

# Статистические критерии

# Статистический критерий-это...

- ...решающее правило, обеспечивающее принятие истинной и отклонение ложной гипотезы с заданной вероятностью (Г.В. Суходольский).

Это правило требуется, чтобы математически обосновать наши выводы

# Виды критериев

## Параметрические

т.е. основанные на расчете параметров генеральной совокупности ( $X$ ,  $\sigma^2$ ).

Достоинства: более мощные и точные.

Трудности:

требуют измерений по шкале интервалов или равных отношений;

только нормальное распределение!;

желательный объем выборки  $N > 50$

# Виды критериев

## Непараметрические

т.е. не включающие в формулу расчета параметров распределения, основанные на оперировании частотами или рангами.

### Достоинства:

- + просты в расчете;
- + применимы на малых выборках ( $N < 10$ );
- + не привязаны к характеру распределения.

Недостатки: менее мощные ( $\beta$ ), имеют табличные ограничения по макс.  $N$

# Выявление различий в уровне исследуемого признака

## U-критерий Манна-Уитни

- **Назначение критерия:** оценка достоверности различий между 2 выборками по уровню признака;
- **Суть критерия:** оценивает зону совпадений значений выборок после сплошного ранжирования.
- **Ограничения критерия:**
  - а)  $N_1 > 2$ ,  $N_2 > 5$  (или каждая  $> 3$ );
  - б)  $N_1$ ,  $N_2$  не более 60

# Выявление различий в уровне исследуемого признака

## U-критерий Манна-Уитни

### Алгоритм подсчета (Е.В. Сидоренко):

- Перенести все данные на отдельные карточки двух цветов (Например,  $n_1$  - синие,  $n_2$  - красные );
- Разложить все карточки по возрастанию значений;
- Приписать каждому значению ранг, начиная с меньшего (Правила ранжирования!)
- Проверить: для всего ряда рангов  $\sum \text{рангов} = \frac{N \times (N+1)}{2}$
- Для каждой выборки **отдельно** посчитать сумму рангов
- Наибольшую сумму рангов обозначить как  $T_x$

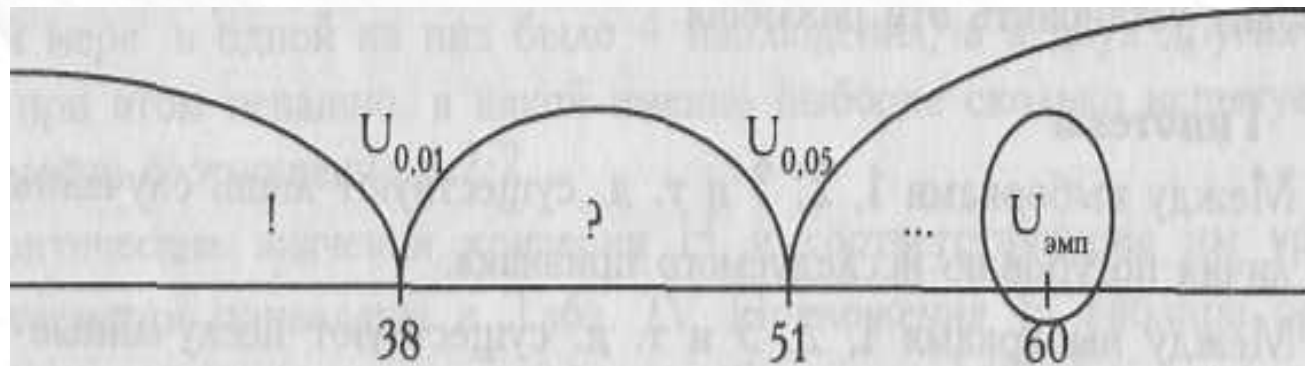
# Выявление различий в уровне исследуемого признака

## U-критерий Манна-Уитни

Алгоритм подсчета (продолжение):

- Считать  $U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x$ ,  
где  $n_x$  — выборка с наибольшей суммой рангов.
- Сопоставить с табличными критическими значениями  $U_{кр}$ .
- Если  $U < U_{кр}$  для  $p=0,01$ , тогда различие значимо

Пример:  
Различий нет



# Выявление различий в уровне исследуемого признака

## H-критерий Крускала-Уоллеса

- **Назначение критерия:** оценка достоверности различий между 3 и более выборками по уровню признака;
- **Суть критерия:** оценивает различия в суммах рангов, полученных каждой выборкой после сплошного ранжирования всех испытуемых.
- **Ограничения критерия:**
  - а)  $N_1 > 2$ ,  $N_2$  и  $N_3 > 4$  (или каждая  $> 3$ );
  - б) упускает различия между отдельными парами выборок



