t-Tests für Mittelwertsunterschiede
1. Unabhängige Stichproben

Frage-stellung	**Wirkt X auf Y?** **Ist Y im Mittel unterschiedlich in der Gruppe in der X=1 und der Gruppe in der X=0?**		
UV und AV	**UV/Prädiktorvariable X: Dichotom (zwei kategorial verschiedene Ausprägungen)** **AV/Kriteriumsvariable Y: kontinuierlich, auf Intervallskalenniveau gemessen**		
Beispiele Inhaltliche Hypothesen	1. In einer Studie wird die Vorhersagekraft einer Variable X mit zwei Ausprägungen auf eine kontinuierliche, auf Intervallskalenniveau gemessene Kriteriumsvariable Y untersucht. Die Hypothese ist, dass Y bei X=1 im Mittel größer/kleiner/ungleich/gleich als bei X=0 ist 2. In einem Experiment wird die UV X between subjects manipuliert. Die Versuchspersonen werden den resultierenden Experimental- und Kontrollbedingungen randomisiert zugeteilt. Die Hypothese ist, dass die AV Y bei Ausprägung X=1 der UV im Mittel größer/kleiner/ungleich/gleich als bei Ausprägung X=0 ist		
Ableitung Vorhersage & Stat. Hypothesen	PH-gerichtet: Y ist im Mittel bei X=1 größer/ kleiner als bei X=0. PH-ungerichtet: Y ist in Mittel bei X=1 ungleich (H_1)/ gleich (H_0) wie bei X=0. ⇨ PV: Vermuteter mittlerer Unterschied in Y zwischen den Gruppen, die durch die Ausprägungen von X gebildet werden. ⇨ Vorhersagekonforme Statistische Hypothesen: Gerichtet: H_1: $\mu_1 - \mu_0 > 0$ H_0: $\mu_1 - \mu_0 \leq 0$ Ungerichtet: H_1: $\mu_1 - \mu_0 \neq 0$; H_0: $\mu_1 - \mu_0 = 0$		
Testplanung per Hand	Determinanten: Alpha, Beta, Effektgröße (Cohens d), N Unabhängige Stichproben gerichtete Testung n= $[(z_{1-\alpha}+z_{1-\beta})^2 * 2]/d^2$ Unabhängige Stichproben ungerichtete Testung n= $[(z_{1-\alpha/2}+z_{1-\beta})^2 * 2]/d^2$		z-Werte (positive Werte) =NORM.S.INV(0,95) = NORM.S.INV(0,975)
Relevante Kennwerte	Mittelwert Y	$M_Y = \frac{1}{n}\sum_{i=1}^{n} y_i$	=MITTELWERT(Bereich)
	Streuung innerhalb jeder Stichprobe (als Schätzer)	$s_Y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - M_Y)^2}$	=STABW.S(Bereich)
	Standardfehler des Mittelwerts	$SE_M = s_Y * \frac{1}{\sqrt{n}}$	=STABW.S(Bereich)/ WURZEL(ANZAHL(Bereich))
	Konfidenzintervall des Mittelwerts (meist 95%)	$t_{(df=n-1,\ Konfidenz)} * SE_M$ Untere Grenze: $M - t * SE_M$ Obere Grenze: $M + t * SE_M$	=T.INV.2S(alpha;n-1)*SE_M =KONFIDENZ.T(alpha;STABW.S(); Anzahl)
	Streuung innerhalb Bedingungen (d.h. innerhalb beider Stichproben)	$s_{iB} = \sqrt{\frac{(n_1-1)s_1^2+(n_0-1)s_0^2}{n_1+n_0-2}} = \sqrt{\frac{s_1^2+s_0^2}{2}}$ bei $n_1=n_0$	
	Standardfehler Mittelwertsdifferenz	$SE_{M_1-M_0} = \sqrt{\frac{(n_1-1)*s_1^2+(n_0-1)*s_0^2}{n_1+n_0-2} * \left(\frac{1}{n_1}+\frac{1}{n_0}\right)}$	

