心理统计
第九讲：估计（Estimation）

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

复习

2

描述统计

3

我们如何描述一个变量?

- 度量的测度
- 集中趋势
- 变异性
- 分布的形状

4

描述样本

统计方法的选择依赖于数据的测度

	等距、等比	顺序或偏态	命名
集中趋势	均值	中数	众数
变异性	标准差	四分位距； 全距	
图	直方图 次数多边图.	棒图	棒图

5

为什么我们需要描述分布?

1. 描述一个组
2. 描述一个个体在组中的位置
   - z 分数
   - $Z = \frac{X-\mu}{\sigma}$
   - 百分位等级
3. 检验推论统计中的前提
   - 方差同质性

6

推论统计

心理学中探究真相的工具

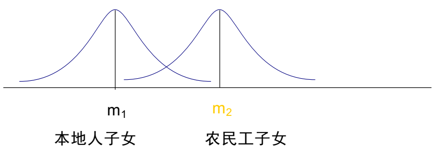
7

假设检验

- 我们想知道两组数据是否有差别.
- 真相在于总体参数,我们不能得到.
- 我们只能从样本的情况猜测总体参数.

8

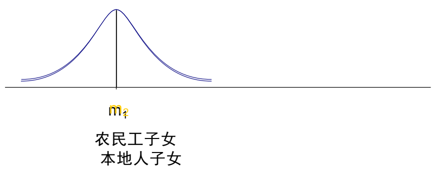
假设检验



农民工子女在克服困难的坚持性上优于本地人子女.

9

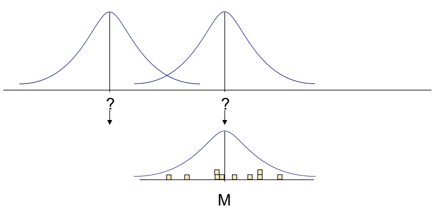
或者



农民工子女在克服困难的坚持性上与本地人子女没有差别.

10

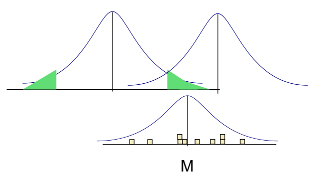
我们收集一个样本



在这个例子中,样本均值是总体均值的一个很好的估计我们会作出正确的结论

11

样本均值分布:虚无分布



如果一个样本均值落在其中一个尾端:

- 它可能是虚无总体的一个非常不常见的样本,
- 或者是从另一个总体中来.

我们作出结论是后者.为什么?

. . .因为我们只有小的概率会犯错误 (.05 or .01)

12

统计结论的步骤

1. 陈述虚无假设和研究假设.
2. 设定标准 找到临界z.
   alpha=.05, 方向性
3. 收集数据并计算 z的观测值.
4. 作出结论.
   如果 z的观测值落在临界区域中,拒绝H_0

13

Step 1 无方向性

- H_0= 无效应 (农民工子女与本地人子女无差别)
  H_1 =有效应 (农民工子女与本地人子女有差别)

14

Step 1 方向性

- H_1=农民工子女在坚持性上高于本地人子女
  H_0=农民工子女在坚持性上不高于本地人子女

15

Step 2 设置标准

- Alpha (α) I 类错误的概率
  (当H_0 正确时拒绝 H_0)
- 确定将 α设置在.05 还是 .01
- 记住正态分布表中.05, .025, .01, or .005对应的z值

16

Step 3 收集数据并计算z的观测值

$$z_{obt} = \frac{M-\mu}{\sigma_M} \quad where \quad \sigma_M = \frac{\sigma}{\sqrt{n}}$$

M = 5.0
μ = 4.0
σ = 4
n = 25

$$\frac{4}{\sqrt{25}} = .8$$

$$\frac{5-4}{.8} = 1.25$$

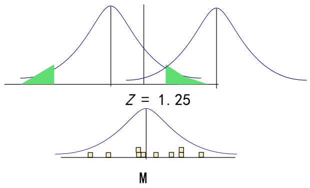
17

Step 4 作出结论

- 如果 z的观测值落在临界区域中,拒绝H_0
- (在无方向性检验中需要更极端的值才能拒绝)
- z_{obs}= 1.25, z_{cri}= 1.96
- 我们能否拒绝H0?

