****

**STATISTICS SSCM 1103**

Assignment

Report on the “relationship between radioactive gold and its length of time it retained in a person’s body.”

Lecturer : Dr. Arifah Bahar

Group Name : Group 5

Group Members :

1. *Lau Mun Hing* (A11SC0240)
2. Tan Kuan Wern (A11SC0191)
3. Muhammad Anas Bin Mohamad Sani (AS090122)
4. Siti Mursyida Abdul Karim (A11SC0142)
5. Zetty Azrah Sutirman (A11SC0249)
6. Choong Jing Yee (A11SC0202)
7. Chan Poh Ching (A11SC0209)

**Introduction**

What is radioactive gold? Gold is used as a drug to treat a small number of medical conditions. Injections of weak solutions of sodium aurothiomalate or aurothioglucose are sometimes used to treat rheumatoid arthritis. Radioactive gold is used in diagnosis. It is injected in a colloidal solution that can be tracked as a beta emitter as it passes through the body. Particles of a radioactive gold isotope are implanted in tissues to serve as a radiation source in the treatment of certain cancers. They may also be used occasionally to treat other diseases such as psoriatic arthritis.

To this very day, linear regression is used in most of the fields in Science. It is used to study the relationship between 2 variables, to see how they correlate with one another. With the data collected and the regression model projected, scientist can determine how strong the variables are associated. From there, scientist can construct an equation, and thus be able to predict the outcome of the dependent variable.

Objective :

1. To learn data analysis on Linear Regression Model.
2. To learn the simulation of data.
3. To find the estimation of the intercept and slope using least squared method.
4. To study the significance of linear relationship.
5. To instil communication and team working skills among members.
6. To implement the knowledge we learnt in Computer Literacy class such as SPSS and Microsoft Excel.

**Data Analysis**

If there is a significant linear relationship between *Days After Injection* (independent variable) and *Serum Gold Percentage Concentration* (dependent variable) , the slope will not equal zero.

H0 : b1 ≠ 0

HA : b1 = 0

The [null hypothesis](http://stattrek.com/Help/Glossary.aspx?Target=Null%20hypothesis), H0 states that the slope is not equal to zero, and the alternative hypothesis, HA states that the slope is equal to zero.

2)i) Data Simulation

|  |  |  |
| --- | --- | --- |
| Random Numbers | Days After Injection | *Serum Gold Concentration (%)* |
| 89 | 1 | 81.5 |
| 57 | 3 | 65.9 |
| 44 | 7 | 40.8 |
| 2 | 6 | 48 |
| 89 | 1 | 81.5 |
| 87 | 2 | 77.4 |
| 38 | 4 | 69 |
| 26 | 3 | 60.7 |
| 61 | 3 | 60.6 |
| 50 | 5 | 57.2 |
| 95 | 5 | 51.3 |
| 54 | 5 | 58.2 |
| 68 | 5 | 53.5 |
| 89 | 1 | 81.5 |
| 41 | 5 | 59.6 |
| 73 | 4 | 69.6 |
| 70 | 3 | 66.9 |
| 10 | 2 | 80.1 |
| 11 | 2 | 75.6 |
| 35 | 4 | 63.7 |
| 81 | 1 | 91.4 |
| 91 | 3 | 69.8 |
| 37 | 5 | 51.5 |
| 43 | 2 | 73.1 |
| 25 | 1 | 83.3 |

2)ii) Estimation of the intercept, b0 and the slope, b1.

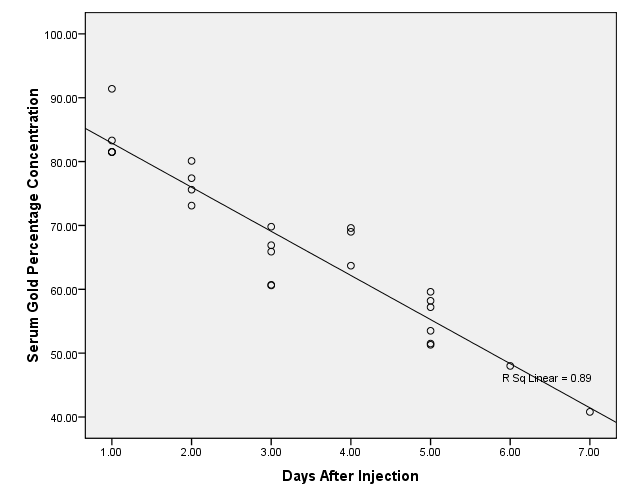
| **Coefficientsa** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 89.808 | 1.890 |  | 47.525 | .000 |
| DAI Days After Injection | -6.910 | .506 | -.944 | -13.662 | .000 |
| a. Dependent Variable: SGC Serum Gold Percentage Concentration | | | | |  |  |

