Stat 443. Time Series and Forecasting.

Concepts at beginning of term: time series data, trend and seasonality, autocorrelation, forecasting rules, assessing and comparing different forecast rules.

Observed data $y_1, \ldots, y_t, \ldots, y_T$ of a variable z at time points $1, \ldots, T$. Realization of a random process $Y_1, \ldots, Y_t, \ldots, Y_T$.

Forecasting :
$\widehat{y}_{t+1|t}$ is the forecast of Y_{t+1} based on $y_1, \ldots, y_t$ (1-step forecast).
$\widehat{y}_{t+h|t}$ is the forecast of Y_{t+h} based on $y_1, \ldots, y_t$ (h-step forecast).

What are some reasonable forecast rules for the time series data examples presented earlier?

Training and holdout subsets

Split $y_1, \ldots, y_T$ into a training set $y_1, \ldots, y_n$ where $1 < n < T$ (say $n/T \approx 0.8$). Let $y_{n+1}, \ldots, y_T$ be the holdout set.

A 1-step forecast rule is based on a function g of past observations

$$\widehat{y}_{t+1|t} = g_t(y_1, \ldots, y_t)$$

and this rule might depend on estimated parameters using the training set. Apply the 1-step forecast rule for $t = n, \ldots, T-1$ to get $\widehat{y}_{n+1|n}, \ldots, \widehat{y}_{T|T-1}$. For example, $\widehat{y}_{n+1|n}$ is a forecast of Y_{n+1} before the observation at time $n+1$ is realized and the forecast error is $y_{n+1} - \widehat{y}_{n+1|n}$ after $y_{n+1}, y_n, \ldots$ are observed.
The root mean square error for the forecast rule is:

$$rmse = \left[(T-n)^{-1} \sum_{i=n+1}^{T} (y_i - \widehat{y}_{i|i-1})^2 \right]^{1/2}$$

Out-of-sample root mean square forecast error

Add superscripts/subscripts when referring to more than one forecasting rule.

If there are M forecast rules based on functions $g_1, \ldots, g_M$, let:

$$\widehat{y}_{i|i-1}^{(m)} = g_t^{(m)}(y_1, \ldots, y_t), \ldots, m = 1, \ldots, M,$$

then their out-of-sample rms forecast errors can be compared:

$$rmse_m = \left[(T-n)^{-1} \sum_{i=n+1}^{T} (y_i - \widehat{y}_{i|i-1}^{(m)})^2 \right]^{1/2}$$

A better forecast rule leads to a smaller rmse. There are assumptions in order that rmse is a good performance measure for out-of-sample prediction

Examples of forecasting rules:

1. persistence (previous observation)
2. average of all observations in training set
3. linear in the previous observation
4. average of all past in the same month in training set (monthly data with annual seasonal cycle) or persistence in the same month
5. exponentially weighted sum of all previous observations
6. linear Holt exponential smoothing
7. Winters additive seasonal smoothing
8. Winters multiplicative seasonal smoothing

Persistence: $\widehat{y}_{t+1|t} = y_t$; use most recent observation as forecast at next time point

- input `train` with size n, `holdout` with size $n_{holdout}$
- mse$\leftarrow$ 0
- fc$\leftarrow$ train[n]; yt$\leftarrow$ holdout[1]; fcerror$\leftarrow$ yt-fc; mse$\leftarrow$ mse+ fcerror^2.
- for i in $2, \ldots, n_{holdout}$:
- yt$\leftarrow$ holdout[i]; fc$\leftarrow$ holdout[i-1];
- fcerror$\leftarrow$ yt-fc; mse$\leftarrow$ mse+ fcerror^2.
- end for
- return rmse=sqrt(mse/$n_{holdout}$)

Average: forecast $\widehat{y}_{t+1|t}$ is the average of all observations in training set

- input `train` with size n, `holdout` with size $n_{holdout}$
- average $\leftarrow n^{-1} \sum_{i=1}^{n}$ train[i]
- mse$\leftarrow$ 0
- for i in $1, \ldots, n_{holdout}$:
- yt$\leftarrow$ holdout[i]; fc $\leftarrow$ average;
- fcerror$\leftarrow$ yt-fc; mse$\leftarrow$ mse+ fcerror^2.
- end for
- return rmse=sqrt(mse/$n_{holdout}$)

Forecast in linear in the previous observation : for example, in R, use `lm` or `lsfit` on $(y_1, y_2), \ldots, (y_{n-1}, y_n)$;

- input `train` with size n, `holdout` with size $n_{holdout}$
- for training set get least squares estimates $\widehat{\beta}_0, \widehat{\beta}_1$ to minimize $S(b_0, b_1) = \sum_{i=2}^{n}(y_i - b_0 - b_1 y_{i-1})^2$.
- yt$\leftarrow$ holdout[1]; fc$\leftarrow$ $\widehat{\beta}_0 + \widehat{\beta}_1$train[$n$];
- fcerror$\leftarrow$ yt-fc; mse$\leftarrow$ fcerror^2.
- for i in $2, \ldots, n_{holdout}$:
- fc$\leftarrow$ $\widehat{\beta}_0 + \widehat{\beta}_1$holdout[$i-1$]; yt $\leftarrow$holdout[i];
- fcerror$\leftarrow$ yt-fc; mse$\leftarrow$ mse+ fcerror^2.
- end for; return rmse=sqrt(mse/$n_{holdout}$)

