Der T-Test für Mittelwertsunterschiede bei “verbundenen Stichproben” / “gepaarten Beobachtungen”

Anwendungsbereich:

1. Wiederholungsmessungen auf Personenebene: Paneluntersuchungen, Begleit-/Evaluationsforschung, Experimente (Vorher- /Nachmessung)

2. Paarvergleich von unterschiedlichen Analyseeinheiten bei einer identischer Erhebungseinheit: Bei einem Ehepaar die separat erhobenen Einstellungen von Ehefrau und Ehemann zum “autoritären Erziehungsstil”

Ansatzpunkt: Differenz der Meßwerte beider Gruppen oder Erhebungszeitpunkte

Testart: Zweiseitig, da über die Meßwertdifferenz kein Apriori-Wissen vorhanden ist.

Nullhpothese: Die Differenz der Meßwerte beider Gruppen/Meßzeitpunkte ist Null. $\mathbf{H_0 : \mu_{Differenz} = 0}$

Alternativhypothese: Die Differenz der Meßwerte beider Gruppen/Meßzeitpunkte ist signifikant von Null verschieden. $\mathbf{H_A : \mu_{Differenz} \neq 0}$

Voraussetzungen:

Fall 1: **Anzahl der Meßwertpaare ≥ 30:**

Hierfür gilt der zentrale Grenzwertsatz, der besagt, daß bei hinreichend großem Stichprobenumfang die Verteilung der Mittelwerte der Meßwertdifferenzen in eine Normalverteilung übergehen. (S. Bortz 1989[3], S.172)

Fall 2: **Anzahl der Meßwertpaare < 30**

Der T-Test ist nur anwendbar, wenn sich die Meßwertdifferenzen annäherungsweise normalverteilen.

Berechnung des empirischen T−Wertes für verbundene Stichproben:

$$T_{pr} = \frac{\bar{x}_{Differenzen} - \mu_0}{\hat{\sigma}_{\bar{x}_{Differenzen}}} \quad , \textit{ wobei } \; H_0\!: \; \mu_0 = 0$$

$$\bar{x}_{Differenzen} = \frac{\sum_{i=1}^{n} D_i}{n}$$

$$\hat{\sigma}_{\bar{x}_{Differenzen}} = \frac{\hat{\sigma}_{Differenzen}}{\sqrt{n}}$$

$$\hat{\sigma}_{Differenzen} = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{x}_{Differenzen})^2}{n - 1}} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}}{n - 1}}$$

Anwendungsbeispiel:

In einer Fabrik wird ein neues Herstellungsverfahren eingeführt. Führt dieses neue Verfahren zu einer statistisch-bedeutsamen Erhöhung der Arbeitsproduktivität ? Um dies zu überprüfen, wurde bei 10 zufällig ausgewählten Arbeitern die von ihnen pro Zeiteinheit gefertigte Stückzahl vor und nach der Einfühung des neuen Produktionsverfahrens erhoben. Mit Hilfe des T-Tests für verbundene Stichproben können wir nun überprüfen, ob dieses neue Verfahren zu einer signifikanten Erhöhung der Produktivität geführt hat. Hierbei betrachten wir die Meßdifferenzen nicht auf der Ebene der Teilstichproben, sondern auf der Ebene der Erhebungseinheiten. In unserem Fall auf der Ebene des einzelnen Arbeiters.

Hierbei soll mit einer Irrtumswahrscheinlichkeit von 1 % über die betriebsweite Einführung des neuen Fertigungsverfahrens entschieden werden.

Tab.1: Analyse der Steigerung der Arbeitsproduktivität

ARBEITER	ALTMETH	NEUMETH	DIFFMETH	DIFF2
01	21	27	-06	36
02	18	31	-13	169
03	20	25	-05	25
04	17	29	-12	144
05	23	26	-03	9
06	25	33	-08	64
07	21	26	-05	25
08	23	30	-07	49
09	19	25	-06	36
10	24	32	-08	64
	Summe =		-73	621

Berechnung der empirischen T–Prüfgröße:

$$\bar{x}_{Differenzen} = \frac{-73{,}0}{10} = -7{,}3$$

$$\hat{\sigma}_{Differenzen} = \sqrt{\frac{621 - \frac{-73^2}{10}}{10-1}} = \sqrt{\frac{621 - 532{,}9}{9}} = \sqrt{9{,}79} = 3{,}13$$

$$\hat{\sigma}_{\bar{x}_{Differenzen}} = \frac{3{,}13}{\sqrt{10}} = \frac{3{,}13}{3{,}16} = 0{,}99$$

$$T_{Pr} = \frac{-7{,}3}{0{,}99} = -7{,}37$$

$$T_{krit\left(\frac{\alpha}{2}=0{,}5;\ (n-1)=9\ F.G.\right)} = 3{,}25$$

Da $|T_{Pr}| > T_{Krit}$ muß H_0 verworfen werden. D.h., das neue Verfahren steigert in signifikanten Ausmaß die Produktivität. Hierbei nehmen wir eine Irrtumswahrscheinlichkeit von 1 % in Kauf.

Allgemeine Bemerkungen zur Verwendung des T-Tests für unabhängige Beobachtungen:

“Zusammenfassend kann man davon ausgehen, daß der t-Test robust ist, wenn gleich große Stichproben aus ähnlichen, möglichst eingipflig-symmetrischen verteilten Grundgesamtheiten verglichen werden. Sind die Stichprobenumfänge deutlich unterschieden, wird die Präzision des t-Testes nicht beeinträchtigt, solange die Varianzen gleich sind. Sind aber weder die Stichprobenumfänge noch die Varianzen gleich, ist mit einem erheblich höheren Prozentsatz von Fehlentscheidungen zu rechnen.”

(Jürgen Bortz: Statistik für Sozialwissenschaftler.
Berlin: Springer, 1989[3] , S. 171f.)