Однородность нескольких выборок

Непараметрические методы составляют одно из наиболее успешных направлений современной статистики. Они широко применимы, быстры в исполнении и легко усваиваются.

Я. Гаек

Диаграмма размахов

UBV (Up Boundary Value) — верхняя выборочная квартиль

LBV (Low Boundary Value) — нижняя выборочная квартиль

Верхний и нижний «усы» ограничивают диапазон невыделяющихся значений

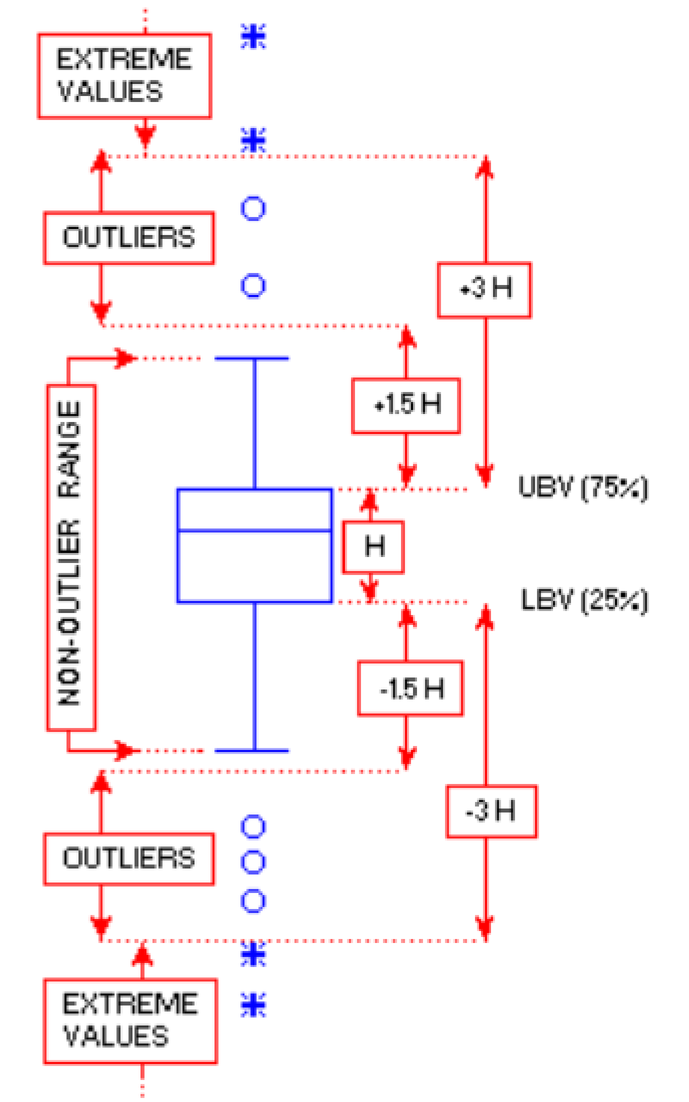
H = UBV − LBV — межквартильный размах

Горизонтальная линия внутри межквартильного диапазона — выборочная медиана

Кружки — возможные «выбросы»

Звёздочки — экстремальные значения или безусловные «выбросы»

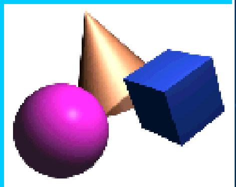
Строки таблицы данных, содержащие экстремальные значения, как правило, рекомендуется исключить из дальнейшего статистического анализа.



Детское восприятие

Данные были получены в результате психологического эксперимента, в ходе которого 36 маленьких детей были случайно разбиты на 3 группы. Каждому ребенку были показаны несколько карточек с изображениями пар нарисованных объектов: геометрических фигур, цветов, машинок и т. п. Задача состояла в том, чтобы выбрать одну из карточек. В случае «правильного» выбора дети получали награду. В первой группе характеристикой, которую необходимо было принимать во внимание, чтобы сделать правильный выбор, являлась *форма* объектов, во второй группе — *цвет*, в третьей группе — *размер*.

Признак PERFRMNC представляет собой число попыток, потребовавшихся детям для выбора карточки, представляющей соответствующую характеристику.

