Universität Bielefeld
Tanja Pasurek
Tutor: Ehsan Abedi

WiSe 2023/24

Foundations of Statistics

Homework 3

Part I. Conditional probability and independence

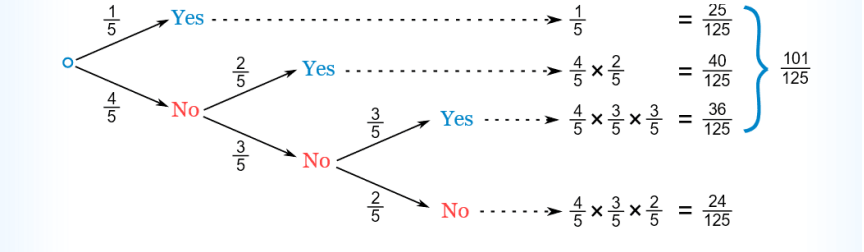
1. Let $A_1, A_2, ..., A_N$ be independent events. Show that the probability that none of the $A_1, A_2, ..., A_N$ occur is less than or equal to

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

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}, $1 \leq i < j \leq N$, are pairwise independent but not independent.

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.



(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.

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

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.

(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.

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?

(b) What is the probability that the selected driver makes a claim in two subsequent years?

(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 $P(A_2|A_1) \geq P(A_1)$.

(d) Find a probability that a claimant is female?

Part II. Discrete random variables

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

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

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 $f(x) := \mathbb{P}(X = x)$.

(b) Calculate the mean $\mathbb{E}(X)$. Is this game fair?

(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
```

(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)`.