Министерство Образования Республики Беларусь

Учреждение образования

«Белорусский государственный университет информатики и радиоэлектроники»

Факультет информационных технологий и управления

Кафедра ИИТ

**ЛАБОРАТОРНАЯ РАБОТА №2**

«Проверка гипотез однородности и согласия»

Выполнили:

студент ФИТиУ

гр. 521703

Сидоров И.С. \_\_\_\_\_\_\_\_\_\_\_\_\_

Султанов К.Г. \_\_\_\_\_\_\_\_\_\_\_\_\_

Проверила:

Степанова М.Д.\_\_\_\_\_\_\_\_\_\_\_\_\_

**Минск, 2008**

**Задание**: изучить методы статистической проверки гипотез.

В лабораторной работе № 2 необходимо

1) проверить гипотезу согласия с помощью критериев Пирсона и

Колмогорова,

2) проверить гипотезы об однородности параметров (математических

ожиданий и дисперсий) двух нормальных распределений.

# Проверка соответствия выбранной модели распределения исходным данным (критерии согласия).

**Задание**: Проверить гипотезу согласия с помощью критериев Пирсона и Колмогорова.

**Данные**: Модельные данные: ξ− случайная величина с нормальным распределением, n  = 100, α = 0,05. Параметры распределения приведены в таблице.

Параметры для моделирования нормально распределенной

# случайной величины

|  |  |
| --- | --- |
| Математическое ожидание  μ | Дисперсия  σ2 |
| 6 | 4 |

# Теоретическая часть

**Критерии согласия**

*Н*0 : (*x*) = 

основаны на использовании различных мер расстояний между анализируемой эмпирической функцией распределения, определяемой по выборке, и гипотетической модельной .

## Критерий Пирсона

Критерий  Пирсона позволяет проверить гипотезу H0, когда значения параметров  неизвестны и данные группированы. Процедура проверки гипотезы состоит из следующих шагов:

1. Диапазон значений исследуемой случайной величины ξ разбивается на *k* взаимно исключающих и непересекающихся интервалов . Длина интервалов разбиения не обязательно одинакова.
2. На основании выборочных данных  строятся статистические оценки  неизвестных параметров , от которых зависит закон распределения *F*.
3. Подсчитывается число наблюдений , попадающих в каждый интервал группирования , *i*= 1,…, *k.*
4. Вычисляются вероятности событий ξ , т.е. вероятности попадания случайной величины ξ в интервал :

 = (; ) – (; ),

1. Вычисляется ожидаемое число наблюдений  в интервале  при условии справедливости гипотезы *Н*0:  = *п*.
2. Вычисляется статистика

 = ,  
которая при верной *Н*0 имеет –распределение с *f = k – s* – 1 степенями свободы.

1. ***Гипотеза*** о том, что исследуемая случайная величина ξ подчиняется закону распределения , ***принимается на уровне значимости α***, если

Δ(α/2) ≤ < Δ(1 – α/2),

1. где Δ(ε) – квантиль уровня ε имеет – распределение с f = k – s – 1 степенями свободы.
2. Если ≥ Δ (1 – α/2), гипотеза Н0 отклоняется, так как выполнение неравенства свидетельствует о слишком большом отклонении исследуемого закона распределения от (х). Случай < Δ (α/2) требует дополнительного исследования. Слишком малые значения статистики критерия говорят о неудачном выборе закона F (завышение числа параметров), нарушении технологии выборочного обследования и т.д.

# Практическая часть

Воспользуемся возможностями программы STATISTICA для выполнения задания.

Введём начальные значения как указано на рис.1.

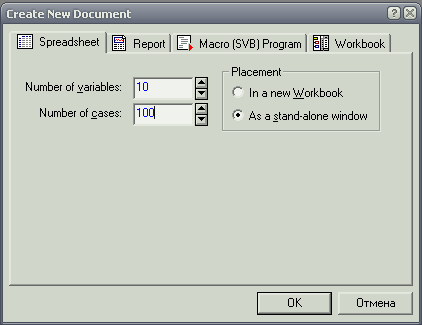
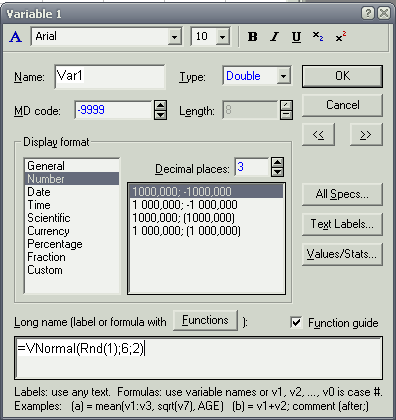
 

Рисунок 1. Ввод исходных данных.

