

**TRƯỜNG ĐẠI HỌC BÁCH KHOA THÀNH PHỐ HỒ CHÍ MINH**

**KHOA KHOA HỌC ỨNG DỤNG**

**BỘ MÔN TOÁN ỨNG DỤNG**



**BÁO CÁO BÀI TẬP LỚN**

**XÁC SUẤT THỐNG KÊ**

**GVHD: NGUYỄN BÁ THI**

**NHÓM 5**

**Sinh viên thực hiện: Lê phước lộc \_ 1511842**

Thành phố Hồ Chí Minh, tháng 4 năm 2017

|  |  |  |
| --- | --- | --- |
| ***Câu 1***: **Bảng sau đây cho ta phân bố thu nhập của hai nhóm tuổi: Nhóm từ 40 – 50 tuổi** | | |
| **và nhóm từ 50 – 60 .trong số các công nhân lành nghề ở Thụy Điển** |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nhóm tuổi | Thu nhập |  |  |  |  |  |
|  | 0 – 1 | 1 – 2 | 2 – 3 | 3 – 4 | 4 – 6 | ≥ 6 |
| 40 – 50 | 7831 | 26740 | 35572 | 20009 | 11527 | 6919 |
| 50 – 60 | 7858 | 20685 | 24186 | 12280 | 6776 | 4222 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Có sự khác nhau về phân bố thu nhập giữa hai nhóm tuổi này trong số các công nhân** | | | | | |
| **hay không? Mức ý nghĩa α = 1%.** |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **1 - Dạng bài:** |  |  | | Đây là dạng bài kiểm định giả thuyết về tỷ lệ | | |   **2- Phương pháp và Các bước giải**  **Bước 1**   |  | | --- | | Tính các tổng sau: | | Tổng các hàng | | Tổng các cột | | Tổng cộng (grand total)  **Bước 2**   |  |  |  | | --- | --- | --- | | Tính tần số lý thuyết |  |  | | Tần số lý thuyết = tổng hàng x tổng cột / tổng cộng | | |   **Bước 3**  Tính xác suất sau:   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **3 - Công cụ giải** | | | |  | | | | Để giải bài này ta dùng hàm CHITEST | | | | | | | | **4 - Tính toán** | | | | | | | |  | | | | | | | | Nhóm tuổi | Thu nhập | | | | | | | |  | | 0 – 1 | 1 – 2 | 2 – 3 | | 3 – 4 | 4 – 6 | | ≥ 6 | Tổng hàng | | 40 – 50 | 7831 | 26740 | 35572 | | 20009 | 11527 | | 6919 | 108598 | | 50 – 60 | 7858 | 20685 | 24186 | | 12280 | 6776 | | 4222 | 76007 | | Tổng cột | 15689 | 47425 | 59758 | | 32289 | 18303 | | 11141 | 184605 |   **Các tần số lý thuyết**   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | 0 -- 1 | 1 -- 2 | 2 -- 3 | 3 -- 4 | 4 -- 6 | ≥ 6 | | 40 - 50 | 9229.4034 | 27898.81179 | 35153.9735 | 18994.72291 | 10767.15 | 6553.941 | | 50 - 60 | 6459.5966 | 19526.18821 | 24604.0265 | 13294.27709 | 7535.853 | 4587.059 |   P(X > ꭓ²) = 3.732E-204  **5 -Kết luận**  Với H0: Mức thu nhập của các lao động trong 2 nhóm tuổi trên là như nhau  Vì P(X > ꭓ²) = 3.7317E-204 < α = 0.01: Bác bỏ H0  Vậy mức thu nhập của các lao động trong 2 nhóm tuổi trên là khác nhau   |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | ***Câu 2*: Điều tra ngẫu nhiên thu nhập của 5 nữ nhân viên và 5 nam nhân viên đang làm việc tại các công ty công ty liên doanh về thu nhập hàng năm của họ thu được kết quả sau:**   |  |  | | --- | --- | | Nữ | Nam | | 9 | 16 | | 12 | 19 | | 8 | 12 | | 10 | 11 | | 16 | 22 |  1. **Với độ tin cậy 95% hãy ước lượng thu nhập trung bình hàng năm của nam và nữ nhân viên.** 2. **Kết quả thu được có phản ánh sự phân biệt đối xử đối với nữ nhân viên hay không. Biết rằng thu nhập hàng năm có phân phối chuẩn.** | | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | **1 - Dạng bài:** | |  |  |  | | **Câu a:** Ước lượng trung bình bài toán 2 mẫu | | | |  | |  |  |  |  |  | | **Câu b:** So sánh giá trị trung bình với phương sai khác nhau   |  |  |  |  | | --- | --- | --- | --- | | **2- Phương pháp và Các bước giải** | | | | |  |  |  |  | | **Câu a:** |  |  |  | | **Bước 1:** Tính trung bình của 2 mẫu |  | | | |  | |  | |  | | **Bước 2:** Tính độ lệch chuẩn mẫu: | | |  | |  |  | | | |  | |  | | **Bước 3:** Tính độ chính xác của ước lượng: | | | | |  | |  | | --- | |  | |  |  | |  |  |  |  | |  |  |  |  | | | | | | |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Câu b:** | |  |  |  |  |  |  |  | |  | **Bước 1**: Đặt giả thiết H0: Mức lương trung bình của nam nhân viên và nữ nhân viên là giống nhau | | | | | | | | |  | **Bước 2**: Tính giá trị thống kê và giá trị theo bản phân phối student | | | | | |  |  | | |  | | --- | |  | | |  |  |  |  |  |  |  | |  | |  |  |  |  |  |  |  | |  | |  |  |  |  |  |  |  | |  | |  |  |  |  |  |  |  | | |  | | --- | |  | | |  |  |  |  |  |  |  | |  | |  |  |  |  |  |  |  | |  | |  |  |  |  |  |  |  |   **Bước 3**: Kết luận     |  |  |  |  |  | | --- | --- | --- | --- | --- | | **3 - Công cụ giải** |  |  |  |  | | *Chọn Tools/Data Analysis…/ t-test: Two-Sample Assuming Equal Variances* | | | | | | |

