

Math statistics

2024-01-23

Содержание

| | |
|--------|---|
| Case A | 1 |
| Case B | 2 |
| Case C | 4 |
| Case D | 5 |
| Case F | 8 |

Case A

Data (первые 6 значений)

| Student | Period | Grade |
|---------|------------------|-------|
| 1 | grade_1st_period | 56 |
| 1 | grade_2nd_period | 71 |
| 2 | grade_1st_period | 61 |
| 2 | grade_2nd_period | 77 |
| 3 | grade_1st_period | 68 |
| 3 | grade_2nd_period | 78 |

(a) Гипотезы:

- H_0 - средняя оценка студентов за первый периоде **не отличается** от средней оценки за второй период

- H_A - средняя оценка студентов за первый периоде **меньше** средней оценки за второй период

(b) Зависимой переменной является **оценка студента**, так как задачей программы была именно изменение оценок студенто в клучшему

(c) В данном случае оценка студента - данные измеренные в относительной шкале оценок, то есть уровень измерения - относительный (**ratio scale level**)

(d) Дизайн эксперимента заключался в измерении показателя до “воздействия” и после него (**pretest-posttest design**). В данном случае были получены данные об оценках студентов в первом периоде аттестации (то есть **до** начала проекта по улучшению оценок) и во втором периоде аттестации, то есть **после** запуска проекта.

(e) Лучшим решением будет проверить одностороннюю гипотезу парным Т-тестом, так как 1. Данные **ДО** и после **ПОСЛЕ** являются зависимыми, потому что измерены у одних и тех же студентов. 2. Хотим проверить гипотезу о неравенстве средних оценок студентов

(f) Т-тест используется для определения того, есть ли статистически значимая разница между средними показателями двух групп. В данном случае две группы - это оценки студентов за первый период и за второй период. В данном случае у нас односторонняя гипотеза, поэтому рассматривается *больше ли* полученное значение статистики, чем некоторое критическое значение, при котором мы можем отвергнуть нулевую гипотезу ($p\text{-value} \leq 0.05$).

Paired t-test

```
ttest = t.test(student_grades$Grade[student_grades$Period == 'grade_1st_period'],
               student_grades$Grade[student_grades$Period == 'grade_2nd_period'],
               paired=TRUE, alternative='less')
```

| t.statistic | p.value |
|-------------|-----------|
| -2.185723 | 0.0301609 |

(g) ДА. Значение статистики и соответствующий *p-уровень значимости* указывает, что мы не можем отклонить альтернативную гипотезу (за пороговое значение берем $p\text{-value} < 0.05$)

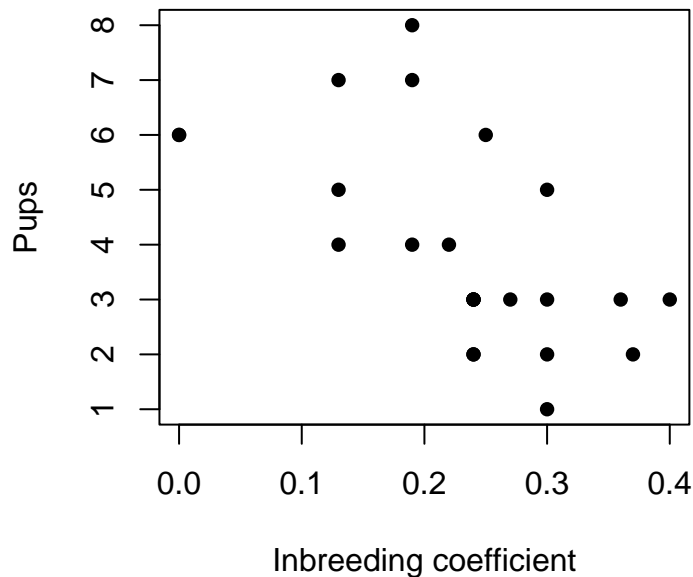
(h) Это означает, что проект по улучшению оценок студентов прошел успешно. Средняя оценка студентов изменилась в большую сторону за второй период аттестации - это подтверждено односторонним статистическим тестом.

Case B

Data (первые 6 значений)

| inbreeding.coefficient | pups |
|------------------------|------|
| 0.00 | 6 |
| 0.00 | 6 |
| 0.13 | 7 |
| 0.13 | 5 |
| 0.13 | 4 |
| 0.19 | 8 |

(a) Изобразить графически:



Виден такой тренд, что при увеличении коэффициента инбридинга уменьшается количество детёнышей.

