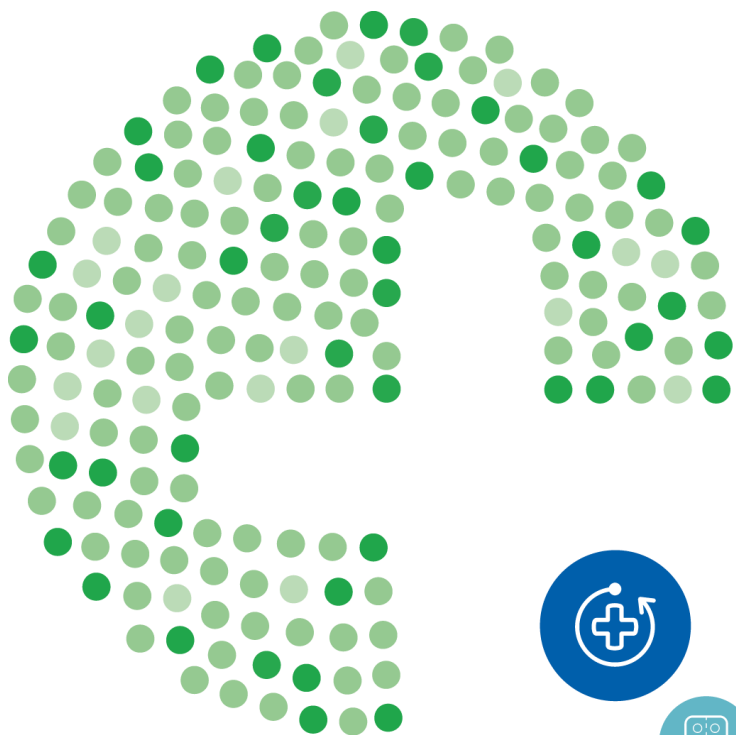


КАК «ПОДКОВАТЬ» ДАННЫЕ

Заместитель заведующего
R&D лабораторией
Музыкин М.А.

m_muzykin@polysan.ru



ICH Harmonised Tripartite Guideline Q1A (R2) Stability testing of new drug substances and products (2003)

ICH Harmonised Tripartite Guideline Q1C Stability testing of new dosage forms (1996)

ICH Harmonised Tripartite Guideline Q1E Evaluation of stability data (2003)

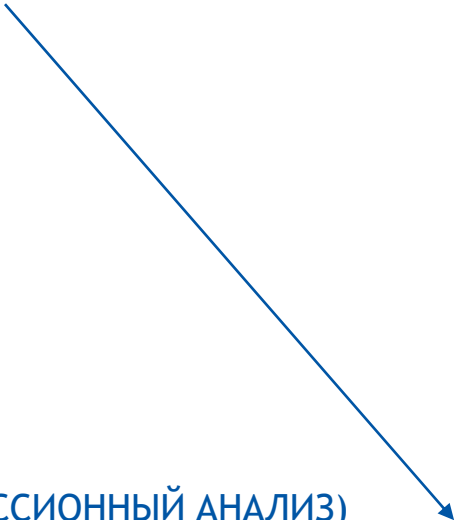
ГОСТ Р 52249-2009 Правила производства и контроля качества лекарственных средств

ОФС.1.1.0009.15 Сроки годности лекарственных средств (ГФ РФ XIII Изд., ч. 1)



Before pooling the data from several batches to estimate a retest period or shelf life, a preliminary statistical test should be performed to determine whether the regression lines from different batches have a common slope and a common time-zero intercept. **Analysis of covariance (ANCOVA)** can be employed, where time is considered the covariate, to test the differences in slopes and intercepts of the regression lines among batches. Each of these tests should be conducted using a **significance level of 0.25** to compensate for the expected low power of the design due to the relatively limited sample size in a typical formal stability study.



- СБОР ДАННЫХ - ПЛАНИРОВАНИЕ
 - ПЕРВИЧНАЯ ОБРАБОТКА ДАННЫХ
 - ГРУБЫЕ ПРОМАХИ
 - РЕЗУЛЬТАТЫ OOS/OOT
 - СТАТИСТИЧЕСКИЙ АНАЛИЗ:
 - ОЦЕНКА ВОЗМОЖНОСТИ ПРОВЕДЕНИЯ
 - АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ (РЕГРЕССИОННЫЙ АНАЛИЗ)
 - АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ (КОВАРИАЦИОННЫЙ АНАЛИЗ - ANCOVA)
- 

- Данные представляют в виде таблиц и графиков
- Числовые значения приводят с точностью, указанной в спецификации для исследования стабильности
- Приводят результаты всех измерений, полученных в контрольной точке, а не средних по точке (это позволяет оценить реальную вариативность данных)
- Число параллельных измерений должно обеспечивать достоверность результата (как правило не менее 3 в каждой точке)



Непосредственно при получении результатов в контрольной точке проводят проверку:

- на наличие грубых ошибок (промахов)
- соответствия спецификации (проверка наличия OOS)

Дополнительно следует оценить отклонение вновь полученных результатов от имеющейся тенденции (OOT)

- Если в результате измерений был получен сомнительный результат, необходим детальный анализ условий эксперимента.
- Если установили, что были нарушены стандартные условия проведения измерений, результат следует отбросить.
- Если сомнительный результат был получен без видимых причин, его допустимо отбросить, но заменить тремя новыми подряд проведенными измерениями.
- Если воспользоваться данными рекомендациями не удастся, следует прибегнуть к статистическим методам. Для этого используют критерий Граббса.

