Week 5:

Transformation of time series, Tests of randomness

Transformations of Time Series

Aim: achieve normality and constant variance

- most of the methods assume that

$$Y_t = Tr_t + S_t + E_t, \quad \mathsf{E} E_t = 0, \quad \mathsf{Var}\, E_t = \sigma^2 = const$$

  and optimality for normal E_t
- prediction intervals: normality

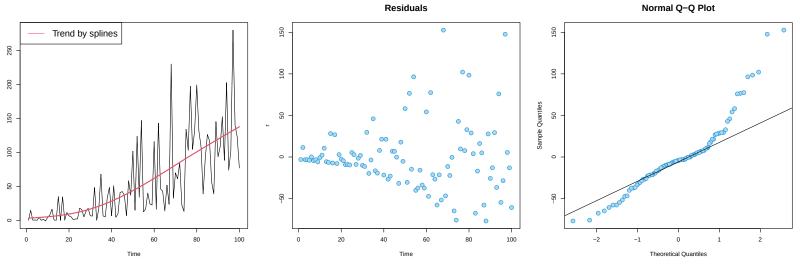
Transformations of Time Series

Aim: achieve normality and constant variance

- most of the methods assume that

$$Y_t = Tr_t + S_t + E_t, \quad \mathrm{E}\, E_t = 0, \quad \mathrm{Var}\, E_t = \sigma^2 = const$$

  and optimality for normal E_t
- prediction intervals: normality



$\rightsquigarrow$ find transformation g such that $g(Y_t)$ satisfies the conditions

Box–Cox

$$g_\lambda(y) = \begin{cases} \frac{(y+c)^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y+c), & \lambda = 0. \end{cases}$$

and use

$$Y_t^\lambda = g_\lambda(Y_t)$$

for a suitable λ and a suitable c

Box–Cox

$$g_\lambda(y) = \begin{cases} \frac{(y+c)^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y+c), & \lambda = 0. \end{cases}$$

and use

$$Y_t^\lambda = g_\lambda(Y_t)$$

for a suitable λ and a suitable c

- g_λ continuous in λ

Box–Cox

$$g_\lambda(y) = \begin{cases} \frac{(y+c)^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y+c), & \lambda = 0. \end{cases}$$

and use

$$Y_t^\lambda = g_\lambda(Y_t)$$

for a suitable λ and a suitable c

- g_λ continuous in λ
- if you intend to fit a regression model $\rightsquigarrow$ it suffices to take $(Y_t + c)^\lambda$ for $\lambda > 0$

Box–Cox

$$g_\lambda(y) = \begin{cases} \frac{(y+c)^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y+c), & \lambda = 0. \end{cases}$$

and use

$$Y_t^\lambda = g_\lambda(Y_t)$$

for a suitable λ and a suitable c

- g_λ continuous in λ
- if you intend to fit a regression model $\rightsquigarrow$ it suffices to take $(Y_t + c)^\lambda$ for $\lambda > 0$

Parameters:

- $c > 0$ such that $Y_t + c > 0$

Box–Cox

$$g_\lambda(y) = \begin{cases} \frac{(y+c)^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y+c), & \lambda = 0. \end{cases}$$

and use

$$Y_t^\lambda = g_\lambda(Y_t)$$

for a suitable λ and a suitable c

- g_λ continuous in λ
- if you intend to fit a regression model $\rightsquigarrow$ it suffices to take $(Y_t + c)^\lambda$ for $\lambda > 0$

Parameters:

- $c > 0$ such that $Y_t + c > 0$
- How to find λ?
  - profile maximum likelihood
  - approximate methods

Box–Cox profile likelihood

Assume that there exists λ such that $g_\lambda(Y_t)$ are independent for $t = 1, \ldots, T$ and

$$g_\lambda(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \sim \mathrm{N}(\mu_t, \sigma^2)$$

where either $\mu_t = Tr_t$ or $\mu_t = Tr_t + S_t$ modelled by a regression model.

Box–Cox profile likelihood

Assume that there exists λ such that $g_\lambda(Y_t)$ are independent for $t = 1, \ldots, T$ and

$$g_\lambda(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \sim \mathrm{N}(\mu_t, \sigma^2)$$

where either $\mu_t = Tr_t$ or $\mu_t = Tr_t + S_t$ modelled by a regression model.

$\hookrightarrow$ derive the density of Y_t (use the transformation theorem)

$$\log f_{Y_t}(y) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(g_\lambda(y) - \mu_t)^2 + (\lambda - 1)\log y$$

Box–Cox profile likelihood

Assume that there exists λ such that $g_\lambda(Y_t)$ are independent for $t = 1, \dots, T$ and

$$g_\lambda(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \sim \mathrm{N}(\mu_t, \sigma^2)$$

where either $\mu_t = Tr_t$ or $\mu_t = Tr_t + S_t$ modelled by a regression model.

