

# Applied Financial Econometrics

## List of Exercises - Effective: 15.4.2018<sup>1</sup>

This list encompasses 35 pages and consists of 51 exercises with 692 points in total.

### Exercises

Number	Exercise	Page
1	( <i>Conditional Expectations</i> )	3
2	( <i>Revision</i> )	3
3	( <i>Working With Data</i> )	4
4	( <i>Some Calculations</i> )	4
5	( <i>First Simulation</i> )	5
6	( <i>AR(3)-Process</i> )	5
7	( <i>AR(3) Construction</i> )	6
8	( <i>Random Walk</i> )	6
9	( <i>Autocovariance of MA- and AR-Processes</i> )	7
10	( <i>Finite Sample Bias in AR Processes</i> )	8
11	( <i>Deterministic vs. Stochastic Trend</i> )	9
12	( <i>Manpower Forecast</i> )	10
13	( <i>Manpower Improvement</i> )	10
14	( <i>Manpower h-Step Forecast</i> )	11
15	( <i>Manpower - Out of Sample</i> )	11
16	( <i>BMW Forecast</i> )	11
17	( <i>BMW Model Check</i> )	12
18	( <i>BMW Improvement</i> )	12
19	( <i>BMW h-Step Forecast</i> )	13
20	( <i>BMW - Unit Root Testing</i> )	13

---

<sup>1</sup> Terms of Use: The contents of this list may be used online for personal use either at home or at an educational institution, but not for commercial purposes; the worksheets may be also printed for personal use offline. For educators, the materials may be used online at educational institutions, either for individual or teacher-guided instruction; teachers are permitted to make copies of a limited number of pages for students for incidental use (not as part of a textbook, etc.), and the following copyright notice must appear on each copy:

© Rolf Tschernig, University of Regensburg, Apr, 2018. All rights reserved.

Hereby, I would like to thank Florian Brezina, Barbara Richter, Joachim Schnurbus, Thomas Schreck, Johannes Vilsmeier and Stefan Rameseder for creating most of the exercises.

21	( <i>Stock Return Forecast</i> )	13
22	( <i>ARIMA Representation</i> )	14
23	( <i>Lag Length Selection</i> )	14
24	( <i>ADF Simulation</i> )	16
25	( <i>Exchange Rate Analysis</i> )	16
26	( <i>Zloty Exchange Rates</i> )	17
27	( <i>Consumer Price Analysis</i> )	18
28	( <i>Maximum Likelihood (Bernoulli)</i> )	21
29	( <i>Maximum Likelihood (Normal)</i> )	21
30	( <i>Maximum Likelihood (GARCH)</i> )	22
31	( <i>Log-Likelihood with Normal Distribution</i> )	22
32	( <i>Log-Likelihood with DGP</i> )	22
33	( <i>Joint Log-Likelihood</i> )	23
34	( <i>Conditional Heteroskedasticity</i> )	23
35	( <i>ARCH(1)</i> )	23
36	( <i>Representations of GARCH(1,1)</i> )	23
37	( <i>AR and GARCH</i> )	24
38	( <i>Siemens</i> )	25
39	( <i>Stock Price Choice to model</i> )	27
40	( <i>Unconditional Expected Value of a stationary Time Series</i> )	28
41	( <i>Asset pricing formula and Sharpe ratio</i> )	29
42	( <i>Expected Wealth</i> )	29
43	( <i>Stochastic discount factor</i> )	29
44	( <i>Stochastic discount factor and risk-neutral probabilities</i> )	30
45	( <i>CAPM</i> )	30
46	( <i>Dynamic regression model</i> )	31
47	( <i>Exponential Family</i> )	32
48	( <i>Inverse Gamma</i> )	33
49	( <i>Bayesian AR(p)</i> )	33
50	( <i>Regression with autocorrelated errors</i> )	34
51	( <i>The Univariate Normal Gibbs Sampler</i> )	35

## 1 Introduction

### 1. Exercise (7 Points) (*Conditional Expectations*)

It is March 31. On April 3 you need to bake a strawberry cake and you would like to know how many strawberries will have grown in your backyard by then. This will solely depend on the weather in the first two days of April.

The following table gives the number of strawberries harvested for all kinds of weather, together with the probability of each type of weather on either day. Probabilities for April 2 are conditional on the weather on April 1.

April 1	prob	April 2	prob	Number of Strawberries
sun	0.5	sun	0.7	10
sun	0.5	rain	0.2	8
sun	0.5	ice	0.1	1
rain	0.3	sun	0.3	7
rain	0.3	rain	0.5	5
rain	0.3	ice	0.2	1
ice	0.2	sun	0.3	2
ice	0.2	rain	0.25	1
ice	0.2	ice	0.45	0

- (2 Points) What is the unconditional probability of a rainy April 2? Is the unconditional probability of icefall on April 2 the same as of icefall on April 1?
- (1 Point) March 31: the weather channel says it will rain on April 1 and the sun will shine on April 2. Assuming the weather forecast is correct, how many strawberries do you expect to harvest?
- (2 Points) How many strawberries do you expect to harvest if it rains on April 1? What if there is sun? What if it ices? Combining these three cases, what do you expect on average? Calculate the variance of strawberries harvested, conditional on the weather on April 1.
- (2 Points) Back to March 31. You do **not** watch the weather channel. How many strawberries do you expect to harvest?

### 2. Exercise (8 Points) (*Revision*)

- (5 Points) Explain all basic regression assumptions (in terms of Wooldridge the MLR assumptions, in terms of Davidson/MacKinnon the A and B assumptions) you are familiar with from introductory courses and show that the OLS estimator (reminder:  $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ ) is unbiased given a subset of these assumptions. When do we need the additional assumptions?
- (3 Points) Explain the role of conditional expectations in regression analysis.

**3. Exercise** (10 Points) (*Working With Data*)

Download the stock market data (file: `ie_data.xls`) from 1871 to present from [here](#) and import it to R or EViews.

The data may cause difficulties, if you try to insert them directly, a preparation in Excel or an editor like Notepad++ may be helpful.

**Attention:** All presentations will be hold with R.

- (a) (4 Points) Perform the following transformations, applying the definitions of real prices and returns. For each of the transformations, document how you went about the task.
- Generate a new series `real_returns` from the series of real prices.
  - Generate the log from both the real price and the real return series.
  - Take the nominal price series and transform it to real prices: which base year do you need to arrive at the same values as Shiller?
- (b) (4 Points) Look at the line graphs for real prices, the log of real prices and the log of real returns
- Compare the two graphs on prices. Pay particular attention to any unusual patterns that you might find in one or both of them.
  - Discuss the graph of the log of real returns: What does it look like? How do you account for anything extraordinary? What does the graph tell you?
- (c) (2 Points) Estimate the AR(1) model  $y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t$  using the real prices series. From what you know about time series: how would you interpret the results?

**4. Exercise** (11 Points) (*Some Calculations*)

Consider an AR(1)-process  $y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t$ . The  $u_t$  are white noise, i.e. mean zero, homoscedastic variance, and uncorrelated over time,  $|\alpha_1| < 1$ .

- (a) (3 Points) First, write  $y_t$  as a function of the parameters, current and lagged errors and the initial value  $y_0$ . Calculate the unconditional expectation of  $y_t$  assuming that the process started a (very) long time ago.
- (b) (3 Points) Now calculate the unconditional variance of  $y_t$ , where again we assume that the process has run since ever.
- (c) (5 Points) Finally, compute the (function of) autocovariances  $\text{Cov}(y_t, y_{t-k})$ , where  $k = 1, \dots, t$ .

## 2 Autoregressive and Moving Average Processes

### 5. Exercise (8 Points) (*First Simulation*)

This exercise probably contains some new content, try to solve it with the help of the lecture notes. Don't panic, without some experience in programming this sheet is not easy, but you will get familiar to the programming-procedures offered by R. If you are not familiar with this software, I recommend you to have a look at [here](#).

To access the help page for a particular function, use the command `?functionname`. Besides basic R commands (e.g. how to define a matrix) you will need the functions `lm`, `rnorm` and `plot`. A loop can be programmed by using a `for` construction, e.g. `for (t in 1:T){}`.

Before starting this sheet, try to become familiar to the procedures and results of the Excel-File for simulating a realisation of an AR(1)-process (see [Homepage](#)).

Consider the AR(1)-process  $y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t$  with  $\{u_t\}$  being white noise with variance 1.

- (a) (1 Point) What is meant by white noise? By Gaussian white noise?
- (b) (4 Points) Use R (or EViews) to create one realization (with 100 observations) of this process for each of the following configurations. Use for simplicity  $y_0 = 0$ . Describe the main insights of each of the graphs.
  - $\alpha_0 = 0$  and  $\alpha_1 = 0.9$
  - $\alpha_0 = 20$  and  $\alpha_1 = 0.75$
  - $\alpha_0 = 0.2$  and  $\alpha_1 = 1$
  - $\alpha_0 = 0$  and  $\alpha_1 = -1.1$

- (c) (3 Points) What changes for each of the realizations if we had a starting value  $y_0 = 50$ ?

### 6. Exercise (21 Points) (*AR(3)-Process*)

Consider an AR(3)-process  $y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + u_t$  with  $\{u_t\}$  being white noise.

- (a) (2 Points) Rewrite the process using a lag polynomial based on the lag operator  $L$ .
- (b) (3 Points) Show that  $(1 - \lambda_1 z)(1 - \lambda_2 z)(1 - \lambda_3 z)$  can be rewritten as the lag polynomial using the lag operator, substituting  $z$  for  $L$ . Why can you substitute  $z$  for  $L$  in the lag polynomial?
- (c) (5 Points) Solve the characteristic equation of the AR(3)-process using  $\alpha_1 = 0.1$ ,  $\alpha_2 = 0.22$  and  $\alpha_3 = -0.04$  and knowing that  $\lambda_1 = 0.2$ . Is the process stationary? Why/why not?
- (d) (5 Points) Use R (or EViews) to simulate 1000 observations of the process given in (c) and then estimate the parameters of the process without a constant.

Repeat the process 10 times, and save the estimation results each time.

- Compare the results to the  $\alpha$ 's given in (c). Do you see anything unexpected? Why do you think this happens?
- Take the average of the results for each parameter. Is it close to the parameters in (c)? Why?

