

РГПУ им. А.И. Герцена

Отчет по лабораторной работе №6

Часть 2

«Проверка статистических гипотез»

Работу выполнили:

Клементьев Алексей

Воложанин Владислав

Лотуга Данила

Сафин Рамаз

Санкт-Петербург
2022

Цель работы: проверить статистические гипотезы о нормальном законе распределения данных, приведенных в решаемых задачах.

Оборудование: ПК, Python.

Задача 1

По результатам $n = 9$ замеров установлено, что выборочное среднее время (в секундах) изготовления детали $\bar{x} = 48$. Предполагая, что время изготовления – нормально распределённая случайная величина с дисперсией $\sigma^2 = 9$, рассмотреть на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean):
    t_score = (avg - null_hypothesis_mean) / (deviation / math.sqrt(n))

    df = n - 1

    t_critical = t.ppf(alpha / 2, df)

    if abs(t_score) < t_critical:
        return f"Отклоняем нулевую гипотезу H0: a = {null_hypothesis_mean}"
    else:
        return f"Не можем отклонить нулевую гипотезу H0: a = {null_hypothesis_mean}"

def main():
    avg = 48
    deviation = 3
    n = 9
    alpha = 0.95
    null_hypothesis_mean = 49

    result = hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Fail to reject the null hypothesis H0: a = 49
```

В данном коде мы вычисляем t-статистику, определяем количество степеней свободы, находим критическое значение t-статистики для двухсторонней гипотезы на заданном уровне значимости и сравниваем полученную t-статистику со значением критической t-статистики. Если значение тестовой статистики находится вне диапазона, заданного критическими значениями, мы отвергаем нулевую гипотезу на данном уровне значимости.

Задача 2

Руководство фирмы утверждает, что размер дебиторского счёта равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счётов и обнаруживает, что средняя арифметическая выборки равна 175 тыс. руб. при среднем квадратическом отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счёта? Принять уровень значимости равным $\alpha = 0,05$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(mu, avg, s, n, alpha=0.05):
    t_stat = (avg - mu) / (s / math.sqrt(n))

    p_value = 2 * (1 - t.cdf(abs(t_stat), n - 1))

    if p_value < alpha:
        return f"Отвергаем нулевую гипотезу H0 = {mu}"
    else:
        return f"Не можем отвергнуть нулевую гипотезу H0 = {mu}"

def main():
    mu = 187.5
    x_bar = 175
    s = 35
    n = 10
    alpha = 0.05

    result = hypothesis_test(mu, x_bar, s, n, alpha)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Не можем отвергнуть нулевую гипотезу H0 = 187.5
```

Применим двусторонний t-тест Стьюдента с уровнем значимости $\alpha = 0,05$ для проверки гипотезы о равенстве выборочного среднего и известного среднего значения.

Вывод программы будет указывать, отвергаем мы нулевую гипотезу или нет. Если размер дебиторского счета может отличаться от утвержденного значения, то может потребоваться дополнительный анализ для выявления действительного размера.

Задача 3

Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии $\overline{S^2} = 0,25$. При уровне значимости $\alpha = 0,1$ выяснить, обеспечивает ли станок требуемую точность.

Код программы:

```
from scipy.stats import chi2

def machine_accuracy(avg_variance, n, alpha, variance):
    chi2_stat = (n - 1) * avg_variance / variance

    chi2_crit = chi2.ppf(1 - alpha, n - 1)

    if chi2_stat < chi2_crit:
        return f"Не может быть отвергнута нулевая гипотеза H0 = {variance}"
    else:
        return f"Отвергается нулевая гипотеза H0 = {variance}"

def main():
    avg_variance = 0.25
    n = 25
    alpha = 0.1
    variance = 0.15

    result = machine_accuracy(avg_variance, n, alpha, variance)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Отвергается нулевая гипотеза H0 = 0.15
```

В этой задаче мы можем использовать критерий хи-квадрат для оценки дисперсии. Если нулевая гипотеза будет отвергнута, то это означает, что станок не обеспечивает требуемую точность.

