# Домашнее задание по разделу математика

**Задача №1**

*В игре «Что? Где? Когда?» в каждом раунде волчок останавливается в секторе номер n, где n равновероятно принимает одно из значений 0, 1,..., 13. При этом играет первый из секторов по часовой стрелке, который ранее не играл. Найдите вероятность того, что после шести раундов сыграют (в любом порядке) секторы 1, 2,..., 6.*

Решение

Для начала введем случайную величину – выпадение определенного сектора, где принимает значения от 1 до 6 (соответствующие выпавшим секторам раунды). Затем попробуем рассчитать следующую вероятность:

.

Рассмотрим разложенную вероятность как произведение вероятностей с конца, то есть с множителя . , так как всего имеется 14 секторов и мы попадаем в один из них. Затем , так как при попадании во втором раунде на первый сектор мы автоматически попадаем на второй сектор, при попадании на второй сектор во втором раунде он задействуется (то есть так как событие несовместные, то складываем с ). По такому же принципу получим, что

Рассмотрим другую расстановку выпавших секторов на разных раундах:

.

Снова рассмотрим последний множитель Он также будет равен , так как при отсутствии каких либо условии мы попадаем в первом раунде на пятый сектор, вероятность чего можно рассматривать как отношение благоприятных исходов (выпал первый сектор) ко всем исходам. Далее предпоследний множитель как и в первом рассмотренном случае равен , так как при условии известного задействованного сектора и его невыпадении его можно причислить к одному из благоприятных исходов, перейдя на следующий сектор.

Таким образом, для всех возможных перестановок выпадающих в каждом раунде секторов можем сказать, что их вероятность будет одинакова, а именно Количество всех перестановок рассчитывается по формуле . **Итого получаем, что конечный ответ равняется: .**

**Задача №2**

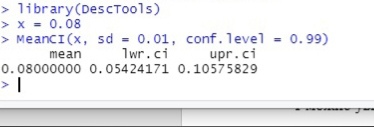
*Аналитик рынка ценных бумаг оценивает среднюю доходность определенного вида акций. Случайная выборка из 16 дней показала, что средняя доходность по акциям данного типа составляет 8% с выборочным средним квадратическим отклонением в 4%. Предполагая, что доходность акции подчиняется нормальному закону распределения, определите 99% -ый доверительный интервал для средней доходности интересующего аналитика вида акций.*

Решение

Итак имеем из условия, что (так как дан 99%-процентный доверительный интервал, который обозначается как ). Задание можно трактовать по-другому: определить доверительный для среднего значения выборки при известном с.к.о. Это означает, что в решении необходимо использовать такое значение как (в случае неизвестного с.к.о. мы бы использовали ). Доверительный интервал для среднего в таком случае имеет вид:

С использованием специальной таблицы или с применением функции *qnorm()* получим, что . Осталось только подставить все данные ранее значения в написанный выше общий вид доверительного интервала:

С использованием языка R проверим правильность вычислений (код скрипта прикреплен в той же папке, где расположен файл с домашним заданием, под названием *ci.R*). На рисунке 1 можно увидеть конечный ответ, соответствующий ранее приведенным вычислениям.



*Рисунок 1 – Результат работы программы ci.R*

**Задача №3**

*Мужчины и женщины по-разному оценивают положительные человеческие качества. Предложили мужчинам и женщинам на основе десятибалльной шкалы (10 баллов – это максимум) оценить важность следующих пяти качеств в представителях противоположного пола:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Качества* | *Ум* | *Доброта* | *Красота* | *Юмор* | *Работоспособность* |
| *Мужчины* | *7* | *8* | *8* | *5* | *7* |
| *Женщины* | *10* | *5* | *3* | *8* | *10* |

*Найдите тесноту связи между этими данными, рассматривая данные, как выборочные наблюдения случайных величин. Сделайте вывод о том, насколько близки или далеки мужчины и женщины в оценках качеств партнеров.*

