

Pivotal

Manish Tripathi <mtripathi@pivotal.io>

Thursday Thunder: Week 5, Learning 5: p-Values

Manish Tripathi <mtripathi@pivotal.io>
To: PDS-ALL <PDS-ALL@pivotal.io>

Fri, Dec 11, 2015 at 12:35 AM

Hi All,



Yeah I know some of you would have started cursing me again on seeing the subject line of this email. But that's fine. I am pretty shameless. So your cursing won't make me stop troubling you all. :D

So topic for today:

What are p-values?.

Here is one definition. Makes quite sense right? :)

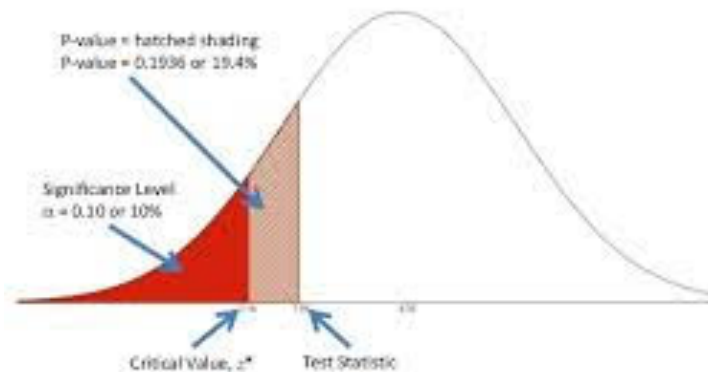
<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001]— HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04]— SIGNIFICANT
0.049	
0.050]— OH CRAP. REDO CALCULATIONS.
0.051	
0.06]— ON THE EDGE OF SIGNIFICANCE
0.07	
0.08]— HIGHLY SUGGESTIVE, SIGNIFICANT AT THE $P < 0.10$ LEVEL
0.09	
0.099]— HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS
≥ 0.1	

On a serious note lets start with what we do in a typical hypothesis testing. Not to get into details of what hypothesis testing is as this post is about p-value and not hypothesis testing, in a nutshell we start with having a hypothesis and we want to test it if that is true or not. So we take a data from a random variable, get it's mean(if testing the mean) and use the sampling distribution of mean to get a test statistic and compare it with a pre-defined level of significance called alpha to decide if we are going to reject the null hypothesis or not.

Hmm.. Ok. So when we just always reject based on a pre-defined alpha value, then why is p-value used?. Don't we always worry about 0.05 level of significance?.

Good questions. So here is the thing what p-value means and why it is needed. Lets assume our data in question has a normal distribution and the sampling distribution of the mean of data is normally distributed too(which it would be if original data is normally distributed. Stupid me :P).

Example 1: Test Statistic and P-Value



A typical definition of p-value you would find everywhere is as follows:

the p-value is defined as the probability of obtaining a result equal to or "more extreme" than what was actually observed, assuming that the hypothesis under consideration is true

I am a bit obtuse and slow in intelligence. So I have always got more confused with that definition. :-(. If you had to explain p-value to a business guy (something you might encounter quite often) I am sure that definition would just confuse more. :-(

So here is how I understand things.

p-value is nothing but the **smallest** alpha value one needs to have to reject a null hypothesis.

Ok. So what does that mean?. What the heck is this alpha. Oh I hate statistics. Why it has to be so confusing. :-)



Ha ha . Yeah it is confusing. Relax. Alpha is the error or risk you are willing to have if you go with your hypothesis. Its the red shaded region in the above figure. And as we know to go with our hypothesis we reject the null hypothesis which means our p-value should be less than the alpha value.

Arrggh.. My head is exploding now. You just said p-value is the smallest alpha. Then how can p-value be lower than alpha?. That's stupid!

:-). Well both are true. For this first lets talk why p-values are used. Imagine you are a statistician and you are testing a hypothesis. Lets say - " 40% of the times Kaushik takes us out for a lunch when asked".

