

两组数值变量比较的假设检验

何平平

北京大学公共卫生学院流行病学与卫生统计学系

学习目的

1. **掌握**统计学方法的基本概念、基本原理和适用条件
2. **熟记**重要的统计学方法的公式
3. **熟练**使用R/SPSS软件完成统计分析，并整理统计结果

一、假设检验的基本思想和一般步骤

(一) 假设检验的基本思想

例： 已知一般健康成年女性血红蛋白的均数为124.7g/L，某医生在某山区随机抽取了20例健康成年女性，测得她们血红蛋白如下，问：该山区健康成年女性血红蛋白均数是否与一般健康成年女性不同？

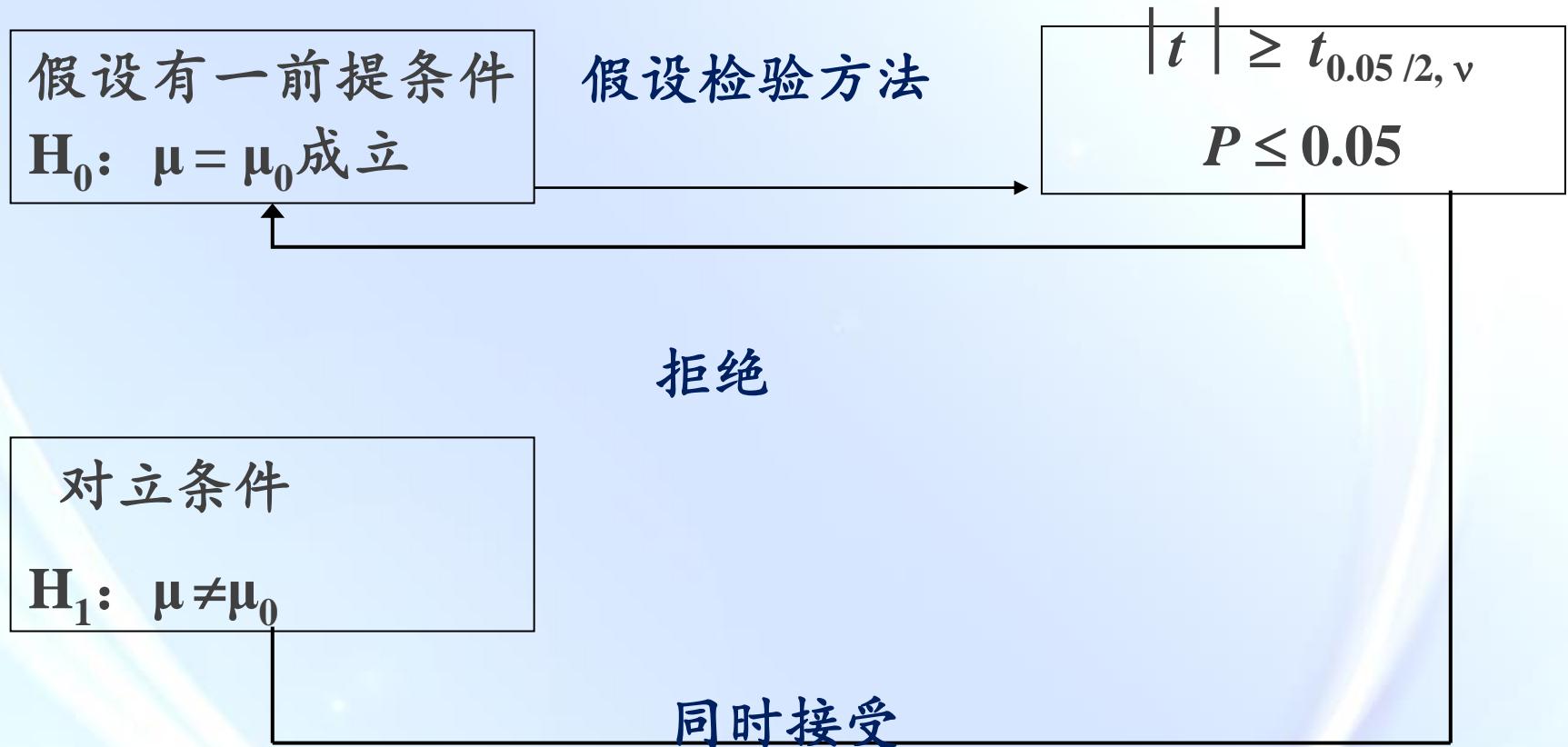
该山区20例健康成年女性血红蛋白(g/L)为： 110 100 120
115 130 120 140 110 125 100 115 125 95 115 110 125
110 115 115 105

