Проверка гипотез. Критерии однородности.

Леонид Иосипой

Курс «Вероятностные модели и статистика»
Центр непрерывного образования, ВШЭ

27 апреля 2021

- Повторение
- Параметрические критерии однородности
- Непараметрические критерии однородности

Повторение

1. Проверка гипотез.

   Правило, позволяющее «принять» или отвергнуть нулевую гипотезу H_0 относительно альтернативы H_1 на основе данных называется статистическим критерием.

   Обычно критерий задается при помощи статистики критерия $T(x_1, \ldots, x_n)$ — функции от данных.

Повторение

1. Проверка гипотез.

   Статистика критерия T должна обладать двумя важными свойствами:

   - при верной H_0 статистика T должна принимать умеренные значения (обычно малые или средние), а при неверной H_0 — экстремальные (обычно большие).
   - при верной H_0 статистика T должна иметь известное нам распределение G_0, а при неверной H_0 — распределение отличное от G_0.

Повторение

1. Проверка гипотез.

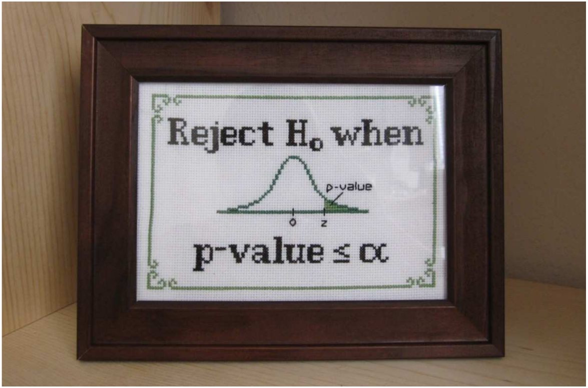
   Достигаемый/Фактический уровень значимости или p-value — это вероятность для статистики T при верной H_0 принять значение $t = T(\mathbf{x})$, которое получилось на данной выборке, или ещё более экстремальное.

   Если для статистики T экстремальными значениями являются большие значения, то это можно записать так:

$$p(\mathbf{x}) = \mathbb{P}(T(\mathbf{X}) \geq t \mid H_0).$$

Повторение

1. Проверка гипотез.



Повторение

1. Проверка гипотез.

   Если значение p-value мало, то данные свидетельствуют против нулевой гипотезы H_0 в пользу альтернативы H_1.

   Если значение p-value недостаточно мало, то данные не свидетельствуют против H_0 в пользу альтернативы H_1.

   Чем ниже p-value, тем сильнее данные свидетельствуют против нулевой гипотезы в пользу альтернативы.

Повторение

1. Проверка гипотез.

Вероятность отвергнуть нулевую гипотезу зависит не только от того, насколько она отличается от истины, но и от параметров задачи, например, размера выборки n.

По мере увеличения n нулевая гипотеза H_0 может сначала приниматься, но потом будут выявлены более тонкие несоответствия выборки гипотезе H_0, и она будет отвергнута.

При помощи инструмента проверки гипотез нельзя доказать справедливость нулевой гипотезы!

Повторение

2. Критерии согласия.

Пусть $X_1, \ldots, X_n$ — выборка из некоторого неизвестного распределения F.

В критериях согласия в качестве нулевой гипотезы H_0 рассматривается гипотеза $F \in \mathcal{F}_\theta$, то есть принадлежность F какому-то параметрическому семейству распределений. Альтернативой H_1 обычно является принадлежность F всем остальным распределениям.

Повторение

2. Критерии согласия.

   Для построения критерия согласия достаточно найти некоторое свойство, которые бы выполнялось для всех распределений из нашего класса и на его основе придумать статистику.

   При этом сколько-то удовлетворительно мажорировать вероятность ошибки второго рода не удается, поскольку вне нашего параметрического семейства есть сколь угодно похожие на наши распределения.

Повторение

2. Критерии согласия.