$\hookrightarrow$ derive the density of Y_t (use the transformation theorem)

$$\log f_{Y_t}(y) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(g_\lambda(y) - \mu_t)^2 + (\lambda - 1)\log y$$

$\hookrightarrow$ independence $\rightsquigarrow$ log-likelihood

$$l(\lambda, \beta, \sigma^2) = const - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}(g_\lambda(Y_t) - \mu_t)^2 + (\lambda - 1)\sum_{t=1}^{n}\log Y_t$$

Box–Cox profile likelihood

Assume that there exists λ such that $g_\lambda(Y_t)$ are independent for $t = 1, \ldots, T$ and

$$g_\lambda(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \sim \mathrm{N}(\mu_t, \sigma^2)$$

where either $\mu_t = Tr_t$ or $\mu_t = Tr_t + S_t$ modelled by a regression model.

$\hookrightarrow$ derive the density of Y_t (use the transformation theorem)

$$\log f_{Y_t}(y) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(g_\lambda(y) - \mu_t)^2 + (\lambda - 1)\log y$$

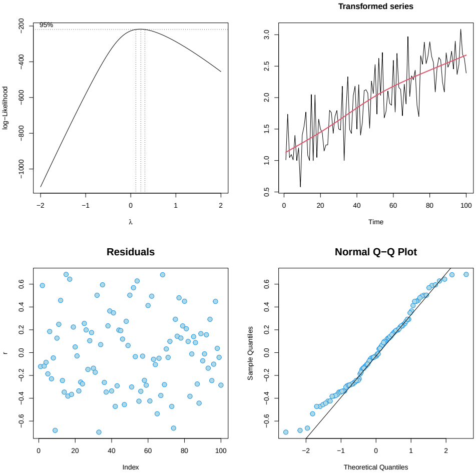
$\hookrightarrow$ independence $\rightsquigarrow$ log-likelihood

$$l(\lambda, \boldsymbol{\beta}, \sigma^2) = const - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n}(g_\lambda(Y_t) - \mu_t)^2 + (\lambda - 1)\sum_{t=1}^{n}\log Y_t$$

$\hookrightarrow$ profile likelihood

$$l(\lambda) = \max_{\boldsymbol{\beta}, \sigma^2} l(\lambda, \boldsymbol{\beta}, \sigma^2) = const - \frac{n}{2}\log SSe(\lambda) + (\lambda - 1)\sum_{t=1}^{n}\log Y_t$$

Box–Cox profile likelihood



- min $Y_t = -0.93 \rightsquigarrow c = 1$, MLE $\rightsquigarrow \widehat{\lambda} = 0.2 \rightsquigarrow g(Y_t) = (Y_t + 1)^{1/5}$
- analyze $\{g(Y_t)\} \rightsquigarrow$ prediction interval for $g(Y_{n+1}) \rightsquigarrow$ prediction interval for Y_{n+1}

Approximate methods for λ

Let Y be a random variable. Taylor expansion of g:

$$g(Y) \approx g(\mathrm{E}\,Y) + g'(\mathrm{E}\,Y)(Y - \mathrm{E}\,Y)$$

so

$$\mathrm{Var}\,g(Y) \approx [g'(\mathrm{E}\,Y)]^2 \mathrm{Var}\,Y \stackrel{!}{=} k^2 = const$$

Approximate methods for λ

Let Y be a random variable. Taylor expansion of g:

$$g(Y) \approx g(\mathrm{E}Y) + g'(\mathrm{E}Y)(Y - \mathrm{E}Y)$$

so

$$\mathrm{Var}\, g(Y) \approx [g'(\mathrm{E}Y)]^2 \mathrm{Var}\, Y \stackrel{!}{=} k^2 = const$$

For g_λ:

$$g'_\lambda(y) = y^{\lambda-1},$$

so

$$\begin{aligned}(\mathrm{E}Y)^{2(\lambda-1)}\mathrm{Var}\, Y &\approx k^2 \\ \sqrt{\mathrm{Var}\, Y} &\approx k(\mathrm{E}Y)^{1-\lambda}\end{aligned}$$

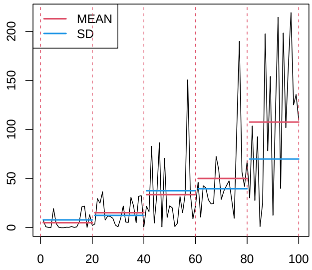
And similar relationship should be observed for the sample counterparts (SD and MEAN)

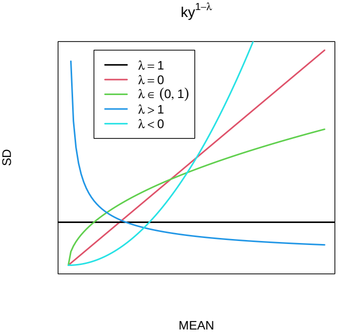
Approximate methods for λ (cont.)

