Comparing Two Groups

The Sign Test

Example: Are there physiological indicators of schizophrenia?

This example was taken from F. Ramsey and D. Schafer, *The Statistical Sleuth*, 2nd edition, Case 2.1.2.[1]

In a 1990 study, researchers used MRI to measure the volumes of various regions of the brain. Measurements were taken on 15 monozygotic twins, where one twin was affected by schizophrenia ("affected") and the other not ("unaffect"). The data we will consider are the volumes of the left hippocampus region in cubic centimetres. You can see the data in Table 1.

Is there evidence that the volume of the left hippocampus is different for people suffering from schizophrenia compared to people who are not?

Pair of twins	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Unaffect	1.94	1.44	1.56	1.58	2.06	1.66	1.75	1.77	1.78	1.92	1.25	1.93	2.04	1.62	2.08
Affected	1.27	1.63	1.47	1.39	1.93	1.26	1.71	1.67	1.28	1.85	1.02	1.34	2.02	1.59	1.97
Difference, d	0.67	−0.19	0.09	0.19	0.13	0.40	0.04	0.10	0.50	0.07	0.23	0.59	0.02	0.03	0.11

Table 1: Left hippocampus volumes for 15 pairs of monozygotic twins where one twin is affected by schizophrenia and the other is not.

Note that these data are an example of matched pairs.

Approach 1: Use a paired t-test.

In a paired t-test, each pair is treated as one observation and the analysis is carried out on the differences.

Let d be the difference in the left hippocampus volume between the unaffected and affected twin in cubic cm.

$$\overline{x}_d = 0.1987, \quad SD(d) = 0.2383, \quad n = 15$$

Let μ_d be the mean for the population of differences in hippocampus volume.

We can test $H_0 : \mu_d = 0$ versus $H_A : \mu_d \neq 0$.

The test statistic is

$$t^* = \frac{\overline{x}_d - 0}{s_d/\sqrt{n}} = \frac{0.1987}{0.2383/\sqrt{15}} = 3.23$$

[1]F. Ramsey and D. Schafer, *The Statistical Sleuth: A Course in Methods of Data Analysis*, 2nd edition, Duxbury, 2002. The data are freely available in the R library `Sleuth2` (see `https://cran.r-project.org/web/packages/Sleuth2/Sleuth2.pdf`).

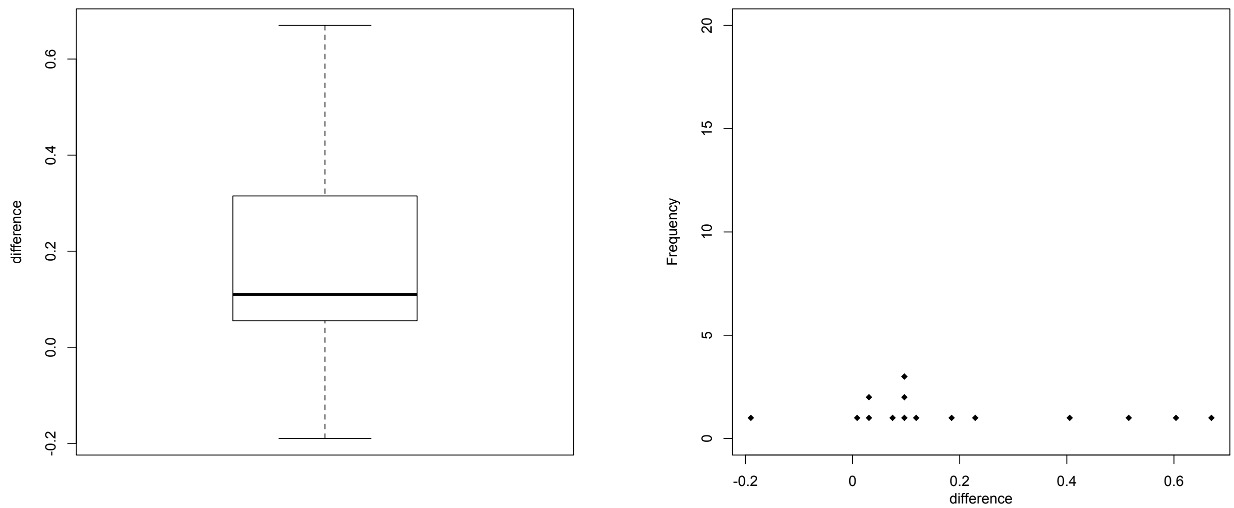

Figure 1: Boxplot and dotplot of difference between unaffected and affected twin in left hippocampus volumes.

We calculate the P-value from a t distribution with 14 degrees of freedom. The P-value is 0.006. (From a t-table, we can estimate that P-value < 0.01.)

We conclude that we have strong evidence that the mean of the differences in left hippocampus volumes between twins with and without schizophrenia is not zero.

A 99% confidence interval for the difference is:

$$\overline{x} \pm t_{0.005} \frac{s_d}{\sqrt{n}} = 0.1986 \pm 2.98 \times \frac{0.2382}{\sqrt{15}} = (0.015,\ 0.382)$$

Note that the confidence interval does not include zero which is consistent with the statistical test.

