十二、 Nonparametric Methods (Chapter 12)

劉仁沛教授
國立台灣大學農藝學研究所生物統計組
國立台灣大學流行病學與預防醫學研究所
國家衛生研究院生物統計與生物資訊組
jpliu@ntu.edu.tw

Learning Objectives

- Learn advantages and disadvantages of nonparametric statistics.
- Nonparametric tests:
  - Testing randomness of a single sample: Run test
  - Testing difference
    - Two independent samples: Mann-Whitney-Wilcoxon Rank Sum test
      - Two-sample z/t test
    - Two dependent samples. Wilcoxon signed rank test
      - Paired sample t test
    - >2 independent samples. Kruskal-Wallis test
      - One-way ANOVA
    - >2 samples with blocking: Friedman test
      - RCBD
  - Correlation: Spearman’s rank correlation

Introduction

- Assumption for t-test or correlation (regression) coefficients
  - Normality
  - Equal variance
  - Independence
- Not all data satisfy these assumptions!

Parametric v.s. Nonparametric statistics

- Parametric statistics mainly are based on assumptions about the population
  - Ex. X has normal population for t-test, or ANOVA.
  - Requires interval or ratio level data.
- Nonparametric statistics depend on fewer assumptions about the population and parameters.
  - “distribution-free” statistics.
  - Most analysis are based on rank.

Advantages and Disadvantages of Nonparametric Techniques

- Advantages
  - There is no parametric alternative
  - Nominal data or ordinal data are analyzed
  - Less complicated computations for small sample size
  - Exact method. Not approximation.
- Disadvantages
  - Less powerful if parametric tests are available.
  - Not widely available and less well know

Wilcoxon Signed-rank Test

Example:

脊柱後側凸病患，其平均 Pimax 值是否小於 110 cm H_2O

H_o：⊠ ⊠ 110 vs. H_a：⊠ < 110

n=9

No power to verify the normality assumption

Wilcoxon Signed-rank Test

Exact Methods for Small Sample($n \leq 30$)

1. Calculate the difference X -110 for each of 9 observations. Differences equal to zero are eliminated, and the number of observations, n, is reduced accordingly.
2. Rank the absolute values of the differences, assigning 1 to the smallest, 2 to the second smallest, and so on. Tied observations are assigned the average of the rank that would have been assigned with no ties.

Wilcoxon Signed-rank Test

3. The indicator variable for sign is 1 if the difference is positive, is 0 if the difference is negative.
4. Multiple the sign and rank of absolute difference. This is called the signed rank.
5. Calculate the sum for the signed ranks, T_{SR}
6. For a two-tailed test, reject the null hypothesis if

   $T_{SR} > W_{1-\alpha/2,\, n}$ or $T_{SR} < W_{\alpha/2,\, n}$

   Note: $W_{\alpha/2,\, n} = n(n+1)/2 - W_{1-\alpha/2,\, n}$

Wilcoxon Signed-rank Test

Methods for Larger Samples (n>30)

Test statistic:

$$z = \frac{T_{SR} - \left[n(n+1)/4 \right]}{\sqrt{[n(n+1)(2n+1)]/24}}$$

Wilcoxon Signed-rank Test

Example: X=pi_{max}

			Rank of	sign*rank	
X	X-110	abs(X-110)	abs(X-110)	sign	abs(X-110)
54.8	-55.2	55.2	3	0	0
62.0	-48.0	48.0	2	0	0
63.3	-46.7	46.7	1	0	0
44.2	-65.8	65.8	4	0	0
40.3	-69.7	69.7	5	0	0
36.3	-73.7	73.3	6	0	0
19.3	-90.7	90.7	9	0	0
24.6	-85.4	85.4	8	0	0
26.6	-83.4	83.4	7	0	0
Sum					0

Wilcoxon Signed-rank Test

N=9<30 □ Exact Method

$T_{SR} = 0$

□=0.05, $T_{0.025,9}$=6 □$T_{SR} = 0 < T_{0.025,9}$=6

Reject H_0 at the 0.05 level.