Задача 4

Расходы сырья x_i и y_j на единицу продукции по старой и новой технологиям приведены в таблице:

	По старой технологии				По новой технологии				
Расходы сырья	x	304	307	308	y	303	304	306	308
Число изделий	nx	1	4	4	ny	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними a_1 и a_2 . Требуется проверить гипотезу $H_0: a_1 = a_2$ против гипотезы $H_1: a_1 \neq a_2$ на уровне значимости $\alpha = 0,1$.

Код программы:

```
import numpy as np
from scipy.stats import t

def hypothesis_test(x, y, nx, ny, alpha=0.1):
    x_mean = np.sum(x * nx) / np.sum(nx)
    y_mean = np.sum(y * ny) / np.sum(ny)

    s2 = (np.sum(nx * (x - x_mean) ** 2) + np.sum(ny * (y - y_mean) ** 2)) /
        (np.sum(nx) + np.sum(ny) - 2)

    t_stat = (x_mean - y_mean) / np.sqrt(s2 * (1 / np.sum(nx) + 1 /
        np.sum(ny)))

    t_critical = t.ppf(1 - alpha / 2, df=np.sum(nx) + np.sum(ny) - 2)

    if np.abs(t_stat) > t_critical:
        return "Отвергаем нулевую гипотезу H0 "
    else:
        return "Не отвергаем нулевую гипотезу H0 "

def main():
    x = np.array([304, 307, 308])
    y = np.array([303, 304, 306, 308])
    nx = np.array([1, 4, 4])
    ny = np.array([2, 6, 4, 1])

    result = hypothesis_test(x, y, nx, ny, alpha=0.1)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

Отвергаем нулевую гипотезу

Нулевая гипотеза (H_0) в данной задаче состоит в том, что средние значения генеральных совокупностей X и Y равны между собой, то есть $\mu_1 = \mu_2$. Средние значения генеральных совокупностей X и Y неизвестны, поскольку мы имеем только выборки из этих совокупностей. В данной задаче мы проводим статистический тест для проверки гипотезы о равенстве средних значений двух генеральных совокупностей на основе выборочных средних значений, поэтому в функции `hypothesis_test()` мы вычисляем выборочные средние значения `x_mean` и `y_mean`.

РГПУ им. А.И. Герцена

Отчет по лабораторной работе №6

Часть 2

«Проверка статистических гипотез»

Работу выполнил:

Клементьев Алексей

Санкт-Петербург
2022

Цель работы: проверить статистические гипотезы о нормальном законе распределения данных, приведенных в решаемых задачах.

Оборудование: ПК, Python.

Задача 1

По результатам $n = 9$ замеров установлено, что выборочное среднее время (в секундах) изготовления детали $\bar{x} = 48$. Предполагая, что время изготовления – нормально распределённая случайная величина с дисперсией $\sigma^2 = 9$, рассмотреть на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean):
    t_score = (avg - null_hypothesis_mean) / (deviation / math.sqrt(n))

    df = n - 1

    t_critical = t.ppf(alpha / 2, df)

    if abs(t_score) < t_critical:
        return f"Отклоняем нулевую гипотезу H0: a = {null_hypothesis_mean}"
    else:
        return f"Не можем отклонить нулевую гипотезу H0: a = {null_hypothesis_mean}"

def main():
    avg = 48
    deviation = 3
    n = 9
    alpha = 0.95
    null_hypothesis_mean = 49

    result = hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Fail to reject the null hypothesis H0: a = 49
```

В данном коде мы вычисляем t-статистику, определяем количество степеней свободы, находим критическое значение t-статистики для двухсторонней гипотезы на заданном уровне значимости и сравниваем полученную t-статистику со значением критической t-статистики. Если значение тестовой статистики находится вне диапазона, заданного критическими значениями, мы отвергаем нулевую гипотезу на данном уровне значимости.

