Basic concepts of statistical analysis

Sasivimol Rattanasiri, Ph.D
Section for Clinical Epidemiology and Biostatistics
Ramathibodi Hospital, Mahidol University
E-mail: sasivimol.rat@mahidol.ac.th

1

Outline of talk

Hypothesis testing:

- Continuous data
  - Independent samples
  - Paired samples
- Categorical data
  - Independent samples
  - Paired samples

2

Steps for Hypothesis testing

Step 1

Generate the null and alternative hypotheses

Null hypothesis

Ho: $\mu_{BMD(calcium+)} = \mu_{BMD(calcium-)}$

Alternative hypothesis

Ha: $\mu_{BMD(calcium+)} \neq \mu_{BMD(calcium-)}$

3

Steps for Hypothesis testing

Step 2

Determine the significance level

The commonly used value of the significance level is 0.05. Usually this value does not exceed 0.10.

4

Steps for Hypothesis testing

Step 3

Select an appropriate test statistic

- The objective of the analysis
- The type of data
- The number of groups of samples
- Independent or dependent groups of samples

5

Steps for Hypothesis testing

Step 4

Calculate the test statistic

- The general form of the test statistic can be expressed as:

$$\text{test statistic} = \frac{\text{observed value - hypothesized value}}{\text{standard error of observed value}}$$

6

Steps for Hypothesis testing

Step 5

Convert the test statistic to *p* value

The *p* value is the probability of obtaining our observed data (or more extreme data) when the null hypothesis is true.

Null hypothesis

Ho: $\mu_{BMD(calcium+)} = \mu_{BMD(calcium-)}$

Alternative hypothesis

Ha: $\mu_{BMD(calcium+)} \neq \mu_{BMD(calcium-)}$

If p=0.30, the probability of obtaining a result that the means of BMD between women who received and who did not receive calcium supplement are not different, when the null hypothesis is true, is 30%.

Steps for Hypothesis testing

Step 6

Draw a conclusion

If the *p* value is less than or equal to significance level, we reject the null hypothesis.

If the *p* value is greater than significance level, we do not reject the null hypothesis.

Null hypothesis

Ho: $\mu_{BMD(calcium+)} = \mu_{BMD(calcium-)}$

Alternative hypothesis

Ha: $\mu_{BMD(calcium+)} \neq \mu_{BMD(calcium-)}$

Significance level	P value	Decision
0.05	0.30	Not reject
0.05	0.01	Reject

Recommendations

- When we reject the null hypothesis, we accept the alternative hypothesis.
- When we do not reject the null hypothesis, we cannot say that the null hypothesis is true, but only that we do not have sufficient evidence to reject it.

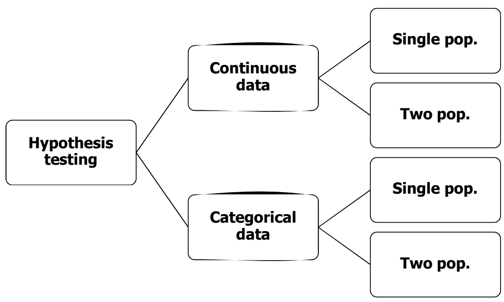

Figure 1 Flow chart for hypothesis testing

Hypothesis testing for continuous data

- ❖ Independent samples
- ❖ Paired samples

13

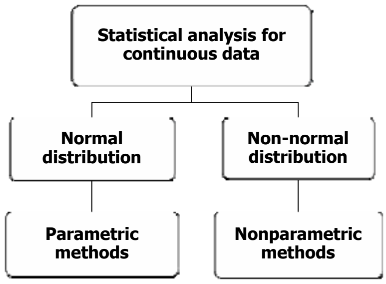

Figure 2 Flow chart for hypothesis testing based upon the distribution of data

14

Statistical tests for two independent groups

Distribution	Parameter	Statistical test
Normal	Mean	Student t-test
Non-normal	Median	Wilcoxon rank-sum test

15

Independent Samples

<u>Class example I</u>

Researchers wanted to test if means/median of weights of HIV patients who received NVP or EFV, are different.

From a study of Monosuthi and et al.