The equation,

Serum Gold Percentage Concentration=-6.91Days After Injection+89.808

The intercept, b0 = 89.808

The slope, b1 = -6.91

2)iii) The linear regression model

2)iv)

| **Correlations** | | | |
| --- | --- | --- | --- |
|  |  | DAI Days After Injection | SGC Serum Gold Percentage Concentration |
| DAI Days After Injection | Pearson Correlation | 1 | -.944\*\* |
| Sig. (2-tailed) |  | .000 |
| N | 25 | 25 |
| SGC Serum Gold Percentage Concentration | Pearson Correlation | -.944\*\* | 1 |
| Sig. (2-tailed) | .000 |  |
| N | 25 | 25 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | |  |

Based on the estimation done earlier we found that,

Serum Gold Percentage Concentration=-6.91Days After Injection+89.808

The intercept, b0 = 89.808

The slope, b1 = -6.91

The slope of a negative number indicates that as *Days After Injection* increases, *Serum Gold Percentage Concentration* decreases.

Based on the linear regression model, we can see that the relationship between the *Days After Injection* and *Serum Gold Percentage Concentration* is linear and negative.

Based on the correlation table, we can see that the Pearson Correlation coefficient is equals to -0.944. This indicates that the relationship between *Days After Injection* and *Serum Gold Percentage Concentration* is strong, negative and linear because its value is near to -1.

| **Correlations** | | | | |
| --- | --- | --- | --- | --- |
|  |  |  | DAI Days After Injection | SGC Serum Gold Percentage Concentration |
| Spearman's rho | DAI Days After Injection | Correlation Coefficient | 1.000 | -.952\*\* |
| Sig. (2-tailed) | . | .000 |
| N | 25 | 25 |
| SGC Serum Gold Percentage Concentration | Correlation Coefficient | -.952\*\* | 1.000 |
| Sig. (2-tailed) | .000 | . |
| N | 25 | 25 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | |  |  |

Based on the Correlations table above, the Spearman Correlation coefficient is -0.952 which is quite similar to Pearson Correlation coefficient of -0.944. this shows that our assumption is correct.

2)v) Validity of the assumption

The negative coefficient shows that the relationship between *Days After Injection* and *Serum Gold Percentage Concentration* is negative. The value of -0.944 from the Pearson Correlation and -0.952 from the Spearman Correlation show that the relationship is strongly linear. Moreover, the probability of being wrong for assuming the relationship of the samples reflects the population is 0. This indicates that the linear relationship done from the sample is equals to population. This supports our H0 which means that there is a linear relationship.

2) vi) Non – linearity

The graph shown is linear.

If nonlinearity exists in the graph, there are few ways to tackle the problem. Firstly, we have to transform the data. Since our dependent variable is from percentage, we transform those data into log form in order to obtain a linear graph. There are also other methods to tackle nonlinearity such as transforming data into their reciprocals, squares, square roots and etc.

**Conclusion**

The negative coefficient shows that the relationship between *Days After Injection* and *Serum Gold Percentage Concentration* is negative. The value of -0.944 from the Pearson Correlation and -0.952 from the Spearman Correlation show that the relationship is strongly linear. Moreover, the probability of being wrong for assuming the relationship of the samples reflects the population is 0.000. This indicates that the linear relationship done from the sample is equals to population.

These results obtained from the analysis support our assumption that the graph is a strong negative linear graph and analysis is correct. We can conclude that there is a significant linear relationship between day injection and serum gold percentage concentration. When the percentage concentration serum gold increase, the less day it retained in a person body. Using the equation we obtained from the data analysis, we can predict the percentage of gold serum concentration in a person’s body using different values for the number of days after injection. An R-squared value of 0.89 indicates that the number of days after injection accounts for 89% as 1 of the factors affecting the percentage of gold serum concentration. There are other factors such as individual body metabolism, living environment, etc.

