

**Лекция 3:
Статистика в клеточной
биологии и в клинических
исследованиях**

Тишков Артем Валерьевич

Никита Николаевич Хромов-Борисов

**Кафедра физики, математики и
информатики ПСПбГМУ им. акад. И.П.
Павлова**

**Упорядоченный посев и пуассонер –
высокоточная техника
количественной микробиологии**

МЕДИЦИНА. XXI ВЕК

№ 2 (11) 2008, с. 92-97

Распределение Пуассона

- Распределение числа событий, происходящих в фиксированном временнóм или пространственном интервале (объеме),
- при условии,
- что эти события независимы и что
- вероятность совпадения (попадания в одну точку пространства) или одновременного наступления дв́ событий пренебрежимо мала.

Симеон Дени Пуассон (*Simeon Denis Poisson*, 21.06.1781—25.04.1840)



ее

Распределение Пуассона

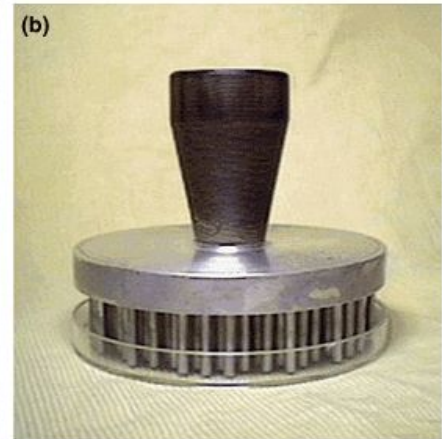
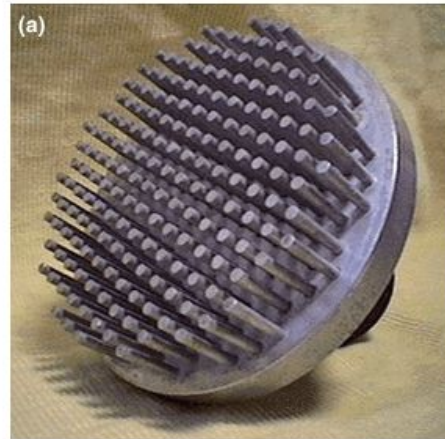
- $P(k) = e^{-\lambda} \lambda^k / k!$
- $e = 2,71828$ – основание натурального логарифма
- $k! = 1 \cdot 2 \cdot \dots \cdot (k-1) \cdot k$ – факториал
- Характеристическое свойство распределения Пуассона – его математическое ожидание (среднее значение) и дисперсия равны друг другу:
 - $E k^* = D k^* = \lambda,$
- т.е. это распределение имеет всего лишь один параметр λ .

Пуассонер, упорядоченный

ПОСЕВ

Н. Н. Хромов-Борисов, Jenifer Saffi , Joao A. P. Henriques

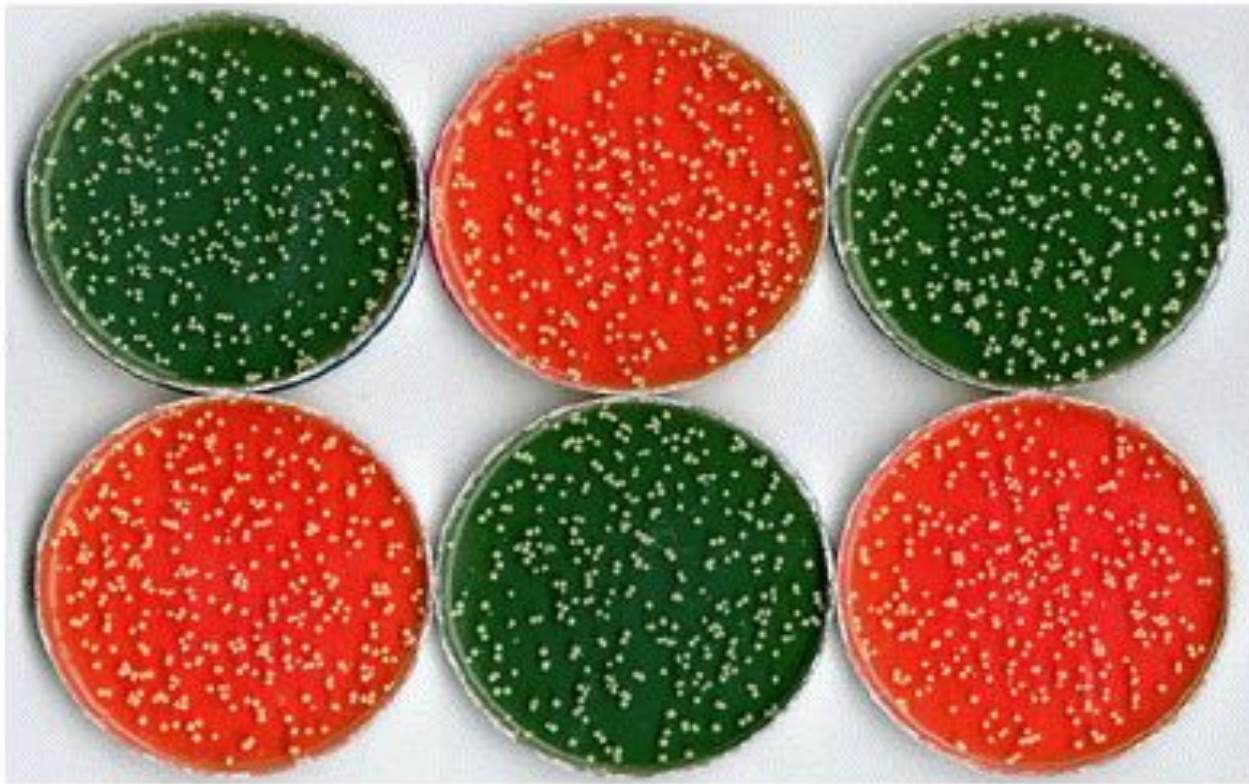
Упорядоченный посев и пуассонер – высокоточная техника количественной микробиологии



Сравнение упорядоченного посева с обычным методом



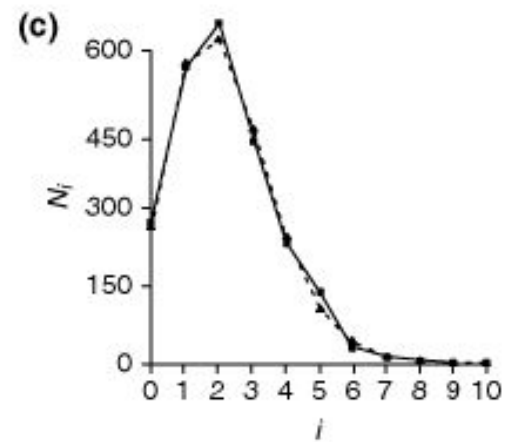
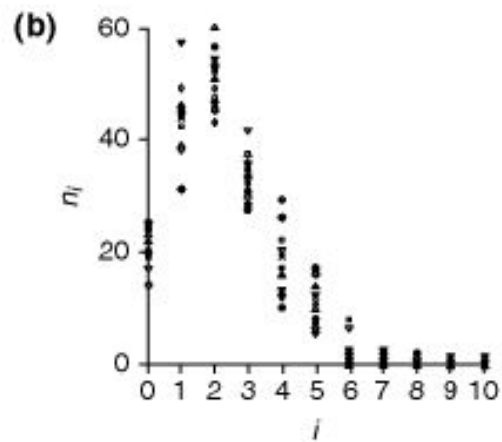
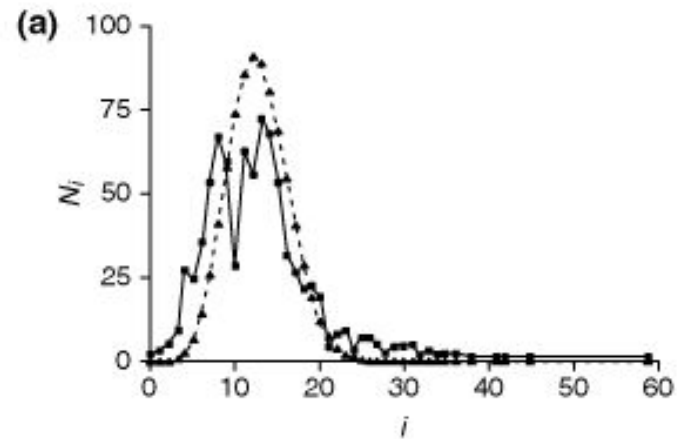
Воспроизводимость



Распределения числа колоний дрожжей на десяти чашках Петри, порожденные пуассонером, и их сравнение с распределением числа колоний, полученных традиционным методом посева

| Число следов (вмятин) с <i>i</i> колониями на каждой чашке Петри: | |
|---|--|
| <i>i</i> | I II III IV V VI VII VIII IX X Все- го |
| 0 | 21 23 19 27 25 18 24 23 26 25 231 |
| 1 | 48 48 57 51 48 43 41 54 45 43 478 |
| 2 | 56 54 54 49 59 49 63 54 47 48 533 |
| 3 | 29 29 35 37 32 38 32 34 36 38 340 |
| 4 | 18 17 13 14 11 23 14 13 20 20 163 |
| 5 | 13 15 5 8 8 7 11 7 11 11 96 |
| 6 | 2 1 2 1 2 8 0 1 2 1 20 |
| 7 | 0 0 0 0 0 1 2 1 0 1 5 |
| 8 | 0 0 0 0 2 0 0 0 0 0 2 |
| 9 | 0 0 1 0 0 0 0 0 0 0 1 |
| 10 | 0 0 1 0 0 0 0 0 0 0 1 |
| Общее число следов | 187 187 187 187 187 187 187 187 187 187 1870 |
| Согласие с распределением Пуассона ^[a] | 0,77 0,98 0,39 0,61 0,47 0,89 0,79 0,51 0,95 1,00 0,97 |
| Общее число колоний на чашке, полученных с применением пуассонера ^[b] | |
| PP = 0,66 | 396 392 378 362 374 437 388 364 394 401 3886 |
| Общее (с повышенной дисперсией) число колоний на чашке, полученное традиционным методом посева ^[c] | |
| PT ≈ 10–32 | 308 443 391 372 341 320 435 381 328 315 3634 |

