BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Introduction to Statistical Methods

ISM Team

BITS Pilani
Pilani Campus

Session 13

Time Series/ Forecasting

(Session 13: 1st / 2nd March 2025)

IMP Note to Self

Time Series

Time Series → Holt Double Exponential Smoothing Technique

- The simple exponential smoothing forecast is a weighted sum of actual value with **α** and the forecast for the same period with (1-**α**). At **α** =1, it is identical to the naïve forecast.
- However the simple exponential smoothing forecasts lag the actual time series values when the time series has a trend.
- The trend makes it difficult for forecast to 'keep up' since there is no account of changes in level from period to period.
- The Holt double exponential smoothing method adds the ability to account for the trend component the time series in addition to the level.

Time Series ➔ Holt Double Exponential Smoothing Technique

- Holt double exponential is a simple exponential smoothing with a added trend components that helps fix the lag issues
- Weights are used to smooth the level and trend.
- Holt smoothed value calculation for the level is exactly same calculation as in simple exponential smoothing
- The trend component is a weighted difference between the current smoothed level and the prior period smoothed level plus prior trend.
- Finally, the Holt double exponential smoothing forecast for the next period is a sum of smoothed level and smoothed trend of the current period.

Time Series → Holt Double Exponential Smoothing Technique

Y_t
Current actual level
α

+

$L_t + T_{t-1}$
Prior forecast level
$1-\alpha$

→

Current smooth level
L_t

+

→

Forecast value

$F_{t+1}=L_t+T_t$

$L_t - L_{t-1}$
Current smoothed level – Prior smoothed level
β

+

T_{t-1}
Prior smoothed trend
$1-\beta$

→

Current smooth trend
T_t

innovate achieve lead

Time Series → Holt Double Exponential Smoothing Technique

Current smoothed level: $L_t = \boldsymbol{\alpha} Y_t + (1 - \boldsymbol{\alpha})(L_{t-1} + T_{t-1})$

Current smoothed Trend: $T_t = \boldsymbol{\beta}(L_t - L_{t-1}) + (1 - \boldsymbol{\beta})\, T_{t-1}$

Forecast next period: $F_{t+1} = L_t + T_t$

Forecast next k period: $F_{t+\mathrm{k}} = L_t + kT_t$

Level smoothing factor: $0 \le \boldsymbol{\alpha} \le 1$

Level smoothing factor: $0 \le \boldsymbol{\beta} \le 1$

Current smoothed level: $L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$

Level smoothing constant

Current data value

Prior period smoothed level

Prior period smoothed trend

Simple exponential smoothing

Time Series ⇨ Holt Double Exponential Smoothing Technique

When $\boldsymbol{\alpha} = 0$ ⇨ $$L_t = (\mathbf{0})Y_t + (1 - (\mathbf{0}))(L_{t-1} + T_{t-1})$$

$$L_t = L_{t-1} + T_{t-1}$$

$$L_t = F_t$$

Smoothed level = Current forecast

Time Series ⇨ Holt Double Exponential Smoothing Technique

When $\alpha = 1$ ⇨ $L_t = (\mathbf{1})Y_t + (1 - (\mathbf{1}))(L_{t-1} + T_{t-1})$

$$L_t = Y_t$$

Smoothed level = Current actual level of the data

Current smoothed trend: $T_t = \boldsymbol{\beta}(L_t - L_{t-1}) + (1 - \boldsymbol{\beta})T_{t-1}$

Trend smoothing constant

Current smoothed level

Prior period smoothed level

Prior period smoothed trend

Time Series ⇒ Holt Double Exponential Smoothing Technique

When $\boldsymbol{\beta}= 0$ ⇒

$$T_t = \boldsymbol{\beta}(L_t - L_{t-1}) + (1 - \boldsymbol{\beta})T_{t-1}$$

$$T_t = (\mathbf{0})(L_t - L_{t-1}) + (1 - (0))T_{t-1}$$

$$T_t = T_{t-1}$$

Smoothed trend = Trend from the last period

innovate achieve lead

Time Series ⇨ Holt Double Exponential Smoothing Technique

When $\boldsymbol{\beta}$= 1 ⇨ $$T_t = \boldsymbol{\beta}(L_t - L_{t-1}) + (1 - \boldsymbol{\beta})T_{t-1}$$

$$T_t = (1)(L_t - L_{t-1}) + (1 - (\mathbf{1}))T_{t-1}$$

$$T_t = L_t - L_{t-1}$$

Current smoothed trend = Current smoothed level – prior smoothed level

Time Series ⇨ Holt Double Exponential Smoothing Technique

Level smoothing constant (α)

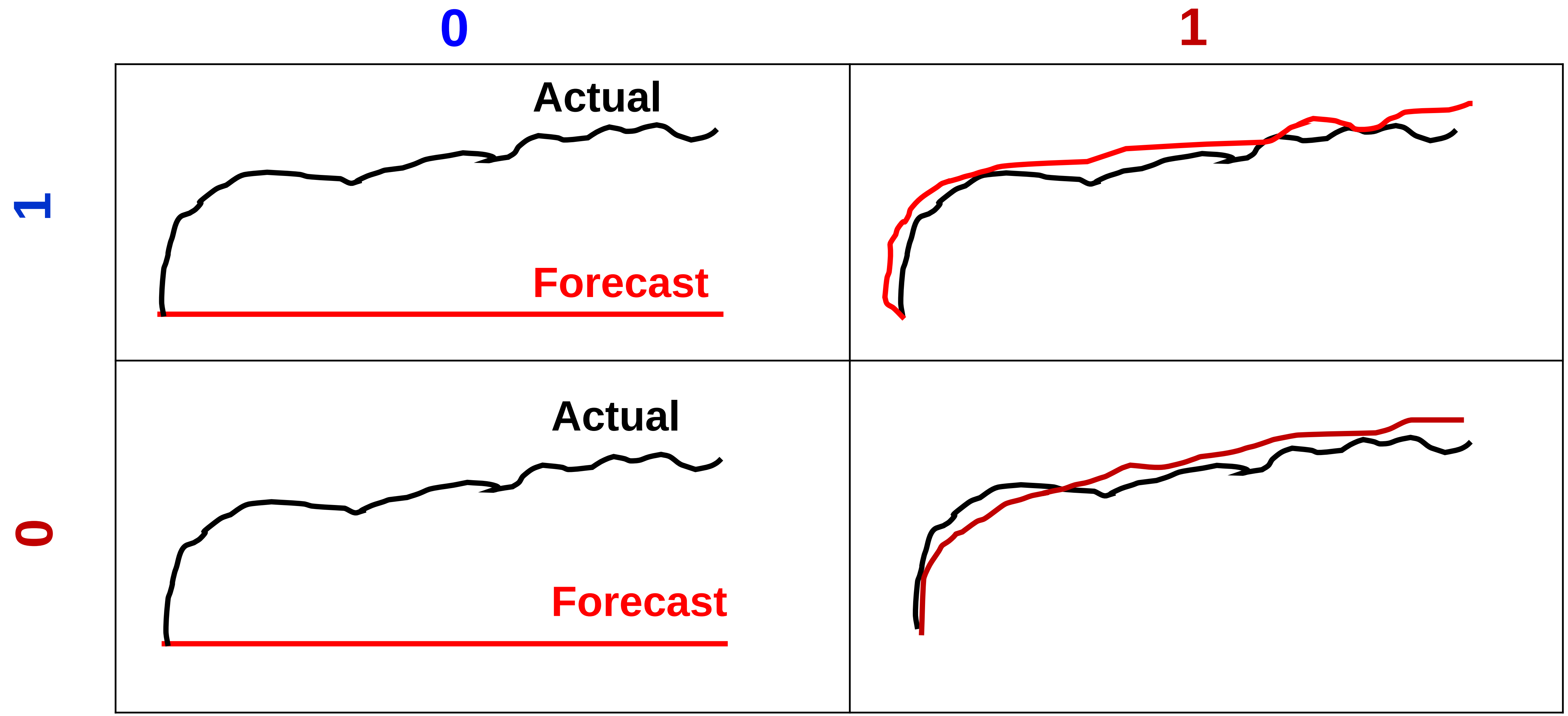
Trend smoothing constant (β)

β \ α	0	1
1	Forecast all the same (flat line) of Y_t	Smoothed level is the same as actual level. Trend is the full difference between smoothed level
0	Forecast all the same (flat line) of Y_t	Same as Naïve fore cast, no trend component

Time Series → Holt Double Exponential Smoothing Technique

Level smoothing constant (α)

Trend smoothing constant (β)



Time Series ➡ Holt Double Exponential Smoothing Technique

Example: Year & GDP

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
GDP	1451	1499	1686	1764	1879	1948	2013	2090	2173	2286	2404
Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	
GDP	2465	2501	2597	2689	2688	2576	2530	2513	2503	2396	

Year	GDP	$\alpha=0.8$	$\beta=0.4$	
		$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$	$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$	$F_{t+1} = L_t + T_t$
2000	1451	1451	0	
2001	1499	1489	15	1451
2002	1686	1650	73	1505
2003	1764	1756	86	1723
2004	1879	1872	98	1842
2005	1948	1952	91	1970
2006	2013	2019	81	2044
2007	2090	2092	78	2101
2008	2173	2172	79	2170
2009	2286	2279	90	2251
2010	2404	2397	101	2369
2011	2465	2472	91	2498
2012	2501	2513	71	2562
2013	2597	2594	75	2584
2014	2689	2685	81	2670
2015	2688	2704	56	2766
2016	2576	2613	-3	2760
2017	2530	2546	-28	2610
2018	2513	2514	-30	2518
2019	2503	2499	-24	2484
2020	2396	2412	-49	2475
			Forecast for the year 2021	**2363**

