# STATISTICS II

**Bachelor's degrees in Economics, Finance and Management**

2nd year/2nd Semester
2025/2026

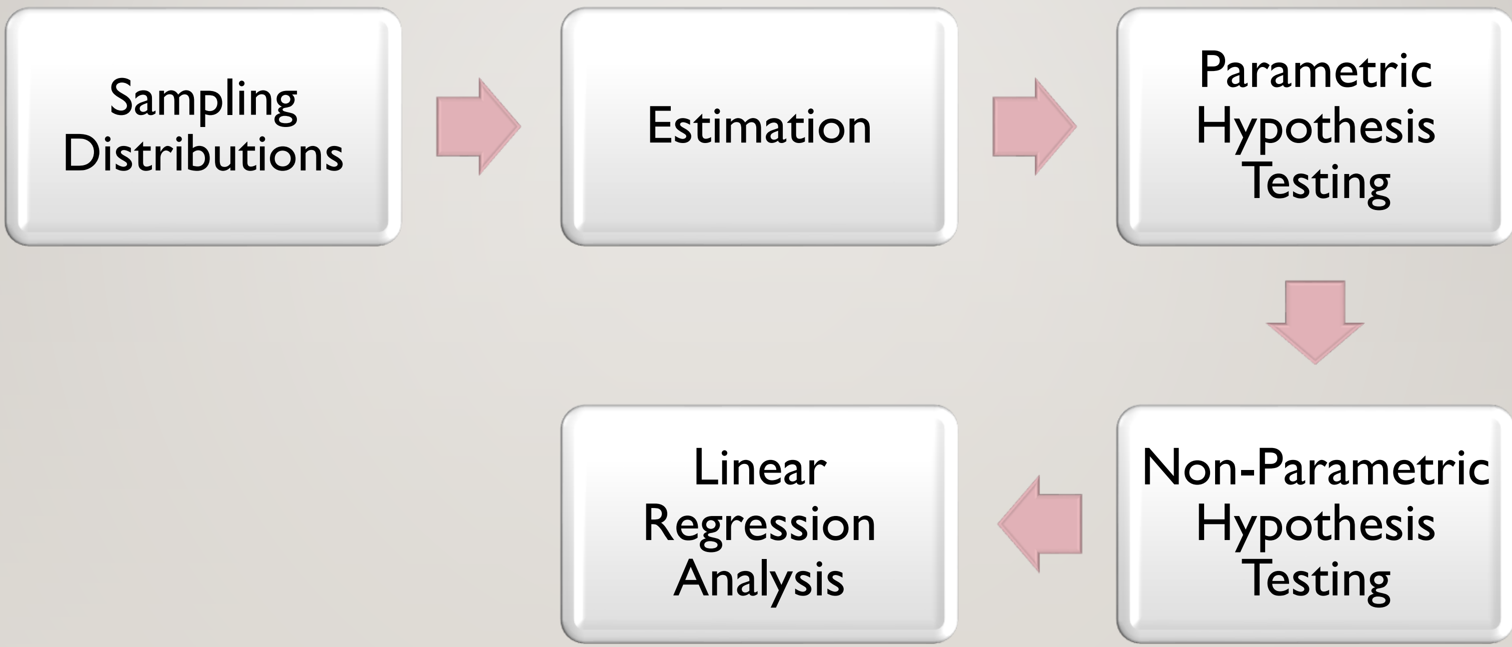
# CONTACT

**Professor**: Elisabete Fernandes
**E-mail**: efernandes@iseg.ulisboa.pt


https://doity.com.br/estatistica-aplicada-a-nutricao


https://basiccode.com.br/produto/informatica-basica/

# PROGRAM


Sampling Distributions
Estimation
Parametric Hypothesis Testing
Non-Parametric Hypothesis Testing
Linear Regression Analysis

# LECTURE 4 HOMEWORK: QUESTIONS AND SOLUTIONS

# EXERCISE 6.48

6.48 A random sample of size $n = 25$ is obtained from a normally distributed population with a population mean of $\mu = 198$ and a variance of $\sigma^2 = 100$.

a. What is the probability that the sample mean is greater than 200?

b. What is the value of the sample variance such that 5% of the sample variances would be less than this value?

c. What is the value of the sample variance such that 5% of the sample variances would be greater than this value?

Newbold et al (2013)

# EXERCISE 6.48 A): SOLUTION

Answer:

**Given:** $\mu = 198,\ \sigma^2 = 100\ (\sigma = 10),\ n = 25.$

**$\bar{X}$ ~ Normal(μ, σ²/n)**

(a) $P(\bar{X} > 200)$

$$\bar{X} \sim N\left(198, \frac{\sigma^2}{n}\right) = N\left(198, \frac{100}{25}\right) = N(198, 4),$$

Standard Deviation of $\bar{X}$

so $\mathrm{sd}(\bar{X}) = 2.$

$$P(\bar{X} > 200) = P\left(Z > \frac{200 - 198}{2}\right) = P(Z > 1) = 1 - \Phi(1) \approx 0.158655\ (\approx 0.158$$

Standard Normal Distribution Table

**Answer (a):** $P(\bar{X} > 200) \approx 0.1587.$

# EXERCISE 6.48 B): SOLUTION

Answer:

**Given:** $\mu = 198,\ \sigma^2 = 100\ (\sigma = 10),\ n = 25.$

**(b) Value $s^2_{\text{low}}$ such that $P(S^2 < s^2_{\text{low}}) = 0.05$**

Let $\chi^2_{0.05,24}$ denote the 5th percentile of $\chi^2_{24}$. Numerically $\chi^2_{0.05,24} \approx 13.8484$. Then

$$s^2_{\text{low}} = \sigma^2 \frac{\chi^2_{0.05,24}}{n-1} = 100 \cdot \frac{13.8484}{24} \approx 57.70.$$

(sample standard deviation $\sqrt{57.70} \approx 7.596$.)

