Mathtrace – user guide

V. Statistics

Contents

V. Statistics

41. Statistics

41. Statistics

Write data as data(…): separate the values by semicolons, spaces or new lines (data(1.5 2.5 3)). A table – a row per observation, a column per variable – goes row by row separated by semicolons; the first row may hold the variables' names. More conveniently: the grid key on the f(x) tab opens a table to fill in, paste into from the clipboard (cells from Excel or Google Sheets) or load from a CSV file, and the "As data" button inserts it for statistics – even a table bigger than the grid.

For one variable, a chip picks what to compute. Descriptive statistics: the arithmetic mean, the median, the mode, the quartiles (the medians of the lower and upper half), the range and the interquartile range, the sample (n − 1) and population (n) variance and standard deviation, the coefficient of variation and the outliers (outside $Q_1 - 1.5 \cdot IQR$ and $Q_3 + 1.5 \cdot IQR$), with a box plot. Frequencies and histogram: the table of absolute, relative and cumulative frequencies – by values for whole numbers with few values, by classes otherwise (their count by Sturges' rule). z-scores: each value standardized by the data's mean and standard deviation, with its T-score.

Two variables (two columns) give correlation and regression: the sums, the covariance, Pearson's correlation coefficient r with the strength of the relationship in words, the least-squares regression line and the coefficient of determination r^2, with a scatter plot. Three or more variables get their descriptive statistics side by side with box plots.

Statistical inference from data: the confidence interval of the mean (chip “Confidence interval of the mean”, 90, 95 or 99 %) by the t distribution; the t-test (chip “t-test”) – one variable against the tested value μ_0, two variables as independent samples (Welch's test, or with equal variances) or paired values (the pairs' differences); the χ^2 test (chip “χ^2 test”) – one column of counts as a goodness-of-fit test to a uniform distribution, a table as a test of independence with Cramér's V. Under the chips you enter the significance level α, the alternative (two-sided, left-sided, right-sided) and μ_0. The steps: the hypotheses, the test statistic, the critical value and region, the p-value and the decision; the graph shows the distribution with the critical region shaded. Write two independent samples of different sizes as named columns – data(A = 23 25 28; B = 20 22 21 25).

Comparing groups: the chip "Rank test (Mann–Whitney, Wilcoxon, Kruskal–Wallis)" compares two groups by the Mann–Whitney test (with "Paired" paired values by Wilcoxon's test) and three or more groups by the Kruskal–Wallis test – by the values' ranks, without assuming a normal distribution; for small samples without ties the p-value is exact. The chip "ANOVA (analysis of variance)" compares the groups' means by one-way analysis of variance with the ANOVA table, the F-test and the effect size η^2. Write the groups in columns, of any lengths (named columns); the graph shows the groups' box plots and the test statistic's distribution. For the χ^2 test of one column of counts, the field "Expected shares" takes your own expected shares – 1:2:1, 0.25 0.5 0.25 or in percent.

The regression of two variables has a model choice under its chip: linear (with the test of the correlation's significance), quadratic (the least squares normal equations), exponential y = a·e^(bx), power y = a·x^b and logarithmic y = a + b·ln x (by linearization through logarithms) – with the coefficient of determination R² and a graph of the fitted curve. The “Statistics and psychometrics” form also has the confidence interval and t-tests from summary values (size, mean, standard deviation) – including the z-test with a known σ and the two-sample t-test.

The normal distribution is in the "Statistics and psychometrics" form (the "Applications" section of the home screen, the Statistics field): the probabilities P(X < x), P(X > x), P(a < X < b) and the quantile – the value below which a given share lies. The steps standardize, z = (x − μ)/σ, and use the distribution function Φ, a negative z as 1 − Φ(|z|) as in the tables; the graph shows the density with the area shaded.

How to enter it

`data(3; 5; 7; 7; 9)`	One variable – the values by semicolons, spaces or one under another.
`data(x y; 1 52; 2 58; 3 61)`	A table row by row, the first row the variables' names (optional).
`data(x = 1 2 3 4; y = 2 4 5 4)`	Named columns, each with its values.

Examples

Descriptive statistics of eight values.

Input: `data(2; 4; 4; 4; 5; 5; 7; 9)`

Result

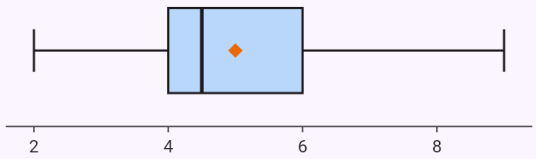
$$\bar{x} = 5$$
$$\tilde{x} = 4.5$$
$$s \approx 2.138$$

Steps

- The count of values $n = 8$, the data sorted:
  $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- The arithmetic mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{40}{8} = 5$
- The median is the middle value of the sorted data:
  $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{4+5}{2} = 4.5$
- The mode (the most frequent value, 3×): $\hat{x} = 4$
- The quartiles are the medians of the lower and the upper half: $Q_1 = 4, \quad Q_2 = \tilde{x} = 4.5, \quad Q_3 = 6$
- The range and the interquartile range
  $R = x_{max} - x_{min} = 9 - 2 = 7$
  $IQR = Q_3 - Q_1 = 6 - 4 = 2$
- The sum of the squared deviations from the mean:
  $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 32$
- The variance – the sample s² and the population σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{32}{7} \approx 4.571$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{32}{8} = 4$
- The standard deviation
  $s = \sqrt{s^2} = \sqrt{4.571} \approx 2.138$
  $\sigma = \sqrt{4} = 2$
- The coefficient of variation:
  $v = \frac{s}{\bar{x}} = \frac{2.138}{5} \approx 0.4276 = 42.8\,\%$
- The fences for outliers
  $Q_1 - 1.5\,IQR = 1, \quad Q_3 + 1.5\,IQR = 9$ – no value lies outside them.
- The sample variance s² divides by n − 1 and estimates the population's variance from a sample – the usual choice. The population σ² divides by n, for data that are the whole population.
- Careful: there are few data, the estimates (of the variance and quartiles especially) are uncertain.

Graph

Box plot (median, quartiles, outliers; ◆ mean)



The outlier 31 in the box plot.

Input: `data(12 15 14 13 16 15 14 31)`

Result

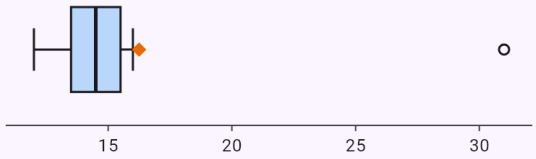
$$\bar{x} = 16.25$$
$$\tilde{x} = 14.5$$
$$s \approx 6.089$$

Steps

- The count of values $n = 8$, the data sorted:
  $12,\ 13,\ 14,\ 14,\ 15,\ 15,\ 16,\ 31$
- The arithmetic mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{130}{8} = 16.25$
- The median is the middle value of the sorted data:
  $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{14 + 15}{2} = 14.5$
- The mode (the most frequent value, 2×): $\hat{x} = 14,\ 15$
- The quartiles are the medians of the lower and the upper half: $Q_1 = 13.5, \quad Q_2 = \tilde{x} = 14.5, \quad Q_3 = 15.5$
- The range and the interquartile range
  $R = x_{max} - x_{min} = 31 - 12 = 19$
  $IQR = Q_3 - Q_1 = 15.5 - 13.5 = 2$
- The sum of the squared deviations from the mean:
  $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 259.5$
- The variance – the sample s² and the population σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{259.5}{7} \approx 37.07$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{259.5}{8} \approx 32.44$
- The standard deviation
  $s = \sqrt{s^2} = \sqrt{37.07} \approx 6.089$
  $\sigma = \sqrt{32.44} \approx 5.695$
- The coefficient of variation:
  $v = \frac{s}{\bar{x}} = \frac{6.089}{16.25} \approx 0.3747 = 37.5\,\%$
- Outliers lie outside the fences
  $Q_1 - 1.5\,IQR = 10.5, \quad Q_3 + 1.5\,IQR = 18.5$: 31
- The sample variance s² divides by n − 1 and estimates the population's variance from a sample – the usual choice. The population σ² divides by n, for data that are the whole population.
- Careful: there are few data, the estimates (of the variance and quartiles especially) are uncertain.