Hint: The information above is not sufficient to simulate the process, since we did not specify the distribution of  $u_t$ . Use e.g. the standard normal distribution for the simulation.

- (e) (2 Points) Discuss the properties of the process in case  $\lambda_2 = 1$ ,  $|\lambda_1| < 1$  and  $|\lambda_3| < 1$ ?
- (f) (4 Points) Now do another 100 simulations. Are the results changing compared to the case of 10 simulations? What would you expect to happen for increasing the replication number to 1000 and 10,000 and further to infinity? What do you think about “increasing the number of replications” versus “increasing the number of observations”?

**7. Exercise (8 Points) (*AR(3) Construction*)**

Construct a stationary and causal AR(3) process  $y_t$  for a given expected value  $\mu = E[y_t] = 20$  and variance  $\gamma_0 = \sigma_y^2 = \text{Var}(y_t) = 10$  and the first three autocorrelations  $\rho_1 = 0.9, \rho_2 = 0.81$  and  $\rho_3 = 0.9$ , i. e.

- (a) (3 Points) calculate the parameters  $c, \alpha_1, \alpha_2$  and  $\alpha_3$ ,
- (b) (3 Points) calculate the variance  $\sigma_u^2$  of the White Noise  $u_t$  and
- (c) (2 Points) calculate the additional autocorrelations  $\rho_4$  and  $\rho_5$ .

**Hint:**

$$\begin{pmatrix} 1 & 0.9 & 0.81 \\ 0.9 & 1 & 0.9 \\ 0.81 & 0.9 & 1 \end{pmatrix}^{-1} = \frac{1}{19} \begin{pmatrix} 100 & -90 & 0 \\ -90 & 181 & -90 \\ 0 & -90 & 100 \end{pmatrix}$$

**8. Exercise (11 Points) (*Random Walk*)**

Write a new program for estimating an AR(1)-process with a random walk.

- (a) (4 Points) Simulate the random walk process 100 times. What do you observe when you take the average of the estimated  $\alpha$ 's? Is this in line with theory? What would you expect for increasing the sample size to infinity?
- (b) (3 Points) Look at the line graphs of five randomly chosen realizations of your simulations each for the AR(3)-process and the AR(1) random walk. What differences do you notice? How does it fit with what you know about random walks?
- (c) (4 Points) Show why the OLS estimator of an AR(1)-process is biased. Use your own words to explain why.

**9. Exercise (33 Points)** (*Autocovariance of MA- and AR-Processes*)

This exercise should give some insights into autocovariances of an MA(2)- and an AR(1)-process respectively.

Consider first the following MA(2)-Process

$$y_t = u_t + 0.9u_{t-1} - 0.3u_{t-2}, \quad u_t \sim i.i.d.(0, 1) \quad (4)$$

- (a) (3 Points) Calculate the variance and the expected value of  $y_t$ .
- (b) (5 Points) Calculate the autocorrelation function for  $k = 1, 2, 3, \dots, 30$ .
- (c) (2 Points) Given your results from above, show that the process is (weakly) stationary.

Consider now the following (weakly) stationary AR(1)-Process

$$y_t = 0.9y_{t-1} + u_t, \quad u_t \sim i.i.d.(0, 1), \quad y_0 = 0, \quad t = 1, 2, 3, \dots \quad (5)$$

- (d) (5 Points) Determine the expected value and the autocorrelation function for the process under the assumption that the moments ( $\mathbf{E}()$ ,  $\mathbf{Var}()$ ,  $\mathbf{Cov}()$ ) are independent of  $t$ .

As a special case, consider the following AR(1)-Process

$$y_t = y_{t-1} + u_t, \quad u_t \sim i.i.d.(0, 1), \quad y_0 = 0, \quad t = 1, 2, 3, \dots \quad (6)$$

- (e) (3 Points) Write the process as a function of the starting value and its error terms. Determine the expected value and the variance.
- (f) (5 Points) Calculate the autocovariance and autocorrelation function. Is the process (weakly) stationary?

Since in reality, you probably do not know what the underlying DGP is, you have to apply a model by figuring out the model type. One (old-fashioned) way is the (partial) autocovariance function.

- (g) (4 Points) Plot the autocovariance and autocorrelation function ( $k = 0 \dots 30$ ) for the processes (1) and (2) in R using the `ARMAacf()` command.
- (h) (4 Points) Define the autocovariance and autocorrelation functions for the process (3) in R using the `function()` command and plot the respective functions for  $k = 0 \dots 30$ . If your functions depend on  $t$ , make two different plots with  $t = 100$  and  $t = 1000$  respectively.
- (i) (2 Points) What do the plots tell you about the characteristics of the underlying processes?

**Remark:** The following table comprises some behaviours of the (theoretical) partial and general autocorrelation function with respect to the type of process.

process type	autocorrelation function	partial autocorrelation function
AR( $p$ )	infinite	finite: $\gamma(k) = 0$ for $k > p$
MA( $q$ )	finite: $\rho(k) = 0$ for $k > q$	infinite
ARMA( $p, q$ )	infinite	infinite

By knowledge of these theoretical properties we expect a similar behavior for the empirical counterparts which however will be empirically hard to distinguish.

**10. Exercise** (10 Points) (*Finite Sample Bias in AR Processes*)

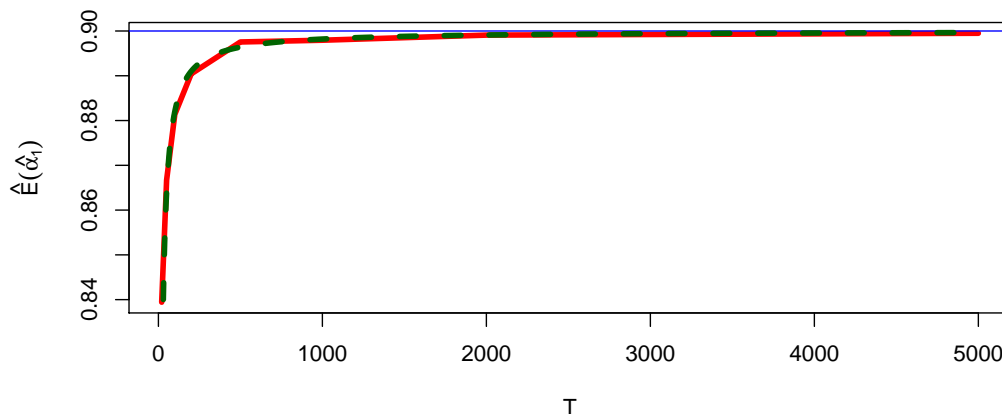
Consider

$$y_t = \alpha_1 y_{t-1} + u_t, \quad u_t \sim IID(0, \sigma^2), \quad |\alpha_1| < 1, \quad t = 1, \dots, T,$$

fulfilling the assumptions (A).

- (2 Points) Write down **explicitly** the OLS estimator  $\hat{\alpha}_1$  for  $\alpha_1$  in the case above. Use the scalar sum notation.
- (2 Points) Is the OLS estimator an unbiased estimator in the case above?
- The following simulation should investigate the bias in finite samples and produces the graph below the code.

```
###
alpha1      <- 0.9
sigma2      <- 1
T           <- c(20, 50, 100, 200, 500, 1000, 2000, 5000)
hurwiczBias <- function(alpha1, T) -alpha1 * 2/ T
N           <- 1000
mc          <- matrix(NA, ncol = length(T), nrow = N)
colnames(mc) <- c(paste0(T))
###
set.seed(1)
for(n in 1:N){
  for(t in T){
    ts.sim      <- arima.sim(list(order=c(1,0,0), ar=alpha1, sd=sqrt(sigma2) ), n=t)
    mc[n, as.character(t)] <- ar.ols(ts.sim, aic=FALSE, order.max=1, demean=FALSE )$ar
  }
}
###
plot(x = T, y = colMeans( mc ), type = "l", col = "red", lwd = 2)
abline(h = alpha1, col = "blue")
xx <- seq(20, 5000, by = 1)
lines(xx, alpha1 + hurwiczBias(alpha1, xx), col = "darkgreen")
```



- (3 Points) How many random variables or shocks are **totally** generated using `arima.sim` for the complete simulation?
- (3 Points) Write down the Hurwicz Bias of the OLS estimator in AR(1) processes for finite samples and interpret it, i. e. approximate

$$E[b] := E[\hat{\alpha}_1 - \alpha_1 \mid y_0, \dots, y_T]$$

and interpret this function depending on all parameters showing up.

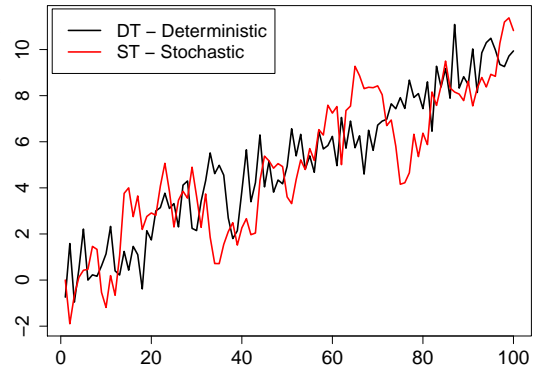


### 11. Exercise (14 Points) (*Deterministic vs. Stochastic Trend*)

This exercise is about distinguishing deterministic (then called trend-stationary) and stochastic (then called difference-stationary) trends (processes). Define  $u_t, v_t \sim WN(0, \sigma^2)$  and consider the two (independent) DGPs

$$\text{DT} \quad y_t = \beta t + u_t \text{ and}$$

$$\text{ST} \quad x_t = \beta + x_{t-1} + v_t.$$



- (a) (2 Points) Calculate  $E[y_t], E[x_t], \text{Var}(y_t)$  and  $\text{Var}(x_t)$ .
- (b) (2 Points) Apply the following R-Code to get a realization of each DGP

```
stochTrend <- function(T, beta , y.start = 0, sigma = 1){
  y <- rep(0, T);  y[1] <- y.start
  for(t in 2:(T)) y[t] <- beta + y[t-1] + rnorm(1, 0 ,sigma)
  return(ts(y))
}
set.seed(5)
sigma <- 1;  T <- 100;  beta <- 0.1
Y_dt <- ts(beta * 1:100 + rnorm(T, 0, sigma))
X_st <- stochTrend(T, beta,sigma = sigma)
plot.ts(cbind(Y_dt, X_st ),plot.type = "single",col = c("black", "red"),lwd = 2,ylab = "")
```

Use this code to generate 100 replications of each DGP and plot them into one graph.

