

Daniel FRISCHEMEIER, Paderborn

## **Wie vergleichen Lehramtsstudierende Verteilungen unter Verwendung der Software TinkerPlots?**

Vergleiche von Verteilungen zweier numerischer Merkmale sind im alltäglichen Leben und in den Medien präsent - eine typische Frage zum Verteilungsvergleich ist beispielsweise (Biehler et al., 2003): „Inwiefern unterscheiden sich Schülerinnen und Schüler hinsichtlich ihrer Zeit am Computer (in Stunden pro Woche)?“ Wenn man solchen Fragestellungen nachgeht, sollte man zum einen mit multivariaten und realen Daten arbeiten (Garfield & Ben-Zvi, 2008), sowie adäquate Software, wie zum Beispiel die Software TinkerPlots (Konold & Miller, 2011) einsetzen. Die Vorzüge der Software TinkerPlots werden u.a. in Biehler (2007) diskutiert. Betrachtet man Lernende beim Durchführen eines Verteilungsvergleiches unter Verwendung einer Software, so eröffnen sich zwei Perspektiven: Zum einen der Blick auf den Einsatz der Software im Verteilungsvergleich-Prozess, zum anderen die stochastischen Aspekte beim Herausarbeiten von Unterschieden und Gemeinsamkeiten zwischen den Verteilungen. Wir wollen im Folgenden die erste Perspektive verfolgen.

### **1. Datenanalyse-Zyklus mit Software**

Biehler (1997, 175) beschreibt einen Zyklus, der den Einsatz von Software in den Verteilungsvergleich-Prozess einordnet und verschiedene Phasen im Datenanalyse-Prozess identifiziert. Biehler unterscheidet hier die vier aufeinander folgenden Phasen „statistical problem“, „problem for the software“, „results of software use“ und „interpretation of results in statistics“ und stellt in der anschließenden empirischen Studie fest, dass Lernende oftmals beim Arbeiten mit Software, mit den von der Software erzeugten Produkten (Kennzahlen, Graphiken), „zufrieden“ sind. Sie streben allerdings keine weitergehenden Beschreibungen oder gar Interpretationen an, oftmals erfolgt ein „Sprung“ von dem realen Problem direkt in die Nutzung der Software.

### **2. Design der Studie**

Wir wollen in einer explorativen Studie überprüfen, ob man ähnliche Beobachtungen auch bei Lernenden, die Verteilungen mit TinkerPlots vergleichen, machen kann. Als Fragestellungen bezüglich der Perspektive „Einsatz der Software im Verteilungsvergleich-Prozess“ formulieren wir: „Welche typischen Phasen können bei der Durchführung eines Verteilungsvergleichs mit TinkerPlots identifiziert werden?“, „Wie groß ist der

In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 385–388).  
Münster: WTM-Verlag

Anteil der einzelnen Phasen am gesamten Verteilungsvergleich-Prozess?“ und „Inwieweit machen die Probanden Schlussfolgerungen und Interpretationen in ihren Daten?“ Zur Beantwortung dieser Fragen, wurden Studierende des Lehramts GHRGe Mathematik beim Durchführen von Verteilungsvergleichen mit TinkerPlots beobachtet. Um den Anforderungen „reale und multivariate“ Daten zu genügen, wurde der Datensatz „Verdienststrukturerhebung 2006“ (kurz: VSE 2006), der eine Zufallsstichprobe (n=861) der ursprünglichen VSE (<http://www.forschungsdatenzentrum.de>) darstellt, verwendet. Dieser Datensatz, der Informationen von 861 Arbeitnehmerinnen und Arbeitnehmern in Deutschland anhand der Variablen „Geschlecht“, „Bruttomonatsverdienst“, „Region“, „Stellung im Beruf“, etc. enthält, wurde in TinkerPlots importiert und den Probanden bereitgestellt. Die Aufgabenstellung „Inwiefern unterscheiden sich Arbeitnehmerinnen und Arbeitnehmer hinsichtlich ihres Bruttomonatsgehalts? Arbeiten Sie Unterschiede in beiden Verteilungen heraus.“ regt einen Verteilungsvergleich an. In diesem Fall den Vergleich der Verteilungen des Merkmals „Bruttomonatsverdienst“, getrennt nach dem Merkmal „Geschlecht“. Die Bearbeitung der Aufgaben erfolgt in Zweiertteams, wobei die Probanden dazu aufgefordert werden, ihre Gedanken, Vorgehensweisen und Intentionen laut zu kommunizieren. Einen Eindruck für eine mögliche, in TinkerPlots erstellte Graphik, die den Unterschied zwischen den beiden Verteilungen herausstellt, gibt die Abbildung 1.

