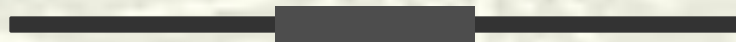


# Планирование исследования. Определение размера выборки



# Планирование научных исследований

---

- Определение целей
  - Определение задач
  - Определение переменных и методов их измерения
  - Выбор дизайна исследования
  - Выбор метода статистической обработки
  - Определение размера выборки
  - Написание протокола исследования и создание ИРК участника
-

# Напомним

---

- Контроль систематической ошибки
    - Дизайн
  - Контроль случайной ошибки
    - Размер выборки
-

# Задача науки – предсказывать будущее



Основные конкуренты – маги и чародеи

# Предсказывать легко

---

- Ручка всегда падает на пол
  - Тело впернутое в воду, выпирает на свободу с силой выпертой воды, телом впернутом туды
  - Но это не всегда так легко...
-

# Индивидуум и группа, проблема предсказаний

---

- Для каждого конкретного человека есть только одно значение
  - Однако у разных людей значения немного различаются и мы не можем объяснить, почему
  - Или они отличаются у одного человека и мы не можем объяснить почему
  - Или можем, но использовать данную информацию для предсказаний будет очень накладно
-

# Проблема предсказаний

---

- Детерминизм – маги и чародеи
  - Стохастический подход – наука
-

# Случайные колебания - вариабельность

---

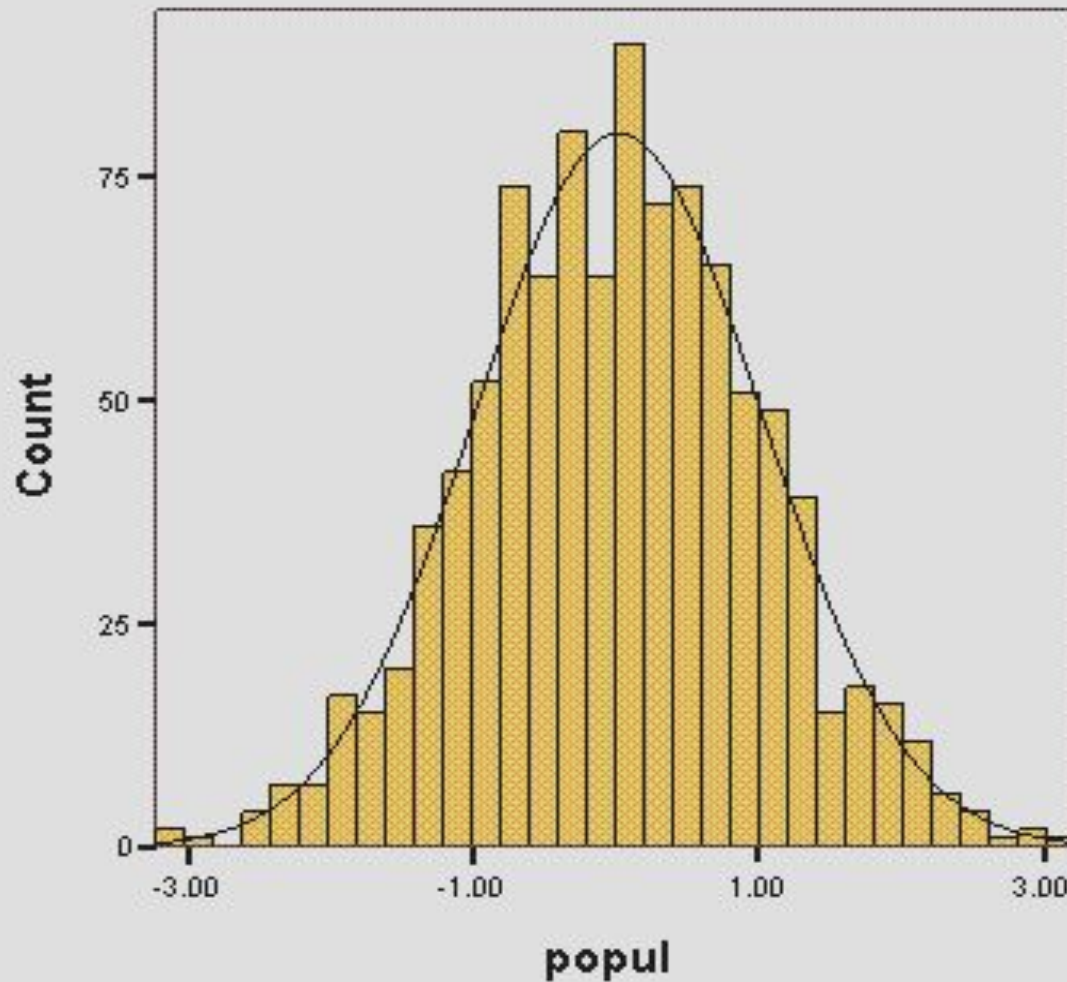
- Есть истинный параметр
  - $\mu$
- И есть множество небольших факторов, которые на него влияют
- Соответственно, наблюдаемое значение
  - $Y_i = \mu + \varepsilon_i$

Если записать все возможные для данной популяции значения  $Y$  и частоту их встречаемости, то получим распределение

---



# Случайная вариабельность - распределение



# Случайные колебания

---

- Неприятный факт: измеренный показатель отличается от истинного, причем будет отличаться всегда
- Однако, колебания случайные, т.е. с одинаковой вероятностью вверх и вниз
- Тогда  $\sum \varepsilon_i = 0$

# Соответственно

---

- Если сложить все значения в популяции
  - $\sum Y_i = \sum (\mu + \varepsilon_i) = N * \mu + \sum (\varepsilon_i) = N * \mu$
  - Соответственно,
  - $\mu = \sum Y_i / N$
- 
- Иными словами, в случае случайной вариабельности популяционное среднее равно истинному значению в популяции
-

# На самом деле...

---

- Мы в реальность предположили, что у нас есть два фактора
    - Истинное значение (которое нас интересует) –  $\mu$
    - Случайная ошибка, варьирующая от наблюдения к наблюдению –  $\varepsilon$
  - Для оценки случайной ошибки надо знать ее распределение.
-

# Распределение фактора случайной ошибки

---

- По определению, среднее значение равно 0
  - Соответственно, остается описать форму распределения (симметричное, несимметричное, колоколообразное, прямоугольное) и ширину разброса
-

# Распределение фактора случайной ошибки

---

- Форма распределения определяет тип
    - Нормальное
    - Экспоненциальное
    - Вейбулла
    - Пуассона и т.д.
  - Ширина разброса определяется стандартным отклонением  $\sigma$ .
-

# Обратите внимание

---

- Истинные параметры (популяционные) обозначаются греческими буквами
  - Выборочные (позднее) - латинскими
-

# Стандартное отклонение

---

- $\sigma = \text{SQRT}(\Sigma(Y_i - \mu)^2 / (n - 1))$

Надо помнить, что иногда  $\sigma$  определяется другими формулами, например в случае распределения Пуассона  $\sigma = \mu$ , в случае большого числа испытаний биномиального распределения  $\sigma = \pi * (1 - \pi)$

---



# Соответственно, описание данных

---

- Истинный параметр
  - $\mu$
- Стандартное отклонение
  - $\sigma$
- Тип распределения

Например,  $N(120,20)$ . Распределение ошибки -  $N(0,\sigma)$

---

# Что нам это дает?

---

- Распределения и теория диагностики
-

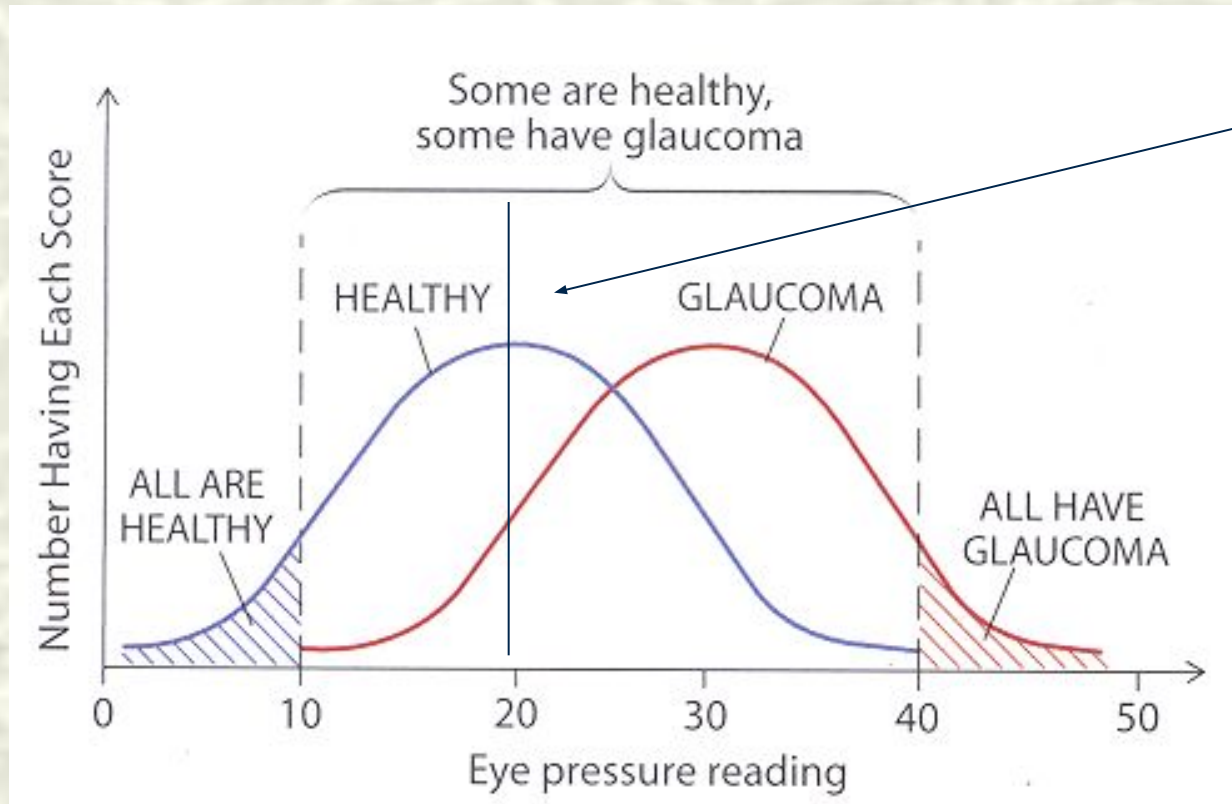
# Конкурирующие гипотезы

---

- Есть заболевание (1) или нет (2)
  - Предположим, известно глазное давление. У пациента может быть глаукома (справедлива гипотеза 2). А может быть и нет (справедлива гипотеза 1).
-

