Lecture 12

Other one-sample tests

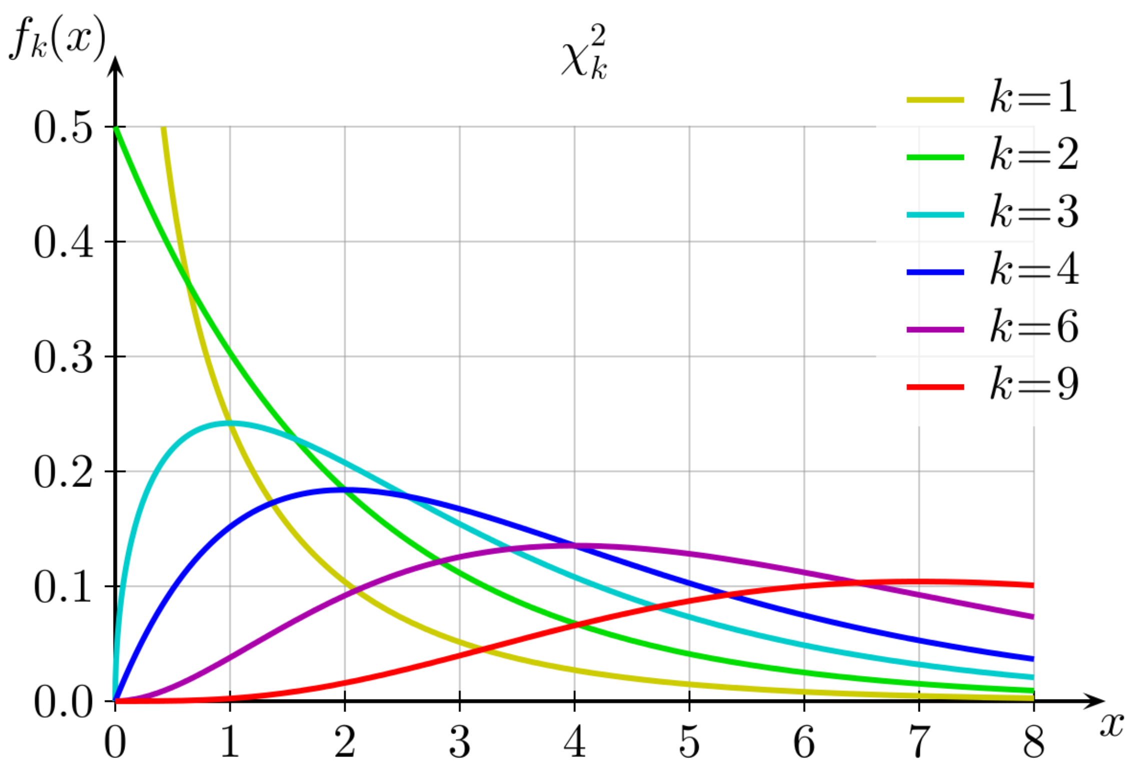
Our hypothesis tests come with certain assumptions. What is we don't meet those assumptions? Suppose we make a normal probability plot of our data and it's not normal? What if our proportion problem doesn't meet our test condition?

Let's look at the continuous case first. As we've seen, one strategy for dealing with distributions that are not normal is to increase the sample size. As the sample size increases, the central limit theorem tells us the sampling distribution will become more normal regardless of the underlying distribution. But what if that isn't possible?

Another option is to perform a transformation of the variables. Sometimes a variable may not be distributed sufficiently normally, but the log of it will be. Applying logarithms, powers (roots or other exponents), and other transformations may improve the normality in some cases. Obviously, this only works when raw data is available, and you may need to test different types of transformations to see if it is able to improve the appearance of your normal probability plot.

We've gotten a glimpse of some computational methods, which we will return to later in the course. There are also non-parametric techniques we will also look at later on. But, we do have some options for dealing with one-sample parameters in specific situations. It is to those cases that we want to turn now.

Two of the additional cases we'll look at use the χ^2 distribution.



The χ^2 distribution, for low degrees of freedom is strongly right skewed. We will encounter it again when we look at tests of independence later in the course, but we can also apply it here for tests on standard deviations, and tests on means from Poisson distributions.

Let's start with the standard deviation test.