In-sample prediction rmse

Example for linear in previous observation.
Training set $y_1, \ldots, y_n$; estimate parameters (μ, ϕ) or (β_0, β_1) using least squares for $(y_1, y_2), \ldots, (y_{n-1}, y_n)$.
Get $\widehat{\beta}_0, \widehat{\beta}_1$ or $\widehat{\mu}, \widehat{\phi}$.
In-sample predictions are $\widehat{y}_2 = \widehat{\beta}_0 + \widehat{\beta}_1 y_1 = \widehat{\mu} + \widehat{\phi}(y_1 - \widehat{\mu}), \ldots, \widehat{y}_n = \widehat{\beta}_0 + \widehat{\beta}_1 y_{n-1} = \widehat{\mu} + \widehat{\phi}(y_{n-1} - \widehat{\mu})$.
In-sample root-mean-square prediction error is $[(n-1)^{-1} \sum_{i=2}^{n} e_i^2]^{1/2}$ where $e_i = y_i - \widehat{y}_i$ is the prediction/forecast error for the ith time point, **after** fitting a model to the training data.
In-sample rmse tends to be smaller than out-of-sample rmse. The latter are better for comparing forecast rules, but statistical software tend to only provide the in-sample rmse.

Note two different parametrization for linear in previous observation. `arima` in R uses the (μ, ϕ) parametrization.

Modifications of forecast rules for seasonality

There are versions of the preceding for time series data with a clear cyclic (seasonal) pattern; for example,
annual cycle for average monthly temperature: forecast rule is average all previous observation in the relevant month, or the last observation in the relevant month (persistence)

Exercises:

1. Implement algorithms in previous pseudo-code into R functions and apply to the data sets in slides.

2. Extend the algorithms and R functions for seasonality.

Corporate stocks : quarterly 1936.01 to 1978.04 1936.01 to 1970.04 as training, quarters 1971.01 to 1978.04 as holdout.

quarter	holdout	expsmo	holt	persist	iid
71.01	9.69	1.73	6.90	10.43	2.87
71.02	0.17	2.46	9.63	9.69	2.87
71.03	-0.59	2.25	6.13	0.17	2.87
71.04	4.66	1.99	3.29	-0.59	2.87
72.01	5.74	2.24	4.17	4.66	2.87
72.02	0.67	2.56	5.26	5.74	2.87
72.03	3.91	2.39	3.24	0.67	2.87
72.04	7.56	2.53	3.61	3.91	2.87
73.01	-4.89	2.99	5.77	7.56	2.87
73.02	-5.77	2.26	0.54	-4.89	2.87
73.03	4.81	1.52	-3.27	-5.77	2.87
73.04	-9.16	1.83	-0.01	4.81	2.87
...					
78.04					
rmse		10.0	12.0	13.7	9.7

Forecasting rules for Vancouver monthly total precipitation 1938–2005 as training set, 2006–2023 as holdout.

yearmon	holdout	addseasonal	multseasonal	persistmon	iidbymon
2006.01	283.6	171.16	190.35	249.6	151.10
2006.02	57.0	96.54	74.10	45.8	114.73
2006.03	92.4	117.55	120.01	132.8	103.88
2006.04	70.0	86.37	81.60	90.2	72.55
2006.05	42.8	67.83	65.61	68.6	57.44
2006.06	54.4	49.38	41.56	49.6	47.67
2006.07	25.2	39.88	36.11	43.6	33.96
2006.08	4.8	37.68	37.15	28.6	37.37
2006.09	39.4	56.03	69.28	53.6	57.45
2006.10	57.8	128.59	143.20	155.4	119.04
2006.11	350.8	162.26	161.34	136.6	158.98
2006.12	146.0	175.60	169.24	160.8	172.26
...					
2023.12					
rmse		45.3	47.6	64.1	44.1

Forecasting rules for monthly CPI in Canada, 1992.01–2018.07 as training set, 2018.08–2025.03 as holdout.

yearmon	holdout	holt	linear	persist	iid
2018.08	105.79	105.77	105.79	105.64	84.16
2018.09	105.72	105.72	105.94	105.79	84.16
2018.10	106.03	105.85	105.87	105.72	84.16
2018.11	105.87	106.17	106.18	106.03	84.16
2018.12	106.11	106.01	106.02	105.87	84.16
2019.01	106.11	106.25	106.26	106.11	84.16
2019.02	106.43	106.24	106.26	106.11	84.16
2019.03	106.82	106.56	106.58	106.43	84.16
2019.04	107.14	106.96	106.97	106.82	84.16
2019.05	107.45	107.28	107.29	107.14	84.16
2019.06	107.37	107.60	107.61	107.45	84.16
2019.07	107.77	107.51	107.53	107.37	84.16
...					
2025.03					
rmse		0.365	0.373	0.458	33.1

Code not provided. The time series data sets used above are at the course web site. Please try to reproduce the results with your own R functions.

For exponential smoothing methods and ARIMA methods, statistical software provide in-sample prediction rmse but not out-of-sample forecast rmse.

Functions can be written to compute out-of-sample forecast rmse making use of some estimated quantities outputted from statistical software.

Take-home message

Any good forecasting method should be better than (a) persistence and (b) average of all previous observations. Use the root mean square forecast error of these two rules as baseline checks for time series data sets used in this course.