Statistics by Hand

An Introductory Course for Psychologists

t-test

Version 3.0

Reminder

- Between-subjects and within-subjects tests.
- Sample and population.
- Wilcoxon matched-pairs test (within-Ss)
- Variance and Standard deviation

Reminder

- Variance
- Standard deviation

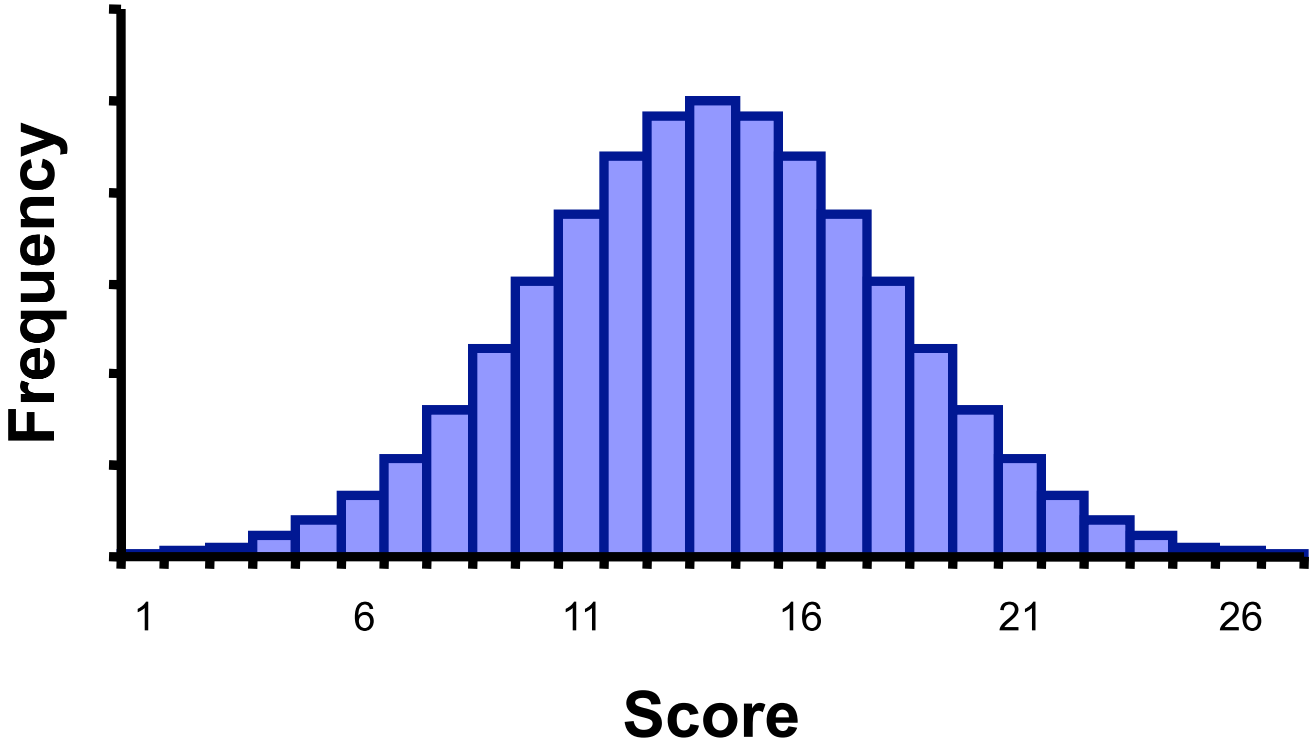
$$s^2 = \frac{\sum \left(X - \overline{X}\right)^2}{N - 1}$$

Variance

$$s = \sqrt{\frac{\sum \left(X - \overline{X}\right)^2}{N - 1}}$$

Standard deviation

"Normal" distribution



- Many types of data are *normally distributed.*
- Height, Astronomical observations, IQ

Z-test

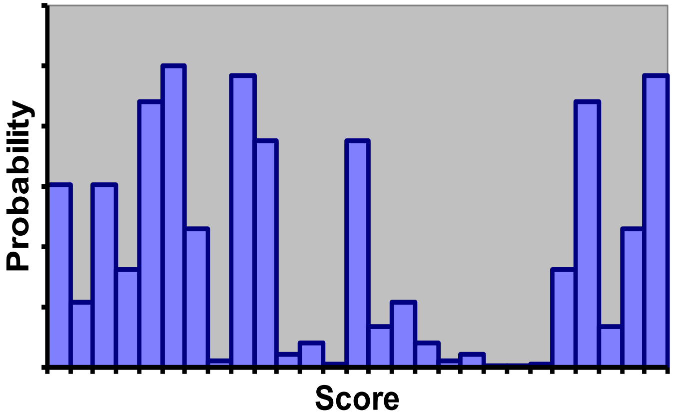
$$z = \frac{X - \mu}{\sigma} = \frac{85 - 100}{15} = -1$$

- Z-test
  - Probability of a score *at least* as high as X coming from a normal distribution with known mean and standard deviation.

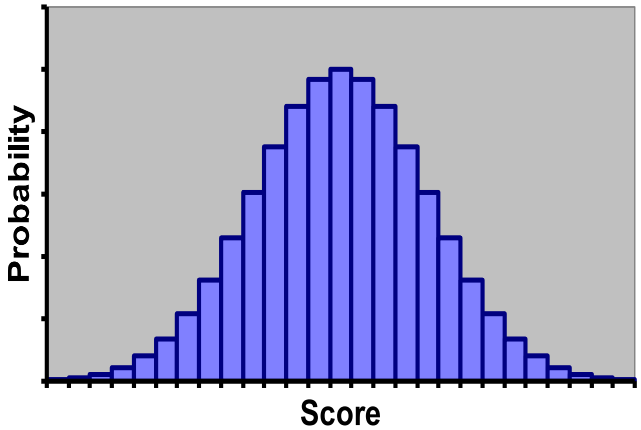
Related samples t-test

- Does the same basic job as a Wilcoxon matched-pairs:
  - Within-subjects comparison of means.
- Is more *powerful* than Wilcoxon but...
- makes some *assumptions* which can be hard to test.
- It's based on the *central limit theorem*.

