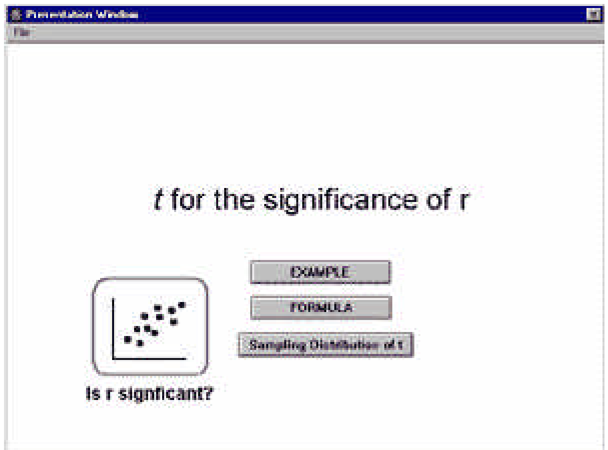
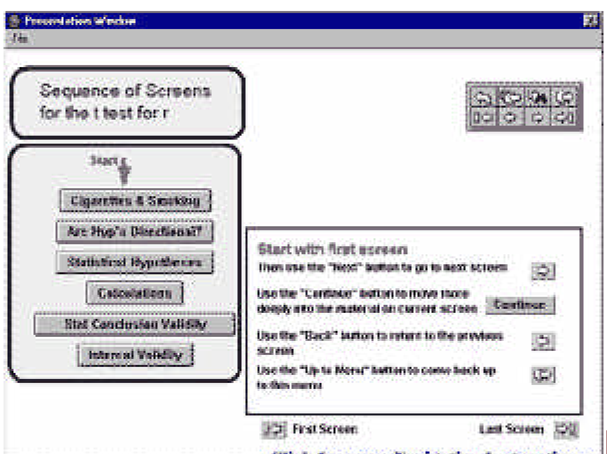
t-test for r

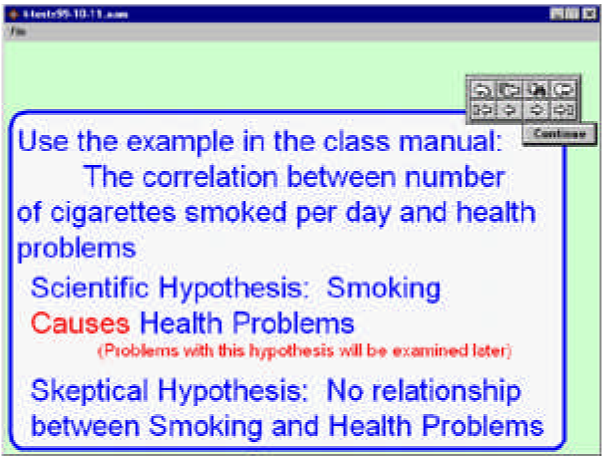


This is the text of the in-class lecture which accompanied the Authorware visual graphics on this topic. You may print this text out and use it as a textbook. Or you may read it online. In either case it is coordinated with the online Authorware graphics.

PRINT: You may print this web page on your local printer if you wish. Then you can read the hard copy of the lecture text as you look at the Authorware graphics.



Example



[Go to Top]

As you can tell from this series of lectures, there are many different kinds of t-tests. The one for testing the significance of the correlation coefficient, is called the t for r.

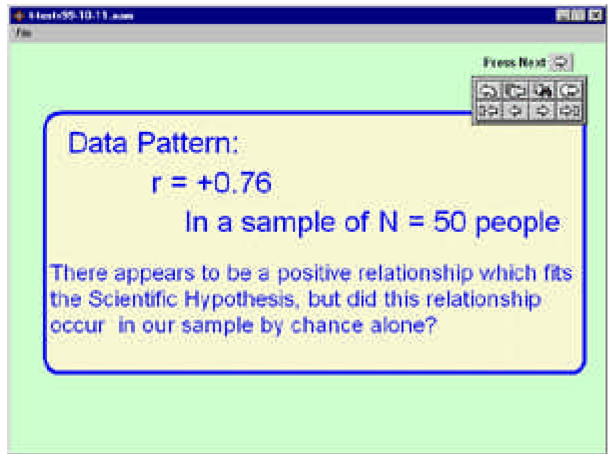
The example we are going to work with involves cigarettes and smoking. This particular t-test, tests the significance of little r, or the correlation coefficient. We have already studied the correlation coefficient so you should be familiar with it in terms of having used it and calculated it a couple of times. It is a descriptive measure of the degree of relationship between two variables and also indicates the direction of the relationship. A positive value of "little r" indicates a direct relationships. A negative r indicates an inverse relationship.

