Questions in the last tutorial

Hypothesis Test SOP

Pick some questiones by students

Pick question by myself : 1 → 3 (+ P-value + Relationship between C.I. & Hypothesis Testing) → 5 ($n<30$) → 6 → 7 → 11 → 12 → 13.

Calc

μ. σ, p.

Statistics
- descriptive statistics
- * Inferential Statistics

Bride : probability → distribution of a r.v.

→ "Sampling distribution" (distribution of a statistic)

statistic (sample) → prameter (population)

point estimation
- $\bar{X} \to \mu$
- $S \to \sigma$
- $\hat{p} \to p$

$\bar{X}$

C.I.
- $\bar{X} \to [\ ,\ ]$
- $\hat{p} \to [\ ,\ ]$

Hypothesis Testing
- H_0. H_1
- test statistic.

Hypothesis Testing SOP:

S0: Get useful information from the question statements /

Natural language → Math language.

S1. H_0 . H_a (what we want to get)

<"=" can only happen in H_0.>

S2: α (type I error) $\alpha = P(\text{reject } H_0 \mid H_0 \text{ is true})$

S3: test statistic

S4: rejection region / critical value /

S5: Conclusion.

In S3.

μ:

- $n \geq 30$
  - σ known → $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$
  - σ unknown → $Z = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$
- $n < 30$ & Data come from Normal (approximate normality)
  - σ known → $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$
  - σ unknown → $t = \dfrac{\bar{X} - \mu_0}{s / \sqrt{n}}$

Differences of Means. (two independent samples)

σ_1, σ_2 Known → $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $(n_1, n_2 > 30)$

σ_1, σ_2 unknown but assume $\sigma_1 = \sigma_2$

→ $t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t_{n_1+n_2-2}$ $S_p = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$

"μ" / σ is unknown / $n \geq 30$ / left-sided : 1, 2, 4.

"μ" / σ is unknown / $n \geq 30$ / two-sided : 3

"μ" / σ is known / $n \geq 30$ / two-sided : 9 → 10 ($n<30$)

"μ" / σ is unknown / $n < 30$ / right-sided : 5

"μ" / σ is unknown / $n < 30$ / two-sided : 8

"μ" / σ is known / $n < 30$ / two-sided : 11 → 12 (σ unknown)

C.I. & Hypothesis Testing : 6

Multi : 7.

$\mu_1 - \mu_2$: 13, 14.

Relationship Between C.I. & Hypothesis Testing.

(Use Q3 as example)

95% C.I. for μ:

$$\left[\bar{X} - Z \cdot \frac{S}{\sqrt{n}} , \bar{X} + Z \cdot \frac{S}{\sqrt{n}} \right]$$

$$\left[2.98 - 1.96 \times \frac{0.82}{\sqrt{75}} , 2.98 + 1.96 \times \frac{0.82}{\sqrt{75}} \right]$$

$$= [-2.79, 3.17]$$

$$\mu_0 = 3.14 \in [-2.79, 3.17]$$

< Same Thing, different way >.

Hypothesis Testing is more flexible → P-value

1. It is hoped that a newly developed pain reliever will more quickly produce perceptible reduction in pain to patients after minor surgeries than a standard pain reliever. The standard pain reliever is known to bring relief in an average of 3.5 minutes with standard deviation 2.1 minutes. To test whether the new pain reliever works more quickly than the standard one, 50 patients with minor surgeries were given the new pain reliever and their times to relief were recorded. The experiment yielded a sample mean of 3.1 minutes and a sample standard deviation of 1.5 minutes. Is there sufficient evidence to indicate, at a 5% level of significance, that the newly developed pain reliever does deliver perceptible relief more quickly?

Standard : $\mu_0 = 3.5$ (min) $\quad \sigma_0 = 2.1$ (min)

Sample : $n = 50 \quad \bar{X} = 3.1 \quad S = 1.5$ → But we don't know its population information.

