Lecture 9-d
Time Series: Forecasting with ARIMA & Exponential Smoothing

 1

Review: Forecasting From ES Models

• **Exponential Smoothing Models** (ES) provide, quick, easy to program ("cheap") forecasts. .

• In general, these models are limited and not optimal, especially compared with Box-Jenkins methods.

• Goal of these models: Suppress the short-run fluctuation by smoothing the series. For this purpose, a weighted average of all previous values works well.

• There are many ES models. We cover two:
- Simple Exponential Smoothing (SES)
- Holt-Winter's Exponential Smoothing (HW ES).

Review: Simple Exponential Smoothing (SES)

- We use the observed time series at time t: $Y_1, Y_2, \ldots, Y_t$.

- **Level equation**: $S_t = \alpha Y_{t-1} + (1-\alpha)S_{t-1}$

where

- α: The smoothing parameter, $0 \le \alpha \le 1$.
- S_t: Value of the smoothed observation at time t –i.e., the forecast.

- The level equation is **recursive**. Using backward substitution:

$$S_{t+1} = \alpha Y_t + (1-\alpha)S_t = \alpha Y_t + (1-\alpha)[\alpha Y_{t-1} + (1-\alpha)S_{t-1}]$$
$$= \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + (1-\alpha)^2 S_{t-1}$$
$$\Rightarrow \; S_{t+1} = c_0 Y_t + c_1 Y_{t-1} + c_2 Y_{t-2} + \cdots$$

where $c_i = \alpha(1-\alpha)^i;\; i = 0, 1, \ldots;\; 0 \le \alpha \le 1.$

Now, S_{t+1} shows the "smoothing" of Y_t, with decreasing weights.

Review: SES − Forecast and Updating

- From the level equation, $S_t = \alpha Y_{t-1} + (1-\alpha)S_{t-1}$

we get: $S_{t+1} = S_t + \alpha(Y_t - S_t)$

$\Rightarrow$ A simple updating forecast: last period forecast + adjustment.

- At time t, For $t+2$, we have (since $S_{t+1} = Y_{t+1}$)

$$S_{t+2} = S_{t+1} + \alpha(Y_{t+1} - S_{t+1}) = S_{t+1}$$

- Then, at time t, the ℓ-step ahead forecast is:

$S_{t+\ell} = S_{t+1}$ $\Rightarrow$ A naive forecast!

Note: SES forecasts are not very interesting after $\ell > 1$.

Review: SES – Forecast and Updating

Example: An industrial firm uses SES to forecast sales:

$$S_{t+1} = S_t + \alpha * (Y_t - S_t)$$

The firm estimates α = **0.25**. The firm observes Y_t = **5** and, last period's forecast, S_t = **3**.

Then, the forecast for time $t+1$ is:

S_{t+1} = **3** + **0.25** * (**5** – **3**) = **3.50**

The forecast for time $t+1$ (& any period after time $t+1$) is:

$S_{t+\ell} = S_{t+1}$ = **3.50** for ℓ > 1.

Later, the firm observes: Y_{t+1} = **4.77**, Y_{t+2} = **3.15**, & Y_{t+3} = **1.85**. Then, the MSE:

MSE = $\frac{1}{3}$ * [(**4.77** – **3.50**)^2 + (**3.15** - **3.50**)^2 + (**1.85** - **3.50**)^2] = **1.486**.

Review: SES – Forecast and Updating

Example (continuation):

Note: If α = **0.75**, then

S_{t+1} = **3** + **0.75** * (**5** – **3**) = **4.50**

A bigger α gives more weight to the more recent observation –i.e., Y_t.

Again, the forecast for time t+1 and any period after time t+1 is:

$S_{t+\ell} = S_{t+1}$ = **4.50** for ℓ > 1.

Review: SES – Selecting α & Initial Values

- The parameter α is often selected as to minimize the MSE.

Numerical Minimization Process:
- Take different α values ranging between 0 and 1.
- Calculate 1-step-ahead forecast errors for each α.
- Calculate MSE for each case.

