Universität Bielefeld
Tanja Pasurek
Tutor: Ehsan Abedi WiSe 2023/24

Foundations of Statistics

Solutions to Homework 3

Part I. Conditional probability and independence

1. Let $A_1, A_2, ..., A_N$ be independent events. Show that the probability that none of the $A_1, A_2, ..., A_N$ occur is less than or equal to

$$\exp\big(-\sum_{n=1}^{N}\mathbb{P}(A_n)\big).$$

Solution: We first note that if A_i and A_j are independent then A_i^c and A_j^c are also independent. Indeed,

$$\begin{aligned}
\mathbb{P}(A_i^c \cap A_j^c) &= \mathbb{P}((A_i \cup A_j)^c)\\
&= 1 - \mathbb{P}(A_i \cup A_j)\\
&= 1 - \mathbb{P}(A_i) - \mathbb{P}(A_j) + \mathbb{P}(A_i \cap A_j)\\
&= 1 - \mathbb{P}(A_i) - \mathbb{P}(A_j) + \mathbb{P}(A_i)\mathbb{P}(A_j)\\
&= [1 - \mathbb{P}(A_i)] - \mathbb{P}(A_j)[1 - \mathbb{P}(A_i)]\\
&= [1 - \mathbb{P}(A_i)][1 - \mathbb{P}(A_j)] = \mathbb{P}(A_i^c)\mathbb{P}(A_j^c).
\end{aligned}$$

Using this, the probability that none of $A_1, A_2, ..., A_N$ occur is given by

$$\begin{aligned}
\mathbb{P}\left(\big(A_1 \cup A_2 \cdots \cup A_N\big)^c\right) &= \mathbb{P}\left(A_1^c \cap A_2^c \cdots \cap A_N^c\right)\\
&= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\cdots\mathbb{P}(A_N^c)\\
&= \Big(1 - \mathbb{P}(A_1)\Big)\Big(1 - \mathbb{P}(A_2)\Big)\cdots\Big(1 - \mathbb{P}(A_N)\Big)\\
&\overset{*}{\leq} \exp(-\mathbb{P}(A_1))\exp(-\mathbb{P}(A_2))\cdots\exp(-\mathbb{P}(A_N))\\
&= \exp\big(-\sum_{n=1}^{N}\mathbb{P}(A_n)\big)
\end{aligned}$$

where the inequality $(*)$ is due to $\exp(-t) \geq 1 - t$ for all $t \in \mathbb{R}$.

2. We roll a die N times. Let A_{ij} be the event that the ith and jth rolls produce the same number. Show that the events A_{ij}, $1 \le i < j \le N$, are pairwise independent but not independent.

Solution: We define a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ as follows

$$\Omega := \Big\{\omega = (\omega_1, \cdots, \omega_N) : \; w_i \in \{1, 2, \cdots, 6\}, \, i \in \{1, 2, \cdots, N\}\Big\},$$
$$\mathcal{A} := 2^\Omega \quad (= \text{power set of } \Omega),$$
$$\mathbb{P} := \text{uniform distribution on } \Omega.$$

Then the event A_{ij} is

$$A_{ij} = \Big\{\omega \in \Omega : \; \omega_i = \omega_j\Big\}, \quad \text{for all } 1 \le i < j \le N.$$

We first compute the probability of A_{ij}

$$\mathbb{P}(A_{ij}) = \frac{6^{N-2} \times 6 \times 1}{6^N} = \frac{1}{6}$$

Step (1). We first show that (A_{ij}) are pairwise independent. Take $i < j$ and $i' < j'$, and observe that

