

Mathematik II für Biologen

Übungsblatt 9 (Abgabe am 19.06.2015)

Aufgabe 31 (Fortsetzung von Aufgabe 11)

(10 Punkte)

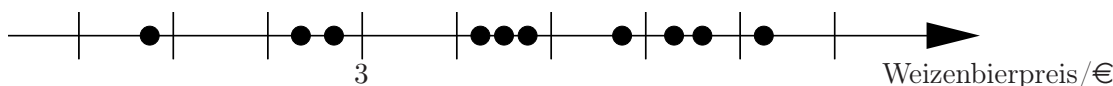
In Aufgabe 11 wurde die Anzahl der google-Treffer für den Ausdruck “ k Katze(n) und h Hund(e)” ermittelt (Tabelle dort). Mit einem χ^2 -Test möchten wir feststellen, ob die Anzahl der Katzen und die Anzahl der Hunde¹⁰ unabhängig voneinander ist. Mit den Daten für $1 \leq h, k \leq 5$ ergibt sich $\chi^2 \approx 70000$ (MATLAB-Berechnung siehe Webforum). Auf dem Signifikanzniveau $\alpha = 5\%$ beträgt der kritische Wert hier ca. 27. (Warum? Und warum verwenden wir nicht auch die Daten für $h, k > 5$?)

- Was folgern wir aus diesem χ^2 -Test über die Abhängigkeit oder Unabhängigkeit der Anzahl der Katzen und der Hunde?
- Führen Sie für Ihr google-Experiment aus Aufgabe 11b (oder für ein neues Experiment, wenn Sie möchten) ebenfalls einen solchen χ^2 -Test durch, und interpretieren Sie das Ergebnis. Geben Sie unbedingt auch den gegoogelten Ausdruck und die Rohdaten an.

Aufgabe 32 (Vorzeichentest)

(10 Punkte)

Eine Kommilitonin behauptet, dass ein Weizenbier in Tübinger Gaststätten im Schnitt 3€ kostet. Sie vermuten aber, dass der Median μ der Verteilung der Weizenbierpreise größer ist. Um dies zu überprüfen, ermitteln Sie die Preise in einigen Gaststätten, stellen die Werte im folgenden Diagramm durch Punkte (•) dar,



und zählen die Anzahl T der Gaststätten, in denen der Weizenbierpreis größer als 3€ war, um einen Vorzeichentest¹¹ durchzuführen.

- Geben Sie Nullhypothese und Alternativhypothese an.
- Welche Werte kann die Teststatistik T theoretisch annehmen? Welchen Wert hat T für Ihre Stichprobe?
- Wie ist T verteilt, wenn die Nullhypothese stimmt? Geben Sie dazu eine Formel für $P[T = k]$ für alle in (a) bestimmten möglichen Werte k von T an.
- Berechnen Sie den zugehörigen p -Wert des Tests.

Aufgabe 33

(10 Punkte)

Die Forschergruppe *Publischi* betreibt Data-Mining.¹² An einem großen Datensatz testen die Wissenschaftler 10 unterschiedliche Nullhypothesen auf dem Signifikanzniveau $\alpha = 5\%$. Eine dieser Nullhypothesen wird verworfen. Die Gruppe *Publischi* veröffentlicht einen Artikel über die zugehörige Alternativhypothese, die durch ihre Arbeit statistisch bewiesen wurde.

Wir nehmen an, dass die Nullhypothesen so unterschiedlich waren, dass die Ausgänge der 10 Hypothesentests unabhängig voneinander waren. Weiter nehmen wir an, dass die Wahrscheinlichkeit, eine wahre Nullhypothese zu verwerfen, jeweils 5% betrug. (Warum?)

- Falls alle Nullhypothesen wahr sind, wie groß war die Wahrscheinlichkeit, dass dennoch mindestens eine von ihnen verworfen wird?
- Kommentieren Sie kritisch das Vorgehen von *Publischi*.



xkcd.com/882/

¹⁰Ob es sich dabei um die Anzahl der Katzen und Hunde eines Tierhalters, in einem Foto oder um irgendetwas anderes handelt, wissen wir nicht. Dazu hätten wir die Daten sorgfältiger ermitteln müssen.

¹¹Das ist nichts weiter, als ein Spezialfall eines Binomialtests, wie Sie gleich sehen werden.