Time Series ⇨ Holts-Winters Triple Exponential Smoothing

Steps in calculation

M: No. of periods/ year (Quarterly (M=4), Monthly (M=12))

Initial seasonal factors:

$$S_1 = \frac{Y_1}{Average\,(Y_1, Y_2, \dots, Y_M)}, \quad S_2 = \frac{Y_2}{Average\,(Y_1, Y_2, \dots, Y_M)},$$

. . .

$$S_M = \frac{Y_M}{Average\,(Y_1, Y_2, \dots, Y_M)}$$

Time Series ➡ Holts-Winters Triple Exponential Smoothing

Steps in calculation

M: No. of periods/ year (Quarterly (M=4), Monthly (M=12))

Initial level:

$$L_{13} = \frac{Y_{13}}{S_1}$$

Initial trend:

$$T_{13} = \frac{Y_{13}}{S_1} - \frac{Y_{12}}{S_{12}}$$

innovate achieve lead

Time Series → Holts-Winters Triple Exponential Smoothing

Level formula

$$L_t = \alpha \frac{Y_t}{S_{t-M}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

Trend formula

$$T_t = \beta(L_t + L_{t-1}) + (1-\beta)T_{t-1}$$

Seasonal factor formula

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-M}$$

Forecast

$$F_{t+1} = (L_t + T_t)S_{t-M+1}$$

Time Series ➡ Holts-Winters Triple Exponential Smoothing

Example on demand data

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Year	2011	2011	2011	2011	2012	2012	2012	2012	2013	2013	2013	2013	2014	2014
Quarter	1	2	3	4	1	2	3	4	1	2	3	4	1	2
Demand	362	385	432	341	382	409	498	387	473	513	582	474	544	582
Period	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Year	2014	2014	2015	2015	2015	2015	2016	2016	2016	2016				
Quarter	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Demand	681	557	628	707	773	592	627	725	854	661				

Period	Year	Quarter	Demand	Level	Trend	Seasonal	Forecast	|Error|	α	β	γ
1	2011	1	362			0.952632			0.2	0.3	0.15
2	2011	2	385			1.013158		Initialization of Seasonal factors			
3	2011	3	432			1.136842					
4	2011	4	341			0.897368					
5	2012	1	382	400.9945	20.99448	0.952632			Initialization of L5 and T5		
6	2012	2	409	418.3288	19.89644	1.007839	427.541436	18.54144			
7	2012	3	498	438.1913	19.88625	1.136789	498.192926	0.192926			
8	2012	4	387	452.7143	18.27726	0.89099	411.064349	24.06435			
9	2013	1	473	476.0971	19.80893	0.958761	448.681393	24.31861			
10	2013	2	513	498.5268	20.59516	1.011018	499.793507	13.20649			
11	2013	3	582	517.6912	20.16593	1.134904	590.13223	8.13223			
12	2013	4	474	536.6842	19.81407	0.889821	479.225172	5.225172			
13	2014	1	544	558.6784	20.4681	0.961006	533.5489	10.4511			
14	2014	2	582	578.4487	20.25875	1.010286	585.527621	3.527621			
15	2014	3	681	598.9761	20.33935	1.13521	679.475612	1.524388			
16	2014	4	557	620.646	20.73852	0.890966	551.080127	5.919873			
17	2015	1	628	643.804	21.46436	0.963173	616.374301	11.6257			
18	2015	2	707	672.175	23.53636	1.016515	672.111514	34.88849			
19	2015	3	773	692.7554	22.64957	1.132303	789.778262	16.77826			
20	2015	4	592	705.2134	19.59211	0.88324	637.401508	45.40151			
21	2016	1	627	710.0391	15.16219	0.951154	698.113019	71.11302			
22	2016	2	725	722.8053	14.44339	1.014493	737.17786	12.17786			
23	2016	3	854	740.642	15.46136	1.135416	834.789143	19.21086			
24	2016	4	661	754.5588	14.99799	0.882156	667.82101	6.82101	k		
25							871.371647		1		
26							692.950477		2		
27							760.498109		3		
28							826.356023		4		

Example on seasonal sale of houses

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Year	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2015	2016	2016	2016	2016
Month	January	February	March	April	May	June	July	August	September	October	November	December	January	February	March	April
Date	01-01-2015	01-02-2015	01-03-2015	01-04-2015	01-05-2015	01-06-2015	01-07-2015	01-08-2015	01-09-2015	01-10-2015	01-11-2015	01-12-2015	01-01-2016	01-02-2016	01-03-2016	01-04-2016
# of houses sold	131646	176938	204653	214838	246787	277672	255973	266235	230914	213793	180825	210587	154183	168895	202480	237613
Period	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Year	2016	2016	2016	2016	2016	2016	2016	2016	2017	2017	2017	2017	2017	2017	2017	2017
Month	May	June	July	August	September	October	November	December	January	February	March	April	May	June	July	August
Date	01-05-2016	01-06-2016	01-07-2016	01-08-2016	01-09-2016	01-10-2016	01-11-2016	01-12-2016	01-01-2017	01-02-2017	01-03-2017	01-04-2017	01-05-2017	01-06-2017	01-07-2017	01-08-2017
# of houses sold	276346	288312	264873	261003	235482	238741	210884	200432	162196	193105	217340	267999	275747	293088	276851	266860
Period	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Year	2017	2017	2017	2017	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Month	September	October	November	December	January	February	March	April	May	June	July	August	September	October	November	December
Date	01-09-2017	01-10-2017	01-11-2017	01-12-2017	01-01-2018	01-02-2018	01-03-2018	01-04-2018	01-05-2018	01-06-2018	01-07-2018	01-08-2018	01-09-2018	01-10-2018	01-11-2018	01-12-2018
# of houses sold	240859	235874	218218	209316	155722	197765	224567	252527	260426	301841	274278	263149	256927	226363	205678	216753
Period	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	
Year	2019	2019	2019	2019	2019	2019	2019	2019	2019	2019	2019	2019	2020	2020	2020	
Month	January	February	March	April	May	June	July	August	September	October	November	December	January	February	March	
Date	01-01-2019	01-02-2019	01-03-2019	01-04-2019	01-05-2019	01-06-2019	01-07-2019	01-08-2019	01-09-2019	01-10-2019	01-11-2019	01-12-2019	01-01-2020	01-02-2020	01-03-2020	
# of houses sold	143530	195802	221326	240711	264338	323832	271417	272680	251292	236077	246599	234858	158800	211854	230158	

Period	Year	Month	Actual values	# of houses sold	Calculation		Seasonal	
1	2015	January	Y1	131646	Y1/Average (Y1, Y2, ..., Y12)	131646/217571.8	0.60506936	S1
2	2015	February	Y2	176938	Y2/Average (Y1, Y2, ..., Y12)	176938/217571.8	0.81323977	S2
3	2015	March	Y3	204653	Y3/Average (Y1, Y2, ..., Y12)	204653/217571.8	0.94062304	S3
4	2015	April	Y4	214838	Y4/Average (Y1, Y2, ..., Y12)	214838/217571.8	0.98743518	S4
5	2015	May	Y5	246787	Y5/Average (Y1, Y2, ..., Y12)	246787/217571.8	1.13427869	S5
6	2015	June	Y6	277672	Y6/Average (Y1, Y2, ..., Y12)	277672/217571.8	1.27623186	S6
7	2015	July	Y7	255973	Y7/Average (Y1, Y2, ..., Y12)	255973/217571.8	1.17649925	S7
8	2015	August	Y8	266235	Y8/Average (Y1, Y2, ..., Y12)	266235/217571.8	1.2236653	S8
9	2015	September	Y9	230914	Y9/Average (Y1, Y2, ..., Y12)	230914/217571.8	1.06132345	S9
10	2015	October	Y10	213793	Y10/Average (Y1, Y2, ..., Y12)	213793/217571.8	0.98263217	S10
11	2015	November	Y11	180825	Y11/Average (Y1, Y2, ..., Y12)	180825/217571.8	0.83110514	S11
12	2015	December	Y12	210587	Y12/Average (Y1, Y2, ..., Y12)	210587/217571.8	0.9678968	S12

innovate achieve lead

Time Series → Holts-Winters Triple Exponential Smoothing

Level	Seasonal
S1	0.605069
S2	0.81324
S3	0.940623
S4	0.987435
S5	1.134279
S6	1.276232
S7	1.176499
S8	1.223665
S9	1.061323
S10	0.982632
S11	0.831105
S12	0.967897

Period	Year	Month	# of houses sold	Level	Trend	Seasonal	Forecast	Error
12	2015	December	**210587**					
13	2016	January	**154183**	**254819**	**37247**	0.605069		
14	2016	February	168895	283627	36403	0.791464	237519	-68624

$$L_{13} = \frac{Y_{13}}{S_1} = \frac{154183}{0.605069} = 254819$$

$$T_{13} = \frac{Y_{13}}{S_1} - \frac{Y_{12}}{S_{12}} = \frac{154183}{0.605069} - \frac{210587}{0.967897} = 37247$$

Period	Year	Month	Date	# of houses sold	Level	Trend	Seasonal	Forecast	Error
13	2016	January	01-01-2016	154183	254819	37247	0.605069		
14	2016	February	01-02-2016	168895	283627	36403	0.791464	237519	-68624
15	2016	March	01-03-2016	202480	309554	35355	0.911971	301028	-98548
16	2016	April	01-04-2016	237613	334482	34313	0.959731	340575	-102962
17	2016	May	01-05-2016	276346	356278	33061	1.098416	418316	-141970
18	2016	June	01-06-2016	288312	372996	31427	1.225905	496887	-208575
19	2016	July	01-07-2016	264873	386494	29634	1.127382	475803	-210930
20	2016	August	01-08-2016	261003	395845	27606	1.167234	509202	-248199

