

LÖSUNG 3B

- Gleichheit der Varianzen kann man mit dem Levene-Test im Menü „Analysieren“, "Deskriptive Statistiken", "Explorative Datenanalyse“ überprüfen (s. Kapitel 9.3.1).

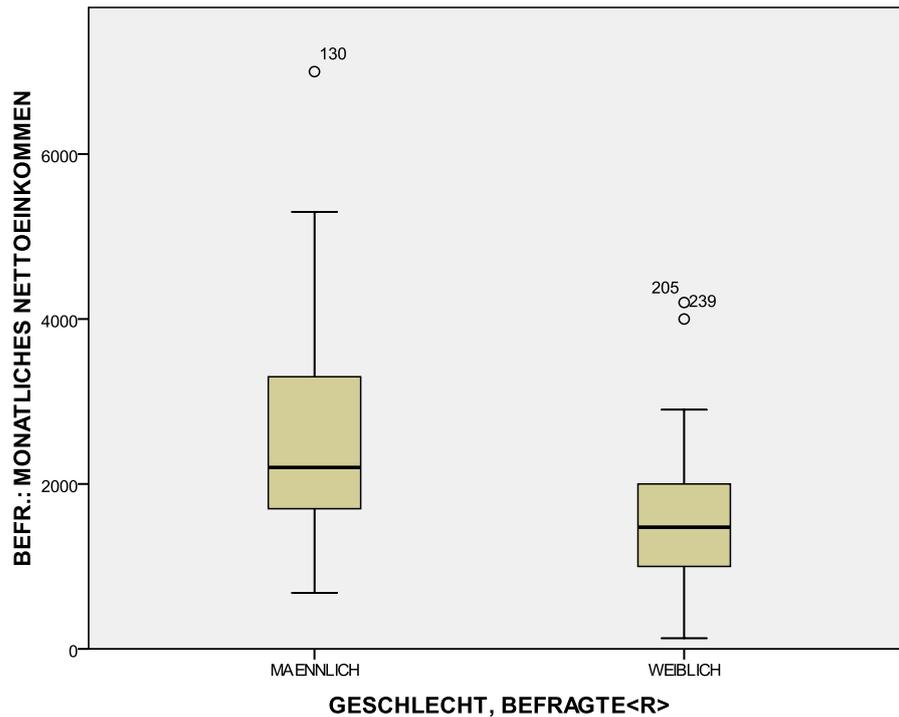
Bei Durchführung des t-Tests für 2 unabhängige Stichproben (im Menü „Analysieren“, "Mittelwerte", s. Lösung zu Aufgabe 6 sowie Kapitel 14.4.2.) wird automatisch ein Levene-Test berechnet. Der Vorteil des Levene-Tests im Menü "Explorative Datenanalyse" besteht aber darin, dass dieser Test mehrere Varianten des Levene-Tests durchführt (basierend auf dem Mittelwert, dem Median und auf dem getrimmten Mittelwert). Zur Absicherung des Ergebnisses des t-Tests sollte man auch die Ergebnisse dieser robusten Levene-Tests zu Rate ziehen.

- Generell sollte aber nicht übersehen werden, dass ein Test auf Gleichheit der Varianzen (analog zur Argumentation zum Test auf Abweichung einer Variable von einer Normalverteilung, s. Lösung zu Aufgabe 3a) nicht unproblematisch ist, weil das Testergebnis stark von der Stichprobengröße beeinflusst wird. Daher sollte man sich zum Vergleichen auch deskriptiver Verfahren bedienen. Man kann die Varianzen berechnen und direkt vergleichen. Auch Boxplots können hilfreich sein. Sie geben einen anschaulichen Vergleich der Größe der Streuung.
- Dazu kann zuerst eine vergleichende Berechnung der Varianz von EINK für Männer und Frauen erfolgen. Eine Möglichkeit dazu besteht darin, eine Gruppenaufteilung vorzunehmen. Wählen Sie „Daten“, „Aufgeteilte Datei“, dort in der Dialogbox „Datei aufteilen“ „Gruppen vergleichen“. Die Variable GESCHL in „Gruppen basierend auf“ ziehen und das Ganze ausführen. Danach berechnen der Varianzen mit „Analysieren“ „Deskriptive Statistiken“, „Deskriptive Statistik...“. Dort EINK in das Feld „Variable(n)“ übertragen. Nach Klicken der Schaltfläche „Optionen...“ öffnet die Dialogbox „Deskriptive Statistik: Optionen“. Dort „Varianz“ wählen. Es zeigt, sich dass die Varianz für Männer mehr als doppelt so groß ist als für Frauen. (Alternativ und einfacher wäre dies mit „Analysieren“, „Mittelwerte vergleichen“, „Mittelwerte“ durchzuführen.)