Пуассоновость



Среднеквадратичное отклонение (стандартная ошибка среднего)

- Поскольку математическое ожидание (среднее значение) и дисперсия распределения Пуассона равны друг другу:
 - $E k^* = D k^* = \lambda$,
- то его среднеквадратичное отклонение есть:
 - $SE = \sqrt{D k^*} = \sqrt{\lambda}$

Элементы планирования экспериментов

Счетная камера Горяева (гемацитометер)

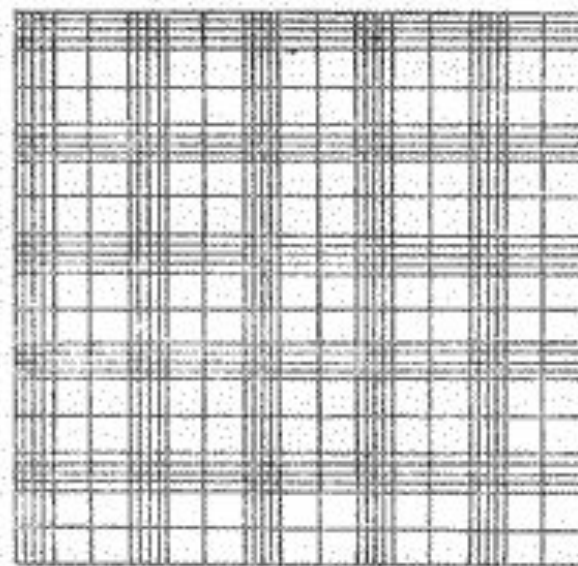
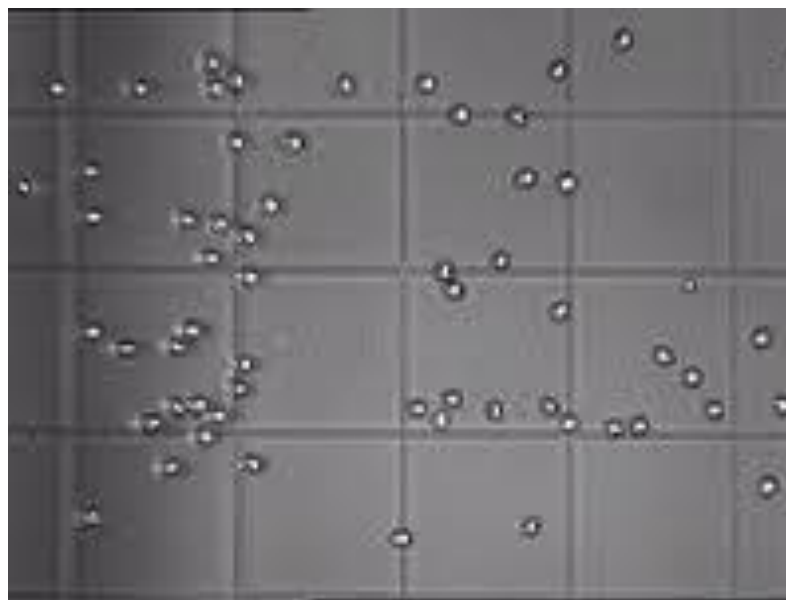
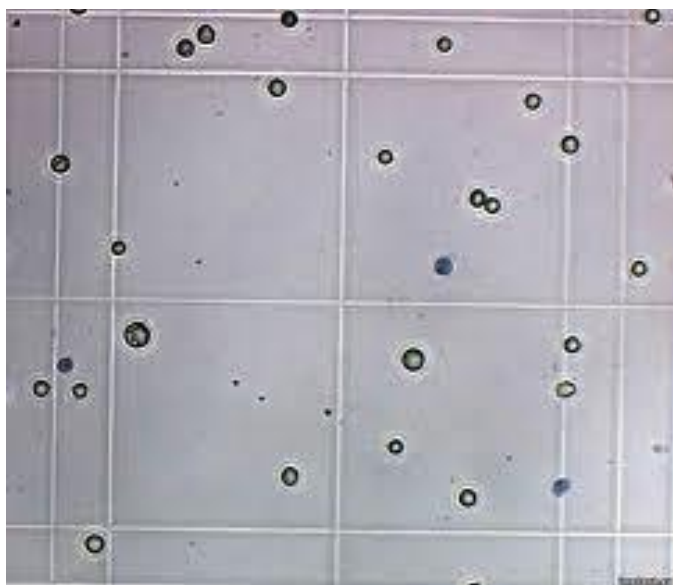
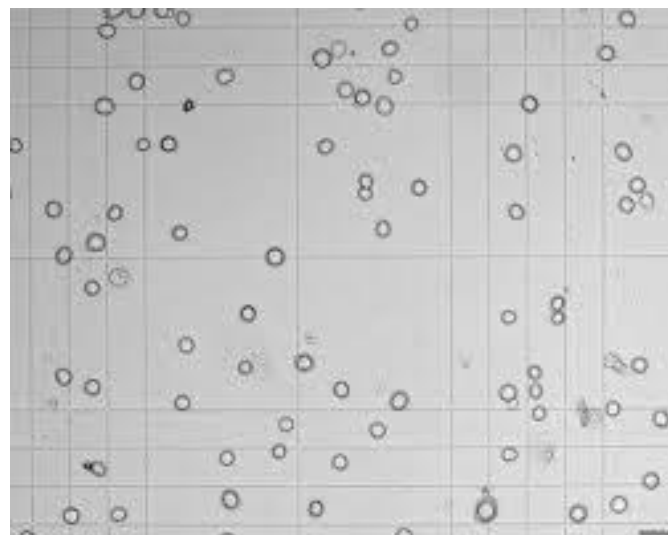
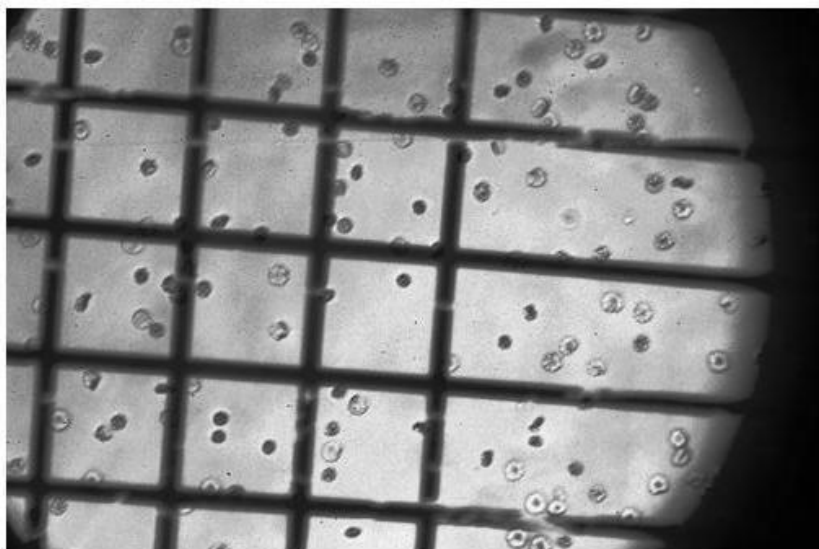


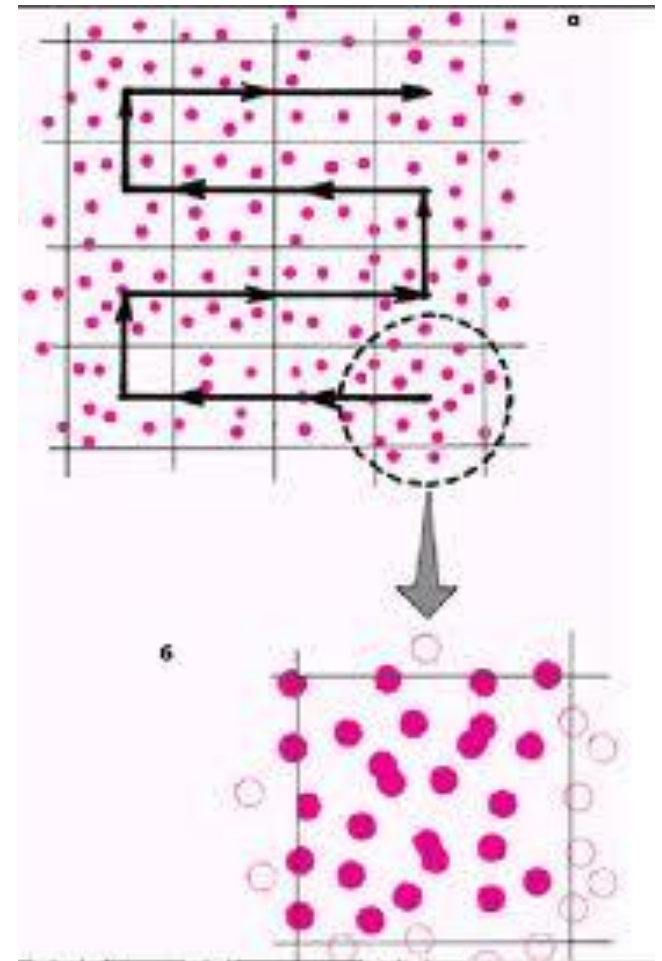
Рис. 1. Сетка камеры Горяева

Клетки в камере Горяева



Как подсчитывать клетки в камере Горяева

- $N \pm \sqrt{N}$
- Сколько клеток надо подсчитать, чтобы относительная ошибка составила 5%?
- Ответ: ~ 400
- Решение:
- $SE = \sqrt{400} = 20$
- $20 : 400 = 0,05$



- **Так сколько же клеток надо подсчитать, чтобы относительная ошибка составила 1%?**
- **Ответ: ~ 10 000**
- **Решение:**
- **$SE = \sqrt{10\ 000} = 100$**
- **$100 : 10\ 000 = 0,01$**

Молитва и сепсис

Leonard Leibovici, Университет Тель-Авива, Израиль



- Основные научные интересы:
- Бактериальные инфекции и антибиотикотерапия;
- Компьютеризация медицинских исследований;
- Медицинская этика;
- Доказательная медицина.
-