Критерий Колмогорова

выборка:	$\mathbf{X} = (X_1, \ldots, X_n)$ $X_i \sim F$, $\underline{F \text{ непрерывно}}$
нулевая гипотеза:	$H_0: \ F = F_0$
альтернатива:	$H_1: \ F \neq F_0$
статистика:	$\sqrt{n} D_n = \sqrt{n} \cdot \sup_{u \in \mathbb{R}} \left\| \widehat{F}_n(u) - F_0(u) \right\|$
нулевое распределение:	$\sqrt{n} D_n \sim K$ – распределение Колмогорова

Повторение

2. Критерии согласия.

Критерий Пирсона (хи-квадрат)

выборка: $\mathbf{X} = (X_1, \ldots, X_n)$

$X_i \sim F$, $\underline{F \text{ дискретно}}$

нулевая гипотеза: $H_0: \; F = F_0$

альтернатива: $H_1: \; F \neq F_0$

статистика: $T_n = \sum_{i=1}^{k} \frac{(\nu_i - np_i)^2}{np_i}$

нулевое распределение: $T_n \sim \chi^2_{k-1}$ — хи-квадрат с $k-1$ степенью свободы (k — количество возможных значений)

Повторение

2. Критерии согласия.

Проверка экспоненциальности (показательности)

- Исключение неизвестного параметра;
- Оценивание параметра интенсивности и его подстановка в критерий Колмогорова (будьте внимательны, меняется предельное распределение!);
- Другие специализированные критерии: Гини, Шапиро-Уилка для экспоненциального случая, Андерсона-Дарлинга и т.д.

Повторение

2. Критерии согласия.

Проверка нормальности

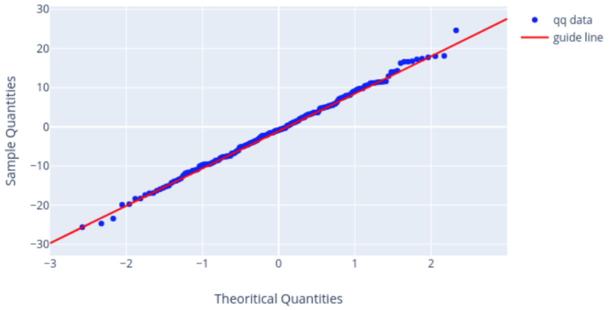
- Исключение неизвестных параметров;
- Оценивание параметров распределения и их подстановка в критерий Колмогорова (будьте внимательны, меняется предельное распределение!);
- Другие специализированные критерии: Шапиро-Уилка, Харке-Бера и т.д.

Повторение

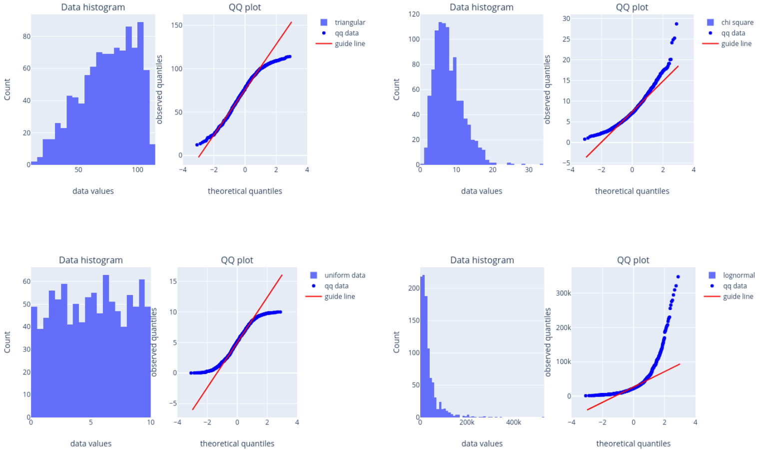
3. Квантильный график(Q-Q Plot).

   Согласие выборки с параметрическим семейством распределений, которое образовано сдвигом/масштабом, можно проверить визуально с помощью квантильного графика (Q-Q Plot).

3. Квантильный график(Q-Q Plot).



3. Квантильный график(Q-Q Plot).



