

## ЛАБОРАТОРНАЯ РАБОТА № 4

### ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ.

#### ОШИБКИ ПЕРВОГО И ВТОРОГО РОДА

Пусть перед нами стоит задача выбора из двух гипотез. В статистике разработаны методы, позволяющие по выборке решать, какая из двух гипотез должна быть отвергнута (в пользу альтернативной). *Statgraphics* дает возможность проверить некоторые из таких гипотез.

### ПРОВЕРКА ПАРАМЕТРИЧЕСКИХ ГИПОТЕЗ О МАТЕМАТИЧЕСКОМ ОЖИДАНИИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ. ОШИБКА ПЕРВОГО РОДА


**Задача.** Требуется

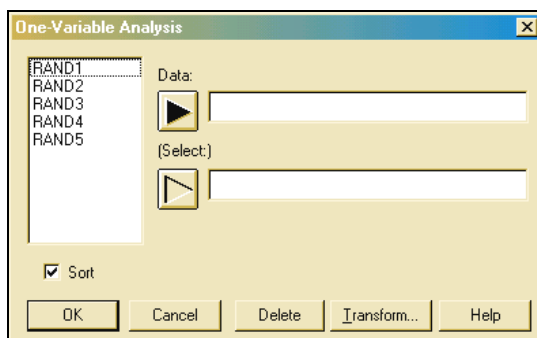
1. Сгенерировать 10 случайных выборок объема  $n=100$  из нормальной генеральной совокупности  $N(0,1)$ . Для получения десяти нужных выборок используйте генератор случайных чисел нормального распределения  $N(0,1)$ .
2. По каждой из полученных выборок проверить на уровне значимости  $\alpha=0,05$  и  $\alpha=0,1$  основную гипотезу  $H_0: m = m_0$ , где  $m_0 = 0$ , при следующих альтернативных гипотезах: а)  $H_a: m \neq 0$ ; в)  $H_a: m < 0$ ; с)  $H_a: m > 0$ ;
3. Выяснить, как часто совершается ошибка первого рода в каждом из этих случаев.


Заготовьте в тетради таблицу следующего вида:

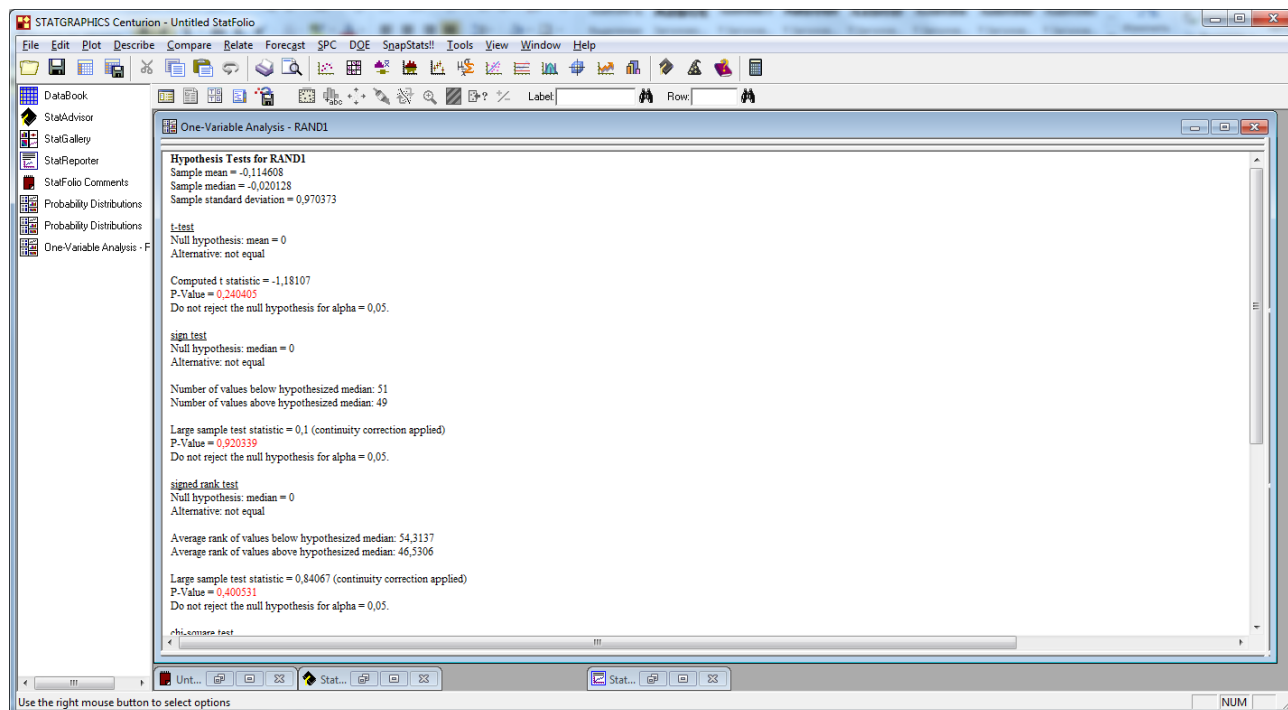
Название выборки	Выбороч- ное сред- нее	а)		в)		с)	
		$\alpha=0,05$	$\alpha=0,1$	$\alpha=0,05$	$\alpha=0,1$	$\alpha=0,05$	$\alpha=0,1$

