МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А.И. ГЕРЦЕНА»



Направление подготовки

09.03.01 – Информатика и вычислительная техника

Профиль «Технологии разработки программного обеспечения»

**Лабораторная работа №6. Часть 3**

**“Проверка статистических гипотез”**

|  | Работу выполнили:  Балаев Жамал,  Васильева Марина,  Иванов Никита,  Шардт Максим,  Рожков Максим  очная форма обучения  курс: 2; группа: ИВТ-1.1 |
| --- | --- |
|  | Научный руководитель:  Профессор Власова Елена Зотиковна |

Санкт-Петербург

2023

## **Содержание**

ЛР [6. Часть 3. Выполнил Балаев Ж. Б. 3](#_nnr3ru9fslfs)

ЛР [6. Часть 3. Выполнил](#_nnr3ru9fslfs)а Васильева М. А. [10](#_l9o1wj9xo6r4)

ЛР [6. Часть 3. Выполнил](#_nnr3ru9fslfs) Иванов Н. Р. [17](#_tqfv80esl2p)

ЛР [6. Часть 3. Выполнил](#_nnr3ru9fslfs) Рожков М. В. [24](#_ao8x14546eh4)

ЛР [6. Часть 3. Выполнил](#_nnr3ru9fslfs) Шардт М. А. [31](#_swro9pot1m6z)

Приложение 1 38

## 

## 

## **Лабораторная работа №6. Часть 3**

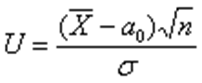
*Выполнил Балаев Ж. Б.*

**Цель работы:** Проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

**Оборудования:** ПК, Visual Studio Code, Microsoft Excel

*Использованные формулы:*

1. Случайная величина, которую рассматривают в качестве статистического критерия

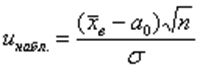


где – случайное значение *выборочной средней*

1. Функция Лапласа для односторонней области (лево- или право-):

1. Функция Лапласа для двусторонней области:

1. Наблюдаемое значение критерия



1. Среднее квадратическое отклонение нормального распределения случайной величины:

1. Статистика гипотезы:

1. Выборочные доли признака:

1. Дисперсия

1. Статистика гипотезы:

,

где - общая выборочная доля признака

1. Статистика критерия:

1. Выборочная доля признака:

*Задание 1:*

Для проверки эффективности новой технологии отобраны две группы рабочих.

В 1-ой группе численность n1 = 50 чел., где применялась новая технология, выборочная средняя выработка составила изделий.

Во 2-ой группе численность n2 = 70 чел., выборочная средняя изделий.

Предварительно установлено, что дисперсия выработки в группах равны соответственно и . На уровне значимости выяснить влияние новой технологии на среднюю производительность.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи мы написали программу, которая по формуле (6) вычисляет статистику гипотезы. Для выбранной конкурирующей области выбирается односторонняя критическая область, значит критическое значение статистики находится по формуле (2).

Если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае получаем, что гипотеза отвергается, т.е. на пятипроцентном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

| from math import sqrt from parse\_data import data  def get\_critical\_area(x, y, a, H):  if 'H1' == H:  t = 1 - 2 \* a  else:  t = 1 - a  return t   def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i     σx2 = 100 σy2 = 74 n1 = 50 n2 = 70 x = 85 y = 78 a = 0.05 h = 'H1'  t = (x - y)/sqrt(σx2 / n1 + σy2 / n2) tkp = get\_tkp(get\_critical\_area(x, y, a, h)) print(f'т.к {t} {["<", ">"][t > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 1



Рис. 1. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 2:*

Контрольная работа по высшей математике выполняли студенты 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решены 58. Во 2-ой группе из 120 задач верно решены 65. На уровне значимости проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи нами была написана программа, которая вычисляет выборочные доли решенных задач по формуле (7), а также статистику по формуле (9).

Аналогично первой задаче: если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае нулевую гипотезу H0 об одинаковой усваиваемости материала принимаем.

| from math import sqrt from parse\_data import data  def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i   n1 = 100 n2 = 120 m1 = 58 m2 = 65 a = 0.02 w1 = m1 / n1 w2 = m2 / n2 p = (m1 + m2) / (n1 + n2)  t = (w1 - w2) / sqrt(p \* (1 - p) \* (1 / n1 + 1 / n2)) tkp = get\_tkp(1 - a)  print(t) print(f'т.к |{t}| {["<", ">"][abs(t) > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 2



Рис. 2. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,02 нулевую гипотезу принимаем.

