两样本比较

主讲人：刘宏志
liuhz@ss.pku.edu.cn

z检验和t检验

z 检验的应用条件：

(1) 样本来自正态总体

(2a) 样本含量 n 较大，或

(2b) n 虽小但总体标准差 σ 已知

t 检验的应用条件：

(1) 总体标准差 σ 未知；

(2) 样本含量 n 较小；

(3) 样本来自正态总体；

(4) 两样本均数比较时方差齐，

即 $\sigma_1^2 = \sigma_2^2$

在 H_0 成立的前提条件下，检验统计量计算公式：

① σ 已知或 σ 未知但n足够大：

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \qquad (\nu = \infty)$$

② σ 未知且n较小：

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \qquad (\nu = n - 1)$$

H_0:样本均值与已知总体均值 μ_0 无差别

配对 t 检验(paired t-test)

- 配对设计：
  - 两组观察对象除了研究因素不同外，其它的可能影响研究结果的因素相同或相似
- 四种情况：
  - 两个同质受试对象分别接受两种不同的处理
  - 同一受试对象分别接受两种不同的处理
  - 同一受试对象接受某种处理的前后数据
  - 同一受试对象的两个不同部位的数据

配对 t 检验(paired t-test)

基本原理:

假设两种处理的效应相同,

即 $\mu_1=\mu_2$,则 $\mu_1 - \mu_2 =0$

（即已知总体均数 $\mu_d=0$）

检验：差数的样本均数 $\bar{d}$ 与所代表的未知总体均数 μ_d 与 0 的比较

目的 ：推断两种处理的效果有无差别或

推断某种处理有无作用

应用条件：差值d服从正态分布

公式：$$t = \frac{\bar{d}}{S_d / \sqrt{n}}$$

上式中 d 表示差值， $\nu = n-1$ （n 为对子数）

例：某医生用A、B两种血红蛋白测定仪器检测了16名健康男子的血红蛋白含量(g/L)检验结果见下表，问两种血红蛋白测量仪器检测结果是否有差别?

两种仪器检测16名男青年血红蛋白含量(g/L)结果

被检测者号	仪器A	仪器B	d	d^2
(1)	(2)	(3)	(4)=(3)-(2)	(5)
1	113	140	27	729
2	125	150	25	625
3	126	138	12	144
4	130	120	−10	100
5	150	140	−10	100
6	145	145	0	0
7	135	135	0	0
8	105	115	10	100
9	128	135	7	49
10	135	130	−5	25
11	100	120	20	400
12	130	133	3	9
13	110	147	37	1369
14	115	125	10	100
15	120	114	−6	36
16	155	165	10	100
合计			$\Sigma d=130$	$\Sigma d^2=3882$

[分析]

- 每人均用两种方法检测血红蛋白，即采用配对方式试验
- 假设两检测方法无差别，则两方法检测值的差应为0
- 由于抽样误差的影响，可导致两方法检测值差值不为0
- 以差值为观察对象，检验差值样本是否来自零总体(μ_d=0)
- 如来自零总体，则两方法检测值相同
- 如不是来自零总体，则表明两方法检测值的不一致不是由抽样误差引起，而是来自不同的总体

(1) 建立检验假设，确定检验水准

$H_0: \mu_d=0$，即两方法检测结果相同

$H_1: \mu_d \neq 0$，即两方法检测结果不同

$\alpha = 0.05$，双侧检验

(2) 选定检验方法，计算检验统计量

差值构成样本与总体之间的比较，可用样本-总体的t检验。依公式计算检验统计量：

$$s_d=\sqrt{\frac{\sum d^2-\frac{(\sum d)^2}{n}}{n-1}}=\sqrt{\frac{3882-\frac{(130)^2}{16}}{16-1}}=13.73(g/L)$$

$$t=\frac{\bar{d}-\mu_{\bar{d}}}{s_{\bar{d}}}=\frac{\bar{d}-0}{s_d/\sqrt{n}}=\frac{8.125}{13.73/\sqrt{16}}=2.367$$

$$\upsilon=n-1=16-1=15$$

（3）确定P值，作出推断结论

以 υ=15， t=2. 367, 查t值表 $t_{0.05/2(15)}$=2. 131。

$t>t_{0.05/2(15)}$, 则$P<0.05$。拒绝H_0，接受H_1, 差异有统计学意义。可认为两种方法检查结果不同。

两独立样本均数的比较

(two-sample test)

