

Измерения в психологии

Измерения в структуре
психологического исследования

Обыденное и научное познание



ГС и выборка

- Репрезентативность
- Случайный отбор и доступность
- Детальное описание выборки и процедуры ее формирования
- Сравнение выборок: независимые и зависимые выборки

Гипотезы: содержательные и статистические

1. Содержательная гипотеза:

утверждение о связи 2-х явлений в ГС.

2. В результате выборочного исследования:

связь в терминах описательных статистик (M_1 и M_2 ; r_{xy} и т.п.).

3. Статистическая гипотеза: утверждение о связи параметров ГС.

\bar{M}_1 \bar{M}_2 – \bar{r}_{xy} параметры.

Основная (нулевая) статистическая гипотеза (H_0) –

утверждение об отсутствии связи в терминах параметров ГС.

Например: $H_0 : \bar{M}_1 = \bar{M}_2$ $H_0 : \bar{r}_{xy} = 0$

4. Определение p -уровня значимости (знч.) и принятие статистического решения:

$p \leq \alpha(0,05)$ H_0 отклоняется (результат статистически достоверен)

$p > \alpha(0,05)$ H_0 не отклоняется (результат статистически не достоверен)

См. «Математические методы...», гл. 7 (стр. 93 – 110)

Теоретическое распределение и p -уровень значимости

Пример: проверка состоятельности тестовой нормы $IQ = 100$.

Выборочные статистики: $N = 100; M = 103; \sigma = 15$

Статистическая гипотеза: $H_0 : \bar{M}_1 = 100$.

Теоретическое распределение соответствует распределению статистики для выборок, многократно извлекаемых из ГС, для которой верна H_0 .

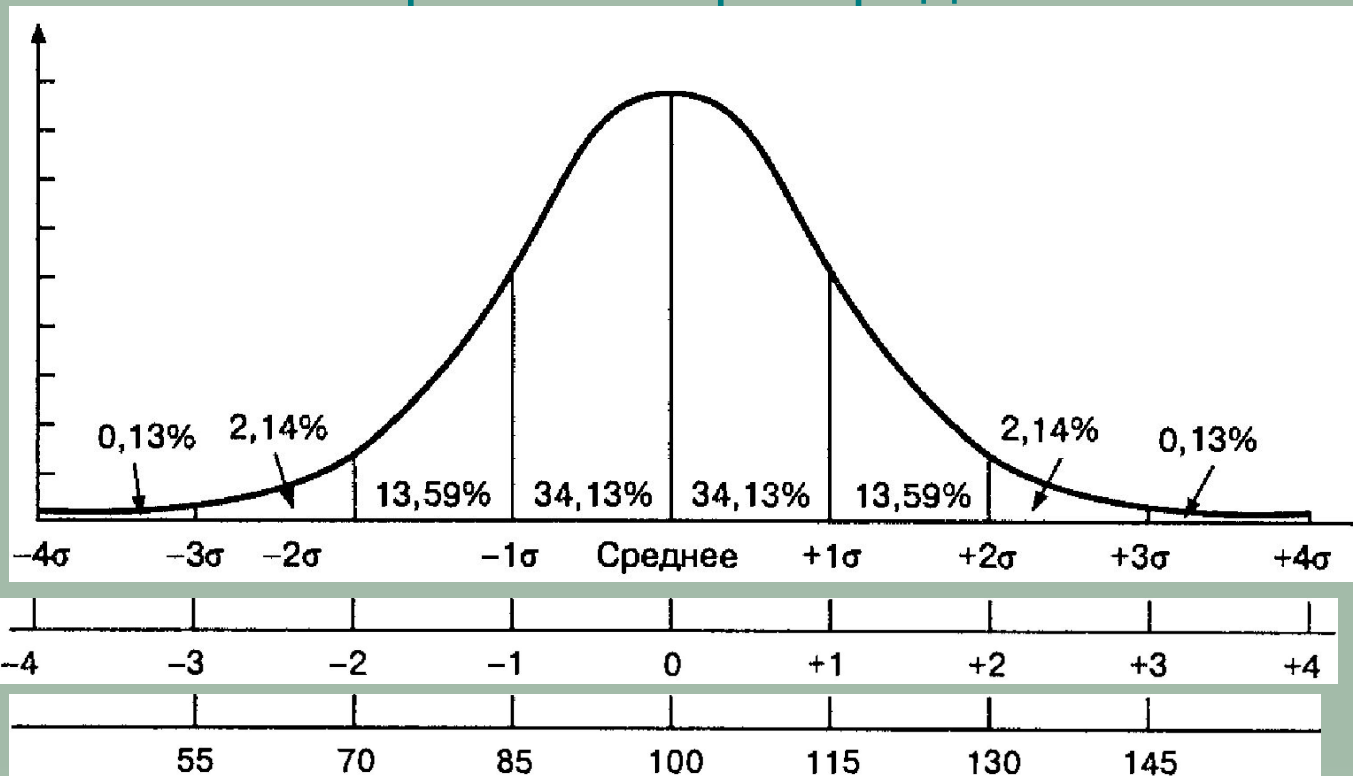
«Центральная предельная теорема»:

распределение выборочных средних из ГС при достаточно большом N является нормальным.

Для распределения выборочных средних:

$$M_{M_x} = \bar{M}_x; \sigma_{M_x} = \sigma_x / \sqrt{N}.$$

Нормальное распределение



$M \pm \sigma$ соответствует $\approx 68\%$ (точно — 68,26%) площади;
 $M \pm 2\sigma$ соответствует $\approx 95\%$ (точно — 95,44%) площади;
 $M \pm 3\sigma$ соответствует $\approx 100\%$ (точно — 99,72%) площади.

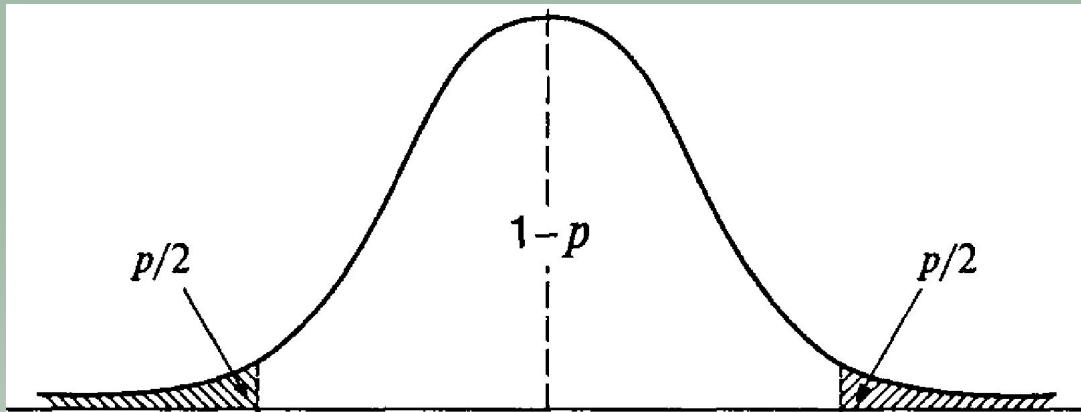
90% всех случаев располагается в диапазоне значений $M \pm 1,64\sigma$;
95% всех случаев располагается в диапазоне значений $M \pm 1,96\sigma$;
99% всех случаев располагается в диапазоне значений $M \pm 2,58\sigma$.