1. divide data into J segments of the same length
2. compute $s_Y(j)$, $\overline{Y}(j)$ for $j = 1, \ldots, J$ from $Y_t + c$
3. plot $(\overline{Y}(j), s_Y(j))$ and try to determine approximate λ from

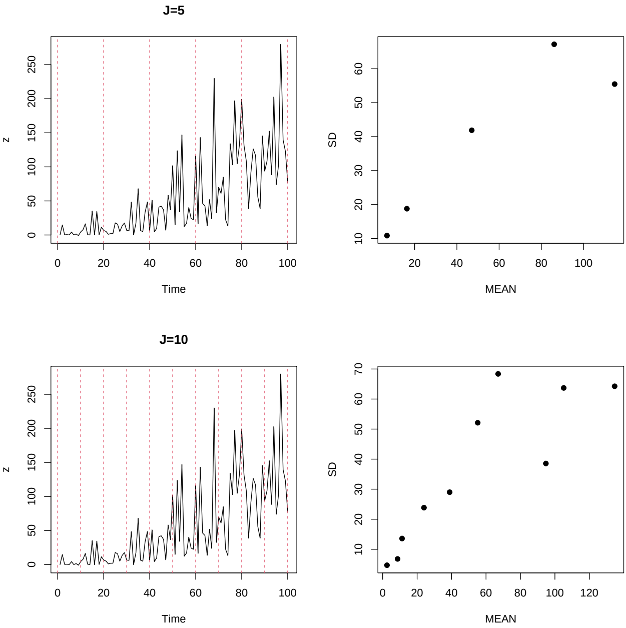
$$s_Y(j) \approx k \cdot (\overline{Y}(j))^{1-\lambda}$$

   for some $k > 0$
4. typically one takes $\widehat{\lambda} \in \{0, 1, 1/2, -1/2\}$





Example



J=5
z
0
50
100
150
200
250
0
20
40
60
80
100
Time
SD
10
20
30
40
50
60
20
40
60
80
100
MEAN
J=10
z
0
50
100
150
200
250
0
20
40
60
80
100
Time
SD
10
20
30
40
50
60
70
0
20
40
60
80
100
120
MEAN

Approximate methods for λ (cont.)

$$s_Y(j) \approx k \cdot (\overline{Y}(j))^{1-\lambda}$$
$$\log[s_Y(j)] \approx \log k + (1-\lambda)\log[\overline{Y}(j)]$$

$\rightsquigarrow$ plot points

$$\left(\log[\overline{Y}(j)], \log[s_Y(j)]\right)$$

and $1 - \lambda$ is the regression slope

Approximate methods for λ (cont.)

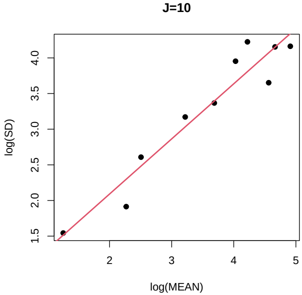
$$s_Y(j) \approx k \cdot (\overline{Y}(j))^{1-\lambda}$$

$$\log[s_Y(j)] \approx \log k + (1-\lambda)\log[\overline{Y}(j)]$$

$\rightsquigarrow$ plot points

$$\left(\log[\overline{Y}(j)], \log[s_Y(j)]\right)$$

and $1-\lambda$ is the regression slope



$\widehat{\lambda} = 1 - 0.77 = 0.23$

Pros and cons of Box-Cox

Pros +	Cons −

Pros and cons of Box-Cox

Pros +

- prediction intervals with exact coverage
- exact statistical tests (if other assumptions satisfied)
- some procedures optimal under normality

Cons −

Pros and cons of Box-Cox

Pros $+$

- prediction intervals with exact coverage
- exact statistical tests (if other assumptions satisfied)
- some procedures optimal under normality

Cons $-$

- point prediction typically biased
- except special cases $\lambda = 0, 1$ no interpretation for the parameters (slope etc) in terms of Y_t

Pros and cons of Box-Cox

Pros +

- prediction intervals with exact coverage
- exact statistical tests (if other assumptions satisfied)
- some procedures optimal under normality

Cons −

- point prediction typically biased
- except special cases $\lambda = 0, 1$ no interpretation for the parameters (slope etc) in terms of Y_t

Most popular transformations

$\hookrightarrow$ $\lambda = 1$: no transformation

$\hookrightarrow$ $\lambda = 0$: $\log$ transformation

Tests of randomness

Tests of randomness

$$H_0 : Y_t \sim \text{iid}$$

against

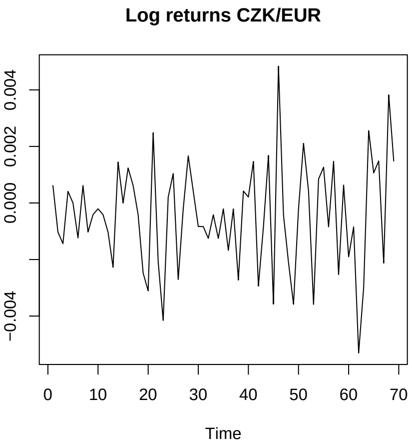
$$H_1 : \text{ either } Y_t \text{ not independent, or } Y_t \text{ not id}$$

Why?

