DISCUSSION

Consider how we tested claims about a population mean μ using **confidence intervals** in the previous Chapter.

1 How is this approach related to the Z-test procedure using a critical region?

2 Is a $(1-\alpha)\times 100\%$ confidence interval for μ equivalent to the acceptance region for a Z-test of H_0: $\mu = \mu_0$ versus H_1: $\mu \neq \mu_0$ with significance level α?

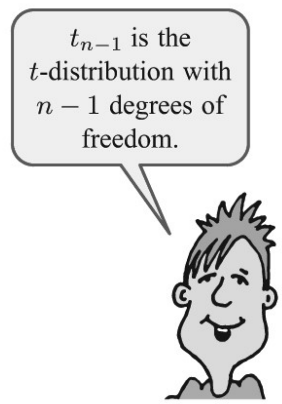
D STUDENT'S t-TEST

When the population variance σ^2 is unknown, we must use the sample variance $S_{n-1}^{\,2}$ to estimate it. We have already seen that when we substitute S_{n-1} for σ in the test statistic of the Z-test, we obtain the statistic $T = \dfrac{\overline{X}_n - \mu_0}{\frac{S_{n-1}}{\sqrt{n}}}$ which has a ***t*-distribution**.

Consider a statistical hypothesis test of H_0: $\mu = \mu_0$ for a normally distributed population with an **unknown** standard deviation. Given a sample of size n, with observed sample mean $\overline{x}$ and sample standard deviation s:

- the **test statistic** is $T = \dfrac{\overline{X}_n - \mu_0}{\frac{S_{n-1}}{\sqrt{n}}}$ which has

  observed value $t = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$
- the **null distribution** is $T \sim t_{n-1}$.



HISTORICAL NOTE

William Sealy Gosset (1876 - 1937) studied chemistry and mathematics at New College, Oxford University. In 1899 he moved to Dublin, Ireland, to work for the brewery of Arthur Guinness & Son. His desire was to improve the production process by selecting the best yielding varieties of barley. Through his study and work with **Karl Pearson** from University College, London, he devised a **test statistic**. His work was published in 1908 in the journal *Biometrika*, but using the name **Student** because of concern by Guinness that other brewers may use his work to their advantage.

William Sealy Gosset

Gosset's test statistic was revised by **Sir Ronald Aylmer Fisher** (1890 - 1962) who recognised the importance of Gosset's work. Fisher called the new statistic t, completing the name **Student's *t*-test**.

The t-distribution is sometimes **Student's *t*-distribution** due to its association with Student's t-test.

STUDENT'S t-TEST FOR THE MEAN OF A POPULATION WITH UNKNOWN VARIANCE

Step 1: State the **null hypothesis** H_0: $\mu = \mu_0$ and **alternative hypothesis** H_1.

Step 2: State the **significance level** α.

Step 3: Calculate the observed value of the **test statistic** $t = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$.

Step 4: Calculate the ***p*-value** using $T \sim t_{n-1}$ as follows:

- If H_1: $\mu > \mu_0$, p-value $= \text{P}(T \geqslant t)$.
- If H_1: $\mu < \mu_0$, p-value $= \text{P}(T \leqslant t)$.
- If H_1: $\mu \neq \mu_0$, p-value $= 2 \times \text{P}(T \geqslant |\,t\,|)$.

Alternatively, find the critical region $\mathcal{C}$ using α and the null distribution.

Step 5: Reject H_0 if p-value $< \alpha$ or if $t \in \mathcal{C}$, otherwise accept H_0.

Step 6: Interpret your decision in the context of the problem. Write your conclusion in a sentence.

Example 4 — Self Tutor

The fat content (in grams) of 30 randomly selected pasties at the local bakery was recorded:

15.1	14.8	13.7	15.6	15.1	16.1	16.6	17.4	16.1	13.9
17.5	15.7	16.2	16.6	15.1	12.9	17.4	16.5	13.2	14.0
17.2	17.3	16.1	16.5	16.7	16.8	17.2	17.6	17.3	14.8

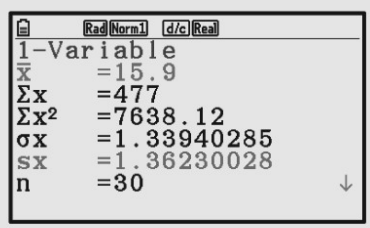
For a mean fat content of pasties made at this bakery μ, conduct a two-tailed t-test of H_0: $\mu = 16$ grams on a 10% level of significance.

