Introduction to Time Series (I)

ZHANG RONG

Department of Social Networking Operations
Social Networking Group
Tencent Company

November 20, 2017

https://zhuanlan.zhihu.com/p/32584136

Outline

Outline

1. Time Series Algorithms
2. Control Chart Theory
3. Opprentice System
4. TSFRESH python package

Time Series

Definition and Methods

Definition of Time Series

A time series is a series of data points indexed in time order.
Methods for time series analysis may be divided into two classes:

- Frequency-domain methods: spectral analysis and wavelet analysis;
- Time-domain methods: auto-correlation and cross-correlation analysis.

Methods of Time Series

Methods for time series analysis may be divided into another two classes:

- Parametic methods
- Non-parametic methods

Moving Average

Moving Average

Let $\{x_i : i \geq 1\}$ be an observed data sequence. A simple moving average (SMA) is the unweighted mean of the previous w data. If the w-days' values are $x_i, x_{i-1}, \ldots, x_{i-(w-1)}$, then the formula is

$$M_i = \frac{1}{w}\sum_{j=0}^{w-1} x_{i-j} = \frac{x_i + x_{i-1} + \cdots + x_{i-(w-1)}}{w}.$$

When calculating successive values, a new value comes into the sum and an old value drops out, that means

$$M_i = M_{i-1} + \frac{x_i}{w} - \frac{x_{i-w}}{w}.$$

Moving Average

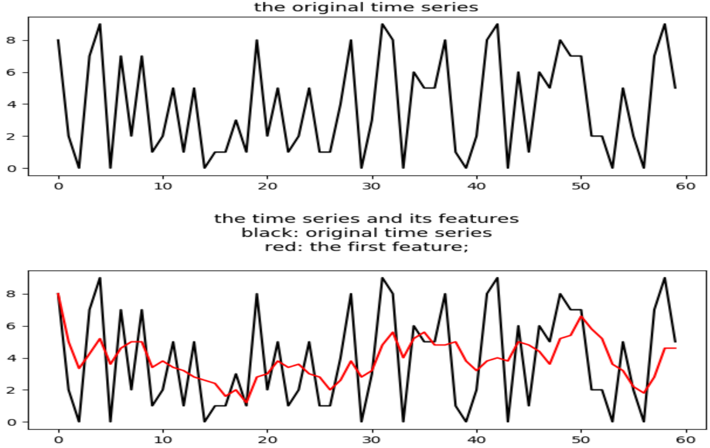


Figure: Moving Average Method for $w = 5$

Cumulative Moving Average

Cumulative Moving Average

Let $\{x_i : i \geq 1\}$ be an observed data sequence. A cumulative moving average is the unweighted mean of all datas. If the w-days values are $x_1, \cdots, x_i$, then

$$CMA_i = \frac{x_1 + \cdots + x_i}{i}.$$

If we have a new value x_{i+1}, then the cumulative moving average is

$$\begin{aligned} CMA_{i+1} &= \frac{x_1 + \cdots + x_i + x_{i+1}}{i+1} \\ &= \frac{x_{i+1} + \text{i}\,\not{n} \cdot CMA_i}{i+1} \\ &= CMA_i + \frac{x_{i+1} - CMA_i}{i+1}. \end{aligned}$$

Weighted Moving Average

Weighted Moving Average

A weighted moving average is the weighted mean of the previous w-datas. Suppose $\sum_{j=0}^{w-1} weight_j = 1$ with all $weight_j \geq 0$, then the weighted moving average is

$$WMA_i = \sum_{j=0}^{w-1} weight_j \cdot x_{i-j}.$$

Weighted Moving Average

A Special Case

In particular, let $\{weight_j : 0 \leq j \leq w-1\}$ be a weight with

$$weight_j = \frac{w-j}{w+(w-1)+\cdots+1} \text{ for } 0 \leq j \leq w-1.$$

In this situation,

$$WMA_i = \frac{wx_i + (w-1)x_{i-1} + \cdots + 2x_{i-w+2} + x_{i-w+1}}{w+(w-1)+\cdots+1}.$$

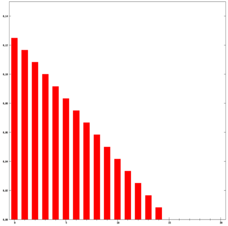


Figure: WMA weights $w = 15$

Weighted Moving Average

A Special Case

Weighted Moving Average

Suppose

$$\begin{aligned} Total_i &= x_i + \cdots + x_{i-w+1}, \\ Numerator_i &= wx_i + (w-1)x_{i-1} + \cdots + x_{i-w+1}, \end{aligned}$$

then the update formulas are

$$\begin{aligned} Total_{i+1} &= Total_i + x_{i+1} - x_{i-w+1}, \\ Numerator_{i+1} &= Numerator_i + wx_{i+1} - Total_i, \\ WMA_{i+1} &= \frac{Numerator_{i+1}}{w + (w-1) + \cdots + 1}. \end{aligned}$$

