

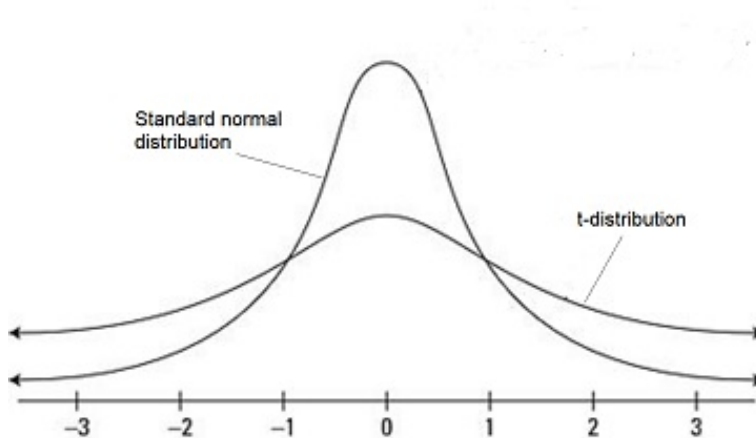
S15: *T-tests*

T-distribution

In conducting a hypothesis test the population standard deviation σ may be unknown. In that case we approximate with the sample standard deviation s when calculating the standard error of the distribution of means.

$$s_{\bar{x}} \approx \sigma_{\bar{x}} \text{ or } \frac{s}{\sqrt{n}} \approx \frac{\sigma}{\sqrt{n}}.$$

A sampling distribution that results from approximating σ with s is called a t-distribution. However, σ is a fixed value and s varies with each sample. This creates a level of uncertainty that slightly distorts the shape of the distribution of sample means. The resulting curve is close to a normal distribution when n is large, but for small n it has heavier tails.



The precise shape of the t-distribution depends on the size of the sample. A t-distribution is always bell shaped but the shape of the curve varies according to the size of the sample. As the sample size increases, a t-distribution becomes more like the z-distribution.



T-tests

A t-test is a hypothesis test which like the z-test is used to make decisions about the similarity of sample and population means. The test statistic is

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Like a z-test it is assumed that the sample data is drawn randomly and that the distribution of sample means follows a normal distribution.

Many students are confused about whether to use a z-test or a t-test to test their hypothesis.

Generally,

1. if σ is known, then a z-test is appropriate
2. if σ is not known, then a t-test is appropriate.

However,

- Many text books only provide t-tables for $n \leq 30$. For larger samples the z-distribution is a reasonable approximation and a z-test may be used even though σ is unknown.
- Some courses provide students with more comprehensive t-tables for n up to 50 or 100. If you have access to relevant tables you should use a t-test whenever σ is not known regardless of the size of the sample.
- If s is known to be a good approximation for σ then a z-test may be appropriate.
- If a computer package such as SPSS, Minitab or EXCEL is being used for the hypothesis test then a t-test should be selected when σ is unknown.

The conventions differ between courses so check what is appropriate for your subject

Steps in a t-test

1. State the hypotheses

$H_0 : \mu_{\bar{x}} = \mu$ (the sample mean is the same as the population mean after allowing for chance variation)

$H_1 : \mu_{\bar{x}} \neq \mu$ (the sample mean is not the same as the population mean after allowing for chance variation)

2. Significance level α is chosen ($\alpha = 0.05 \Rightarrow$ we are defining reasonable as what we can expect 95% of the time)
3. Critical values
Tables or a calculator or a computer are used to find the t-value that corresponds to the chosen significance level. These are called the critical values and depend on the degrees of freedom ($n - 1$) and α .
4. The test statistic is the standardised difference between the sample mean (calculated from the given data) and the known population mean: $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$
5. Decision: Is the result reasonable if H_0 is true?
Is the test statistic more extreme than the critical value?
Yes \Rightarrow Reject H_0 , No \Rightarrow Do not reject H_0
6. Conclusion
There is (if you reject)/is not (if you do not reject) evidence to suggest that . . .

