Non-Parametric Single Sample Tests for the Median

Non-Parametric Statistics

Previously: used methods based on ***normal distribution***

Now: methods that work more generally, without any assumption on the random variable X.

Two basic concepts:

- ***non-parametric statistics*** do not assume the dependence on any parameter.

18.1. Example. The confidence interval for the mean derived previously has the form

$$\overline{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \qquad \text{or} \qquad \overline{X} \pm t_{\alpha/2,n-1} \frac{S}{\sqrt{n}},$$

which uses the parameters $z_{\alpha/2}$ and σ (or $t_{\alpha/2,n-1}$).

In contrast, in the assignments we have studied a non-parametric confidence interval for the median which does not use any parameter that is not directly derived from the random sample.

Non-Parametric Statistics

Two basic concepts:

- ***non-parametric statistics*** do not assume the dependence on any parameter.
- ***distribution-free statistics*** do not assume that X follows any particular distribution (such as the normal distribution).

Although different, both types of methods are loosely referred to as ***non-parametric methods***.

Generally, one uses

- the ***median or other location measure*** instead of the mean;
- the ***interquartile range or other dispersion measure*** instead of the variance.

Sign Test for the Median

Recall that the median of a random variable X is defined as the value M such that

$$P[X \leq M] = P[X \leq M] = 1/2.$$

The ***sign test*** will have a null hypothesis of either the two-tailed or one-tailed form

- $H_0 \colon M = M_0$
- $H_0 \colon M \leq M_0$ or $H_0 \colon M \geq M_0$

and is usually implemented as a ***Fisher test***.

Sign Test for the Median

The idea is simple: Given a random sample $X_1, \dots, X_n$ of size n from X, each measurement has a $1/2$ probability of being smaller than M and a $1/2$ probability of being larger than M.

(We neglect for now the possibility of $X_k = M$.)

If significantly less than one-half of the sample measurements is less than or greater than M_0, this may be taken as evidence to reject H_0.

Given a sample $X_1, \dots, X_n$, define

$$Q_+ = \#\{X_k \colon X_k - M_0 > 0\}, \qquad Q_- = \#\{X_k \colon X_k - M_0 < 0\}.$$

So Q_+ is the number of "positive signs" and Q_- the number of "negative signs." We note that

$$P[Q_- \le k \mid M = M_0] = \sum_{x=0}^{k} \binom{n}{x} \frac{1}{2^n}$$

Sign Test for the Median

18.2. Sign Test. Let $X_1, \dots, X_n$ be a random sample of size n from an arbitrary continuous distribution and let

$$Q_+ = \#\{X_k : X_k - M_0 > 0\}, \qquad Q_- = \#\{X_k : X_k - M_0 < 0\}.$$

We reject at significance level α

- $H_0 \colon M \le M_0$ if $P[Q_- < k \mid M = M_0] < \alpha$,
- $H_0 \colon M \ge M_0$ if $P[Q_+ < k \mid M = M_0] < \alpha$,
- $H_0 \colon M = M_0$ if $P[\min(Q_-, Q_+) < k \mid M = M_0] < \alpha/2$.

Sign Test for the Median

18.3. Example. A certain six-sided die is suspected of being unbalanced. Based on past experience, it is suspected that the median is greater than 3.5. We decide to test the null hypothesis

$$H_0 : M \le 3.5.$$

The die is rolled 20 times, yielding the following results:

X_i	$X_i - M_0$	Sign	X_i	$X_i - M_0$	Sign	X_i	$X_i - M_0$	Sign
5	1.5	$+$	3	-0.5	$-$	4	0.5	$+$
1	-2.5	$-$	6	2.5	$+$	4	0.5	$+$
5	1.5	$+$	2	-1.5	$-$	4	0.5	$+$
4	0.5	$+$	3	-0.5	$-$	3	-0.5	$-$
4	0.5	$+$	5	1.5	$+$	3	-0.5	$-$
6	2.5	$+$	5	1.5	$+$	4	0.5	$+$
6	2.5	$+$	6	2.5	$+$			

Sign Test for the Median

We note that there are 6 negative signs,

$$Q_- = 6.$$

We then find that

$$P[Q_- \leq 6 \mid M = 3.5] = \frac{1}{2^{20}} \sum_{x=0}^{6} \binom{20}{x} = 0.0577.$$

This is the P-value of the test. It would be reasonable to decide not to reject H_0, i.e., the results do not provide convincing evidence that H_0 is false.

Assumptions, Limitations and Issues

Advantages:

- Very flexible, no assumptions on distribution of X.
- Magnitude of $X_i - M_0$ is not needed.

Disadvantages:

- Not very powerful.