The correlation coefficient ranges from minus one through plus one (-1 to +1). A value of zero occurs when thre is no relationship between the variables and minus one (-1) or plus one (+1) indicates perfect relationships. When doing research in the behavioral sciences, the values reach a value of either plus or minus one.

Let's use the example in your class notes outline manual in the t for b section. A research team hypothesizes that cigarette smoking causes health problems. The team does a correlational study on a group of 50 volunteers. The volunteers fill out a form on their smoking habits early in life and their health problems later in life.

The skeptic attitude is that there is no relationship between the smoking variable and the health variable.

The Data Pattern and the PCH of Chance



Suppose that the data yield a positive correlation coefficient of 0.76. This describes a direct (positive) relationship between the two variables. As amount of smoking goes up,the number of health problems goes up.

The skeptic must admit this is consistent with the scientific hypothesis. Therefore a skeptical personwill begin to create plausible hypotheses which compete with the scientific hypothesis to explain the data pattern (r = +0.76). The PCH we will concentrate on here is the PCH of Chance. But there are many others, because correlational studies alone are very poor ways to attempt to support a causal hypothesis (e.g., smoking causes health problems).

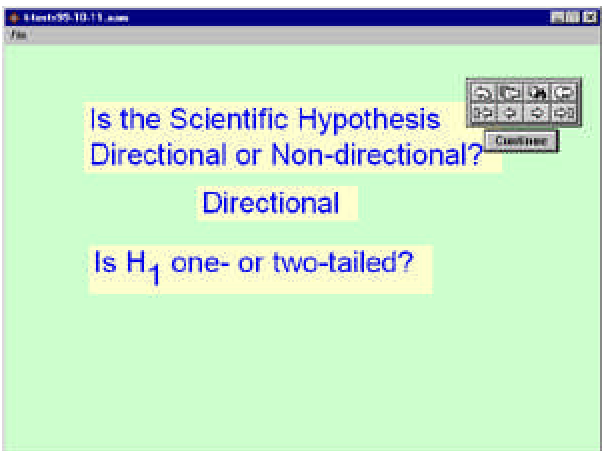
The first step in evaluating the PCH of chance is to determine the relationship between the two variables of interest. The research has done that. They have found a positive correlation between smoking and health problems. In reply, the skeptic will say "*Well how do you know that you didn't just find this relationship by chance? It's just sampling error in your statistic. There appears to be a relationship in your data, but in fact it only occurred by chance alone*."

Now we will review how to evaluate whether or not the value of the correlation coefficient that they found is likely to be due to chance or not.

REVIEW: The scientific hypothesis is that smoking causes health problems. The skeptic says that there's no relationship between smoking and health problems. Note that there are real problems with causal inferences from a correlational study but that will be discussed later. One cannot scientifically take a correlational study and conclude causation between the two variables.

After data collection, you determine the correlation coefficient, and you find a positive correlation of .767 (r = .767) between amount of smoking and later health problems in a sample of 50 people. There appears to be a positive relationship which fits the scientific hypothesis. However, the first question one should determine is - did this relationship occur by chance alone in this sample?

Go to Top



Scientific Hypothesis

The next step is to determine if the scientific hypothesis is directional or not. Try to answer for yourself whether you think its directional or not and *why*.

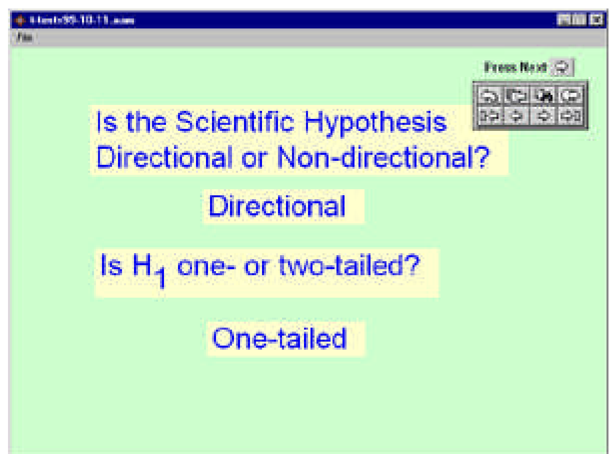
The scientific hypothesis presented here is directional because the researcher is looking for a positive relationship between cigarette smoking and health problems. If smoking causes cancer it follows that as smoking increases then health problems will increase.