16

Steps for Hypothesis Testing

1. Check assumption about normality

For data of weights of patients

	NVP	EFV
Mean	54.82	54.36
Median	54.50	52
SD	8.66	10.60
Skewness	0.33	0.25

We conclude that the weights of patients of two groups have normal distribution.

17

Steps for Hypothesis Testing

2. Generate the null and alternative hypotheses as follows:

H_0: $\mu_{NVP} = \mu_{EFV}$

H_A: $\mu_{NVP} \neq \mu_{EFV}$

3. Set the level of significance

We follow the standard convention and set the level of significance to 0.05.

18

Steps for Hypothesis Testing

4. Select an appropriate statistical test

Question	Answer
No. of samples	2 samples
Characteristic of samples	Independent samples
Distribution of data	Normal distribution
Statistical test	Student t-test

19

The two-sample t-test using STATA program

```
. ttest bw,by(group)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     NVP |      70    54.82286    1.034775    8.657545    52.75854    56.88718
     EFV |      70    54.36429    1.266647    10.59753    51.83739    56.89118
---------+--------------------------------------------------------------------
combined |     140    54.59357    .8150796    9.644152    52.98201    56.20513
---------+--------------------------------------------------------------------
    diff |            .4585714    1.635589               -2.775485    3.692628
------------------------------------------------------------------------------
    diff = mean(NVP) - mean(EFV)                                  t =   0.2804
Ho: diff = 0                                     degrees of freedom =      138

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.6102         Pr(|T| > |t|) = 0.7796 ⇦      Pr(T > t) = 0.3898
```

Steps for Hypothesis Testing

5. Draw a conclusion

-The *p* value is greater than 0.05 which is greater than the level of significance.
-Therefore, we cannot reject the null hypothesis and conclude that the means weight of the NVP and EFV groups are identical.

21

Independent Samples

<u>Class example II</u>

Researchers wanted to test if means/median of CD4 count of HIV patients who received NVP or EFV, are different.

From a study of Monosuthi and et al.

22

Steps for Hypothesis Testing

1. Check assumption about normality

For data of CD4 count of patients

	NVP	EFV
Mean	60.97	63.80
Median	35	29
SD	74.30	74.94
Skewness	2.31	1.45

We conclude that CD4 count of patients of two groups have normal distribution.

23

Steps for Hypothesis Testing

2. Generate null and alternative hypotheses

H_0: $M_{NVP}=M_{EFV}$

H_A: $M_{NVP}\neq M_{EFV}$

3. Set the level of significance to 0.05

24

Steps for Hypothesis Testing

4. Select an appropriate statistical test

Question	Answer
No. of samples	2 samples
Characteristic of samples	Independent samples
Distribution of data	Non-normal distribution
Statistical test	Mann-Whitney test

25

STATA output

```
. ranksum cd4c,by(group)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       group |      obs    rank sum    expected
-------------+---------------------------------
         NVP |       70      5025.5        4935
         EFV |       70      4844.5        4935
-------------+---------------------------------
    combined |      140        9870        9870

unadjusted variance    57575.00
adjustment for ties      -22.54
                     ----------
adjusted variance      57552.46

Ho: cd4c(group==NVP) = cd4c(group==EFV)
             z =   0.377
    Prob > |z| =   0.7060
```

26

Steps for Hypothesis Testing

5. Draw a conclusion

The *p* value from Mann-Whitney test is 0.706 which is greater than the level of significance.

So, we cannot reject the null hypothesis.

Conclusion, the median CD4 count of the NVP group is equal to the median CD4 count of the EFV group.

27

Table 1. Comparison of characteristics between NVP and EFV groups

Characteristics	NVP	EFV	P value
	Mean (SD)	Mean (SD)	
Age (year)			
Weight (kg)			
Height (cm)			
BMI (kg/m^3)			
CD4 count; median (range)			

28

Statistical tests for paired samples

Distribution	Parameter	Statistical test
Normal	Mean	Paired t-test
Non-normal	Median	Wilcoxon matched signed-rank test

29

Paired Samples

Class exmples III

Researchers wanted to test if the mean weights of HIV patients before and after 12 weeks of receiving an antiretroviral therapy regimen are different.