**4 - Tính toán**

|  |  |  |  |
| --- | --- | --- | --- |
| **Câu a:** |  |  |  |
|  |  | Nữ | Nam |
|  |  | 9 | 16 |
|  |  | 12 | 19 |
|  |  | 8 | 12 |
|  |  | 10 | 11 |
|  |  | 16 | 22 |
| Trung bình |  | 11 | 16 |
| Độ lệch chuẩn mẫu | | 3.16227766 | 4.63680925 |
| ɛ |  | 3.926486323 | 5.75735911 |

**Câu b:**

|  |  |
| --- | --- |
| **Nữ** | **Nam** |
| 9 | 16 |
| 12 | 19 |
| 8 | 12 |
| 10 | 11 |
| 16 | 22 |

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | | |
|  |  |  |
|  | *Nữ* | *Nam* |
| Mean | 11 | 16 |
| Variance | 10 | 21.5 |
| Observations | 5 | 5 |
| Hypothesized Mean Difference | 0 |  |
| df | 7 |  |
| t Stat | -1.99204768 |  |
| P(T<=t) one-tail | 0.04331465 |  |
| t Critical one-tail | 1.894578605 |  |
| P(T<=t) two-tail | 0.086629301 |  |
| t Critical two-tail | 2.364624252 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **5 -Kết luận** | | |  | |  | |
| **Câu a** |  | |  | |  | |
| Thu nhập trung bình của các nhân viên nằm trong khoảng : | | | | | | |
|  | **Cận dưới** | | **Cận trên** | |  | |
| **Nữ** | 7.073513677 | | 14.92648632 | |  | |
| **Nam** | 10.24264089 | | 21.75735911 | |  | |
| **Câu b** | |  | |  | |  | |  |  |
| Ta có |t| = -1.99205 < tα/2, do đó không có sự phân biệt đối xử giữa nam và nữ | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Câu 3*: Tính tỉ số tương quan của Y đối với X, hệ số tương quan và hệ số xác định**  **của tập số liệu sau đây. Với mức ý nghĩa α = 5%, có kết luận gì về mối tương quan**  **giữa X và Y (Có phi tuyến không ? Có tuyến tính không).**  **Tìm đường hồi quy của Y đối với X.** | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | |
| **X** | | 0.9 | 1.22 | 1.32 | 0.77 | 1.3 | 1.2 | 0.9 | 1.22 | 1.32 | 0.77 | 1.3 | 1.2 |
| **Y** | | -0.3 | 0.1 | 0.7 | -0.28 | -0.25 | 0.02 | 0.3 | 0.5 | 0.4 | -0.2 | -0.3 | 0.08 |

|  |  |  |
| --- | --- | --- |
| **1 - Dạng bài:** |  |  |
| Đây là bài toán phân tích tương quan và hồi quy | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **2- Phương pháp và Các bước giải** | | | | | | | | | | | | | | | |
|  |  | | | | | | | | | | | | |  |  |
| **Bước 1**: Tính hệ số tương quan | | | | | | | | | | | | | | |  |
|  | |  | | --- | |  | | | | | | | | | | | | | |  |  |
|  |  | | | | | | | | | | | | |  |  |
|  |  | | | | | | | | | | | | |  |  |
|  |  | | | | | | | | | | | | |  |  |