Quantiles of the Wilcoxon Signed Ranks Test Statistic

	$w_{0.005}$	$w_{0.01}$	$w_{0.025}$	$w_{0.05}$	$w_{0.10}$	$w_{0.20}$	$w_{0.30}$	$w_{0.40}$	$w_{0.50}$	$\frac{n(n+1)}{2}$
$n = 4$	0	0	0	0	1	3	3	4	5	10
5	0	0	0	1	3	4	5	6	7.5	15
6	0	0	1	3	4	6	8	9	10.5	21
7	0	1	3	4	6	9	11	12	14	28
8	1	2	4	6	9	12	14	16	18	36
9	2	4	6	9	11	15	18	20	22.5	45
10	4	6	9	11	15	19	22	25	27.5	55
11	6	8	11	14	18	23	27	30	33	66
12	8	10	14	18	22	28	32	36	39	78
13	10	13	18	22	27	33	38	42	45.5	91
14	13	16	22	26	32	39	44	48	52.5	105
15	16	20	26	31	37	45	51	55	60	120
16	20	24	30	36	43	51	58	63	68	136
17	24	28	35	42	49	58	65	71	76.5	153
18	28	33	41	48	56	66	73	80	85.5	171
19	33	38	47	54	63	74	82	89	95	190
20	38	44	53	61	70	83	91	98	105	210
21	44	50	59	68	78	91	100	108	115.5	131
22	49	56	67	76	87	100	110	119	126.5	153
23	55	63	74	84	95	110	120	130	138	176
24	62	70	82	92	105	120	131	141	150	300
25	69	77	90	101	114	131	143	153	162.5	325
26	76	85	99	111	125	142	155	165	175.5	351
27	84	94	108	120	135	154	167	178	189	378
28	92	102	117	131	146	166	180	192	203	406
29	101	111	127	141	158	178	193	206	217.5	435
30	110	121	138	152	170	191	207	220	232.5	465
31	119	131	148	164	182	205	221	235	248	496
32	129	141	160	176	195	219	236	250	264	528
33	139	152	171	188	208	233	251	266	280.5	561
34	149	163	183	201	222	248	266	282	297.5	595
35	160	175	196	214	236	263	283	299	315	630
36	172	187	209	228	251	279	299	317	333	666
37	184	199	222	242	266	295	316	335	351.5	703
38	196	212	236	257	282	312	334	353	370.5	741
39	208	225	250	272	298	329	352	372	390	780
40	221	239	265	287	314	347	371	391	410	820
41	235	253	280	303	331	365	390	411	430.5	861
42	248	267	295	320	349	384	409	431	451.5	903
43	263	282	311	337	366	403	429	452	473	946
44	277	297	328	354	385	422	450	473	495	990
45	292	313	344	372	403	442	471	495	517.5	1035
46	308	329	362	390	423	463	492	517	540.5	1081
47	324	346	379	408	442	484	514	540	564	1128

Mann-Whitney-Wilcoxon Rank Sum Test

- Two independent random samples
- Example: Weight gain at one month by two baby formulas

A: 6.9, 7.6, 7.3, 7.6, 6.8, 7.2, 8.0, 5.5, 7.3

B: 6.4, 6.7, 5.4, 8.2, 5.3, 6.6, 5.8, 5.7 6.2, 7.1

Mann-Whitney-Wilcoxon Rank Sum Test

- Method
  - Rank the observations in the combined sample from the smallest (1) to the largest (n1+n2)
  - In case of ties, use the averaged rank
  - Compute the sum of ranks for each sample

Mann-Whitney-Wilcoxon Rank Sum Test

A Weight	Rank	B Weight	Rank
6.9	11	6.4	7
7.6	16.5	6.7	9
7.3	14.5	5.4	2
7.6	16.5	8.2	19
6.8	10	5.3	1
7.2	13	6.6	8
8.0	18	5.8	5
5.5	3	5.7	4
7.3	14.5	6.2	6
7.1	12		
sum	117		73

Mann-Whitney-Wilcoxon Rank Sum Test

Exact Method for Small Samples($n1+n2 \leq 30$)

Null hypothesis: H_0: The location of population distributions for 1 and 2 are identical.

