

# АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

МОДЕЛИ И ПРОГНОЗИРОВАНИЕ

---

# ИСХОДНЫЕ СТАТИСТИЧЕСКИЕ ДАННЫЕ

- Если процесс регистрации данных происходит для  $n$  объектов по  $p$  признакам и по времени  $t$   $x_i^{(j)}(t_k)$   $j=1, 2, \dots, p; i=1, 2, \dots, n; k=1, 2, \dots, N$ , то говорят о *панельных данных*.

Если рассматривать значения одного признака у одного объекта в равноотстоящие моменты времени, то последовательность  $x(t_1), x(t_2), \dots, x(t_N)$  называют *одномерным временным рядом*.

Если регистрировать значения  $p$  признаков у одного объекта, то говорят о статистическом анализе *многомерного временного ряда*  $X(t) = (x^1(t_k), x^2(t_k), \dots, x^p(t_k)), k=1, 2, \dots, N$

# ИСХОДНЫЕ СТАТИСТИЧЕСКИЕ ДАННЫЕ

Говоря о проблеме прогнозирования на основе одномерных временных рядов, обычно имеется в виду *кратко-* и *среднесрочный* прогноз, поскольку построение долгосрочного прогноза подразумевает обязательное использование методов организации и статистического анализа *специальных экспертных оценок*.

Использование доступных к моменту  $t=N$  наблюдений временного ряда  $x(t)$  для прогнозирования может явиться основой для:

- планирования в экономике, производстве, торговле
- управления и оптимизации социально-экономических процессов
- принятия оптимальных решений в бизнесе
- частичного управления параметрами демографических процессов

# ОСНОВНЫЕ ФАКТОРЫ ВРЕМЕННЫХ РЯДОВ

1. *Долговременные*, формирующие общую тенденцию в изменении анализируемого признака  $x(t)$ . Обычно описывается при помощи монотонной функции  $f(t)$ , называемой *трендом*.
2. *Сезонные*, формирующие периодически повторяющиеся в определенное время года колебания анализируемого признака. Описывается периодической функцией  $\phi(t)$  с периодом, кратным сезонам.
3. *Циклические*, формирующие изменения анализируемого признака, обусловленные действием долговременных циклов экономической, демографической или астрономической природы. Описывается функцией  $\psi(t)$ .
4. *Случайные*, не поддающиеся учету и регистрации. Их воздействие обуславливает *стохастическую природу* анализируемого признака. Обозначается  $\square(t)$

# ОБЩИЕ МОДЕЛИ ВРЕМЕННЫХ РЯДОВ

*Аддитивная форма*

$$x(t) = \lambda_1 f(t) + \lambda_2 \varphi(t) + \lambda_3 \psi(t) + \varepsilon(t), \quad (1)$$

$$\lambda_i = \begin{cases} 1, & \text{фактор участвует в формировании} \\ & \text{уровней ряда } x(t) \\ 0, & \text{в противном случае} \end{cases}$$