Deskriptive Statistik

geschl		N	Varianz
1 MAENNLICH	eink	81	1432672,986
	Gültige Werte (Listenweise)	81	
2 WEIBLICH	eink	62	599961,292
	Gültige Werte (Listenweise)	62	

- Auch ein Vergleich der Boxplots von EINK für Männer und Frauen zeigt die größere Streuung der Einkommen bei den Männern („Grafik“, „Diagrammerstellung“, „Boxplot“ wählen. Durch Doppelklicken auf das Symbol „Einfacher Boxplot“ diesen in die Diagrammvorschau übertragen. Dort GESCHL auf „X-Achse?“ und EINK auf Y-Achse?“ ziehen) (zu Boxplots s. Kapitel 32.13).



- Durchführen des Levene-Test im Menü "Explorative Datenanalyse": „Analysieren“, „Deskriptive Statistiken“, "Explorative Datenanalyse..." wählen. EINK in das Feld für „Abhängige Variablen:“ und GESCHL in das Feld „Faktorliste:“ ziehen, Klicken der Schaltfläche "Diagramme..." öffnet die entsprechende Dialogbox. Dort das Kontrollkästchen "Streubreite vs. mittleres Niveau mit Levene-Test" und den Optionsschalter "Nicht transformiert" wählen.
- Alle Varianten des Levine-Tests zeigen eine signifikante Abweichung von der Annahme gleicher Varianzen der Vergleichsgruppen. In allen Varianten des Tests ist der Wahrscheinlichkeitswert von "Signifikanz" kleiner als das unterstellte Signifikanzniveau in Höhe von $\alpha = 0,05$ (5 %). Die Tests bestätigen das obige Analyseergebnis ungleicher Varianzen.

Test auf Homogenität der Varianz

		Levene-Statistik	df1	df2	Signifikanz
eink BEFR.: MONATLICHES NETTOEINKOMME N	Basiert auf dem Mittelwert	10,165	1	141	,002
	Basiert auf dem Median	6,955	1	141	,009
	Basierend auf dem Median und mit angepaßten df	6,955	1	119,886	,009
	Basiert auf dem getrimmten Mittel	9,233	1	141	,003

- Die Frage, ob es sinnvoll ist, durch eine Transformation einer metrischen Variable Varianzhomogenität zu erreichen, muss immer im Zusammenhang mit der eigentlichen Aufgabenstellung gesehen werden. Durch die Transformation dürfen andere Bedingungen und Voraussetzungen eines statistischen Verfahrens nicht verletzt werden.

Im Fall z.B. der Diskriminanzanalyse werden für die betrachteten Gruppen sowohl gleiche Varianzen/Kovarianzen als auch eine Normalverteilung der metrischen Variablen vorausgesetzt. Wenn eine Transformation der metrischen Variablen zur Angleichung der Varianzen die Bedin-

gung der Normalverteilung nicht verletzt, ist sie sinnvoll. In dem Beispiel zur Diskriminanzanalyse im Buch (s. Kapitel 23) wurden die metrischen Variablen logarithmiert, um approximativ Normalverteilungen zu erreichen.

- EINK hat eine linkssteile Verteilung. Daher sind sowohl eine Logarithmierung als auch Quadratwurzelziehen von EINK Transformationen, die tendenziell die Verteilung in Richtung zu einer Normalverteilung verändern (s. Aufgabe 3a).
- Im Folgenden beschränken wir uns auf die Prüfung der Varianzhomogenität für logarithmierte Werte von EINK. Für den Fall der Quadratwurzelbildung ist analog vorzugehen