Graph

Box plot (median, quartiles, outliers; ◆ mean)



The frequency table of grades and its histogram.

Input: `data(1 2 2 3 3 3 3 4 4 5 1 3 2 4 3)` Frequencies and histogram

Result

$n = 15$

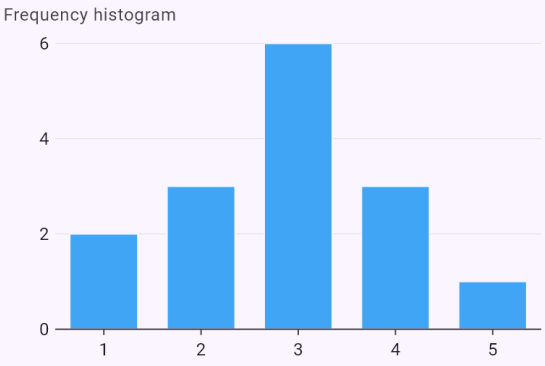
$\hat{x} = 3$

Steps

- The count of values $n = 15$, the data sorted:
  $1,\ 1,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ 4,\ 5$
- The frequency table – absolute n, relative f, cumulative N and F

x_i	n_i	f_i	N_i	F_i
1	2	0.133	2	0.133
2	3	0.200	5	0.333
3	6	0.400	11	0.733
4	3	0.200	14	0.933
5	1	0.067	15	1.000
$\sum$	15	1		

Graph



Continuous data grouped into classes.

Input: `data(12.5 14.1 15.8 17.2 18.9 21.3 22.0 23.7 25.1 26.4 28.8 31.2)` Frequencies and histogram

Result

$n = 12$

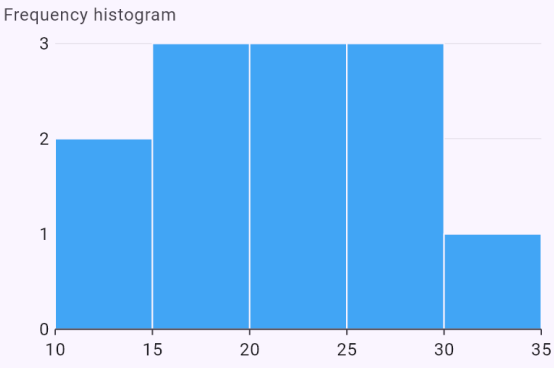
$k = 5, \quad h = 5$

Steps

- The count of values $n = 12$, the data sorted:
  $12.5,\ 14.1,\ 15.8,\ 17.2,\ 18.9,\ 21.3,\ 22,\ 23.7,\ 25.1,\ 26.4,\ 28.8,\ 31.2$
- There are many values – we group them into 5 classes (Sturges' rule)
  $k \approx 1 + \log_2 n = 1 + \log_2 12 \approx 4.585 \Rightarrow k = 5$
  $h \approx \frac{R}{k} = \frac{18.7}{5} \approx 3.74 \Rightarrow h = 5$
- The classes' frequency table (x the class mark)

I_j	x_j	n_j	f_j	N_j	F_j
$[10,\ 15)$	12.5	2	0.167	2	0.167
$[15,\ 20)$	17.5	3	0.250	5	0.417
$[20,\ 25)$	22.5	3	0.250	8	0.667
$[25,\ 30)$	27.5	3	0.250	11	0.917
$[30,\ 35]$	32.5	1	0.083	12	1.000
$\sum$		12	1		

Graph



The values' z-scores and T-scores.

Input: `data(2 4 4 4 5 5 7 9)` z-scores

Result

$\bar{x} = 5$

$s \approx 2.138$

Steps

- The count of values $n = 8$, the data sorted: $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- The arithmetic mean: $\bar{x} = \frac{40}{8} = 5$
- The standard deviation

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2} = \sqrt{\frac{32}{7}} \approx 2.138$$

- Each value standardized

$$z_i = \frac{x_i - \bar{x}}{s} = \frac{x_i - 5}{2.138}$$
$$T_i = 50 + 10 z_i$$

- The values' z-scores and T-scores

x_i	z_i	T_i
2	−1.40	36.0
4	−0.47	45.3
4	−0.47	45.3
4	−0.47	45.3
5	0.00	50.0
5	0.00	50.0
7	0.94	59.4
9	1.87	68.7

- The z-scores use these data's mean and standard deviation. To compare with a population, use the norm's mean and SD (the model "Converting standard scores").

Correlation and the regression line – hours of study and test points.

Input: `data(x y; 1 52; 2 58; 3 61; 4 70; 5 74; 6 79)`

Result

$$r \approx 0.9937$$

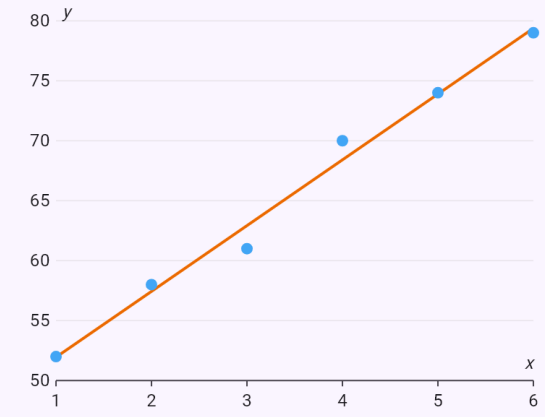
$$\hat{y} = 46.47 + 5.486x$$

Steps

- The count of values $n = 6$, the data sorted
  $[1,\ 52],\ [2,\ 58],\ [3,\ 61],\ [4,\ 70],\ [5,\ 74],\ [6,\ 79]$
- The sums and means
  $\sum x_i = 21$
  $\sum y_i = 394$
  $\sum x_i^2 = 91$
  $\sum y_i^2 = 26\,410$
  $\sum x_i y_i = 1475$
  $\bar{x} = 3.5$
  $\bar{y} = 65.67$
- The standard deviation
  $$s_x = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}} \approx 1.871$$
  $$s_y = \sqrt{\frac{\sum y_i^2 - n\bar{y}^2}{n-1}} \approx 10.33$$
- The covariance
  $$s_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n-1} = \frac{1475 - 6 \cdot 3.5 \cdot 65.67}{5} = 19.2$$
- Pearson's correlation coefficient:
  $r = \frac{s_{xy}}{s_x s_y} = \frac{19.2}{1.871 \cdot 10.33} \approx 0.9937$ – the relationship is very strong, positive (y grows with x).
- The regression line's coefficients by least squares
  $$b = \frac{s_{xy}}{s_x^2} = \frac{19.2}{3.5} \approx 5.486$$
  $$a = \bar{y} - b\bar{x} = 65.67 - 5.486 \cdot 3.5 \approx 46.47$$
- The regression line: $\hat{y} = 46.47 + 5.486x$
- The coefficient of determination $r^2 = 0.9937^2 \approx 0.9874$ – the line explains $98.7\,\%$ of y's variance.
- The test of the null hypothesis ρ = 0 (t distribution with 4 degrees of freedom) – the correlation is statistically significant.
  $$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.9937\sqrt{4}}{\sqrt{1-0.9937^2}} \approx 17.73$$
  $p < 0.0001 \le \alpha = 0.05$
- Correlation describes only a linear relationship: it doesn't say what causes what and may miss a non-linear one. Use the regression line only within the range of the measured x.