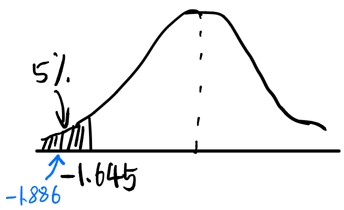
Want to proof : $\mu_1 < \mu_0$

S1: $H_0 : \mu_1 = 3.5 \qquad H_1 : \mu_1 < 3.5$ (left-sided tail test)

S2: $\alpha = 5\%$

S3: $Z = \dfrac{\bar{X} - 3.5}{S / \sqrt{n}} = \dfrac{3.1 - 3.5}{1.5 / \sqrt{50}} = -1.886$

S4: rejection region :



S5: Conclusion : reject H_0. So, newly developed pain reliever deliver perceptible relief more quickly.

2. In the past the average length of an outgoing telephone call from a business office has been 143 seconds. A manager wishes to check whether that average has decreased after the introduction of policy changes. A sample of 100 telephone calls produced a mean of 133 seconds, with a standard deviation of 35 seconds. Perform the relevant test at the 1% level of significance.

$n = 100 \quad \bar{X} = 133 \quad S = 35 \quad \mu_0 = 143 \quad \alpha = 1\%$

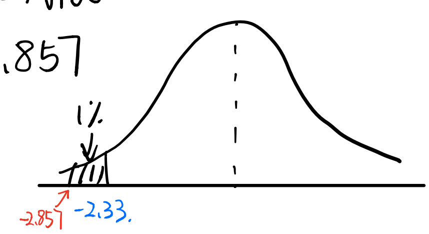
S1. $H_0 : \mu = 143. \qquad H_a : \mu < 143$ (left-sided tail test)

S2 : $\alpha = 1\%.$

S3 : $Z = \dfrac{\bar{X} - 143}{S / \sqrt{n}} = \dfrac{133 - 143}{35 / \sqrt{100}}$ ($n \geq 30$ & σ is unknown)

$= -2.857$

S4 : rejection region :



$Z < -2.33.$

S5 : Conclusion : reject H_0. → decreased.

3. The average household size in a certain region several years ago was 3.14 persons. A sociologist wishes to test, at the 5% level of significance, whether it is different now. Perform the test using the information collected by the sociologist: in a random sample of 75 households, the average size was 2.98 persons, with sample standard deviation of 0.82 person.

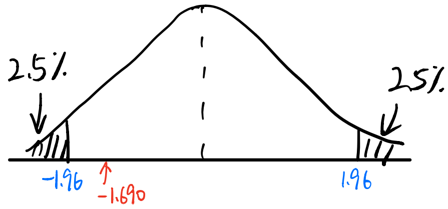
$n = 75 \quad \bar{X} = 2.98 \quad S = 0.82 \quad \mu_0 = 3.14 \quad \alpha = 5\%$

S1: $H_0: \mu = 3.14 \qquad H_a: \mu \neq 3.14$ (two-sided tail test)

S2: $\alpha = 5\%$

S3: $$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{2.98 - 3.14}{0.82/\sqrt{75}} = -1.690$$

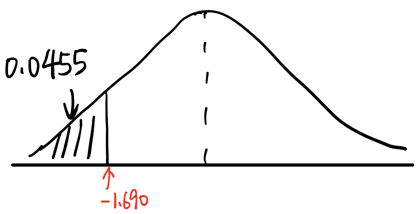
S4: rejection region:



S5: Fail to reject H_0.

Seems not different now.

But, how to calculate P-value?



two tails

$$\text{P-value} = 0.0455 \times 2 = 0.0910.$$

4. The recommended daily calorie intake for teenage girls is 2,200 calories/day. A nutritionist at a state university believes the average daily caloric intake of girls in that state to be lower. Test that hypothesis, at the 5% level of significance, against the null hypothesis that the population average is 2200 calories/day using the sample data of size 36, mean 2150, standard deviation 203.

