**58 Что такое статистическая гипотеза? Какие виды гипотез вы знаете?**

Статистической гипотезой называют любое предположение о виде или свойствах

распределения генеральной совокупности. Мы будем рассматривать две гипотезы: нулевую

(основную) H0 и альтернативную (конкурирующую) H1.

**59 Как принято формулировать нулевую гипотезу?**

Нулевая гипотеза (H0) формулируется таким образом, чтобы она предполагала отсутствие статистически значимого эффекта, различия, связи или влияния между переменными или группами в популяции. Таким образом, H0 может быть сформулирована как утверждение о том, что различия между группами или переменными случайны и не являются статистически значимыми.

Например, для сравнения среднего значения переменной в двух группах нулевая гипотеза может быть сформулирована следующим образом: "Средние значения переменной в двух группах равны".

Формулировка нулевой гипотезы является важной частью статистического анализа, поскольку она определяет, какие результаты будут считаться статистически значимыми. Если результаты исследования противоречат нулевой гипотезе с достаточной уверенностью, то исследователь может сделать вывод о наличии статистически значимого эффекта.

**60 Что такое уровень значимости, как он определяется, как влияет на решение о принятие гипотезы?**

Уровнем значимости гипотезы называют допустимую вероятность совершить ошибку

первого рода, то есть принять неверную гипотезу. Обозначается a.

Уровень значимости a выбирается исследователем до того, как будет проверяться гипотеза.

Значение обычно выбирается небольшим, например, 0,1; 0,05 или 0,01.

Каким образом на основании выборки принимается решение? Для этого необходима

специальная функция, называемая статистикой или критерием. Эта функция зависит от выборки

и потому является случайной функцией.

Множество значений статистики включает:

область принятия гипотезы, то есть множество тех значений статистики, при которых

гипотеза H0 принимается

критическую область, то есть множество тех значений статистики, при которых

гипотеза H0 отклоняется и принимается альтернативная гипотеза:

Выбор уровня значимости влияет на решение о принятии или отвержении гипотезы. Если выбрать более строгий уровень значимости (например, α=0.01), то мы требуем более высокой степени уверенности в отклонении нулевой гипотезы. В этом случае мы будем реже отвергать нулевую гипотезу, что может приводить к ошибке второго рода (false negative). Если выбрать менее строгий уровень значимости (например, α=0.10), то мы будем чаще отвергать нулевую гипотезу, что может приводить к ошибке первого рода (false positive).

**61 Каков порядок обработки данных при тестировании гипотезы о равенстве, какие еще тесте при этом должны быть пройдены, какие требования к данным выдвигаются?**

Обработка данных при тестировании гипотезы о равенстве обычно включает следующие шаги:

Сбор данных: Соберите необходимые данные для анализа. Данные должны быть собраны в соответствии с заданными требованиями и стандартами.

Предварительная обработка данных: Оцените и устраните выбросы, пропуски и ошибки в данных. Преобразуйте данные в нужный формат, если это необходимо.

Проверка на нормальность распределения: Примените соответствующий тест (например, тест Шапиро-Уилка) для проверки нормальности распределения данных. Если данные не распределены нормально, может потребоваться применение непараметрических методов.

Проверка на гомогенность дисперсий: Примените соответствующий тест (например, тест Флигнера-Киллена) для проверки гомогенности дисперсий между группами. Если дисперсии не являются одинаковыми, может потребоваться применение корректированных тестов.

Выбор уровня значимости: Определите уровень значимости, который будет использоваться при проверке гипотезы. Обычно выбирают уровень значимости 0.05.

Выбор статистического теста: Определите статистический тест, который будет использоваться при проверке гипотезы, в зависимости от типа данных и целей исследования.

Проверка гипотезы: Выполните выбранный тест и определите, достигаемый уровень значимости. Если p-значение меньше выбранного уровня значимости, то гипотеза отвергается, иначе гипотеза не может быть отвергнута.

Интерпретация результатов: Интерпретируйте результаты теста и примите решение о принятии или отвержении нулевой гипотезы.

Требования к данным, необходимые для тестирования гипотезы, могут различаться в зависимости от выбранного метода. Однако общими требованиями являются:

Данные должны быть собраны случайным образом.

Данные должны быть независимыми.

Данные должны быть достаточно большими для того, чтобы обеспечить точность результатов.

**5. Для чего вообще нужна гипотеза о равенстве средних?**

Гипотеза о равенстве средних значений двух выборок используется для проверки гипотезы о том,

что средние значения двух выборок равны. Это может быть полезно в различных областях,

например, в медицине для сравнения эффективности двух лекарств или в экономике для

сравнения доходов двух групп людей.

**6.Как тестируются независимые и парные выборки?**

Для тестирования независимых выборок используется t-критерий Стьюдента. Для

тестирования парных выборок используется парный t-критерий Стьюдента или также можно использовать непараметрический тест Флингера-Киллина.