- plot: no presence of any systematic component
- apply this on $\widehat{E}_t = Y_t - \widehat{Tr}_t - \widehat{S}_t - \widehat{C}_t$

H_1 very broad $\rightsquigarrow$ various tests

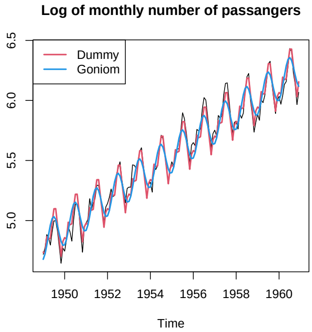
Example I

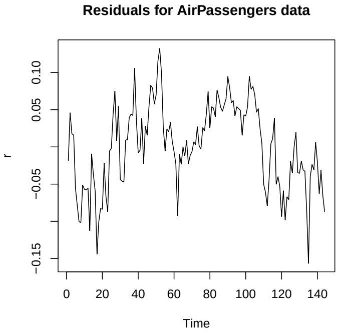


Log returns CZK/EUR
0.004
0.002
0.000
−0.004
0
10
20
30
40
50
60
70
Time

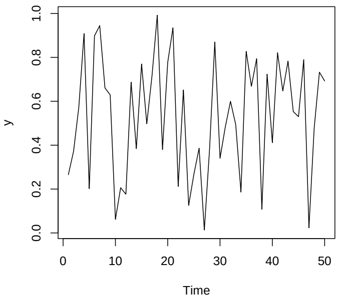
Example II

Air Passengers data: $Y_t = \beta_1 t + \sum_{j=1}^{12} \gamma_j \cdot \mathrm{I}(\mathrm{month}_t = j) + E_t$





Example III: Is my pseudo random generator good?



Setting

Data $Y_1, \ldots, Y_n$

For simplicity: $Y_t \neq Y_{t=1}$ for all t (no ties allowed)

(Is it restrictive?)

Setting

Data $Y_1, \ldots, Y_n$

For simplicity: $Y_t \neq Y_{t=1}$ for all t (no ties allowed)

(Is it restrictive?)

Discussed tests:

1. based on signs of differences
2. based on turning points
3. based on runs (median test)
4. based on Kendall's tau
5. based on Spearman's rho
6. tools based on ACF

Discussion: Usefulness of such tests?

1. Test Based on Signs of Differences

$$V_t = \begin{cases} 1 & Y_t < Y_{t+1} \\ 0 & Y_t > Y_{t+1} \end{cases}$$

Then

$$K_n = \sum_{t=1}^{n-1} V_t$$

is the number of points of growth.

Idea of the test: Reject if K_n differs ”too much” from its expectation under H_0 (i.e. K_n ”too extreme”)

1. Test Based on Signs of Differences

$$V_t = \begin{cases} 1 & Y_t < Y_{t+1} \\ 0 & Y_t > Y_{t+1} \end{cases}$$

Then

$$K_n = \sum_{t=1}^{n-1} V_t$$

is the number of points of growth.

Idea of the test: Reject if K_n differs "too much" from its expectation under H_0 (i.e. K_n "too extreme")

$\hookrightarrow$ either exact or asymptotic distribution of K_n

1. Test Based on Signs of Differences

$$V_t = \begin{cases} 1 & Y_t < Y_{t+1} \\ 0 & Y_t > Y_{t+1} \end{cases}$$

Then

$$K_n = \sum_{t=1}^{n-1} V_t$$

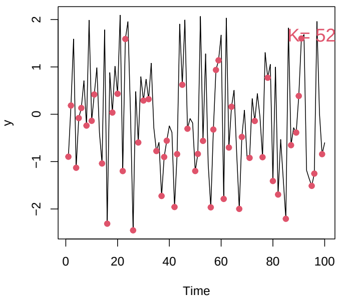
is the number of points of growth.

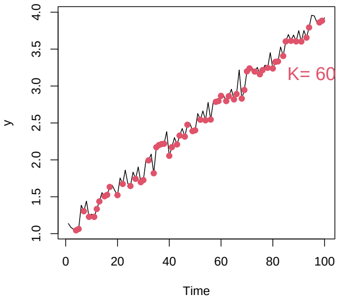
Idea of the test: Reject if K_n differs ”too much” from its expectation under H_0 (i.e. K_n ”too extreme”)

- $\hookrightarrow$ either exact or asymptotic distribution of K_n
- $\hookrightarrow$ K_n is a sum of (dependent) variables $\rightsquigarrow$ CLT might give us asymptotics

