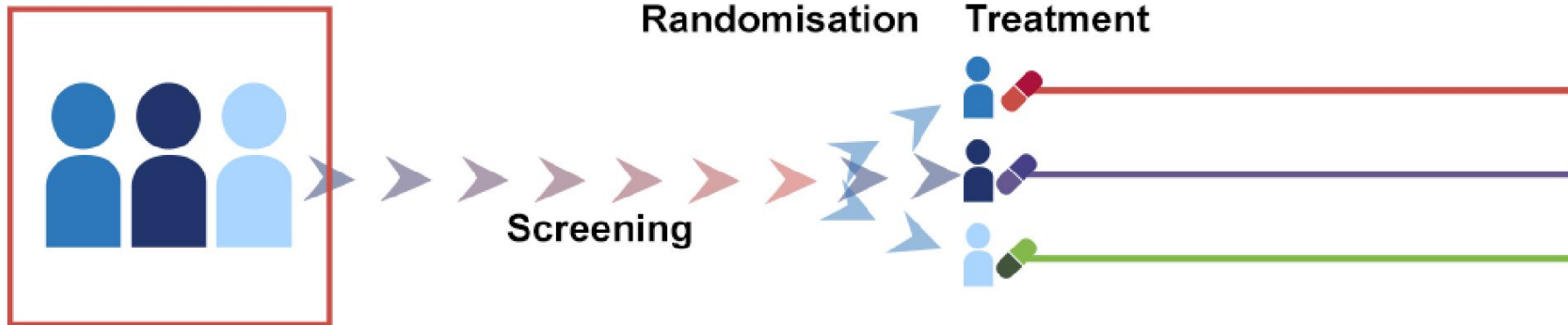


# Представление результатов клинических исследований в зависимости от типа гипотезы

- Заместитель директора по контролю качества, ООО «Статэндокс»
- Доцент кафедры профилактической и семейной медицины ФГБОУ ВО «Уральский государственный медицинский университет» МЗ РФ
- Преподавание статистики и эпидемиологии – 15 лет
- Опыт в клинических исследованиях – 13 лет
- Статистическое сопровождение КИ от концепции до финального отчета – более 500 проектов

# ДИЗАЙНЫ КЛИНИЧЕСКОГО ИССЛЕДОВАНИЯ

## Parallel Trial



## Cross-over Trial



Treatment A



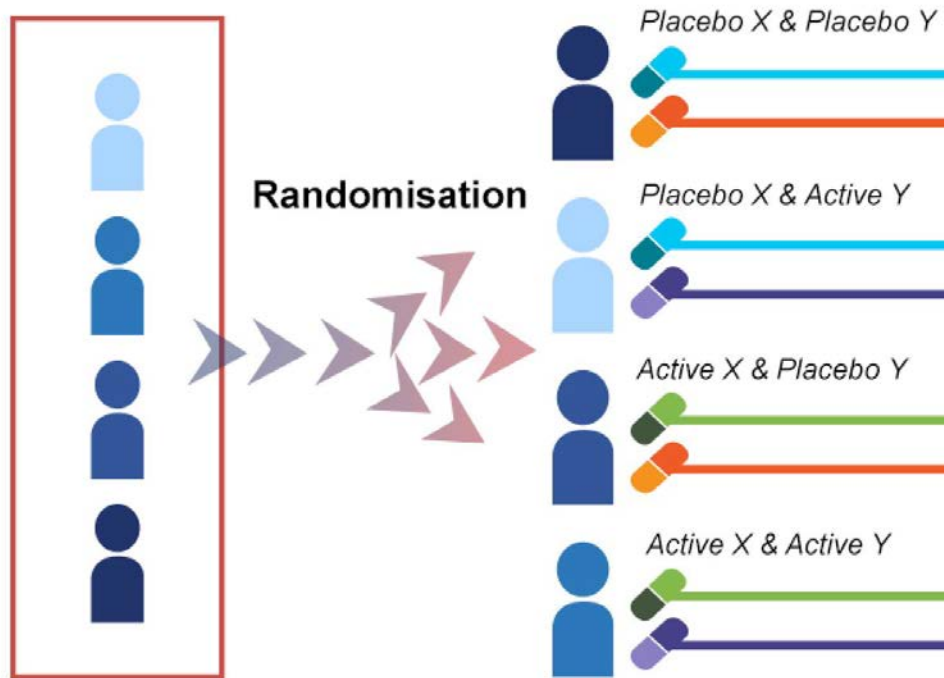
Treatment B



Washout Period



## 2x2 Factorial design



# ДИЗАЙН ПО ТИПУ ТЕСТИРУЕМОЙ ГИПОТЕЗЫ

- *S.-C. Chow et al. Sample Size Calculations in Clinical Research (3<sup>rd</sup> Ed.), 2017:*
  - Сравнение средних:  $\varepsilon = \mu_2 - \mu_1$
  - Сравнение долей:  $\varepsilon = p_1 - p_2$
  - Сравнение времени до наступления события (hazard ratios):  $\varepsilon = \lambda_1 - \lambda_2$
- **Предполагается, что положительное значение  $\varepsilon$  соответствует лучшему исходу, а отрицательное – худшему исходу в **группе изучаемого препарата по сравнению с контрольной группой****



Гипотеза равенства (equality):

$$H_0 : \varepsilon = 0 \quad \text{versus} \quad H_a : \varepsilon \neq 0.$$

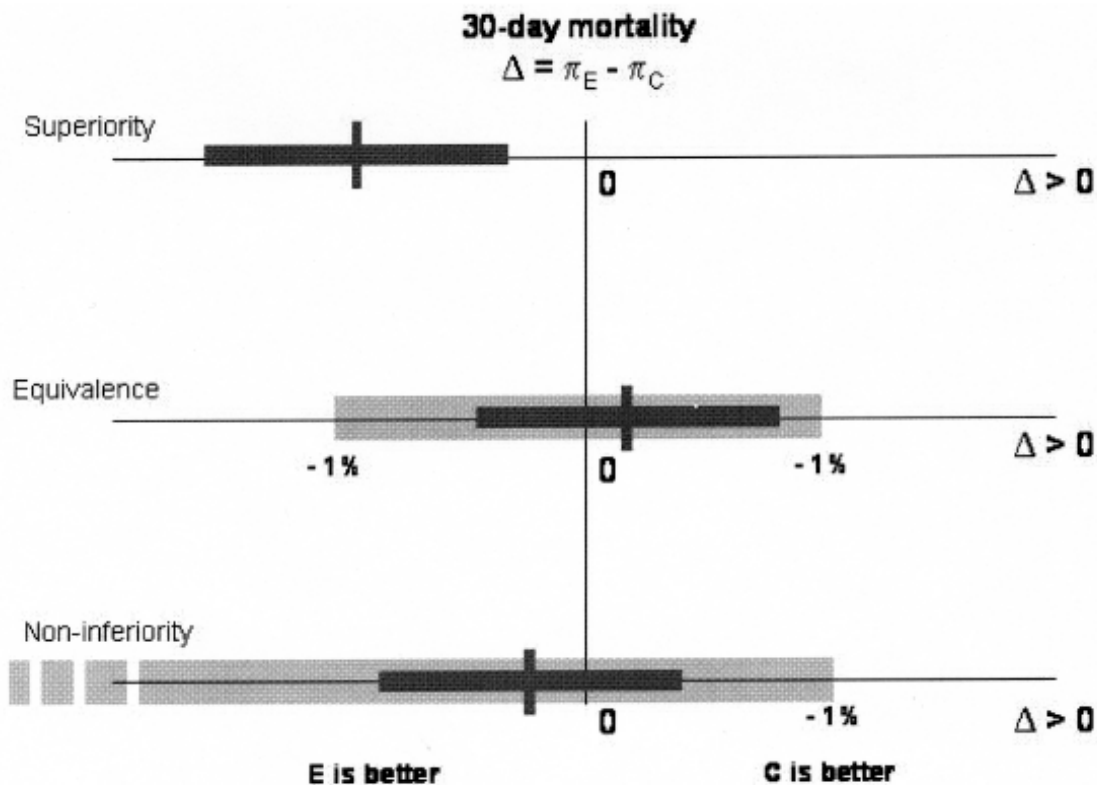
Гипотеза превосходства / не меньшей эффективности (superiority / non-inferiority):

$$H_0 : \varepsilon \leq \delta \quad \text{versus} \quad H_a : \varepsilon > \delta,$$