In a sample of 15 analgesic drug abusers, the standard deviation of serum creatinine is found to be 0.435. The standard deviation of serum creatinine in the general population is 0.40. Determine if the variance of serum creatinine among analgesic abusers differs from the variance of serum creatinine in the general population.

We begin by setting up our hypothesis tests. Since standard deviation and variance are related, a test of standard deviation is equivalent to a test of variance.

$$H_0: \sigma = 0.40$$
$$H_a: \sigma \neq 0.40$$

I have used the standard deviation in my statements since those are the values provided in the problem. If you make your hypothesis statement in terms of variance here, you'll need to remember to square the population value.

$$H_0: \sigma^2 = 0.16$$
$$H_a: \sigma^2 \neq 0.16$$

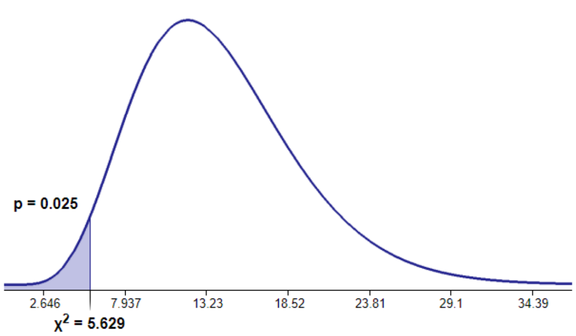
We need to calculate a test statistic. For this test, the test statistic is

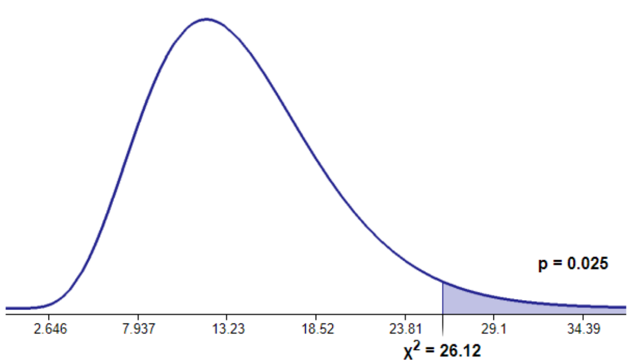
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

In this case, we find $\chi^2 = \frac{(15-1)(0.435)^2}{(0.40)^2} \approx 16.557$ … The degrees of freedom we use with this test is basically the same as we'd use for the t-test, $n - 1$.

Since this is a two-tailed test, we'll need to consider the rejection region on both ends of the distribution $\chi^2_{14,0.025}$ and $\chi^2_{14,0.975}$ as the boundary conditions, or find the P-value for the test.

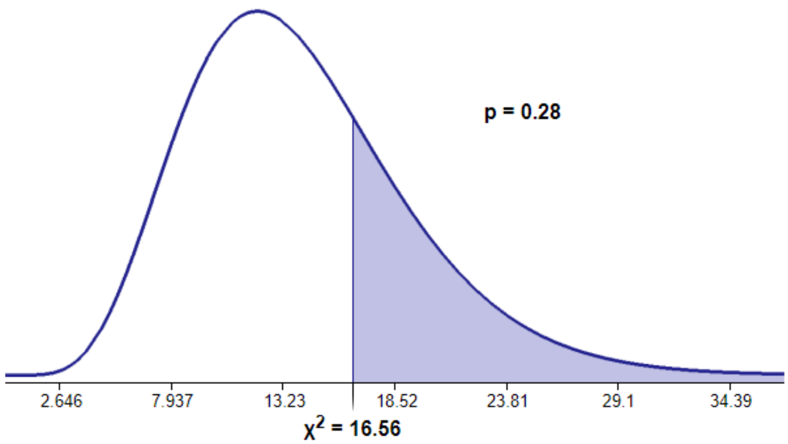
The boundaries are





Using the rejection region method, we see that our test statistic 16.56 falls in between 5.629 and 26.12, and therefore does not fall into the rejection region. We must fail to reject the null hypothesis.

Alternatively, if we look at the P-value, even if this was a one-tailed test, we see that the P-value is still greater than any common significance level.



Our conclusion, in context then, is that we can find no meaningful difference between our data's standard deviation and the general population's standard deviation.

We can also use the χ^2 distribution to conduct a test of means in a Poisson distribution. Consider the following example.

