TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO CUỐI KỲ MÔN XÁC SUẤT THỐNG KÊ ỨNG DỤNG CHO CÔNG NGHỆ THÔNG TIN**

*Người hướng dẫn*: TS NGUYỄN QUỐC BÌNH

*Người thực hiện*: **BÙI PHƯƠNG NAM – 522H0002**

Lớp **: 22H50201**

Khoá  **: K26**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2014**

LỜI CẢM ƠN

Lời đầu tiên, em xin gửi lời cảm ơn đến Thầy Nguyễn Quốc Bình. Trong quá trình học cũng như là làm bài tiểu luận, nhờ có sự giúp đỡ tận tình của Thầy, Thầy đã truyền đạt cho chúng em những kiến thức bổ ích. Thầy đã hướng dẫn chúng em nhiệt tình, nhờ đó mà chúng em có thể hoàn thành báo cáo này.XÁC SUẤT THỐNG KÊ ỨNG DỤNG CHO CÔNG NGHỆ THÔNG TIN là môn học thú vị và nhưng cũng rất phức tạp. Những kiến thức và kỹ năng về môn học này của em vẫn còn nhiều hạn chế. Do đó, bài tiểu luận của em khó tránh khỏi sai sót. Rất mong nhận được những góp ý của Thầy về bài tiểu luận này.

Một lần nữa, em xin chân thành cảm ơn! .

**BÀI TẬP ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là sản phẩm đồ án của riêng tôi và được sự hướng dẫn của TS NGUYỄN QUỐC BÌNH. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*Bùi Phương Nam*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

**Phần xác nhận của GV hướng dẫn**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

**Phần đánh giá của GV chấm bài**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

TÓM TẮT

Bài tiểu luận này tập trung vào khảo sát mức lương khởi điểm của sinh viên mới ra trường trong một lớp học gồm 65 sinh viên. Dữ liệu được thu thập và biểu diễn qua bảng số liệu với các thông tin về giới tính và mức lương khởi điểm của từng sinh viên.

I) Phân tích dữ liệu tổng thể:

* Tính giá trị trung bình và độ lệch chuẩn của mức lương khởi điểm.
* Tính tỉ lệ nữ và nam có mức lương cao hơn mức lương trung bình.
* Tính xác suất chọn ngẫu nhiên người có mức lương thấp hơn và cao hơn mức lương trung bình.
* Xác định số cách lấy mẫu S1 và S2.

II) Lấy mẫu và phân tích:

* Diễn giải phương pháp lấy ngẫu nhiên và điền kết quả vào bảng Google Sheets.
* Tính giá trị trung bình mức lương của tập S1, S2 và so sánh với tổng thể.
* Uớc lượng khoảng tin cậy đối với giá trị trung bình của S1, S2.

III) Phân tích về tương lai:

* Dự đoán tăng lương sau 10 năm làm việc và ảnh hưởng của điều này đối với giá trị trung bình, phương sai, và độ lệch chuẩn của mức lương.

Cuối cùng, bài tiểu luận sẽ kết luận với các nhận định về mức lương khởi điểm của sinh viên mới ra trường, cũng như đưa ra những suy luận về dữ liệu mẫu và ảnh hưởng của việc tăng lương trong tương lai

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc387692905)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN i](#_Toc387692906)

[TÓM TẮT i](#_Toc387692907)

[MỤC LỤC 1](#_Toc387692908)

[DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ 1](#_Toc387692909)

[CHƯƠNG 1 – MỞ ĐẦU 1](#_Toc387692910)

[1.1 Tiểu mục cấp 1 1](#_Toc387692911)

[1.1.1 Tiểu mục cấp 2 1](#_Toc387692912)

[1.1.1.1 Tiểu mục cấp 3 1](#_Toc387692913)

[1.1.1.2 Tiểu mục cấp 3 tiếp theo. 1](#_Toc387692914)

[1.1.2 Tiểu mục cấp 2 tiếp theo 1](#_Toc387692915)

[1.2 Nội dung của chương này 1](#_Toc387692916)

[CHƯƠNG 2 – TỔNG QUAN 1](#_Toc387692917)

[1.1 Trình bày công thức toán học 1](#_Toc387692918)

