РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И ГОСУДАРСТВЕННОЙ СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПРОГНОЗИРОВАНИЕ ВВП С ПОМОЩЬЮ ИЕРАРХИЧЕСКИХ МОДЕЛЕЙ

Касьянова Ксения

ЭО-15-01

Научный руководитель: Демешев Борис Борисович

Goal and tasks:

Goal:

- Improve forecasts of aggregates time series (GDP) with the use of disaggregated time series (GVA by industry)

Tasks:

- Collect data with three-level hierarchical structure: for RF, US and EU
- Grouped Time Series analysis: compare aggregated and disaggregated forecasts
- Clustering of analogous time series to improve forecast quality by finding optimal combination of disaggregated time series

Hypothesis:

- Using three-level structure we can achieve better forecasts for an aggregated TS

- The intractable statistical complexity
- Higher levels may have substantial effects, but without the guidance of well-developed theory unproductive guesswork, data dredging and intractable statistical complications come to the fore
- Technically, a three-level model is a straightforward development of 2-level model

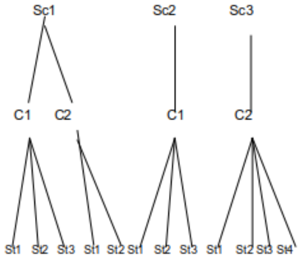


Рис.: Three-level models structure

Данные:

- для eu:
  - ⇒ 3-уровневая: eu28 по странам по a10 (+)
  - ⇒ 2-уровневая: ввп по каждой из стран по a10 (?)
  - ⇒ 2 уровневая: eu28 стран по ввп 28 стран (+)
- для rus:
  - ⇒ 3 уровневая: ввп по врп+налоги и прочее по отраслям (-)
  - ⇒ 2 уровневая: только по врп (-)
  - ⇒ 2 уровневая: только по отраслям (произведенный по оквэд) (+)
  - ⇒ 2 уровневая: только по использованию (+)
- для us:
  - ⇒ 3 уровневая: ввп по штатам и отраслям (?)
  - ⇒ 2 уровневая: ввп по штатам (+)

Data
US:

”Real Gross Domestic Product by Industry”:
Квартальные данные с 2005-01-01 - 2018-04-01 по 50 штатам
Millions of Chained 2012 Dollars
Seasonally Adjusted Annual Rate:

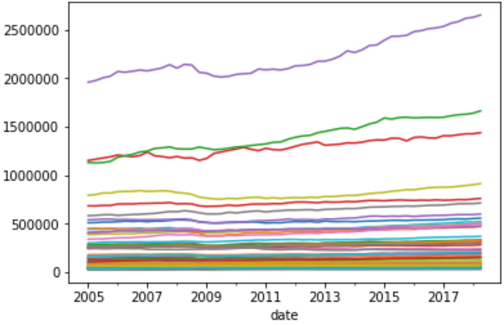


Рис.: Real Gross Domestic Product by State

Data
US:

Разница между ВВП США и суммой ВДС для каждого штата (не больше 2%):

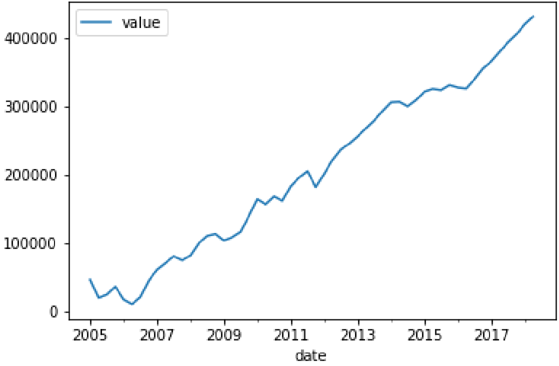


Рис.: Difference between GDP and sum of GVA by states:

Данные по ВРП (в текущих ценах):
для 79 регионов и 3 городов раз в год с 1995,

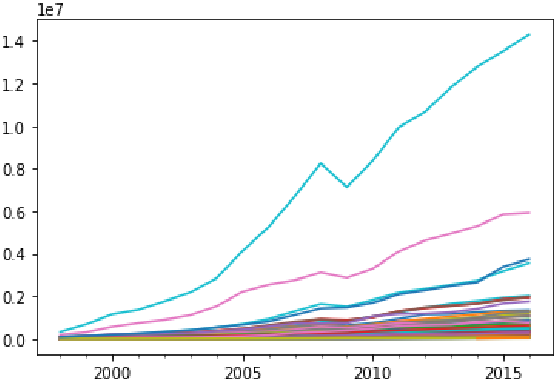


Рис.: ВРП по регионам

Данные (в процентах от ВРП) для 79 регионов и каждой из 12 отраслей:

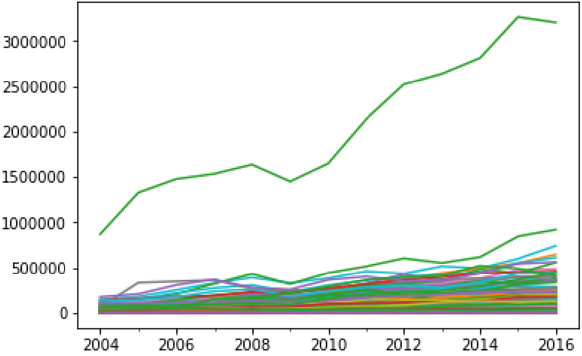
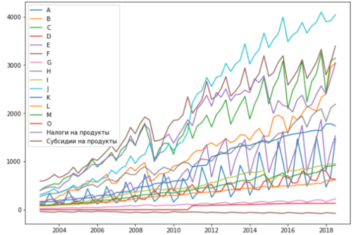


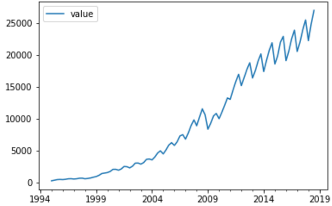
Рис.: ВРП по регионам по отраслям

Data

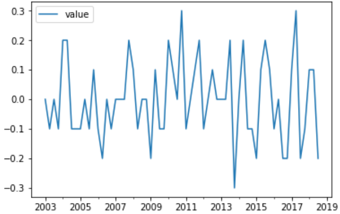
Для России:



ВВП квартальный (производственный метод):



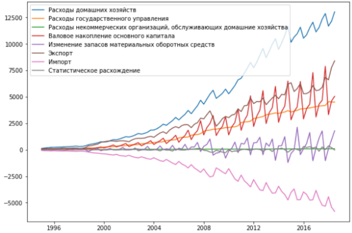
ВВП квартальный



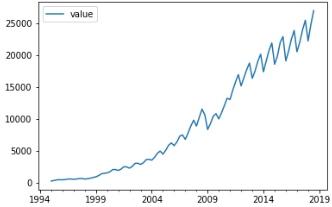
Разница (не более 0.00006 %)

Data

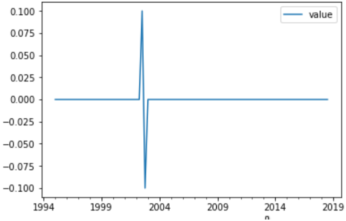
Для России:



ВВП квартальный (по расходам):



ВВП квартальный



Разница (не более 0.00006 %)

Data
ЕС:

Квартальные данные по ЕС по секторам:
2000-01-01 по 2018-07-01

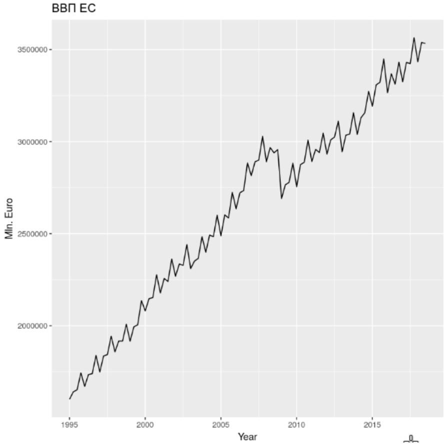


Рис.: ВВП ЕС28

Forecasting :

Aggregated TS:

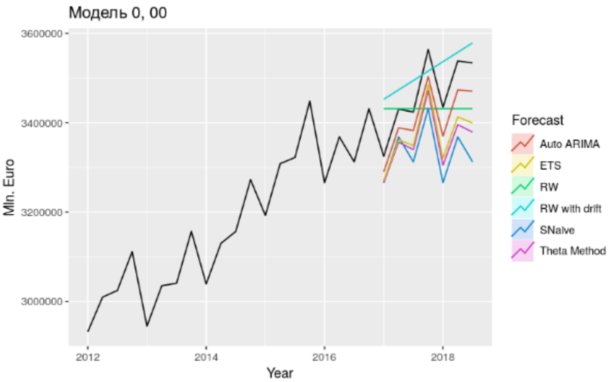


Рис.: Сравнение прогнозов

Forecasting :