Критерии однородности

Теперь перейдем к критериям однородности (A/B-тестам).

Они позволяют проверить гипотезу о том, что две выборки взяты из одного распределения.

Критерии однородности

Примеры:

1. Лабораторных мышей поместили в двухкомнатные клетки. В одной комнате висело зеркало, а в другой — нет. С целью установить, есть ли у мышей какие-то предпочтения насчет зеркал, измерялась доля времени, которое каждая мышь проводила в каждой из своих двух клеток. По данным долям требуется выяснить, имеется ли значимое предпочтение у мышей или отклонения являются чисто случайными и объясняются «естественной» дисперсией выбранной характеристики.

Критерии однородности

Примеры:

2. Первая выборка — это пациенты, которых лечили препаратом А. Вторая выборка — пациенты, которых лечили препаратом Б. Значения в выборках — это некоторая характеристика эффективности лечения (уровень метаболита в крови, температура через три дня после начала лечения, срок выздоровления и т.д.) Требуется выяснить, имеется ли значимое различие эффективности препаратов А и Б, или различия являются чисто случайными и объясняются «естественной» дисперсией выбранной характеристики.

Критерии однородности

Примеры:

3. Первая выборка — это значения некоторой характеристики состояния пациентов, записанные до лечения. Вторая выборка — это значения той же характеристики состояния тех же пациентов, записанные после лечения. Требуется выяснить, имеется ли значимое отличие в состоянии пациентов до и после лечения, или различия чисто случайны.

Критерии однородности

Примеры:

4. Первая выборка — это дни, когда в супермаркете проходила промо-акция типа А (красные ценники со скидкой). Вторая выборка — дни промо-акции типа Б (при покупке двух единиц — подарок). Значения в выборках — это показатель эффективности промо-акции (объём продаж, либо выручка в рублях). Требуется выяснить, какой из типов промо-акции более эффективен.

Критерии однородности

Мы будем различать следующие случаи:

1. Случай одной выборки.
2. Случай двух независимых выборок.
3. Случай двух зависимых выборок.

Критерии однородности

Кроме того, мы будем рассматривать параметрические и непараметрические критерии.

В параметрических критериях делается предположение о том, что выборки взяты из некоторого параметрического семейства распределений, а в непараметрических — нет.

Критерии однородности

Непараметрические критерии менее чувствительные (потому что более общие), но зато они не требуют «идеальных» условий, например, нормальности данных.

Часто бывает так, что при совсем небольших отклонениях от «идеальных» условий непараметрические критерии работают значительно лучше параметрических.

Начнем с параметрических критериев.

Параметрические критерии однородности

1a. Одновыборочный Z-критерий.

выборка: $\mathbf{X} = (X_1, \ldots, X_n)$
$X_i \sim \mathcal{N}(\mu, \sigma^2)$, $\underline{\sigma \text{ известна}}$

нулевая гипотеза: $H_0 : \mu = \mu_0$ (здесь $\mu_0 \in \mathbb{R}$ задано)

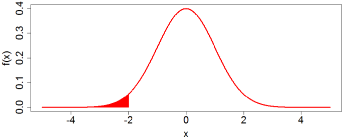
альтернатива: $H_1 : \mu \neq \mu_0$ или $\mu > \mu_0$ или $\mu < \mu_0$

статистика: $Z_n = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$

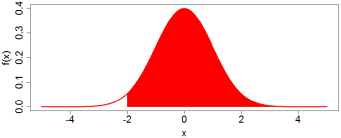
нулевое распределение: $Z_n \sim \mathcal{N}(0, 1)$

Параметрические критерии однородности

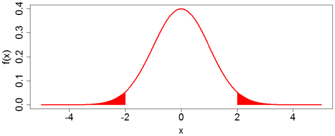
1a. Одновыборочный Z-критерий.



$$H_1 : \mu < \mu_0$$



$$H_1 : \mu > \mu_0$$



$$H_1 : \mu \neq \mu_0$$

Параметрические критерии однородности

1б. Одновыборочный t-критерий.