Illustration

$$V_t = \begin{cases} 1 & Y_t < Y_{t+1} \\ 0 & Y_t > Y_{t+1} \end{cases}$$





Moments of K_n

$$\mathrm{E}K_n = \mathrm{E}\sum_{t=1}^{n-1} V_t = \sum_{t=1}^{n-1} \mathrm{E}V_t = \frac{n-1}{2}$$

because

$$V_t = \mathrm{I}[Y_t < Y_{t-1}] \overset{H_0:iid}{\sim} \mathrm{Alt}(1/2).$$

Moments of K_n

$$\mathrm{E} K_n = \mathrm{E} \sum_{t=1}^{n-1} V_t = \sum_{t=1}^{n-1} \mathrm{E} V_t = \frac{n-1}{2}$$

because

$$V_t = \mathrm{I}[Y_t < Y_{t-1}] \overset{H_0:iid}{\sim} \mathrm{Alt}(1/2).$$

$$\mathrm{Var}\, K_n = \mathrm{Var} \left(\sum_{t=1}^{n-1} V_t \right) = \sum_{t=1}^{n-1} \mathrm{Var}\, V_t + 2 \sum \sum_{s<t} \mathrm{Cov}\,(V_s, V_t)$$

Moments of K_n

$$\mathrm{E}K_n = \mathrm{E}\sum_{t=1}^{n-1} V_t = \sum_{t=1}^{n-1} \mathrm{E}V_t = \frac{n-1}{2}$$

because

$$V_t = \mathrm{I}[Y_t < Y_{t-1}] \overset{H_0:iid}{\sim} \mathrm{Alt}(1/2).$$

$$\mathrm{Var}\, K_n = \mathrm{Var}\left(\sum_{t=1}^{n-1} V_t\right) = \sum_{t=1}^{n-1} \mathrm{Var}\, V_t + 2\sum\sum_{s<t} \mathrm{Cov}\,(V_s, V_t)$$

If $s+1 < t$, then V_s and V_t independent $\rightsquigarrow$ $\mathrm{Cov}\,(V_s, V_t) = 0$.
If $s+1 = t$, then

$$\mathrm{Cov}\,(V_s, V_t) = \mathrm{EI}[Y_s < Y_{s+1} < Y_{s+2}] - \frac{1}{4} \overset{H_0:iid}{=} \frac{1}{6} - \frac{1}{4} = -\frac{1}{12},$$

so

$$\mathrm{Var}\, K_n = \frac{n-1}{4} - 2\frac{n-2}{12} = \frac{n+1}{12}.$$

Asymptotic distribution

It holds that

$$\frac{K_n - \mathrm{E}K_n}{\sqrt{\operatorname{Var} K_n}} = \frac{K_n - \frac{n-1}{2}}{\sqrt{\frac{n+1}{12}}} \xrightarrow{D} \mathrm{N}(0,1).$$

$\hookrightarrow$ Justification: CLT for *m*-dependent processes.

$\hookrightarrow$ Equivalent versions of the test statistic

Test:

$$\text{If} \quad \frac{\left|K_n - \frac{n-1}{2}\right|}{\sqrt{\frac{n+1}{12}}} > u_{1-\alpha/2} \quad \Rightarrow \text{ reject } H_0$$

2. Test Based on Turning Points

$$V_t = \begin{cases} 1 & Y_{t-1} < Y_t, Y_t > Y_{t+1} \text{ or } Y_{t-1} > Y_t, Y_t < Y_{t+1}, \\ 0 & Y_{t-1} < Y_t < Y_{t+1} \text{ or } Y_{t-1} > Y_t > Y_{t+1} \end{cases}$$

and

$$R_n = \sum_{t=2}^{n-1} V_t$$

the total number of upper and lower turning points

Idea of the test: Reject if R_n differs "too much" from its expectation under H_0 (i.e. R_n "too extreme")

- $\hookrightarrow$ tables for exact distribution exist
- $\hookrightarrow$ R_n asymptotically normal (again use CLT for m-dependent)
- $\hookrightarrow$ we need to computed $\mathrm{E}R_n$, Var R_n

Moments of R_n

Now

$$V_t = I[Y_{t-1} < Y_t, Y_t > Y_{t+1} \text{ or } Y_{t-1} > Y_t, Y_t < Y_{t+1}] \overset{H_0:iid}{\sim} \text{Alt}(2/3),$$

so

$$\mathrm{E} R_n = \sum_{t=2}^{n-1} \mathrm{E} V_t = \frac{2(n-2)}{3}.$$

Similar computations as for K_n give

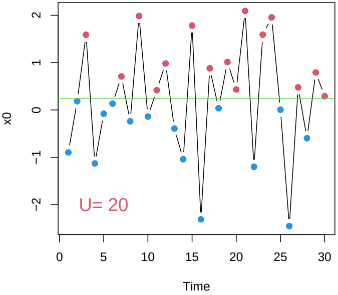
$$\text{Var } R_n = \frac{16n - 29}{90}.$$

Test:

$$\text{If} \quad \frac{|R_n - \mathrm{E} R_n|}{\sqrt{\text{Var } R_n}} > u_{1-\alpha/2} \quad \Rightarrow \text{ reject } H_0$$

