# 《医学统计学》第五次课

## 参数估计、假设检验的基本概念

卫生统计教研室

学习重点：

1. 抽样误差、标准误的概念和计算；
2. 总体均数估计的方法，特别是可信区间的概念和计算。
3. 总体率估计的方法。

统计推断( statistical inference ):

在大多数情况下，研究者并不知道总体的参数，而是在总体中随机抽取一定数量观察单位作为样本进行抽样研究，通过样本指标来说明总体特征，这种从样本获取有关总体信息的过程称为统计推断。

统计推断的两个基本问题:

1.参数估计（parameter estimation）:对总体参数进行估计;

2. 假设检验（test of hypothesis）:用推理的方法来判断某个（或某几个）随机样本是否来源于预先假设的总体。

# 第一节　样本均数的标准误

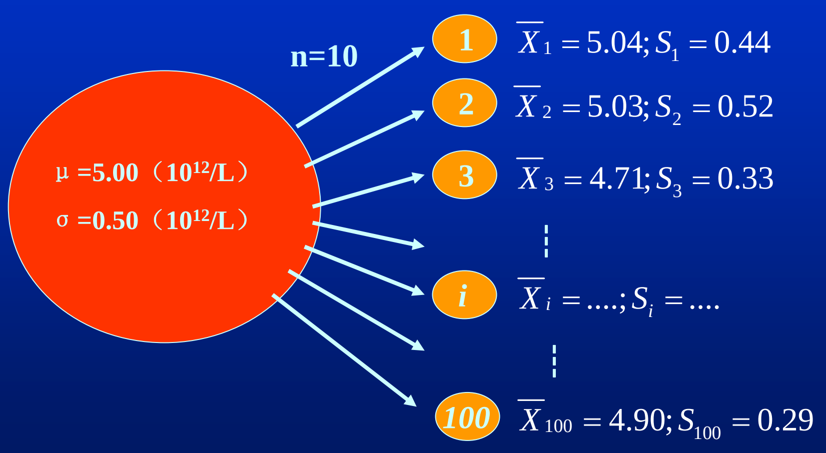
## 一、均数的抽样误差与标准误

1. 抽样误差(sampling　error)：p32.

   由于样本的随机性引起的统计量与参数的差别，或同一总体相同统计量之间的差别。

2. 抽样误差的大小用标准误度量。

## 一、均数的抽样误差与标准误



## 一、均数的抽样误差与标准误

数理统计推理：从正态总体$N$(μ，σ$^2$)中随机抽取样本量为$n$ 的样本，样本均数 $\overline{X} \sim N(\mu,\sigma_{\overline{X}}^2)$，其中 $\sigma_{\overline{X}}$ 为样本均数 $\overline{X}_i$的总体标准差，称为样本均数的标准误。

$$\sigma_{\overline{X}} = \sigma / \sqrt{n} \qquad (7\text{-}1)$$

## 一、均数的抽样误差与标准误

σ通常未知，用样本标准差$S$来估计:

$$S_{\overline{X}} = S / \sqrt{n} \qquad (7\text{-}2)$$

统计上，通常将统计量（如样本均数$\overline{X}$ 、样本率p等）的标准差称为标准误（standard error, SE）.

## 二、样本均数的分布

数理统计的中心极限定理可知:

无论$X$服从何种分布，只要它具有总体均数μ和方差$\sigma^2$，当$n$足够大时(n≥60)，$\overline{X}$的分布近似正态分布$N(\mu, \sigma^2/n)$。

## 三、总体均数的估计

（一）点估计（point estimation）:

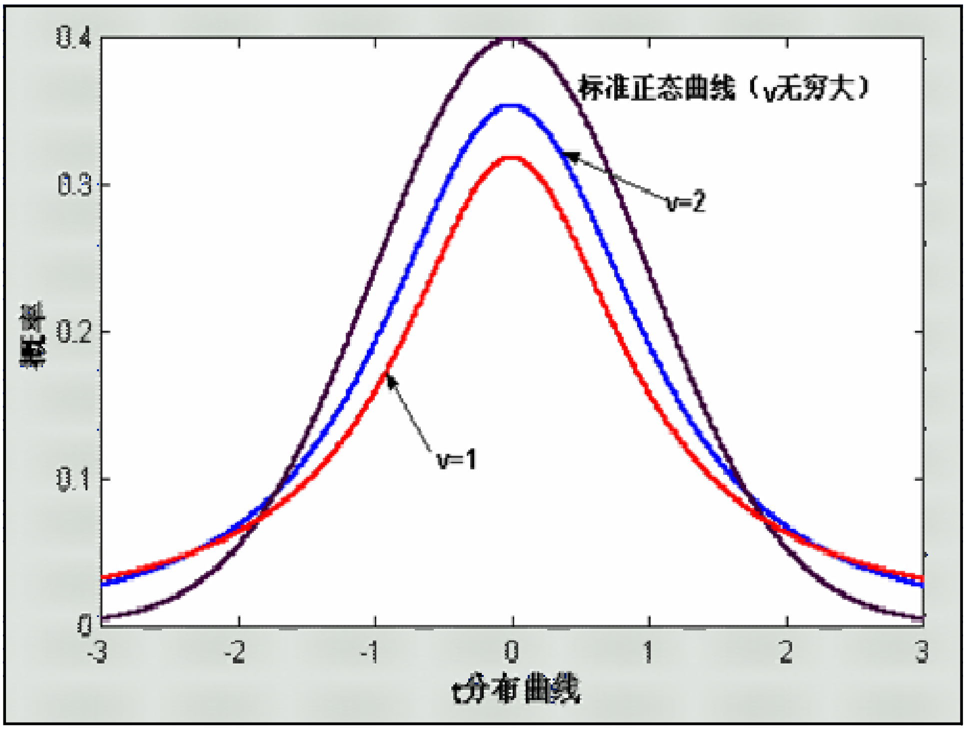
总体均数的点估计就是用样本均数来直接地估计总体均数，即 $\hat{\mu}=\overline{X}$。