Graph

Scatter plot with the regression line



Three variables side by side.

Input: `data(A B C; 12 15 9; 14 18 11; 11 16 10; 15 19 12; 13 17 10)`

Result

$$\bar{x}_{\mathrm{A}} = 13$$
$$\bar{x}_{\mathrm{B}} = 17$$
$$\bar{x}_{\mathrm{C}} = 10.4$$

Steps

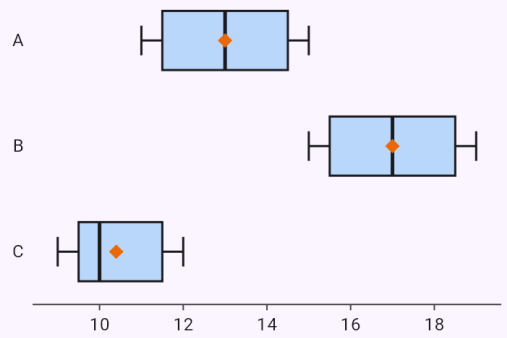
- The descriptive statistics of the variables

	A	B	C
n	5	5	5
$\bar{x}$	13	17	10.4
$\tilde{x}$	13	17	10
Q_1	11.5	15.5	9.5
Q_3	14.5	18.5	11.5
s	1.581	1.581	1.14
σ	1.414	1.414	1.02
x_{min}	11	15	9
x_{max}	15	19	12

- The sample variance s² divides by n − 1 and estimates the population's variance from a sample – the usual choice. The population σ² divides by n, for data that are the whole population.

Graph

Box plot (median, quartiles, outliers; ◆ mean)



95% confidence interval of the mean of ten measurements.

Input: `data(102 98 110 105 99 107 101 104 96 108)` Confidence interval of the mean

Result

$$95\,\%: \ [99.71,\ 106.3]$$

Steps

- Size, mean and sample standard deviation:
  $n = 10, \quad \bar{x} = 103, \quad s \approx 4.595$
- The standard error: $SE = \dfrac{s}{\sqrt{n}} = \dfrac{4.595}{\sqrt{10}} \approx 1.453$
- The quantile for the chosen confidence:
  $t_{0.975}(9) \approx 2.262$
- The confidence interval
  $\bar{x} \pm t \cdot SE = 103 \pm 2.262 \cdot 1.453 = 103 \pm 3.287 \Rightarrow [99.$
- 95 % confidence means: if the sampling were repeated many times, about 95 % of the intervals computed this way would contain the true mean μ. It doesn't mean μ lies in this particular interval with a probability of 95 %.
- For a small sample (n < 30) the t distribution assumes the data come from a roughly normal distribution. With strongly skewed data or outliers a nonparametric test is more suitable.

Does the mean differ from 100? One-sample t-test.

Input: `data(102 98 110 105 99 107 101 104 96 108)` t-test

Result

$$t \approx 2.065$$
$$p \approx 0.06895 > \alpha = 0.05$$

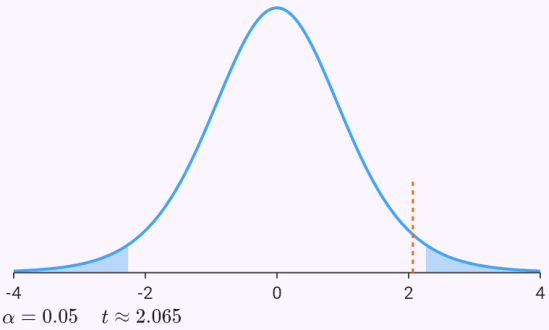
Steps

- Size, mean and sample standard deviation: $n = 10, \quad \bar{x} = 103, \quad s \approx 4.595$
- The null and the alternative hypothesis: $H_0: \mu = 100, \quad H_1: \mu \neq 100$
- The standard error: $SE = \frac{s}{\sqrt{n}} = \frac{4.595}{\sqrt{10}} \approx 1.453$
- The test statistic – under H_0 it has a t distribution with 9 degrees of freedom: $t = \frac{\bar{x} - \mu_0}{SE} = \frac{103 - 100}{1.453} \approx 2.065$
- The critical value at the level α: $t_{0.975}(9) \approx 2.262$, the critical region $|t| \geq 2.262$
- The p-value: $p = 2\,P(T \geq 2.065) \approx 0.06895$
- Since $p \approx 0.06895 > \alpha = 0.05$, we don't reject the null hypothesis – the data don't give enough evidence against it.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.

- For a small sample (n < 30) the t distribution assumes the data come from a roughly normal distribution. With strongly skewed data or outliers a nonparametric test is more suitable.

Graph

t distribution – the critical region shaded, the test statistic dashed



Two independent samples of different sizes – Welch's t-test.

Input: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)` t-test

Result

$$t \approx 3.503$$
$$p \approx 0.005546 \leq \alpha = 0.05$$

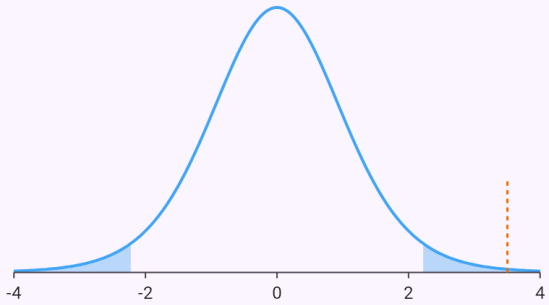
Steps

- Size, mean and sample standard deviation
  $n_1 = 6,\ \bar{x}_1 = 26.5,\ s_1 \approx 2.429$
  $n_2 = 7,\ \bar{x}_2 = 22,\ s_2 \approx 2.16$
- The null and the alternative hypothesis:
  $H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \neq \mu_2$
- The standard error
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2.429^2}{6} + \frac{2.16^2}{7}} \approx 1.285$$
- The Welch–Satterthwaite degrees of freedom
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 10.18$$
- The test statistic – under H_0 it has a t distribution with 10.18 degrees of freedom:
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{26.5 - 22}{1.285} \approx 3.503$$
- The critical value at the level α: $t_{0.975}(10.18) \approx 2.223$, the critical region $|t| \geq 2.223$
- The p-value: $p = 2\,P(T \geq 3.503) \approx 0.005546$
- Since $p \approx 0.005546 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.

- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- Welch's test doesn't assume equal variances of the groups; its degrees of freedom v need not be whole.

Graph

t distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05 \quad t \approx 3.503$

Paired t-test – measurements before and after.