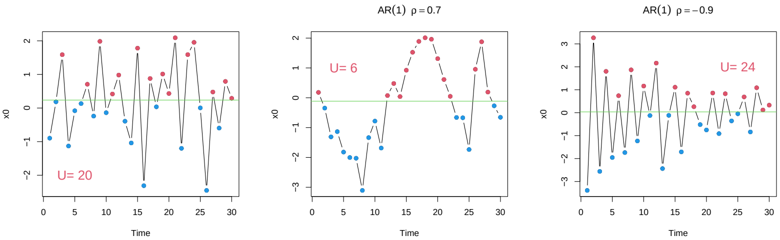
3. Test Based on Runs (Median Test)

- M median of $Y_1, \ldots, Y_n$
- U_n is number of runs



Idea of the test: Reject if U_n ”too extreme”

Illustration



Asymptotic distribution

It is possible to show

$$\mathsf{E}\, U_n = m + 1, \quad \operatorname{Var} U_n = \frac{m(m-1)}{2m-1},$$

where $m = \sum_{t=1}^{n} \mathrm{I}[Y_t > M]$ ($m = n/2$ if n even), and

$$\frac{U_n - \mathsf{E}\, U_n}{\sqrt{\operatorname{Var} U_n}} \xrightarrow{D} \mathrm{N}(0, 1).$$

Reject if

$$\frac{|U_n - \mathsf{E}\, U_n|}{\sqrt{\operatorname{Var} U_n}} > u_{1-\alpha/2}$$

Simulations

$$
\begin{aligned}
\text{IID: } Y_t &\sim \text{iid } \mathrm{N}(0,1), \\
\text{AR: } Y_t &= 0.6 \cdot Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \text{ iid } \mathrm{N}(0,1), \\
\text{LT: } Y_t &= \frac{3}{n} t + \varepsilon_t, \quad \varepsilon_t \text{ iid } \mathrm{N}(0,1), \\
\text{RW: } Y_t &= \sum_{i=1}^{t} \varepsilon_i, \quad \varepsilon_t \text{ iid } \mathrm{N}(0,0.5^2),
\end{aligned}
$$

$N = 1\,000$ replications $\rightsquigarrow$ percentage of rejection

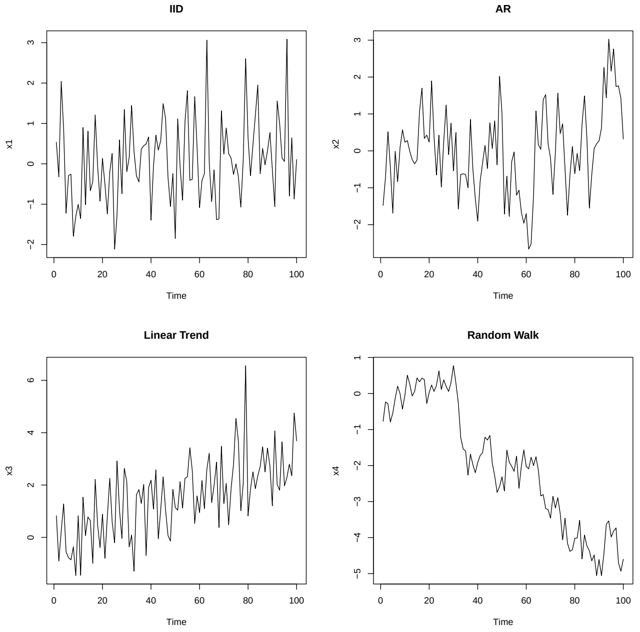
Simulations

$$\begin{aligned}
\text{IID: } Y_t &\sim \text{iid N}(0,1),\\
\text{AR: } Y_t &= 0.6 \cdot Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \text{ iid N}(0,1),\\
\text{LT: } Y_t &= \frac{3}{n} t + \varepsilon_t, \quad \varepsilon_t \text{ iid N}(0,1),\\
\text{RW: } Y_t &= \sum_{i=1}^{t} \varepsilon_i, \quad \varepsilon_t \text{ iid N}(0, 0.5^2),
\end{aligned}$$

$N = 1\,000$ replications $\rightsquigarrow$ percentage of rejection

	K_n			R_n			U_n		
n	50	100	200	50	100	200	50	100	200
IID	5	4	5	6	5	6	6	5	6
AR	6	5	6	43	67	91	79	96	100
LT	7	6	6	6	6	5	58	85	99
RW	24	25	27	78	95	100	100	100	100

$\rightsquigarrow$ back to the critics of the tests....



IID
x1
Time
AR
x2
Time
Linear Trend
x3
Time
Random Walk
x4
Time

Kendall's τ and Spearman's ρ

Consider iid random vectors

$$\begin{pmatrix} U_1 \\ V_1 \end{pmatrix}, \dots \begin{pmatrix} U_n \\ V_n \end{pmatrix}$$

- Pearson's correlation $\rho = \text{cor}(U_i, V_i)$ estimated by

$$\widehat{\rho} = \frac{\sum_{i=1}^n (U_i - \overline{U}_n)(V_i - \overline{V}_n)}{\sqrt{\sum_{i=1}^n (U_i - \overline{U}_n)^2}\sqrt{\sum_{i=1}^n (V_i - \overline{V}_n)^2}}$$

