Grundkurs-Klausur 12.2

HEINZ BÖER, RICARDA-HUCH-GYMNASIUM, GELSENKIRCHEN

Zusammenfassung: *Nach der Vermittlung des in den Richtlinien verlangten Orientierungswissens Stochastik und zusätzlich einer Einführung in den Chi-Quadrat-Test habe ich die Klausur gestellt zu den Schwerpunkten: Parameter-Test, α/β-Fehler, Chi-Quadrat-Test, Vergleich und Kritik der Testergebnisse.*

1 Einleitung

Die Vereinbarung für meine Schule lautet: Bis zum Ende der Jahrgangsstufe 12 ist die Analysis (bis auf Vertiefungen und Erweiterungen in 13.2) abgeschlossen, das Orientierungswissen in Stochastik und Analytischer Geometrie vermittelt. Dann können auf gleichem Wissensstand die 13er-Kurse mit Wiederholern starten und man kann den Schwerpunkt für die 13 sowohl auf die Stochastik als auch auf die Analytische Geometrie legen.

Für meine Kolleg/innen und Schüler/innen habe ich ein Skript zum Orientierungswissen Stochastik geschrieben, in dem nach einer schnellen Hinführung zur Binomialverteilung ins Testen von Hypothesen eingestiegen wird ohne Einführung und Verwendung von Erwartungswert und Standardabweichung, nur durch Rückgriff auf Binomialverteilungen, deren Werte von einem PC-Programm erzeugt werden.

Der Kurs lag im Klausurblock jeweils sehr früh, so dass in der 1. Klausur noch Teile der Funktionsbestimmung aus der Analysis (Straßentrassierungen) vorkamen und der Umgang mit Binomialverteilungen als Stochastik-Teil. Bis zu der unten folgenden 2. Klausur des Halbjahres blieb nach dem Lehrgang zum Orientierungswissen Stochastik noch Zeit, den Chi-Quadrat-Test einzuführen. Das finde ich auch nötig, da er besser zum hier gewählten Schwerpunkt „Medikamententest“ passt und auch real benutzt wird.

Nach der 2. Klausur blieben noch 6 Wochen Zeit für die Vermittlung des Orientierungswissens zur Analytischen Geometrie.

Der behandelte Chi-Quadrat-Test ist nicht Voraussetzung für den Fortgang des – von mir für die 13 gewählten – Schwerpunktes Stochastik. Im neuen Halbjahr 13.1 habe ich die Demoskopie angeschlossen, die wegen der Bundestagswahl gut passte. Dafür sind dann Erwartungswert, Standardabweichung, $\Phi(z)$-Tabelle, ... nötig.

Das Testen nur mit Binomialverteilungen aus dem Orientierungswissen habe ich vertiefend in 13.2 noch einmal aufgegriffen mit Sigma-Umgebungen. Die *n*-Bestimmung für vorgegebene Fehler 1. und 2. Art waren dann auch erst möglich.

Die folgende Klausur markiert den Endpunkt des ersten Unterrichtsabschnitts zur Stochastik, der das Orientierungswissen und den Chi-Quadrat-Test einschließt. Mit den Themen und dem mathematischen Schwierigkeitsgrad liegt sie schon auf Abiturniveau.

Diese Kursthemen-Abfolge und die Schwerpunktsetzungen empfehle ich. Die Schüler/innen sind gut auf ihn eingestiegen und sie sind mit ihm besser klar gekommen als mit der Analysis vorher.

2 Grundkurs 12.2 – Mathematik – Böer – Klausur 2 – 3 U-Stunden

1. Auf Grund langjähriger Erhebungen ist bekannt, dass ein Drittel der Bevölkerung Schnupfen bekommt. Die tägliche Salzwasser-Methode (Kopf zurücklegen, warmes Salzwasser in ein Nasenloch laufen lassen und anschließend wieder herauslaufen lassen; Erfahrene nutzen dafür das andere Nasenloch) testete Richard Ravizza an 147 College-Studenten von der Pennsylvania State University. Er stellte fest, dass die neue Methode zu seltenerer Erkältung führte.
   a) Wie müsste die Testplanung ausgesehen haben?
   b) Welche Ergebnisse würden die Feststellung oben erlauben?

2. Der Zufallszahlen-Generator eines PC wird angezweifelt. Bei 1485 Würfen trat 270-mal die Augenzahl 6 auf. Ist er in Ordnung (α = 7 %)?

