

ÜBUNGEN ZUR WAHRSCHEINLICHKEITSRECHNUNG UND STATISTIK FÜR BIOLOGEN

Blatt 13

Aufgaben 1 bis 5 sollen (ohne Verwendung von R) mit Taschenrechner gelöst werden. Quantiltafeln finden Sie auf der Vorlesungs-Homepage.

1. Aufgabe Im Jahr 2003 wurden im Sichote-Alin-Naturreservat einige männliche, ausgewachsene Sibirische Tiger gewogen und ihr Nachwuchs markiert. 2011 wurden vier der damals markierten männlichen, inzwischen ausgewachsenen Nachkommen aufgespürt und ebenfalls gewogen. In der folgenden Tabelle stehen jeweils die Gewichte von Vater und Sohn untereinander.

Gewicht Vater (2003) in kg	190	220	220	250
Gewicht Sohn (2011) in kg	190	230	240	260

- Anhand der Daten soll untersucht werden, ob sich das mittlere Gewicht der männlichen Tiger im Reservat von der vorherigen zur jetzigen Generation verändert hat. Führen Sie einen geeigneten Test mit Signifikanzniveau $\alpha = 0.05$ durch. Stellen Sie den Rechenweg nachvollziehbar dar und formulieren Sie Ihr Ergebnis in einem adäquaten Satz.
- Um wieviele kg ist ein heute im Reservat lebendes ausgewachsenes Tiger Männchen im Durchschnitt leichter/schwerer als sein Vater? Berechnen Sie für diesen Mittelwert ein 95%-Konfidenzintervall.

2. Aufgabe Angenommen, die Mittelwerte in den in der Tabelle in Aufgabe 1 gegebenen Stichproben stimmen genau mit den entsprechenden Mittelwerten in den Populationen der Jahre 2003 und 2011 überein. Wieviele markierte Tigersöhne von verschiedenen Vätern müssten Sie dann 2011 aufspüren, damit der Unterschied mit fünfundneunzigprozentiger Wahrscheinlichkeit auf einem Signifikanzniveau von $\alpha = 0.01$ statistisch nachgewiesen werden kann?

3. Aufgabe Die Therapie der Influenza mit Neuraminidasehemmern wie Zanamivir oder Oseltamivir sollte innerhalb von 48 Stunden nach Auftreten der ersten Krankheitssymptome beginnen. Da eine laborbasierte Diagnostik zu lange dauern würde, ist es notwendig, kurzfristig eine gesicherte Diagnose zu stellen. Hier bietet sich der Einsatz eines Influenza-Schnelltests an. Einer dieser Schnelltests hat eine Sensitivität von 80% (d.h. er schlägt bei 80% der Infizierten an) und eine Spezifität von 85% (d.h. er schlägt bei 15% der Nicht-Infizierten an). Diese Werte hängen von den genauen Umständen ab. Beispielsweise kann Nahrungsaufnahme kurz vor dem Schnelltest bei einem Rachenabstrich zu negativen Ergebnissen führen, da dann die Anzahl der Viren im Rachenraum verringert ist. Frau Meyer liest einen Bericht über die Krankheit in der Zeitung, macht sich Sorgen und lässt am nächsten Tag einen Influenza-Schnelltest durchführen. Das Testergebnis ist negativ. Wie wahrscheinlich ist es, dass Frau Meyer dennoch mit Influenza infiziert ist? Sehen Sie Frau Meyer hierfür als zufällig aus der Population gegriffene Person an. Nehmen Sie desweiteren an, dass in der Population einer von 8000 mit Influenza infiziert ist. Stellen Sie Ihren Rechenweg mit allen verwendeten Formeln nachvollziehbar dar und formulieren Sie einen Antwortsatz.

4. Aufgabe Für ein Gen, für das vier Allele A, B, C und D unterscheidbar sind, wurden in einer zufälligen Stichprobe von 100 Individuen die verschiedenen Genotypen in folgenden Anzahlen gefunden:

AA	AB	AC	AD	BB	BC	BD	CC	CD	DD
26	11	12	9	10	13	7	4	5	3

- Berechnen Sie das 95%-Wald-Konfidenzintervall für die Häufigkeit des Allels A in der Population.
- Es stellt sich die Frage, ob sich die Population im Hardy-Weinberg-Gleichgewicht befindet. Führen Sie hierzu einen geeigneten χ^2 -Test durch (inklusive Nullhypothese und Antwortsatz). Sie dürfen dabei verwenden, dass der Wert der Teststatistik gleich $X^2 = 14.32$ ist. Erläutern Sie aber, was der Genotyp AA zu diesem X^2 -Wert beiträgt.

5. Aufgabe Wir kommen noch einmal auf die Sibirischen Tiger zurück.

- Berechnen Sie die Regressionsgerade für die Daten aus Aufgabe 1, um aus dem Gewicht eines Tigers von 2003 vorherzusagen, wie schwer seine heute lebenden Söhne sind.
- Welche Vorhersage macht das Modell für das Gewicht des heute lebenden Sohns eines Tigers, der 2003 ausgewachsen war und 220 kg wog?
- Es stellt sich die Frage, ob beim Gewicht bei Tigern die Vererbung überhaupt eine Rolle spielt oder ob die in den Daten beobachtete Abhängigkeit zwischen den Gewichten der Väter und Söhne auch rein zufällig sein könnten. Führen Sie mit den Daten aus Aufgabe 5 einen geeigneten Test mit Signifikanzniveau $\alpha = 0.05$ durch. Stellen Sie den Rechenweg nachvollziehbar dar und formulieren Sie Ihr Ergebnis in einem Satz.

6. Aufgabe Jetzt für alle, die bereits optimal auf die Klausur vorbereitet sind, doch noch eine Aufgabe zur Diskriminanzanalyse, die mit R zu bearbeiten ist: In R erhalten Sie mit `data(iris)` einen Datensatz `iris`, der morphologische Messungen für drei verschiedene Arten von Schwertlilien enthält. Wie zuverlässig lassen sich *Iris versicolor* und *Iris virginica* allein anhand der Blütenblattlänge (`Petal.Length`) unterscheiden, wenn Sie die in der Vorlesung besprochene Methode verwenden, die auf Normalverteilungsannahmen beruht. (Hinweis: die Dichte einer Normalverteilung können Sie in R mit `dnorm` berechnen.)