Полученные результаты: сто случайных величин распределённых по нормальному закону.

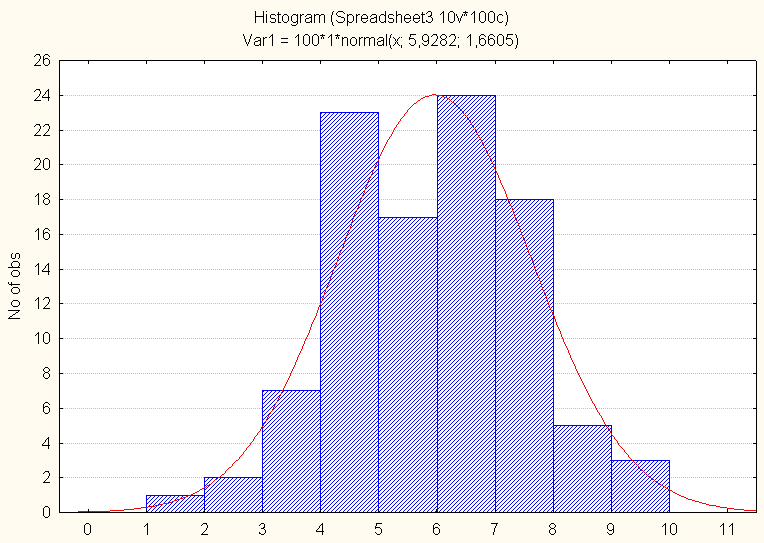


Рисунок 2. Гистограмма полученных результатов.

Для проверки критерия запустим Distribution Fitting из пакета STATISTICA.

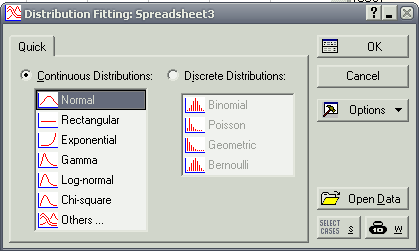


Рисунок 3.

## Результаты проверки гипотезы согласия с помощью критериев -Пирсона и Колмогорова показаны на рис. 4.

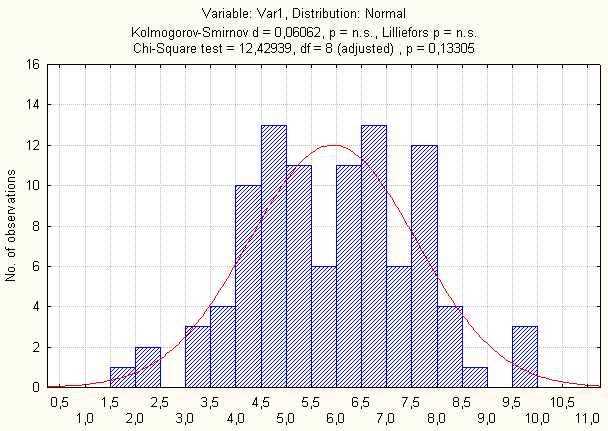


Рисунок 4. Результаты проверки гипотезы согласия.

Применим способ проверки гипотезы, в котором используется решающее правило, основанное на *Р*–значении:

принимается гипотеза 

Для статистики критерия, равной  = 12,429, вычисляется *Р*–значение. Его величина *р* = 0,133, поэтому при уровне значимости α = 0,05 гипотеза H0 о нормальности распределения принимается.

Когда модельное распределение известно полностью и является непрерывным, для проверки гипотезы согласия

*Н*0 : (*x*) = 

целесообразно использовать критерий Колмогорова.

Определим расстояние Колмогорова между эмпирической  и теоретической функциями распределения:

 .

Решающее правило:

принимается гипотеза 

Если верна *Н*0 и *n* ≥ 20, то независимо от *F*0 (*x*) случайная величина *n*1/2 *D* имеет распределение Колмогорова. Это позволяет определить Δ(α).

Для α=0.05, Δ=1.36.

Для проверки этой гипотезы используем критерий Колмогорова. В результате обработки данных имеем *D* = 0.06. Отсюда значение статистики критерия *n*1/2 *D* = = 0,6. при α= 0,05 находим границу области принятия нулевой гипотезы Δ(*α*) = 1,36. Так как *n*1/2 *D* =0, 6 < 1,36, то результаты наблюдений не противоречат гипотезе *Н*0.