两样本均为随机抽样得到的样本

或 采用随机分组得到的样本

t 检验

目的：推断两样本均数分别代表的总体均数 μ_1 与 μ_2 有无差别

适用条件 ：

- 随机抽样的小样本（ σ 未知）
- 两样本来自正态总体
- 两样本的总体方差齐同（ $\sigma_1^2 = \sigma_2^2$ ）

注： 方差齐性的经验判断方法

若 $s_1^2 / s_2^2 \geq 3$

或 $s_1 / s_2 \geq 2$

可怀疑两样本总体方差不等

$s_1^2 / s_2^2 \geq 5$ 可认为两样本总体方差不等

否则可认为两总体方差相等

两样本t检验的统计量在 H_0： $\mu_1 = \mu_2$

的条件下为:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{s_{\overline{x}_1 - \overline{x}_2}}$$

$$\nu = n_1 + n_2 - 2$$

合并标准误差的计算为： $s_{\bar{x}_1-\bar{x}_2} = \sqrt{s_c^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$

两组的共同方差—合并方差s_c^2计算为：

$$s_c^2 = \frac{\left[\sum x_1^2 - \frac{(\sum x_1)^2}{n_1}\right] + \left[\sum x_2^2 - \frac{(\sum x_2)^2}{n_2}\right]}{n_1 + n_2 - 2}$$

$$= \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$$

例：男女大学生的血清谷胱甘肽过氧化酶(GSH-PX)

性别	例 数	均 数	标准差
男	48	96.53	7.66
女	46	93.73	8.23

(1) 建立检验假设，确定检验水准

$H_0: \mu_1 = \mu_2$，即男女的GSH-PX含量两总体均数相同

$H_1: \mu_1 \neq \mu_2$，即男女的GSH-PX含量两总体均数不同

$\alpha = 0.05$，双侧检验

（2）选定检验方法，计算检验统计量

由于两组样本量<100，且方差齐，故选用t检验。

已知：$n_1 = 48, \overline{x_1} = 96.53, s_1 = 7.66$

$n_2 = 46, \overline{x_2} = 9373, s_2 = 8.23$

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{96.53-93.73}{\sqrt{\frac{7.66^2(48-1)+8.23^2(46-1)}{48+46-2}\left(\frac{1}{48}+\frac{1}{46}\right)}} = 1.708$$

$$\upsilon = n_1 + n_2 - 2 = 48 + 46 - 2 = 92$$

北京大学

（3）确定P值，作出推断结论

以 υ= 48 +46 － 2 =92 查 t 界值表，

t =1.708 ＜ $t_{0.05/2(92)}$= 2.000， P ＞ 0.05，

按 α=0.05水准，不拒绝H_0 ， 即差异无统计学意义。可认为男女的GSH-PX含量相同。

z 检验

z 检验是 t 检验的特例，其检验方法与 t 检验方法比较，有以下区别：

① 由于z 检验是大样本资料的检验，故其样本量可以看作无穷大，这时，其样本均数的分布已由 t分布转为正态分布。依此，确定P 值时，理论上$t_{0.05/2,\ v}$ （或$t_{0.01/2,\ v}$）可以用 1.96（ 或 2.58 ）来代替。

②在大样本的情况下，两样本均数比较的合并标准误差，可以简化为 $s^2_{\bar{x}_1-\bar{x}_2}=s^2_{\bar{x}_1}+s^2_{\bar{x}_2}$

即为：

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{s^2_{\bar{x}_1}+s^2_{\bar{x}_2}}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

例： 某地抽查了25～29岁正常人群的红细胞数，测得其结果如下表，问该人群男、女红细胞数是否不同？

某地240名正常人群红细胞数（$\times 10^{12}$/L）

组别	n	$\bar{x}$	s
男	156	4.65	0.55
女	74	4.22	0.44

① 建立检验假设，确定检验水准

$H_0: \mu_1 = \mu_2$，即该地男、女红细胞数相同

$H_1: \mu_1 \neq \mu_2$，即该地男、女红细胞数不同

α=0. 05, 双侧检验

② 选定检验方法，计算检验统计量

由于两样本样本量均>100，故符合z 检验的条件，计算z 值

$$z = \frac{\overline{x_1} - \overline{x_2}}{s_{\overline{x_1} - \overline{x_2}}} = \frac{4.65 - 4.22}{\sqrt{\frac{0.55^2}{156} + \frac{0.44^2}{74}}} = 6.37$$

北京大学

③ 确定P 值，作出推断结论

$z = 6.37 > 1.96$，故$P < 0.05$，拒绝H_0，接受H_1，差异有统计学意义。即可认为该人群男、女红细胞数不同。

两总体均数比较

方差齐性检验

方差齐

方差不齐

t 检验、z检验

t′检验

前提：来自正态总体

方差齐性检验

方差齐性检验的计算公式为：

$$F = \frac{s_1^2\ (较大)}{s_2^2\ (较小)}$$

$$\upsilon_1 = n_1 - 1$$

$$\upsilon_2 = n_2 - 1$$

若两样本是来自同一个正态总体，则它们的方差不应相差过大，其$F \geqslant \rightarrow 1$。

由于抽样误差的存在，其F可能偏离于1，当其偏离过大，超出抽样误差所能引起的范围，则表明方差不齐

例：两组大鼠血糖含量测定结果(mmol/L)