Выходы за пределы спецификации и отклонения от ожидаемой тенденции необходимо фиксировать и оценивать в ходе проведения исследований стабильности и по завершении экспериментов

Критерий OOS - спецификация

Критерии OOT - по уравнению регрессии или визуально-графически

Действия - расследование

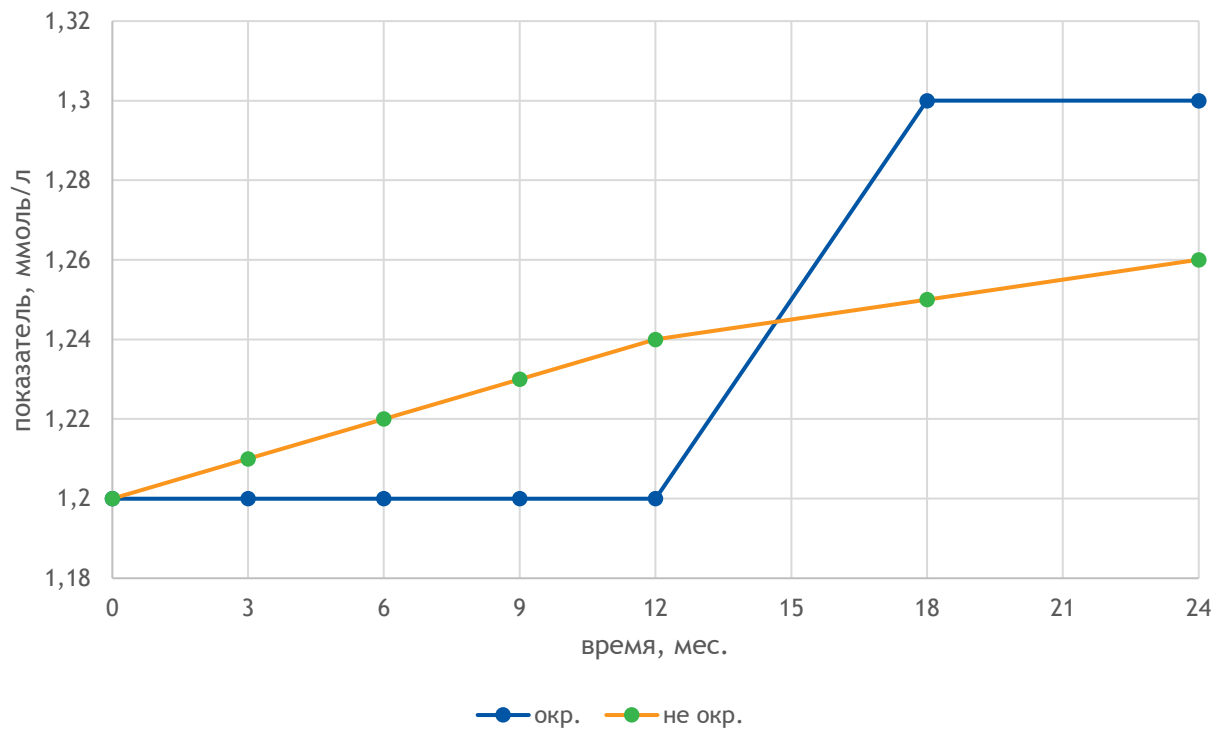
- Основной метод - регрессионный анализ
- Связь показатель/время линейна
- Уровень доверительной вероятности - 95%
- С критериями приемлемости сравнивают верхнее, нижнее или оба граничных значения в зависимости от динамики

ОЦЕНКА ВОЗМОЖНОСТИ ПРОВЕДЕНИЯ СТАТИСТИЧЕСКОГО АНАЛИЗА

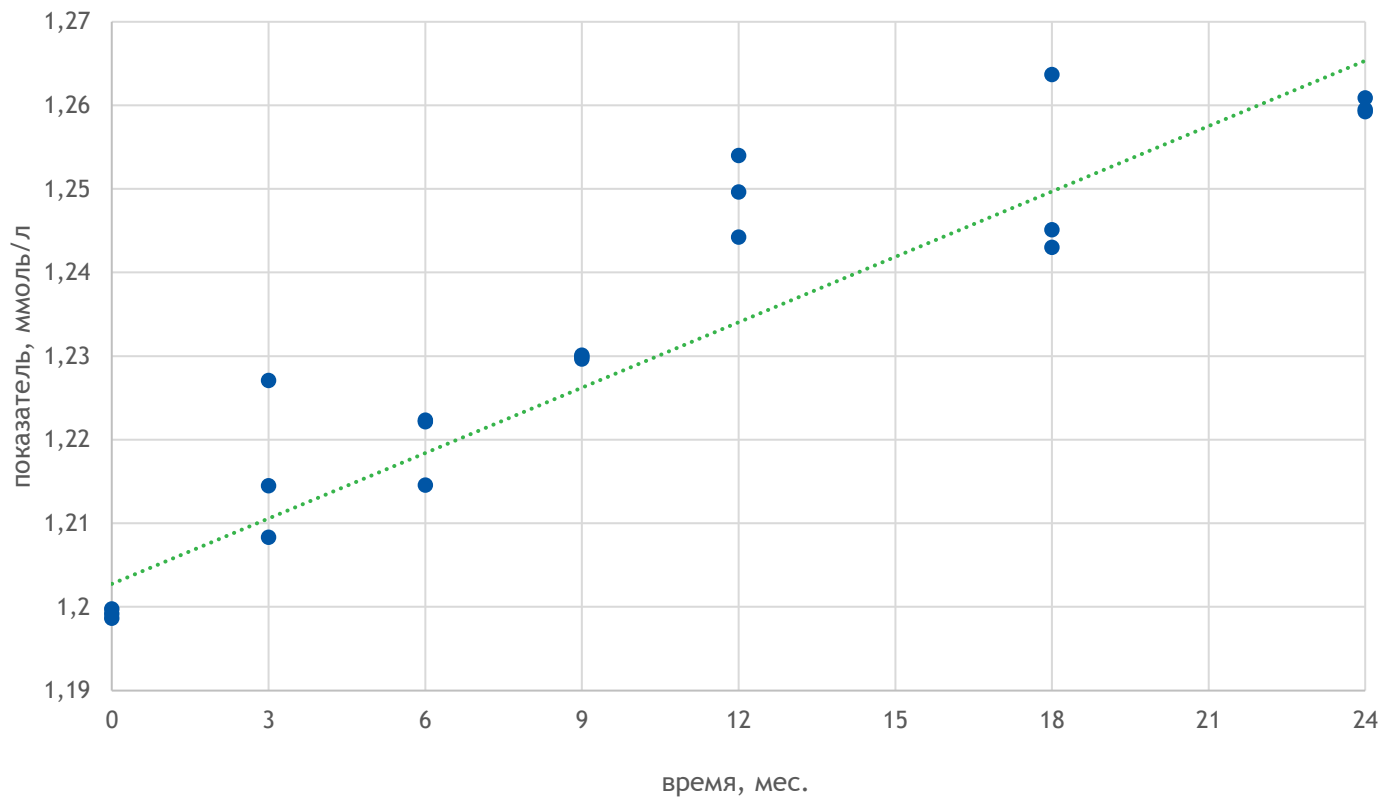
- Достаточное количество данных - не менее трёх определений на каждый параметр математической модели
- Достаточная степень дискретности - количество отличающихся значений должно превышать количество параметров модели как минимум на единицу
- Все данные должны иметь количественный характер



