Introduction to Statistics and Data Science using *eStat*

Chapter 10 Nonparametric Testing Hypothesis

10.1 Nonparametric Test for Location of Single Population

10.1.1 Sign Test

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

10.1 Nonparametric Test for Location of Single Population
- 10.1.1 Sign Test
- 10.1.2 Wilcoxon Signed Rank Sum Test

10.2 Nonparametric Test for Comparing Locations of Two Populations
- 10.2.1 Independent Samples: Wilcoxon Rank Sum Test
- 10.2.2 Paired Samples: Wilcoxon Signed Rank Sum Test

10.3 Nonparametric Test for Comparing Locations of Several Populations
- 10.3.1 Completely Randomized Design: Kruskal-Wallis Test
- 10.3.2 Randomized block design: Friedman Test

10.1 Nonparametric Test for Location Parameter of Single Population

- Make some assumptions about a population distribution and test a population parameter
  ⇨ parametric test.

- Real data may not be appropriate to assume a normal population.
  ⇨ Data may not be continuous but ordinal such as rank.
  ⇨ parametric tests are not appropriate.

- Test by converting data into signs or ranks
  ⇨ distribution-free or nonparametric test
  ⇨ there may be some loss of information about the data.

- If a population can be assumed as a normal distribution, there is no reason to use the nonparametric tests.

10.1 Nonparametric Test for Location Parameter of Single Population

10.1.1 Sign Test

[Example 10.1.1] A bag of cookies is marked with a weight of 200g. Ten bags are randomly selected from several retailers and examined their weights.

203 204 197 195 201 205 198 199 194 207

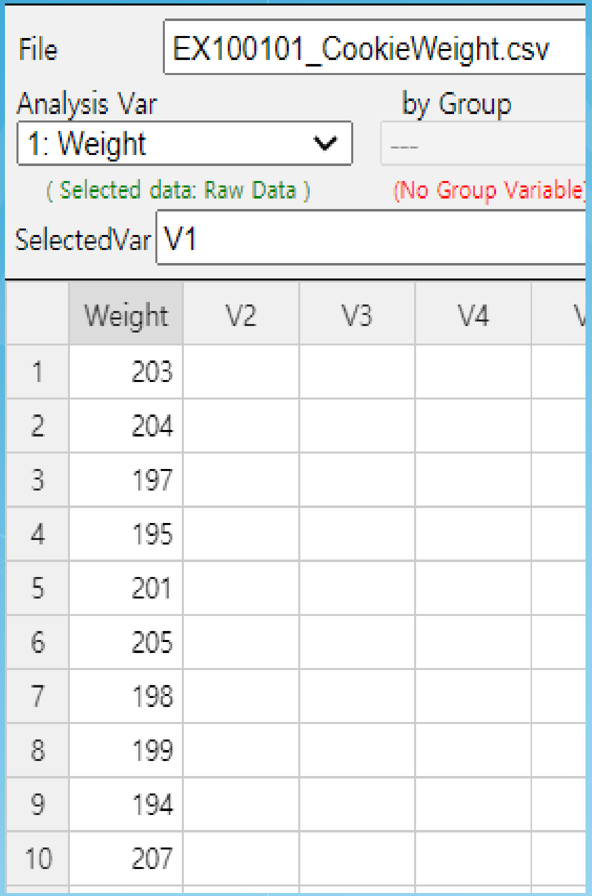
Can you say that there are as many cookies in the bag as weight marked?

1) Draw a histogram of the data to check whether a testing hypothesis using a parametric method can be performed.
2) Test the hypothesis by using a nonparametric method which utilizes the sign data by examining whether data values are smaller or larger than 200 with the significance level of 5%.
3) Check the result of the above test using『eStatU』.