Задача 2

Руководство фирмы утверждает, что размер дебиторского счёта равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счётов и обнаруживает, что средняя арифметическая выборки равна 175 тыс. руб. при среднем квадратическом отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счёта? Принять уровень значимости равным $\alpha = 0,05$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(mu, avg, s, n, alpha=0.05):
    t_stat = (avg - mu) / (s / math.sqrt(n))

    p_value = 2 * (1 - t.cdf(abs(t_stat), n - 1))

    if p_value < alpha:
        return f"Отвергаем нулевую гипотезу H0 = {mu}"
    else:
        return f"Не можем отвергнуть нулевую гипотезу H0 = {mu}"

def main():
    mu = 187.5
    x_bar = 175
    s = 35
    n = 10
    alpha = 0.05

    result = hypothesis_test(mu, x_bar, s, n, alpha)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Не можем отвергнуть нулевую гипотезу H0 = 187.5
```

Применим двусторонний t-тест Стьюдента с уровнем значимости $\alpha = 0,05$ для проверки гипотезы о равенстве выборочного среднего и известного среднего значения.

Вывод программы будет указывать, отвергаем мы нулевую гипотезу или нет. Если размер дебиторского счета может отличаться от утвержденного значения, то может потребоваться дополнительный анализ для выявления действительного размера.

Задача 3

Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии $\overline{S^2} = 0,25$. При уровне значимости $\alpha = 0,1$ выяснить, обеспечивает ли станок требуемую точность.

Код программы:

```
from scipy.stats import chi2

def machine_accuracy(avg_variance, n, alpha, variance):
    chi2_stat = (n - 1) * avg_variance / variance

    chi2_crit = chi2.ppf(1 - alpha, n - 1)

    if chi2_stat < chi2_crit:
        return f"Не может быть отвергнута нулевая гипотеза H0 = {variance}"
    else:
        return f"Отвергается нулевая гипотеза H0 = {variance}"

def main():
    avg_variance = 0.25
    n = 25
    alpha = 0.1
    variance = 0.15

    result = machine_accuracy(avg_variance, n, alpha, variance)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Отвергается нулевая гипотеза H0 = 0.15
```

В этой задаче мы можем использовать критерий хи-квадрат для оценки дисперсии. Если нулевая гипотеза будет отвергнута, то это означает, что станок не обеспечивает требуемую точность.

Задача 4

Расходы сырья x_i и y_j на единицу продукции по старой и новой технологиям приведены в таблице:

	По старой технологии				По новой технологии				
Расходы сырья	x	304	307	308	y	303	304	306	308
Число изделий	nx	1	4	4	ny	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними a_1 и a_2 . Требуется проверить гипотезу $H_0: a_1 = a_2$ против гипотезы $H_1: a_1 \neq a_2$ на уровне значимости $\alpha = 0,1$.

Код программы:

```
import numpy as np
from scipy.stats import t

def hypothesis_test(x, y, nx, ny, alpha=0.1):
    x_mean = np.sum(x * nx) / np.sum(nx)
    y_mean = np.sum(y * ny) / np.sum(ny)

    s2 = (np.sum(nx * (x - x_mean) ** 2) + np.sum(ny * (y - y_mean) ** 2)) /
        (np.sum(nx) + np.sum(ny) - 2)

    t_stat = (x_mean - y_mean) / np.sqrt(s2 * (1 / np.sum(nx) + 1 /
        np.sum(ny)))

    t_critical = t.ppf(1 - alpha / 2, df=np.sum(nx) + np.sum(ny) - 2)

    if np.abs(t_stat) > t_critical:
        return "Отвергаем нулевую гипотезу H0 "
    else:
        return "Не отвергаем нулевую гипотезу H0 "

def main():
    x = np.array([304, 307, 308])
    y = np.array([303, 304, 306, 308])
    nx = np.array([1, 4, 4])
    ny = np.array([2, 6, 4, 1])

    result = hypothesis_test(x, y, nx, ny, alpha=0.1)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

