Chi-Square Distribution

Inference about a Population Variance

If the population from which the sample is selected is (approximately) normally distributed, then

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

has a chi-square distribution with $n-1$ degrees of freedom.

A $(1-\alpha)100\%$ Confidence Interval for σ^2 is of the form

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}}.$$

The confidence interval for σ can be obtained by taking the square root of the two limits of the above interval.

Test about a Population Variance

Null Hypothesis: $H_0: \sigma^2 = \sigma_0^2$

Test Statistic: $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$

Alternative Hypothesis	Rejection region at the significance level α
$H_a: \sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_\alpha$ (upper-tailed test)
$H_a: \sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_\alpha$ (lower-tailed test)
$H_a: \sigma^2 \neq \sigma_0^2$	$\chi^2 > \chi^2_{\alpha/2}$ or $\chi^2 < \chi^2_{(1-\alpha/2)}$ (two-tailed test)

Goodness-of-fit Test

A **Multinomial Experiment** is an experiment with the following properties.

1. It consists of n identical independent trials.
2. Each trial results in one of k possible outcomes.
3. The probabilities p_i of the possible outcomes remain constant for each trial.

The test statistic for a goodness-of-fit test of a multinomial experiment is

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

where:

$O_i =$ observed frequency of category i, and $E_i =$ expected frequency of category $i = np_i$.

In a goodness-of-fit test, the degrees of freedom are $df = k-1$. It is always an upper-tailed test. The number of trials n should be large enough so that $np_i > 5$ for all i.

The p-value is $P(\chi^2 > \chi^2_{\text{obs}})$.