выборка: $\mathbf{X} = (X_1, \ldots, X_n)$

$X_i \sim \mathcal{N}(\mu, \sigma^2)$, $\underline{\sigma \text{ неизвестна}}$

нулевая гипотеза: $H_0: \ \mu = \mu_0$ (здесь $\mu_0 \in \mathbb{R}$ задано)

альтернатива: $H_1: \ \mu \neq \mu_0$ или $\mu > \mu_0$ или $\mu < \mu_0$

статистика: $T_n = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$,

где $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$

нулевое распределение: $T_n \sim \mathbf{T}_{n-1}$

Параметрические критерии однородности

2а. Двухвыборочный Z-критерий, независимые выборки.

выборки: $\mathbf{X} = (X_1, \ldots, X_{n_1}),\ X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$

$\mathbf{Y} = (Y_1, \ldots, Y_{n_2}),\ Y_i \sim \mathcal{N}(\mu_2, \sigma_2^2)$

$\mathbf{X}, \mathbf{Y}$ независимые, σ_1 и σ_2 известны

нулевая гипотеза: $H_0: \ \mu_1 = \mu_2$

альтернатива: $H_1: \ \mu_1 \neq \mu_2$ или $\mu_1 > \mu_2$ или $\mu_1 < \mu_2$

статистика: $$Z_n = \frac{\overline{X} - \overline{Y}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

нулевое распределение: $Z_n \sim \mathcal{N}(0, 1)$

Параметрические критерии однородности

2б. Двухвыборочный t-критерий, независимые выборки.

выборки: $\mathbf{X} = (X_1, \dots, X_{n_1}),\ X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$

$\mathbf{Y} = (Y_1, \dots, Y_{n_2}),\ Y_i \sim \mathcal{N}(\mu_2, \sigma_2^2)$

$\mathbf{X}, \mathbf{Y}$ независимые, σ_1 и σ_2 неизвестны

нулевая гипотеза: $H_0:\ \mu_1 = \mu_2$

альтернатива: $H_1:\ \mu_1 \neq \mu_2$ или $\mu_1 > \mu_2$ или $\mu_1 < \mu_2$

статистика: $$T_n = \frac{\overline{X} - \overline{Y}}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

нулевое распределение: $T_n \approx \mathbf{T}_k$ для некоторого $k \in \mathbb{N}$

Параметрические критерии однородности

2б. Двухвыборочный t-критерий, независимые выборки.

Задача сравнения средних двух нормально распределённых выборок при неизвестных и неравных дисперсиях известна как проблема Беренса-Фишера.

Точного решения этой задачи нет.

Параметрические критерии однородности

2б. Двухвыборочный t-критерий, независимые выборки.

Однако рассмотренная аппроксимация достаточно точна в двух ситуациях:

1. Если выборки одинакового объема $n_1 = n_2$.
2. Если знак неравенства между n_1 и n_2 такой же, как между σ_1 и σ_2, то есть выборка с большей дисперсией имеет больший объем.

Параметрические критерии однородности

3. Двухвыборочный t-критерий, зависимые выборки.

Рассмотрим теперь случай связанных выборок когда элементы X_i и Y_i соответствуют одному и тому же объекту, но измерения сделаны в разные моменты (например, до и после применения лекарства).

Размеры выборок в этом случае должны совпадать:

$$n_1 = n_2 = n.$$

Параметрические критерии однородности

3. Двухвыборочный t-критерий, зависимые выборки.

Рассмотрим выборку, образованную разностями X_i и Y_i:

$$Z_i = Y_i - X_i, \quad i = 1, \ldots, n.$$

Сравнение средних в зависимых выборках ничем не отличается от сравнения среднего разности Z_i с нулём.

Параметрические критерии однородности

3. Двухвыборочный t-критерий, зависимые выборки.