53	2019	May	01-05-2019	264338	241299	-4314	1.041209	248171	16167	
54	2019	June	01-06-2019	323832	241218	-3891	1.177702	274759	49073	
55	2019	July	01-07-2019	271417	238952	-3728	1.076898	254022	17395	
56	2019	August	01-08-2019	272680	236468	-3604	1.106219	258983	13697	
57	2019	September	01-09-2019	251292	235441	-3346	0.981184	226254	25038	
58	2019	October	01-10-2019	236077	234962	-3059	0.91525	210117	25960	
59	2019	November	01-11-2019	246599	240361	-2214	0.80385	180692	65907	
60	2019	December	01-12-2019	234858	240942	-1934	0.891856	210200	24658	
61	2020	January	01-01-2020	158800	242951	-1540	0.578649	136310	22490	
62	2020	February	01-02-2020	211854	245654	-1116	0.757977	180183	31671	
63	2020	March	01-03-2020	230158	246861	-883	0.866844	210197	19961	k
64	2020	April	01-04-2020					256114		1
65	2020	May	01-05-2020					288648		2
66	2020	June	01-06-2020					262990		3
67	2020	July	01-07-2020					269174		4
68	2020	August	01-08-2020					237882		5
69	2020	September	01-09-2020					221089		6
70	2020	October	01-10-2020					193469		7
71	2020	November	01-11-2020					213862		8
72	2020	December	01-12-2020					138246		9

Period	Year	Month	Date	# of houses sold	Level	Trend	Seasonal			
1	2015	January	01-01-2015	131646	S1		0.605069			
2	2015	February	01-02-2015	176938	S2		0.81324			
3	2015	March	01-03-2015	204653	S3		0.940623			
4	2015	April	01-04-2015	214838	S4		0.987435			
5	2015	May	01-05-2015	246787	S5		1.134279			
6	2015	June	01-06-2015	277672	S6		1.276232			
7	2015	July	01-07-2015	255973	S7		1.176499			
8	2015	August	01-08-2015	266235	S8		1.223665	α	0.1	
9	2015	September	01-09-2015	230914	S9		1.061323	β	0.1	
10	2015	October	01-10-2015	213793	S10		0.982632	γ	0.1	
11	2015	November	01-11-2015	180825	S11		0.831105			
12	2015	December	01-12-2015	210587	S12		0.967897	Forecast	Error	
13	2016	January	01-01-2016	154183	254819	37247	0.605069			
14	2016	February	01-02-2016	168895	283627	36403	0.791464	237519	-68624	
15	2016	March	01-03-2016	202480	309554	35355	0.911971	301028	-98548	
16	2016	April	01-04-2016	237613	334482	34313	0.959731	340575	-102962	
17	2016	May	01-05-2016	276346	356278	33061	1.098416	418316	-141970	
18	2016	June	01-06-2016	288312	372996	31427	1.225905	496887	-208575	
19	2016	July	01-07-2016	264873	386494	29634	1.127382	475803	-210930	
20	2016	August	01-08-2016	261003	395845	27606	1.167234	509202	-248199	
21	2016	September	01-09-2016	235482	403293	25590	1.013581	449418	-213936	
22	2016	October	01-10-2016	238741	410291	23731	0.942557	421434	-182693	
23	2016	November	01-11-2016	210884	415993	21928	0.798689	360717	-149833	
24	2016	December	01-12-2016	200432	414837	19619	0.919423	423862	-223430	
25	2017	January	01-01-2017	162196	417817	17955	0.583382	262876	-100680	
26	2017	February	01-02-2017	193105	416594	16038	0.758671	344898	-151793	
27	2017	March	01-03-2017	217340	413200	14094	0.873373	394547	-177207	
28	2017	April	01-04-2017	267999	412489	12614	0.928729	410088	-142089	
29	2017	May	01-05-2017	275747	407697	10873	1.056209	466940	-191193	
30	2017	June	01-06-2017	293088	400621	9078	1.176473	513127	-220039	
31	2017	July	01-07-2017	276851	393287	7437	1.085038	461888	-185037	
32	2017	August	01-08-2017	266860	383514	5716	1.120094	467739	-200879	
33	2017	September	01-09-2017	240859	374070	4200	0.976612	394516	-153657	
34	2017	October	01-10-2017	235874	365468	2920	0.912842	356541	-120667	
35	2017	November	01-11-2017	218218	358871	1968	0.779627	294227	-76009	
36	2017	December	01-12-2017	209316	347522	636	0.887712	331764	-122448	
37	2018	January	01-01-2018	155722	340035	-176	0.57084	203109	-47387	
38	2018	February	01-02-2018	197765	331941	-968	0.742382	257842	-60077	
39	2018	March	01-03-2018	224567	323588	-1706	0.855435	289063	-64496	
40	2018	April	01-04-2018	252527	316885	-2206	0.915547	298941	-46414	
41	2018	May	01-05-2018	260426	307868	-2887	1.035179	332367	-71941	
42	2018	June	01-06-2018	301841	300139	-3371	1.159393	358801	-56960	
43	2018	July	01-07-2018	274278	292369	-3811	1.070346	322004	-47726	
44	2018	August	01-08-2018	263149	283196	-4347	1.101006	323212	-60063	
45	2018	September	01-09-2018	256927	277272	-4505	0.971613	272327	-15400	
46	2018	October	01-10-2018	226363	270288	-4753	0.905306	248993	-22630	
47	2018	November	01-11-2018	205678	265363	-4770	0.779172	207018	-1340	
48	2018	December	01-12-2018	216753	258951	-4934	0.882645	231331	-14578	
49	2019	January	01-01-2019	143530	253758	-4960	0.570318	145003	-1473	
50	2019	February	01-02-2019	195802	250293	-4811	0.746373	184703	11099	
51	2019	March	01-03-2019	221326	246807	-4678	0.859567	209994	11332	
52	2019	April	01-04-2019	240711	244208	-4470	0.92256	221680	19031	
53	2019	May	01-05-2019	264338	241299	-4314	1.041209	248171	16167	
54	2019	June	01-06-2019	323832	241218	-3891	1.177702	274759	49073	
55	2019	July	01-07-2019	271417	238952	-3728	1.076898	254022	17395	
56	2019	August	01-08-2019	272680	236468	-3604	1.106219	258983	13697	
57	2019	September	01-09-2019	251292	235441	-3346	0.981184	226254	25038	
58	2019	October	01-10-2019	236077	234962	-3059	0.91525	210117	25960	
59	2019	November	01-11-2019	246599	240361	-2214	0.80385	180692	65907	
60	2019	December	01-12-2019	234858	240942	-1934	0.891856	210200	24658	
61	2020	January	01-01-2020	158800	242951	-1540	0.578649	136310	22490	
62	2020	February	01-02-2020	211854	245654	-1116	0.757977	180183	31671	
63	2020	March	01-03-2020	230158	246861	-883	0.866844	210197	19961	k

Time Series ➡ Holts-Winters Triple Exponential Smoothing

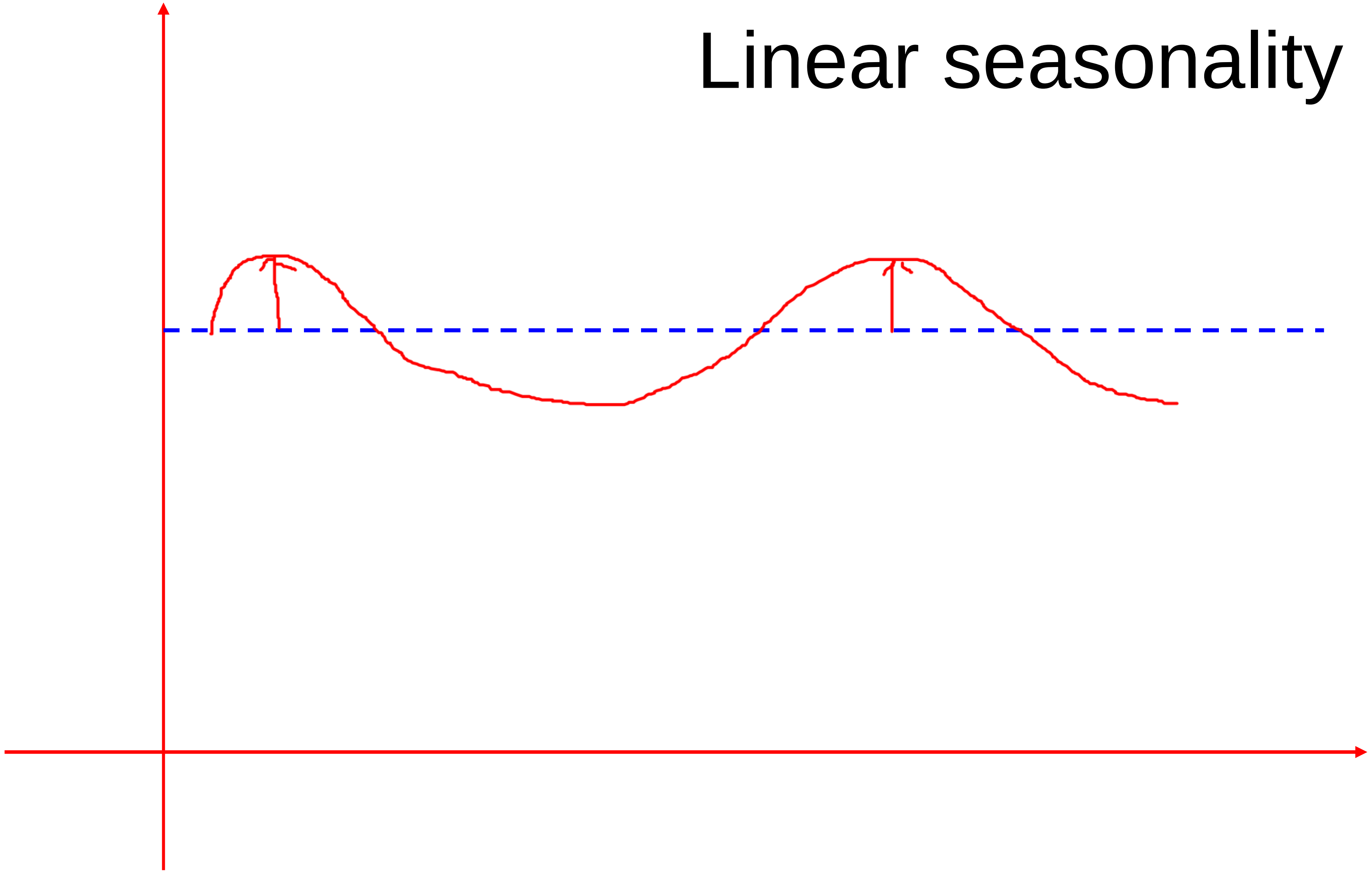
Trend	Seasonality		
	No	Linear	Ratio
No			
Linear			
Ratio			