<table>
<tr>
<td>Test auf Varianz-gleichheit</td>
<td>F-Test
Empirischer F-Wert = Verhältnis der Varianzen
$F_{emp} = \frac{s_1{}^2}{s_0{}^2}$ wobei $s_1^2 > s_0^2$
Freiheitsgrade $df_1=n_1-1$, $df_2=n_0-1$
Vergleich mit F_{krit}
wenn $F_{emp} > F_{krit} \Rightarrow$ Varianzen ungleich</td>
<td>F-Test
=F.TEST(Bereich1; Bereich2)
Kritischer F-Wert (zweiseitig)
=F.INV.RE(Alpha/2;df1;df2)</td>
</tr>
<tr>
<td>Statistischer Test</td>
<td colspan="2">Unabhängige Stichproben
$$t_{emp} = \frac{M_1 - M_0}{SE_{M_1-M_0}} = \frac{M_1 - M_0}{\sqrt{\frac{(n_1-1)*s_1{}^2+(n_0-1)*s_0{}^2}{n_1+n_0-2}*\left(\frac{1}{n_1}+\frac{1}{n_0}\right)}}$$
$$bei\ gleichen\ Varianzen{:}\ t_{emp} = \frac{M_1 - M_0}{s_{iB}*\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_0}\right)}} = d*\frac{1}{\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_0}\right)}}$$
$$bei\ gleichem\ n{:}\ t_{emp} = \frac{M_1 - M_0}{s_{iB}*\sqrt{\frac{2}{n}}}$$
Unabhängige Stichproben und ungleiche Varianzen
$$t_{emp} = \frac{M_1 - M_0}{\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_0{}^2}{n_0}}}$$</td>
</tr>
<tr>
<td>Kritischer t-Wert p-Wert</td>
<td>Unabhängig:
$t_{krit(\alpha;df)}$= ? bei df= N – 2</td>
<td>Kritischer t-Wert (positive Werte)
=T.INV.2S(alpha;df) zweiseitig
=T.INV(1-alpha;df) einseitig

Empirischer p-Wert
=T.VERT.RE(t_{emp};df) einseitig
=T.VERT.2S(t_{emp};df) zweiseitig
=T.TEST(Daten1;Daten2;Seiten;Typ)
Seiten: 1=einseitig; 2= zweiseitig
Typ: 2 = unabhängige Stichproben mit gleicher Varianz
3 = unabhängige Stichproben mit ungleicher Varianz</td>
</tr>
<tr>
<td>Entscheidung SH</td>
<td colspan="2">Entscheidung nach folgendem Schema:
Bei $t_{emp} \geq t_{krit}$ ODER $p \leq$ Alpha aus Testplanung ⇨ Annahme H_1
Bei $t_{emp} < t_{krit}$ ODER $p >$ Alpha aus Testplanung ⇨ Annahme H_0</td>
</tr>
<tr>
<td>Effektstärke Konfidenzintervall</td>
<td>Cohens d
$d = (M_1 - M_0) / s_{iB}$</td>
<td>Standardfehler d
$$SE_d = \hat{\sigma}_d = \sqrt{\frac{n_1+n_0}{n_1*n_0}+\frac{d^2}{2*(n_1+n_0)}}$$
Konfidenzintervall (bei Annahme Normalverteilung)
• Untere Grenze: $d - 1{,}96*SE_d$
• Obere Grenze: $d + 1{,}96*SE_d$</td>
</tr>
</table>

Effektstärke Konventionen	d >= 0,2 ⇨ Kleiner Effekt d >= 0,5 ⇨ Mittlerer Effekt d >= 0,8 ⇨ Großer Effekt d >= 1,2 ⇨ Sehr großer Effekt	
Entscheidungen über PV	Entscheidung nach folgendem Schema: Bei abgeleiteter H_0 und angenommener H_0 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_0 und angenommener H_1 ⇨ PV nicht eingetreten, PH nicht bewährt Bei abgeleiteter H_1 und angenommener H_1 & $d_{emp} > d$ aus Testplanung bzw. Power ≥ 0,8 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_1 und angenommener H_1 & $d_{emp} < d$ aus Testplanung bzw. 0,6 ≤ Power < 0,8 ⇨ PV bedingt eingetreten, PH bedingt bewährt Bei abgeleiteter H_1 und angenommener H_0 ⇨ PV nicht eingetreten, PH nicht bewährt	

