MATH 1700

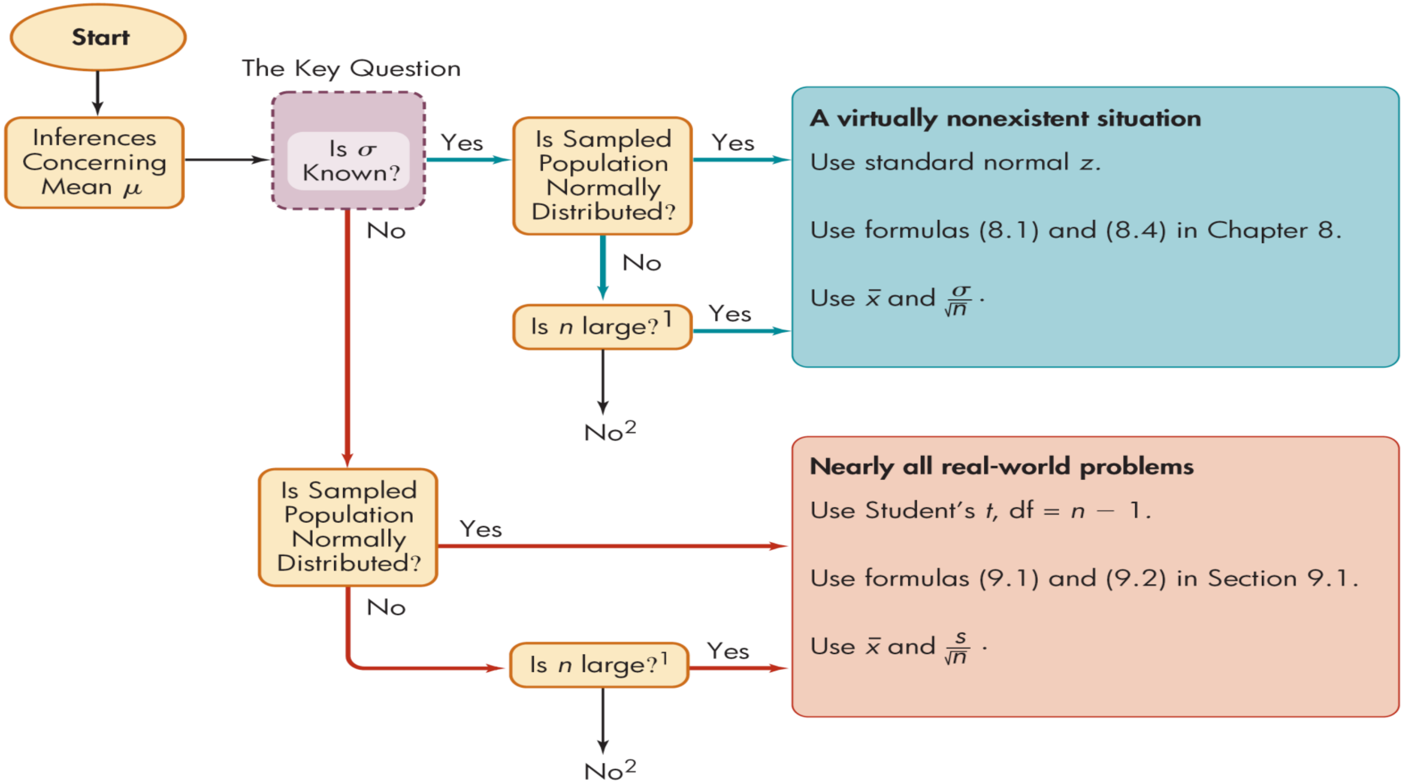
Instructor: Mehdi Maadooliat

Chapter 9A
Inferences Involving One Population

Department of Mathematical and Statistical Sciences

INFERENCE ABOUT THE POPULATION MEAN, μ.

Figure 9.1 presents a diagrammatic organization for the inferences about the population mean.



Do I Use the *z*-Statistic or the *t*-Statistic?

