Stochastikunterricht am Gymnasium aus der Sicht der Universitäten

Hansruedi Künsch
Seminar für Statistik, ETH Zürich

21. Januar 2015

Stochastikunterricht an Universitäten

Natur- und Ingenieurwissenschaften

- Typischerweise wird im ersten Studienjahr Analysis und Lineare Algebra unterrichtet, und im zweiten Jahr Stochastik sowie manchmal noch Analysis III (Partielle Differentialgleichungen) oder Numerik.
- In den Studiengängen Bauing., Umwelting. und Geomatik findet die Stochastikvorlesung bereits im zweiten Semester statt.
- Im Studiengang Biologie an der ETH werden im 2. und 3. Semester Statistik I und II gehalten.
- Im Studiengang Chemie an der ETH wird Lineare Algebra und Statistik in eine Vorlesung im 2. Semester gepackt.
- Je nach Studiengang liegt das Schwergewicht der Stochastikvorlesung mehr bei der Wahrscheinlichkeitstheorie oder bei der Statistik.

Wirtschafts- und Sozialwissenschaften

- Meist wird Statistik im zweiten Semester unterrichtet als Teil eines Moduls über Methodengrundlagen.
- An der Uni Zürich ist das der Fall für Psychologie, Soziologie, Publizistik, Politikwissenschaften sowie für Wirtschaftswissenschaften.
- In diesen Studiengängen ist Statistik oft das einzige mathematische Fach.
- “Viele Studierende sind überrascht, wenn Sie feststellen, welches Gewicht die Statistik in der soziologischen Forschung hat. Viele denken, dass sie die Mathematik nun los sind.”

Was wird in diesen Vorlesungen behandelt ?

- Deskriptive Statistik
- Wahrscheinlichkeitsrechnung: wenig Kombinatorik, diskrete und stetige Verteilungen.
- Schliessende Statistik: Binomialtest, t-Test und Wilcoxon-Test (gepaart und ungepaart), Chiquadrat-Test. Zusammenhang Tests und Vertrauensintervalle.
- Das lineare Regressionsmodell.
- Verwendung von Software (R oder SPSS).
- Behandlung der Stochastik am Gymnasium soll Vorbereitung und nicht Ersatz sein für die Behandlung an der Universität.

Was erwarten die Dozenten vom gymnasialen Mathematikunterricht ?

- Die Fähigkeit, einfache Formeln zu lesen, und sich nicht durch unterschiedliche Bezeichnungen (z.B. t statt x) verwirren zu lassen.
- Klarheit für den Dozenten, was alle StudienanfängerInnen wissen.
- Bereitschaft, sich mit Konzepten auseinanderzusetzen, die schwierig sind und eine präzise Sprache erfordern.
- Grundideen (Rückschluss von Stichprobe auf Population, Schlussfolgerungen sind nicht zu 100% korrekt) wären nützlich, aber ambitiös.
- “Mathematik und Statistik sollten nicht nur als formale Methoden gelehrt werden, sondern insbesondere auch deren Relevanz in Alltag sowie zahlreiche Anwendungsfelder thematisieren.”

Relative Häufigkeiten und Wahrscheinlichkeiten

Wahrscheinlichkeit

- Wenn die Ergebnisse eines Zufallsversuchs alle gleich wahrscheinlich sind, dann gilt

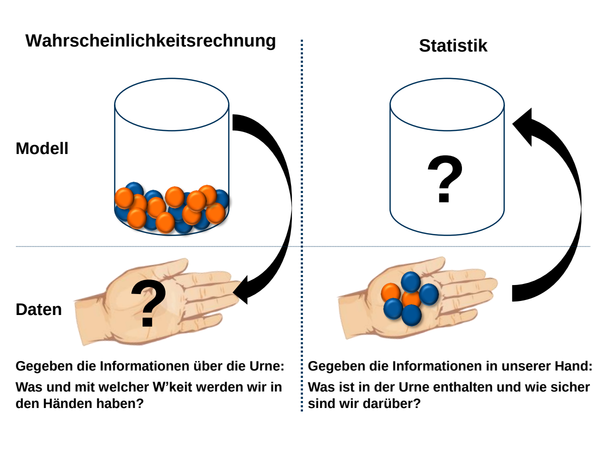
  Wahrscheinlichkeit =
  Anzahl günstige : Anzahl mögliche Ergebnisse.

- Dieser Zugang wird vor allem bei Glücksspielen verwendet.
- In anderen Fällen wird Wahrscheinlichkeit aufgefasst als Idealisierung oder Grenzwert der relativen Häufigkeit bei vielen Wiederholungen des Versuchs.
- Anschaulichere (?) Sprechweise:

  Wahrscheinlichkeit = erwartete relative Häufigkeit

  (mit beliebigem Nenner).

Modelle und Daten



Wie geht der Umkehrschluss in der rechten Figur ?

Verteilung der absoluten Häufigkeit

Die Häufigkeit eines Ereignisses A bei n Wiederholungen ist eine Zufallsvariable X. Sind die Wiederholungen unabhängig, dann ist X binomial$(n, P(A))$-verteilt:

$$\mathbb{P}(X = k) = \binom{n}{k} P(A)^k (1 - P(A))^{n-k}.$$

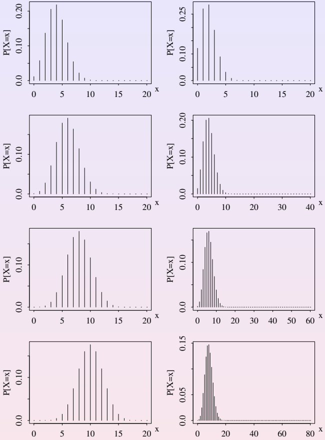
Mit Hilfe eines Taschenrechners oder mathematischer Software kann man zu gegebenem k und $P(A)$ sowohl $\mathbb{P}(X = k)$ als auch

$$\mathbb{P}(X \leq k) = \sum_{j=0}^{k} \mathbb{P}(X = j)$$

berechnen. Damit folgt auch

$$\mathbb{P}(k_1 \leq X \leq k_2) = \mathbb{P}(X \leq k_2) - \mathbb{P}(X < k_1).$$

Die Form der Binomialverteilung



Links: $n = 20$, $P(A) = 0.2, 0.3, 0.4, 0.5$.
Rechts: $n = 20, 40, 60, 80$, $P(A) = 0.1$.