Отвергаем нулевую гипотезу

Нулевая гипотеза (H_0) в данной задаче состоит в том, что средние значения генеральных совокупностей X и Y равны между собой, то есть $\mu_1 = \mu_2$. Средние значения генеральных совокупностей X и Y неизвестны, поскольку мы имеем только выборки из этих совокупностей. В данной задаче мы проводим статистический тест для проверки гипотезы о равенстве средних значений двух генеральных совокупностей на основе выборочных средних значений, поэтому в функции `hypothesis_test()` мы вычисляем выборочные средние значения `x_mean` и `y_mean`.

РГПУ им. А.И. Герцена

Отчет по лабораторной работе №6

Часть 2

«Проверка статистических гипотез»

Работу выполнил:

Воложанин Владислав

Санкт-Петербург
2022

Цель работы: проверить статистические гипотезы о нормальном законе распределения данных, приведенных в решаемых задачах.

Оборудование: ПК, Python.

Задача 1

По результатам $n = 9$ замеров установлено, что выборочное среднее время (в секундах) изготовления детали $\bar{x} = 48$. Предполагая, что время изготовления – нормально распределённая случайная величина с дисперсией $\sigma^2 = 9$, рассмотреть на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean):
    t_score = (avg - null_hypothesis_mean) / (deviation / math.sqrt(n))

    df = n - 1

    t_critical = t.ppf(alpha / 2, df)

    if abs(t_score) < t_critical:
        return f"Отклоняем нулевую гипотезу H0: a = {null_hypothesis_mean}"
    else:
        return f"Не можем отклонить нулевую гипотезу H0: a = {null_hypothesis_mean}"

def main():
    avg = 48
    deviation = 3
    n = 9
    alpha = 0.95
    null_hypothesis_mean = 49

    result = hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Fail to reject the null hypothesis H0: a = 49
```

В данном коде мы вычисляем t-статистику, определяем количество степеней свободы, находим критическое значение t-статистики для двухсторонней гипотезы на заданном уровне значимости и сравниваем полученную t-статистику со значением критической t-статистики. Если значение тестовой статистики находится вне диапазона, заданного критическими значениями, мы отвергаем нулевую гипотезу на данном уровне значимости.

Задача 2

Руководство фирмы утверждает, что размер дебиторского счёта равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счётов и обнаруживает, что средняя арифметическая выборки равна 175 тыс. руб. при среднем квадратическом отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счёта? Принять уровень значимости равным $\alpha = 0,05$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(mu, avg, s, n, alpha=0.05):
    t_stat = (avg - mu) / (s / math.sqrt(n))

    p_value = 2 * (1 - t.cdf(abs(t_stat), n - 1))

    if p_value < alpha:
        return f"Отвергаем нулевую гипотезу H0 = {mu}"
    else:
        return f"Не можем отвергнуть нулевую гипотезу H0 = {mu}"

def main():
    mu = 187.5
    x_bar = 175
    s = 35
    n = 10
    alpha = 0.05

    result = hypothesis_test(mu, x_bar, s, n, alpha)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Не можем отвергнуть нулевую гипотезу H0 = 187.5
```

Применим двусторонний t-тест Стьюдента с уровнем значимости $\alpha = 0,05$ для проверки гипотезы о равенстве выборочного среднего и известного среднего значения.

Вывод программы будет указывать, отвергаем мы нулевую гипотезу или нет. Если размер дебиторского счета может отличаться от утвержденного значения, то может потребоваться дополнительный анализ для выявления действительного размера.

Задача 3

Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии $\overline{S^2} = 0,25$. При уровне значимости $\alpha = 0,1$ выяснить, обеспечивает ли станок требуемую точность.

