Single Sample Tests for the Mean and Variance

Instances of Hypothesis Tests

In this section, we will introduce various test statistics that can be used for either Fisher tests or Neyman-Pearson decision tests. In either case, the emphasis is first on rejecting some null hypothesis at a certain significance level, either directly in a Fisher test or by the test statistic being in a certain critical region. We will also discuss OC curves, as used in Neyman-Pearson tests for most of these tests.

We have already described how to perform tests for the mean based on the normal distribution with known variance (sometimes called ***Z-tests***) and will not repeat these here.

The T-Test

17.1. T-Test. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal distribution and let $\overline{X}$ denote the sample mean, S^2 the sample variance. Let μ be the unknown population mean and μ_0 a null value of that mean. Then any test based on the statistic

$$T_{n-1} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

is called a ***T-test***.

We reject at significance level α

- $H_0 : \mu = \mu_0$ if $|T_{n-1}| > t_{\alpha/2, n-1}$,
- $H_0 : \mu \leq \mu_0$ if $T_{n-1} > t_{\alpha, n-1}$,
- $H_0 : \mu \geq \mu_0$ if $T_{n-1} < -t_{\alpha, n-1}$.

The T-Test

17.2. Example. The breaking strength of a textile fiber is a normally distributed random variable. Specifications require that the mean breaking strength should equal 150 psi. The manufacturer would like to detect any significant departure from this value. Thus, he wishes to test

$$H_0 \colon \mu = 150 \,\text{psi} \qquad\qquad H_1 \colon |\mu - 150 \,\text{psi}| > 2.5 \,\text{psi}$$

A random sample of 15 fiber specimens is selected and their breaking strengths determined. The statistic

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

will follow a T_{14}-distribution. We specify $\alpha = 0.05$, and find $t_{0.025,14} = 2.145$ from Table VI of the textbook. Thus, the critical region is given by $|t| > 2.145$.

The T-Test

The sample mean and variance are computed from the sample data as $\overline{x} = 152.18$ and $s^2 = 16.63$. Therefore, the test statistic is

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{152.18 - 150}{\sqrt{16.63/15}} = 2.07,$$

which does not fall into the critical region, so there is insufficient evidence to reject H_0 at the 5% level of significance.

Note that the T-distribution may be used for $\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$ when a sample is obtained from a normal population. If a sample is obtained from a non-normal population, care must be taken; for large to medium sample sizes ($n \geq 25$) it can be shown that violating the normality assumption does not significantly change α and β. For small sample sizes, a T-test cannot be used and an alternative (non-parametric) test must be employed; such tests will be discussed later.

OC Curves for the T-Test

The OC curves for T-test have a similar appearance to those for the normal distribution. However, when calculating the probability of failing to reject H_0 if $\mu = \mu_0 + \delta$, $\delta > 0$, as we did for the normal distribution, a difficulty occurs. We obtain the quotient of a non-standardized ($\mu \neq 0$) normal distribution with a chi-distribution. This leads to the concept of ***non-central T-distributions***, which we will not go into here.

OC Curves for the T-Test

The OC curves for the T-distribution feature an abscissa whose scale is given by

$$d = \frac{|\mu - \mu_0|}{\sigma},$$

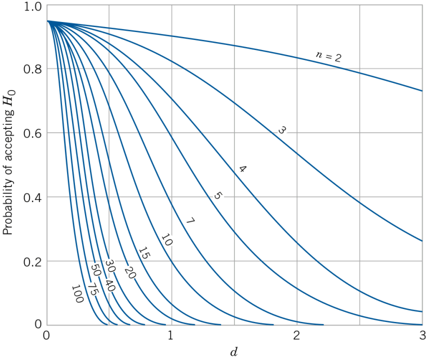
where σ is the ***unknown standard deviation*** of the random variable. We are left with three options:

1. If available, we can use prior experiments to insert a rough estimate for σ.
2. We can express the difference $\delta = |\mu - \mu_0|$ relative to σ, e.g., prescribing $d = \delta/\sigma < 1$ for a small difference in the mean or $d = \delta/\sigma < 2$ for a moderately large difference.
3. We substitute the sample standard deviation s for σ.

OC Curves for the T-Test

17.3. Example.

We return to Example 17.2. If the mean breaking strength of the fiber differs from 150 psi by 2.5 psi or more, we would like to reject the null hypothesis $H_0: \mu = 150\,\text{psi}$ with a probability of at least 0.9. Is the sample size $n = 15$ adequate to assure that the test is this sensitive?



If we use the previously obtained standard deviation $s = \sqrt{16.63} = 4.08$, then

$$d = \frac{|\mu - \mu_0|}{s} = \frac{2.5}{4.08} = 0.61.$$

The OC chart for $n = 15$, $\alpha = 0.05$, two-tailed, then gives $\beta \approx 0.45$. Thus the test is not powerful enough, since $1 - \beta = 0.55 < 0.9$.

