Fachbereich Mathematik und Informatik
der Universität Marburg
Dr. Helga Lohöfer

Sommersemester 2012

**Mathematik für Natur- und Sozialwissenschaftler**

- Blatt 5 -
Abgabe: Mittwoch, den 23.05.2012, 12:10 Uhr in der Vorlesung

**Skript und Aufgabenblätter:** Im Internet unter www.mathematik.uni-marburg.de/~lohoefer/
**Lektürehinweis:** Skript Kap 7.1, 7.2, 8.5, 10.1
**Wichtige Begriffe:** Wahrscheinlichkeit eines Ereignisses E, Erwartungswert $\mu$ und Streuung $\sigma$ einer numerischen Variablen X als **Grenzwerte**. Konfidenzintervall für $\mu$, $\chi^2$-Anpassungstest.

1. Hat sich die Bevölkerung eines Staates verjüngt? Eine repräsentative Erhebung zum Durchschnittsalter der Bevölkerung eines afrikanischen Staates habe im Jahr 1980 bei n = 1600 Personen das Durchschnittsalter $\overline{x} = 34{,}4$ Jahre ergeben, Standardabweichung s = 15,5 Jahre. Eine erneute Erhebung im Jahr 2010 ergab bei n = 2000 Personen das niedrigere Durchschnittsalter $\overline{x} = 32{,}5$ Jahre, Standardabweichung 16,9 Jahre.
   a) Berechnen Sie anhand beider Messreihen jeweils das 95%-Konfidenzintervall für $\mu$ (Eckdaten mit 2 Nachkommastellen). (3)
   b) Haben die beiden Intervalle einen gemeinsamen Durchschnitt? (2)
   c) Welche der beiden Schlussfolgerungen ist richtig, welche falsch:
      (1) Anhand der Erhebungen ist ein Unterschied im Durchschnittsalter der Bevölkerung zwischen 1980 und 2010 nicht feststellbar.
      (2) Das Durchschnittsalter der Gesamtbevölkerung ist von 1980 bis 2010 mit 95% Sicherheit gesunken. (1)

2. Eine zentrale Abiturklausur erbrachte einen landesweiten Notendurchschnitt von $\mu = 9{,}2$. Eine Schule erzielte bei n = 76 Klausurteilnehmern die Durchschnittsnote 8,4 bei einer Standardabweichung von s = 2,1. Unterscheidet sich der lokale Notendurchschnitt vom landesweiten nur rein zufällig? Antwort, wenn möglich, mit Angabe der Sicherhei und Irrtumswahrscheinlichkeit. (3)

3. Bei der Berechnung der Prüfgröße im $\chi^2$-Anpassungstest kann man die Anzahl der Rechenschritte stark reduzieren, wenn man folgende Gleichung ausnutzt:
$$\sum_{i=1}^{r} \frac{(H_i - N_i)^2}{N_i} = \sum_{i=1}^{r} \frac{H_i^2}{N_i} - n$$
Im Detail: $$\frac{(H_1 - N_1)^2}{N_1} + \frac{(H_2 - N_2)^2}{N_2} + \ldots + \frac{(H_r - N_r)^2}{N_r} = \frac{H_1^2}{N_1} + \frac{H_2^2}{N_2} + \ldots + \frac{H_r^2}{N_r} - n$$
Rechnen Sie nach, dass diese Gleichung stimmt. Dabei ist zu benutzen, dass $N_i = n \cdot p_i$ ist. (3)
Bemerkung: Man erkauft die Verringerung der Anzahl der Rechenschritte durch eine erhebliche Vergrößerung der entstehenden Zahlen.

4. In einem Staat wurden in vier aufeinander folgenden Monaten folgende Geburtenzahlen registriert :

| Januar | Februar | März | April |
|---|---|---|---|
| 6200 | 6300 | 6360 | 6500 |

   Überprüfen Sie anhand obiger Daten die Nullhypothese, dass die monatlichen Unterschiede bei den Geburtenzahlen rein zufällig sind (d.h. dass eine Geburt in allen vier Monaten mit derselben Wahrscheinlichkeit $p_i = 0{,}25$ auftritt). Antwort, wenn möglich, mit Angabe der Sicherheit und Irrtumswahrscheinlichkeit. (3)

**Bedingungen für die erfolgreiche Teilnahme:**
- Regelmäßige Teilnahme an der Vorlesung (maximal 2 mal Fehlen).
- Auf mindestens 9 Übungsblätter jeweils mindestens 5 Punkte erhalten, insgesamt jedoch mindestens 75 Punkte. Es sind 11 Übungsblätter mit jeweils mindestens 15 Punkten geplant.
- Dazu das Bestehen einer Klausur.