Time Series ⇨ Smoothing Techniques – Exponential

Time Series ➡ Smoothing Techniques – Exponential

No trend

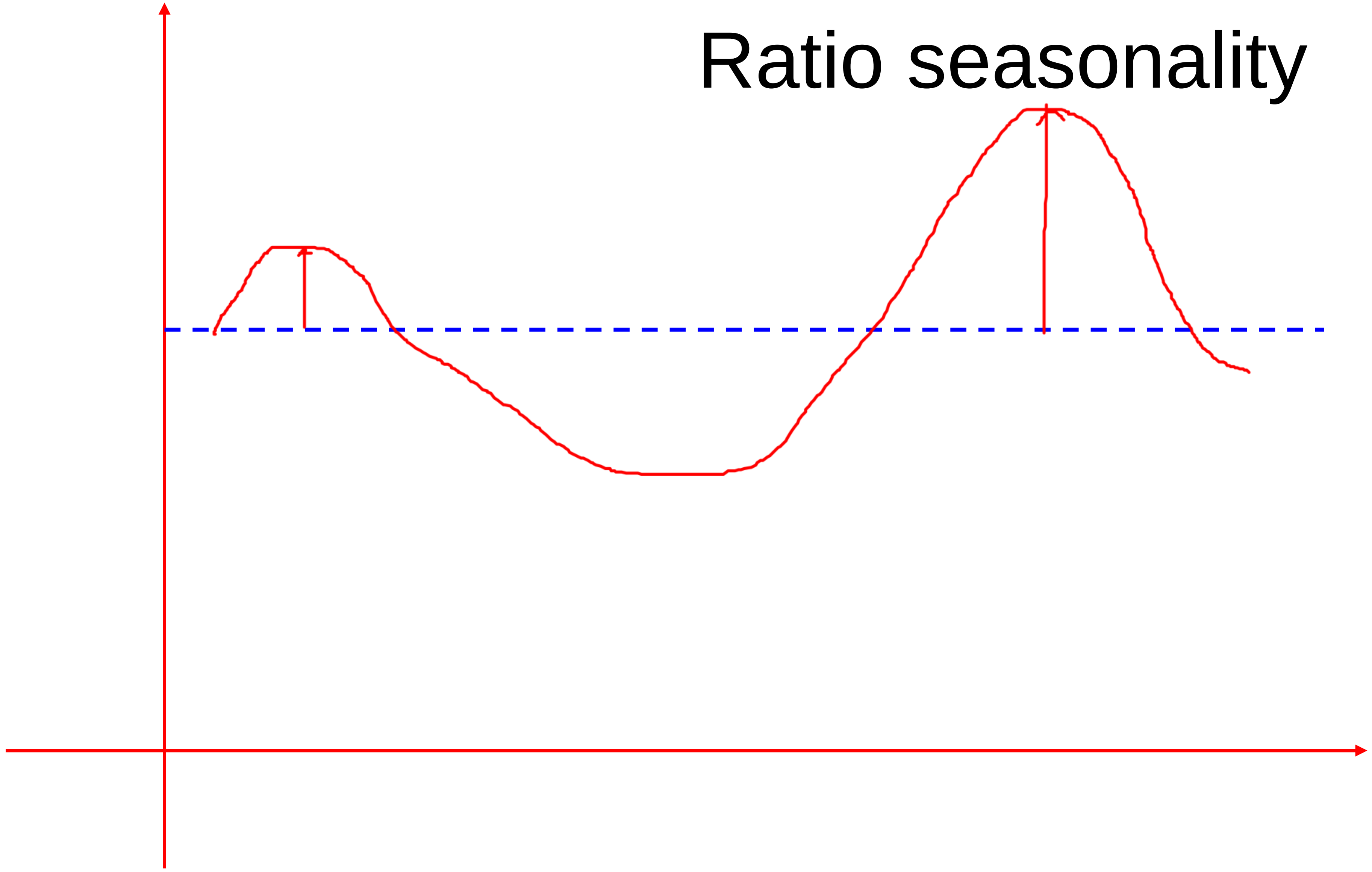
Linear seasonality



Time Series ➡ Smoothing Techniques – Exponential

No trend

Ratio seasonality



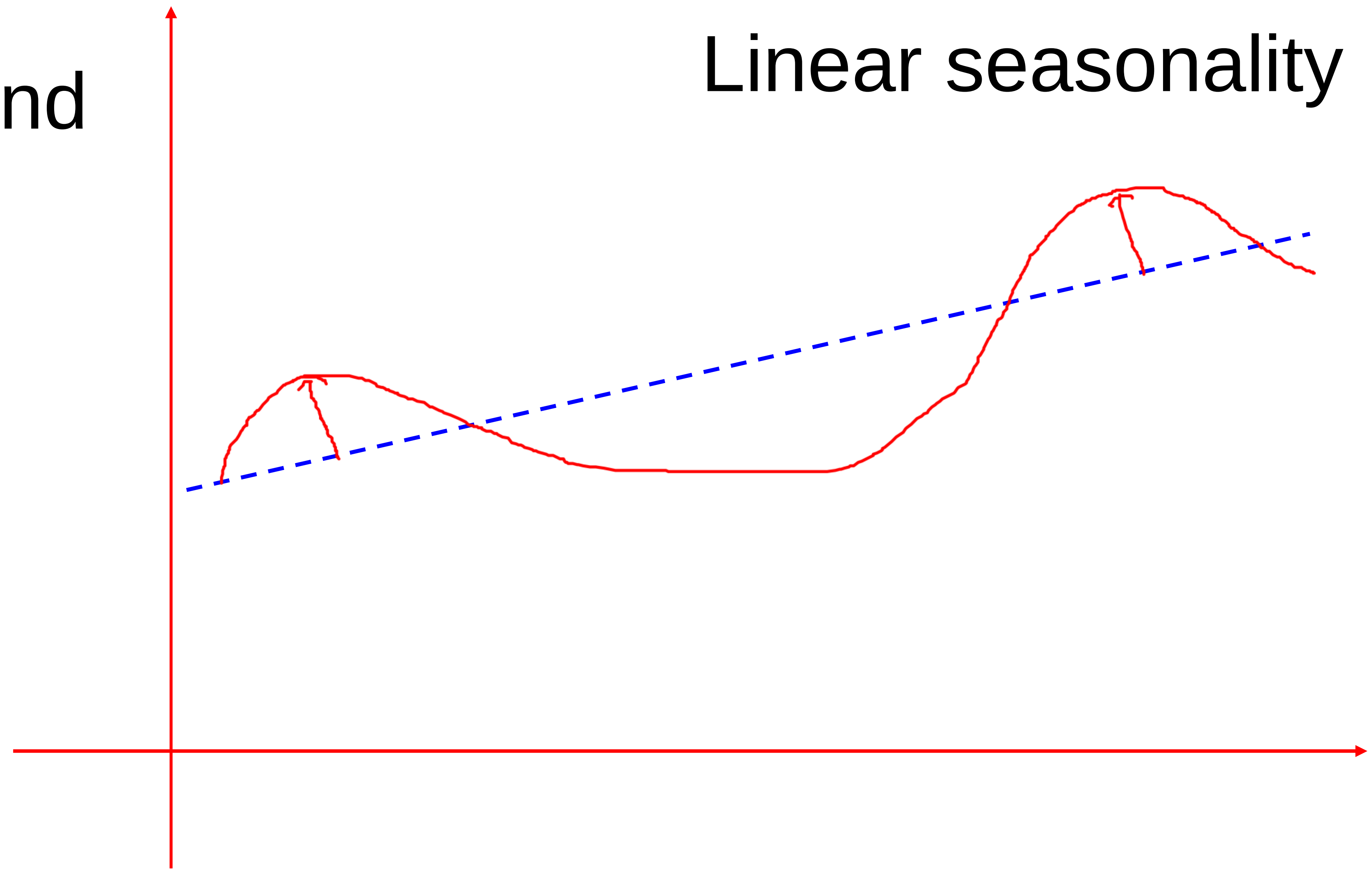
Time Series ➡ Smoothing Techniques – Exponential