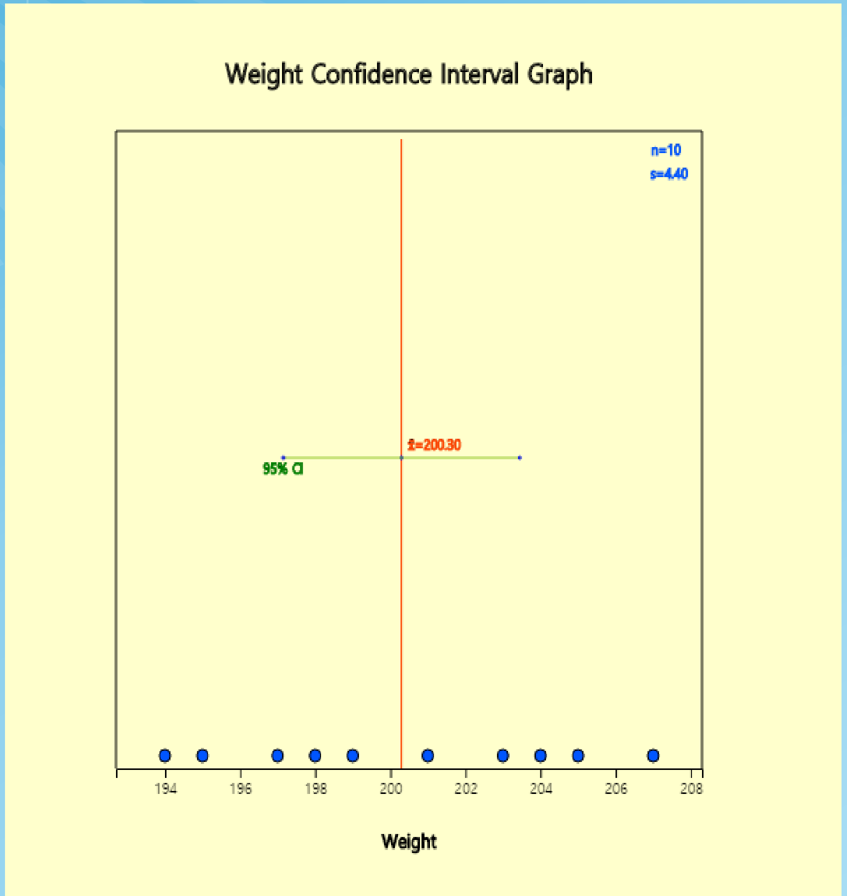
10.1 Nonparametric Test for Location Parameter of Single Population

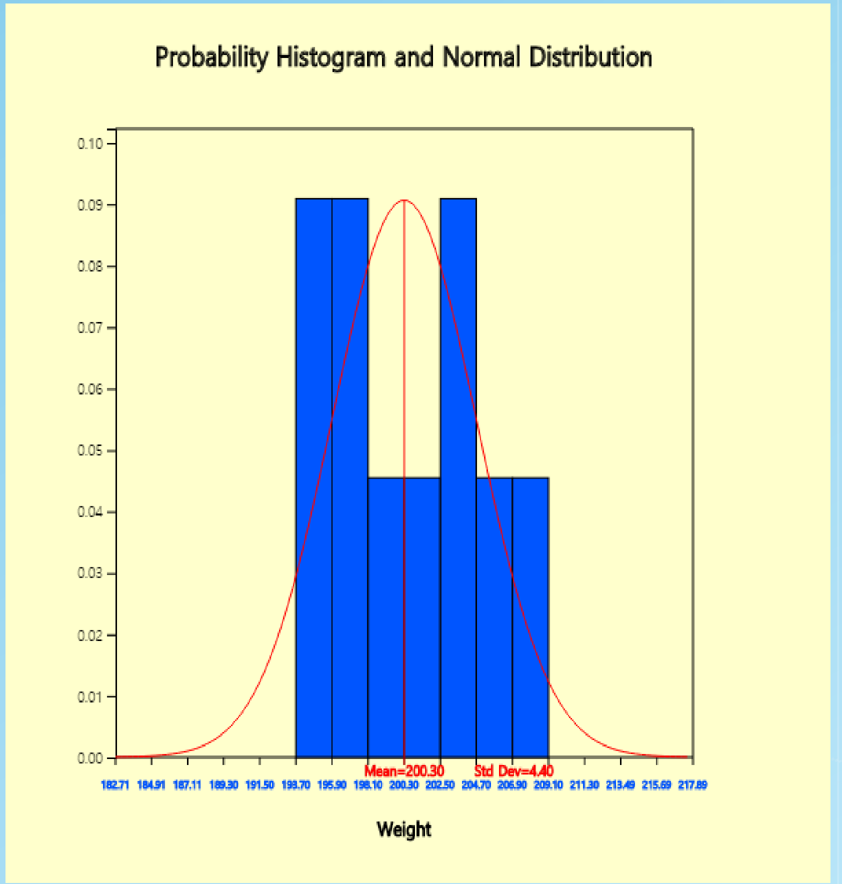
[Answer of Example 10.1.1]