Aggregated TS:

```
accuracy(gdp_rwf, test)
accuracy(gdp_rwfwd, test)
accuracy(gdp_snaive, test)
accuracy(gdp_theta, test)
accuracy(gdp_arima, test)
accuracy(gdp_ets, test)
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	21038.66	93246.38	79262.43	0.8102572	3.103855	0.8108200	-0.66298186	NA
Test set	32817.03	85139.78	65623.51	0.8959751	1.880483	0.6712998	0.04778851	0.8211688

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	9.099129e-10	90841.96	75477.66	-0.04039074	2.964913	0.7721034	-0.6629819	NA
Test set	-5.133762e+04	73862.83	65102.72	-1.51299717	1.899265	0.6659724	-0.4636484	0.6069265

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	80232.22	107828.1	97755.89	3.245085	3.871968	1.00000	0.8540931	NA
Test set	132158.14	143232.6	132158.14	3.791442	3.791442	1.35192	0.5292738	1.532192

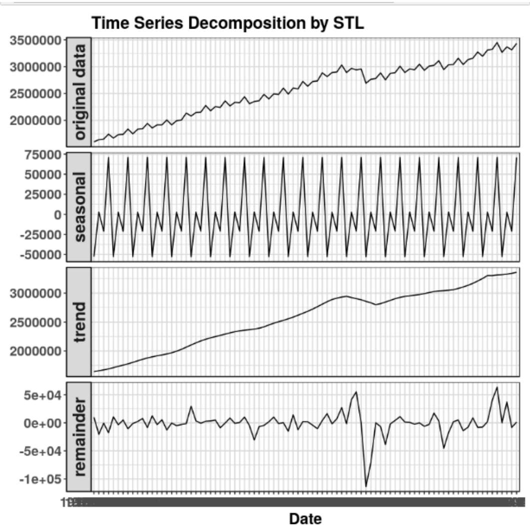
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	18640.08	59189.53	44033.26	0.7022174	1.756951	0.450441	-0.3429432	NA
Test set	105149.04	110592.60	105149.04	3.0218262	3.021826	1.075629	0.5889082	1.176769

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	496.5713	27093.45	18745.11	0.04021781	0.7075616	0.1917542	-0.006078177	NA
Test set	53108.4852	54480.16	53108.49	1.52742685	1.5274268	0.5432766	0.574894662	0.5760742

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	-1002.636	28131.63	19893.79	-0.03510539	0.7573732	0.2035047	0.2537223	NA
Test set	93065.844	97402.36	93065.84	2.67558565	2.6755857	0.9520229	0.5815305	1.034584

Рис.: Сравнение прогнозов

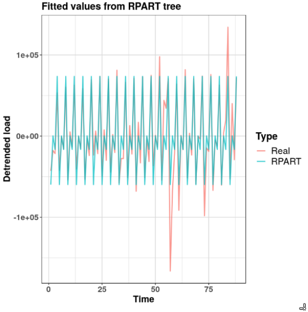
Forecasting :
Aggregated TS:

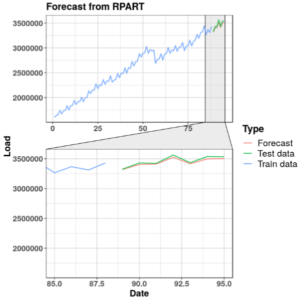


Forecasting :

Aggregated TS:

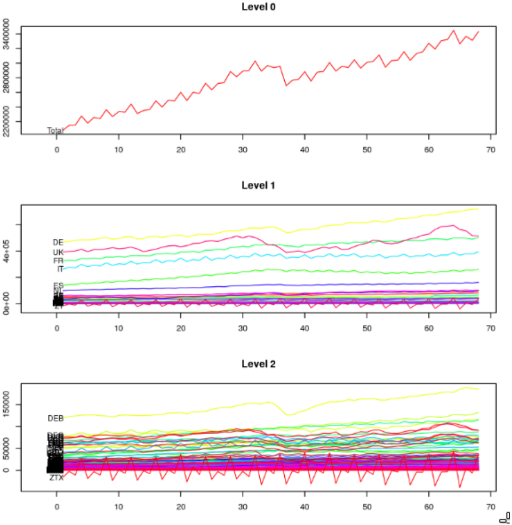
RPART, or CART (Classification and Regression Trees) is recursive partitioning type of binary tree for classification or regression tasks. It performs a search over all possible splits by maximizing an information measure of node impurity, selecting the covariate showing the best split.





	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
Test set	-24051.12	26940	24051.12	-0.6930563	0.6930563	-0.2798248	0.3608793

Forecasting :
Disaggregated TS



Grouped Time Series

Forecasting hierarchical time series

The assumption upon which many of these models are built on, is that by grouping series that behave in a similar way, the idiosyncratic errors within groups will tend to offset each other while the more relevant individual dynamics will be retained to be modelled.