From a study of Monosuthi and et al.

30

Steps for Hypothesis Testing

1. Check assumption about normality

For data about weights of HIV patients:

	Before	After
Mean	54.6	57.3
Median	53	56.7
SD	9.8	10.3
Skewness	0.4	0.3

The weights of patients before and after receiving antiretroviral therapy have normal distributions.

Steps for Hypothesis Testing

2. Generate the null and alternative hypotheses as follows:

H_0: $\mu_{before} = \mu_{after}$ or H_0: $\mu_{difference} = 0$

H_A: $\mu_{before} \neq \mu_{after}$ or H_A: $\mu_{difference} \neq 0$

3. Set the level of significance

We follow the standard convention and set the level of significance to 0.05.

Steps for Hypothesis Testing

4. Select an appropriate statistical test

Question	Answer
No. of samples	2 samples
Characteristic of samples	Paired samples
Distribution of data	Normal distribution
Statistical test	Paired t-test

The paired-t test using STATA program

```
. ttest bw0= bw12

Paired t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     bw0 |     121    54.56694    .8926941    9.819635    52.79947    56.33441
    bw12 |     121    57.31322    .9380435    10.31848    55.45596    59.17048
---------+--------------------------------------------------------------------
    diff |     121   -2.746281    .3710625    4.081688   -3.480959   -2.011603
------------------------------------------------------------------------------
     mean(diff) = mean(bw0 - bw12)                                t =  -7.4011
 Ho: mean(diff) = 0                              degrees of freedom =      120

 Ha: mean(diff) < 0           Ha: mean(diff) != 0           Ha: mean(diff) > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000 ⇦        Pr(T > t) = 1.0000
```

Steps for Hypothesis Testing

5. Draw a conclusion

-*p* value is less than 0.001 which is less than the level of significance.
-So, we reject the null hypothesis and conclude that the mean difference of weights is not equal to zero.
-Alternatively, the mean weights of HIV patients before and after receiving antiretroviral therapy are not equal.

Paired Samples

<u>Class exmples IV</u>

Researchers wanted to test if the median CD4 counts of HIV patients before and after 12 weeks of receiving antiretroviral therapy regimen are different.

From a study of Monosuthi and et al.

Steps for Hypothesis Testing

1. Check assumption about normality

For data about CD4 count of HIV patients:

	Before	After
Mean	62.9	162.7
Median	29	132
SD	74.2	128.6
Skewness	1.7	1.5

The CD4 count of patients before and after receiving antiretroviral therapy have normal distributions.

37

Steps for Hypothesis Testing

2. Generate null and alternative hypotheses

H_0: $M_{before}=M_{after}$ **or** **H_0: $M_{difference}=0$**

H_A: $M_{before}\neq M_{after}$ **H_A: $M_{difference}\neq 0$**

3. Set the level of significance to 0.05

38

Steps for Hypothesis Testing

4. Select an appropriate statistical test

Question	**Answer**
No. of samples	2 samples
Characteristic of samples	Paired samples
Distribution of data	Non-normal distribution
Statistical test	Wilcoxon matched-pairs signed-rank test

39

STATA output

```
. signrank  cd4c0= cd4c12
       Wilcoxon signed-rank test
               sign |      obs   sum ranks    expected
       -------------+---------------------------------
           positive |        7       256.5        3570
           negative |      112      6883.5        3570
               zero |        0           0           0
       -------------+---------------------------------
                all |      119        7140        7140
       unadjusted variance   142205.00
       adjustment for ties       -5.38
       adjustment for zeros       0.00
                             ----------
       adjusted variance     142199.63
       Ho: cd4c0 = cd4c12
                    z =  -8.787
           Prob > |z| =   0.0000
```

40

Steps for Hypothesis Testing

5. Draw a conclusion

The *p* value is less than 0.001 which is less than the level of significance, so we reject the null hypothesis.

Conclusion, the medians CD4 counts before and after receiving antiretroviral therapy regimen are not equal.

