

22S:105 Statistical Methods and Computing

Introduction to Hypothesis Testing

Lecture 14
Mar. 9, 2009

Kate Cowles
374 SH, 335-0727
kcowles@stat.uiowa.edu

Example:

I claim that my husband's resting pulse rate is 45 beats per minute. This is very low and would be typical of either a highly trained athlete or a sick individual.

To test my claim, you wish to measure his resting heart rate on 5 different occasions.

Here, the "population" of interest is all possible measurements of my husband's resting pulse rate. My claim may be interpreted as saying that the mean μ of this "population" of values is 45 beats per minute.

Introduction to Hypothesis Testing

Recall that statistical inference is using data contained in a sample to draw conclusions or make decisions about the entire population from which the sample is taken.

Two main goals of statistical inference

- estimation of unknown population parameters
- testing specific hypotheses about unknown population parameters

The purpose of hypothesis testing is to "assess the evidence provided by data about some claim concerning a population."*

* Moore, D.S. *The Basic Practice of Statistics*

Suppose the measurements you get are:

42 52 43 48 47

The sample mean $\bar{x} = 46.4$. Does this provide evidence against my claim?

We will consider this question by asking what would happen if my claim were true and we repeated the sample of 5 measurements many times.

Suppose first that we knew that the standard deviation of measurements of my husband's resting heart rate was $\sigma = 4$ beats per minute.

- If the claim that $\mu = 45$ is true, the sampling distribution of \bar{x} from 5 measurements is normal with mean $\mu = 45$ and standard deviation

$$\frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{5}} = 1.79$$

- We can judge whether any observed \bar{x} is surprising by finding it on this distribution.

The *alternative hypothesis* is the claim *for* which we are trying to find evidence.

- symbolized H_a

In the example about my husband's heart rate, your alternative hypothesis probably was

$$H_a : \mu > 45$$

The *p-value* of the test is the probability, computed assuming that H_0 is true, that the observed outcome would take a value as extreme as or more extreme than, what we actually observed.

- Small p-values are evidence against the null hypothesis.

Terminology of hypothesis tests

The *null hypothesis* is the statement being tested.

- The test is intended to assess the strength of evidence *against* the null hypothesis.
- Usually is a statement of “no effect,” “no difference,” “nothing going on.”
- The null hypothesis is commonly symbolized as H_0 .
- H_0 is a statement about an unknown population parameter.
- Example:

$$H_0 : \mu = 45$$

The result of a hypothesis test is a decision. The possible outcomes are called

- Rejecting the null hypothesis
- Not rejecting the null hypothesis

Before we carry out the test, we must decide how strong we will require the evidence to be in order for us to reject H_0 . We specify this in terms of a *significance level*.

- The significance level is how small we will require the p-value to be in order to reject H_0 .
- symbol is α
- conventional choices are $\alpha = .05$ and $\alpha = .01$

Example: my husband's resting heart rate

We will choose $\alpha = .05$ as the significance level at which to carry out the test.

To find the p-value of our results, we will standardize \bar{x} so we can use the normal table.

- Remember: the p-value is computed assuming H_0 is true, so the value of μ to use is the value stated in H_0 .

$$\begin{aligned} z &= \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \\ &= \frac{46.4 - 45}{1.79} \\ &= 0.78 \end{aligned}$$

One-sided and two-sided tests of hypotheses

The hypothesis test we just conducted was *one-sided* test. We were interested only in showing that the value of the unknown parameter differed from that given in H_0 in one direction.

$$\begin{aligned} H_0 &: \mu = 45 \\ H_a &: \mu > 45 \end{aligned}$$

We might also have stated the hypotheses this way:

$$\begin{aligned} H_0 &: \mu \leq 45 \\ H_a &: \mu > 45 \end{aligned}$$

According to Table A, the probability of a value this large or larger is 0.218. We would say that for this test result

$$p = 0.218$$

Since this is *larger* than $\alpha = .05$, we cannot reject the null hypothesis. That is, we have decided that the evidence was not sufficient to reject my claim!