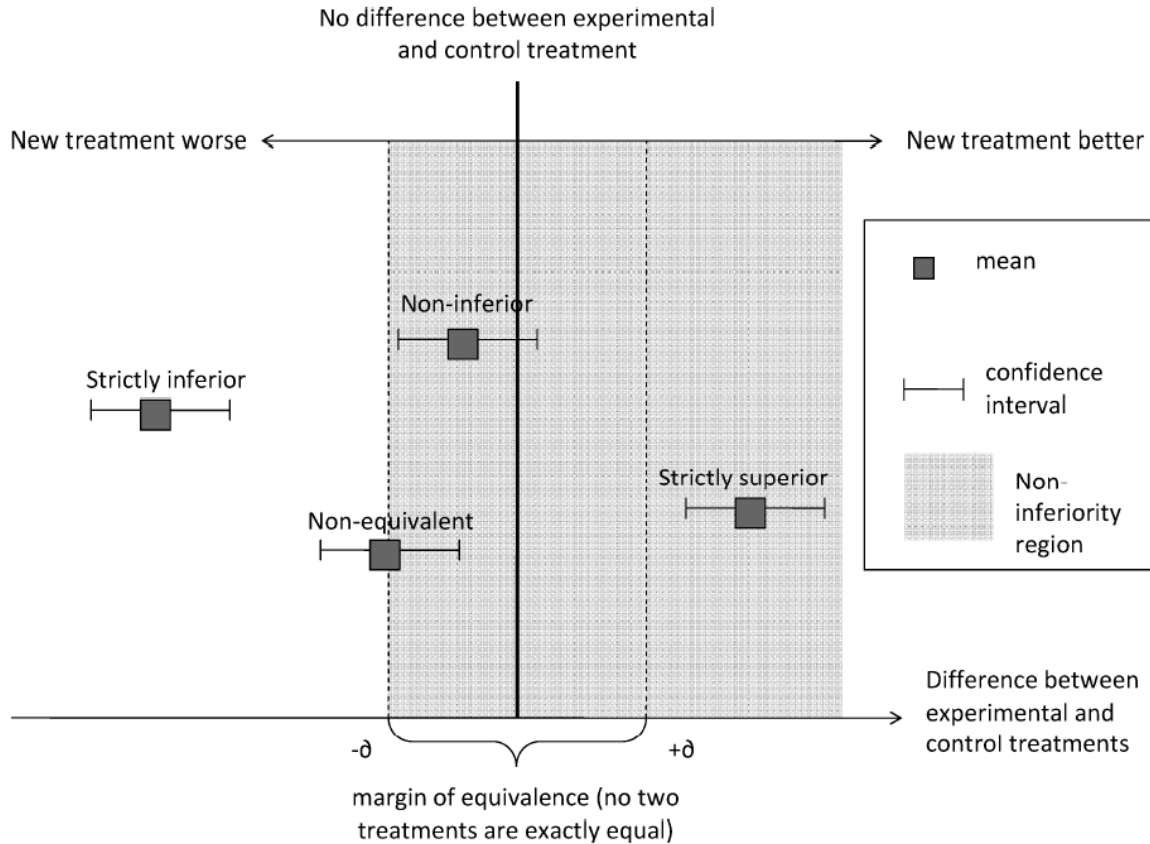
Гипотеза эквивалентности:

$$H_0 : |\varepsilon| \geq \delta \quad \text{versus} \quad H_a : |\varepsilon| < \delta.$$

