ECON207 Session 11-12

Intro to Time Series & Time Series Regressions

Anthony Tay

This Version: 30 Sep 2024

Session 11-12

Session 11-12: Intro to Time Series and Time Series Regressions

- Characteristics of Economic Time Series Data
  - Trends, Seasonality and Cycles
  - Intertemporal Correlations
- Statistical Models and Tools for Describing Such Features
- Covariance-Stationarity and Weak-Dependence
- Regressions with Covariance-Stationary Weakly-Dependent Time Series
- Regressions with Trending and Persistent Series

R Code and Data

We will use the following packages

```
library(fpp3)        # a suite of packages for forecasting, and includes tidyverse
library(patchwork)   #  for plotting
library(ggfortify)   #  for plotting
library(readxl)      #  used to read in excel files
library(mFilter)     #  require for HP filter
library(seasonal)    #  required for seasonal adjustment
library(stargazer)   #  For nice regression output
```

Time series can be stored in several ways in R

We will store our time series in “tsibbles” (a kind of data frame)

R Code and Data

```
ts01 <- read_excel("data\\ts_01.xlsx") %>%    # data in columns, include dates in DATE
  mutate(DATE=yearmonth(DATE)) %>%            # required for conversion to tsibble
  as_tsibble(index=DATE) %>%                  # conversion to tsibble
  mutate(POULTRY_US = POULTRY_US/100000)      # convert POULTRY_US from thousands to 100 million
head(ts01,3)
```

```
# A tsibble: 3 x 7 [1M]
      DATE ELEC_GEN_SG TOUR_SG IP_SG CPI_US DOMEX5_SG POULTRY_US
     <mth>       <dbl>   <dbl> <dbl>  <dbl>     <dbl>      <dbl>
1 1983 Jan        667.  232164  14.3   97.8        NA       3.40
2 1983 Feb        587.  212591  11.4   97.9        NA       3.14
3 1983 Mar        727.  242272  14.5   97.9        NA       3.68
```

```
tail(ts01,3)
```

```
# A tsibble: 3 x 7 [1M]
      DATE ELEC_GEN_SG TOUR_SG IP_SG CPI_US DOMEX5_SG POULTRY_US
     <mth>       <dbl>   <dbl> <dbl>  <dbl>     <dbl>      <dbl>
1 2017 Oct       4504. 1402299  118.   247.     3565.       7.83
2 2017 Nov       4283. 1397330  115.   247.     3582.       7.28
3 2017 Dec       4376. 1569616  120.   247.     3682.       7.04
```

R Code and Data

We will need the following new variables

- **t**: sequence of integers
- **tsq**: sequence of squares
- **d01**, **d02**, ..., **d11**, **d12**: “monthly seasonal dummies”
    - d01 $= 1$ for January data, 0 otherwise
    - d02 $= 1$ for February data, 0 otherwise
    - etc.

R Code and Data

```
ts01 <- ts01 %>% mutate(t=row_number(), tsq = t^2,
                        d01=ifelse(month(DATE)==1,1,0),
                        d02=ifelse(month(DATE)==2,1,0),
                        d03=ifelse(month(DATE)==3,1,0),
                        d04=ifelse(month(DATE)==4,1,0),
                        d05=ifelse(month(DATE)==5,1,0),
                        d06=ifelse(month(DATE)==6,1,0),
                        d07=ifelse(month(DATE)==7,1,0),
                        d08=ifelse(month(DATE)==8,1,0),
                        d09=ifelse(month(DATE)==9,1,0),
                        d10=ifelse(month(DATE)==10,1,0),
                        d11=ifelse(month(DATE)==11,1,0),
                        d12=ifelse(month(DATE)==12,1,0))

# The following string and vector will also be useful for us
seas <- "d02 + d03 + d04 + d05 + d06 + d07 + d08 + d09 + d10 + d11 + d12"
seaslist <- c("d01", "d02", "d03", "d04", "d05", "d06", "d07", "d08", "d09", "d10", "d11", "d12")
```

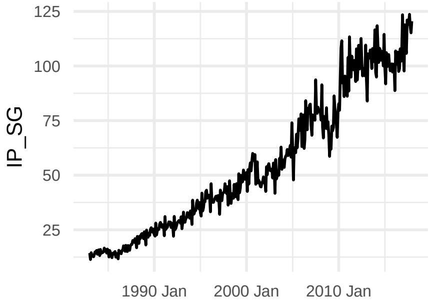
R Code and Data

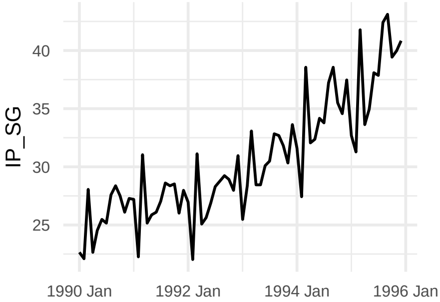
```r
options(width=400)
ts01 %>% select(DATE,t, tsq, seaslist) %>% filter_index(.~"1984M3")
```

```
# A tsibble: 15 x 15 [1M]
       DATE     t   tsq   d01   d02   d03   d04   d05   d06   d07   d08   d09   d10   d11   d12
      <mth> <int> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
 1 1983 Jan     1     1     1     0     0     0     0     0     0     0     0     0     0     0
 2 1983 Feb     2     4     0     1     0     0     0     0     0     0     0     0     0     0
 3 1983 Mar     3     9     0     0     1     0     0     0     0     0     0     0     0     0
 4 1983 Apr     4    16     0     0     0     1     0     0     0     0     0     0     0     0
 5 1983 May     5    25     0     0     0     0     1     0     0     0     0     0     0     0
 6 1983 Jun     6    36     0     0     0     0     0     1     0     0     0     0     0     0
 7 1983 Jul     7    49     0     0     0     0     0     0     1     0     0     0     0     0
 8 1983 Aug     8    64     0     0     0     0     0     0     0     1     0     0     0     0
 9 1983 Sep     9    81     0     0     0     0     0     0     0     0     1     0     0     0
10 1983 Oct    10   100     0     0     0     0     0     0     0     0     0     1     0     0
11 1983 Nov    11   121     0     0     0     0     0     0     0     0     0     0     1     0
12 1983 Dec    12   144     0     0     0     0     0     0     0     0     0     0     0     1
13 1984 Jan    13   169     1     0     0     0     0     0     0     0     0     0     0     0
14 1984 Feb    14   196     0     1     0     0     0     0     0     0     0     0     0     0
15 1984 Mar    15   225     0     0     1     0     0     0     0     0     0     0     0     0
```

Characteristics of Economic Time Series

```
p1 <- ts01 %>% autoplot(IP_SG) + theme_minimal() + xlab("") + theme(text=element_text(size=8))
p2 <- ts01 %>% filter_index("1990M1" ~ "1995M12") %>%
      autoplot(IP_SG) + theme_minimal() + xlab("") + theme(text=element_text(size=8))
p1 | p2
```





Trend, Cycles, Seasonality and other features (e.g., increasing variance)

Elementary Tools

Discuss some elementary tools for describing these characteristics

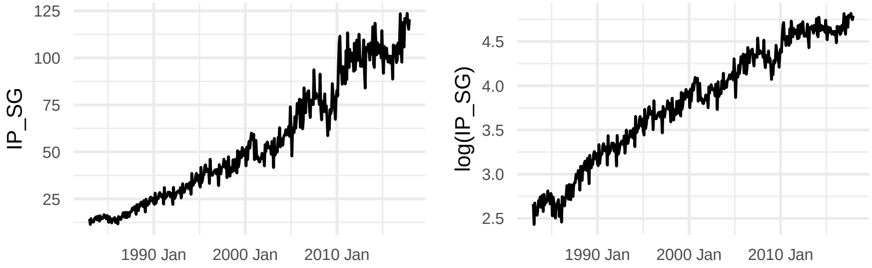
- Transformations
- Elementary Models for
  - Trends
  - Seasonality
  - Cycles
- (Population and Sample) Autocovariance and Autocorrelation Function

Develop vocabulary for describing time series features

A Common Data Transformation

We often work with $\log Y_t$ instead of Y_t

```
p3 <- ts01 %>% autoplot(IP_SG) + theme_minimal() + xlab("") + theme(text=element_text(size=8))
p4 <- ts01 %>% autoplot(log(IP_SG)) + theme_minimal() + xlab("") + theme(text=element_text(size=8))
p3 | p4
```



- Stabilizes increasing variance, $\log Y_t - \log Y_{t-1}$ interpretation as growth rate

Ways to Describe Trend / Deterministic Trend

We often want to

- classify trend: "deterministic" or "stochastic"
- estimate trend (perhaps to remove it, i.e., to detrend data)

Deterministic Trend Models

$$Y_t = f(t) + \epsilon_t \, , \, t = 1, 2, 3, \ldots$$

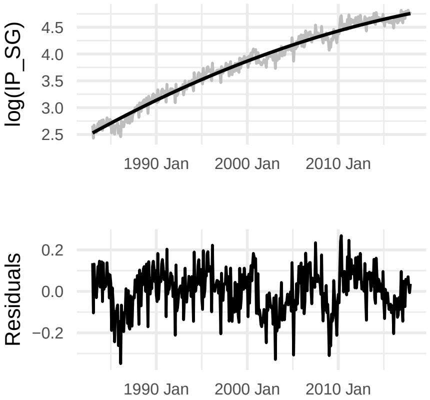
- Deterministic Linear Trend: $Y_t = \beta_0 + \beta_1 t + \epsilon_t \; , \quad t = 1, 2, \ldots$
- Deterministic Quadratic Trend: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon_t \; , \quad t = 1, 2, \ldots$
- Many other possibilities