$$\mathbb{P}(A_{ij} \cap A_{i'j'}) \begin{cases} = \mathbb{P}\left(\{\omega \in \Omega : \omega_i = \omega_j, \, \omega_{i'} = \omega_{j'}\}\right) \\ \quad = \frac{6^{N-4} \times 6 \times 1 \times 6 \times 1}{6^N} = \frac{1}{6^2} & i \ne i', i \ne j', j \ne i', j \ne j' \\ = \mathbb{P}\left(\{\omega \in \Omega : \omega_i = \omega_j = \omega_{j'}\}\right) \\ \quad = \frac{6^{N-3} \times 6 \times 1 \times 1}{6^N} = \frac{1}{6^2} & i = i', i \ne j', j \ne i', j \ne j' \\ = \cdots = \frac{1}{6^2} & i \ne i', i = j', j \ne i', j \ne j' \\ = \cdots = \frac{1}{6^2} & i \ne i', i \ne j', j = i', j \ne j' \\ = \cdots = \frac{1}{6^2} & i \ne i', i \ne j', j \ne i', j = j' \end{cases}$$

So we conclude that in all cases

$$\mathbb{P}(A_{ij} \cap A_{i'j'}) = \mathbb{P}(A_{ij})\mathbb{P}(A_{i'j'})$$

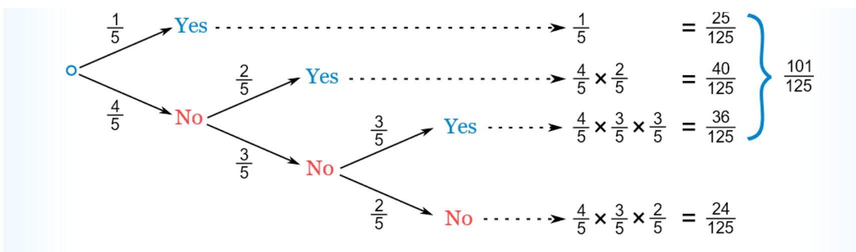
Therefore (A_{ij})s are pairwise independent.

Step (2). We now show that (A_{ij}) are not independent. Take $i < j < j'$ and observe that

$$\mathbb{P}(A_{ij} \cap A_{ij'} \cap A_{jj'}) = \mathbb{P}\left(\{\omega \in \Omega : \omega_i = \omega_j = \omega_{j'}\}\right)$$
$$= \frac{6^{N-3} \times 6 \times 1 \times 1}{6^N} = \frac{1}{6^2} \neq \frac{1}{6^3} = \mathbb{P}(A_{ij})\mathbb{P}(A_{ij'})\mathbb{P}(A_{jj'}).$$

3. (*Friends and random numbers*) 4 friends (**A**lex, **B**lake, **C**hris and **D**usty) each choose a random number between 1 and 5. What is the chance that at least two of them chose the same number?

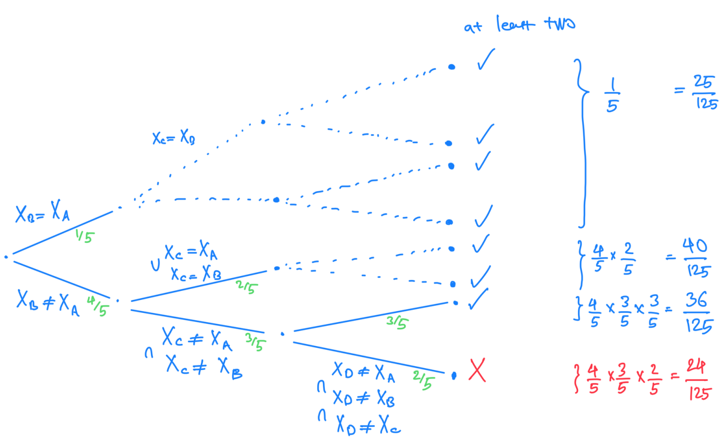
(a) Get the answer p=101/125 using the following **tree diagram** for calculating conditional probabilities.