A recent occupational safety study found 21 bladder cancer deaths observed among rubber workers. Deaths are distributed with a Poisson distribution with an expected value of 18.1. Determine if this is sufficient evidence to think that bladder cancer deaths are more common among rubber workers than the general population.

$$H_0: \mu = 18.1$$
$$H_a: \mu > 18.1$$

Recall that in the Poisson distribution, we have one parameter (λ or in this case μ), which is both the value of the mean and the variance.

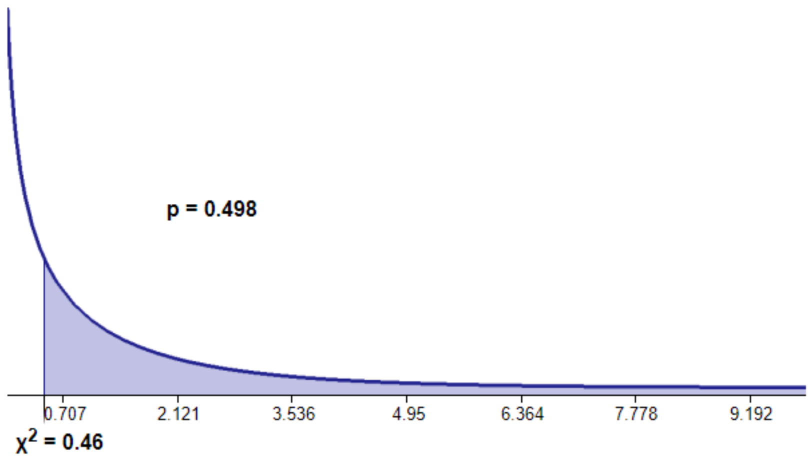
Our test statistic is

$$\chi^2 = \frac{(x - \mu_0)^2}{\mu_0}$$

We can conduct this test if $\mu_0 \geq 10$, and with one degree of freedom.

In this case we have $\chi^2 = \frac{(21-18.1)^2}{18.1} \approx 0.4646 \ldots$

We can find the P-value by looking at the probability to the right of this value.



This is much larger than any common significance level, so we fail to reject the null. This is not enough evidence to think bladder cancer is more common among rubber workers than the general population.

The last case we want to look at is the case where our proportion problem fails our test for the normal approximation. We'll need to use the binomial approximation directly in this case. (Recall that we also had a standard for approximating a binomial distribution as a Poisson distribution. This is an option to consider, but we'll omit it here.)

Example. Based on vital statistics, 20% of all deaths can be attributed to some form of cancer Of the 13 deaths occurring among 55-64 year old males in a nuclear power plant, 5 are from cancer. Is their death rate from cancer higher than the expected rate?

This is a very small sample for a proportion problem, and it fails our npq test. $13(.2)(.8) = 2.08 \ll 10.$ So, we have to look at the exact binomial distribution.

$$H_0: p = 0.2$$
$$H_a: p > 0.2$$

If the population proportion is 0.2, and we are doing 13 trials, we want the probability that we get 5 or more occurrences of cancer. The simplest way to do this is to use the cumulative distribution up to 4, and then subtract from 1.

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.9008 \ldots = 0.09913 \ldots$$

If we are using the standard significance level of $\alpha = 0.05$, then this P-value is too large, and we must fail to reject the null hypothesis. However, if we had set our significance level to 0.10, then this would be small enough to reject the null. This is why we should set our significance level well in advance of conducting our test. It's very tempting at the end to adjust the significance level in "borderline" cases to make a test significant when it should not be.

We can use an exact distribution to do the Poisson problem above if we wanted. As we develop more hypothesis tests, it's useful to consider the benefits and drawbacks of each, and when to apply them: what conditions to check, etc., and how they differ from other similar tests. I encourage you to do this.

Next time, we're going to add the two-sample cases, and there is a great variety of different approaches to the two-sample t-test with special applications.

References:

1. https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineering_and_the_sciences.pdf
2. https://assets.openstax.org/oscms-prodcms/media/documents/IntroductoryStatistics-OP_i6tAI7e.pdf
3. http://sites.msudenver.edu/ngrevsta/wp-content/uploads/sites/416/2020/01/3240chapter7.pdf
4. https://www.albany.edu/~msz03/sta552/ovh/c7_hypothesis.pdf
5. https://en.wikipedia.org/wiki/Chi-squared_distribution
6. http://www.statdistributions.com/chisquare/