Leonard Leibovici **Effects of remote, retroactive intercessory prayer on outcomes in patients with bloodstream infection: randomised controlled trial // *BMJ*, 2001. – Vol. 323. – P. 1450-1451.**

- **Методы**
- **Выборку из 3393 пациентов с заражением крови (с сепсисом) рандомизированно, т.е. случайным образом разбили на две группы – контрольную (1702 пациента) и опытную (1691 пациент).**
- **Перечень имен пациентов во второй группе был передан человеку, который произносил краткую молитву за улучшение здоровья и полное выздоровление всей этой группы целиком.**
- **Пациенты, за которых молились, об этом не знали.**

Основные характеристики двух групп пациентов

Table 1 Baseline characteristics of patients. Values are percentages in each group, unless stated otherwise. None of the differences between the groups was significant

| Characteristic | Intervention group (n=1691) | Control group (n=1702) |
|-----------------------------------|-----------------------------|------------------------|
| Women | 46.3 | 48.5 |
| Median (range) age (years) | 72 (18-101) | 72 (18-99) |
| Source of infection: | | |
| Lungs | 8.3 | 9.4 |
| Urinary tract | 31.3 | 28.9 |
| Intra-abdominal | 9.5 | 8.9 |
| Soft tissues | 7.5 | 7.6 |
| Endocarditis | 3.5 | 3.3 |
| Neutropenic fever | 3.5 | 2.7 |
| Intravascular line | 6.3 | 6.4 |
| Other | 7.8 | 9.6 |
| Unknown | 22.3 | 23.2 |
| Septic shock | 11.2 | 11.8 |
| Neutropenia | 5.7 | 5.8 |
| Infected while in hospital | 40.2 | 41.9 |
| Median (range) creatinine (mg/dl) | 1.2 (0.2-14.8) | 1.2 (0.2-15.0) |
| Median (range) albumin (mg/dl) | 3.8 (1.1-5.1) | 3.8 (1.5-5.0) |

Результаты

| Группа | Умерло | Выжило | Всего |
|-------------|-----------------------------|--------|-------|
| Без молитвы | 514 $0,27^{0,30}_{0,34}$ | 1188 | 1702 |
| С молитвой | 475 $0,25^{0,28}_{0,32}$ | 1216 | 1691 |
| Всего | 989 $0,26^{0,29}_{0,31}$ | 2404 | 3393 |

Проверка независимости (однородности)

| | |
|----------------------|---------------------------|
| Точное P -значение | $P_{\text{exact}} = 0,19$ |
| Бейзов фактор | $BF_{01} = 12,7$ |

Связь между молитвой и смертностью от сепсиса статистически незначима ($P_{\text{val}} = 0,19 > 0,05$). Полученное значение бейзова фактора ($BF_{01} = 12,7$) показывает, что примерно в 13 раз более правдоподобно получить такие данные, когда эта связь действительно отсутствует, чем когда она есть. Молитва, скорее всего, не влияет на смертность при сепсисе.

Основные меры эффекта в таблицах 2x2

- Разность долей (рисков) – ***RD*** (Risk Difference)
- Отношение рисков (долей) – ***RR*** (Risk Ratio)
- Отношение оддсов (шансов за/против) – ***OR*** (Odds Ratio)
- Число подлежащих воздействию – ***NNT*** (Number Needed to Treat)

Таблица 2x2

| Группа | Исход (эффект) | | Всего |
|------------------|----------------------------------|-------------------|--------------|
| | Неблагоприятны й | Благоприятны й | |
| Контроль | <i>a</i> | <i>b</i> | <i>a + b</i> |
| Опыт | <i>c</i> | <i>d</i> | <i>c + d</i> |
| Всего | <i>a + c</i> | <i>b + d</i> | <i>N</i> |
| Разность долей | $RD = [c/(c + d)] - [a/(a + b)]$ | | |
| Отношение долей | $RR = [c/(c + d)] : [a/(a + b)]$ | | |
| Отношение оддов | $OR = ad/bc$ | | |
| Число подлежащих | $NNT = 1/RD$ | | |

Принципы построения бейзовских статистических оценок

Бейзовский Доверительный (правдоподобный) Интервал (ДИ)

$$P(\varphi_L \leq \tilde{\varphi} \leq \varphi_U) = 1 - \alpha$$

$$P(\tilde{\varphi} \leq \varphi_L) = \frac{\alpha}{2}$$

$$P(\tilde{\varphi} \geq \varphi_U) = \frac{\alpha}{2}$$

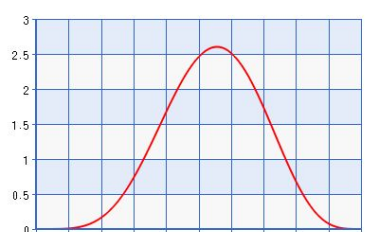
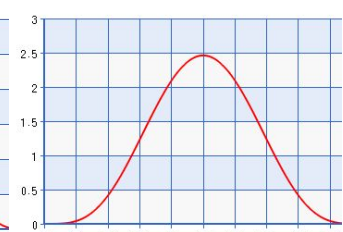
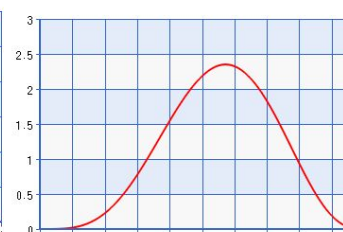
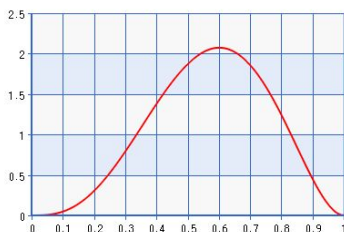
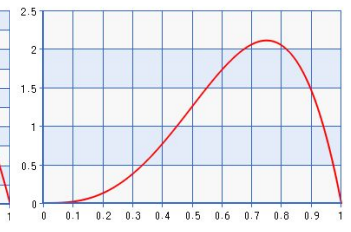
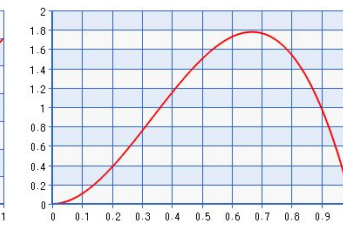
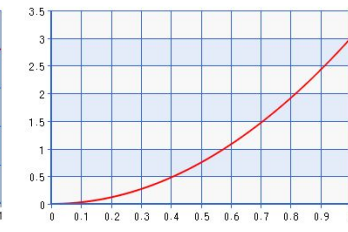
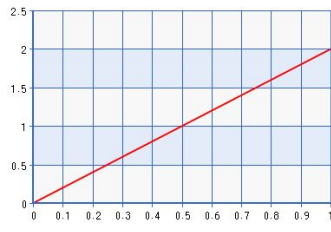
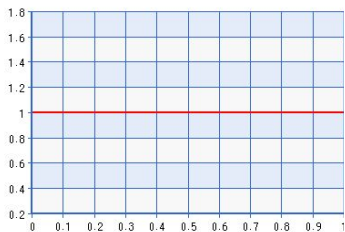
Использованные программы

- Моделирование подбрасывания монет:
- <http://www.random.org/coins/>
- И
- <http://www.random.org/coins/>
- Построение графиков бета-распределения:
- <http://keisan.casio.com/has10/SpecExec.cgi>
- Вычисление байесовских доверительных интервалов для долей:
- Программа LePAC version 2.0.38
- <http://www.univ-rouen.fr/LMRS/Persopage/Lecoutre/PAC.htm>
- И
- http://www.causascientia.org/math_stat/ProportionCI.html

Порождение распределения для доли выпадения орлов $\varphi(H)$

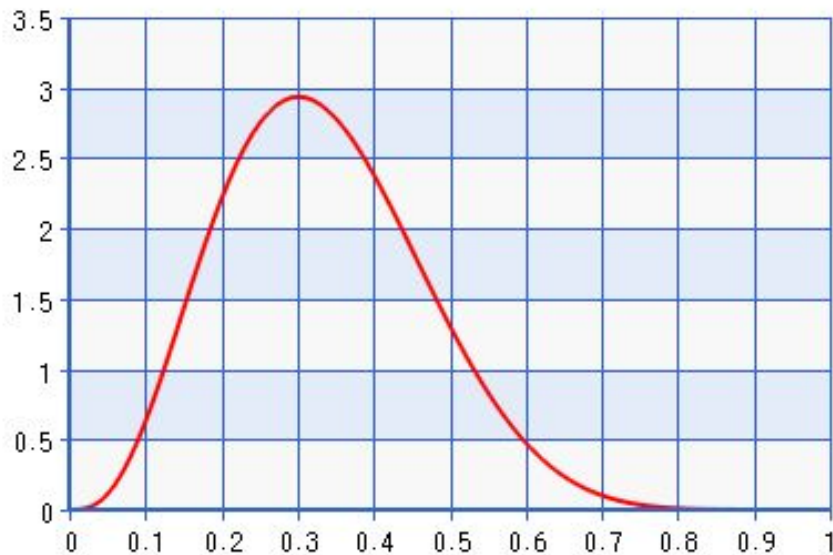
Нет информации

$Beta(a^* = 1, b^* = 1)$



Точечные и интервальные статистические оценки доли выпадения орлов $\varphi(H)$

3 H : 7 T; $n = 10$ *Beta*($a^* = 4, b^* = 8$)

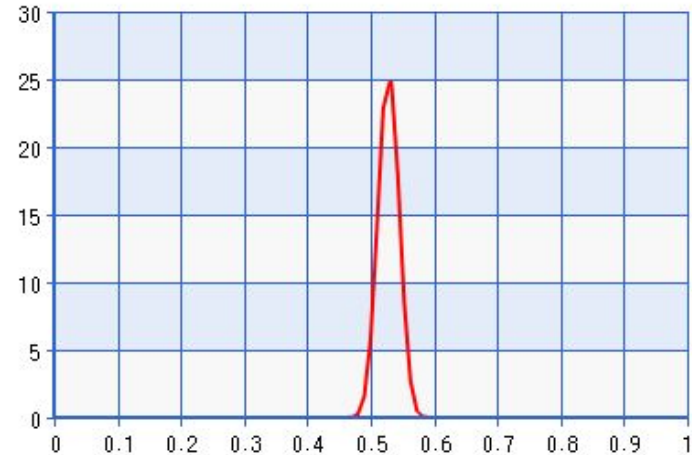
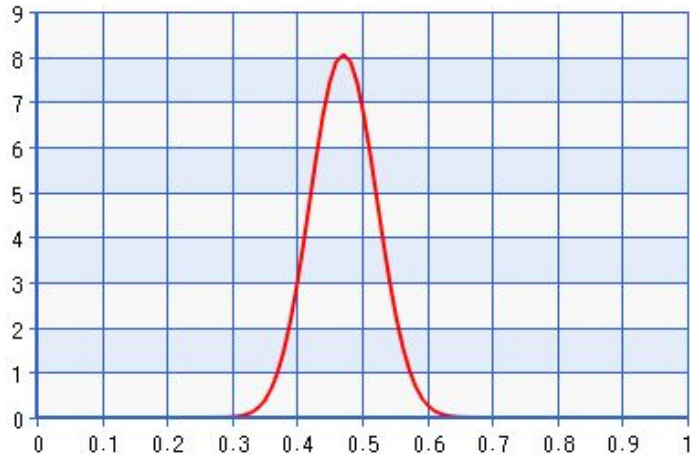