Input: `data(x1 x2; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71)` t-test · Paired

Result

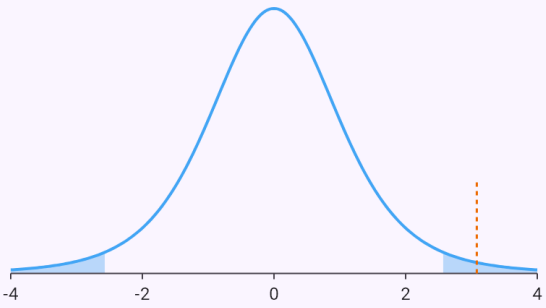
$$t \approx 3.081$$
$$p \approx 0.02743 \leq \alpha = 0.05$$

Steps

- The pairs' differences d (column 1 − column 2): $2,\ 2,\ -1,\ 4,\ 3,\ 3;\ n = 6, \quad \bar{d} \approx 2.167, \quad s_d \approx 1.722$
- The null and the alternative hypothesis: $H_0: \mu_d = 0, \quad H_1: \mu_d \neq 0$
- The standard error: $SE = \frac{s}{\sqrt{n}} = \frac{1.722}{\sqrt{6}} \approx 0.7032$
- The test statistic – under H_0 it has a t distribution with 5 degrees of freedom: $t = \frac{\bar{d} - \mu_0}{SE} = \frac{2.167 - 0}{0.7032} \approx 3.081$
- The critical value at the level α: $t_{0.975}(5) \approx 2.571$, the critical region $|t| \geq 2.571$
- The p-value: $p = 2\,P(T \geq 3.081) \approx 0.02743$
- Since $p \approx 0.02743 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- For a small sample (n < 30) the t distribution assumes the data come from a roughly normal distribution. With strongly skewed data or outliers a nonparametric test is more suitable.
- The paired t-test is a one-sample t-test of the differences (column 1 − column 2) – each pair belongs to the same subject (e.g. before and after).

Graph

t distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05 \quad t \approx 3.081$

Is the die fair? χ² goodness-of-fit test.

Input: `data(8 12 9 11 6 14)` χ² test

Result

$$\chi^2 = 4.2$$
$$p \approx 0.521 > \alpha = 0.05$$

Steps

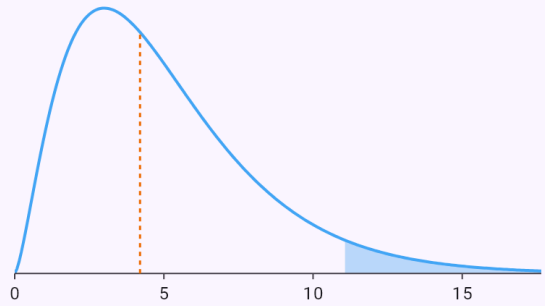
- Null hypothesis: all 6 categories are equally likely ($H_0: p_1 = \cdots = p_6 = \frac{1}{6}$); alternative: at least one differs.
- The expected counts under H_0 $E_i = \frac{N}{k} = \frac{60}{6} = 10$ and the categories' contributions

i	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
1	8	10	0.4
2	12	10	0.4
3	9	10	0.1
4	11	10	0.1
5	6	10	1.6
6	14	10	1.6
$\sum$	60	60	4.2

- The test statistic – under H_0 it has a χ² distribution with 5 degrees of freedom: $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 0.4 + 0.4 + 0.1 + 0.1 + 1.6 + 1.6 = 4.2$
- The critical value at the level α: $\chi^2_{0.95}(5) \approx 11.07$, the critical region $\chi^2 \geq 11.07$
- The p-value: $p = P(\chi^2 \geq 4.2) \approx 0.521$
- Since $p \approx 0.521 > \alpha = 0.05$, we don't reject the null hypothesis – the data don't give enough evidence against it.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- The χ² test works with counts (numbers of cases), not with percentages or means.

Graph

χ² distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05$ $\chi^2 = 4.2$

Test of independence in a 2×2 table.

Input: `data(20 30; 25 15)` χ² test

Result

$$\chi^2 = 4.5$$
$$p \approx 0.03389 \leq \alpha = 0.05$$
$$V \approx 0.2236$$

Steps

- Null hypothesis: the row and the column variable are independent; alternative: they are dependent.
- The expected counts under independence $E_{ij} = \frac{n_{i\cdot}\, n_{\cdot j}}{N}$, the observed ones in brackets

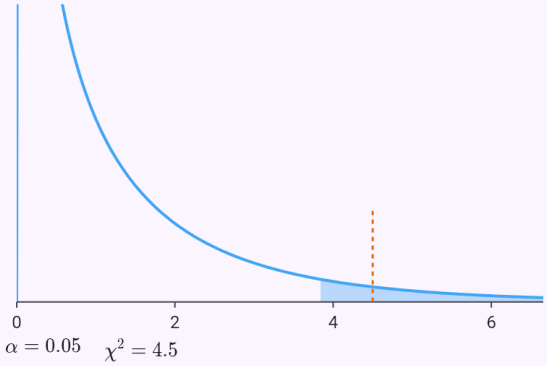
	1	2	∑
1	25 (20)	25 (30)	50
2	20 (25)	20 (15)	40
∑	45	45	90

- The test statistic – under H_0 it has a χ² distribution with 1 degrees of freedom:
$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 1 + 1 + 1.25 + 1.25 = 4.5$$
- The critical value at the level α: $\chi^2_{0.95}(1) \approx 3.841$, the critical region $\chi^2 \geq 3.841$
- The p-value: $p = P(\chi^2 \geq 4.5) \approx 0.03389$
- Since $p \approx 0.03389 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.

- The strength of the association – Cramér's V (0 none, 1 perfect)
$$V = \sqrt{\frac{\chi^2}{N(\min(r,c) - 1)}} = \sqrt{\frac{4.5}{90 \cdot 1}} \approx 0.2236$$
- The χ² test works with counts (numbers of cases), not with percentages or means.

Graph

χ² distribution – the critical region shaded, the test statistic dashed



Two groups by the Mann–Whitney test – small samples, an exact p-value.

Input: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)`

Rank test (Mann–Whitney, Wilcoxon, Kruskal–Wallis)

Result

$$U = 3$$
$$z \approx -2.579$$
$$p \approx 0.009922 \leq \alpha = 0.05$$

Steps

- The null hypothesis: both groups come from the same distribution, the alternative: one tends to have higher values
  $H_0: F_A = F_B$
  $H_1: F_A \neq F_B$

- Both groups' values together sorted and numbered by their rank R (tied values the mean of their ranks)

x	R	
19	1	B
20	2	B
21	3	B
22	4	B
23	5.5	A
23	5.5	B
24	7	B
25	8.5	A
25	8.5	B
26	10	A
27	11	A
28	12	A
30	13	A

- The groups' sizes and rank sums:
  $n_1 = 6, \quad n_2 = 7, \quad R_1 = 60, \quad R_2 = 31$

- The Mann–Whitney U – the number of pairs in which one group's value ranks higher than the other's; the smaller of the two is tested
  $$U_1 = R_1 - \frac{n_1(n_1+1)}{2} = 60 - \frac{6 \cdot 7}{2} = 39$$
  $$U_2 = n_1 n_2 - U_1 = 6 \cdot 7 - 39 = 3$$
  $$U = \min(U_1, U_2) = 3$$

- The test statistic – under H_0 it has the standard normal distribution: $z = \frac{U - \frac{n_1 n_2}{2}}{\sigma_U} = \frac{3 - 21}{6.981} \approx -2.579$

- The critical value at the level α: $z_{0.975} \approx 1.96$, the critical region $|z| \geq 1.96$

- The p-value: $p = 2\,P(Z \geq 2.579) \approx 0.009922$

- Since $p \approx 0.009922 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.

