Einführung in die Ökonometrie — **Wintersemester 16/17**

Elizabeth Heller — **Blatt 7**

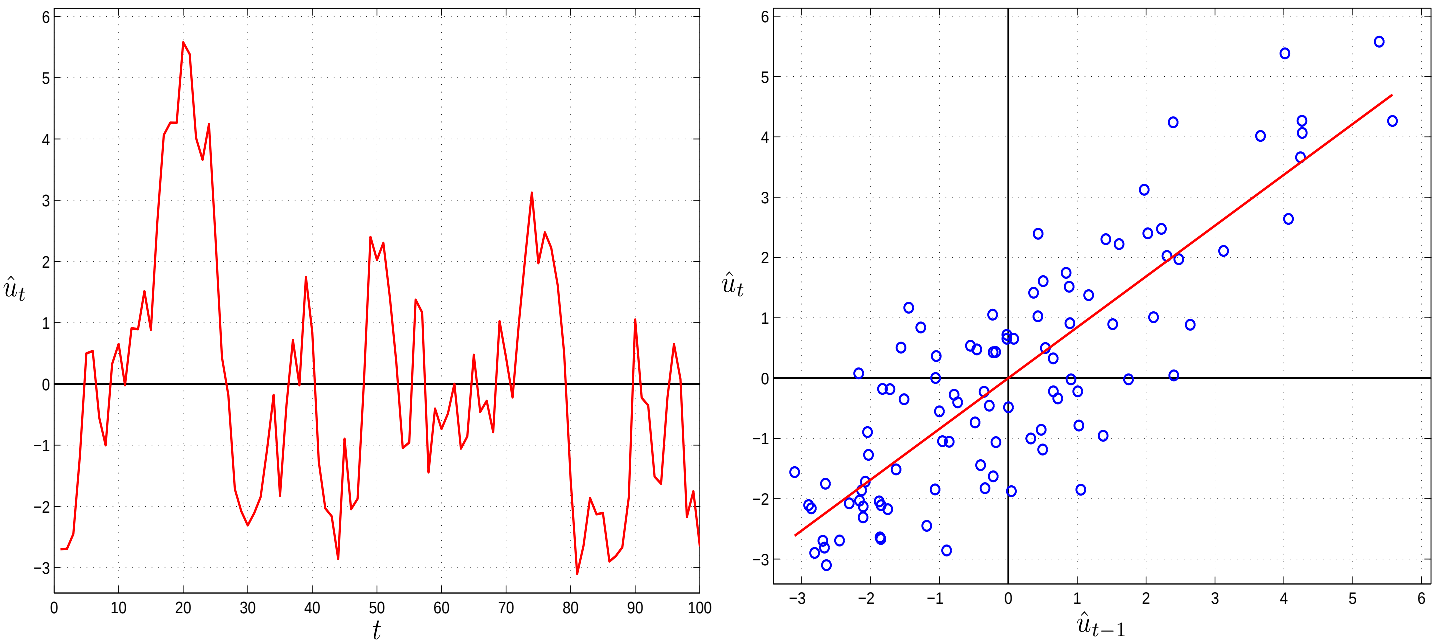
Lösungsvorschläge: Serkan Yener

(a) Für eine Stichprobe von Zeitreihendaten wird die Regressionsgleichung

$$y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t \qquad \text{für} \quad t = 1, \dots, 100$$

geschätzt. In Abbildung 1 wurden die aus der KQ-Schätzung gewonnenen Residuen $\hat{u}_t$ in einem Zeitreihendiagramm (a) bzw. aufeinanderfolgende Residuen in einem Streuungsdiagramm (b) abgetragen.

Abbildung 1: Beispiel für Residuen $\hat{u}_t$ eines linearen Regressionsmodells.



(a) Zeitreihendiagramm der Residuen $\hat{u}_t$ (b) Streuungsdiagramm von $\hat{u}_t$ vs. $\hat{u}_{t-1}$

Welche Eigenschaft der vorliegenden Stichprobe vermuten Sie und was für Probleme gehen damit einher?