$\mathbb{P}(X = k)$ ist venachlässigbar klein ausserhalb von einem (verglichen mit n) engen Bereich um $P(A)$ herum.

Von der Wahrscheinlichkeit zur relativen Häufigkeit

Kennt man $P(A)$, dann kann man für jedes $\alpha \in (0,1)$ zwei Schranken $k_1 < k_2$ bestimmen, so dass

$$\mathbb{P}(X < k_1) \leq \frac{\alpha}{2} < \mathbb{P}(X \leq k_1), \quad \mathbb{P}(X > k_2) \leq \frac{\alpha}{2} < \mathbb{P}(X \geq k_2).$$

Man schneidet auf beiden Seiten einen Bereich mit Wahrscheinlichkeit $\frac{\alpha}{2}$ ab. Beachten Sie, dass k_1 und k_2 von $P(A)$ (und von n und α) abhängen.

Damit gilt mit Wahrscheinlichkeit $\geq 1 - \alpha$

$$\frac{k_1}{n} \leq \frac{X}{n} \leq \frac{k_2}{n}.$$

Nimmt man eine kleine Irrtumswahrscheinlichkeit in Kauf, dann gibt es einen Bereich, in dem die relative Häufigkeit liegt. Dieser Bereich ist i. A. wesentlich kleiner als $[0,1]$.

Von relativer Häufigkeit zur Wahrscheinlichkeit

Als nächstes kehren wir diesen Schluss um: Wenn wir vom einem beobachteten Wert x von X ausgehen, was folgt daraus für $P(A)$?

Es kann sein, dass wir "Pech" gehabt haben: Die Wahrscheinlichkeit $P(A)$ ist grösser als die relative Häufigkeit $\frac{x}{n}$. Wieviel grösser ist noch plausibel ?

Wenn wir eine kleine Irrtumswahrscheinlichkeit zulassen, muss $x \geq k_1$ sein. Das heisst, die Wahrscheinlichkeit des Bereichs von 0 bis x muss grösser als $\frac{\alpha}{2}$ sein. Im Extremfall haben wir also die Gleichung

$$\mathbb{P}(X \leq x) = \frac{\alpha}{2}.$$

Wenn x, n und α gegeben sind, ist das eine Gleichung für $P(A)$. Wir nennen die Lösung $P(A) = p_1 = p_1(x, n, \alpha)$.

Von relativer Häufigkeit zur Wahrscheinlichkeit II

Wenn wir "Glück" gehabt haben, ist $P(A) < \frac{x}{n}$. Analog wie vorher muss $x \leq k_2$ sein. Im Extremfall erhalten wir die Gleichung

$$\mathbb{P}(X \geq x) = \frac{\alpha}{2} \Leftrightarrow \mathbb{P}(X \leq x - 1) = 1 - \frac{\alpha}{2}$$

mit der Lösung $P(A) = p_2 = p_2(x, n, \alpha)$. Zusammengenommen gilt also mit Wahrscheinlichkeit $1 - \alpha$

$$p_2 < P(A) < p_1 .$$

Die Bestimmung von p_1 und p_2 muss numerisch gemacht werden, oder grafisch, indem man $\mathbb{P}(X \leq x)$, bzw. $\mathbb{P}(X \leq x - 1)$ als Funktion von $P(A)$ plottet. Meist genügt eine Genauigkeit auf 2 Stellen.

Anwendung im Unterricht

- Es gibt viele interessante Beispiele, bei denen es darum geht, Wahrscheinlichkeiten aus relativen Häufigkeiten zu schätzen und die verbleibende Unsicherheit zu quantifizieren.
- Eine Liste von solchen Beispielen folgt auf den nächsten zwei Folien.
- Zu jedem Beispiel habe ich weitere Informationen und Hinweise vorbereitet.
- Ziel ist, dass Sie einzeln oder in kleinen Gruppen ein Beispiel genauer anschauen und es wenn möglich für den Unterricht aufbereiten.
- Vorschläge von Ihnen für weitere Beispiele sind ebenfalls willkommen.

Beispiele zur Schätzung von $P(A)$

- Wahrscheinlichkeit von Knabengeburten
- Wahrscheinlichkeit von eineiigen Zwillingen
- Wahrscheinlichkeit von Rechtshändigkeit
- Die Chancen des Torhüters beim Elfmeter
- Geschlechtsselektion durch Spermientrennung
- Genauigkeit von Umfragen
- Umfragen mit heiklen Fragen
- Fang-/Wiederfangmethoden

Beispiele zum Testen auf vorgegebene Werte für $P(A)$

- Mendel'sche Versuche
- Können Elefanten zählen ?
- Tests der Ehrlichkeit von Personen
- Börsenkurse und American Football
- Randomisierte Medikamententests
- Aussersinnliche Wahrnehmung
- Erkennen von echtem Münzwurf
- Beurteilung von Zahlen über Unfälle in verschiedenen Jahren