Plot  $\pm$  standard deviations and the line  $\beta \cdot t$ .

Use the vectors `Y_dt` and `Y_st` from the code above to answer the following questions

- (c) Model both time series with deterministic trends.
- (2 Points) Model  $y_t$  theoretically and practically with a deterministic trend, i. e. what statistic behaviour do the residuals have? Plot them. Plot the (partial) autocorrelation function of  $y_t$  and of the residuals  $\hat{u}_t$ .
  - (2 Points) Model  $x_t$  theoretically and practically with a deterministic trend, i. e. what statistic behaviour do the residuals have? Plot them. Plot the (partial) autocorrelation function of  $y_t$  and of the residuals  $\hat{v}_t^*$ .
- (d) Model both time series with stochastic trends.
- (2 Points) Calculate  $\Delta y_t, \text{Cov}(w_t, w_{t-1})$  for  $w_t := u_t - u_{t-1}$  and  $\rho_{1,w_t}$
  - (2 Points) Model  $y_t$  theoretically and practically with a stochastic trend, i. e. what statistic behaviour do the residuals have? Plot them. Plot the (partial) autocorrelation function of  $y_t$  and of the residuals  $\hat{w}_t$ .
  - (2 Points) Model  $x_t$  theoretically and practically with a stochastic trend, i. e. what statistic behaviour do the residuals have? Plot them. Plot the (partial) autocovariance function of  $y_t$  and of the residuals  $\hat{v}_t$ .

### 3 Applications with Unit Roots and Forecasting

#### 12. Exercise (10 Points) (*Manpower Forecast*)

Your task is to forecast the price of stocks of Manpower Inc for June 2011 using monthly data and starting at May, 1988 and ending at May, 2011.

- (a) (3 Points) Download data from the internet (e.g. from Yahoo Finance). Then correct the data for possible stock splits and import the data into R. (Check if imported correctly). If there is no need for a stock split correction, just explain how this is normally done.
- (b) (1 Point) Plot the time series of the data and judge whether the series is stationary.
- (c) (3 Points) Estimate an AR(5) model and try to determine the roots of the lag polynomial. Use e.g. `polyroot()`. Is any root close to one? Are your results in line with your visual impression of the time series? Are all parameters significant? Be precise on which asymptotic distribution to use.
- (d) (3 Points) Forecast the stock price for June 2011 using your estimated model. Explain which measure you implicitly used to evaluate the prediction error. Is this measure always appropriate for financial time series? If not, why?

#### 13. Exercise (19 Points) (*Manpower Improvement*)

In the last sheet your task was to forecast the price of stocks of Manpower Inc for one period using monthly data and an AR(5) model. Now you should check if you can improve on your previous model.

- (a) (5 Points) Describe the ADF-test. What is the major difference between the ADF- and the KPSS-test? Apply both tests to the price series using R. (Hint: Use e.g. functions `ur.df()` and `ur.kpss()` from the `urca` package. Look at the help pages for information on the specification of the deterministic component and lag length selection.)
- (b) (2 Points) Which assumptions have to hold for the AR model for the OLS estimator to be consistent and asymptotically normal? Explain these assumptions.
- (c) (3 Points) Transform the time series appropriately such that the assumptions for the OLS estimator are met as closely as possible. Use your considerations from the previous exercises.
- (d) (4 Points) Choose a maximum order  $n_{max}$  of your candidate AR( $n$ ) models and select an appropriate order using suitable model selection methods. Justify all your choices.
- (e) (5 Points) Now the task becomes more demanding since you are asked to find the optimal subset model. You should proceed as follows:
  - Start with  $n = n_{max}$ . Save the value of the selection criterion. Next you estimate a smaller model where you leave out that particular lag that exhibits the largest  $p$ -value in the full model. Again save the value of the selection criterion. (Hint: The AIC and SIC can be extracted from an `lm` or `dynlm` object by functions `AIC()` and `extractAIC()`.)

- Continue until the selection criterion does not decrease any more. Document the order in which you dropped lags and the values of the selection criteria at each step. What is the order of the final model?

**14. Exercise (9 Points)** (*Manpower h-Step Forecast*)

Your task in this exercise is the computation of  $h$ -step forecasts up to one year of the stock price of Manpower Inc.

- (5 Points) First, compute your forecast for the logged stock price in Eviews assuming that the DGP of the stock price is a simple random walk with drift. Compute also prediction intervals given a 95% significance level. Second compute your forecast for the stock price itself (after an estimation in logs). Explain the shape of all curves.
- (4 Points) Now compute your forecasts for the stock price and the logged stock price assuming that the DGP of the logged stock price is an AR(1) process with  $-1 < \alpha \leq 1$ . Can you explain the change in the shape of the predictions compared with the random walk case.

**15. Exercise (11 Points)** (*Manpower - Out of Sample*)

It is often useful to check whether an estimated model can predict well out-of-sample. Conduct the following experiment using the Manpower stock price data.

- (2 Points) Using an appropriately selected model order, estimate an AR( $p$ ) model using data only until December 2003.
- (3 Points) Compute 3-period forecasts for each month beginning with March 2004 until May 2009. For forecasting the April 2004 value you do not reestimate the model but you can also use the January 2004 observation.
- (2 Points) Compute the prediction error for each forecast using the available realizations and estimate the MSE based on these out-of-sample prediction errors.
- (4 Points) Compute the MSE based on out-of-sample prediction errors with the estimate of the “true” MSE  $\sigma_y^2(h) = \sigma^2 \sum_{j=0}^{h-1} \phi_j^2$  that you obtain by inserting the parameter estimates based on the sample until December 2003. Try to comment on your results.

**16. Exercise (11 Points)** (*BMW Forecast*)

Your task is to forecast the price of stocks of BMW AG for 02 May 2012 using daily data from 01 January 2003 to 30 April 2012.

- (3 Points) Download data from the internet (e.g. from Yahoo Finance). If necessary, correct the data for possible stock splits and dividends. Import the data into R e.g. using `read.csv()` (Check if imported correctly). If there is no need for a stock split correction, just explain how this is normally done.
- (1 Point) Plot the time series of the data and judge whether the series seems stationary.
- (4 Points) Estimate an AR(4) model and try to determine the roots of the lag polynomial. Is any root close to one? Are your results in line with your visual impression of the time

series? Are all parameters significant? Be precise on which asymptotic distribution to use. Hints: After estimating an  $AR(p)$  model by function `ar.ols()` and assigning the function value to variable `name` you can access the vector of point estimates by `name$ar[, ,1]` and the asymptotic standard errors of the estimator by `name\asy.se.coef\ar`. For more details type `?ar`. Roots of a polynomial are easily computed by `polyroot()`.

- (d) (3 Points) Forecast the stock price for 02 May 2012 using your estimated model. Explain which measure you implicitly used to evaluate the prediction error. Is this measure always appropriate for financial time series? If not, why?

**17. Exercise** (5 Points) (*BMW Model Check*)

In this exercise we will check if predicting the price of stocks of BMW AG works better with an  $AR(1)$  or an  $AR(4)$  model.

- (a) (5 Points) In both models we include a constant term. Predict the 1-day-ahead stock price for 03 January 2011 using data from 01 January 2003 until 30 December 2010. Then, reestimate the model using data until 03 January 2011 and predict the value for 04 January 2011. Then, reestimate the model using data until 04 January 2011 and predict the value for 05 January 2011. Repeat this procedure until you predict the stock price for 30 April 2012. Compare both models using the MSE.

**18. Exercise** (16 Points) (*BMW Improvement*)

In the last sheet your task was to forecast the price of stocks of BMW AG for one period using daily data and an  $AR(4)$  model. Now you should check if you can improve on your previous model.

- (a) (1 Point) Create a time series for the log return on the stock price.
- (b) (2 Points) Choose a maximum order  $p_{max}$  of your candidate  $AR(p)$  models and select an appropriate order using the AIC. Justify the choice of  $p_{max}$ .
- (c) (5 Points) Now, using the same  $p_{max}$  as before, choose the lag length by SIC and HQ. Do these criteria choose the same  $p$  as AIC? Explain.
- (d) (5 Points) Now the task becomes more demanding since you are asked to find the optimal subset model. You should proceed as follows:
- Start with  $p = p_{max}$ . Estimate an  $AR(p_{max})$  and look at the  $p$ -values or  $t$ -statistics of the lags. Estimate then a smaller model where you leave out that particular lag that is least significant.
  - Continue until the largest  $p$ -value is smaller than 0.1. Document the order in which you dropped lags. How does the final model look like?
- (e) (3 Points) Predict the log return for 02 May using the final model.

**19. Exercise** (9 Points) (*BMW h-Step Forecast*)

Your task in this exercise is the computation of  $h$ -step forecasts up to 20 days of the stock price of BMW AG.

- (a) (5 Points) First, compute your forecast for the logged stock price in R assuming that the DGP of the stock price is a simple random walk with drift. Compute also prediction intervals given a 95% significance level, assuming the error terms are normally distributed. Show that normality does not seem to be a bad approximation. Second compute your forecast for the stock price itself (after an estimation in logs). Explain the shape of all curves. (Hint: For simplicity ignore the estimation error of the parameter(s) when computing prediction intervals.)
- (b) (4 Points) Now compute your forecasts for the stock price and the logged stock price assuming that the DGP of the logged stock price is an AR(1) process with  $-1 < \alpha \leq 1$  and deterministic linear trend. Can you explain the change in the shape of the predictions compared with the random walk case? Shortly explain how a confidence interval around the logged stock price should look like.

**20. Exercise** (12 Points) (*BMW - Unit Root Testing*)

This exercise is about unit root testing. You can use functions `ur.df()` and `ur.kpss()` from package `urca` to perform ADF- and KPSS tests, respectively. Also, function `summary()` applied to the object returned by `ur.df()` and `ur.kpss()` might be very useful.

- (a) (4 Points) Perform an ADF test on the time series for the stock price of BMW AG. Explain which pair of hypotheses you are using. Does your result change with the information criterion used to select the lag length (e.g. SIC instead of AIC)?
- (b) (4 Points) Now, perform the ADF test on the time series for the *first difference* of the stock price of BMW AG. Specify the ADF regression and the hypotheses.
- (c) (4 Points) Test the stock price and its first difference for a unit root using the KPSS test. Do your results confirm those of parts a) and b)?