（二）区间估计（interval estimation）:

利用样本信息给出一个区间，并同时给出重复试验时该区间包含总体均数的概率。

**t分布：p96.**

1. t分布是一种小样本分布;
2. t分布是一簇对称的、均数为0的单峰分布曲线，与标准正态曲线相比，顶部稍低而左右两段稍高。自由度$\nu$不同，曲线的形状不同。 $\nu$越大，曲线越接近于标准正态分布曲线。



## t分布

1. t分布是一种小样本分布；
2. t分布是一簇对称的、均数为0的单峰分布曲线，与标准正态曲线相比，顶部稍低而左右两段稍高。自由度$\nu$不同，曲线的形状不同。 $\nu$越大，曲线越接近于正态分布曲线。
3. t界值表：自由度ν取不同值时，t分布单侧或双侧尾部面积α的t分布界值，分别表示为$t_{\alpha,\nu}$和$t_{\alpha/2,\nu}$。附表2

## t分布

4. t统计量：

$$t=\frac{\overline{X}-\mu}{S/\sqrt{n}}\quad,\ \nu=n-1$$

t分布是样本均数的一种分布形式。

## 三、总体均数的估计

（二）区间估计（interval estimation）：

1. σ未知时 按t分布原理，有

$$P(-t_{\alpha/2,\upsilon} < \frac{\overline{X}-\mu}{S_{\overline{X}}} < t_{\alpha/2,\upsilon}) = 1-\alpha \qquad (7\text{-}3)$$

1-α为置信水平或可信度（置信概率），$t_{\alpha/2,\nu}$为t分布$\nu$=n-1，双侧尾部面积为α的t界值。

## 三、总体均数的估计

$$\overline{X} - t_{\alpha/2,\upsilon}S_{\overline{X}} < \mu < \overline{X} + t_{\alpha/2,\upsilon}S_{\overline{X}} \qquad (7\text{-}4)$$

当α=0.05，可信度为100（1-α）%=95%，公式（7-4）为总体均数μ的95%的置信区间或称可信区间（confidence interval,CI）

## 三、总体均数的估计

（二）区间估计：

2. σ已知时，或σ未知但$n$足够大（n>60），按正态分布原理，有

$$P(-u_{\alpha/2} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < u_{\alpha/2}) = 1-\alpha \qquad (7\text{-}5)$$

$$(\overline{X} - u_{\alpha/2}\sigma/\sqrt{n} < \mu < \overline{X} + u_{\alpha/2}\sigma/\sqrt{n}) \qquad (7\text{-}6)$$

## 三、总体均数的估计

（三）可信区间的涵义

1. 与参考值范围的差别

| | 总体均数的可信区间 | 参考值范围 |
|---|---|---|
| 含义 | 由样本统计量推断总体参数时所进行的区间估计。表示该区间在重复试验中，有多大的可能包含总体均数。 | “正常”人 某项指标的波动范围，反映的是总体的特征。表示绝大多数（95%）观察对象某项指标的范围。 |

（三）可信区间的涵义

1. 与参考值范围的差别

| | 总体均数的可信区间 | 参考值范围 |
|---|---|---|
| 计算 | σ 未知：<br>$\overline{X} \pm t_{\alpha/2,\upsilon} S_{\overline{X}}$<br>σ 已知， σ 未知<br>但n足够大（n>60）：<br>$\overline{X} \pm u_{\alpha/2} \sigma / \sqrt{n}$ | 正态分布：<br>$\overline{X} \pm u_{\alpha/2} S$<br>偏态分布：<br>$P_X \sim P_{100-X}$ |

（三）可信区间的涵义

2. 两个要素：

准确度(accuracy) 和精密度（precision）

| | |
|---|---|
| 可信区间包含μ的概率1—α的大小。 | 区间的长度 |
| 概率越大越好 | 区间的长度越小越好 |
| 1. 一定样本量下，二者是矛盾的；<br>2. 可信度固定的前提下，提高精密度的唯一办法是扩大样本量。 | |

