

实用卫生统计学

一、名词解释

参数检验：是一种要求样本来自总体分布型是已知的（如正态分布），在这种假设的基础上，对总体参数（如总体均数）进行统计推断的假设检验。

等级相关：是对等级数据作相关分析，它又称为秩相关，是一种非参数统计方法。

非参数检验：非参数检验是一种不依赖总体分布类型，也不对总体参数（如总体均数）进行统计推断的假设检验。

检验效能：检验效能又称为把握度（ $1 - \beta$ ）。它的含义是：当两总体确实有差别时，按规定的检验水准 α ，能够发现两总体间差别的能力。

四格表资料：两个样本率的资料又称为四格表资料，在四格表资料中两个样本的实际发生频数和实际未发生频数为基本数据，其它数据均可由这四个基本数据推算出来。

时点患病率：指在某时点检查时，接受检查的人群中现患病例所占的比例，分子为新老病人数，又称现患率或流行率。

卫生服务调查统计：它是卫生统计的重要内容之一，卫生服务调查统计是从卫生服务资料的设计、收集，整理、分析的角度，来阐述卫生服务研究的特点、研究方法和注意事项，以便使卫生服务研究工作更具有科学性。

正相关：它是说明具有直线关系的两个变量间，存在有正的相关方向，即当 x 增大时， y 有相应增大的趋势，所算得的相关系数 r 为正值。

秩次：即通常意义上的序号，实际上就是将观察值按顺序由小到大排列，并用序号代替了变量值本身。

二、填空题

1.描述计量资料集中趋势的常用指标有**均数、几何均数、中位数**。

2.统计工作的四个步骤：**计划与设计，收集资料，整理资料，分析资料**。

3.卫生服务调查的分析指标包括**卫生服务需要指标、卫生服务利用指标和卫生服务资源指标**。

4.粗出生率的分母是**某年平均人口数**，分子是**同期内活产总数**。

5.患病率的高低受**疾病的发病率、病程**的影响。

6.收集统计资料的三个基本要求**完整正确及时、要有足够的数量、资料的代表性和可比性**。

7.在对同类事务综合评价的应用实践中,对选用的评价模型进行考察,并不断修改补充,使之具有一定的**科学性、实用性与先进性**,然后推广应用。

8、主要从哪两个方面对原始资料进行检查与核对**数据的取值范围检错、数据间的逻辑关系检错**。

9、若某变量服从正态分布，而且样本量较大，则该材料 95%的个体变量值所在范围为 $\bar{X} \pm 1.96s$ ，99%的个体变量值所在范围为 $\bar{X} \pm 2.58s$ 。

10、设计统计表的横纵标目时，基本要求是**符合逻辑、主谓分明**即 横标目在表中作主语，纵标目作谓语，连贯在一起阅读，可以组成一句完整而通顺的话。

11、参数检验方法，如 u 检验，通常是在假设**样本来自总体分布型是已知的（如正态分布）**的基础上对总体参数进行的检验。

12、评价指标的选取是否合适，直接影响到**综合评价**的结论。指标是不是选取的越多越好呢？太多了，事实上是**重复性的指标**，会相互干扰；太少了，可能所选的指标缺乏**足够的代表性**，会产生比较片面的结果。