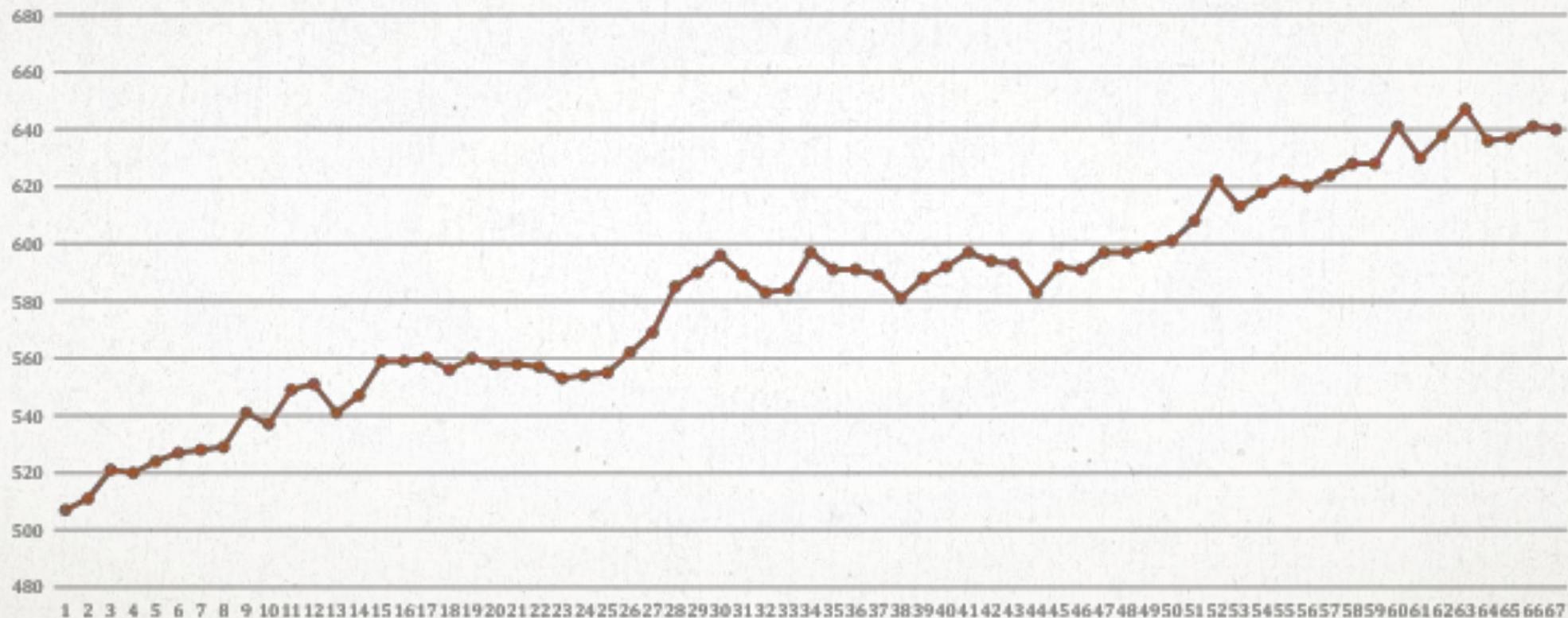
*Мультипликативная форма*

$$x(t) = f(t)^{\lambda_1} * \varphi(t)^{\lambda_2} * \psi(t)^{\lambda_3} * \varepsilon(t)$$

$$\ln x(t) = \lambda_1 \ln f(t) + \lambda_2 \ln \varphi(t) + \lambda_3 \ln \psi(t) + \ln \varepsilon(t)$$

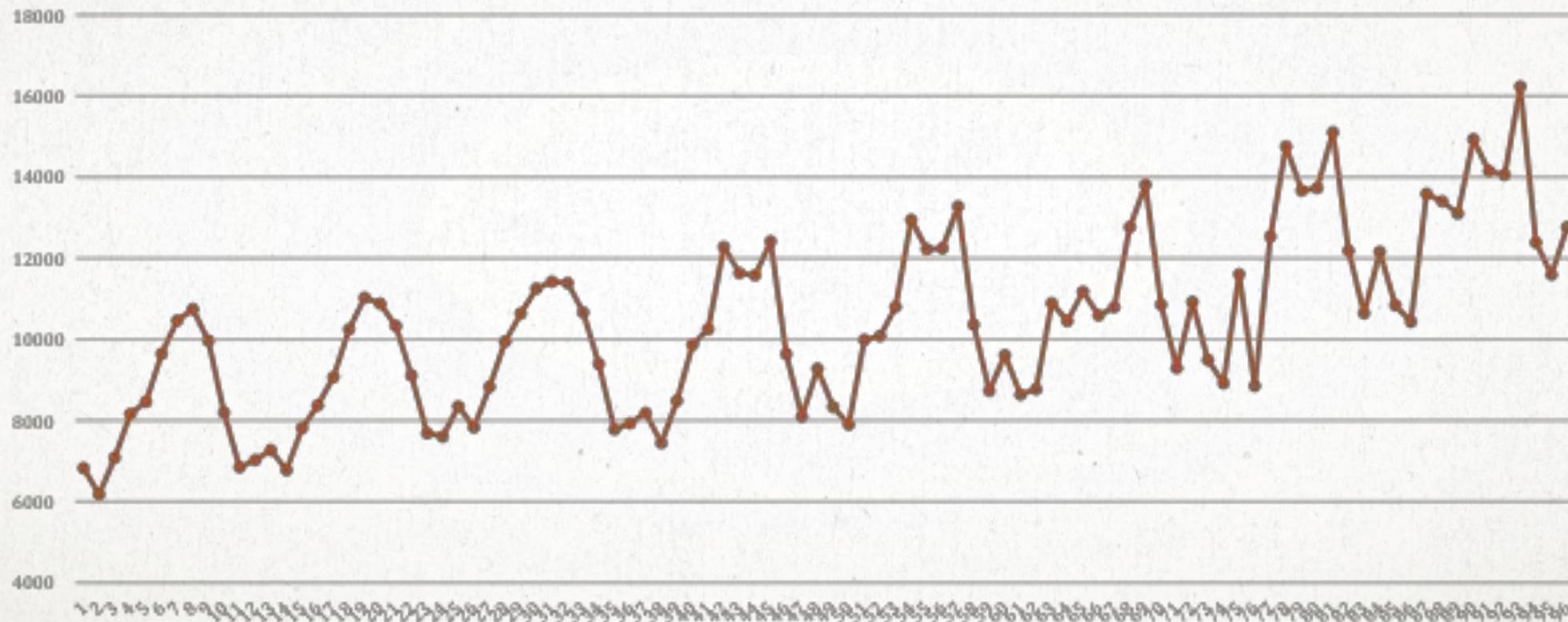
# ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ

