

Test nach Cochran

Inhaltsverzeichnis

Test nach Cochran	2
Lernhinweise	2
Einführung	2
Theorie & Fallbeispiel (1-4)	2
1. Ausgangslage und Fragestellung	2
2. Voraussetzungen, Arbeits- und Alternativhypothesen	4
3. Datenanalyse mit Hilfe von SPSS	4
4. Interpretation der Ergebnisse	6

Test nach Cochran

Lernhinweise

Dieser Lernschritt gilt dem Test nach Cochran, der für Untersuchungsdesigns entwickelt wurde, in deren Rahmen ein dichotom skaliertes Merkmal mehrfach erhoben wird. Wir beschränken uns dabei auf eine empirische Darstellung des Untersuchungsdesigns und die Datenauswertung mit SPSS. Die an einer manuellen Beschreibung Interessierten finden die entsprechenden Angaben in Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer. Kapitel 5.3.1.

Benötigte Vorkenntnisse

- Eigenschaften der Skalenniveaus
- Grundprinzip entscheidungsstatistischer Verfahren

Geschätzte Bearbeitungszeit

Nach Konsultation der in Ihrem Curriculum vorgesehenen Vorbereitungslektüre kann dieser Lernschritt in ca. 10 Minuten bearbeitet werden.

Einführung

Im Bereich der sog. Wirkungsforschung stellt sich ab und zu das Problem, dass ein interessierendes Merkmal in derselben Stichprobe mehrfach erhoben werden muss. Meist soll dabei die Wirkung eines Treatments (d.h. einer Interventions-Massnahme oder einer Behandlung im weitesten Sinne) bezüglich seiner Auswirkungen auf das interessierende Merkmal untersucht werden.

Ist das interessierende Merkmal intervall-, proportional- oder ordinal-skaliert, so stehen uns für die statistische Bearbeitung derartiger Untersuchungsdesigns verschiedene "varianzanalytischen Verfahren" zur Verfügung. Ist das Merkmal aber dichotom skaliert, so sind Varianzanalysen nicht einsetzbar; wir benutzen ein Verfahren, das von Cochran vorgeschlagen wurde. Auf dieses soll im folgenden ganz kurz und sehr summarisch eingegangen werden.

Theorie & Fallbeispiel (1-4)

Inhaltsübersicht

- [1. Ausgangslage und Fragestellung](#)
- [2. Voraussetzungen, Arbeits- und Alternativhypothesen](#)
- [3. Datenanalyse mit Hilfe von SPSS](#)
- [4. Interpretation der Ergebnisse](#)

1. Ausgangslage und Fragestellung

Zur Illustration der Problemstellung und des generellen Vorgehens kann uns das folgende Beispiel dienlich sein:

Test nach Cochran

Im Rahmen der Evaluation eines psychotherapeutischen Trainingsprogramms für Patienten mit gravierenden Schlafstörungen wurde eine Stichprobe von 9 Managerinnen und Managern u.a. bezüglich ihrer subjektiven Einschätzung der Schlafqualität befragt. Die erste Befragung erfolgte vor Beginn des Trainingsprogramms, die sechs weiteren Befragungen jeweils am Montag im Verlauf des 7-wöchigen Trainingsprogramms.

Festgehalten wurde dabei das dichotom skalierte Merkmal "Subjektive Einschätzung der Schlafqualität in der Nacht vom Sonntag auf den Montag" mit den beiden möglichen Ausprägungen "befriedigend" oder "unbefriedigend". Codieren wir die Angaben der Probandinnen und Probanden mit 1 für den Ausprägungsgrad "unbefriedigend" und 2 für "befriedigend", so lassen sich die Erhebungsdaten in einer Tabelle zusammenfassen:

Den Zeilen der Tabelle kann der zeitliche Verlauf der Einschätzungen der Probanden entnommen werden, den Zeilensummen L_i die Anzahl positiver Einschätzungen der einzelnen Personen.

Die Spalten zeigen die Einschätzungen der ganzen Stichprobe zu den jeweiligen Erhebungszeitpunkten, die Spaltensummen T_j die Zahl der positiven Einschätzungen zu den einzelnen Erhebungszeitpunkten.

Soll nun anhand dieser Daten über die "Wirksamkeit" des Trainingsprogramms eine Aussage gemacht werden, so interessiert natürlich die Frage, inwieweit sich die Verteilung der Einschätzungen der Probanden im Verlauf des Trainingsprogramms verändert.