| **Bước 2**: Phân tích tương quan tuyến tính | | | | | | | | | | | | |  | | | | | |  | |  | |  | |  | |  | |
|  | | | |  | | |  | | |  | | |  | | | | | |  | |  | |  | |  | |  | |
| · Giả thuyết: H0 X và Y không có tương quan tuyến tính. | | | | | | | | | | | | | | | | | | | | |  | |  | |  | |  | |
| |  | | --- | | · Áp dụng: "Correlation" để tính hệ số tương quan và hệ số xác định | | | | | | | | | | | | | | | | | | | | | | | |  | |  | |  | |
| · Ta tính: và c là phân vị mức a/2 của phân phối Student với n-2 bậc tự do (n=12) | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|
| · Nếu T > c thì bác bỏ giả thiết H0 | | | | | | | | | | | | |  | | | | | |  | |  | |  | |  | |  | |
| **Bước 3**: Phân tích tương quan phi tuyến | | | | | | | | | | | | |  | | | | | |  | |  | |  | |  | |  | |
|  | | |  | | |  | | |  | |  | | | | | |  | | |  | |  | |  | |  | |
| · Sắp xếp lại bảng số liệu | | | | | | | | |  | |  | | | | | |  | | |  | |  | |  | |  | |
| · Áp dụng: "Anova Single Factor" để tính tỷ số tương quan | | | | | | | | | | | | | | | | | | | |  | |  | |  | |  | |
| · Giả thuyết: Ho là X và Y không có tương quan phi tuyến | | | | | | | | | | | | | | | | | | | |  | |  | |  | |  | |
| |  | | --- | | · Ta tính: và c là phân vị mức a=0,05 của phân bố Fisher bậc tự do (k-2,n-k)  với n=12, k=6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|
|
| · Nếu F<c thì chấp nhận giả thuyết H0 | | | | | | | | | | |  | | | | | |  | | |  | |  | |  | |  | |
| **Bước 4**: Phân tích hồi quy | | | | | | | |  | | | |  | | | | | |
| Áp dụng: "Regression" để tìm đường hồi quy   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **3 - Công cụ giải** | |  |  |  |  | | Để giải quyết tương quan tuyến tính: *Data/Data Analysis…/Correlation* | | | | |  | |  |  |  |  |  |  | | Để giải quyết tương quan phi tuyến: *Data/Data Analysis…/Anova: Single* Factor | | | | | | |  |  |  |  |  |  | | Để giải quyết bài toán hồi quy: *Data/Data Analysis…/Regression* | | | | |  |   **4 - Tính toán** | | | | | | | | | | | | | | | | | |
| X | | Y | | |
| 0.9 | | -0.3 | | |
| 1.22 | | 0.1 | | |
| 1.32 | | 0.7 | | |
| 0.77 | | -0.28 | | |
| 1.3 | | -0.25 | | |
| 1.2 | | 0.02 | | |
| 0.9 | | 0.3 | | |
| 1.22 | | 0.5 | | |
| 1.32 | | 0.4 | | |
| 0.77 | | -0.2 | | |
| 1.3 | | -0.3 | | |
| 1.2 | | 0.08 | | |

