Fachbereich Mathematik und Informatik
der Universität Marburg
Dr. Helga Lohöfer

Wintersemester 2013/14

Grundlagen der Mathematik für Biologen

- Blatt 12 -

Abgabe: Montag, den 27.01.2014, vor der Vorlesung, spätestens14:05 Uhr

Lektüreaufgaben: Skript Kap. 7.1 – 7.3 und 8.1
Thema: Zufallsvariable, Messmethode, Experiment; empirische und ideale Messreihe, die drei statistischen Daten n, $\overline{x}$ und s, ihre Grenzwerte ∞, μ und σ. Normalverteilung, Schnelltest nach David.

1. In einem berühmten Selektionsexperiment (W. Johannsen, Über die Erblichkeit in Populationen und reinen Linien, Jena 1903, S. 19) wurden aus etwa 16000 braunen Bohnen die 25 größten ausgesucht und durch Selbstbefruchtung weitergezüchtet. In der nächsten Generation ergab die Gewichtsmessung nach Auswertung einer Strichliste mit 12 Merkmalsklassen folgende absoluten Häufigkeiten H_j (Intervallbreite Δx jeweils = 0,5g, x_j = Mittelpunkt des j-ten Intervalls) :

j	1	2	3	4	5	6	7	8	9	10	11	12
x_j [g]	2,25	2,75	3,25	3,75	4,25	4,75	5,25	5,75	6,25	6,75	7,25	7,75
H_j	5	18	46	144	127	70	70	63	28	15	8	4

   a) Berechnen Sie die Anzahl n der Messungen sowie die relativen Häufigkeiten $h_j = \frac{H_j}{n}$ und die Werte $p_j = \frac{H_j}{n \cdot \Delta x}$ (j = 1,…,12) der empirischen Dichtefunktion (jeweils 4 Nachkommastellen). (2)

   b) Skizzieren Sie in einem gemeinsamen Koordinatensystem das Histogramm und den Graphen der Dichtefunktion (Punkte verbinden).
   (Maßstab: waagerechte Achse: 2 Kästchen = 0,5 ; senkrechte Achse: 5 Kästchen = 0,100).
   Der Experimentator hatte mit Normalverteilung gerechnet. Liegt sie annähernd vor? (5)

2. Rechnen Sie nach, dass folgende Gleichung stimmt: $\sum \left(x_i - \overline{x}\right)^2 = \sum x_i^2 = -n \cdot \overline{x}^2$

   Im Detail: $\left(x_1 - \overline{x}\right)^2 + \left(x_2 - \overline{x}\right)^2 + \ldots + \left(x_n - \overline{x}\right)^2 = x_1^2 + x_2^2 + \ldots x_n^2 - n \cdot \overline{x}^2$

   Tipp: Es ist $\sum x_i = n \cdot \overline{x}$ (3)

   Bemerkung: Die Berechnung der rechten Seite macht oft erheblich weniger Arbeit als die der linken, dafür können die Summanden wesentlich größer werden..

3. In einem Wasserwerk wurden von einem Brunnen zu verschiedenen Zeiten Proben entnommen und auf ihren Gehalt an Phosphat untersucht. Es ergaben sich folgende Konzentrationen (in $mg \cdot l^{-1}$):

3,2	2,7	2,5	3,3	3,5	3,5	2,6	2,7

   a) Berechnen Sie die statistischen Daten der Messreihe (3 Nachkommastellen). (2,5)

   b) Urteilen Sie mit dem Schnelltest nach David über die Richtigkeit der Hypothese, dass nur schwache äußere Einflüsse zu schwankenden Phosphatkonzentrationen in dem Brunenwasser führten, der Phosphatgehalt dort also normalverteilt ist (wenn möglich, mit Angabe der Sicherheit). (2,5)

Bedingungen für den Erwerb des Übungsscheines:

- Auf mindestens 11 Übungsblätter jeweils mindestens 6 Punkte erhalten, sowie insgesamt mindestens 100 Punkte. Es sind 13 Übungsblätter mit jeweils mindestens 15 Punkten geplant.
- Bestehen einer Klausur. Klausurnote = Modulnote

Klausurtermine:

1. Termin: Mo, 24.02.2014, 10:15 – 12:00, Hans-Meerwein-Str., Hörsaalgebäude Chemie, HS A + B
2. Termin: Do, 10.04.2014, 10:15 – 12:00, Hans-Meerwein-Str., Hörsaalgebäude Chemie, HS A

Bestimmung des Zahlwertes einer Naturkonstanten x

Bei n-facher empirischer Messung der im Zahlwert noch unbekannten reellen Konstante x ergeben sich n teilweise womöglich leicht voneinander abweichende Messwerte $x_1,\ldots,x_n$. Sie haben verglichen mit x den

mittleren Fehler $F_{mitt} = \sqrt{\dfrac{\sum (x_i - x)^2}{n}}$.

Setzt man in die Formel für F_{mitt} bei schon bekannten $x_1,\ldots,x_n$ für x probehalber unterschiedliche Zahlwerte ein, so erweist sich F_{mitt} als eine Funktion der Variablen x. Diese Funktion hat genau ein Minimum. Es wird angenommen bei

$x = \bar{x} = \dfrac{\sum x_i}{n}$ (= arithmetischer Mittelwert der x_i)

Der aufgrund der n Messwerte **bestmögliche Schätzwert für x** ist daher der Mittelwert $\bar{x}$, und der (nur) mittels dieses Schätzwertes erreichbare **minimale mittlere Fehler** ist

$$\sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum x_i^{\,2} - n \cdot \bar{x}^2}{n}}$$

Zufallsvariable x = $f(u_1,\ldots,u_k;\, v_1, \ldots)$, wobei die u_i bei der Messung mit kontrolliert werden, die v_j nicht.

Messmethode = Auswahl der $u_1,\ldots,u_k$ (k = 0 möglich). Beachte: andere Messmethode ⇒ andere Zufallsvariable.

Experiment = planerische Festlegung je eines konkreten Wertes für $u_1,\ldots,u_k$ als Rahmenbedingung für die Messung von x (Im Fall k = 0 gibt es nur ein Experiment zu x, nämlich die Messung von x ohne irgendwelche Rahmenbedingungen).

Messreihe = n-fache empirische Durchführung desselben Experiments. Liefert n Messwerte $x_1,\ldots,x_n$.

Die **drei statistischen Daten** einer Messreihe sind

- n = Anzahl der Messungen = **Länge** der Messreihe
- $\bar{x}$ = **Mittelwert** der x_i
- $s = \sqrt{\dfrac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\dfrac{\sum x_i^{\,2} - n \cdot \bar{x}^2}{n-1}}$ **empirische Standardabweichung** (Achtung: Nenner n – 1).
  s^2 heißt die **empirische Varianz** der Messreihe.

Verlängert man eine Messreihe zum selben Experiment immer weiter, so strebt

- n gegen ∞
- $\bar{x}$ gegen einen Grenzwert μ = **Erwartungswert des Experiments** (=spezifische Konstante des Experiments)
- s gegen einen Grenzwert σ = **Standardabweichung von x** (= spezifische Konstante der Messmethode, also wertgleich für alle Experimente zu x). Folglich strebt
  s^2 gegen σ^2 = **Varianz von x**

Ideale Messreihe = sehr lange Messreihe (n ≥ 3000). Man ordnet ihr als statistische Daten die Grenzwerte zu:

n = ∞, $\bar{x}$ = μ, s = $\bar{x}$