# Статистические гипотезы

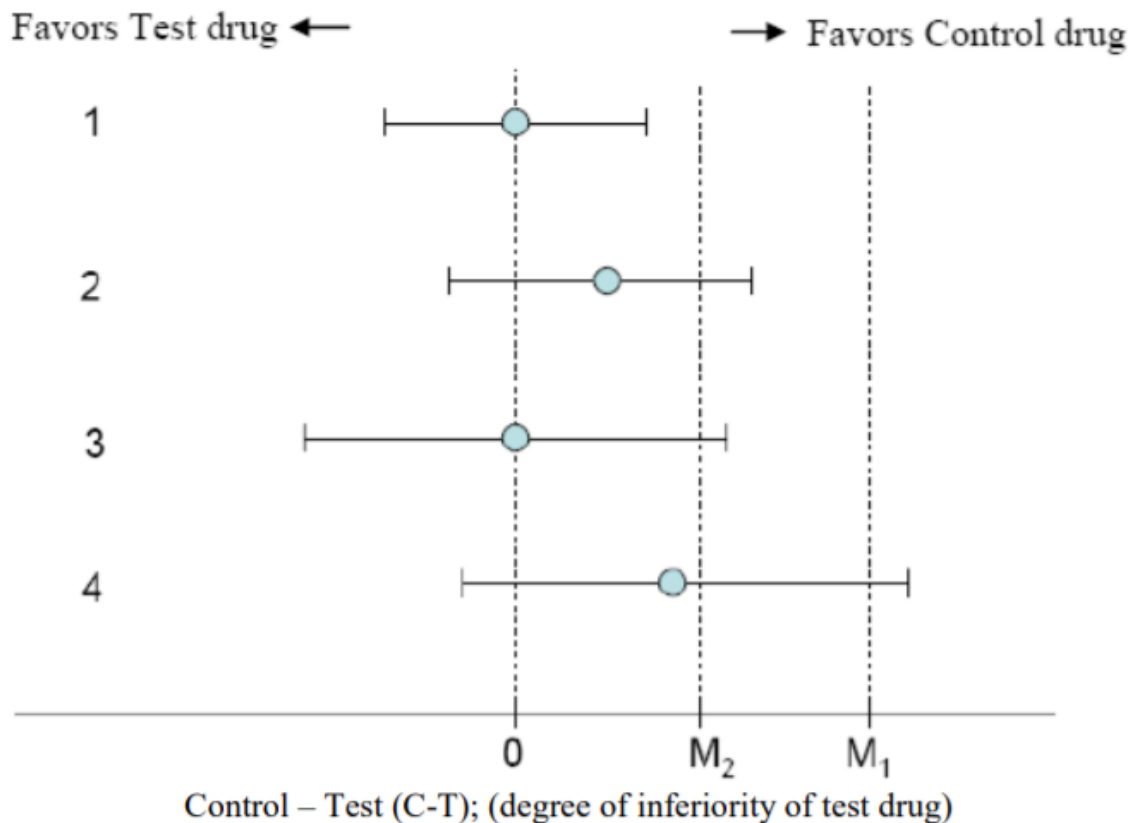


# Статистические гипотезы

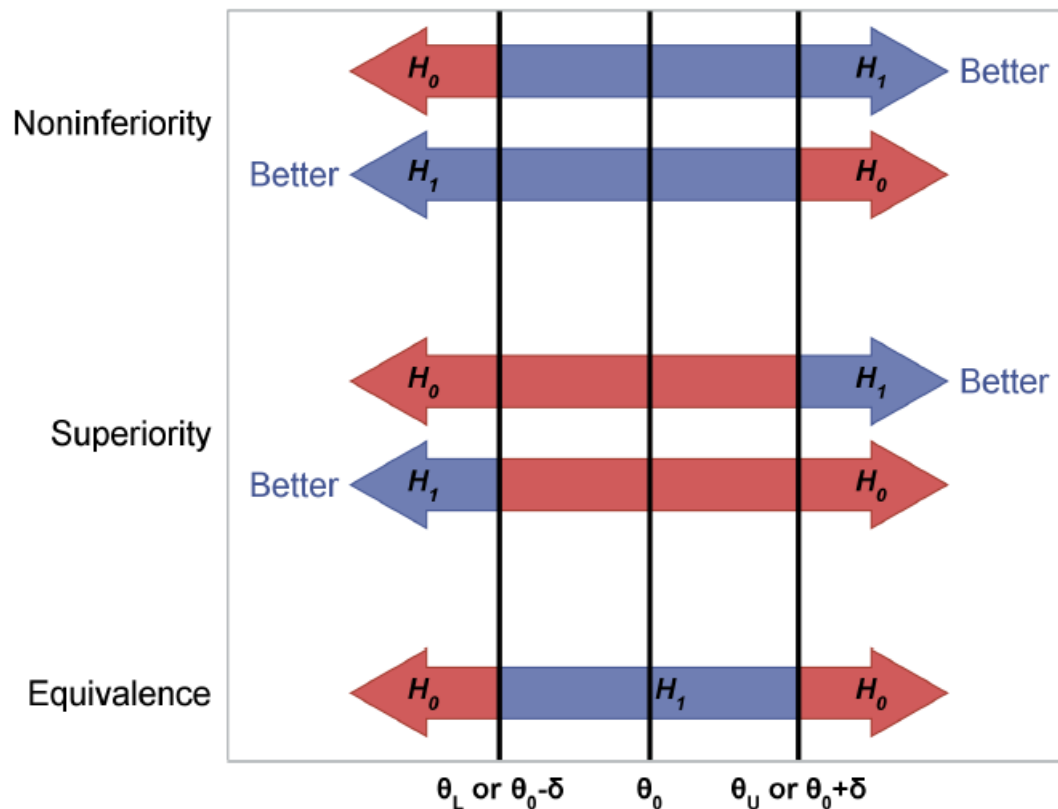


Jeremy Howick. Questioning the Methodologic Superiority of 'Placebo' Over 'Active' Controlled Trials *The American Journal of Bioethics*, 9(9): 34–48, 2009

# Статистические гипотезы



# Статистические гипотезы



- Необходимо четкое, недвусмысленное и однозначное определение следующих положений в протоколе либо плане статистического анализа:
  - Как производится расчет разницы между группами (*опытная – контрольная vs контрольная – опытная*)
  - Клинический смысл разницы (*лучше / хуже ≠ больше / меньше*)
  - Статистическая гипотеза описательно и в математической нотации

**Нулевая гипотеза** состоит в том, что исследуемый препарат ХХХ отличается по эффективности более чем на 0,8 см по шкале ВАШ от препарата сравнения УУУ.

$$H_0: m_A - m_B < -0,8 \text{ или } m_A - m_B > +0,8$$

**Альтернативная гипотеза** состоит в том, что исследуемый препарат ХХХ отличается по эффективности не более чем на 0,8 см по шкале ВАШ от препарата сравнения УУУ.

$$H_A: -0,8 \leq m_A - m_B \leq +0,8$$

Оценка сопоставимости препаратов будет проводиться при помощи двух односторонних тестов (методика TOST, two one-sided tests) при  $\alpha = 0,05$ .

$$H_0 : \delta_1 > \mu_T - \mu_R > \delta_2,$$

$$H_A : \delta_1 \leq \mu_T - \mu_R \leq \delta_2,$$

$$H'_{01} : \mu_T - \mu_R < \delta_1$$

$$H'_{A1} : \mu_T - \mu_R \geq \delta_1$$

и

$$H'_{02} : \mu_T - \mu_R > \delta_2$$

$$H'_{A2} : \mu_T - \mu_R \leq \delta_2$$



# УРОВЕНЬ ЗНАЧИМОСТИ И НАПРАВЛЕННОСТЬ ТЕСТА

- Выявление различий (equality testing):
  - $\alpha = 0,05$
  - 95% доверительный интервал
  - Односторонний или двусторонний вариант
- Превосходство или не меньшая эффективность:
  - Одностороннее тестирование гипотезы при  $\alpha = 0,025$
  - Двусторонний 95% доверительный интервал (*эквивалентно одностороннему 97,5% доверительному интервалу*)
- Эквивалентность:
  - Два односторонних теста при  $\alpha = 0,05$
  - 90% двусторонний доверительный интервал

# НАБОР ДАННЫХ ДЛЯ АНАЛИЗА

## Набор данных для анализа (analysis set)

- **Intention-to-treat:** стандартный набор данных для анализа в исследованиях различий и превосходства
- **Per protocol:** рекомендуемый набор данных для анализа в исследованиях не меньшей эффективности
- При адекватном методе восполнения данных в ИТТ популяции (например, MLE) рекомендуется использовать **комбинированный PP/ИТТ подход** (*Matilde Sanchez M, Chen X. Choosing the analysis population in non-inferiority studies: per protocol or intent-to-treat. Stat Med. 2006 Apr 15;25(7):1169-81*)

# СТАТИСТИЧЕСКИЙ АНАЛИЗ В ЗАВИСИМОСТИ ОТ ГИПОТЕЗЫ

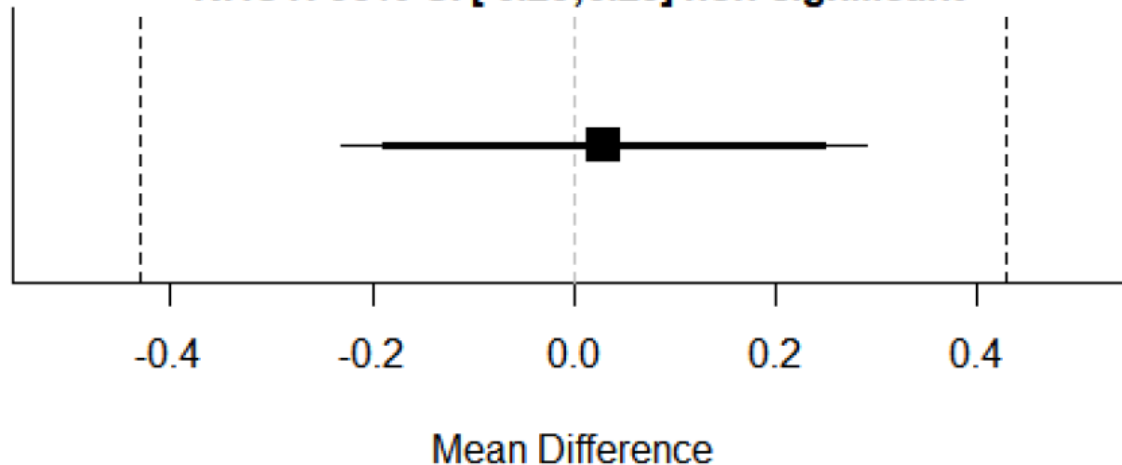
# Постановка проблемы

- Большинство методов «классической» статистики направлены на выявление различий (equality testing)
- Необходимость в оценке эквивалентности, превосходства и не меньшей эффективности предполагает, что нулевая гипотеза содержит в себе соответствующую границу, соответственно, статистические тесты должны учитывать эту границу

- Поправка на оценку ненулевой ожидаемой величины различий были внесены в большинство статистических тестов и реализованы во многих статистических программах (NCSS, R, SAS, Stata и др.):
  - Разница средних – тесты на основе t-распределения
  - Разница долей – тесты на основе z-распределения (большие выборки) либо точного теста Фишера (малые выборки, необходим пересчет разницы в отношение шансов)
  - Разница медиан – тесты на основе z-аппроксимации распределения для критерия Манна-Уитни
  - Разница для времени до наступления события – тесты на основе модели пропорционального риска по Коксу