- Kendall's τ

$$\tau = \text{P}(U_i < V_i) - \text{P}(U_i > V_i)$$

  estimated by

$$\widehat{\tau} = \frac{2}{n(n-1)} \sum_{i<j} \text{sgn}(U_i - U_j)\text{sgn}(V_i - V_j)$$

- Spearman's ρ

$$\rho_S = \text{cor}\left(F_U(U_i), F_V(V_i)\right)$$

  estimated by

$$\widehat{\rho_S} = \frac{\sum_{i=1}^n (R_i - \overline{R}_n)(S_i - \overline{S}_n)}{\sqrt{\sum_{i=1}^n (R_i - \overline{R}_n)^2}\sqrt{\sum_{i=1}^n (S_i - \overline{S}_n)^2}} = 1 - \frac{6}{n^2(n-1)} \sum_{i=1}^n (R_i - S_i)^2,$$

  where R_i and S_i are ranks of U_i and V_i respectively.
- U_i and V_i independent $\rightsquigarrow \rho = \tau = \rho_S = 0$

4. and 5. Tests Based on τ and ρ_S

Idea of the test: Compute correlation between $U_i = Y_i$ and $V_i = i$

$$\widehat{\tau} = \frac{2}{n(n-1)} \sum_{i<j} \operatorname{sgn}(Y_i - Y_j) = \frac{4}{n(n-1)} \sum_{i<j} \mathrm{I}(Y_i - Y_j),$$

$$\widehat{\rho}_S = 1 - \frac{6}{n^2(n-1)} \sum_{i=1}^{n} (R_i - i)^2$$

where $R_1, \ldots, R_n$ are ranks of $Y_1, \ldots, Y_n$

Asymptotic tests: Compare

$$\frac{|\widehat{\tau}|}{\sqrt{\frac{2(2n+5)}{9n(n-1)}}} \quad \text{or} \quad \sqrt{n-1}\,|\widehat{\rho}_S|$$

with $u_{1-\alpha/2}$, and reject for large values

Simulations

$N = 1\,000$ replications $\rightsquigarrow$ percentage of rejection of H_0

	τ			ρ_S		
n	50	100	200	50	100	200
IID	5	5	6	5	5	6
AR	34	29	33	34	30	33
LT	100	100	100	100	100	100
RW	81	85	90	82	85	91

Graphical tools

- plot
- suitable graphical tools from regression
- tools based on sample ACF of $\{Y_t\}$

Graphical tools

- plot
- suitable graphical tools from regression
- tools based on sample ACF of $\{Y_t\}$

Course Stoch. processes II: $\{Y_t\}$ random proces

- ACF

$$\rho_k = \mathrm{cor}(Y_t, Y_{t+k})$$

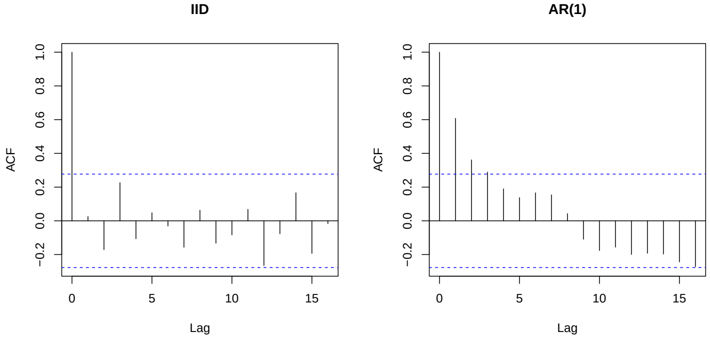
  If $\{Y_t\}$ iid $\rightsquigarrow \rho_k = 0$ for $k \neq 0$

- sample ACF

$$r_k = \frac{\sum_{t=1}^{n-k}(Y_t - \overline{Y}_n)(Y_{t+k} - \overline{Y}_n)}{\sum_{t=1}^{n}(Y_t - \overline{Y}_n)^2}$$

  If $\{Y_t\}$ iid $\rightsquigarrow \sqrt{n} r_k \xrightarrow{D} \mathrm{N}(0, 1)$, i.e. $r_k \mathrel{\dot\sim} \mathrm{N}(0, 1/n)$ for large n

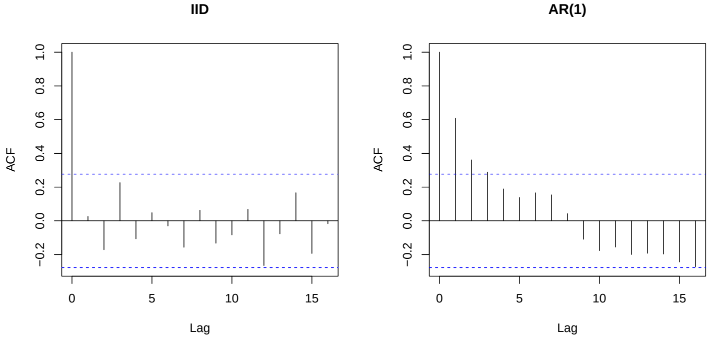
Sample ACF



Horizontal lines:

$$\pm \frac{u_{0.975}}{\sqrt{n}}$$

Sample ACF



Horizontal lines:

$$\pm \frac{u_{0.975}}{\sqrt{n}}$$

Under H_0: r_k lies outside $\left(-\frac{u_{0.975}}{\sqrt{n}}, \frac{u_{0.975}}{\sqrt{n}}\right)$ with asymptotic probability 5% for each $k \geq 1$, independently