Figure 9.1

INFERENCE ABOUT μ WHEN σ IS UNKNOWN

- **Hypothesis Testing**
- $H_0: \mu = \mu_0$ **vs** $H_a: \mu > \mu_0$

  or $\mu < \mu_0$

  or $\mu \neq \mu_0$

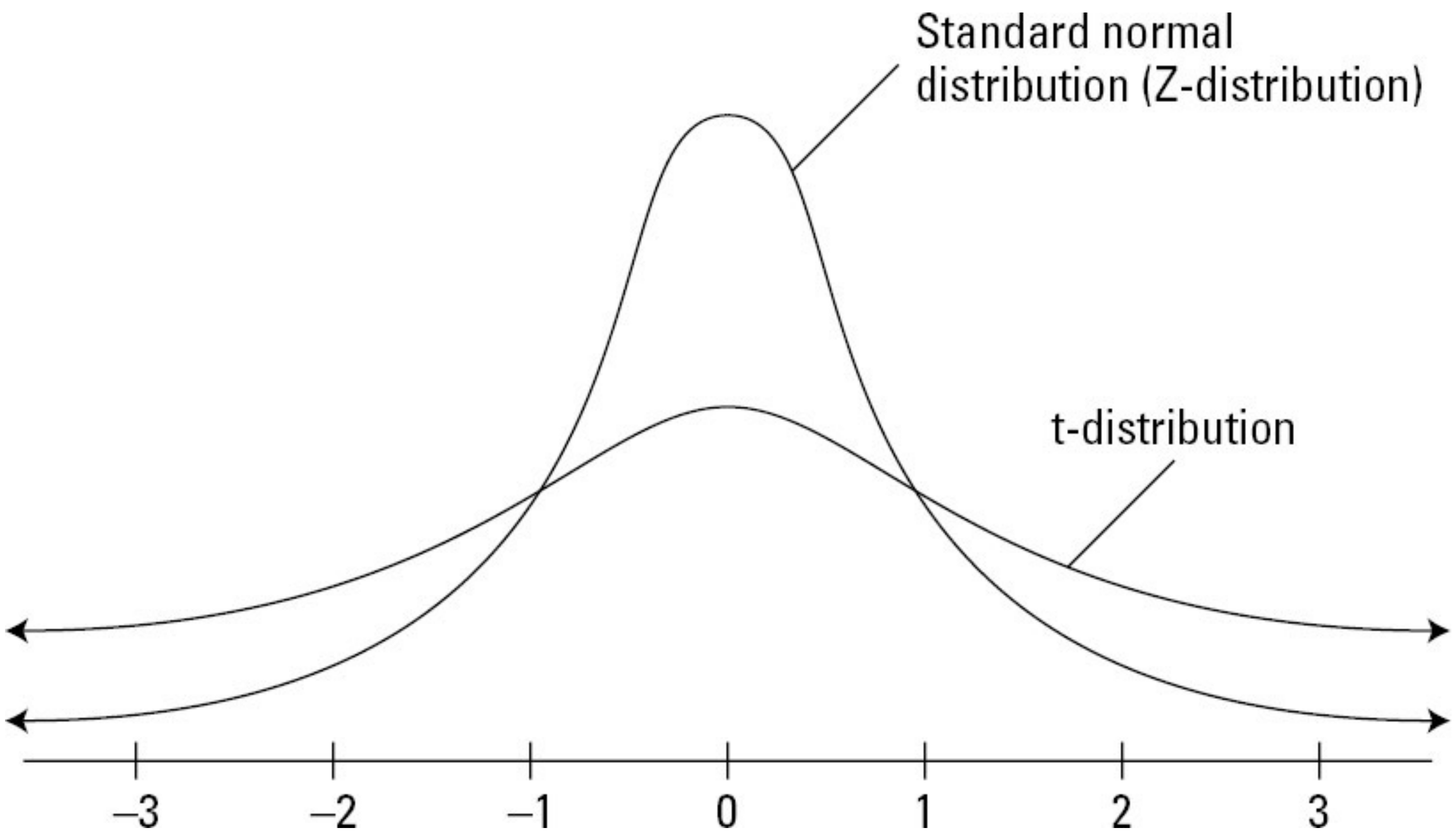
- **Sample:** $\{y_1, y_2, \dots, y_n\}$
- $\bar{y}$ – sample mean, s – sample st.dev.

- **T.S.** $\hat{z} = \frac{\bar{y} - \mu_0}{\frac{s}{\sqrt{n}}}$ (σ replaced by sample st.dev.)

