Assignments – Semester 1 – Basic Statistics

Dr Wan Nor Arifin
Unit of Biostatistics and Research Methodology, Universiti Sains Malaysia.
wnarifin@usm.my

Probability

Probability Q1

In a randomized controlled trial, a researcher wishes to allocate the participants into group A and B. He decides to allocate them into equally-sized groups by using block randomization scheme, with block size of 6.

a) Determine number of possible block combinations. (consisting of two As and two Bs)
    *Hint: In how many number of ways 6 subjects can be allocated into 2 groups A and B?

b) List down the blocks.

c) Calculate the probability of getting block starting with 'B'.

Probability Q2

Briefly describe about permutations and combinations. Give two examples for each in relation to medical and health sciences.

Probability Q3

Briefly explain about conditional probability. Give an example in relation to medical and health sciences.

Probability Q3

A scientist developed a toolkit to test H1N1 infection. Probability of test positive among those infected with H1N1 is 85%, while the probability of test positive among those uninfected is 25%.

A specialist in infectious disease claim that based on his experience, the probability of Abu being infected with H1N1 is 70%.

In the light of this new knowledge, calculate the probability that Abu is infected with H1N1 given that using H1N1 kit, he is tested positive.

Probability Distribution Q1

Suppose in Unhealthy Corporation, it is known that 35% of the staffs are Vape smokers. If a random sample of 30 is selected among the staffs, calculate the probability that:

a) Exactly 11 staffs are smokers.

b) Less than 11 staffs are smokers.

c) 11 or more staffs are smokers.

Probability Distribution Q2

Suppose the number of new Vape smokers in Malaysia is on average 2.5 per day and it was found that the daily distribution follows Poisson distribution.

Calculate the probability that any randomly selected day will be the one

a) 3 new smokers.

b) 4 or less new smokers.

c) 5 or more new smokers.

Based on the same information, calculate the probability that any randomly selected 3 days will be the one with collectively 8 new smokers.

Sampling Distribution

Sampling Distribution Q1

In human population it is known that diastolic blood pressure is normally distributed with mean of 80mmHg and a standard deviation of 30mmHg. Calculate the probability that a random sample of 50 subjects from the population will have:

    a) Mean less than 75 mmHg?

    b) Mean between 75mmHg and 85mmHg?

    c) Mean of more than 85mmHg?

    d) Mean of 85mmHg?

Sampling Distribution Q2

In a Malaysian population, it was found that 360 people out 900 population are hypertensive. Calculate the probability that a random sample of 100 subjects from the population will have:

    a) Prevalence of hypertension of less than 35%.

    b) Prevalence of hypertension of more than 35%.

    c) Prevalence of hypertension between 35% to 45%.

Estimation

Estimation Q1

A researcher collected data on diastolic blood pressure (DBP) from a sample of 80 patients, he found that the mean DBP was 75mmHg with standard deviation of 17mmHg. Estimate with 90%, 95% and 99% confidence the population parameter.

Estimation Q2

In a study, a researcher presented his finding on systolic blood pressure (SBP) collected from 40 patients as 115mmHg (95% CI: 110.52, 119.48). He also reported that the standard deviation for the sample was 14mmHg.

Determine whether he calculated the 95% confidence interval correctly. Justify.

Estimation Q3

In a study on obesity among 321 adolescence in Kelantan, 35% of them were found to be obese. Construct 90%, 95% and 99% confidence interval for the percentage of obesity.

Estimation Q4

In construction of confidence interval for single proportion using standard normal distribution, it follows that it is applicable as long as both np and $n(1-p) > 5$.

In a table, list down and briefly describe the alternatives when we cannot fulfill this prerequisites to calculate confidence interval around a point estimate of proportion.

Hypothesis Testing

Hypothesis Q1

A researcher is going to conduct a study to compare mean fasting blood sugar (FBS) (in mmol/L) between a group of lecturers and a group of medical students. If the researcher is not willing to make assumption whether any of the group has higher or lower mean FBS, state:

    a) Research hypothesis.
    b) Statistical hypothesis.

Hypothesis Q2

A researcher is going to conduct a study to compare proportion of obesity between a group of lecturers and a group of medical students. If the researcher is not willing to make assumption whether any of the group has higher or lower proportion of obesity, state:

    a) Research hypothesis.
    b) Statistical hypothesis.

Hypothesis Testing Q3

Explain the following concepts:

    a) Type I Error (α)

    b) Type II Error (β)

    c) Significance level

    d) Power of a test

    e) *P*-value

Hypothesis Testing Q4

In medical literature, it is known that there is a trend of moving away from the practice of reporting *P*-value to indicate significance of a test towards to reporting of confidence interval (e.g. reporting of independent *t*-test result).

Describe advantages and disadvantages of reporting confidence interval only as compared to the traditional approach of reporting *P*-value.