

# Análise de Séries Temporais: Métricas de acurácia e validação cruzada

Prof. Dr. José Augusto Fiorucci

Universidade de Brasília  
Departamento de Estatística

# Métricas para a qualidade de ajuste

- Métricas tradicionais

- Mean Absolute Error

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

- Mean Squared Error

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2$$

- Root mean squared error

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

# Métricas para a qualidade de ajuste

- Métricas livre de escala
  - Mean Absolute Percentage Error

$$MAPE = 100 \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{|y_t|}$$

- Mean Squared Percentage Error

$$MSPE = \frac{1}{n} \sum_{t=1}^n \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2$$

- Root Mean Squared Percentage Error

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2}$$

# Métricas de acurácia para previsão

- Métricas dependentes de escala

- Mean Absolute Error

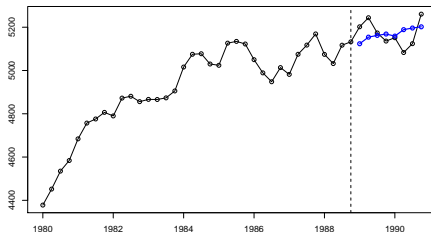
$$MAE = \frac{1}{h} \sum_{i=1}^h |y_{n+i} - \hat{y}_{n+i|n}|$$

- Mean Squared Error

$$MSE = \frac{1}{h} \sum_{i=1}^h (y_{n+i} - \hat{y}_{n+i|n})^2$$

- Root mean squared error

$$RMSE = \sqrt{\frac{1}{h} \sum_{i=1}^h (y_{n+i} - \hat{y}_{n+i|n})^2}$$



# Métricas de acurácia para previsão: MAPE

- Mean Absolute Percentage Error

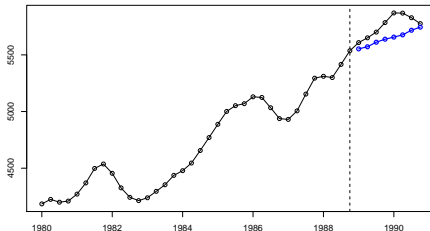
$$MAPE = \frac{100}{h} \sum_{i=1}^h \frac{|y_{n+i} - \hat{y}_{n+i}|}{|y_{n+i}|}$$

- Utilizada nas competições M1 (1982) e M2 (1993))

- Invariante por escala

- Não simétrica,

$$MAPE(Y, \hat{Y}) \neq MAPE(\hat{Y}, Y)$$

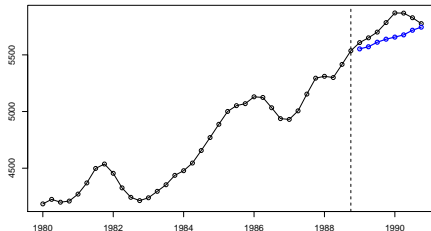


# Métricas de acurácia para previsão: sMAPE

- symmetric Mean Absolute Percentage Error

$$sMAPE = \frac{100}{h} \sum_{i=1}^h \frac{|y_{n+i} - \hat{y}_{n+i}|}{(|y_{n+i}| + |\hat{y}_{n+i}|)/2}$$

- Proposta e utilizada na competição M3 (2000)
- Invariante por escala
- Simétrica
- Favorece erros positivos 😊



# Métricas de acurácia para previsão: MASE

- Mean Absolute Scaled Error

$$MASE = \frac{n - f}{h} \frac{\sum_{i=1}^h |y_{n+i} - \hat{y}_{n+i}|}{\sum_{t=f+1}^n |y_t - y_{t-f}|}$$

sendo  $f$  a frequência da série. Exemplo,  $f = 1$  pra séries anuais e  $f = 12$  para séries mensais.

- Hyndman & Koehler (2006);
- Resolve todos os problemas anteriores;
- Métrica mais aceita atualmente;

# Exemplo

Médias das métricas sMAPE e MASE calculadas para as 3003 séries do banco de dados da M3 (Fiorucci et al., 2016)