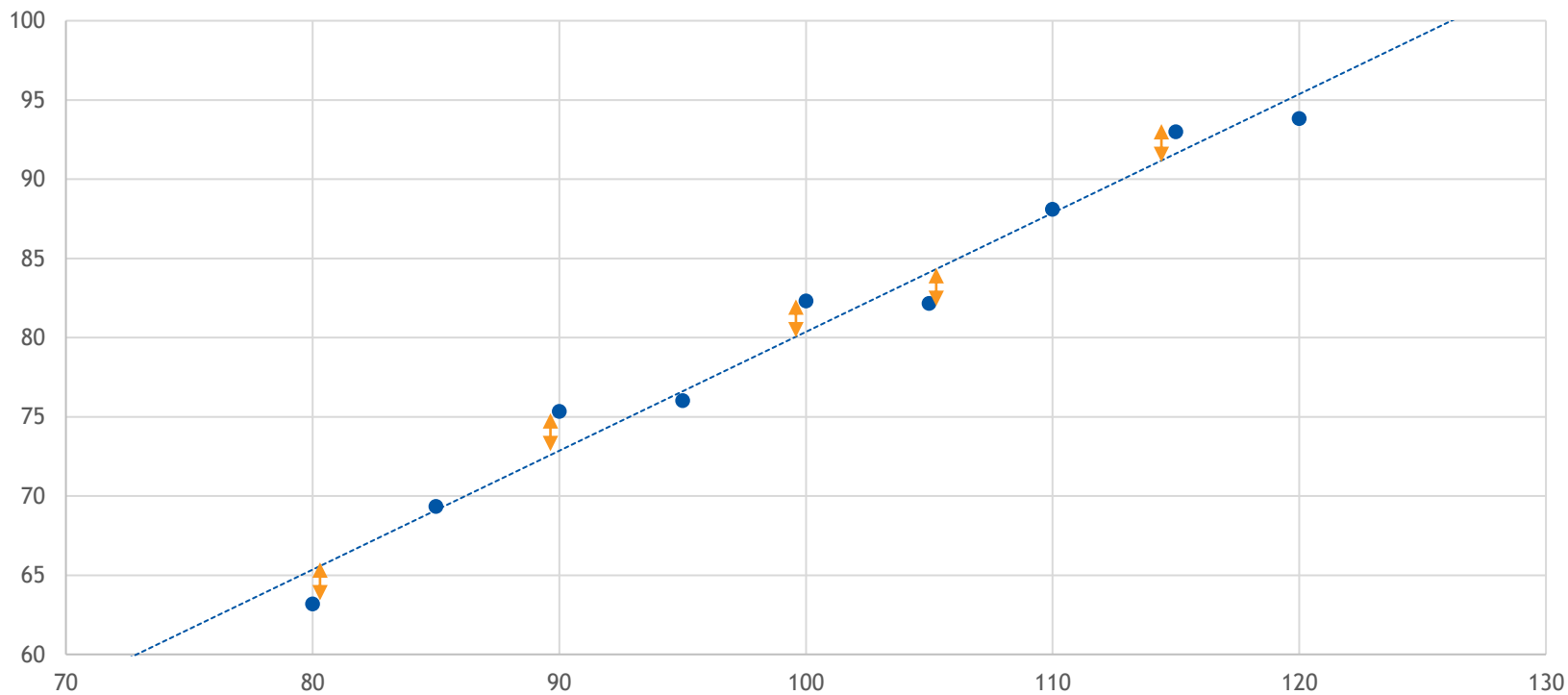
Округлять или не округлять



Средние и «всё хорошо»



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ



Сумма квадратов отклонений точек от прямой должна быть минимальной



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ

При оценке линейной зависимости находят параметры уравнения следующего вида:

$$y = a + b \cdot x$$

Коэффициенты регрессионной зависимости, в соответствии с методом наименьших квадратов, рассчитывают по формулам:

$$b = \frac{n \cdot \sum x_i \cdot y_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}$$

$$a = \frac{\sum y_i - b \cdot \sum x_i}{n}$$



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ

Остаточную дисперсию регрессионной зависимости рассчитывают по формуле:

$$s_0^2 = \frac{\sum (y_i - a - b \cdot x_i)^2}{n - 2}$$

Стандартное отклонение и доверительный интервал свободного члена регрессионной зависимости находят по формулам:

$$s_a = \sqrt{s_a^2} = \sqrt{s_0^2 \cdot \frac{\sum x_i^2}{n \cdot \sum (x_i - \bar{x})^2}}$$

$$\Delta_a = t(P, f) \cdot s_a$$

Условие незначимости:

$$|a| \leq \Delta_a$$



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ

Стандартное отклонение и доверительный интервал углового коэффициента регрессионной зависимости находят по формулам:

$$s_b = \sqrt{s_b^2} = \sqrt{s_0^2 \cdot \frac{1}{\sum(x_i - \bar{x})^2}}$$
$$\Delta_b = t(P, f) \cdot \sqrt{s_b^2}$$