- **The distribution of $\hat{z}$ is no longer N(0, 1).**

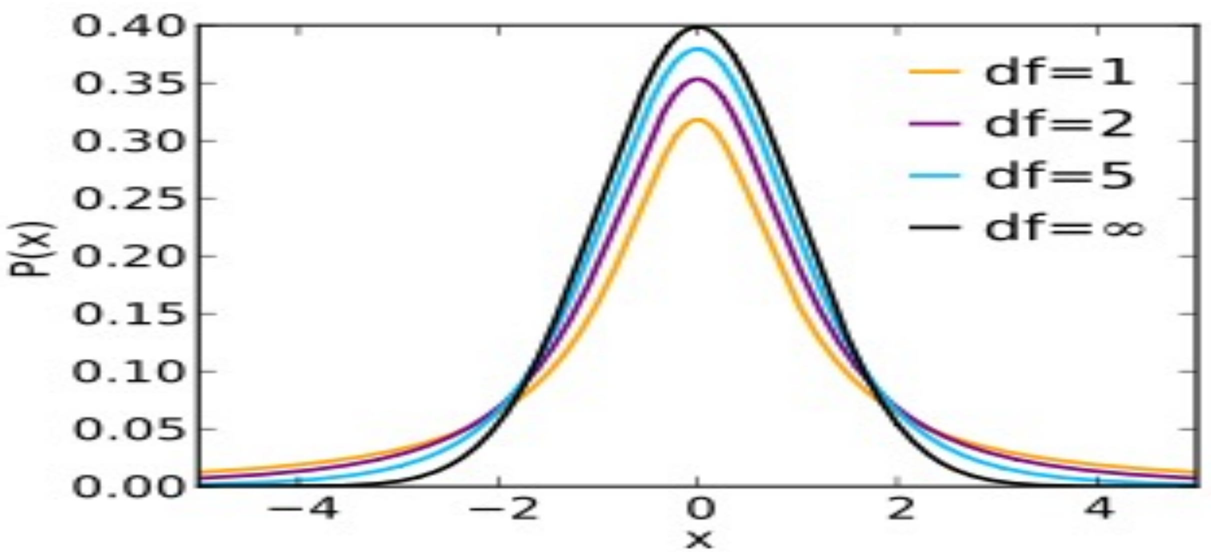
STUDENT T-DISTRIBUTION

- $$t^* = \frac{\bar{y} - \mu}{\frac{s}{\sqrt{n}}}$$
- **The distribution of t is called student-t distribution.**



REMARKS ON T-DISTRIBUTION

1. t-dist. has similar shape as $N(0,1)$, but is flatter than $N(0,1)$.
2. The t-distribution is symmetric around 0 as $N(0,1)$ is.
3. It has a range from $-\infty$ to ∞ as the range of $N(0,1)$.
4. Unlike $N(0,1)$, the t-distribution depends on the **degrees of freedom** df.
5. As the df increases, t-distribution approaches to $N(0,1)$.



- [Applet: t-distribution vs $N(0,1)$]

INFERENCE ABOUT μ WHEN σ IS UNKNOWN

- **Hypothesis Testing**
- $H_0: \mu = \mu_0$ **vs** $H_a: \mu > \mu_0$

  or $\mu < \mu_0$

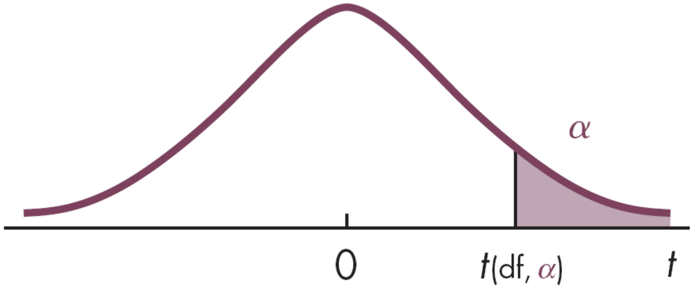
  or $\mu \neq \mu_0$

- **Decision Rule** $\mathrm{df} = n - 1$
  - $H_a: \mu > \mu_0$: **Reject** H_0 **in favor of** H_a **if** $t^* > t(\alpha)$
  - $H_a: \mu < \mu_0$: **Reject** H_0 **in favor of** H_a **if** $t^* < -t(\alpha)$
  - $H_a: \mu \neq \mu_0$: **Reject** H_0 **in favor of** H_a **if** $|t^*| > t(\alpha/2)$

