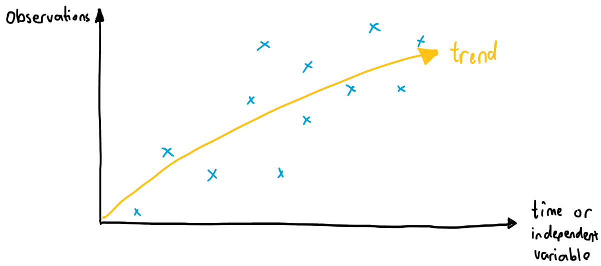
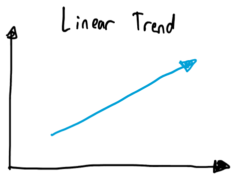
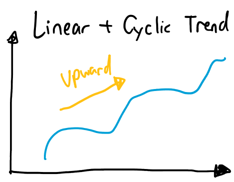
Forecasting

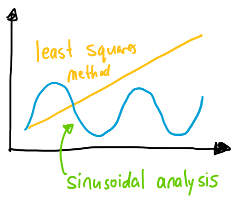
Tuesday, June 11, 2024 5:39 PM

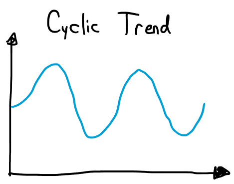
Pre-forecasting Analysis

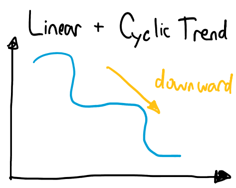












Long-range typically: 2 years / 5 years → & beyond

Medium-range: ~ 6 months → 2 years / 5 years

Short-range: a few hours → ~ 6 months

- There is no sharp boundaries between long, medium, & short horizons

  - Long-range is difficult to forecast!
    - It is often related to government policies

  - Medium-range forecasting:
    - A lot of companies find this method suitable

Medium range forecasting:

- A lot of companies find this method suitable for their needs

- Short-range forecasting:
  - Deals, such matters as: Scheduling operations, human resources planning, overtime shifts, etc.

Mathematical Forecasting Methods

- Linear Regression
- Non-linear regression → Cyclic, Second order forecasting, etc.

- Simple methods
  - linear regression
  - Simple moving average
  - Weighted moving average
  - Exponential Smoothing

Everything will be covered

- Linear Regression (Least Squares Method)

$$\underset{\text{error}}{e} = \underset{\text{observation}}{y_i} - \underset{\text{forecast}}{\hat{y}}$$

$$\underset{\text{sum}}{S} = \sum_{y=1}^{n} (y_i - \hat{y}_i)^2$$

minimize

$$S = (y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + \dots + (y_n - \hat{y}_n)^2$$

$$S = y_1^2 + \hat{y}_1^2 - 2y_1\hat{y}_1 + y_2^2 + \hat{y}_2^2 - 2y_2\hat{y}_2 + \dots + y_n^2 + \hat{y}_n^2 - 2y_n\hat{y}_n$$

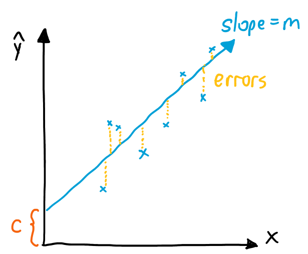
- We remove the constant terms:

$$S = \hat{y}_1^2 - 2y_1\hat{y}_1 + \hat{y}_2^2 - 2y_2\hat{y}_2 + \dots + \hat{y}_n^2 - 2y_n\hat{y}_n$$

- linear regression

- <u>linear regression</u>

$$\hat{y}_i = m x_i + c$$



- Substitute $\hat{y}_i = m x_i + c$ in S:

$$S = (mx_1 + c)^2 - 2y_1(mx_1 + c) + (mx_2 + c)^2 - 2y_2(mx_2 + c) + \ldots + (mx_n + c)^2 - 2y_n(mx_n + c)$$

$$\frac{dS}{dm} = 0 \qquad \frac{dS}{dc} = 0$$

We are trying to find best values of m and c

$$\frac{dS}{dm} = m\left(\sum_{i=1}^{n} x_i^2 + c\left(\sum_{i=1}^{n} x_i\right)\right) - \sum_{i=1}^{n} x_i y_i = 0$$

$$\frac{dS}{dc} = nc + m\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i = 0$$