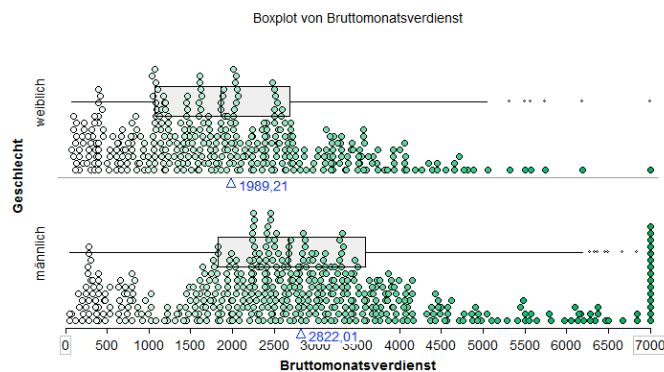


Abbildung 1: mögliche TinkerPlots-Graphik zur VSE 2006

### 3. Durchführung der Studie

Die Teilnehmer der Studie waren GHRGe-Mathematik-Lehramtsstudierende der Universität Paderborn, die u.a. zwei fachliche Lehrveranstaltungen zur Stochastik besucht haben: Eine Vorlesung „Elemente der Stochastik“ sowie ein vertiefendes, an die Vorlesung anschließendes Seminar „Statistisch denken und forschen lernen“ (für Details zu den Lehrveranstaltungen siehe Frischemeier & Biehler (2012)). Nach dem Besuch der beiden Lehrveranstaltungen wurden insgesamt 14 Teilnehme-

rinnen und Teilnehmer (7 Paare) zur Bearbeitung der Aufgabe, welche unter Laborbedingungen stattfand, eingeladen. Die Bildschirmaktivitäten wurden mit Camtasia, die Gesten der Teilnehmer mit einer Videokamera aufgenommen. Die Kommunikation der Paare und die Aktionen mit der Software wurden anschließend transkribiert.

#### 4. Auswertung und Analyse der Daten (Transkripte)

Die Analyse der Transkripte wurde mittels einer strukturierenden, qualitativen Inhaltsanalyse (Mayring, 2010) durchgeführt. Die Kategorien (im Folgenden: Phasen) wurden sowohl deduktiv als auch induktiv ermittelt (Kuckartz, 2012, 69). In einem ersten Schritt haben wir die vier Phasen nach Biehler (1997) als deduktive Grundlage genommen und dann mithilfe der vorliegenden Daten (Transkripte und Videos) induktiv angereichert und verfeinert. Schließlich lassen sich sechs Phasen im Prozess „Verteilungsvergleich mit Software“ identifizieren: *Reales Problem*, *Statistische Aktivität*, *Einsatz der Software*, *AbleSEN der Ergebnisse*, *Schlussfolgerungen* und *Gründe*. Es bleibt eine Restkategorie, die die nicht-kodierbaren restlichen Textstellen enthält. Definitionen sowie Ankerbeispiele zu den Phasen kann man der Tabelle 1 entnehmen. Als Analyseeinheit dienten die für die qualitative Inhaltsanalyse aufbereiteten Transkripte zur Aufgabebearbeitung. Die minimale Kodiereinheit wurde auf ein einzelnes Wort, die maximale Kodiereinheit wurde auf eine Phrase festgesetzt. Einer Textstelle wurde höchstens eine Kodierung zugeordnet.

Phase	Definition	Beispiel
<b>Reales Problem</b>	Formulierung der Untersuchungsabsicht	„Okay, wir sollen jetzt gucken, wie sich männliche und weibliche Arbeitnehmer im Gehalt unterscheiden!“
<b>Statistische Aktivität</b>	Ausdrücken des (Untersuchungs-) Vorhabens auf der statistischen Ebene - Oder: Artikulierung der bevorstehenden Aktivität in TinkerPlots	„Wir könnten die Mittelwerte vergleichen“
<b>Einsatz der Software</b>	Aktive Nutzung der Software TinkerPlots	Es werden die arithmetischen Mittelwerte beider Verteilungen in TinkerPlots berechnet
<b>AbleSEN der Ergebnisse</b>	AbleSEN von Werten oder Graphiken, die mit der Software erstellt wurden	„Das arithmetische Mittel der Verteilung des Merkmals Bruttomonatsverdienst beträgt bei den Frauen ca. 2000€.“
<b>Schlussfolgerungen</b>	Tätigen von Schlussfolgerungen oder Interpretationen	„Das arithmetische Mittel ist bei den Männern höher als bei den Frauen, daher verdienen Männer mehr als Frauen“
<b>Gründe</b>	Angeben von (möglichen) Gründen, die die Erkenntnisse aus den Daten erklären könnten	„Ja... das ist deswegen niedriger, weil da viele Frauen Teilzeit arbeiten“