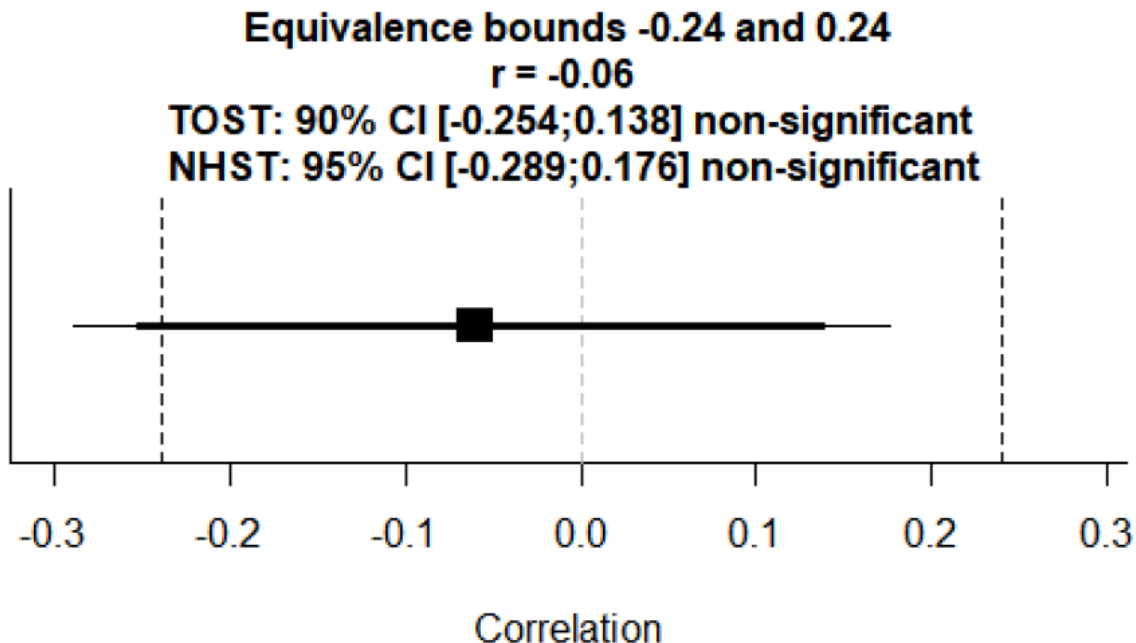
# Пример в R (пакет TOSTER) - средние

**Equivalence bounds -0.429 and 0.429**  
**Mean difference = 0.03**  
**TOST: 90% CI [-0.188;0.248] significant**  
**NHST: 95% CI [-0.23;0.29] non-significant**





# Пример в R (пакет TOSTER) – коэффициент корреляции



# РАСЧЕТ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ

- Стандартные доверительные интервалы для разницы (на основе  $t$ -распределения) – подходят только для параллельных дизайнов и однофакторного анализа
- Доверительные интервалы для least square means [LSM] (средних, полученных методом наименьших квадратов) – расчет на основе mean square errors [MSE] в моделях дисперсионного анализа ANOVA

- Точные доверительные интервалы для разницы долей по методу Newcombe&Wilson, Agresti-Caffo или аналогичным
- Доверительные интервалы на основе аппроксимации z-распределения
- Оценка доверительных интервалов в рамках обобщенной линейной модели (GLM)

- Расчет доверительного интервала для относительного риска (hazard ratio) путем обратного логарифмического преобразования на основе z-аппроксимации
- Расчет доверительного интервала методом Вальда
- Расчет доверительного интервала методом Пето

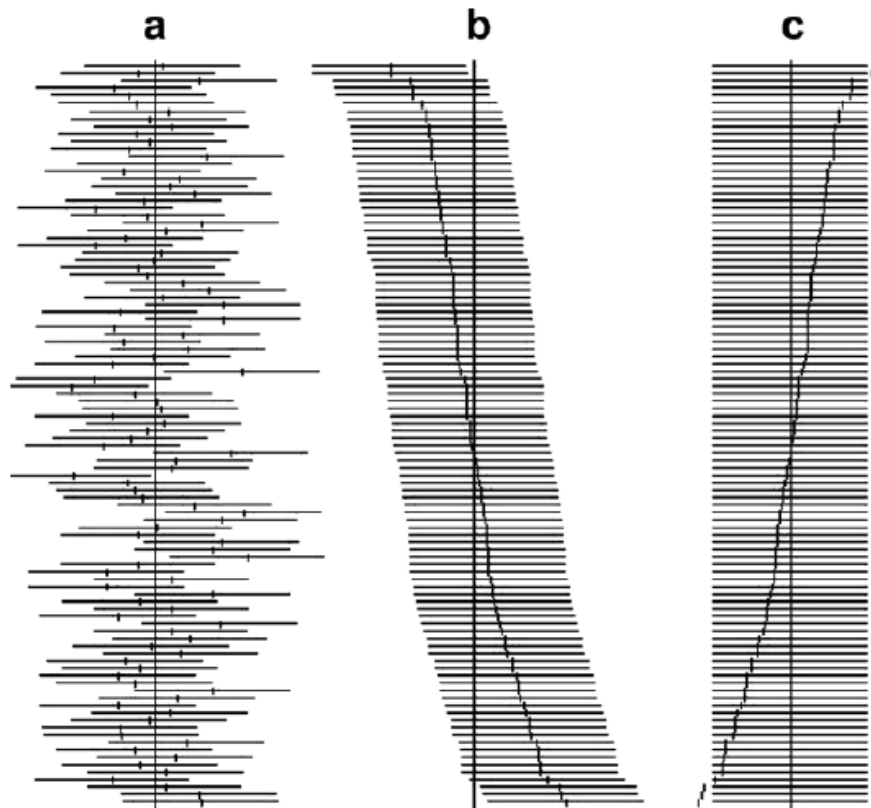
- Для расчета доверительных интервалов для медианы и других квантильных оценок не существует прямых методов
- Аппроксимация не всегда дает адекватные результаты, особенно при малых выборках
- В сложных статистических моделях оценки доверительных интервалов подвержены смещениям

- Resampling – метод математического моделирования выборок, аналогичных исходной, на основе ее параметров (объем, среднее, характер распределения и т.п.)
- Наиболее распространенным методом для resampling является бутстрэппинг (bootstrapping), основным преимуществом которого является отсутствие каких-либо специальных допущений

- Методы resampling позволяют получить любое количество выборок с характеристиками, аналогичными исходной
- Для расчета доверительных интервалов для выборочных статистик обычно используют 10000 смоделированных выборок, что позволяет получить высоко надежную оценку эффекта
- В ряде случаев смоделированные методом bootstrap доверительные интервалы могут использоваться для проверки гипотез независимо от статистических критериев