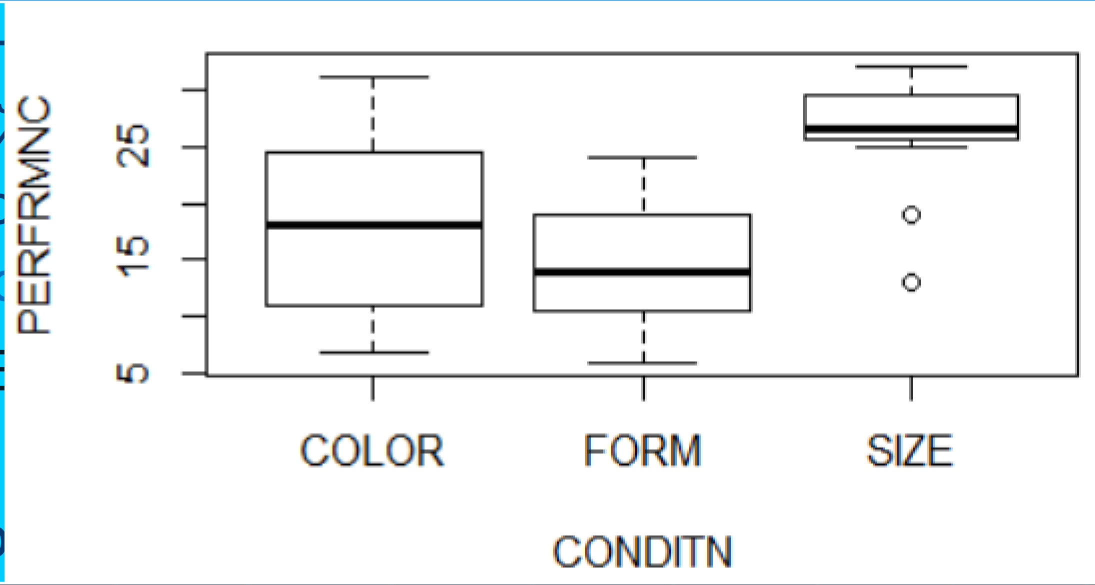


Что важнее: форма, цвет или размер?

1) Импорт

2) Построй ... PERFRMNC для каждо ... столбце CONDITN с

boxplot(PE



3) Какие в ... сти для детей фор ... ек)?

Выводы:

1) Больше всего попыток потребовалось детям, чтобы увидеть различие в размере объектов.

2) В первую очередь дети при сравнении объектов обращают внимание на форму (FORM), затем — на цвет (COLOR), а только потом — на размер (SIZE).

Однофакторная модель

<u>**Данные**</u>. Данные состоят из $\boxed{N = \sum_{j=1}^{k} n_j}$ чисел x_{ij}, по n_j чисел в j-й выборке (обработке), $j = 1, \ldots, k$. Будем считать их реализацией случайных величин (наблюдений) X_{ij}, где

$$\boxed{X_{ij} = \mu + \beta_j + \varepsilon_{ij},} \quad i = 1, \ldots, n_j; \;\; j = 1, \ldots, k.$$

Здесь μ — (неизвестное) общее среднее, β_j — (неизвестный) эффект от воздействия фактора*) для j-й выборки, ε_{ij} — случайная ошибка. Положим $\mu_j = \mu + \beta_j$.

<u>**Допущения:**</u>

Д1. *Все ошибки ε_{ij} независимы.*

Д2. *Все ε_{ij} имеют одинаковое непрерывное (неизвестное) распределение.*

Обработки			
1	2	...	k
x_{11}	x_{12}	...	x_{1k}
x_{21}	x_{22}	...	x_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{n_1 1}$	$x_{n_2 2}$	...	$x_{n_k k}$

Для проверки гипотезы *однородности обработок*

$$H_0\colon \mu_1 = \ldots = \mu_k$$

используются *критерий Краскела—Уоллиса* или (для альтернативы возрастания μ_j) *критерий Джонкхиера—Терпстры*.

*) *Фактором* называется некоторая характеристика, которая оказывает влияние на значения наблюдений. Она одинакова внутри каждой группы, но может меняться от группы к группе. Скажем, в примере 2 ниже в качестве интересующего исследователя фактора выступает *мотивация* (знание цели работы): исследуется ее влияние на производительность при выполнении монотонных производственных операций.