Alternative hypothesis: H_a: The location of the population distributions are shifted in either directions(a two-tailed test).

Mann-Whitney-Wilcoxon Rank Sum Test

3. Test statistics:

   For a two-tailed test, use U,

$$U = T_1 - \frac{n_1(n_1+1)}{2}$$

where T_1 is the rank sums for samples 1.

Mann-Whitney-Wilcoxon Rank Sum Test

4. Rejection rule:

For the two-tailed test and a given value of significance α, reject the null hypothesis of no difference if

$U < w_{\alpha/2,n1,n2}$ or

$U > w_{1-\alpha/2,n1,n2}$

$w_{1-\alpha/2,n1,n2} = n_1 n_2 - w_{\alpha/2,n1,n2}$

Mann-Whitney-Wilcoxon Rank Sum Test

Method for larger Samples (n1+n2>30)

Test Statistics:

$$Z = \frac{U - (n_1 n_2/2)}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}},$$

Reject the null hypothesis is $Z < -z_{\alpha/2}$

or $Z > z_{\alpha/2}$

Mann-Whitney-Wilcoxon Rank Sum Test

Example: $n_1 = 9,\ n_2 = 10,\ n_1 + n_2 = 19$

$$U = 117 - \frac{9(9+1)}{2} = 72$$

Two-tailed test $\alpha = 0.05$

$w_{0.025,9,10} = 21;$

$w_{0.975,9,10} = n_1 n_2 - w_{0.025,9,10} = (9)(10) - 21 = 69.$

Since $U = 72 > w_{0.975,9,10} = 69,$

reject H_0 that no difference exists between two baby formulas on weight gain

Critical Values for the Wilcoxon/Mann-Whitney Test (U)

Nondirectional α=.05 (Directional α=.025)

n_1 \ n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	0	0	0	0	1	1	1	1	1	2	2	2	2
3	-	-	-	-	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	-	-	-	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	-	-	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	-	-	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	-	-	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	-	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	-	0	2	4	7	10	12	15	17	21	23	26	28	31	34	37	39	42	45	48
10	-	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	-	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	-	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	-	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	-	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	-	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	-	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	-	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	-	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	-	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	-	2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

Kruskal-Wallis Test

- K independent samples
- Example: body weights in gram of Wistar rats in a repeated dose toxicity study

Kruskal-Wallis Test

- Data set: body weights in gram of Wistar rats in a repeated dose toxicity study
- Control:
  295.1 277.9 299.4 280.6 285.7 299.2 279.7 277.4 299.2
  287.8 292.0 318.8 280.8 292.9 305.2
- Low dose:
  287.3 289.5 278.4 281.8 264.9 252.0 284.7 268.9 305.6
  295.7 287.6 254.7 292.7 267.9 300.8
- Middle dose:
  247.5 281.1 284.5 295.0 285.9 273.7 244.1 272.7 262.1
  278.8 298.3 298.5 293.5 259.6 275.3
- High dose:
  263.8 255.6 267.2 259.6 238.2 240.4 255.6 255.5 242.5
  296.6 246.0 282.7 254.0 280.6 268.2

12.4 Kruskal-Wallis Test

- Methods:
  - Rank the combined sample from the smallest (1) to the largest(n_1+ n_2+… + n_t)
  - In case of ties, use the averaged rank
  - Compute the sums of ranks for each samples, R_i

12.4 Kruskal-Wallis Test

Methods:

1. Null hypothesis: H_0: The locations of the distributions of all of the $k>2$ populations are identical.
2. Alternative hypothesis: H_a: The locations of at least two of the k frequency distributions differ

12.4 Kruskal-Wallis Test

3. Test statistics:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

4. Rejection region:

Reject H_0 if $H > \chi_{\alpha}^2$ with (k-1) degrees of freedom

Kruskal-Wallis Test

Control:

295.1(49) 277.9(26) 299.4(56) 280.6(30.5) 285.7(38) 299.2(54.5) 279.7(29) 277.4(25) 299.2(54.5) 287.8(42) 292.0(44) 318.8(60) 280.8(32) 292.9(46) 305.2(58); sum = 644.5

Low dose:

287.3(40) 289.5(43) 278.4(27) 281.8(34) 264.9(17) 252.0(7) 284.7(37) 268.9(21) 305.6(59) 295.7(50) 287.6(41) 254.7(9) 292.7(45) 267.9(19) 300.8(57); sum = 506.5

- Middle dose:

  247.5(6) 281.1(33) 284.5(36) 295.0(48) 285.9(39) 273.7(23) 244.1(4) 272.7(22) 262.1(15) 278.8(28) 298.3(52) 298.5(53) 293.5(47) 259.6(13.5) 275.3(24); sum = 443.5

- High dose:

  263.8(16) 255.6(11.5) 267.2(18) 259.6(13.5) 238.2(1) 240.4(2) 255.6(11.5) 255.5(10) 242.5(3) 296.6(51) 246.0(5) 282.7(35) 254.0(8) 280.6(30.5) 268.2(20); sum = 236.0

12.4 Kruskal-Wallis Test

Example:

$n_1 = 15,\ n_2 = 15,\ n_3 = 15,\ n_4 = 15,\ \text{n} = 4\text{x}15 = 60$

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{4} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{60(60+1)} \left[\frac{(644.5)^2}{15} + \frac{(506)^2}{15} + \frac{(443.5)^2}{15} + \frac{(236)^2}{15} \right] - 3(60+1)$$

$$= 18.9266 > \chi^2_{0.05,3} = 7.81$$

Reject H_0 that weights are the same for all groups

Quantiles of the Kruskal-Wallis Test Statistic for Small Sample Sizes

Sample Sizes	$w_{0.90}$	$w_{0.95}$	$w_{0.99}$
2, 2, 2	3.7143	4.5714	4.5714
3, 2, 1	3.8571	4.2857	4.2857
3, 2, 2	4.4643	4.5000	5.3571
3, 3, 1	4.0000	4.5714	5.1429
3, 3, 2	4.2500	5.1389	6.2500
3, 3, 3	4.6000	5.0667	6.4889
4, 2, 1	4.0179	4.8214	4.8214
4, 2, 2	4.1667	5.1250	6.0000
4, 3, 1	3.8889	5.0000	5.8333
4, 3, 2	4.4444	5.4000	6.3000
4, 3, 3	4.7000	5.7273	6.7091
4, 4, 1	4.0667	4.8667	6.1667
4, 4, 2	4.4455	5.2364	6.8727
4, 4, 3	4.773	5.5758	7.1364
4, 4, 4	4.5000	5.6538	7.5385
5, 2, 1	4.0500	4.4500	5.2500
5, 2, 2	4.2933	5.0400	6.1333
5, 3, 1	3.8400	4.8711	6.4000
5, 3, 2	4.4946	5.1055	6.8218
5, 3, 3	4.4121	5.5152	6.9818
5, 4, 1	3.9600	4.8600	6.8400
5, 4, 2	4.5182	5.2682	7.1182
5, 4, 3	4.5231	5.6308	7.3949
5, 4, 4	4.6187	5.6176	7.7440
5, 5, 1	4.0364	4.9091	6.8364
5, 5, 2	4.5077	5.2462	7.2692
5, 5, 3	4.5363	5.6264	7.5429
5, 5, 4	4.5200	5.6429	7.7914
5, 5, 5	4.5000	5.6600	7.9800

統計歷史人物小傳
Frank Wilcoxon (1892-1965)

- PhD in organic chemistry from Cornell in 1924
- Research in fungicide and insecticide from 1924-1943
  - Boyce Thompson Institute for Plant Research
- American Cyanamid from 1943-1957
- More than 70 papers in plant physiology

統計歷史人物小傳
Frank Wilcoxon (1892-1965)

- Apply statistical methods such as t test to plant pathology and discovered they are inadequate.
- He developed the methods based on ranks
- He did not know whether the method is correct or not.
- He sent a manuscript of 4 pages to *Biometrics* in 1945 to let the referees tell him whether the method is correct or not.
- The rest is the history of an new era of nonparametric statistics.

Summary

- Wilcoxon Signed-rank Test
- Mann-Whitney-Wilcoxon Test
- Kruskal-Wallis Test