Linear trend

Time Series ⇨ Smoothing Techniques – Exponential

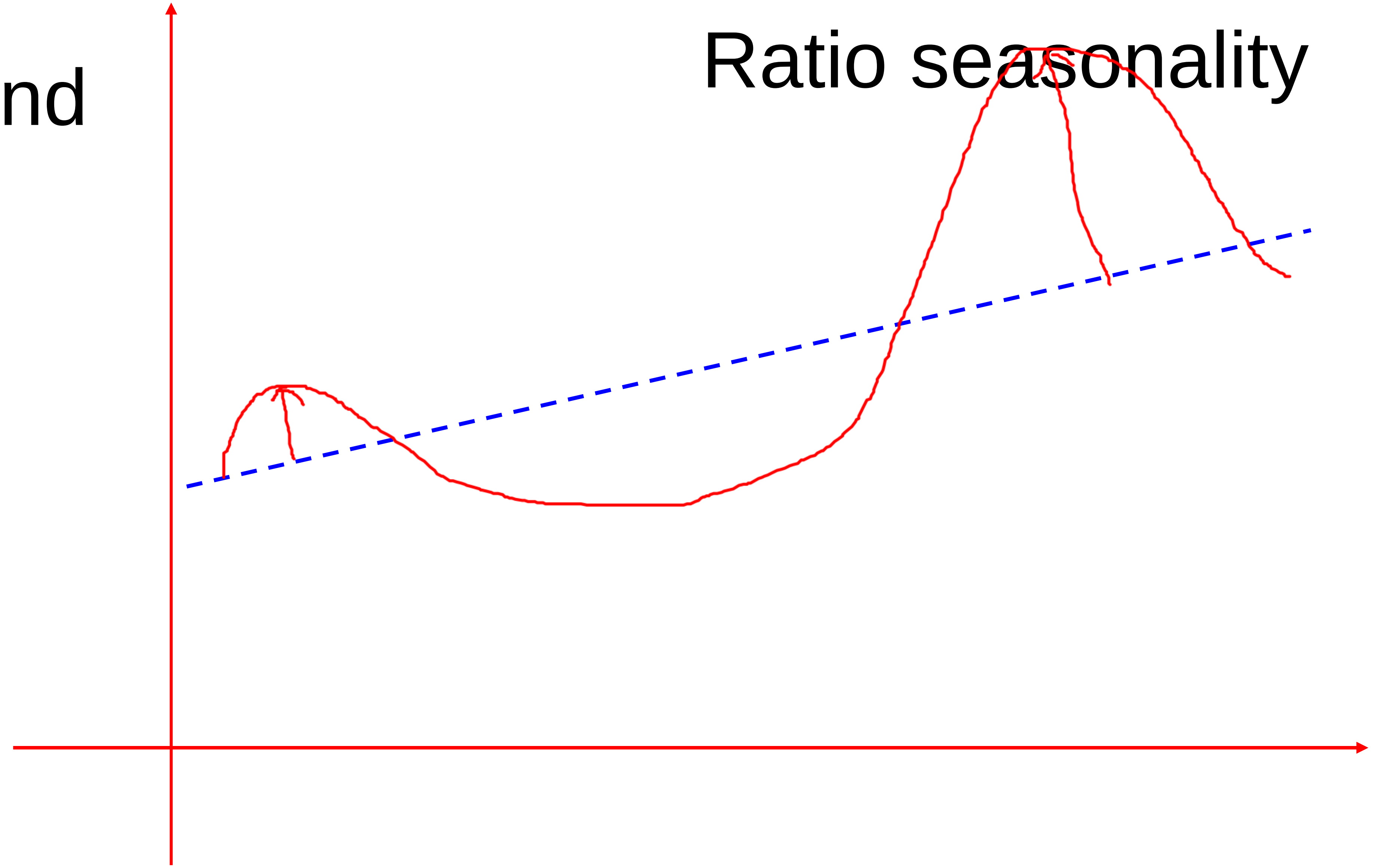
Linear trend

Linear seasonality



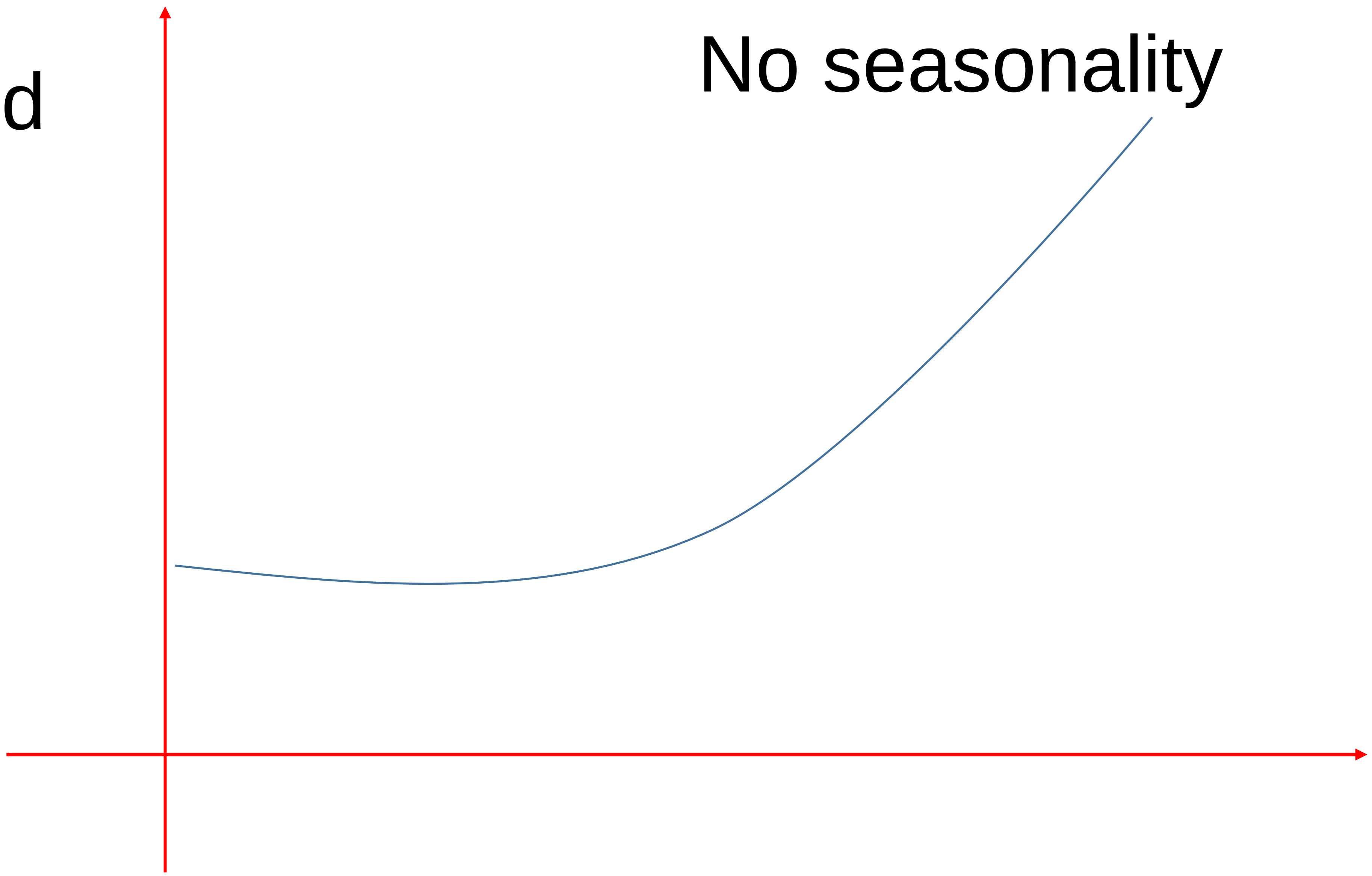
Time Series ➡ Smoothing Techniques – Exponential

