

**STATISTIK 2 (107.325) WS 2014**  
**COMPUTERSTATISTIK (107.258) WS 2014**

**Übung 4**

**9. November 2014**  
**Dutter**

13. Untersuchen Sie den Datensatz „werner\_bcd“ bezüglich den Variablen „CHOLSTRL“ und „ALBUMIN“:
- Korrelation: Von allen Werten der beiden Variablen und (2) von den Werten abhängig von „BRTHPILL“.
  - xy-Plot der Werte: Unterscheidung der Farbe/Symbole bezüglich der Variablen „BRTHPILL“.
  - Grafische Veranschaulichung in jeweils einer Dimension (z.B. Boxplots).  
Bitte um Kommentare.
14. Untersuchen Sie genauer den Datensatz „werner\_bcd“ bezüglich den Variablen „CHOLSTRL“ und „ALBUMIN“:
- t-Tests auf Gleichheit der zugrundeliegenden Mittelwerte von „CHOLSTRL“ der beiden Gruppen von „BRTHPILL“.
  - Analog für „ALBUMIN“.
  - Test auf Gleichheit der Mittelwerte bei zweidimensionaler Betrachtung.
15. In der Inskriptionsstatistik der TU Wien werden Hörerzahlen der Studienrichtung „Technische Mathematik“ (SS2011) mit 559 männliche und 252 weiblich Inskribierte ausgewiesen.
- Frage: Sind diese Zahlen voneinander signifikant verschieden (oder rein zufällig verschieden).
- Lösungsansatz: Annahme: Ein/e Studierender/Studierende ist bezüglich Geschlecht als zufällig zuzuordnen. D.h. die Zufallsvariable  $X$  liefert 0 oder 1, entsprechend Geschlecht männlich oder weiblich. Angenommen, die Wahrscheinlichkeit von weiblich, d.h.  $P(X = 1)$  sei gleich  $p$ . Unter der Annahme der Unabhängigkeit der Studierenden ist  $\sum_{i=1}^n X_i$  binomialverteilt  $Bi(n, p)$ , wobei  $n$  die Gesamtanzahl der betrachteten Studierenden bezeichnet. Die Binomialverteilung kann für große  $n$  und  $p$  nicht allzu verschieden von .5 mit der Normalverteilung mit Mittel  $np$  und Varianz  $p(1-p)n$  gut approximiert werden. (Damit kann der übliche t-Test `t.test` in  $\mathbb{R}$  verwendet werden.)
- Hier ist der Unterschied wohl offensichtlich signifikant. Wie ist es in der Studienrichtung Architektur, wo das Verhältnis 1618 zu 1635 heisst?
- Bitte, rechnen Sie konkrete Signifikanztests für diese Fälle.
16. Nehmen wir an, dass in einer Übungsgruppe nur ein Student und mehrere Studentinnen sitzen. Wie gross muss die Anzahl der weiblichen Studentinnen sein, damit statistisch ein signifikanter Unterschied in der Anzahl erscheint, genauer,

dass die Anzahl weiblicher Teilnehmer signifikant größer als die der männlichen ist? (Wie so häufig, nehmen wir  $\alpha = .05$ .)

Welche andere (ev. bessere) Tests als der approximative t-Test kommen in Frage?

Bitte, stellen Sie die Ausarbeitung in Form eines pdf-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Kurz-Listing des Programmcodes (Funktion) in die TUWEL-Seite

<https://tuwel.tuwien.ac.at/course/view.php?idnumber=107258-2014W>  
bis zum 14. November 2014, 23:45 Uhr.

Bitte den Namen des pdf-File folgendermaßen:

**name\_exer\_4.pdf**

wobei 'name' für den Familiennamen steht.

Empfehlenswert ist es, den Bericht mit „Sweave“ zu erstellen. (Vorlage siehe <http://www.statistik.tuwien.ac.at/public/dutt/vorles/> .)

**Kreuzen Sie außerdem bitte im Internet** jede Nummer des Übungsbeispiels an, das Sie dann in der Übungsstunde an der Tafel (mit Beamer-Unterstützung) vorrechnen wollen und können. Der Termin ist üblicherweise 3 Stunden vor der Übung, d.h. Mo., 11:00. Siehe <http://www.statistik.tuwien.ac.at/cgi-bin/uebbsp/bspstart.cgi> .