5. Under some comditions, we have $W \geq LR \geq LM$. See Engle (1981) "Wald, Likelihood and Lagrange Multiplier Tests in Econometrics," Chap. 13 in *Handbook of Econometrics*, Vol.2, Grilliches and Intriligator eds, North-Holland.

13.2 Example: W, LM and LR Tests

Date file $\Longrightarrow$ cons99.txt (same data as before)

Each column denotes year, nominal household expenditures (家計消費，10 billion yen), household disposable income (家計可処分所得，10 billion yen) and household expenditure deflator (家計消費デフレータ，1990=100) from the left.

1955	5430.1	6135.0	18.1	1970	37784.1	45913.2	35.2	1985	185335.1	220655.6	93.9
1956	5974.2	6828.4	18.3	1971	42571.6	51944.3	37.5	1986	193069.6	229938.8	94.8
1957	6686.3	7619.5	19.0	1972	49124.1	60245.4	39.7	1987	202072.8	235924.0	95.3
1958	7169.7	8153.3	19.1	1973	59366.1	74924.8	44.1	1988	212939.9	247159.7	95.8
1959	8019.3	9274.3	19.7	1974	71782.1	93833.2	53.3	1989	227122.2	263940.5	97.7
1960	9234.9	10776.5	20.5	1975	83591.1	108712.8	59.4	1990	243035.7	280133.0	100.0
1961	10836.2	12869.4	21.8	1976	94443.7	123540.9	65.2	1991	255531.8	297512.9	102.5
1962	12430.8	14701.4	23.2	1977	105397.8	135318.4	70.1	1992	265701.6	309256.6	104.5
1963	14506.6	17042.7	24.9	1978	115960.3	147244.2	73.5	1993	272075.3	317021.6	105.9
1964	16674.9	19709.9	26.0	1979	127600.9	157071.1	76.0	1994	279538.7	325655.7	106.7
1965	18820.5	22337.4	27.8	1980	138585.0	169931.5	81.6	1995	283245.4	331967.5	106.2
1966	21680.6	25514.5	29.0	1981	147103.4	181349.2	85.4	1996	291458.5	340619.1	106.0
1967	24914.0	29012.6	30.1	1982	157994.0	190611.5	87.7	1997	298475.2	345522.7	107.3
1968	28452.7	34233.6	31.6	1983	166631.6	199587.8	89.5				
1969	32705.2	39486.3	32.9	1984	175383.4	209451.9	91.8				

```
          PROGRAM
LINE  *********************************************
|         freq a;
|         smpl 1955 1997;
|         read(file='cons99.txt') year cons yd price;
|         rcons=cons/(price/100);
|         ryd=yd/(price/100);
|         lyd=log(ryd);
|         olsq rcons c ryd;
|         olsq @res @res(-1);
|         ar1  rcons c ryd;
|         olsq rcons c lyd;
|         param a1 0 a2 0 a3 1;
|         frml eq rcons=a1+a2*((ryd**a3)-1.)/a3;
|         lsq(tol=0.00001,maxit=100) eq;
|         a3=1.15;
|         rryd=((ryd**a3)-1.)/a3;
|         ar1  rcons c rryd;
|         end;
***************************************************
```

```
                                  Equation   1
                                  ============
                    Method of estimation = Ordinary Least Squares

Dependent variable: RCONS
Current sample:  1955 to 1997
Number of observations:  43

       Mean of dep. var. = 146270.          LM het. test = .207443 [.649]
  Std. dev. of dep. var. = 79317.2         Durbin-Watson = .115101 [.000,.000]
Sum of squared residuals = .129697E+10  Jarque-Bera test = 9.47539 [.009]
   Variance of residuals = .316335E+08   Ramsey's RESET2 = 53.6424 [.000]
Std. error of regression = 5624.36       F (zero slopes) = 8311.90 [.000]
               R-squared = .995092        Schwarz B.I.C. = 435.051
      Adjusted R-squared = .994972        Log likelihood = -431.289

           Estimated    Standard
Variable  Coefficient     Error       t-statistic   P-value
C         -2919.54      1847.55       -1.58022      [.122]
RYD       .852879       .935486E-02   91.1696       [.000]
```

```
                              Equation   2
                              ============
                Method of estimation = Ordinary Least Squares

Dependent variable: @RES
Current sample:  1956 to 1997
Number of observations:  42

       Mean of dep. var. = -95.5174
  Std. dev. of dep. var. = 5588.52
Sum of squared residuals = .146231E+09
   Variance of residuals = .356662E+07
Std. error of regression = 1888.55
               R-squared = .885884
      Adjusted R-squared = .885884
            LM het. test = .760256 [.383]
           Durbin-Watson = 1.40409 [.023,.023]
              Durbin's h = 1.97732 [.048]
         Durbin's h alt. = 1.91077 [.056]
        Jarque-Bera test = 6.49360 [.039]
         Ramsey's RESET2 = .186107 [.668]
          Schwarz B.I.C. = 377.788
          Log likelihood = -375.919

          Estimated    Standard
Variable  Coefficient    Error       t-statistic   P-value
@RES(-1)  .950693      .053301       17.8362       [.000]
```