*Задание 3:*

Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты 4-х групп первого курса. В 1-ой группе было предложено 105 задач, из которых верно решено 60 задач. Во 2-ой группе всего задач 140 и верно решено 69. В третьей группе было 125 задач, верно решено 63. В 4-ой группе всего задач 160, верно решено 105.

На уровне значимости выяснить, можно ли считать, что различия в усвоении учебного материала студентами 4-х групп первого курса существенны.

*Решение:*

Выдвигаем гипотезу .

Для данной задачи нами была написана программа, которая вычисляет выборочную долю признака по формуле (11) и по формуле (7), а также статистику по формуле (10).

Если , то гипотеза H0 отвергается, в противном случае - принимается. В нашей задаче нулевая гипотеза отвергается, т.е. различие в усвоении материала студентами 4-х групп существенно на уровне значимости .

| def get\_xalpha(x, alpha, k):  from pars\_data import data  return data[alpha][k-1]  pass  n = (105, 140, 125, 160) m = (60, 69, 63, 105) alpha = 0.05 k = len(n) - 1 p = sum(m) / sum(n) w = list(map(lambda m, n: m / n, m, n))  x = (1 / (p \* (1 - p)) \* sum([n[i] \* (w[i] - p) \*\* 2 for i in range(len(n))])) xalpha = get\_xalpha(x, alpha, k)  print(f'Так как {x} {[">", "<"][x < xalpha]} {xalpha}, то на уровне alpha = {alpha} {["гипотеза H0 принимается", "Гипотеза H0 отвергается"][x > xalpha]}') |
| --- |

Листинг кода 3



Рис. 3. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

## **Лабораторная работа №6. Часть 3**

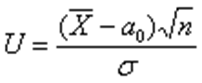
*Выполнила Васильева М. А.*

**Цель работы:** Проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

**Оборудования:** ПК, Visual Studio Code, Microsoft Excel

*Использованные формулы:*

1. Случайная величина, которую рассматривают в качестве статистического критерия

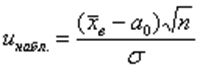


где – случайное значение *выборочной средней*

1. Функция Лапласа для односторонней области (лево- или право-):

1. Функция Лапласа для двусторонней области:

1. Наблюдаемое значение критерия



1. Среднее квадратическое отклонение нормального распределения случайной величины:

1. Статистика гипотезы:

1. Выборочные доли признака:

1. Дисперсия

1. Статистика гипотезы:

,

где - общая выборочная доля признака

1. Статистика критерия:

1. Выборочная доля признака:

*Задание 1:*

Для проверки эффективности новой технологии отобраны две группы рабочих.

В 1-ой группе численность n1 = 50 чел., где применялась новая технология, выборочная средняя выработка составила изделий.

Во 2-ой группе численность n2 = 70 чел., выборочная средняя изделий.

Предварительно установлено, что дисперсия выработки в группах равны соответственно и . На уровне значимости выяснить влияние новой технологии на среднюю производительность.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи мы написали программу, которая по формуле (6) вычисляет статистику гипотезы. Для выбранной конкурирующей области выбирается односторонняя критическая область, значит критическое значение статистики находится по формуле (2).

Если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае получаем, что гипотеза отвергается, т.е. на пятипроцентном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

| from math import sqrt from parse\_data import data  def get\_critical\_area(x, y, a, H):  if 'H1' == H:  t = 1 - 2 \* a  else:  t = 1 - a  return t   def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i     σx2 = 100 σy2 = 74 n1 = 50 n2 = 70 x = 85 y = 78 a = 0.05 h = 'H1'  t = (x - y)/sqrt(σx2 / n1 + σy2 / n2) tkp = get\_tkp(get\_critical\_area(x, y, a, h)) print(f'т.к {t} {["<", ">"][t > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 1



Рис. 1. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 2:*

Контрольная работа по высшей математике выполняли студенты 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решены 58. Во 2-ой группе из 120 задач верно решены 65. На уровне значимости проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи нами была написана программа, которая вычисляет выборочные доли решенных задач по формуле (7), а также статистику по формуле (9).

Аналогично первой задаче: если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае нулевую гипотезу H0 об одинаковой усваиваемости материала принимаем.

| from math import sqrt from parse\_data import data  def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i   n1 = 100 n2 = 120 m1 = 58 m2 = 65 a = 0.02 w1 = m1 / n1 w2 = m2 / n2 p = (m1 + m2) / (n1 + n2)  t = (w1 - w2) / sqrt(p \* (1 - p) \* (1 / n1 + 1 / n2)) tkp = get\_tkp(1 - a)  print(t) print(f'т.к |{t}| {["<", ">"][abs(t) > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 2



Рис. 2. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,02 нулевую гипотезу принимаем.