Deterministic Trend

```
theme1 <- theme_minimal() +
  theme(axis.title=element_text(size=8),
        axis.text=element_text(size=6))
mdl_dqt <- lm(log(IP_SG) ~ t + tsq, data=ts01)
df_plt <- ts01 %>%
  mutate("Fitted"=fitted(mdl_dqt),
         "Residuals"=residuals(mdl_dqt))
p1 <- autoplot(df_plt, log(IP_SG), size=0.5, color='grey') +
  autolayer(df_plt, Fitted, size=0.6) + theme1 + xlab("")
p2 <- autoplot(df_plt, Residuals) +
  theme1 + xlab("")
mdl_dqt %>% summary() %>% coef() %>% round(6)
cat("R-sqr: ", summary(mdl_dqt)$r.squared)
```

```
              Estimate Std. Error   t value Pr(>|t|)
(Intercept)   2.519434   0.015296 164.70896        0
t             0.007774   0.000168  46.33197        0
tsq          -0.000006   0.000000 -15.11436        0
R-sqr:  0.9751655
```

```
p1 / p2
```



Deterministic Trend (Nonparametric Approach)

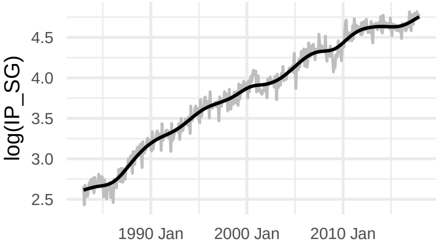
A Nonparametric Approach (HP Filter)

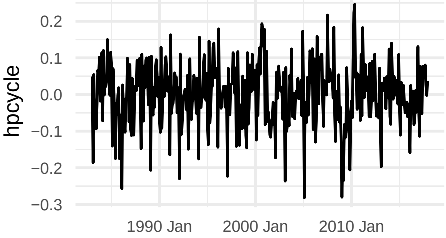
$$\hat{\tau}_t^{hp} = \operatorname{argmin}_{\hat{\tau}_t}\left(\sum_{t=1}^{T}(y_t - \hat{\tau}_t)^2 + \lambda\sum_{t=2}^{T-1}\left[(\hat{\tau}_{t+1} - \hat{\tau}_t) - (\hat{\tau}_t - \hat{\tau}_{t-1})\right]^2\right)$$

```
ts03.ts <- as.ts(ts01)
lipsg <- hpfilter(log(ts03.ts[,'IP_SG']), type="lambda", freq=14400)
hp_dat <- as_tsibble(ts.union("hpcycle"=lipsg$cycle,
                    "hptrend"=lipsg$trend,
                    "log(IP_SG)"=log(ts03.ts[,'IP_SG'])),pivot_longer=F)
p1 <- autoplot(hp_dat, `log(IP_SG)`, color="grey") +
      autolayer(hp_dat, hptrend, size=0.6) +
      theme_minimal() + theme(legend.position = "bottom") + theme1 + xlab("")
p2 <- autoplot(hp_dat, hpcycle) + theme1 + xlab("")
```

Deterministic Trend (Nonparametric Approach)

```
p1 | p2
```



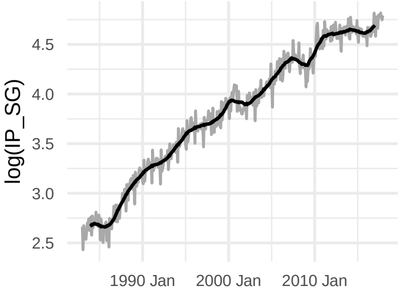


Deterministic Trend (Nonparametric Approach)

Moving Average: $\hat{\tau}_t^{ma} = \frac{1}{2k+1}\sum_{j=-k}^{k} Y_{t+j}, t = k+1, ..., T-k$

```
ma_dat <- ts01 %>%
  select(IP_SG) %>%
  mutate("MA_IP_SG"=as.numeric(NA))
k <- 11
T <- dim(ma_dat)[1]
for (i in (k+1):(T-k)){
  ma_dat[i,"MA_IP_SG"] <- mean(
    log(ma_dat$IP_SG[(i-k):(k+i)]))
}
```

```
autoplot(ma_dat, log(IP_SG), color="darkgray") +
  autolayer(ma_dat, MA_IP_SG, linewidth=0.6) +
  theme1 + xlab("")
```



Stochastic Trends

"Stochastic Trends": Models Y_t as

$$Y_t - Y_{t-1} = \alpha + \epsilon_t \quad \text{i.e.,} \quad Y_t = \alpha + Y_{t-1} + \epsilon_t \quad \text{for all} \quad t$$

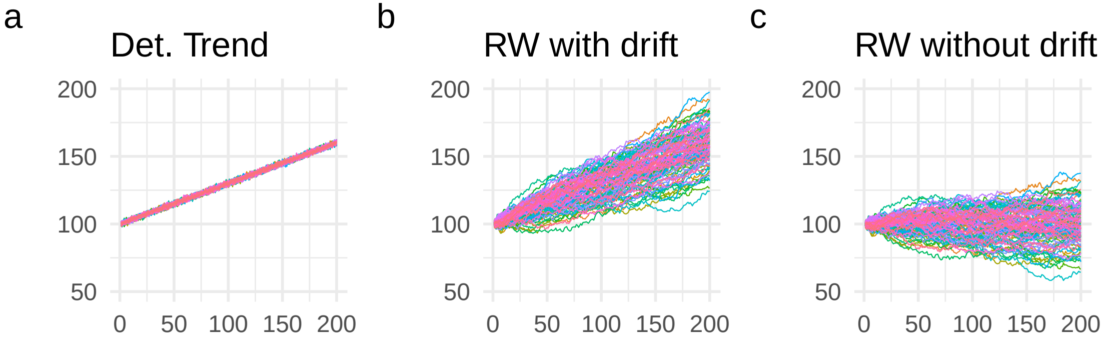
"Random Walk" (with drift if $\alpha \neq 0$)

Essential difference between "stochastic trend" and "deterministic trend"

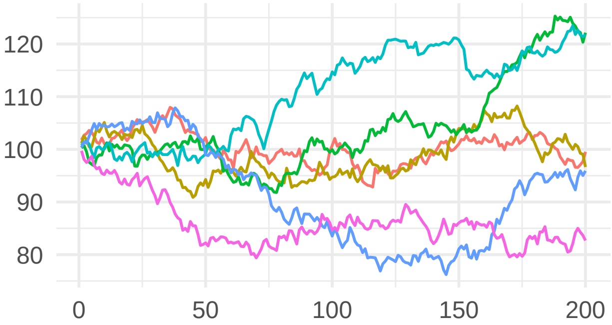
$$\begin{aligned}
Y_1 &= \alpha + Y_0 + \epsilon_1 & \operatorname{var}(Y_1 \mid Y_0) &= \sigma^2 \\
Y_2 &= \alpha + Y_1 + \epsilon_2 = Y_0 + 2\alpha + \epsilon_1 + \epsilon_2 & \operatorname{var}(Y_2 \mid Y_0) &= 2\sigma^2 \\
Y_3 &= \alpha + Y_2 + \epsilon_3 = Y_0 + 3\alpha + \epsilon_1 + \epsilon_2 + \epsilon_3 & \operatorname{var}(Y_3 \mid Y_0) &= 3\sigma^2 \\
&\vdots & &\vdots \\
Y_t &= \alpha + Y_{t-1} + \epsilon_t = Y_0 + \alpha t + \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots + \epsilon_t & \operatorname{var}(Y_t \mid Y_0) &= t\sigma^2
\end{aligned}$$

Stochastic Trends

Simulated det. trends and Random Walks, with unit error variance



Stochastic Trends



Stochastic Trends

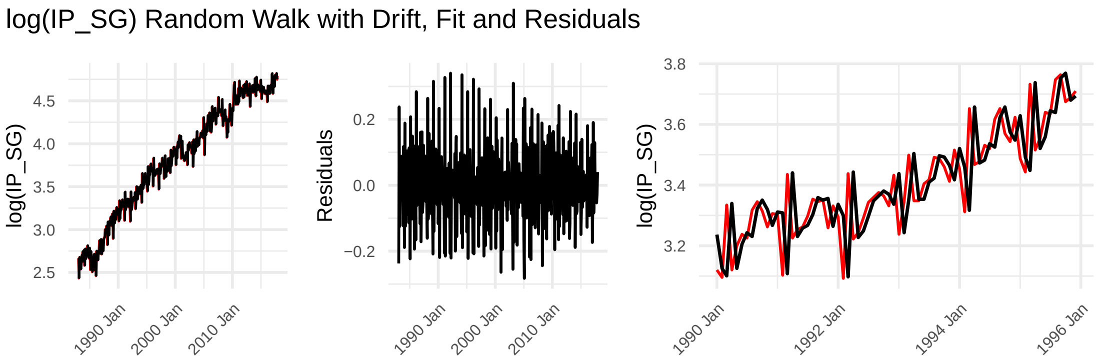
To fit a random walk with drift, just estimate the mean of the first difference:

```
a0 <- mean(diff(log(ts01$IP_SG)))   # Model is Y_{t} = a0 + Y_{t-1} + e_{t}

theme2 <- theme(axis.title = element_text(size=8), axis.text = element_text(size=6),
            axis.text.x = element_text(size=6, angle=45, hjust=1))
df_plt2 <- ts01 %>% mutate(Fitted=lag(log(IP_SG))+a0, Residuals=log(IP_SG)-Fitted)
p1 <- autoplot(df_plt2, log(IP_SG), size=0.4, color="red") +
      autolayer(df_plt2, Fitted, size=0.4) + theme1 + xlab("") + theme2
p2 <- autoplot(df_plt2, Residuals) + theme1 + xlab("") + theme2
df_plt2sub <- df_plt2 %>% filter_index("1990M1" ~ "1995M12")
p3 <- autoplot(df_plt2sub,log(IP_SG), color="red") + theme1 +
      autolayer(df_plt2sub, Fitted, size=0.6) + theme2 + xlab("")
```