|  |  |  |
| --- | --- | --- |
|  | *X* | *Y* |
| X | 1 |  |
| Y | 0.432995932 | 1 |

Hệ số tương quan

|  |  |
| --- | --- |
| R = | 0.432995932 |

Hệ số xác định

|  |  |
| --- | --- |
| R^2 = | 0.187485477 |

|  |  |
| --- | --- |
| T = | 1.519036643 |
| c = | 2.228138852 |
|  |  |

**Phân tích tương quan phi tuyến**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| X | 0.77 | 0.9 | 1.2 | 1.22 | 1.3 | 1.32 |
| Y | -0.28 | -0.3 | 0.02 | 0.1 | -0.25 | 0.7 |
| -0.2 | 0.3 | 0.08 | 0.5 | -0.3 | 0.4 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Anova: Single Factor | |  |  |  |
|  |  |  |  |  |
| SUMMARY |  |  |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |
| 0.77 | 2 | -0.48 | -0.24 | 0.0032 |
| 0.9 | 2 | 0 | 0 | 0.18 |
| 1.2 | 2 | 0.1 | 0.05 | 0.0018 |
| 1.22 | 2 | 0.6 | 0.3 | 0.08 |
| 1.3 | 2 | -0.55 | -0.275 | 0.00125 |
| 1.32 | 2 | 1.1 | 0.55 | 0.045 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 1.007041667 | 5 | 0.201408333 | 3.882570281 | 0.0645508 | 4.38737419 |
| Within Groups | 0.31125 | 6 | 0.051875 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 1.318291667 | 11 |  |  |  |  |

|  |  |  |
| --- | --- | --- |
| |  | | --- | |  | | | 0.76389899 |
|
| |  | | --- | |  | | | 0.87401315 |
|

|  |  |
| --- | --- |
| F | 3.662077708 |
| c | 4.53367695 |

|  |
| --- |
| **Phân tích hồi quy tuyến tính** |
| Giả thiết H0: X và Y hồi quy tuyến tính |
|  |

|  |  |
| --- | --- |
| SUMMARY OUTPUT | |
|  |  |
| *Regression Statistics* | |
| Multiple R | 0.432995932 |
| R Square | 0.187485477 |
| Adjusted R Square | 0.106234025 |
| Standard Error | 0.327281396 |
| Observations | 12 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 0.247160542 | 0.247160542 | 2.307472321 | 0.1597169 |
| Residual | 10 | 1.071131124 | 0.107113112 |  |  |
| Total | 11 | 1.318291667 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | -0.7073118 | 0.516586474 | -1.369203078 | 0.200899808 | -1.858338 | 0.4437146 |
| X | 0.68984661 | 0.454134277 | 1.519036643 | 0.159716927 | -0.322027 | 1.7017208 |

|  |  |
| --- | --- |
| *Lower 95.0%* | *Upper 95.0%* |
| -1.858338183 | 0.443715 |
| -0.32202762 | 1.701721 |