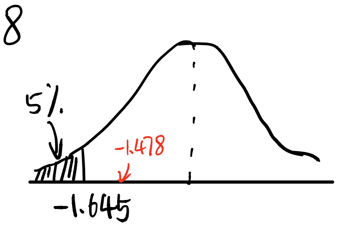
$n = 36 \quad \bar{X} = 2150 \quad S = 203 \quad \alpha = 5\% \quad \mu_0 = 2200$

S1: $H_0: \mu = 2200 \qquad H_a: \mu < 2200$ (left-sided tail test)

S2: $\alpha = 5\%$

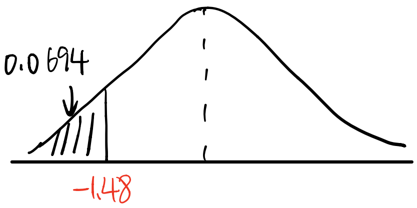
S3: $$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{2150 - 2200}{203/\sqrt{36}} = -1.478$$

S4:



S5: Fail to reject Ho. So, ... lower.

But, how to calculate P-value?



one tail.

P-value = 0.0694

5. Suppose that a delivery company claims that they deliver their packages in 2 days on average, and you suspect it's longer than that. To test this claim, you take a random sample of 10 packages and record their delivery times. You find the sample mean is $\bar{x} = 2.3$ days, and the sample standard deviation is 0.35. Do you have sufficient data to show that the company delivers their packages in more than 2 days, considering significance level 0.05? And what assumptions do you need?

$\mu_0 = 2$ (days) $\quad n=10 \quad \bar{X} = 2.3 \quad S = 0.35$

$n < 30$ If we want to use t-test, we must assume that data come from a normal distribution (or close enough)

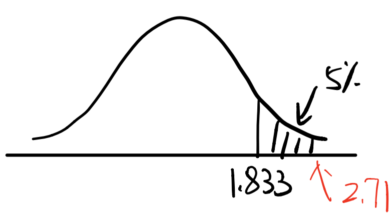
Assumption: Data come from Normal Distribution (or close enough)

S1: $H_0: \mu = 2 \qquad H_1: \mu > 2$ (right-sided tail test)

S2: $\alpha = 5\%$

S3: $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} = \dfrac{2.3-2}{0.35/\sqrt{10}} \approx 2.71 \qquad (\nu = n-1 = 9)$

S4:



S5: reject Ho.

If From P-value aspect:

$t \approx 2.71$ check Table A.4

we can find α is between 0.015 and 0.01.

But it's enough to say that we can reject H_0 and $\mu > 22$.

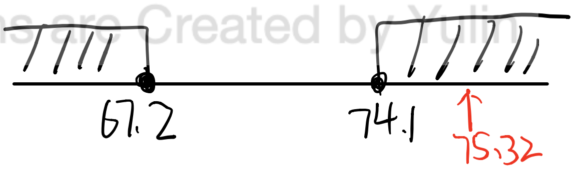
6. An old research has determined the mean age of a certain population to be 70.65 with a 95 % confidence interval of (67.2, 74.1). A new sample yields $\bar{x} = 75.32$.
   (a) What is your conclusion concerning the $H_0 : \mu = 70.65$ with $\alpha = 0.05$?
   (b) Is this a one-sided or a two-sided test?

(a) $\bar{X} = 75.32$.

$H_0 : \mu = 70.65$

$\alpha = 0.05$

rejection region :



$75.32 > 70.65$ reject H_0.

(b) two-sided test.

7. The hypothesis $H_0 : \mu = 15$ is tested against $H_1 : \mu \neq 15$ with $\alpha = 0.05$. A random sample has following properties: $n = 20$, $\bar{x} = 17.5$, $s = 5.9$.
   (a) What assumptions should be made and which distribution should be used? Why?
   (b) What is the range of the P-value that you can find with table A4?
   (c) What is your conclusion regarding the null hypothesis?