$\bar{X} \neq \mu_0$ ，考虑差异来源有两种可能性：

- (1) $\mu = \mu_0$ （同一总体）， $\bar{X} \neq \mu_0$ 仅由抽样误差造成。
- (2) $\mu \neq \mu_0$ （不同总体）， $\bar{X} \neq \mu_0$ 主要由两总体之间的不同造成。

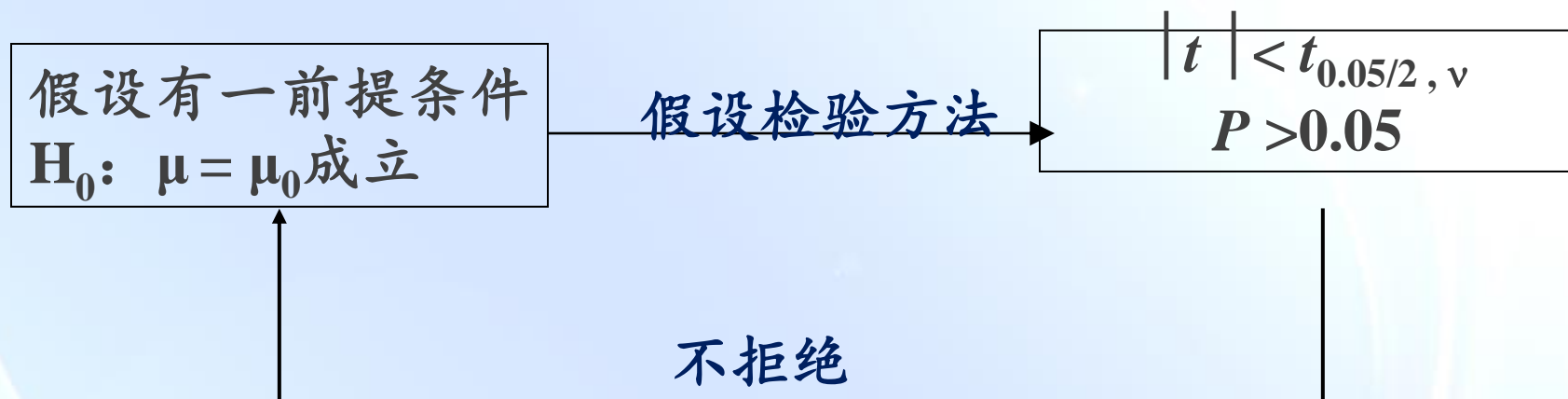
假设检验的目的： $\mu \neq \mu_0$?

小概率原理



P 值：对于 t 检验，是等于或大于现有统计量值 $|t|$ 的概率。

小概率原理



(二) 假设检验的一般步骤

1. 建立检验假设，确定检验水准

(1) 首先要明确指标的类别，是均数的比较还是率的比较。

(2) 建立检验假设。检验假设有两种：①无效假设，又称为零假设 (null hypothesis)，用符号 H_0 表示。②备择假设 (alternative hypothesis)，用符号 H_1 表示。

H_1 是与 H_0 相互对立的假设；当 H_0 被拒绝时，则接受 H_1 。 H_1 可以反映检验为单侧检验还是双侧检验。单侧检验或双侧检验的确定应结合专业知识。如果从专业的角度，判断一种方法的结果不可能低于或高于另一种方法的结果，则可以采用单侧检验（one-sided test）， H_1 为 $\mu > \mu_0$ 或 $\mu < \mu_0$ 。在不能根据专业知识判断两种结果谁高谁低时，则用双侧检验（two-sided test）， H_1 为 $\mu \neq \mu_0$ 。

样本均数与已知总体均数比较的检验假设

	目的	H_0	H_1
双侧检验	是否 $\mu \neq \mu_0$	$\mu = \mu_0$	$\mu \neq \mu_0$
单侧检验	是否 $\mu > \mu_0$	$\mu = \mu_0$	$\mu > \mu_0$
	或是否 $\mu < \mu_0$	$\mu = \mu_0$	$\mu < \mu_0$

(3) 确定检验水准。检验水准，也称为显著性水准 (significance level)，符号为 α 。 α 是预先规定的概率值，它是“是否拒绝 H_0 的界限”。研究者可以根据研究目的规定 α 的大小，通常 α 取 0.05。