- **Hypothesis** $H_0: \mu = 200, \ H_1: \mu \neq 200$


File EX100101_CookieWeight.csv
Analysis Var 1: Weight
by Group ---
(Selected data: Raw Data)
(No Group Variable)
SelectedVar V1

	Weight	V2	V3	V4
1	203			
2	204			
3	197			
4	195			
5	201			
6	205			
7	198			
8	199			
9	194			
10	207			





10.1 Nonparametric Test for Location Parameter of Single Population

[Answer of Example 9.1.1]

• sample data	203	204	197	195	201	205	198	199	194	207
• sign data	+	+	-	-	+	+	-	-	-	+

- Hypothesis $H_0 : M = 200,\ H_1 : M \neq 200$

- if H_0 is correct, the number of + signs may be the most likely to be 5 and 0, 1 or 9, 10 are very unlikely to be present.
- In order to test $H_0 : M = 200$ with 5% significance level, decision rule is as follows:

'If the number of + signs is either 0, 1 (cumulated probability is 0.011) or 9, 10 (cumulated probability from right is 0.011), then reject H_0

10.1 Nonparametric Test for Location Parameter of Single Population

Sign Test

[Hypothesis] $H_0 : M = M_0$ 200

◉ $H_1 : M \neq M_0$ ○ $H_1 : M > M_0$ ○ $H_1 : M < M_0$

[Test Type] Sign Test

Significance Level α = ◉ 5% ○ 1%

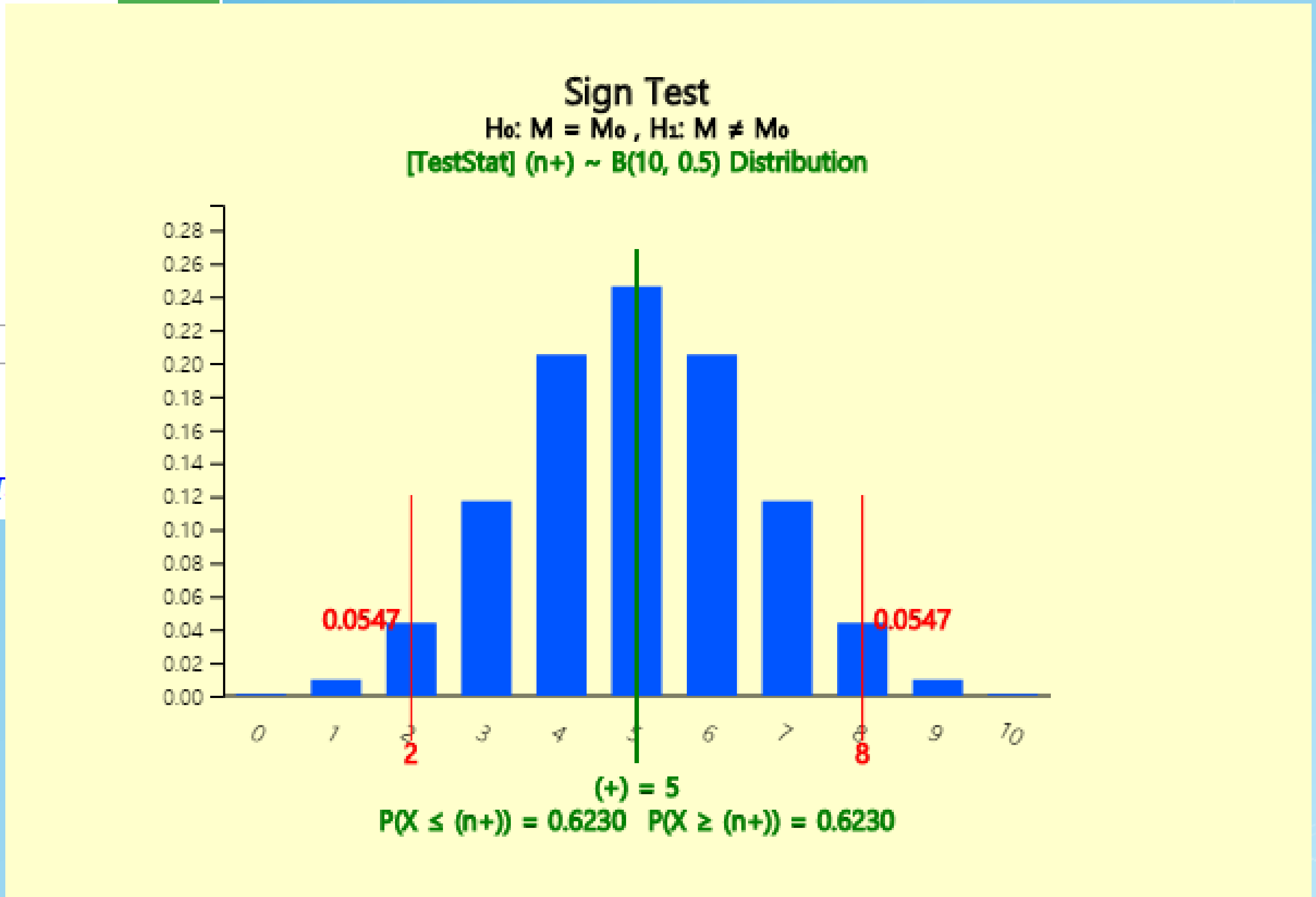
[Sample Data]

Sample 203 204 197 195 201 205 198 199 194 207

[Sample Statistics]

Sign Count (-) = 5 (+) = 5

Execute *If n ≤ 100 Binomial Test, n > 100 Normal Approximation T*