# Выявление различий в уровне исследуемого признака

## H-критерий Крускала-Уоллеса

### Алгоритм подсчета:

- Перенести данные каждой выборки на отдельные карточки определенного цвета;
- Разложить все карточки по возрастанию значений;
- Приписать каждому значению ранг, начиная с меньшего (проверить по Правилам ранжирования!)
- Посчитать сумму рангов каждой выборки, обозначить ее как T1, T2, T3

$$H = \left[ \frac{12}{N \times (N+1)} \sum_{j=\text{выборка}} \left( \frac{T_j^2}{n_j} \right) \right] - 3 \times (N+1)$$

# Выявление различий в уровне исследуемого признака

## **H-критерий Крускала-Уоллеса**

### **Алгоритм подсчета (продолжение):**

- Если хотя бы одна выборка имеет объем  $n > 5$ , **критические значения** по таблицам критерия хи-квадрат ( $\chi^2$ ) для  $df = N - 1$ ;
- Нарисовать ось значимости, отметить  $p = 0.05$  и  $p = 0.01$
- Если рассчитанное значение  $H \geq H_{кр.}$  для  $p = 0.05$ , различие значимо и  $H_0$  отвергается

# Выявление различий в уровне исследуемого признака

## Q-критерий Розенбаума

- непараметрическая оценка различий между двумя выборками по уровню какого-либо признака, количественно измеренного (для выборок с  $N > 11$ );

## S - критерий тенденций Джонкира

- выявляет тенденции изменения признака при переходе от выборки к выборке при сопоставлении 3 и более выборок (объем выборок одинаков, не более 6 выборок,  $N < 10$ )

# Оценка достоверности сдвига

## T-критерий Вилкоксона

- **Назначение критерия:** оценка достоверности изменений показателя выборки в разных условиях, направления и силы сдвига;
- **Суть критерия:** основан на ранжировании абсолютных разностей пар значений зависимых выборок.
- **Ограничения критерия:**
  - а) объем выборки  $5 < N < 50$ ;
  - б) нулевые сдвиги из выборки придется исключить;
  - в) мощнее при значительных сдвигах

# Оценка достоверности сдвига

## T-критерий Вилкоксона

Алгоритм подсчета:

$$T = \sum R_r$$

- Сортировать испытуемых по алфавиту;
- Вычислить разность между показателями «до» и «после»;
- Отдельной колонкой записать модули разностей
- Ранжировать модули разностей по возрастанию (соблюдать Правила ранжирования!)
- Отдельными колонками выписать ранги для + и — сдвигов (пометить те, которые считать нетипичными)
- Считать значение T по формуле, где R<sub>r</sub> - ранговые значения нетипичных сдвигов
- По таблице критических значений определить границы значимости. Сделать статистический вывод

# Оценка достоверности сдвига

## G- критерий знаков

- Установление общего направления сдвига (номинативные и ранговые переменные, незначительные сдвиги;  $5 < (N_1 + N_2) < 300$ );

## Критерий $\chi^2_r$ Фридмана

- Сопоставление показателей, измеренных в 3 или более условиях на одной и той же выборке (не определяет направление изменений;  $N > 2$ ; замеров  $> 3$ )

## L-критерий тенденций Пейджа

Направление изменений 1 выборки от 3 до 6 условий ( $N < 12$ )

# Параметрические критерии

## F-критерий Фишера

**Цель:** сравнение дисперсий 2 независимых выборок

**Ограничения:** измерения по параметрическим шкалам, нормальное распределение признака в генеральной совокупности.

**Гипотезы:**  $H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$       $H_{\text{альт}}: \sigma_1^2 \neq \sigma_2^2$

$$F = \frac{S^2_{\text{большая}}}{S^2_{\text{меньшая}}}$$

Сравнить с  $F_{\text{кр.}}$  для  $df1 = N_{\text{больш}} - 1$  и  $df2 = N_{\text{меньш}} - 1$

Если  $F \leq F_{\text{кр.}}(df1, df2)$  для  $p < 0,01$ , то нулевая гипотеза верна

# Параметрические критерии

**t-критерий Стьюдента** — 1908г., заводы Гиннеса, В.Госсет, оценка процента брака

**Цель:** сравнение средних значений 2 выборок (есть модификации для зависимых, независимых, эмпирической и теоретической выборок).

