

1. Folgende Stichprobe beinhaltet 10 Beobachtungen<sup>2</sup> über (Haushalts-)Einkommen von Personen in Österreich und deren Stichprobengewichte<sup>3</sup>:

	1	2	3	4	5	6	7	8	9	10
eqIncome	26471	23687	12095	7004	35355	11216	4726	15605	21762	13216
weights	493	482	519	515	459	482	581	538	667	616

Die Beurteilung von Ungleichheit von Einkommen (Einkommensschere) ist von zentraler Bedeutung für die Politik. Ein oft verwendeter und beliebter Indikator, die Ungleichheit von oberen und unteren Einkommen zu beurteilen, ist der sogenannte Quintile-Share-Ratio (QSR).

Sei  $q_{0.2}$  und  $q_{0.8}$  das gewichtete 20% und 80% Quantile von den Stichprobenwerten  $x_i$ ,  $i = 1, \dots, n$  und Gewichten  $w_i$ ,  $i = 1, \dots, n$ .

Mit  $I_{\leq q_{0.2}} := \{i \in \{1, \dots, n\} : x_i \leq q_{0.2}\}$  und  $I_{> q_{0.8}} := \{i \in \{1, \dots, n\} : x_i > q_{0.8}\}$  lautet die Schätzung des QSR:

$$\widehat{QSR} := \frac{\sum_{i \in I_{> q_{0.8}}} w_i x_i}{\sum_{i \in I_{\leq q_{0.2}}} w_i x_i}$$

- Schätzen Sie den QSR mit ihren Daten. Da in der Vorlesung und Übung gewichtete Quantile nicht besprochen wurden, dürfen Sie die Quantile ungewichtet berechnen.
- Betrachten Sie die Formel und geben Sie eine Interpretation des QSR.
- Ist der QSR empfindlich gegenüber Ausreißer? Wenn ja/nein: Warum?

(Lösungsblatt:  $\widehat{QSR}$ )

(2 + 1.5 + 1.5)

2. Pärchen aufgepasst! So lautete die Schlagzeile zeitgerecht vor Weihnachten von CNN und in einer großen deutschen Zeitung basierend auf einer empirischen Studie. Die gefährliche Zeit für Beziehungen beginnt laut dieser Studie zwei Wochen vor Weihnachten - ab 10. Dezember. Grundlage dieser Studie ist die Änderung des Beziehungsstatus von 10000 Facebook-Mitgliedern in Facebook.

Die wenigsten Trennungen gibt es laut deren Studie übrigens zu Weihnachten, ein solcher Schritt wird laut CNN als *zu grausam* bewertet.

Von 10000 Facebook-Mitgliedern haben sich 62 Paare in der 'beziehungstechnisch schlimmsten' Woche vor Weihnachten getrennt (z. B. von *verheiratet* oder *in einer Beziehung* zu *Single* oder *es ist kompliziert*). Die Studie berichtet, das im Gegensatz dazu es in der beziehungstechnisch besten Zeit - in der ersten Juliwoche - nur 28 Trennungen gab.

- Schätzen Sie die Anteile von Trennungen und bestimmen Sie die 95%-Konfidenzintervalle zB mit Hilfe einer Approximation durch eine Normalverteilung. Überdecken sich die Konfidenzintervalle?
- Wenn Sie die Überschrift mit dem Inhalt der Studie vergleichen, welche Kritikpunkte fallen Ihnen ein?

(Lösungsblatt: Überdeckung ja/nein)

(3 + 1.5)

3. Abbildung 1 zeigt einen Scatterplot. Welche der folgenden Aussagen sind korrekt?

- Für  $X = 10.6$  wird der Wert von  $Y$  in etwa bei 100 erwartet.
- Die Standardabweichung von  $Y$  ist mindestens 6.
- Die Daten sind standardisiert.
- Die Korrelation ist mindestens 0.8.
- Das Mittel von  $X$  ist maximal 5.

(2.5)

(Lösungsblatt ankreuzen - zB ein Plus wenn zutreffend, ein Minus wenn nicht)

<sup>2</sup>Seriöse Aussagen können erst bei viel größeren Stichproben getätigt werden, aber zur einfacheren Berechnung mit dem Taschenrechner während der Prüfung wurde nur ein kleines Subsample gezogen.

<sup>3</sup>Ein Stichprobengewicht repräsentiert, vereinfacht gesprochen, die erwartete Anzahl von Personen mit gleichen Eigenschaften.

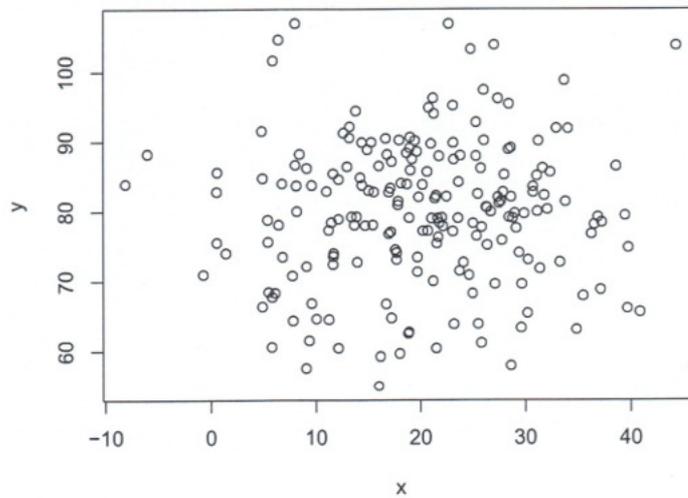


Abbildung 1: Scatterplot

4. Im Rahmen der PISA-Studie wurde in den einzelnen Ländern die durchschnittliche Klassengröße ermittelt. Wir bilden drei Klassen, wobei die Länder der ersten Klasse eine kleine, die Länder der zweiten Klasse eine mittlere und die Länder der dritten Klasse eine hohe Klassengröße besitzen. Aus jeder Klasse wurden jeweils 4 Länder ausgewählt. Die folgende Tabelle zeigt die Ergebnisse im Bereich naturwissenschaftliche Grundbildung dieser Länder:

	1	2	3	4
klein	497	478	478	501
mittel	379	551	552	421
hoch	487	516	532	500

- a) Nehmen Sie an, dass die Daten in den einzelnen Gruppen normalverteilt sind mit der gleichen Varianz  $\sigma^2$ . Stimmen die mittleren Punktezahlen in den drei Gruppen überein (Signifikanzniveau  $\alpha = 0.05$ )? (3)
- b) Zeichnen Sie einen Boxplot von allen Ergebnissen als einzige Messreihe aufgefasst (der Tabelle oben). Geben Sie eine kurze Interpretation ihres Boxplots. (3)

(Lösungsblatt: Wert der Teststatistik)

5. Bei einem Kartrennen nehmen 6 Studenten teil, wobei die Startreihenfolge zufällig ausgewählt wurde. Darunter befinden sich zwei Fahrer von der TU WIEN. Wie gross ist die Wahrscheinlichkeit, dass diese beiden TU-Studenten nebeneinander starten?

(2)

(Lösungsblatt: Wahrscheinlichkeit)