# 第二节　率的标准误

## 一、率的抽样误差与标准误

1. 计数资料，随机变量X服从二项分布，即 $X\sim b(n,\pi)$，n表示观察样本的个数，　π表示阳性结果的发生率。它们是二项分布的两个参数。
2. 率抽样误差：
3. 率的标准误。

## 一、率的抽样误差与标准误

率的标准误：p78.

$$\sigma_p=\sqrt{\frac{\pi(1-\pi)}{n}} \qquad (5\text{-}27)$$

总体率π未知时，以样本率$p$作为π的估计值，则$\sigma_p$的估计值：

$$\hat{\sigma}_p=S_p=\sqrt{p(1-p)/n} \qquad (5\text{-}28)$$

## 二、样本率的分布

数理统计的证明可知：

设$X \sim b(n, \pi)$，从这个总体中抽样所得的样本量为n的样本，各样本率的抽样分布随着n的增大而趋近正态分布，即

$$p \sim N(\pi, \pi(1-\pi)/n)$$

## 三、总体均数的估计

（一）点估计：

$$\hat{\pi} = p$$

（二）区间估计：

当n足够大，且np与n(1－p)均大于5时，p的抽样分布近似正态分布，置信水平为100(1－α)%的总体率可信区间：

$$(p - u_{\alpha/2}\hat{\sigma}_p, p + u_{\alpha/2}\hat{\sigma}_p)$$

$$\hat{\sigma}_p = S_p = \sqrt{p(1-p)/n} \qquad (7\text{-}7)$$

# 第三节　两均数之差的可信区间

## 一、两样本均数之差的分布及标准误

$$\overline{X}_1-\overline{X}_2 \sim N(\mu_1-\mu_2,\sigma_1^2/n_1+\sigma_2^2/n_2) \qquad (7\text{-}8)$$

$$S_{\overline{X}_1-\overline{X}_2}=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \qquad (7\text{-}9)$$

$$\hat{\sigma}_{\overline{X}_1-\overline{X}_2}=\sqrt{\frac{\hat{\sigma}_1^2}{n_1}+\frac{\hat{\sigma}_2^2}{n_2}} \qquad (7\text{-}10)$$

## 二、两总体均数之差的估计

（一）点估计:

$$\hat{\mu}=\hat{\mu}_1-\hat{\mu}_2=\overline{X}_1-\overline{X}_2 \qquad (7\text{-}11)$$

（二）区间估计:

$$(\overline{X}_1-\overline{X}_2)-t_{\alpha/2,\upsilon}S_{\overline{X}_1-\overline{X}_2}<\mu_1-\mu_2<(\overline{X}_1-\overline{X}_2)-t_{\alpha/2,\upsilon}S_{\overline{X}_1-\overline{X}_2}$$

(7-12)

## 第四节　两个率之差的可信区间

### 一、两率之差的分布及标准误

$$p_1 - p_2 \sim N(\pi_1 - \pi_2, \pi_1(1-\pi_1)/n_1 + \pi_2(1-\pi_2)/n_2)$$

$$\hat{\sigma}_{p_1-p_2} = S_{p_1-p_2} = \sqrt{p_c(1-p_c)(\frac{1}{n_1}+\frac{1}{n_2})} \quad (7\text{-}14)$$

合并样本率

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \quad (7\text{-}15)$$

### 一、两率之差的分布及标准误

当$n_1$与$n_2$比较大时，

$$\hat{\sigma}_{p_1-p_2} = S_{p_1-p_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2})} \quad (7\text{-}16)$$

### 二、两总体率之差的估计

# 第八章　假设检验

通过以往大规模调查，已知某地一般新生儿的头围均数为34.50cm，今测得该地某矿区55名新生儿的头围均数为33.89cm，问是否该矿区新生儿的头围总体均数与一般新生儿不同？

分析：

$$\left.\begin{array}{l}\mu_0 = 34.50cm \\ \bar{X} = 33.89cm\end{array}\right\} \text{0.61cm} \Rightarrow$$