t-Tests für Mittelwertsunterschiede 2. Abhängige Stichproben	
Frage-stellung	**Wirkt X auf Y? Ist die mittlere Differenz von Y bei X=1 und X=0 innerhalb der einen Gruppe von Null verschieden?**
UV und AV	**UV/Prädiktor X: Dichotom (zwei kategorial verschiedene Ausprägungen) AV/Kriterium Y: kontinuierlich, auf Intervallskalenniveau gemessen**
Inhaltliche Hypothesen	1. In einer Studie wird die Vorhersagekraft einer Variable X, welche für jede Person zwei unterschiedliche Ausprägungen annimmt, auf eine kontinuierliche, auf Intervallskalenniveau gemessenen Variable Y untersucht. Es wird erwartet, dass die AV bei Ausprägung X=1 der UV im Mittel größer/kleiner/ungleich/gleich als bei Ausprägung X=0 ist 2. In einem Experiment wird die UV X within subjects manipuliert. Die Versuchspersonen nehmen an beiden Bedingungen in randomisierter Reihenfolge teil. Die Hypothese ist, dass die AV Y bei Ausprägung X=1 der UV im Mittel größer/kleiner/ungleich/gleich als bei Ausprägung X=0 ist
Ableitung Vorhersage & Stat. Hypothesen	PH-gerichtet: Y ist im Mittel bei X=1 größer/ kleiner als bei X=0. PH-ungerichtet: Y ist im Mittel bei X=1 ungleich (H_1)/ gleich (H_0) wie bei X=0. ⇨ PV: Vermuteter mittlerer Unterschied zwischen den Messungen t_1 und t_0, die durch die Ausprägungen von X gebildet werden. ⇨ Vorhersagekonforme Statistische Hypothesen: Gerichtet: H_1: $\mu_{Differenzt1\text{-}t0} > 0$ H_0: $\mu_{Differenzt1\text{-}t0} \leq 0$ Ungerichtet: H_1: $\mu_{Differenzt1\text{-}t0} \neq 0$; H_0: $\mu_{Differenzt1\text{-}t0} = 0$

Testplanung per Hand	Determinanten: Alpha, Beta, Effektgröße (Cohens d_{whm}), N Abhängige Stichproben gerichtete Testung $N = [(z_{1-\alpha}+z_{1-\beta})^2 * 2]/d_{whm}^2$ Abhängige Stichproben ungerichtete Testung $N = [(z_{1-\alpha/2}+z_{1-\beta})^2 * 2]/d_{whm}^2$		z-Werte (positive Werte) =NORM.S.INV(0,95) (einseitig) =NORM.S.INV(0,975) (zweiseitig)
Relevante Kennwerte	Differenz Messwerte	$diff_i = x_{i,t1} - x_{i,t0}$	(Differenzen sollten im Mittel positiv sein, für Excel ggf. t1 und t0 vertauschen)
	Mittelwert der Differenzen	$M_{diff} = \frac{1}{N}\sum_{i=1}^{n} diff_i$	=MITTELWERT(Bereich)
	Streuung der Differenzen (als Schätzer)	$s_{diff} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}(diff_i - M_{diff})^2}$	=STABW.S(Bereich)
	Standardfehler des Mittelwerts der Differenzen	$SE_{M_{diff}} = \frac{s_{diff}}{\sqrt{N}}$	=STABW.S(Bereich)/ WURZEL(ANZAHL(Bereich))
	Konfidenzintervall der mittleren Differenz (meist 95%)	$t_{(df=N-1,\ Konfidenz)} * SE_{Mdiff}$	=T.INV.2S(alpha;N-1)*SE_{Mdiff} =KONFIDENZ.T(alpha;STABW.S(); Anzahl)
	Korrelation der Messwerte zwischen den Zeitpunkten t1 und t2	$r_{t1-t0} = \frac{1}{n}\sum_{i}\frac{(y_{t1,i} - \overline{y_{t1}})}{S_{Y_{t1}}} * \frac{(y_{t0,i} - \overline{y_{t0}})}{S_{Y_{t0}}}$	=KORREL(Bereich$_{t1}$;Bereich$_{t0}$)
Statistischer Test	Abhängige Stichproben und gleiche Varianzen $t_{emp} = \frac{M_{diff}}{s_{diff} * \sqrt{\frac{1}{N}}} = \frac{M_{diff}}{SE_{M_{diff}}}$		
Kritischer t-Wert p-Wert	Abhängig: $t_{krit(\alpha;df)}$= ? bei df= N – 1		Kritischer t-Wert (positive Werte) =T.INV.2S(alpha;df) zweiseitig =T.INV(1-alpha;df) einseitig Empirischer p-Wert =T.VERT.RE(t_{emp};df) einseitig =T.VERT.2S(t_{emp};df) zweiseitig =T.TEST(Daten1;Daten2;Seiten;Typ) Seiten: 1=einseitig; 2= zweiseitig Typ: 1 = abhängige Stichproben