Equation 3

============

FIRST-ORDER SERIAL CORRELATION OF THE ERROR
Objective function: Exact ML (keep first obs.)

Dependent variable: RCONS
Current sample: 1955 to 1997
Number of observations: 43

Mean of dep. var. = 146270.	R-squared = .999480
Std. dev. of dep. var. = 79317.2	Adjusted R-squared = .999454
Sum of squared residuals = .145826E+09	Durbin-Watson = 1.38714
Variance of residuals = .364564E+07	Schwarz B.I.C. = 391.061
Std. error of regression = 1909.36	Log likelihood = -385.419

Parameter	Estimate	Standard Error	t-statistic	P-value
C	1672.42	6587.40	.253881	[.800]
RYD	.840011	.027182	30.9032	[.000]
RHO	.945025	.045843	20.6143	[.000]

Equation 4
============

Method of estimation = Ordinary Least Squares

Dependent variable: RCONS
Current sample: 1955 to 1997
Number of observations: 43

Mean of dep. var.	= 146270.	LM het. test	= 2.21031 [.137]
Std. dev. of dep. var.	= 79317.2	Durbin-Watson	= .029725 [.000,.000]
Sum of squared residuals	= .256040E+11	Jarque-Bera test	= 3.72023 [.156]
Variance of residuals	= .624487E+09	Ramsey's RESET2	= 344.855 [.000]
Std. error of regression	= 24989.7	F (zero slopes)	= 382.117 [.000]
R-squared	= .903100	Schwarz B.I.C.	= 499.179
Adjusted R-squared	= .900737	Log likelihood	= -495.418

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
C	-.115228E+07	66538.5	-17.3175	[.000]
LYD	109305.	5591.69	19.5478	[.000]

```
                               NONLINEAR LEAST SQUARES
                               =======================
CONVERGENCE ACHIEVED AFTER  84 ITERATIONS

Number of observations = 43       Log likelihood = -414.362
        Schwarz B.I.C. = 420.004

                              Standard
Parameter  Estimate             Error        t-statistic   P-value
A1         16544.5            2615.60        6.32530       [.000]
A2         .063304            .024133        2.62307       [.009]
A3         1.21694            .031705        38.3839       [.000]

Standard Errors computed from quadratic form of analytic first derivatives
(Gauss)

Equation: EQ
Dependent variable: RCONS

       Mean of dep. var. = 146270.
  Std. dev. of dep. var. = 79317.2
Sum of squared residuals = .590213E+09
   Variance of residuals = .147553E+08
Std. error of regression = 3841.27
               R-squared = .997766
      Adjusted R-squared = .997655
            LM het. test = .174943 [.676]
           Durbin-Watson = .253234 [.000,.000]
```

```
                                 Equation   5
                                 ============
                  FIRST-ORDER SERIAL CORRELATION OF THE ERROR
                  Objective function:  Exact ML (keep first obs.)

Dependent variable: RCONS
Current sample:  1955 to 1997
Number of observations:  43

       Mean of dep. var. = 146270.                 R-squared = .999470
  Std. dev. of dep. var. = 79317.2       Adjusted R-squared = .999443
Sum of squared residuals = .140391E+09        Durbin-Watson = 1.43657
   Variance of residuals = .350977E+07       Schwarz B.I.C. = 389.449
Std. error of regression = 1873.44           Log likelihood = -383.807

                          Standard
Parameter  Estimate         Error         t-statistic    P-value
C          12034.8        3346.47         3.59628        [.000]
RRYD       .140723        .282614E-02     49.7933        [.000]
RHO        .876924        .068199         12.8583        [.000]
```