IBM, 1961.



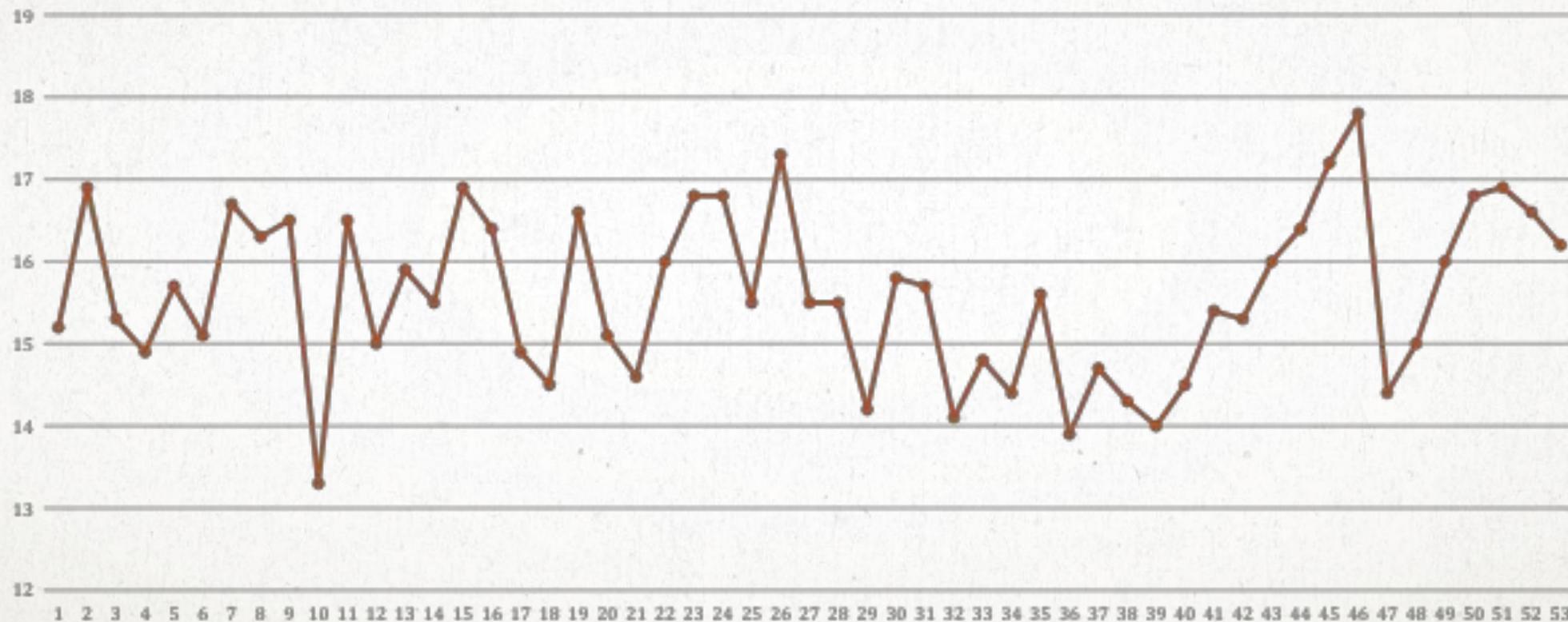
# ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ

Среднемесячные температуры в Москве, 1963-1970 гг.



# ПРИМЕРЫ ВРЕМЕННЫХ РЯДОВ

Средняя температура в Москве, 1884-1939 гг.