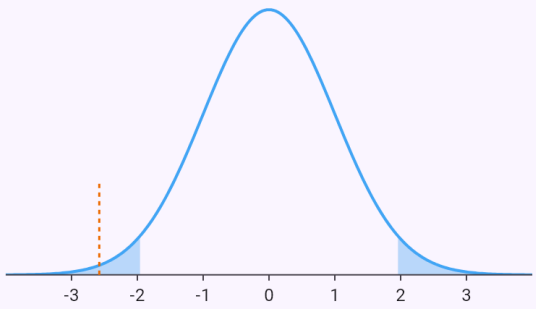
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.

- The effect size r (roughly 0.1 small, 0.3 medium, 0.5 large): $r = \frac{|z|}{\sqrt{N}} = \frac{2.579}{\sqrt{13}} \approx 0.7152$

- The Mann–Whitney test is a rank test: it doesn't assume a normal distribution and suits ordinal data and outliers too. It compares the groups' whole distributions – for similar shapes it is a test of the medians' difference.

Graph

Standard normal distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05 \quad z \approx -2.579$

Paired values before and after by Wilcoxon's test.

Input: `data(pred po; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71; 79 72)`

Rank test (Mann–Whitney, Wilcoxon, Kruskal–Wallis) · Paired

Result

$$T = 1$$
$$z \approx -2.205$$
$$p \approx 0.02743 \leq \alpha = 0.05$$

Steps

- The null hypothesis: the differences' median is zero (no change), the alternative: it isn't: $H_0: \tilde{d} = 0, \quad H_1: \tilde{d} \neq 0$
- The pairs' differences d (1st column − 2nd column): $2,\ 2,\ -1,\ 4,\ 3,\ 3,\ 7$; zero differences left out: $n = 7$
- The non-zero differences sorted by their size (without the sign) and numbered by their rank R; the sign stays with the rank

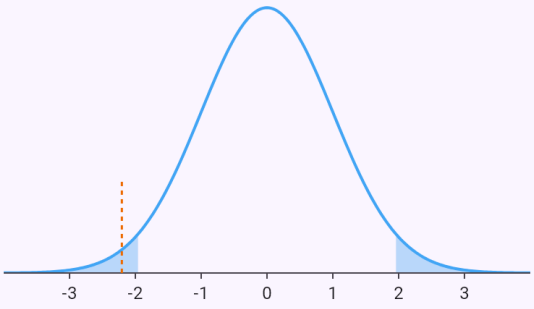
d_i	R	
-1	1	$-$
2	2.5	$+$
2	2.5	$+$
3	4.5	$+$
3	4.5	$+$
4	6	$+$
7	7	$+$

- The rank sums of the positive and the negative differences
  $T^+ = 27$
  $T^- = 1$
  $T = \min\left(T^+, T^-\right) = 1$
- The test statistic – under H_0 it has the standard normal distribution: $z = \dfrac{T - \frac{n(n+1)}{4}}{\sigma_T} = \dfrac{1 - 14}{5.895} \approx -2.205$
- The critical value at the level α: $z_{0.975} \approx 1.96$, the critical region $|z| \geq 1.96$
- The p-value: $p = 2\,P(Z \geq 2.205) \approx 0.02743$
- Since $p \approx 0.02743 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.

- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- The effect size r (roughly 0.1 small, 0.3 medium, 0.5 large): $r = \dfrac{|z|}{\sqrt{n}} = \dfrac{2.205}{\sqrt{7}} \approx 0.8335$
- Wilcoxon's test is the rank counterpart of the paired t-test: it doesn't assume normally distributed differences and takes their signs and sizes into account by their ranks.

Graph

Standard normal distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05 \quad z \approx -2.205$

Three groups by the Kruskal–Wallis test.

Input: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

Rank test (Mann–Whitney, Wilcoxon, Kruskal–Wallis)

Result

$$H \approx 8.263$$
$$p \approx 0.01606 \le \alpha = 0.05$$

Steps

- The null hypothesis: all the groups come from the same distribution, the alternative: at least one differs
  $H_0: F_1 = \cdots = F_3$
  $H_1: \exists\, i, j: F_i \ne F_j$

- All the groups' values together sorted and numbered by their rank R (tied values the mean of their ranks)

x	R	
11	1	A
12	2.5	A
12	2.5	C
13	4.5	A
13	4.5	C
14	6.5	A
14	6.5	C
15	9	A
15	9	B
15	9	C
16	11.5	B
16	11.5	C
17	13	B
18	14	B
19	15	B

- Each group's rank sum
  $R_\mathrm{A} = 23.5\ (n = 5)$
  $R_\mathrm{B} = 62.5\ (n = 5)$
  $R_\mathrm{C} = 34\ (n = 5)$

- The test statistic – under H_0 it has a χ^2 distribution with 2 degrees of freedom
  $$H = \frac{12}{N(N+1)} \sum \frac{R_j^2}{n_j} - 3(N+1) = \frac{12}{15 \cdot 16} \left(\frac{23.5^2}{5} \right.$$
  $$C = 1 - \frac{\sum (t^3 - t)}{N^3 - N} = 1 - \frac{48}{3360} \approx 0.9857$$
  $$H_c = \frac{H}{C} = \frac{8.145}{0.9857} \approx 8.263$$

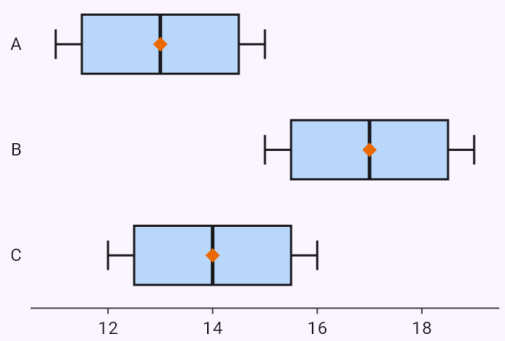
- The critical value at the level α: $\chi^2_{0.95}(2) \approx 5.991$, the critical region $\chi^2 \ge 5.991$
- The p-value: $p = P(\chi^2 \ge 8.263) \approx 0.01606$
- Since $p \approx 0.01606 \le \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.

- The effect size ε² – the share of the ranks' variance explained by the groups:
  $$\varepsilon^2 = \frac{H}{N-1} = \frac{8.263}{14} \approx 0.5902$$

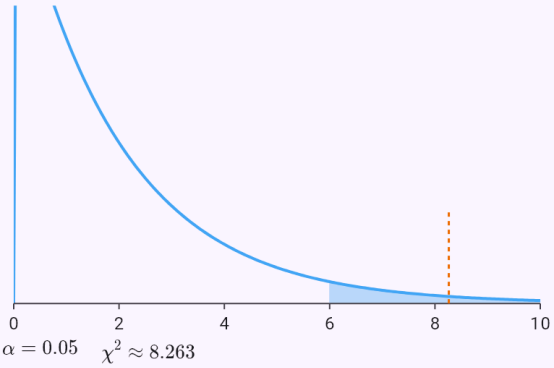
- The Kruskal–Wallis test is the rank counterpart of one-way ANOVA. It only says that some groups differ – which ones, a comparison of the pairs shows (the Mann–Whitney test with Bonferroni's correction).