Key idea: forecast reconciliation

- Ignore structural constraints and forecast every series of interest independently.
- Adjust forecasts to impose constraints.

Existing methods:

- Bottom-up
- Top-down
- Middle-out

An “optimal combination” approach can be advanced by proposing two new estimators based on WLS.
Both now implemented in the hts package

Grouped Time Series

Forecasting hierarchical time series

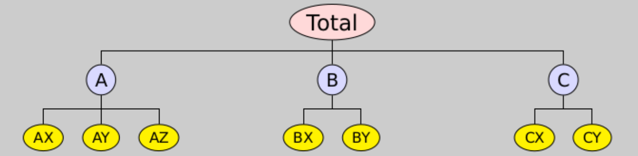


Рис.: Hierarchical structure

Y_t : observed aggregate of all series at time t.
$Y_{X,t}$: observation on series X at time t.

For the hierarchical structure we can write:

$$\begin{bmatrix} y_t \\ y_{A,t} \\ y_{B,t} \\ y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{AA,t} \\ y_{AB,t} \\ y_{AC,t} \\ y_{BA,t} \\ y_{BB,t} \end{bmatrix}$$

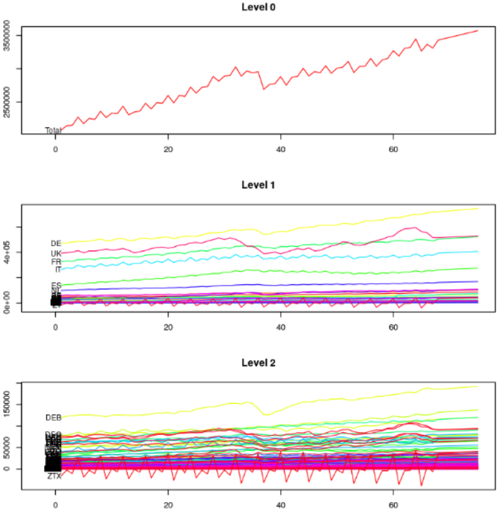
or in more compact notation, where y_t is an n-dimensional vector of all the observations in the hierarchy at time t, S is the summing matrix.

$$y_t = Sb_t$$

b_t : m-dimensional vector of all series at bottom level in time t.

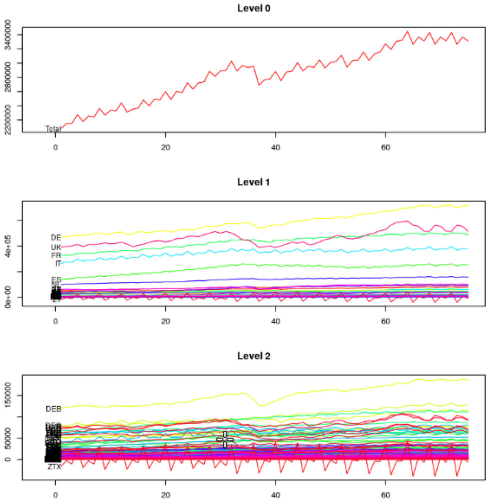
Grouped Time Series

RW with drift

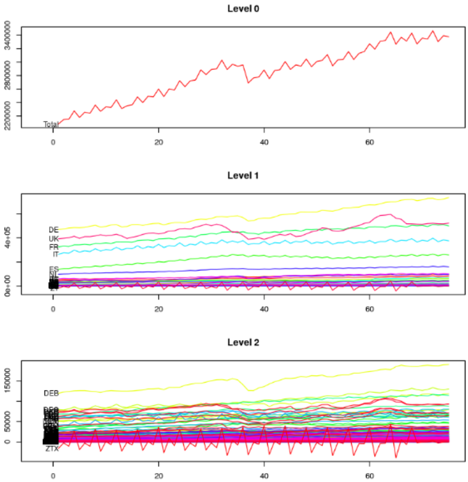


Grouped Time Series

Snaive

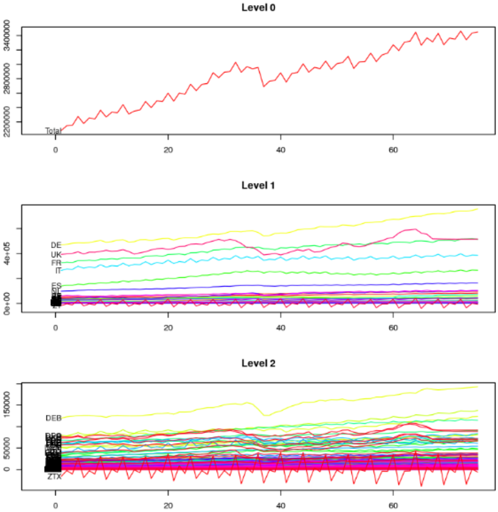


Equivalent to simple exponential smoothing with drift



Grouped Time Series

ARIMA



Grouped Time Series

	ME	RMSE	MAE	MPE	MAPE
Training set	9.099129e-10	90841.96	75477.66	-0.04039074	2.964913
Test set	-5.133762e+04	73862.83	65102.72	-1.51299717	1.899265

	Total
ME	-47804.571429
RMSE	71937.089305
MAE	62579.114286
MAPE	1.826455
MPE	-1.411862

	ME	RMSE	MAE	MPE	MAPE
Training set	80232.22	107828.1	97755.89	3.245085	3.871968
Test set	132158.14	143232.6	132158.14	3.791442	3.791442

	Total
ME	1.321581e+05
RMSE	1.432326e+05
MAE	1.321581e+05
MAPE	3.791442e+00
MPE	3.791442e+00

	ME	RMSE	MAE	MPE	MAPE
Training set	18640.08	59189.53	44033.26	0.7022174	1.756951
Test set	105149.04	110592.60	105149.04	3.0218262	3.021826

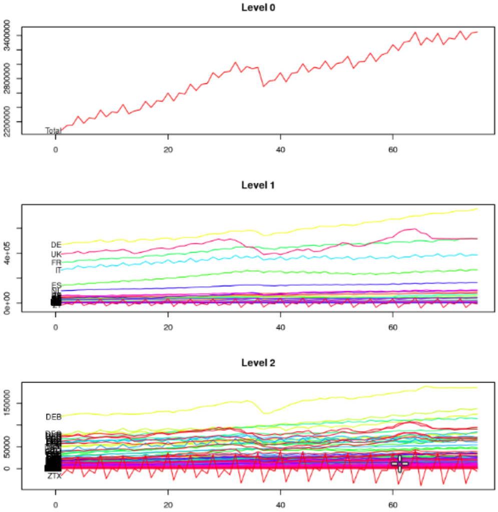
	Total
ME	1.058988e+05
RMSE	1.118594e+05
MAE	1.058988e+05
MAPE	3.041683e+00
MPE	3.041683e+00

	ME	RMSE	MAE	MPE	MAPE
Training set	496.5713	27093.45	18745.11	0.04021781	0.7075616
Test set	53108.4852	54480.16	53108.49	1.52742685	1.5274268

	Total
ME	81657.333946
RMSE	83547.346262
MAE	81657.333946
MAPE	2.348416
MPE	2.348416

Grouped Time Series

ARIMAX



Grouped Time Series

ARIMAX

	ME	RMSE	MAE	MPE	MAPE
Training set	496.5713	27093.45	18745.11	0.04021781	0.7075616