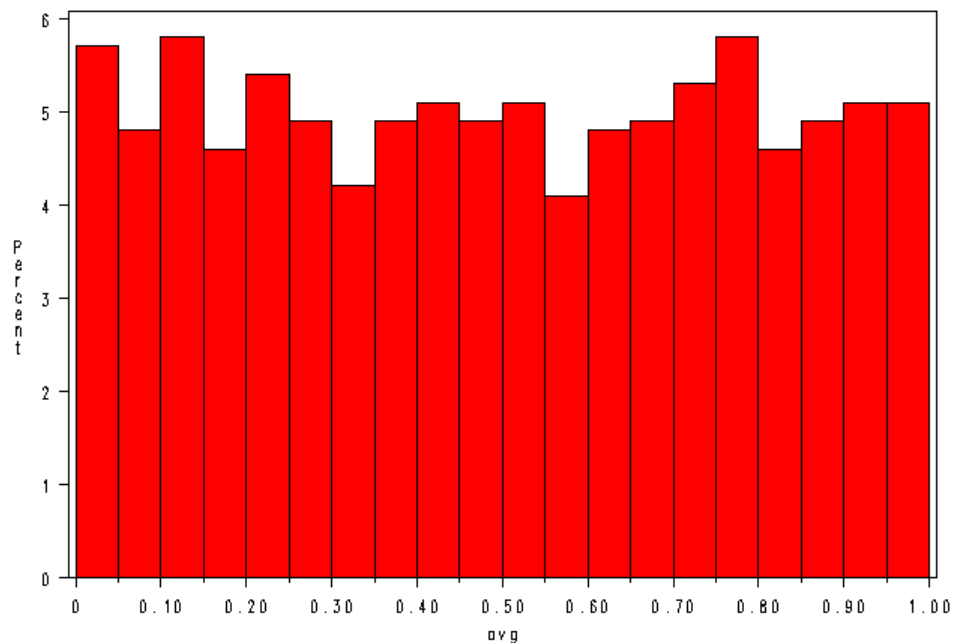
# Пример – моделирование средних



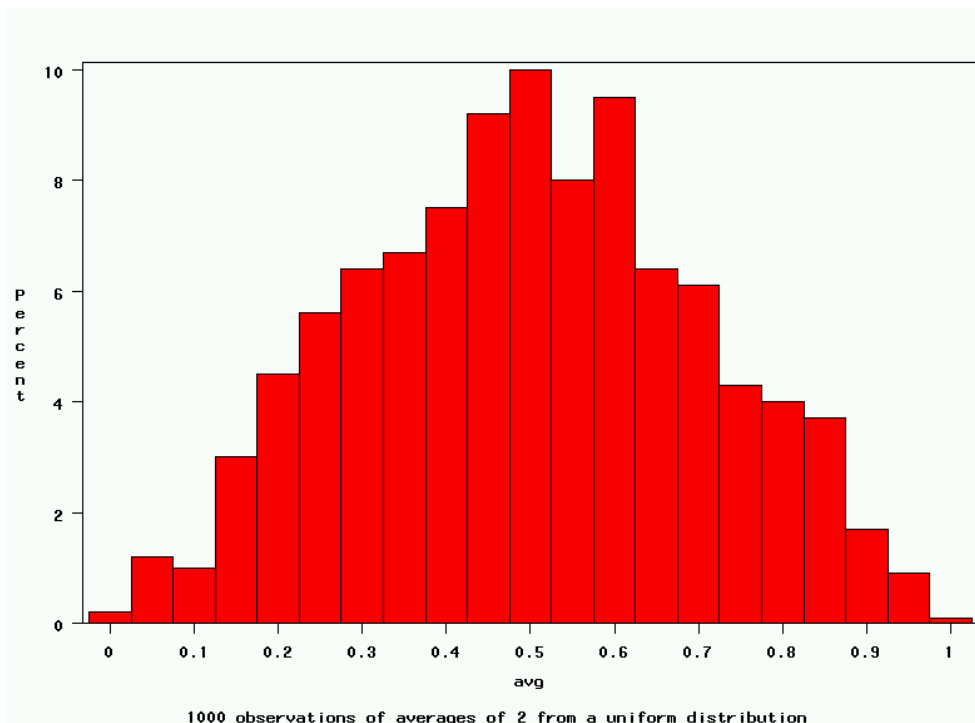
Erik Christensen. Methodology of superiority vs. equivalence trials and non-inferiority trials. *Journal of Hepatology* 46 (2007) 947–954

# Равномерное распределение величин от 0 до 1 со средним 0,5 (исходное распределение)

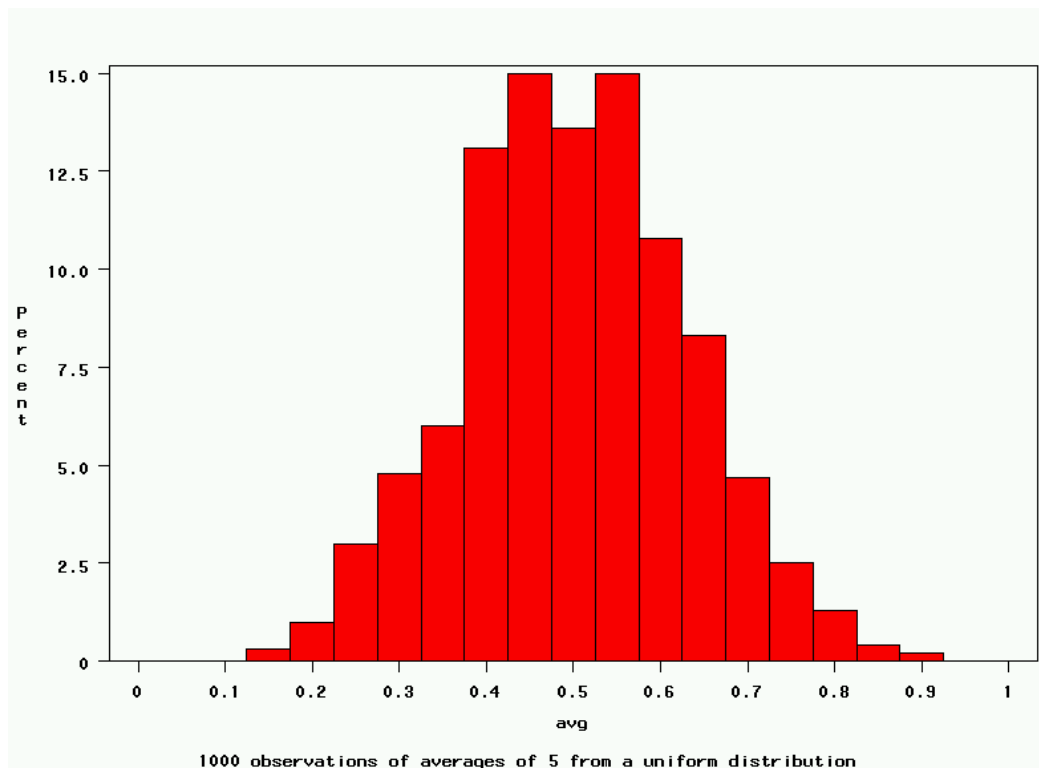
1000 observations of averages of 1 from a uniform dist



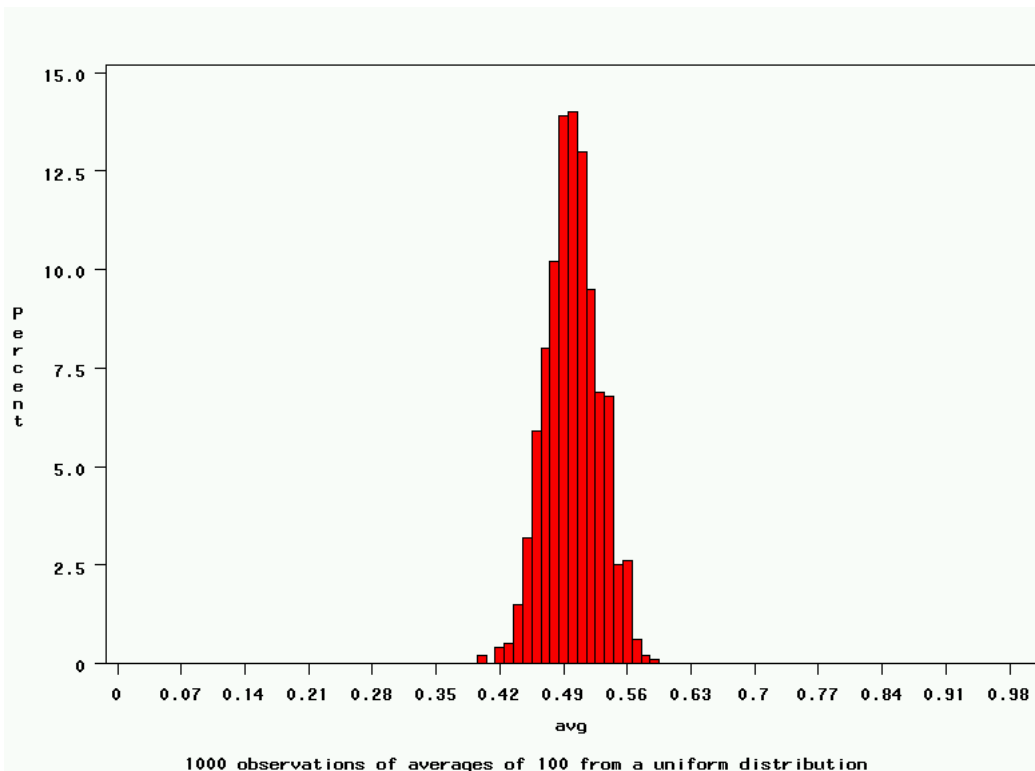
# Выборки с $n=2$ из равномерного распределения