Choose α which has the min MSE: $e_t = Y_t - S_t \Rightarrow \min \sum_{t=1}^{n} e_t^2 \Rightarrow \alpha$

- SES produces a recursive equation, we need initial values, S_1:

– Use the recent past: Set S_1 equal to Y_1. Then, $S_2 = Y_1$.

– Use an average (say, of first 4 or 5 observations).

– Estimate S_1 (similar to the estimation of α.)

7

Review: SES – Forecasting U.S. Dividends

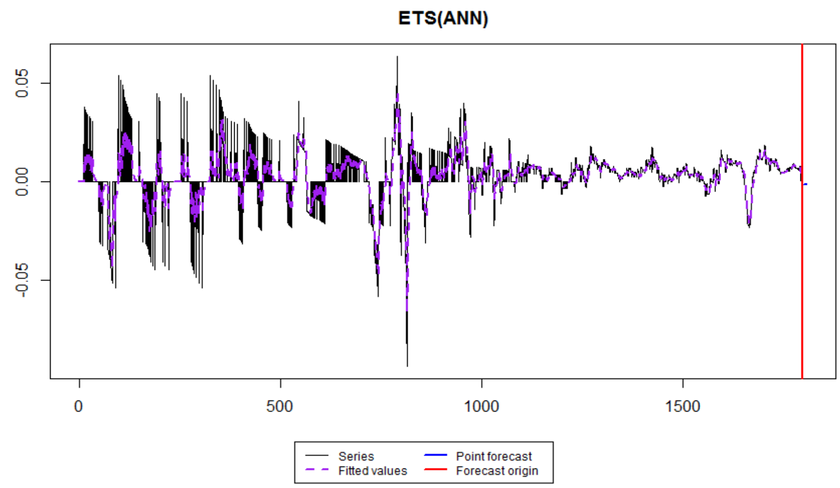
Example 1: We want to forecast log changes in **U.S. monthly dividends** (T=1796) using SES. First, we estimate the model using the R function *HoltWinters*(), which has as a special case SES: set beta=FALSE, gamma=FALSE. We use estimation period *T*=1750.

```
mod1 <- HoltWinters(lr_d[1:1750], beta=FALSE, gamma=FALSE)
> mod1
Holt-Winters exponential smoothing without trend and without seasonal component.

Call:
HoltWinters(x = lr_d[1:1750], beta = FALSE, gamma = FALSE)

Smoothing parameters:
 alpha: 0.289268                    ⇒ Estimated α
 beta : FALSE
 gamma: FALSE

Coefficients:
        [,1]
a 0.004666795                       ⇒ Forecast
```

8

Review: SES – Forecasting U.S. Dividends

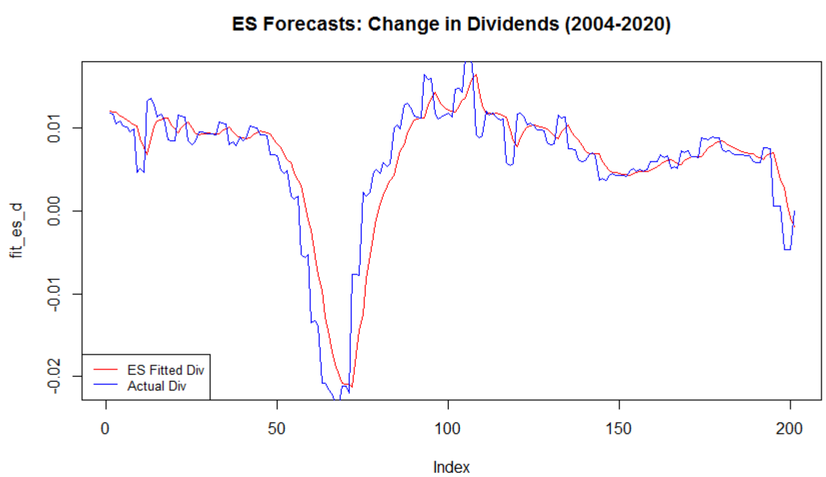
Example 1 (continuation):



9

Review: SES – Forecasting U.S. Dividends

Example 1 (continuation):



10

Review: SES – Remarks

- Some computer programs automatically select the optimal α using a line search method or non-linear optimization techniques. (R does the latter.)

