Wahrscheinlichkeitstheorie und Statistik

Serie 4

Version 1.1 (18. März 2022: Erinnerung an Lernkontrolle hinzugefügt); Version 1 (14. März)

Bitte stellt Fragen in den Übungen und/oder im Forum des Moodle-Kurs (und/oder (anonym) in diesem file `https://docs.google.com/document/d/1CfTkwrNOhTKB8y8cVQW_deUahCejTbQWlm_BzrHIFYA/edit?usp=sharing`)

Bitte stelle sicher, das du die Webseite `https://kahoot.it/` in der Übung am **21. März** öffnen kannst (zB auf deinem smartphone).

Wir empfehlen die Aufgaben selbständig zu lösen und dann im Fach der entsprechenden Übungsgruppe im Raum HG G 53 abzugeben oder selbst mit der Lösung zu vergleichen am besten rechtzeitig vor der Übung am **21. März**.

Vergiss nicht am 28. März pünktlich um spätestens 13:15 in deine Übungsgruppe zur Lernkontrolle zu kommen (im gleichen Raum wo auch sonst immer die Übungsgruppen statt finden, beziehungsweise in LFW C4 für Gruppe 6).

Aufgabe 4.1 Morgen wird es entweder ausschliesslich regnen oder schneien. Die Wahrscheinlichkeit, dass es regnen wird ist $\frac{2}{5}$ und die Wahrscheinlichkeit für Schnee liegt bei $\frac{3}{5}$. Sollte es regnen, dann ist die Wahrscheinlichkeit zu spät zur Vorlesung zu kommen $\frac{1}{5}$, während die Wahrscheinlichkeit bei Schneewetter bei $\frac{3}{5}$ liegt.Wie wahrscheinlich ist es zu spät in der Vorlesung zu erscheinen?

Aufgabe 4.2 (Simpson's paradox).

We are interested in studying the probability of success of a student at an entrance exam for two departments of a university. Consider the following events:

$$\begin{aligned} A &:= \{\text{The student is a man}\} \\ A^c &= \{\text{The student is a woman}\} \\ B &:= \{\text{The student applied for department I}\} \\ B^c &= \{\text{The student applied for department II}\} \\ C &:= \{\text{The student was accepted}\} \\ C^c &= \{\text{The student was not accepted}\} \end{aligned}$$

We assume the following probabilities:

$$\begin{gathered} \mathbb{P}[A] = 0.73, \\ \mathbb{P}[B \mid A] = 0.69, \mathbb{P}[B \mid A^c] = 0.24, \\ \mathbb{P}[C \mid A \cap B] = 0.62, \mathbb{P}[C \mid A^c \cap B] = 0.82, \\ \mathbb{P}[C \mid A \cap B^c] = 0.06, \mathbb{P}[C \mid A^c \cap B^c] = 0.07. \end{gathered}$$

(a) Draw a tree describing the situation with the probabilities associated.

(b) From examining the probabilities in the tree, do you think that women are disadvantaged in the selection process? Why or why not?

(c) Calculate $\mathbb{P}[C \mid A]$ and $\mathbb{P}[C \mid A^c]$, i.e., the acceptance probabilities for men and women. Does this agree with your answer in (b)? Can you explain what is going on?

Aufgabe 4.3 (Monty Hall problem). You are on a game show, and you are given the choice of three doors. Behind one door is a car, behind the others are goats. You pick a door and the host, who knows what is behind the doors, opens another, behind which is a goat. He then asks you, "Do you want to keep your initial chosen door or do you want to switch to the other one?". Assuming that you like cars but not goats, what should you do?

(a) Construct a suitable model where you can answer this question with the help of conditional probabilities.

(b) Try to find an alternative solution (which of course must give the same answer).

Aufgabe 4.4 Es ist sicher, dass ein betimmter Patient p eine der Krankheiten k_1, k_2 or k_3 hat. Um herauszufinden, welche dieser Krankheiten den Patienten befallen haben, werden zwei Tests hintereinander ausgeführt. Diese Tests haben als Ergebnis entweder positiv (+) oder negativ (–). In der Tabelle unten bedeutet + –, dass der erste Test positiv war und der zweite Test negativ, wobei der Patient die Krankheit auch wirklich hatte (analog definieren wir + +, – +, und – –). Wir erhalten bei 10'000 Patienten die folgende Tabelle:

Krankheit	Anzahl von Patienten mit der jeweiligen Krankheit	Testergebnisse + +	+ −	− +	− −
k_1	3'215	2'110	301	704	100
k_2	2'125	396	132	1'187	410
k_3	4'660	510	3'568	73	509
Insgesamt	10'000				

(a) Wie hoch ist die Wahrscheinlichkeit, dass der Patient p die Krankheit k_i, $i = 1, 2, 3$ hat, bevor der Test durchgeführt wird?

(b) Wie hoch ist die Wahrscheinlichkeit, dass der Patient p die Krankheit k_3 hat, unter der Voraussetzung, dass beide Tests positiv waren? Was gilt wenn beide negativ waren?

Weitere interessante Paradoxa zu bedingten Wahrscheinlichkeiten sind hier zu finden: https://towardsdatascience.com/the-false-positive-paradox-f86448a524bc

Wenn du Feedback zum Übungszettel hast, schreibe bitte eine Mail an Jakob Heiss.