3. a) Kommentiere: Für Aufgabe 1 plant jemand einen zweiseitigen Test.
   b) Kommentiere: Wird H_0 nicht verworfen, dann ist H_0 bestätigt.
   c) Wie muss ein Medikamententest angelegt sein? Nenne und begründe die Kriterien.
   d) Erläutere den α-Fehler an Aufgabe 1.

4. Am Naserümpfen beteiligen sich 9 % der Bundesbürger. Da die Ricarda- und Grillo-Schüler/innen als positiv-denkend bekannt sind, vermutet man, dass diese Volksseuche unter ihnen weit weniger verbreitet ist. Für die Jahrgangsstufe 11/12 mit 350 Schüler/innen ergibt sich ein Verwerfungsbereich von $V_{\alpha=5\%} = \{0, 1, ..., 22\}$.
   a) Welchen β-Fehler riskiert man mit diesem Testansatz, wenn am Ricarda und Grillo das Naserümpfen tatsächlich nur halb so häufig ist wie üblich?
   b) Erläutere allgemein die Bedeutung des β-Fehlers.

5. Lancet berichtete im Oktober 1976 von einer Studie über die Überlebensrate von Patienten mit Herzstillstand. Bei sofortigen Wiederbelebungsversuchen durch trainierte Laien haben 27 von 75 Fällen überlebt. Beim Warten auf einen Notarzt haben 43 von 556 Fällen überlebt. Empfohlen wurde daraufhin, in die übliche Erste-Hilfe-Ausbildung auch Wiederbelebungs-Trainings einzubauen. Ist diese Empfehlung okay?
   a) Führe einen linksseitigen Parameter-Test durch.
   b) Führe einen χ^2-Test durch.
   c) Erläutere für den einseitigen Parameter-Test in a) den Zusammenhang von X-Wahl, Seitigkeit des Testes, Formulierung der H_1-Hypothese, Wahl des p-Wertes aus der Tabelle und Auswertung der PC-Binomialverteilungstabelle. Formuliere für a) eine zweite Testmöglichkeit.

6. Von den 1962 Patienten, die ein neues Medikament erhielten, starben 227; in der Placebo-Gruppe mit 2722 Patienten starben 367.
   a) Bei einer Irrtumswahrscheinlichkeit von α = 1 % prüfe mit einem Parameter-Test, ob das neue Medikament zur Markteinführung zugelassen werden soll.
   b) Werte auch mit einem χ^2-Test.
   Zu der (sich in a), b) hoffentlich ergebenden) Verschiedenheit der Wertungen: der exakte Test von Fisher (den wir noch nicht behandelt haben), ergibt $P(X \leq 226) = 2,3\%$ und $P(X \leq 227) = 2,9\%$. Auf dem Hintergrund dieser exakten Vierfelder-Tafel-Wertung
   c) erläutere die Problematik eines Parameter-Testes für Medikamente.
   d) erläutere die Problematik eines χ^2-Testes für Medikamente.

χ^2–Grenzwert	2,71	3,84	5,02	6,63
Anteil der Außerhalb-Werte (α-Fehler)	10 %	5 %	2,5 %	1 %

3 Klausurlösung 2, GK 12.2, 24.05.2002

1. a) H_0: Die Salzwasser-Methode bewirkt nichts. $p = \frac{1}{3}$.
   H_1: Die Salzwasser-Methode senkt die Zahl der Erkälteten. $p < \frac{1}{3}$ (linksseitig).
   $n = 147$; $\alpha = 5\%$
   X: Zahl der Studenten, die am Test teilnehmen und sich erkälten.
   Gesucht ist k, so dass $P(X \leq k)$ letztmalig 5 % unterschreitet.
   $P(X \leq 39) \approx 4,7\%$; $P(X \leq 40) \approx 6,8\%$
   $V = \{0, ..., 39\}$
   Alternative: H_1: Erhöhung der Nicht-Erkälteten-Zahl; $p > \frac{1}{3}$ (rechtsseitig); $P(X \leq 57) \approx 93,1\%$; $P(X \leq 58) \approx 95,1\%$; $V = \{59, 60, ..., 147\}$

   b) Liegt die Zahl der erkälteten Testpersonen unter 40, so sind es signifikant wenig und die Feststellung, dass die Salzwasser-Methode zu seltenerer Erkältung führt, ist erlaubt.