**Answer (b):** $s^2_{\text{low}} \approx 57.70$ (so $S \approx 7.60$).

$\mathbf{Q} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

Chi-Square Distribution Table

**Alternative Solution:**

$P(S^2 < a) = 0.05 \Leftrightarrow P\left(\frac{(n-1)S^2}{\sigma^2} < \frac{(n-1)a}{\sigma^2}\right) = 0.05 \Leftrightarrow$

$P\left(Q < \frac{24 \times a}{100}\right) = 0.05$

Then $\frac{24 \times a}{100} = \chi^2_{0.05;24} = 13.8484$

$\Leftrightarrow$ a = 100 × 13.8484/ 24 = 57.70

# EXERCISE 6.48 C): SOLUTION

Answer:

**Given:** $\mu = 198,\ \sigma^2 = 100\ (\sigma = 10),\ n = 25.$

**(c) Value $s^2_{\text{high}}$ such that $P(S^2 > s^2_{\text{high}}) = 0.05$**

$Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

Let $\chi^2_{0.95,24} \approx 36.4150$. Then

Chi-Square Distribution Table

$$s^2_{\text{high}} = 100 \cdot \frac{36.4150}{24} \approx 151.73.$$

(sample standard deviation $\sqrt{151.73} \approx 12.32$.)

**Answer (c):** $s^2_{\text{high}} \approx 151.73$ (so $S \approx 12.32$).

**Alternative Solution:**

$P(S^2 > a) = 0.05 \Leftrightarrow P\left(\frac{(n-1)S^2}{\sigma^2} > \frac{(n-1)a}{\sigma^2}\right) = 0.05 \Leftrightarrow$

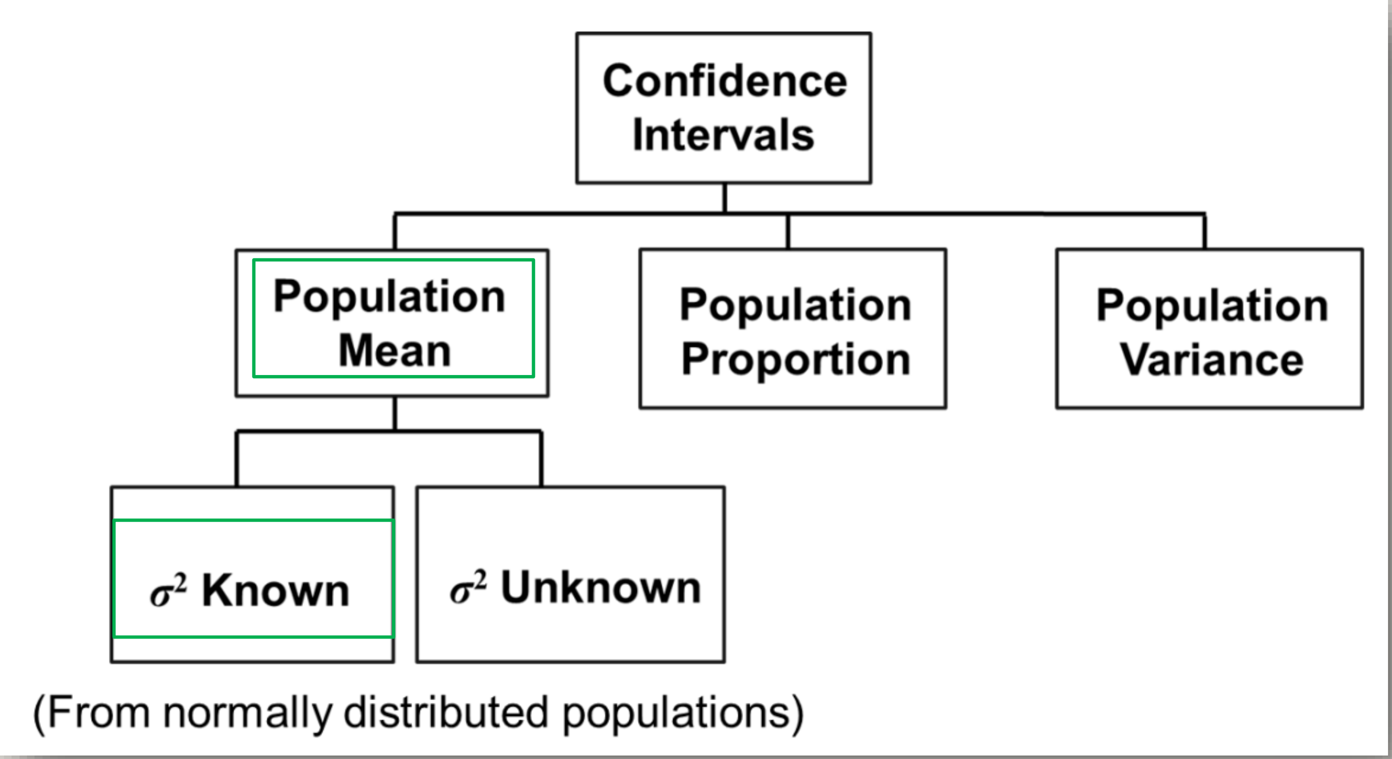
$P\left(Q < \frac{24 \times a}{100}\right) = 0.95$

Then $\frac{24 \times a}{100} = \chi^2_{0.95;24} = 36.4150$

$\Leftrightarrow a = 100 \times 36.4150 / 24 = 151.73$

# LECTURE 5: CONFIDENCE INTERVAL FOR THE MEAN ($\sigma^2$ KNOWN)

# CONFIDENCE INTERVALS WE WILL CONSIDER




Newbold et al (2013)

# CONFIDENCE INTERVAL ESTIMATE FOR THE MEAN ($\sigma^2$ KNOWN)

- Assumptions
  - Population variance $\sigma^2$ is known
  - Population is normally distributed
  - If population is not normal, use large sample
- Confidence interval estimate:

n ≥ 25

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

( $z_{1-\frac{\alpha}{2}}$ is the normal distribution value for a probability of $1-\frac{\alpha}{2}$ in each tail)

Quantile of the standard normal distribution corresponding to probability 1- α/2

Newbold et al (2013)

# CI ESTIMATE FOR THE MEAN ($\sigma^2$ KNOWN): CONFIDENCE LIMITS

- The confidence interval is

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

- The endpoints of the interval are

$$\text{UCL} = \bar{x} + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

Upper confidence limit

$$\text{LCL} = \bar{x} - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

Lower confidence limit

Newbold et al (2013)

# CI ESTIMATE FOR THE MEAN ($\sigma^2$ KNOWN): MARGIN OF ERROR

- The confidence interval,

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

- Can also be written as $\bar{x} \pm ME$
  where *ME* is called the margin of error

$$ME = z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

- The interval width, *w*, is equal to twice the margin of error

**Note:**
The **margin of error** is half the width of the confidence interval. Equivalently, the **width of the interval** is twice the margin of error.