差异可能的原因

1. 抽样误差造成
2. 总体均数不同（来自不同总体）

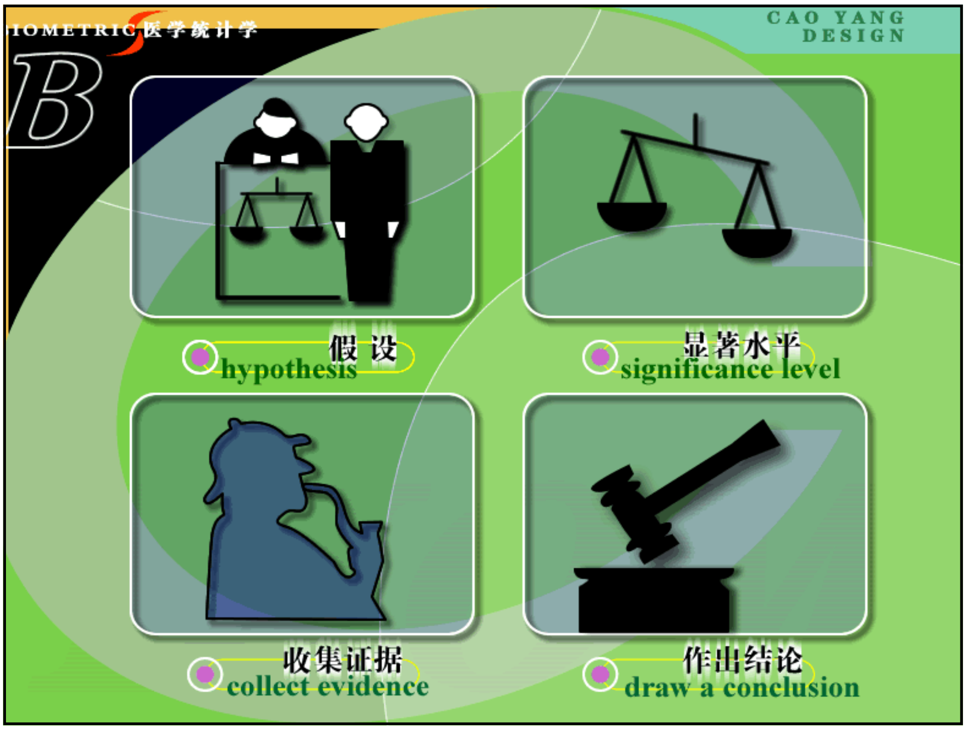
⇓

统计上用相应的统计学检验对原因进行统计推断，作出决策。

假设检验(hypothesis test)

假设检验是依据样本提供的有限信息对总体作推断的统计学方法，是在对研究总体的两种对立的判断之间做选择的决策程序。也称为显著性检验(significance test)。

假设检验的基本逻辑：小概率事件在一次随机抽样试验中不太可能出现。



## 第一节　检验假设与P值

例8-1　通过以往大规模调查，已知某地一般新生儿的头围均数为34.50cm，标准差为1.99cm。为研究某矿区新生儿的发育状况，现从该地某矿区随机抽取新生儿55人，测得其头围均数为33.89cm，问是否该矿区新生儿头围总体均数与一般新生儿不同？

分析差异来源 → 假设检验 → 检验假设

## 第一节　检验假设与P值

根据实验设计和反证法的思想，从所要解决问题的对立面对研究总体的特征建立一个检验假设

研究目的：研究该矿区新生儿的头围总体均数与一般新生儿头围总体均数是否不同

⇓

该矿区新生儿头围总体均数与一般新生儿头围总体均数相同 → 检验假设或无效假设 $H_0$

该矿区新生儿头围总体均数与一般新生儿头围总体均数不同 → 备择假设 $H_1$

## 第一节　检验假设与P值

$H_0$是否成立？

⇩

在$H_0$成立的条件下，样本均数与总体均数的差异能否用抽样误差来解释？

⇩

选择相应的假设检验方法

⇩

计算相应的检验统计量

⇩

确定P值

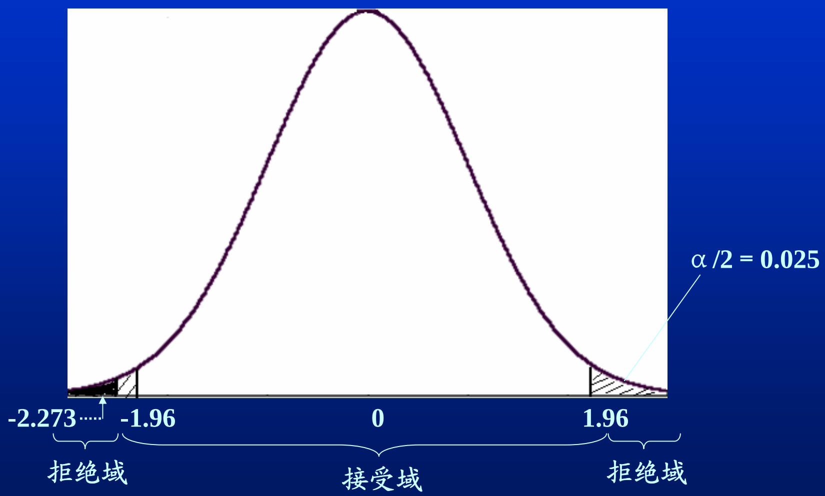
## 第一节　检验假设与P值

$\overline{X}$　　　$u$变换　　　$u$

$$N\left(\mu_0, \sigma_0^2/n\right) \xrightarrow{u=\left(\overline{X}-\mu_0\right)/\left(\sigma_0/\sqrt{n}\right)} N(0,1)$$

$$u=\frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}=\frac{33.89-34.50}{1.99/\sqrt{55}}=-2.273$$