13、发病率的分子为**某年某病新发生的病例数**，患病率的分子为**检查期间现患病例数**。

14、卫生服务调查中常用抽样方法有**系统抽样** 和 **整群抽样**，在卫生服务调查中如果采用抽样方法，就应当正确估计样本量。

15、自然增长率是**粗出生率**与**粗死亡率**之差。

16、统计表是由**标题、标目、线条、数字**四部分构成。

三、单项选择题

1.用某年全市的出生数和婴儿死亡数计算婴儿死亡率,这种方法是属于**统计描述**。

2.比较身高与体重的变异程度,适宜的指标**变异系数**。

3.某日门诊各科疾病分类资料,可作为**计算构成比**的基础。

4.直条图适用于各自独立的分类资料。

5.要减小抽样误差,最切实可行的方法是**适当增加观察例数**。

6.假设检验中的把握度是 $1 - \beta$ 。

7. 直线回归系数假设检验,其自由度为 $n - 2$ 。
8. 可用于偏态或不明分布的资料的各评价指标诸等级分值的确定方法是百分位法。
9. 随机抽取上海市区 120 名男孩作为样本,测得其平均出手体重为 3.20kg,标准差 0.50 kg。则总体均数 95%可信区间的公式是 $(3.20 \pm 1.96 \times 0.50/120)$ 。
10. 下面哪一个不是问卷的顺序字母顺序。
- 11、下面哪一种统计资料的来源不属于经常性的资料调查问卷。
- 12、原始数据分布不明时,表示其集中趋势宜采用中位数。
- 13、随机选取男 200 人,女 100 人为某寄生虫病研究的调查对象,测得其感染阳性率分别为 20%和 15%,则合并阳性率为 18.3 %。
- 14、统计表中资料暂缺或未记录时,其空缺处通常用……表示。
- 15、某假设检验,检验水准为 0.05,经计算 $P > 0.05$,不拒绝 H_0 ,此时若推断有错,其错误的概率为 β , β 未知。
- 16、作配对比较的符号秩和检验时,其统计量是 T 值。
- 17、若直线相关系数 $r = 0.702$,对 r 作假设检验的结果为 $P > 0.05$,则认为两变量无直线相关。
- 18、在掌握有关历史资料的条件下,对于可数量化的指标,常用的筛选评价指标的方法是多元回归法。
- 19、某市为调查本月流感发病情况,选择 n 名居民为观察对象,观察期内有 m 发病,其中有 q 人曾被感染过两次,则期间流感的发病率为 $(m+q)/n$ 。
- 20、非全面调查中,发达国家大多数采用什么抽样调查分层随机抽样方法或多阶段随机抽样方法。

四、简答题

1.标准正态曲线下横轴上下侧尾部面积为 0.5%、2.5%、5%时对应的 u 值分别是?

答:标准正态曲线下横轴上下侧尾部面积为 0.5%、2.5%、5%时对应的 u 值分别是 -2.58、-1.96、-1.64。

2.率的标准化法的基本思想是?直接标准化法的条件是?

答:当不同的人群的总率进行比较时,若其人群的内部构成(如年龄,性别等)存在差异,而这些内部构成对率有影响,为消除构成不同的影响,要按照统一标准构成对两个人群进行校正。这种选择统一构成,然后计算标准化率的方法称为率的标准化法。直接标准化法需要两个条件(1)资料条件,已知实际人群的年龄别(组)率,且各年龄率无明显交叉(2)选择标准,可选择标准人群的年龄组人口数或构成比。

3. 绘制统计表的基本要求是什么?

答:(1)标题不可缺失,要求用一句简明扼要的话来说明表的内容,并写在表的上方(2)横纵标目的设计要符合逻辑,主谓分明,标目的文字应简明,有单位的要注明(3)线条同样要简洁,除基本线条外,应尽量减少不必要的线条(4)数字应准确无误,要求一律采用阿拉伯数字表示,同一指标的数字的位数应一致,位次对齐,表中不宜留空项(5)备注一般不列入表内,而写在表的下方。

4. 假设检验的结论为什么不能绝对化?

答:假设检验的结论不能绝对化,因为无论拒绝 H_0 或不拒绝 H_0 ,假设检验的结论都有可能犯错误的可能。假设检验通常可能发生下面的两类错误。I 型错误:拒绝了实际上成立的 H_0 ,犯 I 型错误的概率是 α , α 通常为 0.05。II 型错误:接受了实际上不成立的 H_0 ,犯 II 型错误的概率是 β ,一般情况下, β 的大小是未知的。

5. 行×列表资料一般包括哪些资料,它们的检验目的有何不同?

答:行×列表资料一般包括多个样本率、多个构成比资料,其基本数据可整理成 R 行 C 列,称为 $R \times C$ 表,又称行×列表。多个样本率或构成比 χ^2 检验目的是推断其总体率或构成比是否不同。对同一样本资料按其两个无序分类变量(行变量和列变量)归纳成双向交叉排列的统计表,其行变量可分为 R 类,列变量可分为 C 类,这种表称为 $R \times C$ 列联表。列联表资料 χ^2 检验的目的是推断两变量(行变量、列变量)之间分布是否相互独立,用列联表的独立性 χ^2 检验。尽管这两种行列表检验目的和检验假设方面有所不同,但计算 χ^2 值和自由度的公式完全相同。

6、绘制统计表的基本原则是什么?

答:绘制统计表有两个基本原则:1)一张表只能有一个中心,要说明什么问题,应该十分明确;2)横纵标目的排列合理,主谓分明。

7、t 检验, $P < 0.001$,是否说明两总体均数之间差别很大,为什么?