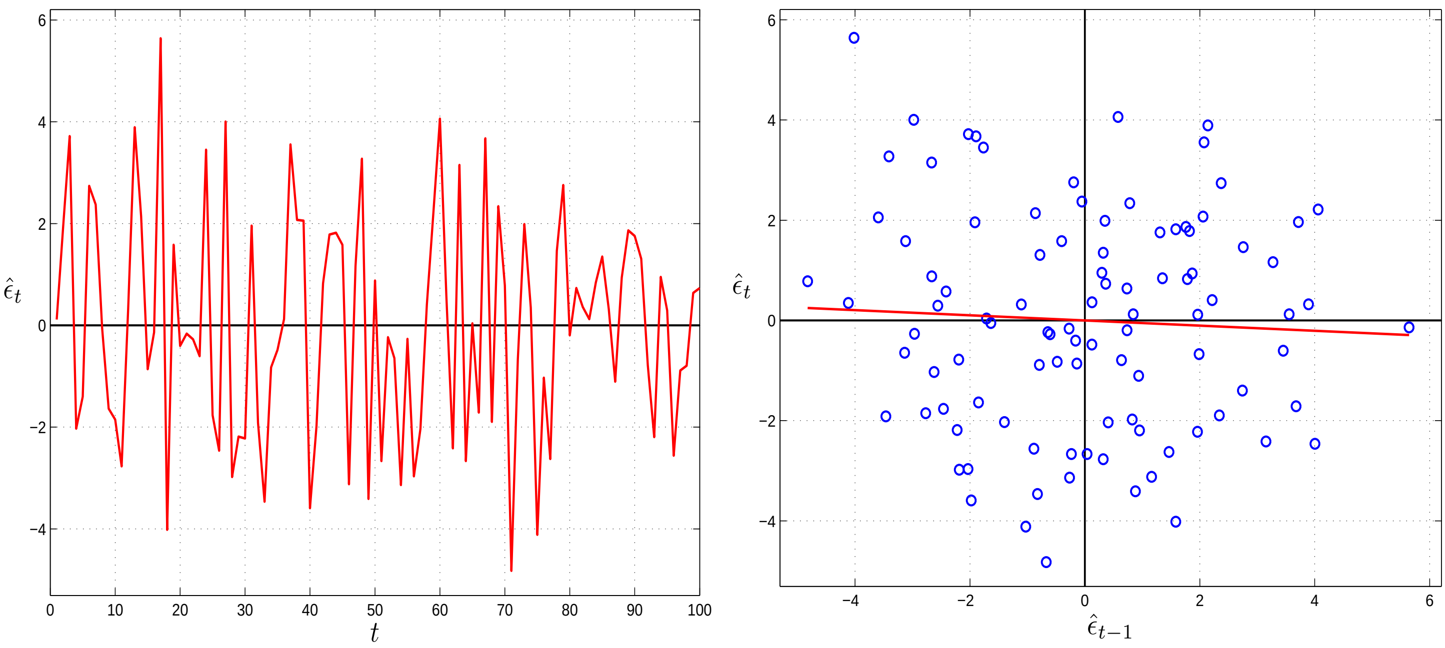
Welche Rolle bei der Erklärung des vermuteten Phänomens spielt die Kennzahl

$$\hat{\rho} = \frac{\sum_{t=2}^{100} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{100} \hat{u}_{t-1}^2} = 0.8428 \ ?$$

Wenn $\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{d} \mathrm{N}(0, 1 - \rho^2)$ gilt, wie könnte ein primitiver Test zum Signifikanzniveau 5% bezüglich des vermuteten Phänomens aussehen und wie würde er ausfallen?

