Main topics of the lab: Exponential smoothing, Holt's and Holt-Winter's methods for time series forecasting. Main measures of accuracy of predictions. Autocorrelations, goodness of fit of a time series model.

If a series does not contain definite trends, one reasonable way of forecasting future values is to use some kind of average of the past values. For series with trend and/or seasonality, we can use some kind of averaging to estimate the slope of a linear trend and also seasonal factors. This leads to a family of exponential smoothing methods.
Let us denote with $\hat{z}_{t+p|t}$ the predicted value of a series for time moment $t+p$, which is computed by using only data at or before the time momet t. When predicting only the next value (the case $p=1$), then we are going to use the simplified notation $\hat{z}_{t+1}$.
The best known time series forecasting methods which are based on exponential smoothin, are the following ones.

- **Exponential smoothing method**, which can be used if there is no trend. Usually the predictions are computed step-by-step, setting $\hat{z}_1 = z_1$ and then computing one step ahead predictions for all other time moments by
$$\hat{z}_{t+1} = \alpha z_t + (1-\alpha)\hat{z}_t,\ t = 1, \ldots, n.$$
The parameter α may be be determined so that the one step ahead prediction errors are (in some sense) as small as possible. A usual measure of smallness is the mean squared prediction error.
- **Holt's method.** This method allows a trend curve which shape changes in time. Method can be used for non-seasonal time series. The prediction formulas are as follows:
$$\hat{z}_{t+p|t} = a_t + b_t\, p,$$
where a_t is the current value of the trend curve computed by
$$a_t = \alpha\, z_t + (1-\alpha)\hat{z}_t = \alpha\, z_t + (1-\alpha)(a_{t-1} + b_{t-1})$$
and the current slope b_t of the trend curve is computed by
$$b_t = \beta\,(a_t - a_{t-1}) + (1-\beta)b_{t-1}.$$
In order to use the method, one should specify a_1, b_1, α, β. Often a_1 and b_1 are determined by fitting a linear function for a relatively small number of first observations (or are specified as $a_1 = z_1,\ b_1 = 0$), the smoothing parameters α and β are usually determined by minimizing a certain measure of prediction accuracy.
- **Holt-Winter's methods** are logical extensions of Holt's method to the case of seasonal data. There are two versions, how trend and seasonality can be combined - they can be added (additive method) or multiplid (multiplicative method). Additive method is suitable, if the magnitude of seasonal fluctuations do not change when the mean of the series changes. For using the multiplicative method, the series should not change the sign and if there are clear increases of the average values of the series, the magnitude of seasonal fluctuations should also increase proportionally. Intuitively, the parameter α describes the smoothness of the trend curve (how quickly the shape of the curve can change), β describes the speed of change in the slope of prediction line and γ describes how quickly the seasonality is allowed to change.

We are going to use the command HoltWinters for applying exponential smoothing methods. By default, this command fits an additive method and determines the values of the parameters by minimizing the mean squared prediction errors. Examples of using the command:
HoltWinters(x)- fits an additive Holt-Winter's method to the series x,
m1=HoltWinters(x,seasonal="multiplicative")-fits a multiplicative HW method to x,
m1=HoltWinters(x,**gamma**=FALSE)-fits Holt's method (no seasonal part) to x,
m1=HoltWinters(x,**beta**=FALSE,**gamma**=FALSE)- fits the exponential smoothing method.
If the parameters α, β and γ found by the command do not seem reasonable (or an error message about convergence problems is issued), it is possible to specify a set of starting parameters for the optimisation

method by the option

optim.start=**c**(alpha=a0,**beta**=b0,**gamma**=g0),

where $a0, b0$ and $g0$ are some concrete parameters between 0 and 1. More information about the command can be found by using the command **help**(HoltWinters). As you can expect, it is possible to produce graphs describing the behaviour of the fitted method by **plot** command, future values can be predicted by the **predict** command and so on.

Exercises:

1. Consider the series of accommodated foreign visitors from Lab 1. Produce graphs with predictions produced by the exponential smoothing, Holt's, and Holt-Winter's methods (both additive and multiplicative), a different graph for each method. Also produce graphs showing the behavior of one step prediction errors (which can be obtained by **residuals**(**model**)).

2. A simple way to compare different methods is to compute a goodness-of-fit measure. Depending on what is the aim of modelling, different measures are appropriate. Some commonly used measures are MAD (Mean Absolute Deviations, also MAE when the word deviations are replaced with errors), RMSE (Root Mean Squared Error, square root of the average of the squares of one step prediction errors) and MAPE (Mean Absolute Percentage Error). Compute the measures for all models fitted in the previous exercise.

3. Use the best method (according to RMSE) for producing predictions together with prediction errors for the next 4 months.

Just by looking at some measures of the goodness of fit we can not determine if a method uses the information hidden in the series in a reasonably efficient way. A good method should use all of the available information and this means that prediction errors should appear to be completely random.

In the context of time series modelling, an important aspect of randomness is the lack of correlation between the observation separated by a fixed time step. Such correlations can be computed by the function acf. Mathematically this function assumes that the time series of prediction error corresponds to a process $(A_t)_{t\in\mathbb{Z}}$ for which the correlations between A_t and A_{t-k} do not depend on time. The command uses the data to estimate such correlations. If some of the correlations (especially for relatively small lags k) are significantly different from 0, the method is not describing correctly the time dependence between observations and should not be used for predicting future values.

Here one should take into account that even in the case of data produced by independent random variables from the same distribution, it may randomly happen that some correlations computed from the generated series may appear to be significantly different from 0. A test which considers a group of autocorrelations together and checks if the assumption of existence of non-zero autocorrelations can be proved is Ljung-Box test. For a good method, the p-value of the test for a reasonably large group of autocorrelations should be significantly larger than 0 (usually the acceptable level is taken to be 0.05).

5. Use the command **rnorm** to form a series from 100 values of independent random variables from the standard normal distribution. Produce the graph of autocorrelations with the command acf. Use the Ljung-Box test (Box.test with type="Ljung−Box") to check the independence of the group of 10 autocorrelations.

6. Check the independence of the errors of methods fitted in the first exercise. When testing a group of autocorrelations of prediction errors, one should actually use the option fitdf =no_of_params when using the command Box.test in order to obtain accurate results.

7. Determine the best exponential smoothing method for predicting the values of the dwelling price index. Check the suitability of the best method.