In der Sprache der Statistiker lautet unsere Frage: Sind die Veränderungen in der Verteilung des dichotom skalierten Merkmals "Subjektive Einschätzung der Schlafqualität" über alle 7 Erhebungen so gravierend, dass bei der Interpretation der Veränderungen der Zufall ausgeschlossen werden kann?

2. Voraussetzungen, Arbeits- und Alternativhypothesen

Unsere Frage lässt sich nach der Idee von Cochran über den simultanen Vergleich der 7 Verteilungen des Merkmals beantworten, wenn die Voraussetzung $n \cdot m > 30$ erfüllt ist (n: Anzahl Probanden; m: Anzahl Erhebungen).

Auf die Besprechung der von Cochran entwickelten Prüfgrösse Q und der zugehörigen Prüfverteilung wollen wir an dieser Stelle verzichten und unser Beispiel direkt mit Hilfe von SPSS bearbeiten. Die an einer manuellen Bearbeitung Interessierten finden Berechnungsformeln für Q in Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer. Kapitel 5.3.1..

- **Prüfung der Voraussetzungen und Wahl des Prüfverfahrens:**

Das interessierende Merkmal ist die dichotom skalierte subjektive Einschätzung der Schlafqualität. Dieses Merkmal wurde an $n = 9$ Patienten zu $m = 7$ Zeitpunkten erhoben.

Uns interessiert die Frage, ob sich die Verteilungen von positiven und negativen Nennungen im Verlauf der 7 Erhebungen so stark unterscheiden, dass bei der Interpretation der beobachteten Veränderungen der Zufall ausgeschlossen werden darf.

Da 7 Verteilungen eines dichotomen Merkmals simultan verglichen werden sollen, ist das Prüfverfahren nach Cochran angezeigt, wenn die einzige formelle Voraussetzung $n \cdot m > 30$ erfüllt ist. Dies ist mit $n = 9$ und $m = 7$ der Fall.

- **Arbeitshypothese H_0 :**

Die 7 Verteilungen unterscheiden sich nur zufällig.

- **Alternativhypothese H_1 :**

Mindestens zwei der 7 Verteilungen unterscheiden sich nicht nur zufällig.

3. Datenanalyse mit Hilfe von SPSS

Wir geben unsere Datentabelle in ein Datenfile **schlaf_quali.sav** ein, wobei wir die Angaben der Patienten wie folgt codieren:

'unbefriedigende Schlafqualität' = 1

'befriedigende Schlafqualität' = 2

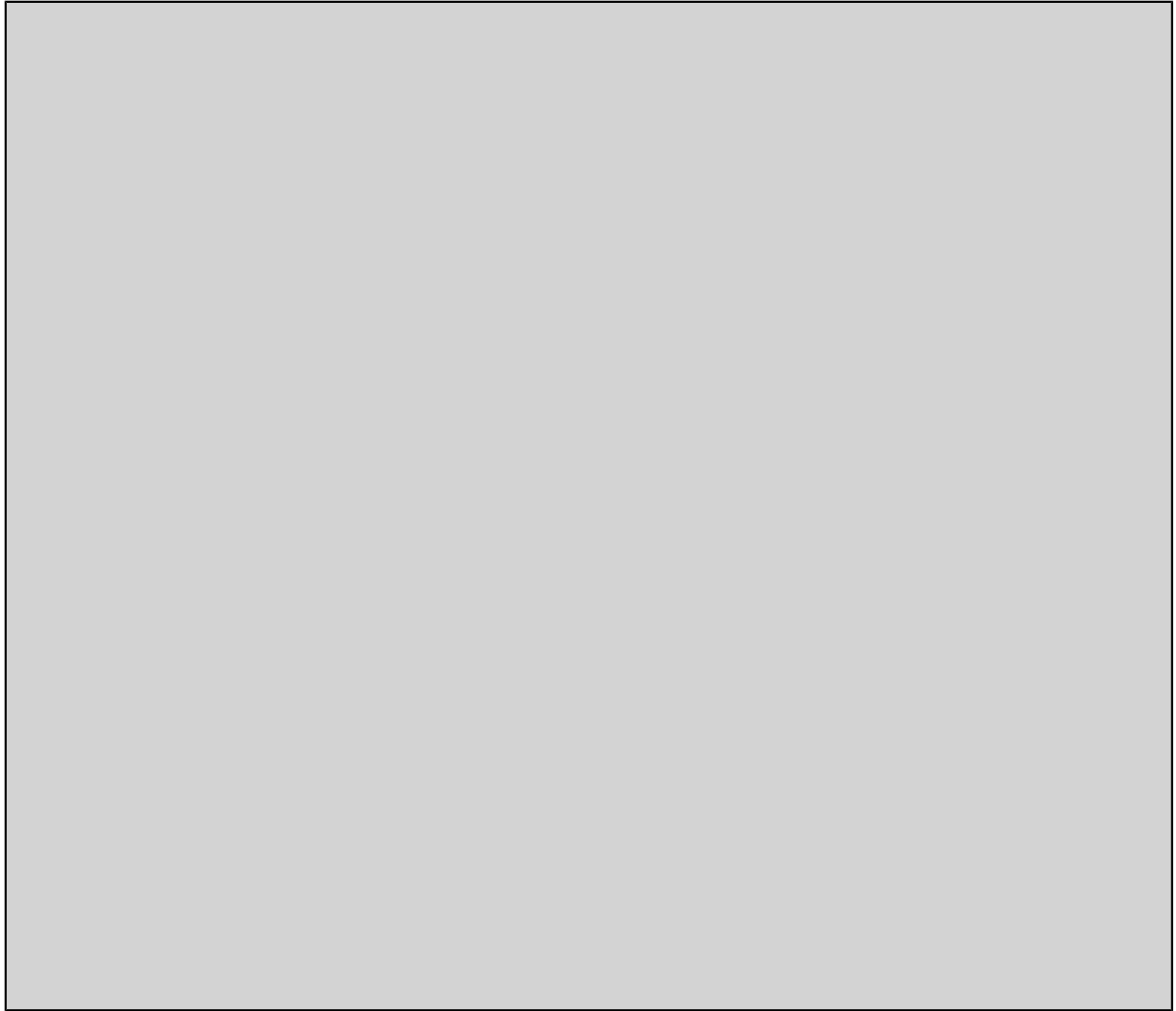
Wir analysieren diese Daten mit den folgenden SPSS-Befehlen:

```
GET FILE 'X:\\SPSS_Dat\\schlaf_quali.sav'.
```

```
LIST prob_nr erheb1 erheb2 erheb3 erheb4 erheb5 erheb6 erheb7.
```

Als erstes lesen wir unser Datenfile **schlaf_quali.sav** ein und lassen uns zur Kontrolle eine Liste der eingegebenen Daten ausgeben.