Lösungsvorschlag: Die beiden Diagramme in Abbildung 1 deuten auf einen gewissen Grad an Verharrungsvermögen der Residuen hin, d.h. positive (negative) Residuen tendieren dazu, auf positive (negative) Residuen zu folgen. Es ist hilfreich sich hierfür als Kontrast die entsprechenden Diagramme von 100 *iid* Zufallsvariablen zu betrachten, deren Erwartungswerte bzw. Varianzen mit dem Stichprobenmittelwert bzw. der -varianz von $\hat{u}_t$ übereinstimmen. In Diagramm (a) von Abbildung 2 ist die Zeitreihe dieser "idealisierten" Störgrößen abgetragen. Offensichtlich besitzt die Zeitreihe der idealisierten Störgrößen keine vergleichbare Persistenz, sondern schwankt erratisch um die Abszisse. Ursächlich hierfür ist die Annahme, dass die "idealisierten" Störgrößen unkorreliert sind. Ebenso zeigt Diagramm (b) in Abbildung 2, dass wohl kein (statistisch signifikanter) linearer Zusammenhang zwischen aufeinander folgenden "idealisierten" Störgrößen existiert.

Abbildung 2: Beispiel für "idealisierte" Residuen $\hat{\epsilon}_t$.
Die "idealisierten" Residuen $\hat{\epsilon}_t$ sind so konstruiert, dass zwar $\mathrm{E}[\,\hat{\epsilon}_t\,] = \bar{\hat{u}}$ (Mittelwert 0) und $\mathrm{Var}[\,\hat{\epsilon}_t\,] = \hat{\sigma}^2_{\hat{u}_t}$ (keine Heteroskedastizität) gilt, jedoch wurde—abweichende von $\hat{u}_t$—$\mathrm{Cov}[\,\hat{\epsilon}_t, \hat{\epsilon}_{t-1}\,] = 0$ (keine Autokorrelation) angenommen.



(a) Zeitreihendiagramm "idealisierter" Störgrößen $\hat{\epsilon}_t$

(b) Streuungsdiagramm von $\hat{\epsilon}_t$ vs. $\hat{\epsilon}_{t-1}$

Diese Persistenz nennt man Autokorrelation und ist in den meisten Fällen ein Resultat einer Modellfehlspezifikation. Diese kann wiederum aus einer falschen funktionalen Form, einem systematischen Messfehler der endogenen Variablen und/oder dem Fehlen einer oder mehrerer (exogener) Variablen entstanden sein. Üblicherweise wird eine fehlende, autokorrelierte Exogene von der Störgröße absorbiert, welche dann durch ein ähnliches, dynamisches Verhalten gekennzeichnet ist. Ist es nicht möglich, diese Variable zu identifizieren oder ihren Einfluss zu eliminieren, so ist die

Annahme unkorrelierter Störgrößen des klassischen, linearen Regressionsmodell aus Aufgabenblatt 1 verletzt. Insbesondere sind die Einträge der Kovarianzmatrix der Störgrößen jenseits der Hauptdiagonalen nun nicht mehr 0. Damit gelten wiederum die Erkenntnisse aus Aufgabenblatt 5, dass der KQ-Schätzer $\hat{\boldsymbol{\beta}}_{KQ}$ zwar noch erwartungstreu und konsistent, jedoch nicht mehr effizient im Sinne vom Gauss-Markov ist.

Für die Formulierung eines effizienten Schätzers benötigt man zunächst ein geeignetes statistisches Modell, das z.B. folgendermaßen lauten könnte:

$$\begin{aligned} y_t &= \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t \\ u_t &= \rho u_{t-1} + \epsilon_t \end{aligned}$$

für $t = 1, \ldots, T$ mit $\boldsymbol{\epsilon} \sim \mathrm{N}(\,\mathbf{0}_{T\times 1}, \sigma^2 \boldsymbol{I}_T\,)$, wobei u_0 als bekannt oder als $u_0 = 0$ angenommen wird. Man bezeichnet dieses Modell als lineares Regressionsmodell mit autokorrelierten Störgrößen 1.Ordnung. Obwohl dieses Modell den einfachsten Fall autokorrelierter Störgrößen darstellt, hat es sich jedoch in ökonomischen Anwendungen als eines der wichtigsten Modelle herausgestellt. Die Schlüsselgröße dieses Modells ist ρ: Der Korrelationskoeffizient, der den linearen Zusammenhang zwischen u_t und u_{t-1} misst. Damit dieses Modell überhaupt wohldefiniert ist, muss man die Restriktion $|\rho| < 1$ auferlegen, da der Störgrößenprozess u_t sonst nicht (schwach) stationär ist. Ein einfacher Test auf Autokorrelation 1. Ordnung kann dann

