**NCD course：”Biostatistics” Assignments (Report)**

*Watch the on-demand lecture\* and answer the following 1 to 9 questions.*

*\*You may skip the exercises parts using the statistical analysis package if necessary.*

1. **t-test, Wilcoxon test**

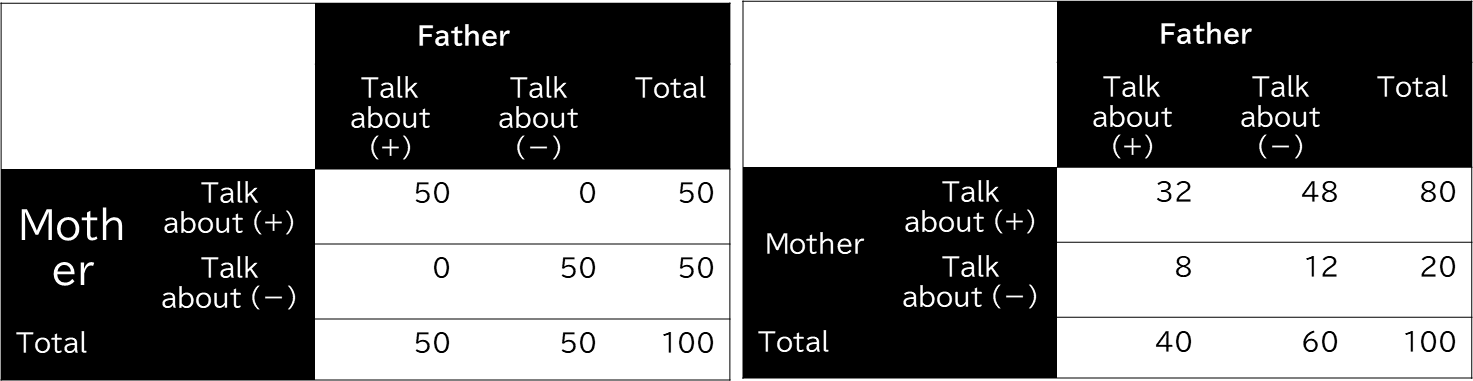
Describe the characteristics of the two tests (t-test and Wilcoxon test) with respect to outliers.

You may use the exercises from the lecture and explain them by showing the results of your analysis.

**2. Describe the analysis method that should be used in the following case study and the reasons for its use.**

We asked junior high school students, "With whom do you feel more comfortable discussing your problems, your father or your mother? The same number of people（n=100）answered each question, and the results are shown in A and B.  
For each of the results of A and B, consider whether there is a difference in whether they would choose their fathers or mothers as their consulting partners.  
Indicate the method of analysis that should be selected and the results of the test.

<A> <B>



<A>

* Consultation rate was 50% for both fathers and mothers.
* In the fathers' consultation group, the consultation rate for mothers is 100%.
* The consultation rate for mothers in the non-father consultation group was 0%.

<B>

* Consultation rate was 40% for fathers and 80% for mothers.
* In the father consultation group, the mother consultation rate was 80%.
* In the father no-consultation group, the consultation rate for mothers was 80%.

**３．Describe what should be noted in the selection of explanatory variable (independent variable) in multiple regression analysis (general linear model).**

**4．Please answer the following questions by placing the appropriate word or phrase in parentheses**

**A Generalized Linear Model consists of three components**

1. An exponential family of **probability distributions**

①General linear model :（ )

②Logistic model: ( )

③Poisson model: ( )

1. A linear predictor
2. **A link function** is to transform an expression into a linear one.

①General linear model :（ )

②Logistic model: ( )

③Poisson model: ( )

**5. Explain the following words and phrases.**

1)Sensitivity

2)Specificity

3)Positive predictive value

4)Negative predictive value

5)Receiver Operating Characteristic(ROC)

**6. Explain what “over-dispersion” often occurs in Poisson regression.**

**7.The progress of the five person (A to E) who participated in the clinical trial with a study duration of 12 months was as follows. Answer Q7-1, Q7-2**

