

Anmerkungen zur Übungsserie 6: Einordnung verschiedener Grafiken

Grafiken für univariate stetige Daten

- Punktdiagramme
 - Überblick über den Bereich, in dem Beobachtungswerte liegen und wie stark sie streuen.
 - Probleme
 - * Zusammenfallende oder sehr nah beieinander liegende Beobachtungswerte sind im Diagramm nicht mehr unterscheidbar, so dass Punkte durch Überdeckung verloren gehen können.
 - * Das Problem ist zu beheben durch Verschiebung der Punktpositionen in die ungenutzte Richtung (vertikal bei horizontalen Punktdiagrammen) durch systematisches Stapeln (*gestapeltes Punktdiagramm*) oder durch zufälliges Verzittern (*verzittertes Punktdiagramm*).
 - Gestapeltes Punktdiagramm gibt Hinweis auf vorliegende Bindungen
- Histogramme
 - Ausgangspunkt: Klasseneinteilung der Beobachtungswerte.
 - Bemerkung
 - * Die Gestalt eines Histogramms hängt stark von der gewählten Klasseneinteilung (und auch des gewählten Gesamtintervalls) ab, deshalb sollte man ggf. etwas experimentieren, um ein möglichst aussagekräftiges Histogramm zu erzeugen.
- Box-Plot
 - Aussage über Lage, Streuung der Daten, Bereich der Datenwerte, ggf. Symmetrie.
 - Keine Aussage über Beobachtungsanzahl, Bindungen, allg. Verteilungsform.
 - Bei einer kleinen Anzahl von Beobachtungswerten sind Box-Plots nicht sehr aussagekräftig.

- Q-Q-Plot
 - Dient unter anderem zum Vergleich der Beobachtungswerte x_1, \dots, x_n mit einer theoretischen Verteilung.
 - Bemerkungen:
 - * Liegen mehrere aufeinanderfolgende Punkte auf einer waagerechten Linie zusammen, entsprechen diese übereinstimmenden Beobachtungswerten. Das Mehrfachauftreten übereinstimmender Beobachtungswerte wird auch als Bindung bezeichnet. Sie können zum Beispiel hindeuten auf
 - gerundete Beobachtungswerte
 - ungenaue Datenerhebung
 - das Ersetzen von fehlenden Werten durch Standardwerte
 - * Ungefähre Gerade \Leftrightarrow Verteilungsmodell passend.
 - * „Treppenstufen“ \Leftrightarrow Bindungen (gleiche Werte).
 - * „gegen S“ \Leftrightarrow Ausreißer? (schwere Verteilungsenden).

Grafiken für univariate diskrete Daten

- Balkendiagramm.
- Kreisdiagramm.

Grafiken für multivariate stetige Daten

- Streudiagramm (auch Scatterplot genannt).
 - Anfällig gegenüber Bindungen, die durch Überlagerung unsichtbar werden.
 - Verzerrung durch Ausreißer.
- Q-Q-Plot zum Vergleich zweier Merkmale.

Graphiken für multivariate diskrete Daten

- Gestapeltes Balkendiagramm.
- Paralleles Balkendiagramm.
- Mosaikplot.

Graphiken für gemischte multivariate Daten

- Paralleles Punktdiagramm.
- Parallele/ (gekerbte oder ungekerbte) Box-Plots.
- Farben und Symbole.