выборки: $\mathbf{X} = (X_1, \ldots, X_n),\ X_i \sim \mathcal{N}(\mu_1, \sigma^2)$
$\mathbf{Y} = (Y_1, \ldots, Y_n),\ Y_i \sim \mathcal{N}(\mu_2, \sigma^2)$

$\mathbf{X}, \mathbf{Y}$ зависимые, σ неизвестна

нулевая гипотеза: $H_0: \ \mu_1 = \mu_2$

альтернатива: $H_1: \ \mu_1 \neq \mu_2$ или $\mu_1 > \mu_2$ или $\mu_1 < \mu_2$

статистика: $T_n = \dfrac{\overline{Z}}{S/\sqrt{n}}$,

где $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \overline{Z})^2$, $Z_i = X_i - Y_i$

нулевое распределение: $T_n \sim \mathbf{T}_{n-1}$

Непараметрические критерии однородности

Теперь перейдем к непараметрическим критериям. Мы не будем предполагать, что данные имеют нормальное распределение.

Достоинством непараметрических критерием является то, что они не требуют нормальности данных, пригодны для выборок малого размера и слабо реагируют на присутствие «выбросов» в данных.

Начнем с модификации известных нам критериев.

Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

Для непрерывных распределений можно модифицировать критерий Колмогорова.

Пусть даны две выборки $\mathbf{X} = (X_1, \ldots, X_{n_1})$ и $\mathbf{Y} = (Y_1, \ldots, Y_{n_2})$ с функциями распределения F_X и F_Y соответственно.

Обозначим их эмпирические функции распределения через $\widehat{F}_{X,n_1}$ и $\widehat{F}_{Y,n_2}$ и рассмотрим статистику

$$D_n = \sup_{u \in \mathbb{R}} \left| \widehat{F}_{X,n_1}(u) - \widehat{F}_{Y,n_2}(u) \right|.$$

Если $F_X = F_Y$, то D_n должна принимать малые значения.

Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

Более точно, оказывается, что при выполнении нулевой гипотезы $F_X = F_Y$, для любого $t > 0$ выполняется

$$\mathbb{P}\left(\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_n \leq t\right) \to K(t),$$

где $K(t)$ — функция Колмогорова.

При $n_1, n_2 \geq 20$ аппроксимация является достаточно точной.

Данный результат можно использовать, чтобы построить критерий однородности.

Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

выборки: $\mathbf{X} = (X_1, \ldots, X_{n_1}),\ X_i \sim F_X$

$\mathbf{Y} = (Y_1, \ldots, Y_{n_2}),\ Y_i \sim F_Y$

$\mathbf{X}, \mathbf{Y}$ независимые; F_X, F_Y непрерывные

нулевая гипотеза: $H_0 :\ F_X = F_Y$

альтернатива: $H_1 :\ F_X \neq F_Y$

статистика: $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_n$

нулевое распределение: $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_n \sim K$ – распределение Колмогорова

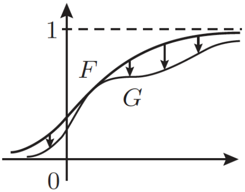
Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

Критерий Колмогорова-Смирнова можно также использовать и при альтернативе доминирования:

$F_X(u) \geq F_Y(u)$ для всех $u \in \mathbb{R}$, причем хотя бы для одного u неравенство строгое.

Мы будем обозначать это как $F_X \succeq F_Y$.



Данное свойство означает, что случайная величина Y стохастически больше случайной величины X, поскольку

$$\mathbb{P}(X \geq u) \leq \mathbb{P}(Y \geq u) \quad \text{для всех } u \in \mathbb{R}.$$

Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

При этом необходимо рассмотреть другую статистику:

$$D_n^+ = \sup_{u \in \mathbb{R}} \left(\widehat{F}_{X,n_1}(u) - \widehat{F}_{Y,n_2}(u) \right).$$

У нее тоже есть предельное распределение (но уже другое):

$$\mathbb{P}\left(\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_n^+ \leq t \right) \to 1 - e^{-2t^2}.$$