Linear trend



Time Series ⇨ Smoothing Techniques – Exponential

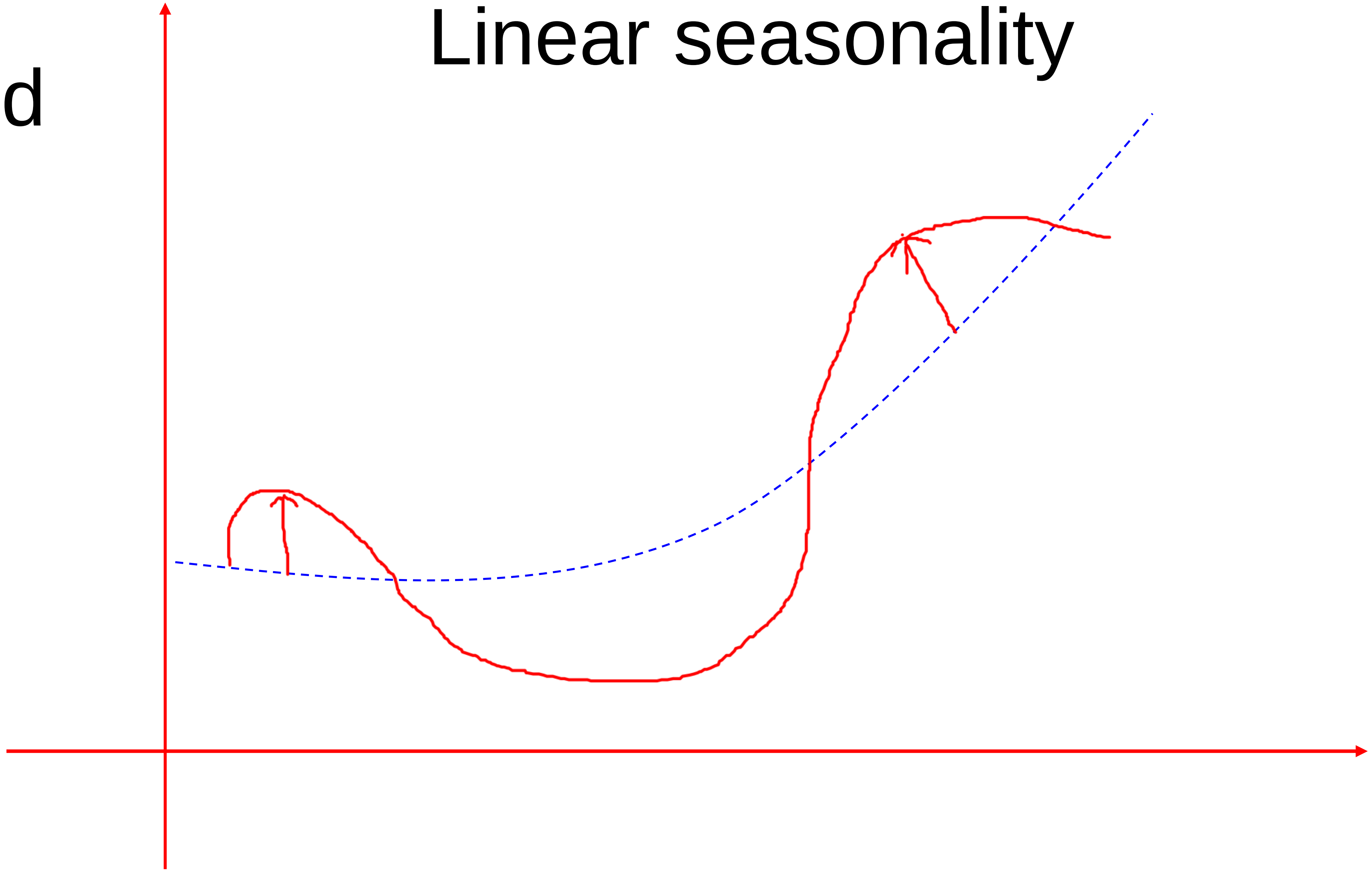
Ratio trend



Time Series → Smoothing Techniques – Exponential

Ratio trend

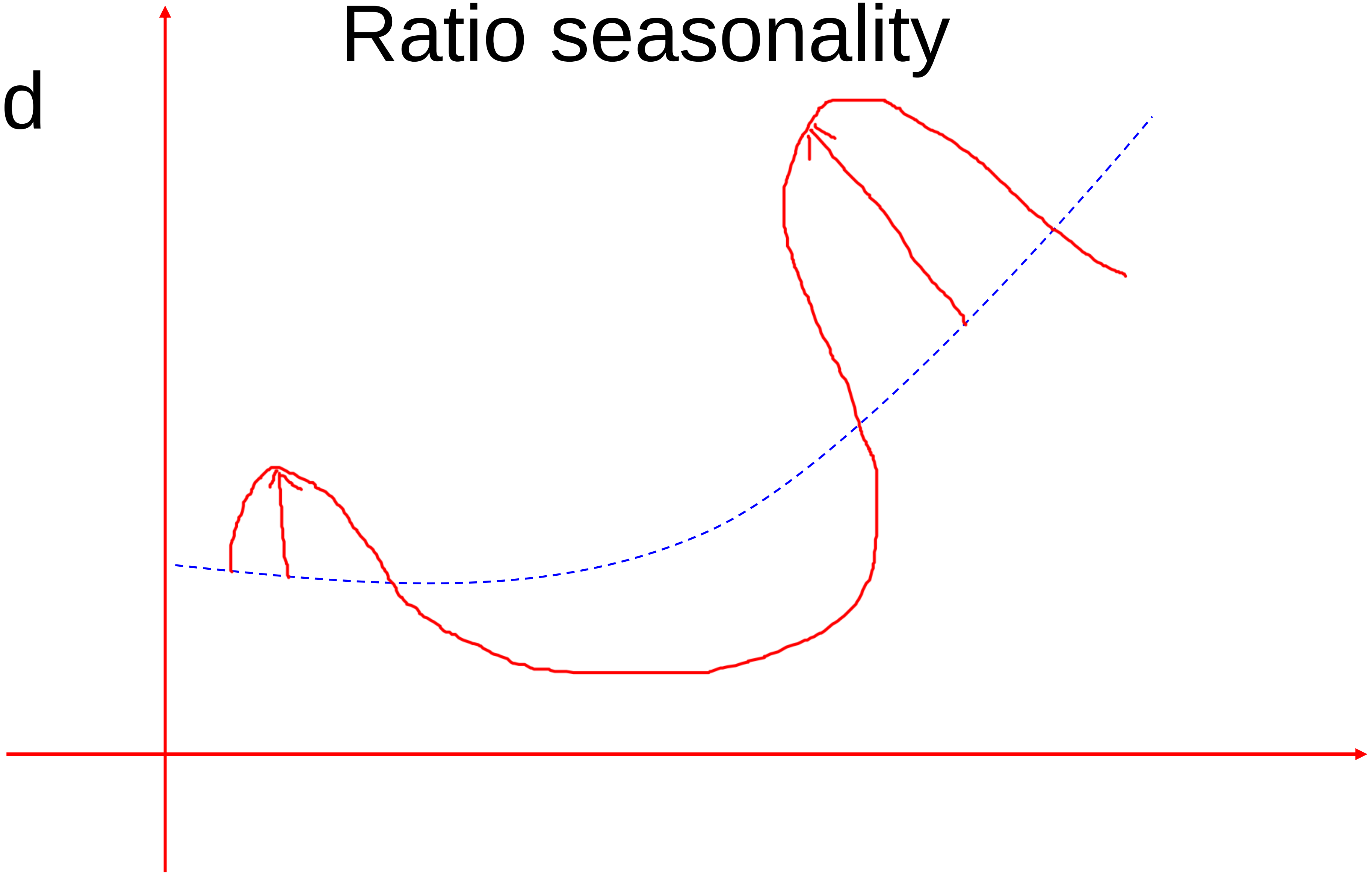
Linear seasonality



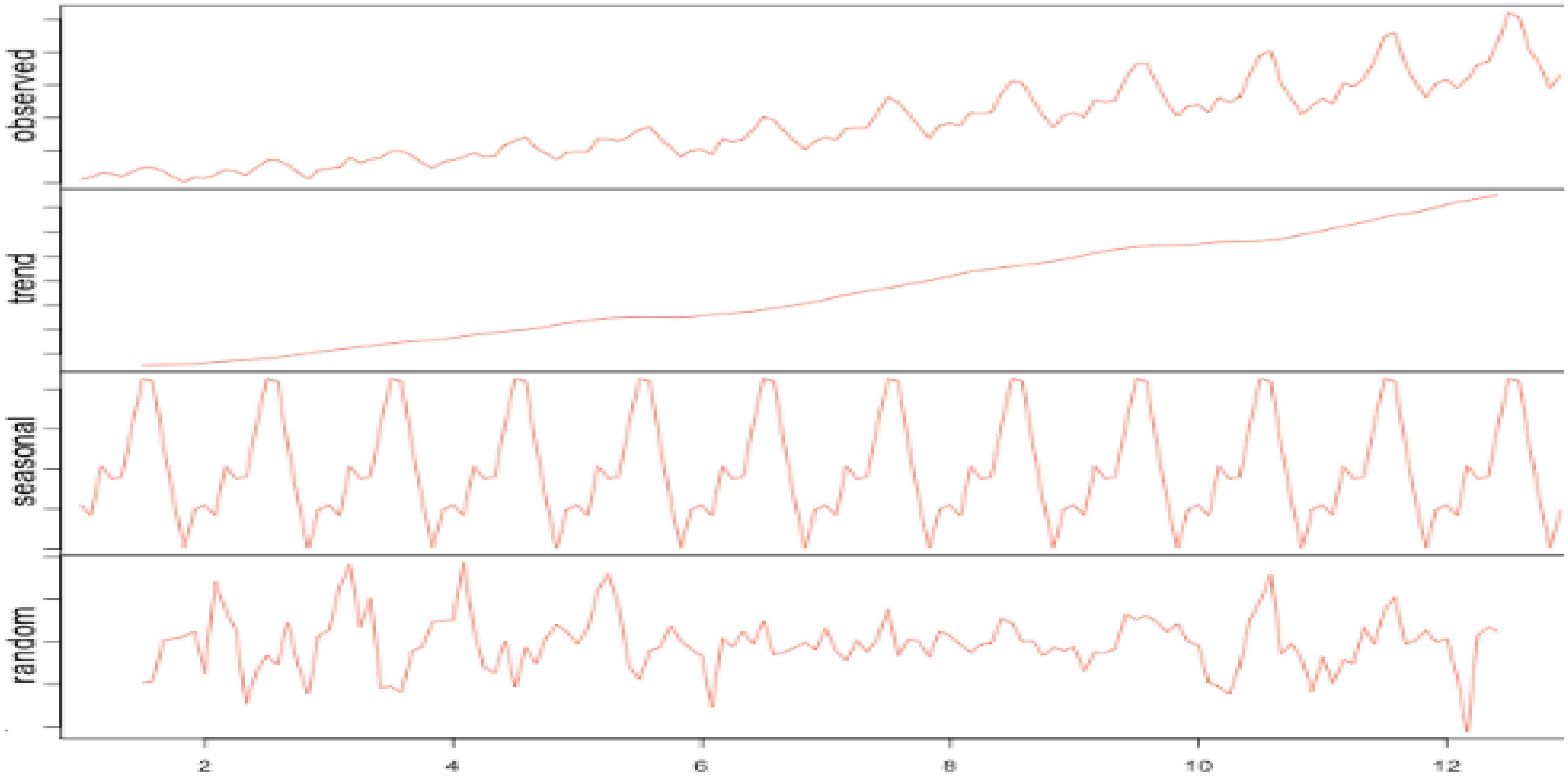
Time Series ⇨ Smoothing Techniques – Exponential

