心理统计
第十六讲：非参数检验

严超赣

Chao-Gan Yan, Ph.D.

yancg@psych.ac.cn
http://rfmri.org/yan

Institute of Psychology, Chinese Academy of Sciences

1

1

参数检验与非参数检验

- 总体分布类型已知，只有有限个参数未知的问题称为参数统计问题，参数统计中的未知总体虽然有无穷多种可能的分布，但所有可能的分布可以由有限个未知参数的变化来描述。
- 总体分布类型未知的问题称为非参数统计问题;非参数统计问题中未知总体的所有可能的分布不能够通过有限个未知参数的变化来描述。
- 数据类型及适用的统计检验
  - 曼-惠特尼U检验
  - 维尔克松T检验
  - 符号检验法
  - 克-瓦式单向方差分析

2

一个例子

- 为比较男孩和女孩的反应时，抽到一个5个男孩和8个女孩的样本
  - 男孩：23 18 29 42 21
  - 女孩：37 56 39 34 26 104 48 25

3

一个例子

- 11个被试来自不同职业，他们的创造力结果如下，问不同职业的创造力是否有显著差异

教师	公司管理	机关行政
128	90	89
114	91	80
103	106	101
92		
85		

4

参数检验与非参数检验

- 参数检验
  - 以明确的总体分布为前提
  - 需要满足某些总体参数的假定条件
- 非参数检验
  - 不依赖于特定的总体分布，无须对总体参数规定条件
  - 观察独立
  - 连续分布

5

数据类型及适用的统计检验

数据类型	统计检验			
等比/等距型	独立样本 t 检验	相关样本 t 检验	方差分析	积差相关
顺序型	曼-惠特尼 U 检验 (The Mann-Whitney U Test)	维尔克松 T 检验 (The Wilcoxon T Test)	克-瓦式单向方差分析	等级相关
类目型	卡方独立性检验	符号检验法	卡方匹配度性检验	卡方独立性检验

6

顺序型量度的数据

- 例：
  - 幼儿园的老师将孩子的成熟度排成名次
  - 人事经理将其员工的创造性排成名次
- 特点：
  - 对评定者的要求较低
  - 不复杂
  - 数据易理解且有吸引力
  - 较易得到

7

如何获得顺序型（等级）量度

- 身高
  - 不必量出高度的绝对测量数据
  - 只需将一组人中最高的挑出来，再将其余的人中最高的挑出来
  - 即只需对两个人谁比较高的判断
- 美貌，才干
  - 不易精确定义和绝对测量
  - 较易对两个人谁比较高进行判断

8

将数值型变量转化成顺序型（等级）量度

- 等级比较简单。
  - “我妹妹1.67米”
  - “我妹妹比我高一点点”
- 原始分数可能违反了特定统计程序的某些假定
  - t检验和方差分析假定数据来自正态分布
  - 如果原始分数违反了正态假定，安全的方法是将原始的等距/等比分数转换成顺序型（等级）量度
- 实验偶尔会产生不确定分数
  - 大鼠可能在规定的时间中走不出迷宫

9

将原始分数转换成顺序型数据

分数	位置	秩
3	1	1.5
3	2	1.5
5	3	3
6	4	5
6	5	5
6	6	5
12	7	7

10

小练习

- 将下列分数转换成顺序性数据
  3、4、4、7、9、9、9、12

11

数据类型及适用的统计检验

数据类型	统计检验			
等比/等距型	独立样本t检验	相关样本t检验	方差分析	积差相关
顺序型	曼-惠特尼 U 检验 (The Mann-Whitney U Test)	维尔克松T检验 (The Wilcoxon T Test)	克-瓦式单向方差分析	等级相关
类目型	卡方独立性检验	符号检验法	卡方匹配度性检验	卡方独立性检验

12

曼-惠特尼U检验 (The Mann-Whitney U Test)

- 用于两个独立样本的检验
- 顺序型数据
  - 如果两个样本的差异是反映真实的总体差异，将两个样本合并后，所有的分数排序为一条线，然后来自一个样本的分数应当集中于线的一端，而来自另一个样本的分数应当集中于线的另一端
  - 如果两个样本没有差异，那么两个样本合并后大分数和小分数应当均匀地混合在一起，因为没有理由假定一组分数会大于另一组