(a) $n = 20 < 30$

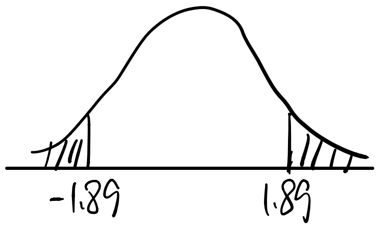
Assumption : Data come from Normal Distribution (or close enough)

We need to use t-distribution, because sample size is smaller than 30, we can't use C.L.T. to approximate t by Z.

$$T = \frac{\bar{X} - 15}{S/\sqrt{n}} = \frac{17.5 - 15}{5.9/\sqrt{20}} \sim t_{19}$$ (t-distribution with $v = 19$ degrees of freedom)

(b) $$t = \frac{\bar{X} - 15}{S/\sqrt{n}} = \frac{17.5 - 15}{5.9/\sqrt{20}} \approx 1.89$$

ν	0.05 $\longleftrightarrow$	0.025
19	1.729	2.093



(two-sided tail test)

$$2 \times 0.025 < \text{P-value} < 2 \times 0.05$$

$$\Rightarrow \text{P-value} \in (0.05, 0.1)$$

(c) Because P-value > 0.05 (α)

So we fail to reject H_0.

8. The average brown trout's I.Q. is 4. Suppose you don't believe this claim. You catch 12 brown trout. A fish psychologist determines the I.Q.s as follows: 5; 4; 7; 3; 6; 4; 5; 3; 6; 3; 8; 5. Conduct a hypothesis test of your belief with $\alpha = 0.05$.

$n = 12$ (< 30) $\qquad \bar{X} = \frac{59}{12}$ $\qquad S = \sqrt{\frac{1}{11}\sum_{i=1}^{12}(X_i - \bar{X})^2} = 1.621$

Assumption: Data come from Normal Distribution (or close enough) / The underlying distribution is not too extreme.

S1. $H_0: \mu = 4$ $\qquad H_1: \mu \neq 4$ (two-sided tail test)

S2: $\alpha = 0.05$

S3: $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} = \dfrac{\frac{59}{12} - 4}{1.621/\sqrt{12}} \approx 1.96$ $\qquad (\nu = 11)$

S4:

ν	0.05	0.025
11	1.796	2.201

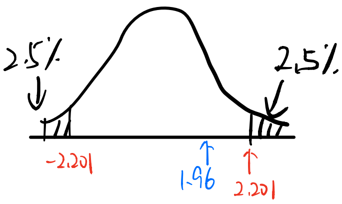
About how to use table,

$$2 \times 0.025 < \text{P-value} < 2 \times 0.05$$

$$\Rightarrow \text{P-value} \in (0.05, 0.1)$$

Becase P-value > 0.05, so we fail to reject H_0.

OR: The acceptance region corresponding to a two-tailed test with $\alpha = 0.05$ is $(-2.201, 2.201)$



Because $t = 1.96 \in (-2.201, 2.201)$, so we fail to reject H_0.

9. A child psychologist says that the average time that working mothers spend talking to their children is 11 minutes per day. Suppose a random sample of 100 working mothers spend on average 11.5 minutes per day talking to their children. The prior research suggests the population standard deviation is 2.3. Do we have evidence that working mothers spend more than 11 minutes per day by talking to their children? Use $\alpha = 0.05$.

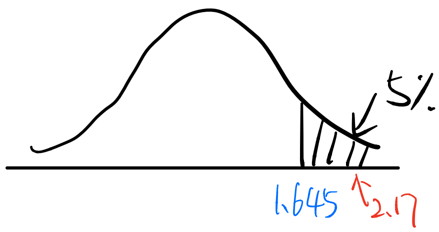
$n = 100 \quad \bar{X} = 11.5 \quad \sigma = 2.3 \quad \alpha = 0.05$

S1: $H_0: \mu = 11 \quad H_1: \mu > 11$ (right-sided tail test)

S2: $\alpha = 0.05$

S3: $z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{11.5 - 11}{2.3 / \sqrt{100}} \approx 2.17$

S4.



$2.17 > 1.645$

S5: reject H_0.

Table A3

z ... 0.07

2.1 ... 0.9850

OR use P-value.

P-value $= 1 - 0.9850 = 0.015 < 0.05$

reject H_0.