[1.2 Trình bày một hình vẽ, sơ đồ 1](#_Toc387692919)

[CHƯƠNG 3 – CƠ SỞ LÝ THUYẾT / NGHIÊN CỨU THỰC NGHIỆM 1](#_Toc387692920)

[3.1 Chèn bảng: 1](#_Toc387692921)

[3.2 Viết tắt 1](#_Toc387692922)

[3.3 Trích dẫn 1](#_Toc387692923)

[3.3.1 Tài liệu tham khảo và cách trích dẫn 1](#_Toc387692924)

[3.3.2 Qui định của Khoa Công nghệ thông tin 1](#_Toc387692925)

**Dữ liệu cho trước**

Trong một khảo sát mức lương khởi điểm của sinh viên vừa ra trường của một lớp học gồm 65 sinh viên, kết quả được đánh số thứ tự như bảng sau.

Đơn vị của mức lương là X00,000 đồng (Ví dụ 27 tức là 2,700,000 VNĐ)

Giới tính: W là nữ, M là nam.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| STT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Giới tính | W | M | M | W | M | M | W | M | M | W |
| Tiền lương | 27 | 61 | 52 | 69 | 88 | 85 | 79 | 99 | 77 | 165 |
| STT | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Giới tính | M | W | M | W | M | M | M | W | M | M |
| Tiền lương | 41 | 83 | 144 | 74 | 143 | 131 | 34 | 59 | 46 | 105 |
| STT | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Giới tính | M | W | W | W | M | M | W | M | M | W |
| Tiền lương | 61 | 118 | 114 | 138 | 24 | 67 | 130 | 56 | 99 | 125 |
| STT | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| Giới tính | M | M | W | W | M | M | W | M | M | M |
| Tiền lương | 87 | 30 | 119 | 40 | 25 | 44 | 123 | 45 | 25 | 94 |
| STT | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
| Giới tính | W | W | W | W | M | M | W | M | M | M |
| Tiền lương | 86 | 128 | 69 | 102 | 91 | 106 | 119 | 139 | 67 | 47 |
| STT | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
| Giới tính | W | W | W | M | M | M | W | W | W | M |
| Tiền lương | 62 | 92 | 124 | 31 | 49 | 68 | 109 | 138 | 105 | 84 |
| STT | 61 | 62 | 63 | 64 | 65 |  |  |  |  |  |
| Giới tính | M | M | W | W | M |  |  |  |  |  |
| Tiền lương | 86 | 66 | 128 | 146 | 59 |  |  |  |  |  |

1. **Yêu cầu**
2. **Tính giá trị trung bình và độ lệch chuẩn của mức lương khởi điểm**

Gọi x­1, x2,… x­65 lần lượt là các lương của 65 sinh viên tương ứng. N = 65 là tổng số sinh viên.

Giá trị trung bình của mức lương khởi điểm tính theo công thức:

Kết quả: 85.0307 85

Độ lệch chuẩn của mức lương được tính theo căn bặc hai của phương sai:

Kết quả: 36.6899 36.69

1. **Tính tỉ lệ nữ có mức lương cao hơn mức lương trung bình**

Khi thống kê số lượng nữ tham gia lao động, ta thu được kết quả:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| W | W | W | W | W |
| 27 | 69 | 79 | 165 | 83 |
| W | W | W | W | W |
| 61 | 118 | 114 | 138 | 99 |
| W | W | W | W | W |
| 86 | 128 | 123 | 138 | 105 |
| W | W | W | W | W |
| 62 | 92 | 69 | 102 | 119 |
| W | W | W | W | W |
| 128 | 146 | 124 | 109 | 40 |
| W | W | W |  |  |
| 74 | 59 | 119 |  |  |

Như vậy, có tổng cộng 28 nữ có mức lương như trên. Tiếp tục lọc ra các kết quả sao cho mức lương trung bình của người tham gia lao động là trên trung bình thì ta có kết quả:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| W | W | W | W | W |
| 165 | 118 | 114 | 138 | 125 |
| W | W | W | W | W |
| 86 | 128 | 102 | 119 | 123 |
| W | W | W | W | W |
| 124 | 109 | 138 | 105 | 146 |
| W | W | W | W |  |
| 130 | 92 | 128 | 119 |  |