13

曼-惠特尼U检验 (The Mann-Whitney U Test)

- 例：两个独立样本，每个n=6
  样本A：27，2，9，48，6，15
  样本B：71，63，18，68，94，8

14

计算曼-惠特尼U检验的步骤

1. 对每种处理条件各得到一个独立的样本，以n_A表示样本A中的被试数目，以n_B表示样本B中的被试数目；
2. 将两个样本合并，将所有被试n_A +n_B排序；
3. 确定来自两个样本的分数在混合排序中是否系统性地聚集在度量的两端？

15

例子

等级	分数	样本		样本A的点数	样本B的点数
1	2	A		6	
2	6	A		6	
3	8		B		4
4	9	A		5	
5	15	A		5	
6	18		B		2
7	27	A		4	
8	48	A		4	
9	63		B		
10	68		B		
11	71		B		
12	94		B		

16

例子

- U_A=30 U_B=6
- U=6
- U_A +U_B = n_A* n_B

17

曼-惠特尼U检验的公式 (The Mann-Whitney U Test)

- $U_A= n_An_B+[n_A（n_A+1）/2]-\sum R_A$
- $U_B= n_An_B+[n_B（n_B+1）/2]-\sum R_B$
- 选择较小的U
- $U_A +U_B = n_A* n_B$

U_A= 6*6+[6（6+1）/2]-27=30
U_B= 6*6+[6（6+1）/2] –51=6
U =6

18

用曼–惠特尼U作假设检验

- 曼-惠特尼检验的虚无假设：两处理之间无系统差异。
- 如果Mann-Whitney U=0，其中一个样本不得分，两个样本无重叠，有最大的差异
- 当两个样本越接近时， Mann-Whitney U越大
- 所以，如果 $U_{obs} <= U_{crit}$，才能拒绝H_0 （与参数检验正好相反）

19

例子

- 为比较男孩和女孩的反应时，抽到一个5个男孩和8个女孩的样本

男孩：23 18 29 42 21

女孩：37 56 39 34 26 104 48 25

20

例：为比较男孩和女孩的反应时，抽到一个 5 个男孩和 8 个女孩的样本：

男孩： 23 18 29 42 21

女孩： 37 56 39 34 26 104 48 25

1 H_0：男孩和女孩的反应时无系统性差异；

H_1：男孩和女孩的反应时有系统性差异。$\alpha = .05$，无方向性

2 n_A=5，n_B=8，查表得 U_{crit}=6（注意如果 $U_{obs} < U_{crit}$，才能拒绝 H_0。）

3 男孩称为样本 A， 女孩称为样本 B

分数	18	21	23	25	26	29	34	37	39	42	48	56	104
秩	1	2	3	4	5	6	7	8	9	10	11	12	13
样本	A	A	A	B	B	A	B	B	B	A	B	B	B
B 的点数				2	2		1	1	1		0	0	0

4 $\Sigma R_A = 1+2+3+6+10=22$；$\Sigma R_B = 4+5+7+8+9+11+12+13=69$

$U_A = n_A n_B + [n_A(n_A+1)/2] - \Sigma R_A = 5*8+5*6/2-22=55-22=33$； $U_B = 40+36-69=7$

5 $U_B = 7 > U_{crit}$，所以接受 H_0. 男孩和女孩的反应时无系统性差异

21

Mann-Whitney U的正态近似

- 当n>20, Mann-Whitney U统计量接近正态分布
- $\mu = n_A n_B/2$
- σ= sqrt（$n_A n_B$（n_A+n_B+1）/12）

$$z = \frac{X-\mu}{\sigma} = \frac{U - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B(n_A+n_B+1)}{12}}}$$

22

例子

- 用上例，n_1=5，n_2=8

  U_A=33，U_B=7 →U=7

  $\mu = n_A n_B/2=20$

  σ = sqrt（$n_A n_B$ （n_A+n_B+1）/12）

  = sqrt （560/12）=6.831

  $$Z = \frac{7-20}{6.831} = \frac{-13}{6.831}$$

  = -1.90

  ABS(Z_{obs}) ＜ABS($Z_{.05/2}$)