1. `Equation 1` vs. `Equation 3` (Test of Serial Correlation)

   `Equation 1` is:

$$\texttt{RCONS}_t = \beta_1 + \beta_2 \texttt{RYD}_t + u_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

   `Equation 3` is:

$$\texttt{RCONS}_t = \beta_1 + \beta_2 \texttt{RYD}_t + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

   The null hypothesis is $H_0 : \ \rho = 0$

Restricted MLE $\Longrightarrow$ `Equation 1`

Unrestricted MLE $\Longrightarrow$ `Equation 3`

The log-likelihood function of `Equation 3` is:

$$\log L(\beta, \sigma_\epsilon^2, \rho) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\epsilon^2) + \frac{1}{2}\log(1-\rho^2) \\ - \frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{n}(\mathtt{RCONS}_t^* - \beta_1\mathtt{CONST}_t^* - \beta_2\mathtt{RYD}_t^*)^2,$$

where

$$\mathtt{RCONS}_t^* = \begin{cases} \sqrt{1-\rho^2}\,\mathtt{RCONS}_t, & \text{for } t = 1, \\ \mathtt{RCONS}_t - \rho\mathtt{RCONS}_{t-1}, & \text{for } t = 2, 3, \cdots, n, \end{cases}$$

$$\mathtt{CONST}_t^* = \begin{cases} \sqrt{1-\rho^2}, & \text{for } t = 1, \\ 1-\rho, & \text{for } t = 2, 3, \cdots, n, \end{cases}$$

$$\mathtt{RYD}_t^* = \begin{cases} \sqrt{1-\rho^2}\,\mathtt{RYD}_t, & \text{for } t = 1, \\ \mathtt{RYD}_t - \rho\mathtt{RYD}_{t-1}, & \text{for } t = 2, 3, \cdots, n. \end{cases}$$

- MLE with the restriction $\rho = 0$ (`Equation 1`) solves:

$$\max_{\beta,\sigma_\epsilon^2} \log L(\beta, \sigma_\epsilon^2, 0)$$

Restricted MLE $\Longrightarrow \tilde{\beta}, \tilde{\sigma}_\epsilon^2$

`Log of likelihood function = -431.289`

- MLE without the restriction $\rho = 0$ (`Equation 3`) solves:

$$\max_{\beta,\sigma_\epsilon^2,\rho} \log L(\beta, \sigma_\epsilon^2, \rho)$$

Unrestricted MLE $\Longrightarrow \hat{\beta}, \hat{\sigma}_\epsilon^2, \hat{\rho}$

`Log of likelihood function = -385.419`

The likelihood ratio test statistic is:

$$
\begin{aligned}
-2\log(\lambda) &= -2\log\Big(\frac{L(\tilde\beta, \tilde\sigma_\epsilon^2, 0)}{L(\hat\beta, \hat\sigma_\epsilon^2, \hat\rho)}\Big) = -2\Big(\log L(\tilde\beta, \tilde\sigma_\epsilon^2, 0) - \log L(\hat\beta, \hat\sigma_\epsilon^2, \hat\rho)\Big) \\
&= -2\Big(-431.289 - (-385.419)\Big) = 91.74.
\end{aligned}
$$

The asymptotic distribution is given by:

$$
-2\log(\lambda) \sim \chi^2(G),
$$

where G is the number of the restrictions, i.e., $G = 1$ in this case.