- We have a recursive equation, we need initial values for S_1. (R takes an average of past observations.)

- This model ignores trends and/or seasonalities. Not very realistic, especially for manufacturing facilities, retail sector, and warehouses.

- Deterministic components, D_t, can be easily incorporated.

- The model that incorporates both a trend and seasonal features is called **Holt-Winter's ES**.

11

Review: Holt-Winters (HW) ES

- In the model for Y_t, in addition to the **level** (S_t), we introduce:
  - **Trend** (T_t) factor
  - **Seasonality** (I_t) factor.

Since we produce "smooth" forecasts for T_t & I_t, this method is also called **triple exponential smoothing**.

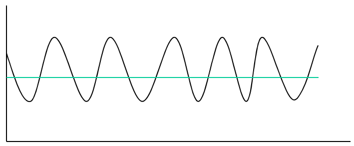
- The h-step ahead forecast combines the forecasts: S_t, T_t & I_{t+h-s}.

- Both, T_t & I_t, can be included as **additively** or **multiplicatively** factors. In this class, we consider an additive trend and the seasonal factor as additive or multiplicative. Then, the h-step ahead forecast:

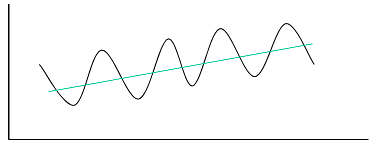
- For the additive model: $\hat{Y}_t(h) = S_t + h\,T_t + I_{t+h-s}$
- For the multiplicative model: $\hat{Y}_t(h) = (S_t + h\,T_t) * I_{t+h-s}$

12

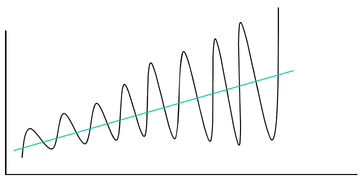
Review: HW ES – Trend & Seasonality



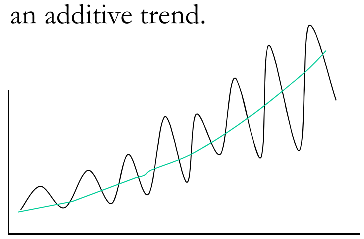
1. No trend and additive seasonal variability.



2. Additive seasonal variability with an additive trend.



3. Multiplicative seasonal variability with an additive trend.



4. Multiplicative seasonal variability with a multiplicative trend.

Note: We will use Model 2 (Additive) and Model 3 (Multiplicative).

Review: HW ES – Additive

• Additive model (additive trend & additive seasonality) forecast:

$$\hat{Y}_t(h) = S_t + h\,T_t + I_{t+h-s}$$

where s is the number of periods in seasonal cycles (=4 for quarters).

• Components:

- **The level**, S_t: A weighted average of "*seasonal adjusted*" Y_t (=$Y_t - I_{t-s}$), and the non-seasonal forecast ($S_{t-1} + T_{t-1}$):

$$S_t = \alpha(Y_t - I_{t-s}) + (1-\alpha)(S_{t-1} + T_{t-1})$$

- **The trend**, T_t: A weighted average of T_{t-1} and the change in S_t.

