Hypothesis Testing

Hypothesis testing allows us to use a sample to decide between two statements made about a Population characteristic.

Population Characteristics are things like “ The mean of a population” or “ the proportion of the population who have a particular property”.

These two statements are called the Null Hypothesis and the Alternative Hypothesis.

Definitions

H_0: The Null Hypothesis
This is the hypothesis or claim that is initially assumed to be true.

H_A: The Alternative Hypothesis
This is the hypothesis or claim which we initially assume to be false but which we may decide to accept if there is sufficient evidence.

This procedure is familiar to us already from the legal system: “Innocent until proven guilty”.

The Null Hypothesis (Innocent) is only rejected in favour of the Alternative Hypothesis (Guilty) if there is sufficient evidence of this “beyond reasonable doubt”.

So to summarise H_0 is the status quo and H_A is what we want to prove using the data we have collected.

H_0 and H_A take the following form:

Null Hypothesis
H_0 : Population Characteristic = Hypothesised Value

Alternative Hypothesis ~ three possibilities

Upper tailed test:
H_A : Population Characteristic > Hypothesised Value

Lower tailed test:
H_A : Population Characteristic < Hypothesised Value

Two tailed test:
H_A : Population Characteristic ≠ Hypothesised Value

NOTE: The same Hypothesised value must be used in the Alternative Hypothesis as in the Null Hypothesis

Type I errors: We reject the Null Hypothesis even though the Null Hypothesis is true.

Type II errors: We do not reject the Null Hypothesis when in fact the Null Hypothesis is false and the Alternative is true.

LEGAL ANALOGY:
Null Hypothesis : Innocent
Alternative Hypothesis: Guilty

Type I error: You are convicted and found guilty even though you are innocent.
Type II error: The jury finds you innocent even though you are guilty of the crime.

Before we embark on a Hypothesis Test we must decide what probability of a Type I error we can live with. We cannot completely eliminate these Type I errors, one reason being that by choosing a rejection region to lower the chance of a Type I error we actually increase the chance of a Type II error.

Definitions

- The probability of a Type I error is called the **Level Of Significance** of the Hypothesis Test and is denoted by α -alpha.
- The probability of a Type II error is denoted by β.

We choose α to be small (.01, .05 or .1) but we cannot completely eliminate the probability of a Type I error, as we mentioned already α and β are related

The only way to reduce β without increasing α is to increase the sample size.

Type I errors are generally considered the more serious so in our testing procedure we control the probability of these errors (α) and usually are unaware of the probability of Type II errors (β).

Rejection Regions

One way of performing a Hypothesis test is to compute a rejection region . If we find that our test statistic (which we measure from our sample) is in this region then we reject our NULL Hypothesis.

The computation of the rejection region is mathematical and involves us using statistical tables like the Normal tables.

What is important is the idea that the rejection region is a region far away from our Null hypothesis. And that it is unlikely that we would observe a sample with a value of the test statistic (for example the sample mean or sample proportion) this far away from the Null Hypothesised value if that Null Hypothesis was true.

For example in a class of 100 people.

Suppose our Null Hypothesis is that the average age of the whole class is 31. Suppose we now observe a value of 20 in a sample we have randomly chosen. That's very unlikely to have happened by chance if our Null Hypothesis was true. It's much more likely that our Null Hypothesis is false. So we decide to reject our Null hypothesis in favour of an alternative, which is that the true average age of the whole class is less than 31.

So the procedure involves us deciding on a rule for when we should reject and when we shouldn't.

We could say that in this example with a NULL of 31, if we find a value below 26 then we will not believe our NULL of 31.

But there's still a chance that the true value could be 31 even though we observe a value below 26.

So maybe we want o reduce the chances of that Type I error. To do that we decide on a new rule it says that we reject if we find a value below 23. Now if we find a sample value below 23 (that's 8 years younger than the hypothesised value of 31) we reject. And this time we are more sure of our decision than when we used 26 as our cut off value.

So formally these regions " below 26" and "below 23" are our rejection regions and it is possible to mathematically construct the intervals so that the Type I error they correspond to is equal to some given value (e.g. 0.1, or 0.05 or 0.01).

Section P-Values

Introduction

The conclusion of a Hypothesis test is dependent on the initial choice of α.

- Remember it was this Significance Level which determined the Rejection Region.
- Remember also that α = Probability of a Type I error.

In these tests we decided what value of α "we could live with" before we conducted the test. Aware as you are that decreasing α increases β, you might choose a different value for α than I would.
Depending on these choices of α we may come to different conclusions about whether to reject the Null Hypothesis.

An alternative approach to Hypothesis testing is not to state any α at the outset but to wait until the Test Statistic has been calculated and then decide whether to reject or not by comparison with α. This way all the calculations can be performed independent of any individual choice of α and at the end each person can make a decision based on their choice of α.

How would we conduct such a test?
Instead of computing a Rejection Region we calculate a P-Value.

IMPORTANT: Small P-values are equivalent to values of the test statistic far away from the NULL. Don't worry about their computation.

Definition

The P-Value is the smallest level of significance at which H_0 can be rejected.

VERY IMPORTANT: How to use P-values

Once the P-value has been determined, the conclusion at any particular level of α results from comparing the P-value to α:

1. P-value $\leq \alpha$ => Reject H_0 at level α

2. P-value $> \alpha$ => Do Not Reject H_0 at level α

These two statements are exactly equivalent to saying

1. Test statistic lies in rejection region, and is far from Null Hypothesised value.
2. Test statistic lies outside rejection region and is close to Null Hypothesised value.

Easy Examples

1. For which of the given P-values would the Null Hypothesis be rejected when performing a level .05 test?

a. .001 b. .021 c. .078 d. .047 e. .148