答:、t 检验, $P < 0.001$,说明差别有统计学意义,可以认为两个总体均数不同;但是假设检验的结论不能直接回答差异的大小,“差别有统计学意义”并不意味着两个总体均数相差很大。差别的大小及差别有无实际意义只能进一步根据专业知识来确定。例如:当样本量足够大时,即使两个样本均数间相差很小也可能得出 $P < 0.05$,此时差别虽具有统计学意义,但不一定有实际意义。

8、作多个样本率或多组构成比资料比较的检验，下结论时应注意什么？

答：作多个样本率或多组构成比资料比较的卡方检验，其结论若是拒绝 H_0 ($P < 0.05$)，只能认为各总体率之间（各总体构成比）总的来说有差别，但不能说明各两组率（或各两组构成比）之间均有差别。

9、为什么计算得到的相关系数 r 和回归系数 b 都要做显著性检验？

答：由于 r 与 b 都是根据样本资料求得的，而在由变量 x 与 y 组成的总体中作随机抽样，获取样本的过程中，不可避免地会存在由抽样误差。因此当算得的 r 与 b 不为零时，有可能实际情况是 x 与 y 之间是不相关或不存在直系回归关系的，只是由于抽样误差， r 与 b 才会不为零。作显著性检验的目的正是对这种情况作出判断，看 r 或 b 是否确实由总体相关系数 ρ 或总体回归系数 β 不为零的总体中抽得。

10、什么是根本死因？为什么死因统计要选择根本死因？

答：WHO 规定，根本死因是指：“（a）引起直接导致死亡的一系列病态事件的那些疾病或损伤，或者（b）造成致命损伤的事故或暴力的情况。”制定根本死因的想法是从预防死亡的角度出发，去寻找带有根本性的、引起一系列疾病并最终导致死亡的那个原因，不管它发生在死前多长时间都应予以记录。根本死因可以是一个明确的疾病诊断，可以是一个无明确诊断的医学情况（如：症状、体征、临床表现），也可以是一个意外的损伤和中毒。

五、分析/计算题

1. 12 名健康成年男子的血清胆固醇（mg/dl）如下：222、142、136、212、129、207、172、150、161、216、174、186，求均数和标准差？

答： $\Sigma X = 2107$
 $\Sigma X^2 = 381911$
 $= 175.58$
 $s = 32.97$

2. 某防疫站观察 3 种药物驱虫的疗效，在服药 7 天后查粪中虫卵的阴转率如表，问三种药物的疗效是否相同？

3 种药物驱虫的疗效比较			
药物	服药人数	阴转人数	阴转率%
甲药	37	28	75.7
乙药	38	18	47.4
丙药	34	10	29.4

附：自由度为 2 时的 χ^2 界值如下

概率	0.50	0.25	0.10	0.05	0.01
χ^2	1.39	2.77	4.61	5.99	9.21

答：这是完全随机设计的三个样本率的比较，该资料可整理为如下表，用行 $\times$ 列表资料的 χ^2 检验。

不同药物驱虫效果比较				
药物	阴转人数	阴转人数	服药人数	阴转率%
甲药	28	9	37	75.7
乙药	18	20	38	47.4
丙药	10	24	34	29.4
合计	56	53	109	51.4

（1）建立检验假设，确定检验水准

H_0 : 三种药物驱虫阴转率相同，即 $\pi_1 = \pi_2 = \pi_3$

H_1 : 三种药物驱虫阴转率不相同或不全相同

$\alpha = 0.05$

(2) 计算统计量 χ^2 值

$\chi^2 = 15.56$

查 χ^2 表, 拒绝 H_0 , 接受 H_1 , 认为三种药物的阴转率不相同。

实用卫生统计学

单选(211)...