Entscheidung über Statis. Hypothese	Entscheidung nach folgendem Schema: Bei $t_{emp} \geq t_{krit}$ ODER $p \leq$ Alpha aus Testplanung ⇨ Annahme H_1 Bei $t_{emp} < t_{krit}$ ODER $p >$ Alpha aus Testplanung ⇨ Annahme H_0	
Effektstärke und Konfidenzintervall	Cohens d_{whm} $d_{whm} = M_{diff} / s_{diff}$ $= \frac{M_1 - M_0}{s_{iB} * \sqrt{1 - r_{t1-t0}}}$ whm: wiederholte Messungen	Standardfehler d_{whm} $SE_d = \widehat{\sigma_d} = \sqrt{\left(\frac{1}{n} + \frac{d^2}{2n}\right) * 2 * (1 - r)}$ wobei r die Korrelation der Messungen r_{t1-t0} ist Konfidenzintervall (bei Annahme Normalverteilung) • Untere Grenze: d – 1,96*SE_d • Obere Grenze: d + 1,96*SE_d
Konventionen Effektstärke		d >= 0,2 ⇨ Kleiner Effekt d >= 0,5 ⇨ Mittlerer Effekt d >= 0,8 ⇨ Großer Effekt d >= 1,2 ⇨ Sehr großer Effekt
Entscheidungen über PV	Entscheidung nach folgendem Schema: Bei abgeleiteter H_0 und angenommener H_0 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_0 und angenommener H_1 ⇨ PV nicht eingetreten, PH nicht bewährt Bei abgeleiteter H_1 und angenommener H_1 & $d_{emp} >$ d aus Testplanung bzw. Power ≥ 0,8 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_1 und angenommener H_1 & $d_{emp} <$ d aus Testplanung bzw. 0,6 ≤ Power < 0,8 ⇨ PV bedingt eingetreten, PH bedingt bewährt Bei abgeleiteter H_1 und angenommener H_0 ⇨ PV nicht eingetreten, PH nicht bewährt	