Step 1: H_0: $\mu = 16$ {the mean fat content is 16 grams}
H_1: $\mu \neq 16$ {the mean fat content is *not* 16 grams}

Step 2: The significance level is $\alpha = 0.1$.

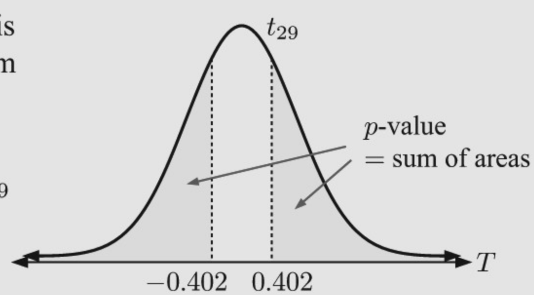
Step 3: The sample size $n = 30$.
Using technology, $\overline{x} = 15.9$ and $s \approx 1.36$.
The observed value of the test statistic is

$$t \approx \frac{15.9 - 16}{\frac{1.36}{\sqrt{30}}} \approx -0.402$$



Step 4: Assuming that a sample size $n = 30$ is large enough for the Central Limit Theorem to apply, we can assume normality.
Since H_1: $\mu \neq 16$,

$$\begin{aligned} p\text{-value} &= 2 \times \text{P}(T \geqslant |\,t\,|) \quad \text{where } T \sim t_{29} \\ &\approx 2 \times \text{P}(T \geqslant |-0.402|) \\ &\approx 0.691 \end{aligned}$$



Step 5: Since p-value $> 0.1 = \alpha$, we do not have enough evidence to reject H_0 in favour of H_1 on a 10% significance level. We therefore accept H_0.

Step 6: Since we have accepted H_0, we cannot conclude that the mean fat content is appreciably different from 16 grams.

Your **graphics calculator** can perform the calculations in a t-test automatically. However, the t-test functionality on most calculator models will only give you the p-value and observed t, but not the critical value(s) of the test.

GRAPHICS CALCULATOR INSTRUCTIONS

Example 5 **Self Tutor**

In 2017, the average house price in a suburb was \$235 000. In 2019, a random sample of 200 houses in the suburb was taken. The sample mean was $\overline{x} = \$215\,000$, and the sample standard deviation was $s = \$30\,000$.

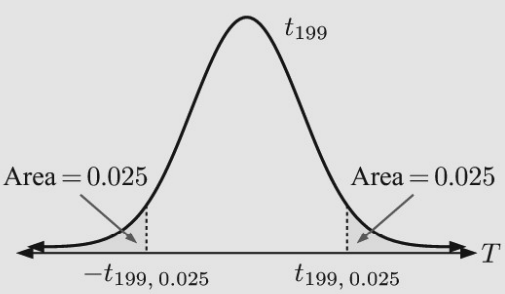
Use a critical value to determine whether there is evidence at the 5% level, that the average house price has changed.

Step 1: Let μ be the population mean of house prices in this suburb.
The hypotheses that should be considered are:
H_0: $\mu = 235\,000$ {the mean house price has stayed the same}
H_1: $\mu \neq 235\,000$ {the mean house price has changed}

Step 2: The significance level is $\alpha = 0.05$.

Step 3: The sample size is $n = 200$.
The observed value of the test statistic is $t = \dfrac{215\,000 - 235\,000}{\frac{30\,000}{\sqrt{200}}} \approx -9.428$

Step 4: The null distribution is t_{199}.
The critical region $\mathcal{C}$ is
$t \leqslant -t_{199,\,0.025}$ or $t \geqslant t_{199,\,0.025}$
Using technology $t_{199,\,0.025} \approx 1.972$.
$\therefore$ $t \leqslant -1.972$ or $t \geqslant 1.972$



Step 5: Since $t \approx -9.428 \in \mathcal{C}$, we reject H_0 on a 5% level of significance.

Step 6: There is sufficient evidence at the 5% level of significance to suggest that $\mu \neq 235\,000$. We conclude that the average house price in 2019 was different from the average house price in 2017.