- 1、12名妇女分别用两种测量肺活量的仪器测最大呼气率(L/min)，比较两种方法检测结果有无差别，可进行（）。-->配对设计 t 检验
- 2、t 检验的作用是（）。-->C.均数的差异由抽样误差所引起的概率大小
- 3、按一定的概率估计总体参数所在的可能范围的方法。-->E.区间估计
- 4、百余例的资料编制频数表时，一般设组数为？（）。-->B.8~15组
- 5、比较8岁男童与18岁男青年的身高的变异程度宜用指标-->变异系数
- 6、比较两家医疗机构近20年来床位数的增加速度，应当使用的统计图是（）。-->D.半对数线图
- 7、比较两种药物疗效时，对于下列情况可作单侧检验（）。已知A药不会优于B药。-->标准差
- 8、比较身高与体重的变异程度，适宜的指标（）。-->D.变异系数
- 9、编制某年某地区的女性简略寿命表需要（λ-->C.出生活产数、年龄组平均人口数和死亡数
- 10、变异系数越大，则以下正确的是（）。-->C.以均数为准变异程度
- 11、表达某地两年几种疾病的患病率，宜绘制的图形是（）。-->B.复式直条图
- 12、表达某地两年几种疾病的患病率，宜绘制哪种图形？（）。-->B.复式直条图
- 13、不受年龄结构影响，能够反映整个人群死亡水平的指标是（）。-->C.标准化死亡率

14、测定某地100名健康成年女性的血红蛋白量，则其总体均数95%置信区间的为（）

$$\bar{X} \pm 1.96 S_{\bar{X}}$$

15、测定某地100名健康成年女性的血红蛋白量，则其总体均数95%置信区间的为（）。

$$A. \chi^2 = \frac{(|b-c|-1)^2}{b+c}$$

16、测量6名健康男子，获得其血清总胆固醇（mmol/L），这种资料是（）。-->A.数值变量资料

17、抽样研究中，设、不变，若逐渐增大样本含量，则样本？（）。

-->B.标准误减小

18、抽样研究中，设样本标准差不变，若逐渐增大样本含量，则会出现下面哪种情况？（）。-->B.标准误减小

19、抽样研究中的样本是（）。-->C.总体中随机抽取的一部分

20、从某市18岁男学生随机抽取20名，测量的身高均数是166.13cm，标准差为5.24cm，则该市18岁男生身高均数的95%可信区间为（）错误提示：小样本需用_t_分布法-->(163.68,168.58) cm

21、粗出生率的分子是（）。-->B.多元回归法

22、等级资料的比较宜采用（）-->D.秩和检验

23、等级资料的两个样本比较秩和检验，其检验统计量T是（）。

-->C.以例数较小者秩和为T

24、动态数列中的定基比指标是？-->相对比

25、动态数列中的绝对增长量指标是？绝对数-->

26、对10名8岁女童，分别用两种测量肺活量的仪器测量最大呼气率(L/min)，比较两种方法检测结果有无差别，可用以下哪项方法（）-->C.配对设计 t 检验

27、对称分布是指（）。-->D.集中位置在正中，左右两侧频数分布大体对称

28、对两地肝癌死亡率比较时，作率的标准化其目的是（）。-->C.除两地人口构成不同的影响

29、对某样品进行测量时，由于仪器未事先校正，造成测量结果普遍偏高，这是（）。-->A.系统误差

30、对某样品进行测量时，由于仪器未事先校正，造成测量结果普遍偏高，这样产生的误差是（）。-->A.系统误差

31、对某样品进行测量时，由于仪器未事先校正，造成测量结果普遍偏高，这种误差是（）。-->A.系统误差

32、对样本相关系数进行检验，得到 $t_r < t_{0.05, n-2}$ ，则下列各项正确的是。-->两变量间存在相关关系

33、对样品进行测量时，由于测量仪器事先未校正，造成测量结果普遍偏高，这种误差是？-->A.系统误差

34、对于率的标准化的理解，不正确的是（）。-->D.校正后得到的总率能更好的反映实际水平

35、对于偏态分布资料，适合对其进行参数估计的方法是（）。-->B.百分位数法

36、对于同一资料，如果 $t_r > t_{0.05, v}$ ，则有（）。-->t b t0 .05.v

37、对于正态分布的资料，适合对其进行参数估计的方法是（）。-->C.Z 分布法

38、对于正态分布的资料，以下适合对其进行参数估计的方法是（）。-->C.Z 分布法

39、多个样本均数比较的假设检验，若总体方差不等且分布呈偏态，宜选用的统计推断方法是（）。-->D.H 检验

40、多组资料比较的秩和检验，确定P值时，可利用查表法的情况正确的是（）。-->D.组数≤3，每组例数≤5

41、反映等比资料集中趋势的指标-->几何均数；

42、反映疾病严重程度的指标是（）。-->B.两周卧床率、两周活动受限率、两周停工（学）率和每人每年因病伤卧床日数

43、反映偏态分布资料的平均水平-->中位数；

44、方差分析中，

当 $F > F_{0.05(v1, v2)}$ ， $P < 0.05$ 时，结果是（）。

答案：B

解析：B.可认为各总体均数不等或不全相等

45、方差分析中， $P < 0.05$ 时，应进一步作下列哪项检验？（）-->C.两两比较的 t 检验

46、方差分析中，当 $F > F_{0.05(v1, v2)}$ ， $P < 0.05$ 时，结果（）。方差分析中，当 $F > F_{0.05(v1, v2)}$ ， $P < 0.05$ 时，结果（）。B.可认为各总体均数不等或不全相等