## Проверка параметрических гипотез

После того, как выборки сгенерированы, в строке меню выберите **Describe**, в раскрывшемся меню выберите **Numeric Data, One-Variable Analyses**, в раскрывшемся окне выберите нужный столбец (их у вас должно быть десять), нажмите кнопку , затем нажмите кнопку ОК.



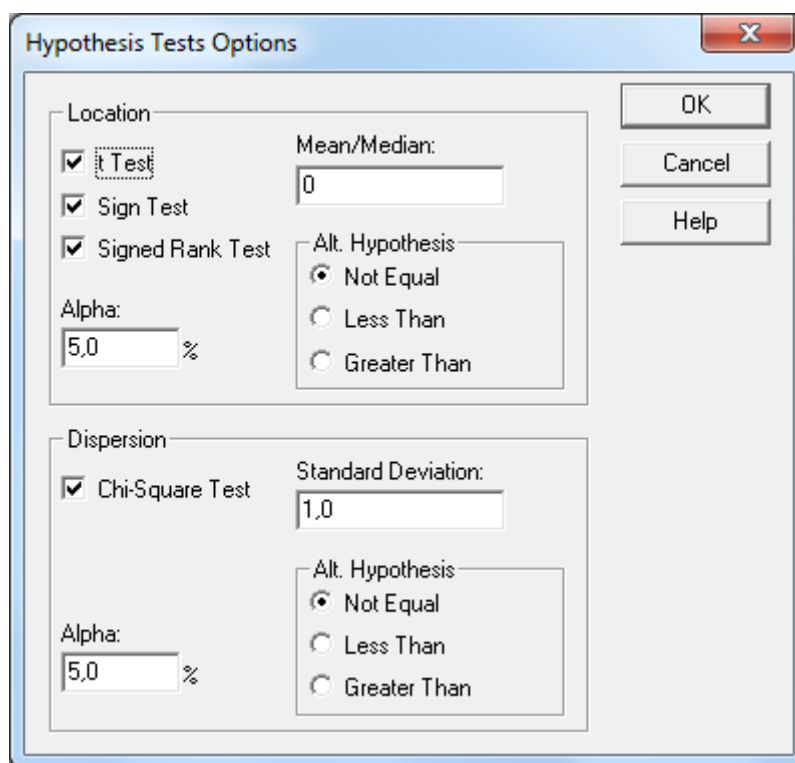
Раскроется окно **Analyses Summary**, нажмите кнопку  **Tables and graphs**, в раскрывшемся окне выберите **Hypothesis Tests**. Раскроется окно **Hypothesis Tests**, в котором можно увидеть примерно следующее:



По умолчанию в качестве нулевой гипотезы выбирается равенство нулю математического ожидания, в качестве альтернативной гипотезы — математическое ожидание не равно нулю. Уровень значимости  $\alpha=0,05$ .  $P\text{-value}=0,240405>0,05$ , поэтому нулевая гипотеза не отвергается при данном уровне значимости. (О

том же говорит результат StatAdvisera «*Do not reject the null hypothesis for alpha =0,05*).