# ЗАДАЧИ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

По имеющейся траектории анализируемого временного ряда  $x(t)$  требуется:

- какие из неслучайных составляющих  $f(t)$ ,  $\varphi(t)$  и  $\psi(t)$  присутствуют в разложении (1)
- Построить «хорошие» оценки для тех неслучайных функций, которые присутствуют в разложении (1)
- Подобрать модель, адекватно описывающую поведение «случайной составляющей  $\varepsilon(t)$ , и статистически оценить параметры этой модели

# ТЕСТИРОВАНИЕ НАЛИЧИЯ/ОТСУТСТВИЯ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

- Если неслучайные составляющие отсутствуют, то ряд состоит из статистически независимых наблюдений, случайно варьирующихся около некоторого постоянного уровня  $a$ , т.е.

$$H_0: M(x(t)) = a = \text{const}$$

Иначе, существует зависимость от времени неслучайной составляющей анализируемого временного ряда

$$H_1: M(x(t)) \neq \text{const}$$

# КРИТЕРИЙ СЕРИЙ, ОСНОВАННЫЙ НА МЕДИАНЕ

- Из элементов временного ряда образуем «серии» плюсов и минусов по правилу

$$x(t) := \begin{cases} +, & x(t) > x_{med} \\ -, & x(t) < x_{med} \end{cases}$$

Члены временного ряда, равные медиане  $x_{med}$  не учитываются

Под «серией» подразумевается последовательность подряд идущих плюсов или минусов (в частности, серия может состоять из одного элемента)

Вычисляют:  $\nu$  – общее число серий;  $\tau$  – протяженность самой длинной серии. Если хотя бы одно из неравенств

$$\begin{aligned} \nu &> \frac{1}{2} (n + 2 - 1.96\sqrt{n - 1}) \\ \tau &< 1.43 \ln(n + 1) \end{aligned}$$

не выполняется, то гипотеза  $H_0$  отвергается с вероятностью ошибки  $\alpha \in (0.05; 0.0975)$

# КРИТЕРИЙ «ВОСХОДЯЩИХ» И «НИСХОДЯЩИХ» СЕРИЙ

- Этот критерий «улавливает» постепенное смещение среднего значения в исследуемом распределении не только монотонного, но и более общего, например периодического, характера.

$$x(t) := \begin{cases} +, & x(t) > x(t-1) \\ -, & x(t) < x(t-1) \end{cases}$$

Если два и несколько идущих друг за другом наблюдений равны между собой, то принимается во внимание только одно из них. Критерий основан на том же соображении, что и предыдущий: если выборка случайна, то в образованной последовательности знаков общее число серий не должно быть слишком малым, а их протяженность – слишком большой. Если хотя бы одно из неравенств

$$v > \frac{1}{3}(2n-1) - 1.96 \sqrt{\frac{16n-29}{90}}$$
$$\tau < \tau_0(n),$$

где  $\tau_0(n < 27) = 5$ ,  $\tau_0(26 < n < 154) = 6$ ,  $\tau_0(153 < n < 1171) = 7$

не выполняется, то гипотеза  $H_0$  отвергается с вероятностью ошибки  $\alpha \in (0.05; 0.0975)$

## КРИТЕРИЙ АББЕ

- Критерий применяется, если выборка подчиняется нормальному закону распределения.

$$Y_{\text{набл}} = \frac{q^2}{s^2}$$
$$q^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (x(i+1) - x(i))^2; \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x(i) - \bar{x})^2; \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x(i)$$

Если окажется, что  $Y_{\text{набл}} < Y_{\text{min}}$ , то гипотеза  $H_0$  отвергается, где

$$Y_{\text{min}} = 1 + \frac{u_\alpha}{\sqrt{n + 0.5(1 + u_\alpha^2)}}$$

$u_\alpha$  –  $\alpha$ -квантиль нормированного нормального распределения

# МЕТОДЫ СГЛАЖИВАНИЯ ВРЕМЕННОГО РЯДА

- *Аналитические* методы основаны на допущении, что известен общий вид неслучайной составляющей разложения (1). Тогда задача выделения неслучайной составляющей (задача *элиминирования* случайных остатков, задача *сглаживания* временного ряда) сводится к задаче построения «хороших» оценок параметров модели. Т.е.  $\hat{f}(t)$  будет представлена в виде формулы  $f(t, \hat{\theta})$  функции известного вида, в которой неизвестные параметры  $\theta = (\theta_0, \theta_1, \dots)$  заменены их статическими оценками.
- *Алгоритмические* методы не имеют допущения о том, что аналитический вид известен. «На выходе» задачи приводится алгоритм расчета оценки  $\hat{f}(t)$  в любой наперед заданной точке  $t$ .

