Кафедра математической статистики ИВМиИТ КФУ 31 декабря 2020 г.

Научно-популярная статья к выпуску подготовил Казанцев А.В.

Автор: Мария Саженкова, магистрант 2 курса кафедры математической статистики

Т-критерий Вилкоксона

Автор статьи – Мария Саженкова – после вручения дипломов бакалавров в июле 2019 г.

Мария Саженкова – бакалавр математики (КФУ, 2019). Выпускная квалификационная работа: Саженкова М.Г. «Свойства процедур множественного тестирования». Научный руководитель – доцент С.В. Симушкин

Оглавление

Введение

Фрэнк Вилкоксон – американский химик и статистик, автор более 60 научных статей [1], а также известных статистических критериев, таких как Т-критерий Вилкоксона, критерий Манна-Уитни-Вилкоксона.

О биографии [2] ученого известно мало фактов. Родиной Вилкоксона (1892-1965) является Ирландия, графство Корк. Образование получил в военном колледже Пенсильвании, а также в Рутгенберском университете и получил степень магистра химии, в Корнеллском университете стал доктором философии по физической химии (раздел химии, наука об общих законах строения, структуры и превращения химических веществ). Статистикой Вилкоксон стал заниматься после прочтения книги Рональда Фишера, которая называется «Statistical Methods for Research Workers» [3].

История возникновения Т-критерия, которому будет посвящена данная работа, неизвестна. Суть этого приема состоит в сравнении связанных между собой выборок с учетом выделенного конкретного признака. При этом важно, чтобы выбранный критерий анализа был оценен в порядковой шкале или, выражаясь языком математического анализа, чтобы признак можно было занумеровать.

В основе метода лежит многоступенчатая обработка материалов: ранжирование, суммирование рангов, определение стабильностей и случайностей, интенсивности. Данная методика позволяет оценить различия между двумя признаками, рядами изменений, которые были выполнены в отношении выборки.

Т-критерий Вилкоксона

Пусть $X^{(n)}$ – выборка из X, $Y^{(m)}$ – выборка из Y. Выборки должны относиться к двум независимым группам наблюдений одной и той же характеристики. Рассматриваемый механизм относится к непараметрическим методам исследований, поэтому перед применением критерия можно не устанавливать факт о равномерном распределении признаков.

Как строится критерий:

Выдвигается гипотеза об однородности выборок, другими словами, гипотеза о том, что $F_Y(x) = F_X(x)$. Если гипотеза отвергается, то принимается альтернатива $F_Y(x) > F_X(x)\ \forall x$ - это означает, что распределение $F_X(x)$ сдвинуто влево относительно $F_Y(x)$. Заранее выбирается уровень значимости α (обычно берется 0,01 или 0,05).

Берутся обе выборки $(x_1, x_2, \dots, x_n)$, $(y_1, y_2, \dots, y_m)$, объединяются исследователем в одну и упорядочиваются по возрастанию:

№	1	2	3	4	5	6	…	n+m-2	n+m-1	n+m
Инф-ция об исх. выборке	x	y	x	x	y	y		y	y	x
Новая выборка	r_{1i}	r_{2i}	r_{3i}	r_{4i}	r_{5i}	r_{6i}	…	$r_{(n+m-2)i}$	$r_{(n+m-1)i}$	$r_{(n+m)i}$

Информация о принадлежности значений исходным выборкам должна быть сохранена. В примере буква i обозначает, что цифра справа - новый индекс в новой объединенной выборке. Этот индекс называется рангом.

Определение. Ранг – порядковый номер элемента.

Чтобы вычислить статику Вилкоксона, необходимо выбрать все ранги первой выборки, то есть выборки X (которая по предположению сдвинута влево), и суммировать их. В формализованном виде:

$$W = \sum_{j=1}^{n} r_j,$$

где $r_1, \dots, r_n$ – ранги всех значений выборки X. Известно, что статистика Вилкоксона асимптотически нормальна [4] с $\mu_W = \frac{n(n+m+1)+1}{2}$, $\sigma_W^2 = \frac{nm(n+m+1)}{12}$.