2. 选定检验方法，计算检验统计量

要根据假设检验的目的、研究设计的类型和样本量的大小等适用条件，选用不同的检验方法和计算相应的统计量。

3. 确定 P 值，作出推断结论

将计算得到的检验统计量与临界值比较，并同时确定 P 值。然后将概率 P 与检验水准 α 比较，从而得出结论。当 $P \leq \alpha$ 时，按所取检验水准 α ，拒绝 H_0 ，接受 H_1 ，可以认为差别有统计学意义，两总体均数不相等；当 $P > \alpha$ 时，按所取的检验水准 α ，不拒绝 H_0 ，差别无统计学意义，尚不能认为两总体均数不相等。

二、均数比较的t检验

(一) 单样本的t检验

(二) 两个独立样本的t检验

(三) 配对设计的t检验

(一) 单样本的t检验

例： 已知一般健康成年女性血红蛋白的均数为124.7g/L，某医生在某山区随机抽取了20例健康成年女性，测得她们血红蛋白如下，问：该山区健康成年女性血红蛋白均数是否与一般健康成年女性不同？

该山区20例健康成年女性血红蛋白(g/L)为：110 100 120
115 130 120 140 110 125 100 115 125 95 115 110 125
110 115 115 105

1. 基本原理

此方法用于一个样本均数 \bar{X} 与已知总体均数 μ_0 的比较。

检验目的是推断 \bar{X} 对应的未知总体均数 μ 是否不同于 μ_0 ,

其检验假设为:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

统计量 t 为

$$t = \frac{\bar{X} - \mu_0}{s / \sqrt{n}} \quad \nu = n - 1$$

当 $P \leq 0.05$ 时，拒绝 H_0 ，认为两总体均数不相等；当 $P > 0.05$ 时，不拒绝 H_0 ，尚不能认为两总体均数不相等。

2. 适用条件

正态性：样本对应的总体服从正态分布。

(二) 两个独立样本的t检验

例：某研究欲考查甘油三酯与血压的关系，在某健康管理中心随机抽取了15例甘油三酯增高的受检者与15例甘油三酯正常的受检者，如下表。试问甘油三酯增高组与甘油三酯正常组的平均收缩压有无不同？

甘油三酯增高组与甘油三酯正常组的收缩压 (mmHg)

甘油三酯增高组 X_1	甘油三酯正常组 X_2
120	120
135	90
125	120
145	110
140	125
160	110
100	125
120	125
145	95
115	140
130	145
100	120
105	105
110	90
140	115

1. 基本原理

令两个独立的样本均数 \bar{X}_1 、 \bar{X}_2 对应的总体均数为 μ_1 、 μ_2 ， μ_1 与 μ_2 是未知的。检验目的是推断 μ_1 与 μ_2 是否不同，其检验假设为：

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

两个独立样本均数的比较

	目的	H_0	H_1
双侧检验	是否 $\mu_1 \neq \mu_2$	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$
单侧检验	是否 $\mu_1 > \mu_2$	$\mu_1 = \mu_2$	$\mu_1 > \mu_2$
	是否 $\mu_1 < \mu_2$	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$

统计量 t 为

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \quad \nu = n_1 + n_2 - 2$$

公式中 $S_{\bar{X}_1 - \bar{X}_2}$ 为两样本均数差值的标准误。

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{s_c^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

公式中 s_c^2 称为合并方差，其计算方法如下：

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

公式中 s_1^2 、 s_2^2 为两样本方差， n_1 、 n_2 为两样本量。

当 $P \leq 0.05$ 时，拒绝 H_0 ，认为两总体均数不相等；当 $P > 0.05$ 时，不拒绝 H_0 ，尚不能认为两总体均数不相等。

2. 适用条件

- 正态性。两个样本对应的总体分别服从正态分布。
- 独立性。两个样本是相互独立的，无任何关联性。
- 方差齐性。两个样本对应的总体方差相等，即 $\sigma_1^2 = \sigma_2^2$ 。

注：

方差齐性是 t 检验的重要假定。若违背了方差齐性假定，则常用统计方法如下：（1）校正 t (t') 检验；（2）先变量变换，满足方差齐性假定后再进行 t 检验；（3）两个独立样本的秩和检验。

方差齐性检验

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

$$F = \frac{S_1^2}{S_2^2}, \quad \nu_1 = n_1 - 1, \quad \nu_2 = n_2 - 1$$

S_1^2 、 S_2^2 是两个样本方差

(教材中 S_1^2 表示数值较大的那个方差，而R语言与之不同)

