# Section 7-5 Estimating a Population Variance

This section covers the estimation of a **population variance** $\sigma^2$ and **standard deviation** $\sigma$.

## Estimator of $\sigma^2$

The **sample variance** $s^2$ is the best point estimate of the **population variance** $\sigma^2$.

## Estimator of $\sigma$

The **sample standard deviation** $s$ is a commonly used **point estimate of** $\sigma$.





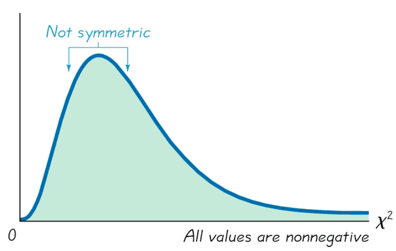
## Construction of confidence intervals for $\sigma^2$

We use the **chi-square distribution**, denoted by Greek character $\chi^2$ (pronounced *chi-square*).

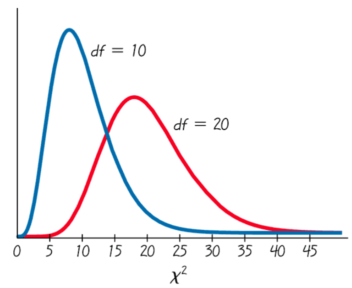
# Properties of the Chi-Square Distribution

1. The chi-square distribution is **not symmetric**, unlike the normal and Student *t* distributions.

**degrees of freedom = *n* – 1**




Chi-Square Distribution




Chi-Square Distribution for df = 10 and df = 20





# Properties of the Chi-Square Distribution

2. The values of chi-square can be zero or positive, but they **cannot be negative**.

3. The chi-square distribution is different for each number of degrees of freedom, which is df = *n* – 1.

   In Table A-4, each critical value of $\chi^2$ corresponds to an area given in the top row of the table, and that area represents the **cumulative area located to the right** of the critical value.

## Example

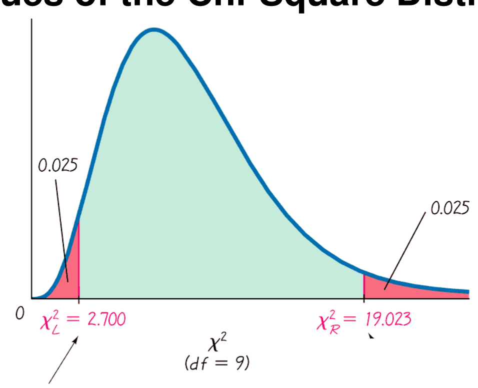
A sample of ten voltage levels is obtained. Construction of a confidence interval for the population standard deviation $\sigma$ requires the left and right critical values of $\chi^2$ corresponding to a confidence level of 95% and a sample size of $n = 10$.

Find the critical value of $\chi^2$ separating an area of 0.025 in the left tail, and find the critical value of $\chi^2$ separating an area of 0.025 in the right tail.




## Example

Critical Values of the Chi-Square Distribution



# Chi-Square Distribution

**TABLE A-4** Chi-Square ($\chi^2$) Distribution

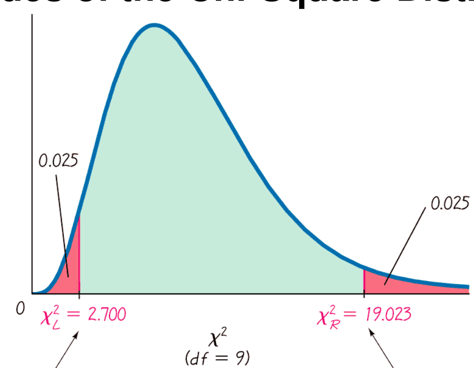
| Degrees of Freedom | 0.995 | 0.99 | 0.975 | 0.95 | 0.90 | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
|---|---|---|---|---|---|---|---|---|---|---|
| | Area to the Right of the Critical Value | | | | | | | | | |
| 1 | — | — | 0.001 | 0.004 | 0.016 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 |
| 3 | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| 4 | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| 5 | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 9.236 | 11.071 | 12.833 | 15.086 | 16.750 |
| 6 | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7 | 0.989 | 1.239 | 1.690 | 2.167 | 2.833 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8 | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| 9 | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |
| 10 | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 |





## Example

### Critical Values of the Chi-Square Distribution




*To obtain this critical value, locate 9 at the left column for degrees of freedom and then locate 0.975 across the top. The total area to the right of this critical value is 0.975, which we get by subtracting 0.025 from 1.*

*To obtain this critical value, locate 9 at the left column for degrees of freedom and then locate 0.025 across the top.*