Stochastic Trends

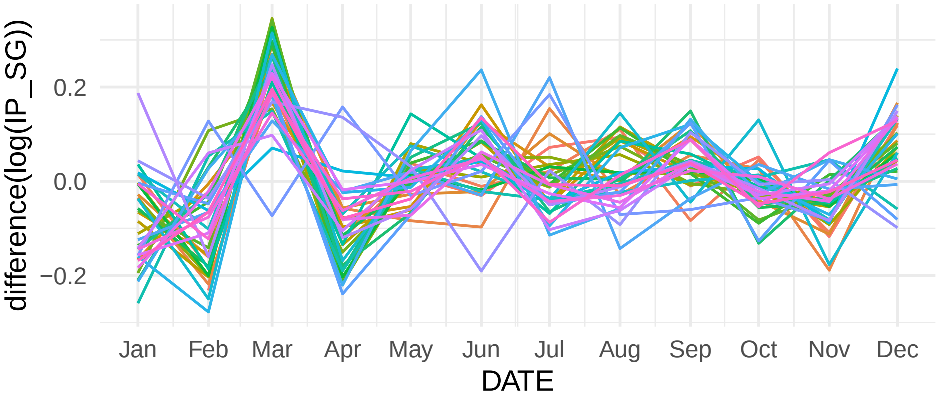
```
(p1 | p2 | p3) + plot_annotation(title="log(IP_SG) Random Walk with Drift, Fit and Residuals",
  theme =  theme(plot.title = element_text(size=10))) +
  plot_layout(widths=c(1,1,1.8))
```



Seasonality

Patterns with regular period, arising from 'mechanical' reasons

```
ts01 %>% gg_season(difference(log(IP_SG))) + theme1
```



Modelling Seasonality with Seasonal Dummies

Three equivalent specifications

$$Y_t = \beta_1 d_{1,t} + \beta_2 d_{2,t} + \cdots + \beta_{12} d_{12,t} + \epsilon_t$$

$$Y_t = \alpha_0 + \alpha_2 d_{2,t} + \cdots + \alpha_{12} d_{12,t} + \epsilon_t$$

$$Y_t = \delta_0 + \delta_2 (d_{2,t} - \tfrac{1}{12}) + \delta_3 (d_{3,t} - \tfrac{1}{12}) + \cdots + \delta_{12} (d_{12,t} - \tfrac{1}{12}) + \epsilon_t$$

Can use in combination with other regressors

Modelling Seasonality with Seasonal Dummies

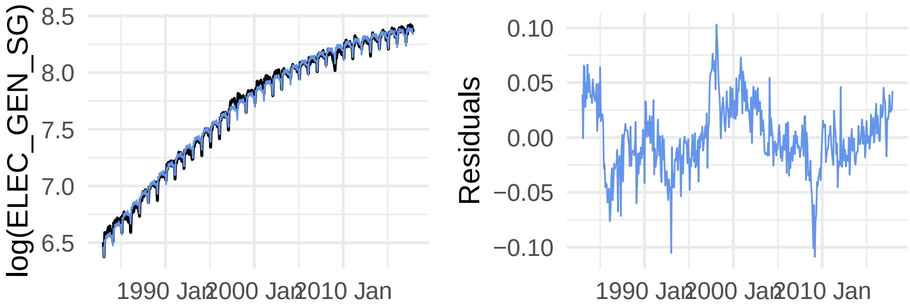
Seasonal Dummies with Quadratic Trend — log *ELEC_GEN_SG*

$$Y_t = \alpha_0 + \alpha_2 d_{2,t} + \cdots + \alpha_{12} d_{12,t} + \beta_1 t + \beta_2 t^2 + \epsilon_t$$

```
mdl_dqt <- lm(as.formula(paste0("log(ELEC_GEN_SG) ~ t + tsq + ", seas)),
              data=ts01)
df_plt1 <- ts01 %>% mutate("Fitted"=fitted(mdl_dqt),
                           "Residuals"=residuals(mdl_dqt))
p1 <- autoplot(df_plt1, log(ELEC_GEN_SG), linewidth=0.5, color="black") +
      autolayer(df_plt1, Fitted, linewidth=0.3, color="cornflowerblue") +
      theme1 + xlab("")
p2 <- autoplot(df_plt1, Residuals, linewidth=0.3, color="cornflowerblue") +
      theme1 + xlab("")
```

Modelling Seasonality with Seasonal Dummies

p1 | p2



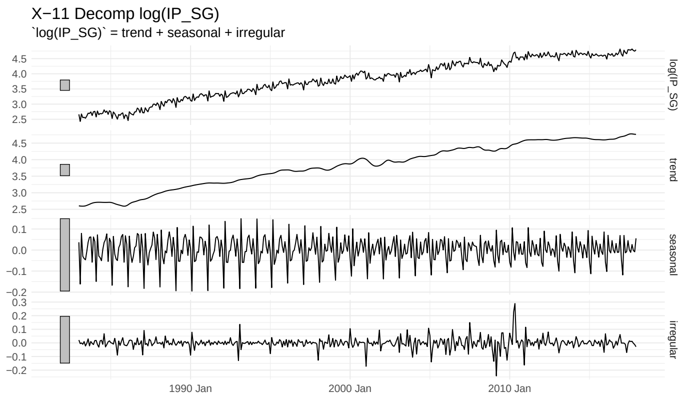
Modelling Seasonality with Seasonal Dummies

- Seasonal dummy approach assumes "very regular" seasonal patterns — may or may not be appropriate
- Other methods (e.g., seasonal ARIMA) — not covered in this course
- Seasonally Adjusted Data
  - Official Statistics Agencies often provide seasonally-adjusted data
  - Typical Method
    - Estimate "Trend-Cycle" nonparametrically (e.g., moving average methods)
    - Estimate Seasonal Component nonparametrically (e.g., "seasonal" m.a. methods)
    - Remove Seasonal Component from original series

Seasonally-Adjusted Series

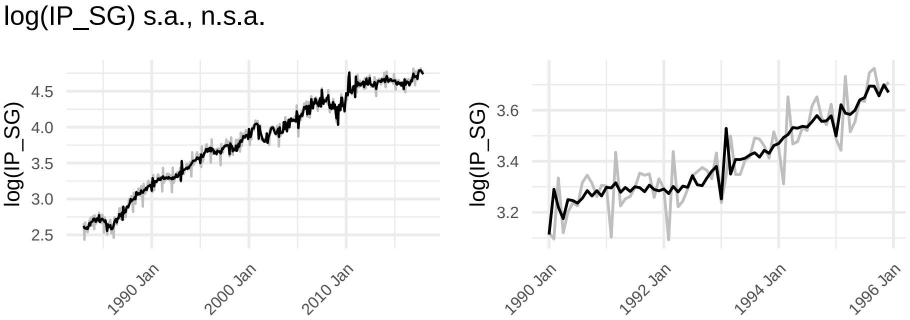
```
ipsg_dcmp <- ts01 %>%
  model(
    x11=X_13ARIMA_SEATS(
        log(IP_SG) ~ x11()
        )
  ) %>%
  components()
p_decomp <- autoplot(ipsg_dcmp) +
  theme1 +
  labs(title="X-11 Decomp log(IP_SG)") +
  xlab("") +
  theme(text=element_text(size=14))
```

```
p_decomp
```



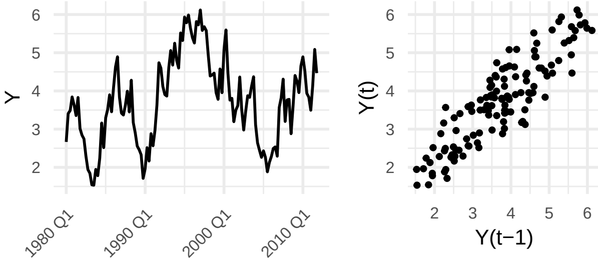
Seasonally-Adjusted Series

```
p2 <- ipsg_dcmp %>% ggplot(aes(x = DATE)) + geom_line(aes(y = `log(IP_SG)`), color="grey", size=0.4) +
  geom_line(aes(y = season_adjust), size=0.4) + theme_minimal() + theme2 + xlab("")
p3 <- ipsg_dcmp %>% filter_index("1990M1" ~"1995M12") %>%
  ggplot(aes(x = DATE)) + geom_line(aes(y = `log(IP_SG)`), size=0.5, color="grey") +
  geom_line(aes(y = season_adjust), size=0.5) + theme_minimal() + theme2 + xlab("")
(p2 | p3) + plot_annotation(title="log(IP_SG) s.a., n.s.a.", theme = theme(title = element_text(size=8)))
```