SPSS output 1

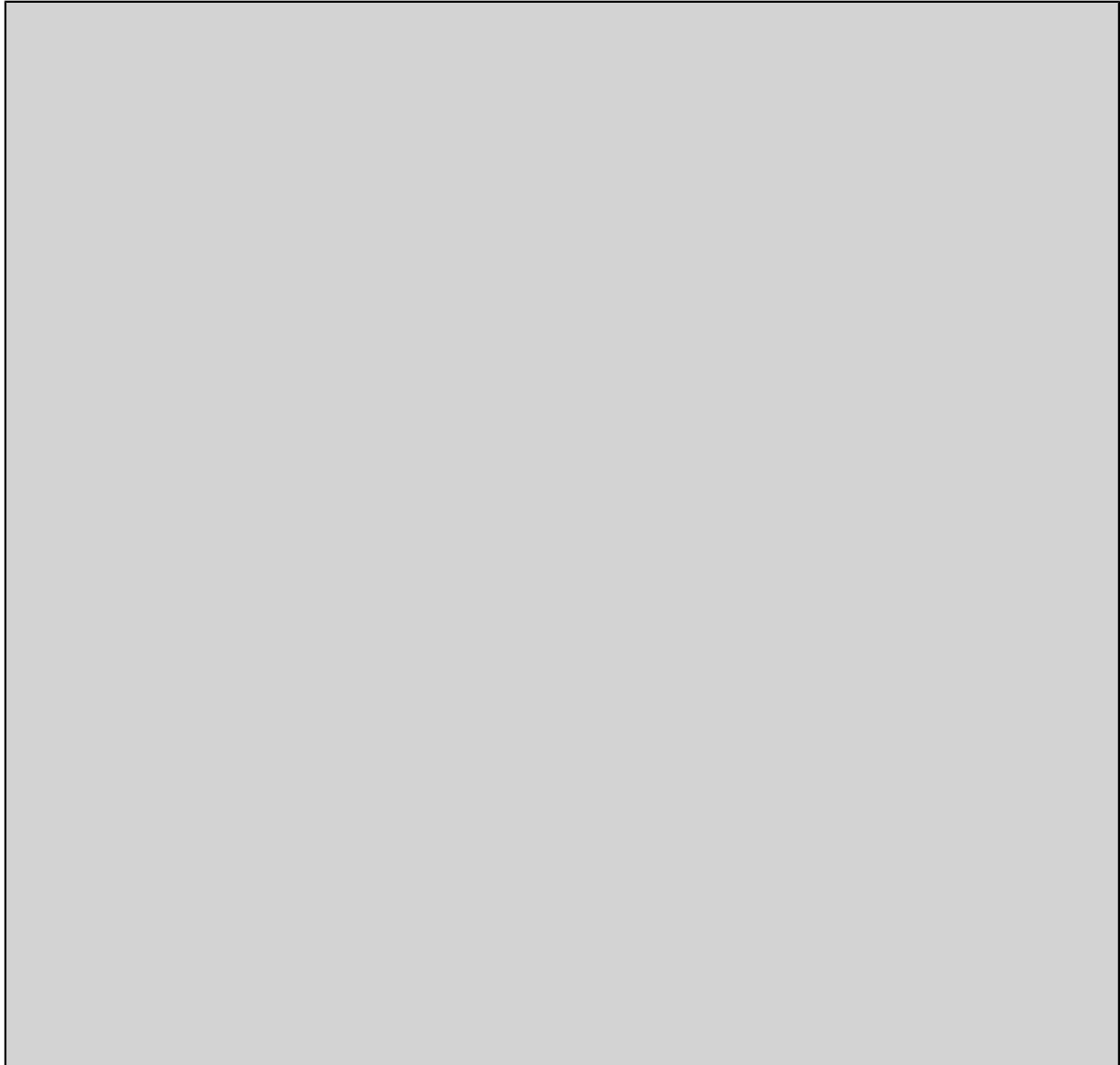


GET FILE 'X:\\SPSS_Dat\\schlaf_quali.sav'.

NPAR TESTS /COCHRAN = **erheb1 erheb2 erheb3 erheb4 erheb5 erheb6 erheb7.**

Als nächstes verlangen wir eine nichtparametrische Datenanalyse nach Cochran (NPAR TESTS /COCHRAN), konkret den simultanen Vergleich der Daten aus den 7 Erhebungen.

SPSS output 2



4. Interpretation der Ergebnisse

Die SPSS-Ausgabe zeigt uns als erstes die Häufigkeitsverteilungen des dichotomen Merkmals 'Schlafqualität' in den 7 Erhebungen. Dann folgen die Kenngrößen zum Test nach Cochran: SPSS nennt den Ausprägungsgrad von Q und die zweiseitige Überschreitungswahrscheinlichkeit dieses Ausprägungsgrades im Rahmen der Prüfverteilung.

Da die Überschreitungswahrscheinlichkeit kleiner ist als 5%, lehnen wir unsere Arbeitshypothese H_0 zu Gunsten der Alternativhypothese H_1 ab. Dies bedeutet, dass sich mindestens zwei der sieben Verteilungen signifikant unterscheiden. Dies legt den formal nicht beweisbaren inhaltlichen Schluss nahe, dass mit einer Irrtumswahrscheinlichkeit $p < 5\%$ angenommen werden darf, dass das zwischen den Erhebungen applizierte Treatment - im gegebenen Beispiel das psychotherapeutische Training - auf die Verteilung des interessierenden Merkmals 'Schlafqualität' Einfluss nimmt.

Test nach Cochran

Was dem Test nach Cochran nicht entnommen werden kann, ist die Zahl der signifikanten Unterschiede, und zwischen welchen Datensätzen diese bestehen. So werden im Anschluss an eine Datenanalyse nach Cochran häufig auch Paarvergleiche durchgeführt, d.h. es werden ausgewählte Datensätze mit dem Verfahren von McNemar gezielt auf signifikante Unterschiede geprüft. Diese plausible Vorgehensweise birgt aber einige Tücken, auf die wir im Rahmen der Besprechung der Varianzanalysen noch detailliert eingehen müssen.