Условие значимости:

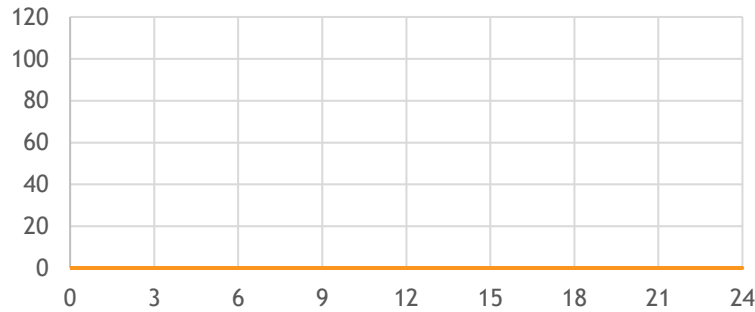
$$|b| > \Delta_b$$

Значимость говорит о наличии связи между величиной показателя и временем.

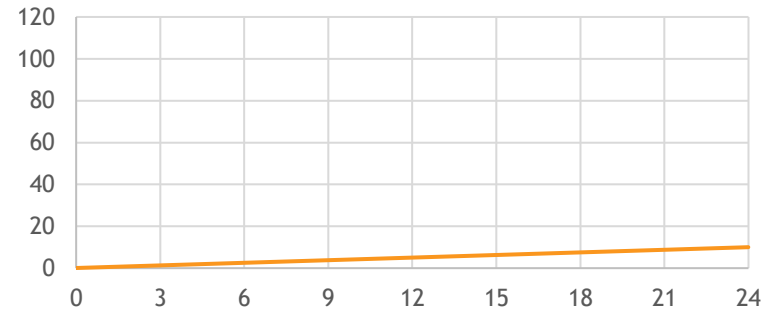


Регрессионный анализ

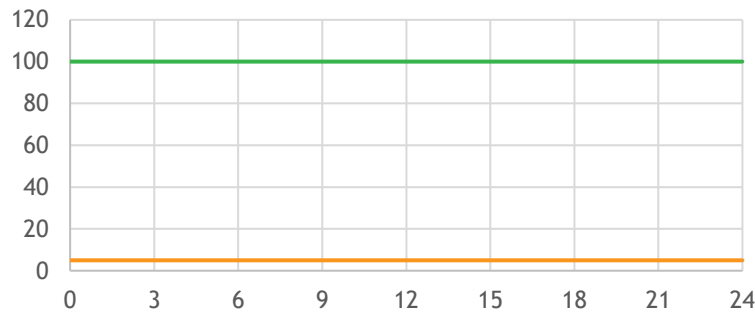
а-незначим, b - незначим



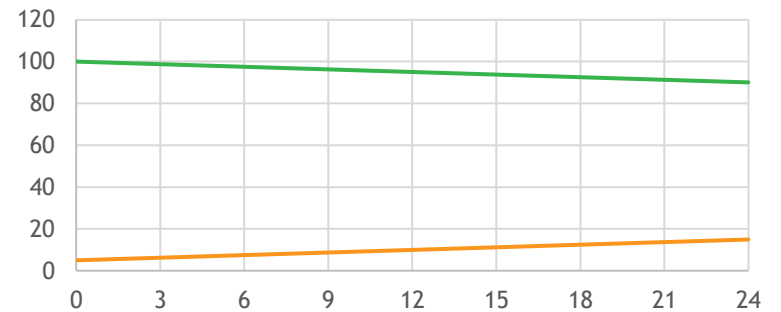
а-незначим, b - значим



а-значим, b - незначим



а-значим, b - значим



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ

a	b	Для чего характерно
незначим	незначим	Следовые количества примесей из субстанций или процесса производства, динамики нет
незначим	значим	Примесь возникает в процессе хранения, необходимо оценить её уровень, определить, позволяет ли такая динамика подтвердить срок годности
значим	незначим	Активные компоненты, содержание которых не изменяется - т.е. они стабильны в процессе хранения; примеси субстанций и/или технологического процесса, не образующиеся в процессе хранения
значим	значим	Активные компоненты, содержание которых меняется или примеси субстанций и/или технологического процесса, содержание которых продолжает расти в процессе хранения. В обоих необходимо определить, позволяет ли такая динамика подтвердить срок годности

Значимость коэффициентов при оценке данных по стабильности позволяет сделать предварительные выводы о процессах, протекающих в процессе хранения или об их отсутствии



Оценка срока годности (экстраполяция):

При прогнозировании нельзя просто продлить линию регрессии до интересующей нас временной отсечки. Любой прогноз характеризует его достоверность. Как и для коэффициентов регрессионной зависимости для самой линии регрессии необходимо построить **доверительный интервал**.

Тип интервала определяется входящей информацией:

- **односторонний** - если заранее известно направление тренда (разрушение активного компонента, рост примесей)
- **двусторонний** - когда направление заранее не определено



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ

Односторонний доверительный интервал представляют в соответствии с формулой:

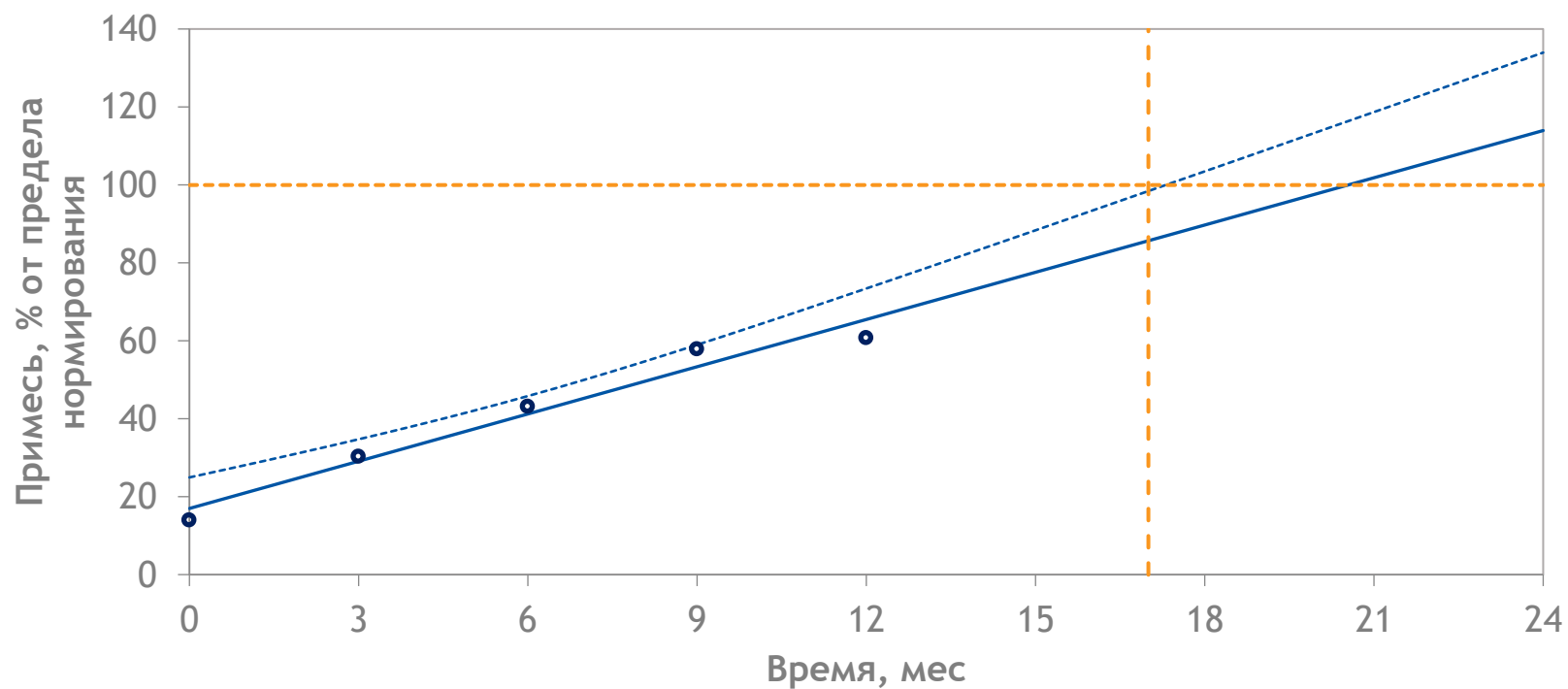
$$y_i + \text{sgn}(b) \cdot \Delta_{y_i} = a + bx + \text{sgn}(b) \cdot s_0 \cdot t(\alpha, n - 2) \cdot \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_{xx}}}$$

Двусторонний доверительный интервал определяется как:

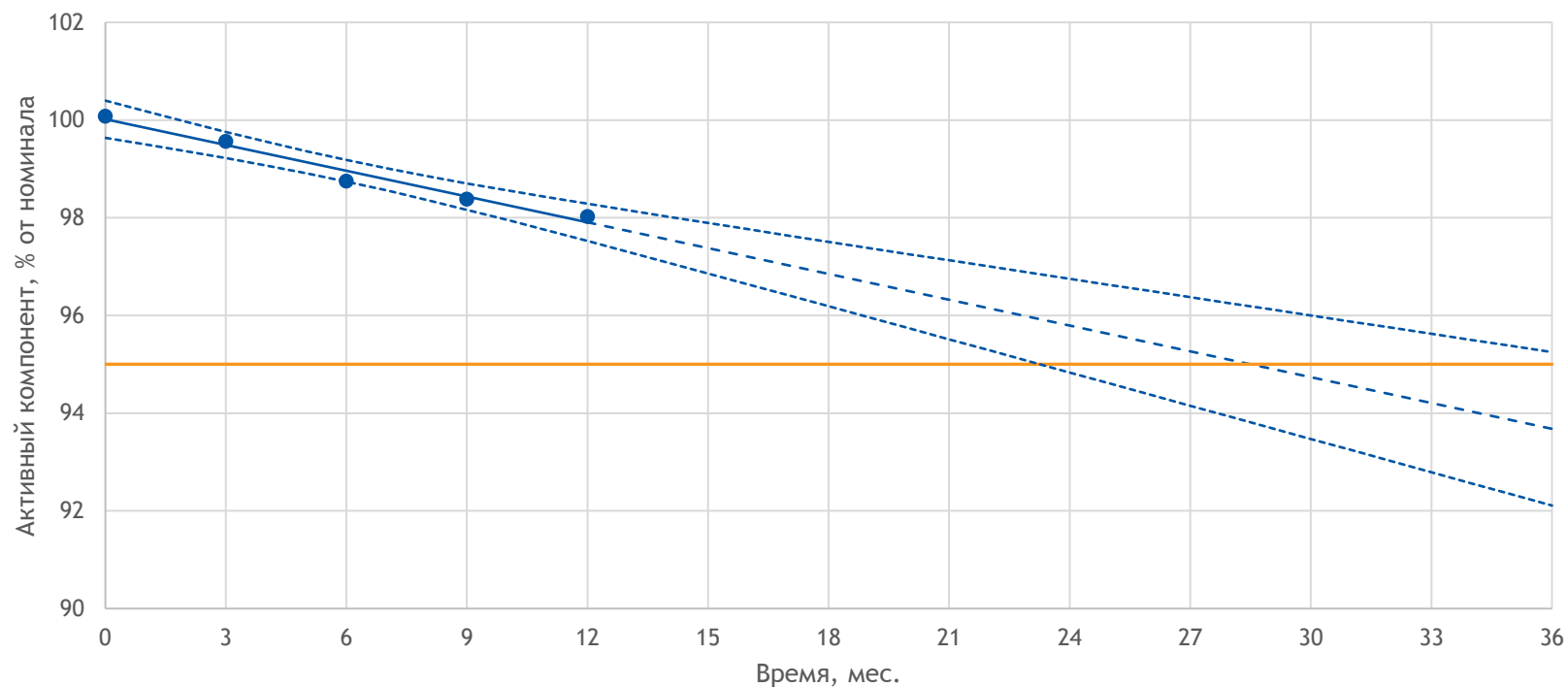
$$y_i \pm \Delta_{y_i} = a + bx \pm s_0 \cdot t(\alpha, n - 2) \cdot \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_{xx}}}$$