# Конкурирующие гипотезы

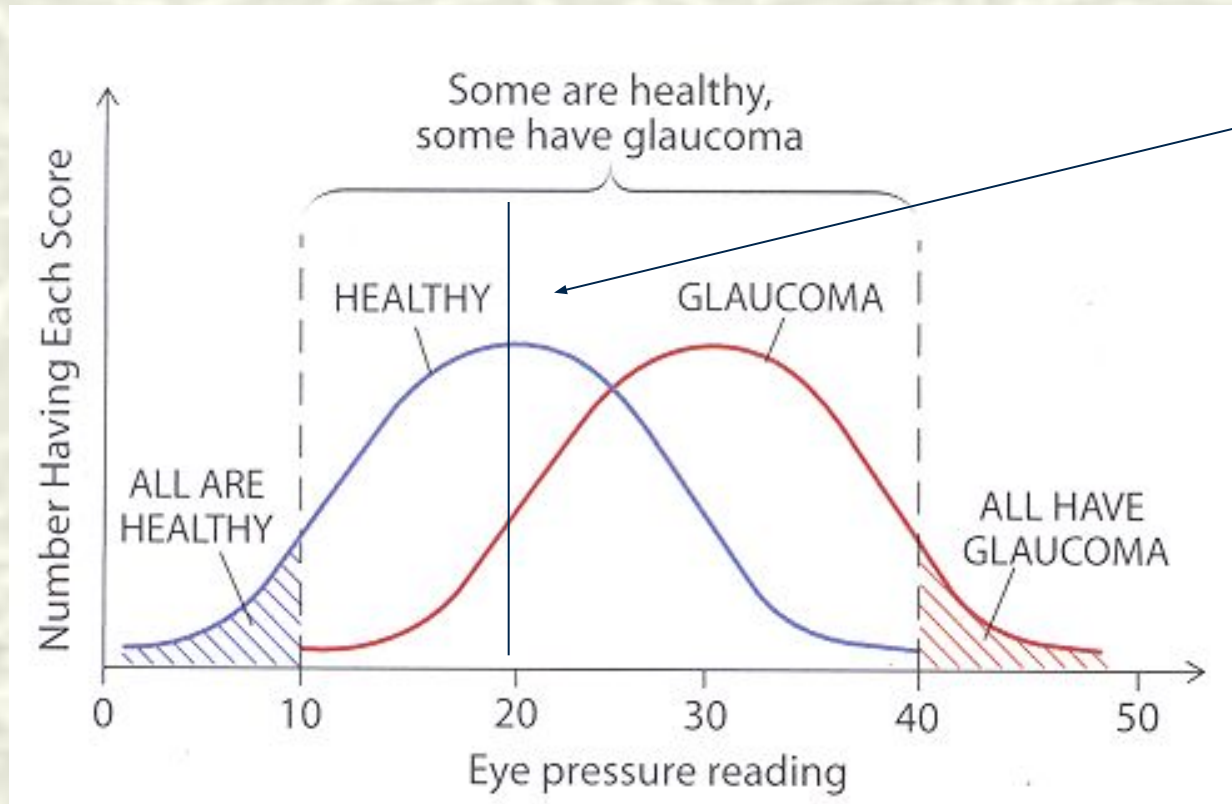
- Известно распределение



Значение  
у  
пациента

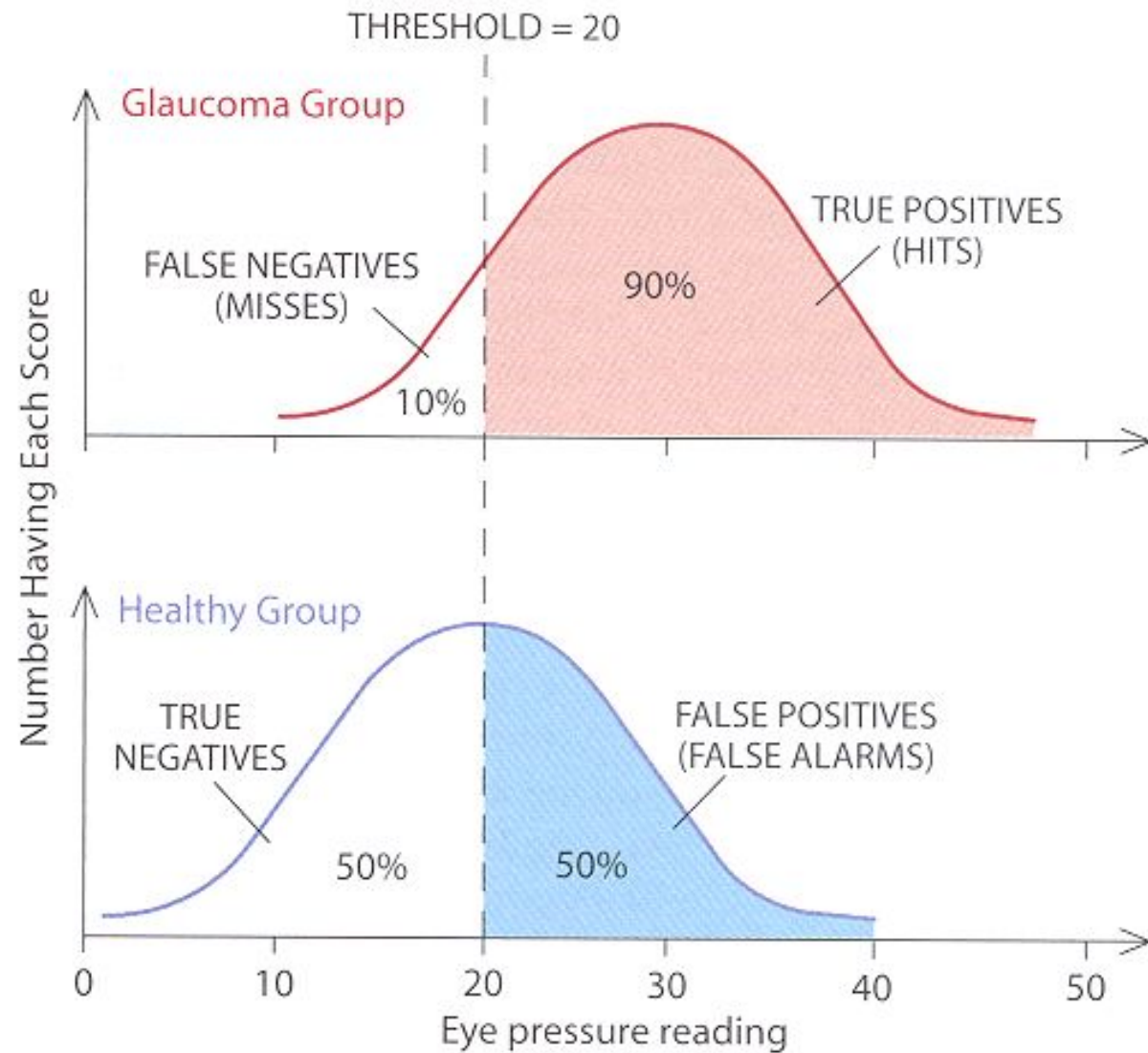
# Конкурирующие гипотезы

- Известно распределение



Значение  
у  
пациента

## 2. Подсчитываем вероятности справедливости каждой гипотезы



# Далее

---

- Два варианта
    - Принимаем решение по каждому случаю индивидуально, на основании рассмотрения другой информации
      - Байсовский подход
    - Заранее устанавливаем границы нормы
      - Фреквентистский подход (стандарт в обычной статистике и планировании эксперимента)
-

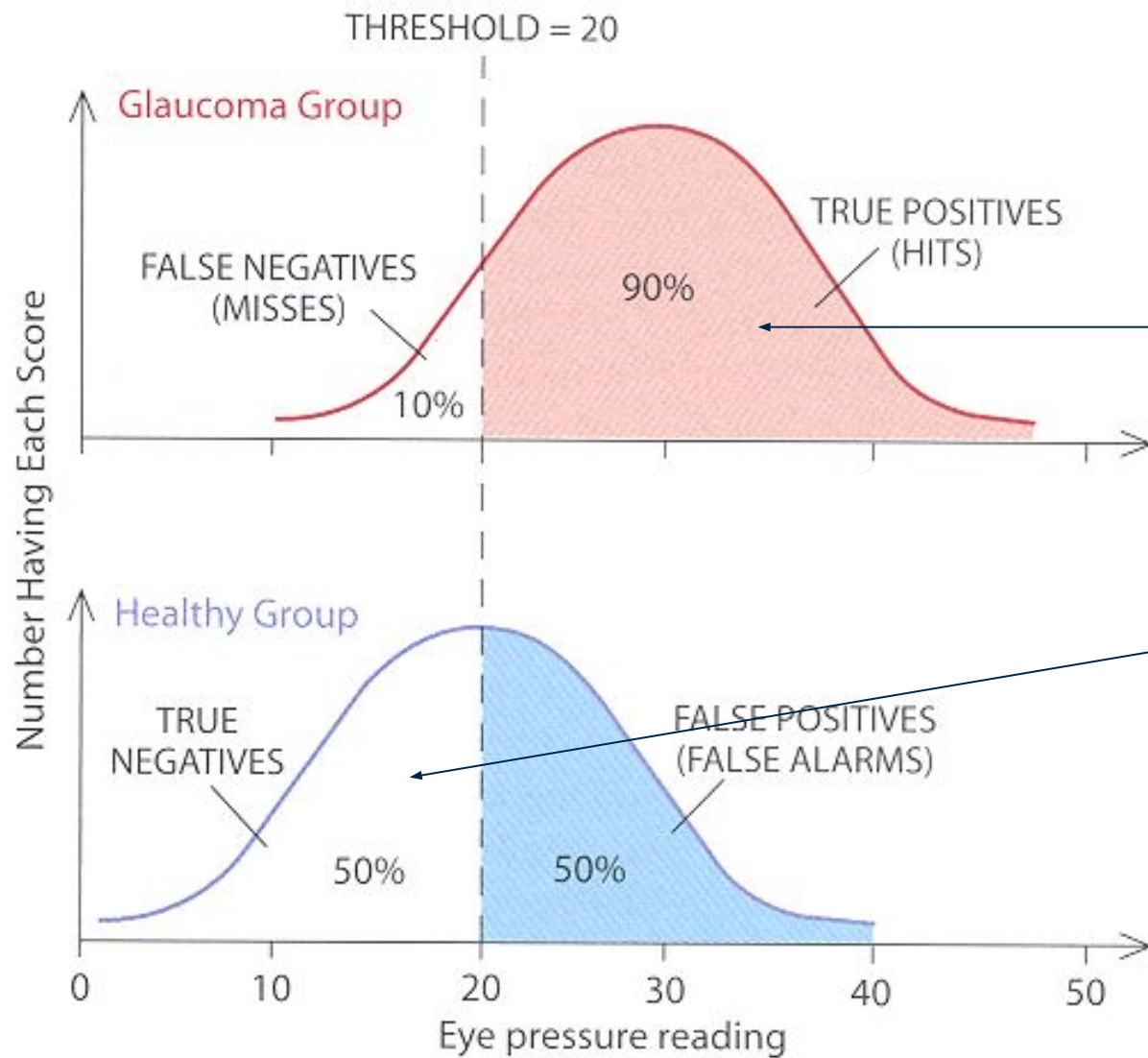
# Установление границ нормы

---

- На самом деле устанавливаем произвольную точку разделения, которая может быть охарактеризована с точки зрения
    - Количества ложноположительных результатов
    - Количества истинно положительных результатов (чувствительность)
    - Количества ложноотрицательных результатов
    - Количества истинно отрицательных результатов (специфичность)
-