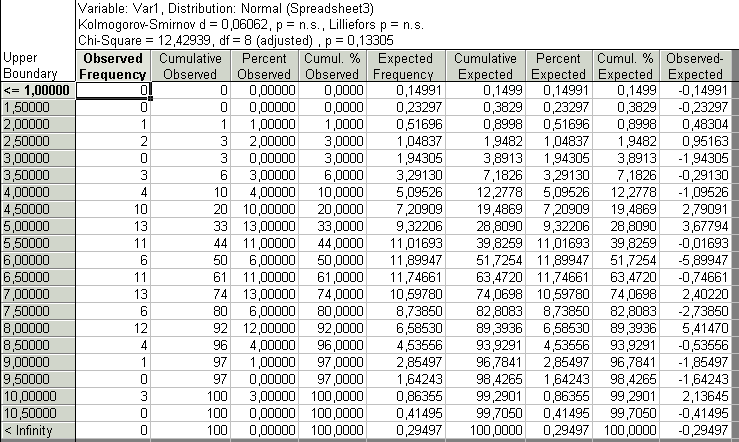


Рисунок 5. Таблица характеризующая нормальное распределение.

# Проверка гипотез однородности

# Теоретическая часть

Гипотезы равенства (однородности) математических ожиданий  и дисперсий  и их двусторонние альтернативы  и можно записать в виде

: =, : ≠;

: = , :  ≠.

Однородность средних значений

**Исходные данные** представлены двумя независимыми случайными выборками  и  объемов  и соответственно:

= (), *i*= 1, 2.

**Предполагается**, что выборки извлечены из нормальных распределений с равными дисперсиями (= ).

Для проверки гипотезы равенства математических ожиданий в двух выборках в этом случае используется t–критерий Стьюдента, или двухвыборочная статистика Стьюдента:

*t* = ( *–* ) / ( (1/+ 1/))1/2 .

В формуле использованы следующие обозначения:

=  *–* выборочное среднее для выборки;

 =  *–* объединенная выборочная дисперсия;

= *–* выборочная оценка дисперсии для выборки  (*i* = 1, 2).

**Решающее правило** имеет вид

принимается гипотеза 

где Δ(ε) *–* квантиль уровня 1*–*α*/2 t*–распределения Стьюдента с +–2 степенями свободы.

**При правосторонней альтернативе** : > **решающее правило** имеет вид

принимается гипотеза 

где Δ(ε) *–* квантиль уровня 1*–*α *t*–распределения Стьюдента с +–2 степенями свободы.

**При левосторонней альтернативе** : < **решающее правило** имеет вид

принимается гипотеза 

где Δ(ε) *–* квантиль уровня α *t*–распределения Стьюдента с +–2 степенями свободы.

Для обеспечения устойчивости решений к нарушению нормальности распределения и равенства дисперсий, разработаны устойчивые критерии, например**, *t*–критерий Уэлча**.

Для проверки гипотез

: – = δ, :–  ≠ δ

(значение δ фиксировано, в случае гипотезы равенства средних δ = 0) при возможных нарушениях условия равенства дисперсий (≠) используется тест вида

принимается гипотеза 

Здесь

*t*1 = ( *–* *–* δ) / (/+/)1/2 – **статистика Уэлча**,

Δ(ε) *–* квантиль уровня 1*–*α*/2 t*–распределения Стьюдента с  степенями свободы,

 =  , .

При правосторонней и левосторонних альтернативах решающие правила аналогичны случаю применения *t*–критерия Стьюдента.

**Однородность дисперсий**

Если необходимо выяснить, за счет чего обнаружилась неоднородность рассматриваемых выборок, то следует дополнительно произвести проверку однородности дисперсий. **Статистический критерий однородности двух выборочных дисперсий** основан на статистике

*F*=  ,

которая в условиях справедливости гипотезы  (2.3) имеет *F–*распределение Фишера с числами степеней свободы числителя и знаменателя, равными соответственно *–*1 и *–*1.

Тест уровня значимости α для проверки гипотезы (2.3) записывается следующим образом:

принимается гипотеза 

где Δ(ε) *–* порог теста, определяемый при верной  (2.3) как квантиль уровняε *F–*распределения Фишера с  *–* 1 и *–* 1 степенями свободы.

Если рассматривается правосторонняя альтернатива  > , то гипотеза  отклоняется при *F >* Δ(1*–*α), а для левосторонней альтернативы  <  гипотеза  отклоняется при *F <* Δ(α).

**Вывод**:

Научились использовать критерий согласия для проверки предположения о согласованности модели с наблюдаемыми данными.

Научились использовать критерий  Пирсона позволяющий проверить гипотезу, когда значения параметров  неизвестны.

А также в случае, когда модельное распределение известно полностью и является непрерывным, для проверки гипотезы согласия целесообразно использовать критерий Колмогорова.