It is important to note that

- The decision about the null hypothesis is not made with certainty but with a level of confidence that the error in the decision is small (for example 5% if $\alpha = 0.05$)
- The decision relates only to rejecting or not rejecting H_0 . H_a is not mentioned in the decision, and we DO NOT accept H_0 or H_a .
- The steps for hypothesis testing may differ from course to course so check with your program.

Example

A manufacturer of batteries claims that on average a battery lasts 200 hours. To test this claim 7 batteries are randomly selected and this sample has an average of only 190 hours with a standard deviation of 9 hours. Does this sample provide evidence at the 1% level of significance that the manufacturers claim is incorrect?

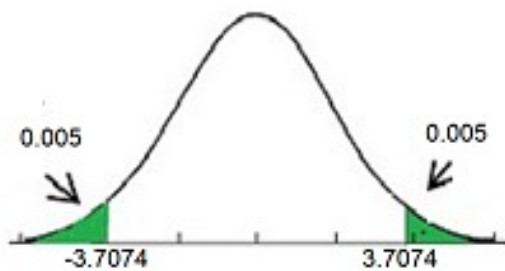
1. State the hypotheses
 $H_0 : \mu_{\bar{x}} = 200$ (the sample mean is the same as the population mean)
 $H_a : \mu_{\bar{x}} \neq 200$ (the sample mean is not the same as the population mean)

2. Significance level $\alpha = 0.01$

3. Critical values

$\alpha = 0.01$ indicates the column and degrees of freedom = $n - 1 = 6$ indicates the row

α (2 tail)	0.1	0.05	0.02	0.01	0.005	0.002	0.001
df							
1	6.3138	12.7065	31.8193	63.6551	127.3447	318.4930	636.0450
2	2.9200	4.3026	6.9646	9.9247	14.0887	22.3276	31.5989
3	2.3534	3.1824	4.5407	5.8408	7.4534	10.2145	12.9242
4	2.1319	2.7764	3.7470	4.6041	5.5976	7.1732	8.6103
5	2.0150	2.5706	3.3650	4.0322	4.7734	5.8934	6.8688
6	1.9432	2.4469	3.1426	3.7074	4.3168	5.2076	5.9589
7	1.8946	2.3646	2.9980	3.4995	4.0294	4.7852	5.4079
8	1.8595	2.3060	2.8965	3.3554	3.8325	4.5008	5.0414
9	1.8331	2.2621	2.8214	3.2498	3.6896	4.2969	
10	1.8124	2.2282	2.7638	3.1693			
11	1.7959	2.2010					



4. Test statistic

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{190 - 200}{\frac{9}{\sqrt{7}}} = -2.94$$

5. Decision:

Is -2.94 more extreme than -3.7074? No \Rightarrow Do not reject H_0

6. Conclusion

There is not evidence to suggest that the average battery life differs from 200 hours.

Exercises

Your answers should be set out and contain all the steps shown above. A brief outline of the main features is given in the answers below.

1. The number of nic nac lollies in a small pack follows a normal distribution and the average is claimed to be 40. A suspicious customer buys 10 packs and counts the nic nacs in each pack. The results are: 38 45 36 40 42 44 35 41 36 40. Use a one sided test with $\alpha = 0.05$ to decide whether the customer's are justified.
Answer: Test statistic = -0.28 do not reject H_0 .
2. A hospital claims a new process will reduce the waiting time for surgery to treat a condition considered as non life threatening. To date the average waiting time has been 15.7 months. A random sample of 33 patients diagnosed with the condition were observed and the time till surgery recorded ($\bar{x} = 13.664, s = 2.544$). Test the hospitals claim at the 1% significance level.
Answer: Test statistic = -0.28 do not reject H_0 .
3. The quality of a pre-natal program for at-risk mothers is to be assessed by comparing the weight of babies born to mothers participating in the program with the historical average of 2800g. The babies of the 25 mothers in the program had a mean birth weight of 3075g and a standard deviation of 300g. Comment on success of the program using a two sided hypothesis test and $\alpha = 0.01$.
Answer: Test statistic = 4.58 reject H_0

The title graphic was sourced from Pixabay:
<https://pixabay.com/photos/pencil-unusual-abnormal-red-4552633/>