41

Table 2. Comparison of the outcomes before and after 12 weeks of receiving therapy

Outcomes	**Before**	**After**	**P value**
	Mean (SD)	**Mean (SD)**	
Weight (kg)			
BMI(kg/m^3)			
ALP			
AST			
CD4 cell/µL; median (range)			
Log VL; median (range)			

42

Statistical tests for comparison continuous data between two groups

Distribution	Parameter	Condition	Statistical test
Normal	Mean	Paired	Paired t-test
	Mean	Independent	Student t-test
Non-normal	Median	Paired	Wilcoxon matched signed-rank test
	Median	Independent	Wilcoxon rank-sum test

43

Hypothesis testing for categorical data

- ❖ Independent samples
- ❖ Paired samples

44

Independent Samples

- A case-control study was conducted to look at the effects of traditional medicine used on the risk of hip fracture.

- Researchers wanted to assess the association between traditional medicine used and the risk of osteoporotic hip fracture.

45

Observed frequencies of the 2x2 table for assessing the association between traditional medicine used and the risk of hip fracture.

Traditional medicine users	Hip fracture		Total
	Yes	No	
Yes	20	8	28
No	208	216	424
Total	228	224	452

46

Statistical analysis

- For independent samples, the Chi-square test is conducted to examine the association between two categorical variables.

- This test is based upon the null hypothesis that the two categorical variables are independent or the proportions of the interested event between two independent groups are not different.

47

Steps for Hypothesis Testing

1. Generate the null hypothesis as follows:

H_0: Traditional medicine used is not associated with the risk of hip fracture.

or

H_0: Proportions of traditional medicine used between two groups of patients, with and without hip fracture are not different.

48

Steps for Hypothesis Testing

2. Generate the alternative hypothesis as follows:

H_a: Traditional medicine used is associated with the risk of hip fracture.

or

H_a: Proportions of traditional medicine used between two groups of patients, with and without hip fracture are different.

Steps for Hypothesis Testing

3. Set the level of significance

   We follow the standard convention and set the level of significance to 0.05.

4. Select an appropriate statistical test

   We are testing a hypothesis about the association between two categorical variables, and the samples are independent, so a statistical test for this case is the Chi-square test.

Chi-square test using STATA

```
. tabulate tredmed hip, chi2 column exact expected

+-------------------+
| Key               |
|-------------------|
|     frequency     |
| expected frequency |
| column percentage  |
+-------------------+

traditiona |     hip fracture
l medicine |        no        yes |     Total
-----------+----------------------+----------
       yes |         8         20 |        28
           |      13.9       14.1 |      28.0
           |      3.57       8.77 |      6.19
-----------+----------------------+----------
        no |       216        208 |       424
           |     210.1      213.9 |     424.0
           |     96.43      91.23 |     93.81
-----------+----------------------+----------
     Total |       224        228 |       452
           |     224.0      228.0 |     452.0
           |    100.00     100.00 |    100.00

          Pearson chi2(1) =   5.2588   Pr = 0.022
```


Steps for Hypothesis Testing

5. Draw a conclusion

-The *p* value is equal to 0.02, we reject the null hypothesis and conclude that traditional medicine used is significantly associated with the risk of hip fracture.

In other words, the proportions of traditional medicine used between two groups of patients, with and without hip fracture are different.

Test for independence – small samples

• The Chi-square test is not an appropriate method of test for independence if the expected value is less than 5 for more than 20% of the total cells.

• The **Fisher's exact test** is an alternative method of test for independence when the requirements of the Chi-square test are not met.

OC used	Cervical cancer		Total
	Yes	No	
Ever used	68	150	218
Never used	254	875	1129
Total	322	1025	1347

The expected frequency of ever used OC and have cervical cancer (n_{11}) can be calculated as:

$$\frac{218 \times 322}{1347} = 52.1$$

OC used	Cervical cancer		Total
	Yes	No	
Ever used	68	150	218
Never used	254	875	1129
Total	322	1025	1347

The expected frequency of ever used OC and did not have cervical cancer (n_{12}) can be calculated as:

$$\frac{218 \times 1025}{1347} = 165.9$$

Observed frequencies of the 2x2 table for assessing the association between HRT and the risk of hip fracture.