We solve these equations simultaneously

independent variable

observation

$$m = \frac{n\sum(xy) - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \underbrace{\frac{\sum(xy) - n\bar{x}\bar{y}}{\sum x^x - n(\bar{x})^2}}_{\text{all data known}}$$

$$\underbrace{c}_{\text{all data known}} = \frac{\sum y - m\sum x}{n} = \bar{y} - m\bar{x}$$

$$\bar{x} = \frac{\sum x}{n} \qquad \bar{y} = \frac{\sum y}{n}$$

- <u>simple moving average</u>
    - We select "n" past observation
    - We divide sum by "n" the results are the weight associated with each observation
    - Say $n=4$ $(\underbrace{O_1, O_2, O_3, O_4}_{\text{observations}})$

$$\underset{\text{forecast}}{F_0} = \frac{O_1 + O_2 + O_3 + O_4}{4} = 0.25\,O_1 + 0.25\,O_2 + 0.25\,O_3 + 0.25\,O_4$$

These are called "weights"

<u>Observations</u> <u>Forecast</u>

Observations	Forecast
O_1	
O_2	
O_3	
O_4	$F_0 = \frac{O_1 + O_2 + O_3 + O_4}{4}$
O_5	$F_1 = \frac{O_2 + O_3 + O_4 + O_5}{4}$
O_6	$F_2 = \frac{O_3 + O_4 + O_5 + O_6}{4}$

- Weighted Moving Average
  - It makes sense to give higher weight to the most recent observation and while moving in distant time, we give less and less for the past data
  - Say $n=4$ as the previous example

$$\sum_{i=1}^{4} w_i = 100\%$$

Previous example:
$\sum w = 0.25 + 0.25 + 0.25 + 0.25 = 1$
$= 100\%$

weights

$$F_0 = 0.6\,O_4 + 0.2\,O_3 + 0.15\,O_2 + 0.05\,O_1$$

$$\sum w_i = 0.6 + 0.2 + 0.15 + 0.05 = 100\%$$

Observations	Weight	Forecast
O_1	0.05	
O_2	0.15	
O_3	0.20	
O_4	0.60	$F_0 = 0.6O_4 + 0.2O_3 + 0.15O_2 + 0.05O_1$
O_5		$F_1 = 0.6O_5 + 0.2O_4 + 0.15O_3 + 0.05O_2$
O_6		$F_2 = 0.6O_6 + 0.2O_5 + 0.15O_4 + 0.05O_3$

- General forecasting idea can also be represented mathematically as:

New forecast = Old forecast + allowance for errors

- For now, forget about where "old forecast" comes from
- One representation is :

Smoothing constant

$$\text{New forecast} = \text{Old forecast} + \alpha \begin{pmatrix} \text{latest observation} \\ -\text{old forecast} \end{pmatrix}$$

$\alpha = 0 \longrightarrow$ ultimate in Stability

$\alpha = 1 \longrightarrow$ ultimate in Sensitivity

$\alpha = 0 \longrightarrow$ New forecast = Old forecast

$\alpha = 1 \longrightarrow$ New forecast = Observation

- Neither make sense. Seems $\alpha = 0.5$ (mid-way) works best!
  But practice & experiments show that best α is between $\boxed{0.1 - 0.2}$. Depending on the case, α could be 0.08 or 0.3.

- <u>Exponential Smoothing</u>

  Let F_0 be the new forecast for the next period

  F_1 be the forecast made 1 period ago

  F_2 be the forecast made 2 periods ago

  etc ...

  Similarly, let:

  O_1 be the latest observation for the present period

  O_2 be the previous observation for 1 period ago

  O_3 be the previous observation for 2 periods ago

  etc ...