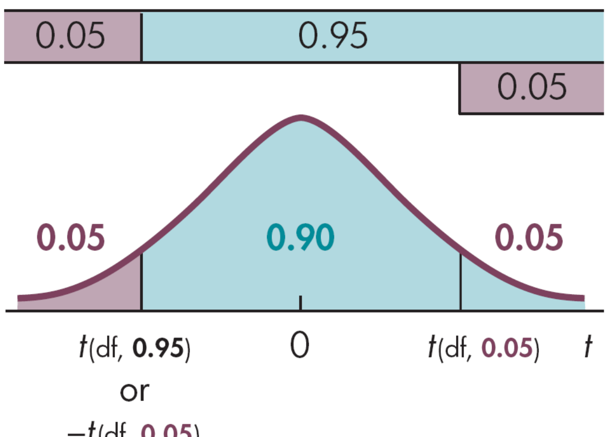
- **Here** $t(A)$ **is a notation for value of** t **so that the area to the right is** A**.**
  - **t-table ("D2L > Useful Links > Z, T and Chi^2 Tables")**
    - $P(t^* \geq t(\alpha))$**, where** t **is a t-distribution with** $\mathrm{df} = n - 1$**.**
  - **t calculator (df)**

USING THE *T*-DISTRIBUTION TABLE

- A notation much like that used with z will be used to identify a critical value. $t(\text{df}, \alpha)$ is the symbol for the value of t with df degrees of freedom and an area of α in the right-hand tail.

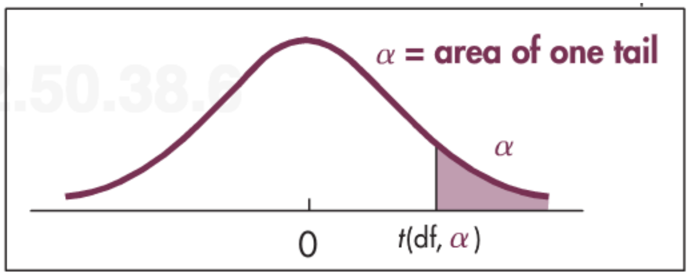


- For the values of t on the left side of the mean, we can use one of two notations.
  - $t(\text{df}, 1-\alpha)$
  - $-t(\text{df}, \alpha)$



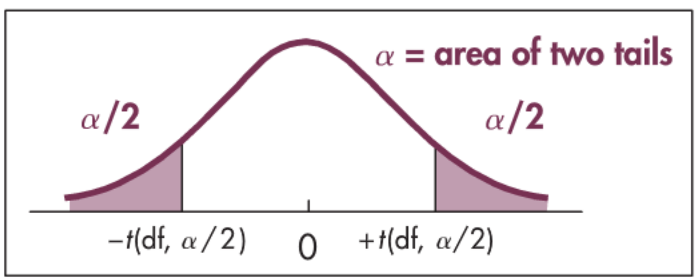
HOW TO USE T-TABLE?

- **Examples:**



  - $\alpha = 0.10, \mathrm{df} = 9$

    $t(\mathrm{df}, \alpha) = t(9, 0.10) = \mathbf{1.38}$



  - $\alpha = 0.02, \mathrm{df} = 14$

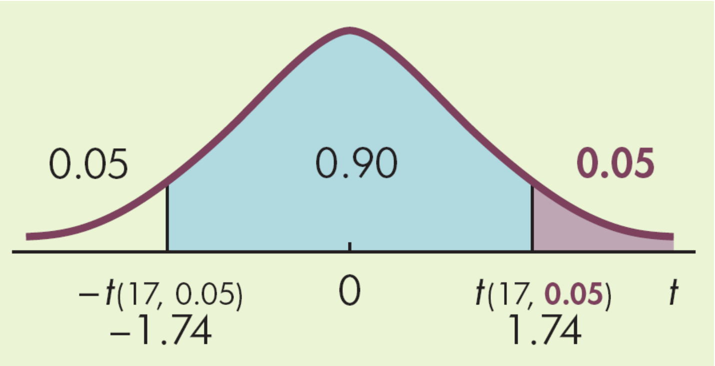
    $t(\mathrm{df}, \alpha/2) = t(14, 0.01) = \mathbf{2.62}$