Критерий Краскела — Уоллиса

Критерий Краскела — Уоллиса применяется для проверки гипотезы H_0 против альтернативы

$$H_1: \text{ не все } \mu_j \text{ равны между собой.}$$

Выполним **следующие шаги**.

1) Ранжируем все N наблюдений вместе от меньшего к большему. Пусть R_{ij} обозначает ранг X_{ij} в этой совместной ранжировке.

2) Положим для $j = 1, \dots, k$

Ранги R_{ij} пробегают все значения от 1 до N.

$$S_j = \sum_{i=1}^{n_j} R_{ij}, \quad R_{\cdot j} = S_j / n_j, \quad R_{\cdot\cdot} = \frac{1}{N} \sum R_{ij} = \frac{1}{N} \frac{N(N+1)}{2} = \frac{N+1}{2}.$$

Таким образом, $R_{\cdot j}$ — это средний ранг наблюдений X_{ij}, относящихся к обработке j, $R_{\cdot\cdot}$ — общий средний ранг.

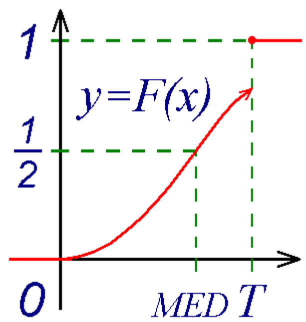
3) Вычислим значение *статистики критерия Краскела — Уоллиса* W, определяемой формулой

$$W = \frac{12}{N(N+1)} \sum_{j=1}^{k} n_j (R_{\cdot j} - R_{\cdot\cdot})^2 \quad \longrightarrow \chi^2_{k-1}$$

Медианный критерий

Медианный тест служит альтернативой критерию Краскела — Уоллиса в однофакторной модели (см. §1 темы 4). Статистикой критерия является

$$M=\sum_{j=1}^{k}\frac{(m_j-n_j/2)^2}{n_j/4},$$



где n_j обозначает размер j-й выборки, $j=1,\ldots,k$;

$$m_j=\sum_{i=1}^{n_j} I_{\{X_{ij}\geqslant MED\}},$$

где MED — выборочная медиана объединенной выборки. Если гипотеза однородности верна, то при увеличении всех n_j распределение статистики M сходится к закону χ^2_{k-1}.

Медианный тест обычно обладает меньшей мощностью по сравнению с Краскела — Уоллиса, однако его можно использовать для цензурированных (например, по времени) наблюдений, для которых не выполняется предположение о непрерывности функции распределения наблюдений.

Пример данных для однофакторной модели

Пример 1. *Содержание влаги в продукте.* 14 образцов некоторого продукта случайным образом разбили на пять групп заданных размеров. Все группы хранились в разных условиях, а после хранения у всех образцов определили содержание влаги. Данные (в %) приведены в следующей таблице (в скобках указаны ранги R_{ij}):

Условия хранения продукта				
1	2	3	4	5
7,8 (7) 8,3 (10,5) 7,6 (6) 8,4 (12) 8,3 (10,5)	5,4 (1) 7,4 (5) 7,1 (3,5)	8,1 (9) 6,4 (2)	7,9 (8) 9,5 (13) 10,0 (14)	7,1 (3,5)
$S_1 = 46$	$S_2 = 9{,}5$	$S_3 = 11$	$S_4 = 35$	$S_5 = 3{,}5$
$R_{\cdot 1} = 9{,}2$	$R_{\cdot 2} = 3{,}17$	$R_{\cdot 3} = 5{,}5$	$R_{\cdot 4} = 11{,}67$	$R_{\cdot 5} = 3{,}5$

Данные содержатся в файле Moisture.txt. Проверьте гипотезу однородности групп по содержанию влаги в продукте: установите и подключите пакет coin, примените функции kruskal_test и median_test с аргументом distribution=approximate(100000)

Множественные сравнения

Когда гипотеза однородности отвергнута, интересно узнать, в каких именно парах выборок есть значимое различие.
Для выявления таких пар можно использовать критерий Данна:

на уровне значимости, не превосходящем α, принять решение $\mu_r \neq \mu_s$, если

$$|R_{\cdot r} - R_{\cdot s}| > C_\alpha \sqrt{N(N+1)(1/n_r + 1/n_s)/12},$$

где $C_\alpha = \Phi^{-1}(1 - \alpha / k / (k-1))$, $\Phi(x)$ — функция распределения закона $N(0, 1)$.