Solution: We define random variables

$$X_A, X_B, X_C, X_D : \Omega \to \{1, 2, 3, 4, 5\}.$$

More precisely, the events on the tree diagram above are as follows



For example, on the first branch, we have

$$\begin{aligned}\mathbb{P}(X_A = X_B) &= \mathbb{P}(X_A = X_B = 1) + \mathbb{P}(X_A = X_B = 2) + \cdots \mathbb{P}(X_A = X_B = 5)\\ &= \mathbb{P}(X_A = 1)\mathbb{P}(X_B = 1) + \cdots + \mathbb{P}(X_A = 5)\mathbb{P}(X_B = 5)\\ &= \frac{1}{5} \times \frac{1}{5} + \cdots + \frac{1}{5} \times \frac{1}{5}\\ &= 5 \times \frac{1}{25} = \frac{1}{5}\end{aligned}$$

and the rest follows by definition of conditional probability. Indeed the chance that at least two of them chose the same number is $p = 101/125$.

(b) But here is something interesting... If we follow the "*No*" path, we can directly calculate the probability $1 - p$ of the complement event and make our life easier. Realise this idea.
Solution: We have

$$\begin{aligned}\mathbb{P}(\{\text{at least two the same}\}) &= 1 - \mathbb{P}(\{\text{at least two the same}\}^c)\\ &= 1 - \mathbb{P}(\{\text{none of them the same}\})\\ &= 1 - \frac{24}{125} = \frac{101}{125} = 80.8\%\end{aligned}$$

(c) Perform a computer simulation in `R` playing this game $n = 1000$ rounds and estimating the probability p. (You can use the function `sample`.)
Solution: `R` code:

```
n <- 1000
c <- 0
set.seed(1)
for (i in 1:n){
  s <- sample(1:5,size=4,replace=TRUE)
  if(length(unique(s)) == 4){
    c <- c + 1 # count how many times they all have different numbers
  }
}
print((n-c)/n*100)
[1] 81.5
```

Remark: In the function `sample()`, if `replace` is FALSE, these probabilities are applied sequentially, that is the probability of choosing the next item is proportional to the weights amongst the remaining items. This is not what we want here.

4. (*Birthday problem*) In a room there are n people. What is the probability that at least two of them have a common birthday?

(a) Give an answer for a year with 365 days, assuming that every day of the year is equally likely to be a birthday.

Solution: we have

$$\begin{aligned} p := \mathbb{P}(\{\text{at least two the same}\}) &= 1 - \mathbb{P}(\{\text{at least two the same}\}^c) \\ &= 1 - \mathbb{P}(\{\text{none the same}\}) \\ &= 1 - \frac{365 \times 364 \times \cdots \times (365 - (n-1))}{365^n} \\ &= 1 - \frac{\frac{365!}{(365-n)!}}{365^n} \end{aligned}$$

(b) Provide a numerical estimation for $n = 3$ and $n = 25$ (e.g. the number of students in a class).

Hint: First calculate the probability of the complement event. Think about the event that no two persons have the same birthday or equivalently that they all have different birthdays.

Solution: we have

$$\begin{aligned} n = 3 \quad &\Rightarrow \quad p = 0.820\% \\ n = 25 \quad &\Rightarrow \quad p = 56.9\% \end{aligned}$$

5. An insurance company insures an equal number of male and female drivers. In any given year the probability that a male driver has an accident involving a claim is α, independently of other years. The analogous probability for females is β. Assume the insurance company selects a driver at random.

(a) What is the probability that the selected driver will make a claim this year?

Solution: In the following we consider the events

$$\begin{aligned} &\text{“}M\text{”} = \{\text{driver is male}\}, && \text{“}F\text{”} = \{\text{driver is female}\}. \\ &\text{“}C\text{ ”} = \{\text{driver makes a claim}\}, && \text{“}N\text{ ”} = \{\text{driver makes no claim}\}. \end{aligned}$$