Area in One Tail	0.25	0.10	0.05	0.025	0.01	0.005
Area in Two Tails						
df	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
3	0.765	1.64	2.35	3.18	4.54	5.84
4	0.741	1.53	2.13	2.78	3.75	4.60
5	0.727	1.48	2.02	2.57	3.36	4.03
6	0.718	1.44	1.94	2.45	3.14	3.71
7	0.711	1.41	1.89	2.36	3.00	3.50
8	0.706	1.40	1.86	2.31	2.90	3.36
9	0.703	1.38	1.83	2.26	2.82	3.25
10	0.700	1.37	1.81	2.23	2.76	3.17
11	0.697	1.36	1.80	2.20	2.72	3.11
12	0.695	1.36	1.78	2.18	2.68	3.05
13	0.694	1.35	1.77	2.16	2.65	3.01
14	0.692	1.35	1.76	2.14	2.62	2.98
15	0.691	1.34	1.75	2.13	2.60	2.95
16	0.690	1.34	1.75	2.12	2.58	2.92
17	0.689	1.33	1.74	2.11	2.57	2.90
18	0.688	1.33	1.73	2.10	2.55	2.88
19	0.688	1.33	1.73	2.09	2.54	2.86
20	0.687	1.33	1.72	2.09	2.53	2.85
21	0.686	1.32	1.72	2.08	2.52	2.83
22	0.686	1.32	1.72	2.07	2.51	2.82
23	0.685	1.32	1.71	2.07	2.50	2.81
24	0.685	1.32	1.71	2.06	2.49	2.80
25	0.684	1.32	1.71	2.06	2.49	2.79
26	0.684	1.31	1.70	2.05	2.47	2.77
27	0.684	1.31	1.70	2.05	2.47	2.77
28	0.683	1.31	1.70	2.05	2.47	2.76
29	0.683	1.31	1.70	2.05	2.46	2.76
30	0.683	1.31	1.70	2.04	2.46	2.75
35	0.682	1.31	1.69	2.03	2.44	2.72
40	0.681	1.30	1.68	2.02	2.42	2.70
50	0.679	1.30	1.68	2.01	2.40	2.68
70	0.678	1.29	1.67	1.99	2.38	2.65
100	0.677	1.29	1.66	1.98	2.36	2.63
df > 100	0.675	1.28	1.65	1.96	2.33	2.58

EXAMPLE 3 – *T-VALUES THAT BOUND A MIDDLE PERCENTAGE*

- **Find the values of the t-distribution that bound the middle 0.90 of the area under the curve for the distribution with $\mathrm{df} = 17$.**

- **Solution:**
- **The middle 0.90 leaves 0.05 for the area of each tail. The value of t that bounds the right-hand tail is $t\ (17, 0.05)\ =\ \mathbf{1.74}$,**



Area in One Tail

	0.25	0.10	0.05	0.025	0.01	0.005
Area in Two Tails						
df	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
3	0.765	1.64	2.35	3.18	4.54	5.84
4	0.741	1.53	2.13	2.78	3.75	4.60
5	0.727	1.48	2.02	2.57	3.36	4.03
6	0.718	1.44	1.94	2.45	3.14	3.71
7	0.711	1.41	1.89	2.36	3.00	3.50
8	0.706	1.40	1.86	2.31	2.90	3.36
9	0.703	1.38	1.83	2.26	2.82	3.25
10	0.700	1.37	1.81	2.23	2.76	3.17
11	0.697	1.36	1.80	2.20	2.72	3.11
12	0.695	1.36	1.78	2.18	2.68	3.05
13	0.694	1.35	1.77	2.16	2.65	3.01
14	0.692	1.35	1.76	2.14	2.62	2.98
15	0.691	1.34	1.75	2.13	2.60	2.95
16	0.690	1.34	1.75	2.12	2.58	2.92
17	0.689	1.33	1.74	2.11	2.57	2.90
18	0.688	1.33	1.73	2.10	2.55	2.88
19	0.688	1.33	1.73	2.09	2.54	2.86
20	0.687	1.33	1.72	2.09	2.53	2.85

EXAMPLE

- Consider a population of hypertension group whose average systolic blood pressure (SBP) is 150. You want to determine whether a new treatment is effective in reducing SBP. A clinical trial was conducted on 25 patients of this population. After 6 months of treatment, SBP was recorded on each subject.