*Задание 3:*

Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты 4-х групп первого курса. В 1-ой группе было предложено 105 задач, из которых верно решено 60 задач. Во 2-ой группе всего задач 140 и верно решено 69. В третьей группе было 125 задач, верно решено 63. В 4-ой группе всего задач 160, верно решено 105.

На уровне значимости выяснить, можно ли считать, что различия в усвоении учебного материала студентами 4-х групп первого курса существенны.

*Решение:*

Выдвигаем гипотезу .

Для данной задачи нами была написана программа, которая вычисляет выборочную долю признака по формуле (11) и по формуле (7), а также статистику по формуле (10).

Если , то гипотеза H0 отвергается, в противном случае - принимается. В нашей задаче нулевая гипотеза отвергается, т.е. различие в усвоении материала студентами 4-х групп существенно на уровне значимости .

| def get\_xalpha(x, alpha, k):  from pars\_data import data  return data[alpha][k-1]  pass  n = (105, 140, 125, 160) m = (60, 69, 63, 105) alpha = 0.05 k = len(n) - 1 p = sum(m) / sum(n) w = list(map(lambda m, n: m / n, m, n))  x = (1 / (p \* (1 - p)) \* sum([n[i] \* (w[i] - p) \*\* 2 for i in range(len(n))])) xalpha = get\_xalpha(x, alpha, k)  print(f'Так как {x} {[">", "<"][x < xalpha]} {xalpha}, то на уровне alpha = {alpha} {["гипотеза H0 принимается", "Гипотеза H0 отвергается"][x > xalpha]}') |
| --- |

Листинг кода 3



Рис. 3. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

## **Лабораторная работа №6. Часть 3**

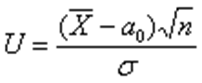
*Выполнил Иванов Н. Р.*

**Цель работы:** Проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

**Оборудования:** ПК, Visual Studio Code, Microsoft Excel

*Использованные формулы:*

1. Случайная величина, которую рассматривают в качестве статистического критерия

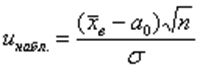


где – случайное значение *выборочной средней*

1. Функция Лапласа для односторонней области (лево- или право-):

1. Функция Лапласа для двусторонней области:

1. Наблюдаемое значение критерия



1. Среднее квадратическое отклонение нормального распределения случайной величины:

1. Статистика гипотезы:

1. Выборочные доли признака:

1. Дисперсия

1. Статистика гипотезы:

,

где - общая выборочная доля признака

1. Статистика критерия:

1. Выборочная доля признака:

*Задание 1:*

Для проверки эффективности новой технологии отобраны две группы рабочих.

В 1-ой группе численность n1 = 50 чел., где применялась новая технология, выборочная средняя выработка составила изделий.

Во 2-ой группе численность n2 = 70 чел., выборочная средняя изделий.

Предварительно установлено, что дисперсия выработки в группах равны соответственно и . На уровне значимости выяснить влияние новой технологии на среднюю производительность.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи мы написали программу, которая по формуле (6) вычисляет статистику гипотезы. Для выбранной конкурирующей области выбирается односторонняя критическая область, значит критическое значение статистики находится по формуле (2).

Если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае получаем, что гипотеза отвергается, т.е. на пятипроцентном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

| from math import sqrt from parse\_data import data  def get\_critical\_area(x, y, a, H):  if 'H1' == H:  t = 1 - 2 \* a  else:  t = 1 - a  return t   def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i     σx2 = 100 σy2 = 74 n1 = 50 n2 = 70 x = 85 y = 78 a = 0.05 h = 'H1'  t = (x - y)/sqrt(σx2 / n1 + σy2 / n2) tkp = get\_tkp(get\_critical\_area(x, y, a, h)) print(f'т.к {t} {["<", ">"][t > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 1



Рис. 1. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 2:*

Контрольная работа по высшей математике выполняли студенты 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решены 58. Во 2-ой группе из 120 задач верно решены 65. На уровне значимости проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи нами была написана программа, которая вычисляет выборочные доли решенных задач по формуле (7), а также статистику по формуле (9).