Плотность бета распределения
Beta($a = 4, b = 8$)

| Статистические оценки | | Ширина ДИ | Уровни доверия | |
|-----------------------|------------|-----------|----------------|-------|
| Точечные | Границы ДИ | | | |
| | | Нижние | Верхние | |
| | 0,11 | 0,61 | 0,50 | 95% |
| 0,32 | 0,07 | 0,69 | 0,62 | 99% |
| | 0,04 | 0,78 | 0,74 | 99,9% |

Точечные и интервальные статистические оценки доли выпадения орлов $\varphi(H)$

47 H : 53 T; $n = 100$; $Beta(a^* = 48, b^* = 54)$ 527 H : 473 T; $n=1000$; $Beta(a^* = 528, b^* = 474)$



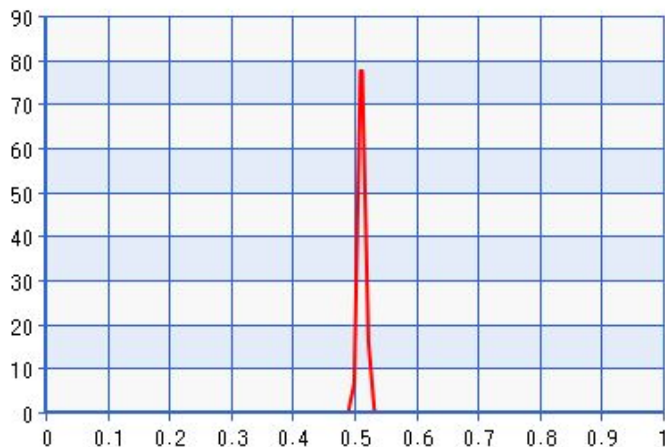
| Статистические оценки | | | | Уровн и дове- рия |
|-----------------------|-------------|--------------|----------------|----------------------------|
| Точеч- ная | Границы ДИ | | Шири- на ДИ | |
| | Ниж- ние | Верх- ние | | |
| | 0,38 | 0,58 | 0,20 | 95% |
| 0,47 | 0,35 | 0,60 | 0,25 | 99% |
| | 0,31 | 0,63 | 0,32 | 99,9% |

| Статистические оценки | | | | Уровн и дове- рия |
|-----------------------|-------------|--------------|----------------|----------------------------|
| Точеч- ная | Границы ДИ | | Шири- на ДИ | |
| | Ниж- ние | Верх- ние | | |
| | 0,50 | 0,56 | 0,06 | 95% |
| 0,53 | 0,49 | 0,57 | 0,08 | 99% |
| | 0,48 | 0,58 | 0,10 | 99,9% |

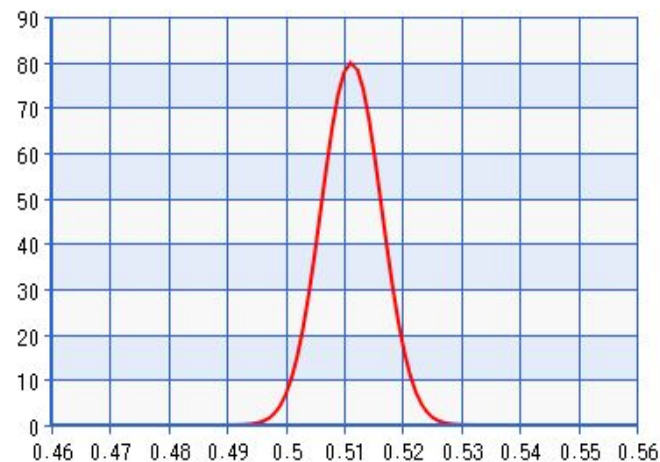
Точечные и интервальные статистические оценки доли выпадения орлов $\varphi(H)$

5111 H : 4889 T; $n = 10\ 000$;

$Beta(a^* = 5112, b^* = 4890)$



Более тонкий масштаб



| Статистические оценки | | | | Уровн и довер ия |
|-----------------------|-------------|--------------|----------------|---------------------------|
| Точеч ная | Границы ДИ | | Шири- на ДИ | |
| | Ниж- ние | Верх- ние | | |
| | 0,501 | 0,521 | 0,020 | 95% |
| 0,511 | 0,498 | 0,524 | 0,026 | 99% |
| | 0,495 | 0,528 | 0,033 | 99,9% |

Оценка доли скончавшихся в контрольной группе, φ_1 в программе LePAC

<http://www.univ-rouen.fr/LMPS/Personage/Lecoutre/PAC.htm>

LePAC version 2.0.41 - [LesProportions 1']

Data PACexpress PAC LesBayesiens LesDistributions Edition Tools/options Screen exit ?

1 group 2 independent groups LesImplications close

Data g1\g2 1\0 prior <- posterior

| cell counts | | |
|-------------|-----|------|
| | 1 | 0 |
| g1 | 514 | 1188 |
| g2 | 475 | 1216 |
| | 989 | 2404 |
| | | 3393 |

beta prior

| | 1 | 0 |
|----|---|------|
| g1 | 1 | 1 |
| g2 | 1 | 1 |
| | 0 | ? |
| | 1 | 1001 |
| | | 0110 |

φ_1 $\delta = \varphi_1 - \varphi_2$ $\tau = \varphi_1 / \varphi_2$ $\varphi_1 / (\varphi_1 + \varphi_2)$
 $(\varphi_1 - \varphi_2) / \varphi_1$ $(\varphi_1 - \varphi_2) / \varphi_2$ $(\varphi_1 - \varphi_2) / [(\varphi_1 + \varphi_2) / 2]$
 $\varphi_1 / (1 - \varphi_1)$ $\omega = [\varphi_1 / (1 - \varphi_1)] / [\varphi_2 / (1 - \varphi_2)]$
 $\varphi_2 / (1 - \varphi_2)$ $\nu = (1 - \omega) / (1 + \omega)$ Q Yule]

Model binomial poisson

$\varphi_1 \sim \beta(515, 1189)$

Statement
 $\Pr(X < x)$ $\Pr(X > x)$
 $\Pr(x_1 < X < x_2)$ $\Pr(X < x_1 \text{ ou } X > x_2)$

Limits Limits

 Probability

Compute Statistics

display limits g1/g2

discrimales limit

distribution probability

Two-sided 0.95 confidence intervals
 Limits 80%, 90%, 95%, 99%, 99.9% ("equal-tailed intervals")

Curve
 p(x)
 $\Pr(X < x)$
 $\Pr(X > x)$

Options

$\varphi_1 \sim \beta(515, 1189)$

mean
 standard deviation

generate a sample

Оценка доли скончавшихся в группе подвергнутых воздействию молитвы φ_2 в программе LePAC

LePAC version 2.0.41 - ['LesProportions 1']

Data PACexpress PAC LesBayesiens LesDistributions Edition Tools/options Screen exit ?

1 group 2 independent groups LesImplications close

Data g1\g2 1\0 prior <- posterior

| cell counts | | 1 | 0 | 1\0 |
|-------------|-----|------|------|-----|
| g1 | 514 | 1188 | 1702 | |
| g2 | 475 | 1216 | 1691 | |
| | 989 | 2404 | 3393 | |

beta prior

| beta prior | | 1 | 0 |
|------------|---|---|-----------------|
| g1 | 1 | 1 | |
| g2 | 1 | 1 | |
| | 0 | ? | 1 1001 0110 |

φ_1 $\delta = \varphi_1 - \varphi_2$
 φ_2 $\tau = \varphi_1 / \varphi_2$ $\varphi_1 / (\varphi_1 + \varphi_2)$
 $(\varphi_1 - \varphi_2) / \varphi_1$ $(\varphi_1 - \varphi_2) / \varphi_2$ $(\varphi_1 - \varphi_2) / [(\varphi_1 + \varphi_2) / 2]$
 $\varphi_1 / (1 - \varphi_1)$ $\omega = [\varphi_1 / (1 - \varphi_1)] / [\varphi_2 / (1 - \varphi_2)]$
 $\varphi_2 / (1 - \varphi_2)$ $\nu = (1 - \omega) / (1 + \omega)$ [Yule]

Model binomial poisson

$\varphi_1 \sim \beta(515, 1189)$

Statement

$\Pr(X < x)$ $\Pr(X > x)$
 $\Pr(x_1 < X < x_2)$ $\Pr(X < x_1 \text{ ou } X > x_2)$

Limits Limits

Probability Probability

Compute Statistics

display limits g1/g2

discrimales limit

distribution probability

Two-sided 0.95 confidence intervals

Limits 80%, 90%, 95%, 99%, 99.9% ("equal-tailed intervals")