Что такое p -уровень

$$H_0 : \bar{M}_1 = 100. \quad N = 100; \quad M = 103; \quad \sigma = 15$$

$$M_{M_x} = \bar{M}_x; \quad \sigma_{M_x} = \sigma_x / \sqrt{N}.$$

$$z = \frac{x_i - M_x}{\sigma_x} = \frac{M - 100}{\sigma_x / \sqrt{N}} = 2.$$



p -уровень значимости – вероятность случайного получения такого (или большего) отклонения от того, что утверждает H_0 .

В данном случае $p = 1 - 0,9544 = 0,046$.

Синонимы: Статистическая значимость – Знч., Significant level – Sig.

Статистический критерий

$$z_3 = \frac{M_x - A}{\frac{\sigma_x}{\sqrt{N}}}$$

- z-критерий.

$$t_3 = \frac{|M - A|}{\sigma/\sqrt{N}}, df = N - 1$$

- t-критерий Стьюдента.

Назначение критерия — проверка статистической гипотезы путем определения p -уровня значимости (вероятности того, что H_0 верна).

Выбор критерия определяется проверяемой статистической гипотезой.

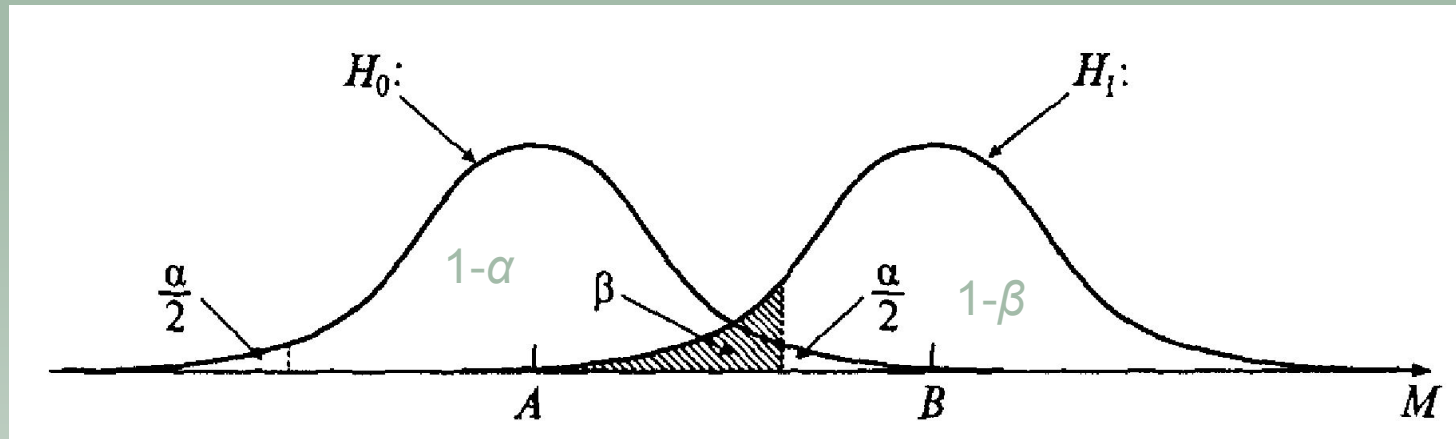
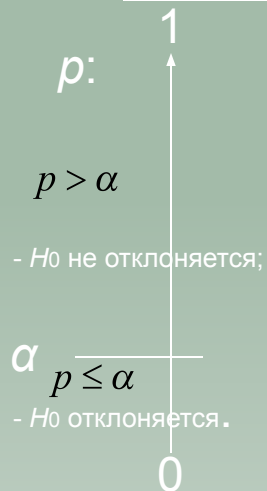
Критерий включает в себя:

- формулу расчета эмпирического значения критерия по выборочным статистикам;
- правило (формулу) определения числа степеней свободы;
- теоретическое распределение для данного числа степеней свободы;
- правило соотнесения эмпирического значения критерия с теоретическим распределением для определения вероятности того, что H_0 верна.

Статистическое решение

В действительности:

Решение исследователя:	H_0 истинна	H_1 истинна
Отклонить H_0 (принять H_1)	Неправильное решение, ошибка I рода, вероятность = α	Правильное решение, вероятность = $1 - \beta$ (мощность или чувствительность критерия)
Принять H_0	Правильное решение, вероятность = $1 - \alpha$ (доверительная вероятность)	Неправильное решение, ошибка II рода, вероятность = β



Традиционная интерпретация уровней значимости при $\alpha = 0,05$

Уровень значимости	Решение	Возможный статистический вывод
$p > 0,1$	Принимается H_0	«Статистически достоверные различия не обнаружены»
$p \leq 0,1$	сомнения в истинности H_0 , неопределенность	«Различия обнаружены на уровне статистической тенденции»
$p \leq 0,05$	значимость, отклонение H_0	«Обнаружены статистически достоверные (значимые) различия»
$p \leq 0,01$	высокая значимость, отклонение H_0	«Различия обнаружены на высоком уровне статистической значимости»