$$\mathrm{H}_0 : \rho = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 : \rho \neq 0$$

lauten. Da der wahre Wert von ρ nicht bekannt ist, muss er durch sein Stichprobenpendant $\hat{\rho}$ ersetzt werden. Unter der vorgegebenen Verteilungsannahme kann die folgende (asymptotische) Teststatistik konstruiert werden:

$$\sqrt{T}\frac{\hat{\rho} - \rho}{\sqrt{1-\rho^2}} = \frac{\hat{\rho} - \rho}{\sqrt{(1-\rho^2)/T}} \overset{\text{asy.}}{\sim} \mathrm{N}(\,0, 1\,)\ .$$

Bei Gültigkeit der Nullhypothese $\mathrm{H}_0 : \rho = 0$ in der Realität folgt die vereinfachte Teststatistik

$$Z := \sqrt{T}\hat{\rho} \overset{\text{asy.}}{\sim} \mathrm{N}(\,0, 1\,)\ .$$

Ähnlich dem Vorgehen bei t-Tests in Aufgabenblatt 2, verwenden wir an dieser Stelle die Symmetrieeigenschaft der Normalverteilung, so dass der ursprüngliche Test äquivalent ist zum Test darüber, ob $|Z|$ im (positiven) kritischen Bereich liegt, also ob $|Z|$ das rechte Quantil $Z_{krit} = \Phi^{-1}(1 - \alpha/2)$ überschreitet.
Hypothese:

$$\mathrm{H}_0 : \rho = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 : \rho \neq 0\ .$$

Teststatistik:

$$Z = \sqrt{T}\hat{\rho} = \sqrt{100} \cdot 0.8428 = 8.428\ .$$

Kritischer Wert:

$$Z_{krit} = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0,975) = 1,960\ .$$

Testentscheidung: H_0 wird abgelehnt, da $Z_{krit} = 1,960 < 8.428 = Z$, d.h. die Störgrößen sind autokorreliert.

(b) Welcher Zusammenhang besteht zwischen dem Test aus Teilaufgabe (a) und dem Durbin-Watson Test? Bestimmen Sie den Wert der Durbin-Watson Teststatistik DW anhand der Informationen aus Teilaufgabe (a)!

Lösungsvorschlag: Der in Teilaufgabe (a) benutzte Test ist ein asymptotischer Test, der in kleinen Stichproben problematisch sein kann. Genauer gesagt, hat dieser Test eine schlechte *Power* gegenüber Autokorrelationstest für endliche Stichproben. Als die Power eines Tests kann man seine Eigenschaft oder Fähigkeit verstehen, die Nullhypothese abzulehnen, wenn sie in der Realität falsch ist. Durbin and Watson (1950) haben die Teststatistik

$$DW = \frac{\sum_{t=2}^{T} \left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\sum_{t=1}^{T} \hat{u}_t^2}$$

zum Testen von Autokorrelation 1. Ordnung in den Residuen vorgeschlagen. Für diese Teststatistik konnten sie eine Wahrscheinlichkeitsverteilung in endlichen Stichproben herleiten. Jedoch gelang es nicht, kritische Werte zur Partitionierung des Wertebereichs in Annahme- und Ablehnungsbereich zu bestimmen, sondern lediglich obere und untere Schranken für DW. Für den Zusammenhang zwischen

$$\hat{\rho} = \frac{\sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{T} \hat{u}_{t-1}^2}$$