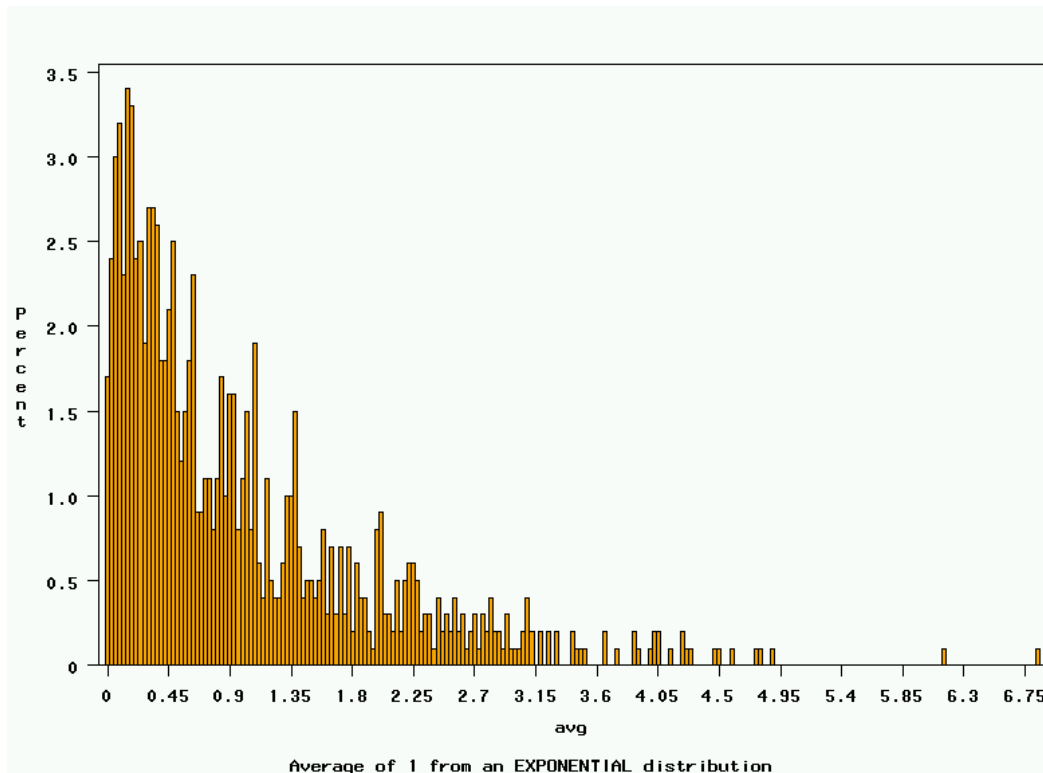
# Выборки с $n=5$ из равномерного распределения



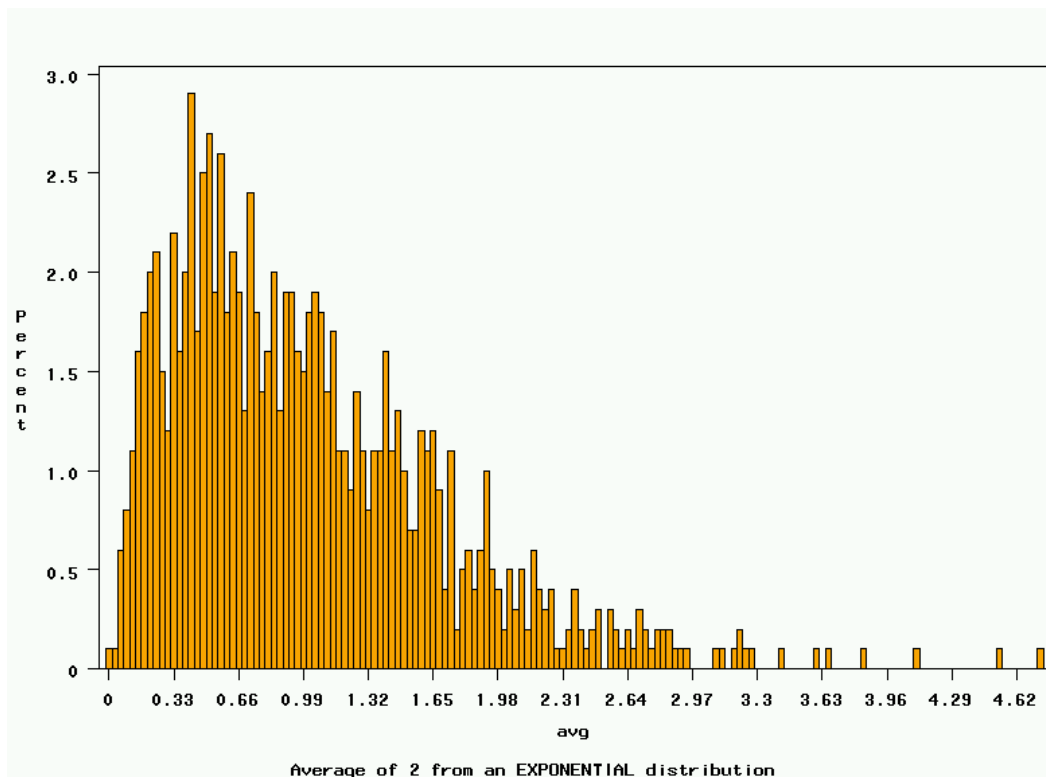
# Выборки с $n=100$ из равномерного распределения



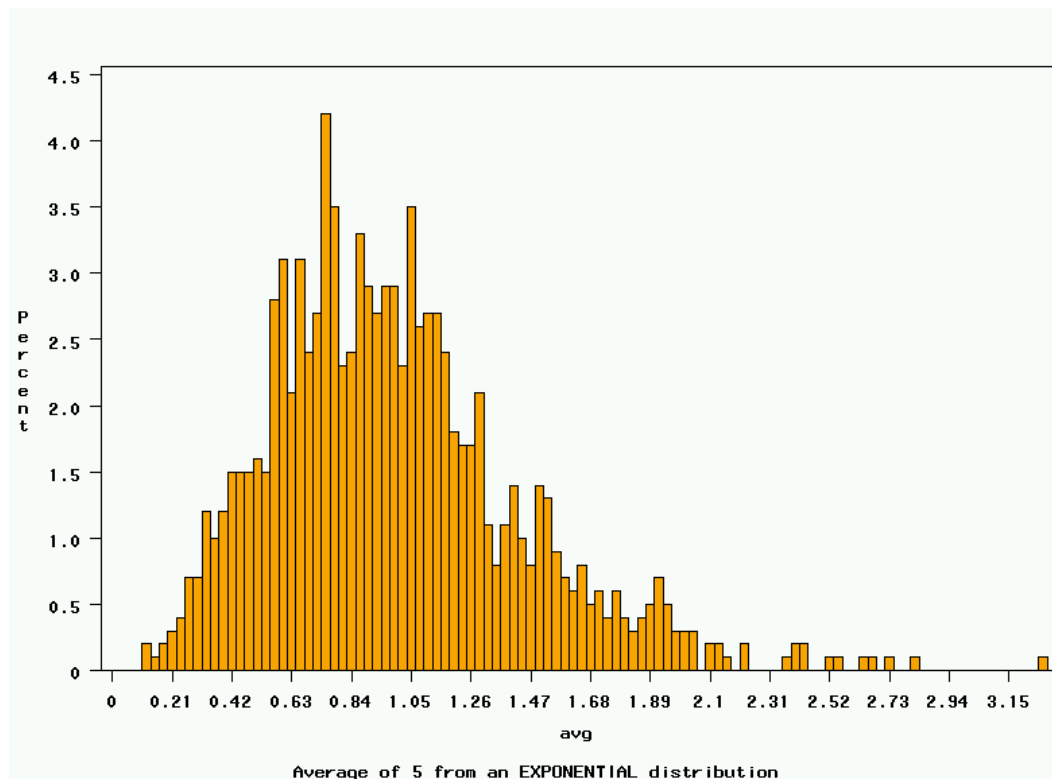
# Экспоненциальное распределение: среднее = 1 (исходное распределение)