$p > 0,1$



$p \leq 0,1$



$p \leq 0,05$



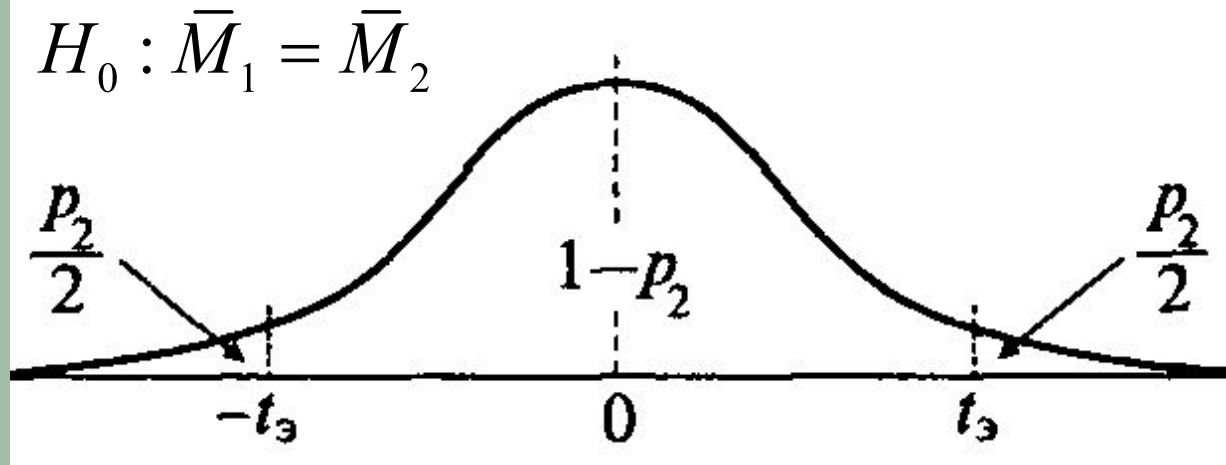
$p \leq 0,01$



$p \leq 0,001$

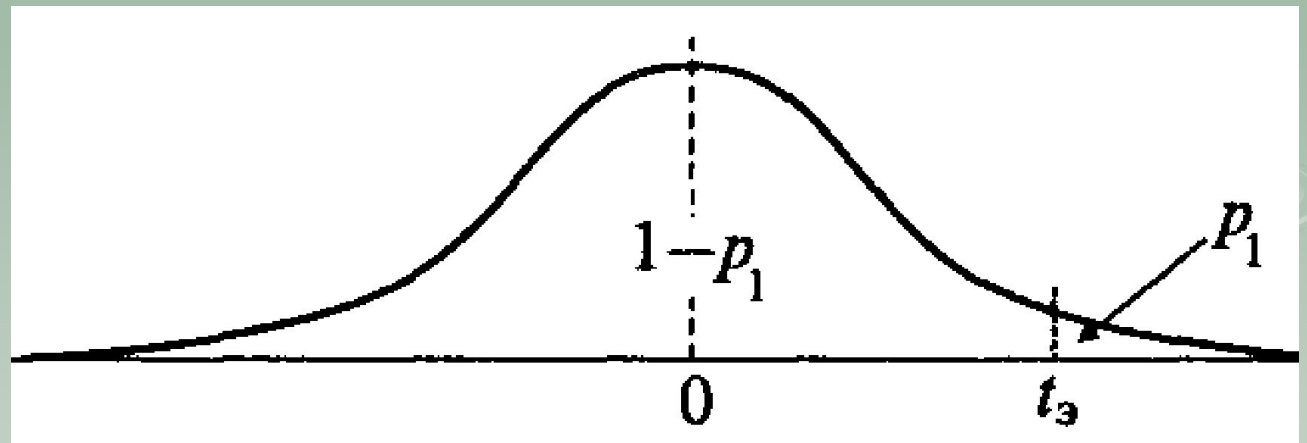
Направленные и ненаправленные альтернативы

Ненаправленная альтернатива:



Направленная альтернатива:

$$p_1 = \frac{p_2}{2}$$

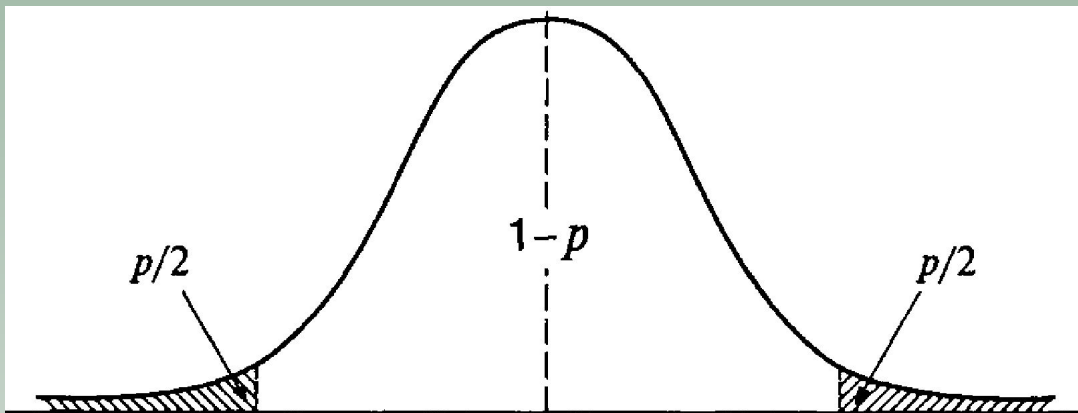


Что такое p -уровень?

Вероятность случайного получения такого (или более неожиданного) отклонения от H_0 на выборке данной численности из ГС, для которой эта H_0 верна

Альтернативные определения:

- Вероятность получения такого эмпирического результата случайно на выборке из ГС, для которой H_0 верна
- Вероятность того, что H_0 верна (связи/различий нет) в ГС, которой репрезентативна данная выборка (?)



Типичные ошибки интерпретации p -уровня значимости

1. Содержательная интерпретация

$$p \leq \alpha(0,05)$$

Отклонение H_0



Подтверждение содержательной гипотезы (а не ее доказательство)