2. Подсчитываем вероятности справедливости каждой гипотезы



чувствительность

специфичность

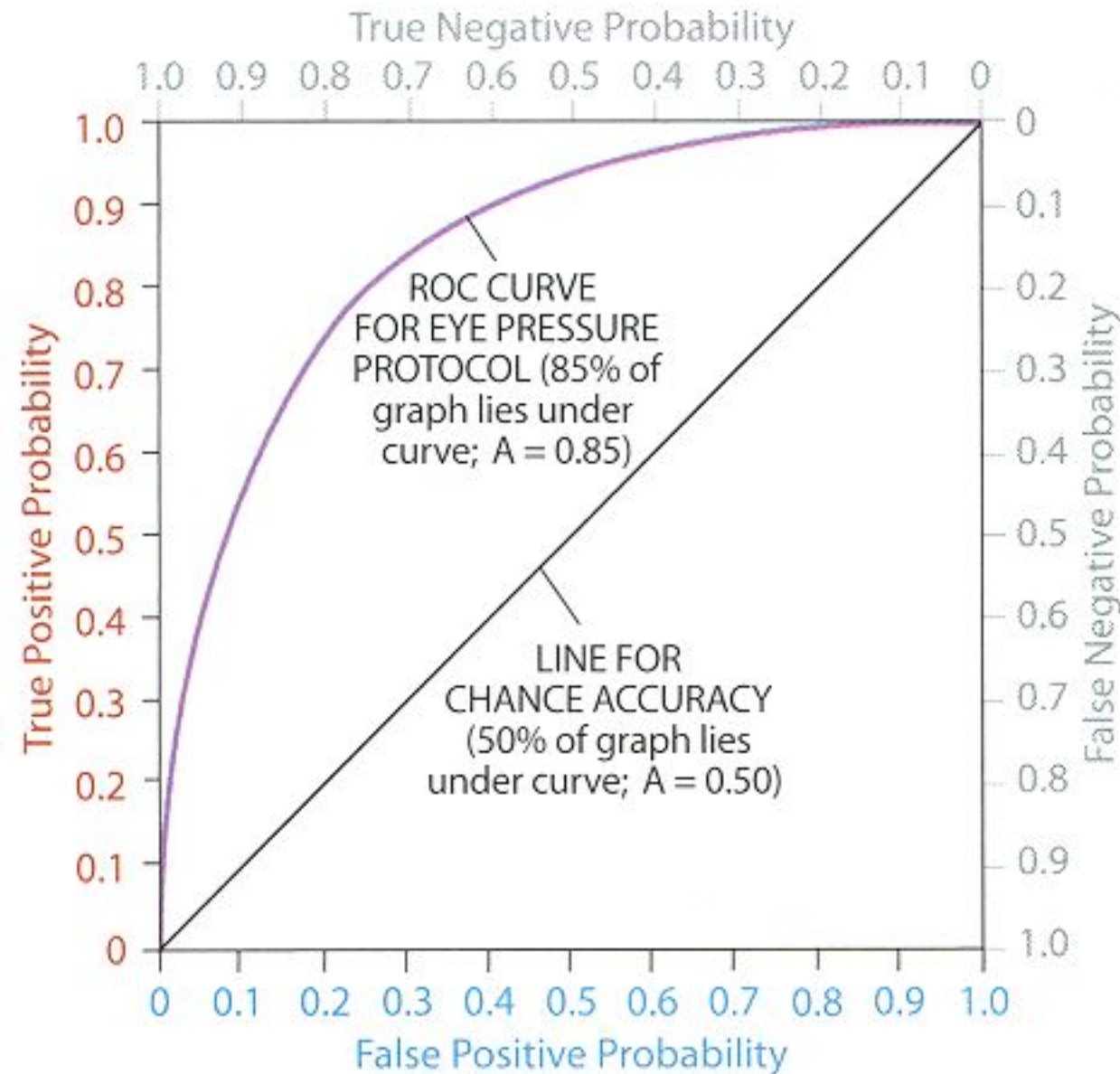
# Se и Sp

- Чувствительность
- Специфичность

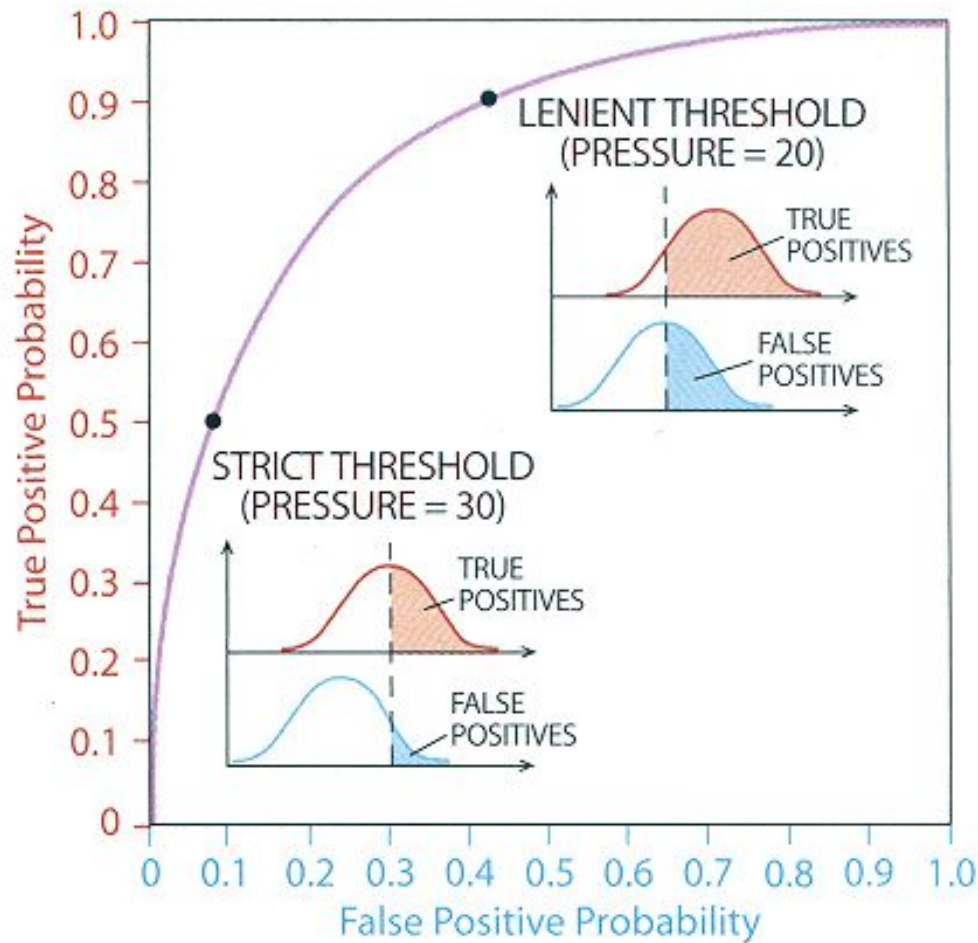
	Есть заболевание	Заболевания нет
Тест +	80	100
Тест-	20	800

$$Se = \frac{TP}{TP + FP} = \frac{80}{80 + 20} = 0.8$$

$$Sp = \frac{TN}{TN + FN} = \frac{800}{800 + 100} = 0,89$$



**3.** Создать характеристическую кривую нанеся на график для каждой диагностической границы точку, соответствующую соотношению истинно положительных против ложно-положительных результатов. Прямая линия - тест аналогичен подбрасыванию монеты. Точность теста определяется как площадь под кривой АОС



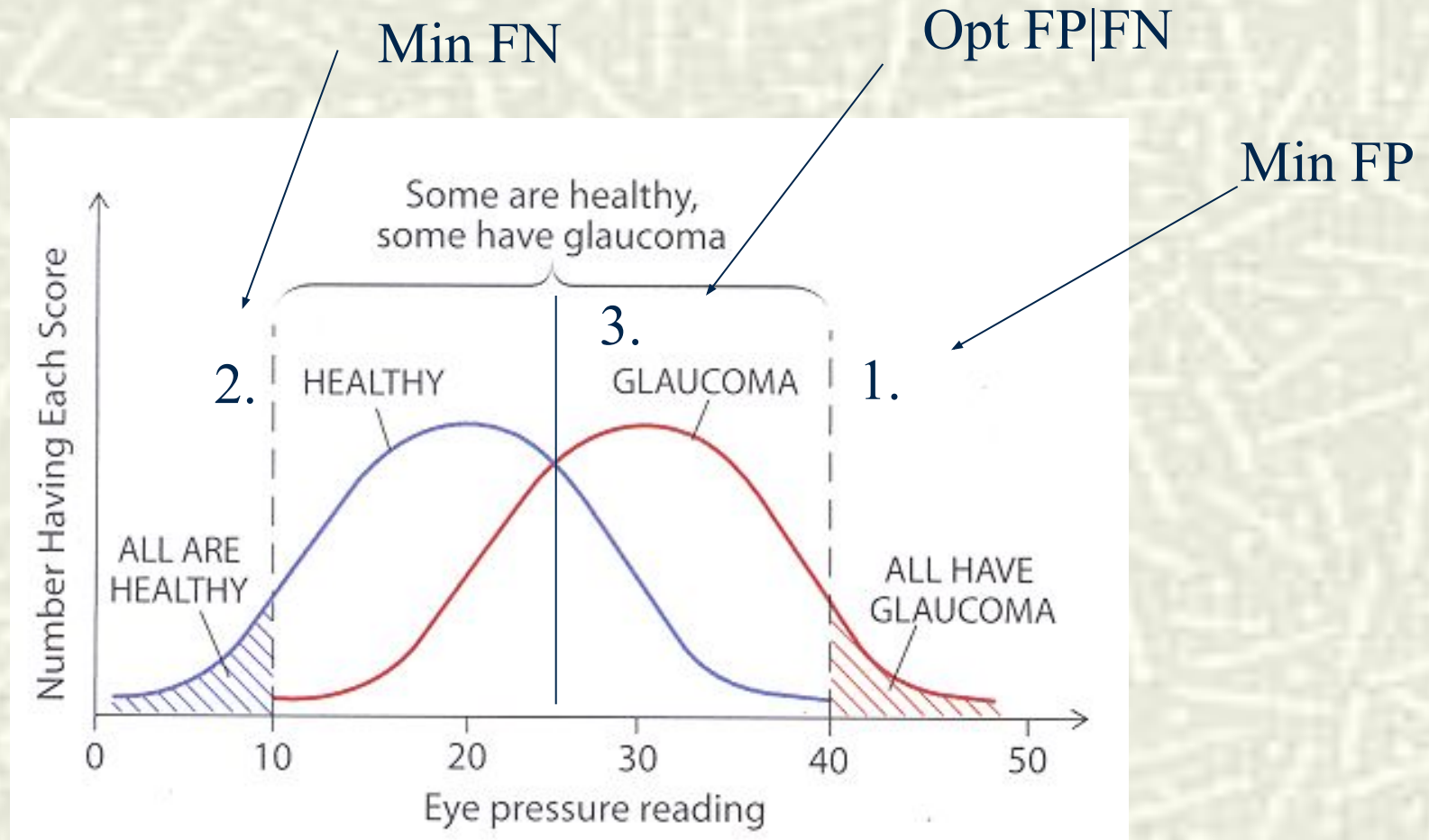
4. Если точность достаточно высока, необходимо выбрать точку разделения таким образом, чтобы было достаточное количество истинно положительных результатов без чрезвычайно большого количества ложно-положительных.

