

Übungen – Einheit 4:

Aufgabe 4.1:

Sie sind daran interessiert, ob sich die Lohneinkommen („income“) von Personen in unterschiedlich großen Haushalten („hhszise“) unterscheiden. Bilden Sie 2 Gruppen und vergleichen Sie die Einkommen von Personen, die in Single-Haushalten leben, mit Personen, die in Mehrpersonen-Haushalten (mit zwei oder mehr Haushaltsmitglieder) wohnen. Formulieren Sie die entsprechende Null- und Alternativhypothese und testen Sie, ob sich die Einkommen der beiden Gruppen signifikant (auf dem 5%-Niveau) voneinander unterscheiden.

Berechnungen: siehe Excel.

Nullhypothese H_0 : Die durchschnittlichen Lohneinkommen von Personen in Single-Haushalten unterscheiden sich nicht von den durchschnittlichen Lohneinkommen in Mehrpersonenhaushalten:
 $income_{single} = income_{Mehrpersonenhaushalte}$

Alternativhypothese H_1 : Die durchschnittlichen Lohneinkommen von Personen in Single-Haushalten unterscheiden sich von den durchschnittlichen Lohneinkommen in Mehrpersonenhaushalten:
 $income_{single} \neq income_{Mehrpersonenhaushalte}$

Aufgabe 4.2:

Sie sind daran interessiert, ob Personen in Single-Haushalten signifikant mehr verdienen („income“) als Personen in Mehrpersonenhaushalten. Formulieren Sie die entsprechende Null- und Alternativhypothese und testen Sie die Hypothesen, indem Sie eine Irrtumswahrscheinlichkeit (alpha) von 5% annehmen.

Berechnungen: siehe Excel.

Nullhypothese H_0 : Die durchschnittlichen Lohneinkommen von Personen in Single-Haushalten sind niedriger (oder gleich) wie die durchschnittlichen Lohneinkommen in Mehrpersonenhaushalten:
 $income_{single} \leq income_{Mehrpersonenhaushalte}$

Alternativhypothese H_1 : Die durchschnittlichen Lohneinkommen von Personen in Single-Haushalten sind höher als die durchschnittlichen Lohneinkommen in Mehrpersonenhaushalten: $income_{single} > income_{Mehrpersonenhaushalte}$

Aufgabe 4.3:

Wählen Sie ein anderes metrisches Merkmal, das Ihnen interessant erscheint, und führen Sie einen einseitigen und einen zweiseitigen Zweistichproben-t-Test durch. Formulieren Sie dabei Ihr Erkenntnisinteresse, sowie Null- und Alternativhypothese!

Aufgabe 4.4:

Im Deutschen Sozio-oekonomischen Panel (GSOEP) geben 2,124 Personen an, eine politische Partei zu unterstützen (Merkmal „pia“). Von diesen 2,124 Personen geben 2,101 Personen auch Auskunft

darüber, welche Partei(en) sie unterstützen (Merkmal „pib“). (Es gibt die – etwas kuriose – Merkmalsausprägung „Several parties“, daher habe ich bei Partei das „en“ in Klammer dazu gesetzt.) Untersuchen Sie die Parteienpräferenzen nach Geschlecht („sex“)!

a) Erstellen Sie eine Kreuztabelle mit absoluten und relativen Häufigkeiten!

Berechnungen: siehe Excel.

b) Es gibt die These, dass Frauen in einem größeren Ausmaß „linke“ Parteien unterstützen als Männer. Finden Sie dafür Evidenz in den Daten? Berechnen Sie die entsprechenden bedingten relativen Häufigkeiten, um diese These zu untersuchen! Finden Sie andere Muster, die Ihnen interessant erscheinen?

Insgesamt gibt es dafür relativ wenig Evidenz: Zwar unterstützen etwas mehr Frauen als Männer die SPD, aber gleiches gilt für die CDU (wobei der Unterschied bei beiden Parteien mit weniger als 1 Prozentpunkt – PP sehr gering ausfällt). Frauen unterstützen die Grünen deutlich stärker als Männer (+4.7 PP), allerdings die (im politischen Spektrum weiter links angesiedelte Partei) „Linke“ weniger stark (-2.3 PP). Sehr weit rechts stehende Parteien (DVU, Rep., NPD) werden von Männern 3 Mal so häufig unterstützt wie von Frauen. Insgesamt zeigt sich, dass Männer stärker als Frauen (rechts- oder links-) extremere Parteien unterstützen.

c) Berechnen Sie eine Kreuztabelle mit den erwarteten absoluten Häufigkeiten.

Berechnungen: siehe Excel.

d) Berechnen Sie das Cramersche Assoziationsmaß V und interpretieren Sie die Stärke des Zusammenhangs zwischen Geschlecht und Parteienpräferenzen!

Berechnungen: siehe Excel.

Interpretation: Das Cramersche Assoziationsmaß V beträgt 0.118. Der Zusammenhang zwischen Geschlecht und Parteienpräferenzen ist daher nur schwach ausgeprägt.

e) Führen Sie einen Chi-Quadrat-Test auf Unabhängigkeit durch! Wie lautet die Null- und wie die Alternativhypothese? Wählen Sie eine passende Irrtumswahrscheinlichkeit (alpha). Kann die Nullhypothese verworfen werden?

Berechnungen: siehe Excel.

Nullhypothese H_0 : Es gibt keinen Unterschied zwischen Männern und Frauen in der Unterstützung politischer Parteien.

Alternativhypothese H_1 : Es gibt einen Unterschied zwischen Männern und Frauen in der Unterstützung politischer Parteien.

Aufgabe 4.5:

Wir sind am Zusammenhang zwischen der Wohnungsgröße und der Miete interessiert, da die Größe einer Wohnung maßgeblich die Miete beeinflussen sollte. Die Miete exklusive Heizkosten („rent“) wird in US-Dollar und die Wohnungsgröße („size“) in square-feet gemessen (warum auch immer). Informationen zur Miethöhe liegen bei 2,362 Personen vor. (Bei Eigentümern wird keine Miete ausgewiesen.)

a) Erstellen Sie ein Streudiagramm, um einen ersten Eindruck des Zusammenhangs zwischen den beiden Merkmalen zu bekommen. Machen Sie ein zweites Streudiagramm mit einer kleineren (Zufalls-) Stichprobe, um die Aussagekraft der Abbildung zu erhöhen.

Berechnungen: siehe Excel.

b) Berechnen Sie die Kovarianz und den Korrelationskoeffizienten zwischen den beiden Variablen. Interpretieren Sie den Korrelationskoeffizienten! Warum ist der Korrelationskoeffizient so deutlich kleiner als 1? Welche anderen Einflussfaktoren könnten hier eine Rolle spielen?

Berechnungen: siehe Excel.

Interpretation: Der Korrelationskoeffizient ist kleiner als 1, weil andere Eigenschaften, die die Höhe der Miete beeinflussen (Lage, Baujahr, Ausstattung) unberücksichtigt bleiben.

Aufgabe 4.6:

Wählen Sie 5 metrisch Variablen aus, zwischen denen Sie einen Zusammenhang vermuten, und erstellen Sie eine Korrelationsmatrix! Finden Sie Ihre Erwartungen bestätigt? Was hat Sie überrascht?