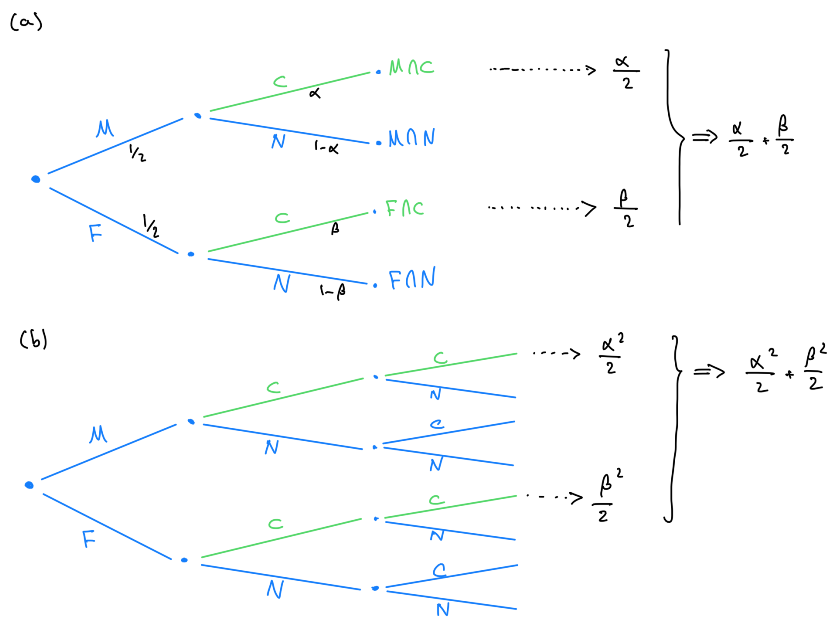
The information we have is

$$\mathbb{P}(M) = \mathbb{P}(F) = \frac{1}{2}$$
$$\mathbb{P}(C|M) = \alpha, \quad \mathbb{P}(C|F) = \beta$$

Using the law of total probability, we obtain

$$\mathbb{P}(C) = \mathbb{P}(C|M)\mathbb{P}(M) + \mathbb{P}(C|F)\mathbb{P}(F)$$
$$= \frac{\alpha}{2} + \frac{\beta}{2}$$

Alternatively, one can get this answer using the tree diagram below



$$\mathbb{P}(C) = \mathbb{P}(C \cap M) + \mathbb{P}(C \cap F)$$
$$= \frac{\alpha}{2} + \frac{\beta}{2}$$

(b) What is the probability that the selected driver makes a claim in two subsequent years?
Solution: again using the tree diagram above, we obtain the answer

$$\frac{\alpha^2}{2} + \frac{\beta^2}{2}$$

(c) Let A_1, A_2 be events that a randomly chosen driver makes a claim in each of the 1st and 2nd years, respectively. Show that $P(A_2|A_1) \geq P(A_1)$.
Solution: We have

$$\mathbb{P}(A_1) = \frac{\alpha+\beta}{2}, \quad \mathbb{P}(A_1 \cap A_2) = \frac{\alpha^2+\beta^2}{2}$$

by part (a) and (b), respectively. Therefore, we get

$$\mathbb{P}(A_2|A_1) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} = \frac{\alpha^2+\beta^2}{\alpha+\beta}.$$

As a result, we observe that

$$\begin{aligned}\mathbb{P}(A_2|A_1) - \mathbb{P}(A_1) &= \frac{\alpha^2+\beta^2}{\alpha+\beta} - \frac{\alpha+\beta}{2} \\ &= \frac{(\alpha-\beta)^2}{2(\alpha+\beta)} \geq 0.\end{aligned}$$

(d) Find a probability that a claimant is female?
Solution: By Bayes' theorem

$$\mathbb{P}(F|C) = \frac{\mathbb{P}(C|F)\mathbb{P}(F)}{\mathbb{P}(C)} = \frac{\beta \times \frac{1}{2}}{\frac{\alpha+\beta}{2}} = \frac{\beta}{\alpha+\beta}.$$

Part II. Discrete random variables

6. Let $X : \Omega \to \mathbb{N}$ be an (integrable) integer-valued random variable. Show that

$$\mathbb{E}(X) = \sum_{n\geq 1}^{\infty} \mathbb{P}(X \geq n).$$

Solution: We have

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=1}^{\infty} k\,\mathbb{P}(X=k) \\ &= \sum_{k=1}^{\infty}\left(\sum_{n=1}^{k} 1\right)\mathbb{P}(X=k) \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{k}\mathbb{P}(X=k) \\ &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty}\mathbb{P}(X=k) \\ &= \sum_{n=1}^{\infty}\mathbb{P}(X\geq k)\end{aligned}$$