# Соответственно, выбираем точку разделения, чтобы

---

- Минимизировать количество ложноположительных
  - Минимизировать количество ложноотрицательных
  - Оптимизировать соотношение ложноположительных и ложноотрицательных значений
-

# Выбор точки разделения



# Выбранная точка разделения

---

- Используется для диагностики.
  - Обратите внимание, что тут, несмотря на вероятности не было статистики. Мы предполагали, что нам известны популяционные значения
-

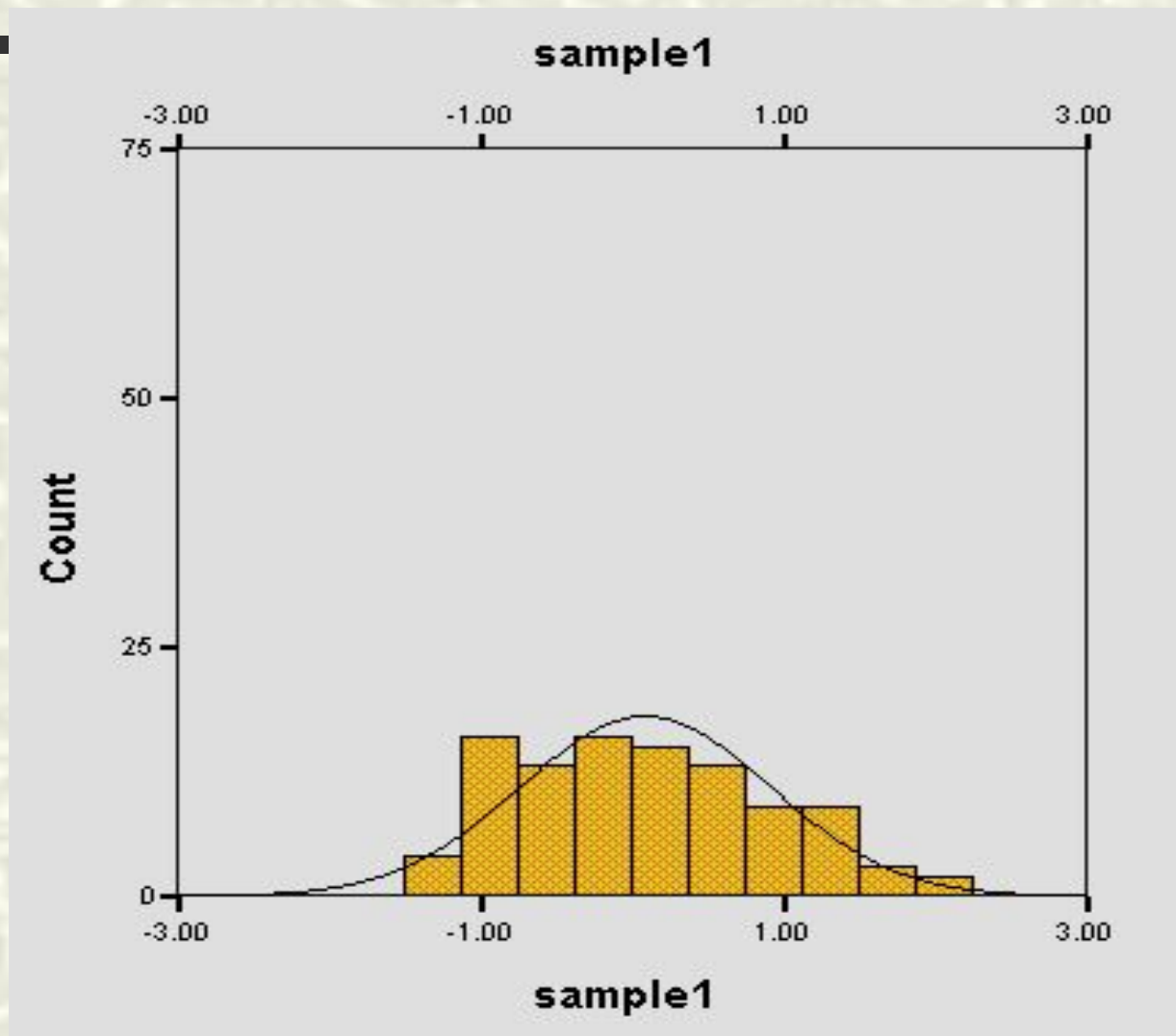
# Однако

---

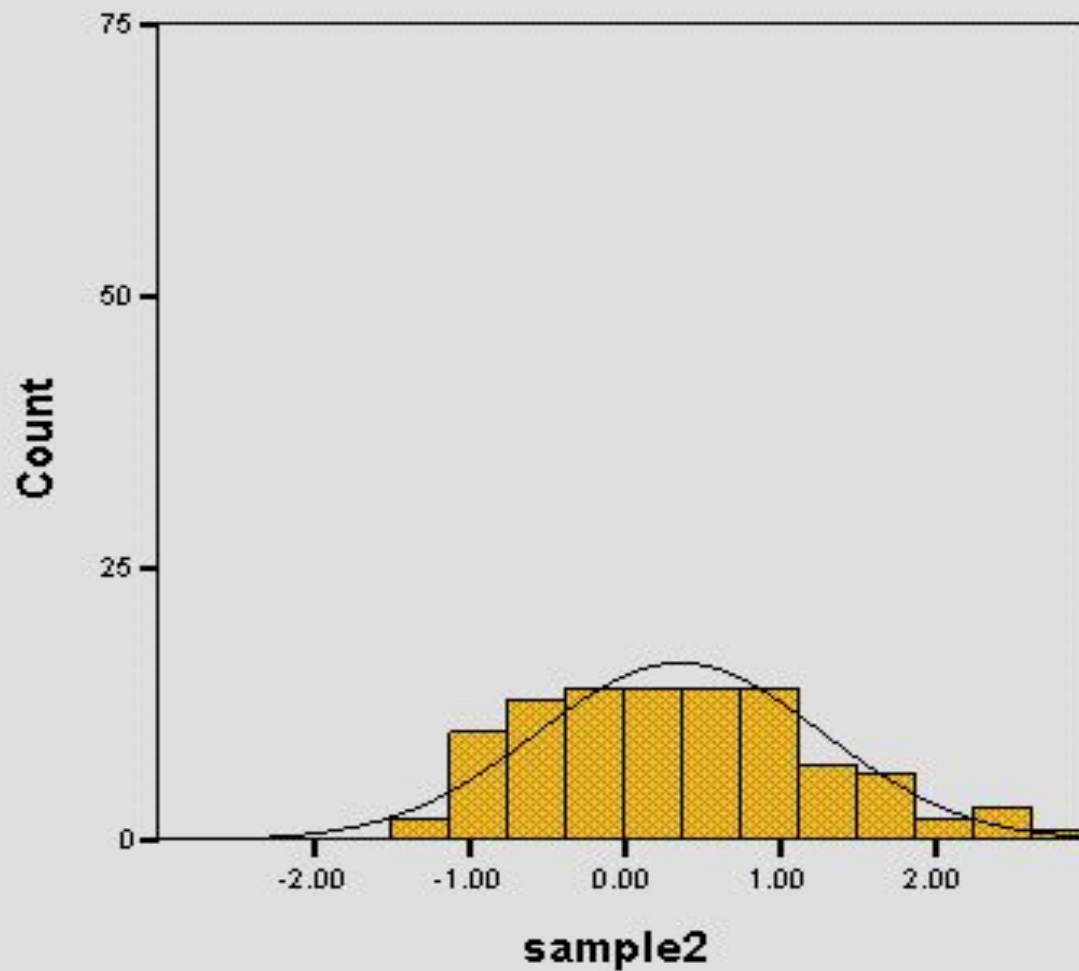
- Мы редко работаем с популяцией (генеральной совокупностью) в целом
    - Это экономически невыгодно
    - Выводы нельзя будет распространять на другие популяции
  - Для работы из популяции берется выборка (выборочная совокупность)
-