|  |  |  |
| --- | --- | --- |
| RESIDUAL OUTPUT | |  |
|  |  |  |
| *Observation* | *Predicted Y* | *Residuals* |
| 1 | -0.08644984 | -0.213550157 |
| 2 | 0.134301072 | -0.034301072 |
| 3 | 0.203285733 | 0.496714267 |
| 4 | -0.1761299 | -0.103870098 |
| 5 | 0.1894888 | -0.4394888 |
| 6 | 0.12050414 | -0.10050414 |
| 7 | -0.08644984 | 0.386449843 |
| 8 | 0.134301072 | 0.365698928 |
| 9 | 0.203285733 | 0.196714267 |
| 10 | -0.1761299 | -0.023870098 |
| 11 | 0.1894888 | -0.4894888 |
| 12 | 0.12050414 | -0.04050414 |

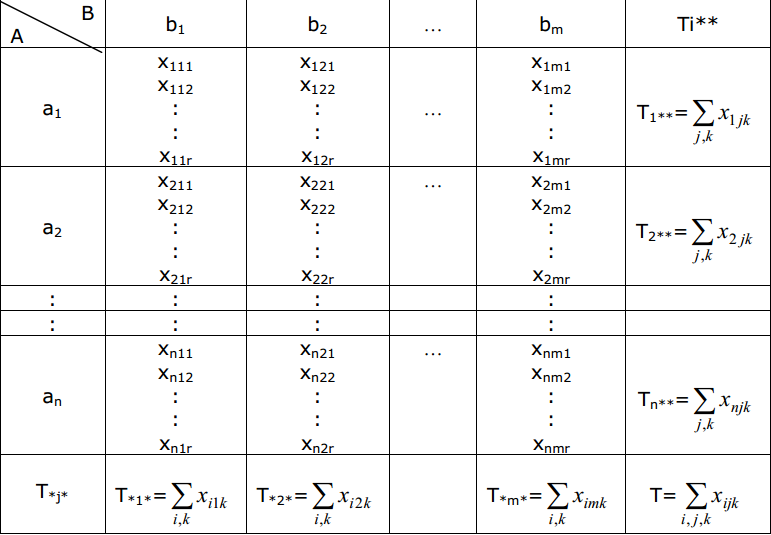
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Tương quan tuyến tính:** | | |  |  |  |  |  |
|  | Hệ số tương quan R = 0.432996 | | | |  |  |  |
|  | Hệ số xác định R^2 = 0.18749 | | | | |  |  |
|  |  |  |  |  |  |  |  |
|  | T < c => chấp nhận giả thiết H0 : X và Y không có tương quan tuyến tính | | | | | |  |
|  |  |  |  |  |  |  |  |
| **Tương quan phi tuyến** | |  |  |  |  |  |  |
|  | F = 3.66208 < c nên chấp nhận giả thiết H0: X và Y không có tương quan phi tuyến | | | | | | |
|  |  |  |  |  |  |  |  |
| **Hồi quy tuyến tính** | |  |  |  |  |  |  |
|  | Sai số tiêu chuẩn của đường hổi quy 0.327281 | | | |  |  |  |
|  |  |  |  |  |  |  |  |
|  | Vậy đường hồi quy Y đối với X: Y = -0.70731 + 0.68947X | | | | |  |  |
|  |  |  |  |  |  |  |  |
|  | Hệ số hồi quy: 0.2009 > 0.05 => Hệ số tự do có ý nghĩa | | | | |  |  |
|  | 0.159717 > 0.05 => Hệ số của X có ý nghĩa | | | |  |  |  |
|  |  |  |  |  |  |  |  |
|  | Vậy phương trình hồi quy tuyến tính này thích hợp vì 0.159717 > 0.05 | | | | |  |  |
| ***Bài 4*: Hãy phân tích tình hình kinh doanh của một số ngành nghề ở 4 quận nội thành trên** | | | | | | | | |
| **cơ sở số liệu về doanh thu của một số cửa hàng với 3 lần quan sát như sau (mức ý nghĩa 10%)** | | | | | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ngành nghề kinh doanh | Khu vực kinh doanh | | | |
|
| Q1 | Q2 | Q3 | Q4 |
| Điện lạnh | 5.7:6.1:6.6 | 3.1:3.3:3.0 | 4.2:4.6: 4.4 | 5.0:5.2:5.5 |
| Vật liệu xây dựng | 5.0:5.2:5.4 | 15.0:14.4:15.6 | 9.5:8.8:9.8 | 17.5:17.2:18 |
| Dịch vụ tin học | 3.8:4.1:3.9 | 1.8:2.0:1.9 | 1.3:1,5:1.4 | 4.8:4.6:4.9 |