18

样本均值分布:虚无分布



- 我们的样本均值没有落在临界区域内.
- 因此这意味着它会在虚无分布内机遇性地发生.
- 接受 H_0.

19

Cohen's *d*

$$d = \frac{Mdifference}{SD}$$

$$d = \frac{5-4}{3.8} = .26$$

$M = 5.0$
$s = 3.8$
$\mu = 4.0$
$\sigma = 4$
$n = 25$

20

效应大小的评价

d的大小	效应大小的评价
0<*d*<0.2	**小的效应 (mean difference <.2 *SD*)**
0.2<*d*<0.8	**中等效应 (mean difference around .5 *SD*)**
***d*>0.8**	**大的效应 (mean difference > .8 *SD*)**

21

在论文中报告结果

- 结果是显著的 (不显著的), z = 观测值 value, p < (.05 or .01) (如果不显著, p >.05 or .01).
- 差异的性质是: 哪个组更好, 报告均值和标准差.
- 效应大小是 ________, 表明 ______________________.

22

在论文中报告结果

尽管农民工子女坚持性分数均值 (M = 5.0, SD = 3.8) 本地人子女坚持性分数(M = 4.0, σ = 4.0), 这种差异未达到显著水平, z = 1.25, p >.05. 效应大小是中等, Cohen's d = .26. 预示着大样本将有可能产生显著的结果.

23

I类错误和II类错误

真相

统计决策

	没有效应 H_0 正确	真实效应 H_0 错误
拒绝 H_0	I类错误	正确
接受 H_0	正确	II类错误

24

如何提高统计效力 (1-β)?

$$z_{obt}=\frac{M-\mu}{\sigma_M}$$

$$\sigma_M=\frac{\sigma}{\sqrt{n}}$$

- 扩大样本量
  - 降低标准误
  - 增加样本均值分布是正态的概率
- 减少测量的误差.
- 增加效应.

25

	样本数据	假设总体参数	样本标准差	标准误	t-统计量
单样本t-检验	$\bar{X}$	μ	$s^2=\frac{SS}{df}$	$\sqrt{\frac{s^2}{n}}$	$t=\frac{\bar{X}-\mu}{s_{\bar{x}}}$
相关样本t-检验	$\bar{D}$	μ_D	$s^2=\frac{SS_D}{df}$	$\sqrt{\frac{s^2}{n}}$	$t=\frac{\bar{D}-\mu_D}{s_{\bar{D}}}$
独立样本t-检验	$\bar{X}_1-\bar{X}_2$	$\mu_1-\mu_2$	$s_p^2=\frac{SS_1+SS_2}{df_1+df_2}$	$\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}$	

$$t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{s_{\bar{x}_1-\bar{x}_2}}$$

26

参数估计

27

估计的实际例子

- 某单位进行了管理改进后，与上月的9个匹配的项目的差值的平均值是6.50美圆，SS=72。统计检验证明该差值在统计上是显著的。被问到节约了多少时，研究者决定用95%的置信区间来报告这个估计值。这个置信区间是多少？

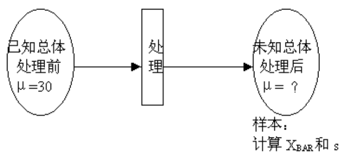
28

估计

- 样本 → 总体
  - 假设检验
  - 估计
- 影响估计的因素
  - 样本容量：2个或100个
  - 估计的类型：点估计或区间估计
    - 6两
    - 0.5两-4斤
    - 在精确度与置信区间，我们有一个trade-off