**21. Exercise** (18 Points) (*Stock Return Forecast*)

You are trying to forecast next year's stock returns. The information available to you comprises data on the companies' earnings and past stock returns. Let the data generating process for the stock returns  $y_t$  (measured at the end of period  $t$ ) be

$$y_t = \beta_1 + \beta_2 x_t + u_t, \quad \beta_2 \neq 0, u_t \sim i.i.d.(0, \sigma_u^2), \quad (7)$$

where  $x_t$  is the change in earnings from the end of period  $t - 1$  until the end of period  $t$ . The DGP for the change in earnings is given by a simple AR(1) process

$$x_t = \nu + \alpha x_{t-1} + v_t, \quad v_t \sim i.i.d.(0, \sigma_v^2), \quad (8)$$

and where  $v_t$  and  $u_s$  are independent for any  $s$  and  $t$ .

- (a) (2 Points) Compute the prediction  $E[y_{T+1}|x_T]$  in terms of equations (7) and (8) given that all parameters are known.

- (b) (1 Point) Compute the prediction error of  $E[y_{T+1}|x_T]$  assuming that all parameters are known.
- (c) (3 Points) Compute the mean squared error of prediction for the predictor  $E[y_{T+1}|x_T]$  given that all parameters are known.
- (d) (3 Points) Compute the prediction  $E[y_{T+1}|y_T]$  explicitly in terms of the parameters of equation (8) given that all parameters are known.
- (e) (3 Points) Compute the prediction error of  $E[y_{T+1}|y_T]$  assuming that all parameters are known. First, describe which additional assumptions/conditions were necessary to “remove” the conditional expectation  $E[u_T|y_T]$  which is left from the last exercise. For the computations, this term can be skipped.
- (f) (4 Points) Compute the mean squared error of prediction for the predictor  $E[y_{T+1}|y_T]$  given that all parameters are known. If you did your calculations correctly, the latter is larger than the mean squared error of prediction of  $E[y_{T+1}|x_T]$ . Can you explain this using your intuition?
- (g) (2 Points) Under which circumstances would you think that in practice (DGP unknown) the prediction based on a simple autoregressive model of order one exhibits a smaller MSE than the one including a regression as well?

**22. Exercise** (9 Points) (*ARIMA Representation*)

- (a) (3 Points) Use the representation  $y_t = \dots$  and lag polynomials to represent the following processes

$$\begin{aligned}
 y_t &= \mu_t + u_t, & u_t &\sim IID(0, \sigma_u^2), \sigma_u^2 > 0 \\
 \mu_t &= \mu_{t-1} + \beta_t + \epsilon_t & \epsilon_t &\sim IID(0, \sigma_\epsilon^2), \sigma_\epsilon^2 > 0 \\
 \beta_t &= g + \beta_{t-1} + \delta_t & \delta_t &\sim IID(0, \sigma_\delta^2), \sigma_\delta^2 > 0, u_t, \epsilon_t, \delta_t \text{ independent}, g \in \mathbb{R}
 \end{aligned}$$

- (b) (3 Points) Define the ARIMA representation of

$$z_t = b + c \cdot t + d \cdot t^2 + a_t, \quad a_t \sim IID(0, \sigma_a^2), b, c, d \in \mathbb{R}$$

- (c) (3 Points) Explain why  $z_t$  can be considered as a special case of  $y_t$ .

**23. Exercise** (45 Points) (*Lag Length Selection*)

This assignment reviews lag length selection by information criteria, unit root tests and forecasting. Only parts d), e), j) and h) require use of R!

- (a) (3 Points) The following R code can be used to simulate a single realization of a specific stochastic process:

```
set.seed(26062011)
filter( rnorm(T, 0, sqrt(5)) , c(0.6, 0.3) , "recursive")
```

State the data generating process. Be precise about error distribution and initial values.

- (b) (3 Points) This stochastic process is stable. Show this by calculating the roots of the lag polynomial.
- (c) (5 Points) This process also fulfills the stationarity conditions. Use this fact to calculate the expected value, the variance and the first order autocovariance.
- (d) (5 Points) Simulate 10000 replications of the process with  $T = 50$ . For every replication, estimate the lag length using the Akaike information criterion. How often does this selection lead to overparameterization? How often is the chosen model misspecified? (Hint: Lag length selection by AIC is readily available in `ar()`)
- (e) (2 Points) Now, repeat part d) for samples of size  $T = 200$ . Do your results change?
- (f) (4 Points) Testing for a unit root can be done using the Augmented Dickey-Fuller-Test or the KPSS-Test. Apply both tests to your first time series. What are the critical values that should be used here (level of significance: 10%, 5%, 1%)?
- (g) (2 Points) Should the test statistic be larger or smaller than the critical value to get a correct decision? Is this different for ADF and KPSS?
- (h) (5 Points) Now, perform the ADF and KPSS test for every realization simulated in part d). Use e.g. functions `ur.df()` and `ur.kpss()` in package `urca`. Consider three different scenarios for the ADF test:
- A lag length: 0
  - B correct lag length
  - C lag length: 10
- For the KPSS test use the implemented procedure for estimating the lag length. How often do we get a test statistic that leads to a correct decision?
- (i) (2 Points) When our decision is wrong, is this a type-I or type-II error?
- (j) (2 Points) Again, use the dataset generated in d) to perform an ADF test on all realizations. But now, use the correct lag length but allow for a deterministic trend under the alternative. Compare your results with those from part h).
- (k) (4 Points) Assume the DGP from part a) is known to us. Rewrite the AR process as an infinite Moving Average process. Calculate (at least) the first five MA parameters.
- (l) (3 Points) Let the last observed data for a given realization be
- |       |      |         |         |         |         |
|-------|------|---------|---------|---------|---------|
| $t$   | $T$  | $T - 1$ | $T - 2$ | $T - 3$ | $T - 4$ |
| $y_t$ | 10.3 | 6.7     | 5.8     | 7.4     | 8.3     |
- Combine this data and your information about the DGP to perform forecasts for  $T + 1$  up to  $T + 5$ .
- (m) (5 Points) For this forecast horizon, calculate the prediction interval for a confidence level of 90%.

**24. Exercise** (12 Points) (*ADF Simulation*)

Consider the influence of the size of  $\alpha_1$  on the Dickey-Fuller test decision for an AR(1)-process  $y_t = \alpha_1 y_{t-1} + u_t$ . Use the packages `fUnitRoots` and `akima`, a seed `set.seed(101)` and the function `pdf()` to get the corresponding  $p$ -values of a Dickey-Fuller distribution.

- (a) (5 Points) Simulate 10000 replications of a random walk

$$y_t = \alpha_1 y_{t-1} + u_t, \quad u_t \sim NID(0, 1), y_0 = 0, \alpha_1 = 1$$

with sample size 50. Use a boxplot for your  $p$ -values and explain your findings.

Hint:

```
boxplot(p_values, main="p-values for ADF test (DGP: random walk)",
ylab="p value", ylim=c(0,1))
abline(h=alpha, col="red")
legend("topleft", col="red", lwd=1, legend=paste0("sign. level ", alpha), bg="white")
```

- (b) (2 Points) Take a look at `p_values` by plotting the histogram. Are the  $p$ -values uniformly distributed? Explain your conclusion.
- (c) (5 Points) Simulate 10000 replications of a stationary AR(1)

$$y_t = \alpha_1 y_{t-1} + u_t, \quad u_t \sim NID(0, 1),$$

with values  $\alpha_1 = 0.09, 0.19, \dots, 0.99$  and sample size 50. Use again a boxplot for your  $p$ -values and explain your findings.

**25. Exercise** (18 Points) (*Exchange Rate Analysis*)

The file `EUR_USD_daily.txt` contains daily observations of the Euro - US Dollar exchange rate beginning in January 1, 1999.

Source: Deutsche Bundesbank: page of daily exchange rates.

- (a) (1 Point) Read the file into R and plot the series.
- (b) (3 Points) Compute daily log returns and their minimum, maximum, mean and variance. Also draw a time series plot and a histogram.
- (c) (2 Points) Check for unit roots in the series of log exchange rate (not log returns) using an ADF-test. Try both reasonable sets of hypotheses.
- (d) (2 Points) Check your results with the KPSS test.
- (e) (3 Points) Plot the standardized residuals of the selected autoregression in log exchange rate and comment on possible violations of the assumptions.
- (f) (5 Points) Check for remaining autocorrelation in the residuals by conducting a Portman-teau test and the Serial Correlation LM test. (Hint: Autocorrelations can be accessed via `acf()`. Look at the help page for details.)
- (g) (2 Points) Are the residuals normally distributed (Hint: Use e.g. function `jarque.bera.test()` from package `tseries`)?



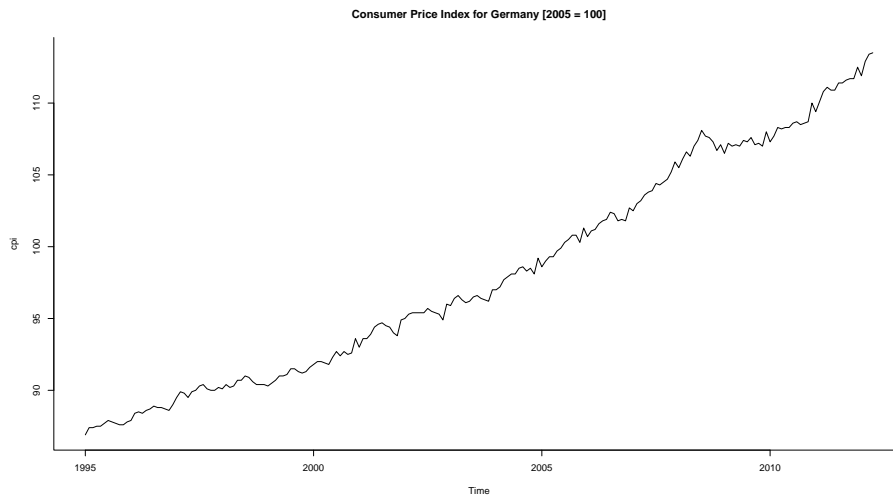
**26. Exercise (17 Points)** (*Zloty Exchange Rates*)

Imagine it is Friday, 08 June 2012, you are flying to Warsaw tomorrow to watch the kick-off of the UEFA EURO 2012. Such a trip should be planned carefully. One very important part in your preparation should be learning what your money is worth in Poland. You decide to have a closer look at the exchange rate Euro - Polish Złoty. Download a daily data series until 01 June 2012 from the web, e.g. the homepage of the ECB.