# Выборки с $n=2$ из экспоненциального распределения

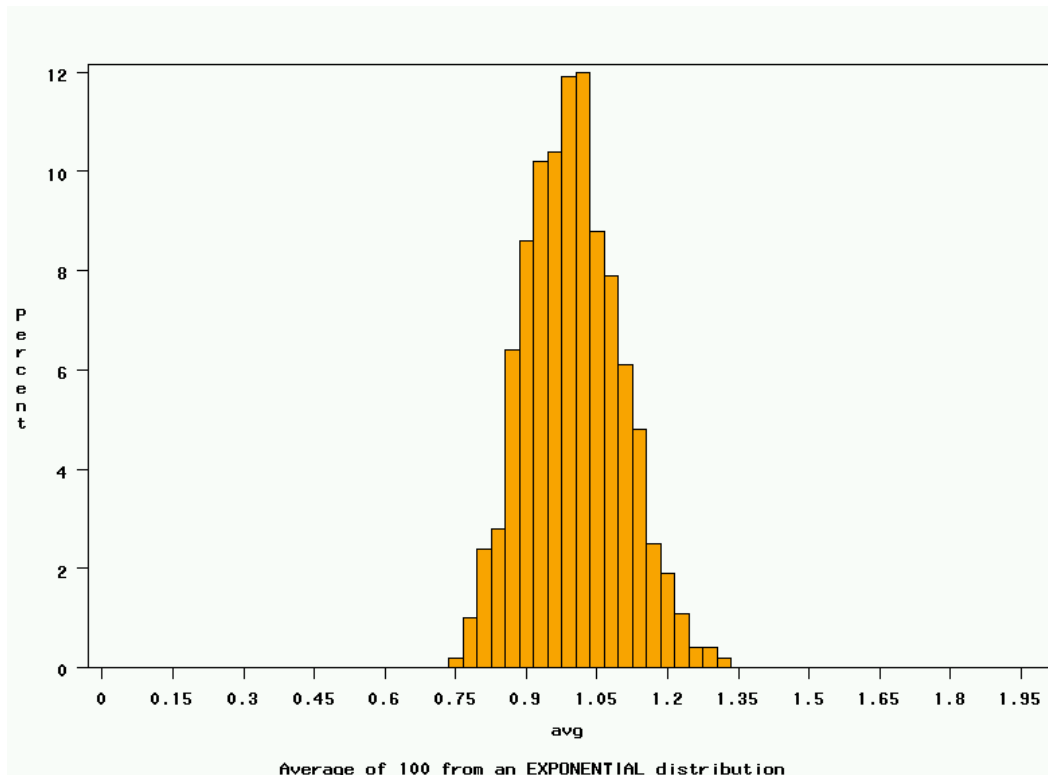


# Выборки с $n=5$ из экспоненциального распределения

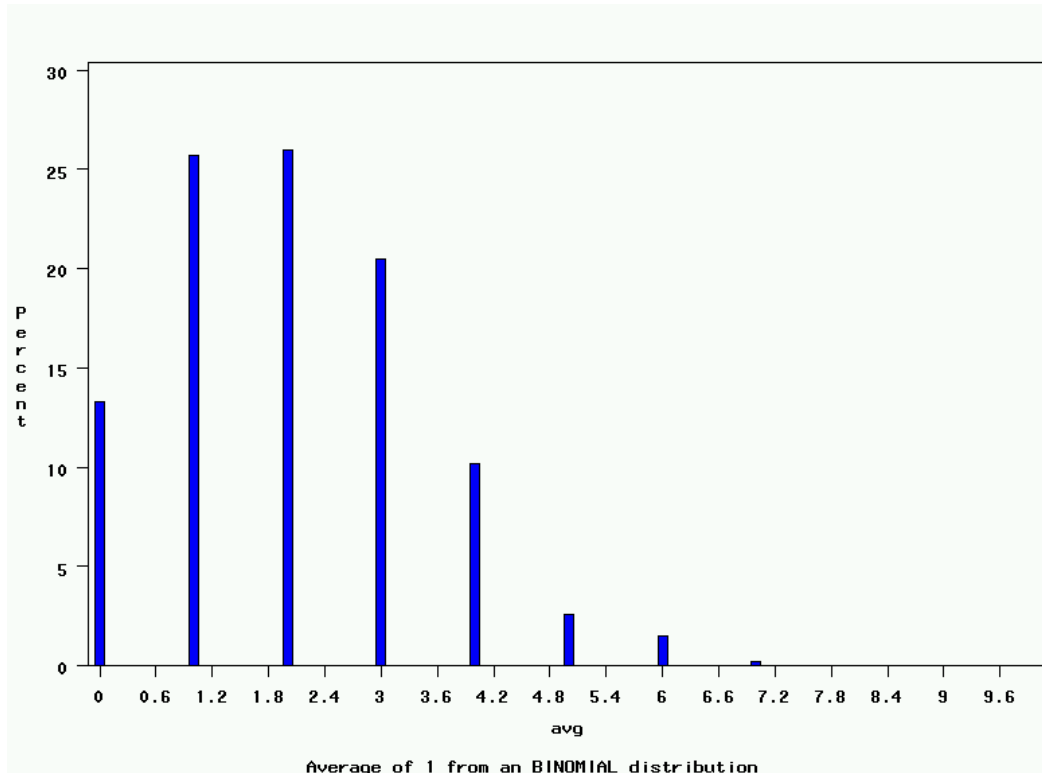




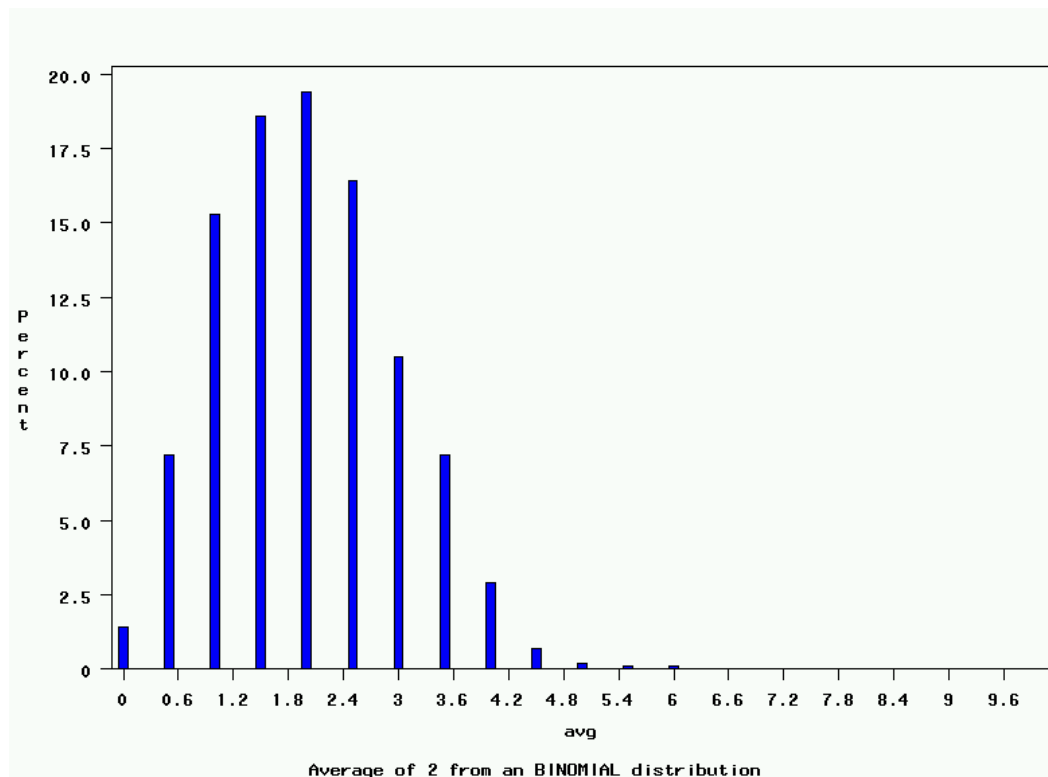
# Выборки с $n=100$ из экспоненциального распределения