Possible Issues:

- In some situations, especially when sampling from a discrete distribution, it may happen that

$$X_i - M_0 = 0.$$

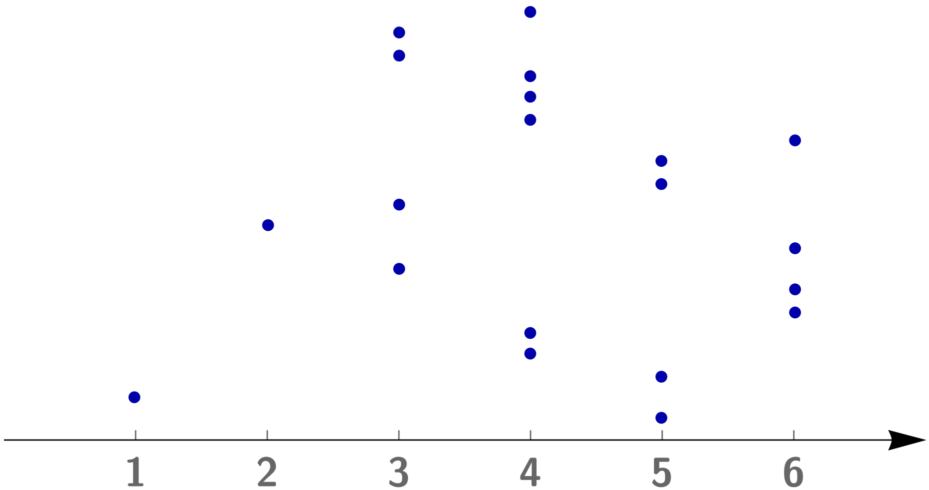
In such case, usual practice is to ***exclude the data from the analysis***.

Wilcoxon Signed Rank Test

The power of the sign test can be increased by taking the magnitude of $X_i - M_0$ into account.

In order to avoid using parameters, Wilxcoxon introduced the notion of ***ranks***: observations are ranked from smallest to largest and instead of considering simply their sign, one analyzes the ***signed rank***.

Frank Wilcoxon (1892-1965). A R Sampson and B Spencer, A conversation with I. Richard Savage, Statistical Science 14 (1999), 126-148.



Wilcoxon Signed Rank Test

This analysis of ranks supposes that the data comes from a distribution that is ***symmetric about its median***. This assumption was not needed for the sign test.

The data is ranked from smallest to largest absolute difference to the null value of the median. In other words, the observation where $|X_i - M_0|$ is smallest will be ranked first and be assigned rank $R_i = 1$, while the observation where $|X_j - M_0|$ is largest will receive rank $R_j = n$.

The signed rank is found by multiplying the rank with -1 if $X_i - M_0 < 0$ and $+1$ if $X_i - M) > 1$.

The positive ranks as well as the negative ranks are summed separately, yielding two statistics W_+ and W_-.

Ties in ranks are assigned the ***average of their ranks***. Hence, the total sum of the ranks is always $n(n+1)/2$.

Test Tables and Normal Approximation

The distribution of the test statistics is complicated; there are tables that give critical values for small sample sizes, typically up to $n \le 20$.

For non-small sample sizes ($n \ge 10$) a normal distribution with parameters

$$\mathrm{E}[W] = \frac{n(n+1)}{4}, \qquad \mathrm{Var}[W] = \frac{n(n+1)(2n+1)}{24}.$$

may be used as an approximation. However, in that case the variance needs to be reduced if there are ties: for each group of t ties, the variance is reduced by $(t^3 - t)/48$.

Wilcoxon Signed Rank Test

18.4. Wilcoxon Signed Rank Test. Let $X_1, \ldots, X_n$ be a random sample of size n from a symmetric distribution. Order the n absolute differences $|X_i - M|$ according to magnitude, so that $X_{R_i} - M_0$ is the R_ith smallest difference by modulus. If ties in the rank occur, the mean of the ranks is assigned to all equal values.

Let

$$W_+ = \sum_{R_i > 0} R_i, \qquad |W_-| = \sum_{R_i < 0} |R_i|.$$

We reject at significance level α

- $H_0 \colon M \le M_0$ if W_- is smaller than the critical value for α,
- $H_0 \colon M \ge M_0$ if W_+ is smaller than the critical value for α,
- $H_0 \colon M = M_0$ if $W = \min(W_+, |W_-|)$ is smaller than the critical value for $\alpha/2$.

Ranking the Results of the Die Rolls

18.5. Example. Returning to the previous Example 18.3, we want to test $H_0\colon M \le 3.5$ and have the following observations, ordered from smallest to largest:

X_i	$X_i - M_0$	R_i		X_i	$X_i - M_0$	R_i	
3	-0.5	-5.5	1	2	-1.5	-13	11
3	-0.5	-5.5	2	5	1.5	$+13$	12
3	-0.5	-5.5	3	5	1.5	$+13$	13
3	-0.5	-5.5	4	5	1.5	$+13$	14
4	0.5	$+5.5$	5	5	1.5	$+13$	15
4	0.5	$+5.5$	6	1	-2.5	-18	16
4	0.5	$+5.5$	7	6	2.5	$+18$	17
4	0.5	$+5.5$	8	6	2.5	$+18$	18
4	0.5	$+5.5$	9	6	2.5	$+18$	19
4	0.5	$+5.5$	10	6	2.5	$+18$	20

$\frac{(10+1)}{2}$

$\frac{11+15}{2}$

$\frac{16+20}{2}$

Finding the *P*-Value

We calculate the sum of the negative ranks,

total rank $\frac{20\times 21}{2} = 210$
w_- 应当比较接近 -105

$$w_- = -5.5 - 5.5 - 5.5 - 5.5 - 13 - 18 = -53.$$

Consulting a table, the critical value for $n = 20$ and $\alpha = 0.05$ is 60. For $\alpha = 0.01$ it is 43. Since $|w_-|$ lies between these values, the P-value of the test is between 1% and than 5%, most likely around 2%-3%.