Như vậy, lúc này lượng dữ liệu đã giảm xuống còn 19. Tỷ lệ cần tìm là:

1. **Tính tỉ lệ nam có mức lương cao hơn mức lương trung bình**

Tương tự cách tính tỷ lệ nữ, ta sẽ lọc ra danh sách người lao động là nam:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| M | M | M | M | M | M |
| 61 | 52 | 88 | 85 | 99 | 77 |
| M | M | M | M | M | M |
| 41 | 144 | 143 | 131 | 34 | 46 |
| M | M | M | M | M | M |
| 61 | 24 | 67 | 56 | 99 | 105 |
| M | M | M | M | M | M |
| 87 | 30 | 25 | 44 | 45 | 25 |
| M | M | M | M | M | M |
| 91 | 106 | 139 | 67 | 47 | 94 |
| M | M | M | M | M | M |
| 31 | 49 | 68 | 84 | 86 | 66 |
| M |  |  |  |  |  |
| 59 |  |  |  |  |  |

Tiếp tục lọc ra những người có mức lương trên trung bình, ta được:

|  |  |  |
| --- | --- | --- |
| M | M | M |
| 88 | 91 | 99 |
| M | M | M |
| 144 | 143 | 131 |
| M | M | M |
| 99 | 105 | 94 |
| M | M | M |
| 87 | 86 | 139 |
| M |  |
| 106 |  |

Như vậy, tổng số nam giới có mức lương trên trung bình là 13. Tỷ lệ cần tìm là:

1. **Bảng dữ liệu trên bao gồm một nhóm nam và một nhóm nữ. Chọn ngẫu nhiên một nhóm biết xác suất chọn là như nhau. Từ nhóm đã chọn, chọn ngẫu nhiên 1 người để phỏng vấn.** 
   1. **Tính xác suất chọn được người có mức lương thấp hơn mức lương trung bình.**

Đối với bài toán này, ta phân loại ra 2 nhóm:

* + - Nhóm nam: 37 người
    - Nhóm nữ: 28 người

Cách giải bài toán gần giống với yêu cầu 1 và 2. Chỉ khác ở chỗ, ta sẽ xét trên mẫu chứ không xét trên tổng thể

Chọn ngẫu nhiên nhóm nữ, ta có 19 cách chọn thỏa mức lương của người đó trên trung bình. Khi đó, xác suất để chọn ra một người có mức lương trên trung bình là:

* 1. **Tính xác suất chọn được một người có mức lương cao hơn mức lương trung bình biết rằng người đó là nam.**

Với trường hợp này, ta sẽ xét trên nhóm nam. Chọn ngẫu nhiên một người để phỏng vấn, ta có 13 cách chọn. Xác suất là:

1. **Gọi S1 là mẫu gồm 30 sinh viên và tập S2 là mẫu gồm 15 sinh viên. Có bao nhiêu cách lấy 1 tập mẫu S1? Có bao nhiêu cách lấy 1 tập mẫu S2?**

Số cách lấy mẫu S1 gồm 30 sinh viên từ 65 sinh viên ban đầu là tổ hợp của 65 chập 30, hay cách

Số cách lấy mẫu S2 gồm 15 sinh viên từ 65 sinh viên ban đầu là tổ hợp của 65 chập 15, hay cách

1. Diễn giải chi tiết phương pháp lấy ngẫu nhiên tập S1, S2 bằng cách chụp màn hình kết quả làm minh chứng

Phương pháp lấy ngẫu nhiên tập S1 được thực hiện bằng ngôn ngữ python. Cụ thể, ta sẽ dùng python để triển khai thuật toán sau:

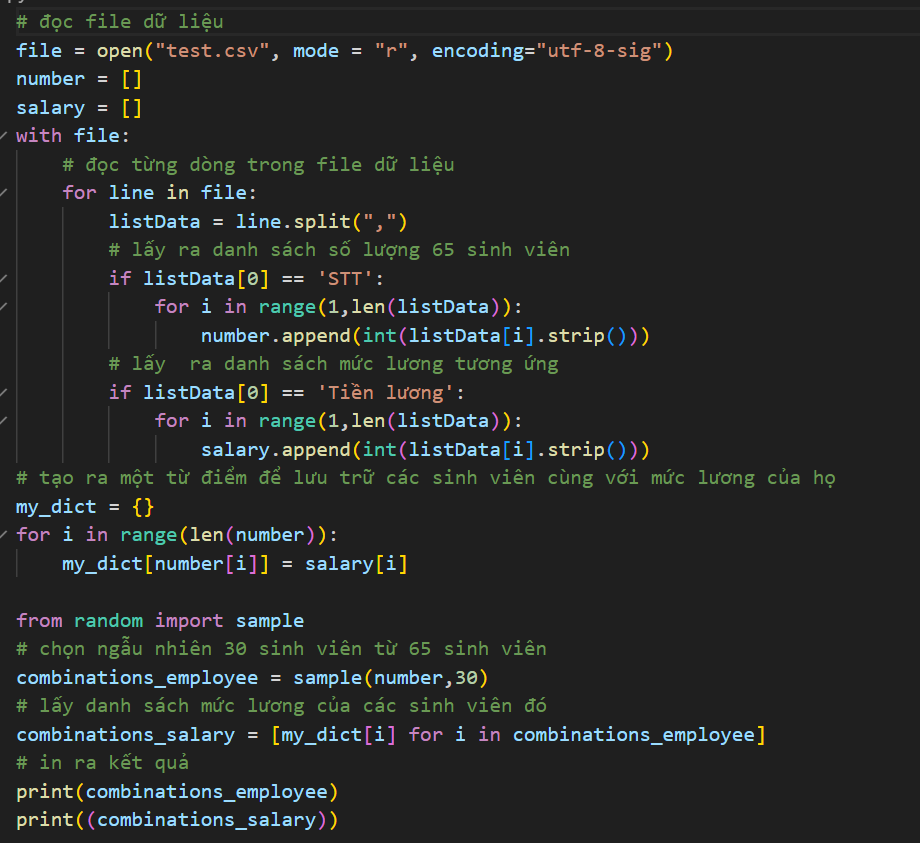
* + Đọc dữ liệu từ bảng trên

Lấy ra danh sách STT và danh sách Tiền Lương

* + Tạo một cấu trúc “từ điển” để lưu từng cặp dữ liệu K-V ứng với STT – Tiền Lương
  + Dùng random để lấy ra ngẫu nhiên 30 từ 65 STT của danh sách STT
  + Từ 30 STT đó, trả về 30 giá trị Tiền Lương tương ứng

Input: Bảng dữ liệu ban đầu đã cho

Code triển khai thuật toán :

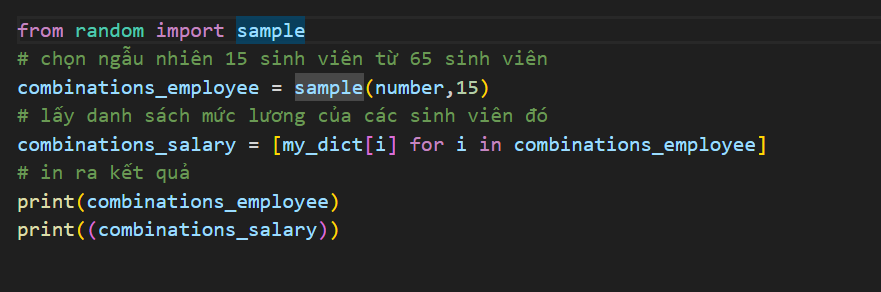


Output:

49, 105, 61, 34, 59, 143, 87, 40, 138, 128, 102, 91, 125, 25, 92,

68, 118, 66, 124, 144, 131, 130, 146, 59, 44, 138, 99, 123, 77, 88

Tương tự với S1, S2 cũng sử dụng thuật toán trên, chỉ khác ở chỗ ta thay đổi số lượng chọn ra là 15.