- Margin of error = (interval width) / 2
- Interval width = 2 × (margin of error)

Newbold et al (2013)

Width of Confidence Interval for the Mean (σ known)

$$\text{Width} = 2 \cdot z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Where:

- $z_{1-\alpha/2}$ = critical value (quantile) from the standard normal distribution
- $\sigma$ = population standard deviation
- $n$ = sample size

# CI ESTIMATE FOR THE MEAN ($\sigma^2$ KNOWN): REDUCING THE MARGIN OF ERROR

$$ME = z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

The margin of error can be reduced if

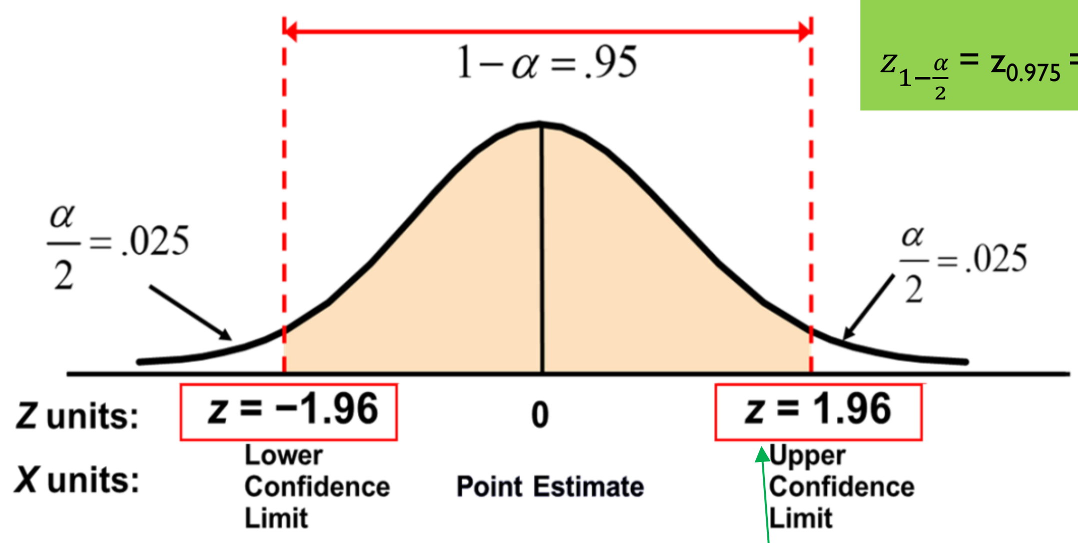
- the population standard deviation can be reduced ($\sigma \downarrow$)
- The sample size is increased ($n \uparrow$)
- The confidence level is decreased, $(1-\alpha)\downarrow$

**Note:**
The **margin of error** decreases when the **standard deviation** is smaller or the **sample size** is larger, and increases when the **critical value** (confidence level) is higher.

Newbold et al (2013)

# Z-QUANTILES FOR A 95% CONFIDENCE INTERVAL

- Consider a 95% confidence interval:




- Find $Z_{0.975} = 1.96$ from the standard normal distribution table

Newbold et al (2013)

$1-\alpha = 0.95 \Leftrightarrow \alpha = 0.05$

$1- \alpha/2 = 0.975$

$z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$

$z_{0.975} = 1.96$ is the quantile of the standard normal distribution corresponding to a cumulative probability of 0.975.

# COMMON LEVELS OF CONFIDENCE

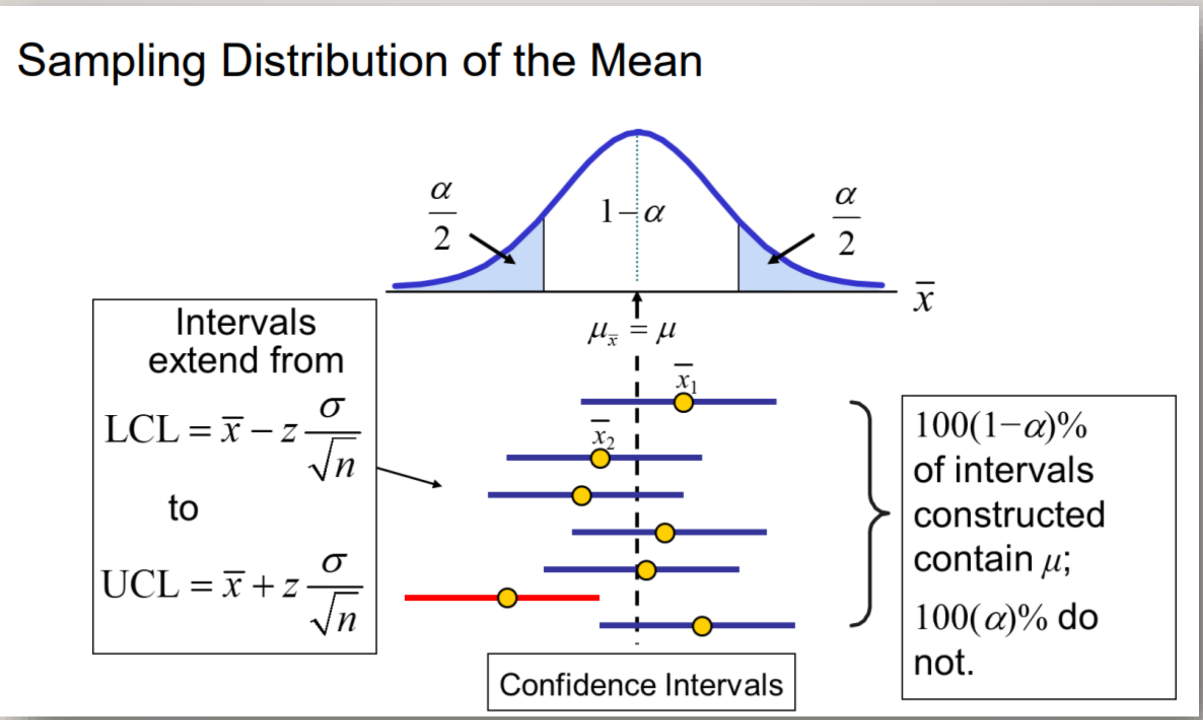
- Commonly used confidence levels are 90%, 95%, 98%, and 99%