Код программы:

```
from scipy.stats import chi2

def machine_accuracy(avg_variance, n, alpha, variance):
    chi2_stat = (n - 1) * avg_variance / variance

    chi2_crit = chi2.ppf(1 - alpha, n - 1)

    if chi2_stat < chi2_crit:
        return f"Не может быть отвергнута нулевая гипотеза H0 = {variance}"
    else:
        return f"Отвергается нулевая гипотеза H0 = {variance}"

def main():
    avg_variance = 0.25
    n = 25
    alpha = 0.1
    variance = 0.15

    result = machine_accuracy(avg_variance, n, alpha, variance)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Отвергается нулевая гипотеза H0 = 0.15
```

В этой задаче мы можем использовать критерий хи-квадрат для оценки дисперсии. Если нулевая гипотеза будет отвергнута, то это означает, что станок не обеспечивает требуемую точность.

Задача 4

Расходы сырья x_i и y_j на единицу продукции по старой и новой технологиям приведены в таблице:

	По старой технологии				По новой технологии				
Расходы сырья	x	304	307	308	y	303	304	306	308
Число изделий	nx	1	4	4	ny	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними a_1 и a_2 . Требуется проверить гипотезу $H_0: a_1 = a_2$ против гипотезы $H_1: a_1 \neq a_2$ на уровне значимости $\alpha = 0,1$.

Код программы:

```
import numpy as np
from scipy.stats import t

def hypothesis_test(x, y, nx, ny, alpha=0.1):
    x_mean = np.sum(x * nx) / np.sum(nx)
    y_mean = np.sum(y * ny) / np.sum(ny)

    s2 = (np.sum(nx * (x - x_mean) ** 2) + np.sum(ny * (y - y_mean) ** 2)) /
        (np.sum(nx) + np.sum(ny) - 2)

    t_stat = (x_mean - y_mean) / np.sqrt(s2 * (1 / np.sum(nx) + 1 /
        np.sum(ny)))

    t_critical = t.ppf(1 - alpha / 2, df=np.sum(nx) + np.sum(ny) - 2)

    if np.abs(t_stat) > t_critical:
        return "Отвергаем нулевую гипотезу H0 "
    else:
        return "Не отвергаем нулевую гипотезу H0 "

def main():
    x = np.array([304, 307, 308])
    y = np.array([303, 304, 306, 308])
    nx = np.array([1, 4, 4])
    ny = np.array([2, 6, 4, 1])

    result = hypothesis_test(x, y, nx, ny, alpha=0.1)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

Отвергаем нулевую гипотезу

Нулевая гипотеза (H_0) в данной задаче состоит в том, что средние значения генеральных совокупностей X и Y равны между собой, то есть $\mu_1 = \mu_2$. Средние значения генеральных совокупностей X и Y неизвестны, поскольку мы имеем только выборки из этих совокупностей. В данной задаче мы проводим статистический тест для проверки гипотезы о равенстве средних значений двух генеральных совокупностей на основе выборочных средних значений, поэтому в функции `hypothesis_test()` мы вычисляем выборочные средние значения `x_mean` и `y_mean`.

РГПУ им. А.И. Герцена

Отчет по лабораторной работе №6

Часть 2

«Проверка статистических гипотез»

Работу выполнил:

Лотуга Данила

Санкт-Петербург
2022

Цель работы: проверить статистические гипотезы о нормальном законе распределения данных, приведенных в решаемых задачах.

Оборудование: ПК, Python.

Задача 1

По результатам $n = 9$ замеров установлено, что выборочное среднее время (в секундах) изготовления детали $\bar{x} = 48$. Предполагая, что время изготовления – нормально распределённая случайная величина с дисперсией $\sigma^2 = 9$, рассмотреть на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean):
    t_score = (avg - null_hypothesis_mean) / (deviation / math.sqrt(n))

    df = n - 1

    t_critical = t.ppf(alpha / 2, df)

    if abs(t_score) < t_critical:
        return f"Отклоняем нулевую гипотезу H0: a = {null_hypothesis_mean}"
    else:
        return f"Не можем отклонить нулевую гипотезу H0: a = {null_hypothesis_mean}"

def main():
    avg = 48
    deviation = 3
    n = 9
    alpha = 0.95
    null_hypothesis_mean = 49

    result = hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Fail to reject the null hypothesis H0: a = 49
```

В данном коде мы вычисляем t-статистику, определяем количество степеней свободы, находим критическое значение t-статистики для двухсторонней гипотезы на заданном уровне значимости и сравниваем полученную t-статистику со значением критической t-статистики. Если значение тестовой статистики находится вне диапазона, заданного критическими значениями, мы отвергаем нулевую гипотезу на данном уровне значимости.

