Konfidenzintervalle

Mittels Hypothesen-Tests

DBW

Gesucht ist der relative Anteil p einer Gruppe in einer Population. Wir machen eine Stichprobe der Größe n und finden dadrinnen H Exemplare aus dieser Gruppe, somit würden wir auf jeden Fall mal vermuten $p \approx h = H/n$. Wir wollen ein Intervall um diesen Wert von h bestimmen. Dazu fixieren wir zuerst eine Irrtumswahrscheinlichkeit α und eine Sicherheit $\gamma = 1 - \alpha$. Typisch wäre $\alpha = 0,05$ und $\gamma = 0,95$.

Für jede Zahl $x \in [0;1]$ betrachten wir die Hypothese $H_0(x): \quad p = x$. Also, für jede Zahl x zwischen 0 und 1 haben wir eine Nullhypothese $H_0(x)$. Dazu können wir dann ein Intervall $I_x = [x - \Delta_x; x + \Delta_x]$ mit

$$\Delta_x = z\sqrt{\frac{x(1-x)}{n}}$$

bestimmen. Da wir h schon haben, können wir also die Hypothese $H_0(x)$ schon mal testen: Falls $h \in I_x$, werden wir $H_0(x)$ nicht ablehnen, falls $h \notin I_x$, werden wir $H_0(x)$ ablehnen.

Wir haben somit eine Methode, um für jede Zahl $x \in [0;1]$ zu bestimmen, ob $H_0(x)$ verworfen wird oder nicht.

Das Konfidenzintervall für p nach der Messung von h ist genau die Menge aller x, für welche $H_0(x)$ nicht verworfen wird.

Wir könnten wir uns das mal etwas anders, aber doch praktisch, so vorstellen: Wir lassen einen Computer folgendes Programm durchlaufen, in dem bei jeder der folgenden Zahlen

0; 0,01; 0,02; 0,03; $\ldots$; 0,99; 1

testen ob h in ihrem Konfidenzintervall liegt, oder nicht. Falls nein, markiert er den Punkt rot, falls ja, markiert er den Punkt grün. Das Konfidenzintervall besteht dann aus den grünen Punkten (und allen Punkten zwischen zwei grünen Punkten). Wenn wir eine größere Genauigkeit wollen, dann können wir auch Schrittgröße 0,001 wählen – es dauert dann halt etwas länger, bis wir das Konfidenzintervall sehen.

Unterschiede?

Wenn du dir die Erklärung vom Konfidenzintervall mal im Buch nachschaust, siehst du, dass beide Methoden äquivalent sind. Es gibt aber feine Unterschiede:

(1) Die Nullhypothesen werden im Buch nicht mit h, sondern mit H gemacht.

(2) In der Formel für I_x benutzen wir, dass wir schon mal annehmen, dass $\sqrt{\frac{p(1-p)}{n}} \approx \sqrt{\frac{h(1-h)}{n}}$. Das ist in Ordnung, da n im Nenner steht, und somit der ganze Ausdruck unter dem Wurzelzeichen klein ist (wir wählen sinnvollerweise n groß).

Durch diese zwei Feinheiten kann es sein, dass beide Methoden nicht genau dasselbe Intervall geben. Erstens werden überall Zahlen gerundet; die z-Werte sind immer bis auf zwei oder drei Dezimalzahlen gerundet, und zweitens mit der in (2) erwähnten Annäherung $\Delta_x \approx \Delta_h$ wird ein weiteres Mal gerundet. Das ist eigentlich kein Problem, aber da wir bei Hypothesen-Testen das meistens mit H und nicht mit h machen, multiplizieren wir sozusagen alles wieder mit n, also mit einer großen Zahl. Somit werden kleine Unterschiede wieder hinaufmultipliziert.

Ein Beispiel

Wir wollen wissen, wie viele Personen in der Stadt Wien mit den öffentlichen Verkehrsmitteln zufrieden sind. Wir machen eine Befragung unter 200 Menschen, und finden 120 Personen, die mit den öffentlichen Verkehrsmitteln zufrieden sind. Wir vermuten also $p \approx 0,6$. Wir werden als Sicherheitsniveau mal $\gamma = 0,95$ nehmen, somit sind unsere z-Werte stets $1,96$.

Zuerst die klassische Methode, so wie es halt nach dem Buch geht: $h = 0,6$ und somit $\frac{h(1-h)}{n} = 0,6 \cdot 0,4/200 = 0,0012$. Daher $\Delta = 1,96 \cdot \sqrt{0,0012} \approx 0,068$. Somit ist $[53,2\%; 66,8\%]$ ein gutes Konfidenzintervall.

Betrachten wir jetzt einige Werte von x und ihre Hypothesen $H_0(x)$, und zwar:

$x = 0,5$; $x = 0,52$; $x = 0,53$; $x = 0,54$; $x = 0,6$; $x = 0,64$; $x = 0,64$ und $x = 0,67$.