7. This exercise is about the casino game *Chuck-a-Luck* (also known as "*Glückswurf*").

This is a game of chance played with 3 standard dice. In the simplest variant, the rules are as follows:

- The player chooses one number, say a, from $\{1, 2, 3, 4, 5, 6\}$.
- The player pays a stake of \$1 and rolls three dice.
- If none of the dice show the number a, the bet is lost.
- If at least one of the dice shows the number a, the player receives the bet back and one additional dollar for each die that shows this number.

(a) Consider a random variable $X =$"player's profit" per game. Determine the probability mass function $f(x) := \mathbb{P}(X = x)$.

Solution (1): The random variable X takes the following values

$$X : \Omega \to \{-1, 1, 2, 3\},$$

where -1 means that the bet is lost.

Define $A := \{1, 2, 3, 4, 5, 6\} \setminus \{a\}$. So $|A| = 5$

X			#	$\mathbb{P}(X = x)$
-1	(i, j, k)	$i, j, k \in A$	$5 \times 5 \times 5 = 125$	$125/216$
1	(a, i, j) (i, a, j) (i, j, a)	$i, j \in A$	$3 \times (5 \times 5) = 75$	$75/216$
2	(a, a, i) (a, i, a) (i, a, a)	$i \in A$	$3 \times 5 = 15$	$15/216$
3	(a, a, a)		1	$1/216$

since the number of all possibilities is $6 \times 6 \times 6 = 216$.

Solution (2): The random variable X takes the following values

$$X : \Omega \to \{-1, 1, 2, 3\},$$

where -1 means that the bet is lost.

Define a new random variable S to be the number of successes

$$S : \Omega \to \{0, 1, 2, 3\}$$

It is clear that

$$S \sim Binomial\Big(n = 3, p = \frac{1}{6}\Big)$$

and we know that

$$\mathbb{P}(S = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Therefore we have

$$\begin{aligned}
\mathbb{P}(X = -1) =& \mathbb{P}(S = 0) = \frac{125}{216} \\
\mathbb{P}(X = +1) =& \mathbb{P}(S = 1) = \frac{75}{216} \\
\mathbb{P}(X = +2) =& \mathbb{P}(S = 2) = \frac{15}{216} \\
\mathbb{P}(X = +3) =& \mathbb{P}(S = 3) = \frac{1}{216}
\end{aligned}$$

(b) Calculate the mean $\mathbb{E}(X)$. Is this game fair?
Solution: We simply compute

$$\mathbb{E}[X] = -1 \times \frac{125}{216} + 1 \times \frac{75}{216} + 2 \times \frac{15}{216} + 3 \times \frac{1}{216} = \frac{-17}{216} = -0.0787 < 0.$$

and conclude that the game is NOT fair. (Don't play this game!)

(c) Now use the `loop` function to simulate the game $n = 10\,000$ and $100\,000$ rounds. In the process we count how much profit we make overall and especially on average per game. You can proceed as follows:

```
nloop<-10000
a<-5
Win<-rep(NA,nloop)
for (k in 1:nloop){
 Dice<-sample(1:6,size=3,replace=TRUE)
 Count_a<-sum(Dice==a)
 Win[k]<-ifelse(Count_a==0,-1,Count_a)
}
sum(Win) ## overall
sum(Win)/nloop  ## on average per game
```

Solution: see output in the next page.

(d) With the following code, you can visualise the development of the average profit over the 100,000 runs.

```
options(scipen=999)
plot(cumsum(Win)/(1:nloop),type="l",bty="n",
ylab="Average Profit",xlab="Number of Rounds")
abline(h=-17/216,col=2,lty=2)
```

Remark: To set the use of *scientific notation* for large numbers ("e notation", e.g. `1e+05` instead of `10000`), you can use the `scipen` option. You can turn it off with `options(scipen = 999)` and back on again with `options(scipen = 0)`.

