

Praktische Übung zu Methoden der medizinischen Wissenschaft (SSM2)

Lernziele

1. Medizinische Daten - Vom Fragebogen zum Datenbankausdruck

Die Präsentation und Übungen für die Studierenden stehen über das Internet zur Verfügung.

Unterschiedliche Arten von Daten verlangen unterschiedliche Darstellungen und Auswertungen. Daher sollen zunächst an einem vorliegenden Datenerhebungsbogen die verschiedenen Arten von medizinischen Daten erarbeitet werden (qualitativ – quantitativ; nominal - ordinal - metrisch; diskret – stetig). In diesem Zusammenhang soll auch die Gestaltung eines Datenerhebungsbogens angesprochen werden (z.B. Identifikation der Patienten, Modularität des Aufbaus, Prinzipien der Verschlüsselung). Begriffe der Qualität von Daten (Vollständigkeit, Richtigkeit, Genauigkeit, Reliabilität, Validität) sollten unter Hinweis auf die Vorlesung diskutiert werden.

Die Haltung von Daten am Computer soll anhand des auf dem Fragebogen basierenden Datensatzes erklärt werden (Einführung in SPSS, Eingabemaske, Labelvergabe, Kodierung). Eine (geringe) Anzahl von Fragebögen wird von den Studierenden in die Datenbank eingegeben. Methoden der Qualitätssicherung (z.B. Source data verification, Plausibilitätsprüfung, Double data entry) und des Datenschutzes (z.B. Anonymisierung) sollen angeschnitten werden. Das Ziel dieser Einheit besteht im Datenbankausdruck des selbst erstellten Datensatzes.

Je nach verbleibender Zeit können noch einige einfache Bearbeitungs- und Auswertungsschritte der Daten unter Hinweis auf die Erklärungen in späteren Übungen durchgeführt werden (Sortieren, Transformieren, Selektieren, Stabdiagramme, Scatterplots, usw.).

2. Tabellen und Diagramme

Für die folgenden Übungen soll ein Datensatz verwendet werden, der über Internet verfügbar ist. Der Begriff der absoluten und relativen Häufigkeit wird anhand von qualitativen und ordinalen Merkmalen erläutert.

Die Vierfeldertafel zur Beschreibung des Zusammenhangs zwischen zwei dichotomen Merkmalen soll zusammen mit den verschiedenen Prozentangaben verstanden und anhand von Beispielen erstellt werden. In diesem Punkt soll bereits auf die statistische Aufgabe des späteren Vergleichs von zwei Anteilen aus einer Vierfeldertafel hingewiesen werden (siehe Punkt 5). Die Erweiterung der Häufigkeitstabellen auf Tafeln mit r Zeilen und c Spalten soll ebenfalls an Beispielen des Datensatzes erarbeitet werden.

Die Probleme der Interpretation bei Häufigkeitstabellen von Merkmalen mit Mehrfachnennungen werden anhand von Beispielen durch den Übungsleiter aufgezeigt. Auch das wichtige Problem der Stichprobeneinheit (Patienten, Aufnahmen auf die Intensivstation, Implantat) wird an Beispielen diskutiert. Die Studierenden sollen Häufigkeitsdarstellungen von qualitativen Daten in graphischer Form von Kreisdiagrammen anhand der gleichen Beispiele erarbeiten. Für ordinale Merkmale wird die graphische Darstellung in Form von Stabdiagrammen erarbeitet. Dabei soll das Konzept auch auf die simultane Darstellung von Stabdiagrammen relativer Häufigkeiten in mehreren Gruppen erweitert werden.

Hinweise auf irreführende graphische Darstellungen sollen gegeben werden („Achsenbruch“, graphische Darstellung ohne den Nullpunkt, 3D-Darstellungen). Hier könnten auch Beispiele aus den Medien von Interesse sein. Je nach verbleibender Zeit könnten Ausgestaltungen der Graphiken (Stapelung, Farben) geübt werden.