<A to E person(n=5)>

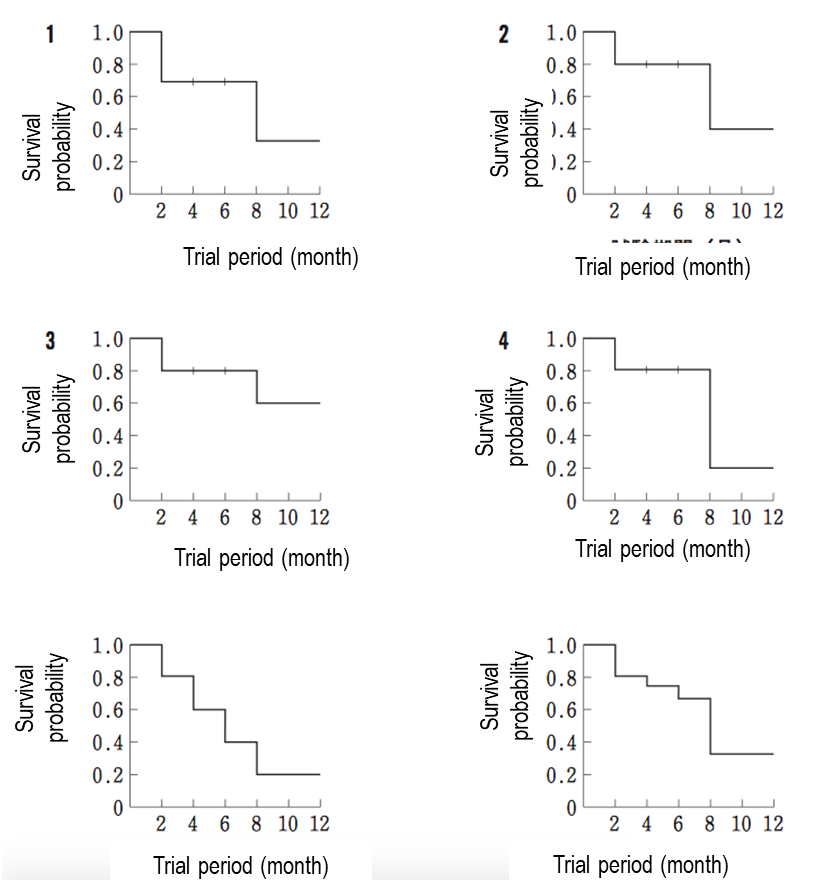
A died after 2 months,

B was lost to follow-up after 4 months,

C was lost to follow-up after 6 months,

D died after 8 months,

E survived until the end of the study at 12 months



**Q7 -1**

**Which of the following is the correct survival curve expressed using the Kaplan-Meier method?**

**Q7-2**

**Construct a life table consisting of survival probability and cumulative failure probability.**

**8．Consider a small prospective cohort study designed to study time to death.**

|  |  |  |
| --- | --- | --- |
| **Participant Identification Number** | **Year of Death** | **Year of Last Contact** |
| 1 |  | 24 |
| 2 | 3 |  |
| 3 |  | 11 |
| 4 |  | 19 |
| 5 |  | 24 |
| 6 |  | 13 |
| 7 | 14 |  |
| 8 |  | 2 |
| 9 |  | 18 |
| 10 |  | 17 |
| 11 |  | 24 |
| 12 |  | 21 |
| 13 |  | 12 |
| 14 | 1 |  |
| 15 |  | 10 |
| 16 | 23 |  |
| 17 |  | 6 |
| 18 | 5 |  |
| 19 |  | 9 |
| 20 | 17 |  |

The study (aging study) involves 20 participants; they are enrolled over a 5 year period and are followed for up to 24 years until they die, the study ends, or they drop out of the study (lost to follow-up).

In the study, there are 6 deaths and 3 participants with complete follow-up (i.e., 24 years).

The remaining 11 have fewer than 24 years of follow-up due to enrolling late or loss to follow-up.

**Q. Calculate and fill in the following tables**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Time, Years | Number at Risk  Nt | Number of Deaths  Dt | Number Censored  Ct | Survival Probability |
| 0 |  |  |  |  |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 5 |  |  |  |  |
| 6 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |
| 11 |  |  |  |  |
| 12 |  |  |  |  |
| 13 |  |  |  |  |
| 14 |  |  |  |  |
| 17 |  |  |  |  |
| 18 |  |  |  |  |
| 19 |  |  |  |  |
| 21 |  |  |  |  |
| 23 |  |  |  |  |
| 24 |  |  |  |  |

**9．Describe the difference between “Principal Components Analysis” and “Factor Analysis”.**

**That is all for the questions.**

**Both of these questions can be answered without the use of the statistical package by viewing the on-demand materials.**

**Submit via Webclass or e-mail(aharada@belle.shiga-med.ac.jp)**