Statistical Inference and Multivariate Analysis (MA324)

Lecture Slides

Lecture 16

Hypothesis Testing

Indian Institute of Technology Guwahati

Jan-May 2023

Few Questions:

- How to decide a **new drug** is better than the existing drug?
- Is the biomarker (with **lower cost**) is equivalent with another one (higher cost) in efficiency?
- You have developed a **novel diagnostic tool** to detect a specific disease. How do you establish it's efficiency?
- How to decide the strength of an **iron-rod** is better than the standard iron-rod available in market?
- How to decide number of **spam mails** is under control in the server?
- How to decide the **non-performing asset of a bank** is under control?
- How to check that your class **has done better in the end-sem compared** to the mid-sem exam in this semester?
- When to **release a tsunami signal** based on the repeated wave signals?
- How does a jury **determine an accused criminal is guilty**?

Testing of Hypothesis:

- In point estimation, we **try to find meaningful guesses** for unknown parameters or parametric functions.
- In testing of hypothesis, we **do not guess** the value of the parametric function. We try to **check if a given statement about parameters is true or not**.
- Let us start with examples.

Example 1:

- Pharmaceutical companies use hypothesis testing **to test if a new drug is efficient**.
- To do so, a group of patients are randomly **divided into two groups**. One of the groups is administered with placebo and the other is administered with the drug. The first and the second groups are called **control group** and **test group**, respectively.
- Assume that the drug is a cough syrup. Let μ_1 denote the **expected number of expectorations per hour** after a patient has used placebo and μ_2 denote the expected number of expectorations per hour after a patient has used the syrup.
- We want to know if $\mu_2 < \mu_1$. In this case, **two expectations are compared**.

- One of the method to attack this problem is to draw RS from both the groups. Let $X_1, X_2, \ldots, X_{n_1}$ denote a RS of size n_1 from the control group. Let $Y_1, \ldots, Y_{n_2}$ denote a RS of size n_2 drawn from test group.
- We want to test if $\mu_2 = \mu_1$ or $\mu_2 < \mu_1$. If $\mu_1 = \mu_2$, then the **new drug is not efficient**. If $\mu_2 < \mu_1$, then the **new drug has some effect**.
- Note that we have taken $\mu_2 < \mu_1$ and we have not considered if $\mu_2 > \mu_1$. The reason for the same is as follows: As we are trying to **check the efficiency of the new drug**, we implicitly assume that $\mu_2 \leq \mu_1$.
- Heuristically, we should compare $\bar{X}$ and $\bar{Y}$.

Example 2:

- Let a coin is tossed 80 times, and head are obtained 55 times. Can we conclude that the **coin is fair based on this data**?

- Let, for $i = 1, 2, \ldots, 80$, X_i be an indicator RV, which takes value 1 if ith toss is a head and takes value zero if the ith toss is a tail. Then, we have a RS $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p)$, where p is the probability of getting a head in a toss.

- We want to test $p = 0.5$ or $p \neq 0.5$. Intuitively, it makes sense to use $\bar{X}$ to check if $p = 0.5$ or not.

- Here, the observed value of $\bar{X}$ is $\bar{x} = 55/80 = 0.6875$. If p is actually equal to 0.5, then, using CLT, we have

$$T = \frac{\sqrt{n}\left(\bar{X} - 0.5\right)}{\sqrt{0.5 \times (1 - 0.5)}} \approx N(0, 1).$$

- Now, if the **number of heads** is **too small** or **too large** (*i.e.*, the value of $\bar{x}$ is not close to 0.5), we should go for **biased coin**.
- If the number of head is **moderate** (*i.e.*, the value of $\bar{x}$ is close to 0.5), we should choose that the **coin is fair**.
- In the first case ($\bar{x}$ not close to 0.5), the absolute observed value of T will be **large**. The absolute observed value of T will be **close to zero** in the second case.
- This discussion suggests that we should **reject the fact that the coin is fair** if $|T| > C$ for some appropriate real constant C.
- Here, the observed value of T is 3.3541, which is **too extreme** with respect to a standard normal distribution as $P\left(|Z| > 3.35\right) \approx 0.0008$, where $Z \sim N(0,\,1)$. Therefore, it is quite **reasonable to reject** the hypothesis $p = 0.5$ based on the data.

Example 3: A coin is tossed 80 times, and head are obtained 35 times. Can we conclude that the coin is significantly fair?

- Here, the observed value of T is -1.1180. Data **do not suggest to reject** the fact that the coin is fair, as the observed value of T is not extreme with respect to a standard normal distribution. Note that $P\left(|Z| > 1.11\right) = 0.267$.

- In the last two examples, we have talked about **extreme or not extreme**. The question is: Which values are considered as extreme and which are not? More precisely, we are rejecting $p = 0.5$ if the observation belong to the set
$$\left\{\boldsymbol{x} : \left|T\left(\boldsymbol{x}\right)\right| > C\right\}.$$

- What **value of C should we choose** so that we can make correct decision? This issues will be discussed as we proceed.