# A research on the relationship between Age(gender) and the incident rate of the Parkinson’s disease

Yundi Zhou(Dino), Hao Huo(Emery), Yushu WU(Nathan)

G12 AP Statistics

Beijing National Day School, Beijing China

Proposal:

1.Research question:

The group will investigate the statistical relationship between age and the indent rate of Parkinson’s disease. We noticed that several factors contributed to the incident of Parkinson’s disease, including age, gender, genetic factor and so on. In this case, we decide to put our focus on the relationship between age and the incident of Parkinson Disease.

2.Background research:

Idiopathic Parkinson's disease is a degenerative neurological disorder classically presenting in old or late middle age, The brain shows characteristic ell loss and depigmentation in pigmented brain stem nuclei. The presence of rounded eosinophilic intra-cytoplasmic inclusions, known as Lewy bodies, in some of the affected neurons is a sine qua non for definitive pathological diagnosis.(N. P. Quinn, 1986)

3.Sampling and experiment design:

Independent Variables: People’s Age

Dependent Variables: The Indent rate of Parkinson Diseases.

Type of Studies：Observational Studies because there is no treatments or placebo in this studies. We will randomly select the clinical history of patients in the neurology departement from the Tiantan Hospital and detect the relationship between the ages and the indent rate of Parkinson Diseases.

Data Selection: We will record the ages of people who had Parkinson Diseases by randomly selecting the clinical history and find the relationship.

Scope of Inference: These data can collect the relationship of people who had Parkinson diseases from all Chinese population.

4. Explanatory Data Analysis:

Sample Size and Power Analysis: Ensure that the sample size is adequate to detect meaningful differences in the incidence rate of Parkinson's disease across different age groups. Conduct a power analysis to determine the required sample size for detecting the effect size of interest with sufficient power.

Random Sampling: Verify that the sample is randomly selected or represents a well-defined population to ensure generalizability of the findings.

Data Quality and Validity: Check for missing data, outliers, or data entry errors. Ensure that the data collected for each participant is accurate and reliable, including age, gender, and diagnostic information.

Data Summaries:

Descriptive Statistics: Calculate summary statistics for age, such as mean, median, standard deviation, and range, to understand the central tendency and variability of age in the sample.

Age Distribution: Plot a histogram or density plot to visualize the distribution of age in the sample. This will provide insights into the shape and skewness of the age distribution.

Incidence

Rate by Age Group: Divide the participants into different age groups and calculate the incidence rate of Parkinson's disease for each group. Summarize the incidence rates using descriptive statistics, such as means or proportions, and create visual representations (e.g., bar chart) to compare the rates across age groups.

Simulations and Test Procedures:

Simulations: Utilize statistical software or programming languages to simulate hypothetical datasets based on known statistical relationships or assumptions. To detect the p-value of the observed counts and how much they deflect from the expected counts.

Statistical Tests: Chi-square of the Independence

Chi-sqaure test Analysis: We search on the internet that people above 50 years old who have Parkinson diseases are 1 percent of the whole elder population. We expect that half of the Parkinson patients are below 70 years old and half them are over 70 years old for both males and females.

Hypothesis Testing: Utilize hypothesis tests, such as chi-square t-tests to determine if there are statistically significant differences in the incidence rate of Parkinson's disease across different age groups.

5. Group Task Assignments and Timeline：

In this project, Dino is responsible for the introduction part and data gathering part, Emery design the study and come up with the conclusion, Nathan will analyze the data. The proposal and data collection will be done by May 18th. The main research part will be done May 25th, and the project will be finished by may 27th.

6. Data:

x

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | Age | Gender |  |
| 1 | 54 | M |  |
| 2 | 70 | M |  |
| 3 | 68 | M |  |
| 4 | 74 | F |  |
| 5 | 61 | M |  |
| 6 | 77 | F |  |
| 7 | 79 | M |  |
| 8 | 63 | M |  |
| 9 | 52 | M |  |
| 10 | 80 | F |  |
| 11 | 81 | M |  |
| 12 | 68 | F |  |
| 13 | 75 | F |  |
| 14 | 67 | M |  |
| 15 | 57 | M |  |
| 16 | 64 | M |  |
| 17 | 74 | F |  |
| 18 | 81 | F |  |
| 19 | 77 | F |  |
| 20 | 76 | M |  |
| 21 | 59 | F |  |
| 22 | 62 | M |  |
| 23 | 70 | M |  |
| 24 | 75 | M |  |
| 25 | 76 | M |  |
| 26 | 71 | F |  |
| 27 | 80 | M |  |
| 28 | 59 | M |  |
| 29 | 69 | F |  |
| 30 | 75 | F |  |
| 31 | 82 | M |  |
| 32 | 77 | M |  |
| 33 | 59 | F |  |
| 34 | 51 | M |  |
| 35 | 64 | M |  |
| 36 | 69 | M |  |
| 37 | 72 | F |  |
| 38 | 74 | M |  |
| 39 | 80 | F |  |
| 40 | 75 | M |  |
| 41 | 61 | M |  |
| 42 | 60 | M |  |
| 43 | 71 | F |  |
| 44 | 70 | F |  |
| 45 | 74 | M |  |
| 46 | 64 | F |  |
| 47 | 55 | M |  |
| 48 | 71 | M |  |
| 49 | 82 | M |  |
| 50 | 59 | M |  |
| 51 | 68 | F |  |
| 52 | 77 | F |  |
| 53 | 74 | M |  |
| 54 | 72 | M |  |
| 55 | 75 | M |  |
| 56 | 68 | M |  |
| 57 | 65 | M |  |
| 58 | 70 | F |  |
| 59 | 54 | M |  |
| 60 | 68 | F |  |
| 61 | 62 | F |  |
| 62 | 78 | F |  |
| 63 | 75 | M |  |
| 64 | 74 | M |  |
| 65 | 78 | F |  |
| 66 | 69 | M |  |
| 67 | 65 | M |  |
| 68 | 70 | M |  |
| 69 | 72 | F |  |
| 70 | 80 | F |  |
| 71 | 81 | M |  |
| 72 | 79 | M |  |
|  |  | M |  |

The data is collected by Dino’s father, the neurological doctor who help us to collect the data from the Tiantan Hospital from the department of neurology.

7. References:

Quinn, N. P., and F. A. Husain. “Parkinson’s Disease.” *British Medical Journal (Clinical Research Edition)*, vol. 293, no. 6543, 1986, pp. 379–82. *JSTOR*, http://www.jstor.org/stable/29524136. Accessed 18 May 2023.