For a paired t-test to be appropriate, certain conditions must hold. In particular:

1. Each pair of observations is independent of each other.
2. The differences follow a normal distribution in the population or the sample size is sufficiently large so that the mean of the differences is approximately normally distributed.

Figure 1 gives a boxplot and a dotplot of the differences in left hippocampus volumes between the unaffected and affected twins. Although there are only 15 observations, there is some indication that the distribution is right-skewed and thus we may not be confident that the second condition for the paired t-test is satisfied.

Approach 2: The Sign Test

The sign test is an example of a "non-parametric" procedure. Non-parametric procedures don't rely on the data or the test statistic having (or even approximately having) a particular distribution. In particular, normality is not required and the test is not affected by outliers.

Disadvantages of using non-parametric procedures:

1. Non-parametric methods are generally less powerful than parametric methods such as the t-test if the assumptions of the t-test hold.

2. The statement of the hypotheses must be changed to apply a non-parametric test (e.g., instead of testing mean, might be testing the median).

3. For non-parametric methods, there is often not a corresponding estimate of the size of the differences, such as a confidence interval.

The sign test is a non-parametric test about the centre of a distribution. It is used as an alternative to the one-sample t-test (including the matched pairs t-test).

We will now apply the sign test to the schizophrenic twins example:
Let $p =$ the probability that the unaffected twin has a larger left hippocampus volume than the affected twin.
The null hypothesis of "no effect" or "no difference" is equivalent to $p = 1/2$.
So we test $H_0: \ p = 1/2$ versus $H_A: \ p \neq 1/2$.
For our 15 pairs of twins, in 14 the unaffected twin had a larger left hippocampus volume (and smaller for 1 pair). That is, we have 14 positive differences and 1 negative difference.
Our estimate of p is $\hat{p} = 14/15$.
Let $X =$ the number of pairs (twins) for which the unaffected twin had a larger left hippocampus volume. X is the number of twins for which $d > 0$. X is our test statistic.
When H_0 is true, X has a Binomial(15, 0.5) distribution.

The P-value is the probability of observing the data we observed, or a value more extreme. Because we are carrying out a two-sided test, this is the probability of observing few or many positive differences. We have observed one negative and 14 positive differences. More extreme than this would be 15 positive differences. The other extreme would be observing 14 negative differences and one positive difference or 15 negative differences. So our P-value is

$$\text{P-value} = P(X \geq 14) + P(X \leq 1) = P(X = 14) + P(X = 15) + P(X = 1) + P(X = 0)$$

where $X \sim \text{Binomial}(15, 0.5)$.

Figure 2 gives the probabilities for a Binomial(15, 0.5) distribution. We can see that the probability of observing 0, 1, 14, or 15 is very small. Calculating the P-value gives 0.00098. Since our P-value is very small we have strong evidence against the null hypothesis. That

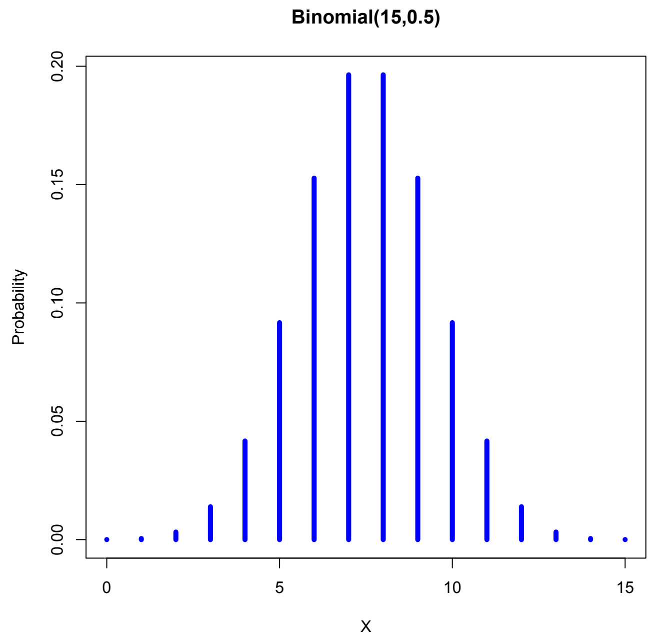

Figure 2: Probabilities for the Binomial(15, 0.5) distribution.

is, we have strong evidence that it is not the case that both members of a set of twins are equally likely to have the larger left hippocampus.

Some notes on the sign test:

1. The hypotheses for a sign test can be written in terms of the median of the distribution. Let $\tilde{\mu}$ be the median of the differences. Then the null hypothesis of "no difference" is $H_0: \ \tilde{\mu} = 0$.
   Why is this equivalent to $H_0 : \ p = 0.5$? Because when $p = 0.5$ we expect half of the differences to be positive and half to be positive, and then the median of the distribution of the differences is 0.

   So for our twins example, we can conclude that the median of the differences in the left hippocampus between the schizophrenic and unaffected twin is not 0.

2. When carrying out the sign test, any differences of 0, if they occur, are ignored and then n is the number of non-zero differences.