Exponential Weighted Moving Average

Exponential Weighted Moving Average

Suppose $\{Y_t : t \geq 1\}$ is an observed data sequence, the exponential weighted moving average series $\{S_t : t \geq 1\}$ is defined as

$$S_t = \begin{cases} Y_1, & t = 1 \\ \alpha \cdot Y_{t-1} + (1-\alpha) \cdot S_{t-1}, & t \geq 2 \end{cases}$$

- $\alpha \in [0, 1]$ is a constant smoothing factor.
- Y_t is the observed value at a time period t.
- S_t is the value of the EMWA at any time period t.

Exponential Weighted Moving Average

Moreover, from above definition,

$$S_t = \alpha[Y_{t-1} + (1-\alpha)Y_{t-2} + \cdots + (1-\alpha)^k Y_{t-(k+1)}] + (1-\alpha)^{k+1} S_{t-(k+1)}$$

for any suitable $k \in \{0, 1, 2, \cdots\}$. The weight of the point Y_{t-i} is $\alpha(1-\alpha)^{i-1}$.

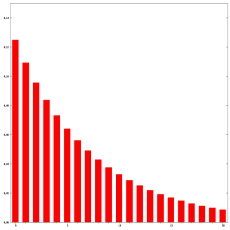


Figure: EMA weights $k = 20$

Exponential Weighted Moving Average

Exponential Weighted Moving Average

Suppose $\{Y_t : t \geq 1\}$ is an observed data sequence, the alternated exponential weighted moving average series $\{S_t : t \geq 1\}$ is defined as

$$S_{t,alternate} = \begin{cases} Y_1, & t = 1 \\ \alpha \cdot Y_t + (1 - \alpha) \cdot S_{t-1,alternate}, & t \geq 2 \end{cases}$$

Here, we use Y_t instead of Y_{t-1}.

Exponential Weighted Moving Average

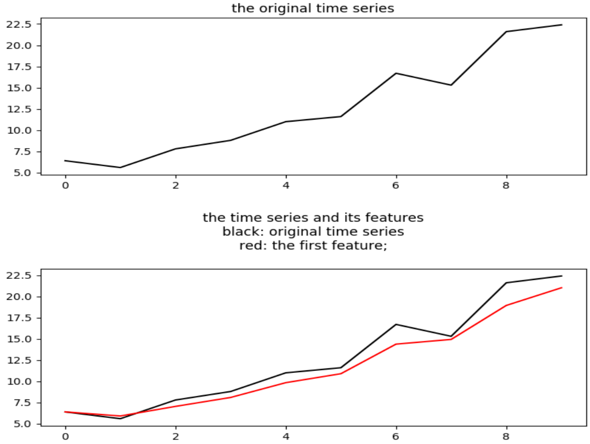


Figure: Exponential Weighted Moving Average Method for $\alpha = 0.6$

Double Exponential Smoothing

Double Exponential Smoothing

Suppose $\{Y_t : t \geq 1\}$ is an observed data sequence, there are two equations associated with double exponential smoothing:

$$\begin{aligned} S_t &= \alpha Y_t + (1-\alpha)(S_{t-1} + b_{t-1}), \\ b_t &= \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1}, \end{aligned}$$

where $\alpha \in [0, 1]$ is the data smoothing factor and $\beta \in [0, 1]$ is the trend smoothing factor.

Double Exponential Smoothing

Double Exponential Smoothing

Here, the initial values are $S_1 = Y_1$ and b_1 has three possibilities:

$$
\begin{aligned}
b_1 &= Y_2 - Y_1, \\
b_1 &= \frac{(Y_2 - Y_1) + (Y_3 - Y_2) + (Y_4 - Y_3)}{3} = \frac{Y_4 - Y_1}{3}, \\
b_1 &= \frac{Y_n - Y_1}{n - 1}.
\end{aligned}
$$

Forecast

- The one-period-ahead forecast is given by $F_{t+1} = S_t + b_t$.
- The *m*-period-ahead forecast is given by $F_{t+m} = S_t + mb_t$.