Non-directional Hypothesis

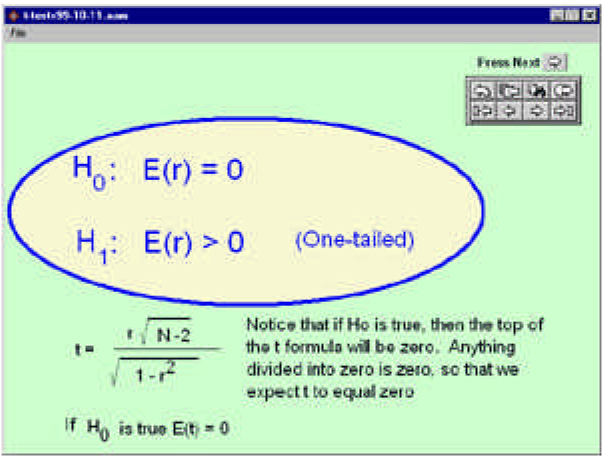
In contrast, a non-directional hypothesis would be - *There is some sort of relationship between cigarettes and smoking but we don't know whether it'll be positive or negative relationship. Perhaps people who smoke will have less health problems, perhaps they will have more.*

The smoking example doesn't fit very well with a non-directional hypothesis. Maybe if you were you were doing a preliminary study which evaluated an herbal tea that claimed to be good for health, you might wonder whether this particular herbal tea is good for health or not. Perhaps you know that some herbal remedies have been shown to actually be bad for people and had to be withdrawn from the market. Others are probably good for people and useful. In early research, a researcher would be more likely to hypothesize that there will be a relationship between health and ingestion of a substance but doesn't know if it will be a positive one or a negative one.

Directions and Tails



In this example examining cigarette smoking and health problems, we are expecting a directional relationship -- more smoking more health problems. This conclusion leads us to a the nest question -- Is it a one- or two-tailed test? You may remember from the t-test homework that directional hypotheses go with one-tailed tests.



Statistical Hypotheses

The null hypothesis in the statistical model corresponds to the PCH of Chance. The skeptic thinks that any relationship in this sample of 50 volunteers is just due to chance alone and thus is expecting there to be no relationship between the two variables. In other words, the skeptic expects that smoking and health problems are unrelated, and therefore the correlation coefficient will be equal to zero.

The null hypothesis for the t for r will be expressed as the expected value of r is zero [E(r) = 0].

The scientist on the other hand is expecting a correlation coefficient above zero, a positive correlation between the amount of smoking and health problems, our two variables. So the alternative hypothesis, H1, is that the expected value of r is greater than zero.

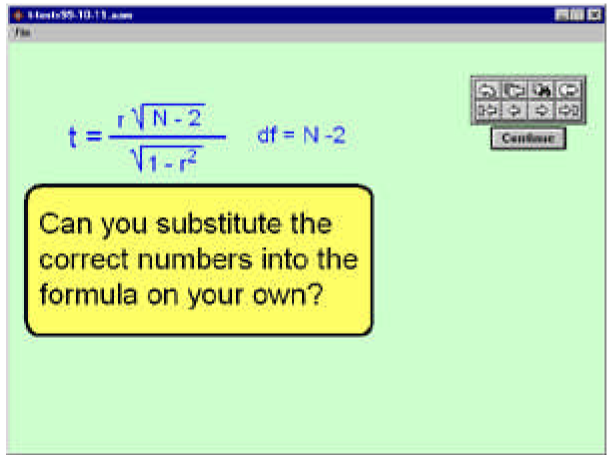
Expected Value of t

Notice the t formula shown at the bottom of the graph. Do not write it down yet, we will do that in the next graphic. Notice that if you substitute r = 0 into the t formula (as H0 expects) then the value of t will be zero. As usual, the null hypothesis is going to expect the value of t to be zero.

