

# Übung 6

Rudolf Dutter & Matthias Templ  
Computerstatistik (107.258) WS 2007  
28. November 2007



## 1 Daten

Für diese Übung werden wiederum die EU-SILC Daten aus dem Jahre 2004 verwendet.

## 2 Übungsbeispiele

Bsp 17 Varianzanalyse (Modell 1): Variable P120040 gibt die Zufriedenheit der befragten Personen bezüglich Ihrem persönlichen Einkommen an. Testen Sie, ob mind. eine Gruppe aus Variable P120040 sich bezüglich dem tatsächlichen „mittleren“ Brutto-Einkommen (signifikant) unterscheidet, d.h. kann die Nullhypothese (die mittleren Brutto-Einkommen bzgl. aller Faktoren sind gleich) verworfen werden? Verwenden Sie dazu Funktion `anova()` oder Funktion `aov()`. Interpretieren Sie Ihr Resultat.

Bsp 18 Erzeugen Sie parallele Boxplots des Brutto-Einkommens gesplittet nach Variable P120040. Transformieren Sie Ihre Daten zuerst auf eine symmetrischere Verteilung um die Gruppen besser vergleichen zu können. Zum Transformieren der Werte können Sie die Funktion `box.cox()` aus dem Paket `car` verwenden. Die *box-cox* Transformation ist folgendermaßen definiert:

$$y_i^{(\lambda)} = \begin{cases} \frac{(y_i+c)^\lambda - 1}{\lambda} & , \text{ wenn } \lambda \neq 0 \\ \log(y_i + c) & , \text{ wenn } \lambda = 0 \end{cases}$$

Wählen Sie dabei passende Parameter für die *box-cox* Transformation.

Hinweis: Parallele Boxplots kann man von Variablen V1, V2 und V3 folgendermaßen erzeugen:

```
boxplot(V1,V2,V3)
```

Hat man ein (transformiertes) ‘data.frame’ namens ‘eu1’, mit der Variablen “pek\_g” und einem Faktor, z.B. “P120040”, (dessen Levels sinnvoll geordnet sind) dann funktioniert in eleganter Weise `boxplot(pek_g ~ P120040, data = eu1)`

## 3 Abgabe

Bitte senden Sie die Ausarbeitung in Form eines *pdf*-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Listing des Programmcodes (Funktion) an Ihren jeweiligen Gruppenleiter bis 5. Dezember 2007

r.dutter@tuwien.ac.at  
oder  
m.templ@tuwien.ac.at

Bitte den Namen des *pdf*-File wieder folgendermaßen:

**name\_UE\_6.pdf**

wobei ‘name’ für den Familiennamen steht.

Bitte beachten Sie, dass Sie zusätzlich Ihre gelösten Beispiele für die Übung ankreuzen müssen.