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ



АНАЛИЗ ДАННЫХ ДЛЯ ОДНОЙ СЕРИИ РЕГРЕССИОННЫЙ АНАЛИЗ



Экстраполировать можно на срок до $2 \cdot x_{\max}$, но не более чем на 12 мес. (ICH Q1E)



АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

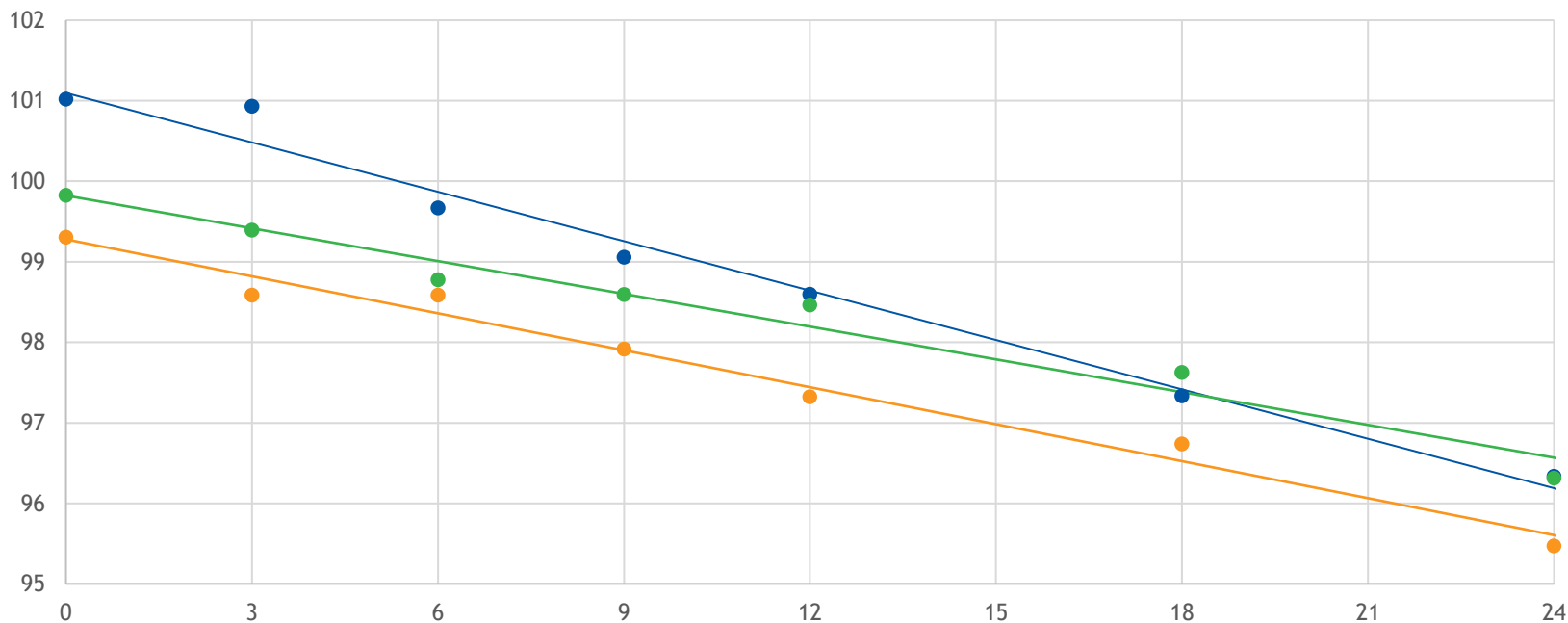
Суть метода: оценка равенства нескольких векторных случайных величин (по сути это смесь дисперсионного и регрессионного анализов)

Порядок действий:

- На первом этапе проверяем равенство наклонов - строим модель с параллельными «индивидуальными» трендами и оцениваем насколько такая модель окажется менее точной по сравнению с индивидуальными трендами.
- На втором этапе проверяем возможность построения единого тренда для всех совокупностей, оцениваем насколько эта модель окажется менее точной по сравнению с моделью с параллельными трендами