Аналогично первой задаче: если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае нулевую гипотезу H0 об одинаковой усваиваемости материала принимаем.

| from math import sqrt from parse\_data import data  def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i   n1 = 100 n2 = 120 m1 = 58 m2 = 65 a = 0.02 w1 = m1 / n1 w2 = m2 / n2 p = (m1 + m2) / (n1 + n2)  t = (w1 - w2) / sqrt(p \* (1 - p) \* (1 / n1 + 1 / n2)) tkp = get\_tkp(1 - a)  print(t) print(f'т.к |{t}| {["<", ">"][abs(t) > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 2



Рис. 2. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,02 нулевую гипотезу принимаем.

*Задание 3:*

Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты 4-х групп первого курса. В 1-ой группе было предложено 105 задач, из которых верно решено 60 задач. Во 2-ой группе всего задач 140 и верно решено 69. В третьей группе было 125 задач, верно решено 63. В 4-ой группе всего задач 160, верно решено 105.

На уровне значимости выяснить, можно ли считать, что различия в усвоении учебного материала студентами 4-х групп первого курса существенны.

*Решение:*

Выдвигаем гипотезу .

Для данной задачи нами была написана программа, которая вычисляет выборочную долю признака по формуле (11) и по формуле (7), а также статистику по формуле (10).

Если , то гипотеза H0 отвергается, в противном случае - принимается. В нашей задаче нулевая гипотеза отвергается, т.е. различие в усвоении материала студентами 4-х групп существенно на уровне значимости .

| def get\_xalpha(x, alpha, k):  from pars\_data import data  return data[alpha][k-1]  pass  n = (105, 140, 125, 160) m = (60, 69, 63, 105) alpha = 0.05 k = len(n) - 1 p = sum(m) / sum(n) w = list(map(lambda m, n: m / n, m, n))  x = (1 / (p \* (1 - p)) \* sum([n[i] \* (w[i] - p) \*\* 2 for i in range(len(n))])) xalpha = get\_xalpha(x, alpha, k)  print(f'Так как {x} {[">", "<"][x < xalpha]} {xalpha}, то на уровне alpha = {alpha} {["гипотеза H0 принимается", "Гипотеза H0 отвергается"][x > xalpha]}') |
| --- |

Листинг кода 3



Рис. 3. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

## **Лабораторная работа №6. Часть 3**

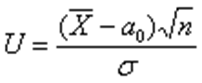
*Выполнил Рожков М. В.*

**Цель работы:** Проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

**Оборудования:** ПК, Visual Studio Code, Microsoft Excel

*Использованные формулы:*

1. Случайная величина, которую рассматривают в качестве статистического критерия

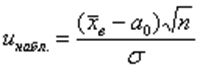


где – случайное значение *выборочной средней*

1. Функция Лапласа для односторонней области (лево- или право-):

1. Функция Лапласа для двусторонней области:

1. Наблюдаемое значение критерия



1. Среднее квадратическое отклонение нормального распределения случайной величины:

1. Статистика гипотезы:

1. Выборочные доли признака:

1. Дисперсия

1. Статистика гипотезы:

,

где - общая выборочная доля признака

1. Статистика критерия:

1. Выборочная доля признака:

*Задание 1:*

Для проверки эффективности новой технологии отобраны две группы рабочих.

В 1-ой группе численность n1 = 50 чел., где применялась новая технология, выборочная средняя выработка составила изделий.

Во 2-ой группе численность n2 = 70 чел., выборочная средняя изделий.

Предварительно установлено, что дисперсия выработки в группах равны соответственно и . На уровне значимости выяснить влияние новой технологии на среднюю производительность.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи мы написали программу, которая по формуле (6) вычисляет статистику гипотезы. Для выбранной конкурирующей области выбирается односторонняя критическая область, значит критическое значение статистики находится по формуле (2).

Если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае получаем, что гипотеза отвергается, т.е. на пятипроцентном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

| from math import sqrt from parse\_data import data  def get\_critical\_area(x, y, a, H):  if 'H1' == H:  t = 1 - 2 \* a  else:  t = 1 - a  return t   def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i     σx2 = 100 σy2 = 74 n1 = 50 n2 = 70 x = 85 y = 78 a = 0.05 h = 'H1'  t = (x - y)/sqrt(σx2 / n1 + σy2 / n2) tkp = get\_tkp(get\_critical\_area(x, y, a, h)) print(f'т.к {t} {["<", ">"][t > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 1



Рис. 1. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 2:*

Контрольная работа по высшей математике выполняли студенты 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решены 58. Во 2-ой группе из 120 задач верно решены 65. На уровне значимости проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи нами была написана программа, которая вычисляет выборочные доли решенных задач по формуле (7), а также статистику по формуле (9).