Notice also that if H1 is correct the value of t should be greater than zero.

Go to Top

Substitute Into the Formula



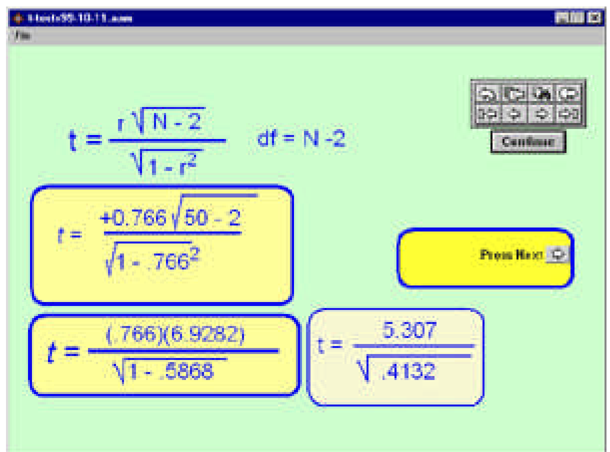
Go To Top

Here is the t for r formula. it is r times the square root of n minus 2, divided by the square root of one minus r squared. It is fairly simple t formula, assuming that you have the value r of course.

N, of course, is the number of research participants. The degrees of freedom equal N minus 2.

What I'd like you to do now is to substitute into the formula. Remember r = .766 and the number of subjects was 50.

Calculations



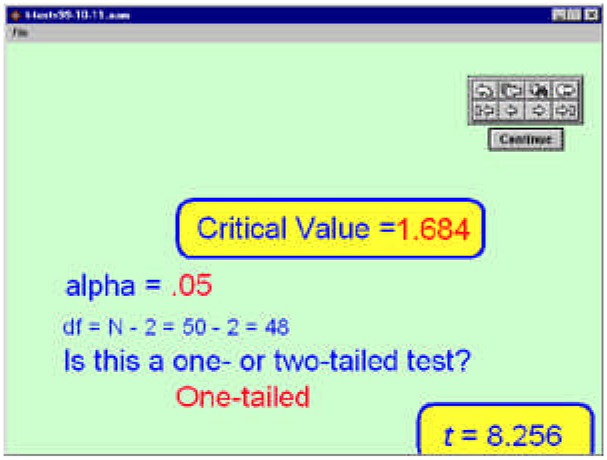
Here is my substitution which you can check against yours. r is equal to .766. Next you take .766 times the square root of 50 minus 2, all that divided by the square root of one minus .766 squared.

Jot down whatever arithematic steps that you find useful.

The t came out to be 8.256. This is our calculated t; we had to calculate it using a formula to get it.

Degrees of freedom are N minus 2. N = 50 so 50 - 2 = 48.

Go to Top



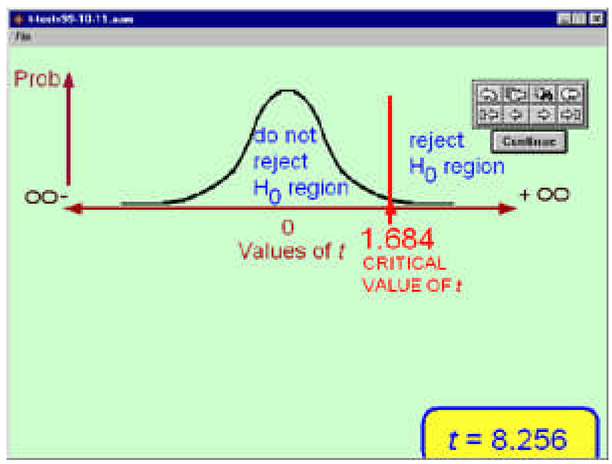
Critical Value

Let's assume that we use alpha equal to .05. Now we know the alpha and the degrees of freedom.

Next we have to decide if this test is one-tailed or two-tailed. We've already discussed the issue and decided it is one-tailed.

So using the t tables, the critical value is 1.684. That assumes you look up a one-tailed test, with alpha set to .05, and 48 degrees of freedom.

The Sampling Distribution of t



As we've discussed before the sampling distribution of of t is slightly different from by very similar to the normal distribution. H0 expects t to be zero.

