

Zu Abschnitt 3.4.2: Statistische Tests auf Gruppenebene

Statistische Tests für Mittelwerte: Realnormvergleiche

Beispiel 2: Ist die Klasse 7d besser als der Bundesdurchschnitt?

Während also der Bundesdurchschnitt im PISA-Fähigkeitsbereich „Problemlösen“ bei $\bar{x} = 513$ liegt ($\hat{\sigma}_x = 95$), hat die Klasse 7d des Heinrich-Heine-Gymnasiums nun einen Wert von $\bar{x} = 520$ erreicht. Nun stellt sich die Frage: Kann diese Klasse signifikant besser Probleme lösen als der Bundesdurchschnitt?

Diese Frage lässt sich mit einem „Ein-Gruppen-*t*-Test“ beantworten (s. Eid et al., 2013; Abschn. 10.1). Der Unterschied zum Ein-Gruppen-*z*-Test ist, dass dort die Populationsstandardabweichung bekannt sein muss, während diese beim Ein-Gruppen-*t*-Test aus den Daten geschätzt werden kann. In unserem Beispiel ist die Klasse 7d die Stichprobe; die geschätzte Populationsstandardabweichung beträgt hier $\hat{\sigma}_x = 95$, demnach beträgt die geschätzte Populationsvarianz $\hat{\sigma}_x^2 = 9025$. Auch hier benötigen wir *keine Originaldaten*; der Test kann sehr leicht „von Hand“ mit Hilfe der entsprechenden Formel berechnet werden. Der Wert der Prüfgröße berechnet sich wie folgt:

$$t = \frac{\bar{x} - \mu}{\hat{\sigma}_{\bar{x}}}$$

Dabei sind

- $\bar{x}$ der Mittelwert, der gegen einen Realwert (hier: den PISA-Bundesdurchschnitt) getestet werden soll (in unserem Beispiel $\bar{x} = 520$),
- μ ist nun der empirische Realwert, d.h. der PISA-Bundesdurchschnitt ($\mu = 513$) und
- $\hat{\sigma}_{\bar{x}}$ die aus der Stichprobe geschätzte Standardabweichung der Stichprobenkennwerteverteilung von Mittelwerten aus Stichproben der Größe n

Den Wert für $\hat{\sigma}_{\bar{x}}$ müssen wir zunächst noch „von Hand“ ermitteln. Die Formel lautet:

$$\hat{\sigma}_{\bar{x}} = \sqrt{\frac{\hat{\sigma}_x^2}{n}}$$

Dabei sind

- $\hat{\sigma}_x^2$ die aus der Stichprobe geschätzte Populationsvarianz (hier: $\hat{\sigma}_x^2 = 9025$),
- n = die Stichprobengröße, d.h. die Anzahl Schüler in der Klasse 7d. Sagen wir der Einfachheit halber, diese sei $n = 25$.

Damit berechnet sich $\sigma_{\bar{x}}$ zu

$$\hat{\sigma}_{\bar{x}} = \sqrt{\frac{\hat{\sigma}_x^2}{n}} = \sqrt{\frac{9025}{25}} = 19$$

Damit berechnet sich der Wert der Prüfgröße T zu:

$$t = \frac{\bar{x} - \mu}{\hat{\sigma}_{\bar{x}}} = \frac{520 - 513}{19} = \frac{7}{19} \approx 0,368$$

Die Quantile unter der t -Verteilung hängen von der Anzahl der Freiheitsgrade (df) ab. Diese berechnen sich hier zu

$$df = n - 1$$

Für unser Datenbeispiel mit $n = 25$ Schülern benötigen wir eine t -Verteilung mit $df = 24$ Freiheitsgraden. Die Wahrscheinlichkeit für den ermittelten (oder jeden größeren) t -Wert beträgt damit etwa $p = 0,36$. Auf dem 5%-Niveau ist die Abweichung vom Realwert demnach nicht signifikant. Man kann also nicht sagen, dass die Klasse 7d im PISA-Test signifikant besser ist als der Bundesdurchschnitt.