组别	例数	均数	标准差
硫酸氧钒	12	6.5	1.34
空白对照	8	13.7	4.21

(1) 建立检验假设，确定检验水准

H_0: $\sigma_1^2 = \sigma_2^2$，即两组大鼠血糖含量总体方差相等

H_1: $\sigma_1^2 \neq \sigma_2^2$，即两组大鼠血糖含量总体方差不等

$\alpha = 0.05$, 双侧检验

(2) 选定检验方法，计算检验统计量

$$F = \frac{s_1^2}{s_2^2} = \frac{4.21^2}{1.34^2} = 9.87$$

$$\nu_1 = n_1 - 1 = 8 - 1 = 7$$

$$\nu_2 = n_2 - 1 = 12 - 1 = 11$$

(3) 确定P值，作出推断结论

以 $\upsilon_1=7$， $\upsilon_2=11$，$F=9.87$ 查附表6 ，F 界值表，有 $9.87>3.01=F_{0.05,(7,11)}$，故 $P<0.05$。按 $\alpha=0.05$水准，拒绝H_0，接受H_1 差异有统计学意义。故可认为两组大鼠血糖含量总体方差不齐。(故该资料不可直接用方差相等的两样本的t 检验)

t' 检验 — 近似t检验

- 基本思想：
  - 在方差不齐的情况下进行比较
  - 样本均数的分布曲线由t分布转化为t'分布
  - t'分布较复杂，故用t分布的临界值估计t'分布的临界值，即对临界值校正后依t检验进行分析
- Cochran & cox 法： 对临界值校正
- Satterthwaite 法
- welch 法

Satterthwaite 法、welch 法：对自由度校正

北京大学

Cochran & cox 法

计算公式： $$t' = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$t'_a = \frac{s_{\overline{x}_1}^2 t_{a,\upsilon_1} + s_{\overline{x}_2}^2 t_{a,\upsilon_2}}{s_{\overline{x}_1}^2 + s_{\overline{x}_2}^2}$$

$$\upsilon_1 = n_1 - 1$$
$$\upsilon_2 = n_2 - 1$$

例：请检验两组大鼠血糖含量是否相同?

硫酸氧钒组： $n_1 = 12, \overline{x_1} = 6.5, s_1 = 1.34$

空白对照组： $n_2 = 8, \overline{x_2} = 13.7, s_2 = 4.21$

（1）建立检验假设，确定检验水准

$H_0: \mu_1 = \mu_2$，即两总体的血糖值相同

$H_1: \mu_1 \neq \mu_2$，即两总体的血糖值不同

$\alpha = 0.05$，双侧检验

(2) 选定检验方法，计算检验统计量

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{13.7 - 6.5}{\sqrt{\dfrac{4.21^2}{8} + \dfrac{1.34^2}{12}}} = 4.6817$$

$$\upsilon_1 = n_1 - 1 = 8 - 1 = 7$$

$$\upsilon_2 = n_2 - 1 = 12 - 1 = 11$$

$$t'_{0.05/2} = \frac{s_{\bar{x}_1}^2 t_{\alpha,\upsilon_1} + s_{\bar{x}_2}^2 t_{a,\upsilon_2}}{s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2} = \frac{\dfrac{4.21^2}{8} \times 2.365 + \dfrac{1.34^2}{12} \times 2.201}{\dfrac{4.21^2}{8} + \dfrac{1.34^2}{12}} = 2.355$$

北京大学

(3) 确定P值，作出推断结论

以 $t' = 4.6817 > t'_{0.05/2} = 2.355$ ，得P <0. 05。

按 α =0. 05水平，拒绝H_0，接受H_1，有统计学意义。即可认为两组大鼠血糖含量不同。

北京大学

Satterthwaite 法

自由度校正的计算公式为：

$$\upsilon = \frac{(s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2)^2}{\dfrac{s_{\bar{x}_1}^4}{n_1 - 1} + \dfrac{s_{\bar{x}_2}^4}{n_2 - 1}} = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

例：请检验两组大鼠血糖含量是否相同?

