

Lsn37

Clark

Admin

Recall, if n is large enough we can rely on the CLT to give the distribution of

$$\frac{\mu - \bar{X}}{\frac{\sigma}{\sqrt{n}}}$$

However, note that we often find ourselves in a situation where perhaps n isn't really that big, so perhaps we cannot say that $S \rightarrow^p \sigma$. In that case, though, all is not lost, if we make the additional assumption that $Y_1, \dots, Y_n \sim N(\mu, \sigma^2)$, then we can use

as our test statistic. In that case, our rejection region is for extreme values of $T = t$.

For instance, a chemical process has produced, on average, 800 tons of chemical per day. The daily yields for the past week are 785, 805, 790, 793, and 802. If we are interested in answering the question as to whether the data indicate that the average yield is less than 800 tons, what would our null and alternative hypothesis be?

What is our test statistic (with distribution under H_0)?

Assuming we are only willing to risk $\alpha = .05$, our rejection region can be found from

```
qt(.05,4)
```

```
## [1] -2.131847
```

Graphically this is:

Our realization of our test statistic is:

Our conclusion is:

What about if we assume the CLT holds? How would the problem above change?

We can do something similar if we want to compare two population means. Now our test statistic is:

Again we reject for really high or really low values of t .

Two methods for teaching reading were applied to two randomly selected groups of elementary students and then compared on the basis of a reading comprehension test given at the end of the learning period. The sample means and variances computed from the test scores are $\bar{Y}_1 = 64, S^2 = 52, n = 11$ and $\bar{Y}_2 = 69, S^2 = 71, n = 14$. What is the p value for testing if there is a difference in the mean scores for the populations associated with the two teaching methods?

We can also conduct test for the population variance. For example if we have a random sample, $Y_1, Y_2, \dots, Y_n \sim N(\mu, \sigma^2)$. Suppose we want to test $H_0 : \sigma^2 = \sigma_0^2$ versus various alternatives. Here we can use:

As our test statistic. Our picture becomes a bit different though.

The manufacturer of a machine to package soap powder claimed that her machine could load cartons at a given weight with a range of no more than .4 ounce. The mean and variance of a sample of eight 3-pound boxes were found to equal 3.1 and .018 respectively. Thes the hypothesis that the variance of the population of weight measurements is $\sigma^2 = .01$ against the alternative that $\sigma^2 > .01$

The next two lessons are hard and can be a bit confusing, so I want to just chat a bit about what's going on. Recall that α and β are the probabilty of committing Type I and Type II errors, respectively. Power is $1 - \beta$. A good test would be one that has a low α and a low β . What we're going to seek to do is find, for a fixed α the test that has the smallest β or in other words, the largest power. This might seem weird and we've only really given you one way of constructing a test. But the Neyman-Pearson Lemma say sthat if we want to test $H_0 : \theta = \theta_0$ vs $H_a : \theta = \theta_a$ what we should do is form the ration of

and find the value k such that the test has the desired value of α and that test is the most powerful. For example.

Suppose Y represents a single observation from a distribution that has pdf $f(y) = \theta y^{\theta-1}$. If we want to test $H_0 : \theta = 2$ vs $H_a : \theta = 1$ we would first for the ratio of:

So therefore, the most powerful test is when $2y < k$ for some k . If we want the most powerful test for $\alpha = .05$, we would have to find the value of k such that $P(y < k/2) = .05$. So we calculate:

So the Neyman-Pearson lemma gives the general form of the rejection region, but we still need to solve for k for a given α . Note that this is only for simple vs simple hypothesis, if we have the more common $H_0 : \theta = \theta_0$ vs $H_a : \theta > \theta_0$ we don't really have a theorem that can help us. However, if it turns out that when we form the ratio:

and the rejection region does not depend on the value of θ_a then we can actually say more than a Most Powerful test, we call this the Uniformly Most Powerful Test or the UMP test. For example, if we have $\theta \in \{1/4, 1/2, 3/4\}$ and a Most Powerful test of size 0.05 for $H_0 : \theta = 1/4$ vs $H_a : \theta = 1/2$ is ALSO the Most Powerful Test of size 0.05 of $H_0 : \theta = 1/4$ vs $H_a : \theta = 3/4$, then we can say that our test is the UMP test of $H_0 : \theta = 1/4$ vs $H_0 : \theta > 1/4$.

Look at Example 10.23. Read it. Struggle with it. Cry a little. Then we'll talk. . . .