- (a) (1 Point) Plot the series.
- (b) (3 Points) Compute daily log returns and their minimum, maximum, mean and variance. Also draw a time series plot and a histogram.
- (c) (2 Points) Check for unit roots in the series of log exchange rate (not log returns) using an ADF-test. Try both reasonable sets of hypotheses.
- (d) (2 Points) Check your results with the KPSS test.
- (e) (3 Points) Plot the standardized residuals of an autoregression model for the log exchange rate and comment on possible violations of underlying assumptions.
- (f) (5 Points) Check for remaining autocorrelation in the residuals by conducting a Portman-teau test and the Serial Correlation LM test.
- (g) (1 Point) Are the residuals normally distributed?

## 27. Exercise (29 Points) (*Consumer Price Analysis*)

In this exercise we want to analyse the development of consumer prices in Germany. We use a data series with monthly data for the consumer price index, `cpi`. The time series plot is shown in the following figure:



- (a) The following output shows results for the ADF test on `cpi`, where lag length has been chosen by AIC and maximum lag length  $p_{max}$  was set to 12.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
    Min       1Q   Median       3Q      Max
-0.87526 -0.17517  0.00171  0.17748  0.78224

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.277417    1.764666   1.857  0.0649 .
z.lag.1        -0.038161    0.020938  -1.823  0.0700 .
tt              0.005377    0.002652   2.027  0.0441 *
z.diff.lag1    -0.165139    0.064332  -2.567  0.0111 *
z.diff.lag2     0.025816    0.065549   0.394  0.6942
z.diff.lag3    -0.016678    0.065795  -0.253  0.8002
z.diff.lag4    -0.040819    0.065835  -0.620  0.5360
z.diff.lag5    -0.036121    0.065771  -0.549  0.5836
z.diff.lag6    -0.055858    0.065736  -0.850  0.3966
z.diff.lag7     0.058860    0.065673   0.896  0.3713
z.diff.lag8     0.012590    0.065589   0.192  0.8480
z.diff.lag9    -0.033371    0.065302  -0.511  0.6100
z.diff.lag10  -0.023401    0.065177  -0.359  0.7200
```

```

z.diff.lag11 -0.048475  0.065216  -0.743  0.4583
z.diff.lag12  0.568803  0.064012   8.886 6.46e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Residual standard error: 0.2834 on 180 degrees of freedom

Multiple R-squared: 0.456, Adjusted R-squared: 0.4137

F-statistic: 10.78 on 14 and 180 DF, p-value: < 2.2e-16

Value of test-statistic is: -1.8226 3.8014 2.7716

Critical values for test statistics:

```

      1pct  5pct 10pct
tau3 -3.99 -3.43 -3.13
phi2  6.22  4.75  4.07
phi3  8.43  6.49  5.47

```

- i. (2 Points) Write down the regression equation used in the testing procedure.
  - ii. (3 Points) What are the null and alternative hypothesis and what is your test decision? Explain the test decisions depending on the assumptions of an underlying  $t$ -distribution (cf. regression output) and Dickey-Fuller distribution.
  - iii. (3 Points) Judge whether deterministic terms, lag selection criteria and maximum lag length are reasonably chosen.
- (b) (1 Point) The KPSS test is used to check the result from the ADF test. The output is:

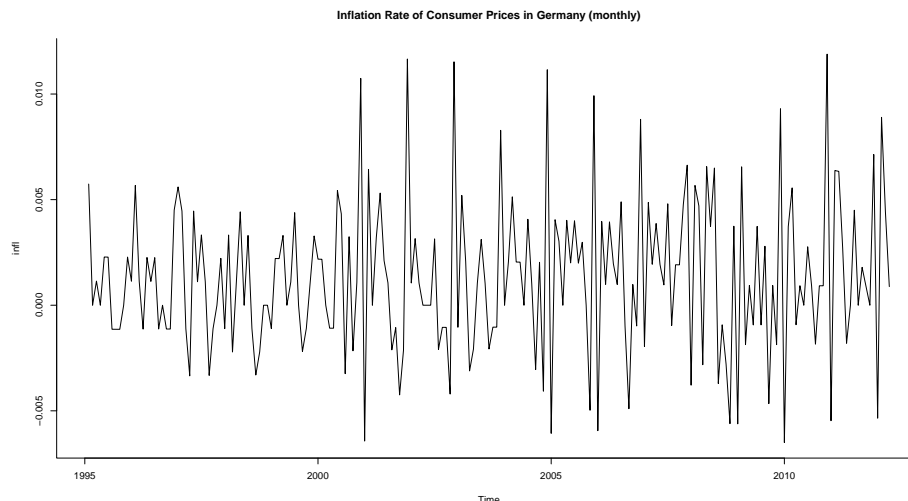
```

#####
# KPSS Unit Root Test #
#####
Test is of type: tau with 14 lags.
Value of test-statistic is: 0.3227
Critical value for a significance level of:
      10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216

```

Does this test confirm the result from the ADF test?

- (c) (2 Points) The monthly inflation rate, `infl`, is calculated as the first difference in `log cpi`. The time series plot is shown below:



What would be the reasonable pair of hypotheses to use in an ADF test on `infl`? How many lags should (at least) be included in the test regression?

- (d) (5 Points) Fitting an AR model for `infl` leads to following output:

Call:

```
ar.ols(x = infl, demean = FALSE, intercept = TRUE)
```

Coefficients:

```

      1      2      3      4      5      6      7      8      9
-0.1857  0.0021 -0.0406 -0.0708 -0.0355 -0.0538  0.0449  0.0001 -0.0432
     10     11     12
-0.0169 -0.0371  0.5633
```

Intercept: 0.001157 (0.0004321)

Order selected 12 sigma<sup>2</sup> estimated as 7.557e-06

The conventional standard error is roughly the same for all parameter estimators (except the one for the constant term). It is about 0.062. Use this to judge which lags have significant explaining power. What might the explanation for their significance?

- (e) (3 Points) Using a smaller model for `infl` yields the following output:

Time series regression with "ts" data:

Start = 1996(2), End = 2012(4)

Call:

```
dynlm(formula = infl ~ 1 + L(infl, c(1, 12)))
```

Coefficients:

```

      (Intercept)  L(infl, c(1, 12))1  L(infl, c(1, 12))12
      0.0008118      -0.1889144      0.5928478
```

Calculate the roots of the estimated lag polynomial and find out whether they lie inside or outside the unit circle (You will have to use a computer for this!).

- (f) (3 Points) Using the small model from the previous part, calculate the expected value of the (assuming) stationary time series and use this to calculate the *yearly* inflation rate.
- (g) (4 Points) Check the residuals from the smaller model for autocorrelation of order 1 to 12. Use the estimated serial correlation coefficients in the following table:

$k$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\hat{\rho}_k$	0.0083	0.0766	-0.0523	-0.0220	-0.0374	-0.0216	0.0706	0.0274	0.0452	-0.0163	0.0709	-0.1441	-0.0480	-0.1106	0.0006

- (h) (3 Points) First four moments of *standardized* residuals are:

order	moment
1	0.0000
2	0.9949
3	-0.1178
4	3.4228

Are residuals normally distributed?

## 4 Maximum Likelihood, ARCH and GARCH

### 28. Exercise (13 Points) (*Maximum Likelihood (Bernoulli)*)

After picking a bucket full of cherries, you realize that out of 500 cherries 375 are inhabited by worms.

- (4 Points) Calculate the probability  $\hat{p}$  for finding a worm in a cherry that maximizes the probability of your cherry-sample. Try both the regular likelihood version and the log-likelihood.
- (2 Points) What is the difference between the probability and the likelihood function?
- (2 Points) Calculate the Fisher information matrix.
- (5 Points) This part is about implementing the (numerical) ML estimation using a quasi-Newton algorithm of the problem above. The code

```
set.seed(4)                # starting point for reproduction
n      <- 500              # number of observations
prob   <- 0.75            # true parameter
data   <- rbinom(n, 1, prob)
```

generates a sample of size 500 with 375 worms inhabited cherries („ones“) and 125 others („zeros“). Define in R the binomial log likelihood `bin.ll` with two input variables the parameter  $p$  as `p` and the corresponding data set `data` and output value the log likelihood value. Then use `optim` to optimize the function numerically. Plot the function on  $(0,1)$  and check if your optimized value is really the maximum value. Use then the optional `hessian` parameter in the `optim` function to get the standard errors of your estimator.

### 29. Exercise (9 Points) (*Maximum Likelihood (Normal)*)

- (2 Points) Let  $Z$  be a normally distributed random variable with mean  $\mu_z$  and variance  $\sigma_z^2$ . Write down the probability density function of this variable.
- (2 Points) Let  $\mathbf{Z}$  be the set of  $T$  variables which are independent and have the same distribution as  $Z$ . Give the joint distribution of  $\mathbf{Z}$  (in terms of univariate marginal distributions **and** in the multivariate case).
- (2 Points) Let  $Y$  be a *conditionally* normally distributed random variable with  $E(Y|X = x, \alpha) = \alpha x$  and  $\text{Var}(Y|X = x) = \text{Var}(Y) = \sigma_y^2$ . Write down the (univariate, conditional) probability density function of this variable.
- (3 Points) Let  $\mathbf{Y}$  be the set of  $T$  variables which have the same marginal distribution as  $Y$ . A specific random variable in this set is called  $Y_t$  and depends on  $Y_{t-1}$  and not on  $Y_{t-2}$ . Give the joint distribution of  $\mathbf{Y}$ .