Ratio trend

Ratio seasonality



Time Series ⇨ Decomposition view



Time Series ⇨ Definition of Stochastic Process

- Consider the sequence of random variables Y_0, Y_1, Y_2, … which forms a family of random variables **$\{Y_t, t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$**

where Ω - Sample space

$\mathcal{F}$ - all collection of subsets A of Ω and

P – probability measure

- T – index set

Index set:

- The index set **T** is a collection of **all time functions** that can result from random experiment, usually the index **T** denote time. $\mathbf{Y_t}$ are **independent and identically distributed (iid) random variables** with mean μ_t and variance σ_t^2. Each realization of this process gives an ensemble or a data set.
- iid rvs: if each random variable has the same probability distribution as the others and all are mutually independent

Time Series ⇨ Definition of Stochastic Process

- Examples of index sets:

  (1) T = (-∞, ∞) or T = [0, ∞). In this case Y_t is a continuous time stochastic process.

  (2) T = {0, ±1, ±2, ….} or T = {0, 1, 2, …}. In this case Y_t is a discrete time stochastic process.

- We use uppercase letter $\{Y_t\}$ to describe the process. A time series, $\{Y_t\}$ is a realization or sample function from a certain process.

- We use information from a time series to estimate parameters and properties of process $\{Y_t\}$.

innovate achieve lead

Time Series → Probability distribution of the process

- For any stochastic process with index set *T*, its probability distribution function is uniquely determined by its finite dimensional distributions.
- The *k* dimensional distribution function of a process is defined by

$$F_{Y_{t_1},Y_{t_2},\ldots,Y_{t_k}} = P(Y_{t_1} \le y_{t_1}, Y_{t_2} \le y_{t_2}, \ldots, Y_{t_k})$$

for any t_1, t_2, …, $t_k \in T$ and any real numbers y_1, …, y_k .

- The distribution function tells us everything we need to know about the process $\{Y_t\}$.

Time Series ⇨ Definition of Stochastic Process

In stochastic time series,

$$Y_t,\ t \in Z = \{0, \pm 1, \pm 2, \ldots\}$$

is a family of random variables, Y_t denoting the value of the characteristic of interest at time t.

Thus, $\mathbf{Y} = (y_1, y_2, \ldots, y_n)'$ is seen as a realized value of the random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ with joint probability density function $f_{\mathbf{Y}}(y)$.

Time Series ⇨ Definition of Stochastic Process

The joint distribution function of a finite random variables

$$\{Y_{t_1}, \ldots Y_{t_n}\}, t_1 < t_2 < \ldots < t_n$$

from the collection $\{Y_t, t \in T\}$ is

$$F_{Y_{t_1}, \ldots Y_{t_n}}(y_1, \ldots, y_n) = P[Y_{t_1} \leq y_{t_1}, \ldots, P[Y_{t_n} \leq y_{t_n}],$$

$$(y_{t_1}, \ldots, y_{t_n}) \in R^n$$

Time Series ➡ **Stationary Stochastic Process in Time series**

A special kind of stochastic process is based on the assumption that the process is in a particular state of equilibrium. This type of assumption is called stationarity.

A stochastic process is strictly stationary if its properties are unaffected by a change of origin.

Time Series ➡ Moments of Stochastic Processes

- We can describe a stochastic process via its moments, ie., $E(Y_t)$, $E(Y^2_t)$, $E(Y_t, Y_s)$, etc. We often use the first two moments.
- The mean function of the process is $E(Y_t) = \mu_t$.
- The variance function of the process is $V(Y_t) = \sigma^2_t$.

Time Series ⇨ Moments of Stochastic Processes

- The Covariance function between Y_t, and Y_s is

$$Cov\ (Y_t, Y_s) = E((Y_t - \mu_t)\ (Y_s - \mu_s))$$

- The Correlation function between Y_t, and Y_s is

$$\rho\ (Y_t, Y_s) = \frac{Cov\ (Y_t, Y_s)}{\sqrt{\sigma_t^2}\sqrt{\sigma_s^2}}$$

- These moments are functions of time

Thus, a time series stochastic process $\{Y_{t_1}, Y_{t_2}, \ldots Y_{t_k}\}$ is said to be strictly stationary, if the joint distribution of k observations $Y_{t_1}, \ldots Y_{t_k}$ made at time $t_1, \ldots, t_k$ is same as that of k+h observations $Y_{t_{1+h}}, Y_{t_{2+h}}, \ldots Y_{t_{k+h}}$ made at time points $t_{1+h}, \ldots, t_{k+h}$ for any h. That is,

$$F_{Y_{t_1}, \ldots Y_{t_k}}(y_{t_1}, \ldots, y_{t_k}) = F_{Y_{t_{1+h}}, Y_{t_{2+h}}, \ldots Y_{t_{k+h}}}(y_{t_{1+h}}, y_{t_{2+h}}, \ldots, y_{t_{k+h}})$$

Time Series ⇨ Stationary Stochastic Process in Time series

- If $\{Y_t\}$ is a strictly stationary process and $E(Y^2_t) < \infty$ then, the mean function is a constant and the variance function is also a constant.
- Moreover, for a strictly stationary process with first two moments finite, the covariance function, and the correlation function depends only on the time difference *s*.
- A trivial example of a strictly stationary process is a sequence of independent and identically distributed (iid) random variables.

Time Series ➡ Stationary Stochastic Process in Time series

That is, $F_{Y_{t_1}, \ldots Y_{t_k}}(y_{t_1}, \ldots, y_{t_k}) = F_{Y_{t_{1+h}}, \ldots Y_{t_{k+h}}}(y_{t_{1+h}}, \ldots, y_{t_{k+h}})$

for all possible non-empty finite distinct sets

$(t_1, \ldots, t_k)$ and $(t_{1+h}, \ldots, t_{k+h})$ in the index set T and all $y_{t_1}, \ldots, y_{t_k}$ in the range of random variables Y_t.

Note: For reliable prediction to be made, the time series should be stationary (No systematic change such as seasonality, trend i.e. only random fluctuations)

Time Series ➡ Stochastic Process

- Strict stationarity is too strong of a condition in practice. It is often difficult assumption to assess based on an observed time series $Y_1, \ldots, Y_k$.
- In time series analysis we often use a weaker sense of stationarity in terms of the moments of the process.
- A process is said to be *n*th-order weakly stationary if all its joint moments up to order *n* exists and are time invariant, i.e., independent of time origin.

Time Series ⇨ Stochastic Process

- For example, a second-order weakly stationary process will have constant mean and variance, with the covariance and the correlation being functions of the time difference along.
- A strictly stationary process with the first two moments finite is also a second-ordered weakly stationary. But a strictly stationary process may not have finite moments and therefore may not be weakly stationary.

When k = 1, strict stationary implies that the pdf $f(y_t)$ is the same for all t, is say f(y). The stochastic process f(y) has a constant mean

$$\mu = E(Y_t) = \int_{-\infty}^{\infty} y f(y) dy$$

and a constant variance

$$\sigma^2 = E(Y_t - \mu)^2 = \int_{-\infty}^{\infty} (y-\mu)^2 \mathrm{f(y)dy}$$

provided the mean and variance exists.

The mean (μ) and variance (σ²) are estimated as

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

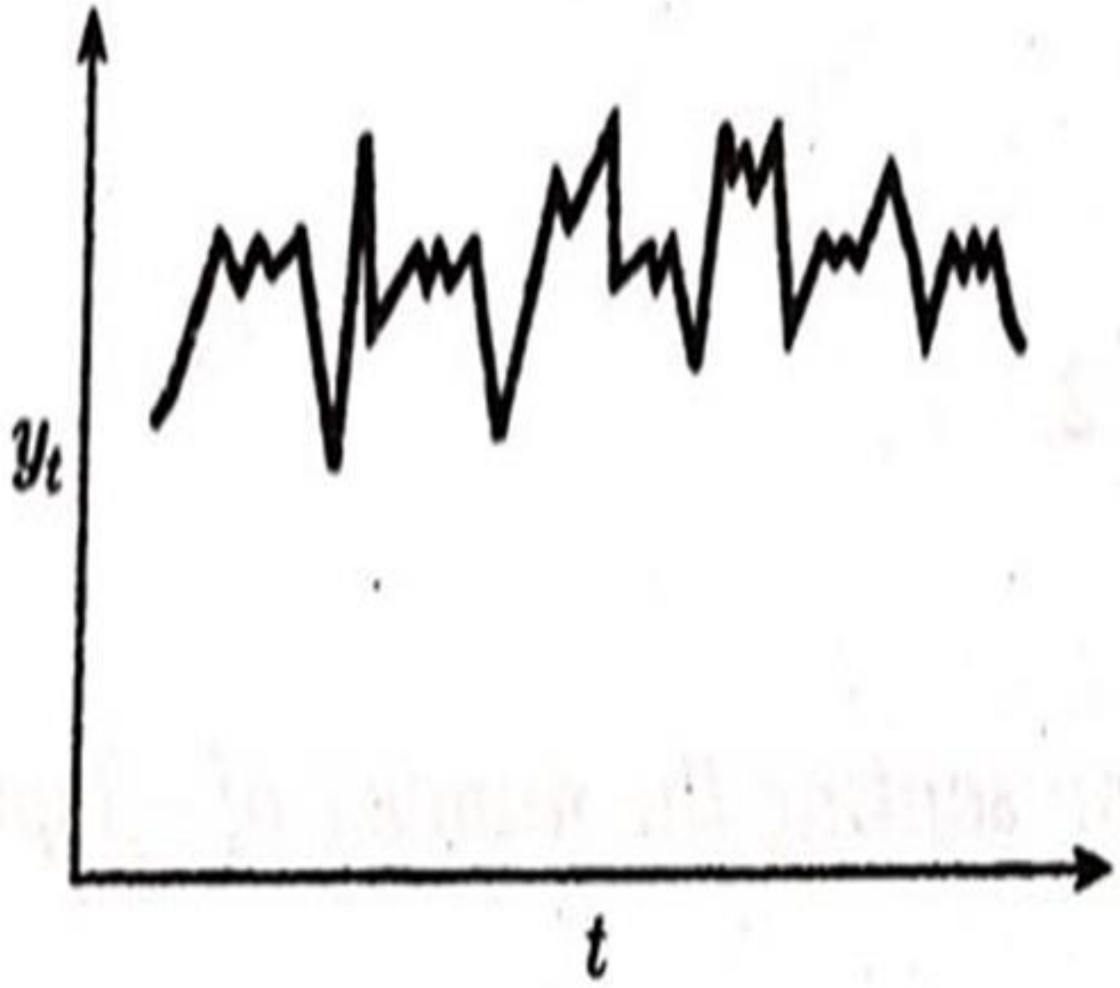
$$\hat{\sigma}_y^2 = S_y^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2$$