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

- **The seasonality**, I_t: A weighted average of seasonal index of s last year, I_{t-s}, and the current seasonal index ($Y_{t-1} - S_{t-1} - T_{t-1}$):

$$I_t = \gamma(Y_t - S_{t-1} - T_{t-1}) + (1-\gamma)I_{t-s}$$

14

Review: HW ES – Additive

- Then, the model for the h-step ahead forecast

$$\hat{Y}_t(h) = S_t + h\,T_t + I_{t+h-s}$$

has three equations:

Level: $S_t = \alpha(Y_t - I_{t-s}) + (1-\alpha)(S_{t-1} + T_{t-1})$

Trend: $T_t = \beta\,(S_t - S_{t-1}) + (1-\beta)\,T_{t-1}$

Seasonal: $I_t = \gamma(Y_t - S_{t-1} - T_{t-1}) + (1-\gamma)I_{t-s}$

- We have only three smoothing parameters:

  α = level coefficient

  β = trend coefficient

  γ = seasonality coefficient

15

Review: HW ES – Multiplicative

- In the multiplicative seasonal case (with an additive trend), we have the h-step ahead forecast:

$$\hat{Y}_t(h) = (S_t + h\,T_t) * I_{t+h-s}$$

- Details for *multiplicative* seasonality –i.e., Y_t/I_t– and *additive* trend

- The forecast, S_t, now shows the average Y_t adjusted ($\frac{Y_t}{I_{t-s}}$).
- The trend, T_t, is a weighted average of T_{t-1} and the change in S_t.
- The seasonality is also a weighted average of I_{t-s} and the Y_t/S_t.

- Then, the model has three equations:

$$S_t = \alpha\,\frac{Y_t}{I_{t-s}} + (1-\alpha)\,(S_{t-1} + T_{t-1})$$
$$T_t = \beta\,(S_t - S_{t-1}) + (1-\beta)\,T_{t-1}$$
$$I_t = \gamma\,\frac{Y_t}{S_t} + (1-\gamma)\,I_{t-s}$$

16

Review: HW ES – Multiplicative

• We think of (Y_t/S_t) as capturing *seasonal effects.*
s = # of periods in the seasonal cycles
(s = 4, for quarterly data; s = 12, for monthly)

• Again, we have only three parameters:
α = smoothing parameter
β = trend coefficient
γ = seasonality coefficient

• Q: How do we determine these 3 parameters?
- Ad-hoc method: α, β and γ can be chosen as values between
$0.02 < \alpha, \gamma, \beta < 0.2$
- Optimal method: Minimization of the MSE, as in SES.

17

Review: HW ES – Multiplicative

Example: An industrial firm uses HW ES to forecast sales next two quarters (h = 1, 2, & 3; with s = 4):

$$\hat{Y}_t(h) = \hat{Y}_{t+h} = (S_t + h\,T_t) * I_{t+h-s}$$

with S_t, T_t, & I_t factors given by:

$$S_t = \alpha \frac{Y_t}{I_{t-s}} + (1-\alpha)\,(S_{t-1} + T_{t-1})$$
$$T_t = \beta\,(S_t - S_{t-1}) + (1-\beta)\,T_{t-1}$$
$$I_t = \gamma \frac{Y_t}{S_t} + (1-\gamma)\,I_{t-s}$$

The firm estimates: $\alpha = 0.25$; $\beta = 0.1$; & $\gamma = 0.4$. It observes $Y_t = 5$; last quarter's smoothed forecasts: $S_{t-1} = 3$, $T_{t-1} = 1.2$; & last year's seasonal factors: $I_{t-4} = 1.1$, $I_{t-3} = 0.7$, $I_{t-2} = 1.2$, & $I_{t-3} = 0.8$.

• Components forecasts:

$$S_t = 0.25 \frac{5}{1.1} + (1 - 0.25) * (3 + 1.3) = 4.2864$$

Review: HW ES – Multiplicative

Example (continuation): (α = 0.25; β = 0.1; & γ = 0.4.)

$$S_t = 0.25 * \frac{5}{1.1} + (1 - 0.25) * (3 + 1.2) = 4.2864$$

$$T_t = 0.1 * (4.2864 - 3) + (1 - 0.1) * 1.2 = 1.2086$$

$$I_t = 0.4 * \frac{5}{4.2864} + (1 - 0.4) * 1.1 = 1.1266$$

The forecast for h = 1 (next quarter) is:

$$\hat{Y}_{t+1} = (4.2864 + 1.2086) * 0.7 = 4.8125$$

The forecast for h = 2 & 3 are:

$$\hat{Y}_{t+2} = (4.2864 + 2 * 1.2086) * 1.2 = 7.8475.$$

$$\hat{Y}_{t+3} = (4.2864 + 3 * 1.2086) * 0.8 = 6.1329.$$

Review: HW ES – Initial Values

- Initial values for algorithm

- Similar to the SES, we can use an average (using a year of past observations), use the recent past, or estimate the initial values by minimizing the MSE or MAE.