硫酸氧钒组：$n_1 = 12, \overline{x_1} = 6.5, s_1 = 1.34$

空白对照组：$n_2 = 8, \overline{x_2} = 13.7, s_2 = 4.21$

(1) 建立检验假设，确定检验水准

$H_0: \mu_1 = \mu_2$，即两总体的血糖值相同

$H_1: \mu_1 \neq \mu_2$，即两总体的血糖值不同

$\alpha = 0.05$, 双侧检验

(2) 选定检验方法，计算检验统计量

$$t' = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{13.7 - 6.5}{\sqrt{\frac{4.21^2}{8} + \frac{1.34^2}{12}}} = 4.6817$$

$$\upsilon' = \frac{(s_{\bar{x}_1}^2 + s_{\bar{x}_2}^2)^2}{\frac{s_{\bar{x}_1}^4}{n_1 - 1} + \frac{s_{\bar{x}_2}^4}{n_2 - 1}} = \frac{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^2}{\frac{(\frac{s_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{s_2^2}{n_2})^2}{n_2 - 1}} = 7.9542 \approx 8$$

(3) 确定P值，作出推断结论

以 $t^{'}=4.6817>t_{0.05/2,8}=2.306$，得$P$ <0.05。按 α =0.05水平，拒绝H_0，接受H_1，差异有统计学意义。即可认为两组大鼠血糖含量不同。

两总体均数比较

方差齐性检验

方差齐

方差不齐

t 检验、z检验

t'检验

前提：来自正态总体

两独立样本差别的秩和检验
Wilcoxon rank sum test

两独立样本秩和检验计算表

A样本		B样本	
观察值	秩次	观察值	秩次
7	4	3	1
14	6	5	2
22	10	6	3
36	11	10	5
40	13	17	7
48	14	18	8
63	15	20	9
98	16	39	12
n_1=8	秩和 R_1=89	n_2=8	秩和 R_2=47

假定：两组样本的总体分布形状相同

基本思想

如果两总体分布相同 → 两样本来自同一总体

↓

任一组秩和不应太大或太小

↓

T 与平均秩和 $n_0(1+N)/2$ 应相差不大

$$T=\begin{cases}\text{较小例数组的秩和}, n_1 \neq n_2 \\ \min(R_1, R_2), n_1 = n_2\end{cases}$$

$$N = n_1 + n_2$$

$$n_0 = \min(n_1, n_2)$$

北京大学

(1) 提出假设H_0：两样本来自相同总体；

H_1：两样本来自不同总体（双侧） 或H_1：样本A高于样本B（单侧）

(2) 编秩：两样本混合编秩次，求得R_1、R_2、T。

相同观察值（即相同秩，ties），不同组----平均秩次。

(3) 确定P值作结论：

①小样本：**查表法 （威尔科克森和曼恩-惠特尼检验临界值表）**

如果T位于检验界值区间内，$P > \alpha$ ，不拒绝H_0；否则，拒绝H_0

本例T =47，取 α=0.05，查表得双侧检验界值区间（49，87），T 位于区间外，P<0.05，因此在 α=0.05的水平上，拒绝H_0，接受H_1。

②大样本：**正态近似法**

$$u = \frac{|T - n_0(N+1)/2|}{\sqrt{n_1 n_2 (N+1)/12}}$$

本例$u = 2.205 > \mu_{0.05/2} = 1.96$

配对设计资料的秩检验 (Wilcoxon signed rank test)

家兔号	A照射	B照射	B-A	秩次
1	39	55	16	10
2	42	54	12	9
3	51	55	4	3
4	43	47	4	3
5	55	53	-2	-1
6	45	63	18	11
7	22	52	30	12
8	48	44	-4	-3
9	40	48	8	6
10	45	55	10	8
11	40	32	-8	-6
12	49	57	8	6
合计				T=10 （68）

1．H_0：差值的总体中位数=0，
H_1：差值的总体中位数≠0；
α=0.05

2．求差值；绝对值从小到大编秩次
（i）绝对值相等者取平均秩次；
（ii）将差值的正负标在秩次之前；
（iii）零差值时不参与编秩

3．分别求正负秩次之和，以绝对值较小者为*R*值

4．根据统计量*R*确定对应的*P*值
（i）小样本时，查表
(ii) 大样本时，正态近似

北京大学

小样本（5≤n≤ 50）时，查表(威尔科克森符号秩检验表）

若统计量T值在上、下界值范围内，其P值大于相应的概率水平

本例：T=10，n=12，查表，双侧检验的界值区间（13，65），T位于区间外，得$P<0.05$，拒绝H_0，接受H_1，故认为A，B两种照射方式造成的急性皮肤损伤程度不同，B照射的损伤程度比A照射严重。

北京大学

t检验 vs. 秩和检验

适用条件

t检验：　a.样本所在总体呈正态分布

　b.各总体方差要齐

秩和检验：　a. 不满足正态和方差齐性条件的小样本资料

　b. 总体分布类型不明的小样本资料

　c. 单向有序列联表资料

　d. 各种资料的初步分析