N):
draw.append(random.sample(list(deck),k))
# draw


freq=np.ones([1,N])
i=0
for i in np.arange(13)+1:
freq= np.vstack((freq, np.sum( draw ==i, axis=1)))
freq=freq[1:,:] # 13*N freqency table

# Probability of exactly two pairs
TwoP = np.sum(np.sum(freq==2, axis=0)==2)/N

# Probability of exactly three pairs
ThreeP = np.sum(np.sum(freq==2, axis=0)==3)/N
# Probability of two three of a kinds
TwoTok = np.sum(np.sum(freq==3, axis=0)==2)/N
# Conditional probabilities
# Count hands with two pair in the first four cards
draw=np.array(draw)
drawF = draw[:,[0,1,2,3]]

freqf=np.ones([1,N])
i=0
for i in np.arange(13)+1:
freqf= np.vstack((freqf, np.sum(drawF == i, axis=1)))
freqf=freqf[1:,:] # 13*N freqency table

TwoPF = np.sum(np.sum(freqf==2, axis=0)==2)/N

# Count hands with one pair in the last two cards
# given that there were two pair in the first four cards
drawL =draw[:,[4,5]]
# drawL.shape
freqL=np.ones([1,N])
i=0
for i in np.arange(13)+1:
freqL= np.vstack((freqL, np.sum(drawL == np.ones([N,2])*i, axis=1)))
freqL=freqL[1:,:] # 13*N freqency table

TcDf = np.sum((np.sum(freqL==2, axis=0)==1 )*( np.sum(freqf==2, axis=0)==2))/(TwoPF*N)
# Count hands with Three Pair given that we did not get
# two pair in the first four cards
TiDf  = np.sum((np.sum(freq==2, axis=0)==3 )*( np.sum(freqf==2, axis=0)!=2))/N

from tabulate import tabulate
print(tabulate([["D", TwoP],["T", ThreeP],["Df",TwoPF],["T|Df",TcDf], \
["T&Dfc", TiDf]], headers=['Name', 'Simulation']))
```