und DW formt man zunächst die Durbin-Watson Teststatistik um

$$
\begin{aligned}
DW &= \frac{\sum_{t=2}^{T}\big(\hat{u}_t - \hat{u}_{t-1}\big)^2}{\sum_{t=1}^{T}\hat{u}_t^2} = \frac{\sum_{t=2}^{T}\big(\hat{u}_t^2 + \hat{u}_{t-1}^2 - 2\hat{u}_t\hat{u}_{t-1}\big)}{\sum_{t=1}^{T}\hat{u}_t^2} \\
&= \frac{\sum_{t=2}^{T}\hat{u}_t^2 + \hat{u}_1^2 - \hat{u}_1^2 + \sum_{t=2}^{T}\hat{u}_{t-1}^2 + \hat{u}_T^2 - \hat{u}_T^2 - 2\sum_{t=2}^{T}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{T}\hat{u}_t^2} \\
&= \frac{\sum_{t=1}^{T}\hat{u}_t^2 - \hat{u}_1^2 + \sum_{t=1}^{T}\hat{u}_t^2 - \hat{u}_T^2 - 2\sum_{t=2}^{T}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{T}\hat{u}_t^2} \\
&= \frac{2\sum_{t=1}^{T}\hat{u}_t^2}{\sum_{t=1}^{T}\hat{u}_t^2} - 2\frac{\sum_{t=2}^{T}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=1}^{T}\hat{u}_t^2} - \frac{\hat{u}_1^2 + \hat{u}_T^2}{\sum_{t=1}^{T}\hat{u}_t^2} \\
&= 2 - 2\frac{\sum_{t=2}^{T}\hat{u}_t\hat{u}_{t-1}}{\sum_{t=2}^{T}\hat{u}_{t-1}^2}\frac{\sum_{t=2}^{T}\hat{u}_{t-1}^2}{\sum_{t=1}^{T}\hat{u}_t^2} - \frac{\hat{u}_1^2 + \hat{u}_T^2}{\sum_{t=1}^{T}\hat{u}_t^2} \\
&= 2 - 2\hat{\rho}\frac{\sum_{t=2}^{T}\hat{u}_{t-1}^2}{\sum_{t=1}^{T}\hat{u}_t^2} - \frac{\hat{u}_1^2 + \hat{u}_T^2}{\sum_{t=1}^{T}\hat{u}_t^2}
\end{aligned}
$$

und betrachtet den Fall, dass der Stichprobenumfang T groß ist, dann gilt die Approximation

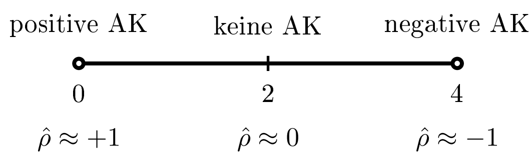
$$\widetilde{DW} \approx 2 - 2\hat{\rho} = 2(1-\hat{\rho}) \; ,$$

weil dann

$$\frac{\sum_{t=2}^{T}\hat{u}_{t-1}^2}{\sum_{t=1}^{T}\hat{u}_t^2} \approx 1 \qquad \text{und} \qquad \frac{\hat{u}_1^2 + \hat{u}_T^2}{\sum_{t=1}^{T}\hat{u}_t^2} \approx 0$$

gilt. Demnach ergibt sich $\widetilde{DW} = 2(1 - 0.8428) = 0.3144$. Aus dieser Beziehung lassen sich auch die Extreme der DW-Teststatistik erkennen, denn, da wir ρ auf das Interval $-1 < \rho < +1$ beschränkt haben, sollte $0 < DW, \widetilde{DW} < 4$ gelten. Vgl. Abbildung 3.

Abbildung 3: Durbin-Watson Teststatistik & Stichprobenautokorrelation.



(c) Beschreiben Sie das grundlegende Problem des Durbin-Watson Tests! Führen Sie den Durbin-Watson Test zum Signifikanzniveau 5% anhand von $\widetilde{DW}$ und $DW = 0.3083$ durch! Welche Hypothesen sind insbesondere für die Praxis relevant?