Graph

Box plot (median, quartiles, outliers; ◆ mean)



χ² distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05$ $\chi^2 \approx 8.263$

Three groups' means by one-way ANOVA.

Input: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

ANOVA (analysis of variance)

Result

$$F \approx 8.667$$
$$p \approx 0.004687 \leq \alpha = 0.05$$
$$\eta^2 \approx 0.5909$$

Steps

- The null hypothesis: all the groups' means are equal, the alternative: at least two differ

  $H_0: \mu_1 = \cdots = \mu_3$

  $H_1: \exists\, i, j: \mu_i \neq \mu_j$

- The groups' sizes, means and standard deviations and the grand mean $\bar{x} \approx 14.67$

	n_j	$\bar{x}_j$	s_j
A	5	13	1.581
B	5	17	1.581
C	5	14	1.581
$\sum$	15	14.67	

- Analysis of variance – the sum of squares between the groups (B) and within them (W), the degrees of freedom and mean squares

  $$SS_B = \sum n_j (\bar{x}_j - \bar{x})^2$$
  $$SS_W = \sum \sum (x_{ij} - \bar{x}_j)^2$$

	SS	df	MS	F
B	43.33	2	21.67	8.667
W	30	12	2.5	
$\sum$	73.33	14		

- The test statistic – under H_0 it has the F distribution with 2 and 12 degrees of freedom:

  $$F = \frac{MS_B}{MS_W} = \frac{21.67}{2.5} \approx 8.667$$

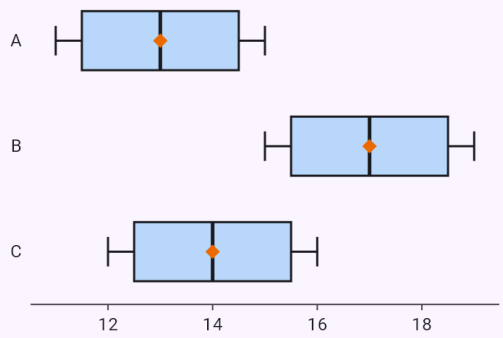
- The critical value at the level α: $F_{0.95}(2;\, 12) \approx 3.885$, the critical region $F \geq 3.885$
- The p-value: $p = P(F \geq 8.667) \approx 0.004687$
- Since $p \approx 0.004687 \leq \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.
- The effect size η² – the share of the variance explained by the differences between the groups (roughly 0.01 small, 0.06 medium, 0.14 large):

  $$\eta^2 = \frac{SS_B}{SS_T} = \frac{43.33}{73.33} \approx 0.5909$$

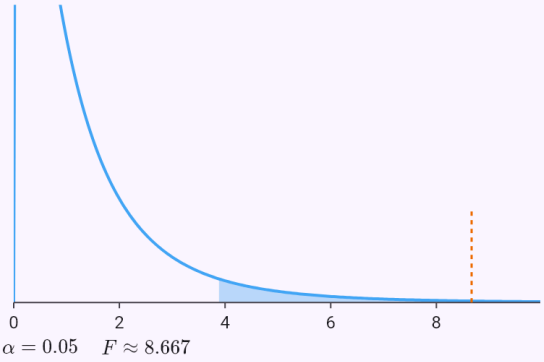
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- ANOVA assumes independent groups, a roughly normal distribution in each of them and similar variances. Where that doesn't hold (small groups, outliers, ordinal data), use the Kruskal–Wallis test.
- Rejecting H_0 only says that at least two means differ. Which ones, a comparison of the pairs shows (post hoc tests, e.g. t-tests with Bonferroni's correction).

Graph

Box plot (median, quartiles, outliers; ◆ mean)



F distribution – the shaded part is the critical region, dashed the test statistic



$\alpha = 0.05$ $F \approx 8.667$

A goodness-of-fit test against the expected shares 1 : 2 : 1.

Input: `data(30 50 20)` χ² test · Expected shares 1:2:1

Result

$$\chi^2 = 2$$
$$p \approx 0.3679 > \alpha = 0.05$$

Steps

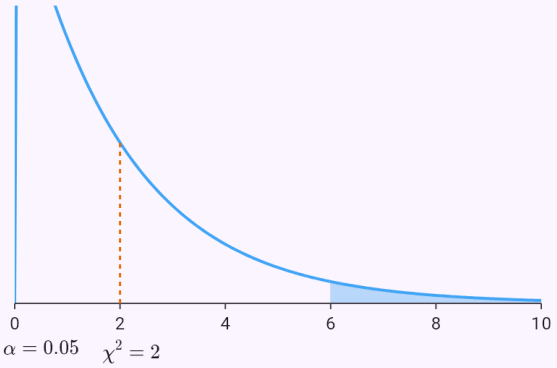
- The null hypothesis: the categories have the given shares, the alternative: at least one share differs: $H_0: p_1 = 0.25,\ p_2 = 0.5,\ p_3 = 0.25$
- The expected counts by the given shares $E_i = N \cdot p_i, \quad N = 100$ and the categories' contributions

i	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
1	30	25	1
2	50	50	0
3	20	25	1
$\sum$	100	100	2

- The test statistic – under H_0 it has a χ² distribution with 2 degrees of freedom: $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 1 + 0 + 1 = 2$
- The critical value at the level α: $\chi^2_{0.95}(2) \approx 5.991$, the critical region $\chi^2 \geq 5.991$
- The p-value: $p = P(\chi^2 \geq 2) \approx 0.3679$
- Since $p \approx 0.3679 > \alpha = 0.05$, we don't reject the null hypothesis – the data don't give enough evidence against it.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- The χ² test works with counts (numbers of cases), not with percentages or means.

Graph

χ² distribution – the critical region shaded, the test statistic dashed



Quadratic regression.

Input: `data(x y; 1 2.1; 2 3.9; 3 9.2; 4 15.8; 5 25.1; 6 36.2)`

Correlation and regression · Quadratic

Result

$$\hat{y} = 1.73 - 0.9354x + 1.116x^2$$

$$R^2 \approx 0.9997$$

Steps

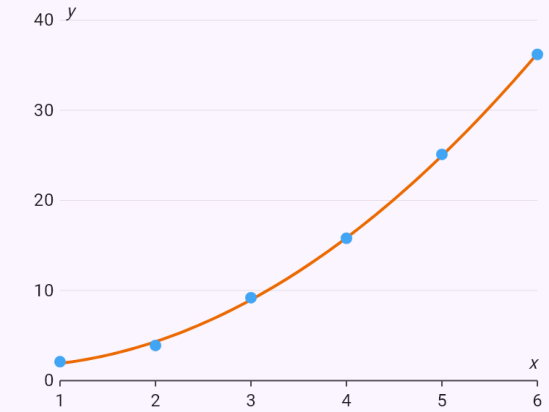
- The count of values $n = 6$, the data sorted
  $[1,\ 2.1],\ [2,\ 3.9],\ [3,\ 9.2],\ [4,\ 15.8],\ [5,\ 25.1],\ [6,\ 36.2]$
- The sums and means
  $$\sum x_i = 21$$
  $$\sum x_i^2 = 91$$
  $$\sum x_i^3 = 441$$
  $$\sum x_i^4 = 2275$$
  $$\sum y_i = 92.3$$
  $$\sum x_i y_i = 443.4$$
  $$\sum x_i^2 y_i = 2284$$
- The least squares normal equations for ŷ = a + bx + cx²
  $$6a + 21b + 91c = 92.3$$
  $$21a + 91b + 441c = 443.4$$
  $$91a + 441b + 2275c = 2284$$
- The regression curve: $\hat{y} = 1.73 - 0.9354x + 1.116x^2$
- The coefficient of determination – the curve explains $99.97\,\%$ of y's variance.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0.3012}{874.5} \approx 0.9997$$

- The regression curve is reasonable only within the range of the measured x – outside it (extrapolation) the estimate may be very inaccurate.