## 第一节　检验假设与P值




标准正态分布下$H_0$的接受域与拒绝域示意图

## 第一节　检验假设与P值

**检验水准α**：是预先认为确定的，表示拒绝了实际上是成立的$H_0$的概率大小，或者在拒绝$H_0$做出有差别结论时可能犯错误的最大概率，也称为显著性水准。

α的大小可根据研究目的确定，一般常取α = 0.05或0.01

检验统计量 ⟹ 确定P值

**P值含义**：是指在$H_0$规定的总体中进行随机抽样，得到大于等于（或小于等于）依据现有样本信息计算得到的检验统计量的概率

## 第二节　假设检验的基本步骤

### 一、建立假设，确定检验水准

$H_0$: μ = 34.50(该矿区新生儿的头围总体均数与当地一般新生儿相同)

$H_1$: μ ≠ 34.50(该矿区新生儿的头围总体均数与当地一般新生儿不同)

α = 0.05

### 二、选择适当的假设检验方法，计算相应的检验统计量

$$u=\frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}=\frac{33.89-34.50}{1.99/\sqrt{55}}=-2.273$$

### 三、确定P值，下结论

根据计算得到的检验统计量，查相应的界值表获得相应的P值

若P ≤ α，拒绝$H_0$，接受$H_1$，下有差别的结论

若P > α，不拒绝$H_0$，根据目前资料尚不能认为有差别

已知$|u|=2.273$ 由标准正态分布的u界值得$u_{0.05/2}=1.96$，因2.273 > $u_{0.05/2}$，故P < 0.05

## 第二节　假设检验的基本步骤

### 假设检验的结论：统计结论、专业结论

统计结论：若P ≤ α，拒绝$H_0$，接受$H_1$，差别有统计学意义
若P > α，不拒绝$H_0$，根据目前资料尚不能认为差别有统计学意义

专业结论：根据统计结论对实际问题中的总体均数是否不同以及差异的方向做出专业推断

按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义（统计结论），故可以认为矿区新生儿的头围总体均数与一般新生儿不同，矿区新生儿的头围小于一般新生儿（专业结论）


## 第三节　大样本均数的假设检验

一、单样本均数的*u*检验

二、两均数比较的*u*检验

均数比较*u*检验的适用条件:

1.样本量：单样本数据要求每组例数≥60；两样本数据例数之和≥60，且例数基本均等

2.要求两总体方差已知

3.理论要求：为随机样本，组间有可比性

## 一、单样本均数的u检验

研究目的：推断样本所代表的未知总体均数$\mu$与已知总体均数$\mu_0$有无差别

已知总体均数$\mu_0$一般为理论值、标准值或经大量观察所得的稳定值

检验统计量计算公式:

$$u = \frac{\overline{X} - \mu_0}{\sigma_0 / \sqrt{n}}$$

## 一、单样本均数的u检验

例8-2 1995年已知某地20岁应征男青年平均身高为168.5cm。2003年在当地20岁应征男青年中随机抽取85人，平均身高为171.2cm，标准差为5.3cm，问2003年当地20岁应征男青年身高与1995年相比是否不同？

1.建立假设，确定检验水准α

$H_0$: μ = 168.5（2003年当地20岁应征男青年的身高与1995相比无变化）

$H_1$: μ ≠ 168.5（2003年当地20岁应征男青年的身高与1995相比有变化）

α = 0.05

2. 计算检验统计量$u$

$$u = \frac{\overline{X} - \mu_0}{\hat{\sigma} / \sqrt{n}} = \frac{171.2 - 168.5}{5.3 / \sqrt{85}} = 4.70$$

3.确定P值，下结论

检验界值$u_{0.05/2} = 1.96$，因$4.70 > u_{0.05/2}$，故$P < 0.05$。按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义，即2003年当地20岁应征男青年的身高与1995相比有变化，比1995年增高了。

## 二、两均数比较的u检验

完全随机设计：将受试对象完全随机地分配到两个组中，两组分别接受不同的处理。

目的：推断两总体均数是否相同。

两样本均数比较的u检验的基本原理：

$$\bar{X}_1-\bar{X}_2 \quad\Downarrow\quad N\left(\mu_1-\mu_2,\hat{\sigma}^2_{\bar{X}_1-\bar{X}_2}\right) \xrightarrow{u=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\hat{\sigma}_{\bar{X}_1-\bar{X}_2}}} N(0,1) \quad u\Downarrow$$

检验统计量计算公式：

$$u=\frac{\bar{X}_1-\bar{X}_2}{\hat{\sigma}_{\bar{X}_1-\bar{X}_2}}=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1}+\frac{\hat{\sigma}_2^2}{n_2}}}$$

## 二、两均数比较的u检验

例8-3 为比较某药治疗流行性出血热的疗效，将72名流行性脑炎患者随机分为两组，两组样本量、均数、标准差分别为$n_1$=32，$\bar{X}_1$=29 ，$S_1$=1.9；$n_2$=40，$\bar{X}_2$=52，$S_2$=2.7。问试验组和对照组的平均退热天数有无差别？

