**Introduction**

Statistics is a subject which will enable us to converting all data into useful information to let us to figure out something important from it. To make a complete statistics, we shall go through three main processes to succeed it. The processes are collecting data, summarizing data and interpreting data. We also can conclude that statistics is a science that related with the collection, classification, analysis and interpretation of information or data. Therefore, in the process of making this assignment, we will use some statistics skill that we learnt before to complete this assignment.

In medieval times, gold was often seen as beneficial for the health, in the belief that something so rare and beautiful could not be anything but healthy. Some gold salts do have anti-inflammatory properties and are used as pharmaceuticals in the treatment of arthritis and other similar conditions. However, only salts and radioisotopes of gold are of pharmacological value, as elemental (metallic) gold is inert to all chemicals it encounters inside the body. In modern times, injectable gold has been proven to help to reduce the pain and swelling of rheumatoid arthritis and tuberculosis.

Linear relationship is widely used in biological, behavioral and social sciences to describe possible relationship between variables. It ranks as one of the most important tools used in this field.

This research is to studying that whether there is a relationship between radioactive gold (195 Au-aurothiomalate) and its length of time it retained in a person’s body.

**Data Analysis**

Original data:

There are 100 patient to be chosen:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Days after Injection | Serum Gold % Concentration | Days after Injection | Serum Gold % Concentration | Days after Injection | Serum Gold % Concentration |
| 6 | 44.2 | 3 | 69.4 | 7 | 27.2 |
| 6 | 48.0 | 4 | 63.7 | 5 | 53.5 |
| 1 | 90.1 | 2 | 81.9 | 1 | 82.5 |
| 6 | 41.2 | 5 | 51.5 | 3 | 66.9 |
| 2 | 72.5 | 4 | 69.0 | 7 | 37.7 |
| 5 | 51.2 | 3 | 78.6 | 7 | 35.4 |
| 3 | 65.7 | 4 | 68.7 | 4 | 69.6 |
| 2 | 75.0 | 5 | 59.6 | 4 | 68.4 |
| 6 | 46.1 | 2 | 73.1 | 4 | 64.1 |
| 2 | 80.1 | 2 | 73.1 | 4 | 69.1 |
| 2 | 75.6 | 7 | 40.8 | 4 | 64.0 |
| 2 | 74.6 | 2 | 76.7 | 4 | 62.6 |
| 2 | 75.3 | 1 | 83.4 | 1 | 83.1 |
| 3 | 65.0 | 2 | 76.6 | 5 | 58.0 |
| 7 | 36.6 | 5 | 58.0 | 1 | 91.4 |
| 4 | 69.1 | 2 | 80.7 | 1 | 86.7 |
| 5 | 64.3 | 5 | 57.2 | 2 | 78.7 |
| 2 | 77.1 | 2 | 80.2 | 2 | 73.7 |
| 3 | 64.6 | 3 | 69.2 | 4 | 61.9 |
| 7 | 45.6 | 5 | 59.9 | 4 | 64.5 |
| 3 | 79.4 | 5 | 58.2 | 2 | 77.4 |
| 5 | 50.4 | 3 | 62.1 | 2 | 77.7 |
| 2 | 73.7 | 3 | 63.7 | 1 | 81.5 |
| 1 | 81.2 | 3 | 65.9 | 1 | 93.8 |
| 1 | 83.3 | 2 | 81.6 | 3 | 69.8 |
| 3 | 60.7 | 7 | 31.7 | 4 | 63.2 |
| 3 | 63.2 | 3 | 63.0 | 5 | 55.8 |
| 2 | 79.0 | 3 | 60.6 | 3 | 67.2 |
| 7 | 43.3 | 6 | 44.5 | 5 | 51.3 |
| 2 | 70.9 | 5 | 58.3 | 5 | 54.5 |
| 7 | 28.0 | 1 | 81.7 | 5 | 59.3 |
| 6 | 40.8 | 7 | 23.1 | 6 | 49.8 |
| 5 | 57.0 | 4 | 61.2 | 5 | 57.1 |
|  |  |  |  | 5 | 58.1 |

1. Data simulation

25 samples chosen is given below:

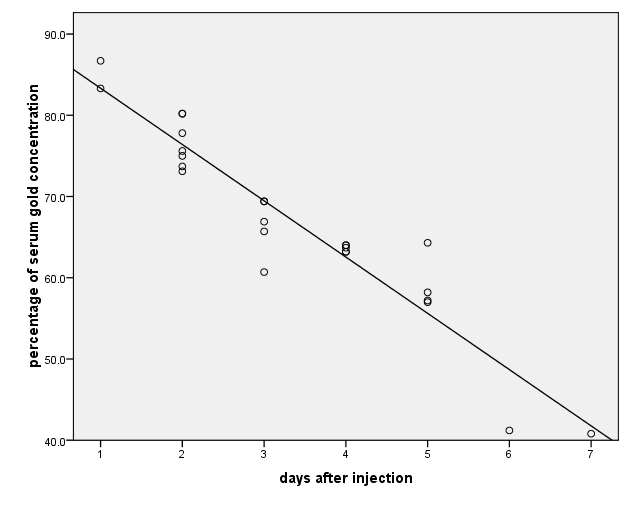
|  |  |  |  |
| --- | --- | --- | --- |
| Days after Injection | Serum Gold % Concentration | Days after Injection | Serum Gold % Concentration |
| 2 | 73.7 | 5 | 58.2 |
| 3 | 69.4 | 4 | 64.0 |
| 5 | 64.3 | 2 | 73.1 |
| 4 | 64.0 | 2 | 75.6 |
| 3 | 65.7 | 2 | 77.8 |
| 5 | 57.0 | 4 | 63.2 |
| 3 | 69.4 | 3 | 66.9 |
| 1 | 86.7 | 2 | 75.0 |
| 7 | 40.8 | 2 | 80.2 |
| 4 | 63.7 | 3 | 60.7 |
| 6 | 41.2 | 2 | 80.2 |
| 5 | 57.2 | 4 | 63.2 |
| 1 | 83.3 |  |  |

1. Estimation of the intercept, b0, and the slope, b1.

b0=90.256

b1=-6.9278

(iii) r= -0.9483



1. To test the significance of linear relationship:

Hypothesis

Ho = β1 = 0

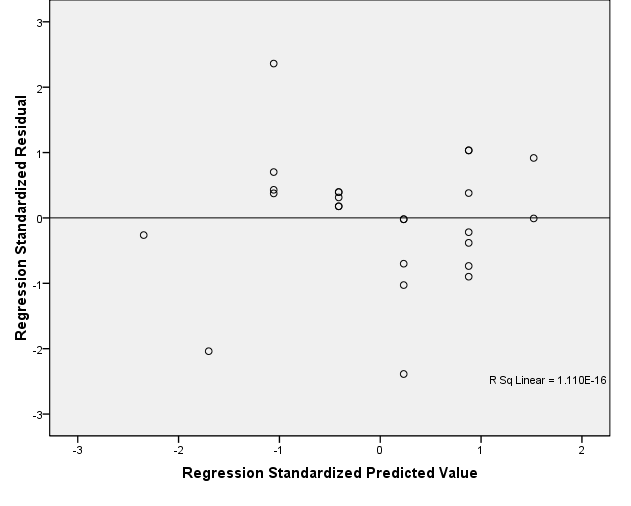
H1 = β1 ≠ 0

Model

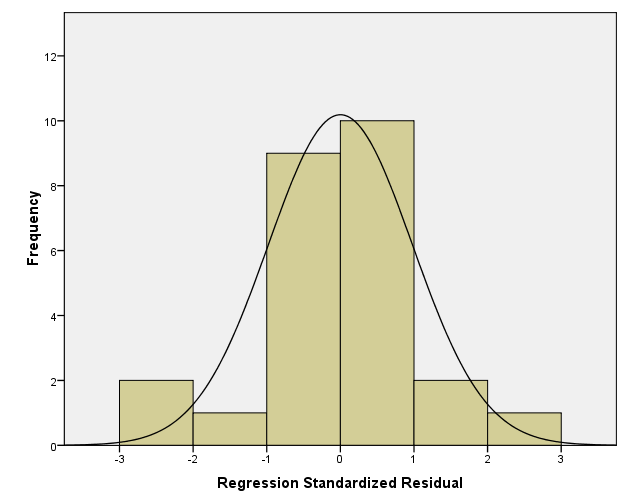
Validity of assumption:

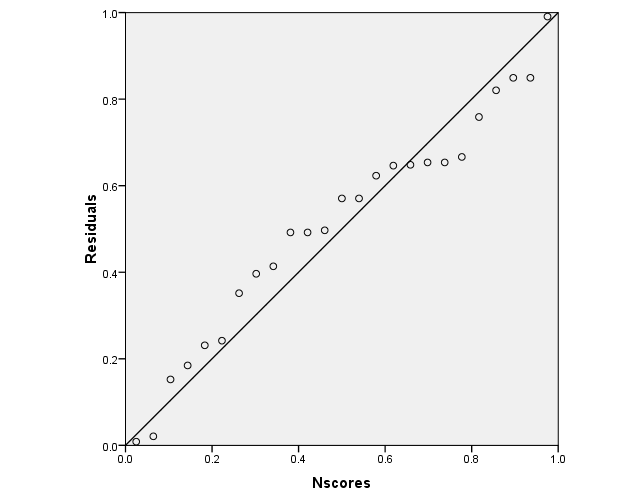
Independence Assumption: These data were not collected over time, so there is no reason to said that the Percentage of Serum Gold Concentration influence the Percentage of Serum Gold Concentration of another.

