Hypothesis Tests for a Population Mean

Shengping Yang PhD

I am sending you an excel file with results from my blood pressure study. Do these data fit a normal distribution? How should the data be analyzed? Do patients who take 50 mg/d thiazide diuretics have systolic blood pressure lower than 160 mmHg?

These are typical questions in statistical analysis. First, let us see what is called a normal distribution.

1. The normal distribution

Normal distributions are continuous probability distributions that are bell shaped and symmetric, with probability density function $f(x) = \dfrac{1}{(\sigma\sqrt{2\pi})} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, where the two parameters μ and σ are the mean and standard deviation, respectively.

Figure 1. Normal distributions
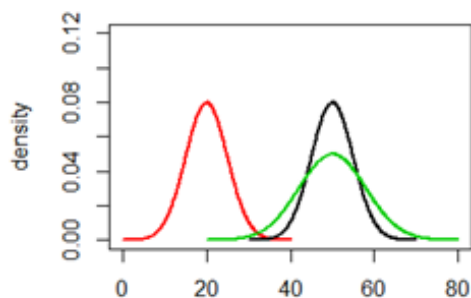
Normal distributions are very important in making statistical inferences because they provide a reasonable approximation to the distribution of many different variables. There are many different normal distributions that are distinguished by their mean and standard deviation. The *mean* of a normal distribution describes where the distribution is centered, and the
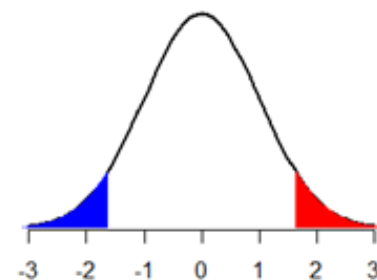
standard deviation describes how much the distribution spreads out around the center. Figure 1 illustrates how mean and standard deviation of a normal distribution determine the normal curve. For example, the normal curves in black and red have the same standard deviation but different means, thus the spreads of the two curves are the same, but the centers of the distributions are different. On the other hand, the black and green curves have the same mean, but different standard deviations.

2. The standard normal distribution

Normal distribution with $\mu=0$ and $\sigma=1$ is called the standard normal distribution; the letter z is widely used to represent a variable whose distribution is standard normal. The standard normal distribution is important because we can always translate our problem of finding a probability based on some other normal distribution into an "equivalent" problem that involves finding an area under the standard normal curves.

Figure 2. Standard normal

Converting a normal distribution with mean μ and standard deviation σ to a standard normal can be

Corresponding author: Shengping Yang
Contact Information: shengping.yang@ttuhsc.edu
DOI: 10.12746/swrccc2014.0205.064

done by using $z=(x-\mu)/\sigma$. The standard normal curve is useful in characterizing extreme values, e.g., the largest 5%, the smallest 5% and the most extreme 10% (include both the largest and smallest 5% because the standard normal distribution is symmetric). As we can see from Figure 2, the z curve area to the left of -1.645 (shaded in blue) is equal to 0.05, i.e. $P(z<-1.645)\approx0.05$. In other words, in a long sequence of observations from a standard normal distribution, approximately 5% of the observed z values will be less than -1.645. Similarly, approximately 5% of the observed z values will be greater than 1.645. As a result, the most extreme 10% of the z values are those either less than -1.645, or greater than 1.645.

3. The null and alternative hypotheses

Built upon what we have described above, a test of hypotheses can be performed to decide between two competing claims about a population characteristic using data collected from such a population. The basic idea of hypothesis testing is that we start with proposing a null hypothesis (H_0), which is a claim about a population characteristic that is initially assumed to be true. The alternative hypothesis (H_a) is the competing claim. The hypothesis H_0 will be rejected only if the sample evidence strongly suggests that H_0 is false. In general, the null hypothesis will have the form of

H_0: population characteristic = hypothesized value, where the hypothesized value is a specified number relevant to a study.

The alternative hypothesis could have one of following three forms depending on the objectives of a study.

H_a: population characteristic < hypothesized value
or
H_a: population characteristic > hypothesized value
or
H_a: population characteristic ≠ hypothesized value

In the blood pressure study, the corresponding null and alternative hypotheses will be:

H_a: μ=160
H_a: μ<160   (μ is the population mean)

4. Type I and type II errors

After the hypotheses have been formulated, a test procedure will need to be used to determine whether H0 should be rejected. Recall that a hypothesis testing is a method that uses sample data to decide between two competing claims about a population characteristic. Therefore, unless such a decision is made based on the entire population, the risk of error is inevitable. In fact, there are two types of errors that can occur when making a decision in a hypothesis testing.

Type I error (α) – the error of rejecting H_0 when H_0 is true

Type II error (β) – the error of failing to reject H_0 when H_0 is false

The natural question here is why not keep both α and β as small as possible, i.e., equal to 0? The answer is that when we try to use sample data (incomplete information) to make an inference about a population, this is the price we have to pay. More specifically, to achieve a small type I error, the test procedure will require very strong evidence against H0, thus null hypothesis is unlikely to be rejected - the consequence of which is an increased type II error. Therefore, the best approach is to achieve a compromise between a small type I error and a small type II error, and the rule of thumb is to use a procedure with the maximum acceptable type I error based on the assessment of the consequences of types I and II errors. In fact, a type I error of 0.05 and 0.01 are commonly used in practical problems.

In the blood pressure study, we can pre-specify the type I error as 0.05.

5. Hypothesis tests for a population mean

Depending on the distribution of the population, the sample size, as well as the objectives of a study, test statistics used in a hypothesis testing can be different.

In the blood pressure study, the objective is to test whether the true average blood pressure for

those patients who take 50 mg/d thiazide diuretics is lower than 160 mmHg.

Now assuming either sample size is large or the distribution of systolic blood pressure for those patients is approximately normal, there will be two scenarios:

(A) The standard deviation of the population σ is known, we can use $z = \dfrac{\bar{x} - \mu}{(\sigma/\sqrt{n})}$ as the test statistic (note that if systolic blood pressure follows a normal distribution, then the sample mean blood pressure also follows a normal distribution, or if sample size is large, then follows a normal distribution by Central Limit Theorem).

(B) Since it is very rare that σ is known, we can use $t = \dfrac{\bar{x} - \mu}{(s/\sqrt{n})}$ (with appropriate degree of freedom) as the test statistic. Note that s can be estimated from the sample.

Since this is a lower-tailed test, the *p* value (assuming that the null hypothesis is true, the probability of obtaining a test statistic at least as extreme as the one that was actually observed) is the area corresponding to the left of the computed *z/t* value. If the *p* value is less than the pre-specified type I error, we reject H_0 at the 0.05 level of significance and conclude that there is sufficient evidence that the systolic blood pressure of patients who take 50 mg/d thiazide diuretics is lower than 160 mmHg.

References

1. Mendenhall W, Beaver RJ, and Beaver BM. Introduction to Probability and Statistics. Pacific Grove: Brooks/Cole Publishing. 2003. 205-409. Print.
2. Peck R, Olsen C, and Devore J. Introduction to Statistics and Data Analysis. Pacific Grove: Brooks/Cole Publishing. 2001. 346-515. Print.