47、方差分析中，当 $P < 0.05$ 时，应进一步作下列检验（）。-->q 检验

48、非全面调查中，发达国家大多数采用什么抽样调查（）。-->D.分层随机抽样方法或多阶段随机抽样方法

49、根据某地7岁~12岁小学生的近视情况的调查资料，对患者的年龄分布进行统计描述，宜绘制哪种图形？（）-->直方图

50、关于变异系数，下面哪个说法是错误的？-->变异系数的单位与原始数据相同

51、关于变异系数，下面哪个说法是错误的？（）。-->A.变异系数的单位与原始数据相同

52、关于标化后的总死亡率，以下正确的是（）。-->D.仅仅作作为比较的基础，它反映了一种相对水平

53、关于标准差，下面哪个说法是正确的（）。-->B.标准差必定大于或等于零

54、关于标准差，下面哪个说法是正确的？-->同一资料的标准差一定比其均数小

55、关于标准差，下面哪个说法是正确的？（）B.标准差必定大于或等于零

56、关于等级相关的叙述，不正确的是（）。-->D.等级相关适用于任何类型的资料

57、关于构成比指标，描述正确的是（）。-->B.表示事物内部各部分的比重大小

58、关于构成比指标，描述正确的是（）。B.构成比表示事物内部各部分的比重大小死亡

59、关于构成比指标，描述正确的是（）。-->B.构成比表示事物内部各部分的比重大小

60、关于率，描述正确的是？-->率是说明某现象发生的频率或强度的指标

61、关于相对数，下列哪一个说法是错误的（）。-->D.率与构成比意义虽然不同，但性质相近。可以混用

62、关于相对数，下列哪一个说法是错误的？-->率与构成比意义不同，但性质相近，经常可以互用

63、关于相对数，下列哪一个说法是错误的？（）。-->D.率与构成比意义虽然不同，但性质相近，经常可以混用

64、关于相对数，下列说法错误的是（）。-->D.率与构成比意义不同，但性质相近，经常可以互用

65、关于相关系数，下列说法错误的是（）。-->D.相关系数与回归系数的符号相同，且呈正比关系1

66、关于相关系数的假设检验正确的是（）。-->以上均可

67、横轴为连续性变量的生1段，同时要求各组段的组距相等，这种统计图为（）。-->C.直方图

68、计算麻疹疫苗接种后血清检查的阳转率，分母应该是（）-->C.麻疹疫苗接种人数

69、假设检验中的把握度是？（一）。-->D.1-p

70、假设已知某地35岁以上正常成年男性血压收缩压的总体均数为120.2mmHg，标准差为11.2mmHg，后者反映的是（）。-->A.个体变异

71、将某地居民的性别、年龄结合起来分组，研究不同性别、年龄别的住院率，这样得到的统计表属于（）。-->B.复合表

72、均数和标准差的关系是（）。-->D.标准差越小，均数对各变量值的代表性越好

73、可以用来描述近视率与年龄关系的统计图是（）。-->C.散点图

74、可用于偏态或不明分布的资料的各评价指标诸等级分值的确定方法是？-->D.百分位法

75、来自正态总体且方差齐性的多个样本均数比较时，通常选择的统计方法是（）。-->D.方差分析

76、两变量分布类型未知时可进行的相关分析可用（）-->B.等级相关分析

77、两变量未知分布型，应用（）。-->校正的等级相关系数

78、两个大样本均数比较的u检验（ $\alpha=0.05$ 的双侧检验），（u）=2.26统计结论是？-->A.P005

79、两个样本回归系数进行比较，得到差别无显著性的结论，以下各项正确的是（）。-->B.两样本所代表的总体回归系数相等的可能性较大

80、两样本均数假设检验的目的是判断（）。-->两总体均数是否相等

81、两组资料中，对于回归系数6较小的一组，下面说法正确的是（）。-->D.r可能较大也可能小

82、满足参数检验的计量资料如果用秩和检验，拒绝H0时可能会（）。-->B.增加II类错误的概率

83、描述对称分布资料的个体观察值的离散趋势-->标准差

84、描述对称分布或正态分布资料的平均水平-->均数；

85、描述末端无确定值资料的离散程度-->四分位数间距；

86、描述偏态分布资料个体观测值的变异程度，宜用以下哪个指标？（）。-->D.四分位数间距

87、描述偏态分布资料个体观察值的离散趋势-->四分位数间距；

88、描述一组偏态分布资料的变异度时，最适宜采用的指标是？（）。-->C.四分位数间距

89、某护士记录了50名婴儿出生体重的测定结果，小于2500克15人，介于2500克和3999克的有25人，大于4000克的有10人，此资料属于（）。-->有序分类变量资料

90、某年某地6岁的男燕身高服从正态分布，其均数为115.0CM，标准差为10CM，下面说法正确的是？（）。-->C.2.5%的6岁男儿童的身高大于134.6CM

91、某年内两所规模、等级相近的医院对5种类型的肝炎病人的治疗中，几乎都是甲医院的治愈率高于乙医院，但是总的治愈率与此截然相反，原因可能是？-->两个医院各型肝炎病人构成不同