**30. Exercise** (8 Points) (*Maximum Likelihood (GARCH)*)

Write down the likelihood function of  $y_t$  conditional on  $y_{t-1}, y_{t-2}, \dots$  and  $\frac{u_t}{\sigma_t} = \xi_t \sim IN(0, 1)$  for a model with

- (a) (2 Points) AR(1) mean equation and ARCH(1) variance equation and
- (b) (3 Points) mean equation  $y_t = \alpha_0 + \alpha_1 y_{t-1} + \delta \sigma_{t-1} + u_t$  and GARCH(1,1) variance equation.
- (c) (3 Points) AR(2) mean equation and TGARCH(1,1) variance equation and

**31. Exercise** (15 Points) (*Log-Likelihood with Normal Distribution*)

Let  $Z$  be a random variable following the distribution  $N(\mu, \sigma^2)$ . A realization of this random variable is denoted by  $z$ .

- (a) (2 Points) State the density  $f(z|\theta_0)$ . What are the elements in  $\theta_0$ ?
- (b) (1 Point) Consider the case, where  $\mu = 2$  and  $\sigma^2 = 3$ . Sketch the graph of the density function.
- (c) (2 Points) Consider three different realizations of  $Z$ , namely  $z_1 = 1, z_2 = 2, z_3 = 3$ . For each  $z_j$  calculate the density  $f(z_j|\theta_0)$
- (d) (2 Points) For which  $z$  is the density maximized?
- (e) (3 Points) Consider a random sample of  $T$  realizations of  $Z$ . State the joint density, the likelihood and log-likelihood for the sample.
- (f) (5 Points) Maximize the log-likelihood with respect to the parameters in  $\theta$ .

**32. Exercise** (11 Points) (*Log-Likelihood with DGP*)

Let the DGP be given by

$$p_t = \alpha_0 + \alpha_1 p_{t-1} + u_t, \quad (12)$$

where  $u_t$  is Gaussian White Noise with variance  $\sigma^2$ . Assume that  $|\alpha_1| < 1$ . The process started in the infinite past with  $p_{-\infty} = 0$ .

- (a) (2 Points)  $p_t$  denotes a random variable. What is its unconditional distribution? What is its distribution conditional on past values  $p_{t-1}, p_{t-2}, \dots$ ?
- (b) (2 Points) State the conditional density/likelihood function  $f(p_t|p_{t-1}, p_{t-2}, \dots; \alpha_0, \alpha_1, \sigma^2)$  and the log-likelihood.
- (c) (2 Points) For a sample  $(p_1, p_2, \dots, p_T)$  show that the joint unconditional density

$$f(p_1, p_2, \dots, p_T | \alpha_0, \alpha_1, \sigma^2)$$

can be decomposed into the product of conditional density functions. Give the conditional log-likelihood function of this sample.

- (d) (5 Points) Find the Maximum Likelihood estimators for parameters  $\alpha_0, \alpha_1$  and  $\sigma^2$ . For this purpose, calculate the first derivatives of the conditional log-likelihood with respect to  $\alpha_0, \alpha_1$  and  $\sigma^2$  and find their roots (i.e. impose first order condition that first derivatives need to be zero). Then solve for the parameters such that they only depend on observable values.

**33. Exercise** (8 Points) (*Joint Log-Likelihood*)

Let  $(y_1, y_2, \dots, y_T)$  be a random sample from distribution  $N(\mu, 1)$ . The log-likelihood for random variable  $y_t$  is then

$$l_t = -\frac{1}{2} \log 2\pi - \frac{1}{2}(y_t - \mu)^2$$

- (a) (3 Points) Find the joint log-likelihood function and its first and second derivative with respect to  $\mu$ . How does the log-likelihood function look like?
- (b) (5 Points) Instead of directly estimating  $\mu$  using the first order conditions, we use the BHHH algorithm. For a sample  $\mathbf{y} = (6, 5, 4, 5.2, 4.8)$  start with  $\mu_0 = 1$  and iterate on the update function until your estimate converges (use e.g. a step length of 0.5). How many iteration steps do you need until convergence?

**34. Exercise** (4 Points) (*Conditional Heteroskedasticity*)

- (a) (4 Points) Give an overview of the different conditional heteroskedasticity models and suggest which one to choose in certain situations.

**35. Exercise** (13 Points) (*ARCH(1)*)

Consider the following DGP:

$$\begin{aligned} r_t &= \mu + u_t \\ u_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \gamma_0 + \gamma_1 u_{t-1}^2 \end{aligned}$$

where  $\epsilon_t \sim N(0, 1) \forall t$  and independent of  $\sigma_s \forall s$ ,  $|\gamma_1| < 1$  and  $E[\sigma_t] < \infty$ .

- (a) (3 Points) Calculate the unconditional moments  $\mu = E(r_t)$  and  $\text{Var}(r_t)$ , assuming that the processes have run since ever.
- (b) (1 Point) Is  $u_t$  normally distributed? Is  $\sigma_t^2$  normally distributed?
- (c) (3 Points) Let  $\mu = 0.05$ ,  $\gamma_0 = 0.00125$  and  $\gamma_1 = 0.5$ . You observe  $r_T = 0.06$ . Calculate the conditional mean and variance of  $r_{T+1}$ .
- (d) (3 Points) Simulate a realization of this process with at least 100 observations. Have a look at the autocorrelation function for  $r_t$  and  $r_t^2$ .
- (e) (3 Points) Use your simulated series as a sample and estimate the parameters of the model.

**36. Exercise** (10 Points) (*Representations of GARCH(1,1)*)

Consider a GARCH(1,1) process where  $\xi_t \sim N(0, 1) \forall t$  and  $\sigma_t$  are independent of each other.

- (a) (5 Points) Derive the mean, (unconditional) variance, and the conditional fourth moment of the process, assuming it is stationary. Write down all major steps in your calculations.  
**Hint:** Assume that  $E(\sigma_t)$  is finite.
- (b) (2 Points) Rewrite the GARCH(1,1) into an ARCH( $\infty$ ) process.
- (c) (3 Points) Derive the ARMA representation of  $u_t^2$ . Therefore introduce  $e_t = u_t^2 - \sigma_t^2$ . What is the role of the GARCH and ARCH parameters, respectively?

**37. Exercise (36 Points) (AR and GARCH)**

Download daily stock prices of ThyssenKrupp AG, e.g. using

```
library(tseries)
dataset <- get.hist.quote(instrument = "TKA.DE", start = "2003-01-01", end = "2012-06-29")
```

- (a) (2 Points) Create a log return series. Specify a suitable  $AR(p)$ -Model for the log return series and estimate this model by OLS.
- (b) (3 Points) Check the correlogram of the residuals and squared residuals. What does this graphic tell you in general and what are your conclusions in the present case?
- (c) (3 Points) Regardless of your conclusions, perform an ARCH-LM test. How do you choose  $q$  and why? Describe the test in general and explain the results you get here.
- (d) (5 Points) Estimate a GARCH(1,1)-model using function `ugarchfit()` in package `rugarch`. Use the output to check whether the model for mean and variance are adequate.
- (e) (2 Points) Forecast values for the return series and its conditional variance for the next ten days. (Hint: Use e.g. function `ugarchforecast()` with your fitted model from part (d).)
- (f) Explain the output of the `ugarchfit()` function for

```
specAlt <- ugarchspec(variance.model = list(garchOrder = c(2,2)),
                      mean.model = list(armaOrder = c(6,0), include.mean = TRUE))
(fitAlt <- ugarchfit( specAlt, r))
```

in detail, i. e. give explanations for the following tests and values.

Hint: Read the [Introduction to the rugarch Package](#).

- i. (1 Point) What models are estimated for the mean and variance equation?
- ii. (2 Points) Explain the parameters. How are they obtained and what changes are in the `robust` version?
- iii. (2 Points) Define the information criteria in terms of the log-likelihood function.
- iv. (3 Points) Define the hypothesis' and the test statistic of the Weighted Ljung-Box test on standardized (squared) residuals.
- v. (3 Points) Explain the results concerning the Weighted Ljung-Box test on standardized (squared) residuals.
- vi. (2 Points) Explain the Weighted ARCH-LM test and its result.
- vii. (2 Points) Explain Nyblom's Parameter Stability Test.
- viii. (2 Points) Explain briefly in your own words the `Sign Bias Test`.
- ix. (2 Points) Explain briefly in your own words the `Adjusted Pearson Goodness-of-Fit Test`.
- x. (2 Points) Use the `plot` function onto your fit object and explain the results.



**38. Exercise (20 Points)** (*Siemens*)

This exercise will be a recap of many issues we dealt with this semester. We will analyse the time series of the stock price of Siemens AG shares. You can find the data [here](#).

- (a) (1 Point) From financial theory we might expect that the price follows a random walk. Define a random walk.
- (b) (4 Points) If the series follows a random walk, what does this imply for forecasting the price of the stock and price-change on the stock? What is different, when the time series follows a stationary AR(1) process or a unit root AR(2) process?
- (c) (4 Points) Next you see two outputs. Describe what hypotheses have been tested and interpret the test results.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression drift
Call:
lm(formula = z.diff ~ z.lag.1 + 1)
Residuals:
    Min       1Q   Median       3Q      Max
-30.1671 -0.7242 -0.0106  0.7336 12.7046
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.194281   0.098585   1.971  0.0488 *
z.lag.1      -0.002918   0.001370  -2.130  0.0332 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.911 on 3002 degrees of freedom
Multiple R-squared:  0.001509,    Adjusted R-squared:  0.001177
F-statistic: 4.538 on 1 and 3002 DF,  p-value: 0.03324

Value of test-statistic is: -2.1302 2.2708
Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.43 -2.86 -2.57
phi1  6.43  4.59  3.78
and
#####
# KPSS Unit Root Test #
#####
Test is of type: mu with 9 lags.
Value of test-statistic is: 3.2849
Critical value for a significance level of:
      10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

- (d) (2 Points) Interpret the following output:

```
Call:
ar.ols(x = price, aic = FALSE, order.max = 1)
Coefficients:
      1
0.9971
Intercept: -0.002179 (0.03485)
Order selected 1 sigma^2 estimated as 3.649
```