Lösungsvorschlag: Die Durbin-Watson Teststatistik unterscheidet sich auf fundamentale Weise von den bisher verwendeten Teststatistiken. Während für die Verteilung der t- und F-Teststatistiken lediglich die Anzahl der Freiheitsgrade ausschlaggebend waren, hängt die Verteilung der Durbin-Watson Teststatistik direkt von den

Regressoren $\boldsymbol{X}$ ab. Um dies zu erkennen, definiert man zunächst die Matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & & \vdots \\ 0 & -1 & 2 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix},$$

so dass die Durbin-Watson Teststatistik umgeschrieben werden kann als

$$DW = \frac{\hat{\boldsymbol{u}}'\boldsymbol{A}\hat{\boldsymbol{u}}}{\hat{\boldsymbol{u}}'\hat{\boldsymbol{u}}} = \frac{(\boldsymbol{M}\boldsymbol{u})'\boldsymbol{A}(\boldsymbol{M}\boldsymbol{u})}{(\boldsymbol{M}\boldsymbol{u})'(\boldsymbol{M}\boldsymbol{u})} = \frac{\boldsymbol{u}'\boldsymbol{M}\boldsymbol{A}\boldsymbol{M}\boldsymbol{u}}{\boldsymbol{u}'\boldsymbol{M}\boldsymbol{u}}\ .$$

Durbin and Watson (1950) haben dann gezeigt, dass sich dieser Ausdruck als

$$DW = \frac{\sum_{t=1}^{T-k} \lambda_t z_t^2}{\sum_{t=1}^{T-k} z_t^2}$$

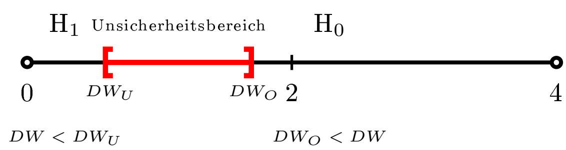
formulieren lässt, wobei $\lambda_1, \ldots, \lambda_{T-k}$ die von Null verschiedenen Eigenwerte der Matrix $\boldsymbol{MAM}$ und $z_1, \ldots, z_{T-k}$ unabhängige $\mathrm{N}(0, \sigma^2)$-Zufallsvariablen sind. Während man den Nenner wieder (nach Standardisierung) als eine $\chi^2(T-k)$-verteilte Zufallsvariable begreifen kann, gilt dies nicht für den Zähler. Außerdem kann man zeigen, dass Zähler und Nenner nicht mehr unabhängig voneinander sind, weshalb sich der übliche "F-Verteilungsverdacht" hier nicht bestätigt. Jedoch ist die Abhängigkeit der Eigenwerte λ_t von $\boldsymbol{X}$ ein wesentlich schwerwiegenderes Problem. Für einen gegebenen Datensatz $\boldsymbol{X}$ lassen sich mit den in Durbin and Watson (1950) und Durbin and Watson (1971) diskutierten Verfahren die Verteilung von DW in endlichen Stichproben approximieren. Bei der heutigen Rechnerleistung ist es kein Problem mehr diese Verteilung und die entsprechenden kritischen Werte exakt für einen gegebenen Datensatz zu berechnen. In den 1950er Jahren war dies allerdings noch ein gravierendes Problem, so dass Durbin und Watson sich darauf beschränken mussten, für "extreme" Realisationen der exogenen Variablen in $\boldsymbol{X}$ eine untere Schranke DW_U und eine obere Schranke DW_O für DW zu berechnen. Diese Schranken wurden in Durbin and Watson (1951) erstmals veröffentlicht und befinden sich teilweise auch in Tabelle A.6 im Skript.