Задача 2

Руководство фирмы утверждает, что размер дебиторского счёта равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счётов и обнаруживает, что средняя арифметическая выборки равна 175 тыс. руб. при среднем квадратическом отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счёта? Принять уровень значимости равным $\alpha = 0,05$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(mu, avg, s, n, alpha=0.05):
    t_stat = (avg - mu) / (s / math.sqrt(n))

    p_value = 2 * (1 - t.cdf(abs(t_stat), n - 1))

    if p_value < alpha:
        return f"Отвергаем нулевую гипотезу H0 = {mu}"
    else:
        return f"Не можем отвергнуть нулевую гипотезу H0 = {mu}"

def main():
    mu = 187.5
    x_bar = 175
    s = 35
    n = 10
    alpha = 0.05

    result = hypothesis_test(mu, x_bar, s, n, alpha)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Не можем отвергнуть нулевую гипотезу H0 = 187.5
```

Применим двусторонний t-тест Стьюдента с уровнем значимости $\alpha = 0,05$ для проверки гипотезы о равенстве выборочного среднего и известного среднего значения.

Вывод программы будет указывать, отвергаем мы нулевую гипотезу или нет. Если размер дебиторского счета может отличаться от утвержденного значения, то может потребоваться дополнительный анализ для выявления действительного размера.

Задача 3

Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии $\overline{S^2} = 0,25$. При уровне значимости $\alpha = 0,1$ выяснить, обеспечивает ли станок требуемую точность.

Код программы:

```
from scipy.stats import chi2

def machine_accuracy(avg_variance, n, alpha, variance):
    chi2_stat = (n - 1) * avg_variance / variance

    chi2_crit = chi2.ppf(1 - alpha, n - 1)

    if chi2_stat < chi2_crit:
        return f"Не может быть отвергнута нулевая гипотеза H0 = {variance}"
    else:
        return f"Отвергается нулевая гипотеза H0 = {variance}"

def main():
    avg_variance = 0.25
    n = 25
    alpha = 0.1
    variance = 0.15

    result = machine_accuracy(avg_variance, n, alpha, variance)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Отвергается нулевая гипотеза H0 = 0.15
```

В этой задаче мы можем использовать критерий хи-квадрат для оценки дисперсии. Если нулевая гипотеза будет отвергнута, то это означает, что станок не обеспечивает требуемую точность.

Задача 4

Расходы сырья x_i и y_j на единицу продукции по старой и новой технологиям приведены в таблице:

	По старой технологии				По новой технологии				
Расходы сырья	x	304	307	308	y	303	304	306	308
Число изделий	nx	1	4	4	ny	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними a_1 и a_2 . Требуется проверить гипотезу $H_0: a_1 = a_2$ против гипотезы $H_1: a_1 \neq a_2$ на уровне значимости $\alpha = 0,1$.

Код программы:

```
import numpy as np
from scipy.stats import t

def hypothesis_test(x, y, nx, ny, alpha=0.1):
    x_mean = np.sum(x * nx) / np.sum(nx)
    y_mean = np.sum(y * ny) / np.sum(ny)

    s2 = (np.sum(nx * (x - x_mean) ** 2) + np.sum(ny * (y - y_mean) ** 2)) /
        (np.sum(nx) + np.sum(ny) - 2)

    t_stat = (x_mean - y_mean) / np.sqrt(s2 * (1 / np.sum(nx) + 1 /
        np.sum(ny)))

    t_critical = t.ppf(1 - alpha / 2, df=np.sum(nx) + np.sum(ny) - 2)

    if np.abs(t_stat) > t_critical:
        return "Отвергаем нулевую гипотезу H0 "
    else:
        return "Не отвергаем нулевую гипотезу H0 "

def main():
    x = np.array([304, 307, 308])
    y = np.array([303, 304, 306, 308])
    nx = np.array([1, 4, 4])
    ny = np.array([2, 6, 4, 1])

    result = hypothesis_test(x, y, nx, ny, alpha=0.1)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

Отвергаем нулевую гипотезу

Нулевая гипотеза (H_0) в данной задаче состоит в том, что средние значения генеральных совокупностей X и Y равны между собой, то есть $\mu_1 = \mu_2$. Средние значения генеральных совокупностей X и Y неизвестны, поскольку мы имеем только выборки из этих совокупностей. В данной задаче мы проводим статистический тест для проверки гипотезы о равенстве средних значений двух генеральных совокупностей на основе выборочных средних значений, поэтому в функции `hypothesis_test()` мы вычисляем выборочные средние значения `x_mean` и `y_mean`.

РГПУ им. А.И. Герцена

Отчет по лабораторной работе №6

Часть 2

«Проверка статистических гипотез»

Работу выполнил:

Сафин Рамаз

Санкт-Петербург
2022

Цель работы: проверить статистические гипотезы о нормальном законе распределения данных, приведенных в решаемых задачах.

Оборудование: ПК, Python.

Задача 1

По результатам $n = 9$ замеров установлено, что выборочное среднее время (в секундах) изготовления детали $\bar{x} = 48$. Предполагая, что время изготовления – нормально распределённая случайная величина с дисперсией $\sigma^2 = 9$, рассмотреть на уровне 0,95 гипотезу $H_0: a = 49$, против конкурирующей гипотезы $H_1: a \neq 49$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean):
    t_score = (avg - null_hypothesis_mean) / (deviation / math.sqrt(n))

    df = n - 1

    t_critical = t.ppf(alpha / 2, df)

    if abs(t_score) < t_critical:
        return f"Отклоняем нулевую гипотезу H0: a = {null_hypothesis_mean}"
    else:
        return f"Не можем отклонить нулевую гипотезу H0: a = {null_hypothesis_mean}"

def main():
    avg = 48
    deviation = 3
    n = 9
    alpha = 0.95
    null_hypothesis_mean = 49

    result = hypothesis_test(avg, deviation, n, alpha, null_hypothesis_mean)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Fail to reject the null hypothesis H0: a = 49
```

В данном коде мы вычисляем t-статистику, определяем количество степеней свободы, находим критическое значение t-статистики для двухсторонней гипотезы на заданном уровне значимости и сравниваем полученную t-статистику со значением критической t-статистики. Если значение тестовой статистики находится вне диапазона, заданного критическими значениями, мы отвергаем нулевую гипотезу на данном уровне значимости.

Задача 2

Руководство фирмы утверждает, что размер дебиторского счёта равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счётов и обнаруживает, что средняя арифметическая выборки равна 175 тыс. руб. при среднем квадратическом отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счёта? Принять уровень значимости равным $\alpha = 0,05$.

Код программы:

```
import math
from scipy.stats import t

def hypothesis_test(mu, avg, s, n, alpha=0.05):
    t_stat = (avg - mu) / (s / math.sqrt(n))

    p_value = 2 * (1 - t.cdf(abs(t_stat), n - 1))

    if p_value < alpha:
        return f"Отвергаем нулевую гипотезу H0 = {mu}"
    else:
        return f"Не можем отвергнуть нулевую гипотезу H0 = {mu}"

def main():
    mu = 187.5
    x_bar = 175
    s = 35
    n = 10
    alpha = 0.05