The 1% upper probability point of $\chi^2(1)$ is 6.635.

$$
91.74 > 6.635
$$

Therefore, $H_0:\ \rho = 0$ is rejected.

There is serial correlation in the error term.

2. `Equation 1` (Test of Serial Correlation $\longrightarrow$ Lagrange Multiplier Test)

`Equation 2` is:

$$\texttt{@RES}_t = \rho \texttt{@RES}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2),$$

where $\texttt{@RES}_t = \texttt{RCONS}_t - \hat{\beta}_1 - \hat{\beta}_2 \texttt{RYD}_t$, and $\hat{\beta}_1$ and $\hat{\beta}_2$ are OLSEs.

The null hypothesis is $H_0: \ \rho = 0$

```
@RES(-1)   .950693        .053301        17.8362        [.000]
```

Therefore, the Lagrange multiplier test statistic is $17.8362^2 = 318.13 > 6.635$.

$H_0: \ \rho = 0$ is rejected.

3. `Equation 3` (Test of Serial Correlation $\longrightarrow$ Wald Test)

`Equation 3` is:

$$\texttt{RCONS}_t = \beta_1 + \beta_2 \texttt{RYD}_t + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma^2_\epsilon)$$

The null hypothesis is $H_0 : \rho = 0$

```
RHO          .945025         .045843         20.6143          [.000]
```

The Wald test statistics is $20.6143^2 = 424.95$, which is compared with $\chi^2(1)$.

4. `Equation 1` vs. `NONLINEAR LEAST SQUARES` (Choice of Functional Form – linear):

`NONLINEAR LEAST SQUARES` estimates:

$$\texttt{RCONS}_t = a1 + a2\frac{\texttt{RYD}_t^{a3} - 1}{a3} + u_t.$$

When $a3 = 1$, we have:

$$\texttt{RCONS}_t = (a1 - a2) + a2\texttt{RYD}_t + u_t,$$

which is equivalent to `Equation 1`.

The null hypothesis is $H_0 :\ a3 = 1$, where $G = 1$.

- MLE with $a3 = 1$ MLE (`Equation 1`)

  `Log of likelihood function = -431.289`

- MLE without $a3 = 1$ (`NONLINEAR LEAST SQUARES`)

  `Log of likelihood function = -414.362`

The likelihood ratio test statistic is given by:

$$-2\log(\lambda) = -2\big(-431.289 - (-414.362)\big) = 33.854.$$

The 1% upper probability point of $\chi^2(1)$ is 6.635.

$$33.854 > 6.635$$

$H_0 : \ a3 = 1$ is rejected.

Therefore, the functional form of the regression model is not linear.

5. `Equation 4` vs. `NONLINEAR LEAST SQUARES` (Choice of Functional Form – log-linear):

In `NONLINEAR LEAST SQUARES`, i.e.,

$$\mathtt{RCONS}_t = a1 + a2\frac{\mathtt{RYD}_t^{a3} - 1}{a3} + u_t,$$

if $a3 = 0$, we have:

$$\mathtt{RCONS}_t = a1 + a2\log(\mathtt{RYD}_t) + u_t,$$

which is equivalent to `Equation 3`.

The null hypothesis is $H_0 : \ a3 = 0$, where $G = 1$.