1.建立假设，确定检验水准α

$H_0$: $\mu_1=\mu_2$； $H_1$: $\mu_1\neq\mu_2$。 α = 0.05

2. 计算检验统计量u

$$u=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\hat{\sigma}_1^2/n_1+\hat{\sigma}_2^2/n_2}}=\frac{2.9-5.2}{\sqrt{1.9^2/32+2.7^2/40}}=-4.23$$

3.确定P值，下结论

检验界值$u_{0.05/2}$ = 1.96，因4.23 > $u_{0.05/2}$，故P < 0.05。按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义，故可以认为两组的退热天数的总体均数不相等，疗效不同，即试验组的平均退热天数比对照组短。

## 第四节　大样本率的假设检验

一、单样本率的$u$检验
二、两个率比较的$u$检验

率的$u$检验的适用条件:

1.$n$较大，如每组例数大于60例
2.样本$p$或1-$p$均不接近于100％和0
3.$np$和$n$(1-$p$)均大于5

## 一、单样本率的$u$检验

研究目的：推断样本所代表的未知总体率π与已知总体率$\pi_0$有无差别

已知总体率$\pi_0$一般为理论值、标准值或经大量观察所得的稳定值。

单样本率的$u$检验基本原理：在$\pi=\pi_0$，且为大样本的条件下，样本率p近似服从均数为$\pi_0$，方差为$\sigma_p^2$的正态分布。

检验统计量计算公式:

$$u=\frac{p-\pi_0}{\sigma_p}=\frac{p-\pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

## 一、单样本率的*u*检验

例8-4 已知某地40岁以上成年男性高血压患病率为8.5％，经健康教育数年后，随机抽取该地成年男性1000名，查出高血压患者55例，患病率为5.5％。问健康教育前后该地成年男性高血压患病率是否有所变化？

1.建立假设，确定检验水准α

$H_0$: $\pi$ = 8.5%；$H_1$: $\pi \neq$ 8.5 ％。 α = 0.05

2.计算检验统计量*u*

$$u=\frac{p-\pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}=\frac{0.055-0.085}{\sqrt{0.085(1-0.085)/1000}}=-3.402$$

3.确定P值，下结论

检验界值$u_{0.05/2}$ = 1.96，因3.402 > $u_{0.05/2}$，故P < 0.05。按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义，故可以认为健康教育后该地成年男性高血压患病率有所降低。

## 二、两个率比较的*u*检验

目的：推断两个总体率是否不同。

两个率比较的*u*检验基本原理：在 $\pi_1 = \pi_2$成立且为大样本的条件下，两样本率的差值（$p_1$-$p_2$）近似服从均数为 $\pi_1 - \pi_2$=0, 方差为$\sigma^2_{p_1-p_2}$的正态分布。

检验统计量计算公式：

$$u=\frac{p_1-p_2}{\hat{\sigma}_{p_1-p_2}}=\frac{p_1-p_2}{\sqrt{p_c(1-p_c)(1/n_1+1/n_2)}}$$

$$p_c=\frac{X_1+X_2}{n_1+n_2}$$

## 二、两个率比较的$u$检验

例8-5 某医院用黄芪注射液和胎盘球蛋白进行穴位注射治疗小儿支气管哮喘病人，黄芪注射液治疗117例，有效103例；胎盘球蛋白治疗55例，有效49例。试比较两种疗法有效率有无差别？

1.建立假设，确定检验水准α

$H_0$: $\pi_1 = \pi_2$；$H_1$: $\pi_1 \neq \pi_2$。 $\alpha = 0.05$

2. 计算检验统计量$u$

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{103 + 49}{117 + 55} = 0.8837$$

$$u = \frac{p_1 - p_2}{\sqrt{p_c(1-p_c)(1/n_1 + 1/n_2)}} = \frac{0.8803 - 0.8909}{\sqrt{0.8837(1-0.8837)(1/117 + 1/55)}} = -0.0054$$

3.确定P值，下结论

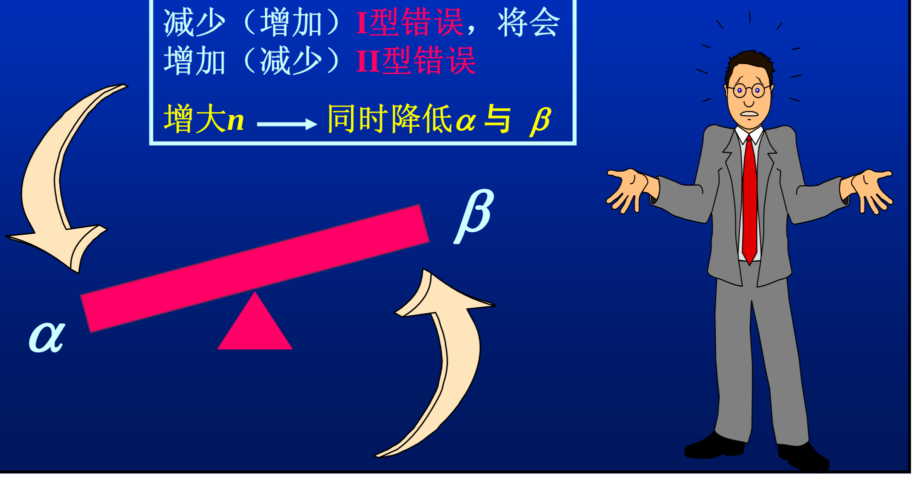
检验界值$u_{0.05/2} = 1.96$，因$0.0054 < u_{0.05/2}$，故$P > 0.05$。按$\alpha = 0.05$水准，不拒绝$H_0$，差别无统计学意义，故尚不能认为两疗法治疗小儿支气管哮喘疗效有差别。