$$p > \alpha(0,05)$$

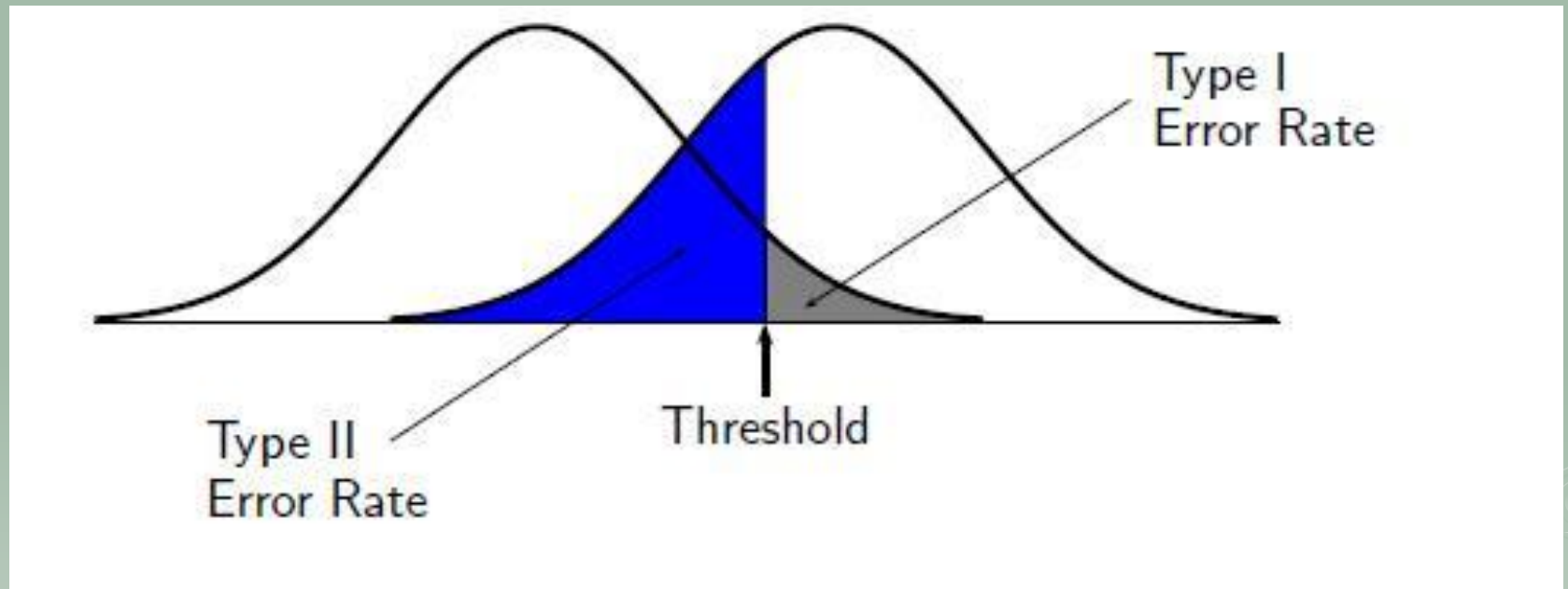
Принятие H_0



НЕ подтверждение содержательной гипотезы (а не ее опровержение)

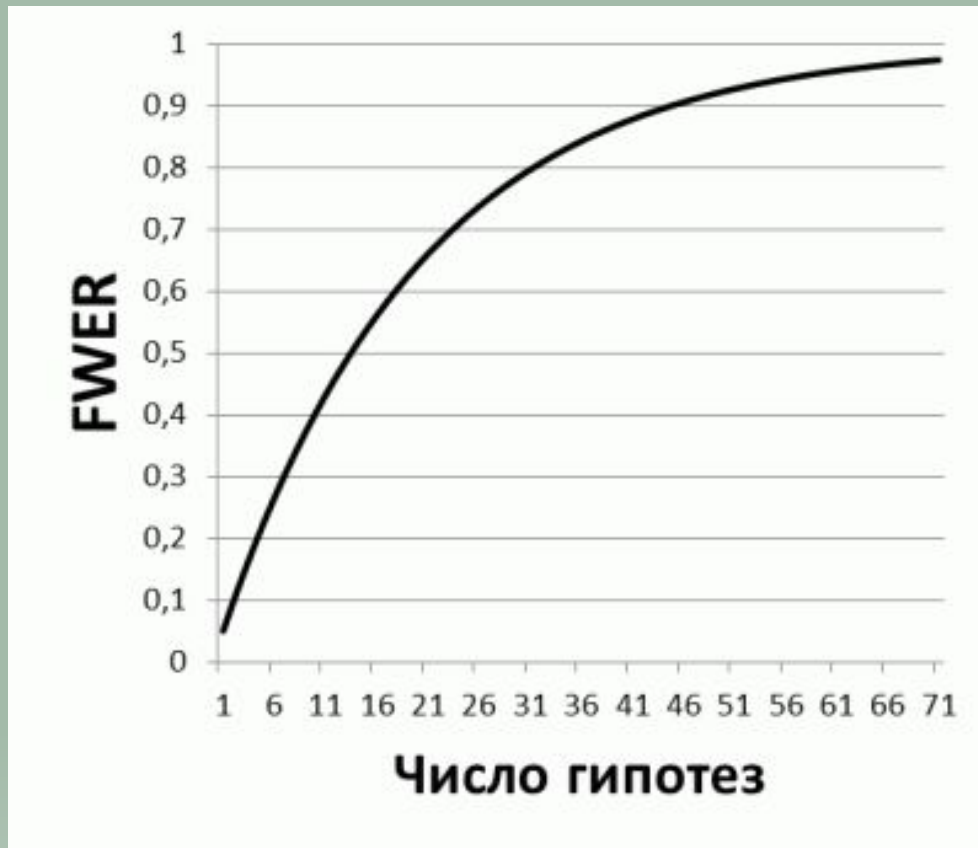
2. Игнорирование проблемы множественной проверки значимости

Однократная проверка значимости



Вероятность ошибочного отклонения H_0 (когда она верна)

Number of genes tested (N)	False positives incidence	Probability of calling 1 or more false positives by chance ($100(1-0.95^N)$)
1	1/20	5%
2	1/10	10%
20	1	64%
100	5	99.4%



$FWER = 1 - (1 - \alpha)^n$ где n – кол-во проверяемых гипотез.

FWER – вероятность ошибочного отклонения H_0 , т.е. вероятность ошибочного признания результата статистически достоверным.

Иначе говоря – это вероятность того, что результаты исследования не соответствуют действительности.

Учет множественности статистических проверок

- I. Применение многомерных методов
- II. Поправка Бонферрони для семейства n гипотез: для n гипотез каждый p -уровень умножаем на n , перед сравнением с α
- III. Поправка Benjamini & Hochberg (1995; 2000) для семейства n гипотез:
 - 1) Упорядочиваем все p от min до max (i – текущий номер p в ряду);
 - 2) Для каждого i вычисляем: $p^*n/i = p_{\text{корр.}}$;
 - 3) Если $p_{\text{корр.}} \leq \alpha$ – результат статистически достоверен!