¹²vgl. z.B. <http://de.wikipedia.org/wiki/Data-Mining>

Aufgabe 34

(10 Punkte)

Das Pharmaunternehmen ANTIQUARTIS preist das neue Mittel PASTOFEBRIL gegen Weidefieber bei Kühen an. Ein Landwirt probierte dieses Mittel an seiner Kuh Thekla aus, die an Weidefieber erkrankt war. Daraufhin wurde Thekla gesund. Nun ist auch die Kuh Elma an Weidefieber erkrankt.

Geben Sie aufgrund der Beobachtung an Thekla ein 95%-Vertrauensintervall für die Wahrscheinlichkeit p an, dass auch Elma gesund wird, wenn sie mit PASTOFEBRIL behandelt wird. Führen Sie dazu einen einseitigen Binomialtest durch (ohne jedoch die in der Vorlesung angegebenen Faustregeln anzuwenden, da $n = 1$ nun wirklich nicht groß ist). Testen Sie hierbei einseitig $H_0 : p = p_0$ gegen $H_A : p > p_0$ und erinnern Sie sich daran, dass das Vertrauensintervall aus denjenigen p_0 besteht, für die H_0 nicht verworfen werden kann. (Die Alternative H_A entspricht der Behauptung von ANTIQUARTIS und der Hoffnung des Landwirtes, dass PASTOFEBRIL tatsächlich wirkt.)

Aufgabe 35

(10 Zusatzpunkte)

Sie bearbeiten in einer Klausur eine Multiple-Choice-Aufgabe bestehend aus fünf Fragen mit jeweils fünf möglichen Antworten, von denen immer genau eine richtig ist. Für jede richtig beantwortete Frage werden Ihnen 2 Punkte gutgeschrieben, für jede falsch beantwortete Frage werden Ihnen 2 Punkte abgezogen, unbeantwortete Fragen geben weder Punktgutschriften noch -abzüge. Falls sich eine negative Gesamtpunktzahl für die Aufgabe ergibt, wird sie mit Null Punkten gewertet.

Die Antwort auf drei Fragen wissen Sie sicher, bei den anderen zwei Fragen haben Sie keine Ahnung – Sie entscheiden sich, bei diesen zwei Fragen zu raten. Sei X die von Ihnen in der gesamten Aufgabe erreichte Punktzahl.

- Welche Werte kann X annehmen?
- Berechnen Sie für jeden der Werte aus Teil a die Wahrscheinlichkeit, dass er angenommen wird.
- Bestimmen Sie nun auch den Erwartungswert $E[X]$, die Varianz $\text{Var}(X)$ sowie die Standardabweichung $\sigma(X)$.
- War es eine gute Idee zu raten?

Aufgabe 36

(10 Punkte)

Eine Suspension von Bakterien (Keimen) wird auf 100 Petri-Schalen abgesetzt, pro Petri-Schale 0,2 ml. Die Keimzahl in der Suspension sei 100/ml. Nach einer gewissen Zeit werden die Bakterienkolonien, von denen jede aus einem Keim entsteht, unter dem Mikroskop ausgezählt.

- Wie groß ist die erwartete (durchschnittliche) Anzahl λ_0 von Bakterienkolonien pro Petri-Schale?

Im folgenden nehmen wir an, dass die Anzahl X von Bakterienkolonien pro Petri-Schale poissonverteilt ist mit einem Parameter, der λ genannt wird.

- Für einen anderen Versuch braucht man Petri-Schalen, die genau eine Bakterienkolonie enthalten. Wie stark muss obengenannte Suspension verdünnt werden, damit die erwartete Anzahl solcher Schalen maximal ist? (Für welches λ ist die Wahrscheinlichkeit maximal, dass eine Schale genau eine Bakterienkolonie enthält?)
- In einem anderen Fall möchte man die Suspension möglichst stark verdünnen, aber so, dass man erwarten kann, dass 80 der 100 Schalen mindestens eine Bakterienkolonie enthalten. Wie stark darf man maximal verdünnen? (Ab welchem λ ist die Wahrscheinlichkeit, für eine fest vorgegebene Petri-Schale keine Bakterienkolonie zu erhalten, $\leq 20\%$?)
- Wie gross ist für $\lambda = 0,8$ die Wahrscheinlichkeit, dass eine nicht-leere Petri-Schale genau eine Bakterienkolonie enthält? (Salopp gefragt: Wieviele Prozent der Petri-Schalen, die mindestens eine Bakterienkolonie enthalten, werden genau eine enthalten? HINWEIS: Dies ist eine bedingte Wahrscheinlichkeit.)
- Wie klein muss λ sein, damit die Wahrscheinlichkeit, dass eine nicht-leere Petri-Schale genau eine Bakterienkolonie enthält, gleich $\frac{3}{4}$ ist?