  接受H0，男孩和女孩的反应时并无显著差异

23

曼–惠特尼U检验的统计前提

- 不要求正态分布
- 不要求方差同质
- 要求观察独立
- 要求变量是连续的，即较少相同的等级

24

数据类型及适用的统计检验

数据类型	统计检验			
等比/等距型	独立样本 t 检验	相关样本 t 检验	方差分析	积差相关
顺序型	曼-惠特尼 U 检验 (The Mann-Whitney U Test)	维尔克松 T 检验 (The Wilcoxon T Test)	克-瓦式单向方差分析	等级相关
类目型	卡方独立性检验	符号检验法	卡方匹配度性检验	卡方独立性检验

25

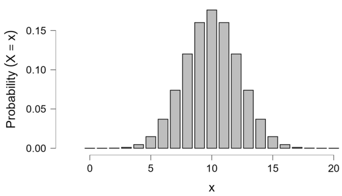
符号检验

- 符号检验：
  - 比较两个相关（配对）样本的差异，数据来自顺序量表，它将中位数作为集中趋势的量度，两样本每对数据之差用正，负号表示，做单样本的二项分布检验判断正负号数是否存在显著性差异。利用正，负号的数目对某种假设做出判定的非参数检验方法。

26

符号检验

- $p = \Pr(X > Y)$,
- H_0: $p = 0.50$，即对每对数据(x_i, y_i)，$p(x_i > y_i) = p(y_i > x_i) = 0.50$
- W是样本中$y_i - x_i > 0$的数量，假设H_0 为真，则$W \sim B(n, 0.5)$



概率分布函数

27

符号检验

- 某研究者测定噪声刺激前后15名成人被试的心率变化，问：噪声对心率有无显著影响？

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
刺激前	61	70	68	73	85	81	65	62	72	84	76	60	80	79	71
刺激后	75	79	85	77	84	87	88	76	74	81	85	78	88	80	84
差值	−14	−9	−17	−4	1	−6	−23	−14	−2	3	−9	−18	−8	−1	−13
符号	−	−	−	−	+	−	−	−	−	+	−	−	−	−	−

28

符号检验

- 1. 提出假设
  - H_0：噪声刺激前后被试的心率差值 d 总体的中位数 = 0
  - H_1：噪声刺激前后被试的心率差值 d 总体的中位数 ≠ 0

29

符号检验

- 2. 根据虚无假设H_0所提供的前提条件，选择合适的统计模型。
  - 样本量n = 15的配对样本。

30

符号检验

- 3. 规定显著性水平α，α确定后，否定域也随之被确定了。
  - 指定小概率的最大限为0.05。

31

符号检验

- 4. 计算检验统计量的值。
  - 计算差值并赋以符号，$d > 0$，记为“+”，$d < 0$,记为“-”，$d = 0$记为“0”，统计各符号个数，记为n_+, n_-, n_0,$N = n_+ + n_-$，检验的统计量为k，$k = \min(n_+, n_-)$

32

符号检验

- 4. 计算检验统计量的值。
  - $n_+ = 2, n_- = 13, n_0, N = n_+ + n_- = 15$，检验的统计量为k，$k = \min(n_+, n_-) = n_+ = 2$

33

符号检验

- 5.做出决策。根据显著性水平α和检验统计量的分布，查相应的统计表，确定接受域和否定域的临界值，用计算出的统计量值与临界值做比较，从而做出拒绝或不拒绝虚无假设的决策。
  - $N = 15$, $k_{0.05(2)} = 2$, $p < 0.05$
  - 拒绝虚无假设，噪声刺激对心率的影响达到了0.05的显著性水平。