Curve

$p(x)$
 $\Pr(X < x)$
 $\Pr(X > x)$

Options

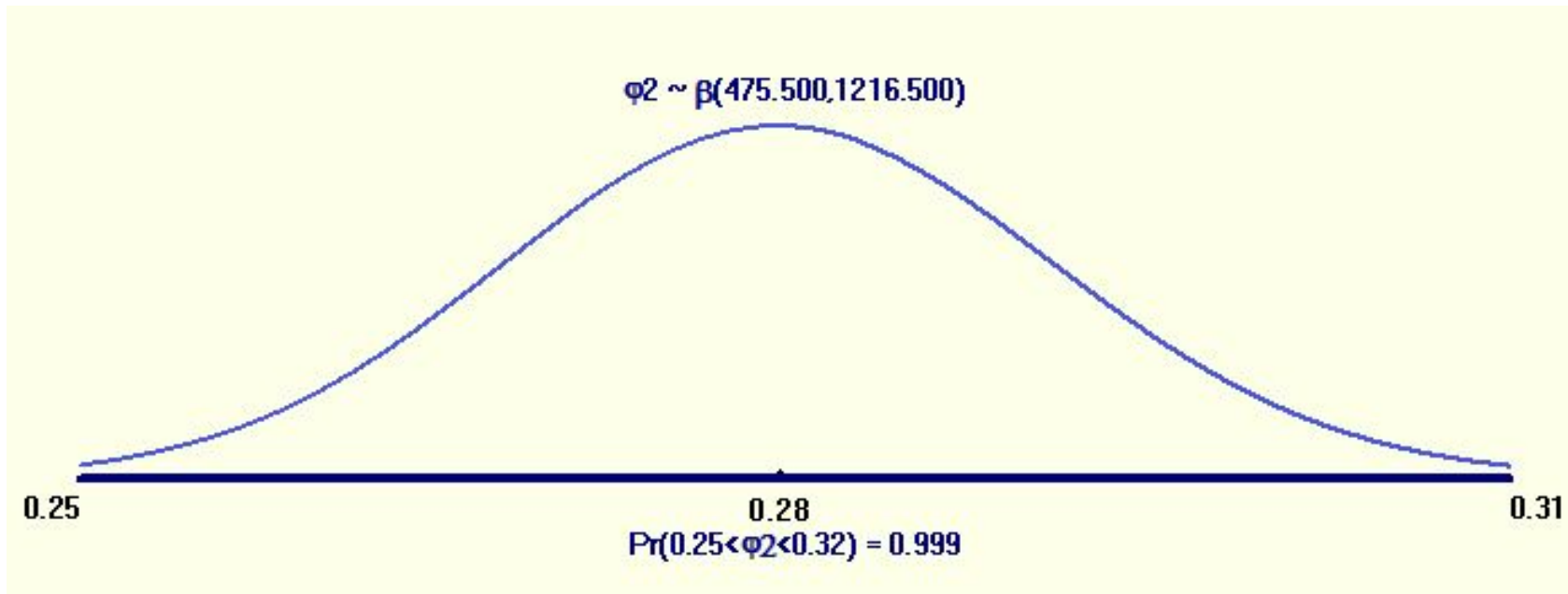
$\varphi_1 \sim \beta(515, 1189)$

mean

standard deviation

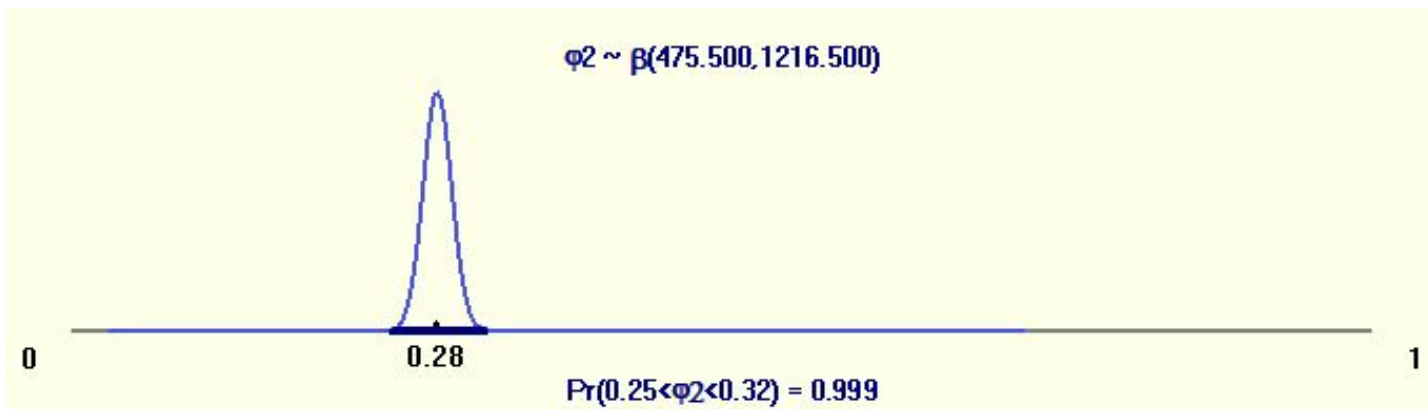
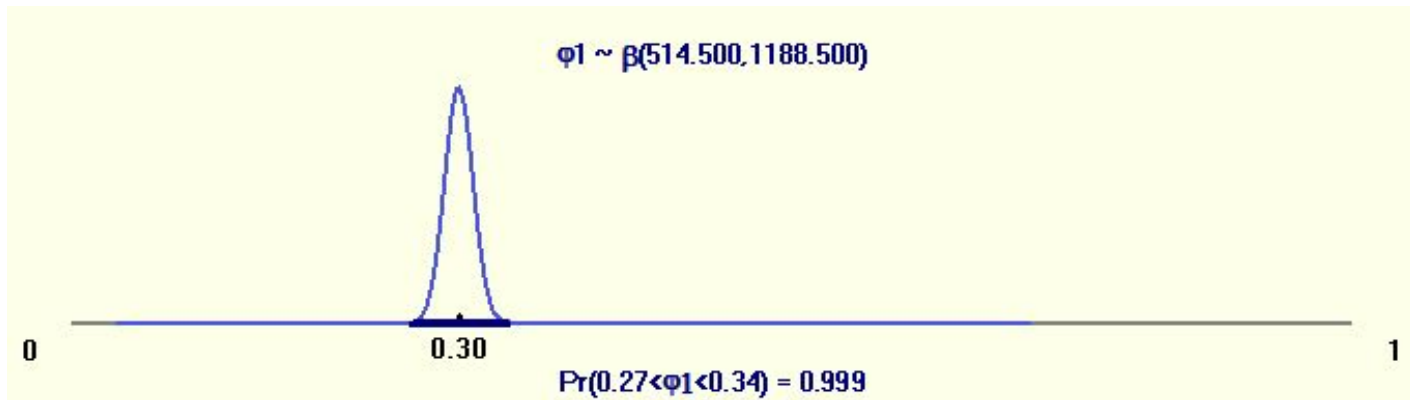
generate a sample

Плотность распределения и 99,9%-й ДИ для доли скончавшихся в группе подвергнутых воздействию молитвы, φ_2



$$\varphi_2 = {}_{0,25}0,28_{0,32}$$

Плотности распределения для долей скончавшихся от сепсиса в группах пациентов, подвергнутых (φ_1) и не подвергнутых молитве (φ_2)



Оценка неизвестной разности долей

$$RD_{\text{unkn}} = \delta = \varphi_1 - \varphi_2 \text{ в программе LePAC}$$

LePAC version 2.0.41 - ['LesProportions 1']

Data PACexpress PAC LesBayesiens LesDistributions Edition Tools/options Screen exit ?

1 group 2 independent groups LesImplications close

Data g1\g2 1\0 prior <- posterior

| cell counts | | | |
|-------------|-----|------|------|
| | 1 | 0 | 1\0 |
| g1 | 514 | 1188 | 1702 |
| g2 | 475 | 1216 | 1691 |
| | 989 | 2404 | 3393 |

beta prior

| | | 1 | | 0 | |
|--|----|----|----|------|------|
| | g1 | g2 | g1 | g2 | g1 |
| | 1 | 1 | 1 | 1 | 1 |
| | 0 | ? | 1 | 1001 | 0110 |

$\varphi_1 \sim \beta(515, 1189) \mid \varphi_2 \sim \beta(476, 1217)$

Model binomial poisson

Statement

Pr(X<x) Pr(X>x) Limits
 Pr(x1<X<x2) Pr(X<x1 ou X>x2) Probability

Compute Statistics

display limits g1/g2

decimals limit 2 Two-sided 0.999 confidence intervals
 distribution 3 probability 3 Limits 80%, 90%, 95%, 99%, 99.9% ("equal-tailed intervals")

CurVe

p(x) Pr(X<x) Pr(X>x)

Options

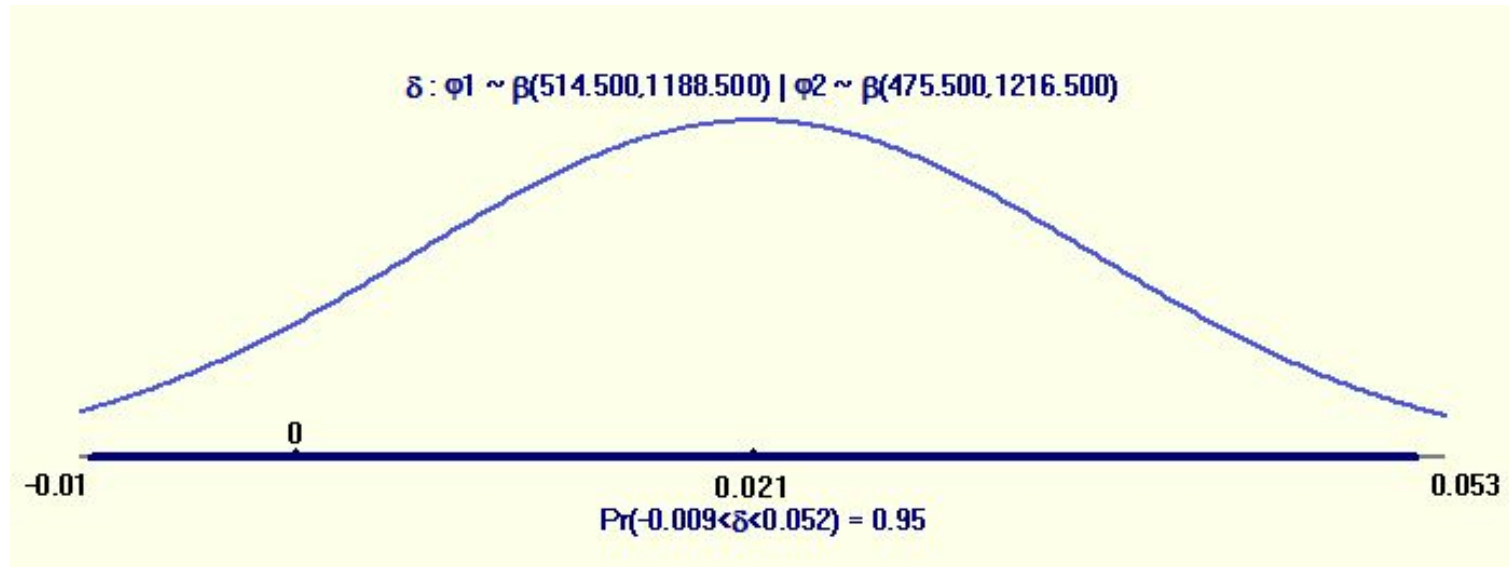
$\delta : \varphi_1 \sim \beta(515, 1189) \mid \varphi_2 \sim \beta(476, 1217)$