Double Exponential Smoothing

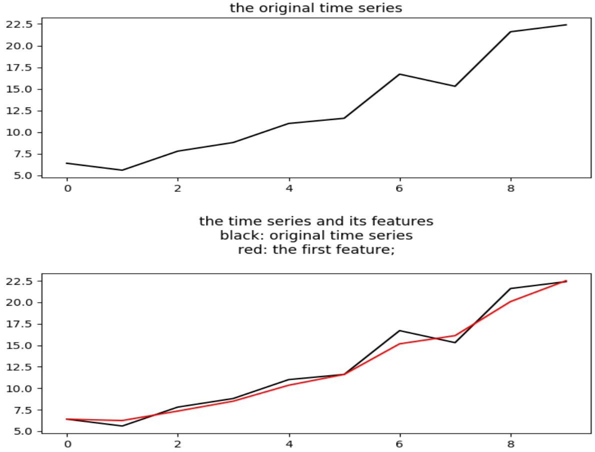


Figure: Double Exponential Smoothing for $\alpha = 0.6$ and $\beta = 0.4$

Triple Exponential Smoothing

Multiplicative Seasonality

Triple Exponential Smoothing (Multiplicative Seasonality)

Suppose $\{Y_t : t \geq 1\}$ is an observed data sequence, then the triple exponential smoothing is

$$
\begin{aligned}
S_t &= \alpha \frac{Y_t}{c_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}), \quad \text{Overall Smoothing} \\
b_t &= \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1}, \quad \text{Trend Smoothing} \\
c_t &= \gamma \frac{Y_t}{S_t} + (1-\gamma)c_{t-L}, \quad \text{Seasonal Smoothing}
\end{aligned}
$$

where $\alpha \in [0, 1]$ is the data smoothing factor, $\beta \in [0, 1]$ is the trend smoothing factor, $\gamma \in [0, 1]$ is the seasonal change smoothing factor.

Triple Exponential Smoothing

Multiplicative Seasonality

Forcast

The *m-period-ahead forecast* is given by
$F_{t+m} = (S_t + mb_t)c_{(t-L+m) \mod L}$.

Triple Exponential Smoothing

Initial values are

$$
\begin{aligned}
S_1 &= Y_1, \\
b_0 &= \frac{(Y_{L+1} - Y_1) + (Y_{L+2} - Y_2) + \cdots + (Y_{L+L} - Y_L)}{L}, \\
c_i &= \frac{1}{N} \sum_{j=1}^{N} \frac{Y_{L(j-1)+i}}{A_j}, \forall i \in \{1, \cdots, L\}, \\
A_j &= \frac{\sum_{i=1}^{L} Y_{L(j-1)+i}}{L}, \forall j \in \{1, \cdots, N\}.
\end{aligned}
$$

Triple Exponential Smoothing

Additive Seasonality

A.K.A. Holt-Winters

Triple Exponential Smoothing (Additive Seasonality)

Suppose $\{Y_t : t \geq 1\}$ is an observed data sequence, then the triple exponential smoothing is

$$\begin{aligned} S_t &= \alpha(Y_t - c_{t-L}) + (1-\alpha)(S_{t-1} + b_{t-1}), \text{ Overall Smoothing} \\ b_t &= \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1}, \text{ Trend Smoothing} \\ c_t &= \gamma(Y_t - S_{t-1} - b_{t-1}) + (1-\gamma)c_{t-L}, \text{ Seasonal Smoothing} \end{aligned}$$

where $\alpha \in [0,1]$ is the data smoothing factor, $\beta \in [0,1]$ is the trend smoothing factor, $\gamma \in [0,1]$ is the seasonal change smoothing factor.

The *m*-period-ahead forecast is given by
$F_{t+m} = S_t + mb_t + c_{(t-L+m) \mod L}$.

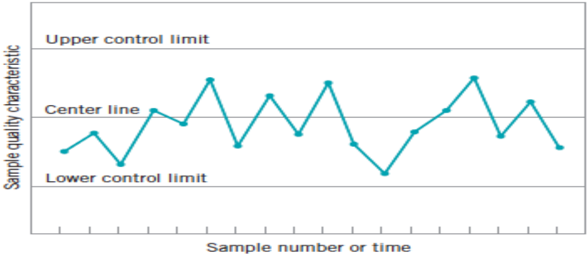
Outline

Definition of Control Chart Theory

Control Chart

The control chart is a graphical display of a quality characteristic that has been measured from a sample versus the sample number or time.

- Center Line: the average value of the quality characteristic
- Upper Control Limit (UCL) and Lower Control Limit (LCL): two horizontal lines.