Непараметрические критерии однородности

1. Критерий Колмогорова-Смирнова

выборки: $\mathbf{X} = (X_1, \ldots, X_{n_1}),\ X_i \sim F_X$

$\mathbf{Y} = (Y_1, \ldots, Y_{n_2}),\ Y_i \sim F_Y$

$\mathbf{X}, \mathbf{Y}$ независимые; F_X, F_Y непрерывные

нулевая гипотеза: $H_0:\ F_X = F_Y$

альтернатива: $H_1:\ F_X \succeq F_Y$

статистика: $\sqrt{\frac{n_1 n_2}{n_1+n_2}} D_n^+$

нулевое распределение: $\sqrt{\frac{n_1 n_2}{n_1+n_2}} D_n^+ \sim K^+$ – указано выше (не имеет названия)

Непараметрические критерии однородности

2. Критерий хи-квадрат

Для дискретных распределений можно модифицировать критерий хи-квадрат.

Допустим, что имеется k независимых между собой выборок размеров n_i из дискретных распределений F_i, $i = 1, \ldots, k$.

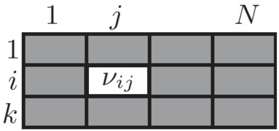
Необходимо проверить гипотезу однородности
$H_0: \; F_1 = F_2 = \ldots = F_k$
против альтернативы
H_1 : хотя бы один из F_i значимо отличается от других.

Непараметрические критерии однородности

2. Критерий хи-квадрат

Как обычно, будем работать с частотами:

- подсчитаем ν_{ij} — количество элементов a_j в i-й выборке для всех возможных пар индексов (i, j);
- запишем их в таблицу сопряженности размера $k \times N$, которую и будем анализировать в дальнейшем.



Непараметрические критерии однородности

2. Критерий хи-квадрат

Если гипотеза H_0 верна, то:

- ожидаемое количество наблюдений в ячейке с индексами i и j равно $n_i p_j$, где p_j — вероятность получить a_j;
- естественной оценкой для p_j служит общая по всем выборкам частота значений a_j: $\widehat{p}_j = \frac{\nu_{1,j}+\ldots+\nu_{k,j}}{n}$.
- В этих обозначениях, можно доказать, что статистика

$$T_n = \sum_{i=1}^{k} \sum_{j=1}^{N} \frac{(\nu_{ij} - n_i \widehat{p}_j)^2}{n_i \widehat{p}_j}$$

сходится к распределению хи-квадрат $\chi^2_{(k-1)(N-1)}$.

Непараметрические критерии однородности

2. Критерий хи-квадрат

выборки: $\mathbf{X_1} = (X_{1,1}, \dots, X_{1,n_1}),\ X_1 \sim F_1$

$\dots$

$\mathbf{X_k} = (X_{k,1}, \dots, X_{k,n_k}),\ X_k \sim F_k$

$\mathbf{X_1}, \dots, \mathbf{X_k}$ независимые

нулевая гипотеза: $H_0:\ F_1 = F_2 = \dots = F_k$

альтернатива: $H_1:$ есть различия

статистика: T_n

нулевое распределение: $T_n \sim \chi^2_{(k-1)(N-1)}$ – хи-квадрат с $(k-1)(N-1)$ степенями свободы

Непараметрические критерии однородности

3. Критерий Манна-Уитни

Критерий Манна-Уитни (или ранговых сумм Уилкоксона) применяется для проверки гипотезы однородности против альтернативы доминирования.

Данный критерий был предложен Уилкоксоном для выборок одинакового размера. Манн и Уитни обобщили его на случай выборок разного размера.

Непараметрические критерии однородности

3. Критерий Манна-Уитни

Напомним, что по любой выборке $X_1, \ldots, X_n$ всегда можно сопоставить вариационный ряд, то есть упорядочить её по неубыванию:

$$X_{(1)} \leq \ldots < \underbrace{X_{(j_1)} = \ldots = X_{(j_2)}}_{\text{связка размера } j_2 - j_1 + 1} < \ldots \leq X_{(n)}.$$

Рангом наблюдения X_i называется:

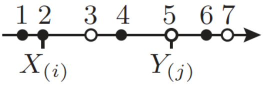
- если X_i не попадает в связку, то его позиция в вариационном ряду;
- если X_i попадает в связку, то $(j_1 + j_2)/2$; то есть в связке все объекты получают одинаковый средний ранг.