Modelling Cycles

```
ts02 <- read_excel("data\\ts_02.xlsx") %>%
        mutate(Period=yearquarter(Period), Y_1=lag(Y,1)) %>% as_tsibble(index=Period)
p1 <- ts02 %>% autoplot(Y) + theme1 + theme2 + theme(aspect.ratio = 0.7) + xlab("")
p2 <- ts02 %>% ggplot(aes(x=Y_1, y=Y)) + geom_point(size=0.6) + theme_minimal() + ylab("Y(t)") +
      xlab("Y(t-1)") + theme(axis.title = element_text(size=8),
                             axis.text = element_text(size=6), aspect.ratio = 1)
p1 | p2
```



Modelling Cycles

Scatterplot suggests something like:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t$$

may be suitable for modeling cycles

"Autoregression of Order 1" or "AR(1)" — Member of ARIMA class of models

- ARIMA models covered in detail in other courses
- We limit ourselves to AR(1) in this course

Modelling Cycles

We have already seen an AR(1)

- Random Walk: $Y_t = \beta_0 + Y_{t-1} + \epsilon_t$, i.e., $\beta_1 = 1$
- Used to describe "stochastic trend"

For cycles, we use the "covariance-stationary AR(1)"

- $Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t$, i.e., $|\beta_1| < 1$
- We will explain the term "covariance-stationary" in a moment
- We first cover a little bit of Time Series theory

A Little TS Theory

- Stochastic Processes
- Autocovariance and Autocorrelation Functions (Population and Sample)
- Covariance-Stationary Processes
  - Trend-Stationary Processes
  - Difference-Stationary Process
- Weakly Dependent Series / Persistent Series

Stochastic Processes

Throughout, think of the "population" as a "data generating process",

- e.g., think of economy as a "machine" that generates data on output, prices, interest rates, etc. over time
- more precisely, this "machine" generates *sequences of random variables*
  - $Y_1, Y_2, \ldots, Y_t, \ldots$
  - $X_1, X_2, \ldots, X_t, \ldots$
  - etc.
- we call such sequences of random variables as **stochastic processes**
- Our time series data are realizations of these stochastic processes

Stochastic Processes

A **times series model** is a stylized description of a stochastic process

- Linear deterministic trend model: $Y_t = \beta_0 + \beta_1 t + \epsilon_t$, $\epsilon \overset{iid}{\sim} (0, \sigma^2)$
- Random walk with drift: $Y_t = \beta_0 + Y_{t-1} + \epsilon_t$, $\epsilon \overset{iid}{\sim} (0, \sigma^2)$

Important to understand how various stochastic processes behave:

- unconditional means $E(Y_t)$ and unconditional variances $Var(Y_t)$
- conditional means $E(Y_t \mid Y_{t-1}, \dots)$ and conditional variances $Var(Y_t \mid Y_{t-1}, \dots)$

Stochastic Processes

For stochastic processes, **intertemporal correlations**, i.e., "dynamics", are important

- autocovariance and autocorrelations, for each t and for $k = 1, 2, \ldots$

$$\begin{aligned}\gamma_{t,k} &= Cov(Y_t, Y_{t-k}) \\ &= E((Y_t - E(Y_t))(Y_{t-k} - E(Y_{t-k})))\end{aligned}$$

$$\rho_{t,k} = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t-k})}}$$

We call these "autocovariance functions" and "autocorrelation functions"

The term "serial correlation" and "autocorrelation" are synonymous

Stochastic Processes (White Noise)

Examples

White Noise: ϵ_t where for all t,

- $E(\epsilon_t) = 0$
- $Var(\epsilon_t) = \sigma^2$
- $Cov(\epsilon_t, \epsilon_s) = 0$ for all $t \neq s$

If ϵ_t and ϵ_s are independent for all $s \neq t$, we call ϵ_t an "independent white noise" process

Covariance Stationary Processes

A **Covariance Stationary** process Y_t is one where

- $E(Y_t)$ is the same finite constant for all t
- $Var(Y_t)$ is the same finite constant for all t
- $Cov(Y_t, Y_{t-k})$ is, for any k, the same finite constant for all t
  - May be different for different k
  - but for any given k, same for all t, i.e.,

$$\gamma_{t,k} = Cov(Y_t, Y_{t-k}) = E((Y_t - E(Y_t))(Y_{t-k} - E(Y_{t-k}))) = \gamma_k$$

Covariance Stationary Processes

Since for covariance-stationary processes,

$$Cov(Y_t, Y_{t-k}) = \gamma_{k,t} = \gamma_k \text{ and } Var(Y_t) = \gamma_{0,t} = \gamma_0$$

we have

$$Autocorr(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0}$$

Covariance-Stationary Processes are "stable" processes

- Expect to see steady fluctuations around a constant value (no trend, no changes in variance)
- Autocorrelation patterns not changing over time (we elaborate on this soon)

Covariance Stationary Processes

Example A white noise process is covariance stationary:

$$E(Y_t) = 0,\, Var(Y_t) = \sigma^2,\, Cov(Y_t, Y_s) = 0 \ \text{ for all } \ t, s, t \neq s$$

Example A (pure) deterministic trend process is not covariance stationary, since

$$Y_t = \beta_0 + \beta_1 t + \epsilon_t\,,\ \epsilon_t \sim WN(0, \sigma^2) \ \implies \ E(Y_t) = \beta_0 + \beta_1 t \ \text{ depends on } \ t$$

A deterministic trend process is said to be "trend-stationary"

A **trend-stationary** process is one that is:

$$Y_t = f(t) + \text{ cov. stat. process}$$

Covariance Stationary AR(1)

Example A covariance-stationary process is a **covariance-stationary AR(1)** if it satisfies

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t \, , \; |\beta_1| < 1 \, , \; \epsilon \sim WN(0, \sigma^2)$$

Properties:

Mean: $E(Y_t) = \dfrac{\beta_0}{1-\beta_1}$

- *Proof*:
  $E(Y_t) = \beta_0 + \beta_1 E(Y_{t-1}) + E(\epsilon_t) \;\Rightarrow\; E(Y_t) = \beta_0 + \beta_1 E(Y_t) \;\Rightarrow\; E(Y_t) = \dfrac{\beta_0}{1-\beta_1}$
- If $\beta_0 = 0$, then $E(Y_t) = 0$

Covariance-Stationary AR(1)

Variance: $Var(Y_t) = \dfrac{\sigma^2}{1-\beta_1^2}$

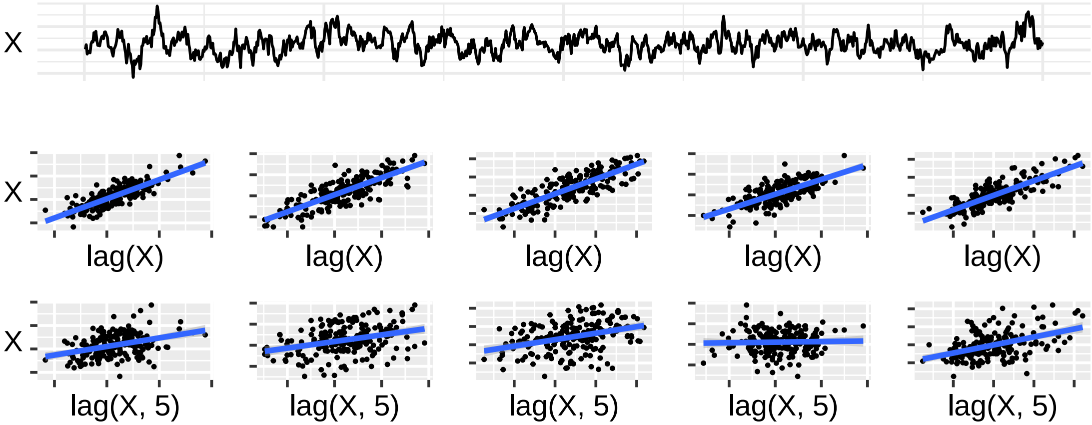
Autocovariance Fn.: $\gamma_k = Cov(Y_t, Y_{t-k}) = E((Y_t - E(Y_t))(Y_{t-k} - E(Y_{t-k})))$

$$Cov(Y_t, Y_{t-1}) = \frac{\sigma^2\beta_1}{1-\beta_1^2}\ ,\quad Cov(Y_t, Y_{t-2}) = \frac{\sigma^2\beta_1^2}{1-\beta_1^2}\ ,\quad \ldots,\quad Cov(Y_t, Y_{t-k}) = \frac{\sigma^2\beta_1^k}{1-\beta_1^2}$$

Autocorrelation Fn.: $\rho_1 = \beta_1\ ,\ \rho_2 = \beta_1^2\ ,\ \ldots,\ \rho_k = \beta_1^k\ ,\ \ldots$

Since $|\beta_1| < 1$, ACF decays with k, $\rho_k \to 0$ as $k \to \infty$ (“Weak Dependence”)

Covariance-Stationary AR(1)



Covariance-Stationary AR(1)

Measuring Intertemporal Correlations:

Sample Autocovariance Function:

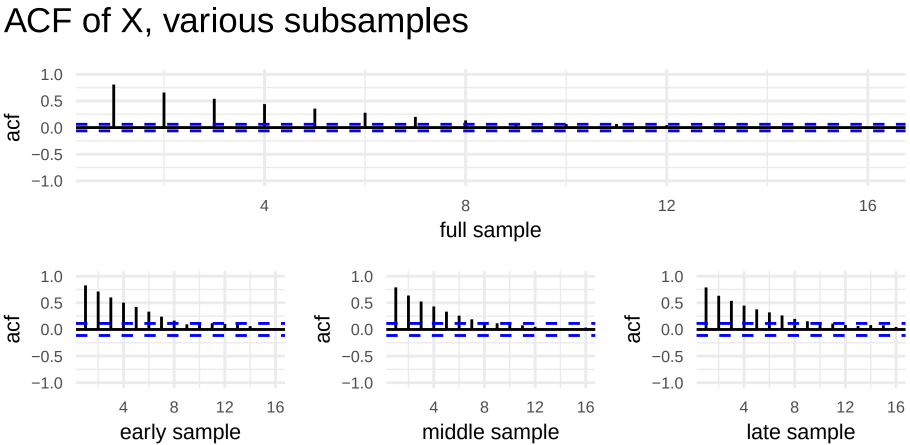
$$\hat{\gamma}_k = \frac{1}{T}\sum_{t=k+1}^{T}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})\,,\ k = 0, 1, 2, ...$$

$\hat{\gamma}_0$ is the sample variance (biased version)

Sample Autocorrelation Function:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}\,,\ \ k = 0, 1, 2, 3, ...$$

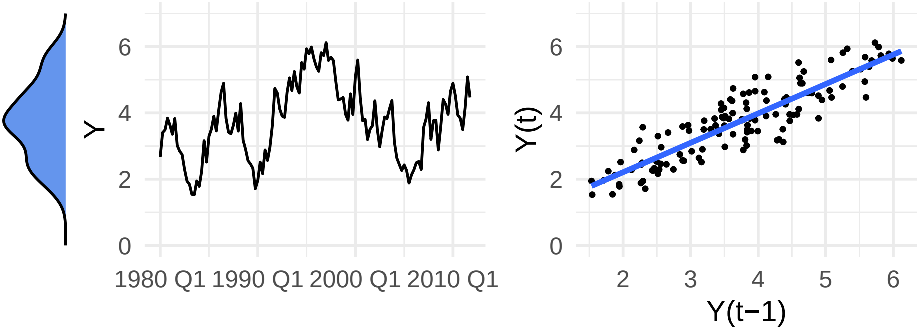
Covariance-Stationary AR(1)



Covariance-Stationary AR(1)

Conditional Mean: $E(Y_t \mid Y_{t-1}, \ldots) = \beta_0 + \beta_1 Y_{t-1}$

Conditional Variance: $Var(Y_t \mid Y_{t-1}, \ldots) = \sigma^2$



AR(1) with structural break

Example *AR(1) with Structural Break* Suppose

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t \text{ for all } t < \tau$$
$$Y_t = \beta_0^* + \beta_1^* Y_{t-1} + \epsilon_t \text{ for all } t \geq \tau$$

where $\beta_0 \neq \beta_0^*$ or $\beta_1 \neq \beta_1^*$ (or both), though both β_1 and β_1^* are less than one in absolute value

AR(1) with structural break

- before and after break, process is covariance stationary
- but when treated as a whole, process is *not* covariance-stationary

Random Walk without Drift Again

Example Random Walk without Drift

$$Y_t = Y_{t-1} + \epsilon_t = Y_0 + \epsilon_t + \epsilon_{t-1} + \cdots + \epsilon_1 \ , \quad \epsilon \sim WN(0, \sigma^2)$$

Unconditional variance depends on how we treat Y_0

- If we treat Y_0 as fixed (easiest), then $Var(Y_t) = t\sigma^2$
- As process evolves, i.e., as $t \to \infty$, the variance becomes infinity

Furthermore: $\rho_{t,k} = Cor(Y_t, Y_{t-k}) = \dfrac{(t-k)\sigma^2}{\sqrt{t\sigma^2}\sqrt{(t-k)\sigma^2}} = \sqrt{\dfrac{(t-k)}{t}}$

- As process evolves, i.e., as $t \to \infty$, we see that for any fixed k, $\rho_{t,k} \to 1$
- Not covariance-stationary, nor weakly dependent, but "highly persistent"

Difference-Stationary Processes

Random Walk (with / without drift) are examples of "difference-stationary processes"

$$Y_t = \beta_0 + Y_{t-1} + \epsilon_t \,, \;\; \epsilon \sim WN(0, \sigma^2)$$

- Y_t not covariance-stationary
- but $\Delta Y_t = Y_t - Y_{t-1}$ is stationary

For RW:

$$\Delta Y_t = Y_t - Y_{t-1} = \beta_0 + \epsilon_t \,, \;\; \epsilon_t \sim WN(0, \sigma^2)$$

Summary

- Trend
  - Deterministic Trend (including non-parametric) / Stochastic Trend
  - Both not covariance-stationary
- Seasonality
  - Seasonal Dummy Model
  - Not covariance-stationary
  - It is possible to have (other kinds) of seasonal processes that *are* stationary
- Cycles
  - We looked at covariance-stationary AR(1)
  - There are many other kinds of covariance-stationary processes

Combining Elements

We often have to mix and match elements to describe time series data

e.g. linear trend with zero-mean AR(1) errors

$$Y_t = \beta_0 + \beta_1 t + \epsilon_t \,, \;\; \epsilon_t = \rho\epsilon_{t-1} + u_t \,, \; |\rho| < 1 \,, \; u_t \sim WN(0, \sigma_u^2)$$

This is a trend-stationary process (which is not covariance-stationary)

$$Y_t = \underbrace{\beta_0 + \beta_1 t}_{\text{deterministic trend}} + \underbrace{\epsilon_t}_{\text{stationary AR(1)}} \nearrow \quad \epsilon_t = \rho\epsilon_{t-1} + u_t \,, \; |\rho| < 1 \,, \; u_t \sim WN(0, \sigma_u^2)$$

Such a process contains deterministic trend and covariance-stationary cycles

Combining Elements

We can estimate the following in a number of ways:

$$Y_t = \beta_0 + \beta_1 t + \epsilon_t \ , \ \ \epsilon_t = \rho\epsilon_{t-1} + u_t \, , \ |\rho| < 1 \, , \ u_t \sim WN(0, \sigma_u^2)$$

Estimate by MLE, or note that

$$\begin{aligned}
Y_t &= \beta_0 + \beta_1 t + \epsilon_t \\
\rho Y_{t-1} &= \rho\beta_0 + \rho\beta_1(t-1) + \rho\epsilon_{t-1} \\
Y_t - \rho Y_{t-1} &= \beta_0 - \rho\beta_0 + \beta_1 t - \rho\beta_1(t-1) + \epsilon_t - \rho\epsilon_t \\
Y_t - \rho Y_{t-1} &= (\beta_0 - \rho\beta_0 + \rho\beta_1) + (1-\rho)\beta_1 t + u_t \\
Y_t &= \beta_0^* + \beta_1^* t + \beta_2^* Y_{t-1} + u_t
\end{aligned}$$

which can be estimated by OLS

Example

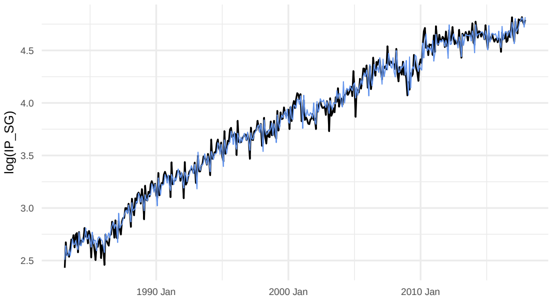
```r
df_ipsg <- ts01 %>%
  select(c(IP_SG, t, tsq, seaslist)) %>%
  mutate(IP_SG_1= lag(IP_SG,1)) %>%
  drop_na()
formula <- paste0(
  "log(IP_SG) ~ log(IP_SG_1) + t + tsq + ",
  seas) %>% as.formula
mdl_ipsg <- lm(formula, data=df_ipsg)
```

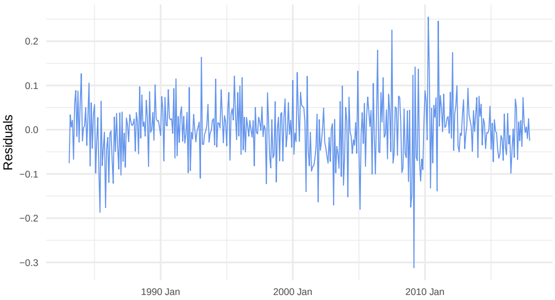
```r
mdl_ipsg %>% summary() %>% coefficients() %>%
  round(4)
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.8431     0.0987  8.5457   0.0000
log(IP_SG_1)   0.6390     0.0380 16.8206   0.0000
t              0.0028     0.0003  8.9616   0.0000
tsq            0.0000     0.0000 -6.4109   0.0000
d02           -0.0438     0.0169 -2.5876   0.0100
d03            0.2355     0.0180 13.0548   0.0000
d04           -0.0154     0.0167 -0.9241   0.3560
d05            0.0551     0.0169  3.2519   0.0012
d06            0.1024     0.0169  6.0405   0.0000
d07            0.0647     0.0167  3.8761   0.0001
d08            0.0806     0.0167  4.8168   0.0000
d09            0.1116     0.0167  6.6841   0.0000
d10            0.0641     0.0166  3.8532   0.0001
d11            0.0264     0.0167  1.5841   0.1139
d12            0.1350     0.0169  7.9961   0.0000
```

Combining Elements

```
df_ipsg_plt <- df_ipsg %>% mutate("Fitted"=fitted(mdl_ipsg), "Residuals"=residuals(mdl_ipsg))
p1 <- autoplot(df_ipsg_plt, log(IP_SG), linewidth=0.5, color="black") +
      autolayer(df_ipsg_plt, Fitted, linewidth=0.3, color="cornflowerblue") + theme1 + xlab("")
p2 <- autoplot(df_ipsg_plt, Residuals, linewidth=0.3, color="cornflowerblue") + theme1 + xlab("")
p1 | p2
```





Time Series Regressions

Time Series Regressions

$$Y_t = \beta_0 + \beta_1 X_{t1} + \dots \beta_k X_{tk} + \epsilon_t$$

Issues include

- Generally cannot assume data are iid
- Have to account for trend, seasonality, cycles
- Previously assumed ϵ_t iid, perhaps no longer appropriate
- May have to consider dynamic specification
- Key assumption for unbiasedness may not hold

Time Series Regressions (Dynamic Specifications)

- Static Regression: $Y_t = \alpha_0 + \alpha_1 X_t + \epsilon_t$
- Distributed Lag Models: $Y_t = \alpha_0 + \beta_0 X_t + \beta_1 X_{t-1} + \ldots \beta_q X_{t-q} + \epsilon_t$
- Autoregressions: $Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t \,, \; |\alpha_1| < 1$
- Autoregressive Distributed Lag (ARDL) models

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_q X_{t-q} + \epsilon_t$$

How should we interpret the parameters of such models?