Alternatively, we may use the normal distribution with mean $\mu = n(n+1)/4 = 105$ and variance

$$\sigma^2 = \frac{n(n+1)(2n+1)}{24} - \frac{10^3 - 10}{48} - 2 \cdot \frac{5^3 - 5}{48}.$$

Then

$$z = \frac{w_- - \mu}{\sigma} = -1.977$$

and we find that $P[Z < -1.977] = 0.024$.

Conclusion of the Test

It may reasonable to reject H_0 based on this P-value. There is some evidence that the die does not follow a symmetric distribution with median less than or equal to 3.5.

In practice, we may come to several conclusions:

- the die results follow a non-symmetric distribution; or
- the die results follow a symmetric distribution, but the median is greater than 3.5.

This example features many ties between results. In general, the power if the signed rank test is reduced if there are very many ties and some modified tests have been recently proposed to improve this.

Nevertheless, we obtained a result with a smaller P-value than we did for the Sign Test.

Assumptions, Limitations and Issues

Advantages:

- Fairly powerful; may even be used as an alternative to the T-test without much loss of power.

Disadvantages:

- Assumes a symmetric distribution around the median.

Possible Issues:

- As in the sign test, observations where $X_i - M_0 = 0$ are discarded.
- Some authors prefer to use a modified but equivalent version of the test, where all positive and negative ranks are added together. The test is equivalent, but different tables need to be used.

Hypothesis Tests with Mathematica

We can use Mathematica for calculating test statistics. Suppose we have the following data:

```
data := {41.50, 41.38, 42.24, 41.85, 41.76,
  42.08, 41.62, 42.16, 41.71, 41.44}
```

We want to test $H_0: \mu \le \mu_0 = 41.5$, assuming a known variance of $\sigma^2 = 0.1$. The Z-test statistic is

```
x̄ := Mean[data]; n := Length[data]; σ0 := √0.1 ; μ0 := 41.5;
```

$$Z = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}}$$

```
2.74
```

We can then find a P-value for the test:

```
1 - CDF[NormalDistribution[0, 1], Z]

0.00307196
```

Hypothesis Tests with Mathematica

Mathematica also has many standard tests built-in. The previous test can be performed as follows:

```
Needs["HypothesisTesting`"]

ZTest[data, 0.1, 41.5, "TestDataTable",
 AlternativeHypothesis -> "Greater"]
```

	Statistic	P–Value
Z	2.74	0.00307196

The corresponding two-tailed test $H_0: \mu = \mu_0$ would yield:

```
ZTest[data, 0.1, 41.5, "TestDataTable",
 AlternativeHypothesis -> "Unequal"]
```

	Statistic	P–Value
Z	2.74	0.00614392

Hypothesis Tests with Mathematica

Of course, there are also T-tests:

```
TTest[data, 41.5, "TestDataTable",
 AlternativeHypothesis -> "Unequal"]
```

	Statistic	P-Value
T	2.8439	0.0192801

The sign test and the Wilcoxon signed rank test are also implemented:

```
SignTest[data, 41.5, "TestDataTable",
 AlternativeHypothesis -> "Unequal"]
```

	Statistic	P-Value
Sign	7	0.179687

```
SignedRankTest[data, 41.5, "TestDataTable",
 AlternativeHypothesis -> "Unequal"]
```

	Statistic	P-Value
Signed-Rank	41.5	0.028263

Hypothesis Tests with Mathematica

The chi-squared test is called the Fisher ratio test in Mathematica:

```
FisherRatioTest[data, 0.1, "TestDataTable",
 AlternativeHypothesis -> "Unequal"]
```

	Statistic	P-Value
Fisher Ratio	8.3544	0.997726

We can verify the result by hand:

$$\chi := \sqrt{\frac{(\mathbf{n}-\mathbf{1})\ \mathbf{Variance}[\mathbf{data}]}{\sigma_0{}^2}}\ ;\ \chi^2$$

8.3544

$$\mathbf{2}\ \left(\mathbf{1} - \mathbf{CDF}\left[\mathbf{ChiSquareDistribution}[\mathbf{n}-\mathbf{1}],\ \chi^2\right]\right)$$

0.997726

Note the behavior of the cumulative distribution function for the chi-squared distribution and the doubling of the *P*-value!