Output:

61, 74, 128, 106, 59, 45, 83, 69, 125, 25, 52, 94, 143, 119, 92

1. Tính giá trị trung bình mức lương khởi điểm của tập S1, S2. Bạn có kết luận gì về các giá trị này so với giá trị trung bình của tổng thể. (0.5 điểm)

Giá trị trung bình của S1 sẽ được tính bằng công thức ở yêu cầu 1 áp dụng lên kết quả ở yêu cầu 6. Kết quả: 94.46666666666667

Nhận xét: giá trị trung bình mẫu cao hơn so với tổng thể vì việc lựa chọn mẫu mang tính ngẫu nhiên. Trong trượng hợp trên, ta đã chon ra một mẫu mà tại đó số lượng người có mức lương trên trung bình nhiều hơn so với phần còn lại và điều đó làm cho trung bình mẫu cũng tăng theo

Tương tự, giá trị trung bình của S2 được ghị nhận là: 85.0

Nhận xét: Đây có vẻ là một sự trùng hợp khi bản thân nó bằng chính trung bình của tổng thể. Để có cái nhìn chính xác hơn, ta có thể tính thêm cả độ lệch chuẩn để đánh giá mẫu dữ liệu. Kết quả là: 34.29077343626333 – không chênh lêch quá 2.5 đợn vị so với độ lệch chuẩn của tổng thể (36.69). Điều này có nghĩa là độ phân tán của dữ liệu mẫu so với độ phân tán của dữ liệu gốc gần như giống nhau. Dù không chắc chắn nhưng ta có thể nhận định rằng mẫu mà ta chọn có phân phối dữ liệu gần với phân phối của tổng thể.

1. **Uớc lượng khoảng tin cậy đối của giá trị trung bình của tập S1, S2 với độ tin cậy là 80% và 95%. Bạn có kết luận gì về khoảng giá trị này so với giá trị trung bình của tổng thể và so với nhau.**

Giả sử phân phối của tổng thể là phân phối chuẩn N(, trong đó ta đã biết và là giá trị cần ước lượng khoảng

Xét tập S1 (n = 30):

* + Trung bình mẫu:
  + Gọi α là xác suất mà không rơi vào khoảng tin cậy. Theo đề bài, độ tin cậy là 100(1- α)% = 80% ⬄ α = 20

Bằng phương pháp tra bảng ngược ta được

* + Độ chính xác:

Khoảng tin cậy cần tìm ()

Xét tập S2 (n = 15):

* + Trung bình mẫu:
  + Gọi α là xác suất mà không rơi vào khoảng tin cậy. Theo đề bài, độ tin cậy là 100(1- α)% = 95% ⬄ α = 0.0 5

Bằng phương pháp tra bảng ngược ta được

* + Độ chính xác:

Khoảng tin cậy cần tìm (

Kết luận:

* Cả 2 khoảng tin cậy phù hợp với trung bình tổng thể.
* Giới hạn trên của cả 2 khoảng này khá gần nhau trong khi giới hạn dưới của chúng chênh lệch gần 19.4 đơn vị.

1. **Với mức ý nghĩa 0.05, hãy thực kiểm định giả thuyết trái, phải, hai bên của giá trị trung bình của S1 so với giá trị trung bình của tổng thể. Kết luận của kiểm định giả thuyết có đúng với thực tế hay không**?

Giả sử phân phối của tổng thể là phân phối chuẩn N(, trong đó ta đã biết và . Xét tập S1:

* + n = 30 =>
  + trung bình mẫu cần kiểm định:
  + α = 0.05 => α/2 = 0.025 => thống kê kiểm định: TS = =

|  |  |  |  |
| --- | --- | --- | --- |
|  | Kiểm định trái | Kiểm định phải | Kiểm định 2 bên |
| Giả thuyết |  |  |  |
| Miền bác bỏ  (chưa biết ) |  |  |  |
| Dữ liệu thực tế | => TS = 1.41 | | |
| Miền bác bỏ  (đã biết ) | {1.413 < -1.65} (vô lý) | {1.413 > 1.65} (vô lý) | {1.413 > 1.96} (vô lý) |
| Kết luận tạm thời | chấp nhận h0, bác bỏ h1 | chấp nhận h0, bác bỏ h1 | chấp nhận h0, bác bỏ h1 |
| p-value |  |  |  |
| Kết luận chính thức | chấp nhận h0 , bác bỏ h1 | chấp nhận h0, bác bỏ h1 | chấp nhận h0, bác bỏ h1 |