Central limit theorem - Basic idea



- Start with a population.
- It can have any distribution.
- Take many samples, each of size *N*
- Take the *mean* of each sample.
- The distribution of those means **always** approaches normal for as *N* increases.



Central limit theorem – In full

- Part 1: If there's no difference between the groups in the population as a whole (the null hypothesis), then the mean difference in a sample will, on average, be zero.
- Part 2: The standard deviation of these mean differences in the population can be estimated as $s / \sqrt{N}$
  - s = Standard deviation of the sample
  - N = sample size
- Part 3: The distribution of these mean differences is near-normal if N is large (regardless of how the population itself is distributed).

Applying CLT

Auditory RT	230	300	250	240	330	260	280	250	300	270
Visual RT	260	230	220	250	220	260	210	310	270	240

Auditory RT	320	240	320	330	280	290	280	310	330	320
Visual RT	270	260	210	390	200	290	310	280	170	260

- Reaction time to auditory and visual warning signals (within-subjects design)
- Part 1: If there's no difference between the groups in the population as a whole (the null hypothesis), then the mean difference in a sample will, on average, be zero.

Applying CLT

Auditory RT	230	300	250	240	330	260	280	250	300	270
Visual RT	260	230	220	250	220	260	210	310	270	240
	-30	70	30	-10	110	0	70	-60	30	30
Auditory RT	320	240	320	330	280	290	280	310	330	320
Visual RT	270	260	210	390	200	290	310	280	170	260
	50	-20	110	-60	80	0	-30	30	160	60

- Take differences as before
- Calculate the mean (do not ignore sign)
- Mean difference = 31. **Significant?**

Applying CLT via Z-test

Part 3: The distribution of these mean differences is near-normal if N is large (regardless of how the population itself is distributed).

• Assuming N is large enough, Part 3 allows use of the Z-test (which assumes normality)

$$Z = (X - \mu) / \sigma$$

Part 1: If there's no difference between the groups in the population as a whole (the null hypothesis), then the mean difference in a sample will, on average, be zero.

$$Z = (X - 0) / \sigma$$

Applying CLT via Z-test

$$Z = (X - 0) / \sigma$$

– Part 2: The standard deviation of these mean differences in the population can be estimated as
  s / √N
  - s = Standard deviation of the sample
  - N = sample size

$$Z = X \;/\; (s / \sqrt{N})$$

Calculation

Difference	X - mean	(X - mean)2
-30	-61	3721
70	39	1521
30	-1	1
-10	-41	1681
110	79	6241
0	-31	961
70	39	1521
-60	-91	8281
30	-1	1
30	-1	1
50	19	361
-20	-51	2601
110	79	6241
-60	-91	8281
80	49	2401
0	-31	961
-30	-61	3721
30	-1	1
160	129	16641

Mean	31
Sum (X - mean)2	65980
N-1	19
Variance	3472.6316
Std. Dev.	58.929039

$$Z = X / (s / \sqrt{N})$$

$$= 31 / (58.9 / \sqrt{20})$$

$$= 2.35$$

$p < 0.05$

Experiment worked.

Sample bias problem

Part 2: The standard deviation of these mean differences in the population can be estimated as $s / \sqrt{N}$ s = Standard deviation of the sample, N = sample size.

- Estimates of s.d. from a sample are slightly biased, in the sense that most of the time our estimate will be lower than the true (population) value.
- William Gossett created a modified Z table that corrects for this problem.
- $t = 2.35$, d.f. = $N - 1$ $p < 0.05$.

Related samples *t*-test

- 1. Calculate the difference between each pair.

Auditory RT	230	300	250	240	330	260	280	250	300	270
Visual RT	260	230	220	250	220	260	210	310	270	240
	-30	70	30	-10	110	0	70	-60	30	30
Auditory RT	320	240	320	330	280	290	280	310	330	320
Visual RT	270	260	210	390	200	290	310	280	170	260
	50	-20	110	-60	80	0	-30	30	160	60

- 2. Calculate the mean difference

Difference	X - mean	(X - mean)2
-30	-61	3721
70	39	1521
30	-1	1
-10	-41	1681
110	79	6241
0	-31	961
70	39	1521
-60	-91	8281
30	-1	1
30	-1	1
50	19	361
-20	-51	2601
110	79	6241
-60	-91	8281
80	49	2401
0	-31	961
-30	-61	3721
30	-1	1
160	129	16641

Mean	31
Sum (X - mean)2	65980
N-1	19
Variance	3472.6316
Std. Dev.	58.929039

3. Calculate the standard deviation of the differences.
4. Calculate the standard error:

standard error = s / √N

= 58.93 / √20 = 13.2

Related samples *t*-test

- 5. Divide the mean difference by the standard error.

  t = 31 / 13.2 = 2.35

- 6. Calculate d.f. (= N-1)
- 7. If *t exceeds* the appropriate value in the table then the result is significant.

t-test or Wilcoxon?

- Use a t-test if...
  - N is large (more than 30)
    - (because then CLT part 3 will be correct)
  - or, you know the population is roughly normal (symmetrical with only one peak)
    - Use histogram to assess
    - N.B.: histogram can't be done for N<10
  - Otherwise..
    - Use a Wilcoxon

Problems

- *T-1*
- *T-2*