

7A Mean of a Normal Distribution with Unknown Variance

For a random sample $X_1, X_2, \dots, X_n$ from a normal distribution with mean μ and unknown variance σ^2 :

Degrees of freedom of the t-distribution, and what happens as $v \rightarrow \infty$:

Shape of the t-distribution:

Percentage points of the t-distribution (the value of t exceeded with the probability shown):

v	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169

1. The random variable X has a t -distribution with 10 degrees of freedom. Determine values of t for which:

a) $P(X > t) = 0.025$

b) $P(X < t) = 0.95$

c) $P(X < t) = 0.025$

d) $P(|X| > t) = 0.05$

e) $P(|X| < t) = 0.98$

2. The random variable Y has a t_4 -distribution. Determine:

a) $P(Y > 3.747)$

b) $P(Y < -2.132)$

100(1 - α)% confidence interval for the mean μ of a normal distribution with unknown variance:

3. A sample of 6 trout taken from a fish farm were caught and their lengths in centimetres were measured. The lengths of the fish were as follows:

26.8 26.0 25.8 25.5 24.3 24.6

Assuming that the lengths of trout are normally distributed, find a 90% confidence interval for the mean length of trout in the fish farm.

4. The percentage starch content of potatoes is normally distributed with mean μ . In order to assess the mean value of the starch content, a random sample of 12 potatoes is selected and their starch content measured. The percentages of starch contents obtained were as follows:

23.2 20.3 18.6 20.0 20.8 21.6 19.4 18.7 22.1 19.5 21.3 22.6

Find a 95% confidence interval for the mean.

7B Hypothesis Test for the Mean of a Normal Distribution with Unknown Variance

Test statistic and its distribution under H_0 :

Formulae for $\bar{X}$ and s^2 :

Steps for the test:

- 1
- 2
- 3
- 4
- 5
- 6
- 7

1. A shopkeeper sells jars of jam. The weights of the jars of jam are normally distributed with a mean of 150 g. A customer complains that the mean weight of 8 jars she had bought was only 147 g. An estimate for the standard deviation of the weights of the 8 jars of jam calculated from the 8 observations was 2 g.
 - a) Test, at the 5% significance level, whether 147 g is significantly less than the quoted mean.

- b) Discuss whether the customer has cause for complaint.

2. The temperatures ($^{\circ}\text{C}$) were measured at noon on 10 days during the month of March in West Cumbria. The readings were:

12.8 11.4 12.9 15.1 15.4 13.5 14.9 15.0 16.0 15.8

Using a 5% significance level, test whether or not this is an increase over the previous year when the average noon day temperature was 13.5°C .

3. A concrete manufacturer tests cubes of its concrete at regular intervals, and their compressive strengths in N m^{-1} are determined. The mean value of the strengths is required to be 0.47 N m^{-1} . A new supplier of cement offers to supply the firm at a cheaper rate than the present supplier, and a trial bag of cement is used to make 12 concrete cubes. Upon testing, these cubes are found to have strengths (x) such that $\sum x = 5.52$ and $\sum x^2 = 2.542$. Assume that the strengths are normally distributed.

- a) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the use of the new cement has altered the mean strength of the concrete.

- b) In the light of your conclusion to the test in part a, what would you recommend the manufacturer to do?

7C The Paired t-Test

Paired data:

What you assume about the differences:

Test statistic for a paired t-test:

1. In an experiment to test the effects of alcohol on the reaction times of people, a group of 10 students took part in the experiment. The students were asked to react to a light going on by pushing a switch that would switch it off again. Their reaction times were automatically recorded. After the students had each drunk one pint of beer the experiment was repeated. The results are shown below.

Student	A	B	C	D	E	F	G	H	I	J
Before (s)	0.8	0.2	0.4	0.6	0.4	0.6	0.4	0.8	1.0	0.9
After (s)	0.7	0.5	0.6	0.8	0.8	0.6	0.7	0.9	1.0	0.7
Difference										

Test, at the 5% significance level, whether or not the consumption of a pint of beer increased the students' reaction times.

2. In order to compare two methods of measuring the hardness of metals, readings of Brinell hardness were taken using each method for 8 different metal specimens. The resulting Brinell hardness readings are given in the table below:

Material	Method A	Method B	d
Aluminium	29	31	
Magnesium alloy	64	63	
Wrought iron	104	105	
Duralumin	116	119	
Mild steel	138	140	
70/30 brass	156	156	
Cast iron	199	200	
Nickel chrome steel	385	386	

Use a paired t-test, at the 5% level of significance, to test whether or not there is a difference in the readings given by the two methods.

7D Difference Between Means of Two Independent Normal Distributions

Pooled estimate of the common variance σ^2 : $s_p^2 =$

1. A random sample of 15 observations is taken from a population and gives an unbiased estimate for the population variance of 9.47. A second random sample of 12 observations is taken from a different population that has the same population variance as the first population, and gives an unbiased estimate for the variance as 13.84. Calculate an unbiased estimate of the population variance σ^2 using both samples.

The three assumptions needed when the sample sizes are small:

Distribution of the test statistic for the difference between two means:

Confidence interval for $\mu_x - \mu_y$ when the variances are equal but unknown:

2. In a survey on the petrol consumption of cars, a random sample of 12 cars with 2-litre engines was compared with a random sample of 15 cars with 1.6-litre engines. The following results show the consumption, in suitable units, of the cars:

2-litre cars: 34.4, 32.1, 30.1, 32.8, 31.5, 35.8, 28.2, 26.6, 28.8, 28.5, 33.6, 28.8

1.6-litre cars: 35.3, 34.0, 36.7, 40.9, 34.4, 39.8, 33.6, 36.7, 34.0, 39.2, 39.8, 38.7, 40.8, 35.0, 36.7

Calculate a 95% confidence interval for the difference between the two mean petrol consumption figures. You may assume that the variables are normally distributed and that they have the same variance.

7E Hypothesis Test for the Difference Between Means

Steps for a test of the difference between two means (variances equal but unknown):

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9

1. Two groups of students, X and Y, were taught by different teachers. At the end of their course, a random sample of students from each class was selected and given a test. The test results out of 50 were as follows:

Group X: 40 37 45 34 30 41 42 43 36

Group Y: 38 43 36 45 35 44 41

The headteacher wishes to find out if there is a significant difference between the results for these two groups.

- a) Write down any assumptions that need to be made in order to conduct a difference of means test on this data.

- b) Assuming that these assumptions apply, test at the 10% level of significance whether or not there is a significant difference between the means.

2. A random sample of the heights, in cm, of sixth form boys and girls was taken with the following results:

Boys' heights: 152, 148, 147, 157, 158, 140, 141, 144

Girls' heights: 142, 146, 132, 125, 138, 131, 143

- a) Carry out a two-sample t-test at the 5% significance level on these data to see whether the mean height of boys exceeds the mean height of girls by more than 4 cm.

- b) State any assumptions that you have made.