Does the Plot Thicken? Condition: Neither the original scatterplot nor the residual scatterplot shows any changes in the spread about the line.



Nearly Normal Condition: A histogram of the residuals is unimodal and symmetric. The Normal probability plot of the residuals is quite straight, indicating that the Normal model is reasonable for the errors.





Mechanics

= 57.76

= - 400.12

= 3081.99

= 0.4831

= - 14.3403

p-value = P (tn-2 > ttest)

reject H0 if p-value < α = 0.05

p-value = P (t25-2 < - 14.3403) < 0.01 < 0.05 = α

we should reject H0.

(vi) We should convert all data in y-axis to be log(x) or 1/x to make the graph or scatterplot to become near to linear if nonlinearity exists.

**Conclusion**

A data analysis was made for the study of relationship between radioactive gold(195 Au-aurothiomalate) and its length of time it retained in a person’s body. The data used for the data analysis is based on the serum gold concentrations found in randomly sampled blood taken from patients who were given an initial dose of 50mg.

From the data analysis made, it has been found out that there is a linear relationship between the radioactive gold(195 Au-aurothiomalate) and its length of time it retained in a person’s body. The scatterplot of radioactive gold(195 Au-aurothiomalate) against length of time shows a strong and negative linear relationship. As a conclusion, we can say that the gold concentration decreases as the length of time it retained in a person’s body increases.

**REFERENCE**

1. Applied Statistics for Engineers and Physical Scientists, Johannes Ledolter & Robert V. Hogg, Pearson Education, Inc, 2010.
2. STATISTICS WORKBOOK For Sciences and Engineering, Zalina Mohd Daud, Maizah Hura Ahmad, Robiah Adnan, Shariffah Suhaila Syed Jamaludin, Fadhilah Yusof, Ismail Mohamad, Jabatan Matematik Fakulti Sains University Teknologi Malaysia, 2005.
3. INTRO STATS Third Edition, Richard D. De Veaux, Paul F. Velleman, David E. Bock, Pearson Education, Inc, 2009.

Appendix: Meeting Minute

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| FIRST MEETING | |  | |  | |  | |  | |
| Date | Time | | Venue | | Minutes | | Person In | |
|  |  | |  | |  | | Charge | |
|  |  | |  | | 1.Introduction & Discussion | |  | |
|  |  | |  | | -A brief introduction among  all group members to | |  | |
| 15/12/2011 | 1500-1700 | | C08-413 | | understand the task better. | |  | |
|  |  | |  | | The task was discussed | |  | |
|  |  | |  | | between group members. | |  | |
|  |  | |  | | 2.Data simulation | | NG SIEW HUI | |
|  |  | |  | | -Random sample of 25 pairs  of data were chosen. | |  | |
|  |  | |  | | 3.Calculation | |  | |
|  |  | |  | | -All calculations required for | |  | |
|  |  | |  | | the data analysis is done. | |  | |
|  |  | |  | | 4.Meeting dismissed | |  | |

SECOND MEETING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DATE | TIME | VENUE | MINUTE | PERSON IN |
|  |  |  |  | CHARGE |
|  |  |  | 1.Distribution of task |  |
| 20/12/2011 | 1400-1500 | C08-413 | -Everyone is given a task to be | Fahim |
|  |  |  | completed by three days |  |
|  |  |  | 2.Discussion |  |
|  |  |  | -Suitable method to solve the |  |
|  |  |  | question were discussed. |  |
|  |  |  | 3.Meeting dismissed |  |

THIRD MEETING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date | Time | Venue | Minutes | Person In |
|  |  |  |  | Charge |
|  |  |  | 1.Task completion |  |
|  |  |  | -Each member hand in |  |
| 20/12/2011 | 1000-1200 | C08-413 | their work based on the |  |
|  |  |  | task distributed to them. |  |
|  |  |  | 2.Compilation | Sarah |
|  |  |  | -All the work were |  |
|  |  |  | compiled together to |  |
|  |  |  | form a complete task  set. |  |
|  |  |  | 3.Meeting dismissed |  |