92、某人为了研究某药物的疗效，用该药治疗10人，其中6人治愈，其疗效最好表示为？-->6月10日

93、某日门诊各科疾病分类资料，可作为（）。-->C.计算构成比的基础

94、某市为调查本月流感发病情况，选择n名居民为观察对象，观察期内有m发病，其中有q人曾被感染过两次，则期间流感的发病率为（）。-->C. (m+q)/n

95、某医师用A药治疗9例病人，治愈7人，用B药治疗10例病人，治愈1人，比较两药疗效时，适宜的统计方法是？（）。-->B.直接计算概率法

96、某医院用某种新疗法治疗某病患者，治疗结果见下表，请问该资料的类型是（）。

治疗效果	治愈	显效	好转	恶化	死亡
治疗人数	15	45	6	4	0

C. 有序分类变量资料

97、某医院用某种新疗法治疗某病患者，治疗结果见下表，请问该资料的类型是？-->有序分类变量资料

98、某组资料共5例，变量的平方和以及变量的和分别是190和30，则均数和标准差分别为？-->6、1.58

99、配对计量资料，差值分布不接近正态分布，宜用何种检验（）。-->B.秩和检验

100、配对计量资料，差值分布不接近正态分布，应选用以下哪种方法？（）。-->C.配对设计差值的符号秩和检验

101、配对计数资料的卡方检验公式中，b是指（）。-->B.表示甲乙(+、-)的对子数

102、配对设计的符号秩和检验中，其检验假设H0为（）。-->C.差值总体中位数等于零 Md=0

103、配对资料比较的Wilcoxon符号秩和检验，其检验统计量T是（）。-->以秩和较小者为秩和 T

104、配对资料的差别的卡方检验，其备择假设是（）。-->D.B≠C

105、配对资料设计的秩和检验，确定P值时，可利用查表法的样本例数的范围为？（）-->5≤n≤50

106、偏态分布数据，描述集中趋势宜用的指标是（）-->D.中位数

107、偏态分布数值资料，经对数变换后仍呈偏态分布，其集中趋势宜用（）。-->C.中位数

108、如果能用t检验处理的资料而用秩和检验，其缺点是（）。-->A.检验效率降低

109、若要通过样本进行统计推断，样本应该是（）。-->C.总体中随机抽取的一部分

110、三个率比较的卡方检验，若P<0.05，则结论是（）。-->C.至少有两个总体率有差别

111、三行四列表作Z2检验，当有4个格子的理论频数1<T<5时，应采取的方法是（）。-->B.增加样本例数

112、散点图中散点密集于一条直线，且呈水平分布，可初步判断两变量为（）。-->C.零相关

113、是指用样本指标（称为统计量）估计总体指标（称为参数）。-->C.参数估计；

114、双侧Z检验， $\alpha=0.05$ 对应的Z界值是（）。-->B.1.96

115、四格表 χ^2 检验中， $\chi^2 < \chi^2_{0.05, 1}$ ，可认为以下哪项正确？（）-->两总体率相同

116、四格表中的一个实际频数为1时，以下成立的是（）。-->B.还不能确定是否可做X2检验

117、四格表资料， $n>40$ ，有一个理论频数小于5但大于1，其他理论频数都大于5，此数据应做何种检验？（）-->必须用校正的 χ^2 检验

118、四格表资料的卡方检验，其校正条件是（）。-->D.有一个理论数小于5大于且n40

119、四格表资料的卡方检验的校正条件是？（）。-->A.有理论数大于1

120、四行四列表作X2，检验，当有2个格子的理论频数T<1时，以下哪项可取？（）B.应进行适当的合并