0.02
Pr(-0.03 < δ < 0.07) = 0.999

generate a sample

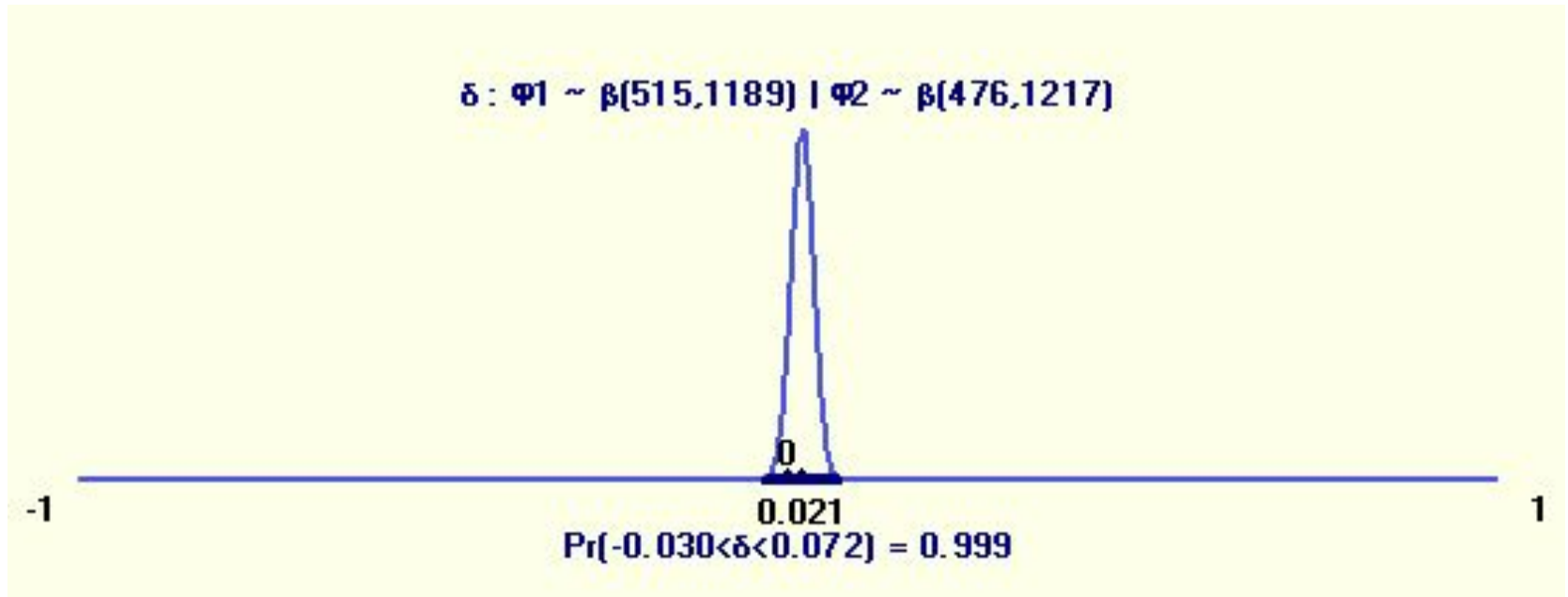
mean
 standard deviation

Плотность распределения и 95%-й ДИ для оцениваемой разности долей $RD_{\text{unkn}} = \delta = \varphi_1 - \varphi_2$



$$RD = \begin{matrix} & & 0,021 & & \\ -0,009 & & & & 0,052 \end{matrix}$$

Плотность распределения для оцениваемой
разности долей $\delta = \varphi_1 - \varphi_2 = RD$ в допустимых
границах от -1 до +1



95%, 99% и 99,9% ДИ для оцениваемой разности долей $RD_{\text{unkn}} = \delta = \varphi_1 - \varphi_2$

| Уровень доверия | ДИ | | |
|-----------------|---------|---------|--------|
| | Границы | | Ширина |
| | Нижняя | Верхняя | |
| 95% | - 0,009 | 0,052 | 0,061 |
| 99% | - 0,019 | 0,061 | 0,080 |
| 99,9% | - 0,030 | 0,072 | 0,102 |

Когда доли равны ($\varphi_1 = \varphi_2$), то их разность равна нулю: $RD = \delta = \varphi_1 - \varphi_2 = 0$.

Все три полученных ДИ для оцениваемой разности долей RD_{unkn} содержат значение $RD = 0$.

Это дает нам основание утверждать, что, скорее всего, оцениваемое этими интервалами неизвестное нам значение RD_{unkn} статистически не отличается от нуля и, соответственно, первая и вторая доли статистически одинаковы.

Основной вывод: Молитва, скорее всего, не влияет на смертность при сепсисе.

Что такое отношение рисков, $RR = \tau$?

- Это есть отношение двух условных вероятностей (долей), например, доли скончавшихся в контрольной группе φ_1 к доле скончавшихся в опытной группе φ_2 :

- $RR = \varphi_1 / \varphi_2$

Оценка неизвестного отношения долей (рисков) $RR_{\text{unkn}} = T = \varphi_1 / \varphi_2$ в программе LePAC

LePAC version 2.0.41 - [LesProportions 1']

Data PACexpress PAC LesBayesiens LesDistributions Edition Tools/options Screen eXit ?

1 group 2 independent groups LesImplications cIUse

| Data | | g1\g2 | 1\0 |
|-------------|-----|-------|------|
| cell counts | | | |
| | 1 | 0 | |
| g1 | 514 | 1188 | 1702 |
| g2 | 475 | 1216 | 1691 |
| | 989 | 2404 | 3393 |

beta prior

| prior ← posterior | |
|-------------------|------------------------|
| g1 | 1/2 1/2 |
| g2 | 1/2 1/2 |
| 0 | ? 1 1001 0110 |

φ_1 $\delta = \varphi_1 - \varphi_2$
 φ_2 $\tau = \varphi_1 / \varphi_2$ $\varphi_1 / (\varphi_1 + \varphi_2)$
 $(\varphi_1 - \varphi_2) / \varphi_1$ $(\varphi_1 - \varphi_2) / \varphi_2$ $(\varphi_1 - \varphi_2) / [(\varphi_1 + \varphi_2) / 2]$
 $\varphi_1 / (1 - \varphi_1)$ $\omega = [\varphi_1 / (1 - \varphi_1)] / [\varphi_2 / (1 - \varphi_2)]$
 $\varphi_2 / (1 - \varphi_2)$ $\nu = (1 - \omega) / (1 + \omega)$ Q Yule

Model binomial poisson

$\varphi_1 \sim \beta(514.500, 1188.500) \mid \varphi_2 \sim \beta(475.500, 1216.500)$

sTatement

Pr(X<x) Pr(X>x)
 Pr(x1<X<x2) Pr(X<x1 ou X>x2)

Limits: 0.97, 1.19

Probability: 0.95

Compute

Statistics: display limits g1/g2

display limits: Two-sided 0.95 confidence intervals

Limits: 80%, 90%, 95%, 99%, 99.9% ("equal-tailed intervals")

CurYe

p(x)
 Pr(X<x)
 Pr(X>x)

Options

generate a sample: 10000

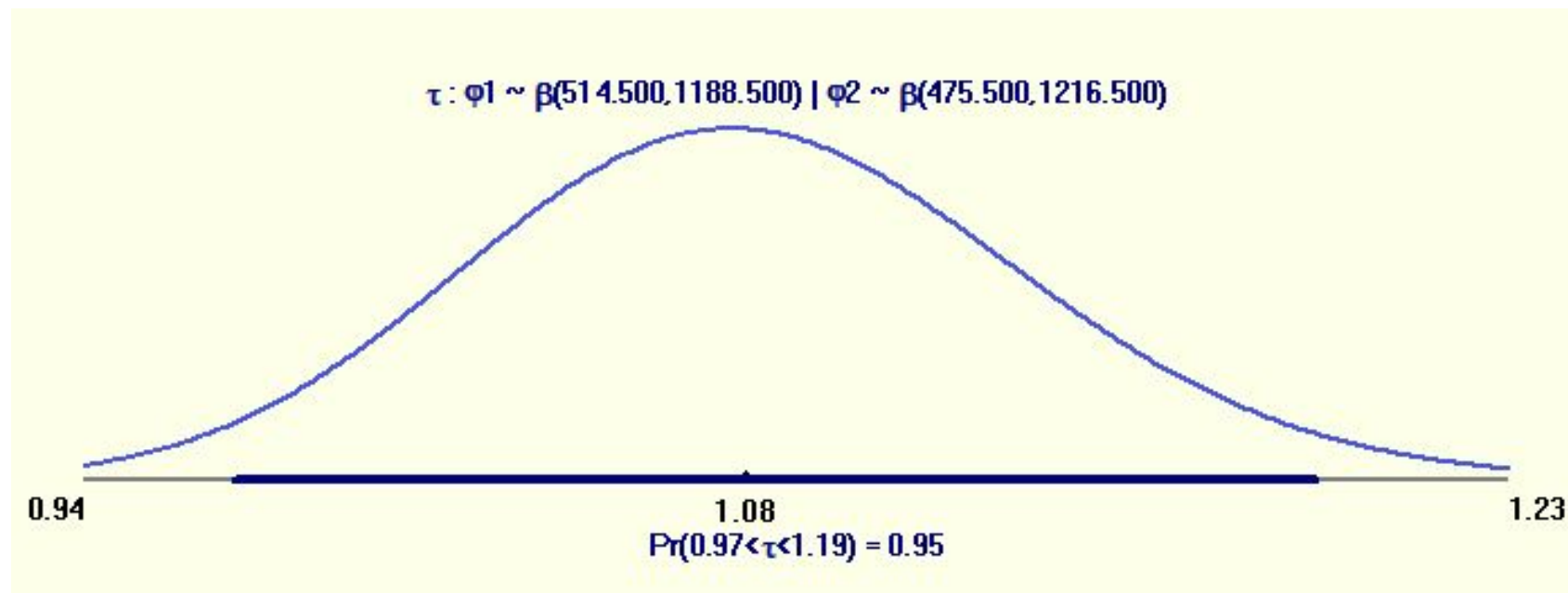
mean: 1.077

standard deviation: 0.058

$\tau : \varphi_1 \sim \beta(514.500, 1188.500) \mid \varphi_2 \sim \beta(475.500, 1216.500)$