Time Series Regressions (Dynamic Specifications)

Effect of one-unit one-period "impulse" in X_t

$$
\begin{aligned}
Y_t &= \alpha_0 + \boxed{\beta_0} X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_q X_{t-q} + \epsilon_t \\
Y_{t+1} &= \alpha_0 + \beta_0 X_{t+1} + \boxed{\beta_1} X_t + \beta_2 X_{t-1} + \ldots + \beta_q X_{t-q+1} + \epsilon_{t+1} \\
Y_{t+2} &= \alpha_0 + \beta_0 X_{t+2} + \beta_1 X_{t+1} + \boxed{\beta_2} X_t + \ldots + \beta_q X_{t-q+2} + \epsilon_{t+2} \\
&\vdots \\
Y_{t+q} &= \alpha_0 + \beta_0 X_{t+q} + \beta_1 X_{t+q-1} + \beta_2 X_{t+q-2} + \ldots + \boxed{\beta_q} X_t + \epsilon_{t+q} \\
Y_{t+q+1} &= \alpha_0 + \beta_0 X_{t+q+1} + \beta_1 X_{t+q} + \beta_2 X_{t+q-1} + \ldots + \beta_q X_{t+1} + \epsilon_{t+q+1}
\end{aligned}
$$

Coefficients are called "dynamic multipliers"

Time Series Regressions (Dynamic Specifications)

Effect of a permanent shift in X_t

$$
\begin{aligned}
Y_t &= \alpha_0 + \boxed{\beta_0} X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_q X_{t-q} + \epsilon_t \\
Y_{t+1} &= \alpha_0 + \boxed{\beta_0} X_{t+1} + \boxed{\beta_1} X_t + \beta_2 X_{t-1} + \ldots + \beta_q X_{t-q+1} + \epsilon_{t+1} \\
Y_{t+2} &= \alpha_0 + \boxed{\beta_0} X_{t+2} + \boxed{\beta_1} X_{t+1} + \boxed{\beta_2} X_t + \ldots + \beta_q X_{t-q+2} + \epsilon_{t+2} \\
&\vdots \\
Y_{t+q} &= \alpha_0 + \boxed{\beta_0} X_{t+q} + \boxed{\beta_1} X_{t+q-1} + \boxed{\beta_2} X_{t+q-2} + \ldots + \boxed{\beta_q} X_t + \epsilon_{t+q} \\
Y_{t+q+1} &= \alpha_0 + \boxed{\beta_0} X_{t+q+1} + \boxed{\beta_1} X_{t+q} + \boxed{\beta_2} X_{t+q-1} + \ldots + \boxed{\beta_q} X_{t+1} + \epsilon_{t+q+1}
\end{aligned}
$$

We refer to $\beta_0 + \beta_1 + \cdots + \beta_q$ as "long-run cumulative dynamic multiplier"

Time Series Regressions (Dynamic Specifications)

Interpretation of AR(1)?

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t \,, \;\; |\alpha_1| < 1$$

- A tool for describing "stable cycles"
- β_1 is the autocorrelation of Y_t at lag one
- Can be viewed as "reduced form" expression of cyclical behavior implied by economic interactions
- Member of the ARMA class of models (not covered in this course)
- A useful forecasting tool

Time Series Regressions (Dynamic Specifications)

Consider ARDL(1,1)

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \epsilon_t$$

This implies an "Infinite Distributed Lag Structure". Assume $|\alpha_1| < 1$.

- Lag: $Y_{t-1} = \alpha_0 + \alpha_1 Y_{t-2} + \beta_0 X_{t-1} + \beta_1 X_{t-2} + \epsilon_{t-1}$
- Substitute in ARDL

$$Y_t = \alpha_0(1+\alpha_1) + \alpha_1^2 Y_{t-2} + \beta_0 X_t + (\beta_1 + \alpha_1\beta_0) X_{t-1} + \alpha_1\beta_1 X_{t-2} + \epsilon_t + \alpha_1\epsilon_{t-1}$$

- Repeat with Y_{t-2}, then Y_{t-3}, and so on
- Y_t depends on X_t and infinite number of lags of X_t

Key Assumption for Consistency

Consider simple linear regression

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Key assumption for unbiasedness / consistency is $E(\epsilon_i \mid X_1, X_2, \dots, X_n) = 0$

In time series context, this assumption becomes

$$E(\epsilon_t \mid X_T, X_{T-1}, \dots, X_1) = 0$$

which turns out often to be too strong

Key Assumption for Consistency

E.g. 1: Regressions with lagged dependent variable, such as the AR(1)

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t \,,\; \epsilon_t \overset{iid}{\sim} (0, \sigma^2)\,,\; t = 2, 3, ..., T\,.$$

- Assumption $E(\epsilon_t \mid X_T, X_{T-1}, ..., X_1) = 0$ is $E(\epsilon_t \mid Y_T, Y_{T-1}, ..., Y_1) = 0$
- but this is impossible since ϵ_t must be correlated with Y_t

E.g. 2: Regressions $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$ where noise term may contain information that can predict future X, implies correlation between ϵ_t and future X_s, $s > t$

Key Assumption for Consistency

The weaker assumption

$$E(\epsilon_t \mid X_t) = 0 \qquad \text{"contemporaneous exogeneity"}$$

is much more likely to hold

E.g., for AR(1)

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \epsilon_t$$

this assumption becomes $E(\epsilon_t \mid Y_{t-1}) = 0$, which is possible

Key Assumption for Consistency

If variables are covariance-stationary and weakly dependent, and if contemporaneous exogeneity holds, then

- although OLS estimator is still biased, it will nonetheless be consistent

$$\hat{\beta}_1^{ols} = \beta_1 + \frac{(1/T)\sum_{t=1}^{T}(X_t - \overline{X})\epsilon_t}{(1/T)\sum_{t=1}^{T}(X_t - \overline{X})^2}$$

As long as $cov(X_t, \epsilon_t) = 0$ and variables are cov. stationary weakly dependent, a CLT guarantees that numerator of second term converges to zero

Time Series Regressions, Autocorrelations in noise terms

We continue with simple linear regression case

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

- Assume that noise term satisfies contemporaneous exogeneity
- Assume (for the moment) that $X_t \neq Y_{t-1}$

With time series data, the default and Heteroskedasticity-robust standard errors formulas are often not appropriate

- in numerator of second term of RHS in previous slide, variance of sum is not sum of variance

OLS Properties – Estimator Variances

We give consistent variance estimators for the general k-regressor case $Y_t = X_{t*}\beta + \epsilon_t$ where X_{t*} is the time-t row vector of regressors

Conditional homoskedasticity and iid sample: $\widehat{Var}(\hat{\beta}) = \widehat{\sigma^2}\left(\sum_{t=1}^{T} X_{t*}^{\mathrm{T}} X_{t*}\right)^{-1}$

Conditional heteroskedasticity and iid sample:

$$\widehat{Var}(\hat{\beta}) = \left(\sum_{t=1}^{T} X_{t*}^{\mathrm{T}} X_{t*}\right)^{-1} \left(\sum_{t=1}^{T} \hat{\epsilon}_t^2 X_{t*}^{\mathrm{T}} X_{t*}\right) \left(\sum_{t=1}^{T} X_{t*}^{\mathrm{T}} X_{t*}\right)^{-1}$$

Conditional Heteroskedasticity and Correlation in $X_{t*}\epsilon_t$

$$\widehat{Var}(\hat{\beta}) = \left(\sum_{t=1}^{T} X_{t*}^{\mathrm{T}} X_{t*}\right)^{-1} \left(\sum_{t=1}^{T} \hat{\epsilon}_t^2 X_{t*}^{\mathrm{T}} X_{t*} + \sum_{v=1}^{q}\left(1 - \frac{v}{q+1}\right)(X_{t*}^{\mathrm{T}} X_{t-v,*} + X_{t-v,*}^{\mathrm{T}} X_{t*})\epsilon_t\epsilon_{t-v}\right) \left(\sum_{t=1}^{T} X_{t*}^{\mathrm{T}} X_{t*}\right)^{-1}$$

