**BÀI TẬP:**

Mở tệp **SMOKE\_SDH.sav. thực hiện các yêu cầu sau:**

1, Cho biết:

* *Số lượng biến:16*
* *Số lượng bản ghi:337*

2, Sử dụng Test kiểm định để kiểm tra các biến sau đây trong dữ liệu có phân phối chuẩn hay không?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tên biến | Test: | | | Kết luận |
| Statistic | df | Sig |
| Tuổi |  |  |  |  |
| Cân nặng (Cannang kg) |  |  |  |  |
| Chiều cao (Chieucao m) |  |  |  |  |
| Số điếu thuốc hút trong một ngày (NUMCIGAR) |  |  |  |  |
| Cân nặng sau bỏ thuốc (WEIGHT\_BT) |  |  |  |  |

3,Tạo biến **PLHT** (Phân loại hút thuốc) và sử dụng lệnh để điền dữ liệu cho biến này theo quy định sau đây:

|  |  |  |
| --- | --- | --- |
| **So dieu thuoc hut** | **PLHT** | **Số lượng** |
| Dưới 15 | 1 | *27* |
| Từ 15 đến 25 | 2 | *45* |
| Trên 25 | 3 | *10* |

4, Thực hiện lệnh và điền kết quả vào các bảng sau:

**Bảng 1: Đặc điểm của Tuổi và Giới (compare mean)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Nam** | | **Nữ** | | **Chung** | |
| n | Tuổi (+SD) | n | Tuổi (+SD) | n | Tuổi (+SD) |
| **122** | **43.35 ± 17.075** | **215** | **47.32 ± 19.439** | **337** | **45.88 ± 18.690** |

**Bảng 2: Đặc điểm Hút thuốc theo Chủng tộc**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | Chung toc | | | Tổng |
| 1 | 2 |  |
| Hút thuốc | Hút thuốc | n | 73 | 9 |  | 82 |
| % hút thuốc theo Chủng tộc | 25.3 | 19.1 |  | 24.4 |
| Không hút thuốc | n | 216 | 38 |  | 254 |
| % hút thuốc theo Chủng tộc | 74.7 | 80.9 |  | 75.6 |
| Total | | n | 289 | 47 |  | 336 |

**Bảng 3: Đặc điểm Tình trạng hút thuốc theo Giới tính**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | | Giới tính | | Tổng |
| Nam | Nu |
| Tình trạng  hút thuốc | Hút thuốc | n | 36 | 47 | 82 |
| % hút thuốc theo giới tính | 29.5 | 21.5 | 24.4 |
| Không hút thuốc | n | 86 | 168 | 254 |
| % hút thuốc theo giới tính | 70.5 | 78.5 | 75.6 |
| Total | | n | 122 | 214 | 336 |

5*,* Trong khoảng tin cậy 95%, **Chủng tộc** có ảnh hưởng đến **Hút thuốc** hay không? Vì sao?

*(Có yêu cầu tỷ lệ 95% -> dùng OR. Nếu không có -> kiểm định 4 bước)*

*Lệnh thực hiện: Analyze -> Descriptive Statistic -> Crosstab -> Statistic -> chọn Chi-square, Risk. Chọn rows là SMOKE, Column là RACE*

Nhận xét:

OR = 1.427 # 1

Trong khoảng tin cậy 95% OR từ 0.658 đến 3.093 có chứa 1 nên chủng tộc không ảnh hưởng đến tình trạng hút thuốc; và sự khác nhau là không có ý nghĩa thống kê.

6, Nhận xét mối quan hệ (độc lập/phụ thuộc) giữa **Hút thuốc** và **Hôn nhân** trong khoảng tin cậy 95%.

Lệnh thực hiện:

Nhận xét:

*1. Giả thuyết H0: hút thuốc và hôn nhân độc lập;*

*H1 hút thuốc ở hôn nhân phụ thuộc nhau*

*2. Khi bình phương QS: 9.985 (mức ý nghĩa Asymptotic Sig (2-sided) = 0.076*

*3. Khi bình phương lý thuyết: (bậc tự do =1, mức ý nghĩa 0.05): 11*

*4. Kết luận: Khi bình phương QS = 9.985 < LT 11 -> Chấp nhận H0: hút thuốc và tình trạng hôn nhân độc lập với nhau*

7, So sánh **Tuổi** trung bình giữa các **Chủng tộc** trong khoảng tin cậy 95%.

Lệnh thực hiện:Analyze -> Nonparametric Test -> Legacy Dialogs -> 2 Independent Samples

Nhận xét:

8, So sánh **Cân nặng (kg)** trung bình giữa **nhóm hút thuốc** trong khoảng tin cậy 95%.

Lệnh thực hiện:

Nhận xét:

9, **Chiều cao (m)** và **Cân nặng (kg)** có tương quan tuyến tính với nhau hay không? Nếu có, hãy viết phương trình hồi quy.

Lệnh thực hiện:

10, Vẽ biểu đồ biểu diễn theo các yêu cầu sau đây:

- Biểu diễn tỷ lệ phần trăm H**út thuốc** theo **Giới tính**.

- Biểu diễn **Tuổi** trung bình giữa các **Chủng tộc**.

- Biểu diễn tỷ lệ phần trăm **Tình trạng hút thuốc** theo **Giới tính**.

11, Lưu OUTPUT vào thư mục của Học viên.

*1. Giả thuyết H0: hút thuốc và chủng tộc độc lập;*

*H1 hút thuốc ở chủng tộc phụ thuộc nhau*

*2. Khi bình phương QS: 0.52 (mức ý nghĩa Asymptotic Sig (2-sided) = 0.471*

*3. Khi bình phương lý thuyết: (bậc tự do =1, mức ý nghĩa 0.05): 3.84*

*4. Kết luận: Khi bình phương QS = 0.52 < LT 3.84 -> Chấp nhận H0: hút thuốc và chủng tộc độc lập với nhau*