- $\bar{y} = 147.2, \quad s = 5.5$

- Is there a sufficient evidence at $\alpha = 0.05$ that the new treatment is effective?
- Assume that the distribution of SBP is normal.

EXAMPLE CONT'D

- $H_0: \mu = 150 \;\; vs. \;\; H_a: \mu < 150$
- **TS:** $t^* = \frac{\bar{y} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{147.2 - 150}{\frac{5.5}{\sqrt{25}}} = -2.55$
- **Decision Rule: Reject H_0 in favor of H_a if** $t^* < -t(\text{df}, \alpha)$
- $\text{df} = n - 1 = 25 - 1 = 24, \;\; \alpha = 0.05$
- $t(\text{df}, \alpha) = 1.71$
- **Reject H_0 in favor of H_a if** $t^* < -1.71$.

Area in One Tail	0.25	0.10	0.05	0.025	0.01	0.005
Area in Two Tails						
df	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
19	0.688	1.33	1.73	2.09	2.54	2.86
20	0.687	1.33	1.72	2.09	2.53	2.85
21	0.686	1.32	1.72	2.08	2.52	2.83
22	0.686	1.32	1.72	2.07	2.51	2.82
23	0.685	1.32	1.71	2.07	2.50	2.81
24	0.685	1.32	1.71	2.06	2.49	2.80
25	0.684	1.32	1.71	2.06	2.49	2.79
26	0.684	1.31	1.70	2.05	2.47	2.77
27	0.684	1.31	1.70	2.05	2.47	2.77

- **Conclusion: Is $t^* < -1.71$? Yes, since $t^* = -2.55$.**
  Thus we reject H_0, and we have sufficient evidence to conclude that the new treatment is effective.

P−VALUE APPROACH

- $H_0: \mu = 150 \;\; vs. \;\; H_a: \mu < 150$,   **TS:** $t^* = -2.55$

Area in One Tail	0.25	0.10	0.05	0.025	0.01	0.005
Area in Two Tails — **df**	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
19	0.688	1.33	1.73	2.09	2.54	2.86
20	0.687	1.33	1.72	2.09	2.53	2.85
21	0.686	1.32	1.72	2.08	2.52	2.83
22	0.686	1.32	1.72	2.07	2.51	2.82
23	0.685	1.32	1.71	2.07	2.50	2.81
24	0.685	1.32	1.71	2.06	2.49	2.80
25	0.684	1.32	1.71	2.06	2.49	2.79
26	0.684	1.31	1.70	2.05	2.47	2.77
27	0.684	1.31	1.70	2.05	2.47	2.77

- $\text{p−value} = P(t < -2.55)$
- $0.005 < \text{p−value} < 0.01$
- **Since** $\text{p−value} < \alpha = 0.05$, **we reject** H_0 **in favor of** H_a**. We have sufficient evidence to conclude that the new treatment is effective.**

- P−value **Formula:**
  - $H_a: \mu > \mu_0$, $\text{p−value} = P(t > t^*)$
  - $H_a: \mu < \mu_0$, $\text{p−value} = P(t < t^*)$
  - $H_a: \mu \neq \mu_0$: $\text{p−value} = 2 * P(t > |t^*|)$

ESTIMATION OF μ USING A CONFIDENCE INTERVAL

- $100(1-\alpha)\%$ **Confidence Interval of μ is**

Confidence Interval for Mean

$$\bar{x} - t(\text{df}, \alpha/2)\left(\frac{s}{\sqrt{n}}\right) \text{ to } \bar{x} + t(\text{df}, \alpha/2)\left(\frac{s}{\sqrt{n}}\right), \text{ with df} = n - 1 \qquad \textbf{(9.1)}$$

- **Assumption: Either**
  - $n \geq 30$
  - **the sample is drawn from a normal population.**

EXAMPLE 4 – *CONFIDENCE INTERVAL FOR μ WITH σ UNKNOWN*

A random sample of 20 weights is taken from babies born at Northside Hospital. A mean of $6.87\ lb$ and a standard deviation of $1.76\ lb$ were found for the sample. Estimate, with 95% confidence, the mean weight of all babies born in this hospital. Based on past information, it is assumed that weights of newborns are normally distributed.

Solution:

Step 1 The Set-Up:

Describe the population parameter of interest.