34

例：一位研究者选取了36 个学生考察他们在两种教学方法下的结果，数据如下：

学生	教学方法 A	教学方法 B	差异	学生	教学方法 A	教学方法 B	差异
001			A<B	019			A<B
002			A>B	020			A<B
003			A<B	021			A>B
004			A<B	022			A<B
005			A>B	023			A<B
006			A<B	024			A>B
007			A<B	025			A<B
008			A<B	026			A<B
009			A>B	027			A<B
010			A<B	028			A>B
011			A<B	029			A<B
012			A>B	030			A<B
013			A<B	031			A>B
014			A<B	032			A<B
015			A>B	033			A>B
016			A>B	034			A<B
017			A>B	035			A<B
018			A<B	036			A>B

35

符号检验法 (The Sign Test)

- 点数整个样本(n)中正的差异的数目 ， 然后用p=q=1/2的二项检验

36

符号检验法例题

1 H_0: 两种教学方法下的结果无系统性差异；p=q=1/2

H_1：两种教学方法下的结果有系统性差异。p≠q

α＝.05

2 pn=qn=18, z_{crit}=+-1.96

3 z＝(X-pn)/sqrt(npq)=(23-18)/ sqrt(36*1/2*1/2)=5/3=1.67 < +1.96

4 接受 H_0：两种教学方法下的结果无系统性差异

37

数据类型及适用的统计检验

数据类型	统计检验			
等比/等距型	独立样本 t 检验	相关样本 t 检验	方差分析	积差相关
顺序型	曼-惠特尼 U 检验 (The Mann-Whitney U Test)	维尔克松 T 检验 (The Wilcoxon T Test)	克-瓦式单向方差分析	等级相关
类目型	卡方独立性检验	符号检验法	卡方匹配度性检验	卡方独立性检验

38

维尔克松T检验 (The Wilcoxon T Test)

- 检验重复测量设计的两种处理条件之间的差别
- 对差异分数样本的绝对值进行排序
  - 虚无假设认为两处理间没有显著差异，如果虚无假设正确，样本数据之间的差异都是由机会造成。所以正的和负的差异应当均匀地混合在一起
  - 相反，如果两处理间有系统差异，会造成持续出现正的差异或持续出现负的差异

39

维尔克松T检验 (The Wilcoxon T Test)

被试	处理1	处理2	差异	等级
1	18	43	+25	6
2	9	14	+5	2
3	21	20	-1	1
4	30	48	+18	5
5	14	21	+7	3
6	12	4	-8	4

- 将差异分数排序，忽略正负号 (+ 或 -)，然后分别计算正的差异分数的秩次和 以及负的差异分数的秩次和。
- Wilcoxon T 就是较小的那个和
- 如果 $T_{obs} < T_{crit}$, 才能拒绝H_0
- 如果T=0，所有的差异值都是正的（负的）

40

维尔克松T检验

学生	教学方法 A	教学方法 B
001	8	24
002	12	10
003	15	19
004	31	52
005	26	20
006	32	40
007	19	29

41

例：一位研究者选取了 7 个学生考察他们在两种教学方法下的结果。数据如下：

学生	教学方法 A	教学方法 B	差异	差异绝对值排序
001	8	24	+16	6
002	12	10	-2	1
003	15	19	+4	2
004	31	52	+21	7
005	26	20	-6	3
006	32	40	+8	4
007	19	29	+10	5

1 H_0：两种教学方法下的结果无系统性差异；

H_1：两种教学方法下的结果有系统性差异。

α＝.05

2 n=7, α＝.05

查表得 T_{crit}=2（注意如果 $T_{obs} < T_{crit}$, 才能拒绝 H_0。

3 计算差异和差异绝对值排序

ΣR_+= 6+2+7+4+5＝24；ΣR_-＝1+3＝4； T＝4

4 T＝4 > T_{crit}. 所以接受 H_0. 两种教学方法下的结果无系统性差异

42

相同的等级和0分数

- 在Wilcoxon 检验中，有两类相同的等级
  - 一个被试在处理1和处理2中所得的分数相同，得到的差异分数为0
  - 两个或以上的被试得到相同的差异分数（无论正负号）
- 一些统计学家主张去掉那些差异分数为0的被试，将样本容量相应减少
- 另一种程序是将0差异分数均匀地分配在正负两组中，但这种作法增大T值，使H0更难被拒绝