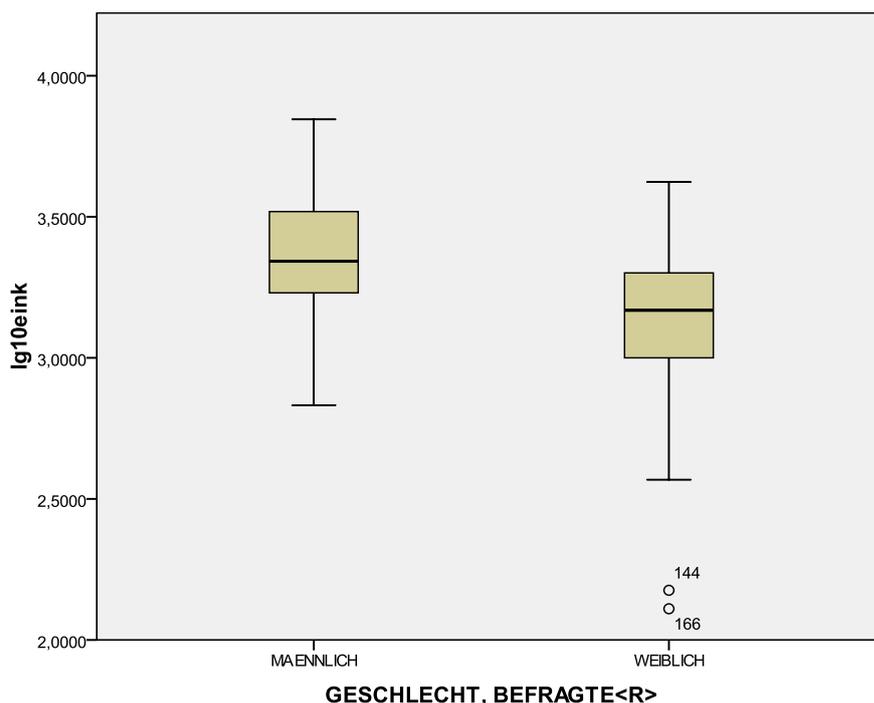
Die Logarithmierung kann mittels „Transformieren“ (s. Lösung zu Aufgabe 3a) vorgenommen werden oder alternativ in der Dialogbox „Explorative Datenanalyse: Diagramme“. Im letzteren Fall wird neben dem Kontrollkästchen "Streubreite vs. mittleres Niveau mit Levene-Test", der Optionsschalter „Transformiert“ aktiviert und in dem Eingabefeld „Exponent“ „Natürl. Log“ aus der Drop-Down-Liste gewählt (ob EINK mit dem Logarithmus zur Basis 10 oder alternativ der natürlichen Basis e transformiert wird, ist für das Testergebnis unerheblich).

- Ein Vergleich der Varianzen der logarithmierten Einkommen für die beiden Gruppen zeigt, dass die Varianzen sich angenähert haben.

Deskriptive Statistik

geschl		N	Varianz
1 MAENNLICH	lg10eink	81	,045
	Gültige Werte (Listenweise)	81	
2 WEIBLICH	lg10eink	62	,071
	Gültige Werte (Listenweise)	62	

- Ein Vergleich der Boxplots zeigt, dass für den mittleren Bereich der Verteilung von LOG10EINK sich die Streuungen der beiden Gruppen nur wenig unterscheiden.



- Führt man mit für die Variable LG10EINK den Levene-Test durch: „Analysieren“, "Explorative Datenanalyse" (in der Unterdialogbox "Explorative Datenanalyse: Diagramme" wird für "Streuung vs. mittleres Niveau mit Levene-Test" "Nicht transformiert" gewählt), ergibt sich das in der Tabelle unten angegebene Ergebnis. Es zeigt sich, dass die Werte für "Signifikanz" größer sind als durch das Signifikanzniveau von $\alpha = 0,05$ festgelegt. Damit wird die H_0 -Hypothese, gleiche Varianzen in den Vergleichsgruppen, nicht abgelehnt. Der Test bestätigt das durch den Vergleich der Boxplots gewonnene Ergebnis.

Test auf Homogenität der Varianz

		Levene- Statistik	df1	df2	Signifikanz
lg10eink	Basiert auf dem Mittelwert	,566	1	141	,453
	Basiert auf dem Median	,464	1	141	,497
	Basierend auf dem Median und mit angepaßten df	,464	1	123,241	,497
	Basiert auf dem getrimmten Mittel	,451	1	141	,503