Tabelle 1: Definitionen und Ankerbeispiele zu den Phasen

#### 5. Ergebnisse

Nach einem vollständigen Materialdurchlauf wurden Häufigkeitsauswertungen (bezüglich) der einzelnen Kodierungen durchgeführt.

	Reales Problem	Statistische Aktivität	Einsatz der Software	AbleSEN der Ergebnisse	Schlussfolgerungen	Gründe	Rest	Gesamt
Rel. H'keit	0.0172	0.1838	0.3407	0.2525	0.0858	0.0269	0.0931	1.0000
(Abs. H'keit)	(7)	(75)	(139)	(103)	(35)	(11)	(38)	(408)

Tabelle 2: Häufigkeit der Kodierungen zu den einzelnen Phasen

Die Tabelle 2 zeigt die Häufigkeiten der Kodierungen zu den einzelnen Phasen. Es fällt auf, dass die Anteile der Kodierungen am höchsten in den Phasen *Einsatz der Software* (ca. ein Drittel der Kodierungen) und *Ablese der Ergebnisse* (ca. ein Viertel der Kodierungen) sind. Beide Phasen machen zusammen fast 60% des gesamten Prozesses aus. *Schlussfolgerungen* haben einen geringen Anteil (ca. 8,6%) an Kodierungen. Sehr gering ist der Anteil auch bei den Phasen *Gründe* und *Reales Problem* mit ca. 2,7% und 1,7%. Der Anteil der Kodierungen zur Phase der *statistischen Aktivität* beträgt mit ca. 18,3 % der Kodierungen ungefähr die Hälfte des Anteils der Kodierungen zur Phase *Einsatz der Software*. Diese Tatsachen bestätigen Biehlers Beobachtungen (s.o.) und könnten ebenfalls als Indiz dafür gewertet werden, dass Lernende die Ergebnisse der Software oft nicht für ihre weiteren Schlussfolgerungen nutzen und die Neigung haben direkt in die Software „einzutauchen“.

## 6. Ausblick

Für eine zukünftige, differenziertere Auswertung wird eine Unterscheidung der Anteile der Kodierungen hinsichtlich der einzelnen Paare vorgenommen. Des Weiteren wird der „Verteilungsvergleich-Prozess“ auch unter der stochastischen Perspektive (Herausarbeiten von Unterschieden und Gemeinsamkeiten zwischen den Verteilungen) betrachtet werden.

## Literatur

- Biehler, R. (1997). Students' difficulties in practising computer supported data analysis - Some hypothetical generalizations from results of two exploratory studies. In J. Garfield & G. Burrill (Eds.): *Research on the Role of Technology in Teaching and Learning Statistics*. Voorburg: International Statistical Institute. 169-190.
- Biehler, R., Kombrink, K., & Schweynoch, S. (2003). MUFFINS – Statistik mit komplexen Datensätzen – Freizeitgestaltung und Mediennutzung von Jugendlichen. *Stochastik in der Schule*, 23(1), 11-25.
- Biehler, R. (2007). TINKERPLOTS: Eine Software zur Förderung der Datenkompetenz in Primar- und früher Sekundarstufe. *Stochastik in der Schule*, 27(3), 34-42
- Frischemeier, D. & Biehler, R. (2012). Statistisch denken und forschen lernen mit der Software TinkerPlots. In Kleine, M. und Ludwig, M. (Eds.): *Beiträge zum Mathematikunterricht 2012*, WTM: Münster. 257-260.
- Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York: Springer.
- Konold, C. & Miller, C. (2011). TinkerPlots 2.0. Emeryville, CA: KeyCurriculumPress.
- Kuckartz, U. (2012): *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. 1. Auflage. Weinheim und Basel.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 11. Auflage. Weinheim: Beltz.