10. Consider the previous exercise, but now with a small change: The random sample does not include 100 working mothers, but only 10 working mothers. Will your conclusion change?

$n = 10 < 30$.

Assumption: Data comes from Normal Distribution (or not too extreme)

Use t-test. $T = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{11.5 - 11}{2.3 / \sqrt{10}} \approx 0.6875 \quad \nu = 9$

Check Table A4:

ν \ α	0.30	0.20
9	0.543	0.883

One-Side tail test: P-value $\in (0.20, 0.30)$

P-value $> 0.05 \Rightarrow$ Fail to reject H_0

11. The Achenbach Child Behaviour Checklist is designed so that scores from normal children are Normal with mean $\mu = 50$ and standard deviation $\sigma = 10$. We are given a sample of $n = 5$ children under stress with an average score of $\bar{x} = 56.0$. Is there evidence that children under stress show an abnormal behaviour? You may assume that σ still is known to be equal to 10. Use level of significance 0.05.

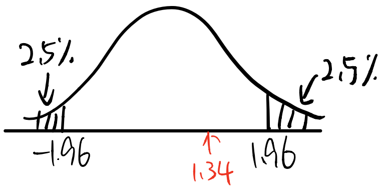
$n = 5\ (< 30)$ $\quad \bar{X} = 56$ $\quad \mu_0 = 50$ $\quad \sigma = 10$ $\quad \alpha = 0.05.$

S1: $H_0: \mu = 50$ $\quad H_1: \mu \neq 50.$ (two-sided tail test)

S2: $\alpha = 0.05$

S3: $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{56 - 50}{10/\sqrt{5}} \approx 1.34$

S4:



$1.34 \in (-1.96, 1.96)$

S5: Fail to reject H_0.

Not show an abnormal behavior.

12. Solve the previous exercise, assuming that we do not know σ, but are given a sample of $n = 5$ children under stress with an average score of $\bar{x} = 56.0$ and standard deviation $s = 8.5$.

$n = 5\ (< 30)$ $\quad S = 8.5$ $\quad \bar{X} = 56.$ $\quad \mu_0 = 50$

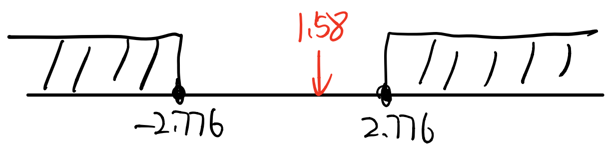
S1: $H_0: \mu = 50$ $\quad H_1: \mu \neq 50$

S2: $\alpha = 5\%$

S3: $t = \dfrac{\bar{X} - \mu_0}{S / \sqrt{n}} = \dfrac{56 - 50}{8.5/\sqrt{5}} \approx 1.58$ $\quad (\nu = 4)$

S4: acceptance region for two-sided t_4-distribution with $\alpha = 0.05$ is $(-2.776, 2.776)$

ν	α
	0.025 ← two-sided.
4	2.776



S5: Fail to reject Ho.

13. A manufacturer claims that the average tensile strength of thread A exceeds the average tensile strength of thread B by at least 12 kilograms. To test this claim, 50 pieces of each type of thread were tested under similar conditions. Type A thread had an average tensile strength of 86.7 kilograms with a standard deviation of 6.28 kilograms, while type B thread had an average tensile strength of 77.8 kilograms with a standard deviation of 5.61 kilograms. Test the manufacturer's claim using a 0.05 level of significance, using $H_0 : \mu_A - \mu_B \geq 12$ vs $H_1 : \mu_A - \mu_B < 12$.