I find it worthwhile for students to draw out the t probability distribution, which is symmetrical around zero, the value that H0 predicts.

[Go to Top]

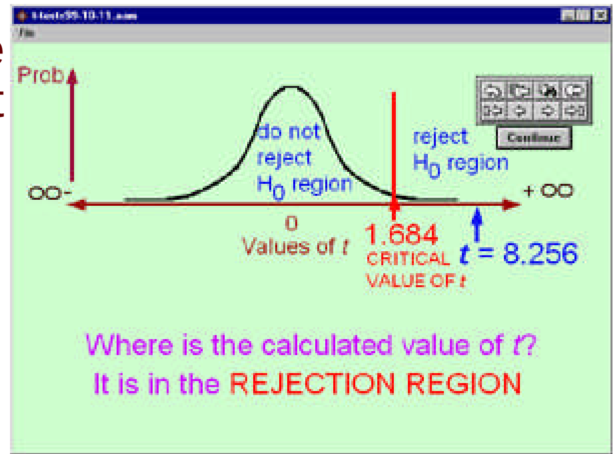
Hypothesis testing Logic

Let's review the hypothesis testing logic. H0 is claiming that this t statistic should come out to be zero or very near zero. You wouldn't want to hold H0 to an exact prediction, but HO expects t to be somewhere in the neighborhood of zero. H0 is part of the general prediction that there is no relationship between smoking and health problems so you're only getting a chance value of r. A chance value of r shouldn't be very far from zero, therefore, the t value ought not to be very far from zero.

Our calculated value was close to 8. Eight seems a long way from zero, but the question remains, is it a long way or not? How do I evaluate whether or not an actual calculated value of my test statistic is a long way from zero? In general, the answer comes from evaluating whether 8 is very close to zero where its probability is extremely high, or does 8 lie in one of the tails some distance from zero where it probability is very low? If 8 has a very low probability of occurring given that H0 i true and t should be zero, then we'll reject H0.

We chose alpha equal to .05, so that means the probability is .05 or 1 in 20, that b chance alone we would get a t value of 1.6 or larger.

Compare the Critical t to the Calculated t



The calculated value of t is farther from zero than is the critical value of t. So the calculated t falles in the rejection region. If H0 is true and if I get out in this rejection region, what I'm saying is that this occurrence is a very a improbable event, a very improbable value of calculated t. That assumes that H0 IS correct. Instead of believing that H0 is correct I will reject it.

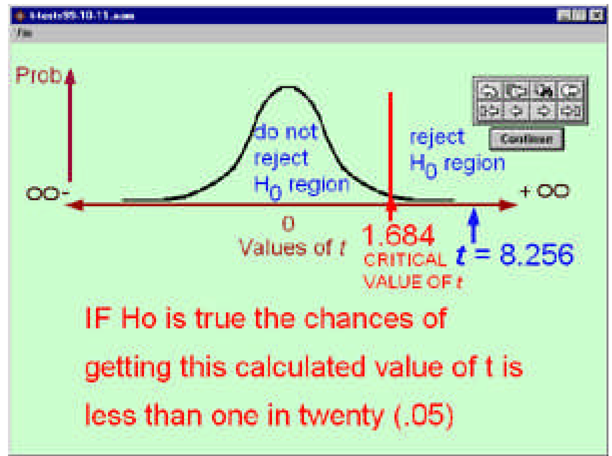
Go to Top

Significance

So we put the calculated value on the number line and we see it falls in the rejection region. It is farther away from zero than the critical value, and so it is in the rejection region. The consequences of its placement here is that we reject H0 . The mechanics are really simple.

When we reject H0 we say that the result was "signicant." Signicance is another way of saying that we are able to rject H0.

Alpha



I'm going to repeat some things but I want you to keep thinking in terms of probability. If H0 is true, the value of t *ought* to be zero or near zero, and it *ought* to be in the region to the left of the red line where most of the probability lies.

The chances of getting a value to the right of the red line just by chance alone has low probability. The probability of getting in the rejection region by chance alone is the area under the curve out there beyond the red line. The probabilities out there in the tail are very tiny probabilities. So if H0 is true, the probability of calculated t falling in the rejection region is less than 1 in 20, or .05.