3σ Control Chart

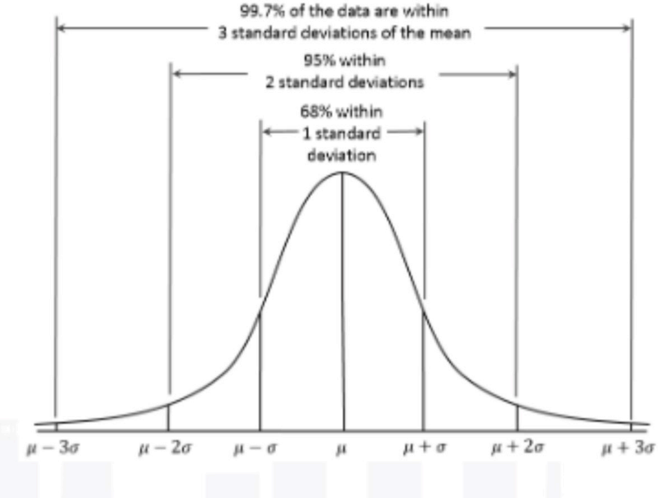
Simplest Control Chart

3σ Control Chart

Suppose that w is a sample statistic that measures some quality characteristic, the mean of w is μ_w and the standard deviation of w is σ_w. Then the center line, the upper control limit and the lower control limit becomes:

$$\begin{aligned} \text{UCL} &= \mu_w + L\sigma_w \\ \text{Center line} &= \mu_w \\ \text{LCL} &= \mu_w - L\sigma_w \end{aligned}$$

where L is the "distance" of the control limits from the center line, expressed in standard deviation units. In particular, if $L = 3$, then it is the 3σ control chart.


99.7% of the data are within
3 standard deviations of the mean
95% within
2 standard deviations
68% within
1 standard
deviation
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$

3σ Control Chart

Simplest Control Chart

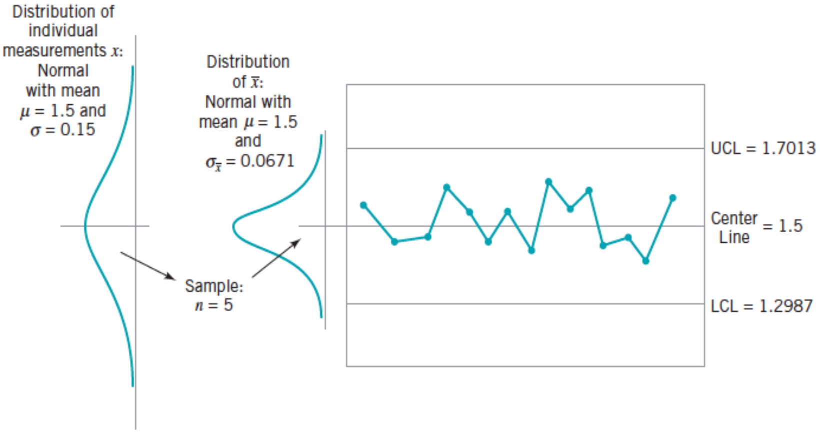


Figure: How the control chart works

The Cumulative Sum Control Chart

CUSUM Control Chart

Let x_i be the i-th observation on the process $\{x_i : 1 \leq i \leq n\}$, $\{x_i : 1 \leq i \leq n\}$ has a normal distribution with mean μ and standard deviation σ. The cumulative sum control chart is calculated by, for all $1 \leq i \leq n$,

$$C_i = \sum_{j=1}^{i} (x_j - \mu_0) = C_{i-1} + (x_i - \mu_0),$$

where $C_0 = 0$ and μ_0 is the target for the process mean.

- If $|C_i|$ exceed the decision interval H, then the process is considered to be out of control.
- The decision interval H is 3σ or 5σ.

The Cumulative Sum Control Chart

Data for the Cusum Example

Data for the Cusum Example

Sample, i	(a) x_i	(b) $x_i - 10$	(c) $C_i = (x_i - 10) + C_{i-1}$
1	9.45	−0.55	−0.55
2	7.99	−2.01	−2.56
3	9.29	−0.71	−3.27
4	11.66	1.66	−1.61
5	12.16	2.16	0.55
6	10.18	0.18	0.73
7	8.04	−1.96	−1.23
8	11.46	1.46	0.23
9	9.20	−0.80	−0.57
10	10.34	0.34	−0.23
11	9.03	−0.97	−1.20
12	11.47	1.47	0.27
13	10.51	0.51	0.78
14	9.40	−0.60	0.18
15	10.08	0.08	0.26
16	9.37	−0.63	−0.37
17	10.62	0.62	0.25
18	10.31	0.31	0.56
19	8.52	−1.48	−0.92
20	10.84	0.84	−0.08
21	10.90	0.90	0.82
22	9.33	−0.67	0.15
23	12.29	2.29	2.44
24	11.50	1.50	3.94
25	10.60	0.60	4.54
26	11.08	1.08	5.62
27	10.38	0.38	6.00
28	11.62	1.62	7.62
29	11.31	1.31	8.93
30	10.52	0.52	9.45

The Cumulative Sum Control Chart

The first 20 of these observations were drawn at random from a normal distribution with $\mu = 10$ and standard deviation $\sigma = 1$. They are plotted on a Shewhart control chart.