EXERCISE 30D

1 A population has unknown variance σ^2. A sample of size 24 has sample mean $\overline{x} = 17.14$ and sample standard deviation $s = 4.365$. We are required to test the hypothesis H_0: $\mu = 18.5$ against H_1: $\mu \neq 18.5$.

a Find:
  i the test statistic **ii** the null distribution **iii** the p-value.

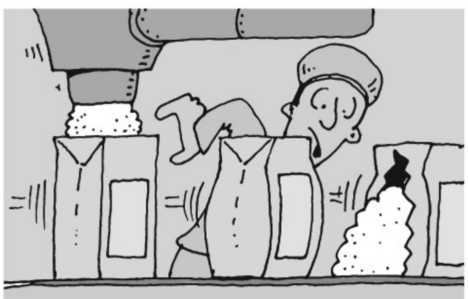
b What decision should be made at a 5% level using:
  i the p-value **ii** the critical region?

2 A store claims that the mean price of a bottle of wine has fallen over the last 12 months. Records show that 12 months ago the mean price was \$13.45 for a 750 mL bottle. A random sample of prices of 389 different bottles of wine is now taken from the store. The sample mean is \$13.30 with sample standard deviation \$0.25. Use a p-value to test the store's claim at a 2% level of significance.

3 A machine is used to fill bottles with 500 mL of vinegar. A sample of ten random measurements of the volumes put in different bottles had mean 499 mL and standard deviation 1.2 mL. Assuming that the volumes of vinegar are normally distributed, test at the 1% level whether the machine is underfilling the bottles.

4 The length of screws produced by a machine is normally distributed. The machine is supposed to produce screws with mean length $\mu = 2.00$ cm. A quality controller selects a random sample of 15 screws. She finds that the mean length of the 15 screws is $\overline{x} = 2.04$ cm with sample standard deviation $s = 0.09$ cm. Does this justify the need to adjust the machine on a 2% level of significance?

5



A machine packs sugar into 1 kg bags. It is known that the masses of the bags of sugar are normally distributed. A random sample of eight filled bags was taken and the masses of the bags measured to the nearest gram. Their masses in grams were:

1001, 998, 999, 1002, 1001, 1003, 1002, 1002.

Perform a test at the 1% level, to determine whether the machine under-fills the bags.

6 A market gardener claims that the carrots in his field have a mean weight of more than 50 grams. A prospective buyer will purchase the crop if the market gardener's claim is true. To test this she pulls 20 carrots at random, and finds that their individual weights in grams are:

57.6	34.7	53.9	52.5	61.8	51.5	61.3	49.2	56.8	55.9
57.9	58.8	44.3	58.3	49.3	56.0	59.5	47.0	58.0	47.2

a Explain why it is reasonable to assume that the carrots' weights are normally distributed.

b Determine whether the buyer will purchase the crop using a 5% level of significance.

7 The management of a golf club claims that the mean income of its members is more than €95 000, and its members can therefore afford to pay increased fees. To show that this claim is invalid, the members seek the help of a statistician. From a random sample of 113 club members, the statistician finds an average income of €96 318 with standard deviation €14 268.

a Test at a 0.02 significance level whether the management's claim should be rejected.

b If the conclusion made by the statistician is incorrect, what type of error is made?

E PAIRED t-TESTS

We are often interested in comparing sets of results for the same group of individuals.

For example, we might consider:

- race times for a class of students at the start and finish of the athletics season
- test results for students in a class at the start and end of a semester.

In this situation we choose a random sample of individuals from the population, then use the data for this group to provide two sets of results for comparison. The data can therefore be **matched in pairs** for each individual in the group.

Notice that the two sets of results are not independent, since they come from the same sample group. However, the scores in each set of results must be independent for the analysis to be meaningful.

Suppose we have 2 sets of results X and Y with population means μ_1 and μ_2 respectively.