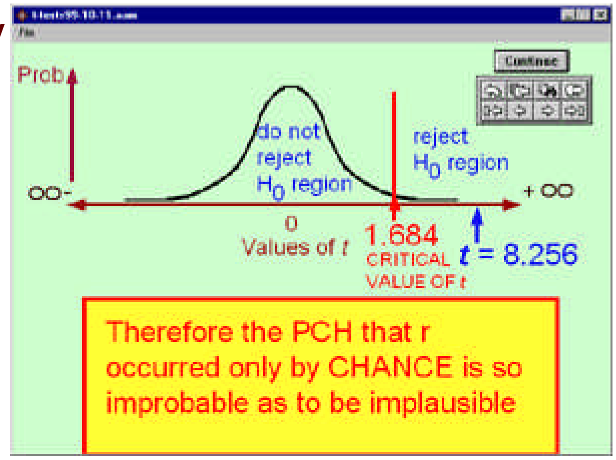
So I'll say it is so improbable I would get a t value so large if H0 is true, I'm going to reject H0.

Of course, it is remotely possible that by chance alone you could flip a coin one hundred times and get all heads. Anything could happen by chance alone. But what i'm saying is that the probability I'm wrong when I reject H0 is very small, .05 or 5%.

Alpha is thejust the probability that I am wrong when I reject H0.

Go to Top

Statistical Conclusion Validity



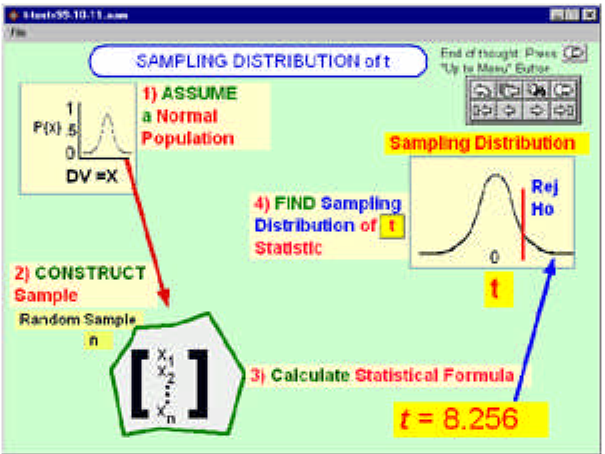
If H0 (which is based on chance) is highly improbable, then the PCH that r occurred only by chance alone is so improbable that it is implausible.

So rejecting H0 in the realm of statistics has implications for the PCH of Chance in the realm of science.

To review, we've rejected both statistical null hypothesis, H0, and with it the scientific skeptic's criticism that this r occurred by chance alone. Chance is no longer a plausible competing hypothesis. It becomes an implausible competing hypothesis.

The validity with which we can argue against chance as a PCH is what we mean by statistical conclusion validity.

Go to Top



Summary of the 4-step Process

This is a summary of the whole process from the normal population from which we sample our data to rejecting H0 using the sampling distribution of t.

In steps 1 and 2, we assume that we have a sample that comes from some kind of normal population. The t-test always assumes that the original population was normal. Our research project amounts to taking a sample of data from 50 people.

In step 3 we calculate a statisical formula on the data. In this case we calculate the t for r.

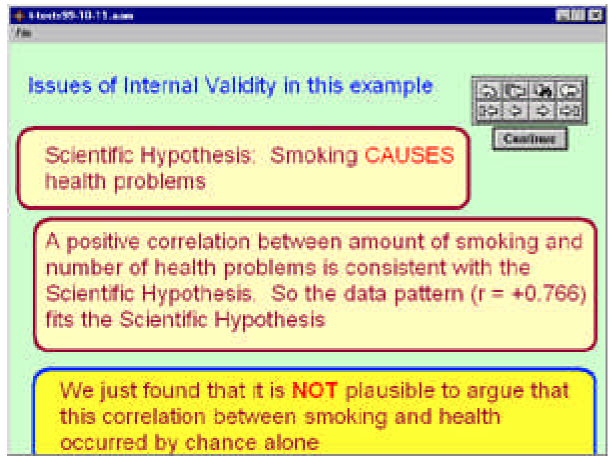
Through a lot of complex math (which we can't do in this class) we know that the sampling distribution of t for r is a t distribution. Moreover, if H0 is true, the t distribution ought to be centered around zero.