**Ограничения:** нормальное распределение в выборках; предварительное сравнение дисперсий с помощью F-критерия Фишера.

**Гипотезы:**  $H_0: M_1 = M_2 = X$      $H_{\text{альт}}: M_1^2 \neq M_2^2$

**Два случая:** при равенстве генеральных дисперсий и при их неравенстве



# Параметрические критерии

## t-критерий Стьюдента

Дисперсии равны  $\sigma_1^2 = \sigma_2^2$

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2|}{s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \sqrt{\frac{df_1 \cdot s_1^2 + df_2 \cdot s_2^2}{df_1 + df_2}}$$

Сравнить с  $t_{\text{крит.}}$  для  
 $df = n_1 + n_2 - 2$

Если  $t < t_{\text{крит}}$  для  $p < 0,01$ , то гипотеза  $H_0$  верна

## t-критерий Стьюдента

Дисперсии неравны

$$|t| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Найти  $df$  по формуле:

$$\frac{1}{df} = \frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1}$$

Где  $c = \frac{s_1^2/n_1}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$  и сравнить

# Параметрические критерии

Основной принцип критерия:

$$t = (\text{наблюдаемое} - \text{ожидаемое}) / \text{s.e.}$$

**Одновыборочный t-критерий:** сравнить среднее выборки со средним генеральной совокупности

**Независимый 2-выборочный t-критерий:** сравнить средние 2 независимых выборок

**T-критерий для 2 зависимых выборок:** сравнить изменение среднего в выборке «до» и «после»

# Многофункциональные критерии

## $F$ - критерий (угловое преобразование) Фишера

**Назначение критерия:** решать задачи сопоставления уровней исследуемого признака, сдвигов в значениях исследуемого признака и сравнения распределений;

**Суть критерия:** определяет долю (%) наблюдений в данной выборке, которая характеризуется интересующим исследователя эффектом.

### Ограничения и возможности критерия:

- а) измерения могут быть сделаны по любой шкале;
- б) оценивает 2 выборки!;
- в)  $N$  каждой выборки  $> 5$ .

# Многофункциональные критерии

## $\varphi$ - критерий Фишера

### Алгоритм подсчета

1. Определить значения признака, говорящие о наличии эффекта (в сложных случаях использовать критерий  $\lambda$  Колмогорова-Смирнова)
2. Составить и заполнить таблицу:  
  
1 выборка —  $n_1$  есть эффект —  $n_2$  нет эффекта  
2 выборка —  $n_3$  есть эффект —  $n_4$  нет эффекта

# Многофункциональные статистические критерии

## $\varphi$ - критерий Фишера

### Алгоритм подсчета

3. Определить по каждой выборке процентные доли испытуемых, у которых «есть эффект», записать %.
4. Проверить, не равняется ли одна из сопоставляемых процентных долей нулю. Если да, использовать  $\chi^2$  -критерий
5. Определить по таблицам величины углов  $\varphi_1$  и  $\varphi_2$  для каждой из сопоставляемых процентных долей. Обозначить **больший % как угол  $\varphi_1$**

# Многофункциональные статистические критерии

## $\varphi$ - критерий Фишера

### Алгоритм подсчета

6. Посчитать значение  $\varphi$  — критерия по формуле:

$$\varphi = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

Где  $n_1$  и  $n_2$  — объем выборок

7. Сравнить полученное значение с критическими:

$$\varphi_{\text{эмп}} = \begin{cases} 1,64 (\rho \leq 0,05) \\ 2,31 (\rho \leq 0,01) \end{cases}$$

8. Если  $\varphi_{\text{эмп}} \geq \varphi_{\text{кр}}$ ,  $H_0$  отвергается (различия статистически значимы).

# Многофункциональные критерии

## Биномиальный $t$ -критерий

**Цель:** сопоставления частоты встречаемости какого-либо эффекта в выборке с теоретической или заданной частотой его встречаемости; для  $5 < N < 300$ ;

## $\chi^2$ - критерий Пирсона

**Цель:** а) сопоставление эмпирического распределения признака с теоретическим; б) сопоставление двух, трех или более эмпирических распределений одного и того же признака.

**Ограничения:**  $N > 30$  (чем больше, тем лучше); неперекрывающиеся разряды признака; требуется поправка на непрерывность

# Проверка характера распределения

## 1) Критерий Колмогорова-Смирнова:

сравнение двух распределений, сравнение эмпирического и теоретического распределений.

## 2) Критерий Шапиро-Уилка:

сравнение распределения выборки с нормальным.