    result = hypothesis_test(mu, x_bar, s, n, alpha)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Не можем отвергнуть нулевую гипотезу H0 = 187.5
```

Применим двусторонний t-тест Стьюдента с уровнем значимости $\alpha = 0,05$ для проверки гипотезы о равенстве выборочного среднего и известного среднего значения.

Вывод программы будет указывать, отвергаем мы нулевую гипотезу или нет. Если размер дебиторского счета может отличаться от утвержденного значения, то может потребоваться дополнительный анализ для выявления действительного размера.

Задача 3

Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена оценка дисперсии $\overline{S^2} = 0,25$. При уровне значимости $\alpha = 0,1$ выяснить, обеспечивает ли станок требуемую точность.

Код программы:

```
from scipy.stats import chi2

def machine_accuracy(avg_variance, n, alpha, variance):
    chi2_stat = (n - 1) * avg_variance / variance

    chi2_crit = chi2.ppf(1 - alpha, n - 1)

    if chi2_stat < chi2_crit:
        return f"Не может быть отвергнута нулевая гипотеза H0 = {variance}"
    else:
        return f"Отвергается нулевая гипотеза H0 = {variance}"

def main():
    avg_variance = 0.25
    n = 25
    alpha = 0.1
    variance = 0.15

    result = machine_accuracy(avg_variance, n, alpha, variance)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

```
Отвергается нулевая гипотеза H0 = 0.15
```

В этой задаче мы можем использовать критерий хи-квадрат для оценки дисперсии. Если нулевая гипотеза будет отвергнута, то это означает, что станок не обеспечивает требуемую точность.

Задача 4

Расходы сырья x_i и y_j на единицу продукции по старой и новой технологиям приведены в таблице:

	По старой технологии				По новой технологии				
Расходы сырья	x	304	307	308	y	303	304	306	308
Число изделий	nx	1	4	4	ny	2	6	4	1

Предполагается, что генеральные совокупности X и Y имеют нормальные распределения с одинаковыми дисперсиями и средними a_1 и a_2 . Требуется проверить гипотезу $H_0: a_1 = a_2$ против гипотезы $H_1: a_1 \neq a_2$ на уровне значимости $\alpha = 0,1$.

Код программы:

```
import numpy as np
from scipy.stats import t

def hypothesis_test(x, y, nx, ny, alpha=0.1):
    x_mean = np.sum(x * nx) / np.sum(nx)
    y_mean = np.sum(y * ny) / np.sum(ny)

    s2 = (np.sum(nx * (x - x_mean) ** 2) + np.sum(ny * (y - y_mean) ** 2)) /
        (np.sum(nx) + np.sum(ny) - 2)

    t_stat = (x_mean - y_mean) / np.sqrt(s2 * (1 / np.sum(nx) + 1 /
        np.sum(ny)))

    t_critical = t.ppf(1 - alpha / 2, df=np.sum(nx) + np.sum(ny) - 2)

    if np.abs(t_stat) > t_critical:
        return "Отвергаем нулевую гипотезу H0 "
    else:
        return "Не отвергаем нулевую гипотезу H0 "

def main():
    x = np.array([304, 307, 308])
    y = np.array([303, 304, 306, 308])
    nx = np.array([1, 4, 4])
    ny = np.array([2, 6, 4, 1])

    result = hypothesis_test(x, y, nx, ny, alpha=0.1)
    print(result)

if __name__ == "__main__":
    main()
```

Результат работы программы:

Отвергаем нулевую гипотезу

Нулевая гипотеза (H_0) в данной задаче состоит в том, что средние значения генеральных совокупностей X и Y равны между собой, то есть $\mu_1 = \mu_2$. Средние значения генеральных совокупностей X и Y неизвестны, поскольку мы имеем только выборки из этих совокупностей. В данной задаче мы проводим статистический тест для проверки гипотезы о равенстве средних значений двух генеральных совокупностей на основе выборочных средних значений, поэтому в функции `hypothesis_test()` мы вычисляем выборочные средние значения `x_mean` и `y_mean`.