3. Darstellung von Messdaten: univariat

- Histogramme

Die graphische Darstellung von Messvariablen mittels Histogrammen wird durch die Erstellung einer Rangliste erarbeitet. Die Wahl der Klassenbreite wird über die unterschiedliche Darstellung des gleichen Datensatzes (mit unterschiedlichen Intervallgrenzen) diskutiert. Das Konzept soll auf die Darstellung von Histogrammen relativer Häufigkeiten in mehreren Gruppen erweitert werden. In diesem Punkt soll auf das Problem des späteren Vergleichs von Mittelwerten zweier Gruppen verwiesen werden (siehe dazu Kapitel 6).

Statistische Lage- und Streuungsmaße werden für das Beispiel von Mittelwert und der Standardabweichung berechnet. Neben der formelmäßigen Darstellung werden die Maße erklärt (z.B. über den Schwerpunkt und die durchschnittliche quadratische Abweichung) und an verschiedenen Beispielen graphisch durch den Übungsleiter interpretiert. Auch der Einfluss von Ausreißern (einzelne Werte, die stark von allen anderen abweichen) auf die berechneten Mittelwerte und Standardabweichungen werden anhand eines Datensatzes veranschaulicht.

Der ‚Standard Error of the mean‘ (SEM) als Streuungsmaß nicht für die Einzeldaten, sondern als Maß für die „Unsicherheit“ für den Mittelwert selbst wird erklärt und an einem Beispiel von den Studierenden errechnet. Es erfolgt ein Hinweis auf Konfidenzintervalle als Schätzung eines Bereichs für den unbekanntem Mittelwert aus der Stichprobe. Die Darstellung von Verlaufskurven über den Mittelwert und SEM wird an einem Beispiel durch den Übungsleiter veranschaulicht. Der Einfluss von systematisch fehlenden Werten auf die Mittelung von Verlaufskurven wird diskutiert.

- Median, Quartile, Boxplot

Am Beispiel des Medians wird über die Erstellung einer Rangliste ein Lagemaß eingeführt, das robust gegenüber Ausreißern ist. Zur Erklärung wird eine Stichprobe mit ungeradem Stichprobenumfang verwendet. An Beispielen wird gezeigt, dass der Median unverändert bleibt, wenn einzelne Werte weit von den anderen stark abweichen. Durch einen kleinen Schritt der Verallgemeinerung werden Quartile definiert. An Beispielen wird auch deren Robustheit gegenüber einzelnen Ausreißern

erarbeitet. Der Interquartilsabstand (IQR) wird als Streuungsmaß eingeführt. Dieser charakterisiert die Breite eines Bereichs, in dem die zentralen 50% der Daten liegen.

Median, Quartile, Minimum und Maximum bilden die Basis für die graphische Darstellung in Form von Boxplots. Dieses Konzept wird erklärt, von den Studierenden am Rechner umgesetzt und auf eine Gegenüberstellung mehrerer Gruppen erweitert.

4. Darstellung von Messdaten: bivariat

- Regression, Korrelation

Das Problem der graphischen Darstellung von zwei an einer Person gemessenen Variablen wird anhand eines Streudiagrammes eingeführt.

Der Korrelationskoeffizient als Maß für den linearen Zusammenhang wird definiert und anhand verschiedener Beispiele zusammen mit dem Streudiagramm von den Studierenden berechnet. Die Unabhängigkeit des Korrelationskoeffizienten bzgl. der Messeinheit wird vom Übungsleiter anhand von Beispielen erklärt. Der Einfluss von stark abweichenden einzelnen Beobachtungen auf den Wert des Korrelationskoeffizienten wird anhand verschiedener Beispiele dokumentiert. Schließlich wird auf das spätere Problem der statistischen Prüfung eines zwischen den Variablen bestehenden Zusammenhangs verwiesen (siehe Kapitel 6).

Die Berechnung einer Regressionsgeraden zur Bestimmung des Wertes einer Variablen aus dem Wert einer anderen Variablen wird demonstriert. Die Regressionsgerade wird von den Studierenden anhand von Beispielen berechnet und gezeichnet. Die Abhängigkeit des Wertes des Regressionskoeffizienten von den Merkmalseinheiten und seine Interpretation werden diskutiert. An einem Beispiel wird demonstriert, wie durch die Transformation eines Merkmales ein linearer Zusammenhang erreicht werden kann.

