



# 1. Übungszettel zur LV Computational Health Informatics (WiSe 2020/21)

Ausgabe Mi, 28. Oktober 10 Uhr

**Abgabe der Lösungen ist bis Di, 3. November 14 Uhr möglich**

→ [exercise-chi@chi.uni-hannover.de](mailto:exercise-chi@chi.uni-hannover.de)

## **Aufgabe 1.1: Vorhersagewerte eines diagnostischen Verfahrens und Prävalenz (7 Punkte)**

Nehmen Sie an, dass ein diagnostischer Test eine Sensitivität von 99% und eine Spezifität von 99,5% hat. Die Wahrscheinlichkeit, dass eine infizierte Person fälschlicherweise ein negatives Ergebnis erhält, ist also 1%. Die Wahrscheinlichkeit, dass eine nicht-infizierte Person ein falsch positives Ergebnis erhält, ist 0,5%. Interessanter als diese Wahrscheinlichkeiten sind in der Praxis aber die **Vorhersagewerte**, also die Wahrscheinlichkeiten, dass ein Testergebnis den richtigen Krankheitsstatus anzeigt:

- **Positiver Vorhersagewert:** Wahrscheinlichkeit, dass eine Person mit einem positiven Testergebnis wirklich infiziert ist
- **Negativer Vorhersagewert:** Wahrscheinlichkeit, dass eine Person mit einem negativen Testergebnis tatsächlich nicht infiziert ist

Nehmen Sie an, die Krankheitshäufigkeit (Prävalenz) ist sehr gering: 0,00001. Ohne zu rechnen: Was erwarten Sie, wie gut ist bei der gegebenen Sensitivität und Spezifität der positive und negative Vorhersagewert (nahe bei 1 oder wie weit darunter)?



Überprüfen Sie Ihre Vermutung an einer Beispielrechnung mit 100.000 Personen, indem Sie die Werte in der folgenden Tabelle berechnen:

	positiver Befund	negativer Befund	Summe
Infiziert			
Nicht infiziert			
Summe			

Berechnen Sie aus diesen Zahlen den positiven und negativen Vorhersagewert über die relativen Häufigkeiten:

- Positiver Vorhersagewert = Anteil von Infizierten mit richtigem (positiven) Befund relativ zu allen Personen mit einem positiven Befund
- Negativer Vorhersagewert = Anteil von Nicht-Infizierten mit richtigem (negativen) Befund relativ zu allen Personen mit einem negativen Befund

Haben Sie mit Ihrer Schätzung richtig gelegen? Falls nicht: Woran kann es gelegen haben?

Berechnen Sie die Tabelle und die beiden Vorhersagewerte noch einmal für größere Prävalenzen von 0,001 und 0,1. Ergibt sich ein wesentlicher Unterschied? Was schließen Sie für die Praxis aus diesen Berechnungen bzw. Ergebnissen?

## Aufgabe 1.2: Medizinische Studien (4 Punkte)

Es soll mit einer Studie untersucht werden, ob ein Zusammenhang zwischen Rauchgewohnheiten und dem Auftreten von Lungenkrebs besteht. Geben Sie den Typ und die Bezeichnung (bspw. Diagnosestudie, siehe Vorlesung) der im Folgenden kurz beschriebenen möglichen medizinischen Studien an:

- a. Es werden zwei Stichproben verglichen, bestehend aus nicht-erkrankten und an Lungenkrebs erkrankten Personen bezüglich ihrer Rauchgewohnheiten.
- b. Es werden große Populationen über mehrere Jahre beobachtet, die aus Nichtrauchern, schwachen, mäßigen und starken Rauchern bestehen. Ziel ist es herauszufinden, bei welchen Personen wie häufig Lungenkrebs auftritt.



- c. Man gibt jeder Person aus einer großen Population zufällig vor, ob und wieviel diese rauchen soll und wartet dann auf das Eintreten der Zielgröße.
- d. Man setzt Laborratten Zigarettenrauch aus (bzw. auch nicht) und wartet dann auf das Eintreten der Zielgröße.