43

例子

学生	前测	后测	差异分数	去掉 0 的等级	保留 0 的等级
A	18	18	0	-	1.5
B	24	24	0	-	1.5
C	31	30	-1	1	3
D	28	24	-4	2	4
E	17	24	+7	3	5
F	16	24	+8	4	6
G	15	26	+11	5.5	7.5
H	18	29	+11	5.5	7.5
I	20	36	+16	7	9
J	9	28	+19	8	10

44

例子

- 将差异分数为0的被试A，B去掉
- N=8，alpha=0.05，Tcrit=3
- 正的差异分数的等级分别是3，4，5.5，5.5，7，8，所以
  ∑R+=33
- 负的差异分数的等级分别是1和2，所以
  ∑R-=3
  Wilcoxon T=3
  因为Tobs<= Tcrit
  所以拒绝虚无假设，学生的前后测分数有显著差异

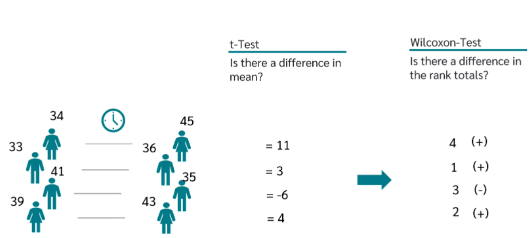
45

符号秩次检验

- 符号秩次检验（signed rank test，Wilcoxon test）：经过改进的符号检验，多用于配对样本检验，适用条件与符号检验一致，除了比较差值符号，还比较各对差值大小的秩次高低，剔除d = 0的该对数据，先按差值的绝对值大小排秩次，再将正负号标在秩次前，比较正秩和负秩的秩次和。

46

符号秩次检验



https://www.youtube.com/watch?v=NZsL2eDQiDQ&t=689s

47

符号秩次检验

- 某研究者测定噪声刺激前后15名成人被试的心率变化，问：噪声对心率有无显著影响?

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
刺激前	61	70	68	73	85	81	65	62	72	84	76	60	80	79	71
刺激后	75	79	85	77	84	87	88	76	74	81	85	78	88	80	84
差值	−14	−9	−17	−4	1	−6	−23	−14	−2	3	−9	−18	−8	−1	−13
秩次	4	7	3	11	14	10	1	4	13	12	7	2	9	14	6
符号	−	−	−	−	+	−	−	−	−	+	−	−	−	−	−

48

符号秩次检验

- 1. 提出假设
  - H_0:噪声刺激前后被试的心率差值 d 总体的中位数 = 0
  - H_1:噪声刺激前后被试的心率差值 d 总体的中位数 ≠ 0

49

符号秩次检验

- 2. 根据虚无假设H_0所提供的前提条件，选择合适的统计模型。
  - 样本量n = 15的配对样本。

50

符号秩次检验

- 3. 规定显著性水平α， α确定后，否定域也随之被确定了。
  - 指定小概率的最大限为0.05。

51

符号秩次检验

- 4. 计算检验统计量的值。
  - 编秩次，定符号。
  - 确定正的秩次和T_+, T_-
  - $T = \min(T_+, T_-)$

52

符号秩次检验

- 5. 做出决策。根据显著性水平α和检验统计量的分布，查相应的统计表，确定接受域和否定域的临界值，用计算出的统计量值与临界值做比较，从而做出拒绝或不拒绝虚无假设的决策。
- T ＜ $T_{0.05(15)}$， $P < 0.05$，拒绝虚无假设，噪声对心率有显著影响

53

数据类型及适用的统计检验

数据类型	统计检验			
等比/等距型	独立样本 t 检验	相关样本 t 检验	方差分析	积差相关
顺序型	曼-惠特尼 U 检验 (The Mann-Whitney U Test)	维尔克松 T 检验 (The Wilcoxon T Test)	克-瓦式单向方差分析	等级相关
类目型	卡方独立性检验	符号检验法	卡方匹配度性检验	卡方独立性检验