In specifying null and alternative hypotheses:

- There must be no overlap in the range of values included in the two hypotheses.
- All possible values of the unknown population parameter must be covered in one or the other of the two hypotheses.

Two-sided hypothesis tests

Example: We wish to compare fasting serum cholesterol levels in persons over 21 living in a group of islands in the South Pacific with typical levels found in the U.S.

We know that levels in adults over 21 in the US are approximately normally distributed with

- mean 190 mg/dl
- standard deviation 40 mg/dl.

We have no idea what the relative levels of serum cholesterol are on the islands as compared with the U.S.

The hypotheses for our *two-sided* test are:

$$\begin{aligned} H_0 : \mu &= 190 \\ H_a : \mu &\neq 190 \end{aligned}$$

Before we look at our data, we will decide on the *significance level* α for our test. Let us choose $\alpha = .05$.

We then perform blood tests on 100 adults from the islands and find that the sample mean level $\bar{x} = 181.5$ mg/dl.

To carry out our hypothesis test, we note that, if H_0 is true, the sampling distribution of \bar{x} is normal with

$$\begin{aligned} \mu &= 190 \\ \sigma_{\bar{x}} &= \frac{40}{\sqrt{100}} = 4 \end{aligned}$$

We will assume that the levels on the islands are normally distributed with

- unknown mean μ
- known standard deviation 40 mg/dl

We will standardize the value of \bar{x} that we observed to find out how likely we would have been to get a value as extreme as what we got, or more extreme, if H_0 were true.

$$\begin{aligned} z &= \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \\ &= \frac{181.5 - 190}{4} \\ &= -2.125 \end{aligned}$$

We must find out what area under the standard normal curve lies

- to the left of -2.125
- *and* to the right of 2.125

The answer is $.017 + .017 = .034$.

This is the p - *value* for the test. Since $p < .05$ we reject the null hypothesis and conclude that serum cholesterol levels are different among adult residents of the Pacific Islands than among adults in the U.S.

$$= \frac{181.5 - 190}{3.81}$$

$$= -2.231$$

One sample t-tests

If we don't know the population standard deviation, then we

- estimate it with the sample standard deviation s
- compute a t statistic rather than a z statistic
- compare to a t distribution with the appropriate degrees of freedom

Example: If we do *not* assume that we know σ for serum cholesterol levels among residents of the Pacific Islands.

From the sample of 100 adults, we compute

$$s = 38.1 \text{ mg/dl}$$

We then compute

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

We try to use Table C to find the area to the left of -2.231 and to the right of 2.231 under a t curve with 99 degrees of freedom.

The closest we can come is that under a t curve with 100 degrees of freedom, the area in one tail would be between .01 and .02.

Thus we conclude that the p-value is somewhere between .02 and .04.

SAS can do a much better job for us! It would provide a p-value of .0279.

Thus, if we had chosen $\alpha = .05$, we would reject the null hypothesis.

Types of error in hypothesis testing

$$H_0 : \mu = \mu_0$$

$$H_a : \mu \neq \mu_0$$

True state of the world

	H_0 is false	H_0 is true
Reject H_0	Correct!	Type I error
Do not reject H_0	Type II error	Correct!

$$\alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$$

$$\beta = P(\text{fail to reject } H_0 \mid H_0 \text{ is false})$$

or, put another way

α = probability of making Type I error

β = probability of making Type II error

power $(1 - \beta)$ = probability of correctly rejecting H_0 when it is false; depends on our definition of H_a

Return to the example of my husband's resting heart rate.

- What value of \bar{x} would have been required in order to reject

$$H_0 : \mu = 45$$

in favor of

$$H_a : \mu > 45$$

if $\alpha = .05$?

For a standard normal, $z = 1.645$ cuts off the upper .05 area.

The corresponding value for the sampling distribution of \bar{x} if H_0 is true is

$$\bar{x} = \mu + z\sigma$$

$$= 45 + 1.645(1.79)$$

$$= 47.9$$