| Confidence Level | Confidence Coefficient, $1-\alpha$ | Quantiles $z_{1-\frac{\alpha}{2}}$ |
|---|---|---|
| 80% | .80 | 1.28 |
| 90% | .90 | 1.645 |
| 95% | .95 | 1.96 |
| 98% | .98 | 2.33 |
| 99% | .99 | 2.58 |
| 99.8% | .998 | 3.08 |
| 99.9% | .999 | 3.27 |

Newbold et al (2013)

**Note:**
•The most commonly used confidence levels are **90%, 95%, and 99%**.

# INTERVALS AND LEVEL OF CONFIDENCE




Newbold et al (2013)

**Note:**
•The **confidence level** is the probability that the interval, not the parameter, **contains the true population value**.

# CI ESTIMATE FOR THE $\mu$ ($\sigma^2$ KNOWN): EXAMPLE

- A sample of 11 circuits from a large normal population has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is 0.35 ohms.
- Determine a 95% confidence interval for the true mean resistance of the population.

Newbold et al (2013)

# CI ESTIMATE FOR THE $\mu$ ($\sigma^2$ KNOWN): EXAMPLE

- A sample of 11 circuits from a large normal population has a mean resistance of 2.20 ohms. We know from past testing that the population standard deviation is .35 ohms.
- Solution:

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

$$= 2.20 \pm 1.96\left(\frac{.35}{\sqrt{11}}\right)$$

$$= 2.20 \pm .2068$$

$$1.9932 < \mu < 2.4068$$

n = 11 (sample size)
$\bar{x}$ = 2.20 (sample mean)
σ = 0.35 (population standard deviation)
1 – α = 0.95 (confidence level), then α = 0.05 and $z_{1-\alpha/2} = z_{0.975}$ = 1.96 (see previous slide)

Newbold et al (2013)

$CI_{0.95}(\mu)$ = (1.9932; 2.4068)

- We are 95% confident that the true mean resistance is between 1.9932 and 2.4068 ohms
- Although the true mean may or may not be in this interval, 95% of intervals formed in this manner will contain the true mean

# EXERCISE 7.13

7.13 A personnel manager has found that historically the scores on aptitude tests given to applicants for entry-level positions follow a normal distribution with a standard deviation of 32.4 points. A random sample of nine test scores from the current group of applicants had a mean score of 187.9 points.

a. Find an 80% confidence interval for the population mean score of the current group of applicants.
b. Based on these sample results, a statistician found for the population mean a confidence interval extending from 165.8 to 210.0 points. Find the confidence level of this interval.

Newbold et al (2013)

# EXERCISE 7.13 A): SOLUTION

Answer:

Given: $\sigma = 32.4,\ n = 9,\ \bar{x} = 187.9.$

Because the population standard deviation $\sigma$ is known and the population is assumed normal, we use the normal (z) distribution.

**(a) 80% confidence interval for $\mu$**

$$IC_{(1-\alpha)}(\mu) = \left(\bar{x} - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)$$

Standard error:

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{32.4}{\sqrt{9}} = \frac{32.4}{3} = 10.8.$$

Critical value for an 80% two-sided interval: $Z_{0.90} = 1.28155$

$1\text{-}\alpha = 0.80 \Leftrightarrow \alpha = 0.20$

$1\text{-}\ \alpha/2 = 0.90$

$z_{1-\frac{\alpha}{2}} = z_{0.90} = 1.28155$

Standard Normal Distribution Table

Margin of error:

$$ME = Z_{0.90}\ SE = 1.28155 \times 10.8 \approx 13.8427.$$

Confidence interval:

$$\bar{x} \pm ME = 187.9 \pm 13.8427$$

$$\boxed{80\%\ \mathrm{CI}:\ [174.06,\ 201.74]\ \text{(approximately)}}$$

# EXERCISE 7.13 B): SOLUTION

Answer:

Given: $\sigma = 32.4,\ n = 9,\ \bar{x} = 187.9.$

Because the population standard deviation $\sigma$ is known and the population is assumed normal, we use the normal (z) distribution.

$$IC_{(1-\alpha)}(\mu) = \left(\bar{x} - z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right)$$

**(b) Confidence level for the interval $[165.8,\ 210.0]$**

The interval is centered at $\bar{x} = 187.9$. Margin of error:

$$ME = 210.0 - 187.9 = 22.1.$$

Compute the corresponding z:

$$z_{1-\frac{\alpha}{2}} = \frac{ME}{SE} = \frac{22.1}{10.8} \approx 2.0463.$$

The central probability is $P(-z < Z < z) = 2\Phi(z) - 1$. Using $\Phi(2.0463) \approx 0.979636$,

$$\text{Confidence level} = 2(0.979636) - 1 \approx 0.95927 \approx 95.93\%.$$

**Alternative Solution:**

ME = $z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$ = (210.0 – 165.8)/2 = 22.1

Then $z_{1-\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$ = 22.1 ⇔ $z_{1-\frac{\alpha}{2}}$= 22.1 × $\sqrt{n}$/σ=2.0463

Using Standard Normal Distribution Table:

$z_{0.9798} \sim$ 2.05 then $1 - \frac{\alpha}{2} = 0.9798$

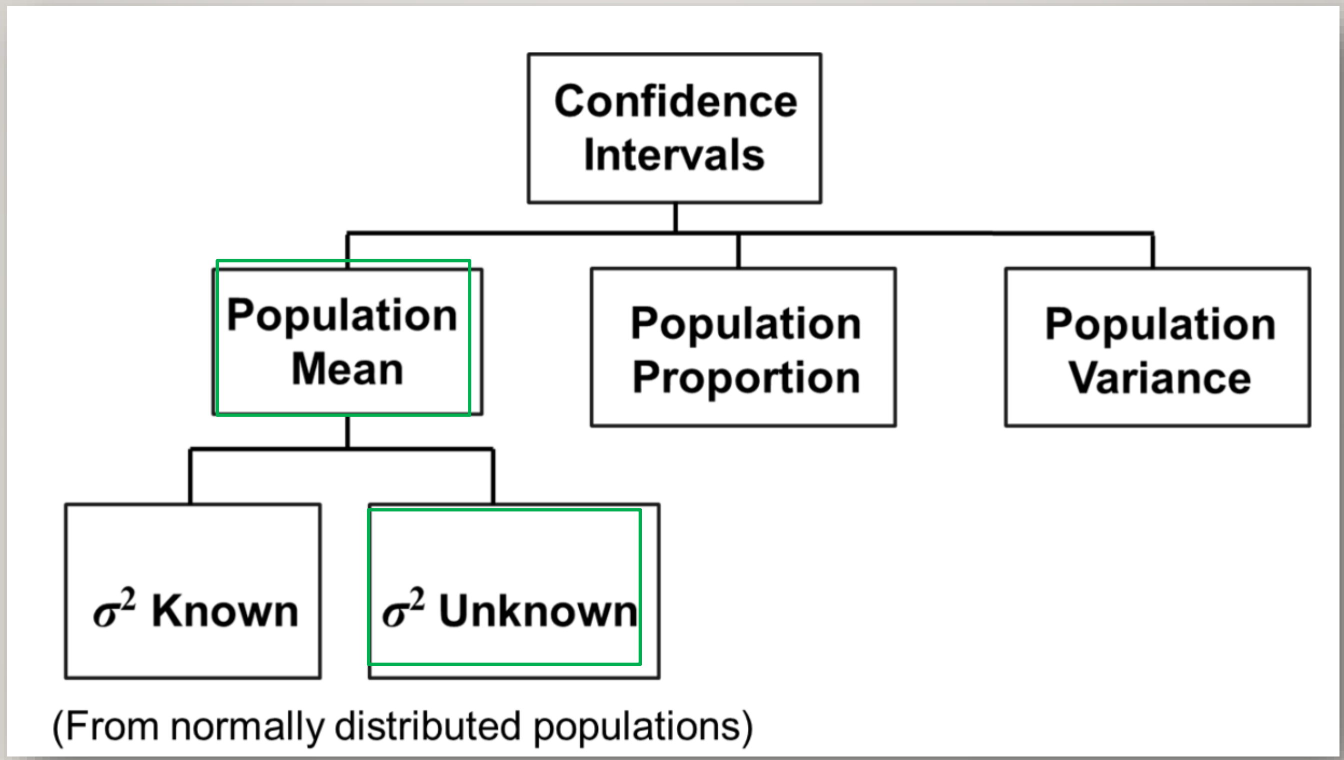
⇔ α = 0.0404

1- α = 1-0.0404= 0.9596 (95.96%)

**Note:**
The **margin of error** is half the width of the confidence interval.

# LECTURE 5: CONFIDENCE INTERVAL FOR THE MEAN ($\sigma^2$ UNKNOWN)

# CONFIDENCE INTERVAL ESTIMATION FOR THE MEAN ($\sigma^2$ UNKNOWN)




Newbold et al (2013)

# CONFIDENCE INTERVAL ESTIMATION FOR THE MEAN ($\sigma^2$ UNKNOWN)

- If the population standard deviation $\sigma$ is unknown, we can substitute the sample standard deviation, *s*
- This introduces extra uncertainty, since *s* is variable from sample to sample
- So we use the *t* distribution instead of the normal distribution

Newbold et al (2013)

# CONFIDENCE INTERVAL ESTIMATION FOR THE MEAN ($\sigma^2$ UNKNOWN)

$$CI_{(1-\alpha)}(\mu) = \left(\bar{x} - t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}\right)$$

- Assumptions
  - Population standard deviation is unknown
  - Population is normally distributed
  - If population is not normal, use large sample
- Use Student's *t* Distribution
- Confidence Interval Estimate:

$$\bar{x} \pm t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}$$

$t_{1-\frac{\alpha}{2};n-1}$ is the critical value of the *t* distribution with *n* − 1 d.f.

and an area of $\frac{\alpha}{2}$ in each tail: $P\left(t_{n-1} > t_{1-\frac{\alpha}{2};n-1}\right) = \frac{\alpha}{2}$

# CI ESTIMATE FOR THE MEAN ($\sigma^2$ UNKNOWN): MARGIN OF ERROR

- The confidence interval,

$$CI_{(1-\alpha)}(\mu) = \left(\bar{x} - t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}\right)$$

$$\bar{x} \pm t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}$$

- Can also be written as $\bar{x} \pm ME$

  where *ME* is called the margin of error:

$$\text{MSE} = t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}$$

Newbold et al (2013)

# CI ESTIMATE FOR THE $\mu$ ($\sigma^2$ UNKNOWN): EXAMPLE

A random sample of $n$ = 25 has $\bar{x} = 50$ and $s$ = 8. Form a 95% confidence interval for $\mu$

– d.f. = $n$ − 1 = 24, so $t_{1-\frac{\alpha}{2};n-1} = t_{0.975;24}$ = 2.0639

The confidence interval is

$$CI_{(1-\alpha)}(\mu) = \left(\bar{x} - t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}\right)$$

$$\bar{x} \pm t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}$$

n = 25 (sample size)
$\bar{x} = 50$ (sample mean)
s = 8 (sample standard deviation)
1 – α = 0.95 (confidence level), then
$t_{0.975;24}$ = 2.0639 (see Student´s table)

$$50 \pm (2.0639)\frac{8}{\sqrt{25}}$$

$$46.698 < \mu < 53.302$$

$CI_{95\%}$ (μ) = (46.698, 53.302)

Newbold et al (2013)

# EXERCISE 7.28

7.28 A business school placement director wants to estimate the mean annual salaries 5 years after students graduate. A random sample of 25 such graduates found a sample mean of $42,740 and a sample standard deviation of $4,780. Find a 90% confidence interval for the population mean, assuming that the population distribution is normal.

