**ДЗ-6. Применение статистических критериев. Проверка гипотез**.

*Часть 1*.

Изучено.

*Часть 2.*

Условия для применения теста Стьюдента:

Выборки должны происходить из нормально распределенных совокупностей. На размер выборки ограничений не накладывается.

Изображение выглядит как текст

Автоматически созданное описаниеДля c1 проверим выборку на “нормальность”

Изображение выглядит как текст, экран, табло

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описание Данные c1 можем вполне считать нормально распределенными.

Проверим гипотезы для 2-ух предполагаемых средних(6800 ближе к правде, чем 5000).

Изображение выглядит как текст

Автоматически созданное описание

Изображение выглядит как текст

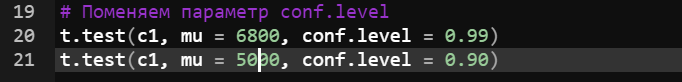
Автоматически созданное описаниеПо данным тестам, мы принимаем гипотезу с mu = 6800, и отклоняем с mu = 5000.

По данным тестам, мы принимаем гипотезу с mu > 6800, и отклоняем с mu < 5000.

Изображение выглядит как текст

Автоматически созданное описание В данном случае мы принимаем обе гипотезы: mu > 6800, mu > 5000(p-value = 1 это серьезный аргумент). При этом заметим, что mu > 6800 + mu < 6800 ~= 1, что согласуется с теорией вероятности.

Изображение выглядит как текст

Автоматически созданное описаниеПопробуем поменять доверительный интервал(параметр conf.level и посмотрим, что изменится).

Заметим, что величины t, df и p-value остались прежними. Изменился лишь только доверительный интервал, который нам показывает где сосредоточена основная “вероятность.”

Реализуем оценку мощности критериев:Изображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описание

Изображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеПо данным оценкам мощности можно понять насколько маленькое изменение насколько правильно будет принято или отвергнуто. Напрашивается тривиальный вывод, чем меньше дельта, тем меньше мощность оценки. Чем больше выборка, тем больше оценка мощности

Применим ранговые тесты Уилкоксона-Манна-Уитни.

В отличие от теста Стьюдента тест Уилкоксона-Манна-Уитни не требует проверку на нормальную распределенность. Однако для последнего есть свои ограничения: 5 < n < 25 и он очень плохо работает с повторяющимися данными.

Изображение выглядит как текст

Автоматически созданное описание Изображение выглядит как текст

Автоматически созданное описаниеСравнение с результами теста Стьюдента:

mu = 5000: (student)p-value 5.088\*e^(-5) < 0.95  
 (WMW) p-value 0.001 < 0.95

Оба теста отвергают гипотезу mu = 5000.

mu = 6800: (student)p-value 0.826 > 0.95  
 (WMW) p-value 0.8339 > 0.95

Оба теста принимают гипотезу mu = 6800. При этом

(student)p-value ~= (WMW) p-value, что дает нам дополнительную уверенность в правильности этих тестов.

Изображение выглядит как текст

Автоматически созданное описаниеПрименение теста Уилкоксона-Манна-Уитни таким образом дает нам информацию о том, что выборки c1 и c2 имеют одинаковое статистическое выборочное среднее.(p-value > 0.5 в обоих случаях гипотеза принимается). Стоит отметить, что p-value во втором случае больше, чем в первом. Из этого можно сделать вывод, что скорее всего выборки c1 и c2 зависимы.

Тесты Фишера, Левене, Бартлетта, Флигнера-Килина.

Изображение выглядит как текст

Автоматически созданное описание

Изображение выглядит как текст

Автоматически созданное описаниеТест Фишера поддерживает гипотезу об одинаковой дисперсии у с1 и с2(что изначально так и задавалось). Для проверки гипотез на однородность дисперсии используется тесты Левене, Бартлетта, Флигнера-Килина. Посмотрим, что они скажут по поводу имеющихся данных из библиотеки Mass genotype(эти данные я буду использовать как свои данные). Тест Левене выдает p-value 0.09, на основании чего можно сделать вывод(хоть и на грани) об однородности дисперсии веса для разных групп мышат.

Изображение выглядит как текст

Автоматически созданное описаниеТесты Бартлетта и Флигнера- Киллена говорят нам примерно о том же, что и тест Левене(значение p-value примерно одинаковое).

Вывод: Если взять стандартный доверительный уровень 0.95, то гипотезу об однородности дисперсий можно принять.

*Часть 3. Свои данные.*

Изображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеВ качестве своих данных я возьму данные genotype из библиотеки Mass, но для них я проверю другой фактор(Mother).

Изображение выглядит как текст

Автоматически созданное описаниеP-value > 0.05 во всех случаях => можем принять все 3 гипотезы.

Изображение выглядит как текст

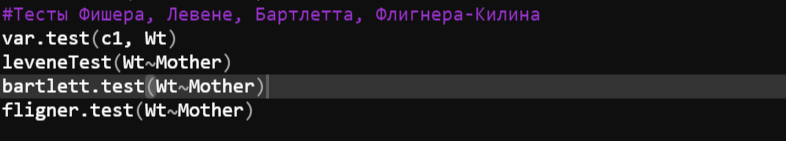
Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описание Изображение выглядит как текст

Автоматически созданное описание Тест на мощность t-критерия говорит нам о том, что изменение с точностью до грамма будет давать нам не очень хороший критерий(p-value = 0.15), c точностью до 2 грамм – лучше(p-value = 0.46), с точностью до 4 грамм – отличный критерий(p-value = 0.96).

Изображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеИзображение выглядит как текст

Автоматически созданное описаниеP-value = 0.4461 => гипотеза о среднем равном 55 принимается, однако заметим что на этих же данных с этими же параметрами тест Стьюдента давал p-value = 0.33. Такая разница в p-value может быть обусловлены тем, что данные не на столько “нормальны”, как в первом случае.

В тесте Фишера я взял 2 случайные выборки из нормального распределения с разными параметрами. Результат закономерен p-value < 0.05, гипотеза об одинаковости дисперсий отвергается.

А вот тесты Левене, Бартлетта, Флигнера-Килина дали p-value > 0.05 следовательно гипотезу об однородности дисперсий можно принять. Более того в среднем p-value данных тестов для wt при факторе Mother больше, чем p-value для wt при факторе Litter. Из этого можно сделать вывод, что разброс веса мышат зависит от того, кто их вскармливал меньше, чем от того, кто их родил.