- (e) (5 Points) Let `resid` denote the series of residuals from the estimation performed in part d). Describe and interpret the following outputs:

```
Time series regression with "ts" data:
Start = 10, End = 3005
Call:
dynlm(formula = resid ~ 1 + L(price, 1) + L(resid, 1:8), data = data)

Residuals:
      Min       1Q   Median       3Q      Max
-29.8520  -0.7396  -0.0037   0.7460  12.7395

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.0240465  0.1001637  -0.240 0.810291
L(price, 1)    0.0003299  0.0013975   0.236 0.813428
L(resid, 1:8)zoo.coredata.x...tt..1 -0.0182665  0.0183405  -0.996 0.319350
L(resid, 1:8)zoo.coredata.x...tt..2 -0.0059445  0.0183390  -0.324 0.745850
L(resid, 1:8)zoo.coredata.x...tt..3 -0.0675748  0.0183365  -3.685 0.000233 ***
L(resid, 1:8)zoo.coredata.x...tt..4  0.0058990  0.0183495   0.321 0.747868
L(resid, 1:8)zoo.coredata.x...tt..5  0.0231830  0.0183423   1.264 0.206361
L(resid, 1:8)zoo.coredata.x...tt..6  0.0234171  0.0182961   1.280 0.200682
L(resid, 1:8)zoo.coredata.x...tt..7 -0.0155948  0.0182922  -0.853 0.393985
L(resid, 1:8)zoo.coredata.x...tt..8 -0.0165020  0.0182788  -0.903 0.366708
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.901 on 2986 degrees of freedom
Multiple R-squared: 0.006927, Adjusted R-squared: 0.003933
F-statistic: 2.314 on 9 and 2986 DF, p-value: 0.01367
```

and

```
Time series regression with "ts" data:
Start = 10, End = 3005
Call:
dynlm(formula = resid^2 ~ 1 + L(resid^2, 1:8), data = data)

Residuals:
      Min       1Q   Median       3Q      Max
-79.19  -2.56  -2.19  -1.05  900.59

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.22814   0.39184   5.686 1.42e-08 ***
L(resid^2, 1:8)zoo.coredata.x...tt..1  0.05545   0.01823   3.041 0.002375 **
L(resid^2, 1:8)zoo.coredata.x...tt..2  0.05108   0.01826   2.798 0.005174 **
L(resid^2, 1:8)zoo.coredata.x...tt..3  0.06349   0.01826   3.477 0.000515 ***
L(resid^2, 1:8)zoo.coredata.x...tt..4  0.04376   0.01829   2.393 0.016784 *
L(resid^2, 1:8)zoo.coredata.x...tt..5  0.02135   0.01829   1.167 0.243105
L(resid^2, 1:8)zoo.coredata.x...tt..6  0.04350   0.01825   2.383 0.017224 *
L(resid^2, 1:8)zoo.coredata.x...tt..7  0.02152   0.01825   1.179 0.238443
L(resid^2, 1:8)zoo.coredata.x...tt..8  0.08417   0.01822   4.619 4.02e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 19.79 on 2987 degrees of freedom
Multiple R-squared: 0.03041, Adjusted R-squared: 0.02782
F-statistic: 11.71 on 8 and 2987 DF, p-value: < 2.2e-16
```

- (f) (4 Points) Also interpret the following output:

```

Call:
garch(x = resid, order = c(0, 3))
Model:
GARCH(0,3)
Residuals:
      Min       1Q   Median       3Q      Max
-12.497243  -0.468437  -0.006638   0.491119   8.789050

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
a0    1.07730    0.01203   89.558 <2e-16 ***
a1    0.15490    0.01570    9.865 <2e-16 ***
a2    0.35581    0.01030   34.558 <2e-16 ***
a3    0.30092    0.01967   15.299 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:
      Jarque Bera Test

data: Residuals
X-squared = 33464.9, df = 2, p-value < 2.2e-16
      Box-Ljung test
data: Squared.Residuals
X-squared = 0.0187, df = 1, p-value = 0.8912

```

### 39. Exercise (10 Points) (*Stock Price Choice to model*)

- (a) (5 Points) For a stock price series of your choice find an adequate model for the mean and variance equation. Justify your choices by presenting test outcomes etc. (Use a dataset which contains daily data for at least two years.)
- (b) (5 Points) Compare it with your neighbour. Exchange applied descriptives (i. e. statistics, plots) and statistical tests. Discuss test assumptions and power. Forecast your stock prices with different measures (losses and horizons) and compare your models depending on information criteria and your forecast measures.

## 5 Long Run

### 40. Exercise (7 Points) (*Unconditional Expected Value of a stationary Time Series*)

Let the stationary series  $y_t$  be generated by

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t,$$

where  $u_t \sim WN(0, \sigma_u^2)$  and  $|\alpha_1| < 1$ .

- (a) (2 Points) Calculate expected value and variance of  $y_t$ .
- (b) (5 Points) Calculate the variance of the mean estimator  $\hat{\mu} = \frac{1}{T} \sum_{t=1}^T y_t$  and show that it converges for  $T \rightarrow \infty$ .

## 6 Finance

### 41. Exercise (7 Points) (*Asset pricing formula and Sharpe ratio*)

Below you find information on three assets  $A$ ,  $B$ ,  $C$  with payoffs  $x^A$ ,  $x^B$ , and  $x^C$ , respectively:

$$R^f = 1.05$$

$$E[x^A] = 3.5, \quad Cov(m, x^A) = 2, \quad Var(x^A) = 5$$

$$E[x^B] = 5.2, \quad Cov(m, x^B) = 2, \quad Var(x^B) = 5$$

$$E[x^C] = 4.4, \quad Cov(m, x^C) = -2, \quad Var(x^C) = 5$$

- (3 Points) Determine the prices for each of the assets using a well-known pricing formula and explain your results.
- (2 Points) Calculate the Sharpe ratio for each asset.
- (2 Points) Which properties share all assets on the mean-variance frontier?

### 42. Exercise (3 Points) (*Expected Wealth*)

- (3 Points) You estimated the mean yearly return of an asset and obtained 10% with standard deviation 8%. You checked the residuals and found no correlation. Given an initial wealth of 100 units, compute the wealth that you expect in twenty years from now. Comment on possible approximations that you have to do.

### 43. Exercise (6 Points) (*Stochastic discount factor*)

- (3 Points) Derive the stochastic discount factor for the following utility functions:

- $u(c) = \ln(c)$
- $u(c) = kc^\alpha l^{1-\alpha}$ , with  $k$  a constant,  $c$  the level of consumption and  $l$  all other factors, such as leisure, that might enter a person's utility function. For simplicity  $l$  is kept fixed.
- $u(c) = -\frac{1}{2}(c - c^*)^2$

Also check the behavior of  $u'(c)$  for increasing  $c$ .

- (3 Points) Explain the effect of an anticipated change in consumption (increase  $c_{t+1}$  while  $c_t$  and  $\beta$  remain unchanged) on the stochastic discount factor and on the price of an asset. Give the economic intuition as well as the mathematical line of reasoning. Are the effects the same for all three utility functions?

**44. Exercise** (5 Points) (*Stochastic discount factor and risk-neutral probabilities*)

Consider the following simple artificial setup for a single investor with the following values

$$\beta = 0.6$$

$$u(c) = \frac{1}{2}c^{1/2}$$

$$c_t = e_t = e_{t+1} = 4$$

and the following state conditions:

state	probability	payoff
$s$	$\pi(s)$	$x_{t+1}(s)$
1	$\frac{1}{2}$	12
2	$\frac{1}{2}$	0

- (3 Points) Compute for each state the stochastic discount factors and the prices of each contingent claim.
- (2 Points) Compute the risk-neutral probabilities and the price of the asset and comment on all results.

**45. Exercise** (6 Points) (*CAPM*)

Use the dataset in [capm.RData](#) for this exercise. It contains the (dividend and stock-split adjusted) stock prices of seven DAX companies, the DAX series itself and a three-month money market rate (EURIBOR). Data are for Jan. 2003 until Jun. 2012.

- (1 Point) Create a series of log returns for every stock price series.
- (2 Points) Create excess return series for all stock price series and the DAX using the interest rate series as approximation for the riskless return. (Hint: Interest rates are *par annum*. Divide by 12 to get monthly interest rates.)
- (3 Points) For every excess return series estimate the CAPM model by regressing the excess return of the stock on a constant and the excess return of the DAX. Test whether the constant/pricing error is significant.

## 7 Summary

### 46. Exercise (27 Points) (*Dynamic regression model*)

Consider a dynamic regression model

$$y_t = X_t' \boldsymbol{\beta} + u_t, \quad u_t \sim IID(0, \sigma_u^2), \quad t = 1, \dots, T,$$

with  $X_t' = (x_{t1} \cdots x_{tk})$  and  $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$ .

(a) Now you observe residual autocorrelation (not necessarily DGP error autocorrelation) in a dynamic regression model. This is usually taken as a signal of misspecification for which these possibilities can be the reasons:

- Autoregressive errors in the DGP.
- Dynamic misspecification.
- Omitted variables and non-modelled structural shifts.
- Misspecified functional form.

i. (3 Points) What are the consequences of autocorrelation in the DGP error?

ii. (3 Points) What are the consequences of dynamic misspecification?

iii. (3 Points) What are the consequences of omitted variables and non-modelled structural shifts?

iv. (2 Points) What are the consequences of a misspecified functional form?

(b) How can we test the model formulation or misspecification for the case of ...

i. (2 Points) ... no autocorrelation?

ii. (2 Points) ... no heteroskedasticity?

iii. (2 Points) ... for correct functional form?

iv. (2 Points) ... for normality of the error term?

(c) (8 Points) Empirical application:

Download US data for aggregate consumption  $c_t$  in the private sector, disposable income  $y_t$  and wealth including owner occupied housing  $w_t$ . Do a small empirical analysis of the first differences  $\Delta$  of these variables by specifying a model for

$$E[\Delta c_t \mid \Delta y_t, \Delta w_t]$$

using the tests we just discussed.

## 8 Bayesian Estimation Methods

### 47. Exercise (14 Points) (*Exponential Family*)

One of the most important classes of probability density functions (p.d.f.'s) is the exponential family with important members like the normal, exponential, log-normal,  $\gamma$ ,  $\chi^2$ ,  $\beta$ , Dirichlet, Bernoulli, categorical, Poisson, geometric, inverse Gaussian and many more. Since the p.d.f.s of the  $k$ -parametric members of the exponential family are of the form

$$f(x; \boldsymbol{\theta}) = a(\boldsymbol{\theta}) \cdot b(x) \cdot \exp\left(\sum_{i=1}^k c_i(\boldsymbol{\theta})d_i(x)\right), \quad \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^k, x \in \mathbb{R}$$

for suitable functions  $a, c_i : \mathbb{R}^k \rightarrow \mathbb{R}, b, d : \mathbb{R} \rightarrow \mathbb{R}$  and where  $\text{supp}(x)$  does not depend on  $\boldsymbol{\theta}$  (=regular), one can deduce many rules and symmetries. One of the most important theorems from the Bayesian point of view is the following (for  $k = 1$ ):

"The prior distribution  $p(\theta) \propto C(\theta)^a \exp(\phi(\theta)b)$  is conjugate to the exponential family distribution likelihood, i. e. of the the exponential family."