Graph

Scatter plot with the regression curve



Exponential regression by linearization.

Input: `data(x y; 0 2.0; 1 2.7; 2 3.7; 3 5.0; 4 6.8; 5 9.1)` Correlation and regression · Exponential

Result

$$\hat{y} = 2.002\, e^{0.3042x}$$

$$R^2 \approx 0.9998$$

Steps

- The count of values $n = 6$, the data sorted $[0,\ 2],\ [1,\ 2.7],\ [2,\ 3.7],\ [3,\ 5],\ [4,\ 6.8],\ [5,\ 9.1]$
- Taking logarithms turns the model into a line in the variables (u, v): $y = a\, e^{bx} \Rightarrow \ln y = \ln a + bx$; the transformed pairs $[0,\ 0.6931],\ [1,\ 0.9933],\ [2,\ 1.308],\ [3,\ 1.609],\ [4,\ 1.91$
- The regression line's coefficients by least squares

$$b = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{\sum (u_i - \bar{u})^2} = \frac{5.324}{17.5} \approx 0.3042$$

$$\ln a = \overline{\ln y} - b\bar{x} = 1.455 - 0.3042 \cdot 2.5 \approx 0.6943$$

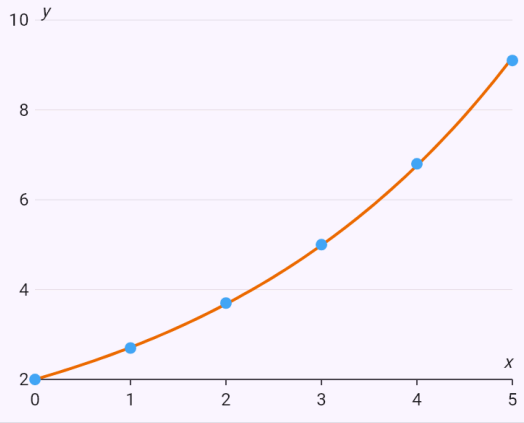
- The coefficient a by the back transformation: $a = e^{0.6943} \approx 2.002$
- The regression curve: $\hat{y} = 2.002\, e^{0.3042x}$
- The coefficient of determination – the curve explains $99.98\ \%$ of y's variance.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0.006563}{35.95} \approx 0.9998$$

- The coefficients are fitted by least squares to the transformed data (logarithms), not to the original values – the result may differ slightly from a nonlinear regression in statistical software.
- The regression curve is reasonable only within the range of the measured x – outside it (extrapolation) the estimate may be very inaccurate.

Graph

Scatter plot with the regression curve



Right-sided t-test from summary values: n = 25, x̄ = 104, s = 15, μ₀ = 100.

Input: `model(ttest; find=test[one]; n=25[one]; mean=104[one]; sd=15[one]; mu0=100[one];`

```
alpha=0.05[one]; alternative=greater; sigma=sample)
```

Result

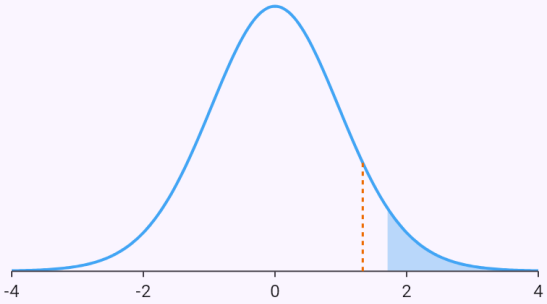
$$t \approx 1.333$$
$$p \approx 0.09747 > \alpha = 0.05$$

Steps

- Find t; given:
  $n = 25, \quad \bar{x} = 104, \quad s = 15, \quad \mu_0 = 100, \quad \alpha = 0.05$
- A t-test checks whether the population mean differs from a given value μ_0. The test statistic $t = (\bar{x} - \mu_0)/(s/\sqrt{n})$ has a t distribution with $n - 1$ degrees of freedom under the null hypothesis; the p-value says how likely such or a more extreme result would be if H_0 held. With a known σ it is a z-test. The model's equation: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
- The null and the alternative hypothesis:
  $H_0: \mu = 100, \quad H_1: \mu > 100$
- The standard error: $SE = \frac{s}{\sqrt{n}} = \frac{15}{\sqrt{25}} = 3$
- The test statistic – under H_0 it has a t distribution with 24 degrees of freedom:
  $t = \frac{\bar{x} - \mu_0}{SE} = \frac{104 - 100}{3} \approx 1.333$
- The critical value at the level α: $t_{0.95}(24) \approx 1.711$, the critical region $t \geq 1.711$
- The p-value: $p = P(T \geq 1.333) \approx 0.09747$
- Since $p \approx 0.09747 > \alpha = 0.05$, we don't reject the null hypothesis – the data don't give enough evidence against it.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- For a small sample ($n < 30$) the t distribution assumes the data come from a roughly normal distribution. With strongly skewed data or outliers a nonparametric test is more suitable.

Graph

t distribution – the critical region shaded, the test statistic dashed



Welch's t-test of two groups from summary values.

Input: `model(twosample; find=test[one]; n1=30[one]; mean1=52[one]; sd1=8[one]; n2=28[one];`

```
mean2=47[one]; sd2=10[one]; alpha=0.05[one]; variances=welch; alternative=two)
```

Result

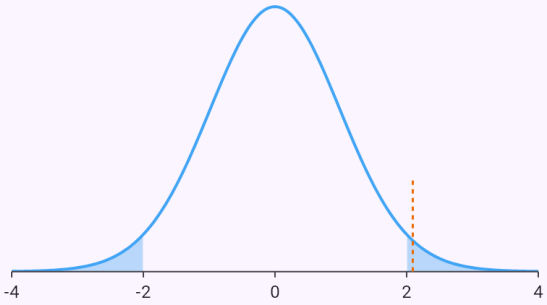
$$t \approx 2.093$$
$$p \approx 0.04124 \le \alpha = 0.05$$

Steps

- Find t; given
  $n_1 = 30, \quad \bar{x}_1 = 52, \quad s_1 = 8, \quad n_2 = 28, \quad \bar{x}_2 = 47,$
- Comparing the means of two independent groups. Welch's test doesn't assume equal variances and is the recommended choice today; the classic t-test with a pooled variance assumes both groups have the same variance. The groups' sizes, means and standard deviations are enough. The model's equation:
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$
- The null and the alternative hypothesis:
  $H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \ne \mu_2$
- The standard error
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{8^2}{30} + \frac{10^2}{28}} \approx 2.388$$
- The Welch–Satterthwaite degrees of freedom
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 51.71$$
- The test statistic – under H_0 it has a t distribution with 51.71 degrees of freedom:
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{52 - 47}{2.388} \approx 2.093$$
- The critical value at the level α: $t_{0.975}(51.71) \approx 2.007$, the critical region $|t| \ge 2.007$
- The p-value: $p = 2\,P(T \ge 2.093) \approx 0.04124$
- Since $p \approx 0.04124 \le \alpha = 0.05$, we reject the null hypothesis – the result is statistically significant.
- The p-value is the probability of a result at least as extreme as the observed one if H_0 holds – not the probability that H_0 holds. Not rejecting H_0 doesn't mean H_0 is true; statistical significance says nothing about the size or practical importance of a difference.
- Welch's test doesn't assume equal variances of the groups; its degrees of freedom ν need not be whole.