## Confidence Interval for Estimating a Population Variance

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





## Confidence Interval for Estimating a Population Standard Deviation

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

## Confidence Interval for Estimating a Population Standard Deviation

For a sample of 10 voltages:

123.3 , 123.5, 123.7, 123.4, 123.6, 123.5, 123.5, 123.4, 123.6, 123.8

The standard deviation s= 0.15

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

$$\frac{(10-1)(0.15)^2}{19.023} < \sigma^2 < \frac{(10-1)(0.15)^2}{2.700}$$

$$0.010645 < \sigma^2 < 0.075000.$$

$$0.10 \text{ volt} < \sigma < 0.27 \text{ volt}.$$





## Definitions

**In statistics, a hypothesis is a claim or statement about a property of a population.**

**A hypothesis test is a standard procedure for testing a claim about a property of a population.**

## Main Objectives

We will study hypothesis testing for

1. population proportion $p$
2. population mean $\mu$
3. population standard deviation $\sigma$





## Example

**Claim**: the XSORT method of gender selection **increases** the likelihood of having a baby girl.

**This is a claim about proportion (of girls)**

**To test this claim 10 couples (volunteers) were subject to XSORT treatment.**

**If 4 or 5 or 6 have girls, the method probably does not increase the likelihood of a girl.**

**If 9 or 10 couples have girls, the method is probably increases the likelihood of a girl.**

## Rare Event Rule for Inferential Statistics

**If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.**




# Components of a Formal Hypothesis Test

## Null Hypothesis: $H_0$

- The **null hypothesis** (denoted by $H_0$) is a statement that the value of a population parameter (such as *proportion, mean*, or *standard deviation*) is **equal to** some claimed value.
- We test the null hypothesis directly.
- Either reject $H_0$ or fail to reject $H_0$ (in other words, accept $H_0$ ).




## Alternative Hypothesis: $H_1$

- The **alternative hypothesis** (denoted by $H_1$) is the statement that the parameter has a value that **somehow differs** from the null hypothesis.
- The symbolic form of the alternative hypothesis must use one of these symbols:

  $\neq$, $<$, $>$. (not equal, less than, greater than)

## Example 1

**Claim: the XSORT method of gender selection increases the likelihood of having a baby girl.**

**We express this claim in symbolic form: $p > 0.5$ (here $p$ denotes the proportion of baby girls)**

**Null hypothesis must say "equal to", so**

$$H_0 : p = 0.5$$

**Alternative hypothesis must express difference:**

$$H_1 : p > 0.5$$

**Original claim is now the alternative hypothesis**





## Example 1 (continued)

**We always test the null hypothesis.**

**If we reject the null hypothesis, then the original clam is <u>accepted</u>.**

**Final conclusion would be: XSORT method increases the likelihood of having a baby girl.**

**If we fail to reject the null hypothesis, then the original clam is <u>rejected</u>.**

**Final conclusion would be: XSORT method does not increase the likelihood of having a baby girl.**

## Example 2

**Claim: for couples using the XSORT method the likelihood of having a baby girl is 50%**

**Express this claim in symbolic form: $p=0.5$**
**(again $p$ denotes the proportion of baby girls)**

**Null hypothesis must say "equal to", so**

$$H_0 : p=0.5$$

**Alternative hypothesis must express difference:**

$$H_1 : p\neq 0.5$$

**Original claim is now the null hypothesis**




## Example 2 (continued)

**If we reject the null hypothesis, then the original claim is rejected.**

**Final conclusion would be: for couples using the XSORT, the likelihood of having a baby girl is not 0.5**

**If we fail to reject the null hypothesis, then the original claim is accepted.**

**Final conclusion would be: for couples using the XSORT the likelihood of having a baby girl is indeed equal to 0.5**

## Example 3

**<u>Claim</u>: for couples using the XSORT method the likelihood of having a baby girl is <u>at least 0.5</u>**

**Express this claim in symbolic form: $p \geq 0.5$**
**(again $p$ denotes the proportion of baby girls)**

**Null hypothesis must say "equal to", so**
**$H_0 : p = 0.5$ (this agrees with the claim!)**
**Alternative hypothesis must express difference:**
**$H_1 : p < 0.5$**
**Original claim is now the null hypothesis**




## Example 3 (continued)

**If we reject the null hypothesis, then the original clam is rejected.**

**Final conclusion would be: for couples using the XSORT, the likelihood of having a baby girl is less 0.5**

**If we fail to reject the null hypothesis, then the original claim is accepted.**

**Final conclusion would be: for couples using the XSORT the likelihood of having a baby girl is indeed at least 0.5**