29

估计的基本实验情境



30

假设检验与估计：统计意义和实际意义

- 假设检验涉及yes-no的决策，不能给出更多的信息
- 一种新药治疗高胆固醇
  - 用药前水平：225
  - 用药后水平：210
  - 假设检验：统计上显著

31

估计的逻辑

- 估计的逻辑与假设检验不同
  - 假设检验：试图否定虚无假设。
  - 估计： 对于总体参数的值作有根据的猜测。
- 何时需要作估计
  - 想了解总体的基本信息，但不能测量到所有个体，所以抽取一个样本
  - 如果已经知道处理有效应，进而想知道效应有多大
  - 经过假设检验后拒绝了H_0
    - "我们拒绝处理没有造成差别，但我们希望知道到底有多少差别。"

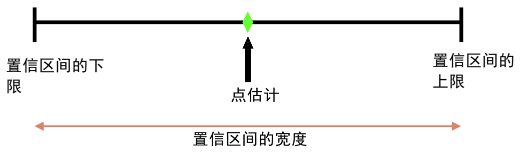
32

两类总体均值估计

- 均值的点估计 *（point estimates）*: 用单一数值作为未知数量的估计
- 均值的区间估计 *（interval estimates）*： 用某一数值的范围作为未知数量的估计.
  - 置信区间（confidence intervals；C.I.）当一个区间与一个特定的置信度 (或概率)一起出现时，称为置信区间.
- 两类估计都由同样的方程所决定，其差别是对于点估计，只计算一个单一的数值，但对于区间估计,应计算两点之间的区间。

33

点估计和区间估计



34

1. 均值的点估计

- 总体均值的最佳单一值估计 ：
  - 如果我们可以得到所有可能随机的样本，那么最佳的估计就是样本均值分布的均值

35

点估计的缺点

- 一个特定的点估计不容易准确,因为它只是许许多多点估计中的一个.
- 点估计没有提供与所估计参数有多相近的点估计值的概率信息.
- 区间估计比点估计提供更多的有关总体特征的信息
- 这些都可被区间估计所克服

36

估计的逻辑步骤

- *Step 1*: 开始作出 “合理的” 估计z (或 t) 值在估计中应该是什么.
- 对于点估计, z = 0, 恰好在中间
  对于区间估计, z (或 t) 值取决于估计要求的置信度
- *Step 2*: 接受关于 z (或 t) 值"合理的" 估计, 然后将其代入公式 以求得未知的总体参数.

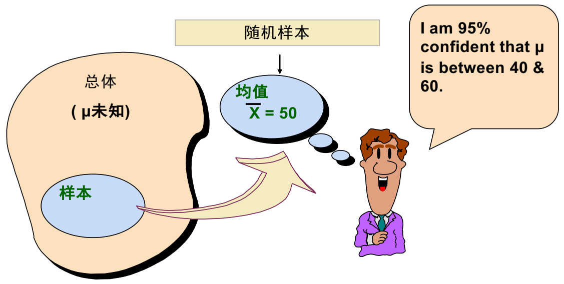
37

2. 均值的区间估计

- 如何有更大的机会使估计准确? 我们可以用区间估计。
- 仍然考虑样本均值分布. 以z分数为例, 取±1 z单位. 这是我们发现大约 68% 的均值落入这个范围. 所以我们可以相当确信总体均值会在这个范围内。

38

估计过程



39

一般公式

- 置信区间的一般公式:

点估计 ± (临界值)(标准误)

40

置信度, (1-α)

- 假定置信度= 95%
- 亦写作 (1 - α) = 0.95
- 解释:
  - 从长远看, 95% 的置信区间包含那个真正的参数.