The null hypothesis might state that the means are equal:

$$H_0\colon\ \mu_1 = \mu_2$$

or equivalently, that the *difference* between the means is zero:

$$H_0\colon\ \mu_1 - \mu_2 = 0.$$

A randomly selected sample of 10 individuals is chosen, and the results from these individuals are used to form samples for X and Y:

x_i	5	6	0	5	7	2	5	8	8	4
y_i	1	4	0	8	4	1	4	5	5	2

For each individual in the sample, we can calculate the difference d between its x-value and its y-value:

x_i	5	6	0	5	7	2	5	8	8	4
y_i	1	4	0	8	4	1	4	5	5	2
$d_i = x_i - y_i$	4	2	0	-3	3	1	1	3	3	2

These values d_i are observed values of the random variable $D = X - Y$.

$$\begin{aligned} \text{Notice that}\quad \overline{D} &= \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) \\ &= \frac{1}{n}\sum_{i=1}^{n}X_i - \frac{1}{n}\sum_{i=1}^{n}Y_i \\ &= \overline{X} - \overline{Y} \\ \text{so}\quad \mathrm{E}(\overline{D}) &= \mathrm{E}(\overline{X} - \overline{Y}) \\ &= \mathrm{E}(\overline{X}) - \mathrm{E}(\overline{Y}) \\ &= \mu_1 - \mu_2 \end{aligned}$$

So, the population mean difference $\mu_D = \mu_1 - \mu_2$ is the population mean of the differences d_i.

We can therefore use a t-test to test $H_0\colon\ \mu_D = 0$, using the differences d_i to calculate the test statistic and p-value.

For the example above, the test statistic is $t = \dfrac{\overline{d} - 0}{\frac{s_D}{\sqrt{n}}} = \dfrac{1.6 - 0}{\frac{2.01}{\sqrt{10}}} \approx 2.52$.

We would then calculate the p-value using the t-distribution with $10 - 1 = 9$ degrees of freedom.

Alternatively, we could find the critical region $\mathcal{C}$ using α and the null distribution.

EXERCISE 30E

1 The performance of 20 elite sprinters was monitored over a twelve month period. Below is the best time in seconds for each athlete in 100 m trials at the start and end of the year.

Athlete	A	B	C	D	E	F	G	H	I	J
Start	10.3	10.5	10.6	10.4	10.8	11.1	9.9	10.6	10.6	10.8
End	10.2	10.3	10.8	10.1	10.8	9.7	9.9	10.6	10.4	10.6

Athlete	K	L	M	N	O	P	Q	R	S	T
Start	11.2	11.4	10.9	10.7	10.7	10.9	11.0	10.3	10.5	10.6
End	10.8	11.2	11.0	10.5	10.7	11.0	11.1	10.5	10.3	10.2

a Calculate the difference $d_i =$ end time $-$ start time for each athlete.

b State the hypotheses that should be considered for the population mean difference μ_D if we are interested in whether:

i the athletes have improved

ii there is any change in the athletes' performance.

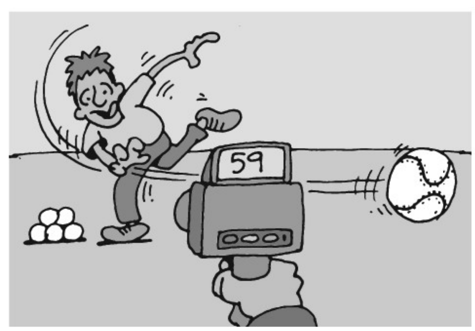
c Conduct a t-test at a 5% level of significance to test whether the athletes have improved.

2 A mathematics tutor claims to significantly increase students' test results with a week of tutoring. To test this claim, 12 students were tested prior to receiving tutoring, and their results recorded. However, the students were not given the answers or their results. After a week of tutoring, the students repeated the test to see whether they had improved. The results were:

Before tutoring	15	17	25	11	28	20	23	34	27	14	26	26
After tutoring	20	16	25	18	28	19	26	37	31	13	27	20

Conduct a paired t-test at a 5% level of significance to test the tutor's claim.

3 A group of 12 year old children were asked to throw a baseball as fast as they could. A radar was used to measure the speed of each throw. One year later, the same group was asked to repeat the experiment. The results are shown below, with the speeds given in km h^{-1}.



Child	A	B	C	D	E	F	G	H	I	J	K
Age 12	76	81	59	67	90	74	78	71	69	72	82
Age 13	79	82	66	72	93	76	77	82	75	77	86

a Calculate the difference in speed d_i for each child's throws.

b A sports commission report suggests that an average throwing speed difference of 5 km h^{-1} is expected between these ages.

Conduct a hypothesis test at a 5% level of significance to determine whether the report's claim is valid.

Hint: The null value for μ_D does not have to be zero.