Ist eine Zeitreihe negativ autokorreliert ($-1 < \rho < 0$), so besitzt sie die Tendenz zu periodischen Schwankungen, was bei ökonomischen (außer bei nicht-saisonbereinigten) Daten eher selten zu beobachten ist. Von einem theoretischen Standpunkt aus betrachtet, sollten sich die meisten ökonomischen Variablen langsam verändern, da sie diese Veränderungen das Ergebnis von graduellen Anpassungsprozessen sind. Dieses Verharrungsvermögen oder Persistenz lässt sich eher durch positive Autokorrelation erklären. Da sich dies auch mit unserer graphischen Analyse des gegebenen Datensatzes deckt, lautet der einseitige Hypothesentest demnach

$$\mathrm{H}_0 : \rho = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 : \rho > 0 \, .$$

Ist für den Durbin-Watson Test kein exakter kritischer Wert vorhanden, sondern lediglich die untere und obere Schranke, so schiebt sich ein zusätzlicher "Unsicherheitsbereich" zwischen Annahme- und Ablehnungsbereich. Vgl. Abbildung 4

Abbildung 4: Durbin-Watson Teststatistik & Stichprobenautokorrelation.



Für einen Stichprobenumfang $T = 100$ und zwei exogene Regressoren (ohne Konstante) $k' = 2$ ergibt sich aus Tabelle A.6 eine untere Schranke von $DW_U = 1.63$ und eine obere Schranke von $DW_O = 1.72$. Damit gilt

$$\widetilde{DW} = 0.3144 \quad < \quad 1,63 = DW_U$$
$$DW = 0.3083 \quad < \quad 1,63 = DW_U$$

und wir lehnen die Nullhypothese "keine Autokorrelation in den Störgrößen" zum Signifikanzniveau von 5% ab.

(d) Nennen Sie weitere Probleme des Durbin-Watson Tests! Welche Abhilfen kennen Sie?

Lösungsvorschlag: Außer der Freiheit von Autokorrelation setzt der Durbin-Watson Test die Gültigkeit aller Annahmen des linearen Regressionsmodells voraus—inklusive der Normalverteilung der Störgrößen. Die Annahme normalverteilter Störgrößen lässt sich mit Hilfe des Zentralen Grenzwertsatzes umgehen, wenn man gewillt ist, statt eines Tests für endliche Stichproben einen asymptotischen Test zu verwenden, wie z.B. den Test von Breusch (1978) und Godfrey (1978) oder den Test von Ljung and Box (1978). Dabei ist der Ljung-Box Test problematisch, da die Herleitung aus einem reinen Zeitreihenmodell ohne exogene Regressoren erfolgte.

Ein weiterer Nachteil des Durbin-Watson Tests ist, dass die Regressionsgleichung zwingend mit Absolutglied geschätzt worden sein muss, was bei den Tests von Breusch-Godfrey und Ljung-Box nicht sein muss.

Weiter darf die Regressormatrix auch keine verzögerte endogene Variable $(y_{t-1}, y_{t-2}, \ldots)$ enthalten. Dies kann ein wesentlicher Nachteil sein, weil oftmals die Autokorrelation in den Störgrößen durch das Einführen einer verzögerten Endogenen beseitigt werden kann, deren Fehlen dann als eine Fehlspezifikation verstanden werden kann. Ein Test auf Autokorrelation der neuen Störgrößen ist dann mit dem Durbin-Watson Test nicht mehr valide. Durbin (1970) hat mit seinem h-Test eine mögliche Lösung dieses Problems bereit gestellt. Interessanterweise sind der Breusch-Godfrey Test und Durbin's h-Test asymptotisch equivalent.

Schließlich ist der Durbin-Watson Test nur für das Testen auf Autokorrelation 1. Ordnung angelegt. Damit lässt sich nicht bestimmen, ob lediglich Autokorrelation 1. Ordnung oder aber auch höherer Ordnung vorliegt. Für den Fall saisonaler Autokorrelation, welche z.B. in nicht-saisonbereinigten Quartalsdaten zu finden ist, hat Wallis (1972) die Durbin-Watson Teststatistik erweitert, um auf Autokorrelation 4. Ordnung zu testen.