41

μ的置信区间(σ已知)

- 前提
  - 总体标准差σ已知
  - 总体正态分布
  - 如果总体不是正态,用一个大样本
- 置信区间的估计值

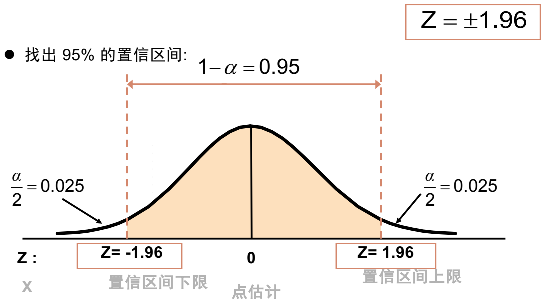
$$\bar{X} \pm Z\frac{\sigma}{\sqrt{n}} \qquad Z=\frac{X-\mu}{\sigma_X}$$

- $\bar{X}$ 是点估计值
  Z 尾端α/2是是正态分布临界值
  $\sigma/\sqrt{n}$ 是标准误

42

找出临界值, Z

- 找出 95% 的置信区间:



43

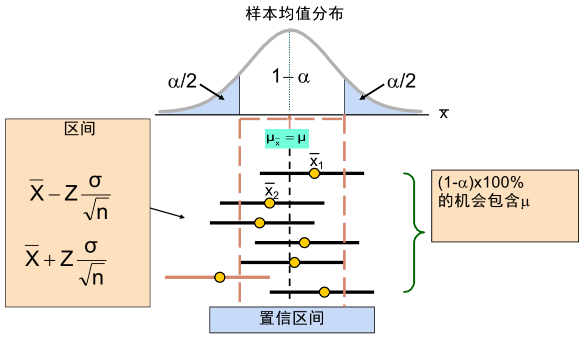
常用的置信度

- 常用的置信度90%, 95%, and 99%

Confidence Level	*Confidence Coefficient,* $1-\alpha$	*Z value*
80%	0.80	1.28
90%	0.90	1.645
95%	0.95	1.96
98%	0.98	2.33
99%	0.99	2.58
99.8%	0.998	3.08
99.9%	0.999	3.27

44

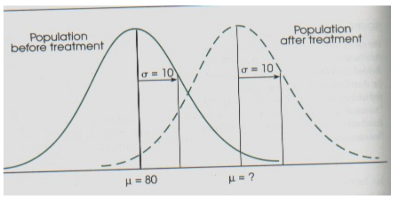
置信区间和置信度



45

用z 分数作区间估计

- 总体标准差已知，总体均值未知
- 被估计的总体常常接受过处理
  - 大学生的平均阅读成绩是80分，参加阅读训练班后学生的成绩有所提高，提高了多少？



46

用z 分数作区间估计

$$Z = \frac{X-\mu}{\sigma_X}$$

- 用同样的公式, 得到区间的最小值和最大值.

  $\mu = X \pm Z(\sigma_X)$

- 估计第一步是要确定置信度

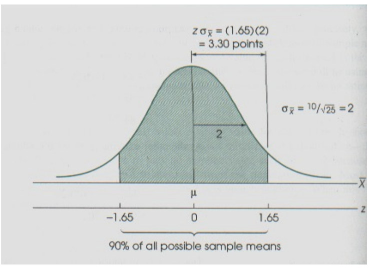
47

z检验例

- X = 85, σ = 10, n = 25，求90%置信区间
- 需要找出样本均值周围90%所对应的两个 z分数，即 双尾各 5%, 所以 z分数是 ±1.65.
- m1 = X + (z)(σ_X) = 85 + (1.65)(10/sqrt (25))= 88.30
  m2 = X - (z)(σ_X) = 85 - (1.65)(10/sqrt (25)) = 81.70

48

z检验例



49

方差未知情况下总体均值的估计

- 如果总体σ未知, 我们可以用样本s来替换它.
- 这引入了额外的不确定性,因为S 因样本不同而不同
- 所以我们用大样本,并且应用中心极限定理
- $\mu = X \pm t_{crit}(S_X)$

$$t = \frac{X - \mu}{s_X}$$

50

单一样本 t 统计量的估计

- 例： X = 2.83, SS= 6, n = 16，求99%置信区间

51

单一样本 t 统计量的估计

- 例： X = 2.83, SS= 6, n = 16，求99%置信区间
- S_{X} = sqrt （SS/df/n） = sqrt（ 6/15/16） =0.158
- t_{crit}= 2.947
- $\mu_1 = X$ -(tcrit)(S_X)=2.83-2.947*0.16=2.36
- $\mu_2 = X$ +(tcrit)(S_X) 2.83+2.947*0.16=3.30