- We need at least one complete season of data to determine the initial estimates of I_{t-s}.

- R uses an average.

20

HW ES: Example – Log U.S. Vehicles Sales

Example: We want to forecast log U.S. monthly vehicle sales with HW. We use the R function *HoltWinters*().

```
l_car_18 <- l_car[1:512]
l_car_ts <- ts(l_car_18, start = c(1976, 1), frequency = 12)     # convert lr_d in a ts object
hw_d_car <- HoltWinters(l_car_18, seasonal="additive")
> hw_d_car
Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:
HoltWinters(x = lr_d_ts, seasonal = "additive")

Smoothing parameters:
 alpha: 0.4355244               ⇒ Estimated smoothing parameter
 beta : 0.009373815             ⇒ Estimated trend parameter ≈ 0 (no trend)
 gamma:0.3446495                ⇒ Estimated seasonal parameter
```

21

HW ES: Example – Log U.S. Vehicles Sales

Example (continuation):

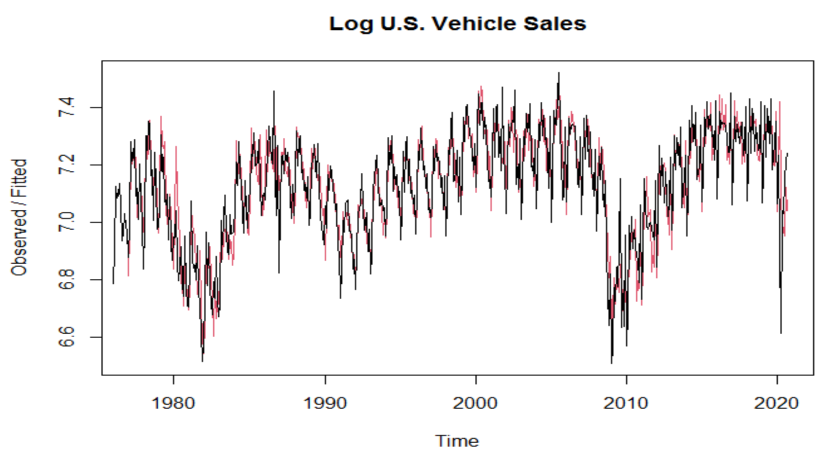
```
> hw_d_car

Coefficients:
          [,1]
a    7.177857555             ⇒ forecast for level
b    0.0001100345            ⇒ forecast for trend
s1  -0.075314457             ⇒ forecast for seasonal month 1
s2  -0.084468361             ⇒ forecast for seasonal month 2
s3   0.049447067
s4  -0.273299309
s5  -0.138251757
s6  -0.026603921
s7  -0.144953062
s8   0.079214066
s9   0.037899454
s10  0.020477134
s11  0.089309775
s12 -0.012530316
```