Time Series ➡ Stationary Stochastic Process in Time series

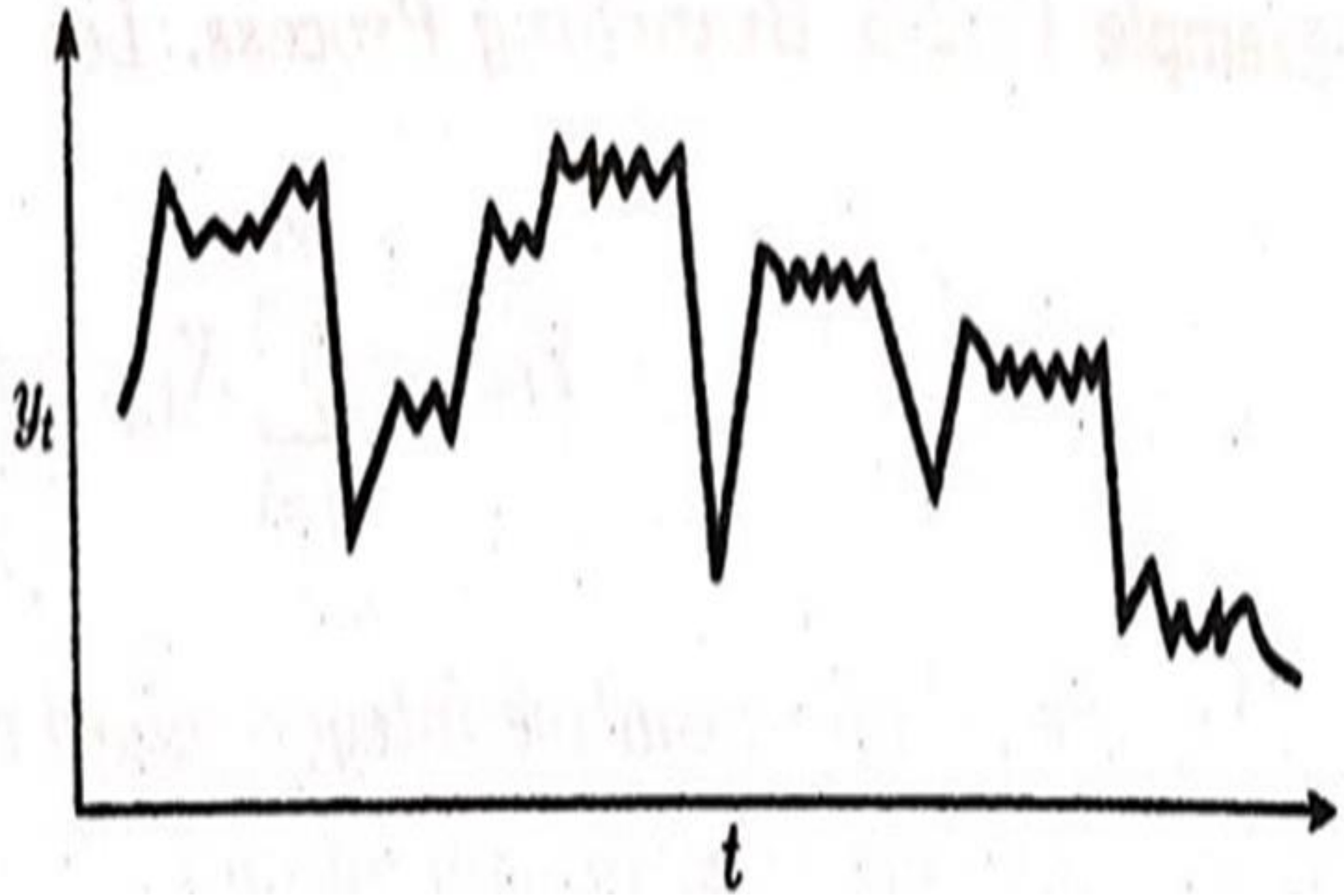
Thus, a stationary process remains at equilibrium about a common mean value.

However, in industry, business (ex. Stock price) and economics many time series are better represented as non-stationary and in particular, having no natural mean.

Time Series → Stationary Stochastic Process in Time series



Stationary time series



Non-stationary time series

Time Series ➡ Stationary Stochastic Process in Time series

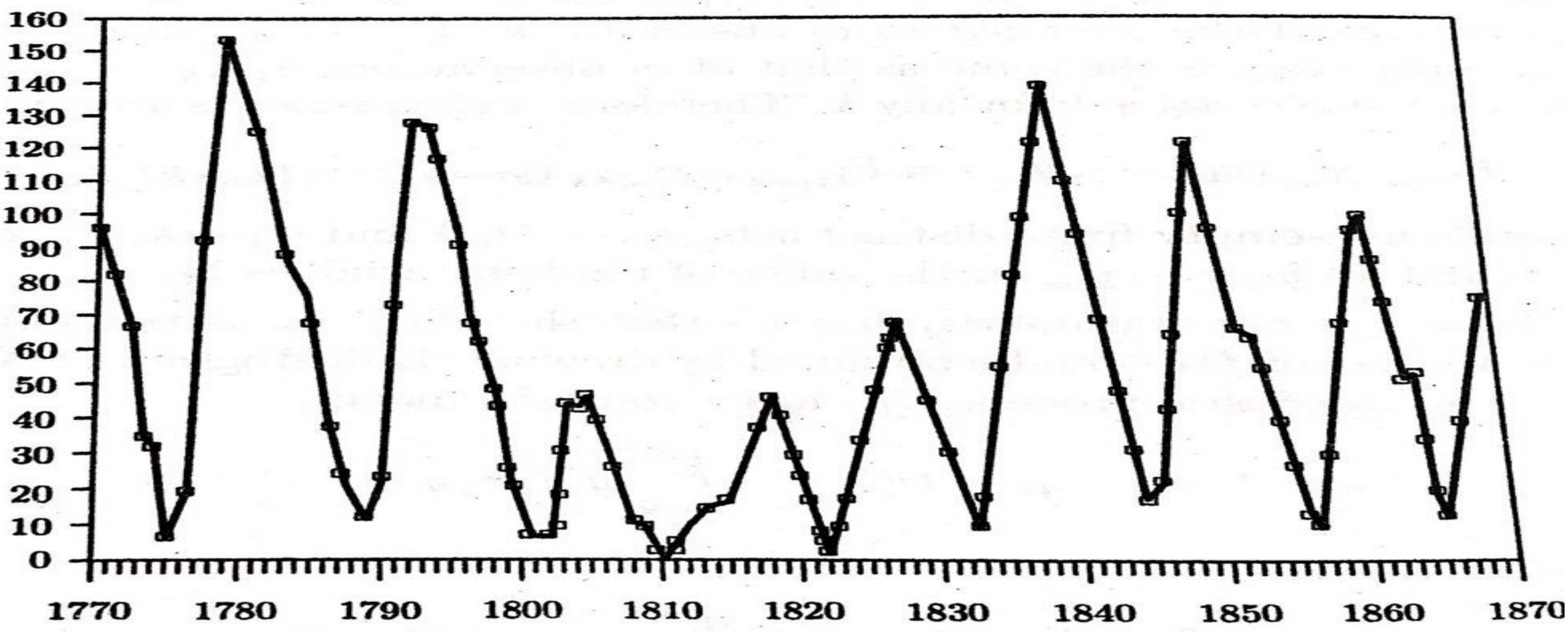


Fig 15.3 The Wolfer sunspot numbers, 1770-1869.
(Source: Box and Jenkins, 1976)

Time Series ⇨ Stationary Stochastic Process in Time series

However, many non-stationary time series can be so modified that the reduced to time series obeys the original series.

The two main component which cause lack of stationarity are trend and seasonality. In fitting the stationary time series model, we therefore, assume that the trend and seasonality have been eliminated from the original series.

Thank you