Consider now the gamma p.d.f., i. e.  $X \sim \gamma(\lambda, r)$ , which is given by

$$f_X(x; \lambda, r) = \frac{\lambda^r}{\Gamma(r)} \cdot x^{r-1} \exp(-\lambda x), \quad x \in \mathbb{R}^+, r \in \mathbb{R}_0^+, \lambda \in \mathbb{R}^+$$

with shape  $r$ , scale  $r$  and  $\Gamma(x) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  the gamma function.

- (3 Points) State the gamma function  $\Gamma(x)$ , its basic properties and the idea behind. Plot the gamma function and the gamma p.d.f. State the first moments of the gamma p.d.f.
- (3 Points) Show for known parameter  $r$  that the gamma p.d.f. constitutes an exponential family member ( $k = 1$ ) and then show for known parameter  $\lambda$  the same.
- (2 Points) Show that the gamma p.d.f. constitutes an 2-exponential family member.
- (2 Points) Show that the gamma p.d.f. is a conjugate prior for the Poisson p.d.f.
- (2 Points) Given data  $X_i \sim NID(\mu, \sigma^2), i = 1, \dots, n$ , show that a  $N(m, s^2)$  prior is conjugate for this likelihood (mean unknown, variance known) to another normal. What are its first two moments?
- (2 Points) Show that the  $\beta(\alpha_1, \alpha_2)$  p.d.f. is a conjugate prior for a binom( $1, p$ ) likelihood. State the posterior distribution.



**48. Exercise** (5 Points) (*Inverse Gamma*)

Consider the inverse gamma p.d.f., i. e.  $Y \sim \gamma^{-1}(\lambda, r)$ , which is given by

$$f_Y(y; \lambda, r) = \frac{\lambda^r}{\Gamma(r)} \cdot y^{-r-1} \exp(-\lambda/y), \quad x \in \mathbb{R}^+, r, \lambda \in \mathbb{R}^+.$$

- (a) (2 Points) Show that  $Y \sim \gamma^{-1}(\lambda, r)$  where  $Y := \frac{1}{X}$  with  $X \sim \gamma(a, b)$ .

**Hint:** Use a probability integral transformation.

- (b) (3 Points) Derive the first and second moments of the gamma distribution using the moment generating function (m.g.f.)  $M_Y(t) := E[\exp(tY)] = \int_{-\infty}^{\infty} \exp(ty) f_Y(y) dy$ .

Alternatively, you can calculate  $E[Y^n]$  directly.

**Hint:** If  $Y$  has a m.g.f.  $M_Y(t)$ , then  $E[Y^n] = M_Y^{(n)}(0)$  where  $M_Y^{(n)}(0) := \frac{d}{dt} M_Y(t) |_{t=0}$ , that is, the  $n$ -th (non-centered) moment is equal to the  $n$ -th derivative of  $M_Y(t)$  evaluated at  $t = 0$ .

**49. Exercise** (42 Points) (*Bayesian AR(p)*)

Consider the following Gaussian AR( $p$ )-process

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \epsilon_t, \quad \epsilon_t | \mathbf{X}_t, \sigma^2, \boldsymbol{\alpha} \sim NID(0, \sigma^2), \quad t = 1, \dots, T$$

and for given hyperparameters  $\mathbf{m}, \mathbf{M}, a$  and  $b$  the priors  $\mathbb{R}^{p+1} \ni \boldsymbol{\alpha} | \sigma^2 \sim N(\mathbf{m}, \sigma^2 \cdot \mathbf{M}^{-1})$  and  $\sigma^2 \sim \gamma^{-1}(a, b)$ .

- (a) (3 Points) Write the model as a multiple linear regression model  $\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$  given initial values for  $y_i, i = -p + 1, \dots, 0$ .
- (b) (2 Points) What is the conditional distribution of  $\mathbf{y} | \mathbf{X}, \boldsymbol{\alpha}, \sigma^2$ ? What is the likelihood of our data  $(\mathbf{y}, \mathbf{X})$ ?
- (c) (7 Points) Derive the full posterior of  $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \sigma^2)$ , i. e.  $p(\boldsymbol{\theta} | (\mathbf{y}, \mathbf{X}))$  and the full conditional posteriors  $p(\boldsymbol{\alpha} | (\mathbf{y}, \mathbf{X}); \sigma^2)$  and  $p(\sigma^2 | (\mathbf{y}, \mathbf{X}); \boldsymbol{\alpha})$ .

**Hint:** Remember the normal inverse gamma distribution and do not consider initial value distributions.

- (d) The following parts are about the implementation of a Gibbs sampler to sample from the joint posterior  $(\boldsymbol{\alpha}, \sigma^2)$  for an Gaussian AR(1).

- i. (2 Points) Sketch the ideas of Gibbs Sampling / MCMC / Metropolis Hastings.

**Hint:** There is a nice article by [Casella and George \(1992\)](#).

- ii. (1 Point) Apply the formula from above onto  $p = 1$  and state the Gibbs sampling procedure theoretically.
- iii. (3 Points) Implement a Gibbs sampler for a normal AR(1) without drift/intercept and trend in R and apply it onto  $\mathbf{y}$  generated via:

```
n          <- 100
y0.true   <- 0.0
rho.true  <- 0.95 # close to non-stationarity
```

```
sig2.true <- 1.0
set.seed(1255)
y <- arima.sim(n = n, list(ar = c(rho.true), sd = sqrt(sig2.true)))
```

- iv. (2 Points) Estimate the unknown parameters of an AR(1) model using the posterior means as Bayesian point estimates and compare them with the true population values.
  - v. (2 Points) Use the corresponding function in the package `MCMCpack` and replicate your results.
  - vi. (2 Points) Estimate the process above with OLS and Maximum Likelihood, compare your estimates again.
  - vii. (2 Points) Analyze the MCMC output using R-tools from the package `coda`.
  - viii. (3 Points) Forecast  $k$ -steps using the Bayesian forecasting mechanism. Apply a HPD-Intervall (Highest Posterior Density Interval) onto your forecasts.
  - ix. (3 Points) Derive the marginal posteriors of  $\boldsymbol{\alpha}$  and  $\sigma^2$ .
  - x. (2 Points) Compare visually **and** quantitatively the MCMC-approximated marginal posteriors with the exact marginals from (ix)
  - xi. (2 Points) Sketch the idea how to implement a truncated Normal prior to restrict the AR(1)-coefficient to stationarity.
  - xii. (2 Points) Compare your Bayesian forecast and HPD-Intervall with the frequentist forecast and prediction intervall (OLS or MLE).
- (e) (4 Points) Summarize the differences and (dis-)advantages of the frequentist and the Bayesian approach.

**50. Exercise** (10 Points) (*Regression with autocorrelated errors*)

Consider the simple regression model

$$y_t = \beta_1 + \beta_2 x_t + \epsilon_t \text{ with } \epsilon_t = \phi \epsilon_{t-1} + \nu_t, \quad \nu_t \sim NID(0, \sigma^2), \quad t = 1, \dots, T.$$

- (a) (1 Point) Write the model in matrix form  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  and  $\boldsymbol{\varepsilon} = \mathbf{E}\boldsymbol{\phi} + \boldsymbol{\nu}$ .
- (b) (1 Point) Use a conjugate prior distribution  $\boldsymbol{\beta} \mid \phi, \sigma^2 \sim N(\mathbf{b}_0, \mathbf{A}_0)$  with known hyperparameters to derive the full conditional posterior of  $\boldsymbol{\beta}$ .
- (c) (2 Points) Use a conjugate prior  $\phi \mid \boldsymbol{\beta}, \sigma^2 \sim N(\mathbf{c}_0, \mathbf{B}_0) \cdot I_{\text{supp}(\phi)}$  with indicator function  $I_{\text{supp}(\phi)}$  used to constrain the roots of  $\Phi(L) = 0$  to be outside the unit circle. State the full conditional posterior of  $\phi$ .
- (d) (2 Points) Use a conjugate prior  $\sigma^2 \mid \boldsymbol{\beta}, \phi \sim \gamma^{-1}(\nu_0/2, \delta_0/2)$  for  $\sigma^2$ . Derive the full conditional posterior of  $\sigma^2$ .
- (e) (2 Points) State a Gibbs sampling approach to sample from  $(\boldsymbol{\beta}, \phi, \sigma^2)$ .
- (f) (2 Points) Write an R-program to implement the algorithm from before.

**51. Exercise** (12 Points) (*The Univariate Normal Gibbs Sampler*)

Consider  $X_1, \dots, X_n \sim NID(\mu, \sigma^2) = NID(\boldsymbol{\theta})$  from a Bayesian point of view:

$$\underbrace{\pi(\boldsymbol{\theta} | \mathbf{x})}_{\text{posterior}} \propto \underbrace{f(\mathbf{x} | \boldsymbol{\theta})}_{\text{likelihood}} \cdot \underbrace{\pi(\boldsymbol{\theta})}_{\text{prior}}$$

- (a) (1 Point) State the likelihood function  $f(\mathbf{x} | \boldsymbol{\theta})$ .
- (b) (1 Point) State a noninformative prior.
- (c) (2 Points) Calculate the joint posterior distribution  $\pi(\boldsymbol{\theta} | \mathbf{x})$ .
- (d) (3 Points) Calculate both conditional posterior distributions  $\pi(\mu | \sigma^2, \mathbf{x})$  and  $\pi(\sigma^2 | \mu, \mathbf{x})$ .

**Hint:** Define  $\tau = 1/\sigma^2$ .

- (e) (2 Points) Write down the pseudo code of a Gibbs sampler.
- (f) (3 Points) Implement your Gibbs Sampler in **R**. Plot the typical Bayes-Gibbs inference plots (acf and convergence). Additionally, plot the marginal and joint posteriors.

**Hint:** For diagnostics, you can use the R-package `coda`.