

Selbststudium Sitzung 2

Vera Husfeldt

Chur, 18. November 2022

1 Variable Alter

Erstellen Sie auf der Basis des Datensatzes DataLit eine neue Variable „Alter“ und berechnen Sie folgende Kennwerte:

- Modalwert
- Median
- Arithmetischer Mittelwert
- Spannweite
- Quartilsabstand
- Varianz
- Standardabweichung
- Schiefe
- Kurtosis

Überprüfen Sie zudem die Normalverteilungsannahme für diese neue Variable mithilfe des Augenscheins über das Histogramm mit Linie und mit dem QQ-Plot.

2 Berechnung von z-Werten

Auf der Grundlage von Daten aus einer Schweizer Schülererhebung wird aus verschiedenen Variablen (z.B. Angaben zum Beruf der Eltern, zur Elternausbildung sowie zur Anzahl von Büchern zu Hause) ein Index zur sozialen Herkunft erstellt. Dieser Index erscheint in einer neu gebildeten numerischen Variable im Datensatz, gibt also für jeden Fall in diesem Datensatz einen Skalenwert zur sozialen Herkunft an. Die neu gebildete Skala läuft von 0 (Wert mit der geringsten Ausprägung) bis 10 (Wert mit der höchsten Ausprägung).

Für die gesamte Schweiz liegt der arithmetische Mittelwert auf dieser Skala bei 5.6 und die Standardabweichung beträgt 1.8 (Zahlen sind von mir frei erfunden!). Die Verteilung entspricht einer Normalverteilung.

Der Mittelwert der Verteilung der Bündner Schülerinnen und Schüler liegt etwas tiefer als in der Gesamtschweiz, nämlich bei 5.1 mit einer Standardabweichung von 2.

2.1 Aufgabenstellung

- Wo in der Verteilung der Schweiz liegt der Bündner Mittelwert, bzw. wie viele Schweizer Schülerinnen und Schüler haben bzgl. der sozialen Herkunft einen tieferen Wert als die oder der typische Bündner Schülerin bzw. Schüler?
- In einem Bündner Ort beträgt der Mittelwert auf der Skala zur sozialen Herkunft 5.6, er ist also genauso hoch wie in der Gesamtschweiz. Was ist zur Lage dieses Wertes bezogen auf die Verteilung in Graubünden zu sagen? Also: Wie viele Schülerinnen und Schüler in GR liegen mit ihrem Wert darunter?