$Pr(0.97 < \tau < 1.19) = 0.95$

Плотность распределения и 95%-й ДИ для оцениваемого отношения долей (рисков) $RR_{\text{unkn}} = T = \varphi_1 / \varphi_2$



$$RR = {}_{0,97}1,08_{1,19}$$

95%, 99% и 99,9% ДИ для оцениваемого отношения долей $RR_{\text{unkn}} = T = \varphi_1 / \varphi_2$

| Уровень доверия | ДИ | | |
|-----------------|---------|---------|--------|
| | Границы | | Ширина |
| | Нижняя | Верхняя | |
| 95% | 0,97 | 1,19 | 0,061 |
| 99% | 0,94 | 1,23 | 0,080 |
| 99,9% | 0,90 | 1,28 | 0,102 |

Когда доли равны ($\varphi_1 = \varphi_2$), то их отношение равно единице:

$$RR = T = \varphi_1 / \varphi_2 = 1.$$

Все три полученных ДИ для оцениваемого отношения долей RR_{unkn} содержат значение $RR = 1$.

Это дает нам основание утверждать, что, скорее всего, оцениваемое этими интервалами неизвестное нам значение RR_{unkn} статистически не отличается от 1, соответственно, первая и вторая доли статистически одинаковы.

Основной вывод: Молитва, скорее всего, не влияет на смертность при сепсисе.

Что такое «отношение шансов», *OR*?

- Это «трехэтажное» отношение:
- 1. Вероятность есть отношение количества исходов k , благоприятствующих данному событию (A) к общему количеству исходов N :
 - $P(A) = k / N$
- 2. Шансы (Odds) суть ставки за и против, т. е. отношение вероятности данного события $P(A)$ к вероятности противоположного события $P(\text{non}A) = 1 - P(A)$:
 - $\text{Odds} = P(A) : [1 - P(A)] = k / (N - k)$
- 3. Отношение шансов (*OR* – Odds Ratio) есть отношение шансов за и против события A к шансам за и против события B :
 - $OR = \{P(A) / [1 - P(A)]\} : \{P(B) / [1 - P(B)]\}$

Оценка неизвестного отношения однов (шансов за/против)

$OR_{\text{unkn}} = \omega = [\varphi_1 / (1 - \varphi_1)] : [\varphi_2 / (1 - \varphi_2)]$ в программе LePAC

LePAC version 2.0.41 - [LesProportions 1']

Data PACeXpress PAC LesBayesiens LesDistributions Edition Tools/options Screen eXit ?

1 group 2 independent groups LesImplications cLOse

| Data | | g1\g2 | 1\0 |
|-------------|-----|-------|------|
| cell counts | | | |
| | 1 | 0 | |
| g1 | 514 | 1188 | 1702 |
| g2 | 475 | 1216 | 1691 |
| | 989 | 2404 | 3393 |

| beta prior | | 1 | 0 |
|------------|-----|------|------|
| g1 | 1/2 | 1/2 | |
| g2 | 1/2 | 1/2 | |
| | 0 | ? | 1 |
| | | 1001 | 0110 |

φ_1 $\delta = \varphi_1 - \varphi_2$
 φ_2 $\tau = \varphi_1 / \varphi_2$ $\varphi_1 / (\varphi_1 + \varphi_2)$
 $(\varphi_1 - \varphi_2) / \varphi_1$ $(\varphi_1 - \varphi_2) / \varphi_2$ $(\varphi_1 - \varphi_2) / [(\varphi_1 + \varphi_2) / 2]$
 $\varphi_1 / (1 - \varphi_1)$ $\omega = [\varphi_1 / (1 - \varphi_1)] / [\varphi_2 / (1 - \varphi_2)]$
 $\varphi_2 / (1 - \varphi_2)$ $\nu = (1 - \omega) / (1 + \omega)$ Q Yule

Model binomial poisson

$\varphi_1 | (1 - \varphi_1) \sim \beta_{11}(514.500, 1188.500) \mid \varphi_2 | (1 - \varphi_2) \sim \beta_{11}(475.500, 1216.500)$

sStatement
 Pr(X<x) Pr(X>x)
 Pr(x1<X<x2) Pr(X<x1 ou X>x2)

Limits: 0.96, 1.28
 Probability: 0.95

Compute Statistics
 display limits g1/g2

Males: limit 2, distribution 3, probability 2
 Two-sided 0.95 confidence intervals
 Limits 80%, 90%, 95%, 99%, 99.9% ("equal-tailed intervals")

CurYe
 p(x)
 Pr(X<x)
 Pr(X>x)

Options

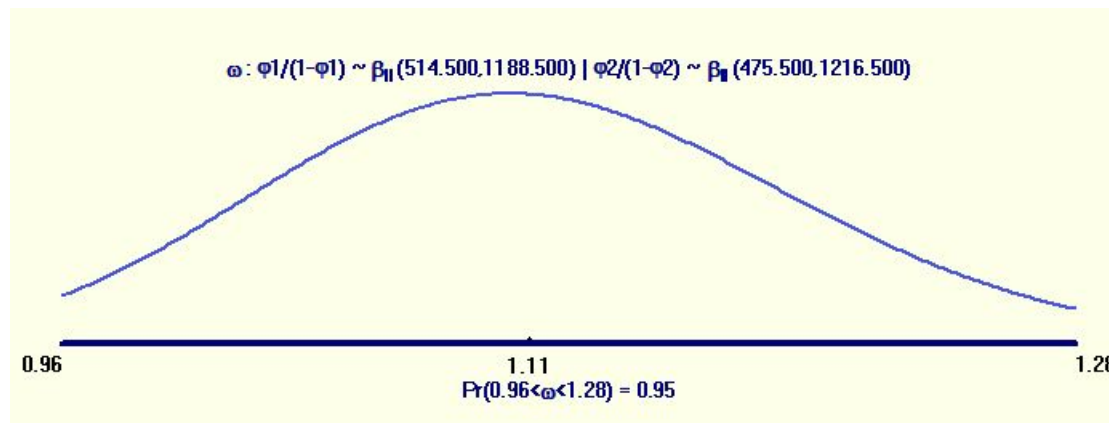
$\omega : \varphi_1 / (1 - \varphi_1) \sim \beta_{11}(514.500, 1188.500) \mid \varphi_2 | (1 - \varphi_2) \sim \beta_{11}(475.500, 1216.500)$

mean: 1.111
 standard deviation: 0.084

generate a sample
 10000

Плотность распределения и 95%-й ДИ для оцениваемого отношения оддов (шансов за/против),

$$OR_{\text{unkn}} = \omega = [\varphi_1 / (1 - \varphi_1)] : [\varphi_2 / (1 - \varphi_2)]$$



$$OR = {}_{0,96}1,11_{1,28}$$

95%, 99% и 99,9% ДИ для оцениваемого отношения оддов (шансов за/против) $OR = \omega = [\varphi_1 / (1 - \varphi_1)] : [\varphi_2 / (1 - \varphi_2)]$

| Уровень доверия | ДИ | | |
|-----------------|---------|---------|--------|
| | Границы | | Ширина |
| | Нижняя | Верхняя | |
| 95% | 0,96 | 1,28 | 0,061 |
| 99% | 0,91 | 1,35 | 0,080 |
| 99,9% | 0,86 | 1,42 | 0,102 |

Когда доли равны, то отношение оддов равно единице: $OR = \omega = [\varphi_1 / (1 - \varphi_1)] : [\varphi_2 / (1 - \varphi_2)] = 1$.

Все три полученных ДИ для оцениваемого отношения оддов OR_{unkn} содержат значение $OR = 1$.

Это дает нам основание утверждать, что, скорее всего, оцениваемое этими интервалами неизвестное нам значение OR_{unkn} статистически не отличается от 1, соответственно, первая и вторая доли статистически одинаковы.

Основной вывод: Молитва, скорее всего, не влияет на смертность при сепсисе.

Результаты

- Смертность в опытной группе была примерно на 2% ниже, чем в контрольной, однако наблюдаемое различие между долями φ_1 и φ_2 является статистически незначимым, т.е. оказывается **кажущимся**.
- $\varphi_1 = {}_{0,27}^{0,30}_{0,34}$
- $\varphi_2 = {}_{0,25}^{0,28}_{0,32}$
- $RD = \delta = \varphi_1 - \varphi_2 = {}_{-0,030}^{0,021}_{0,072}$ содержит значение **0**.
- $RR = \tau = \varphi_1 / \varphi_2 = {}_{0,90}^{1,07}_{1,28}$
- $OR = \omega = [\varphi_1(1 - \varphi_1)] / [\varphi_2(1 - \varphi_2)] = {}_{0,86}^{1,11}_{1,42}$ – оба содержат значение **1**.