10.1 Nonparametric Test for Location Parameter of Single Population

Table 10.1.1 Decision rule of the sign test

Type of Hypothesis	Decision Rule Test Statistic n_+ = 'number of plus sign data'
1) $H_0 : M = M_0$ $H_1 : M > M_0$	If $n_+ > B(n,0.5)_\alpha$, then reject H_0, else accept H_0
2) $H_0 : M = M_0$ $H_1 : M < M_0$	If $n_+ < B(n,0.5)_{1-\alpha}$, then reject H_0, else accept H_0
3) $H_0 : M = M_0$ $H_1 : M \neq M_0$	If $n_+ < B(n,0.5)_{1-\alpha/2}$ or $n_+ > B(n,0.5)_{\alpha/2}$, then reject H_0, else accept H_0

If the observed value is the same as M_0?

If any of the observations has the same value as M_0, they are not used in the sign test. In other words, reduce n.

10.1 Nonparametric Test for Location Parameter of Single Population

Table 10.1.2 Decision rule of the sign test (large sample case)

Type of Hypothesis	Decision Rule Test Statistic: n_+ = 'number of plus sign data'
1) $H_0 : M = M_0$ $H_1 : M > M_0$	If $\dfrac{n_+ - 0.5n}{\sqrt{0.25n}} > z_\alpha$, then reject H_0, else accept H_0
2) $H_0 : M = M_0$ $H_1 : M < M_0$	If $\dfrac{n_+ - 0.5n}{\sqrt{0.25n}} < -z_\alpha$, then reject H_0, else accept H_0
3) $H_0 : M = M_0$ $H_1 : M \neq M_0$	If $\left\| \dfrac{n_+ - 0.5n}{\sqrt{0.25n}} \right\| > z_{\alpha/2}$, then reject H_0, else accept H_0

Thank you