Newbold et al (2013)

# EXERCISE 7.28: SOLUTION

✅ Answer:

**Given data**

$$n = 25, \quad \bar{x} = 42{,}740, \quad s = 4{,}780$$

Because the **population standard deviation** is *unknown* and the **sample size is small (n = 25)**, we use the **t-distribution** with

$$df = n - 1 = 24$$

n = 25 (sample size)
$\bar{x} = 42.740$ (sample mean)
s = 4.780 (sample standard deviation)
$1 - \alpha = 0.90$ (confidence level),
then $t_{1-\alpha/2;\, n-1} = t_{0.95;24} = 1.711$ (see Student´s table)

**Compute the standard error**

$$SE = \frac{s}{\sqrt{n}} = \frac{4780}{\sqrt{25}} = \frac{4780}{5} = 956$$

**Find the critical t-value for a 90% confidence interval**

m the t-table (or calculator):

$$t_{0.95;24} = 1.711$$

Student´s t Distribution Table

# EXERCISE 7.28: SOLUTION

Answer:

**Given data**

$$n = 25, \quad \bar{x} = 42{,}740, \quad s = 4{,}780$$

Because the **population standard deviation** is *unknown* and the **sample size is small (n = 25)**, we use the **t-distribution** with

$$df = n - 1 = 24$$

n = 25 (sample size)
$\bar{x} = 42.740$ (sample mean)
s = 4.780 (sample standard deviation)
1 – α = 0.90 (confidence level),
then $t_{1-\alpha/2;\, n-1} = t_{0.95;24} = 1.711$ (see Student´s table)

**Compute the margin of error**

$$ME = t_{0.95;24} \times SE = 1.711 \times SE = 1.711 \times 956 = 1636.82$$

**Construct the confidence interval**

$$CI_{(1-\alpha)}(\mu) = \left(\bar{x} - t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}, \bar{x} + t_{1-\frac{\alpha}{2};n-1} \times \frac{s}{\sqrt{n}}\right)$$

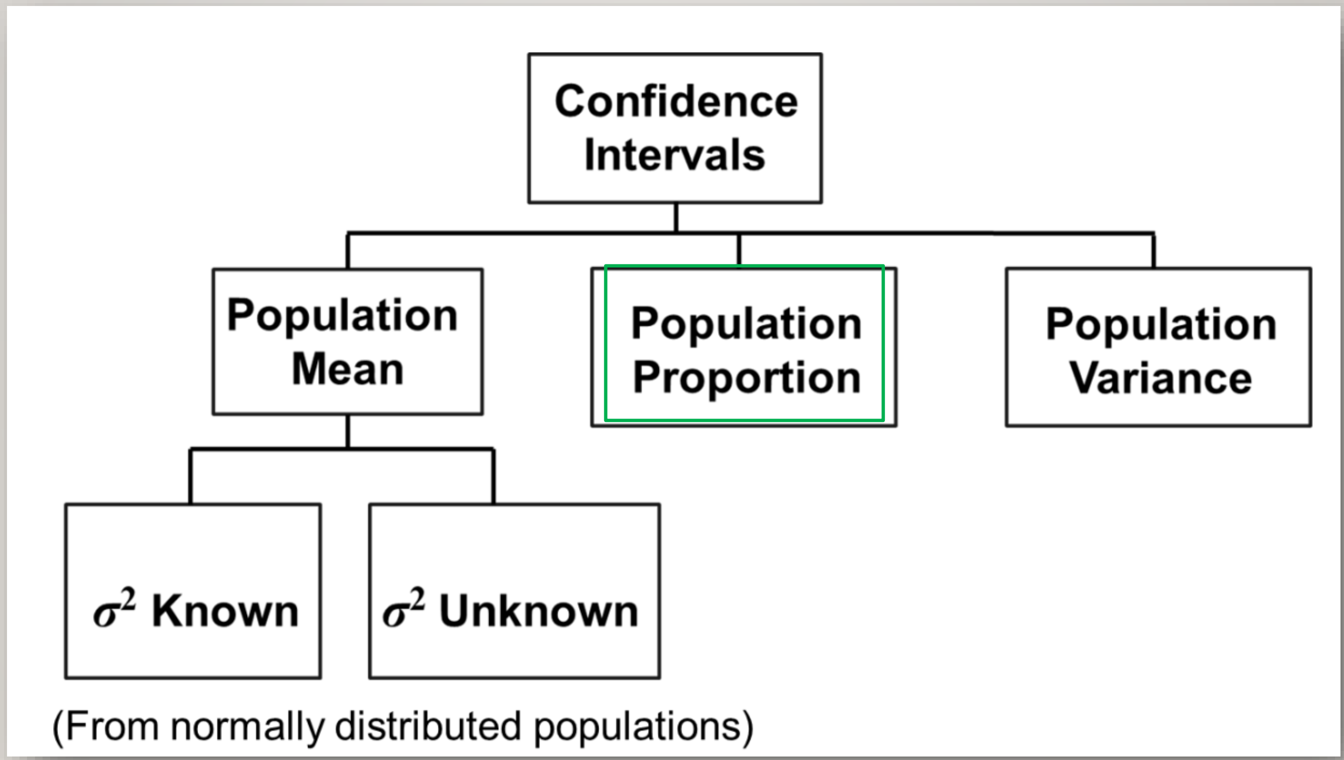
$$\bar{x} \pm ME = 42{,}740 \pm 1{,}637$$

$$\boxed{90\%\ CI : [41{,}103,\ 44{,}377]}$$

$CI_{90\%}(\mu) = (41.103, 44.377)$

# LECTURE 5: CONFIDENCE INTERVAL FOR THE PROPORTION

# CONFIDENCE INTERVAL ESTIMATION FOR POPULATION PROPORTION




Newbold et al (2013)

# CONFIDENCE INTERVAL ESTIMATION FOR POPULATION PROPORTION

- An interval estimate for the population proportion ($P$) can be calculated by adding an allowance for uncertainty to the sample proportion $(\hat{p})$

Newbold et al (2013)