In step 4, we set up a rejection region based on the sampling distribution of t. In our example data we found that our calculated t falls beyond the critical value out in the rejection region.

That's the summary of the whole process - you have original population, the sampl data, a calculated statistic, and finally the sampling distribution which is the basis of our probability argument.

Go to Top

Limitations of Statistical Conclusion Validity



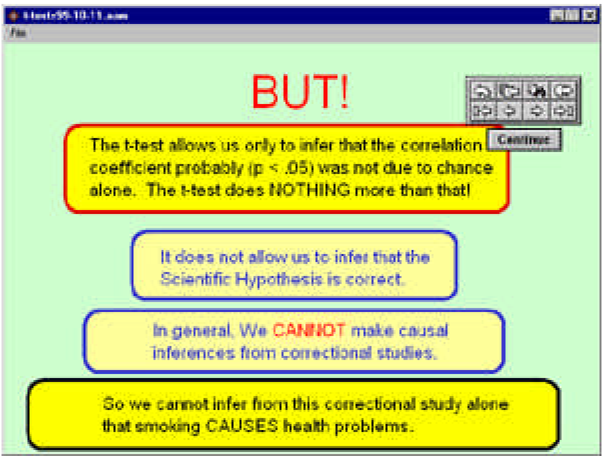
Internal validity is a topic usually discussed in research methods, but when discussing correlation I think it is important to be clear about what we can and cannot say based on our statistical findings.

What we've accomplished so far is to discredit the PCH of chance. It is highly unlikely that this correlation coefficient of .766 occurred by chance alone. This leads one to conclude that something is going on between cigarettes health problems.

From the point of veiw of the scientific hypothesis, what is going on is that smoking CAUSES health problems. A positive correlation between amount of smoking and the number of health problems is certainly *consistent* with the scientific hypothesis that smoking causes health problems. Based on this hypothesis one would expect that as the amount of smoking increases the amount of health problems will also increase. Moreover, we determined that it is not plausible to argue that this correlation occurred by chance alone. So we have argu against the PCH of chance.

The important point here is THAT IS ALL these statistical procedures do. They jus do that one thing. They just rule out the PCH of chance.

After Chance is eliminated as a PCH then the real interesting scientific arguments and discussions begin. Let's take a quick look at some of the major issues involve in these arguments.



Significance and Proof

The t-test allowed us to discredit the PCH chance, but nothing more. It does not allo us to infer the scientific hypothesis is cor Obtaining a significant result doesn't prov the scientific hypothesis. The most glarin case of this is when people attempt to conclude that a causal hypothesis is true based on a significant correlation.

Both consumers and producers of media make this mistake frequently when discussing research, so be alert for this mistake when you hear about a study on radio or television.

When a news story or adverisment reports that there are significantly less sympt after one treatment than another, it only means that the difference between the treatments is unlikely to be due to chance alone.

Go to Top

Correlation and Causation

If you recall the Correlation lecture, we argued that a correlation alone is not evidence of causation. Just because cigarettes and health problems are correlated does not BY ITSELF mean that cigarettes cause health problems. This is an important point. Let's go over it again.

We cannot make causal inferences from the correlational studies. Just because a variable is correlated with another variable does NOT mean that it causes the second variable. For example, there is a positive correlation between the number o ice cream cones purchased in New York City, and the number of deaths in Bombay India. However, that does not mean that if you buy an ice cream cone in New York City you're killing somebody in Bombay. There is not any causal relationship between your purchase and death's of other people in Bombay.

In fact, fluctuations in both ice cream sales and mortality rates are caused by a third variable: Global weather patterns. It turns out that when it is hot in New Yor it tends to be hot in Bombay. Extreme heat stresses people and so people who are very ill or weak are more likely to die when it is very hot. And when it is very hot, people buy more ice cream cones.