Непараметрические критерии однородности

3. Критерий Манна-Уитни

Вычислим статистику V_n критерия Манна-Уитни:

1. Обозначим через R_j ранг порядковой статистики $Y_{(j)}$, $j = 1, \ldots, m$, в вариационном ряду, построенном по объединенной выборке $(X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2})$.
2. Положим $V_n = R_1 + \ldots + R_{n_2}$.



Непараметрические критерии однородности

3. Критерий Манна-Уитни

Если верна гипотеза однородности, то значения $Y_{(j)}$ должны быть рассеяны по всему вариационному ряду; напротив, достаточно большое значение V_n указывает на тенденцию преобладания Y_j над X_i, что свидетельствует в пользу справедливости гипотезы доминирования.

При выполнении гипотезы однородности V_n распределена также как $Z_1 + \ldots + Z_{n_2}$, где Z_i — выборка без возвращения из чисел $\{1, \ldots, n_1 + n_2\}$. Для больших n_1, n_2 можно использовать сходимость к нормальному распределению.

Непараметрические критерии однородности

3. Критерий Манна-Уитни

выборки: $\mathbf{X} = (X_1, \ldots, X_{n_1}),\ X_i \sim F_X$

$\mathbf{Y} = (Y_1, \ldots, Y_{n_2}),\ Y_i \sim F_Y$

$\mathbf{X}, \mathbf{Y}$ независимые

нулевая гипотеза: $H_0: \ F_X = F_Y$

альтернатива: $H_1: \ F_X \succeq F_Y$ или $F_X \preceq F_Y$ или $F_X \neq F_Y$

статистика: V_n

нулевое распределение: указано на предыдущем слайде

Непараметрические критерии однородности

Перейдем теперь к проверке однородности для случая зависимых выборок. Напомним, что в этой модели выборки должны быть одинакового размера.

Как было в случае параметрических критериев, мы будем работать с приращениями

$$Z_i = Y_i - X_i, \quad i = 1, \ldots, n.$$

Непараметрические критерии однородности

Разложим каждое из них на две части:

$$Z_i = \theta + \varepsilon_i, \quad i = 1, \ldots, n,$$

где θ — интересующий нас эффект воздействия (систематический сдвиг), а ε_i — случайные ошибки, включающие в себя влияние неучтенных факторов на Z_i.

Мы будем дополнительно далее предполагать, что $\varepsilon_1, \ldots, \varepsilon_n$ независимы и имеют непрерывные (вообще говоря, разные) распределения с равной нулю медианой.

Непараметрические критерии однородности

4. Критерий знаков

Самым простым непараметрическим критерием для проверки однородности двух зависимых выборок является критерий знаков.

Статистикой критерия является величина

$$S_n = \sum_{i=1}^{n} \mathbf{I}_{\{Z_i>0\}}.$$

Тогда при верной гипотезе однородности S_n будет иметь биномиальное распределение $\mathbf{B}_{n,1/2}$, а при альтернативе — биномиальное распределение $\mathbf{B}_{n,p}$ с $p \neq 1/2$. Для больших n можно использовать сходимость к нормальному закону.

Непараметрические критерии однородности

4. Критерий знаков

выборки: $\mathbf{X} = (X_1, \ldots, X_n)$, $X_i \sim F_X$
$\mathbf{Y} = (Y_1, \ldots, Y_n)$, $Y_i \sim F_Y$

$\mathbf{X}, \mathbf{Y}$ зависимые

нулевая гипотеза: $H_0: \ \theta = 0$

альтернатива: $H_1: \ \theta \neq 0$ или $\theta > 0$ или $\theta < 0$

статистика: S_n

нулевое распределение: $S_n \sim \mathbf{B}_{n,1/2}$ для малых выборок
нормальное для больших выборок

Непараметрические критерии однородности

5. Критерий знаковых рангов Уилкоксона

Сделаем дополнительное предположение, что случайные величины $\varepsilon_1, \ldots, \varepsilon_n$ имеют одинаковое распределение, симметричное относительно нуля (то есть медианы).