t-Tests für Mittelwertsunterschiede
3. Eine Stichprobe

Frage-stellung	**Gibt es einen Unterschied zwischen dem Mittelwert der Variable Y in einer Gruppe und dem bekannten (oder vermuteten) Mittelwert in einer bestimmten Population?**		
UV und AV	**UV/Prädiktorvariable X: kategorial mit einer Ausprägung** **AV/Kriteriumsvariable Y: kontinuierlich, auf Intervallskalenniveau gemessen**		
Inhaltliche Hypothesen	1. In einer Studie wird untersucht, ob Y im Mittel in einer Gruppe anders ist, als in einer bestimmten Population. Es wird erwartet, dass Y im Mittel größer/kleiner/ungleich/gleich als in der Vergleichspopulation ist.		
Ableitung Vorhersage & Stat. Hypothesen	PH-gerichtet: Y ist in der untersuchten Gruppe größer/kleiner als in der Vergleichspopulation. PH-ungerichtet: Y ist in der untersuchten Gruppe anders als in der Vergleichspopulation. ⇨ PV: Vermutete Abweichung des Mittelwerts in der Gruppe von der Vergleichspopulation. ⇨ Vorhersagekonforme Statistische Hypothesen: Gerichtet: H_1: $\mu_{Gruppe} > \mu_{vergleichspopulation}$ H_0: $\mu_{Gruppe} \leq \mu_{vergleichspopulation}$ Ungerichtet: H_1: $\mu_{Gruppe} \neq \mu_{vergleichspopulation;}$ H_0: $\mu_{Gruppe} = \mu_{vergleichspopulation}$		
Testplanung per Hand	Determinanten: Alpha, Beta, Effektgröße (Cohens d), N Eine Stichprobe gerichtete Testung N= $[(z_{1-\alpha}+z_{1-\beta})^2*2]/d^2$ Eine Stichprobe ungerichtete Testung N= $[(z_{1-\alpha/2}+z_{1-\beta})^2*2]/d^2$		z-Werte (positive Werte) =NORM.S.INV(0,95) (einseitig) =NORM.S.INV(0,975) (zweiseitig)
Relevante Kennwerte	Mittelwert Y	$M_Y = \frac{1}{n}\sum_{i=1}^{n} y_i$	=MITTELWERT(Bereich)
	Wenn Streuung in Vergleichspopulation bekannt, ist Schätzung unnötig (z.B. Intelligenz) σ_Y		
	Bei Schätzung der Streuung Y	$\hat{\sigma}_Y = s_Y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - M_Y)^2}$	=STABW.S(Bereich)
	Standardfehler des Mittelwerts wenn σ bekannt	$SE_M = \sigma_Y * \frac{1}{\sqrt{n}}$	= σ / WURZEL(ANZAHL(Bereich))
	Standardfehler des Mittelwerts wenn σ geschätzt	$SE_M = s_Y * \frac{1}{\sqrt{n}}$	=STABW.S(Bereich)/ WURZEL(ANZAHL(Bereich))
	Konfidenzintervall der mittleren Differenz (meist 95%)	$t_{(df=N-1,\ Konfidenz)} * SE_{Mdiff}$	=T.INV.2S(alpha;N-1)*SE_M =KONFIDENZ.T(alpha;STABW.S(); Anzahl)

Statistischer Test	Abhängige Stichproben und gleiche Varianzen $t_{emp} = \frac{M_Y - \mu}{SE_M}$	
Kritischer t-Wert p-Wert	$t_{krit(\alpha;df)}$= ? bei df= N – 1	Kritischer t-Wert (positive Werte) =T.INV.2S(Alpha;df) zweiseitig =T.INV(1-Alpha;df) einseitig Empirischer p-Wert =T.VERT.RE(t_{emp};df) einseitig =T.VERT.2S(t_{emp};df) zweiseitig
Entscheidung SH	Entscheidung nach folgendem Schema: Bei $t_{emp} \geq t_{krit}$ ODER $p \leq$ Alpha aus Testplanung ⇨ Annahme H_1 Bei $t_{emp} < t_{krit}$ ODER $p >$ Alpha aus Testplanung ⇨ Annahme H_0	
Effektgröße/ Konventionen	Cohens d d = (M-μ)/ σ	d = 0,2 ⇨ Kleiner Effekt d = 0,5 ⇨ Mittlerer Effekt d = 0,8 ⇨ Großer Effekt
Entscheidungen über PV	Entscheidung nach folgendem Schema: Bei abgeleiteter H_0 und angenommener H_0 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_0 und angenommener H_1 ⇨ PV nicht eingetreten, PH nicht bewährt Bei abgeleiteter H_1 und angenommener H_1 & d_{emp} > d aus Testplanung bzw. Power ≥ 0,8 ⇨ PV eingetreten, PH bewährt Bei abgeleiteter H_1 und angenommener H_1 & d_{emp} < d aus Testplanung bzw. 0,6 ≤ Power < 0,8 ⇨ PV bedingt eingetreten, PH bedingt bewährt Bei abgeleiteter H_1 und angenommener H_0 ⇨ PV nicht eingetreten, PH nicht bewährt	