(b) Вычислить корреляцию

```
cor(wolves$inbreeding.coefficient, wolves$pups)
```

```
## [1] -0.6077184
```

Корреляция между коэффициентом инбридинга и количеством детёнышей равна ~ -0.61 , что указывает на умеренную **отрицательную корреляцию** между показателями. Это говорит о том, что в основном при увеличении коэффициента инбридинга, количество детёнышей снижается.

(c) Гипотезы для тестирования

- H_0 : корреляция = 0
- H_A : корреляция $\neq 0$

Используем **T-тест** Стьюдента

```
crtest = cor.test(wolves$inbreeding.coefficient, wolves$pups, method = 'pearson')
```

| t.statistic | p.value |
|-------------|----------|
| -3.5893 | 0.001633 |

Значение статистики = -3.59, *p-уровень значимости* = 0.001633.

Таким образом, мы не можем отвергнуть альтернативную гипотезу, что корреляция между коэффициентом инбридинга и количеством детёнышей не равна нулю (она равна -0.61).

Case C

Data (первые 6 значений)

| Patient.number | first.meeting | One.month.later |
|----------------|---------------|-----------------|
| 1 | Non-compliant | Compliant |
| 2 | Non-compliant | Compliant |
| 3 | Compliant | Compliant |
| 4 | Non-compliant | Non-compliant |
| 5 | Non-compliant | Compliant |
| 6 | Non-compliant | Non-compliant |

(a) Гипотезы:

- H_0 - соотношения количества пациентов, соблюдающих и несоблюдающих режим приема лекарств, в первый день и через месяц после визитов социального работника **равны**

- H_A - соотношения количества пациентов, соблюдающих и несоблюдающих режим приема лекарств, в первый день и через месяц после визитов социального работника **не равны**

(b) Зависимой переменной является **соблюдение режима приема лекарств пациента** (*соблюдает или нет*), так как именно это стремится изменить программа - чтобы те, кто не соблюдал режим, стали его соблюдать.

(c) Соблюдение режима приема лекарств пациента - номинальные данные (**nominal level**). Есть только две категории - соблюдает режим и не соблюдает.

(d) Измерение до "воздействия" (регулярных визитов соц.работника в течение месяца) и после него (**pretest-posttest design**). В данном случае каждому пациенту ставится в соответствие категория соблюдения режима принятия лекарств в самом начале программы и **через месяц**

(e) Данные можно представить в виде таблицы сопряженности:

```
##           After month
## First day  Complaint Non-complaint
## Complaint      2         1
## Non-complaint  6         3
```

(f) Так как данные *в первый день* и *через месяц* зависимые, то применить обычный тест Хи-квадрат нельзя, поэтому для такого случая зависимых данных воспользуемся **McNemar test** (*с поправкой Yates*)

```
mctest = mcnemar.test(data)
```

| McNemar_chi_squared | p.value |
|---------------------|---------|
| 2.285714 | 0.13057 |

(g) HET

(h) Результаты теста не позволяют отвергнуть нулевую гипотезу о том, что доля соблюдающих режим приема лекарств пациентов изменилась. Часть пациентов, действительно стали соблюдать режим. Однако есть и те, кто как соблюдал/не соблюдал, так и продолжил соблюдать/не соблюдать режим. Также есть и пациент, который наоборот перестал соблюдать режим. Таким образом, можно поставить под сомнение эффективность данной медицинской программы.

Case D

Data (первые 6 значений)

| condition | stretch | treatment |
|--------------|---------|-----------|
| BothWrithing | 36.7 | bw |
| BothWrithing | 81.1 | bw |
| BothWrithing | 66.7 | bw |
| BothWrithing | 66.7 | bw |
| BothWrithing | 44.4 | bw |
| BothWrithing | 54.4 | bw |

(a) Полная модель:

$$y = \beta_1 I_{ow} + \beta_2 I_{bw} + \beta_0, \text{ где}$$

- y - переменная отклика, процент времени, в течение которого мышь демонстрировала *поведение растягивания*, то есть испытывала дискомфорт(**stretch**),
- I_{ow} - индикатор (*dummy variable*), принимающий значение 1, если мышь была помещена к мыши-компаньону, которой не делали инъекций, и 0 в противном случае,
- I_{bw} - индикатор (*dummy variable*), принимающий значение 1, если мышь была помещена к мыши-компаньону, которой также была сделана инъекция, и 0 в противном случае,
- β_0 - базовое значение, средний процент времени дискомфорта для **изолированных** мышей,
- β_1 - коэффициент для I_{ow} , разница в среднем проценте времени дискомфорта между мышами, помещенными с **неинъектированным компаньоном** и изолированными мышами,
- β_2 - коэффициент для I_{bw} , разница в среднем проценте времени дискомфорта между мышами, помещенными с **инъектированным компаньоном** и изолированными мышами,

```
full_model = lm(stretch ~ condition, data = mice)
summary(full_model)
```

```
##
## Call:
## lm(formula = stretch ~ condition, data = mice)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -37.194 -9.339  2.806  9.506 29.031
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    37.194     4.220   8.814 8.06e-11 ***
## conditionOneWrithing  -1.825     6.411  -0.285 0.77741
## conditionBothWrithing 20.856     6.560   3.179 0.00289 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.4 on 39 degrees of freedom
## Multiple R-squared:  0.255, Adjusted R-squared:  0.2168
## F-statistic: 6.674 on 2 and 39 DF, p-value: 0.003216
```