После вычисления статистики Вилкоксона для проверки гипотезы однородности определяется p-значение:

$$p = P_0(P \leq w) \sim \Phi\left(\frac{w - \mu_W}{\sigma_W}\right),$$

Если $p \leq \alpha$, гипотеза однородности отвергается, принимается альтернатива, если $p > \alpha$, гипотеза однородности принимается.

Критерий Вилкоксона обычно применяется в 2 случаях:

1.1) когда распределение выборок неизвестно, но есть предположение о том, что оно не подчиняется нормальному закону;

1.2) когда выдвигается гипотеза однородности распределений выборок.

Когда применять критерий бесполезно:

2.1) если в ходе испытаний элементы в выборке принимают конкретный набор значений, к примеру (-1, 0, 1);

2.2) если в выборке много одинаковых значений.

<u>Пример 1.1-1.2</u> Пусть имеется 2 выборки, сгенерированные случайным образом из экспоненциального распределения,

$X = (x_1, \dots, x_{500}), Y = (y_1, \dots, y_{1000}) \sim F_X(x) = \begin{cases} 1 - e^{-\lambda x}, x \geq 0, \\ 0, \quad\quad\quad x < 0, \end{cases}$ где $\lambda = 1$. Из выборки Y вычтем также сгенерированную случайным образом выборку $Q =$

$(q_1, \dots, q_{1000}) \sim \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b, \\ 1, & x > b, \end{cases}$ где $a = 0, b = 0.1$. Проверим гипотезу однородности выборок, H_0: $F_Y(y) = F_X(x)$.

Построим совместную выборку, найдем $W = \sum_{j=1}^{n1} r_j = 383844$,

$\Phi\left(\frac{w-\mu_W}{\sigma_W}\right) == 1.086702$,

p = $\alpha_{\text{крит}}$ = 1-$P(W < w)$=0.14,

$\alpha_{\text{крит}} = 0.14 > \alpha = 0.05 \rightarrow$ гипотеза однородности выборок принимается.

Пример 2.1-2.2 Пусть выборки принимают конкретный набор значений, Например $x = (-1, 0, 1, 1, 0, -1, -1, -1, 1), y = (0, -1, 1, 0, 0, 0, 0, 0, 1, 1, 1)$.

Обработаем выборку, чтобы вычислить ранги, получим:

```
      [,1] [,2]
 [1,] "x"  "-1"
 [2,] "x"  "-1"
 [3,] "x"  "-1"
 [4,] "x"  "-1"
 [5,] "y"  "-1"
 [6,] "x"  "0"
 [7,] "x"  "0"
 [8,] "y"  "0"
 [9,] "y"  "0"
[10,] "y"  "0"
[11,] "y"  "0"
[12,] "y"  "0"
[13,] "y"  "0"
[14,] "x"  "1"
[15,] "x"  "1"
[16,] "x"  "1"
[17,] "y"  "1"
[18,] "y"  "1"
[19,] "y"  "1"
[20,] "y"  "1"
```