Что такое *NNT* – количество подлежащих воздействию?

- **NNT – Number Needed to Treat**
- **Среднее количество пациентов, которых надо подвергнуть (данному) воздействию, дабы предотвратить один неблагоприятный исход**
- **(или получить один дополнительный благоприятный исход)**
- **по сравнению с контрольной группой (без данного воздействия).**

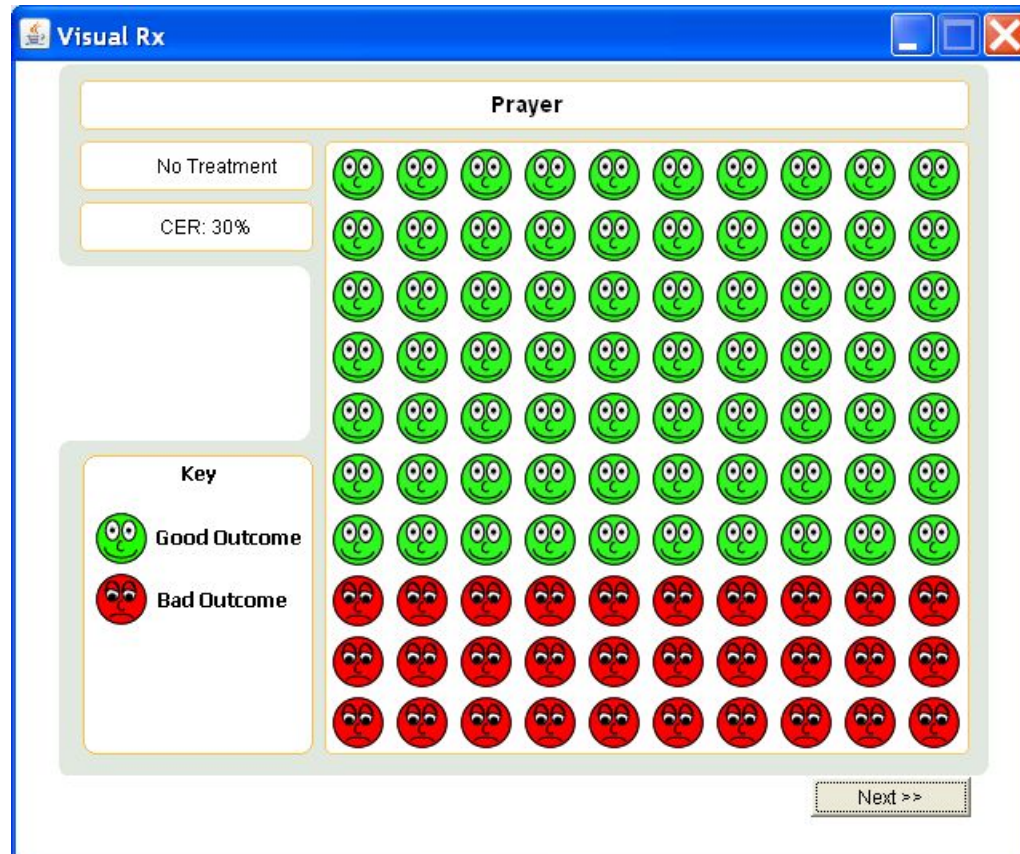
Прочувствуйте разницу

- **Утверждение:**
- **«необходимо подвергнуть данному воздействию 50 пациентов, чтобы предотвратить один неблагоприятный исход»**
- **информативнее и понятнее, нежели:**
- **«данное воздействие снижает риск неблагоприятного исхода на 0,02»**

- Относительные меры эффекта OR , RR , часто приводят к впечатляющим цифрам, даже когда абсолютные эффекты воздействия (RD) оказываются малыми
- Примеры:
 - 1. $\varphi_1 = 0,6$; $\varphi_2 = 0,1$; $RR = 6$; $OR = 13,5$;
 - $RD = 0,5$; $NNT = 2$
 - 2. $\varphi_1 = 0,06$; $\varphi_2 = 0,01$; $RR = 6$; $OR = 110,06$; **НО**
 - $RD = 0,05$ и $NNT = 20$

Программа Visual Rx

<http://www.nntonline.net/visualrx/>



Вербальные шкалы

Надежность доверительных интервалов (ДИ)

| Уровень значимости α | Уровень доверия $100(1 - \alpha)\%$ | Степень надёжности ДИ |
|--------------------------------|--|-----------------------|
| 0,05 | 95% | Низкая |
| 0,01 | 99% | Средняя |
| 0,001 | 99,9% | Высокая |

Возможные словесные интерпретации для градаций Se и Sp

| | |
|-------------------|--|
| $Se = P(T+ D+)$ | Чувствительность «позитивов» к наличию болезни |
| 0,0 – 0,5 | Практически бесполезная |
| 0,5 – 0,7 | Низкая |
| 0,7 – 0,9 | Средняя |
| 0,9 – 1,0 | Высокая |
| $Sp = P(T- D-)$ | Специфичность «негативов» в отношении отсутствия болезни |

Возможные словесные интерпретации для градаций *PPV* и *NPV*

| $PPV = P(D+ T+)$ | Способность «позитивов» предсказывать наличие болезни |
|--------------------|--|
| 0,0 – 0,5 | Практически бесполезная |
| 0,5 – 0,7 | Низкая |
| 0,7 – 0,9 | Средняя |
| 0,9 – 1,0 | Высокая |
| $NPV = P(D- T-)$ | Способность «негативов» предсказывать отсутствие болезни |

Принятые словесные интерпретации для градаций $LR[+]$ и $LR[-]$

| | |
|------------|---|
| $LR[+]$ | Повышение посттестовых шансов за/против наличия болезни у субъекта с позитивом по сравнению с претестовыми шансами за/против наличия у него болезни |
| 1 – 3 | Практически ничтожное |
| 3 – 10 | Малое |
| 10 – 33 | Среднее |
| 33 – 100 | Высокое |
| 100 – 1000 | Очень высокое |
| > 1000 | Практически идеальное |
| $LR[-]$ | Повышение посттестовых шансов за/против отсутствия болезни у субъекта с негативом по сравнению с претестовыми шансами за/против отсутствия у него болезни |

Словесные интерпретации для градаций AUC

| Интервал AUC | Способность диагностического теста распознавать наличие или отсутствие болезни |
|--------------|---|
| 1,0 – 0,9 | Отличная |
| 0,8 – 0,9 | Хорошая |
| 0,7 – 0,8 | Удовлетворительная |
| 0,6 – 0,7 | Посредственная |
| 0,5 – 0,6 | Неудовлетворительная |

Традиционная интерпретация значений P_{val} и шкала Michelin

| Значение P_{val} | Статистическая значимость | Шкала Мишлена |
|--------------------|---------------------------|---------------|
| $> 0,05$ | Незначимо | |
| $0,05 - 0,01$ | Умеренно значимо | * |
| $0,01 - 0,001$ | Значимо | ** |
| $< 0,001$ | Высоко значимо | *** |

Калибровка P -значений

| P -значение | Нижняя граница для вероятности нулевой гипотезы $P(H_0)$ | Верхняя граница для вероятности воспроизведения P_{repr} |
|---------------|--|---|
| 0,05 | > 30% | < 50% |
| 0,01 | > 10% | < 73% |
| 0,001 | > 2% | < 90% |

Для наглядности значения в таблице округлены до первой значащей цифры. Более точно значения для $P(H_0)$ (сверху вниз) равны 29%, 11% и 1,8%.

Posavac E.J. Using p values to estimate the probability of statistically significant replication // Understanding Statistics, 2002. – Vol. 1. – No. 2. – P. 101-112.

Интерпретация убедительности Бейзовых факторов, BF_{10} и BF_{01}

| BF_{01} | Свидетельство в пользу гипотезы H_0 против гипотезы H_1 |
|-----------|--|
| >100 | Убедительное |
| 30 – 100 | Очень сильное |
| 10 – 30 | Сильное |
| 3 – 10 | Умеренное (слабое) |
| 1 – 3 | Пренебрежимо малое |
| BF_{10} | Свидетельство в пользу гипотезы H_1 против гипотезы H_0 |

Интерпретация стандартизированного размера эффекта по Коуэну d_c

<http://www.sportsci.org/resource/stats/>

| Размер эффекта, d_c | Градация эффекта |
|-----------------------|-----------------------|
| 0 – 0,2 | Ничтожный |
| 0,2 – 0,5 | Малый |
| 0,5 – 1,0 | Средний |
| 1,0 – 2,0 | Большой |
| 2,0 – 4,0 | Очень большой |
| 4,0 - ∞ | Исключительно большой |

Словесная интерпретация для градаций модуля разности долей $|RD|$ и для числа субъектов, подлежащих воздействию NNT

| $ RD $ | NNT | Интерпретация клинического эффекта |
|--------------|-----------|--|
| $< 0,05$ | >20 | Ничтожный |
| $0,05 - 0,1$ | $10 - 20$ | Малый |
| $0,1 - 0,2$ | $5 - 10$ | Умеренный |
| $0,2 - 0,5$ | $2 - 5$ | Высокий |
| $> 0,5$ | < 2 | Очень высокий |

Словесная интерпретация
(вербальная шкала) градаций для отношения долей
RR

| RR | Интерпретация клинического эффекта |
|------------------|---|
| 1,0 – 3,0 | Практически ничтожный |
| 3,0 – 10 | Слабый |
| 10 – 33 | Умеренный |
| 33 – 100 | Сильный |
| > 100 | Очень сильный |

**Словесная интерпретация
(вербальная шкала) градаций для отношения
шансов *OR***

| <i>OR</i> | Интерпретация силы статистической связи |
|------------------|--|
| 1 – 1,5 | Практически ничтожная |
| 1,5 – 3,5 | Очень слабая |
| 3,5 – 9,0 | Слабая |
| 9,0 – 32 | Умеренная |
| 32 – 360 | Сильная |
| > 360 | Практически идеальная |

Спасибо за внимание!
Слайды доступны для всех

**Никита Николаевич Хромов-
Борисов**

**Кафедра физики, математики и
информатики ПСПбГМУ им. акад. И.П.
Павлова**

Nikita.KhromovBorisov@gmail.com

8-952-204-89-49