Пример: корреляционная матрица

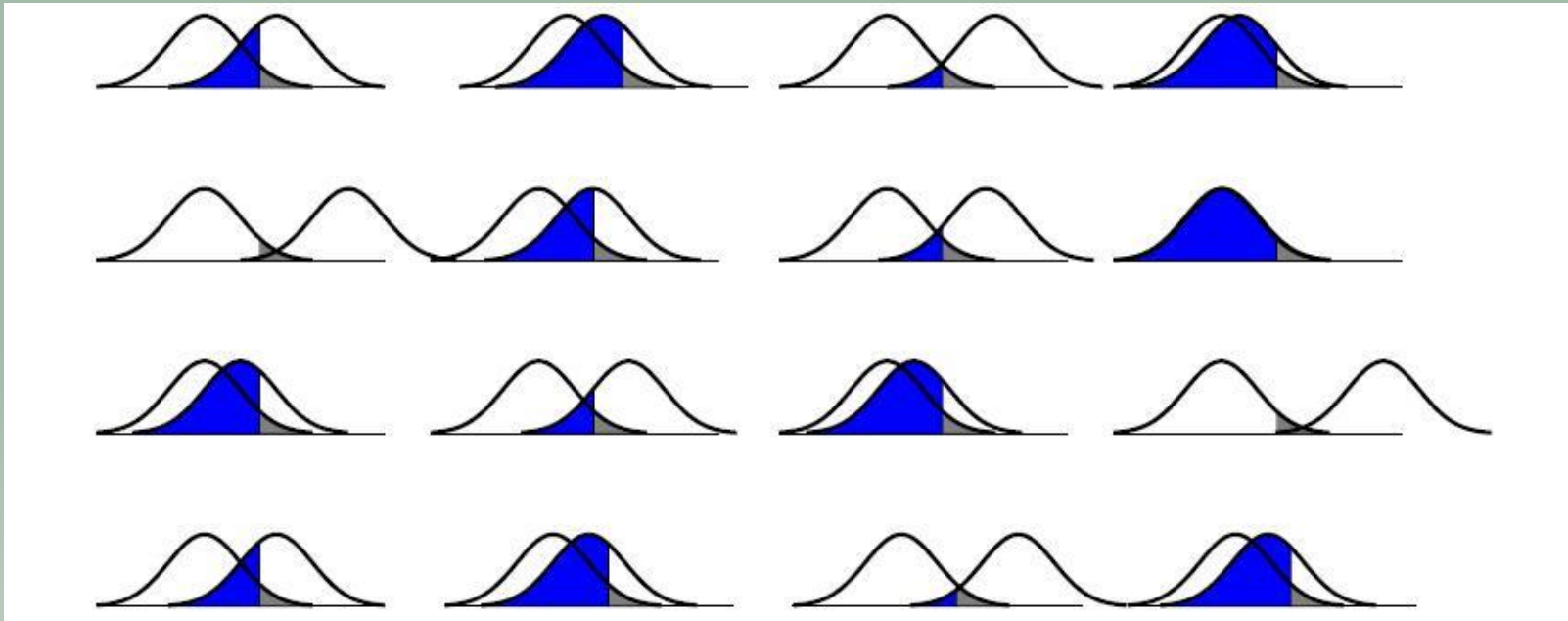
		Корреляции				
		тест1	тест2	тест3	тест4	тест5
тест1	Корреляция Пирсона	1	,436**	-,051	-,134	-,201*
	Знч.(2-сторон)		,000	,617	,185	,047
	N	100	100	100	100	100
тест2	Корреляция Пирсона	,436**	1	-,196	-,153	,015
	Знч.(2-сторон)	,000		,050	,127	,886
	N	100	100	100	100	100
тест3	Корреляция Пирсона	-,051	-,196	1	,441**	,483**
	Знч.(2-сторон)	,617	,050		,000	,000
	N	100	100	100	100	100
тест4	Корреляция Пирсона	-,134	-,153	,441**	1	,475**
	Знч.(2-сторон)	,185	,127	,000		,000
	N	100	100	100	100	100
тест5	Корреляция Пирсона	-,201*	,015	,483**	,475**	1
	Знч.(2-сторон)	,047	,886	,000	,000	
	N	100	100	100	100	100

** Корреляция значима на уровне 0.01 (2-сторон).

* Корреляция значима на уровне 0.05 (2-сторон).

Последствия коррекции многократной проверки значимости

При одном и том же пороге принятия/отклонения H_0 (α – серая штриховка) вероятность того, что верна альтернативная гипотеза (β – синяя штриховка), может быть любой.



Поправка на многократность проверки непредсказуемо увеличивает вероятность β .

Рекомендации

- Минимизировать кол-во измерений за счет увеличения их надежности и валидности
- Применять многомерные методы для большого кол-ва измерений
- Обязательно применять коррекцию р-уровня значимости при многократной проверке

Модель исследования



Измерения и шкалы

- Объекты (случаи)
- Свойства и их признаки
- Переменные – результаты измерений
- Измерения в психологии - косвенные
- Шкала измерения (С.Стивенс):
соотношение между свойствами чисел и
измеряемым свойством; задается
измерительной операцией.


Измерительные шкалы

- Номинативная (номинальная, наименований, неколичественная). Операция – классификация. Пол, хобби, должность...
- Порядковая (ранговая) – количественная, неметрическая. Операция – упорядочивание («больше – меньше»). Место в турнире...
- Интервальная (метрическая). Равным разностям между числами соответствуют равные разности в измеряемом свойстве (единицы измерения, «равноинтервальность»). Температура по С, летоисчисление от р.х. ...
- Абсолютная (метрическая). «Равноинтервальная» + ноль (отсутствие измеряемого свойства). Вес, длина, время...

Исходные данные в SPSS

ex01.sav [Наборданных1] - Редактор данных IBM SPSS Statistics


Файл Правка Вид Данные Преобразовать Анализ Прямой маркетинг Графика Сервис Окно Справка



	№	пол	класс	вуз	хобби	тест1	тест2	тест3	тест4	тест5	отметка1	отметка2
1	1	МУЖ	Б	ЕСТ...	искус...	6	7	13	10	14	3,90	4,20
2	2	МУЖ	А	ЕСТ...	спорт	8	9	10	11	11	3,55	3,95
3	3	МУЖ	В	ТЕХН	компь...	10	6	10	8	9	3,75	4,65
4	4	ЖЕН	В	ГУМ	компь...	13	9	10	12	6	3,85	3,95

ex01.sav [Наборданных1] - Редактор данных IBM SPSS Statistics

Файл Правка Вид Данные Преобразовать Анализ Прямой маркетинг Графика Сервис



	Имя	Тип	Ширина	Десяти...	Метка	Значения
1	№	Числовая	3	0		Нет
2	пол	Числовая	3	0		{1, ЖЕН}...
3	класс	Числовая	3	0		{1, А}...
4	вуз	Числовая	3	0		{1, ГУМ}...
5	хобби	Числовая	3	0		{1, спорт}...
6	тест1	Числовая	3	0	счет в уме	Нет
7	тест2	Числовая	3	0	числовые ряды	Нет
8	тест3	Числовая	3	0	словарь	Нет
9	тест4	Числовая	3	0	осведомленн...	Нет
10	тест5	Числовая	3	0	кратковремен	Нет