121、随机测量某地7岁150名男童体重，其均数是22.6kg，标准差为3.2kg，则该地7岁男童身高均数的95%可信区间为（）错误提示：大样本需用正态分布近似法(21.14, 23.11)kg-->(22.09, 23.11) kg

122、随机抽取北京8岁男童100名作样本，测得其平均出生体重为3.20kg，标准差为0.5kg。则总体均数95%置信区间的公式是

123、随机选取男200人，女100人为某寄生虫病研究的调查对象，测得其感染阳性率分别为20%和15%，则合并阳性率为（）。-->B.18.3%

124、随着测定次数的增加，正负误差可以相互抵偿，误差的平均值将逐渐趋向于零，是指（）。-->C.随机测量误差

125、随着测定次数的增加，正负误差可以相互抵偿，误差的平均值将逐渐趋向于零，是指哪种误差？（）。-->C.随机测量误差

126、缩写NIR是指（）。-->A.自然增长率

127、缩写NRR代表的是（）。-->A.净再生率

128、通常是均数的标准差，它反映了均数之间的离散程度；也反映了均数抽样误差的大小。-->B.标准误；

129、统计表的基本结构包括（）。-->A.标题、标目、线条、数字

130、统计表中不应出现的项目是（）。-->D.备注

131、统计表中某处无数字，对应表中应填写以下哪项？（）C.—

132、统计表中暂缺或未记录，对应表中应填写的是（）。-->A.-----

133、统计表中资料暂缺或未记录时，表示其空缺处的通常做法是（）。-->B.某年平均人口数

134、统计推断的内容是？（）。-->D.假设检验，用样本估计相应总体指标

135、完全随机设计的方差分析中的组间均方是（）。-->反映了随机误差和可能存在的处理因素的综合结果。

136、完全随机设计的两个大样本的均数比较（总体标准差未知），可考虑？-->D.以上都可

137、完全随机设计的两样本比较秩和检验，其检验统计量 T 是（）。-->以例数较小者秩和为 T

138、完全随机设计方差分析中从总体变异中分出组间变异和组内变异是指？（）。-->B.从总离均差平方和分出组间离均差平方和组内离均差平方

139、下列不属于经常性来源的资料是（）。-->D.调查问卷

140、下列各项正确的是（）。-->r 越大，散点图越集中于直线上

141、下列各项中正确的是（）。-->r 0，一定有 b 0

142、下列关于假设检验的描述，正确的是（）。-->D.统计学得出差异有显著性的结论，并不一定专业上有意义

143、下列关于相对比计算的叙述，正确的是（）。-->D.对公式中的两指标的性质量明确限制，相对比的用途可以很广

144、下列哪个算式可出现负值？（）。--> $\sum XY - (\sum X)(\sum Y) / n^2$

145、下列说法正确的是（）。-->C.回归系数不能反映两变量的关系是否密切

146、下列统计方法中不属于参数检验的是（）。-->D.秩和检验

147、下列有关等级相关的说法，错误的是（）。-->D.等级相关分析是用相关系数 r 来说明两变量相关关系的密切程度与相关方向

148、下面不是反映疾病的频率指标是（）。-->C.两周每千人患病日数

149、下面的变量中，属于分类变量的是（）。-->B.血型

150、下面的评价指标的权重估计方法中，哪一种是客户定权法？（）。-->D.相关系数法

151、下面哪个指标是样本指标？（）。-->D.