22

HW ES: Example – Log U.S. Vehicles Sales

Example (continuation):

```
plot(hw_d_car)
```



SES: Forecasting Log U.S. Vehicles Sales

Example (continuation): Now, we forecast one-step ahead forecasts

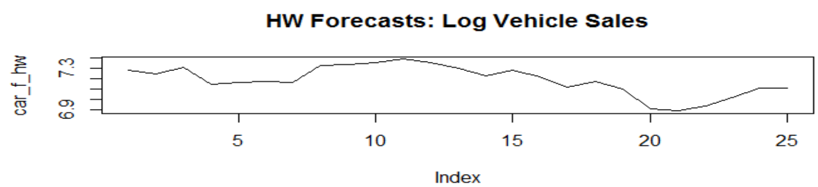
```
T_last <- nrow(hw_d_car$fitted)
h <- 25
ses_f_hw <- matrix(0,h,1)
 alpha <- 0.4355244
 beta <- 0.009373815
 gamma <- 0.3446495
y <- l_car
T <- length(l_car)
sm <- matrix(0,T,1)
Tr <- matrix(0,T,1)
I <- matrix(0,T,1)
T1 <- T-h+1
a <- T1
sm[a-1] <- 7.177857555
Tr[a-1] <- -0.000309358
I[501:512] <- c(-0.075314457,-0.084468361,0.049447067,-0.273299309,-0.138251757, -
0.026603921, -0.144953062,0.079214066,0.037899454,0.020477134,0.089309775,-
0.012530316)
```

24

SES: Forecasting Log U.S. Vehicles Sales

Example (continuation):

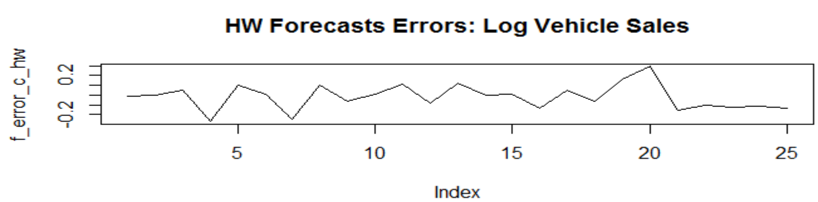
```
while (a <= T) {
        sm[a] = alpha * y[a-1] +  (1-alpha) * sm[a-1]
        Tr[a] = beta * (sm[a] - sm[a-1]) + (1 - beta) * Tr[a-1]
        I[a] = gamma * (y[a] - sm[a]) + (1 - gamma) * I[a - 12]
a <- a + 1
}
hh <- c(1:h)
car_f_hw <- sm[T1:T] + hh*Tr[T1:T] + I[T1:T]
car_f_hw
f_error_c_hw<- car_f_hw - y[T1:T]
plot(car_f_hw, type="l", main ="SES Forecasts: Log Vehicle Sales")
```



SES: Forecasting Log U.S. Vehicles Sales

Example (continuation):

```
plot(f_error_c_hw, type="l", main ="SES Forecasts Errors: Log Vehicle Sales")
```




```
MSE_hw <- sum(f_error_c_hw^2)/h
> MSE_hw
[1] 0.01655964
```

26

Evaluation of forecasts – Accuracy measures

• The mean squared error (*MSE*) and mean absolute error (*MAE*) are the most popular accuracy measures:

$$\text{MSE} = \frac{1}{m}\sum_{i=T+1}^{T+m}(\hat{y}_i - y_i)^2 = \frac{1}{m}\sum_{i=T+1}^{T+m} e_i{}^2$$

$$\text{MAE} = \frac{1}{m}\sum_{i=T+1}^{T+m}|\hat{y}_i - y_i| = \frac{1}{m}\sum_{i=T+1}^{T+m}|e_i|$$

where m is the number of out-of-sample forecasts.

• But other measures are routinely used:

- Mean absolute percentage error (*MAPE*) = $\frac{100}{T-(m-1)}\sum_{i=T+1}^{T+m}|\frac{\hat{y}_i - y_i}{y_i}|$

- Absolute *MAPE (AMAPE*) = $\frac{100}{T-(m-1)}\sum_{i=T+1}^{T+m}|\frac{\hat{y}_i - y_i}{\hat{y}_i + y_i}|$

Remark: There is an asymmetry in MAPE, the level y_i matters.

Evaluation of forecasts – Accuracy measures

- % correct sign predictions (PCSP) = $\frac{1}{T-(m-1)}\sum_{i=T+1}^{T+m} z_i$

where z_i = 1 if $(\hat{y}_{i+l} * y_{i+l}) > 0$
= 0, otherwise.