54

克-瓦式单向方差分析

- $H=12/（N（N+1））\sum R^2/n-3（N+1）$
  R：每一组数据的等级和

55

例子

- 11个被试来自不同职业，他们的创造力结果如下，问不同职业的创造力是否有显著差异

教师	公司管理	机关行政
128	90	89
114	91	80
103	106	101
92		
85		

56

例子

教师	R1	公司	R2	机关	R3
128	11	90	4	89	3
114	10	91	5	80	1
103	8	106	9	101	7
92	6				
85	2				
ΣR	37		18		11

$H=(12\times(37^2/5+18^2/3+112/3))/(11\times12)-3\times(11+1)=38.38-36=2.38$

H0.05=4.533

接受虚无假设，三组之间没有差异。

57

例子

- 15个被试在3种处理条件的排序如下,问3种处理条件有没有系统的区别

处理1	处理2	处理3
3	1	9
4	2	11
7	5	13
10	6	14
12	8	15

58

例子

- R1=36; R2=22; R3=62
- H=12/15*16(36*36/5+22*22/5+62*62/5-3*16=0.05*1124.8-48=8.24

59

费里德曼双向方差分析

$$\chi_r^2=\frac{12}{nk(k+1)}\sum R^2-3n(k+1)$$

60

非参数检验的优缺点

- 优点：
  - 一般不需要严格的假设前提。
  - 稳定性，对个别较大的偏离数据不太敏感。
  - 运算简单，容易理解。
  - 适用于小样本，无分布样本，数据污染样本，混杂样本，预实验。

61

非参数检验的优缺点

- 缺点：
  - 未能充分利用资料的全部信息，对数据变化不敏感，与参数检验相比，犯二类错误β的概率大一点。
  - 对于大样本数据，如不采用适当的近似计算，运算会十分庞杂。
  - 不能处理变量间的交互作用。

62

参数检验 VS 非参数检验

自变量	因变量	参数检验	非参数检验
定量变量	定量变量	皮尔逊相关	斯皮尔曼等级相关
定性（分类）变量	定性（分类）变量	皮尔逊相关	卡方检验
两组定性变量	来自不同样本的定量变量	独立样本t检验	秩和检验，中位数检验
两组定性变量	来自同一样本的定量变量	配对样本t检验	符号检验，符号秩次检验
三组或三组以上的定性变量	定量变量	方差分析	Kruskal-Walls H检验
三组或三组以上的定性变量	两个或两个以上的定量变量	多元方差分析	ANOSIM(Analysis of similarities)

63

置换检验

- 为了缓解当下年轻人普遍的焦虑状况，某临床心理学家发明了一种新的疗法。他想要对疗法的效果进行测试，于是选取了10名有焦虑症状的被试，在确保这10名被试接受测试前的焦虑水平同质的情况下，将这两组被试随机分为两组，一组是治疗组，一组是控制组。接受治疗后，使用焦虑自评量表(SAS)进行测量，两组得分如下：

治疗组	34	46	47	50	43
控制组	49	56	48	60	53

64

置换检验

- 我们是否可以使用之前学到的检验方法来对治疗效果进行检验？
  - z检验
  - t检验
  - f检验

65

置换检验

- 无论是z检验，t检验，还是F检验，都有一条最基本的假定——总体呈正态分布。
- 但是我们并不知道这些被试的总体的分布情况。那么我们该怎么对治疗效果进行检验呢？

66

置换检验

- 置换检验(Permutation test)是Fisher于20世纪30年代提出的一种基于大量计算，利用样本数据的随机排列，进行统计推断的方法。
- 因其对总体分布自由，应用较为广泛，特别适用于总体分布未知的小样本资料，以及某些难以用常规方法分析资料的假设检验问题。

67

置换检验

- 在介绍置换检验的原理之前，我们先以t检验为例，回顾一下参数检验的基本原理。

68

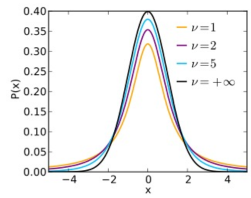
参数检验

- 假定治疗组和控制组总体呈正态分布；
- 虚无假设为治疗组和控制组之间没有差异；
- 使用独立样本t检验的方法来计算两组差异的t值以及对应的自由度；
- 将计算得到的t值与对应自由度下的t分布进行比较，得到p值，判断p值是否小于显著性水平；
- 若p值是否大于显著性水平，则没有证据能够拒绝虚无假设，若p值是否小于显著性水平，则有证据能够拒绝虚无假设。