Практическое задание 1

1) Установите пакет pgirmess, подключите его, поставив перед ним «галку» на вкладке Packages
2) Найдите все пары выборок в каждой из таблиц Kruskal и Moisture, значимо различающиеся на уровне $\alpha = 0{,}05$, с помощью функции kruskalmc [Kruskal multiple comparisons]
3) Перебором найдите с точностью 0,01 уровень значимости, при котором критерий Данна обнаруживает значимые различия в таблице Moisture

Вычисление контрастов

Контрастом Δ_{ij} называют разность уровней фактора в *i*-й и *j*-й выборках, т. е. $\Delta_{ij} = \mu_i - \mu_j$. *Уточнённые оценки контрастов* определяются следующим образом:

сначала вычисляются $k(k-1)/2$ *первичных оценок* контрастов

$$V_{rs} = MED\{X_{ir} - X_{js},\ 1 \le i \le n_r,\ 1 \le j \le n_s\}, \quad 1 \le r < s \le k,$$

$$V_{rr} = 0, \quad V_{sr} = -V_{rs}.$$

Затем вычисляются взвешенные суммы $W_r = \frac{1}{N}\sum_{s=1}^{k} n_s V_{rs},\ 1 \le r \le k.$

Уточнённые оценки контрастов определяются как $\hat{\Delta}_{rs} = W_r - W_s.$

Практическое задание 2* (выполните его дома, если захотите)

1) Напишите программу для вычисления уточнённых оценок контрастов
2) Перекодируйте коды групп из 1-го столбца таблицы Kruskal с помощью команды as.numeric в числа 1, 2, 3
3) Вычислите уточнённые оценки контрастов для групп из таблицы Kruskal

Типы критериев

$\mathbf{P}(T > t_0) = \alpha_0$

Точные. Для выборки заданного размера n известно распределение статистики критерия. На основе этого распределения вычисляется фактический уровень значимости α_0 (p-level).

[Например, критерий Стьюдента из темы 4.]

$\mathbf{P}(T > t_0) \to \alpha_0$

Асимптотические. При увеличении размера выборки распределение статистики критерия сходится к некоторому известному закону. Фактический уровень значимости вычисляется приближённо.

[Например, критерий Колмогорова из темы 3.]

$\mathbf{P}(T > x_{1-\alpha}) \le \alpha$

Консервативные. Известна только оценка сверху α на уровень значимости. Для заданного α наблюдаемое значение статистики критерия сравнивается с критическим $x_{1-\alpha}$.

[Например, критерий Данна.]

Классический однофакторный дисперсионный анализ (ANOVA)

Analysis of Variance

Предполагается, что наблюдения нормально распределены:

$$X_{ij} \sim \mathcal{N}(\mu_j, \sigma_j^2) \; (i = 1, \ldots, n_j, j = 1, \ldots, k),$$

причём дисперсии σ_j^2 одинаковы для всех j. Проверяется гипотеза

$$\mu_1 = \ldots = \mu_k.$$

Рассмотрим суммы

Распределение Фишера

$V_{tot} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - X_{..})^2$ — общая изменчивость (*total*),

$V_{int} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - X_{.j})^2$ — изменчивость внутри выборок (*interior*),

$V_{out} = \sum_{j=1}^{k} n_j (X_{.j} - X_{..})^2$ — изменчивость между выборками (*outer*).