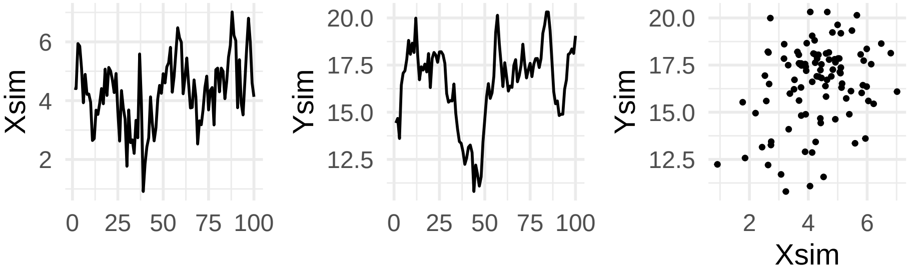
“Heteroskedasticity and Autocorrelation Consistent” or (HAC) var-cov matrix estimators (several kinds, the above is “Newey-West”)

OLS Properties

Simulation Example: $\{X_t, Y_t\}_{t=1}^{100}$ where

$$X_t = 0.8 + 0.8X_{t-1} + \epsilon_t \, , \; \epsilon_t \overset{iid}{\sim} N(0,1)$$

$$Y_t = 0.8 + 0X_t + u_t \, , \; u_t = 0.95u_{t-1} + v_t \, , \; v_t \overset{iid}{\sim} N(0,1)$$



OLS Properties

```
mdlsim <- lm(Ysim~Xsim, data=df)
cat("OLS with Default Standard Errors\n")
mdlsim %>% lmtest::coeftest()
```

```
OLS with Default Standard Errors

t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) 14.24286    0.79751 17.8592 < 2.2e-16 ***
Xsim         0.53330    0.17928  2.9748  0.003692 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

OLS Properties

```r
mdlsim <- lm(Ysim~Xsim, data=df)
cat("OLS with Heteroskedasticity-Robust S.E.\n")
lmtest::coeftest(mdlsim, vcov=sandwich::vcovHC(mdlsim, type="HC2"))
```

```
OLS with Heteroskedasticity-Robust S.E.

t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)
(Intercept) 14.24286    0.85027 16.7509  < 2e-16 ***
Xsim         0.53330    0.18137  2.9404  0.00409 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Not much difference is s.e., since errors are not heteroskedastic

OLS Properties

```
mdlsim <- lm(Ysim~Xsim, data=df)
cat("OLS with Heteroskedasticity and Autocorrelation (HAC) Robust S.E.\n")
lmtest::coeftest(mdlsim, vcov=sandwich::NeweyWest)
```

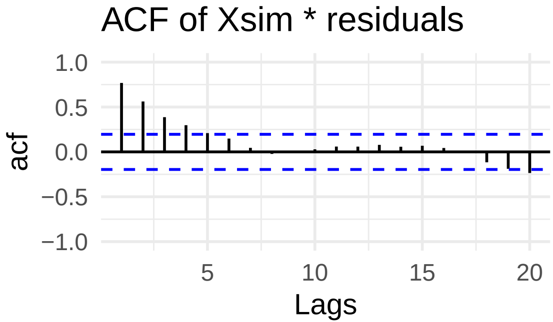
```
OLS with Heteroskedasticity and Autocorrelation (HAC) Robust S.E.

t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) 14.24286    2.06874  6.8848 5.551e-10 ***
Xsim         0.53330    0.34783  1.5332    0.1284
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

OLS Properties

```
df <- df %>%
  mutate("res"=resid(mdlsim), "Xres"=Xsim*res)
ACF(df, Xres) %>%
  autoplot() + theme_minimal() + ylim(-1,1) +
  xlab("Lags") + ggtitle("ACF of Xsim * residuals")
```



- this ACF suggests that the estimator standard errors in third regression are appropriate
- standard errors in first two regressions too small since they ignore correlations in residuals

Time Series Regressions with AR Errors

We now consider the special case of time series regressions with a particular kind of autocorrelation in the error term

"**Regression with (zero-mean) AR(1) errors**"

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \, , \, \epsilon_t = \rho\epsilon_{t-1} + u_t \, , \, |\rho| < 1 \, , \, u_t \overset{iid}{\sim} (0, \sigma^2)$$

How to estimate? A few options:

- Continue with OLS, use HAC variance estimators to get s.e. (ok, but not efficient)
- "Cochrane-Orcutt procedure" (a kind of "Generalized Least Squares)
- Transformation into "Dynamically Complete" ARDL

Time Series Regressions with AR Errors

Cochrane-Orcutt and ARDL aprroaches use the following transformation:

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \, , \, \epsilon_t = \rho\epsilon_{t-1} + u_t \, , \, |\rho| < 1 \, , \, u_t \overset{iid}{\sim} (0, \sigma^2)$$

$$\begin{aligned} \Rightarrow \quad \rho Y_{t-1} &= \rho\beta_0 + \rho\beta_1 X_{t-1} + \rho\epsilon_{t-1} \\ \Rightarrow \quad Y_t - \rho Y_{t-1} &= (1-\rho)\beta_0 + \beta_1(X_t - \rho X_{t-1}) + \epsilon_t - \rho\epsilon_{t-1} \\ \Rightarrow \quad Y_t^* &= \beta_0^* + \beta_1 X_t^* + u_t \end{aligned}$$

where $Y_t^* = Y_t - \rho Y_{t-1}$ and $X_t^* = X_t - \rho X_{t-1}$

- Transformed regression is regression without autocorrelation or heteroskedasticity
- If $E(\epsilon_t \mid X_T, X_{T-1}, \dots, X_1) = 0$ then all requirements for Gauss-Markov Theorem are met in the *transformed regression*, and OLS estimator from transformed regression will be best linear unbiased

Time Series Regressions with AR Errors

But ρ is unknown, and must be estimated

"Cochrane-Orcutt" suggestion:

- Estimate $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$, get $\hat{\epsilon}_t$
- Run regression $\hat{\epsilon}_t = \rho \hat{\epsilon}_{t-1} + u_t$, get $\hat{\rho}$
- Compute $Y_t^* = Y_t - \hat{\rho} Y_{t-1}$ and $X_t^* = X_t - \hat{\rho} X_{t-1}$
- Estimate regression $Y_t^* = \beta_0^* + \beta_1 X_t^* + u_t$ using OLS

$$\hat{\beta}_1^{gls} = \frac{\sum_{t=2}^{T} (X_t^* - \overline{X^*})(Y_t^* - \overline{Y^*})}{\sum_{t=2}^{T} (X_t^* - \overline{X^*})^2}$$

Time Series Regressions with AR Errors

Alternative: since

$$Y_t - \rho Y_{t-1} \;=\; (1-\rho)\beta_0 + \beta_1(X_t - \rho X_{t-1}) + u_t$$

Estimate ARDL version

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + u_t$$

although this is not exactly the same as the original

- original has 3 parameters
- ARDL version has 4 parameters
- To make them exactly the same, have to restrict $\alpha_1\alpha_2 + \alpha_3 = 0$

Time Series Regressions with AR Errors

Both approaches can be extended to multiple regressors, and also higher-ordered AR processes, e.g.,

$$Y_t = \beta_0 + \beta_1 X_{1t} + \ldots \beta_k X_{kt} + \epsilon_t \,, \;\; \epsilon_t = \rho_1 \epsilon_{t-1} + \cdots + \rho_p \epsilon_{t-p} + u_t$$

We omit discuss of higher-ordered AR processes in this course

Most researchers nowadays will either

- use OLS with HAC standard errors
- use ARDL approach, adding lags of Y_t and regressors until residuals do not indicate autocorrelations

Testing for Autocorrelation

To check for autocorrelation in noise terms

- check sample a.c.f. of regresson residuals

A more formal approach:

- regress

$$\hat{\epsilon}_t \text{ on } \hat{\epsilon}_{t-1}, \dots, \hat{\epsilon}_{t-p}, X_{1t}, \dots, X_{kt}$$

  where $X_{1t}, \dots, X_{kt}$ are the regressors in X_t

- test for significance of the coefficients on the lagged residuals

Regression with Non-Stationary Series

Regression with Non-Stationary Series

- Regressions on trending and seasonal series
- Regressions on persistent series (containing random walk characteristics)

Regression with Deterministic Trend

If trend (and seasonality) are deterministic, they should be included in the regression

- Otherwise you will almost always get a significant regression, regardless of variables (basically an omitted variable situation)

E.g., Let Y_t is *log IP_SG* and X_t is *POULTRY_US* and consider the regressions

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 t + \beta_3 t^2 + \epsilon_t$$

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 t + \beta_3 t^2 + \text{seasonal dummies} + \epsilon_t$$

```
mdl1 <- lm(log(IP_SG) ~ POULTRY_US, data=ts01)
mdl2 <- lm(log(IP_SG) ~ POULTRY_US + t + I(tsq/1000), data=ts01)
formula <- paste0("log(IP_SG) ~ POULTRY_US + t + I(tsq/1000) + ", seas) %>% as.formula()
mdl3 <- lm(formula, data=ts01)
stargazer(mdl1, mdl2, mdl3, type="text", omit = seaslist, omit.stat = c("f", "ser", "adj.rsq"))
```

Regression with Deterministic Trend

	Dependent variable: log(IP_SG) (1)	(2)	(3)
POULTRY_US	0.445***	0.045***	0.002
	(0.010)	(0.013)	(0.016)
t		0.007***	0.008***
		(0.0004)	(0.0004)
I(tsq/1000)		-0.004***	-0.006***
		(0.001)	(0.001)
Constant	1.065***	2.388***	2.484***
	(0.062)	(0.042)	(0.051)
Observations	420	420	420
R2	0.833	0.976	0.982

- t^2 divided by 1000 for scaling purposes
- Models 1 and 2, no seasonal dummies
- Seasonal dummies in Model 3, but omitted in output
- Both log IP_SG and POULTRY_US clearly not related
  - POULTRY significant in regressions without for trend and seasonality
- Note high R^2 in all three regressions (always the case with trending series)

Regression with Deterministic Trend

E.g., $Y_t \sim$ log(ELEC_SG), $X_t \sim$ log(IP_SG)

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

vs. $$Y_t = \beta_0 + \beta_1 X_t + \beta_2 t + \beta_3 t^2 + seas.\,dummies + \epsilon_t$$