The Chi-Squared Test

17.4. Chi-Squared Test. Let $X_1, \ldots, X_n$ be a random sample of size n from a normal distribution and let S^2 denote the sample variance. Let σ^2 be the unknown population variance and σ_0^2 a null value of that variance. Then a test for the variance based on the statistic

$$\chi_{n-1}^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

is called a ***chi-squared test***. We reject at significance level α

- $H_0: \sigma = \sigma_0$ if $\chi_{n-1}^2 > \chi_{\alpha/2,n-1}^2$ or $\chi_{n-1}^2 < \chi_{1-\alpha/2,n-1}^2$,
- $H_0: \sigma \leq \sigma_0$ if $\chi_{n-1}^2 > \chi_{\alpha,n-1}^2$,
- $H_0: \sigma \geq \sigma_0$ if $\chi_{n-1}^2 < \chi_{1-\alpha,n-1}^2$.

The Chi-Squared Test

It is important to be aware of the following difficulty:

- The T-distribution can be used in the presence of large sample sizes for the distribution of the sample mean even if the underlying distribution is non-normal.
- It is, however, not possible to approximate the χ^2_{n-1} statistic in this way if the distribution is non-normal, regardless of sample size! Therefore, normality of the data must first be tested, and if the data is non-normal, other methods must be used.

The Chi-Squared Test

17.5. Example. One random variable studied while designing the front-wheel-drive half-shaft of a new model automobile is the displacement (in millimeters) of the constant velocity (CV) joints. With the joint angle fixed at 12°, 20 simulations were conducted, resulting in the following data:

6.2	1.9	4.4	4.9	3.5
4.6	4.2	1.1	1.3	4.8
4.1	3.7	2.5	3.7	4.2
1.4	2.6	1.5	3.9	3.2

For these data, $\overline{x} = 3.39$ and $s = 1.41$. Engineers designing the front-wheel-drive half-shaft claim that the standard deviation in the displacement of the CV shaft is less than 1.5 mm. Do these data support this contention?

The Chi-Squared Test

We can translate the described situation into a Fisher test for $H_0\colon \sigma \geq 1.5$, which is equivalent to testing

$$H_0\colon \sigma^2 \geq 2.25.$$

From Table IV we obtain $\chi^2_{1-0.05,19} = 10.1$. Hence the test will have a P-value of less than 0.05 if

$$\frac{(n-1)s^2}{\sigma_0^2} < 10.1.$$

The observed value of the test statistic is

$$\frac{19 \cdot 1.41^2}{2.25} = 16.79.$$

Since this value is greater than 10.1, there is no evidence to reject H_0 at the 5% level of significance.

OC Curves for the Chi-Squared Test

The abscissa parameter for the OC curves for the two-tailed chi-squared test is

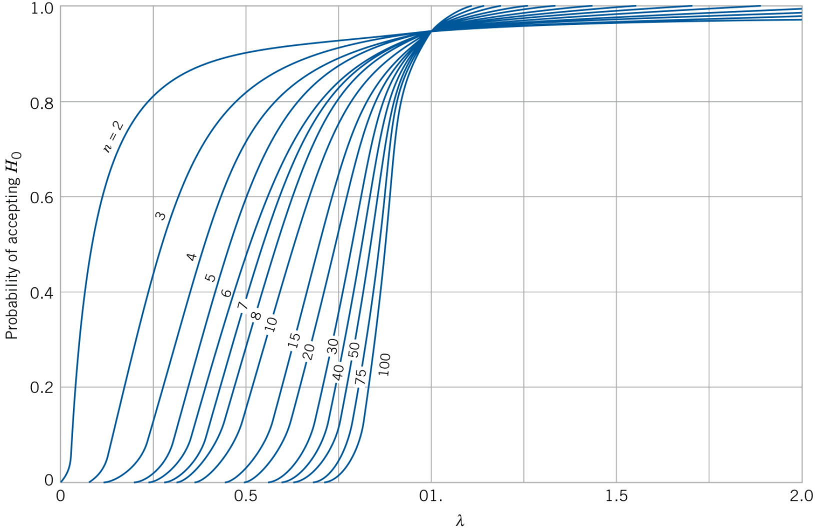
$$\lambda = \frac{\sigma}{\sigma_0}.$$

Note that the OC curves for the left- and right-tailed chi-squared distributions are distinct!

17.6. Example. Returning to Example 17.5, the engineers concerned are dissatisfied that H_0 was not rejected. A second test (this time of Neyman-Pearson type) is to be performed to establish that the standard deviation is less than $\sigma_0 = 1.5\,\mathrm{mm}$.

1. If we want to preset $\alpha = 0.05$, what is the critical region for the test at a sample size $n = 20$?
2. If $n = 20$, what true value of σ is necessary so that the test will have a power of $1 - \beta = 0.9$?
3. For $\alpha = 0.05$, make a statement on the sample size necessary to ensure that H_0 is rejected with 90% probability if $\sigma = 1.35$.

OC Curves for the Chi-Squared Test



OC Curves for Tests on the Variance

1. From the table for the χ^2_{19} distribution we see that $P[\chi^2_{1-0.05,19} \le 10.1] = 0.05$, so the critical region for the variance is

$$\frac{(n-1)s^2}{\sigma_0^2} < 10.1 \qquad \Leftrightarrow \qquad s^2 < \frac{2.25 \cdot 10.1}{19} = 1.20$$

   i.e., $s < 1.09$.

2. For $n = 20$, the line intersects the horizontal rule $\beta = 0.1$ at $\lambda = 0.6$. This means that

$$\sigma < 0.6\sigma_0 = 0.9$$

   is necessary for H_0 to be rejected 90% of the time.

3. The graph shows that a sample size significantly larger than $n = 100$ would be necessary.