Data	Methods	Yearly	Quarterly	Monthly	Other	All	Time (min)
sMAPE metric (%)							
Original data	Naive	17.88	11.32	18.18	6.30	16.58	0.31
	SES	17.78	10.83	16.14	6.30	15.07	0.22
	Damped	17.07	10.96	16.25	<b>4.30</b>	15.02	1.05
	ETS	16.89	9.69	14.07	4.34	13.28	38.36
	ARIMA	17.62	9.99	15.30	4.54	14.27	23.48
Seasonally adjusted data	Naive	17.88	10.02	16.76	6.30	15.38	0.62
	SES	17.78	9.77	14.17	6.30	13.53	0.55
	Damped	17.07	9.79	13.96	<b>4.30</b>	13.24	1.52
	ETS	16.89	9.80	14.04	4.34	13.28	32.87
	ARIMA	17.62	9.74	15.68	4.54	14.49	21.40
	STheta	16.74	9.23	13.83	4.93	13.05	7.37
	DOTM	<b>15.94</b>	<b>9.28</b>	<b>13.74</b>	<b>4.58</b>	<b>12.90</b>	13.59
MASE metric							
Original data	Naive	3.17	1.46	1.17	3.09	1.50	0.31
	SES	3.17	1.41	1.09	3.10	1.43	0.22
	Damped	2.92	1.37	1.10	1.81	1.36	1.05
	ETS	2.83	1.18	0.86	<b>1.79</b>	1.15	38.36
	ARIMA	2.99	1.17	0.88	1.87	1.19	23.48
Seasonally adjusted data	Naive	3.17	1.25	1.04	3.09	1.37	0.62
	SES	3.17	1.24	0.93	3.10	1.29	0.55
	Damped	2.92	1.17	0.88	1.81	1.17	1.52
	ETS	2.83	1.18	0.87	<b>1.79</b>	1.16	32.87
	ARIMA	2.99	1.15	0.89	1.87	1.19	21.40
	STheta	2.77	1.12	0.86	2.28	1.16	7.37
	DOTM	<b>2.59</b>	1.12	<b>0.85</b>	<b>1.94</b>	<b>1.12</b>	13.59

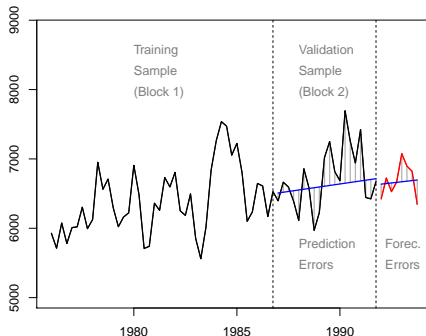


# Validação cruzada:

- Ideia básica: o melhor modelo é o modelo que gera as melhores previsões;
- A série é dividida em duas partes:
  - Treinamento: parte em que o modelo é ajustado
  - Teste: parte utilizada para verificar a acurácia do modelo
- Vários modelos são ajustados na parte de treinamento
- O modelo que obter as melhores previsões para a parte de teste (de acordo com uma métrica) é escolhido como o melhor

# Validação cruzada: avaliação com origem fixa

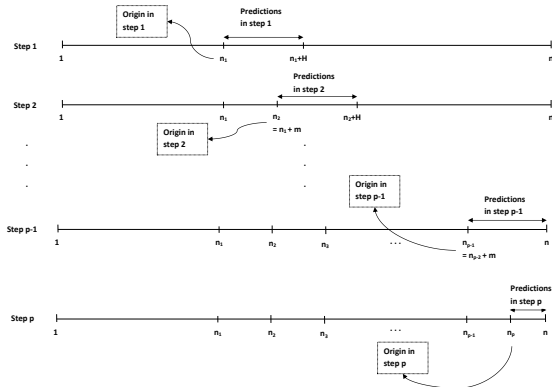
- Apenas uma origem e uma aplicação do processo;



# Validação cruzada: avaliação por janela deslizante

## • GROE: Generalized Rolling Origin Evaluation

- Processo que permite  $p$  origens com regiões de teste sobrepostas;
- Origens para previsão:  $n_1, n_2, \dots, n_p$ ;
- Função `groe()` do pacote *forecTheta* e função `tsCV()` do pacote *forecast*.



# Exemplo: avaliação por janela deslizante

```
require(forecTheta)
require(forecast)
require(Mcomp)
data(M3)
y <- M3[[1]]$x
## avaliação com janela deslizante
## origens: n1=n-7, n2=n-5, n3=n-3, n4=n-1
## retorna a soma dos erros absolutos
groey(y, ses, g="AE", n1=length(y)-7,m=2,H=5)
```

```
[1] 12648.21
```

```
groey(y, holt, g="AE", n1=length(y)-7,m=2,H=5)
```

```
[1] 4051.717
```

```
groey(y, function(y,h) forecast(ets(y), h), g="AE",n1=length(y)-7,m=2,H=5)
```

```
[1] 10701.33
```

```
groey(y, function(y,h) forecast(auto.arima(y), h), g="AE", n1=length(y)-7,m=2,H=5)
```

```
[1] 3111.654
```