当两样本所属总体方差不等时，Satterthwaite近似法

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$
$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

(三) 配对设计的t检验

1. 基本原理

此方法常用于三种情况：（1）按照某些特征，先将两个受试对象配成对子，再对同一对子中的两个个体分别给予处理，以比较两种处理是否有差异。（2）把同一受试对象分成两部分（如同一份标本分成两部分），再分别给予两种不同处理，以比较两种处理是否有差异。（3）自身对照设计。对同一受试对象处理前后的比较。

例：某研究人员将小学一年级的新生按学生是否经过幼儿教育分为幼儿园儿童组和散居儿童组。并将两组儿童配对，配对时使智商（IQ）、家庭经济水平及一般健康状况等尽可能相近。经过一段时间学习后，进行考试，成绩见下表。试问：幼儿园儿童考试成绩与散居儿童有无不同？

幼儿园及散居儿童考试成绩

配对号 (1)	幼儿园儿童组 (2)	散居儿童组 (3)	差值 d (4) = (2) - (3)
1	83	78	5
2	74	74	0
3	67	63	4
4	64	66	-2
5	70	68	2
6	67	63	4
7	81	77	4
8	64	65	-1
合计	—	—	16

以配对设计为例，令每一对子的两个观测变量为 X 、 Y ，差值 $d = X - Y$ ， d 的样本均数为 \bar{d} ， \bar{d} 对应的总体均数为 μ_d ， μ_d 是未知的。检验目的是推断 μ_d 是否不等于0，其检验假设为：

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

统计量 t 为

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} \quad \nu = n - 1$$

当 $P \leq 0.05$ 时，拒绝 H_0 ，认为两总体均数不相等；当 $P > 0.05$ 时，不拒绝 H_0 ，尚不能认为两总体均数不相等。

2. 适用条件

正态性：差值 d 对应的总体服从正态分布。

注：

上述情况得到的两个样本是相关、不独立的，这一点与两个独立样本的 t 检验不同。

例：尿铅测定长期以来用湿式热消化法—双硫脲法，后改用硝酸—高锰酸钾冷消化法，试就下表资料，说明两种方法测得结果有无不同？

16例患者的尿样分别用两种方法测定尿铅 ($\mu\text{mol/L}$)

患者号 (1)	冷消化法 (2)	热消化法 (3)	差值 d (4) = (2) - (3)
1	2.41	2.80	-0.39
2	12.07	11.24	0.83
3	2.90	3.04	-0.14
4	1.64	1.83	-0.19
5	2.75	1.88	0.87
6	1.06	1.45	-0.39
7	3.23	3.43	-0.20
8	0.77	0.92	-0.15
9	3.67	3.81	-0.14
10	0.37	0.37	0.00
11	4.49	4.01	0.48
12	0.19	0.29	-0.10
13	5.16	4.44	0.72
14	0.16	0.27	-0.11
15	5.45	5.41	0.04
16	0.04	0.07	-0.03
合计	—	—	1.10

例：某研究者欲研究控制饮食是否对高血脂病人有疗效，对18名高血脂病人控制饮食一年，观察了这18名病人控制饮食前后血清胆固醇的变化。问控制饮食前后血清胆固醇有无变化？

编号	控制饮食前 胆固醇(mg%)	控制饮食后 胆固醇(mg%)	差值 d
(1)	(2)	(3)	(4) = (2) - (3)
5	195	225	-30
6	150	150	0
7	210	173	37
8	300	268	32
9	184	193	-9
10	173	186	-13
11	212	210	2
12	207	200	7
13	250	200	50
14	150	157	-7
15	205	193	12
16	164	239	-75
17	286	286	0
18	203	167	36
合计			60

注：

配对设计 t 检验的统计效率高于两个独立样本 t 检验。

例：测量10例育龄妇女在服用OC（口服避孕药）前、后的收缩压，计算得服用OC前、后收缩压差值的均数为4.80mmHg，标准差为4.566mmHg。问：服用OC前、后的收缩压是否有变化？

例：测量8例服用OC的育龄妇女的收缩压，均数为132.86mmHg，标准差为15.34mmHg；21例未服用OC的育龄妇女，收缩压均数为127.44mmHg，标准差为18.23mmHg。问：服用OC及不服用OC的两组收缩压是否有差异？