μ, the mean weight of newborns at Northside Hospital

EXAMPLE 4 – *SOLUTION*

Step 2 The Confidence Interval Criteria:

a. Check the assumptions.

Past information indicates that the sampled population is normal.

b. Identify the probability distribution and the formula to be used.

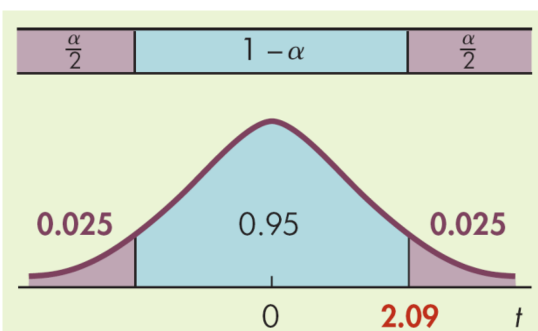
The value of the population standard deviation, σ, is unknown. Student's t-distribution will be used with formula (9.1).

c. State the level of confidence: $1 - \alpha = 0.95$.

EXAMPLE 4 – *SOLUTION*

- **Step 3 The Sample Evidence:**
  - **Collect the sample information:**
  - $n = 20$, $\bar{x} = 6.87$, **and** $s = 1.76$.

- **Step 4 The Confidence Interval:**
  - **a. Determine the confidence coefficients.**
  - ✓ **Since** $1 - \alpha = 0.95$, $\alpha = 0.05$, **and therefore** $\alpha/2 = 0.025$.
  - ✓ **Also, since** $n = 20$, $\text{df} = n - 1 = 19$.
  - ✓ $t(\text{df}, \alpha/2) = t(19, 0.025) = 2.09$



Area in One Tail	0.25	0.10	0.05	0.025	0.01	0.005
Area in Two Tails						
df	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
3	0.765	1.64	2.35	3.18	4.54	5.84
4	0.741	1.53	2.13	2.78	3.75	4.60
16	0.690	1.34	1.75	2.12	2.58	2.92
17	0.689	1.33	1.74	2.11	2.57	2.90
18	0.688	1.33	1.73	2.10	2.55	2.88
19	0.688	1.33	1.73	2.09	2.54	2.86
20	0.687	1.33	1.72	2.09	2.53	2.85

EXAMPLE 4 – *SOLUTION*

- **Step 4 The Confidence Interval:**

b. Find the maximum error of estimate.

✓ $E = t(\text{df}, \alpha)\frac{s}{\sqrt{n}}$:

✓ $E = t(19{,}0.025)\frac{1.76}{\sqrt{20}}$:

✓ $E = (2.09)(0.394)$:

✓ $E = 0.82$

c. Find the lower and upper confidence limits.

✓ $\bar{x} - E$ **to** $\bar{x} + E$: $6.87 - 0.82$ **to** $6.87 + 0.82$

✓ $\mathbf{6.05}$ **to** $\mathbf{7.69}$

- **Step 5 The Results: State the confidence interval.**
  - ✓ **6.05 to 7.69 is the 95% confidence interval for μ.**
  - ✓ **That is, with 95% confidence we estimate the mean weight of babies born at Northside Hospital to be between 6.05 and 7.69 lb.**

SUMMARY (SLIGHT CHANGES IN NOTATIONS)

MARQUETTE UNIVERSITY
Be The Difference.

Case 1: One Variable

	Numerical Variable, σ known	**Numerical Variable,** σ unknown
Parameter of Interest:	Mean, μ	Mean, μ
Confidence Interval Formula:	$\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$	$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ with df = $n - 1$
Name of Hypothesis Test, H_0	One Sample Z Test, $H_0 : \mu = \mu_0$	One Sample T Test, $H_0 : \mu = \mu_0$
Test Statistic Formula:	$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ with df = $n - 1$
***p*-value:**	$H_a : \mu \neq \mu_0$, p-value = $2P(Z \geq \lvert z \rvert)$ $H_a : \mu > \mu_0$, p-value = $P(Z \geq z)$ $H_a : \mu < \mu_0$, p-value = $P(Z \leq z)$	$H_a : \mu \neq \mu_0$, p-value = $2P(T \geq \lvert t \rvert)$ $H_a : \mu > \mu_0$, p-value = $P(T \geq t)$ $H_a : \mu < \mu_0$, p-value = $P(T \leq t)$

QUESTIONS?

- ANY QUESTION?