$\bar{X}_A = 86.7$ $\quad S_A = 6.28$ $\quad \bar{X}_B = 77.8$ $\quad S_A = 5.61$

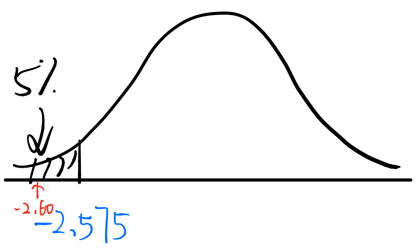
$n_A = 50$ $\quad n_B = 50$ $\quad \alpha = 0.05$

S1: $H_0 : \mu_A - \mu_B = 12$ (or $\geq$) $\quad H_1 : \mu_A - \mu_B < 12$ (left-sided)

S2: $\alpha = 0.05$

S3: $$Z = \frac{(\bar{X}_A - \bar{X}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} = \frac{(86.7 - 77.8) - 12}{\sqrt{\frac{6.28^2 + 5.61^2}{50}}} \approx -2.60$$

S4:



$Z = -2.60 < -2.575$

S5: reject Ho.

14. A study was conducted to see if increasing the substrate concentration has an appreciable effect on the velocity of a chemical reaction. With a substrate concentration of 1.5 moles per liter, the reaction was run 15 times, with an average velocity of 7.5 micromoles per 30 minutes and a standard deviation of 1.5. With a substrate concentration of 2.0 moles per liter, 12 runs were made, yielding an average velocity of 8.8 micromoles per 30 minutes and a sample standard deviation of 1.2. Is there any reason to believe that this increase in substrate concentration causes an increase in the mean velocity of the reaction of more than 0.5 micromole per 30 minutes? Use a 0.01 level of significance and assume the populations to be approximately normally distributed with equal variances.

$n_1 = 15$ $\quad \bar{X}_1 = 7.5$ $\quad S_1 = 1.5$

$n_2 = 12$ $\quad \bar{X}_2 = 8.8$ $\quad S_2 = 1.2$

$$S_p = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{14 \times 1.5^2 + 11 \times 1.2^2}{15+12-2} \approx 1.89$$

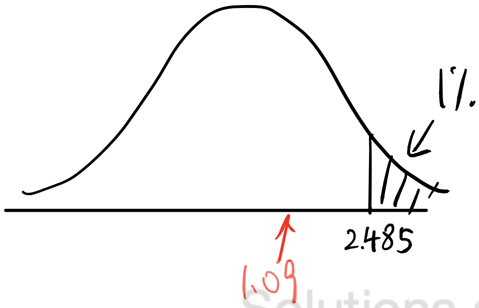
S1: $H_0: \mu_2 - \mu_1 = 0.5$ $\quad H_1: \mu_2 - \mu_1 > 0.5$ (right-sided)

S2: $\alpha = 0.01$

S3: $$t = \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(8.8 - 7.5) - 0.5}{1.89\sqrt{\frac{1}{15} + \frac{1}{12}}} \approx 1.09$$

$$\nu = n_1 + n_2 - 2 = 25$$

S4:



$1.09 < 2.485.$

S5: Fail to reject H_0.

Tutorial 9; 22-05-23

1. It is hoped that a newly developed pain reliever will more quickly produce perceptible reduction in pain to patients after minor surgeries than a standard pain reliever. The standard pain reliever is known to bring relief in an average of 3.5 minutes with standard deviation 2.1 minutes. To test whether the new pain reliever works more quickly than the standard one, 50 patients with minor surgeries were given the new pain reliever and their times to relief were recorded. The experiment yielded a sample mean of 3.1 minutes and a sample standard deviation of 1.5 minutes. Is there sufficient evidence to indicate, at a 5% level of significance, that the newly developed pain reliever does deliver perceptible relief more quickly?

2. In the past the average length of an outgoing telephone call from a business office has been 143 seconds. A manager wishes to check whether that average has decreased after the introduction of policy changes. A sample of 100 telephone calls produced a mean of 133 seconds, with a standard deviation of 35 seconds. Perform the relevant test at the 1% level of significance.

3. The average household size in a certain region several years ago was 3.14 persons. A sociologist wishes to test, at the 5% level of significance, whether it is different now. Perform the test using the information collected by the sociologist: in a random sample of 75 households, the average size was 2.98 persons, with sample standard deviation of 0.82 person.