Solution: see output in next page.

```r
  # ----- part (a)
dbinom(0, size =3, prob=1/6)*216
```

```
## [1] 125
```

```r
dbinom(1, size =3, prob=1/6)*216
```

```
## [1] 75
```

```r
dbinom(2, size =3, prob=1/6)*216
```

```
## [1] 15
```

```r
dbinom(3, size =3, prob=1/6)*216
```

```
## [1] 1
```

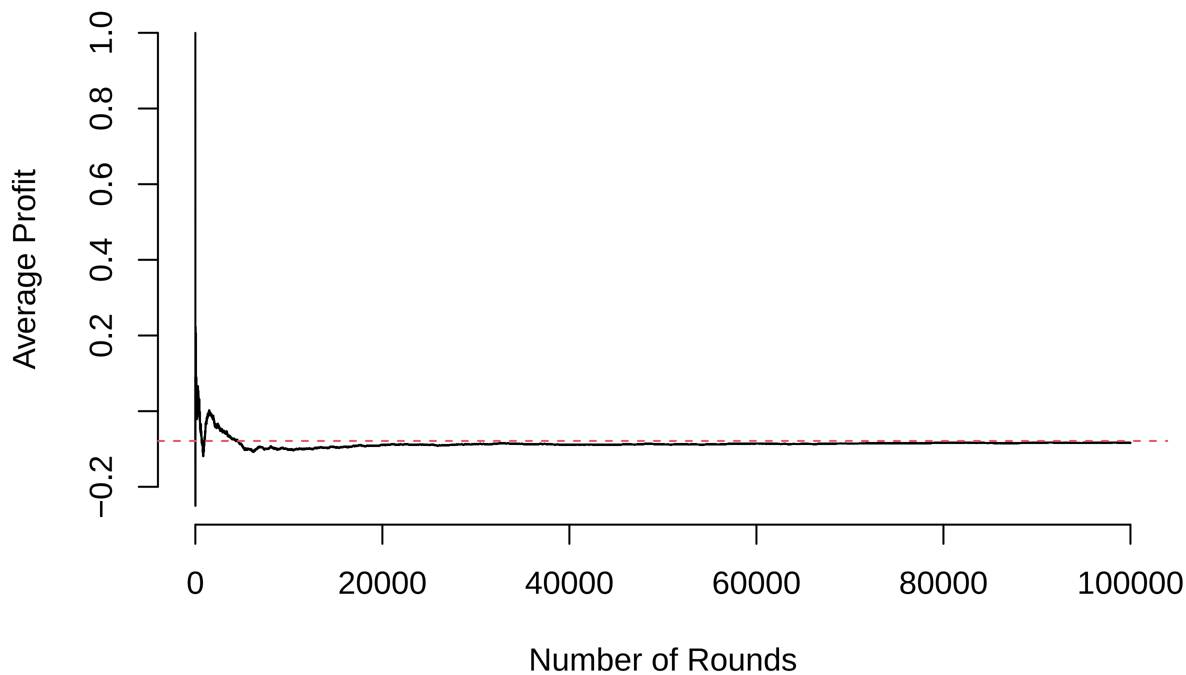
```r
# part (c)
set.seed(2)
nloop<-100000 # 10000 or 100000
a<-5
Win<-rep(NA,nloop)
for (k in 1:nloop){
  Dice<-sample(1:6,size=3,replace=TRUE)
  Count_a<-sum(Dice==a)
  Win[k]<-ifelse(Count_a==0,-1,Count_a)
  }
sum(Win) ## overall
```

```
## [1] -8379
```

```r
sum(Win)/nloop ## on average per game
```

```
## [1] -0.08379
```

```r
# part (d)
options(scipen=999)
plot(cumsum(Win)/(1:nloop),type="l",bty="n",
     ylab="Average Profit",xlab="Number of Rounds")
abline(h=-17/216,col=2,lty=2)
```



Average Profit
1.0
0.8
0.6
0.4
0.2
−0.2
0
20000
40000
60000
80000
100000
Number of Rounds