

StatWiki - ein interaktives statistisches Lexikon für Lehre

Im Rahmen der Einführung von Bachelor- und Masterstudiengängen an den deutschen Universitäten erhält man eine klare Zweiteilung des Studiums. Insbesondere in den Masterstudiengängen kann man daher nicht das gleiche statistische Vorwissen voraussetzen. Um die Lücken der Studenten aufzufüllen setzen wir eine Wiki Software (MediaWiki) ein, die mit Hilfe der statistischen Programmiersprache R erweitert wird. Aus den eingebetteten R Programmen werden sowohl Tabellen als auch Grafiken erzeugt, die im Wiki angezeigt werden. Mit Hilfe von Formularen können Variablen an R Programme übergeben werden, so dass auch Interaktion möglich ist.



Das StatWiki ist, genau wie die Wikipedia (www.wikipedia.org), eine Sammlung statistischer Begriffe. Im Gegensatz zur Wikipedia sind jedoch derzeit bestimmte Bereiche der Statistik nicht berücksichtigt, z.B. die Zeitreihenanalyse. Des Weiteren ist die Zielrichtung eine andere: Wir gehen davon aus, dass an allen deutschen Universitäten mehr oder weniger die gleichen statistischen Grundlagen in den Sozial- und Wirtschaftswissenschaften gelehrt werden, d.h. jeder Begriff muss nicht ausführlich erklärt, sondern kann in stark verkürzter Form dargestellt werden.

Da die Grundlage ein Wiki ist, kann jeder registrierte Benutzer, auch und insbesondere Studenten, die Seiten verändern, so dass in der Zukunft auch längere Erklärungen zu schwierigeren Themen möglich sind. Obwohl das StatWiki eher auf Studenten des Masterprogrammes zielt, setzen wir es auch im Bachelorgrundlagenstudium ein, z.B. indem wir „Kochrezepte“ integrieren.

Das StatWiki umfasst derzeit ca. 500 Begriffe, davon sind ca. 300 Weiterleitungen. Die hohe Zahl von Weiterleitungen ergibt sich zum einen aus synonymen Begriffen, z.B. Alpha Fehler, Fehler 1. Art, und zum anderen aus der Zusammenfassung von Begriffen auf einer Seite, z.B. allen Begriffen der Testtheorie auf der Seite „Testtheorie“.

Tabellen und Grafiken

Im Grundlagenstudium der Statistik werden die Studenten in der Regel bei den Verteilungsmodellen mit umfangreichen Tabellen konfrontiert. Im StatWiki werden die Standard-Verteilungstabellen mit R Programmen erzeugt. Die Grafik am Fuß zeigt auf deren linken Seite oben die Seite zur Binomialverteilung. Die Formeln wurden mit LaTeX mit Hilfe des <math> Tag als Grafik eingebettet. Am Ende der Seite findet sich die Tabelle mit den Werten der Verteilungsfunktion (siehe links mitte). Das in die Seite eingebettete R Programm wird links unten gezeigt. Spezielle entwickelte R Makros erlauben die Ausgabe von HTML Code. Die rechte Seite der Grafik unten zeigt die Einbettung von Grafiken im TeachWiki, welches dieselben Programme wie das StatWiki verwendet.

Formulare

Ursprünglich wurden Grafiken und Tabellen direkt in die Webseiten eingebettet. Der Wunsch nach Interaktivität, z.B. durch Angabe von Parametern, würde jedoch bei einer Änderung ein Neuzeichnen der gesamten Seite erfordern. Besonders im TeachWiki, das Seminararbeiten von Studenten enthält, wurden längere Seiten notwendig. Daher wurden die Ausgaben der R Programme generell in ein <iframe> eingebettet, welches nur eine Aktualisierung desselben notwendig macht.

Die Grafik unten zeigt ein Formular, mit dem bei der Binomialverteilung der Wert p eingestellt werden kann. Das Zusammenspiel von <Rform> und <R> Tag ermöglicht die Formulare.