- MLE with $a3 = 0$ (`Equation 3`)

  `Log of likelihood function = -495.418`

- MLE without $a3 = 0$ (`NONLINEAR LEAST SQUARES`)

  `Log of likelihood function = -414.362`

The likelihood ratio test statistic is:

$$-2\log(\lambda) = -2\Big(-495.418 - (-414.362)\Big) = 162.112 > 6.635.$$

Therefore, $H_0 : \ a3 = 0$ is rejected.

As a result, the functional form of the regression model is not log-linear, either.

6. `Equation 1` vs. `Equation 5` (Simultaneous Test of Serial Correlation and Linear Function):

`Equation 5` is:

$$\mathtt{RCONS}_t = a1 + a2\frac{\mathtt{RYD}_t^{a3} - 1}{a3} + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

The null hypothesis is $H_0:\ a3 = 1,\ \rho = 0$

Restricted MLE $\Longrightarrow$ `Equation 1`

Unrestricted MLE $\Longrightarrow$ `Equation 4`

Remark: In Lines 14–16 of `PROGRAM`, we have estimated `Equation 4`, given $a3 = 0.00, 0.01, 0.02, \cdots$.

As a result, $a3 = 1.15$ gives us the maximum log-likelihood.

The likelihood ratio test statistic is:

$$-2\log(\lambda) = -2\Big(-431.289 - (-383.807)\Big) = 94.964.$$

$-2\log(\lambda) \sim \chi^2(2)$ in this case.

The 1% upper probability point of $\chi^2(2)$ is 9.210.

$$94.964 > 9.210$$

$H_0:\ a3 = 1, \rho = 0$ is rejected.

`Equation 3` vs. `Equation 5` vs. (Taking into account serially correlated errors, Choice of Functional Form – linear):

The null hypothesis is $H_0:\ a3 = 1,\ \rho = 0$

From `Equation 3`,

```
Log likelihood = -385.419
```

From `Equation 5`,

```
Log likelihood = -383.807
```

$$2(-383.807 - (-385.419)) = 3.224 < 6.635.$$

$H_0:\ a3 = 1$ is not rejected, given $\rho \neq 0$.

Thus, if serial correlation is taken into account, the regression model is linear.

14　Unit Root (単位根) and Cointegration (共和分)

Textbooks

・J.D. Hamilton (1994) *Econometric Analysis*

沖本・井上訳 (2006)『時系列解析 (上・下)』

・A.C. Harvey (1981) *Time Series Models*

国友・山本訳 (1985)『時系列モデル入門』

・沖本竜義 (2010)『経済・ファイナンスデータの計量時系列分析』

14.1 Unit Root (単位根) Test (Dickey-Fuller (DF) Test)

1. Why is a unit root problem important?

   (a) Economic variables increase over time in general.

      One of the assumptions of OLS is stationarity on y_t and x_t.

      This assumption implies that $\dfrac{1}{T}X'X$ converges to a fixed matrix as T is large.

      That is, asymptotic normality of OLS estimator does not hold.

   (b) In nonstationary time series, the unit root is the most important.

      In the case of unit root, OLSE of the first-order autoregressive coefficient is consistent.

      OLSE is $\sqrt{T}$-consistent in the case of stationary AR(1) process, but OLSE is T-consistent in the case of nonstationay AR(1) process.

(c) A lot of economic variables increase over time.

It is important to check an economic variable is trend stationary (i.e., $y_t = a_0 + a_1 t + \epsilon_t$) or difference stationary (i.e., $y_t = b_0 + y_{t-1} + \epsilon_t$).