Test set	53108.4852	54480.16	53108.49	1.52742685	1.5274268

	Total
ME	81657.333946
RMSE	83547.346262
MAE	81657.333946
MAPE	2.348416
MPE	2.348416

	Total
ME	81627.482399
RMSE	83516.597308
MAE	81627.482399
MAPE	2.347558
MPE	2.347558

Forecast-error clustering

5. Forecast-error clustering

Ignoring the common factor and interdependencies will tend to make forecasting errors cluster instead of cancelling out. The dissimilarity measure the correlations of the out-of-sample forecasting errors for the most recent periods. Specifically, for each component i we fit $x_{i,t-p+1} = a_i + \rho x_{i,t-p} + e_{i,t}$, where p is the number of periods that are evaluated for the measure. With the model, we generate forecasts from $t - p + 1$ to t and calculate the corresponding forecasting errors as $\hat{x}_{i,s|s-1} - x_{i,s}$ for $s = t - p + 1$ to t and collect them in $\hat{e}_i^t$. With this, the dissimilarity measure is defined as:

$$FC_{x_i,x_j} = 1 - abs\left(\frac{cov(\hat{e}_i^t, \hat{e}_j^t)}{\sigma_{\hat{e}_i^t}\sigma_{\hat{e}_i^t}}\right)$$

The Diebold-Mariano Test

The loss associated with forecast i is assumed to be a function of the forecast error, e_{it}, and is denoted by $g(e_{it})$.

A problem with these loss function is that they are symmetric functions (squared-error loss, absolute error loss)

When it is more costly to underpredict y_t than to overpredict it, the following loss function can be used:

$$g(e_{it}) = \exp(\lambda e_{it}) - 1 - \lambda e_{it}$$

We define the loss differential between the two forecasts by

$$d_t = g(e_{1t}) - g(e_{2t})$$

and say that the two forecasts have equal accuracy if and only if the loss differential has zero expectation for all t.'

The Diebold-Mariano Test

So, we would like to test the null hypothesis

$$H_0 : E(d_t) = 0, \forall t$$

versus the alternative hypothesis

$$H_1 : E(d_t) \neq 0$$

The null hypothesis is that the two forecasts have the same accuracy. The alternative hypothesis is that the two forecasts have different levels of accuracy

Suppose that the forecasts are h(> 1)-step-ahead. In order to test the null hypotesis that the two forecasts have the same accuracy, Diebold-Mariano utilize the following statistic

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \sim N(0; 1)$$

where

$f_d(0) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_d(k)$ is the spectral density of the loss differential at frequency 0, $\gamma_d(k)$ is the autocovariance of the loss differential at lag k