Список переменных

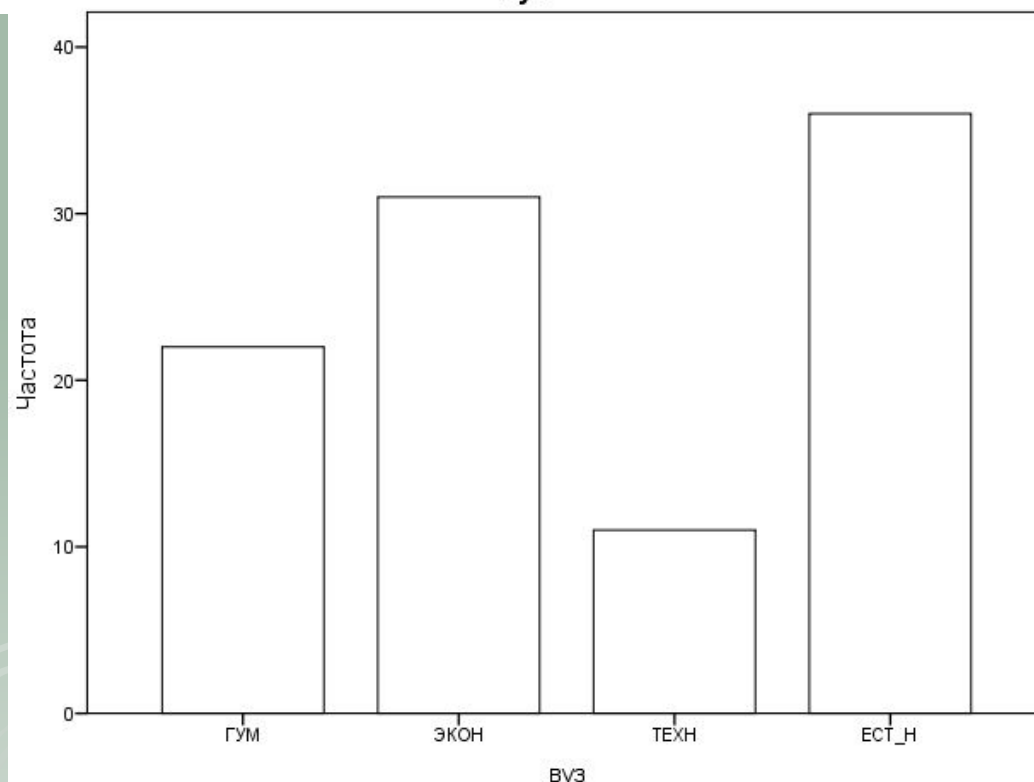
№	Имя	Название	Тип	Диапазон
1	№		ном	1–100
2	пол		ном	1 – жен, 2 – муж
3	класс		ном	1– А, 2 – Б, 3 - В
4	вуз	Предполагаемый для поступления ВУЗ	ном	1 — гуманитарный; 2 — экономический, 3 — технический; 4 — естественнонаучный
5	хобби	Внешкольные увлечения	ном	1 — спорт; 2 — компьютер; 3 — искусство
6	тест1	Счет в уме	кол	1 – 20
7	тест2	Числовые ряды	кол	1 – 20
8	тест3	Словарный запас	кол	1 – 20
9	тест4	Осведомленность	кол	1 – 20
10	тест5	Кратковременная вербальная память	кол	1 – 20
11	отметка1	Средний балл отметок за 10-й класс	кол	3–5
12	отметка2	Средний балл отметок за 11-й класс	кол	3–5
13	тест_ср	Среднее по тестам 1 – 5	кол	1 – 20

Распределения частот (номинальные переменные)

Таблица
распределения:

		Частота	Процент	Валидный процент	Кумулятивный процент
Валидные	ГУМ	22	22,0	22,0	22,0
	ЭКОН	31	31,0	31,0	53,0
	ТЕХН	11	11,0	11,0	64,0
	ЕСТ_Н	36	36,0	36,0	100,0
	Итого	100	100,0	100,0	

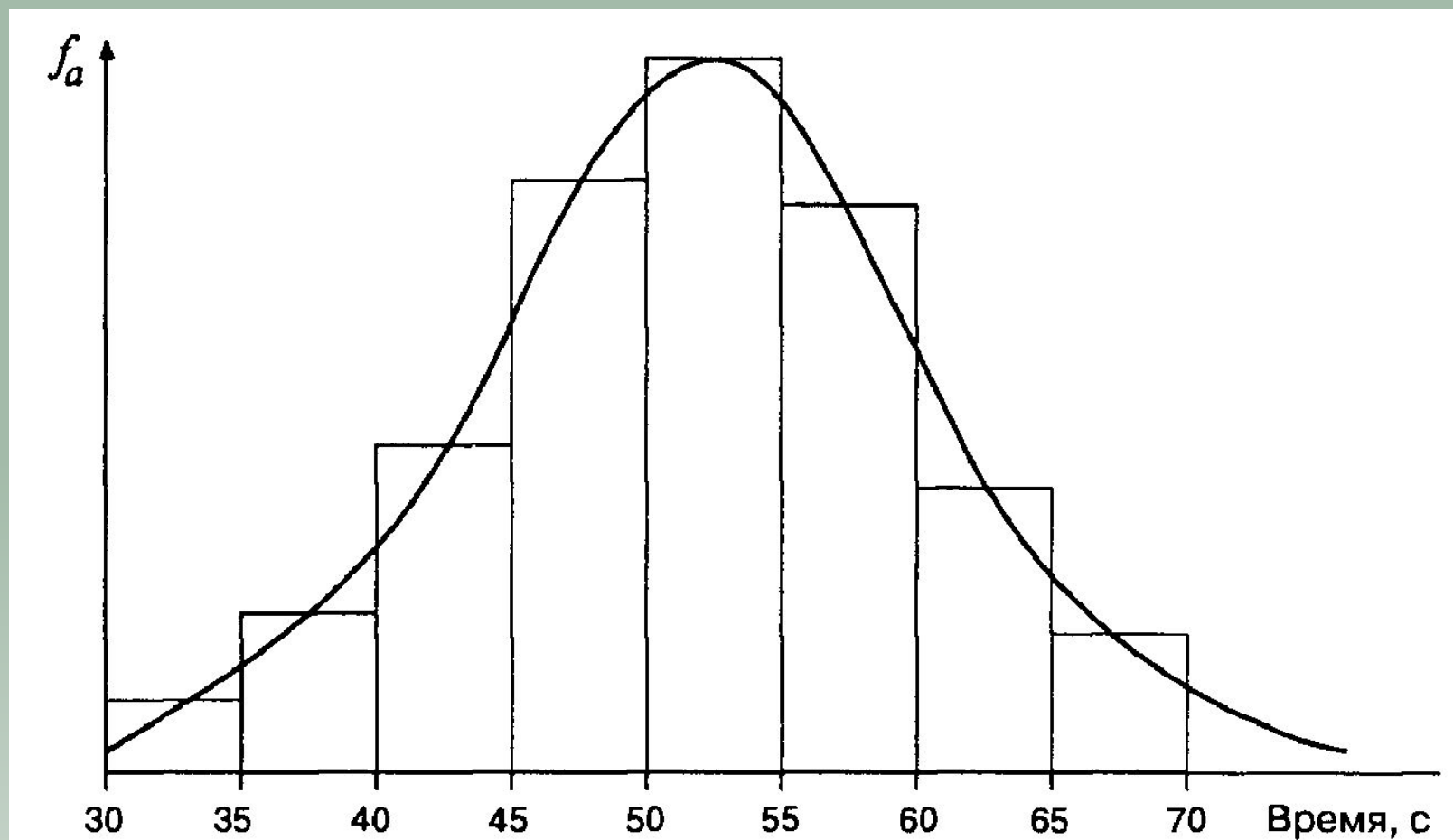
График распределения –
столбиковая диаграмма:



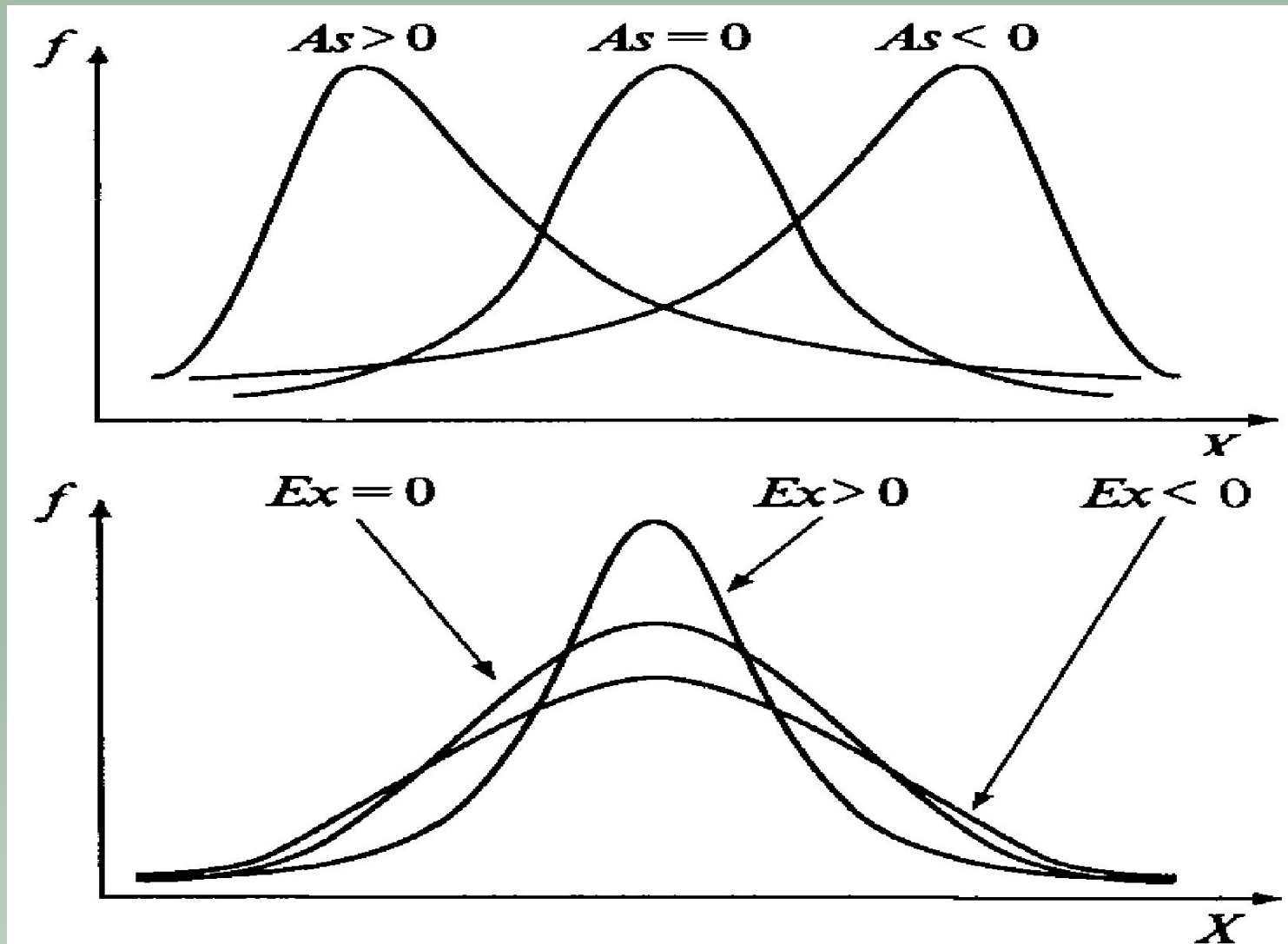
Распределения частот (количественные переменные)