## 第五节 检验水准与两类错误

### $\alpha$ 与 $\beta$间的关系

减少（增加）I型错误，将会增加（减少）II型错误

增大$n$ ⟶ 同时降低$\alpha$与 $\beta$



## 第五节 检验水准与两类错误

### 假设检验的两类错误

| 真实情况 | 假设检验结论：拒绝$H_0$，接受$H_1$ | 假设检验结论：不拒绝$H_0$ |
|---|---|---|
| $H_0$成立 | Ⅰ型错误α | 推断正确(1-α) |
| $H_0$不成立 | 推断正确(1-β) | Ⅱ型错误β |

Ⅰ型错误(type Ⅰ error): 拒绝了实际上是成立的$H_0$。 “弃真”

Ⅱ型错误(type Ⅱ error)：不拒绝实际上是不成立的$H_0$。“存伪”

1-β:称为检验效能或检验功效，又称为把握度(power of a test)，即两总体确有差别时，按检验水准α，能检出有差别的能力

1-α:称为可信度(confidence level)，即重复抽样时，样本区间包含总体参数(μ)的概率

CAO YANG DESIGN

BIOMETRICS 医学统计学

## 两类错误

### I型错误

真实情况是$H_0$，经检验被拒绝了。犯这类错误的概率是α。

### II型错误

真实情况不是$H_0$，经检验却被接受了。犯这类错误的概率是β。

## 第六节　单侧检验与双侧检验

例8-2 1995年已知某地20岁应征男青年平均身高为168.5cm。2003年在当地20岁应征男青年中随机抽取85人，平均身高为171.2cm，标准差为5.3cm，问2003年当地20岁应征男青年身高与1995年相比是否增高了？

1.建立假设，确定检验水准α

$H_0$: μ = 168.5（2003年当地20岁应征男青年的身高与1995相比无变化）

$H_1$: μ > 168.5（2003年当地20岁应征男青年的身高与1995相比有所增加）

单侧检验　α = 0.05

2. 计算检验统计量$u$

$$u = \frac{\overline{X} - \mu_0}{\hat{\sigma} / \sqrt{n}} = \frac{171.2 - 168.5}{5.3 / \sqrt{85}} = 4.70$$

3.确定P值，下结论

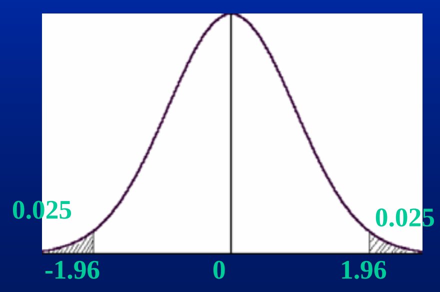
检验界值$u_{0.05}$ = 1.645，因4.70 > $u_{0.05}$，故P < 0.05。按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义，即2003年当地20岁应征男青年的身高比1995年增高了。

## 第六节　单侧检验与双侧检验

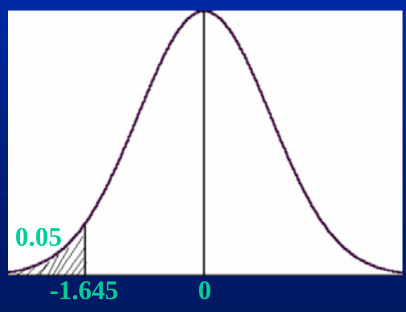
例8-3 为比较某药治疗流行性出血热的疗效，将72名流行性脑炎患者随机分为两组，两组样本量、均数、标准差分别为$n_1$=32, $\overline{X}_1$ = 2.9，$S_1$=1.9; $n_2$=40, $\overline{X}_2$=5.2，$S_2$=2.7。问试验组和对照组的平均退热天数有无差别？

双侧检验　　$H_0$: $\mu_1 = \mu_2$；　$H_1$: $\mu_1 \neq \mu_2$

单侧检验　　$H_0$: $\mu_1 = \mu_2$；　$H_1$: $\mu_1 < \mu_2$




双侧$u$检验的检验水准α




单侧$u$检验的检验水准α

## 第六节　单侧检验与双侧检验

例8-4 已知某地40岁以上成年男性高血压患病率为8.5％，经健康教育数年后，随机抽取该地成年男性1000名，查出高血压患者55例，患病率为5.5％。问健康教育后该地成年男性高血压患病率是否有所降低?