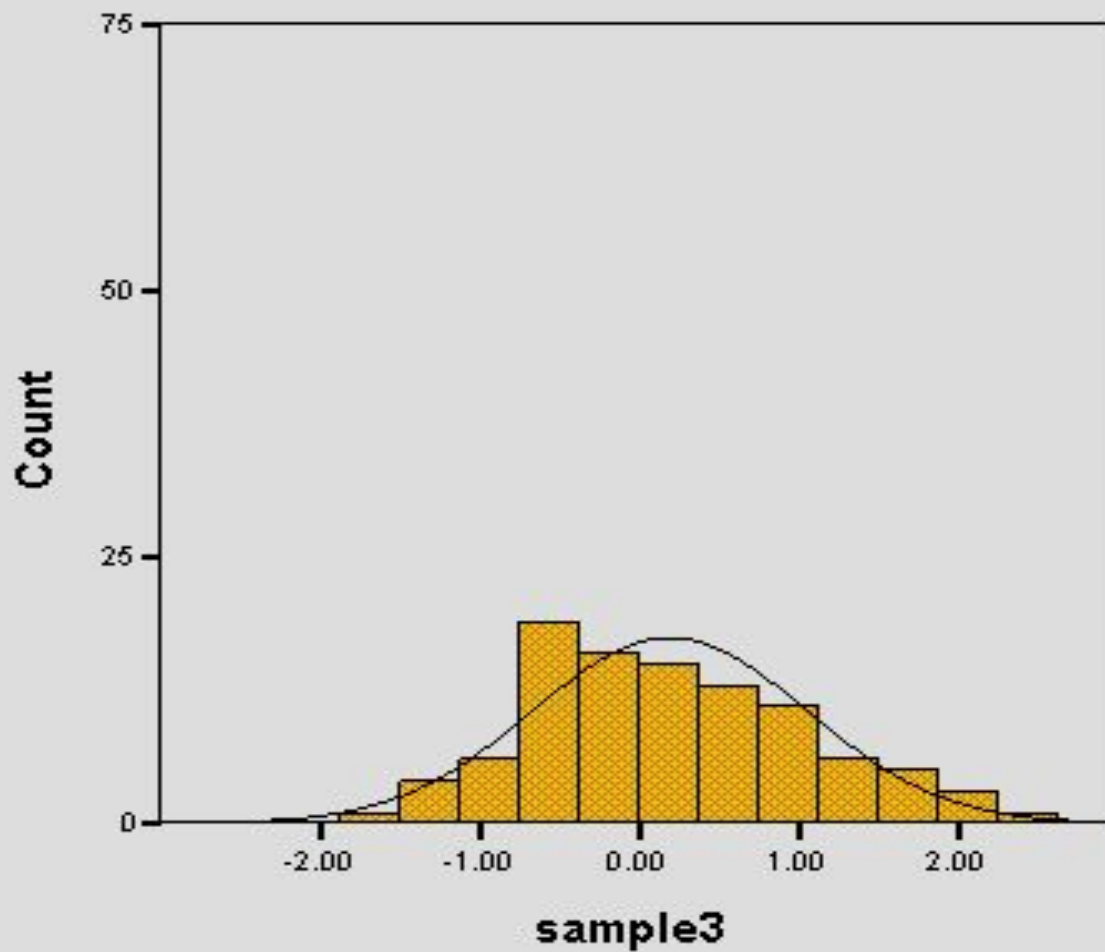
# Выборка 1



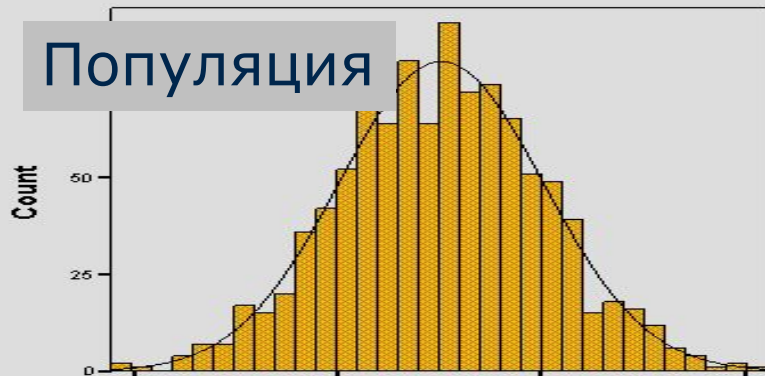
# Выборка 2



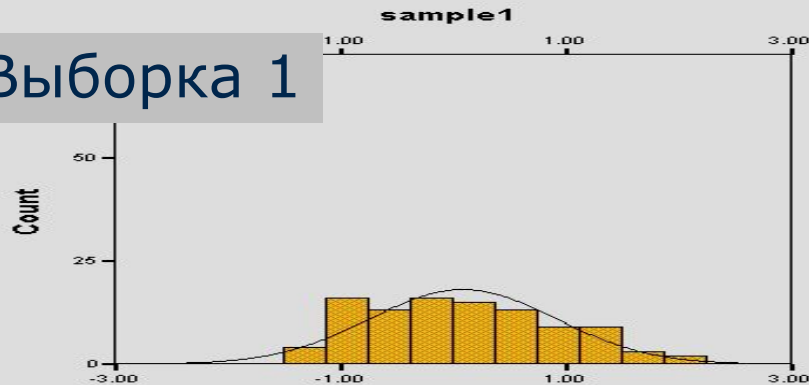
# Выборка 3



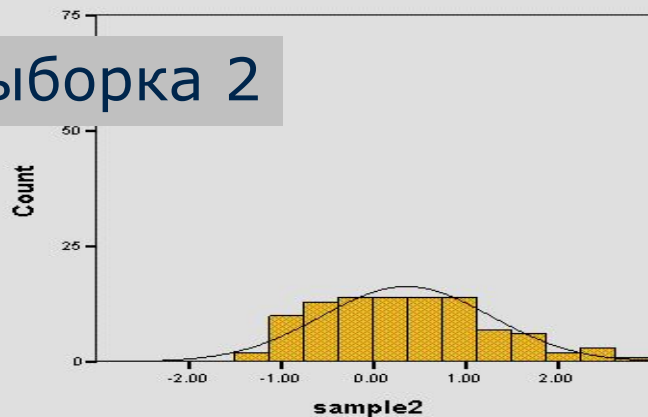
Популяция



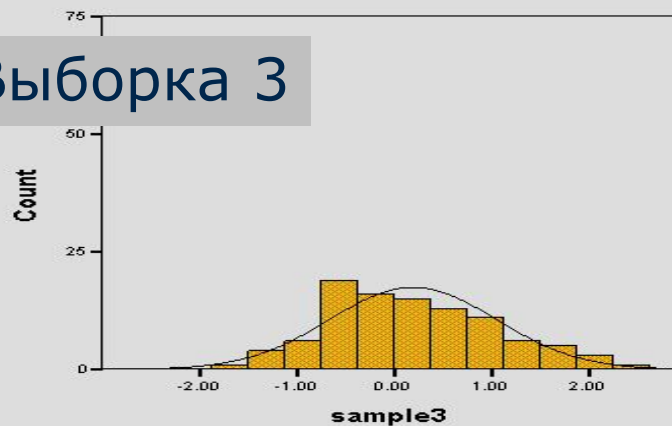
Выборка 1



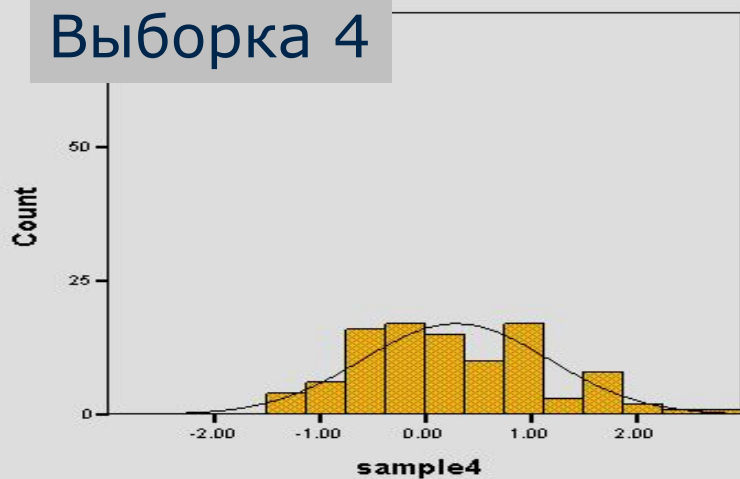
Выборка 2



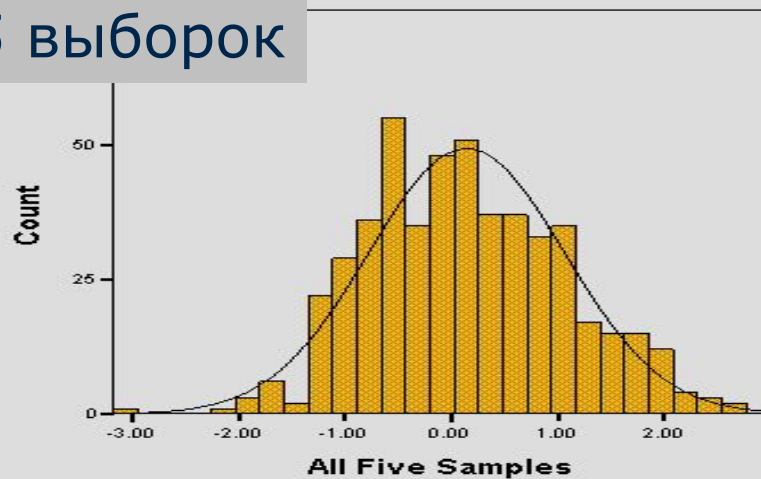
Выборка 3



Выборка 4



5 выборок



# Соответственно

---

- Для каждой выборки тоже можно рассчитать среднее (выборочное) и дисперсию (выборочную)
  - Однако в разных выборках они уже совпадать не будут!
-

Есть популяция из 100000 объектов с интересующим нас показателем, распределенным  $N(120,20)$

Сделаем 10 случайных выборок по 10 элементов в каждой. Результат - оценка средних (у вас получится немного другой):

131,1	125,7
109,2	122,4
119,8	116,6
116,3	128,0
115,6	118,9

Ни в одном случае среднее не равно 120!

Оценка стандартного отклонения:

25,8	11,6
11,8	15,0
18,0	24,4
19,2	17,3
12,5	21,7

Опять-таки ни в одном случае стандартное отклонение не равно популяционным 20. В двух случаях оно почти в два раза ниже! (т.е. если мы используем это значение, то будем считать, что 95% значений находятся в интервале 100-140, а на самом деле этот интервал 80-160).

# Соответственно,

---

- Средние и дисперсии превращаются в случайные величины
    - Для того, чтобы их отличать популяционные показатели обозначают греческими буквами
      - $\mu$ ,  $\sigma$
    - А выборочные – латинскими
      - $M$ ,  $s$
-



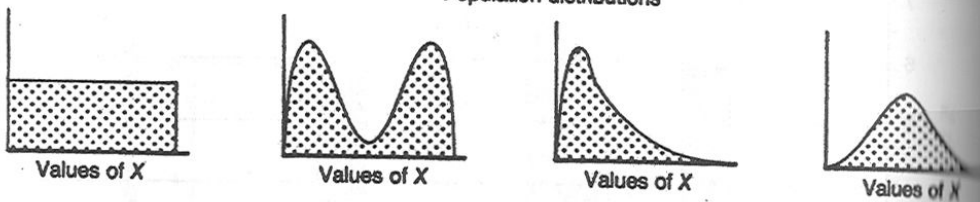
- Выборочные показатели меняются от исследования к исследованию
- Однако они уже результат усреднения
  - $m = \Sigma(\mu + \varepsilon_i) / N$ 
    - Это не равно  $\mu$ , поскольку  $\Sigma(\varepsilon_i)$  не равно нулю, но чем больше  $N$  тем оно будет ближе
- Поэтому суммарные выборочные показатели варьируют меньше, чем исходные данные