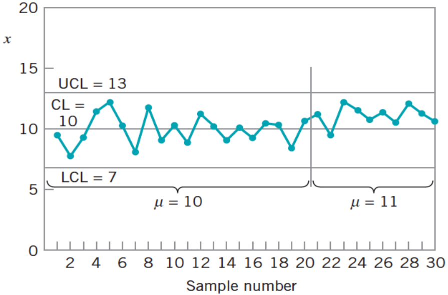


Figure: A Shewhart control chart for the data

The Cumulative Sum Control Chart

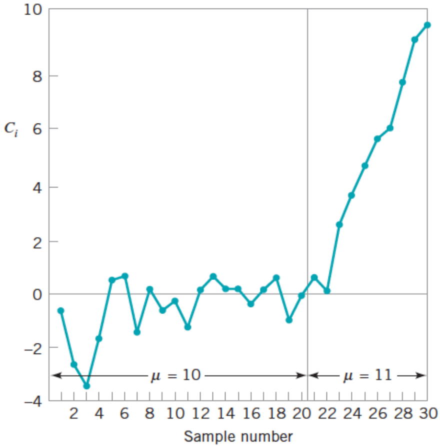


Figure: Plot of the cumulative sum from column (c) in above table

The Cumulative Sum Control Chart

Comparison to three-sigma control limit

Difference

- Three-sigma control limit: one or more points beyond a three-sigma control limit
- CUSUM control limit: it is a good choice when small shifts are important.

The Tabular or Algorithmic Cusum

Let x_i be the i-th observation on the process $\{x_i : 1 \leq i \leq n\}$, it has mean μ_0 and standard deviation σ. The statistics C^+ and C^- are computed as follows:

$$\begin{aligned} C_i^+ &= \max\left[0, x_i - (\mu_0 + K) + C_{i-1}^+\right] \\ C_i^- &= \max\left[0, (\mu_0 - K) - x_i + C_{i-1}^-\right] \end{aligned}$$

where $C_0^+ = C_0^- = 0$. K is the reference value, is calculated as

$$K = \frac{|\mu_1 - \mu_0|}{2}, \text{ where } \mu_1 = \mu_0 + \delta\sigma \text{ and } \delta = 1.$$

If either C_i^+ or C_i^- exceed the decision interval $H = 5\sigma$, the process is considered to be out of control. Here δ and H are parameters.

The Tabular or Algorithmic Cusum

		(a)			(b)		
Period i	x_i	$x_i - 10.5$	C_i^+	N^+	$9.5 - x_i$	C_i^-	N^-
1	9.45	−1.05	0	0	0.05	0.05	1
2	7.99	−2.51	0	0	1.51	1.56	2
3	9.29	−1.21	0	0	0.21	1.77	3
4	11.66	1.16	1.16	1	−2.16	0	0
5	12.16	1.66	2.82	2	−2.66	0	0
6	10.18	−0.32	2.50	3	−0.68	0	0
7	8.04	−2.46	0.04	4	1.46	1.46	1
8	11.46	0.96	1.00	5	−1.96	0	0
9	9.20	−1.3	0	0	0.30	0.30	1
10	10.34	−0.16	0	0	−0.84	0	0
11	9.03	−1.47	0	0	0.47	0.47	1
12	11.47	0.97	0.97	1	−1.97	0	0
13	10.51	0.01	0.98	2	−1.01	0	0
14	9.40	−1.10	0	0	0.10	0.10	1
15	10.08	−0.42	0	0	−0.58	0	0
16	9.37	−1.13	0	0	0.13	0.13	1
17	10.62	0.12	0.12	1	−1.12	0	0
18	10.31	−0.19	0	0	−0.81	0	0
19	8.52	−1.98	0	0	0.98	0.98	1
20	10.84	0.34	0.34	1	−1.34	0	0
21	10.90	0.40	0.74	2	−1.40	0	0
22	9.33	−1.17	0	0	0.17	0.17	1
23	12.29	1.79	1.79	1	−2.79	0	0
24	11.50	1.00	2.79	2	−2.00	0	0
25	10.60	0.10	2.89	3	−1.10	0	0
26	11.08	0.58	3.47	4	−1.58	0	0
27	10.38	−0.12	3.35	5	−0.88	0	0
28	11.62	1.12	4.47	6	−2.12	0	0
29	11.31	0.81	5.28	7	−1.81	0	0
30	10.52	0.02	5.30	8	−1.02	0	0