1.建立假设，确定检验水准α

$H_0$: $\pi$ = 8.5%；$H_1$: $\pi$<8.5％。单侧检验　α = 0.05

2.计算检验统计量$u$

$$u=\frac{p-\pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}=\frac{0.055-0.085}{\sqrt{0.085(1-0.085)/1000}}=-3.402$$

3.确定P值，下结论

检验界值$u_{0.05}$ = 1.65，因3.402 > $u_{0.05}$，故P < 0.05。按α = 0.05水准，拒绝$H_0$，接受$H_1$，差别有统计学意义，故可以认为健康教育后该地成年男性高血压患病率有所降低。

CAO YANG DESIGN

## 单侧检验与双侧检验

### 单侧检验

若只考虑被检验的参数大于(或小于)已知参数的情况，这样的备择假设称为单侧假设，检验称为单侧检验。

### 双侧检验

被检验的参数大于和小于已知参数的情况均考虑，这样的备择假设称为双侧假设，检验称为双侧检验。

BIOMETRIC 医学统计学

第七节 假设检验的统计意义与实际意义

一、严密的研究设计是假设检验结论正确的前提

实验设计和抽样设计的好坏，直接关系到研究结果的可靠性。

二、假设检验的统计意义

**1.**假设检验结论的正确性是以概率作为保证的

**2.P**值的含义

**3.**统计结论的表述

**4.**假设检验与可信区间的区别与联系

三、假设检验的实际意义

第七节 假设检验的统计意义与实际意义

假设检验的基本逻辑:

小概率事件在一次随机抽样试验中不太可能出现

若P≤α，拒绝$H_0$，接受$H_1$，下有差别的结论

若P＞α，不拒绝$H_0$，根据目前资料尚不能认为有差别

## 第七节　假设检验的统计意义与实际意义

P值含义：是指在$H_0$规定的总体中进行随机抽样，得到大于等于（或小于等于）依据现有样本信息计算得到的检验统计量的概率。

P值的大小不仅与总体间差别的大小有关，更与样本量和抽样误差的大小有关

下结论时需注意：
不能认为P值越小，总体参数间的差别越大。
假设检验只能做出拒绝$H_0$或不拒绝$H_0$的定性判断，但不能给出总体参数间的差别大小和方向。

## 第七节　假设检验的统计意义与实际意义

假设检验与可信区间的区别与联系：

可信区间用于推断总体均数的范围，假设检验用于推断总体均数间是否相等。
可信区间具有假设检验的主要功能，但是并不能完全代替假设检验。
可信区间只能在预先确定的检验水准下进行计算，而假设检验能够获得一个确切的概率P值。

## 第七节　假设检验的统计意义与实际意义

假设检验的结论与样本量有关，当样本量足够大时，标准误趋向于0，无论两者相差多少，都能够得到足以拒绝$H_0$的统计量和$p$值。

对假设检验结果的实际意义或临床意义的判定，一定要结合专业知识。当专业上和统计学上都有意义时，试验结果才具有实用价值。

## SAS应用

### 单一总体均数的可信区间

```
/*例5-1*/
data prg5_1;
n=10; mean=166.95; std=3.64;
t=tinv(0.975,n-1);
in=t*std/sqrt(n);
lclm=mean-in;
uclm=mean+in;
run;
proc print;
var lclm uclm;run;
```

## 两总体均数相差的可信区间

程序5-2

```
data prg5_2;
n1=29; n2=32; m1=20.10; m2=16.89;
s1=7.02;s2=8.46;
sc2=(s1**2*(n1-1)+s2**2*(n2-1))/(n1+n2-2);
st=sqrt(sc2*(1/n1+1/n2));
t=tinv(0.975,n1+n2-2);
in=t*st;
lclm=abs(m1-m2)-in;
uclm=abs(m1-m2)+in;
proc print;
var lclm uclm;
run;
```

## 小　结

一、假设检验是依据样本提供的有限信息对总体作推断的统计学方法，是对研究总体的两种对立的判断之间做选择的决策程序。

二、假设检验的基本步骤：建立检验假设，确定检验水准→计算相应的检验统计量→确定p值→下结论。

三、假设检验的基本逻辑是：小概率事件在一次抽样中不太可能出现。

四、假设检验有两类错误。

## 小 结

五、假设检验与相应的可信区间估计既能提供等价的结果，又有各自不同的功能。

六、假设检验方法很多，每种方法均有相应的应用条件。应综合考虑研究目的、设计类型、变量类型及样本含量等要素后才能选择合适的假设检验方法。

七、参数估计不同于统计描述，它完成的是由样本统计量推断总体参数的工作。在抽样过程中存在抽样误差，我们用标准误来度量。对总体参数进行估计有点估计和区间估计两个估计方法。区间估计根据标准误的大小和样本的分布规律计算得出。