- % correct direction change predictions (PCDP)= $\frac{1}{T-(m-1)}\sum_{i=T+1}^{T+m} z_i$

where z_i = 1 if $(\hat{y}_{i+l} - y_i) * (y_{i+l} - y_i) > 0$
= 0, otherwise.

Remark: We value forecasts with the right direction (sign) or forecast that can predict turning points. For stock investors, the sign matters!

• MSE penalizes large errors more heavily than small errors, the sign prediction criterion, like MAE, does not penalize large errors more.

Evaluation of forecasts – Accuracy measures

Example: We compute MSE and the % of correct direction change (PCDC) predictions for the one-step forecasts for U.S. monthly vehicles sales based on the SES and HW ES models.

```
> MSE_ses
[1] 0.027889

> MSE_hw
[1] 0.01655964
```

• We calculate PCDC with following script for HW & SES:

```
bb_hw <- (car_f_hw - y[(T1-1):(T-1)]) * (y[T1:T] - y[(T1-1):(T-1)])
indicator_hw <- ifelse(bb_hw > 0,1,0)     # ifelse ("if else") produces a 1 if condition is true
pcdc_hw <- sum(indicator_hw)/h
> indicator_hw
 [1] 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 0 1 1 1 1 1 1 0 0 0
> pcdc_hw
[1] 0.76
```

Evaluation of forecasts – Accuracy measures

Example (continuation):

```
bb_s <- (ses_f_c - y[(T1-1):(T-1)]) * (y[T1:T] - y[(T1-1):(T-1)])
indicator_s <- ifelse(bb_s > 0,1,0)
pcdc_s <- sum(indicator_s)/h
> indicator_s
 [1] 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 0 0 0
> pcdc_s
[1] 0.76
```

Note: Same percentage of correct direction change (PCDC) predictions, but the sequence of correct predictions is not the same.

Evaluation of forecasts – DM Test

• To determine if one model predicts better than another, we define the loss differential between two forecasts:

$$d_t = g(e_t^{M1}) - g(e_t^{M2})$$

where g(.) is the forecasting loss function, M1 and M2 are two competing sets of forecasts –could be from models or something else.

• We only need $\{e_t^{M1}\}$ & $\{e_t^{M2}\}$, not the structure of M1 or M2. In this sense, this approach is "*model-free*."

• Typical (symmetric) loss functions: $g(e_t) = e_t^2$ & $g(e_t) = |e_t|$.

• But other g(.)'s can be used: $g(e_t) = \exp(\lambda\, e_t^2) - \lambda\, e_t^2$ $(\lambda>0)$.

Note: This is a more general test than MGN: It works for any loss function, not just MSE.

Evaluation of forecasts – DM Test

• Then, we test the null hypotheses of equal predictive accuracy:

$$H_0: E[d_t] = 0$$
$$H_1: E[d_t] = \mu \neq 0.$$

- Diebold and Mariano (1995) assume $\{e_t^{M1}\}$ & $\{e_t^{M2}\}$ is covariance stationarity and other regularity conditions (finite Var[d_t], independence of forecasts after ℓ periods) needed to apply CLT. Then,

$$\frac{\bar{d} - \mu}{\sqrt{Var[\bar{d}]/T}} \xrightarrow{d} N(0,1), \qquad \bar{d} = \frac{1}{m}\sum_{i=T+1}^{T+m} d_i$$

• Then, under H_0, the DM test is a simple *z-test*:

$$DM = \frac{\bar{d}}{\sqrt{\widehat{V}ar[\bar{d}]/T}} \xrightarrow{d} N(0,1)$$

Evaluation of forecasts – DM Test

where $\hat{V}ar[\bar{d}]$ is a consistent estimator of the variance, usually based on sample autocovariances of d_t:

$$\hat{V}ar[\bar{d}] = \gamma(0) + 2\sum_{j=k}^{\ell}\gamma(j)$$

• There are some suggestion to calculate small sample modification of the DM test. For example, :

$$\text{DM}^* = \text{DM}/\{[T + 1 - 2\,\ell + \ell\,(\ell - 1)/T]/T\}^{1/2} \sim t_{T-1}.$$

where ℓ-step ahead forecast. If time-varying volatility (ARCH) is suspected, replace ℓ with $[0.5\,\sqrt{(T)}] + \ell$.