Portmanteau tests

Box-Pierce, Ljung- Box, Q-test

Idea of the test:

- $\hookrightarrow$ fix K
- $\hookrightarrow$ If $\{Y_t\}$ iid, then $\sqrt{n}r_1, \ldots, \sqrt{n}r_K$ asymptotically $N(0, 1)$ and independent

Portmanteau tests

Box-Pierce, Ljung- Box, Q-test

Idea of the test:

$\hookrightarrow$ fix K

$\hookrightarrow$ If $\{Y_t\}$ iid, then $\sqrt{n}r_1, \ldots, \sqrt{n}r_K$ asymptotically $\mathrm{N}(0,1)$ and independent

$\hookrightarrow$ Take

$$Q = n \sum_{k=1}^{K} r_k^2$$

and it should be asymptotically χ_K^2

Portmanteau tests

Box-Pierce, Ljung- Box, Q-test

Idea of the test:

- fix K
- If $\{Y_t\}$ iid, then $\sqrt{n}r_1, \ldots, \sqrt{n}r_K$ asymptotically N$(0, 1)$ and independent
- Take
$$Q = n\sum_{k=1}^{K} r_k^2$$
and it should be asymptotically χ_K^2

Test: Reject if $Q > q_{K,1-\alpha}$ for $q_{1-\alpha}$ quantile of χ_K^2

Portmanteau tests

Box-Pierce, Ljung- Box, Q-test

Idea of the test:

- $\hookrightarrow$ fix K
- $\hookrightarrow$ If $\{Y_t\}$ iid, then $\sqrt{n}r_1, \ldots, \sqrt{n}r_K$ asymptotically N$(0, 1)$ and independent
- $\hookrightarrow$ Take

$$Q = n \sum_{k=1}^{K} r_k^2$$

and it should be asymptotically χ^2_K

Test: Reject if $Q > q_{K,1-\alpha}$ for $q_{1-\alpha}$ quantile of χ^2_K

Small sample improvement:

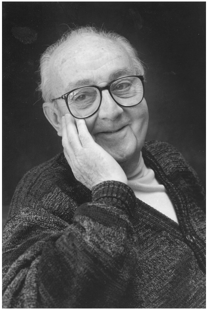
$$Q^* = n(n+2) \sum_{k=1}^{K} \frac{r_k^2}{n-k}$$

If $\{Y_t\}$ are residuals from an ARMA model $\rightsquigarrow$ modify the degrees of freedom

Box-Jenkins methodology

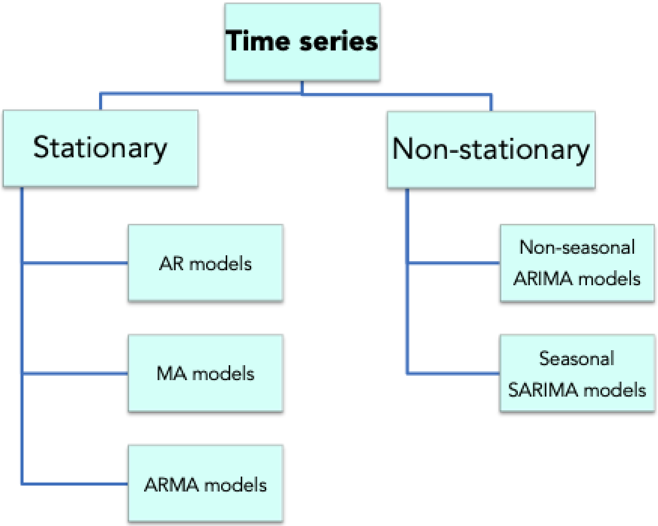
Box-Jenkins methodology

- AutoRegressive Integrated Moving Average (ARIMA) models
- 1970s, popularized by Box and Jenkins
- rely on autocorrelation patterns in the data



George E. P. Box
1919 – 2013

Gwilym M. Jenkins
1932 – 1982


Time series
Stationary
Non-stationary
AR models
MA models
ARMA models
Non-seasonal ARIMA models
Seasonal SARIMA models

Notions and definitions

Time series $\{Y_t\}$

- strict stationarity
- (weak) stationarity
- white noise WN
- autocovariance function $\{\gamma_k\}$
- autocorrelation function (ACF) $\{\rho_k\}$
- partial autocorrelation function (PACF) $\{\rho_{kk}\}$

Sample counterparts

- sample mean
- sample autocovariance function $\{c_k\}$
- sample ACF $\{r_k\}$
- sample PACF $\{r_{kk}\}$

Practical recommendation: $n > 50$, $k < n/4$