Consider k-step ahead prediction for both cases.

$$\text{(Trend Stationarity)} \qquad y_{t+k|t} = a_0 + a_1(t+k)$$

$$\text{(Difference Stationarity)} \qquad y_{t+k|t} = b_0 k + y_t$$

2. **The Case of $|\phi_1| < 1$:**

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \qquad y_0 = 0, \qquad t = 1, \cdots, T$$

Then, OLSE of ϕ_1 is:

$$\hat{\phi}_1 = \frac{\displaystyle\sum_{t=1}^{T} y_{t-1} y_t}{\displaystyle\sum_{t=1}^{T} y_{t-1}^2}.$$

In the case of $|\phi_1| < 1$,

$$\hat{\phi}_1 = \phi_1 + \frac{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow \phi_1 + \frac{\mathrm{E}(y_{t-1}\epsilon_t)}{\mathrm{E}(y_{t-1}^2)} = \phi_1.$$

Note as follows:

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow \mathrm{E}(y_{t-1}\epsilon_t) = 0.$$

By the central limit theorem,

$$\frac{\overline{y\epsilon} - \mathrm{E}(\overline{y\epsilon})}{\sqrt{\mathrm{V}(\overline{y\epsilon})}} \longrightarrow N(0,1)$$

where

$$\overline{y\epsilon} = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t.$$

$$
\begin{aligned}
\mathrm{E}(\overline{y\epsilon}) &= 0, \\
\mathrm{V}(\overline{y\epsilon}) &= \mathrm{V}\big(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t\big) = \mathrm{E}\Big(\big(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t\big)^2\Big) \\
&= \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T}\sum_{s=1}^{T} y_{t-1}y_{s-1}\epsilon_t\epsilon_s\Big) = \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T} y_{t-1}^2\epsilon_t^2\Big) = \frac{1}{T}\sigma_\epsilon^2\gamma(0),
\end{aligned}
$$

where $\gamma(\tau) = \mathrm{Cov}(y_t,\ y_{t-\tau}) = \mathrm{E}((y_t - \mathrm{E}(y_t))(y_{t-\tau} - \mathrm{E}(y_{t-\tau})))$. Therefore,

$$
\frac{\overline{y\epsilon}}{\sqrt{\sigma_\epsilon^2\gamma(0)/T}} = \frac{1}{\sigma_\epsilon\sqrt{\gamma(0)}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0,1),
$$

which is rewritten as:

$$
\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0, \sigma_\epsilon^2\gamma(0)).
$$

Using $\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \mathrm{E}(y_{t-1}^2) = \gamma(0)$, we have the following asymptotic distribution:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{\dfrac{1}{\sqrt{T}}\displaystyle\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\left(0, \frac{\sigma_\epsilon^2}{\gamma(0)}\right) = N\left(0, 1 - \phi_1^2\right).$$

Note that $\gamma(0) = \dfrac{\sigma_\epsilon^2}{1 - \phi_1^2}$.

3. In the case of $\phi_1 = 1$, as expected, we have:

$$\sqrt{T}(\hat{\phi}_1 - 1) \longrightarrow 0.$$

That is, $\hat{\phi}_1$ has the distribution which converges in probability to $\phi_1 = 1$ (i.e., degenerated distribution).

Is this true?

4. **The Case of $\boldsymbol{\phi_1 = 1}$:** $\Longrightarrow$ Random Walk Process

$y_t = y_{t-1} + \epsilon_t$ with $y_0 = 0$ is written as:

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \cdots + \epsilon_1.$$

Therefore, we can obtain:

$$y_t \sim N(0, \sigma_\epsilon^2 t).$$

The variance of y_t depends on time t. $\Longrightarrow$ y_t is nonstationary.

5. Remember that $\hat{\phi}_1 = \phi_1 + \dfrac{\sum y_{t-1}\epsilon_t}{\sum y_{t-1}^2}$.

   (a) First, consider the numerator $\sum y_{t-1}\epsilon_t$.

   We have $y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2$.

Therefore, we obtain:

$$y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2).$$

Taking into account $y_0 = 0$, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2.$$

Divided by $\sigma_\epsilon^2 T$ on both sides, we have the following:

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2.$$

From $y_t \sim N(0, \sigma_\epsilon^2 t)$, we obtain the following result:

$$\left(\frac{y_T}{\sigma_\epsilon\sqrt{T}}\right)^2 \sim \chi^2(1).$$