|  |  |
| --- | --- |
| **1 - Dạng bài: bài toán phân tích phương sai hai yếu tố có lặp.** |  |
| Yêu cầu: xác định xem doanh có khác nhau giữa mỗi địa điểm kinh doanh hay không |  |

Giả thuyết : HA là doanh thu không phụ thuộc vào ngành nghề kinh doanh.  
 HB là doanh thu không phụ thuộc vào khu vực kinh doanh.  
 HAB là doanh thu giữa ngành nghề kinh doanh và khu vực kinh doanh không có tương tác với nhau.

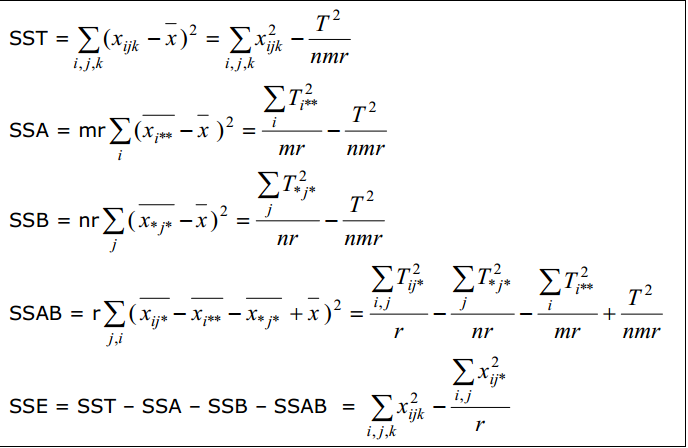
|  |  |  |
| --- | --- | --- |
| **2- Phương pháp và Các bước giải** | | |
| **Bước 1:** | Tính các giá trị cần thiết |  |



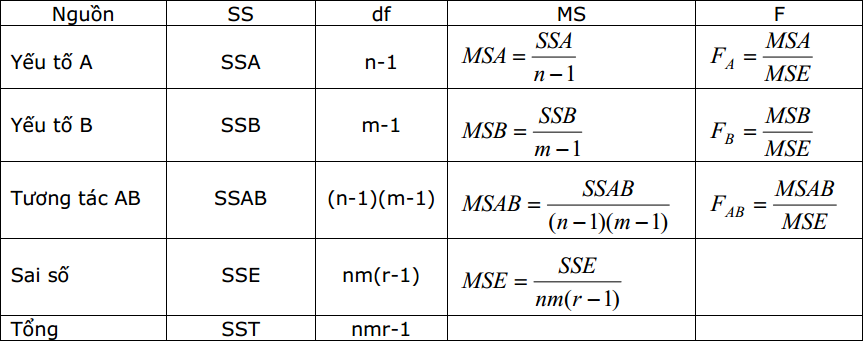
Ta cần tính



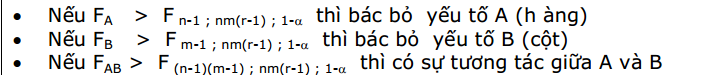
và suy ra



**Bước 2:** Tính các giá trị thống kê



|  |  |
| --- | --- |
| **Bước 3:** | Kết luận |



|  |  |  |  |
| --- | --- | --- | --- |
| **3 - Công cụ giải** |  |  |  |
| Chọn Tools\Data Analysis…\Anova: Two Factor With Replication | | | |
|  | | | |