# Центральная предельная теорема

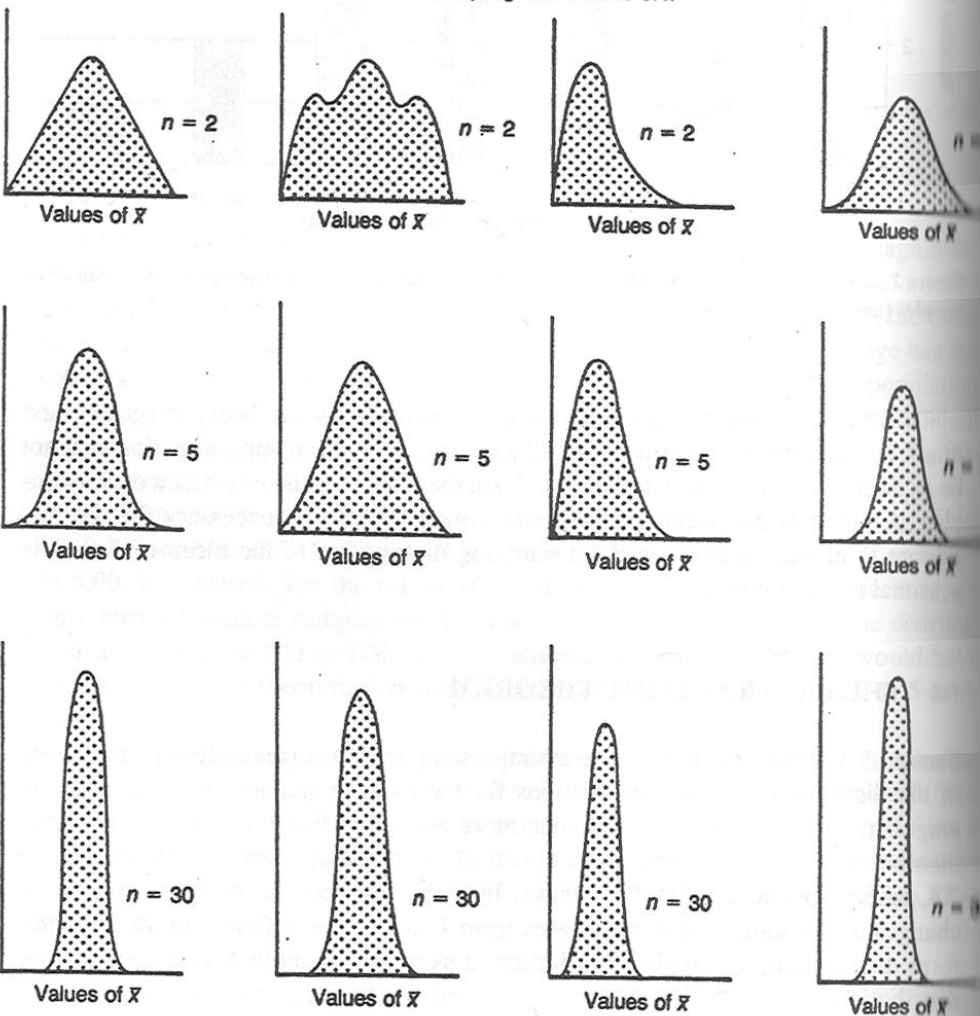
---

- Вне зависимости от формы исходного распределения, если размер выборки достаточно велик, распределение выборочных средних подчиняется нормальному закону и дисперсия выборочных средних в  $n$  раз меньше, чем дисперсия индивидуальных значений в популяции
-

Population distributions



Sampling distributions of  $\bar{x}$



# Центральная предельная теорема

В случае большого  
количества наблюдений  
распределение выборочных  
средних становится  
нормальным

Figure 7.4. The effect of shape of population distribution and sample size on the distribution of means of random samples. (Source: Kuzma, J. W. *Basic Statistics for the Health Sciences*. Mountain View, California: Mayfield Publishing Company, 1984, Figure 7.3, p. 82.)

# Нормальное распределение

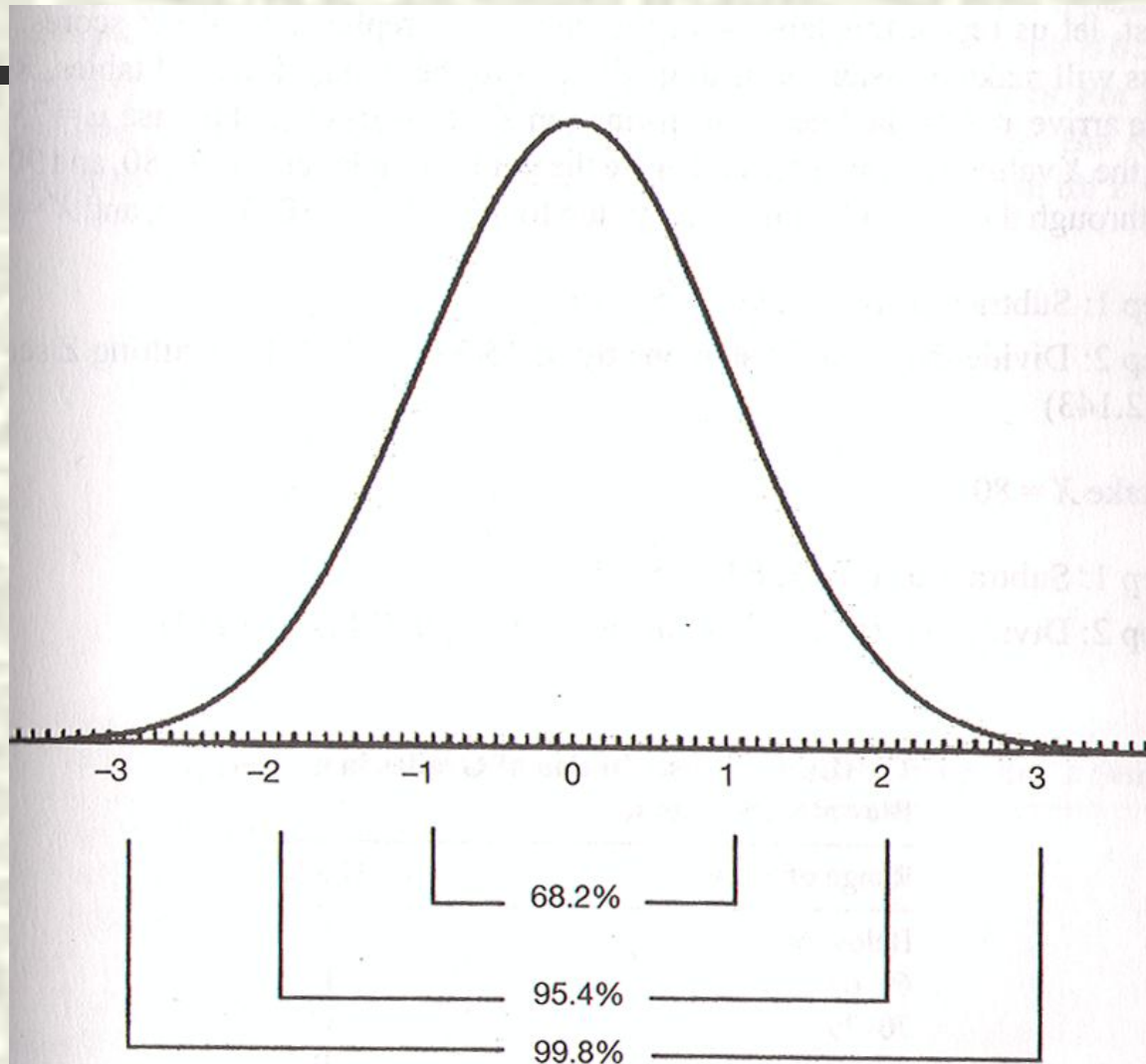


Figure 6.1. The standard normal distribution.

# Поскольку распределение нормальное

---

- Можно сделать ряд выводов о возможном распределении выборочных средних.
  - Если будет проведено много испытаний, то
    - 68% выборочных средних будут находиться в пределах  $1 \sigma$  от популяционного среднего
    - 95% выборочных средних будут находиться в пределах  $2 \sigma$  от популяционного среднего
-

# Напомним, для того, чтобы описать распределение необходимо

---

- Истинный параметр
  - Стандартное отклонение
  - Распределение
  
  - При достаточном количестве наблюдений распределение известно (по ЦПТ – нормальное).
-

# Еще раз вернемся к результатам исследования

---

- Есть выборочное среднее.
  - Есть конкурирующие гипотезы
    - Препарат не работает, отклонения связаны только со случайными факторами (нулевая гипотеза)
    - Препарат работает (альтернативная гипотеза)
  - Надо выбрать одну (как в случае с диагностикой)
-

# Для описания обеих гипотез надо

---

- Знать истинный параметр
  - Для нулевой гипотезы он прост:
    - $\mu=0$
  - Для альтернативной его надо определить
- Стандартное отклонение
  - Надо определить, стандартное отклонение для выборочного среднего ( $m$ ) по ЦПТ в корень квадратный из размера выборки меньше, чем популяционное



# Приемлемое количество FP|FN

---

- Т.е. допустимую ошибку
  - Ошибка I типа
    - Вероятность признать справедливость альтернативной гипотезы, когда это не так (обозначается  $\alpha$ ), АКА уровень достоверности
  - Ошибка II типа
    - Вероятность признать справедливость нулевой гипотезы, когда это не так (обозначается  $\beta$ )
    - Чаще используют  $1-\beta$ , называемую мощностью – вероятность найти различия, если они действительно существуют
-

# Откуда это все брать?

---

- Дисперсия
    - Пилотное исследование
    - Литературные данные
    - Ориентировочные методы (диапазон/4)
  - Предполагаемый эффект
    - Литературные данные
    - Минимальные клинически значимые различия
-

# Ошибки

---

- Обычно уровень ошибки I типа фиксируется в данной отрасли
  - 5%, 1%, 0,1%
- Мощность исследования редко допускается ниже 80%

# Итак

---

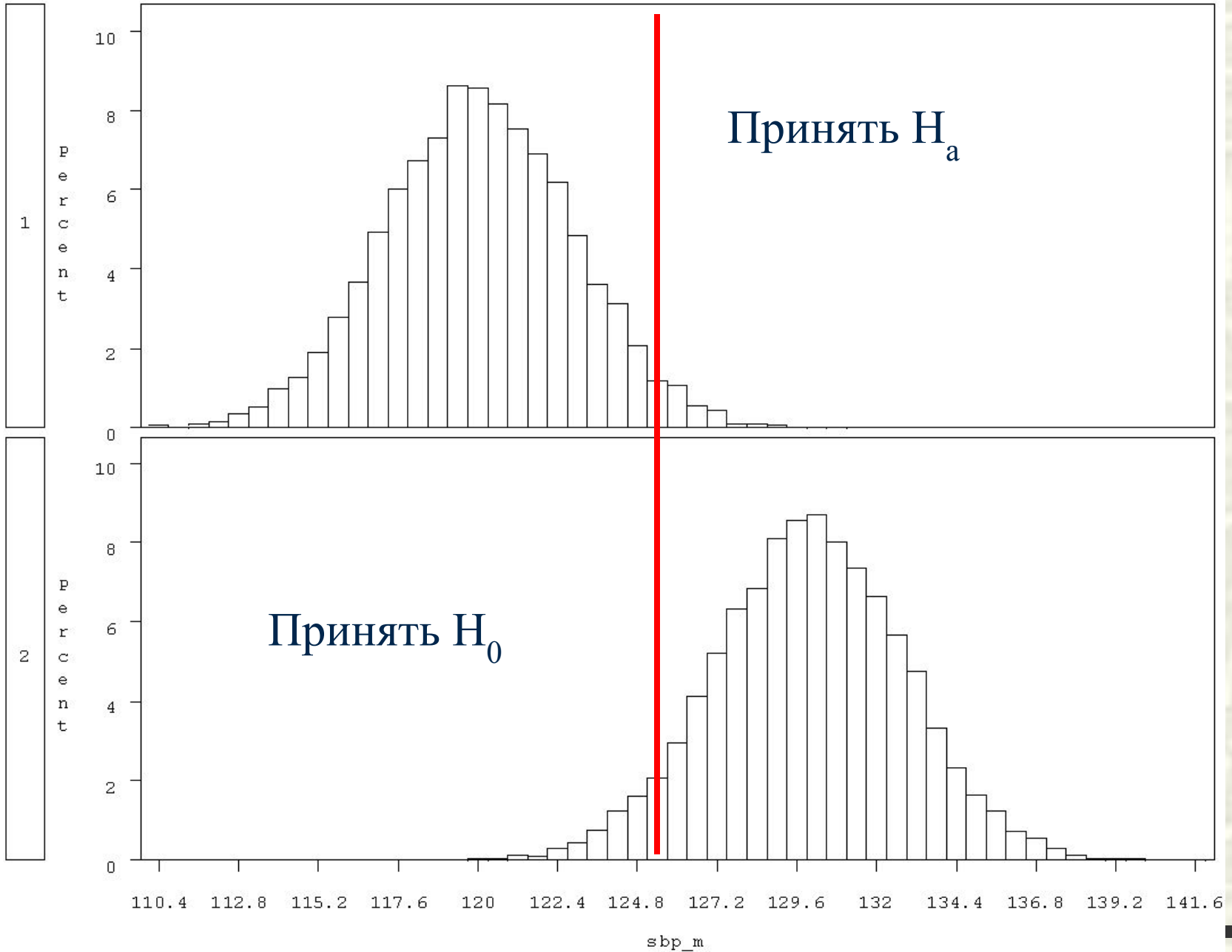
- Зная
    - Истинные параметры
    - Стандартное отклонение (в популяции)
    - Форму распределения
    - Приемлемые значения ошибок
  