Mittels kleiner Berechnung finden wir, dass $nx(1-x) = 200 \cdot x \cdot (1-x)$ für all diese Werte größer als 9 ist, also werden wir die Nullhypothesen mit der Normalverteilung behandeln. Wir werden dabei die Nullhypothesen mit dem Wert $H = 120$ testen, statt alles auf Anteile herunterzurechnen (also nicht mit $0,6$ rechnen). Wir bezeichnen EH den Erwartungswert der zufriedenen Personen; diese hängt von der Hypothese ab, und zwar $EH = 200 \cdot x$. Da wir mit dem Wert 120 und nicht mit dem Anteil $0,6$ testen, haben wir $\Delta_x = z\sqrt{nx(1-x)}$.

(a) $\Delta_{0,5} = z\sqrt{200 \cdot 0,5 \cdot 0,5} \approx 13,86$ und der Erwartungswert beträgt jetzt $EH = 200 \cdot 0,5$; somit wird $H_0(0,5)$ abgelehnt, da 120 mehr als $100 + 13,86$ ist.

(b) $\Delta_{0,52} = 13,85$ und EH ist jetzt $0,52 \cdot 200 = 104$ und weil $104 + 13,85$ noch immer kleiner als 120 ist, wird auch der Wert $0,52$ abgelehnt, also $H_0(0,52)$ wird abgelehnt.

(c) $\Delta_{0,53} = 13,83$ und $EH = 0,53 \cdot 200 = 106$ und jetzt ist 120 noch immer zu groß, aber es ist knapp dran! Trotzdem wird $H_0(0,53)$ abgelehnt. Es wird also kritisch sein - eine spannende Rundungssache - ob wir 0,53 im Konfidenzintervall aufnehmen oder nicht.

(d) $\Delta_{0,54} = 13,81$ und $EH = 0,54 \cdot 200 = 108$ und jetzt ist $120 \in [108 - 13,81; 108 + 13,81]$. Die Nullhypothese $H_0(0,54)$ wird somit nicht abgelehnt. Wir sollten $0,54$ also in unserem Konfidenzintervall aufnehmen.

(d) $\Delta_{0,6} = 13,58$ und $EH = 120$, sodass der Wert 120 wirklich im Intervall $[120 - 13,58; 120 + 13,58]$ liegt. Freilich wird der Wert 0,6 auch im Konfidenzintervall sein.

(e) $\Delta_{0,64} = 13,30$ und $EH = 128$ und da $128 - 13,30 = 114,70 < 120 < 128 + 13,30 = 141,30$ werden wir $H_0(0,64)$ nicht ablehnen. Somit sollte $0,64$ im Konfidenzintervall sein.

(f) $\Delta_{0,66} = 13,13$ und $EH = 132$, sodass 120 noch im Intervall $[132 - 13,13; 132 + 13,13] = [118,87; 145,13]$ liegt un die Hypthese $H_0(0,66)$ wird somit nicht abgelehnt.

(g) $\Delta_{0,67} = 13,04$ und $EH = 134$, und wir sehen $120 < 134 - 13,04$ und daher wird $H_0(0,67)$ abgelehnt.

Zusammenfassend: Wir sehen, dass wir als Konfidenzintervall nehmen würden $[54\%; 66\%]$. Wir haben hierbei natürlich nur auf ganze Prozente geschaut. Würden wir die Werte $x = 0,531$, $x = 0,532$ und so weiter auch noch nehmen, würden wir ein besseres Zusammenpassen mit dem obigen Konfidenzintervall $[53,2; 66,8]$ sehen. Zu beachten ist aber Folgendes: Die erste ganze Zahl größer als 53,2 ist 54 und die erste ganze Zahl kleiner als 66,8 ist 66. Durch die etwas gröbere Behandlug haben wir also die Grenzen des Konfidenzintervalls „nach innen“ gerundet. Wenn man sich kurz überlegt, wie Hypothesen Testen funktioniert, ist das auch sehr logisch.

Übung: Wiederhole obiges Prozedere für die Zahlen $x = 0,531$, $x = 0,532$, bis $x = 0,539$ und auch für $x = 0,661$, $x = 0,662$ bis $x = 0,669$ und bestimme auf diese Weise ein etwas genaueres Konfidenzintervall für den gesuchten Anteil der Zufriedenen.

Übung: Im obigen Beispiel siehst du, dass die Werte von Δ_x nicht so variieren. Kannst du erklären warum? Hinweis: $\Delta_{x_1} - \Delta_{x_2} \approx \frac{\mathrm{d}}{\mathrm{d}x}\Delta_x \cdot (x_1 - x_2)$ und wenn wir $z\sqrt{nx(1-x)}$ nach x differenzieren finden wir

$$\frac{\mathrm{d}}{\mathrm{d}x}\Delta_x = \frac{zn(1-2x)}{2\sqrt{nx(1-x)}} .$$