Тогда $V_{tot} = V_{int} + V_{out}$. Статистика критерия $R = \dfrac{\frac{1}{k-1} V_{out}}{\frac{1}{N-k} V_{int}} \sim F_{k-1, N-k}.$

Инвариантна к сдвигу и масштабированию

Критика ANOVA

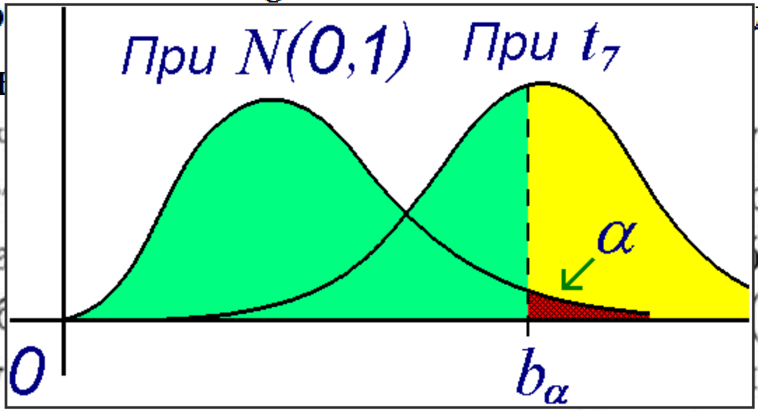
При проверке однородности нескольких независимых *нормальных* выборок с помощью ANOVA предполагается, что дисперсии наблюдений во всех выборках *одинаковы*. Для контроля этого обычно применяется **критерий Бартлетта**, статистикой которого служит отношение взвешенных среднего арифметического и среднего геометрического выборочных дисперсий:

$$B = \left(\frac{1}{N}\sum_{j=1}^{k} n_j S_j^2\right) \Big/ \sqrt[N]{\prod_{j=1}^{k} (S_j^2)^{n_j}}.$$

Оказалось, что э … даже

к небольшим от … от нормального.

Обозначим … уровня α статистики Бартлет … справедливости гипотезы совпа … 3 г. установил, что даже для … $(0, 1)$ на весьма близкое к нему … §7 темы 3) эта



вероятность меняется драматически. На рисунке изображен сдвиг плотности статистики B при такой замене. Например, для $k = 10$ получаем следующую картину: статистик, уверенный в том, что использует уровень значимости 5%, в действительности имеет дело с уровнем 49%!

Предполагалось, что отклонения от идеальных моделей можно игнорировать как несущественные; что статистические процедуры, оптимальные в строгой модели, останутся примерно таковыми и в приближенной модели. К сожалению, оказалось, что такие надежды зачастую не имеют под собой никакой почвы; даже безобидные отклонения часто имеют следствием эффекты гораздо более сильные, нежели это предвидело большинство статистиков.

Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. «Робастность в статистике»

Критерий Джонкхиера — Терпстры

Если исследователь надеется выявить значимое возрастание уровня интересующего его фактора от выборки к выборке, надо применять не критерий Краскела — Уоллиса, а более чувствительный *критерий Джонкхиера — Терпстры*.

Он используется для проверки гипотезы однородности H_0 против *альтернативы возрастания уровня фактора*

$$H_2\colon\ \mu_1 \leqslant \ldots \leqslant \mu_k,$$

где хотя бы одно из неравенств строгое.

1) Вычисляются $k(k-1)/2$ статистик Манна — Уитни U_{rs}, $1 \leqslant r < s \leqslant k$, где

$$U_{rs} = \sum_{i=1}^{n_r} \sum_{l=1}^{n_s} I_{\{X_{ir} < X_{ls}\}}. \qquad (1)$$

2) В качестве *статистики критерия Джонкхиера* берется

$$J = \sum_{\boxed{r<s}} U_{rs} = \sum_{r=1}^{k-1} \sum_{s=r+1}^{k} U_{rs}. \qquad (2)$$

Если верна H_0, то статистика $J^* = (J - \mathbf{M}J)/\sqrt{\mathbf{D}J}$ имеет асимптотическое распределение $\mathcal{N}(0,\, 1)$.

L-критерий (линейно растущие веса)

В однофакторной модели против альтернативы H_2 возрастания уровня фактора также применяется L-критерий, статистика L которого имеет вид

$$L = \frac{1}{\sqrt{N}} \sum_{j=1}^{k} \left(j - \frac{k+1}{2} \right) \left(R_{\cdot j} - \frac{N+1}{2} \right),$$

где $R_{\cdot j}$ — средний ранг наблюдений из j-й выборки, $N = n_1 + \ldots + n_k$ — общее число наблюдений в k выборках. Таким образом, L — сумма центрированных средних рангов с *линейно* растущими весами, что объясняет название критерия: *linear* (англ.) — линейный.