  Then :

$$F_0 = F_1 + \alpha (O_1 - F_1) \quad (1)$$

Rewrite:

$$F_0 = \alpha O_1 + (1-\alpha) F_1 \qquad (2)$$

But, F_1 itself was determined in the same way as F_0:

$$F_1 = \alpha O_2 + (1-\alpha) F_2 \qquad (3)$$

Sub (3) into (2)

$$F_0 = \alpha O_1 + (1-\alpha)\{\alpha O_2 + (1-\alpha) F_2\} \qquad (4a)$$

Or:

$$F_0 = \alpha O_1 + \alpha(1-\alpha) O_2 + (1-\alpha)^2 F_2 \qquad (4b)$$

But the same argument applies to F_2 as it did for F_1

$$F_2 = \alpha O_3 + (1-\alpha) F_3 \qquad (5)$$

Sub (5) into (4b)

$$F_0 = \alpha O_1 + \alpha(1-\alpha) O_2 + (1-\alpha)^2\{\alpha O_3 + (1-\alpha) F_3\} \qquad (6)$$

Again and again indefinitely ...

$$F_0 = \alpha\{O_1 + (1-\alpha) O_2 + (1-\alpha)^2 O_3 + (1-\alpha)^3 O_4 + \dots + (1-\alpha)^k O_{k+1}\} + \underbrace{(1-\alpha)^n F_n}_{0 \text{ as } n \to \infty} \qquad (7)$$

But equation (1) is the same as equation (7)! Which one would you use?

$F_0 =$

needs only:

- F_1 old forecast
- O_1 latest observation
- α smoothing constant

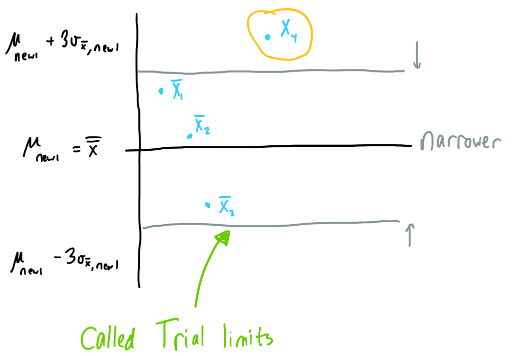
- In equation (7) we see individual terms where the weights decrease by a factor $(1-\alpha)$, hence the term "exponential smoothing"

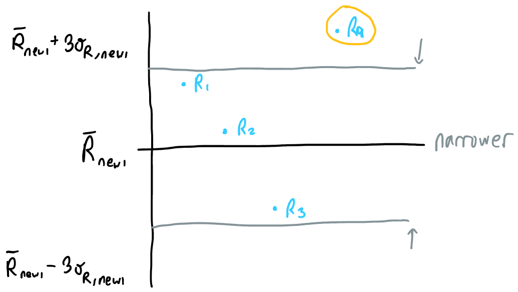
- Similar to weighted moving average:

$$\text{New forecast} = O_1 w_1 + O_2 w_2 + O_3 w_3 + \dots + O_n w_n$$

Observation   Weight

$$\sum_{i=1}^{n} w_i = 1$$





- We recalculate the μ as follows

$$\mu_{new2} = \frac{\sum_{i=1}^{25} \bar{X}_i - \bar{X}_7 - \bar{X}_1 - \bar{X}_4}{22}$$

- Recalculate the new $\bar{R}$

$$\bar{R}_{new2} = \frac{\sum_{i=1}^{25} R_i - R_7 - R_{10} - R_{16} - R_4}{21}$$

- Recalculate σ

$$\sigma_{new2} = \frac{\bar{R}_{new2}}{d_2} \quad \text{— same as before}$$

$$\sigma_{\bar{x},new2} = \frac{\sigma_{new2}}{\sqrt{n}}$$

$$\sigma_{\bar{R},new2} = d_3\, \sigma_{new2}$$

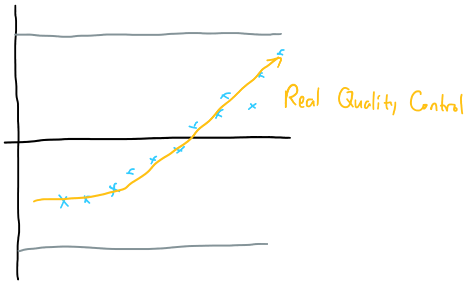
The initial charts are called Trial Limits

- If in the second round all data fall inside the limits for both charts, we call the charts Final Charts
- If again some data points fall outside the limits we conclude that the "factors of production" have some issues, we try to find the problem area(s), and fix them. Take 25 new samples from the production line and repeat the process (method).
- We cannot "blame" the operator if other factors of production are not suitable. The issue(s) called:

time and repeat the process (method).

- We cannot "blame" the operator if other factors of production are not suitable. The issue(s) called: "an assignable cause of variation"

(pp. 85) $\bar{X}$ Chart



Points show a trend: future points will likely fall outside a limit