HRT	Hip fracture		Total
	Yes	No	
Yes	1	2	3
No	213	214	427
Total	214	216	430

Expected frequencies of the 2x2 table for assessing the association between HRT and the risk of hip fracture.

HRT	Hip fracture		Total
	Yes	No	
Yes	1.5	1.5	3
No	212.5	214.5	427
Total	214	216	430

Statistical tests for hypothesis testing about the association between two categorical variables

%Exp freq. less than 5	Statistical test
Greater than 20%	Fisher's exact test
Less than or equal to 20%	Chi-square test

Table 3. Comparison of risk factors between CKD and non-CKD groups

Factors	CKD	Non-CKD	P value
	N (%)	N (%)	
Sex			
Male			
Female			
Hypertension			
Yes			
No			

Paired samples

- Researchers wanted to compare the pain relief (yes/no) by two different analgesics in the same subjects.
- In a matched case-control study in which investigators matched case to control patients with BMI , They wanted to assess the association between HRT and the hip fracture.

2x2 contingency table for paired samples

Cases	Controls		
	Exposed	Unexposed	Total
Exposed	a	b	a+b
Unexposed	c	d	c+d
Total	a+c	b+d	n

- a,b,c,d are the frequencies of pairs for each combination.
- b and c are the discordant pairs.

61

Statistical test for paired samples

Statistical test for testing the association between two categorical variables in paired samples is the McNemar's test which can be calculated as:

$$\chi^2 = \frac{(|b-c|)^2}{b+c}$$

If the number of discordant pairs is less than 20, the exact McNemar's test is required.

62

Class example

- For a matched case-control study, the patients were assigned to pairs matched on body mass index (BMI).

- Researchers wanted to assess the association between hormone replacement therapy (HRT) and the hip fracture.

63

2x2 contingency table of 372 matched pairs for assessing the association between HRT and hip fracture

Cases	Controls		
	HRT+	HRT-	Total
HRT+	102	50	152
HRT-	100	120	220
Total	202	170	372

64

Null hypothesis

H_0: HRT is not associated with the risk of hip fracture.

OR

H_0: Paired proportions of patients who received HRT between two groups of patients, with and without hip fracture are not different.

65

Alternative hypothesis

H_a: HRT is associated with the risk of hip fracture.

OR

H_a: Paired proportions of patients who received HRT between two groups of patients, with and without hip fracture are different.

66

STATA command: individual data

```
. mcc case control

                 | Controls               |
Cases            |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
         Exposed |       102          50  |        152
       Unexposed |       100         120  |        220
-----------------+------------------------+------------
           Total |       202         170  |        372

McNemar's chi2(1) =     16.67    Prob > chi2 = 0.0000
Exact McNemar significance probability       = 0.0001

Proportion with factor
        Cases       .4086022
        Controls    .5430108     [95% Conf. Interval]
                   ---------     --------------------
        difference -.1344086     -.2001631  -.0686541
        ratio       .7524752      .656141   .8629532
        rel. diff. -.2941176     -.4547495  -.1334858

        odds ratio        .5      .3487202   .7089431   (exact)
```

67

STATA command: summary data

```
. mcci 102 50 100 120

                 | Controls               |
Cases            |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
         Exposed |       102          50  |        152
       Unexposed |       100         120  |        220
-----------------+------------------------+------------
           Total |       202         170  |        372

McNemar's chi2(1) =     16.67    Prob > chi2 = 0.0000
Exact McNemar significance probability       = 0.0001

Proportion with factor
        Cases       .4086022
        Controls    .5430108     [95% Conf. Interval]
                   ---------     --------------------
        difference -.1344086     -.2001631  -.0686541
        ratio       .7524752      .656141   .8629532
        rel. diff. -.2941176     -.4547495  -.1334858

        odds ratio        .5      .3487202   .7089431   (exact)
```

68

Draw conclusions

-The *p* value is less than 0.001 which is less than the level of significance.

- We reject the null hypothesis and conclude that HRT is significantly associated with the hip fracture.

- In other words, the paired proportions of HRT between two groups of patients, with and without hip fracture are different.

69