(b) Нулевая модель:

В нулевой модели мы предполагаем, что **разницы** в проценте времени дискомфорта между тремя группами мышей (изолированных, помещенных с неинъецированным компаньоном и помещенных с инъецированным компаньоном) **нету**:

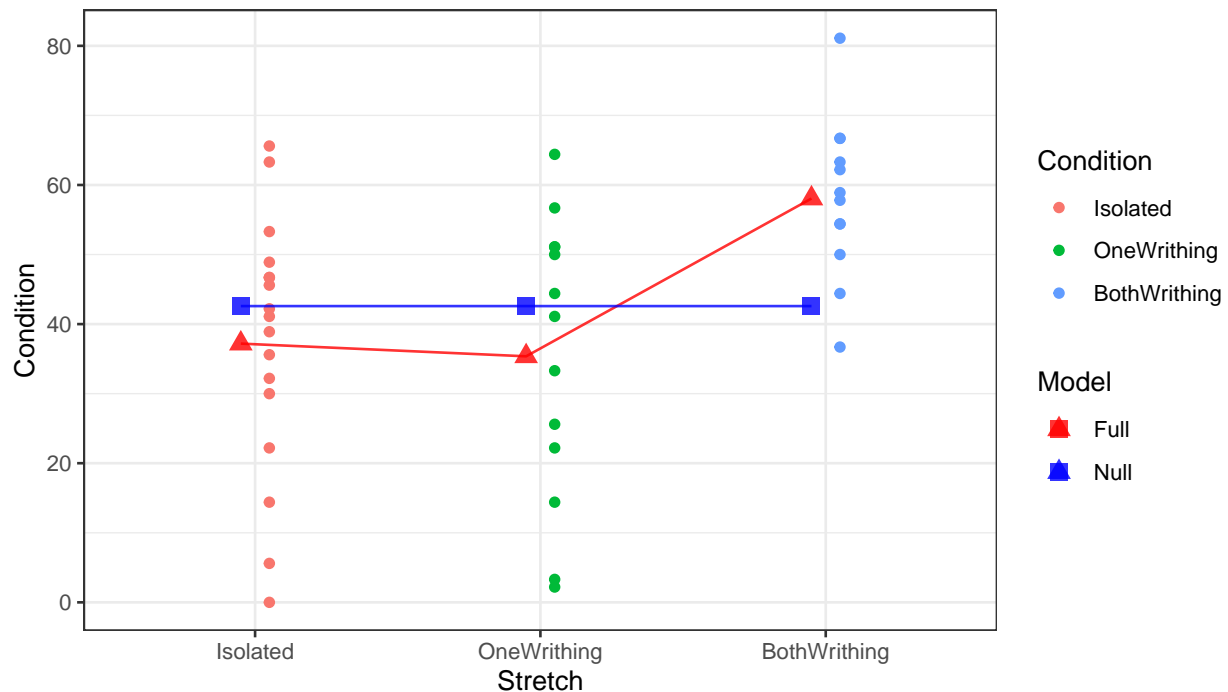
$y = \beta_0$, где

- β_0 - базовое значение, средний процент времени дискомфорта

```
null_model = lm(stretch ~ 1, data = mice)
summary(null_model)
```

```
##
## Call:
## lm(formula = stretch ~ 1, data = mice)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -42.588 -10.113   3.562  13.537  38.512
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  42.588     3.034  14.04  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 19.66 on 41 degrees of freedom
```

(с) График



В полной модели наблюдается такой паттерн, что **мышь “эмпатична”**, то есть процент времени дискомфорта увеличивается, если мышь-компаньон тоже подверглась инъекции.

(с) Статистически протестируем полную модель

То есть сравним ее с нулевой.

