

Для анализа критериев качественных признаков были выбраны (согласно 5му варианту): weakness.1 и headache.1 из набора данных про алкогольно-абсисентный синдром.

Критерий независимости хи-квадрат

Будем рассматривать в качестве основной гипотезы гипотезу о том, что признаки независимы.

Pearson's Chi-squared test

```
data: tbl  
X-squared = 4.4266, df = 4, p-value = 0.3513
```

По результатам применения критерия независимости хи-квадрат видно, что $p\text{-value} > 0.05$ (которое мы принимаем по-умолчанию), поэтому нет оснований отвергать гипотезу о том, что признаки независимы.

Точный критерий Фишера

Для той же основной гипотезы рассмотрим критерий Фишера.

Fisher's Exact Test for Count Data

```
data: tbl  
p-value = 0.3571  
alternative hypothesis: two.sided
```

Он также показал, что нет оснований отвергать гипотезу о том, что признаки независимы.

Коэффициенты неопределённости

При помощи коэффициентов неопределённости продемонстрируем слабую связь между признаками.

Воспользовавшись вспомогательной функцией (см. скрипт) вычислим соответствующие значения энтропии для признаков в отдельности и в совокупности.

```
> entropy.a  
[1] 1.341632  
> entropy.b  
[1] 1.196509  
> entropy.ab  
[1] 2.426871
```

По результатам вычислений можно судить о том, какое количество информации требует этот признак для его уточнения (в эксперименте).

Используя полученные значения, вычисляем коэффициенты неопределённости.

J	8.76787011981674
J.ab	9.29959001567376
J.ba	8.29366566192157

Из полученных коэффициентов неопределённости можно говорить о том, что признаки слабо связаны, поскольку, например, о признаке А по признаку В можно узнать лишь 9.3% информации. Предыдущие критерии подтверждают догадку.

Критерий Мак-Немара

Для проверки критериев независимости качественных признаков для зависимых выборок был выбран (соответственно варианту) признак: chest.pain.

Для того, чтобы применить критерий Мак-Немара, необходимо построить соотв. матрицу 2x2 (алгоритм в скрипте). В результате получаем следующую таблицу сопряжённости признаков:

	[,1]	[,2]
[1,]	9	0
[2,]	13	1

Рассмотрим основную гипотезу о том, что изменения в признаке случайны.

```
McNemar's Chi-squared test with continuity correction
```

```
data: ctable
```

```
McNemar's chi-squared = 11.077, df = 1, p-value = 0.0008741
```

По результатам применения метода Мак-Немара видно, что следует отклонить основную гипотезу ($p\text{-value} < 0.05$). Это можно интерпретировать, как повод отказаться от гипотезы о том, что боль в груди равновероятно может увеличиться или уменьшиться в период прохождения лечения.

Критерий Кохрена

Поскольку критерий Кохрена применяется для признаков, имеющих лишь две градации. Данные были модифицированы обобщением всех степеней тяжести боли в груди.

```
> cohranStat = s * (s-1) * multiplier; cohranStat  
[1] 44.8  
> cohranPValue=1-pchisq(cohranStat, s-1); cohranPValue  
[1] 1.869837e-10
```

В результате вычислений (алгоритм которых описан в скрипте) видим, что мы имеем все основания отклонить гипотезу о том, что изменения в период прохождения лечения (3 дня) случайны.