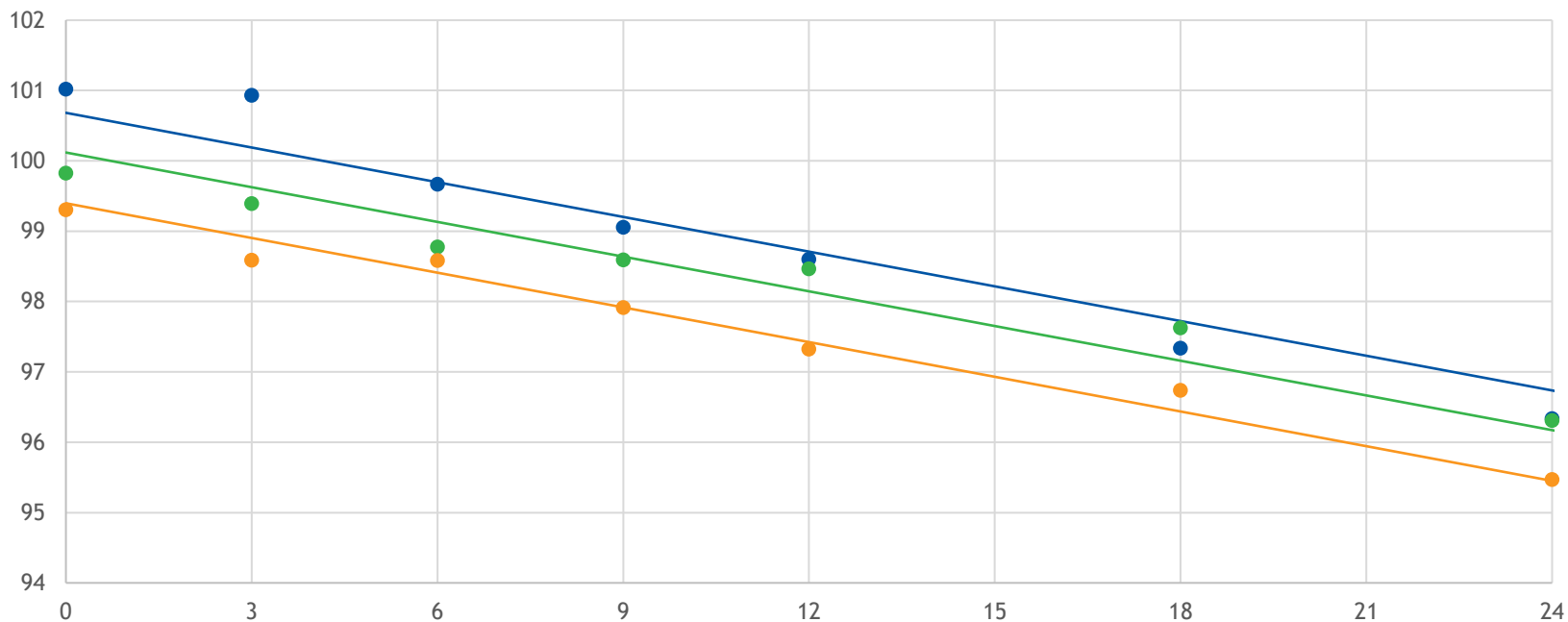


Шаг 0 - Индивидуальные тренды



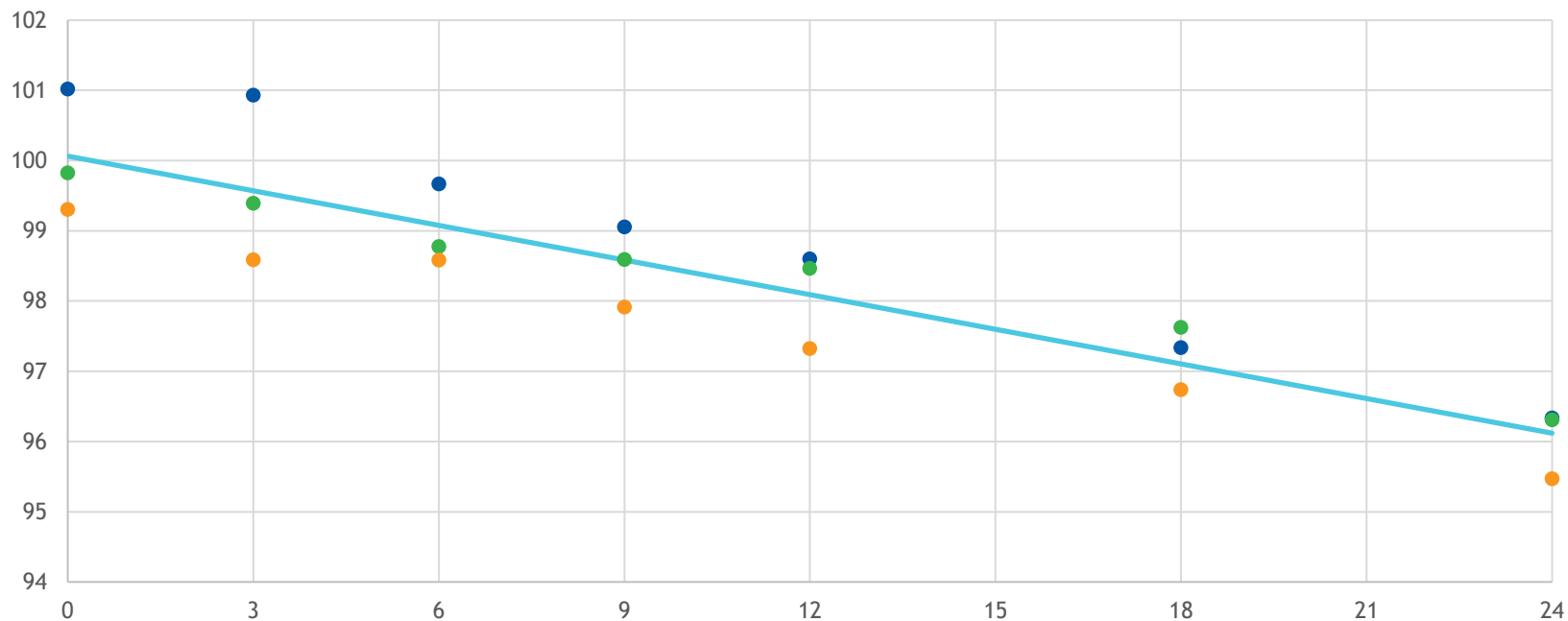
АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Шаг 1 - Общий наклон



АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Шаг 2 - Общий тренд



Проверка гипотезы о равенстве наклонов

N - общее число точек, K - общее число серии,

i - номер серии, j - номер временной точки в серии

Вспомогательные суммы:

Индивидуальные тренды	Параллельные тренды
$SS_{xx}(i) = \sum (x_i - \bar{x})^2$	$SS_{xx}(W) = \sum_{i=1}^K S_{xx}(i)$
$SS_{yy}(i) = \sum (y_i - \bar{y})^2$	$SS_{yy}(W) = \sum_{i=1}^K S_{yy}(i)$
$SS_{xy}(i) = \sum (x_i - \bar{x})(y_i - \bar{y})$	$SS_{xy}(W) = \sum_{i=1}^K S_{xy}(i)$

АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Остаточные суммы квадратов отклонений для каждой (i) серии составят:

$$SSE(i) = SS_{yy}(i) - \frac{(SS_{xy}(i))^2}{SS_{xx}(i)}$$

Совместная сумма квадратов составит:

$$SSE = \sum_{i=1}^K SSE(i)$$

Сумма квадратов отклонений, обусловленная регрессией составит:

$$SSW = SS_{yy}(W) - \frac{(SS_{xy}(W))^2}{SS_{xx}(W)}$$

Сумма квадратов отклонений, обусловленная разницей в наклонах составит:

$$SS_b = SSW - SSE$$

Нулевая гипотеза о равенстве наклонов отвергается при условии:

$$F = \frac{SS_b}{K-1} \cdot \frac{N-2K}{SSE} > F(\alpha, K-1, N-2K)$$



АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Величина тангенса для модели с общим углом составит:

$$b = \frac{SS_{xy}(W)}{SS_{xx}(W)}$$

Условие значимости тангенса:

$$|b| \leq t(\alpha, n - 2) \cdot \sqrt{\frac{SSW}{(N - 2K) \cdot SS_{xx}(W)}}$$



Проверка гипотезы о равенстве пересечений

Вспомогательные расчёты:

Средние	Суммы
$\bar{x} = \frac{1}{N} \sum_{i=1}^K \sum_{j=1}^{n_i} x_{ij}$	$SS_{xx}(T) = \sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$
$\bar{y} = \frac{1}{N} \sum_{i=1}^K \sum_{j=1}^{n_i} y_{ij}$	$SS_{yy}(T) = \sum_{i=1}^K \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$
	$SS_{xy}(T) = \sum_{i=1}^K \sum_{j=1}^{n_i} (x_{ij} - \bar{x})(y_{ij} - \bar{y})$

АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Общая сумма квадратов отклонений составит:

$$SST = SS_{yy}(T) - \frac{\left(SS_{xy}(T)\right)^2}{SS_{xx}(T)}$$

Сумма квадратов отклонений, обусловленная разницей в точках пересечения составит:

$$SS_a = SST - SSW$$

Нулевая гипотеза о равенстве точек пересечения отвергается при условии:

$$F = \frac{SS_a}{K - 1} \cdot \frac{N - K - 1}{SSW} > F(\alpha, K - 1, N - K - 1)$$



АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Величина тангенса для обобщенной модели составит:

$$b = \frac{SS_{xy}(T)}{SS_{xx}(T)}$$

Условие значимости тангенса:

$$|b| \leq t(\alpha, n - 2) \cdot \sqrt{\frac{SST}{(N - 2) \cdot SS_{xx}(T)}}$$



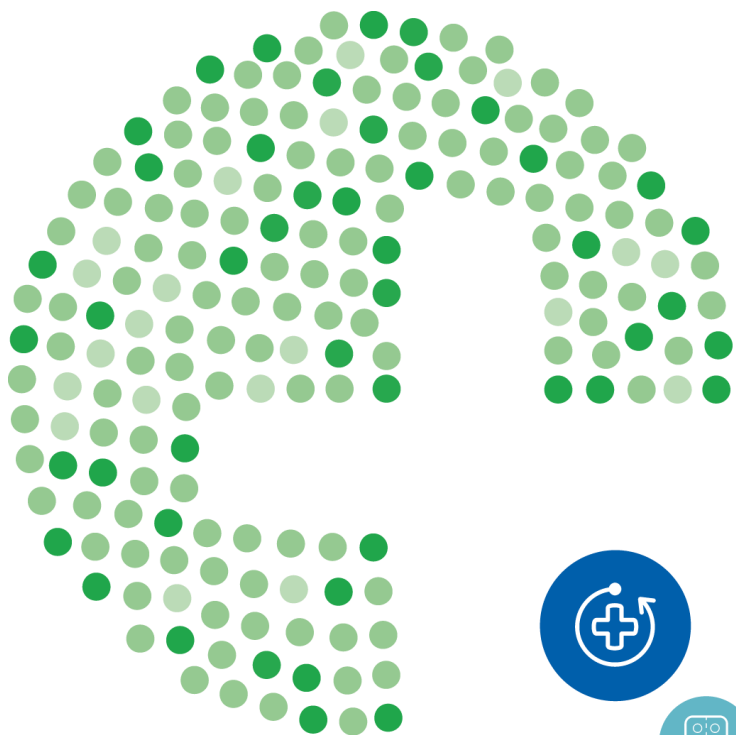
АНАЛИЗ ДЛЯ НЕСКОЛЬКИХ СЕРИЙ КОВАРИАЦИОННЫЙ АНАЛИЗ

Параметры расчёта доверительного интервала для нескольких серий

$$y_i \pm \Delta_{y_i} = a + bx \pm s_0 \cdot t(\alpha, n - 2) \cdot \sqrt{\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_{xx}}}$$

Показатель	Результат ковариационного анализа	
	равны только наклоны	равны наклоны и точки пересечения
a	$a_i = \bar{y}_i - b \cdot \bar{x}_i$	$a = \bar{y} - b \cdot \bar{x}$
b	$\frac{SS_{xy}(W)}{SS_{xx}(W)}$	$\frac{SS_{xy}(T)}{SS_{xx}(T)}$
\bar{x}	$\frac{1}{K} \sum_{i=1}^K \frac{\sum_{j=1}^{n_i} x_j}{n_i}$	$\frac{1}{N} \sum_{i=1}^K \sum_{j=1}^{n_i} x_{ij}$
n	N	
s_0	$\sqrt{\frac{SSW}{N - 2K}}$	$\sqrt{\frac{SST}{N - 2}}$
SS_{xx}	$SS_{xx}(W)$	$SS_{xx}(T)$





БЛАГОДАРЮ ЗА ВНИМАНИЕ!



наука
производство
продвижение
ЧЕЛОВЕК

НАУЧНО-ТЕХНОЛОГИЧЕСКАЯ
ФАРМАЦЕВТИЧЕСКАЯ
ФИРМА «ПОЛИСАН»

www.polysan.ru