69

参数检验

- 参数检验的核心就是计算统计量在该统计量分布中的位置，t检验是基于t分布来进行计算，z检验是基于正态分布进行计算，F检验是基于F分布来进行计算。



t分布示意图
https://en.wikipedia.org/wiki/Student%27s_t-distribution

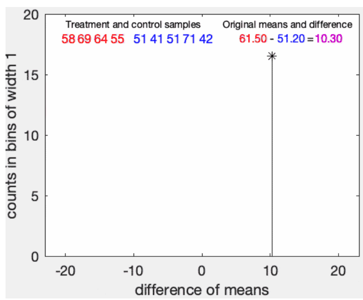
70

置换检验

- 除了上节课提到的不对参数分布做要求的非参数检验方法外，置换检验通过将样本数据随机排列来构造新的分布。
- 那么置换检验是怎样构造新的分布的呢？

71

置换检验的示意图



https://commons.wikimedia.org/wiki/File:Permutation_test_example_animation.gif

72

置换检验的原理

- 首先，我们做出假设：

$$H_0: \mu_{\text{treatment}} = \mu_{\text{control}}$$
$$H_1: \mu_{\text{treatment}} \neq \mu_{\text{control}}$$

  - 虚无假设表明新的治疗方法没有效果，
  - 备择假设表明新的治疗效果存在效果

73

置换检验的原理

- 然后，我们计算两组均值之差

d = -9.2

治疗组	34	46	47	50	43
控制组	49	56	48	60	53

74

置换检验的原理

34	46	47	50	43
49	56	48	60	53

↓ 将原数据进行重新随机排列

46	48	34	60	49	53	47	56	50	43

↓ 将重新排列后的结果分为两组

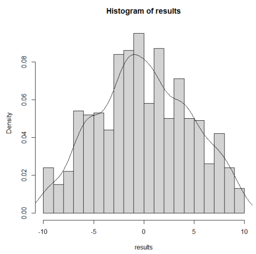
46	48	34	60	49
53	47	56	50	43

d = -2.4 计算两组均值的差值

75

置换检验的原理

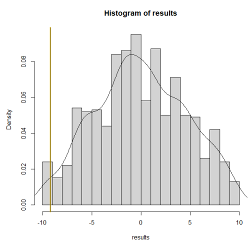
- 重复该过程i次，每一次我们都会得到一个均值的差值d_i. 最后我们会得到i个d_i形成的新分布。



76

置换检验的原理

- 最后，我们将真实的治疗组与控制组之间的差值d 的相对位置在新分布中表示出来，并计算小于等于该值的差值d_i数量占所有d_i的比例。



$p = 0.015$

- 因此我们提供了证据支持新的治疗效果对缓解焦虑有效。

77

置换检验的应用

- 一个可视化的例子：羊毛产量会受到清洁剂的影响吗？

https://www.jwilber.me/permutationtest/

78

置换检验的应用

- 在心理学研究中，由于临床数据和神经信号的数据多为小样本数据，且一般都不服从正态分布，故置换检验能发挥很大的作用。

79

作业

1. 将下列数据转换成等级数据
   14，3，4，0，3，5，14，3
2. 一个实验用两个样本，一个n=25，另一个n=10，得到 Mann-Whitney U=50，问另一个样本的U值是多少？
3. 为检验一种风湿药的疗效，得到10个病人服药后握力的差异分数为+3，+46，+16，-2，+38，+14，0，-8，+25，+41。在.05水平检验差异是否达到显著
4. 发展心理学家观察了4岁男童和女童与其它儿童交往的频率，问其是否有显著性差异
   男童分数： 8　17　14　21
   女童分数：18　25　23　21　34　28　32　30　13

80