

AUFGABE 3A

Für viele statistische Methoden wird vorausgesetzt, dass eine untersuchte Variable normalverteilt ist. Dies sollte vor Anwendung der Methode geprüft werden.

Wenn die Bedingung der Normalverteilung für eine Variable grob verletzt wird, wird häufig versucht, durch Transformation der Variable näherungsweise eine Normalverteilung zu erreichen. Im Anwendungsbeispiel zur Diskriminanzanalyse in Kapitel 23 z.B. wurden die metrischen Variablen logarithmiert um näherungsweise eine Normalverteilung zu erreichen

Für z.B. den t-Test auf Differenz von Mittelwerten einer metrischen Variable x für zwei unabhängige Stichproben (s. Aufgabe 6) wird für die Stichprobenverteilung von \bar{x} eine Normalverteilung vorausgesetzt. Bei relativ kleinen Stichproben ist dies nur zu erwarten, wenn die Stichprobe aus einer normalverteilten Grundgesamtheit stammt. Daher sollte man prüfen, ob x in der Stichprobe annähernd normalverteilt ist.

In z.B. der Regressionsanalyse wird für Signifikanzprüfungen von Regressionskoeffizienten vorausgesetzt, dass die Residualvariable im Grundgesamtheitsmodell normalverteilt ist (s. Kapitel 18.1.2). Für die empirische Residualvariable eines Regressionsmodells sollte daher geprüft werden, ob sie annähernd normalverteilt ist.

Prüfungen auf Normalverteilung haben also einen hohen Stellenwert.

In dieser Aufgabe sollen Sie prüfen, ob die Variable EINK (Datei ALLBUS90.SAV) näherungsweise normalverteilt ist.

Prüfen Sie des Weiteren, ob durch Logarithmierung oder durch Quadratwurzelziehen von EINK näherungsweise eine Normalverteilung erreicht werden kann.