

Gedächtnisklausur Multivariate-Verfahren 18.02.2021 (WISE 20/21) bei Leonhart

1. Grundlagen/Datenvorbereitung:

1a. Bei einer Variable werden die fehlenden Werte mit der Mittelwertersetzung geschätzt. Hat das einen Einfluss auf die Korrelation dieser Variable x mit einer weiteren Variable y?

1b. Was sagt der p-Wert in der Inferenzstatistik aus?

2. Faktorenanalyse:

Gegen sei eine Hauptkomponentenanalyse mit 40 Variablen. In der quadrierten Faktorladungsmatrix kann man bestimmte Werte ablesen.

2a. Welchen Wert kann man über die Spaltensumme ablesen und was sagt dieser inhaltlich aus?

2b. In welchem Wertebereich liegt dieser Wert?

3. Regression

Nennen Sie einen der vielen Kennwerte, der Multikollinearität erfasst. Wie kann man mit dem Problem der Multikollinearität bei einer multiplen Regression umgehen?

4. Multiple logistische Regression & Diskriminanzanalyse:

Sie müssen sich zwischen einer Diskriminanzanalyse und multinominalen logistischen Regression entscheiden. Sie haben vier Gruppen. Nennen Sie jeweils einen Punkt, der für das jeweilige Verfahren spricht.

5. Manova:

1. Variable Probandenmotivation, erhoben als "niedrig", "mittel", "hoch". Was würde dafür sprechen die Variable als Faktor vs. Als Kovariate aufzunehmen?

2. Welchen Typ der Quadratsummenzerlegung muss man bei leeren Zellen wählen?

6. Clusteranalyse:

a. Welche Methode führt zu gleich großen Gruppen + welches Skalenniveau wird dabei vorausgesetzt

b. Häufigstes verwendetes Distanzmaß nennen

7. SEM:

Modell, mit 2 Faktoren mit jeweils 4 Variablen, wobei bei zwei Faktoren zu 0.87 korreliert haben. Die SMC waren überwiegend über 0.4, die Fehler der manifesten Variablen haben zu ca. 0.3 korreliert. Hohe Korrelation zwischen den Faktoren (.87)

a) globale und lokale Passung bewerten (CMIN war signifikant, CMIN/df-Verhältnis war groß...)

b) Was würden sie kritisch besonders kritisch bewerten?

c) Welche Restriktionen hätte es für das Modell aus 1a), wenn man die latenten Konstrukte unabhängig annehmen würde

8. HLM:

Jeweils ein konkretes Variablen-Beispiel machen, bei welchen Variablen eine Zentrierung Sinn machen würde und wann nicht. Art der Zentrierung benennen.

SPSS / R Teil:

- Output Diskriminanzanalyse zur Vorhersage der Remission bei Krebspatientinnen (2 Gruppen - 1 Funktion).
 - a) Was ist besonders bei diesem schrittweisen Vorgehen? (*Insgesamt war bei dem Beispiel nur sehr wenig Varianz aufgeklärt*)
 - b) Bewerten Sie das gesamte Modell. Was ist bei dem Bericht und der Diskussion kritisch zu erwähnen?
- Gleiche Daten: Output logistische Regression (Vorhersage von Rückfällen nach Brustkrebs):
 - a) Regressionsgleichung aufstellen
 - b) bewerten, ob man das Modell im klinischen Kontext anwenden würde