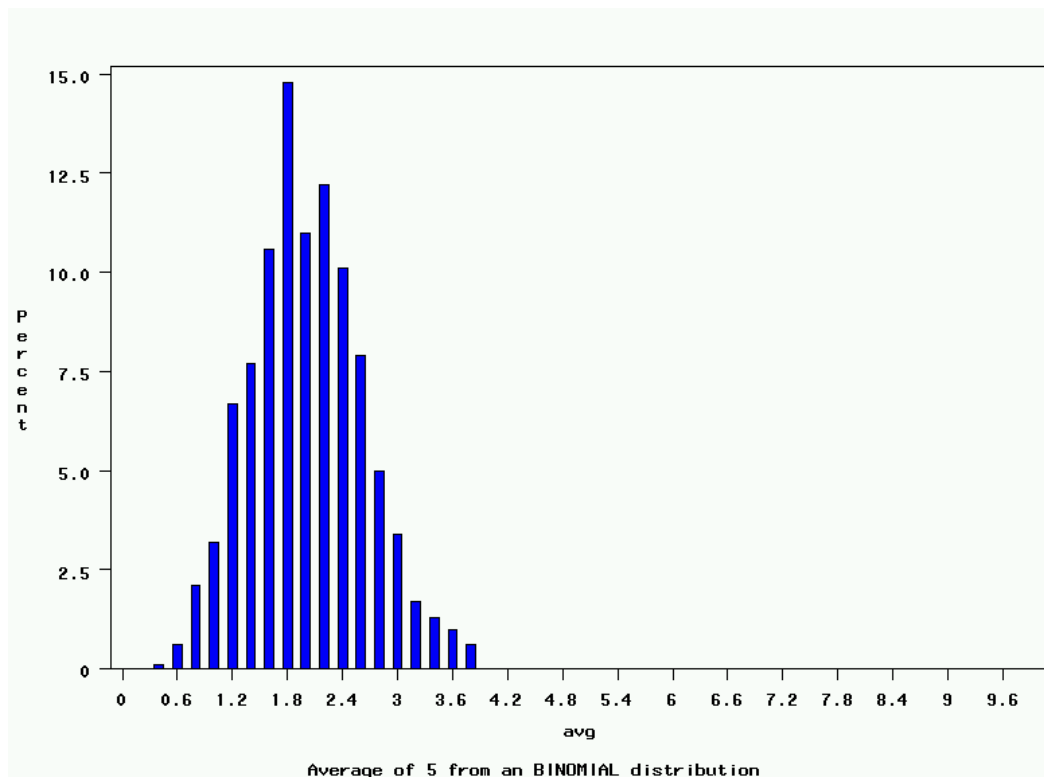
# Биномиальное распределение вероятностей, среднее = 2 (исходное распределение)



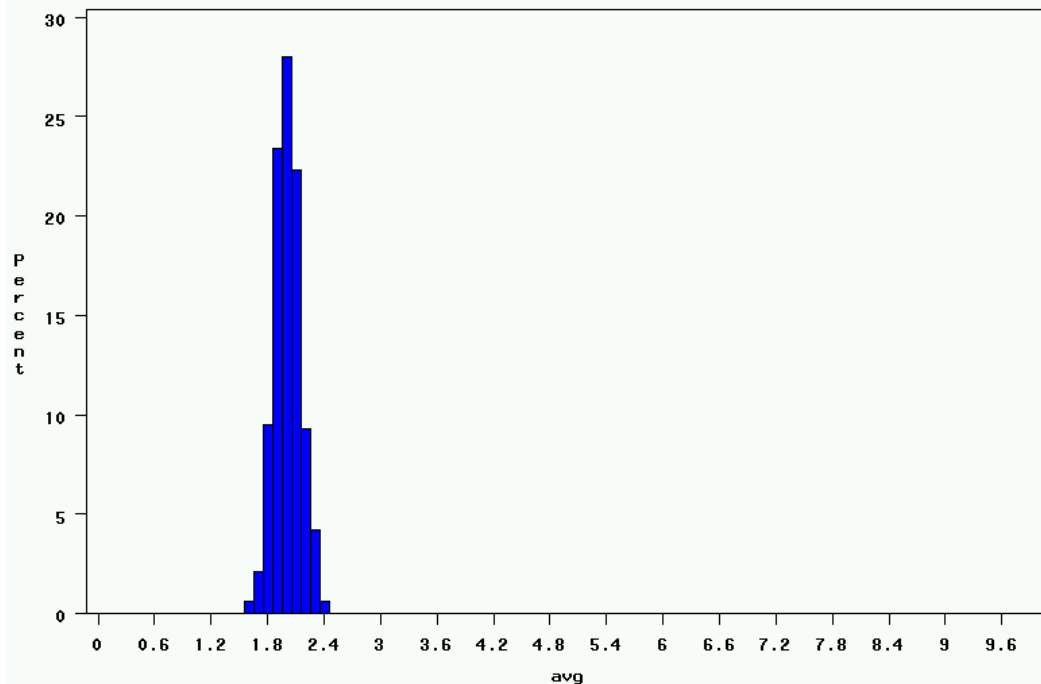
# Выборки с $n=2$ из биномиального распределения



# Выборки с $n=5$ из биномиального распределения



# Выборки с $n=100$ из биномиального распределения



Average of 100 from an BINOMIAL distribution

# РЕГУЛЯТОРНЫЕ АСПЕКТЫ

Необходимо представить подробное описание статистического анализа каждой первичной переменной эффективности. Минимальные требования к детализированному отчету:

- статистическая модель, лежащая в основе анализа.
- формулировка исследуемого клинического предположения, выраженная в конкретных статистических терминах, например, в форме нулевой и альтернативной гипотез;
- статистические методы, использованные для оценки эффекта, построения доверительных интервалов и др.
- допущения, лежащие в основе статистических методов
- критерий значимости, выборочное распределение критерия значимости при справедливости нулевой гипотезы, значения критерия значимости, уровень значимости (т.е. р-значение) и промежуточные сводные данные в формате, позволяющем статистику уполномоченного органа быстро и легко верифицировать результаты анализа. Необходимо указать, являются ли р-значения одно- или двусторонними. Необходимо представить обоснования использования односторонних критериев

- анализ, основанный на применении критерия Стьюдента (t-критерий) должен содержать значения **t-статистики**, связанные с ней степени свободы, **p-значение**, величины **двух выборок (групп)**, среднее и дисперсию каждой из выборок, **объединенную оценку дисперсии**.
- Документация многоцентровых исследований, проанализированная с помощью дисперсионного анализа должна включать следующие минимальные сведения: **таблицу дисперсионного анализа с выделением центров, лекарственных препаратов, эффекта взаимодействия этих факторов, остаточную и общую дисперсии**.



- Для исследований с **перекрестным дизайном** документация должна включать сведения о последовательностях включения пациентов, пациентах внутри последовательностей, исходных значениях в начале каждого периода, отмывке и длительности отмывки, выпадениях в течение каждого периода, лекарственных препаратах, периодах, **дисперсии, обусловленной взаимодействием этих факторов (лекарственных препаратов и периодов), остаточной и общей дисперсиях.** Для каждого источника изменчивости, исключая общую дисперсию, в таблице необходимо указать степени свободы, сумму квадратов, среднее квадратичное, соответствующий критерий F, р-значение и среднее значение суммы квадратов

# КРАТКОЕ РЕЗЮМЕ

- Четко определяем расчет разницы, формулировку статистических гипотез и графическое представление данных
- Определяем популяцию для анализа
- Выбираем/уточняем статистические критерии и метод расчета доверительных интервалов
- Представляем данные единообразно для всего отчета
- Информация по всем конечным точкам должна быть представлена в графическом и текстовом/табличном виде с учетом требований регулятора

Спасибо за  
внимание!