Note: If $\{e_t^{M1}\}$ & $\{e_t^{M2}\}$ are perfectly correlated, the numerator and denominator of the DM test are both converging to 0 as $T \to \infty$.

⇒ Avoid DM test when this situation is suspected (say, two nested models.) Though, in small samples, it is OK.

Evaluation of forecasts – DM Test

Example: Code in R

```
dm.test <- function (e1, e2, h = 1, power = 2) {
d <- c(abs(e1))^power - c(abs(e2))^power
 d.cov <- acf(d, na.action = na.omit, lag.max = h - 1, type = "covariance", plot = FALSE)$acf[, , 1]
 d.var <- sum(c(d.cov[1], 2 * d.cov[-1]))/length(d)
 dv <- d.var                                   #max(1e-8,d.var)
 if(dv > 0)
   STATISTIC <- mean(d, na.rm = TRUE) / sqrt(dv)
 else if(h==1)
   stop("Variance of DM statistic is zero")
 else
 {
   warning("Variance is negative, using horizon h=1")
   return(dm.test(e1,e2,alternative,h=1,power))
 }
   n <- length(d)
 k <- ((n + 1 - 2*h + (h/n) * (h-1))/n)^(1/2)
 STATISTIC <- STATISTIC * k
 names(STATISTIC) <- "DM"
}
```

Evaluation of forecasts – DM Test

Example: We compare the SES and HW forecasts for the log of U.S. monthly vehicle sales. We use the *dm.test* function, part of the forecast package.

```
library(forecast)
> dm.test(f_error_c_ses, f_error_c_hw, power=2)

        Diebold-Mariano Test

data:  f_error_c_sesf_error_c_hw
DM = 1.6756, Forecast horizon = 1, Loss function power = 2, p-value = 0.1068
alternative hypothesis: two.sided

> dm.test(f_error_c_ses,f_error_c_hw, power=1)

        Diebold-Mariano Test

data:  f_error_c_sesf_error_c_hw
DM = 1.94, Forecast horizon = 1, Loss function power = 1, p-value = 0.064
alternative hypothesis: two.sided
```

Note: Cannot reject H_0: $MSE_{SES} = MSE_{HW}$ at 5% level

Evaluation of forecasts – DM Test: Remarks

- The DM tests is routinely used. Its "model-free" approach has appeal. There are model-dependent tests, with more complicated asymptotic distributions.

- The loss function does not need to be symmetric (like MSE).

- The DM test is based on the notion of unconditional –i.e., on average over the whole sample- expected loss.

- Following Morgan, Granger and Newbold (1977), the DM statistic can be calculated by regression of d_t, on an intercept, using NW SE. But, we can also condition on variables that may explain d_t. We move from an unconditional to a conditional expected loss perspective.

Combination of Forecasts

• Idea – from Bates & Granger (*Operations Research Quarterly,* 1969):

- We have different forecasts from R models:

$$\hat{Y}_T^{M1}(\ell), \hat{Y}_T^{M2}(\ell), \quad \dots, \hat{Y}_T^{MR}(\ell)$$

• Q: Why not combine them?

$$\hat{Y}_T^{Comb}(\ell) = \omega_{M1}\hat{Y}_T^{M1}(\ell) + \omega_{M2}\hat{Y}_T^{M2}(\ell) + \dots + \omega_{MR}\hat{Y}_T^{MR}(\ell)$$

• Very common practice in economics, finance and politics, reported by the press as "consensus forecast." Usually, as a simple average.

• Q: Advantage? Lower forecast variance. Diversification argument.

Intuition: Individual forecasts are each based on partial information sets (say, private information) or models.

37