Аналогично первой задаче: если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае нулевую гипотезу H0 об одинаковой усваиваемости материала принимаем.

| from math import sqrt from parse\_data import data  def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i   n1 = 100 n2 = 120 m1 = 58 m2 = 65 a = 0.02 w1 = m1 / n1 w2 = m2 / n2 p = (m1 + m2) / (n1 + n2)  t = (w1 - w2) / sqrt(p \* (1 - p) \* (1 / n1 + 1 / n2)) tkp = get\_tkp(1 - a)  print(t) print(f'т.к |{t}| {["<", ">"][abs(t) > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 2



Рис. 2. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,02 нулевую гипотезу принимаем.

*Задание 3:*

Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты 4-х групп первого курса. В 1-ой группе было предложено 105 задач, из которых верно решено 60 задач. Во 2-ой группе всего задач 140 и верно решено 69. В третьей группе было 125 задач, верно решено 63. В 4-ой группе всего задач 160, верно решено 105.

На уровне значимости выяснить, можно ли считать, что различия в усвоении учебного материала студентами 4-х групп первого курса существенны.

*Решение:*

Выдвигаем гипотезу .

Для данной задачи нами была написана программа, которая вычисляет выборочную долю признака по формуле (11) и по формуле (7), а также статистику по формуле (10).

Если , то гипотеза H0 отвергается, в противном случае - принимается. В нашей задаче нулевая гипотеза отвергается, т.е. различие в усвоении материала студентами 4-х групп существенно на уровне значимости .

| def get\_xalpha(x, alpha, k):  from pars\_data import data  return data[alpha][k-1]  pass  n = (105, 140, 125, 160) m = (60, 69, 63, 105) alpha = 0.05 k = len(n) - 1 p = sum(m) / sum(n) w = list(map(lambda m, n: m / n, m, n))  x = (1 / (p \* (1 - p)) \* sum([n[i] \* (w[i] - p) \*\* 2 for i in range(len(n))])) xalpha = get\_xalpha(x, alpha, k)  print(f'Так как {x} {[">", "<"][x < xalpha]} {xalpha}, то на уровне alpha = {alpha} {["гипотеза H0 принимается", "Гипотеза H0 отвергается"][x > xalpha]}') |
| --- |

Листинг кода 3



Рис. 3. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

## **Лабораторная работа №6. Часть 3**

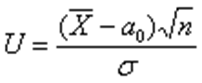
*Выполнил Шардт М. А.*

**Цель работы:** Проверить статистическую гипотезу о нормальном законе распределения данных, приведенных в решаемой задаче.

**Оборудования:** ПК, Visual Studio Code, Microsoft Excel

*Использованные формулы:*

1. Случайная величина, которую рассматривают в качестве статистического критерия

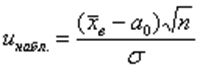


где – случайное значение *выборочной средней*

1. Функция Лапласа для односторонней области (лево- или право-):

1. Функция Лапласа для двусторонней области:

1. Наблюдаемое значение критерия



1. Среднее квадратическое отклонение нормального распределения случайной величины:

1. Статистика гипотезы:

1. Выборочные доли признака:

1. Дисперсия

1. Статистика гипотезы:

,

где - общая выборочная доля признака

1. Статистика критерия:

1. Выборочная доля признака:

*Задание 1:*

Для проверки эффективности новой технологии отобраны две группы рабочих.

В 1-ой группе численность n1 = 50 чел., где применялась новая технология, выборочная средняя выработка составила изделий.

Во 2-ой группе численность n2 = 70 чел., выборочная средняя изделий.

Предварительно установлено, что дисперсия выработки в группах равны соответственно и . На уровне значимости выяснить влияние новой технологии на среднюю производительность.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи мы написали программу, которая по формуле (6) вычисляет статистику гипотезы. Для выбранной конкурирующей области выбирается односторонняя критическая область, значит критическое значение статистики находится по формуле (2).