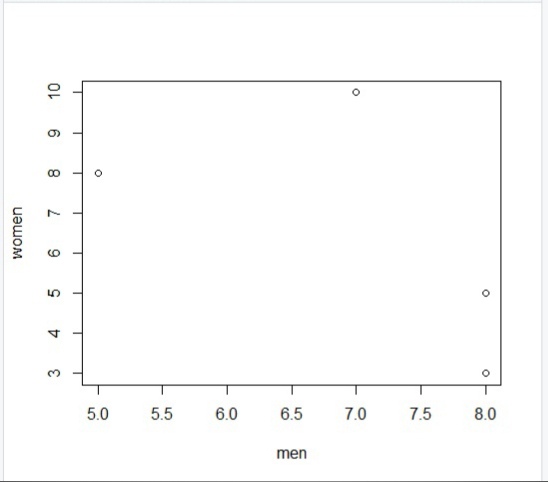
Решение

Так как в условии сказано, что данные нужно рассматривать как выборочные наблюдения случайных величин, то таблицу можно представить в следующем виде:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Качества* |  |  |  |  |  |
| *Мужчины* | *7* | *8* | *8* | *5* | *7* |
| *Женщины* | *10* | *5* | *3* | *8* | *10* |

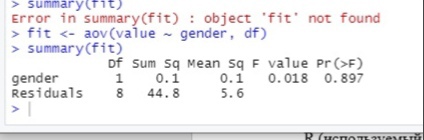
То есть все происходящее можно рассмотреть как задачу о сравнении двух групп, где независимая переменная представлена в качестве номинативной переменной (мужчина или женщина), а зависимая переменная представлена в качестве количественной переменной. Условие задачи, на мой взгляд, поставлено немного некорректно, так как непонятно, что означают численные значения, представленные в таблице: то ли это минимальное значение показателя при опросах, то ли максимальное, а то и среднее. Поэтому я идеализирую результаты, поставив условие, что у нас выходят довольно красивые целые значения оцениваемых качеств.

Для расчета тесноты связи можно было бы использовать классический линейный коэффициент корреляции Пирсона, однако у нас, во-первых, довольно мало значения для использования данного метода, которые к тому же, во-вторых, не распределены линейным образом (на рисунке 2 с использованием языка R представлен график, очевидно демонстрирующий отсутствие линейной зависимости между двумя рассматриваемыми группами).



*Рисунок 2 – Отсутствие нормальности распределения*

А значит для сравнения двух групп лучше всего подойдет дисперсионный анализ либо t-тест. В данном случае t-тест также вряд ли будет уместно использовать ввиду малого количества данных, поэтому лучше всего воспользоваться его непараметрическим аналогом U критерием Манна-Уитни, который состоит в ранжировании значений каждой из групп и сравнении значений, расположенных посередине. Для начала воспользуемся дисперсионным анализом для нахождения тесноты взаимосвязи в рамках рассматриваемых значений. На рисунке 3 представлено проведение данного анализа с использованием языка R (используемый код прикреплен в той же папке, что и данный файл с домашним заданием, под названием *tests.R*).

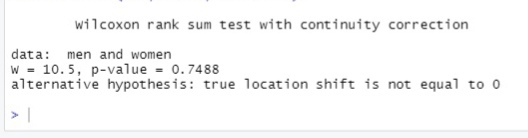


*Рисунок 3 – Проведение однофакторного дисперсионного анализа в R*

Как видим, *F-value* у нас имеет достаточно малое значение, что свидетельствует о том, что теснота взаимосвязи в основном обусловлена внутригрупповыми взаимосвязями, а не межгрупповыми, так как *F-value* рассчитывается по следующей формуле:

Где – межгрупповая сумма квадратов, – внутригрупповая сумма квадратов, а и – соответствующие степени свободы. **Таким образом, можно сделать вывод о том, что две группы мало связаны; в основном тесно связаны качества, оценки для которых дают отдельно мужчины и отдельно женщины.** Более того обратив внимание на *p-value* мы можем увидеть, что оно достаточно большое (0,897 > 0,05), исходя из чего можно сделать вывод о том, что мы не можем опровергнуть нулевую гипотезу о том, что средние значения для двух выборок (мужчины и женщины) сильно разнятся, **а значит мужчины и женщины достаточно близки в оценках качеств партнеров.**

Также второе утверждение можно подкрепить использованием непараметрического U критерия Манна-Уитни. Опять-таки воспользуемся языком программирования R (скрипт прикреплен в том же файле *tests.R).* Результат данного теста представлен на рисунке 4.



*Рисунок 3 – Применение непараметрического U критерия Манна-Уитни*

Наше значение *p-value >* 0,05, следовательно, мы не можем опровергнуть гипотезу о неравенстве средних показателей для двух выборок.