SHOULD STUDIES PUBLISHED IN ACADEMIC JOURNALS USE P-VALUES?

The use of p-values in recent years has really become controversial to determine statistical significance in the academic circle. According to these detractors, p-values are often over-emphasized or in some cases even abused. So, what exactly are these p-values and what's the controversy all about? A *P* value is the probability of observing a hypothetical parameter at least as extreme as the one observed due to chance alone. When we calculate a p-value, we're actually putting to the test what's known as the null hypothesis. P-value actually calculates the probability of the Null Hypothesis happening. Say for example, there was a weight loss study run out on a group of 50 people. 25 in the treatment group and 25 in the control group. Then the null hypothesis would be that **there was no difference** in the mean weight loss between the 2 groups. And a p-value < 0.05 would provide some evidence of some difference between the 2 groups. Now coming to the controversy. Most academics believe that Businesses and some times, researchers, abuse this cutoff value of 0.05 of the p-value to prove that they have indeed uncovered something unusual or remarkable. This is done via p-hacking where these people conduct an analysis and then reconduct it with some minor alterations until a statistically significant finding is obtained. Let's consider a study where we see the use of p-values and how 'important' they are to the study.

A study on the significance of predictor variables of the outcome of a One day International(ODI) cricket match was carried out for the 8 major teams in the cricketing world. (Link to the study is in the description below). Predictor variables like Home-Field Advantage(HF), Coin-toss result(),Bat first or second, and Day Or Night game were evaluated for their quantifiable significance to the outcome of a game using Logistic Regression and 'Classification and Regression Tree'. P-values were important to this study. They are given in the bracket for each team below the regression coefficients.

Table 1

Estimated home-field advantage, effects of coin-toss win, day vs day-night games, and bat-first vs bat-second from model (1). Corresponding *p*-values are reported in the bracket

Team	HM	TS	BF	DN
Australia	0.676	–0.372	0.136	0.419
	(0.01)	(0.17)	(0.62)	(0.16)
India	0.513	0.232	–0.361	0.004
	(0.03)	(0.32)	(0.12)	(0.99)
South Africa	0.796	–0.066	–0.214	0.522
	(<0.01)	(0.81)	(0.43)	(0.05)
Sri Lanka	0.806	0.419	–0.231	0.693
	(<0.01)	(0.10)	(0.37)	(0.01)
England	0.145	0.005	–0.575	0.566
	(0.57)	(0.98)	(0.02)	(0.03)
New Zealand	1.014	0.251	–0.658	0.694
	(<0.01)	(0.38)	(0.02)	(0.02)
Pakistan	0.498	–0.012	0.086	0.007
	(0.11)	(0.97)	(0.76)	(0.98)
West Indies	0.254	0.225	–0.176	0.453
	(0.43)	(0.39)	(0.51)	(0.22)

Based on the reported p-values, South Africa, Sri Lanka and New Zealand have the most significant home-field advantage. Here we encounter the common problems associated with p-values in statistical analysis. There is nothing quantifiable associated with the

variables. I mean, what I would want to know would be something like suppose home field advantage is a statistically significant variable, then what percentage of games would a team win if it is playing at home. P-value tell us nothing about that. Confidence intervals are meant for those kind of things and the author fails to emphasize on any confidence intervals in the study. He is not alone. This is a mistake a lot of people looking to show some breakthrough idea or trying to prove a point through their research commit. India has a p-value of 0.03 and Sri Lanka has a p-value of <0.01 for home field advantage. This doesn't tell us that Sri Lanka is more likely to win at home as compared to India. It just tells that home field advantage is statistically significant in deciding game outcomes for India and Sri Lanka. But by how much? I don't know.. The author doesn't either. Otherwise he would have mentioned something about confidence intervals. Also the author doesn't take into account latent variables like form of a player into account. These might be affecting the outcome variable and thereby producing a low p-value.

So, to conclude, p-values are not sole conclusive evidences that can help us pick statistically significant variables in deciding an outcome. They can just be treated as the initial evidence that something interesting might be happening if we get a p-value<0.05. But we definitely need to accompany our p-values with confidence intervals because confidence intervals relate directly to our response OR Outcome variable which is the variable of interest. Even after that, we should wait for someone else to reproduce the same results before jumping to any strong conclusions because our conclusions might be partly based on some latent variables that we did not take into account.