Technische Grundlagen

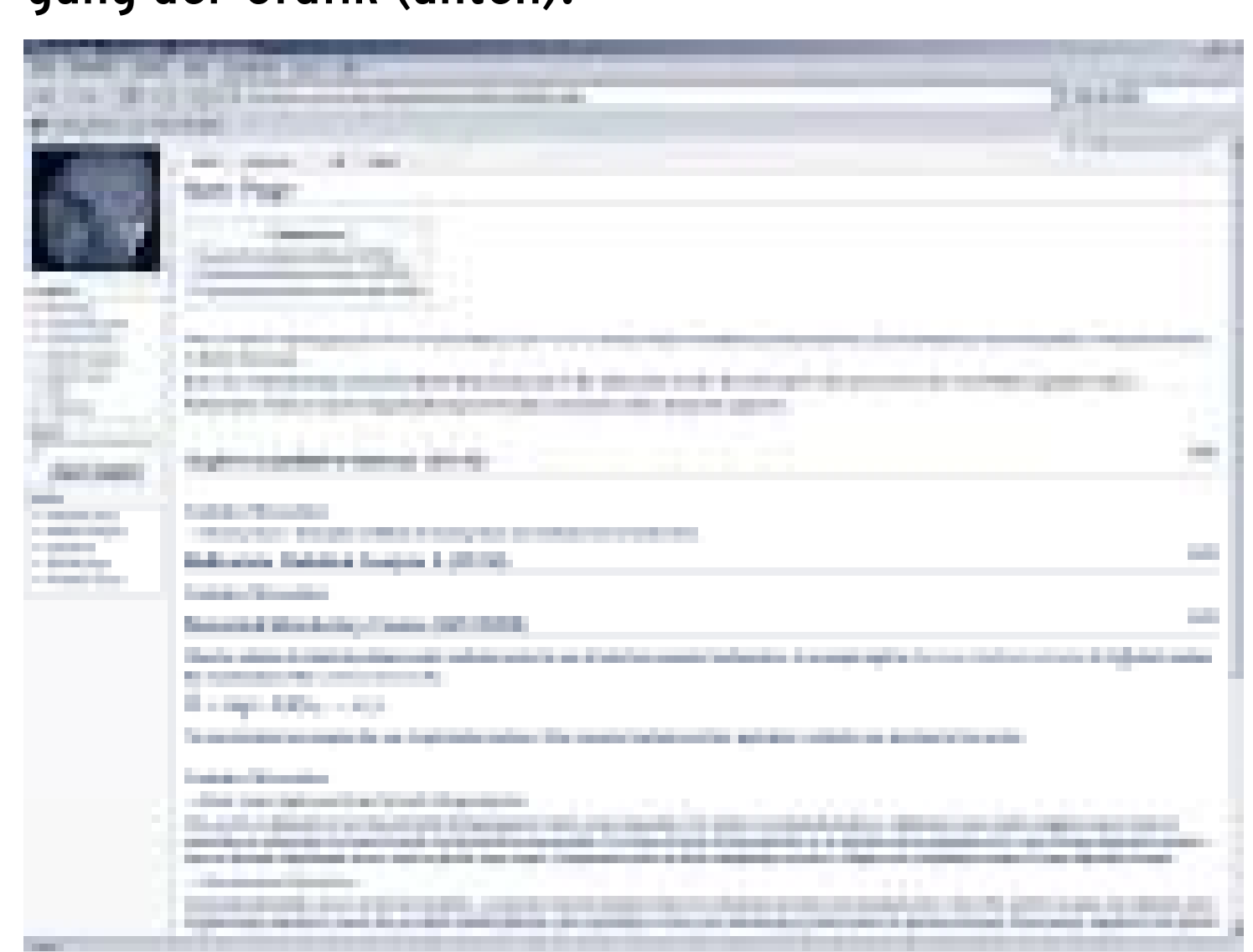
Die technische Grundlage des StatWiki ist das MediaWiki, welches auch für die Wikipedia genutzt wird. Ursprünglich haben wir die Kombination Zope, ZWiki und LatexWiki (ZZL) genutzt, jedoch ist die Entwicklung des LatexWiki 2005 beendet worden. Die Einbettung von Formeln mittels LaTeX ist jedoch unumgänglich für statistische Lehrinhalte.

Der modulare Aufbau des MediaWiki erlaubt die Erweiterung desselben durch PHP Programme. Bereits im Rahmen des Multimediaförderprogrammes 2004 wurden für die ZZL Version Pythonskripte für die Einbettung von R Programmen, die Tabellen und Grafiken erzeugen, entwickelt.

TeachWiki

Das englisch- und deutschsprachige TeachWiki, das auf den gleichen Grundlagen funktioniert wie das StatWiki, nutzen wir für Abschlussarbeiten der Studenten in den verschiedenen Lehrveranstaltungen des Lehrstuhles Statistik.

Die Grafik unten rechts zeigt einen Ausschnitt eines Textes zur Behandlung von fehlenden Daten (oben), das Ergebnis von 500 Simulationen für verschiedene Parameter eines Datensatzes bei einer bestimmten Ersetzungsmethode der fehlenden Werte der Daten (mitte) und das eingebettete R Programm zur Erzeugung der Grafik (unten).



HU Berlin | Institut für Statistik und Ökonometrie
Siegbert Klinke | sigbert@wiwi.hu-berlin.de

Multimedia-Förderprogramm 2005

