

Antworten zu Kapitel 12: Unterschiede zwischen zwei abhängigen Stichproben

(1) Wie lauten die Voraussetzungen für die Anwendung des t -Tests für abhängige Stichproben?

Die Anwendung des t -Tests für abhängige Stichproben ist an folgende Voraussetzungen geknüpft: (1) Die Messwerte innerhalb der Teilpopulationen müssen voneinander unabhängig sein. Anders ausgedrückt: Zwischen den Messwertpaaren darf es keine systematischen Einflüsse oder Abhängigkeiten geben. (2) Die Differenzvariable D muss in der Population normalverteilt sein. Die Normalverteilungsannahme impliziert, dass stetige Variablen vorliegen müssen.

(2) Was bedeutet die Aussage, es dürfe beim t -Test für abhängige Stichproben zwar »zwischen den Stichproben« Abhängigkeiten geben, nicht aber »innerhalb der Stichproben«? Erläutern Sie diese Aussage an einem Beispiel.

In zwei Schulklassen werden zwei unterschiedliche Trainings zur Verbesserung des Leseverständnisses durchgeführt. Wenn sich nun die Schüler innerhalb der Klassen über das Training austauschen und sich dadurch gegenseitig beeinflussen, indem sie einander beispielsweise motivieren oder demotivieren, dann wirkt sich das auf ihre Mitarbeit bei den Übungen und damit auch auf die Wirkung des Trainings aus. Die Abhängigkeit, die so zwischen den Messwerten innerhalb der beiden Stichproben entsteht, ist unsystematisch und lässt sich nicht modellieren, da man nicht weiß, wie sich die Schüler gegenseitig beeinflussen haben.

(3) Wieso hat der t -Test für abhängige Stichproben unter Umständen eine größere Teststärke als der t -Test für unabhängige Stichproben?

Betrachtet man die Definition der Effektgröße δ'' , so lässt sich leicht nachvollziehen, weshalb der t -Test für abhängige Stichproben eine größere Teststärke hat als der t -Test für unabhängige Stichproben. Die Effektgröße δ'' ist definiert als der Quotient aus dem Populationsmittelwert der Differenzen μ_D und der Standardabweichung der Differenzvariablen D . Diese Standardabweichung wird – bei ansonsten gleichen Stichprobenvarianzen – umso geringer, je größer die Kovarianz σ_{12} ist. Daraus folgt, dass bei gleichem Differenzmittelwert μ_D die Effektgröße δ'' umso größer wird, je größer σ_{12} ist. Je ähnlicher sich die Messwertpaare sind, desto kleiner ist die Standardabweichung der Differenzvariablen, und desto größer ist die Effektgröße. Da die Effektgröße mit der Teststärke in einem direkten Zusammenhang steht, hat der t -Test für abhängige Stichproben umso mehr Power, je größer die Kovarianz ist.

(4) Beschreiben Sie die Logik, die dem Vorzeichentest zugrunde liegt.

Beim Vorzeichentest wird nur das Vorzeichen der Differenz $x_{m1} - x_{m2}$ betrachtet. Man bestimmt für jedes Messwertpaar, welcher der beiden Messwerte größer ist. Für den Fall, dass die Mediane der beiden abhängigen Stichproben nicht voneinander abweichen ($\eta_1 = \eta_2$), erwartet man genauso viele Differenzen mit positivem Vorzeichen wie solche mit negativem Vorzeichen. Wenn die Mediane der beiden abhängigen Stichproben voneinander abweichen, erwartet man hingegen mehr Differenzen mit positivem als mit negativem Vorzeichen oder umgekehrt.

(5) Welcher Test ist teststärker: der Vorzeichentest oder der Wilcoxon-Vorzeichen-Rangtest?

Der Vorzeichentest ist teststärker, da er eine kleinere Stichprobe benötigt, um einen gleich großen Effekt aufzudecken als der Wilcoxon-Vorzeichen-Rangtest.

(6) Beschreiben Sie die Logik, die dem McNemar-Test zugrunde liegt. Begründen Sie, wieso die Prüfgröße beim McNemar-Test exakt binomialverteilt ist.

Die Nullhypothese des McNemar-Tests besagt, dass die Kategorienwechsel über die Bedingungen bzw. die Messzeitpunkte hinweg zufällig sind. Im Falle von zwei Kategorien und zwei Bedingungen (bzw. Messzeitpunkten) müsste also die Wahrscheinlichkeit, von der einen in die andere Bedingung zu wechseln, für beide Richtungen gleich sein und die Vierfeldertafel auf Populationsebene sollte eine symmetrische Struktur aufweisen. Je ungleicher die Veränderungen in die eine oder andere Richtung sind, desto eher spricht dies gegen die Nullhypothese. Der McNemar-Test überprüft dann, ob die Wahrscheinlichkeit, die beobachtete Zelhäufigkeit oder eine noch größere Abweichung von der Erwartung unter der Nullhypothese zu finden, kleiner als das vorgegebene Signifikanzniveau α ist. Die Prüfgröße beim McNemar-Test ist binomialverteilt. Die beiden Parameter, die die Binomialverteilung unter der Nullhypothese beschreiben, sind s , also die Anzahl der Personen, die gewechselt haben, und die Wahrscheinlichkeit π , die unter der Nullhypothese $\pi_0 = 0,50$ beträgt.