```
mdl4 <- lm(log(ELEC_GEN_SG) ~ log(IP_SG), data=ts01)
formula <- paste0("log(ELEC_GEN_SG) ~ log(IP_SG) + t + I(tsq/1000) + ", seas) %>%
            as.formula()
mdl5 <- lm(formula, data=ts01)
```

Regression with Deterministic Trend

```r
mdl4 %>% summary() %>% coefficients() %>% round(4)
```

```
            Estimate Std. Error  t value Pr(>|t|)
(Intercept)   4.5102     0.0290 155.5907        0
log(IP_SG)    0.8361     0.0075 111.5611        0
```

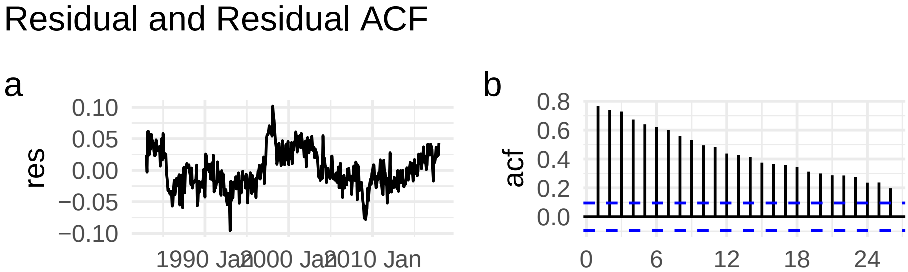
```r
mdl5 %>% summary() %>% coefficients() %>% round(4)
```

```
             Estimate Std. Error  t value Pr(>|t|)
(Intercept)    6.2345     0.0429 145.3252   0.0000
log(IP_SG)     0.0888     0.0170   5.2209   0.0000
t              0.0077     0.0001  54.8456   0.0000
I(tsq/1000)   -0.0089     0.0002 -58.6607   0.0000
d02           -0.0880     0.0076 -11.5928   0.0000
d03            0.0330     0.0076   4.3205   0.0000
d04            0.0240     0.0074   3.2424   0.0013
d05            0.0640     0.0074   8.6535   0.0000
d06            0.0313     0.0074   4.2070   0.0000
d07            0.0552     0.0074   7.4411   0.0000
d08            0.0457     0.0074   6.1534   0.0000
d09            0.0146     0.0075   1.9333   0.0539
d10            0.0458     0.0075   6.1422   0.0000
d11           -0.0009     0.0074  -0.1280   0.8982
d12           -0.0035     0.0075  -0.4607   0.6452
```

Regression with Deterministic Trend

Model is dynamically incomplete

```
dat <- ts01 %>% select() %>% mutate("res"=residuals(mdl5))
p1 <- dat %>% autoplot(res) + theme_minimal() + xlab("")
p2 <- dat %>% ACF(res) %>% autoplot() + theme_minimal() + xlab("Lag")
(p1 | p2) + plot_annotation(tag_levels = 'a',
  title="Residual and Residual ACF ")
```



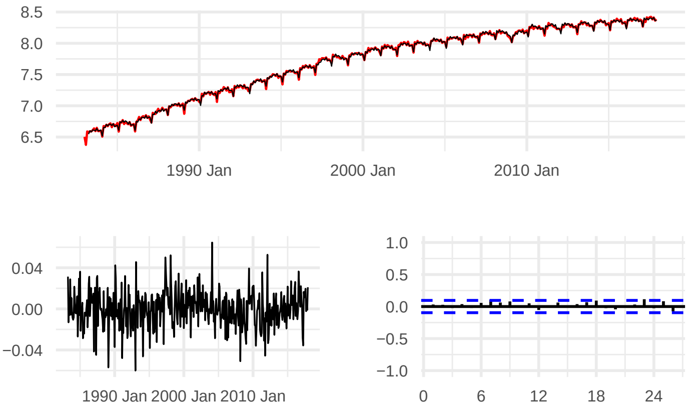
Regression with Deterministic Trend

$$Y_t = \alpha_0 + \beta_0 X_t + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \gamma_3 Y_{t-3} + seas.\ dummies + \delta_1 t + \delta_2 t^2 + \epsilon_t$$

```
formula <- paste0("log(ELEC_GEN_SG) ~ log(IP_SG) + ",
                  "log(lag(ELEC_GEN_SG,1)) + log(lag(ELEC_GEN_SG,2)) + ",
                  "log(lag(ELEC_GEN_SG,3)) + t + tsq + ", seas) %>%
          as.formula()
mdl6 <- lm(formula, data=ts01)
dat <- ts01 %>% select() %>% filter_index("1983M4"~.) %>%
  mutate("fit"=fitted(mdl6), "res"=residuals(mdl6))
p0 <- ts01 %>% autoplot(log(ELEC_GEN_SG), color="red", size=0.4) +
      autolayer(dat, fit, size=0.2) + theme1 + xlab("") + ylab("")
p1 <- autoplot(dat, res, size=0.3) + theme1 + xlab("") + ylab("")
p2 <- ACF(dat, res) %>% autoplot() + theme1 +
  ylim(c(-1,1)) + xlab("") + ylab("")
```

Regression with Deterministic Trend

```
p0 / (p1 | p2)
```



Regression with Persistent Series (A Warning)

We end with a remark on regression with persistent series

Simulate 200 pairs of random walks $\{X_t^{(r)}, Y_t^{(r)}\}_{t=1}^{100}$, $r = 1, 2, ..., 200$:

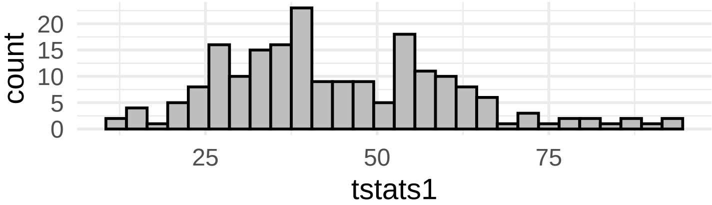
$$X_t^{(r)} = \alpha_X + X_{t-1}^{(r)} + u_t^{(r)}$$
$$Y_t^{(r)} = \alpha_Y + Y_{t-1}^{(r)} + v_t^{(r)}$$

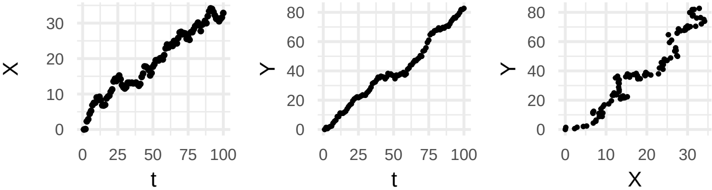
where $u_t^{(r)}$ and $v_t^{(r)}$ are independent Normal$(0,1)$ noise terms, $\alpha_X = 0.5$ and $\alpha_Y = 0.8$.

For each replication r

- regress $Y_t^{(r)}$ on $X_t^{(r)}$, with intercept
- collect the t-statistic on the coefficient of $X_t^{(r)}$

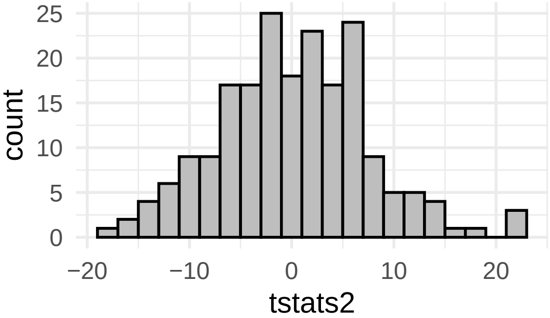
Regression with Persistent Series (A Warning)





Regression with Persistent Series (A Warning)

Repeat with $\alpha_X = \alpha_Y = 0$ (i.e., no drift)



Regression with Persistent Series

- Regressions with persistent series not always spurious, can give very good results
- Simple (maybe incomplete) solution for persistent series is to take first differences (i.e., transform to stationarity)
- Stochastic vs Deterministic Trend?
- Will have to leave further details to “next course”

Roadmap

- (Previous) Session 1: Statistics Review
- (Previous) Session 2: Simple Linear Regression
- (Previous) Session 3: Estimator Standard Errors; Multiple Linear Regression
- (Previous) Session 4: Matrix Algebra
- (Previous) Session 5: OLS using Matrix Algebra
- (Previous) Session 6: Hypothesis Testing
- (Previous) Session 7: Prediction
- (Previous) Session 8: Instrumental Variable Regression
- (Previous) Session 9: GLS / Panel Data Regressions
- (Previous) Session 10: Maximum Likelihood Estimation / Limited Dep. Var. Models
- **This Session 11-12: Introduction to Time Series / Time Series Regressions**