

Antworten zu Kapitel 10: Tests auf Abweichung von einem fixen Wert

(1) Wie lauten die Voraussetzungen für die Anwendung des Einstichproben-Gauß-Tests und des Einstichproben- t -Tests?

Der Einstichproben-Gauß-Test setzt voraus, dass das Merkmal X in der Population normalverteilt ist und der Erwartungswert $E(X) = \mu$ und die Standardabweichung σ_X in der Population bekannt sind. Der Einstichproben- t -Test setzt ebenfalls voraus, dass Merkmal in der Population normalverteilt und der Erwartungswert $E(X) = \mu$ bekannt ist, die Standardabweichung σ_X in der Population muss hier jedoch nicht bekannt sein; sie muss lediglich aus den Daten geschätzt werden können.

(2) Worin besteht der Unterschied zwischen dem Vorzeichentest und dem Wilcoxon-Vorzeichen-Rangtest?

Der Wilcoxon-Vorzeichen-Rangtest nutzt mehr Informationen aus den Daten als der Vorzeichentest, er hat aber auch strengere Voraussetzungen als der Vorzeichentest. Während der Vorzeichentest lediglich voraussetzt, dass die Variable X mindestens ordinalskaliert ist, setzt der Wilcoxon-Vorzeichen-Rangtest voraus, dass die Variable X metrisches Skalenniveau aufweist und in der Population stetig und symmetrisch verteilt ist.

(3) Wie lautet die Prüfgröße für den Wilcoxon-Vorzeichen-Rangtest, und was bedeutet sie?

Die Prüfgröße für den Wilcoxon-Vorzeichen-Rangtest lautet W^+ . Sie wird aus der Summe der Rangplätze aller positiven Differenzen zwischen x_m und η_0 berechnet und mit dem Erwartungswert der Prüfgröße W^+ bei Geltung der Nullhypothese verglichen.

(4) Wie ist die Effektgröße v beim χ^2 -Test auf Abweichung einer Stichprobenvarianz von einem fixen Wert definiert? Handelt es sich hier um eine standardisierte oder eine unstandardisierte Größe?

Die Effektgröße v bildet der Quotient aus der Stichprobenvarianz und dem fixen Wert für die Populationsvarianz unter der Nullhypothese: $v = \sigma_X^2 / \sigma_0^2$. Je mehr v von 1 abweicht, desto größer ist der Effekt. Da v über verschiedene Untersuchungen und Metriken hinweg vergleichbar ist, handelt es sich – wie bei Cohen's δ – um eine standardisierte Effektgröße.

(5) Erläutern Sie, wieso der χ^2 -Test auf Abweichung einer Stichprobenvarianz von einem fixen Wert $df = n - 1$ Freiheitsgrade hat.

Die Anzahl der Freiheitsgrade entspricht der Anzahl der quadrierten standardnormalverteilten Variablen, die über Personen von $m = 1$ bis n hinweg aufsummiert werden. Diese Summe ist jedoch nicht beliebig; sie liegt aufgrund der Standardisierung jedes einzelnen Wertes fest und muss immer $n - 1$ betragen. Daher gibt es entsprechend auch nur $n - 1$ Summanden, die frei variieren können.

(6) Erläutern Sie, was beim Binomialtest auf Abweichung einer Wahrscheinlichkeit von einer theoretischen (fixen) Wahrscheinlichkeit der kritische Wert s_{krit} bedeutet. Konkret: Was bedeutet ein Wert von $s_{\text{krit}} = 40$ bei $n = 60$?

Ein kritischer Wert s_{krit} von 40 bei 60 untersuchten Personen bedeutet, dass von den 60 untersuchten Personen 40 eine Merkmalsausprägung von $x = 1$ hätten aufweisen müssen, um die Nullhypothese verwerfen zu können.

- (7) Was bedeutet beim Binomialtest auf Abweichung einer Wahrscheinlichkeit von einer theoretischen (fixen) Wahrscheinlichkeit ein Populationseffekt von $\gamma = 0,5$ inhaltlich? Wie ist die Effektgröße definiert? Ist sie standardisiert? Was bedeuten positive und negative Werte?**

Wenn man davon ausgeht, dass $\pi_0 = 0,5$ ist, bedeutet ein Wert von $\gamma = +0,5$, dass die Wahrscheinlichkeit für $x = 1$ in der Population gleich 1 ist, denn $\gamma = \pi - \pi_0$. Positive Werte bedeuten, dass die Wahrscheinlichkeit für $x = 1$ in der Population höher ist als unter der Nullhypothese erwartet wird, und negative Werte bedeuten, dass die Wahrscheinlichkeit für $x = 1$ in der Population niedriger ist als unter der Nullhypothese erwartet wird. Die Effektgröße ist standardisiert, da sie über unterschiedliche Untersuchungen und Variablen hinweg vergleichbar ist.

- (8) Erläutern Sie, wieso der χ^2 -Test für ein kategoriales Merkmal $df = k - 1$ Freiheitsgrade hat.**

Die Häufigkeitsverteilungen in den $k - 1$ Kategorien können frei variieren, die absolute Häufigkeit in der letzten Kategorie ist allerdings festgelegt, da sich die absoluten Häufigkeiten in allen Kategorien zu n aufaddieren.

- (9) Erläutern Sie, was beim KS-Anpassungstest die Prüfgröße D_{\max} bedeutet. Spricht ein hoher Wert für D_{\max} eher für oder gegen die Annahme, das Merkmal X sei in der Population normalverteilt?**

Die Prüfgröße D_{\max} des KS-Anpassungstests ist definiert als die größte absolute Abweichung zwischen der empirischen Verteilungsfunktion $F(x)$ und der Verteilungsfunktion, die man unter der Nullhypothese erwarten würde. Ein hoher Wert für D_{\max} spricht eher gegen die Annahme, dass das Merkmal X in der Population normalverteilt ist. Je höher der Wert für D_{\max} ist, desto stärker weicht die Verteilungsfunktion $F(x)$ des Merkmals von der erwarteten Verteilung (z. B. der Normalverteilung) ab.

- (10) Worin besteht der Unterschied zwischen dem KS-Anpassungstest und dem Lilliefors-Test?**

Der Lilliefors-Test kann im Gegensatz zum KS-Anpassungstest auch dann verwendet werden, wenn Populationsmittelwert und Populationsstandardabweichung unbekannt sind. Allerdings ist der Lilliefors-Test strenger als der KS-Anpassungstest. Abweichungen von der Normalverteilung werden beim Lilliefors-Test eher signifikant.