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

- Реализуются в рамках моделей регрессии, в которых в роли зависимой переменной выступает  $x(t)$ , генерирующая анализируемый временной ряд, а единственной объясняющей переменной является время  $t$ .

Рассматривается модель  $x(t) = f(t, \theta) + \varepsilon(t)$ ,  $t=1,2,\dots,N$

в которой общий вид функции  $f(t, \theta)$  известен, но неизвестны параметры  $\theta$ . Если ошибки  $\varepsilon(t)$  взаимно некоррелированы, то оценки параметров могут быть получены с помощью МНК:

$$\hat{\theta} = (X^T X)^{-1} X^T Y$$

В случае коррелированности ошибок необходимо применять ОМНК. В некоторых случаях необходима техника статистического анализа нелинейных моделей регрессии.

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

На практике различают четыре основных типа экономического роста:

- I – постоянный рост (с постоянным или близким к нему абсолютным цепным приростом);
- II – увеличивающийся рост (с увеличивающимся абсолютным цепным приростом);
- III – уменьшающийся рост (с уменьшающимся абсолютным цепным приростом);
- IV – рост с качественными изменениями динамических характеристик на протяжении исследуемого периода.

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

Для каждого типа роста наиболее часто в практике экономических исследований встречаются следующие виды функций трендов.

*I тип роста*

- Линейная функция:  $f(t) = \alpha_0 + \alpha_1 t$ .
- Линейно-гиперболическая функция:  $f(t) = \alpha + \beta t + \gamma / t$ , где  $\beta > 0$ ;  $\gamma > 0$ .
- Линейно-логарифмическая функция 2-го порядка:

$$f(t) = \alpha_0 + \alpha_1 \ln(t) + \alpha_2 \ln^2(t), \quad \text{где } \alpha_1 > 0; \alpha_2 > 0.$$

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

*II тип роста*

- Показательная функция:  $f(t) = \alpha(1+\beta)^t$ , где  $\alpha > 0; \beta > 0$ .
- Парабола 2-го порядка:  $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2$ , где  $\alpha_1 > 0; \alpha_2 > 0$ .
- Парабола 3-го порядка:  $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$ ,

где  $\alpha_1 > 0; \alpha_2 > 0; \alpha_3 > 0$ .

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

- *III тип роста*

- Степенная функция:  $f(t) = \alpha t^\beta$ , где  $\alpha > 0$ ;  $0 < \beta < 1$ .

- Линейно-логарифмическая функция:  $f(t) = \alpha_0 + \alpha_1 \ln(t)$ , где  $\alpha_1 > 0$ .

- Парабола 2-го порядка:  $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2$ , где  $\alpha_1 > 0$ ;  $\alpha_2 > 0$ .

- Гипербола 1-го порядка:  $f(t) = \alpha_0 + \alpha_1 / t$ , где  $\alpha_1 < 0$ .

- Гипербола 2-го порядка:  $f(t) = \alpha_0 + \alpha_1 / t + \alpha_2 / t^2$ , где  $\alpha_1 < 0$ ;  $\alpha_2 < 0$ .

- Модифицированная экспонента:  $f(t) = \alpha + \beta e^{-t}$ , где  $\beta < 0$ .

# АНАЛИТИЧЕСКИЕ МЕТОДЫ ОЦЕНКИ НЕСЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

*W* тип роста

- Линейно-логарифмическая ф-ция 2-го порядка:  $f(t) = \alpha_0 + \alpha_1 \ln(t) + \alpha_2 \ln^2(t)$ , где  $\alpha_1 > 0$ ;  $\alpha_2 > 0$ .
- Парабола 3-го (и более высоких) порядков:  $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$ , где  $\alpha_1 > 0$ ;  $\alpha_2 > 0$ ;  $\alpha_3 > 0$
- Логистическая функция:  $f(t) = \frac{\alpha}{1 + \beta * e^{-\gamma t}}$ , где  $\alpha > 0$ ;  $\beta > 0$ ;  $\gamma > 0$
- Первая функция Торнквиста:  $f(t) = \frac{\alpha t}{\beta + t}$ , где  $\alpha > 0$ ;  $\beta > 0$ .
- Кривая Гомперца:  $f(t) = \alpha \beta^{\gamma t}$ , где  $\alpha > 0$ ;  $\beta > 0$ ;  $\gamma > 0$ .

**КОНЕЦ ЛЕКЦИИ**

---