The Tabular or Algorithmic Cusum

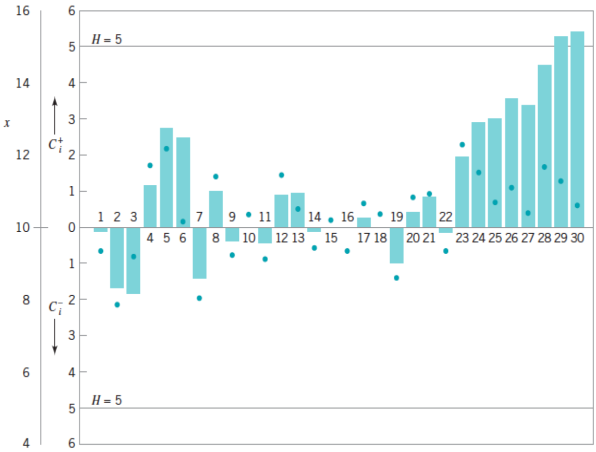


Figure: CUSUM status charts for the above example

EWMA Control Chart

Exponentially Weighted Moving Average Control Chart

EWMA

The exponentially weighted moving average is defined as

$$z_i = \lambda x_i + (1-\lambda) z_{i-1},$$

where $0 \leq \lambda \leq 1$ is a constant and $z_0 = \mu_0$. Here, μ_0 is the process target. In particular, $z_0 = \overline{x}$.

Moreover,

$$z_i = \lambda \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^i z_0,$$

and the sum of their weights is one. That means

$$\lambda \sum_{j=0}^{i-1} (1-\lambda)^j + (1-\lambda)^i = 1.$$

EWMA Control Chart

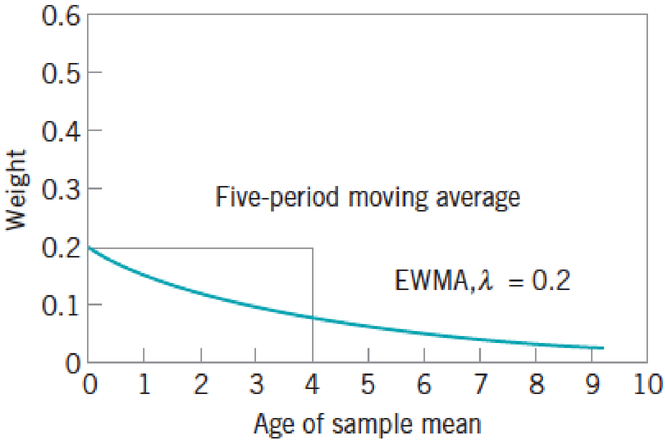


Figure: Weights of past sample means

EWMA Control Chart

Exponentially Weighted Moving Average Control Chart

EWMA Control Chart

If the observations x_i are independent random variables with variance σ^2, then the variance of z_i is

$$\sigma_{z_i}^2 = \sigma^2 \left(1 - (1-\lambda)^{2i}\right)[1 - (1-\lambda)^{2i}].$$

The EWMA control chart is

$$\begin{aligned} UCL &= \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2i}]}, \\ \text{Center Line} &= \mu_0, \\ LCL &= \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2i}]}. \end{aligned}$$

EWMA Control Chart

Exponentially Weighted Moving Average Control Chart

EWMA Control Chart

Note that the term $[1-(1-\lambda)^{2i}]$ tends to 1 as i tends to ∞, then the simplified EWMA control chart is

$$\begin{aligned} UCL &= \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}}, \\ \text{Center Line} &= \mu_0, \\ LCL &= \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}}. \end{aligned}$$

EWMA Control Chart

EWMA Calculations for Example 9.2

Subgroup, i	* = Beyond Limits x_i	EWMA, z_i	Subgroup, i	* = Beyond Limits x_i	EWMA, z_i
1	9.45	9.945	16	9.37	9.98426
2	7.99	9.7495	17	10.62	10.0478
3	9.29	9.70355	18	10.31	10.074
4	11.66	9.8992	19	8.52	9.91864
5	12.16	10.1253	20	10.84	10.0108
6	10.18	10.1307	21	10.9	10.0997
7	8.04	9.92167	22	9.33	10.0227
8	11.46	10.0755	23	12.29	10.2495
9	9.2	9.98796	24	11.5	10.3745
10	10.34	10.0232	25	10.6	10.3971
11	9.03	9.92384	26	11.08	10.4654
12	11.47	10.0785	27	10.38	10.4568
13	10.51	10.1216	28	11.62	10.5731
14	9.4	10.0495	29	11.31	10.6468*
15	10.08	10.0525	30	10.52	10.6341*