  - Мы можем оценить принадлежность результата к тому или иному классу... Но у нас ведь нет варьирующего параметра?
-

# Нет, есть

---

- Поскольку стандартное отклонение выборочного показателя зависит от размера выборки, это единственный показатель, который может варьировать при планировании исследования
-

# правило



# Квантили распределения

---

- Значения распределения, для которых известны вероятности наличия значений больших данного называются квантилями
- Для нормального распределения  $N(0,1)$  обозначаются буквой  $z$
- Соответственно  $z_{0.975}=1.96$
- Очевидно, что любое нормальное распределение можно привести к виду  $N(0,1)$  если рассчитать  $z=(Y_i-\mu)/\sigma$

# Соответственно

- Но:  $N(0, \sigma/\sqrt{n})$ 
  - $z\alpha = (\mu_0 + \text{граница}) / \sigma/\sqrt{n}$
- На:  $N(X, \sigma/\sqrt{n})$ 
  - $z\beta = (\text{та-граница}) / \text{граница}$
- Отсюда, поскольку граница-то одна
- $++++0+++||++++X+++++$
- граница



# Наиболее общая формула

---

- $N = (z\alpha + z\beta)^2 * \sigma^2 / (\mu_0 - \mu_a)^2$

# Общая формула, количественный результат

Если  $\sigma$  известна

$$n = \left( \frac{z_{1-\alpha/2} + z_{1-\beta}}{\delta} \right)^2 2\sigma^2$$

Если  $\sigma$  неизвестна

$$n = \left( \frac{t_{(1-\alpha/2, 2(n-1))} + t_{(1-\beta, 2(n-1))}}{\delta} \right)^2 2s^2$$

# Таблица, количественный результат

Размер эффекта	Мощность 80%		Мощность 90%	
	n/группу	n/группу	n/группу	n/группу
	(z test)	(t test)	(z test)	(t test)
1.5	7	9	10	11
1.0	16	17	22	23
0.9	20	21	26	27
0.8	25	26	33	34
0.7	33	34	43	44
0.6	44	45	59	60
0.5	63	64	85	86
0.45	78	79	104	105
0.4	99	100	132	132
0.35	129	129	172	172
0.3	175	175	234	234
0.25	252	252	337	337
0.2	393	393	526	526
0.15	698	698	934	935
0.1	1570	1570	2102	2102

# Общая формула, качественный результат

$$N = [z_{1-\alpha/2} \sqrt{pq / \pi_1 (1 - \pi_1)} + z_{1-\beta} \sqrt{p_1 q_1 / \pi_1 + p_0 q_0 / (1 - \pi_1)}]^2 / (p_1 - p_0)^2$$

где  $q_1=1-p_1$ ,  $q_0=1-p_0$ ,  $p=(p_0+p_1)/2$ ,  $q=1-p$ , а  $\pi_1$  - это пропорция пациентов в группе получающих активное лекарство.

Если наблюдений мало:

$$N' = N / 4 [1 + \sqrt{1 + 2 / (\pi_1 (1 - \pi_1) N (p_1 - p_0))}]^2$$

# Таблица, качественный результат

а) Размеры выборки (численность одной группы) для достижения 80% мощности ( $\alpha=0,05$  двухсторонняя).

$P_0$	$P_1$									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	
5%	474	88	43	27	19	14	11	8	7	
10%		219	72	38	25	17	13	10	8	
15%	726	945	134	57	33	22	15	11	9	
20%	219		313	91	45	28	19	13	10	
25%	113	1134	1291	165	66	36	23	16	11	
30%	72	313		376	103	49	29	19	13	
35%	51	151	1417	1511	183	70	37	22	15	
40%	38	91	376		408	107	49	28	17	
45%	30	62	176	1574	1605	186	68	35	20	
50%	25	45	103	408		408	103	45	25	

б) Размеры выборки (численность одной группы) для достижения 90% мощности ( $\alpha =0,05$  двухсторонняя).

$P_0$	$P_1$									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	
5%	621	114	55	34	23	17	13	10	7	
10%		286	92	49	31	21	15	11	9	
15%	957	1252	174	73	42	27	19	14	10	
20%	286		412	119	58	35	23	16	11	
25%	146	1504	1714	216	85	46	28	19	13	
30%	92	412		496	134	63	36	23	15	
35%	65	198	1882	2008	240	90	47	28	18	
40%	49	119	496		538	140	63	35	21	
45%	38	80	230	2092	2134	244	88	44	25	
50%	31	58	134	538		538	134	58	31	

# Общая формула, время до наступления события

$$N=D/P$$

$$D = (z_{1-\alpha/2} + z_{\beta})^2 / [\pi_1(1 - \pi_1)\theta_1^2]$$

$\theta$  – логарифм ОР  
 $\pi$  – пропорция в группе лечения

$$P = 1 - \frac{1}{6} [S(f) + 4S(0.5t + f) + S(t + f)]$$

$t$  - это время, которое требуется для рекрутирования пациента и  $f$  - это время наблюдения для последнего, введенного в исследование пациента.

Во многих случаях функция выживаемости аппроксимируется экспоненциальным распределением (иными словами  $S_0(t) = \exp(-\lambda_0 t)$  и  $S_1(t) = \exp(-\lambda_1 t)$ ). Параметры  $\lambda_0$  и  $\lambda_1$  - это риски, которые могут быть получены путем указания средней медианной выживаемости.

Медианное время выживаемости для контрольной и тестовой групп составляет два и три года, соответственно. Тогда  $\lambda_0 = 0,347$   $\lambda_1 = 0,231$  и логарифм отношения риска  $\theta_1 = \log(0,347 / 0,231) = 0,407$ .

Для исследования с равным разделением на группы требуемое количество событий для того, чтобы достигнуть 90% мощности составляет  $D = 256$ .

Если исследование будет рекрутировать пациента в первый год и будет иметь трехлетний период наблюдения после того, как последний человек был введен в исследование, тогда  $P = 0,628$  и требуемый размер выборки может быть оценен как  $N = 256 / 0,628 = 408$ .

# Упрощенные формулы

---

- Только t-тест (для качественных и количественных) – достаточно большой размер выборки (более 30 объектов)
- Уровень ошибки I типа фиксирован на 5%
- Размер эффекта d
  - $d^2 = (M_1 - M_2)^2 / \sigma^2$  – количественный показатель
  - $d^2 = (p_1 - p_2)^2 / \sigma^2$  – качественный показатель,
    - $\sigma^2 = p * (1 - p)$

# Желаемая оценка $d$

---

- Согласно Cohen (1988) если размер эффекта не превышает 0,2, говорят о слабом эффекте терапии, если он оказывается равным 0,5 - говорят об эффекте средней силы и если он превышает 0,8 - то говорят о большом эффекте действия препарата.
-



# Дизайн до-и-после (одна группа)

$$n = \frac{C}{d^2}$$

Сила	$\beta$	C
80%	0,20	7,85
85%	0,15	9,0
90%	0,10	10,5

# Две группы наблюдений

$$n = \frac{2 * C}{d^2}$$

Сила	$\beta$	C
80%	0,20	7,85
85%	0,15	9,0
90%	0,10	10,5

# Пропорции

$$n = \frac{2 \cdot C \cdot p \cdot (1 - p)}{(p_2 - p_1)^2}$$

$$p = \frac{(p_1 + p_2)}{2}$$

Сила	$\beta$	C
80%	0,20	7,85
85%	0,15	9,0
90%	0,10	10,5

# Кроме того

- Есть программы

The POWER Procedure  
Two-sample t Test for Mean Difference

Фиксированные элементы сценария

Distribution	Normal
Method	Exact
Alpha	0.02
Group 1 Mean	120
Group 2 Mean	130
Standard Deviation	20
Nominal Power	0.8
Number of Sides	2
Null Difference	0

Вычислено N Per Group

Actual Power	N Per Group
0.802	82

# SAS

---

```
PROC POWER;
```

```
TWOSAMPLEMEANS
```

```
ALPHA=0.02
```

```
STDDEV=20
```

```
GROUPMEANS= (120 130)
```

```
POWER=0.8
```

```
NPERGROUP=.
```

```
;
```

```
RUN;
```

---

# SAS

Statement	Description
PROC POWER	invokes the procedure
MULTREG	tests of one or more coefficients in multiple linear regression
ONECORR	Fisher's $z$ test and $t$ test of (partial) correlation
ONESAMPLEFREQ	tests of a single binomial proportion
ONESAMPLEMEANS	one-sample $t$ test, confidence interval precision, or equivalence test
ONEWAYANOVA	one-way ANOVA including single-degree-of-freedom contrasts
PAIREFREQ	McNemar's test for paired proportions
PAIREDMEANS	paired $t$ test, confidence interval precision, or equivalence test
TWOSAMPLEFREQ	chi-square, likelihood ratio, and Fisher's exact tests for two independent proportions
TWOSAMPLEMEANS	two-sample $t$ test, confidence interval precision, or equivalence test
TWOSAMPLESURVIVAL	log-rank, Gehan, and Tarone-Ware tests for comparing two survival curves
PLOT	displays plots for previous sample size analysis

# Stata

```
. sampsi 130 120, sd(20) alpha(0.02) power(0.8)
```

```
Estimated sample size for two-sample comparison of means
```

```
Test Ho: m1 = m2, where m1 is the mean in population 1  
and m2 is the mean in population 2
```

```
Assumptions:
```

```
alpha = 0.0200 (two-sided)  
power = 0.8000  
m1 = 130  
m2 = 120  
sd1 = 20  
sd2 = 20  
n2/n1 = 1.00
```

```
Estimated required sample sizes:
```

```
n1 = 81  
n2 = 81
```

# R

```
> pwr.t.test(d=(130-120)/20, power=0.8, sig.level=0.02, type="two.sample", alternative="two.sided")
```

```
Two-sample t test power calculation
```

```
      n = 81.65515
      d = 0.5
sig.level = 0.02
  power = 0.8
alternative = two.sided
```

```
NOTE: n is number in *each* group
```

R использует размер эффекта  $d = (M_1 - M_2) / \sigma$



# R

<a href="#">cohen.E5</a>	Conventional effects size
<a href="#">ES.h</a>	Effect size calculation for proportions
<a href="#">ES.w1</a>	Effect size calculation in the chi-squared test for goodness of fit
<a href="#">ES.w2</a>	Effect size calculation in the chi-squared test for association
<a href="#">pwr</a>	Basic power calculations pwr
<a href="#">pwr.2p.test</a>	Power calculation for two proportions (same sample sizes)
<a href="#">pwr.2p2n.test</a>	Power calculation for two proportions (different sample sizes)
<a href="#">pwr.anova.test</a>	Power calculations for balanced one-way analysis of variance tests
<a href="#">pwr.chisq.test</a>	power calculations for chi-squared tests
<a href="#">pwr.f2.test</a>	Power calculations for the general linear model
<a href="#">pwr.norm.test</a>	Power calculations for the mean of a normal distribution (known variance)
<a href="#">pwr.p.test</a>	Power calculations for proportion tests (one sample)
<a href="#">pwr.r.test</a>	Power calculations for correlation test
<a href="#">pwr.t.test</a>	Power calculations for t-tests of means (one sample, two samples and paired samples)
<a href="#">pwr.t2n.test</a>	Power calculations for two samples (different sizes) t-tests of means

# Другой тип задач

---

- Точность выводов (оценка распространенности)

# Оценка параметров

---

- Средние (количественные переменные)
  - Доли (качественные переменные)
-

# Оценка параметров

---

- Определяем возможные пределы колебаний интересующего нас параметра
  - Оценить распространенность с точностью  $\pm 5\%$
  - Оценить уровень АД с точностью  $\pm 5$  мм рт. ст.