# CONFIDENCE INTERVALS FOR THE POPULATION PROPORTION

- Recall that the distribution of the sample proportion is approximately normal if the sample size is large, with standard deviation

$$\sigma_P = \sqrt{\frac{P(1-P)}{n}}$$

$$Z = \frac{\hat{p} - P}{\sqrt{\frac{P(1-P)}{n}}} \sim \text{Normal}(0, 1)$$

- We will estimate this with sample data:

$$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Newbold et al (2013)

# CONFIDENCE INTERVAL ENDPOINTS

- The confidence interval for the population proportion is given by

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- where

  $z_{1-\frac{\alpha}{2}}$ is the standard normal value for the level of confidence desired

  - $\hat{p}$ is the sample proportion
  - $n$ is the sample size
  - $n \geq 25$

Newbold et al (2013)

$$CI_{(1-\alpha)}(p) = \left(\hat{p} - z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}(\mathbf{1}-\hat{p})}{\boldsymbol{n}}}, \hat{p} + z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}(\mathbf{1}-\hat{p})}{\boldsymbol{n}}}\right)$$

# CONFIDENCE INTERVALS FOR THE POPULATION PROPORTION: EXAMPLE

- A random sample of 100 people shows that 25 are left-handed.
- Form a 95% confidence interval for the true proportion of left-handers

Newbold et al (2013)

# CONFIDENCE INTERVALS FOR THE POPULATION PROPORTION: EXAMPLE

- A random sample of 100 people shows that 25 are left-handed. Form a 95% confidence interval for the true proportion of left-handers.

$$\hat{p} \pm z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\frac{25}{100} \pm 1.96\sqrt{\frac{.25(.75)}{100}}$$

$$0.1651 < P < 0.3349$$

n = 100 (sample size)
$\hat{p} = 25/100$ (sample proportion)
1 – α = 0.95 (confidence level), then $z_{1-\frac{\alpha}{2}} = z_{0.975}$ = 1.96 (see Standard Normal Distribution Table)

$CI_{95\%}$ (p) = (0.1651, 0.3349)

$CI_{95\%}$ (p) = (16.51%, 33.49%)

Newbold et al (2013)

**Interpretation:**

- We are 95% confident that the true proportion of left-handers in the population is between 16.51% and 33.49%.
- Although the interval from 0.1651 to 0.3349 may or may not contain the true proportion, 95% of intervals formed from samples of size 100 in this manner will contain the true proportion.

# EXERCISE 7.40

7.40 Suppose that the local authorities in a heavily populated residential area of downtown Hong Kong were considering building a new municipal swimming pool and leisure center. Because such a development would cost a great deal of money, it first of all needed to be established whether the residents of this area thought that the swimming pool and leisure center would be a worthwhile use of public funds. If 243 out of a random sample of 360 residents in the local area thought that the pool and leisure center should be built, determine with 95% confidence the proportion of all the local residents in the area who would support the proposal.

Newbold et al (2013)

# EXERCISE 7.40: SOLUTION

Answer:

$$CI_{(1-\alpha)}(p) = \left(\hat{p} - z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}(\mathbf{1}-\hat{p})}{\boldsymbol{n}}}, \hat{p} + z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}(\mathbf{1}-\hat{p})}{\boldsymbol{n}}}\right)$$

**Sample information:**

- Sample size $n = 360$
- Number who support the proposal $x = 243$

**1. Sample proportion**

$$\hat{p} = \frac{x}{n} = \frac{243}{360} = 0.675$$

**2. Standard error**

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.675)(0.325)}{360}} = \sqrt{0.000609375} \approx 0.0247$$

**3. 95% Confidence Interval**

Use $z = 1.96$ for 95% confidence:

$$\hat{p} \pm 1.96 \cdot SE = 0.675 \pm 1.96(0.0247)$$

$$= 0.675 \pm 0.0484$$

$$\text{Lower limit} \approx 0.6266 \quad \text{and} \quad \text{Upper limit} \approx 0.7234$$

n = 300 (sample size)
$\hat{p} = 243/360$ (sample proportion)
1 – α = 0.95 (confidence level),
then $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$ (see Standard Normal Distribution Table)

**Final Answer (95% CI):**

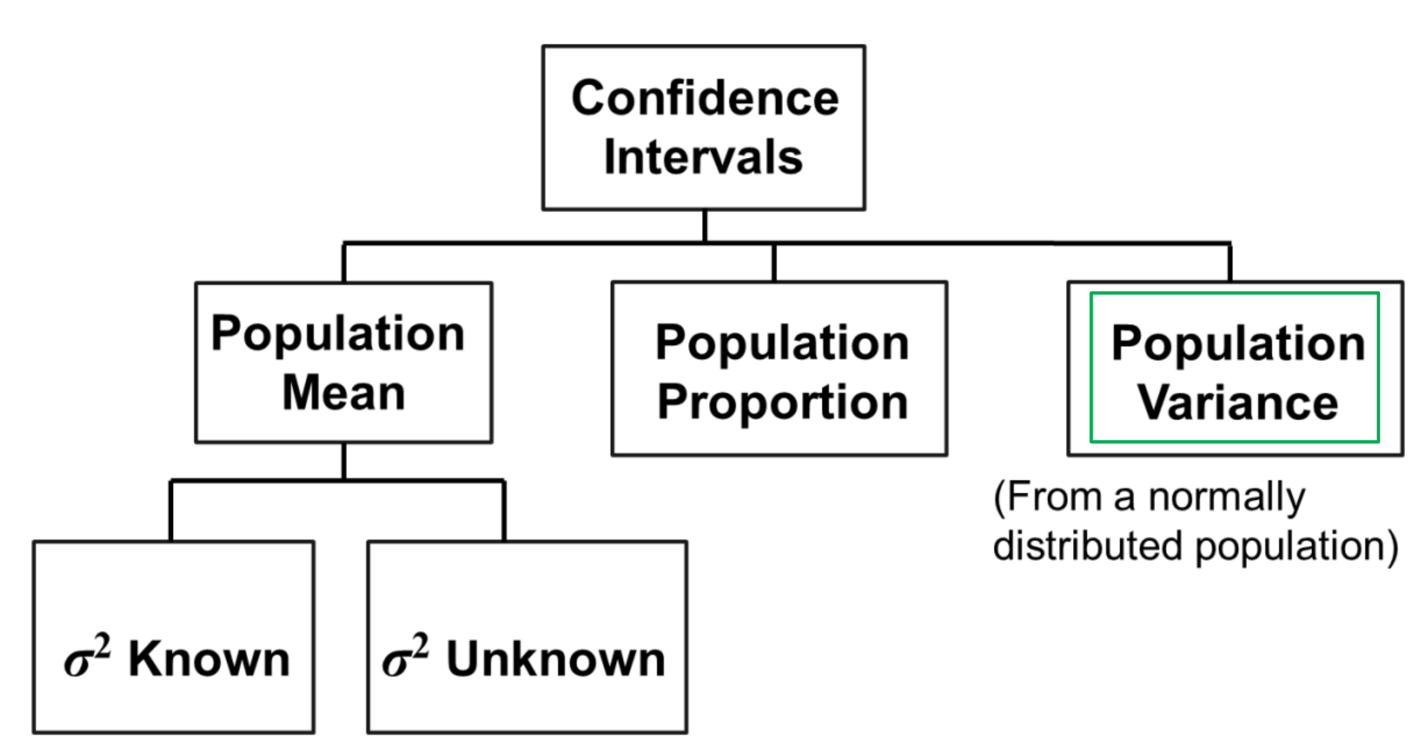
$$\boxed{0.63 \leq p \leq 0.72} \quad \text{(approximately)}$$