Известно, что при справедливости гипотезы H_0 однородности выборок $\mathbf{M}L = 0$ и

$$\mathbf{D}L = \frac{N+1}{12} \sum_{j=1}^{k} \frac{1}{n_j} \left(j - \frac{k+1}{2} \right)^2,$$

причём статистика $L^* = L / \sqrt{\mathbf{D}L}$ имеет асимптотическое распределение $\mathcal{N}(0, 1)$ при увеличении размеров всех выборок: $n_j \to \infty$, $n_j/N \to p_j$, $0 < p_j < 1$, $j = 1, \ldots, k$.

Пример данных для альтернативы возрастания уровня фактора

Пример 2. *Роль мотивации.* П. Хандел в 1969 г. исследовал влияние чистой мотивации (знания цели работы) на выполнение монотонных производственных операций (вытачивание металлических заготовок определенных форм и размеров). 18 мужчин были случайным образом разделены на 3 группы. Рабочие, попавшие в контрольную группу A, не имели информации о требуемой производительности, в группе B они получили лишь общее представление о том, что должны делать, наконец, в группе C рабочие имели точную информацию о задании и могли контролировать себя по графику, лежащему перед ними. В таблице приведены числа заготовок, обработанных каждым из рабочих за время эксперимента (в скобках указаны ранги R_{ij}).

Группа A	Группа B	Группа C
40 (5,5)	38 (2,5)	48 (18)
35 (1)	40 (5,5)	40 (5,5)
38 (2,5)	47 (17)	45 (15)
43 (10,5)	44 (13)	43 (10,5)
44 (13)	40 (5,5)	46 (16)
41 (8)	42 (9)	44 (13)
$S_1 = 40{,}5$	$S_2 = 52{,}5$	$S_3 = 78$
$R_{\cdot 1} = 6{,}75$	$R_{\cdot 2} = 8{,}75$	$R_{\cdot 3} = 13$

Данные содержатся в файле Motivat.txt. Проверьте гипотезу однородности групп:

1) Постройте диаграммы размахов для всех групп на одном рисунке

2) Установите пакет NSM3 и подключите его, примените функцию pJCK

Двухфакторная модель

__Данные__. В каждом из n блоков содержится по одному наблюдению x_{ij} на каждую из k обработок. Будем считать наблюдения реализацией случайных величин X_{ij} в модели

Блоки	Обработки			
	1	2	...	k
1	x_{11}	x_{12}	...	x_{1k}
2	x_{21}	x_{22}	...	x_{2k}
⋮	⋮	⋮	⋮	⋮
n	x_{n1}	x_{n2}	...	x_{nk}

$$\boxed{X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij},} \quad i = 1, \ldots, n; \quad j = 1, \ldots, k.$$

Здесь μ — неизвестное общее среднее, α_i — эффект блока i (неизвестный мешающий параметр), β_j — эффект обработки j (интересующий нас параметр), ε_{ij} — случайная ошибка.

Пусть справедливы те же самые, что и в однофакторной модели, допущения Д1 и Д2.

Для проверки гипотезы

$$H_0\colon\ \beta_1 = \ldots = \beta_k$$

используются критерии Фридмана или (для альтернативы возрастания β_j) Пейджа.

Критерий Фридмана

Блоки	Обработки			
	1	2	$\dots$	k
1	Q_{11}	Q_{12}	$\dots$	Q_{1k}
2	Q_{21}	Q_{22}	$\dots$	Q_{2k}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	Q_{n1}	Q_{n2}	$\dots$	Q_{nk}

Для проверки гипотезы

$$H_0 \colon \beta_1 = \ldots = \beta_k$$

против альтернативы

$$H_1 \colon \text{не все } \beta_j \text{ равны между собой}$$

применяется *критерий Фридмана.*

1) Отдельно для каждого i-го блока (строки таблицы) ранжируем k наблюдений внутри него от меньшего к большему. Обозначим через Q_{ij} ранг X_{ij} в совместной ранжировке $X_{i1}, \ldots, X_{ik}$.