# Эти колебания

---

- Определяют ширину доверительного интервала.
- Соответственно, задача определить нужную ширину доверительного интервала (точнее, полуширину)
- Известно, что полуширина интервала равна

$$w = 1.96 * \frac{\sigma}{\sqrt{n}}$$

$$w = t * \frac{s}{\sqrt{n}}$$

---

# Соответственно

---

$$n = 3.84 * \frac{\sigma^2}{w^2}$$

$$n = t^2 * \frac{s^2}{w^2}$$

---

# Пример

---

- Надо определить средний уровень артериального давления в группе детей в возрасте 5 лет с точностью  $\pm 5$  мм рт. ст.
  - Известно, что популяционное стандартное отклонение 20 мм рт. ст.
  - Необходимо обследовать  
 $3,84 * 20^2 / 5^2 = 62$
  - ребенка
-

# Распространенность

---

- Известно, что в этом случае

- $\sigma^2 = p^*(1-p)$

- Поэтому

$$n = 3.84 \frac{p^*(1-p)}{w^2}$$

- Учитывая, что дисперсия максимальна при  $p=0,5$ , можно оценить максимальное количество обследованных

$$n = \frac{0.96}{w^2} \approx \frac{1}{w^2}$$

---



# Пример

---

- Оценить распространенность с точностью  $\pm 2\%$
- Необходимо обследовать
  - $n = 1 / (0,02)^2 = 2500$
- ПАЦИЕНТОВ

# Пример

---

- Необходимо оценить распространенность хронического гастрита с точностью  $\pm 3\%$ . Ожидаемая распространенность 12%.
  - $n = 3.84 * 0.12 * (1 - 0.12) / (0.03)^2 = 451$
  - Необходимо обследовать 451 человека
-

# Можно сделать таблицу

	w (%)									
p (%)	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
5	7296	1824	811	456	292	203	149	114	90	73
10	13824	3456	1536	864	553	384	282	216	171	138
15	19584	4896	2176	1224	783	544	400	306	242	196
20	24576	6144	2731	1536	983	683	502	384	303	246
25	28800	7200	3200	1800	1152	800	588	450	356	288
30	32256	8064	3584	2016	1290	896	658	504	398	323
35	34944	8736	3883	2184	1398	971	713	546	431	349
40	36864	9216	4096	2304	1475	1024	752	576	455	369
45	38016	9504	4224	2376	1521	1056	776	594	469	380
50	38400	9600	4267	2400	1536	1067	784	600	474	384

# Таким образом

---

- Для описательных исследований размер выборки определяется исходя из желаемой точности оценки популяционного параметра и variability данных.
-

# Итак

---

- Для каждой задачи
    - Указать зависимые переменные, их тип
    - Оценить разброс данных
    - Выбрать метод статистического анализа
    - Установить вероятность ошибки I типа
    - Определить желаемый результат
    - Установить вероятность ошибки II типа
    - Рассчитать размер выборки
-

# Переменные

---

- Соответственно, изучаемые переменные
    - Переменные исхода – отклика, результата, зависимые
    - Переменные воздействия – вмешивающиеся, влияющие, независимые
-

# Показатели исхода

---

- Те показатели, которые являются индикаторами наступления исхода (смерть, развитие ИМ) или интересующими нас показателями, влияние на которые мы изучаем (зависимые переменные)
- Обычно для каждой задачи имеется только один основной показатель исхода (зависимая переменная) и может быть несколько вторичных (дополнительных)

# Показатели исхода

---

- Показатели исхода (зависимые переменные) всегда измеряемые показатели и исследователь должен понимать, каким образом он/она будут измерять эти показатели
  - Переменные исхода
    - Первичные
    - Вторичные
-



# Переменные исхода

---

- Первичные

- Первичная переменная – предоставляет наиболее клинически значимые и прямые доказательства для цели исследования.
  - В исследовании может быть только одна первичная переменная
  - Обычно это переменная эффективности
  - Другие потенциальные первичные переменные
    - Безопасность/переносимость
    - Качество жизни
    - Экономические показатели
  - Отбор переменной производится на основании принятых норм и стандартов в данной области.
  - Необходимо использовать надежные и достоверные переменные, которые использовались в ранних исследованиях или опубликованной литературе

Размер выборки оценивается по первичной переменной

---

# Переменные исхода

---

- В протоколе должно быть дано точное определение первичной переменной, которое будет использоваться в статистическом анализе
- Смертность – не первичная переменная
  - Сравнение
    - Доли умерших?
    - Распределения времен дожития?
- Эффект терапии тоже не первичная переменная
  - Сравнение
    - Возник исход/нет
    - Время по первого возникновения
    - Скорость возникновения события (количество на длительность)

# Переменные исхода

---

- Вторичные переменные
    - Поддерживающие показатели в дополнение к первичной переменной или
    - Показатели эффекта для вторичных задач исследования
    - Количество вторичных переменных должно быть ограничено и они должны быть четко описаны в протоколе
-

# Композитные переменные

---

- Объединение нескольких переменных в одну, с использованием четко прописанного алгоритма
  - Позволяет избавиться от проблемы множественного сравнения не раздувая ошибку I типа
  - Метод должен быть описан в протоколе
  - Надо оценить валидность и надежность переменной
-

# Переменные глобальной оценки

- Переменные для оценки «общей» эффективности или «общей» безопасности
- Обычно имеют субъективный компонент. Надо представить следующую информацию
  - Соответствие основной цели исследования
  - Основания для оценки надежности и валидности
  - Как собранные данные используются для оценки по глобальной шкале
  - Как оцениваются пациенты с пропущенными данными
- Если исследователь опирается на объективные показатели, они также должны быть включены в анализ как дополнительные первичные или важные вторичные переменные
- Переменные глобальной полезности включают оценку врачом пользы и риска назначения терапии.
  - Смешивают два разных показателя
  - Не рекомендуются как первичные переменные

# Множественные первичные переменные

---

- В некоторых случаях необходимо иметь несколько первичных переменных (диапазон эффектов)
  - Необходимо спланировать сравнения
  - Указать, какой минимум или все должны давать доказательства успеха исследования
  - Необходимо объяснить эффект на ошибку I типа и описать методы контроля ошибки I типа
-

# Суррогатные переменные

---

- Непрямые показатели, которые коррелируют с интересующим клиническим исходом
- По возможности не должны использоваться
  - Нет уверенности в том, что они являются предиктором клинического исхода
  - Могут не давать адекватной оценки клинического эффекта, которая может быть сравнена с нежелательными явлениями

# Категоризированные переменные

---

- Дихотомизация интервальных или ординальных переменных, а также иное снижение размерности шкалы
  - Иногда полезно, если есть клиническое обоснование
  - Должно быть четкое предварительное описание в протоколе
  - Надо учитывать потерю мощности
-



# Переменные

---

- Показатели, которые измеряются в исследовании
  - Исходя из характеристик измерительного прибора выделяют переменные, измеряющиеся
    - Номинальной шкалой
    - Ординальной шкалой
    - Интервальной шкалой
    - Шкалой отношений
-

# Переменные

---

- Количество информации, которая содержится в переменных зависит от типа шкалы, при помощи которой происходит измерение
- Соответственно, надо использовать наиболее точный, т.е. информативный, из доступных методов
- Кроме того, следует помнить, что измеренную с большей точностью переменную можно огрубить, а вот наоборот сделать не получится
  - Разные методы статобработки предназначены для переменных, измеренных разными шкалами:
    - Точный тест Фишера: номинальные
    - Тест Викоковсона: ординальные
    - Т-тест: интервальные



Рост мощности

---

# Переменные

---

- Поскольку шкала напрямую определяет мощность статистического теста, то чем более точный измерительный метод используется, тем меньше надо единиц наблюдения
  - Чем более надежен измерительный инструмент, тем выше его точность и тем меньше надо единиц наблюдения
  - Поэтому с теоретической точки зрения количественные (интервальные) переменные предпочтительнее все других
-

# Однако...

---

- Целый ряд исходов, интересующих пациента бинарен
  - Это номинальная переменная
  - Поскольку ДМ считает только важные для пациента исходы, то заменять их на количественные показатели только с целью повышения эффективности исследования нельзя
  - Но можно адекватно операционализировать исследовательский вопрос
    - «Доктор, умру ли я?»
    - «Да, абсолютно точно»
    - «Вопрос только в том, когда»
  - Замена бинарного показателя (жизнь/смерть) на интервальный (времена дожития)
-

# Для удобства

---

- Переменные, измеренные при помощи разных шкал, имеют дополнительные названия
  - Качественные (номинальные)
  - Полуколичественные (ординальные)
  - Количественные (интервальные)
- Кроме того, для компьютерного ввода необходимо учесть, как будем кодировать переменные
  - Число (любые)
    - Количество разрядов числа, наличие десятичной запятой и число знаков после запятой (для ИРК)
  - Текст (всегда номинальные)

Лучше всего все кодировать числами

# Показатели исхода

---

- Показатели исхода (зависимые переменные) всегда измеряемые показатели и исследователь должен понимать, каким образом он/она будут измерять эти показатели
- На что следует обратить внимание
  - Тип переменной (качественная/количественная)
  - Пределы колебаний
  - Точность измерения
    - Ошибка измерения должна быть значительно меньше диапазона изменений переменной

# Пять этапов формирования групп исследования

---

- Определить популяцию, сформулировав:
    - Критерии включения
    - Критерии исключения
  - Определить размер групп
    - Определить размер эффекта
    - Квантифицировать разброс данных
    - Установить вероятности ошибок I и II типа
    - Вычислить размер группы
  - Определить способ разделения
    - Случайный
  - Сформировать группы
  - Сравнить группы
-