52

两个独立样本 t 统计量的估计

- 两个独立样本 t 统计量的置信区间的估计:

$$X_1 - X_2 \pm t_{\alpha/2} \cdot s_{(\overline{X}_1 - X_2)}$$

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{s_{\overline{X}_1 - \overline{X}_2}} = \frac{(\overline{X}_1 - \overline{X}_2)}{s_{\overline{X}_1 - \overline{X}_2}}$$

53

两个独立样本 t 统计量的估计

- 例: 8个长子的心理测验分数X_1 = 11.4, SS_1= 26
 8个次子的心理测验分数X_2 = 13.9, SS_2=30
 求均值差异80%的置信区间

54

两个独立样本 t 统计量的估计

- 例: 8个长子的心理测验分数X_1 = 11.4, SS_1= 26
  8个次子的心理测验分数X_2 = 13.9, SS_2=30
  求均值差异80%的置信区间
  - S_p^2= （26+30）/14=4
  - $s_{X_1-X_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ =1
  - t_{crit}=1.345
  - $X_1 - X_2$ = 13.9 -11.4=2.5
  - $\mu_1 = (X_1 - X_2) - (t_{crit})(s_{X_1-X_2})$=2.5-1.345=1.155
  - $\mu_2 = (X_1 - X_2) - (t_{crit})(s_{X_1-X_2})$=2.5+1.345=3.845

55

两个相关样本 t 统计量的估计

- 两个相关样本: $D \pm (t_{crit})(S_D)$

$$t = \frac{\overline{D} - \mu_D}{s_{\overline{D}}}$$

56

两个相关样本 t 统计量的估计

- *n*=9，D = 6.50， SS=72，求此差异95%的置信区间

57

两个相关样本 t 统计量的估计

- *n*=9，D = 6.50， SS=72，求此差异95%的置信区间
- s=sqrt（SS/df）=3
- S_D =3/sqrt（n）=3/sqrt（9）=1
- t_{crit}=2.306
- μ 1= D - (t_{crit}) (S_D)=6.50-2.306*1=4.194
- μ 2= D + (t_{crit}) (S_D)=6.50+2.306*1=8.806

58

估计的实际例子

- 某单位进行了管理改进后，与上月的9个匹配的项目的差值的平均值是9.50美圆，SS=72。被问到节约了多少时，研究者决定用99%的置信区间来报告这个估计值。这个置信区间是多少？

59

影响置信区间宽度的因素

1. **置信度**
2. **样本容量**
3. **样本方差**
4. **样本均值的标准误**

60

确定样本容量

确定样本容量

对于均值

取样误差

$$\bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$$

$$e = Z\frac{\sigma}{\sqrt{n}}$$

61

确定样本容量

确定样本容量

对于均值

$$e = Z\frac{\sigma}{\sqrt{n}} \Rightarrow \text{解方程得到} \Rightarrow n = \frac{Z^2\,\sigma^2}{e^2}$$

62

确定样本容量

(continued)

- 为确定样本容量, 我们必须知道:
  - 置信度 $(1 - \alpha)$, 它决定临界 Z 值
  - 可接受的取样误差, e
  - 标准差, σ

63

确定样本容量例题

如果 $\sigma = 45$, 估计均值差异在 ± 5 而置信度在90%,至少需要多大样本?

$$n = \frac{Z^2\,\sigma^2}{e^2} = \frac{(1.645)^2(45)^2}{5^2} = 219.19$$

所需样本量 n = 220

(总要入位)

64

Thanks for your attention!

65

65