2) Положим для $j = 1, \ldots, k$

$$T_j = \sum_{i=1}^{n} Q_{ij}, \quad Q_{\cdot j} = T_j/n, \quad Q_{\cdot\cdot} = \frac{1}{nk} \cdot n \ldots$$

Здесь $Q_{\cdot j}$ — это средний ранг по всем n блокам наблюдений, относящихся к j-й обработке (столбцу таблицы), $Q_{\cdot\cdot}$ — средний ранг по всей таблице.

3) В качестве *статистики критерия Фридмана* возьмем

$$F = \frac{12n}{k(k+1)} \sum_{j=1}^{k} (Q_{\cdot j} - Q_{\cdot\cdot})^2 \qquad \longrightarrow \chi^2_{k-1}$$

Алгоритм
критерия
реализован
функцией
friedman.test

Данные содержатся в файле Animals.txt

Интеллект животных

Д. Хэбб и К. Уильямс разработали тест эстакадного лабиринта для сравнительной оценки “сообразительности” животных. Он состоит из 12 заданий. В таблице даны средние числа ошибок при выполнении этих заданий крысами, кроликами и кошками (в скобках указаны ранги Q_{ij} внутри каждой строки).

Номер задания	$T_1 = 26$, Крысы	$T_2 = 29$, Кролики	$T_3 = 17$ Кошки
1	1,5 (2)	1,7 (3)	0,3 (1)
2	1,1 (2)	1,5 (3)	1,0 (1)
3	1,8 (1)	8,1 (3)	3,6 (2)
4	1,9 (3)	1,3 (2)	0,0 (1)
5	4,3 (3)	4,0 (2)	0,6 (1)
6	2,0 (1)	4,6 (2)	5,5 (3)
7	8,4 (3)	4,0 (2)	1,0 (1)
8	6,6 (3)	5,1 (2)	3,1 (1)
9	2,4 (2)	2,5 (3)	0,1 (1)
10	6,5 (2)	6,9 (3)	1,6 (1)
11	2,6 (2)	2,5 (1)	4,3 (3)
12	6,5 (2)	6,8 (3)	1,0 (1)

В языке R алгоритм критерия Фридмана реализован функцией friedman.test [используйте as.matrix]

Есть ли различие?

Классический двухфакторный дисперсионный анализ

Двухфакторный дисперсионный анализ. Пусть ошибки ε_{ij} распределены по закону $\mathcal{N}(0, \sigma^2)$. В этом случае оптимальным критерием для проверки H_0 является *F-критерий двухфакторного дисперсионного анализа*, основанный на статистике

$$\left[\frac{n}{k-1}\sum_{j=1}^{k}(X_{\cdot j} - X_{\cdot\cdot})^2\right] \Big/ \left[\frac{1}{(n-1)(k-1)}\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij} - X_{i\cdot} - X_{\cdot j} + X_{\cdot\cdot})^2\right],$$

$$\text{где}\quad X_{i\cdot} = \frac{1}{k}\sum_{j=1}^{k}X_{ij},\quad X_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n}X_{ij},\quad X_{\cdot\cdot} = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}X_{ij},$$

имеющей при H_0 распределение Фишера $F_{k-1,(n-1)(k-1)}$.

Можно доказать, что справедливо следующее тождество:

$$\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij} - X_{\cdot\cdot})^2 = k\sum_{i=1}^{n}(X_{i\cdot} - X_{\cdot\cdot})^2 + n\sum_{j=1}^{k}(X_{\cdot j} - X_{\cdot\cdot})^2 +$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij} - X_{i\cdot} - X_{\cdot j} + X_{\cdot\cdot})^2.$$

Оно показывает, что *общая изменчивость* распадается на части, обусловленные влиянием *эффектов блоков*, *эффектов обработок*, и часть, связанную с *изменчивостью самих данных*.

Критерий Пейджа

Иногда обработки упорядочены естественным образом, например, по интенсивности стимулов, сложности заданий и т. п. Критерий Пейджа, в отличие от критерия Фридмана, учитывает информацию, содержащуюся в предполагаемой упорядоченности.