Для того чтобы изменить параметры, щелкните в этом окне правой кнопкой, в появившемся контекстном меню выберите ***Pane Options***, раскроется диалоговое окно ***Hypothesis Tests Options***, в котором вы можете выбирать нужные вам значения **среднего (*mean*)**, уровня значимости (*alpha*), а также варианты альтернативных гипотез (***Not Equal*** — не равно, ***Less Than*** — меньше чем, ***Greater Than*** — больше чем). Выбрав, нажмите ОК.



Результаты видны в окне ***Hypothesis Tests (t-test)***. Занесите результаты для  $\alpha=0,05$  и  $\alpha=0,1$  в колонку *A* (плюс — принимаем гипотезу, минус – отвергаем). Обратите внимание, в этом же окне можно проверять тесты для дисперсии.

Измените тип альтернативной гипотезы ( $m < m_0$ ), занесите результат в колонку *B*. В столбец *C* занесите результат для альтернативной гипотезы третьего типа ( $m > m_0$ ). Аналогично заполните таблицу для остальных девяти выборок.

В этом примере известно, что основная гипотеза верна, так как выборки извлечены из генеральной нормальной совокупности с  $m = 0$ . Подсчитайте для каждого из вариантов *A*, *B* и *C*, в скольких случаях из десяти критерий Стьюдента, заложенный в *Statgraphics*, отверг основную гипотезу  $m = m_0$ , т. е. совер-

шил ошибку первого рода. Сравните частоту этого события в зависимости от  $\alpha$ . Обратите внимание на величину разности между  $\bar{x}$  и  $m_0$  и на то, как это повлияло на результаты проверки гипотез, а также на то, что в некоторых случаях результат проверки основной гипотезы зависит от вида альтернативной. Результаты покажите преподавателю.

### **ОШИБКИ ВТОРОГО РОДА. СВЯЗЬ ВЕРОЯТНОСТЕЙ ОШИБОК ПЕРВОГО И ВТОРОГО РОДА.**

От вас требуется:

1. По выборкам, полученным ранее, проверить на уровне значимости  $\alpha$  основную гипотезу  $H_0: m=m_0$ , где  $m_0 = 0,2$ , при альтернативной гипотезе:  $H_1: m < 0,2$  для двух значений  $\alpha = 0,05$  и  $\alpha = 0,1$ .
2. Выяснить, как часто совершается ошибка второго рода в каждом из двух случаев.

Заготовьте в тетради таблицу следующего вида:

Название выборки	$\alpha = 0,05$	$\alpha = 0,1$

Заполните таблицу аналогично тому, как вы делали это в первом задании. Не забывайте при необходимости изменять содержимое полей, в которых задаются  $m_0$ , тип альтернативной гипотезы и  $\alpha$ .

В данном задании известно, что альтернативная гипотеза  $H_0: m < 0,2$  верна, поскольку  $m=0$ . Для каждого из двух случаев подсчитайте, сколько раз не отвергается основная гипотеза  $m=0,2$ ; т. е. совершается ошибка второго рода. Как зависит частота ошибки второго рода от вероятности ошибки первого рода ( $\alpha$ )? *Покажите результаты преподавателю.*

### **ЗАДАНИЕ**

Предложена диета для похудения. В результате двухнедельного применения этой диеты масса тела пациентов изменилась следующим образом:

Масса, кг	1	2	3	4	5	6	7	8	9	10
До диеты	68	80	92	81	70	79	78	68	57	76
После диеты	60	84	87	79	74	71	72	67	57	60

Можно ли рекомендовать эту диету? ( $\alpha=0,05$ )

Сведите эту задачу к проверке гипотезы о равенстве математического ожидания нулю. В каком случае применение критерия Стьюдента строго обосновано? При каком условии его можно использовать приближенно? **Почему?**

### ВОПРОСЫ К ЛАБОРАТОРНОЙ РАБОТЕ

1. Что такое статистическая гипотеза?
2. Что такое простая гипотеза?
3. Что такое параметрическая гипотеза? Приведите пример.
4. Что такое ошибка первого рода? второго рода при проверке статистических гипотез?
5. Что такое критическая область?
6. Что такое наилучшая критическая область (область принятия решения)?
7. Что происходит с вероятностью ошибки второго рода при уменьшении вероятности ошибки первого рода? Что такое непараметрическая гипотеза? Приведите пример.
8. Можно ли по выборочному среднему спрогнозировать, в каких случаях гипотеза  $H_0$  будет отвергаться?