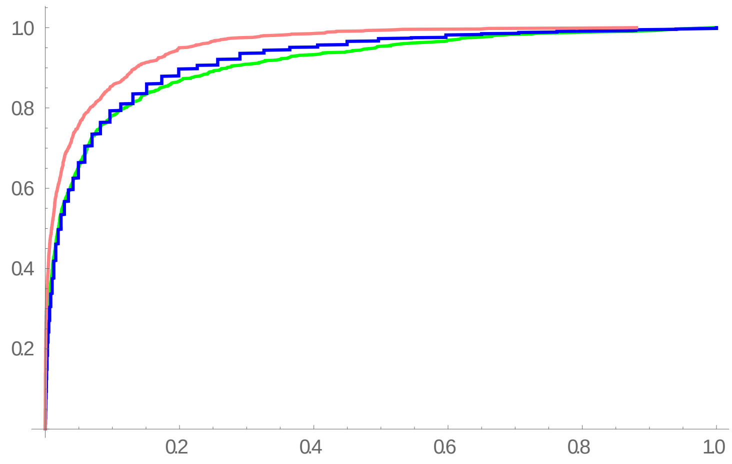
В данной ситуации нет очевидного ответа, каким образом присваивать ранги. На картинке мы видим, что $rang(x_1) = 1, rang(x_6) = 2, rang(x_7) = 3, rang(x_8) = 4, rang(y_2) = 5$. В то же время, совершенно справедливо присвоить , $rang(y_2) = 1$. В теории математической статистики до сих пор нет четкого ответа на вопрос о том, как в данном случае распределять ранги. Ответ на этот вопрос очень важен, поскольку именно от вычисления рангов зависит конечный результат, отвержение или принятие гипотезы.

Сравнение с другими критериями

Данный критерий можно сравнивать с параметрическими критериями, например, с критерием Стьюдента, также часто сравнивают Т-критерий Вилкоксона с критерием Крамера-Уэлча. Рассмотрим

Пример 3. Пусть имеется Логистическое распределение с параметрами μ и σ и три критерия: Стьюдента, Вилкоксона, Крамера-Уэлча. Требуется сравнить их мощности.

Для данного эксперимента у Логистической модели были выбраны следующие параметры: $\mu x = 0, \mu y = 2, \sigma = 1$, после чего для каждого критерия по отдельности выполнялся следующий алгоритм: генерировались две случайные выборки с заданными параметрами длины nx и ny. Далее вычислялись статистика и p-значения. Методом стохастического моделирования данный алгоритм осуществлялся $m = 1000$ раз, в результате чего был получен вектор p-значений. Стоит заметить, что для проверки мощности критериев выборки специально генерировались с разными средними. После того, как было получено три вектора p-значений (для каждого критерия) строился один общий график, на котором зеленым цветом изображался вектор p-значений критерия Стьюдента, синим вектор p-значений критерия Вилкоксона, красным вектор p-значений критерия Крамера-Уэлча.



Эксперимент показал, что все три критерия имеют примерно одинаковые мощности. График иллюстрирует, что критерий Крамера-Уэлча имеет наибольшую мощность. Данный график был получен при $nx = 10$ и $ny = 15$ для каждого критерия, стоит заметить, что вышеуказанные объемы выборок являются достаточно малыми.

По сравнению с критерием Стьюдента критерий Вилкоксона имеет преимущество, так как критерий Стьюдента хорошо работает только при одинаковых дисперсиях у двух выборок [5].

Заключение

Таким образом, использование Т-критерия Вилкоксона позволяет упорядочить имеющиеся данные, разбить их на две группы и проанализировать. Результаты, получены при помощи этого приема, признаются достоверными и обоснованными, отражают общую тенденцию изучаемого явления ил признака. Чаще всего им пользуются студенты при выполнении курсовых, дипломных работ и отчетов по практике, а также соискатели ученых степеней при доказательстве выдвинутой гипотезы или проверке действия некоего закона в современных условиях.

Список литературы

1 Karas, J. & Savage, I.R. (1967) *Publications of Frank Wilcoxon (1892—1965).* Biometrics 23(1): 1-10

2 Bradley, R.A. (1966) *Obituary: Frank Wilcoxon.* Biometrics 22(1): 192—194

3 Fisher R. A. Statistical methods for research workers //Breakthroughs in statistics. – Springer, New York, NY, 1992. – С. 66-70.

4 Симушкин С. В. Теоретические аспекты заданий курсового проекта по математической статистике //Казань: Изд-во КГУ. – 2004. – С. 43-45Симушкин С. В. и др. Методы теории вероятностей. ЧI Многомерные модели. Математические основания. – 2016.

5 Симушкин С. В. и др. Методы теории вероятностей. ЧI Многомерные модели. Математические основания. – 2016.