2. H_0: Der Zufallszahlengenerator ist in Ordnung. $p = \frac{1}{6}$.
   H_1: Der Zufallszahlengenerator ist nicht in Ordnung. $p \neq \frac{1}{6}$ (beidseitig).
   $n = 1485; p = 0,167; \alpha = 7\%$
   X: Zahl der 6en im Versuch.
   $P(X \leq 221) = 3,1\%$; $P(X \leq 222) = 3,7\%$; $P(X \leq 273) = 96,5\%$; $P(X \leq 274) = 96,6\%$
   $\alpha = 7\%$ muss auf beide Seiten mit 3,5 % aufgeteilt werden.
   $V = \{0, 1, ..., 221, 275, 276, ..., 1485\}$
   Da $270 \notin V$, wird H_0 nicht verworfen und H_1 nicht akzeptiert. Der Zufallsgenerator gilt als in Ordnung.

3. a) Bei Medikamententests ist aus ethischen Gründen prinzipiell nur ein einseitiger Test erlaubt.
   Da zudem im Text von „seltenerer Erkältung" die Rede ist, muss ein linksseitiger Test gemacht werden.

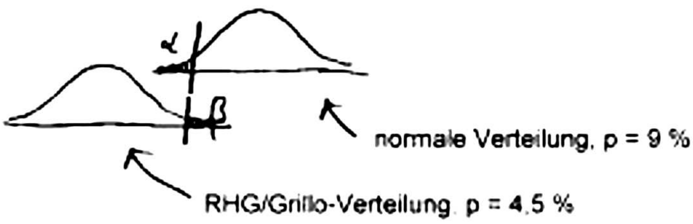
   b) Mit einem Zufallsversuch lässt sich ein genauer p-Wert (H_0: $p = \frac{1}{3}$) nicht positiv bestätigen. Viele p-Werte in der Nähe von $\frac{1}{3}$ wären auch mit dem Testergebnis verträglich und würden nicht zum Verwerfen von H_0 führen.

   c) Der Medikamententest muss randomisiert sein, damit die Gruppe, die das Medikament erhält, nicht un/bewusst „günstig" zusammengestellt wird. Er muss kontrolliert sein, d. h. es muss eine Kontrollgruppe geben. Er muss für die Patienten „blind" sein, d. h. sie dürfen nicht wissen, ob sie das Medikament oder ein Placebo erhalten, denn das Wissen um die Medikamenteneinnahme könnte schon zur Heilung führen. Um Einflussnahmen des Arztes auszuschalten, darf auch er nicht wissen, ob er das Medikament oder das Placebo gibt.

   d) In 5 % der Fälle kann es passieren, dass H_0 tatsächlich zutrifft und auf Grund einer Zufallsschwankung trotzdem ein seltenes Ergebnis von kleiner als 40 Erkälteten eintritt. Dann würde die Salzwasser-Methode als helfend akzeptiert, obwohl sie wirkungslos ist.

4. a) $\beta = P_{350;4,5\%}(X > 22) = 1 - P(X \leq 22) = 1 - 95,3\% = 4,7\%$.
   Bei 4,5 % Naserümpfhäufigkeit liefert der Test einen β-Fehler von 4,7 %.

   b)



   Trifft $H_1 (p < 9\%)$ tatsächlich zu, so kann auf Grund einer Zufallsschwankung trotzdem ein hoher Wert in der Stichprobe (über 22 Naserümpfer) auftreten. Dann wird irrtümlich H_0 nicht verworfen, H_1 nicht akzeptiert.

5. a)

	gestorben	überlebt	
Wiederbelebung	48	27	75
keine Wiederbelebung	513	43	556
	561	70	631

H_0: Wiederbelebung bewirkt keine Änderung der Gestorbenenzahl; $p = \frac{513}{556} \approx 0,923$
H_1: Wiederbelebung senkt die Gestorbenenzahl, $p < \frac{513}{556}$, linksseitig.
X: Zahl der Gestorbenen trotz Wiederbelebung; $n = 75$
Gesucht ist k, so dass letztmalig P $(X \leq k)$ die 5-%-Grenze unterschreitet.
$P(X \leq 64) = 2,8\%$; $P(X \leq 65) = 6,2\% \rightarrow V = \{0, 1, ..., 64\}$
Da $48 \in V$, wird H_0 verworfen, H_1 akzeptiert. Die Wiederbelebung senkt die Zahl der Gestorbenen.

b)

	gestorben	überlebt	
Wiederbelebung	0,1057 $\hat{=}$ 66,7	0,0132 $\hat{=}$ 8,3	75 $\hat{=}$ 0,1189
keine Wiederbelebung	0,7834 $\hat{=}$ 494,3	0,0977 $\hat{=}$ 61,7	556 $\hat{=}$ 0,8811
	561 $\hat{=}$ 0,8891	70 $\hat{=}$ 0,1109	