$CI_{95\%}(p) = (0.63, 0.72)$

So, we are 95% confident that **between 63% and 72%** of all local residents would support building the swimming pool and leisure center.

# LECTURE 6: CONFIDENCE INTERVAL FOR THE VARIANCE

# ESTIMATION FOR THE POPULATION VARIANCE




Newbold et al (2013)

# CONFIDENCE INTERVALS FOR THE POPULATION VARIANCE

- Goal: Form a confidence interval for the population variance, $\sigma^2$
  - The confidence interval is based on the sample variance, $s^2$
  - Assumed: the population is normally distributed

Newbold et al (2013)

# CONFIDENCE INTERVALS FOR THE POPULATION VARIANCE

The random variable

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

$$Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

follows a chi-square distribution with $(n - 1)$ degrees of freedom

Where the chi-square value $\chi^2_{n-1,1-\alpha}$ is the number for which

$$P\left(\chi^2_{n-1} > \chi^2_{n-1,1-\alpha}\right) = \alpha$$

Newbold et al (2013)

# CONFIDENCE INTERVALS FOR THE POPULATION VARIANCE

The $100(1-\alpha)\%$ confidence interval for the population variance is given by

$$\text{LCL} = \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2,\, n-1)}}$$

$$\text{UCL} = \frac{(n-1)s^2}{\chi^2_{(\alpha/2,\, n-1)}}$$

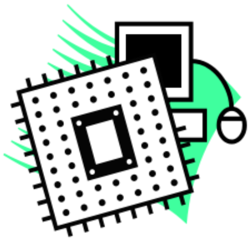
Newbold et al (2013)

$$CI_{(1-\alpha)}(\sigma^2) = \left( \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} \right)$$

# CONFIDENCE INTERVALS FOR THE POPULATION VARIANCE: EXAMPLE

You are testing the speed of a batch of computer processors. You collect the following data (in Mhz):

| | |
|---|---|
| **Sample size** | **17** |
| **Sample mean** | **3004** |
| **Sample std dev** | **74** |



Assume the population is normal.

Determine the 95% confidence interval for $\sigma_x^2$

Newbold et al (2013)

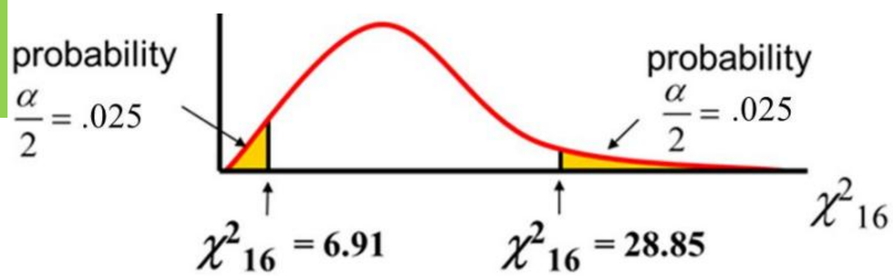
# FINDING THE CHI-SQUARE VALUES

- $n$ = 17 so the chi-square distribution has ($n$ − 1) = 16 degrees of freedom
- $\alpha = 0.05,$ so use the the chi-square values with area 0.025 in each tail:

$$\chi^2_{(1\text{-}\alpha/2,\, n\text{-}1)} = \chi^2_{(0.975,\, 16)} = 28.85$$

$$\chi^2_{(\alpha/2,\, n\text{-}1)} = \chi^2_{(0.025,\, 16)} = 6.91$$

n = 17 (sample size)
s = 74 (sample standard deviation)
α = 0.05 (significance level)
1 – α = 0.95 (confidence level)
$\chi^2_{(0.025,\, 16)}$ = 6.91
$\chi^2_{(0.975,\, 16)}$ = 28.85 (see Chi-Square Distribution Table)




Newbold et al (2013)

# CALCULATING THE CONFIDENCE LIMITS

- The 95% confidence interval is

$$CI_{(1-\alpha)}(\sigma^2) = \left( \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}, \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2},n-1}} \right)$$

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

$$\frac{(17-1)(74)^2}{28.85} < \sigma^2 < \frac{(17-1)(74)^2}{6.91}$$

$$3037 < \sigma^2 < 12680$$

$CI_{95\%}\ (\sigma^2) = (3037, 12680)$

Converting to standard deviation, we are 95% confident that the population standard deviation of CPU speed is between 55.1 and 112.6 Mhz

Newbold et al (2013)

$CI_{95\%}\ (\sigma) = (55.1, 112.6)$

# EXERCISE 7.49

7.49 A manufacturer is concerned about the variability of the levels of impurity contained in consignments of raw material from a supplier. A random sample of 15 consignments showed a standard deviation of 2.36 in the concentration of impurity levels. Assume normality.

a. Find a 95% confidence interval for the population variance.
b. Would a 99% confidence interval for this variance be wider or narrower than that found in part a?

Newbold et al (2013)

# THANKS!

**Questions?**