Переменная x – время решения тестовой задачи

$$N = 40; x_{\min} = 32; x_{\max} = 67.$$

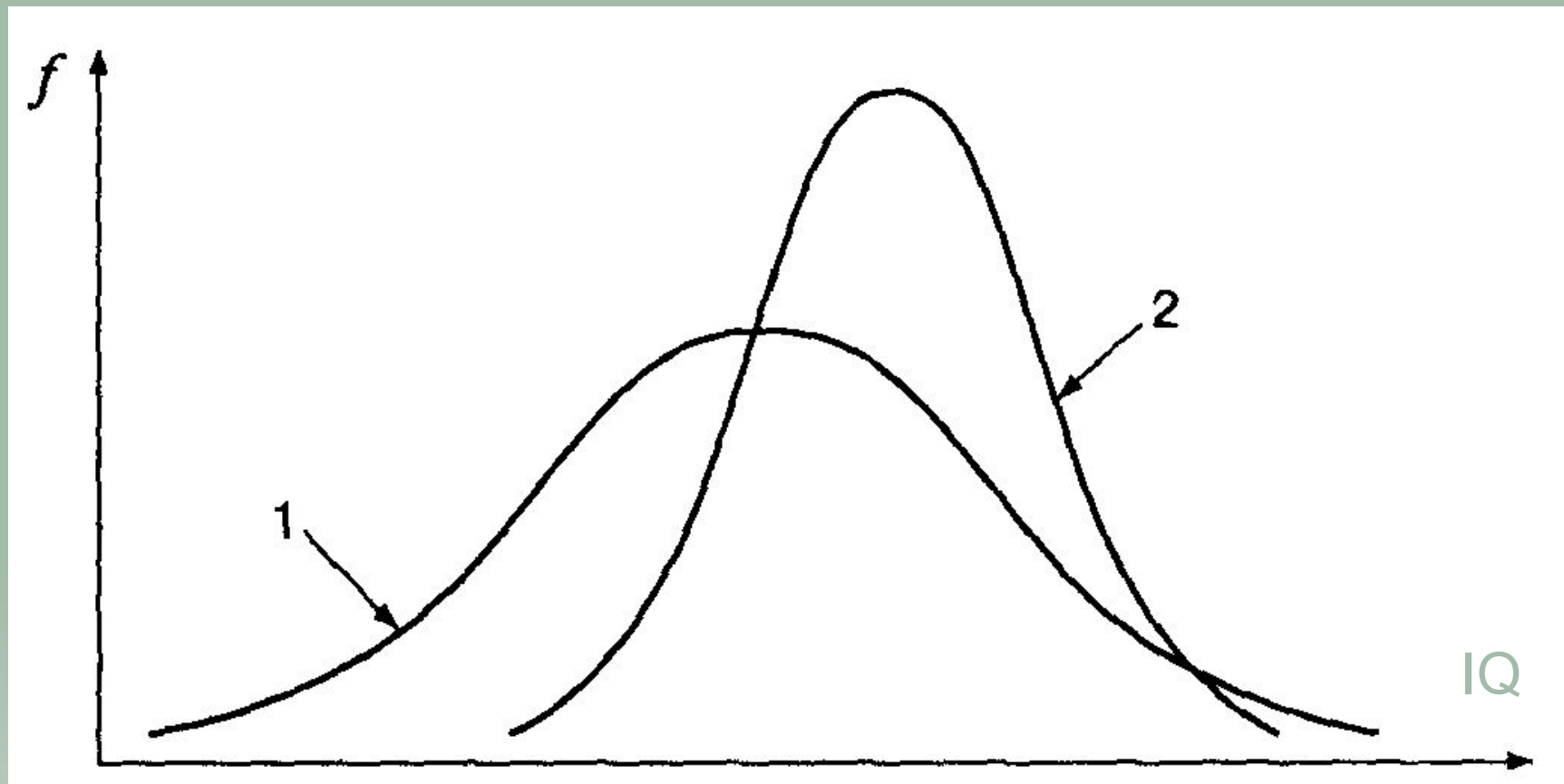


Виды графиков распределения



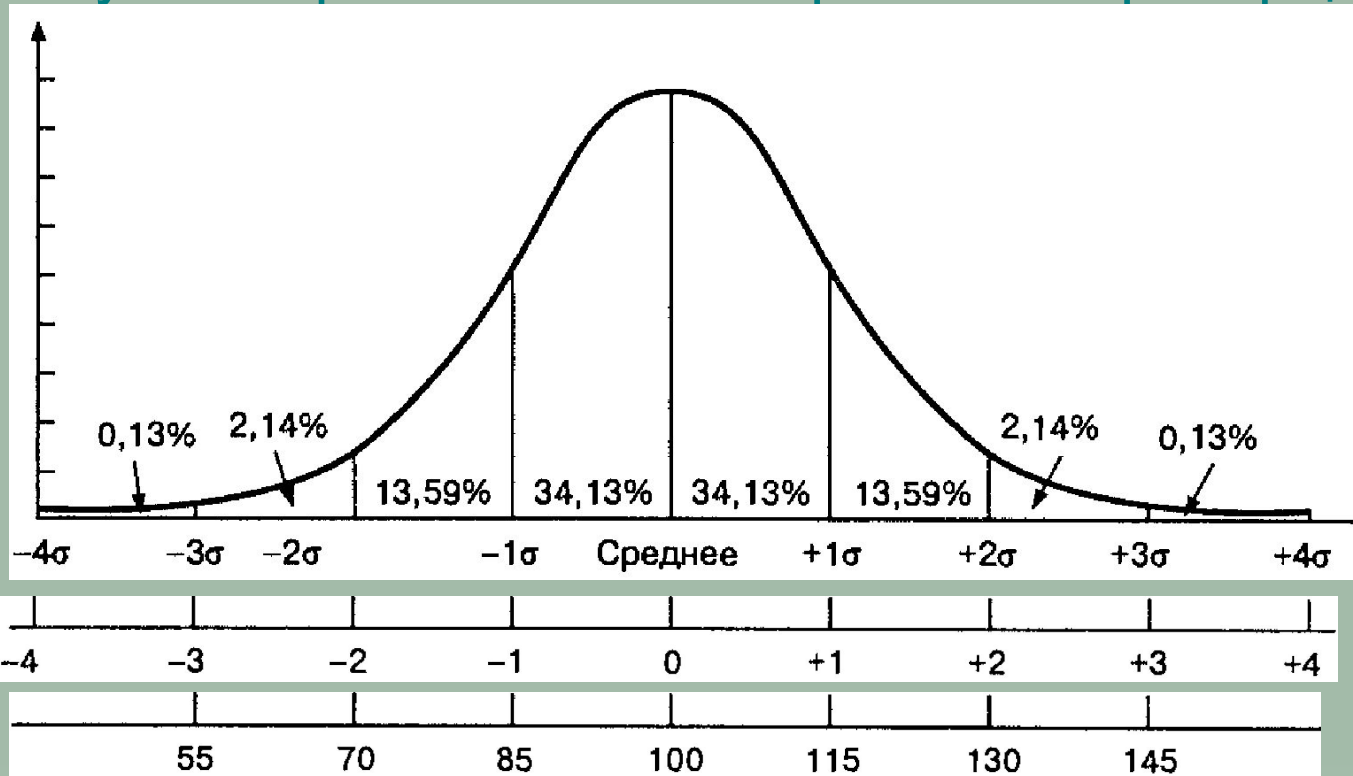
Если X – кол-во правильно решенных задач, как интерпретировать?

Интерпретация графиков распределения



1, 2 – юноши... девушки... ?

«Постулат нормальности»: нормальное распределение



$M \pm \sigma$ соответствует $\approx 68\%$ (точно — 68,26%) площади;
 $M \pm 2\sigma$ соответствует $\approx 95\%$ (точно — 95,44%) площади;
 $M \pm 3\sigma$ соответствует $\approx 100\%$ (точно — 99,72%) площади.

90% всех случаев располагается в диапазоне значений $M \pm 1,64\sigma$;
95% всех случаев располагается в диапазоне значений $M \pm 1,96\sigma$;
99% всех случаев располагается в диапазоне значений $M \pm 2,58\sigma$.

Проверка нормальности

Зачем?

1. Визуально, по графику распределения и с контролем выбросов.
2. По критериям асимметрии и эксцесса.
3. По статистическим критериям нормальности: а) К-С и К-С с поправкой Лилиефорса; б) Шапиро-Уилка

Пример описательных статистик (SPSS)

Описательные статистики

	N	Среднее	Стд. отклонение	Асимметрия		Эксцесс	
	Статистика	Статистика	Статистика	Статистика	Стд. ошибка	Статистика	Стд. ошибка
счет в уме	100	10,25	2,653	-,093	,241	-,363	,478
числовые ряды	100	10,35	2,768	-,127	,241	-,435	,478
словарь	100	11,96	2,857	,261	,241	1,285	,478
осведомленность	100	11,51	3,040	,198	,241	,254	,478
кратковременная память	100	11,84	2,915	,032	,241	-,106	,478
N валидных (целиком)	100						