Условие строгой симметрии относительно медианы является почти столь же нереалистичным, как и предположение, что распределение величин Z_i в точности нормально.

Как правило, надежно проверить симметрию можно лишь по выборке из нескольких сотен наблюдений.

Непараметрические критерии однородности

5. Критерий знаковых рангов Уилкоксона

Критерий знаковых рангов Уилкоксона основан на статистике

$$W_n = U_1 R_1 + \ldots + U_n R_n,$$

где

- $U_i = \mathbf{I}_{\{Z_i > 0\}}$;
- R_i — ранги величин $|Z_i|$ в ряду $|Z_1|, \ldots, |Z_n|$.

Известно, что при выполнении гипотезы однородности W_n сходится к нормальному распределению. А для малых выборок можно посчитать распределение W_n явно (с помощью метода Монте-Карло).

Непараметрические критерии однородности

5. Критерий знаковых рангов Уилкоксона

выборки: $\mathbf{X} = (X_1, \dots, X_n),\ X_i \sim F_X$
$\mathbf{Y} = (Y_1, \dots, Y_n),\ Y_i \sim F_Y$

$\mathbf{X}, \mathbf{Y}$ зависимые

нулевая гипотеза: $H_0: \ \theta = 0$

альтернатива: $H_1: \ \theta \neq 0$ или $\theta > 0$ или $\theta < 0$

статистика: W_n

нулевое распределение: табличное для малых выборок
нормальное для больших выборок

Непараметрические критерии однородности

Резюме

- При проверке гипотезы однородности крайне важно определиться, с каким из двух случаев мы имеем дело: двумя реализациями независимых между собой наблюдений или парными повторными наблюдениями.
- Критерии, применимые для зависимых выборок, можно использовать и для независимых. Однако при этом игнорируется важная информация о совместной независимости, что снижает чувствительность методов.
- В свою очередь применение критериев, предполагающих независимость выборок, в зависимом случае представляет собой грубую ошибку.

Непараметрические критерии однородности

Оценка параметра сдвига

Если гипотеза однородности H_0 отвергнута в пользу альтернативы доминирования, можно дополнительно оценить параметр «сдвига».

В качестве оценки сдвига можно взять:

- в случае независимых выборок — выборочную медиану попарных приращений

$$\mathrm{MED}\{X_i - Y_j,\ \ i = 1, \ldots, n_1, j = 1, \ldots, n_2\}.$$

- в случае зависимых выборок — выборочную медиану приращений

$$\mathrm{MED}\{Z_i,\ \ i = 1, \ldots, n\}.$$

Непараметрические критерии однородности

Критика критериев Стьюдента (Z- и t-критериев)

Эффективность критериев Стьюдента быстро уменьшается при отклонении от строгой нормальности.

Для иллюстрации этого факта рассмотрим модель Тьюки с функцией распределения $F_\varepsilon(u) = (1-\varepsilon)\Phi(u) + \varepsilon\Phi(u/3)$, где $\Phi(u)$ — функция распределения $\mathcal{N}(0, 1)$.

Непараметрические критерии однородности

Критика критериев Стьюдента (Z- и t-критериев)

В следующей таблице показано изменение эффективности $E(F_\varepsilon)$ критерия Манна-Уитни относительно критерия Стьюдента при небольшом «утяжелении хвостов».

ε	0	0,01	0,03	0,05	0,08	0,10	0,15
$E(F_\varepsilon)$	0,955	1,009	1,108	1,196	1,301	1,373	1,497

Более того, оказывается, что для произвольного гладкого симметричного распределения $E(F_\varepsilon) \geq 0.864$, и существуют распределения, на которых $E(F_\varepsilon)$ сколь угодно велика.

Спасибо за внимание!