H_0: Wiederbelebung und Überleben sind unabhängig voneinander.
H_1: Wiederbelebung und Überleben hängen voneinander ab.
α = 5 %; $n = 631$

$$\chi^2 = \frac{(66,7-48)^2}{66,7} + \frac{(8,3-27)^2}{8,3} + \frac{(494,3-513)^2}{494,3} + \frac{(61,7-43)^2}{61,7} = 53,7 > 3,84$$

Da der χ^2-Wert über der 5-%-Grenze von 3,84 liegt, wird H_0 verworfen, H_1 akzeptiert: Wiederbelebung hat Einfluss auf das Überleben. An den Daten sieht man, dass sie die Überlebenden-Zahl erhöht.

c) Wird H_1 mit „Senkung der Gestorbenenzahl“ formuliert, so passt wie oben ein linksseitiger Test und im PC muss 5 % unterschritten werden. p aus der Vergleichsgruppe wird berechnet als Anteil der Verstorbenen und X aus der Versuchsgruppe zählt die Anzahl der Gestorbenen.
Wird H_1 mit „Erhöhung der Überlebenden-Zahl“ formuliert, so passt ein rechtsseitiger Test und im PC muss 95 % überschritten werden. p aus der Vergleichsgruppe bezeichnet den Anteil der Überlebenden und X aus der Versuchsgruppe die Anzahl der Überlebenden.

6. a)

	gestorben	überlebt	
Medikament	227	1735	1962
Placebo	367	2355	2722
	594	4090	4684

H_0: Das Medikament wirkt nicht; $p = \frac{367}{2722}$
H_1: Das Medikament senkt die Zahl der Verstorbenen; $p < 13,48\%$; linksseitig
$\alpha = 1\%; n = 1962$
X: Zahl der Gestorbenen, die das Medikament erhalten
$P(X \leq 229) = 0,9\%$; $P(X \leq 230) = 1,1\% \rightarrow V = \{0, 1, ..., 229\}$
Da $227 \in V$, wird H_0 verworfen, H_1 akzeptiert. Das Medikament senkt die Zahl der Sterbefälle.
Alternative: H_1: Erhöhung der Überlebendenzahl; $p > \frac{2355}{2722}$; $P(X \leq 1731) = 98,9\%$;
$P(X \leq 1732) = 99,1\%; V = \{1733, ..., 1962\}$

b)

	gestorben	überlebt	
Medikament	0,053 $\hat{=}$ 248,8	0,3658 $\hat{=}$ 1713,3	0,4189 $\hat{=}$ 1962
Placebo	0,0737 $\hat{=}$ 345,1	0,5074 $\hat{=}$ 2376,7	0,5811 $\hat{=}$ 2722
	0,1268 $\hat{=}$ 594	0,8732 $\hat{=}$ 4090	4684

H_0: Medikamentengabe und Überlebendenzahl sind unabhängig.
H_1: Das Medikament hat Einfluss.
$\alpha = 1\%, n = 4684$

$$\chi^2 = \frac{(248-277)^2}{248,8} + \frac{(1713,3-1735)^2}{1713,3} + \frac{(345,1-367)^2}{345,1} + \frac{(2376,7-2355)^2}{2376,7} = 3,77.$$

Da $P(\chi^2 \geq 6,63) = 1\%$ und $P(\chi^2 \geq 3,84) = 5\%$, ist H_0 weder auf dem 1 %- noch auf dem 5 %-Niveau zu verwerfen. H_1 wird nicht akzeptiert. Medikamentengabe und Überleben sind unabhängig voneinander.

c) Der Parameter-Test ist für 4-Felder-Tafel-Daten insofern nicht geeignet, als der p-Wert aus der Vergleichsgruppe nicht sicher ist. Die doppelten Zufallsschwankungen (p unsicher und zufällige Testgruppenzusammensetzung) machen den Test unzuverlässig. Hier: Die signifikante Abweichung bei $\alpha = 1\%$ stellt sich als nicht korrekt heraus.

d) Ein χ^2-Test ist immer zweiseitig und deshalb für natürlich einseitige Medikamententests nicht geeignet. Mit $\chi^2 = 3,77$ liegt er etwa bei $\alpha = 6\%$. Auf die Beidseitigkeit verteilt, entspricht das etwa einem einseitigen α-Wert von rund 3 %, was der einseitige Fisher-Test bestätigt.

Anschrift des Verfassers
Heinz Böer — Ricarda Huch-Gymnasium
Bahnhofstr. 72 — Schultestr. 50
48301 Appelhülsen — 45888 Gelsenkirchen
boeer.hamers@t-online.de