Для проверки гипотезы однородности H_0

$$H_0 \colon \beta_1 = \ldots = \beta_k$$

против *альтернативы возрастания эффектов обработок*

$$H_2 \colon \beta_1 \leqslant \ldots \leqslant \beta_k,$$

где хотя бы одно из неравенств строгое, вычисляется *статистика критерия Пейджа*

$$L = \sum_{j=1}^{k} jT_j = T_1 + 2T_2 + \ldots + kT_k,$$

Взвешенная сумма T_j

где $T_j = \sum_{i=1}^{n} Q_{ij}$ — сумма рангов по всем n блокам (строкам), относящихся к j-й обработке (столбцу).

Если верна H_0, то распределение стандартизованной статистики $L^* = (L - \mathbf{M}L)/\sqrt{\mathbf{D}L}$ сходится к $\mathcal{N}(0,\,1)$ при $n \to \infty$.

Зависимость прочности волокон хлопка от количества удобрения

Пример 3. Прочность волокон хлопка. В опыте, описанном в книге Cochran W. G., Cox G. M. “*Experimental Designs*”, изучалось влияние количества калийного удобрения, вносимого в почву, на разрывную прочность волокон хлопка. При $n = 3$ блоках использовалось $k = 5$ уровней удобрений. С каждой делянки отбирался один образец хлопка, на котором производилось 4 измерения показателя прочности по Прессли. В таблице приведены средние по этим четырем замерам, а в круглых скобках — ранги Q_{ij} внутриблочного ранжирования.

Блоки	Калийное удобрение (кг/га)				
	163	122	82	61	41
1	7,46 (2)	7,17 (1)	7,76 (4)	8,14 (5)	7,62 (3)
2	7,68 (2)	7,57 (1)	7,73 (3)	8,15 (5)	8,00 (4)
3	7,21 (1)	7,80 (3)	7,74 (2)	7,87 (4)	7,93 (5)
	$T_1 = 5$	$T_2 = 5$	$T_3 = 9$	$T_4 = 14$	$T_5 = 12$

Данные содержатся в файле Cotton.txt. Проверьте критерием Пейджа гипотезу об отсутствии влияния количества удобрения на прочность нити против альтернативы убывания прочности с ростом количества удобрения. Для проверки используйте функцию pPage из пакета NSM3 [используйте команду as.matrix]

Правильный выбор критерия

$$\mathbf{P}\left(\bigcup A_{ij}\right) \le \sum \mathbf{P}(A_{ij})$$

Зачем нужны многовыборочные критерии, когда есть двухвыборочные?

Казалось бы, можно применить такую стратегию:

на диаграмме размахов увидеть две наиболее отличающиеся выборки и проверить их однородность с помощью какого-нибудь двухвыборочного критерия.

1) Однако при этом отбрасывается информация о разбросе наблюдений, содержащаяся в других выборках, и поэтому происходит потеря в чувствительности.

2) Кроме того, меняется уровень значимости, так как распределение максимального различия средних отличается от распределения различия средних выборок с заранее фиксированными номерами.

Пусть $X_1, \ldots, X_n$ — случайные числа. Тогда

$\mathbf{M}|X_n - X_1| = 1/3$,

$\mathbf{M}|X_{(n)} - X_{(1)}| \to 1$ при $n \to \infty$.

Если выборок много (7), то и различных пар тоже много ($7 * 6 / 2 = 21$). Вероятность, что хотя бы в одной из пар случайно будет наблюдаться неоднородность, можно только оценить суммой соответствующих вероятностей, вычисленных для одной пары (консервативность).

Главное в теме

- Перед применением критериев полезно построить диаграммы размахов для выявления «выбросов» или существенного различия в межквартильных размахах выборок
- Важно научиться различать случаи, когда следует применять однофакторную модель (обычно она используется для независимых групп), а когда — двухфакторную (обычно она применяется для повторных наблюдений), а также различать упорядоченные и неупорядоченные альтернативы
- После выявления общей неоднородности, следует выполнить множественные сравнения выборок (обработок) и вычислить контрасты между значимо различающимися выборками
- Дисперсионный анализ (ANOVA) рекомендуется использовать только как *вспомогательный инструмент*. На практике его обычно вполне могут заменить непараметрические критерии