4. The recommended daily calorie intake for teenage girls is 2,200 calories/day. A nutritionist at a state university believes the average daily caloric intake of girls in that state to be lower. Test that hypothesis, at the 5% level of significance, against the null hypothesis that the population average is 2200 calories/day using the sample data of size 36, mean 2150, standard deviation 203.

5. Suppose that a delivery company claims that they deliver their packages in 2 days on average, and you suspect it's longer than that. To test this claim, you take a random sample of 10 packages and record their delivery times. You find the sample mean is $\bar{x} = 2.3$ days, and the sample standard deviation is 0.35. Do you have sufficient data to show that the company delivers their packages in more than 2 days, considering significance level 0.05? And what assumptions do you need?

6. An old research has determined the mean age of a certain population to be 70.65 with a 95 % confidence interval of $(67.2, 74.1)$. A new sample yields $\bar{x} = 75.32$.
   (a) What is your conclusion concerning the $H_0 : \mu = 70.65$ with $\alpha = 0.05$?
   (b) Is this a one-sided or a two-sided test?

7. The hypothesis $H_0 : \mu = 15$ is tested against $H_1 : \mu \neq 15$ with $\alpha = 0.05$. A random sample has following properties: $n = 20$, $\bar{x} = 17.5$, $s = 5.9$.
   (a) What assumptions should be made and which distribution should be used? Why?
   (b) What is the range of the P-value that you can find with table A4?
   (c) What is your conclusion regarding the null hypothesis?

8. The average brown trout's I.Q. is 4. Suppose you don't believe this claim. You catch 12 brown trout. A fish psychologist determines the I.Q.s as follows: 5; 4; 7; 3; 6; 4; 5; 3; 6; 3; 8; 5. Conduct a hypothesis test of your belief with $\alpha = 0.05$.

9. A child psychologist says that the average time that working mothers spend talking to their children is 11 minutes per day. Suppose a random sample of 100 working mothers spend on average 11.5 minutes per day talking to their children. The prior research suggests the population standard deviation is 2.3. Do we have evidence that working mothers spend more than 11 minutes per day by talking to their children? Use $\alpha = 0.05$.

10. Consider the previous exercise, but now with a small change: The random sample does not include 100 working mothers, but only 10 working mothers. Will your conclusion change?

11. The Achenbach Child Behaviour Checklist is designed so that scores from normal children are Normal with mean $\mu = 50$ and standard deviation $\sigma = 10$. We are given a sample of $n = 5$ children under stress with an average score of $\bar{x} = 56.0$. Is there evidence that children under stress show an abnormal behaviour? You may assume that σ still is known to be equal to 10. Use level of significance 0.05.

12. Solve the previous exercise, assuming that we do not know σ, but are given a sample of $n = 5$ children under stress with an average score of $\bar{x} = 56.0$ and standard deviation $s = 8.5$.

13. A manufacturer claims that the average tensile strength of thread A exceeds the average tensile strength of thread B by at least 12 kilograms. To test this claim, 50 pieces of each type of thread were tested under similar conditions. Type A thread had an average tensile strength of 86.7 kilograms with a standard deviation of 6.28 kilograms, while type B thread had an average tensile strength of 77.8 kilograms with a standard deviation of 5.61 kilograms. Test the manufacturer's claim using a 0.05 level of significance, using $H_0 : \mu_A - \mu_B \geq 12$ vs $H_1 : \mu_A - \mu_B < 12$.

14. A study was conducted to see if increasing the substrate concentration has an appreciable effect on the velocity of a chemical reaction. With a substrate concentration of 1.5 moles per liter, the reaction was run 15 times, with an average velocity of 7.5 micromoles per 30 minutes and a standard deviation of 1.5. With a substrate concentration of 2.0 moles per liter, 12 runs were made, yielding an average velocity of 8.8 micromoles per 30 minutes and a sample standard deviation of 1.2. Is there any reason to believe that this increase in substrate concentration causes an increase in the mean velocity of the reaction of more than 0.5 micromole per 30 minutes? Use a 0.01 level of significance and assume the populations to be approximately normally distributed with equal variances.