**4 - Tính toán**

Chuyển đổi bảng về dạng

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Q1 | Q2 | Q3 | Q4 |
| Điện lạnh | 5.7 | 3.1 | 4.2 | 5 |
| 6.1 | 3.3 | 4.6 | 5.2 |
| 6.6 | 3 | 4.4 | 5.5 |
| Vật liệu xây dựng | 5 | 15 | 9.5 | 17.5 |
| 5.2 | 14.4 | 8.8 | 17.2 |
| 5.4 | 15.6 | 9.8 | 18 |
| Dịch vụ tin học | 3.8 | 1.8 | 1.3 | 4.8 |
| 4.1 | 2 | 1.5 | 4.6 |
| 3.9 | 1.9 | 1.4 | 4.9 |

**Dùng công cụ**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Anova: Two-Factor With Replication | | |  |  |  |
|  |  |  |  |  |  |
| SUMMARY | Q1 | Q2 | Q3 | Q4 | Total |
| *Điện lạnh* |  |  |  |  |  |
| Count | 3 | 3 | 3 | 3 | 12 |
| Sum | 18.4 | 9.4 | 13.2 | 15.7 | 56.7 |
| Average | 6.13333333 | 3.133333333 | 4.4 | 5.233333333 | 4.725 |
| Variance | 0.20333333 | 0.023333333 | 0.04 | 0.063333333 | 1.391136 |
|  |  |  |  |  |  |
| *Vật liệu xây dựng* |  |  |  |  |  |
| Count | 3 | 3 | 3 | 3 | 12 |
| Sum | 15.6 | 45 | 28.1 | 52.7 | 141.4 |
| Average | 5.2 | 15 | 9.366666667 | 17.56666667 | 11.78333 |
| Variance | 0.04 | 0.36 | 0.263333333 | 0.163333333 | 25.50697 |
|  |  |  |  |  |  |
| *Dịch vụ tin học* |  |  |  |  |  |
| Count | 3 | 3 | 3 | 3 | 12 |
| Sum | 11.8 | 5.7 | 4.2 | 14.3 | 36 |
| Average | 3.93333333 | 1.9 | 1.4 | 4.766666667 | 3 |
| Variance | 0.02333333 | 0.01 | 0.01 | 0.023333333 | 2.129091 |
|  |  |  |  |  |  |
| *Total* |  |  |  |  |  |
| Count | 9 | 9 | 9 | 9 |  |
| Sum | 45.8 | 60.1 | 45.5 | 82.7 |  |
| Average | 5.08888889 | 6.677777778 | 5.055555556 | 9.188888889 |  |
| Variance | 0.98111111 | 39.34194444 | 12.22027778 | 39.58361111 |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Sample | 519.77056 | 2 | 259.885278 | 2549.283379 | 1.12E-28 | 2.5383319 |
| Columns | 102.05417 | 3 | 34.0180556 | 333.6920981 | 1.08E-19 | 2.3273897 |
| Interaction | 214.79833 | 6 | 35.7997222 | 351.1689373 | 3.71E-22 | 2.0351316 |
| Within | 2.4466667 | 24 | 0.10194444 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 839.069722 | 35 |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **5 -Kết luận** |  |  |  |  |
| Từ kết quả ta có các kết luận sau | |  |  |  |
| FA > F crit : doanh thu khác nhau theo từng ngành | | |  |  |
| FB > F crit : doanh thu khác nhau theo địa điểm kinh doanh | | | |  |
| FC > F crit : có sự tương tác giữa ngành và địa điểm kinh doanh trên doanh thu | | | | |