EWMA Control Chart

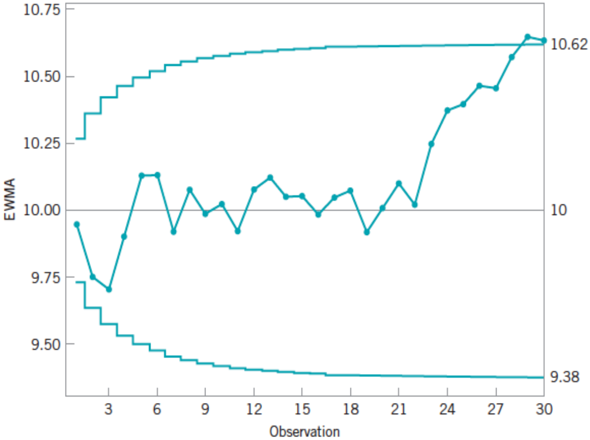


Figure: The EWMA control chart for the above example

Moving Average Control Chart

Moving Average Control Chart

Suppose that the observations are $\{x_1, \cdots, x_n\}$, then the moving average of span w at time i is defined as

$$M_i = \frac{x_i + x_{i-1} + \cdots + x_{i-w+1}}{w}.$$

The variance of the moving average M_i is

$$V(M_i) = \frac{1}{w^2} \sum_{j=i-w+1}^{i} V(x_j) = \frac{1}{w^2} \sum_{j=i-w+1}^{i} \sigma^2 = \frac{\sigma^2}{w}.$$

Moving Average Control Chart

Moving Average Control Chart

Let μ_0 be the target value and the moving average control chart for M_i are

$$\begin{aligned} UCL &= \mu_0 + \frac{3\sigma}{\sqrt{w}}, \\ \text{Center Line} &= \mu_0, \\ LCL &= \mu_0 - \frac{3\sigma}{\sqrt{w}}. \end{aligned}$$

The control procedure would consist of calculating the new moving average M_i as each observation x_i becomes available, plotting M_i on a control chart with upper and lower control limits, and concluding that the process is out of control if M_i exceeds the limits.

Moving Average Control Chart

Moving Average Chart for Example 9.3

Observation, i	x_i	M_i	Observation, i	x_i	M_i
1	9.45	9.45	16	9.37	10.166
2	7.99	8.72	17	10.62	9.996
3	9.29	8.91	18	10.31	9.956
4	11.66	9.5975	19	8.52	9.78
5	12.16	10.11	20	10.84	9.932
6	10.18	10.256	21	10.9	10.238
7	8.04	10.266	22	9.33	9.98
8	11.46	10.7	23	12.29	10.376
9	9.2	10.208	24	11.5	10.972
10	10.34	9.844	25	10.6	10.924
11	9.03	9.614	26	11.08	10.96
12	11.47	10.3	27	10.38	11.17
13	10.51	10.11	28	11.62	11.036
14	9.4	10.15	29	11.31	10.998
15	10.08	10.098	30	10.52	10.982

Moving Average Control Chart

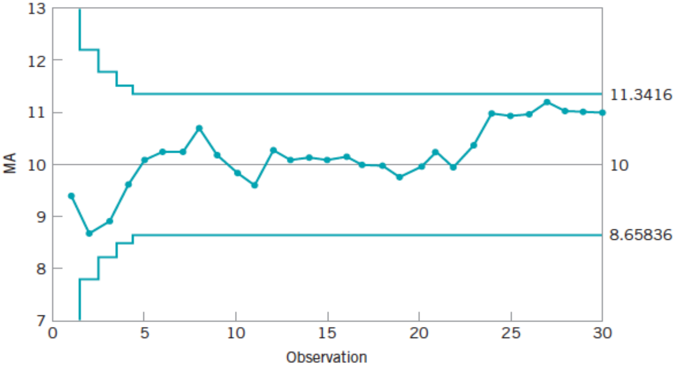


Figure: Moving average control chart with $w = 5$

Multivariate Data Control Chart

The Multivariate Process Monitoring and Control

Multivariate Data Control Chart

- Hotelling T^2 Control Chart
- The Multivariate EWMA Control Chart
- Regression Adjustment
- Principal Components Method
- Partial Least Squares

Outline

TSFRESH python package

TSFRESH python package

- tsfresh is used to to extract characteristics from time series.
- Paper: Time Series Feature extraction based on scalable hypothesis tests
- Spend less time on feature engineering
- Automatic extraction of 100s of features

TSFRESH python package

TSFRESH python package

Let $\{x_1, \cdots, x_n\}$ be a time series, some features are

- max, min, median, mean μ, variance σ^2, standard deviation σ,
- range is maximum minus minimum
- skewness is the third standardized moment:

$$skewness = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^3,$$

- kurtosis is the fourth standardized moment:

$$kurtosis = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^4.$$

TSFRESH python package

TSFRESH python package

Let $\{x_1, \cdots, x_n\}$ be a time series, some features are

- absolute energy: $E = \sum_{i=1}^{n} x_i^2$,
- absolute sum of changes: $E = \sum_{i=1}^{n-1} |x_{i+1} - x_i|$,
- aggregate autocorrelation:

$$\frac{1}{n-1}\sum_{\ell=1}^{n}\frac{1}{(n-\ell)\sigma^2}\sum_{t=1}^{n-\ell}(x_t-\mu)(x_{t+\ell}-\mu),$$