Для этого хорошо подойдет F-статистика (ANOVA):

```
F_test = anova(full_model, null_model)
```

| F_statistic | p.value |
|-------------|-----------|
| 6.673586 | 0.0032162 |

Таким образом, результаты теста позволяют отвергнуть нулевую гипотезу, что **разницы** в проценте времени дискомфорта между тремя группами мышей (изолированных, помещенных с неинъектированным компаньоном и помещенных с инъектированным компаньоном) **нету**. Мы принимаем альтернативную гипотезу, что **разница есть**.

Проведем пост-хок тест

Пост-хок тест проводим, чтобы узнать, какая группа мышей статистически значимо отличается от остальных - то есть статистически значимо отличается средний процент времени дискомфорта. Для этого воспользуемся тестом Тьюки:

```
Tu_test = TukeyHSD(aov(stretch ~ condition, data=mice))
```

| | diff | lwr | upr | p adj |
|--------------------------|-----------|------------|----------|-----------|
| OneWrithing-Isolated | -1.824887 | -17.443482 | 13.79371 | 0.9563649 |
| BothWrithing-Isolated | 20.855882 | 4.872769 | 36.83900 | 0.0079593 |
| BothWrithing-OneWrithing | 22.680769 | 5.710639 | 39.65090 | 0.0064731 |

Из результатов теста видно, что группа, где мышь была помещена к мыши-компаньону, которой также была сделана инъекция, статистически значимо отличается от остальных. *Скорректированный p-уровень значимости* в попарных сравнениях **Изолированная vs. С инъецированным компаньоном** и **С НЕинъецированным компаньоном vs. С инъецированным компаньоном** оказался менее 0.05. Это означает, что среднее время дискомфорта в сравниваемых группах статистически значимо отличается.

Процент времени *дискомфорта* в случае, когда обеим мышам была сделана инъекция, больше примерно на ~21% чем в случаях, когда мышь была изолирована или когда компаньону не была сделана инъекция.

Таким образом, действительно тестируемые мыши были более эмпатичны и время их дискомфорта было больше, если вторая мышь так же “страдала”.

Case F

Data

| week | anxiety |
|------|---------|
| 0 | 34 |
| 1 | 31 |
| 2 | 28 |
| 3 | 33 |
| 4 | 35 |
| 5 | 29 |
| 6 | 28 |

Week 0 \equiv Baseline

(a) Гипотеза

Гипотеза состоит в том, что значение тревожности девушки уменьшалось с течением времени. В случае верности гипотезы можно будет говорить об эффективности терапии.

(b) Зависимой переменной является значение тревожности, так как именно это значение в процессе терапии девушка стремится уменьшить.

(c) В данном случае значение тревожности измерялось в относительной шкале, то есть уровень измерения - относительный (**ratio scale level**)

(d) В исследовании исследовалось изменение показателя тревожности со временем - в течение нескольких недель. Такие исследования называют **лонгитюдными (longitudinal research design)**.

(e) Чтобы изучить изменение тревожности со временем, подберем модель линейной регрессии, где зависимая переменная *тревожность*, а независимая - время (номер недели)

$$anxiety = \beta_1 * week + \beta_0$$

Сравним эту модель с нулевой моделью, где уровень тревожности будет одинаковым (средним) на протяжении всего времени

$$anxiety = \beta_0$$

Таким образом, мы хотим проверить есть ли какой-то тренд в значении тревожности со временем.

То есть проверяем значимо ли статистически коэффициент (β_1) при номере недели (*week*) отличен от нуля. Для этого используем **F-статистику**

(f) F-статистика используется для сравнения степени соответствия между полной и нулевой моделями в регрессионном анализе. F-статистика рассчитывается, как отношение средней квадратичной ошибки нулевой модели на среднюю квадратичную ошибку полной модели. Полученное значение сопоставляется с F-распределением со степенями свободы, равными разнице количества коэффициентов моделей. Если полученная статистика больше критического значения, мы отвергаем нулевую гипотезу о том, что полная модель не обеспечивает лучшего соответствия, чем нулевая модель. То есть мы делаем вывод, что полная модель лучше соответствует данным, чем нулевая модель.

Модели:

```
full = lm(anxiety ~ week, data=anxiety)
null = lm(anxiety ~ 1, data=anxiety)
```

F-тест

```
F_test = anova(full,null)
```

| F.statistic | p.value |
|-------------|-----------|
| 0.9382819 | 0.3772082 |

(g) НЕТ. Полученные результаты теста не позволяют отвергнуть нулевую гипотезу о том, что средний уровень тревожности в течение нескольких недель был одинаков. Действительно, из данных было видно, что в течение шести недель тревожность девушки и поднималась и понижалась примерно в равной степени.

(h) Таким образом, можно говорить о неэффективности терапии, которую получает девушка. Уровень тревожности в среднем остается не изменяется с течением времени, что было показано статистически.