Graph

t distribution – the critical region shaded, the test statistic dashed



$\alpha = 0.05 \quad t \approx 2.093$

The normal distribution: P(X < 85) for N(100; 15²).

Input: `model(normal; find=below[one]; mu=100[one]; sigma=15[one]; x=85[one])`

Result

$$P(X < 85) \approx 0.1587 = 15.87\,\%$$

Steps

- Find P; given: $\mu = 100, \quad \sigma = 15, \quad x = 85$
- Many quantities (height, test performance, measurement error) vary symmetrically around the mean μ, the less often the farther from it – the normal (Gaussian) distribution N(μ; σ²) with the standard deviation σ. Probabilities are computed by standardizing, z = (x − μ)/σ, and the distribution function Φ of the standard normal N(0; 1): P(X < x) = Φ(z). A quantile is the value below which a given share p lies. The model's equation

$$X \sim N(\mu,\, \sigma^2)$$

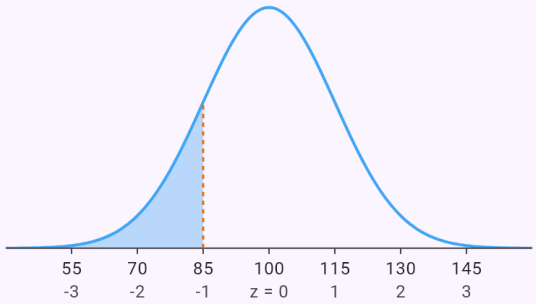
$$Z = \frac{X - \mu}{\sigma} \sim N(0,\, 1)$$

- Expressing the sought quantity: $z = \dfrac{x - \mu}{\sigma}$
- Substituting: $z = \dfrac{85 - 100}{15} = -1$
- Substituting

$$P(X < 85) = \Phi(-1) = (1 - \Phi(1)) \approx (1 - 0.8413) \approx (0$$

Graph

Normal density, the shaded area is the probability sought



$P \approx 0.1587 \quad x = 85$

The share of people between 160 and 185 cm tall with μ = 170, σ = 8.

Input: `model(normal; find=between[one]; mu=170[one]; sigma=8[one]; a=160[one]; b=185[one])`

Result

$$P(160 < X < 185) \approx 0.864 = 86.4\,\%$$

Steps

- Find P; given: $\mu = 170, \quad \sigma = 8, \quad a = 160, \quad b = 185$
- Many quantities (height, test performance, measurement error) vary symmetrically around the mean μ, the less often the farther from it – the normal (Gaussian) distribution N(μ; σ²) with the standard deviation σ. Probabilities are computed by standardizing, z = (x − μ)/σ, and the distribution function Φ of the standard normal N(0; 1): P(X < x) = Φ(z). A quantile is the value below which a given share p lies. The model's equation

$$X \sim N(\mu,\, \sigma^2)$$

$$Z = \frac{X - \mu}{\sigma} \sim N(0,\, 1)$$

- Expressing the sought quantity:

$$z_a = \frac{a - \mu}{\sigma}, \quad z_b = \frac{b - \mu}{\sigma}$$

- Substituting:

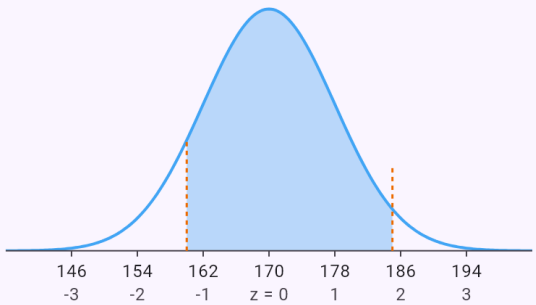
$$z_a = \frac{160 - 170}{8} = -1.25, \quad z_b = \frac{185 - 170}{8} = 1.875$$

- Substituting

$$P(160 < X < 185) = \Phi(1.875) - \Phi(-1.25) \approx 0.9696$$

Graph

Normal density, the shaded area is the probability sought



$P \approx 0.864 \quad a \quad b$

A quantile: the score below which 95 % of the population lie.

Input: `model(normal; find=quantile[one]; mu=100[one]; sigma=15[one]; p=95[pct])`

Result

$x \approx 124.7$

Steps

- Find x; given: $\mu = 100, \quad \sigma = 15, \quad p = 95\,\%$
- Converting the units: $p = 95\,\% = 0.95$
- Many quantities (height, test performance, measurement error) vary symmetrically around the mean μ, the less often the farther from it – the normal (Gaussian) distribution N(μ; σ²) with the standard deviation σ. Probabilities are computed by standardizing, z = (x − μ)/σ, and the distribution function Φ of the standard normal N(0; 1): P(X < x) = Φ(z). A quantile is the value below which a given share p lies. The model's equation

$$X \sim N(\mu,\, \sigma^2)$$

$$Z = \frac{X - \mu}{\sigma} \sim N(0,\, 1)$$

- Expressing the sought quantity

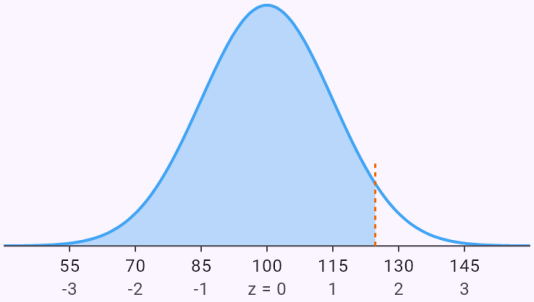
$$P(X < x) = \Phi\left(\frac{x - \mu}{\sigma}\right) = p \Rightarrow x = \mu + z_p\sigma, \quad z_p = $$

- Substituting

$$z_{0.95} = \Phi^{-1}(0.95) \approx 1.645, \quad x = 100 + 1.645 \cdot 15 \approx 1$$

Graph

Normal density, the shaded area is the probability sought



$p = 0.95 \quad x \approx 124.7$

Known limitations

- The quartiles are the medians of the halves (the school method); statistical software (Excel, SPSS) uses other formulas and may get slightly different values for small data.
- Comparing pairs of groups after ANOVA or the Kruskal–Wallis test (post hoc) and multi-way ANOVA aren't available yet – compare the pairs by a t-test or the Mann–Whitney test with Bonferroni's correction; nonlinear regressions are computed by linearization.
- Data are entered by hand, the grid or a CSV file only – the app doesn't recognize a table from a photo yet.

Generated from the Mathtrace app