You want to test if this hypothesis is true or not. Or you want to test if this 40% value is close to the population proportion. Remember your hypothesis is as above. That is the alternate hypothesis. The null hypothesis here is "Kaushik doesn't take us out 40% of the times"

Wait. Why is alternate hypothesis the one you thought of and why it's opposite is null hypothesis. Aren't we interested in finding if our hypothesis is true or not? Then why are we testing it's opposite as null?.

Very Good question. And something which has bothered me quite a lot. So there was this famous philosopher called Carl Popper. He made an argument known as Popperian argument which is that in Science , the philosophy is that you cant prove you are right. You can prove that you are not wrong. For example, Einstein came up with the theory of Relativity. No one at that time could prove that his theory is right. The way to check if he is right is to check if he is wrong anywhere. And if you fail to prove he is wrong that means he is right. So people kept on finding ways to figure out where Theory of Relativity will fail. That was the Null Hypothesis. And they kept on rejecting it as they couldn't find a way where it will fail. So by proving he is not wrong, they proved he is right.

Ooooh. So that is why in hypothesis testing we always say , You reject the null hypothesis or fail to reject the alternate hypothesis instead of "Accepting the alternate hypothesis".

Bingo. Now you are thinking :-).

So coming to our point. Lets say as a Statistician you come to a conclusion that based on the test statistic which would be proportion of times Kaushik took us out for lunch on the past sample data, and you reject the null hypothesis(Kaushik doesn't take us out 40% of the times) at a significance level of 0.05, then it's a good conclusion. But, what if I want to know what is going to be the result at 0.10 or 0.20 significance level. Will your test statistic still fall beyond this new alpha, critical value?. One can't know because all the statistician did used a pre-defined alpha value and took a decision of rejecting or accepting a hypothesis.

That's where p-value concept is used. *p-value is that smallest or minimum alpha value we need to reject a null hypothesis.* And what can be that minimum value?. It is the value at which we have the test statistic. In fact test statistic is nothing but the quantile of the p-value.

So now when you tell someone my p-value of test statistic is x, it immediately tells them that they need at least that alpha or significance level before they would be in a position to reject a null hypothesis. That means if another statistician decides to choose a different alpha value and if it is less than the p-value of the test statistic calculated by the first statistician, then they can't say "*Kaushik takes us out 40% of the times for lunch when asked*" i.e they fail to reject the null hypothesis-Kaushik doesn't take us out 40% of the times for lunch when asked.

Hmm..This makes sense. So then how do you explain p-value to a business person.

Well most people go with the definition I mentioned above. For me that is more confusing as am a bit dumb to understand things. So if I explain, it would be that p-value is the minimum risk they would have in going with their hypothesis or business decision. That means if lets say FCA business head decides that only 1% of the car have a brake problem then the minimum risk or error they can have in their judgement is given by the p-value.

Cool. I think I understand this a bit now. But you need to cut your rant short!! 🙄

Ok. Sorry will try it from next time. So now since you feel you understand a bit about p-value and quantiles, here is a very basic question for this week.

Question:

RELATED KEYWORDS: [IIT-Bombay](#) | [IIT-B](#) | [IIM-A](#) | [Common-Admission-Test](#) | [CAT-Tests](#) | [Cat](#)

3 from IIT-Bombay score a 100 percentile in CAT

Vinamrata Borwankar, TNN | Dec 29, 2014, 01:19AM IST

[f](#) Like [Share](#) 3.1k [t](#) Tweet [G+](#) 2 [in](#) Share 1

MUMBAI: Three boys from the Indian Institute of Technology, Bombay are among the 16 candidates who scored a 100 percentile in the recent Common Admission Test (CAT) 2014, an entrance exam for admission to India's premier Indian Institutes



Does this news sound a bit dubious and hard to believe theoretically?. If yes, then why. If no, then why. Again, rules of the game remain the same. First one to respond would get a coffee from me. :-)

Manish