

## Perspective

## The Life of P-‘value’

Shekhawat KS<sup>1\*</sup> and Chauhan A<sup>2</sup><sup>1</sup>Department of Public Health Dentistry, Srinivas Institute of Dental Sciences, Rajiv Gandhi University of Health Sciences, India<sup>2</sup>Department of Oral Biology, Faculty of Dentistry, Manipal University, India**\*Corresponding author:** Shekhawat KS, Department of Public Health Dentistry, Rajiv Gandhi University of Health Sciences, Srinivas Institute of Dental Sciences, Mukka, Karnataka:574164, India**Received:** November 18, 2016; **Accepted:** November 22, 2016; **Published:** November 25, 2016

## Perspective

Post-graduate students and research scholars are often seen scratching their scalps over, what to them seems like the most important question of the millennium, Are the results statistically significant? The emotions and expectation associated with their results under the banner of, “statistically significant” is much more than the bewilderment and excitement on winning a lottery ticket. Significant results makes the post-graduate students feel more confident when confronted by their guides, and few even take it as good omen, that guarantees publication of their research work [..... which is not always true].

Now how and when do we say that results are statistically significant?

The general notion is:

Post-graduate students and research scholars (hence forth referred as students) often ask their respective statisticians regarding P-value and the usual reply is, ‘take P-value as 0.05’. When requested to explain the importance and use of P-value, statisticians (and other part time bio-statisticians) think for a while and then reply in a calm tone, “if P-value is less than 0.05, the results are significant”, and if more than 0.05 it indicates that the happiness of the student is short-lived. The criticism received by student from their research guide is often more than criticism received by any cricket team on losing a match to a rival team.

Jokes apart, a P-value are just the end result of a long process. The long process is usually the research work carried out which involves in simple words: data collection, data entry & analysis and interpretation of data. A couple of stages preceding to data collection also needs to be recognized and understood, which at times both the students and their respective guides fail to enlighten themselves with. The enlightenment needs to be *not* on P-value but on factors which decide the P value.

So continuing with the problem.....what is P-value?

Even though the exact terminology may differ, one generally agreed upon definition is that we are ultimately testing the null hypothesis against a level of significance ( $\alpha$ ) set by the researcher

[1]. P-values are the probabilities of obtaining an effect atleast ‘as extreme as’ the one in your sample data, assuming the truth of the null hypothesis [2]. Testing the null hypothesis. Now what is Null hypothesis?

At this point we have to introduce some terms regarding hypothesis testing..... (Which are for some students like bouncers in a cricket match and way over your helmet?). Hypothesis, null hypothesis and alternate hypothesis.

- Hypothesis (H) means assumption. These assumptions are based on previous observations/research. But since every assumption is influenced by number of factors, we need to be sure whether these factors are actually *THE FACTORS*. These assumptions are made with respect to time, place and person.

- Null Hypothesis (H<sub>0</sub>) is the step-brother of Hypothesis, so the step-brother quotes the exact opposite. For e.g., if hypothesis says, Pandavas of Mahabharata were considered to be good beings, then null hypothesis states, Pandavas of Mahabharata were not considered to be good beings. The null hypothesis indicates no association between the investigated factors or characteristics. Then it becomes the obligation of the researcher to subject the null hypothesis to hypothesis testing. Once the hypothesis is tested, based on the P value we either reject or accept the null hypothesis.

- On the other hand, once we accept the null hypothesis, it is termed as the alternative hypothesis that indicates an association between the investigated variables.

It is synonymous with the Judiciary system, where the judge does not oversee the proceedings with a predetermined mind that the convict is guilty. Rather the judge assumes that the convict is 100 percent innocent, and then the charges on the convict are either approved or rejected. Only if overwhelming evidence of the person’s guilt can be shown, is the judge expected to declare the person guilty-otherwise the person is considered innocent [3]. Depending on the degree of charges proved, the judge announces the verdict. The crux here is to understand, that the step brother (null hypothesis) is given importance and then attempts (statistical tests) are made which can either accept or reject the step brother which is based on P value.

## Hypothesis testing

When a hypothesis is being tested, (e.g. jury trial) the responses can be in any of the following domain:

1. Person is innocent [answer is right, True positive]
2. Person is guilty [answer is wrong, True negative]
3. Person is innocent, but judge finds guilty [False Positive] = Type I error, ( $\alpha$ ) alpha error
4. Person is guilty, but judge finds innocent [False Negative] = Type II error, ( $\beta$ ) beta error

[Few select journals also mandate their authors to provide the

power of the study. The power of the study is same as the sensitivity of the study or  $1 - \beta$  (One minus Beta)].

Regardless of the statistical technique and/or the type of design used, P-values are often reported in clinical research. The researcher will recognize it as, “.sig” in the print out. This P value is compared with a priori alpha ( $\alpha$ ), which serves as a basis for rejecting or accepting the null hypothesis. This level of significance is designated by the researcher [1].

Most students thank (*their respective gods, and probably make offerings*) when they get an opportunity to put, asterisk (\*) if the P values are below 0.05 ( $\alpha$ -level, generally recommended). They even follow the unintentional trickery of using: \*P<0.05, \*\*P<0.01 and \*\*\*P<0.001. Such asterisks, goads the researcher into conclusions that their findings are more significant and hence more powerful [1] (few scientific journals do not support this concept and rather insist on other factor like effect size and confidence interval [4,5]. There are also few journals available that encourage those results which are, ”not statistically significant”).

But there are few unlucky ones who might not have the chance to put asterisk (\*) under “.sig” (*probably, due to an astronomical misalignment of planets.....as believed by a segment of population*). Let’s say for e.g. the P-value what they get is 0.15 (as against 0.05, as set by the researcher). The p-value of 0.15, means that the observed difference (between the variables or factors) can be attributed to chance by 15%. In such cases, the null hypothesis is not rejected, rather it becomes the alternate hypothesis. Also, failing to reject the null hypothesis ( $P>0.05$ ) does not provide corroborating evidence

for the non-existence of the phenomenon. In other words, failing to reject null hypothesis does not prove the null hypothesis (...and nevertheless, it is not an offence if your P value is more than 0.05).

Just to be clear...

P-value is one of the widely used statistical terms in decision making in biomedical research, which assists the investigators to conclude about the significance of a research [5]. This communication is not ‘for’ and/or ‘against’ P value, rather a perspective. The main objective behind this write up was to provide a gist of P-value so that young researchers and post-graduate students have a clear concept and can easily dodge the questions by their guide on P-value during their research work.

PS: The euphemism and examples given are just to create an environment of humor so that readers can relax, smile a bit and understand P-values. Further reading is recommended for in-depth understanding.

## References

1. Glaser DN. The controversy of significance testing: Misconceptions and alternatives. *American Journal of Critical Care*. 1999; 8: 291-296.
2. Schervish MJ. P Values: What they are and what they are not. *Am Stat*. 1996; 50: 203-206.
3. Ghermann J. Course, Statistic 50 content. Syllabus, Hypothesis testing, summer. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Intern Med*. 1988; 108: 258-265.
4. Panagiotakos DB. The Value of P-Value in Biomedical Research. *The Open Cardiovascular Medicine Journal*. 2008; 2: 97-99.