- autocorrelation: parameter is lag ℓ,

$$\frac{1}{(n-\ell)\sigma^2}\sum_{t=1}^{n-\ell}(x_t-\mu)(x_{t+\ell}-\mu).$$

TSFRESH python package

TSFRESH python package

Let $\{x_1, \cdots, x_n\}$ be a time series, some features are

- count above mean, count below mean
- variance larger than standard deviation
- first location of maximum, first location of minimum
- last location of maximum, last location of minimum
- has duplicate, has duplicate max, has duplicate min
- longest strike above mean, longest strike below mean

TSFRESH python package

TSFRESH python package

Let $\{x_1, \cdots, x_n\}$ be a time series, some features are

- mean change: $\sum_{i=1}^{n-1}(x_{i+1} - x_i)/n = (x_n - x_1)/n$
- mean second derivative central:

$$\frac{1}{n}\sum_{i=1}^{n-2}\frac{1}{2}(x_{i+2} - 2 \cdot x_{i+1} + x_i)$$

- percentage of reoccurring data points to all data points
- percentage of reoccurring values to all values
- ratio value number to time series length
- sum of reoccurring data points
- sum of reoccurring values

TSFRESH python package

Initialization of Time Series

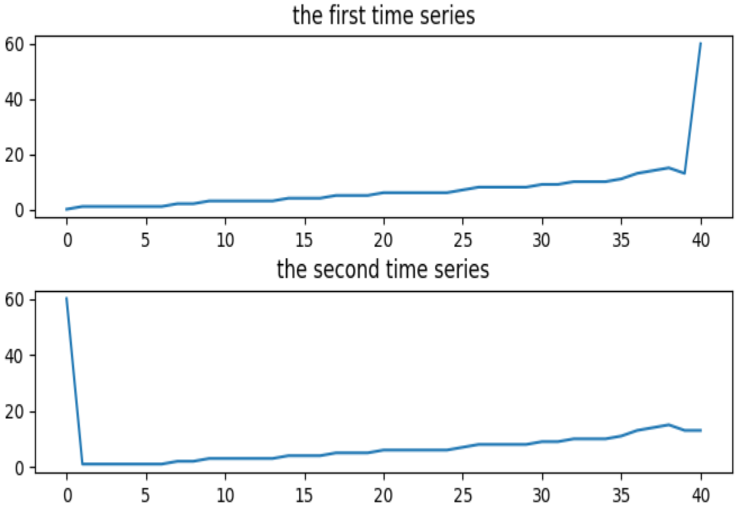
Let $\{x_1, \cdots, x_n\}$ be a time series, some initialization methods are, for $1 \leq i \leq n$,

$$\begin{aligned}
y_i &= \frac{x_i}{mean(\{x_i : 1 \leq i \leq n\})}, \\
y_i &= \frac{x_i}{median(\{x_i : 1 \leq i \leq n\})}, \\
y_i &= \frac{x_i}{\max - \min}, \\
y_i &= \frac{x_i}{(\max - \min)/10},
\end{aligned}$$

where max and min denotes the maximum and minimum value of the time series, respectively.

TSFRESH python package

Example of Two Lists



TSFRESH python package

Features of the Above Two Lists

nonParametersFeatures	th	value_list1	value_list2
feature	0	60	60
feature	1	0	1
feature	2	7.19512195122	7.51219512195
feature	3	85.0350981559	84.493753718
feature	4	9.22144772559	9.1920483962
feature	5	4.71450748799	4.67091571882
feature	6	26.5796091617	26.2452662595
feature	7	6.0	6.0
feature	8	5609	5778
feature	9	64	75
feature	10	1	1
feature	11	15	16
feature	12	26	25
feature	13	0.975609756098	0.0
feature	14	0.0	0.0243902439024
feature	15	1.0	0.0243902439024
feature	16	0.0243902439024	0.170731707317
feature	17	True	True
feature	18	False	False
feature	19	False	True
feature	20	15	15
feature	21	26	25
feature	22	1.6	1.875
feature	23	1.5	-1.175
feature	24	0.589743589744	0.756410256641
feature	25	0.625	0.666666666667
feature	26	0.853658536585	0.878048780488
feature	27	0.39024390243902444	0.36585365853658536
feature	28	188	201
feature	29	61	61
feature	30	295	308
feature	31	60	59

Thank you for watching!

https://zhuanlan.zhihu.com/p/32584136

ZHANG RONG
zr9558@gmail.com
zr9558.wordpress.com