Kết luận: Tất cả các kiểm định ở trên đều không đúng với thực tế vì trong khi

1. **Giả sử tập dữ liệu gồm 65 sinh viên trên là tập mẫu trong một tổng thể là hữu hạn, tuy nhiên do nhiều điều kiện khác nhau nên không thể tiến hành thu thập dữ liệu của tổng thể. Hãy thực hiện lại các bước tính toán theo yêu cầu trong câu 8, 9.**

Trong bài toán này, ta thấy phân phối của tổng thể vẫn là phân phối chuẩn N(, nhưng ta không biết cả 2 tham số của nó là gì

Câu 8:

Xét tập S1:

* + Trung bình mẫu: ;
  + Độ lệch mẫu:
  + n = 30
  + Gọi α là xác suất mà không rơi vào khoảng tin cậy. Theo đề bài, độ tin cậy là 100(1- α) % = 80%
* α/2 = 0.1

Khoảng tin cậy: (

Kết luận: dù chưa biết trung bình tổng thể nhưng ta có thể xác định nó rơi vào khoảng với một độ tin cậy là 80%

Xét tập S2:

* + Trung bình mẫu:
  + Độ lệch mẫu: 34.29
  + n = 15
  + Gọi α là xác suất mà không rơi vào khoảng tin cậy. Theo đề bài, độ tin cậy là 100(1- α) % = 95%
* α/2 = 0.025
* 2.145

Khoảng tin cậy : (

Kết luận: tương tự như trên, dù chưa biết trung bình tổng thể nhưng ta có thể xác định nó rơi vào khoảng với một độ tin cậy là 95%

Câu 9:

Xét tập S1:

* + Trung bình mẫu
  + n = 30
  + độ lệch mẫu : s = 37.124
  + α = 0.05 =>

α/2 = 0.025 =>

Gọi µ là trung bình của tổng thể, ước lượng hợp lý cực đại cho µ

* + thống kê kiểm định: TS = =

MLE : μ = x ̅= 94.46666666666667

|  |  |  |  |
| --- | --- | --- | --- |
|  | Kiểm định trái | Kiểm định phải | Kiểm định 2 bên |
| Giả thuyết |  |  |  |
| Miền bác bỏ  (chưa biết ) |  |  |  |
| Dữ liệu thực tế (*μ)* | MLE : *μ* = > TS = 0 | | |
| Miền bác bỏ  (đã biết ) |  |  |  |
| Kết luận tạm thời | Không bác bỏ h0 | Không bác bỏ h0 | Không bác bỏ h0 |
| p-value | = 0.1 |  |  |
| Kết luận chính thức | Chấp nhận h0  Bác bỏ h1 | Không có đủ chứng cứ  Để chấp nhận hay bác bỏ | Không có đủ chứng cứ  Để chấp nhận hay bác bỏ |

Kết luận: Trong bài toán trên, ta quy tổng thể thành một mẫu dữ liệu nhỏ nhưng không biết chính xác tổng thể chứa nó có bộ tham số nào. Điều này đã gây khó dễ và ta chỉ có thể ước lượng khoảng tin cậy cho bộ tham số của tổng thể nhưng không thể kiểm định chính xác giả thuyết đặt ra cho nó

**TÀI LIỆU THAM KHẢO**

[CoderSchool | Top Coding School In Vietnam](https://www.coderschool.vn/blog/hypothesis-testing-la-gi-cac-buoc-thuc-hien-hypothesis-testing-trong-phan-tich-du-lieu)

**PHỤ LỤC**

Phần này bao gồm những nội dung cần thiết nhằm minh họa hoặc hỗ trợ cho nội dung luận văn như số liệu, biểu mẫu, tranh ảnh. . . . nếu sử dụng những câu trả lời cho một *bảng câu hỏi thì bảng câu hỏi mẫu này phải được đưa vào phần Phụ lục ở dạng nguyên bản* đã dùng để điều tra, thăm dò ý kiến; **không được tóm tắt hoặc sửa đổi**. Các tính toán mẫu trình bày tóm tắt trong các biểu mẫu cũng cần nêu trong Phụ lục của luận văn. Phụ lục không được dày hơn phần chính của luận văn