$\hat{f}_d(0) = \frac{1}{2\pi} \sum_{k=-(T-1)}^{T-1} I(\frac{k}{h-1})\hat{\gamma}_d(k)$ is a consistent estimate of $f_d(0)$

$\hat{\gamma}_d(k) = 1/T \sum_{t=|k|+1}^{T} (d_t - \bar{d})(d_{t-|k|} - \bar{d})$

$$I(\tfrac{k}{h-1}) = \begin{cases} 1, & if\ |\frac{k}{h-1}| \leq 1 \\ 0, & otherwise \end{cases}$$

The Diebold-Mariano Test

Harvey, Leybourne, and Newbold (1997) (HLN) suggest that improved small-sample properties can be obtained by:

1. making a bias correction to the DM test statistic, and
2. comparing the corrected statistic with a Student-t

distribution with (T-1) degrees of freedom, rather than the standard normal.

$$HLN = DM\sqrt{(n+1-2h+h(h-1))/n} \sim T(n-1)$$

- The Diebold-Mariano test should not be applied to situations where the competing forecasts are obtained using two nested models, since at the population level, if the null hypothesis of equal predictive accuracy is true, the forecast errors from the competing models are exactly the same and perfectly correlated, which means that the numerator and denominator of a Diebold-Mariano test are each limiting to zero as the estimation sample and prediction sample grow.
- However, when the size of the estimation sample remains finite, parameter estimates are prevented from reaching their probability limits and the Diebold-Mariano test remains asymptotically valid even for nested models, under some regularity assumptions (see Giacomini and White 2003).

Shrinkage is implicit in Bayesian inference and penalized likelihood inference, and explicit in James–Stein-type inference. In contrast, simple types of maximum-likelihood and least-squares estimation procedures do not include shrinkage effects, although they can be used within shrinkage estimation schemes.

Stein’s paradox, in decision theory and estimation theory, is the phenomenon that when three or more parameters are estimated simultaneously, there exist combined estimators more accurate on average (that is, having lower expected mean squared error) than any method that handles the parameters separately.
The best guess about the future is usually obtained by computing the average of past events. Stein’s paradox defines circumstances in which there are estimators better that the arithmetic average
Stein’s paradox **modern generalization, the Bayesian hierarchical model**.
Bayesian hierarchical model can improve overall estimation accuracy, thereby improving our confidence in the assessment results, especially for standard compliance assessment of waters with **small sample sizes.**