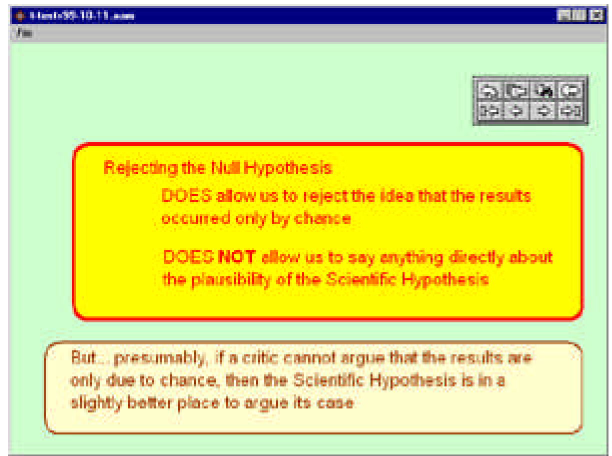
The third variable, weather, causes changes in both mortality rates and ice cream sales. This third variable accounts for the small but statistically significant correlation between ice cream sales in New York City and deaths in Bombay, India

For many reasons such as third variable confoundings (see Correlation lecture), it i hazardous, if not impossible, to infer causality from correlational studies.

If there is causation between two variables one would expect to find a correlation between them. But the converse is not true. A significant correlation between variable "X" and variable "Y" does not necessarily imply that X causes Y.

In other words, correlation is consistent with causation even while it does not establish causation.

Go to Top



Causation: Smoking and Health

So we cannot infer from the correlational study alone that smoking causes health problems. That causation cannot be inferred from correlation alone was the source of a lot of difficulty for the scientific medical community investigating the relationship between smoking and health problems in the 1950's and early 60's.

For reasons that we will go into in more depth in Research Methods, causation is more appropriately inferred from a well controlled experimental study than from a correlational study. Since, for ethical reasons, researchers can't run an experimental study with human subjects (such as assigning somebody when they're 7 years old to be either in the smoking group or non-smoking group) they were never able to run an experimental study with humans. All researchers were able to do were correlational studies. There was great controversy about the causal link between smoking and health.

Researchers eventually did experimental studies with mice. They found a very strong causal relationship between smoking and cancer; then they generalized their results to humans.

The point is that REJECTING H0 DOES ALLOW us to reject the idea that the results have occurred by chance alone.

REJECTING H0 DOES NOT allow us to say anything directly about the plausibility of the scientific hypothesis.

NOTE: The scientific community, based on numerous sophisticated studies, has strongly concluded that smoking cigarettes do in fact cause health problems. The point we are making here is that on a single correlational study alone, that conclusion would be unwarranted.

An Absurd Study

Suppose someone made the scientific hypothesis that buying ice cream in New York City kills people in Bombay and then found a significant correlation between ice cream sales in New York and deaths in Bombay. Suppose this person calculated the t for r and it was significant. Suppose the the researcher rejected H0 and came to the statistical conclusion that this result did not occur by chance alone. So far s good.

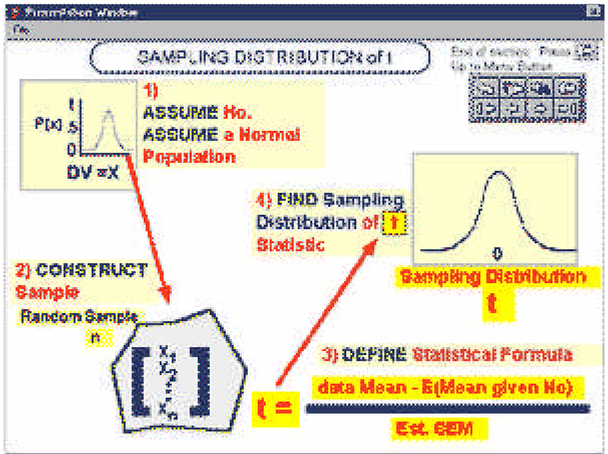
But if the scientisti made the conlcusion that therefor buying an ice cream cone i New York City causes a death in Bombay, s/he would have gone beyond the logica limitations of the study.

Rejecting H0 only allows you to say this obtained correlation isn't due to chance. That's all.

I use this absurd example to make the point that whereas this hypothesis testing logic allows you to reject chance, it does not in any way confirm the scientific hypothesis. A researcher would have to do that on a basis of other kinds of scientific argument.

Go to Top

Go to Top



SAMPLING DISTRIBUTION OF t

Review: As usual, the sampling distribution of the test statistic (t in this case) is derived from the original population (normal in this case), a sample of size generating a certain number of degrees of freedom (N = 2 in this case), and a formula calculated on the sample data (t for r, in this case).

We use the sampling distribution of t to find our rejection regions and to make our decision about Ho.

[Go to Top]