Если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае получаем, что гипотеза отвергается, т.е. на пятипроцентном уровне значимости делаем вывод, что новая технология позволяет повысить среднюю выработку рабочих.

| from math import sqrt from parse\_data import data  def get\_critical\_area(x, y, a, H):  if 'H1' == H:  t = 1 - 2 \* a  else:  t = 1 - a  return t   def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i     σx2 = 100 σy2 = 74 n1 = 50 n2 = 70 x = 85 y = 78 a = 0.05 h = 'H1'  t = (x - y)/sqrt(σx2 / n1 + σy2 / n2) tkp = get\_tkp(get\_critical\_area(x, y, a, h)) print(f'т.к {t} {["<", ">"][t > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 1



Рис. 1. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 2:*

Контрольная работа по высшей математике выполняли студенты 2-х групп. В 1-ой группе было предложено 100 задач, из которых правильно решены 58. Во 2-ой группе из 120 задач верно решены 65. На уровне значимости проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

*Решение:*

Исходя из условия задачи получаем:

1. Проверяемая гипотеза
2. Конкурирующая гипотеза:

Для данной задачи нами была написана программа, которая вычисляет выборочные доли решенных задач по формуле (7), а также статистику по формуле (9).

Аналогично первой задаче: если фактически наблюдаемое значение статистики t больше критического tкр, определенного на уровне значимости α, то гипотеза H0 отвергается. В другом же случае делается вывод, что нулевая гипотеза H0 не противоречит имеющимся наблюдениям.

В нашем случае нулевую гипотезу H0 об одинаковой усваиваемости материала принимаем.

| from math import sqrt from parse\_data import data  def get\_tkp(f):  for i in data:  if f < i:  return data[[i, prev][(abs(prev - f) < (abs(i - f)))]]  prev = i   n1 = 100 n2 = 120 m1 = 58 m2 = 65 a = 0.02 w1 = m1 / n1 w2 = m2 / n2 p = (m1 + m2) / (n1 + n2)  t = (w1 - w2) / sqrt(p \* (1 - p) \* (1 / n1 + 1 / n2)) tkp = get\_tkp(1 - a)  print(t) print(f'т.к |{t}| {["<", ">"][abs(t) > tkp]} {tkp}, то {["Гипотеза H0 отвергается", "Гипотеза H0 принимается"][t < tkp]}') |
| --- |

Листинг кода 2



Рис. 2. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,02 нулевую гипотезу принимаем.

*Задание 3:*

Контрольную работу по высшей математике по индивидуальным вариантам выполняли студенты 4-х групп первого курса. В 1-ой группе было предложено 105 задач, из которых верно решено 60 задач. Во 2-ой группе всего задач 140 и верно решено 69. В третьей группе было 125 задач, верно решено 63. В 4-ой группе всего задач 160, верно решено 105.

На уровне значимости выяснить, можно ли считать, что различия в усвоении учебного материала студентами 4-х групп первого курса существенны.

*Решение:*

Выдвигаем гипотезу .

Для данной задачи нами была написана программа, которая вычисляет выборочную долю признака по формуле (11) и по формуле (7), а также статистику по формуле (10).

Если , то гипотеза H0 отвергается, в противном случае - принимается. В нашей задаче нулевая гипотеза отвергается, т.е. различие в усвоении материала студентами 4-х групп существенно на уровне значимости .

| def get\_xalpha(x, alpha, k):  from pars\_data import data  return data[alpha][k-1]  pass  n = (105, 140, 125, 160) m = (60, 69, 63, 105) alpha = 0.05 k = len(n) - 1 p = sum(m) / sum(n) w = list(map(lambda m, n: m / n, m, n))  x = (1 / (p \* (1 - p)) \* sum([n[i] \* (w[i] - p) \*\* 2 for i in range(len(n))])) xalpha = get\_xalpha(x, alpha, k)  print(f'Так как {x} {[">", "<"][x < xalpha]} {xalpha}, то на уровне alpha = {alpha} {["гипотеза H0 принимается", "Гипотеза H0 отвергается"][x > xalpha]}') |
| --- |

Листинг кода 3



Рис. 3. Результат выполнения программы

Также дополнительно был написан код для анализа данных таблицы, благодаря которому мы автоматически получаем критическое значение статистики.

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

***Приложение 1.***

Предлагаем вам ознакомиться с исходным кодом для решения задач:

<https://github.com/Cplusx3/AD/tree/main/%D0%9B%D0%B0%D0%B1%D0%BE%D1%80%D0%B0%D1%82%D0%BE%D1%80%D0%BD%D0%B0%D1%8F%20%D1%80%D0%B0%D0%B1%D0%BE%D1%82%D0%B0%206%20%D1%87%D0%B0%D1%81%D1%82%D1%8C%203>

Также данные задачи были решены в Excel’e:

<https://docs.google.com/spreadsheets/d/1FnuBI9cT3sbQDu8lBP2NaQHWdoJwPXFd-aV--i8trE4/edit?usp=sharing>