**Klausurtermine:**
1. Klausur: Do, 19.07.12, 11:00 – 12:45, Hörsaalgebäude Chemie Lahnberge, Hans-Meerwein-Str., HS A + B
2. Klausur: Do, 11.10.12, 11:00 – 12:45, Hörsaalgebäude Chemie Lahnberge, Hans-Meerwein-Str., HS B

**Wahrscheinlichkeit eines Ereignisses:**

Ist $H_i$ die absolute und $h_i = \frac{H_i}{n_i}$ die prozentuale Häufigkeit, mit der eine messbare Variable bei $n_i$ Messungen das Ereignis E als Ergebnis liefert, und gilt $n_1 < n_2 < \ldots < n_i < \ldots$ , so strebt die Folge ($h_i$) immer gegen einen **Grenzwert** $p = p(E) = \lim h_i$ . Wegen $0 \leq h_i \leq 1$gilt auch $0 \leq p \leq 1$. p ist die **Wahrscheinlichkeit** von E.

**Erwartungswert und Standardabweichung einer Variablen X:**

Ist X eine **reellwertige (oder ganzzahlige)** Zufallsvariable, betrachtet man zu X viele Messreihen unterschiedlicher Längen $n_i$ mit jeweiligen Mittelwerten $\overline{x}_i$ und Standardabweichungen $s_i$, (mit stets derselben Methode gewonnen!) und gilt $n_1 < n_2 < \ldots < n_i < \ldots$ , so folgt immer:

1. Die Folge der Mittelwerte ( $\overline{x}_i$ ) strebt gegen einen **Grenzwert**

   $\mu = \lim \overline{x}_i$ , den „Erwartungswert **von X**“ (ist X ganzzahlig, so wird oft $\lambda$ statt $\mu$ geschrieben).

2. Die Folge der Standardabweichungen ($s_i$) strebt gegen einen **Grenzwert**

   $\sigma = \lim s_i$ , die „Standardabweichung **von X**“.

$\sigma^2 = \lim s^2$ heißt die Varianz **von X**.

Die Grenzwerte p, $\mu$ und $\sigma$ sind theoretische garantierte Größen, hängen allein von E bzw. X ab, nicht von den konkreten Messreihen, ihr Zahlwert ist aber bei konkretem E bzw. X immer nur durch Schätzung bestimmbar (im Unterschied zum exakt berechenbaren „dynamischen Gleichgewicht“ $\hat{y}$ , s. Aufgabenblatt 4).

**Der $\chi^2$-Anpassungstest (**siehe Internetskript 10.1)**:**

Jede Messung von X liefere genau eines von r Ergebnissen $E_1, \ldots , E_r$.

Nullhypothese $H_0$ : „Die Ergebnisse $E_1, \ldots , E_r$ besitzen die Wahrscheinlichkeiten $p_1 = p(E_1), \ldots , p_r = p(E_r)$“

Eine empirische Messreihe zu X liefere das Ergebnis $E_i$ mit der absoluten Häufigkeit $H_i$ ($i = 1,2,\ldots, r$).

Der $\chi^2$-Anpassungstest überprüft, ob die Nullhypothese mit den Messergebnissen verträglich ist:

- Berechne $n = H_1 + H_2 + \ldots + H_r$
- berechne $N_i = n \cdot p_i$ und sichere, dass $n \cdot p_i \geq 5$ gilt für alle $i = 1,2,\ldots, r$
- berechne die Prüfgröße $\chi^2 = \sum_{i=1}^{r} \frac{(H_i - N_i)^2}{N_i}$
- schlage in der **$\chi^2$-Tabelle** nach für $f = r – 1$ (im Allgemeinen) bzw. für $f = r – 3$ (bei Normalverteilung) oder $f = r – 2$ (bei Poisson- oder Binomialverteilung)
- <u>Auswertung</u>:
  Ist $\chi^2 \leq$ tab-$\chi^2$(95%), so wird die Nullhypothese **angenommen,** ohne dass man sagen kann, mit welcher Sicherheit sie richtig ist oder mit welcher Irrtumswahrscheinlichkeit dieseAkzeptanz behaftet ist.

  Ist $\chi^2 >$ tab-$\chi^2$(95%), so spricht nur noch 5% Wahrscheinlichkeit für die Richtigkeit der Nullhypothese. Sie wird dann abgelehnt, weil mit 95% Sicherheit bzw. 5% Irrtumswahrscheinlichkeit falsch.

  Ist $\chi^2 >$ tab-$\chi^2$(99%), so wird die Nullhypothese sogar mit 99% Sicherheit bzw. 1% Irrtumswahrscheinlichkeit abgelehnt.

  Ist $\chi^2 >$ tab-$\chi^2$(99.9%), so wird die Nullhypothese mit 99,9% Sicherheit bzw. 0,1% Irrtumswahrscheinlichkeit abgelehnt.