**Reference**

* **BASIC STATISTICS** FOR SCIENCES AND EDUCATION, FADHILAH YUSOF, ZALINA MOHD DAUD, MAIZAH HURA AHMAD, ROBIAH ADNAN, ZARINA MOHD KHALID, ARIFAH BAHAR, NORHAIZA AHMAD, PEARSON MALAYSIA SDN. BHD. 2009.
* **MATHEMATICS AND STATISTICS FOR TECHNOLOGIES,** HENRY GEORGE CURNING, C J ANSON, NEW YORK, 1967.
* **INTRODUCTORY STATISTICS,** JOHN A. INGRAM, CAMMINGS PUBLISHING COMPANY 1974.
* **STATISTICS FORMULAE AND TABLES,** DR. MUHAMMAD H. LEE, UNIVERSITY TECHNOLOGY MALAYSIA.
* **APPLIED STATISTICS FOR ENGINEERS AND PHYSICAL SCIENTISTS,** JOHANNES LEDOLTER, ROBERT V.HOGG, PEARSON EDUCATION, INC, 2010.
* <http://geology.com/minerals/gold/uses-of-gold.shtml>
* <http://www.rheumatology.org.au/downloads/gold230811.pdf>
* <http://en.cnki.com.cn/Article_en/CJFDTOTAL-XDYQ200903025.htm>

**Appendix**

Group 5

Minutes for group discussion 1

Members:

1. Lau Mun Hing (A11SC0240)
2. Tan Kuan Wern (A11SC0191)
3. Muhammad Anas Bin Mohammad Sani (AS090122)
4. Siti Mursyida Abdul Karim (A11SC0142)
5. Zetty Azrah Sutirman (A11SC0249)
6. Choong Jing Yee (A11SC0202)
7. Chan Poh Ching (A11SC0209)

Reports:

1. Leader explains the details of the statistics assignment to all members.
2. He also assigns works to each member.

**Introduction**- Tan Kuan Wern

**Data analysis**- Chan Poh Ching , Choong Jing Yee , Zetty Azrah Sutirman, Siti Mursyida Abdul Karim

**Conclusion**- Muhammad Anas Bin Mohammad Sani

**Reference and Appendices**-Lau Mun Hing

1. He sets the date for our next meeting.

* Tuesday, 20th of Dec 2011
* Further discussion
* Check for mistakes
* Solve any problem that we may encounter.

1. Meeting is adjourned at 1620.

Minutes for group discussion 2

Members:

1. Lau Mun Hing (A11SC0240)
2. Tan Kuan Wern (A11SC0191)
3. Muhammad Anas Bin Mohammad Sani (AS090122)
4. Siti Mursyida Abdul Karim (A11SC0142)
5. Zetty Azrah Sutirman (A11SC0249)
6. Choong Jing Yee (A11SC0202)
7. Chan Poh Ching (A11SC0209)

Reports:

1. Problem Solving:

We discussed about problems we accounted during the process of work such as how to tackle non-linearity.

1. Checking for mistakes:

We went through our work once more to identify any mistakes or errors.

1. Introduction and conclusion:

We discussed about suitable points for introduction and conclusion.

1. Meeting was adjourned at 11.30am.

Minutes for group discussion 3

Members:

1. Lau Mun Hing (A11SC0240)
2. Tan Kuan Wern (A11SC0191)
3. Muhammad Anas Bin Mohammad Sani (AS090122)
4. Siti Mursyida Abdul Karim (A11SC0142)
5. Zetty Azrah Sutirman (A11SC0249)
6. Choong Jing Yee (A11SC0202)
7. Chan Poh Ching (A11SC0209)

Reports:

1. Finalizing of assignment:

We combined each member’s work to obtain a complete set of assignment.

1. Reviewing of assignment:

We checked the page sequences arrangement, spelling errors and did final changes to parts that are necessary.

1. Soft copies of the assignment are given to each member for reference purposes. A hard copy will be printed out to be handed in for assessment. A copy of assignment will be uploaded to the E-portfolio.

Summary of group work

Firstly, we had a group discussion to distribute works to each member. There were no arguments or disagreements among members. Everyone was very cooperative and they carried out their works accordingly.

We had three group meetings. Each meeting minutes was uploaded to the E-portfolio by different group members. The assignment was fully done on time, thanks to the members’ cooperation. We solved problems that we encountered and also discovered mistakes made during the process of working together. All group members contributed ideas during each discussion. We also discussed the ideas whether we accepted the ideas or not.

We tried to work together to do the assignment perfectly as we hoped. After doing the assignment, we found that we know more about linear regression model including the assumptions and linearity of the graph. We feel thankful because Dr.Arifah Bahar gave us a good chance to learn more through this statistics assignment.