5. Der Vergleich von Häufigkeiten: Der CHI-Quadrat-Test (χ^2 -Test)

Das Prinzip des Vergleichs von zwei Häufigkeiten wird am Beispiel von Anteilen überlebender Patienten in zwei unterschiedlichen Diagnosegruppen in Form einer Vierfeldertafel aus der 2. Übung erarbeitet.

Die „skeptische“ Ausgangshypothese (Nullhypothese), die es zu widerlegen gilt, wird formuliert als: „Die zugrunde liegenden Überlebenswahrscheinlichkeiten für die beiden Diagnosen sind gleich“. Die (zweiseitige) Alternative ergibt sich als das Komplement, „Die beiden Überlebenswahrscheinlichkeiten sind ungleich“. Die unter der Gleichheitshypothese erwartete Verteilung der Anzahl der verstorbenen Patienten auf die beiden Gruppen in dieser Tafel soll erarbeitet werden. Es soll an dem Beispiel in heuristischer Weise eine Größe erarbeitet werden, die eine Quantifizierung der Abweichung zwischen den tatsächlich beobachteten Anzahlen und diesen erwarteten Anzahlen erlaubt. Die Formel für diese als Summe aller quadrierten standardisierten Abweichungen resultierende Prüfgröße (χ^2 -Wert) wird angegeben.

Das in der Vorlesung eingeführte Konzept des kritischen Bereichs, in dem die Zweifel an der Gültigkeit der Nullhypothese zu groß werden, wird für diese Prüfgröße

wiederholt. Die Entscheidungstafel des statistischen Tests wird gezeigt. Die kritischen Grenzen für diesen χ^2 -Wert werden unter verschiedenen Vorgaben für das Signifikanzniveau α des Tests tabellarisch gegenübergestellt und diskutiert. Das Testergebnis wird softwaremäßig berechnet und der gefundene p-Wert anhand der Tabelle der kritischen Grenzen für den χ^2 -Wert diskutiert. An einem theoretischen Beispiel mit gleichen beobachteten Anteilen überlebender Patienten und steigendem Stichprobenumfang wird das Problem der statistischen Power des χ^2 -Tests pragmatisch erklärt.

Das Konzept des Vergleichs von zwei Häufigkeiten wird auf den Vergleich von Häufigkeitsverteilungen in größeren rc -Tafeln verallgemeinert. Der analog zum Vierfeldertest gebildete χ^2 -Wert wird für ein Beispiel softwaremäßig berechnet und das Ergebnis zusammen mit den nun für die größere Tafel geänderten kritischen Grenzen diskutiert.

6. Andere statistische Tests

- ***Der t-Test zum Vergleich der Mittelwerte in unabhängigen Gruppen***
- ***Der t-Test zum Paarvergleich von zwei Mittelwerten***
- ***Die Prüfung auf Vorliegen einer Korrelation zwischen zwei Messvariablen***

In Analogie zur 5. Übung werden die entsprechenden (zweiseitigen) Hypothesen für die drei Probleme in verbaler Form definiert. Die Teststatistiken für den t-Test werden gezeigt und pragmatisch erklärt. An Beispielen werden die Tests softwaremäßig umgesetzt. Dabei sollen die Ergebnisse in Verbindung mit den graphischen Darstellungen (Histogramme aus der 3. Übung) der beobachteten Verteilungen diskutiert werden.

Der Vergleich von Mittelwerten in mehreren unabhängigen Gruppen wird eher im Black-Box-Verfahren an einem Beispiel umgesetzt. Auch der Test des Korrelationskoeffizienten und, nach Maßgabe der verbleibenden Zeit, ein Konfidenzintervall für den Anstieg einer Regressionsgeraden werden ohne Hinweis auf den methodischen Hintergrund an Beispielen berechnet. Allerdings werden auch dabei die Ergebnisse immer in Verbindung mit den graphischen Darstellungen der beobachteten Werte (Streudiagramm aus der 4. Übung) diskutiert.