152、下面哪一个不是问卷的顺序？（）。-->C.字母顺序

153、下面说法正确的是（）。-->r 值大小与 b 值大小无关

154、下面说法正确的是（）。-->D.经 r 的显著性检验，拒绝 H₀。可认为两变量存在相关

155、下面有关病人的变量中，属于分类变量的是（）。-->性别

156、线图适用于（）。-->B.连续性资料

157、性质相同的两份资料做回归分析，在 $a=0.05$ 水平， $t_{0.01}$ ， $(n-2) > t_{b1} > t_{0.05}$ ， $(n-2)$ ， $t_{b2} > t_{0.01}$ ， $(n-2)$ ，则下列各项正确的是（）。-->b 1 和 b 2 的大小无法确定

158、样本相关系数 $r=0$ ，说明（）。-->两变量间的关系不确定

159、样本相关系数抽样于（）。-->C.总体相关系数

160、要表示某校 18 岁女生体重与肺活量的相关关系，宜绘制下列哪一种统计图？（）。-->C.散点图

161、要减小抽样误差，最切实可行的方法是（）。-->A.适当增加观察例数

162、一偏态分布数值资料读袭习数变换仍呈偏态分布，其集中趋势宜用（）。-->C.中位数

163、一组变量值，其大小分别为 10，15，11，6，13，24，23，9，其中位数是（）。-->B.12

164、医学人口统计应属于卫生统计学中的哪部分内容？（）。-->A.健康统计

165、已知 $r=1$ ，则一定有（）。--> $S_{yx}=0$

166、已知 $t_{b1}=t_{b2}$ ，则一定有（）。--> $t_{r1}=t_{r2}$

167、已知双侧 $t_{0.05/2, 18}=2.101$ ，若 $t=2.78$ ，则 t 出现的概率（）。-->P 0.05 布分布资料的个体观察值的离散趋势

168、已知双侧 $F_{0.05/2, 18}=2.101$ ，若 $t=2.78$ ，则 F 出现的概率为（）。-->C.P<0.05

169、以下不是配对 χ^2 检验前提条件的是（）。-->D.样本来自正态分布总体，样本量较小，配对设计数值变量资料

170、以下关于非参数检验的描述错误的是（）。-->A.非参数检验的检验效能高于参数检验

171、以下关于统计资料的要求的选项错误的是（）。-->D.研究对象越多越好，资料就更完整

172、以下检验方法中，不属于非参数检验方法的是（）。-->A.Z 检验

173、以下哪项不是配对 χ^2 检验前提条件？（）。-->D.样本来自正态分布总体，样本量较小，配对设计数值变量资料

174、以下哪项公式适用于偏态分布资料的参数估计？（）。
 $P_x \sim P_{100-x}$ (双侧)

175、以下哪项可以用以指明统计表内数字含义的？（）。-->标题

176、以下哪项是方差分析的使用条件（）。-->A.样本独立随机，来自正态分布总体，各总体的方差相等

177、以下哪项是样本均数的标准误？（）。
 $S_{\bar{x}}$

178、以下哪项是总体均数的标准误的计算公式？（）。
 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

179、以下哪项是总体率的标准误的计算公式？（）。
 $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$

180、以下哪项为配对 χ^2 检验的计算公式？（）
 $A. \chi^2 = \frac{(|b-c|-1)^2}{b+c}$

181、以下哪项用以说明统计表的内容？（）。-->备注

182、以下是总体均数的标准误的计算公式为（）。
 $A. \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

183、以下指标中，哪项指标是构成比？（）
 $A. 残疾百分比 = \frac{\text{某种(类)残疾数}}{\text{所有残疾数}} \times 100\%$

184、以下指标中，哪项指标是绝对数？-->平均人口数 = (初期人口数 + 期末人口数) / 2

185、以下指标中，哪项指标是率？-->某病发病率 = 该年(时期)新发生的某病的病例数 / 同年(时期)内可能发生某病的平均人口数 × K

186、以下指标中，哪项指标是相对比？-->每千人口护士数 = 本地区护士人口总数 / 调查地区人口数 × 1000‰

187、以下指标中，是绝对数的是（）。
 $c. \text{平均人口数} = \frac{\text{期初人口数} + \text{期末人口数}}{2}$

188、以样本统计量来直接代表总体参数的方法。-->D.点估计；

189、用触诊和 X 摄片对 100 名妇女做乳腺检查，触诊有 40 名阳性，X 摄片有 70 名阳性，两种均阳性者为 10 名，两种方法检查均为阴性的人数是（）。-->20

190、用触诊和 X 摄片对 100 名妇女做乳腺检查，触诊有 40 名阳性，X 摄片有 70 名阳性，两种均阳性者为 10 名，宜选用下面哪种方法进行假设检验？-->配对 χ^2 检验

191、用均数和标准差可全面描述下列哪种资料的分布特征？-->正态分布资料

192、用某年全市的出生数和婴儿死亡数计算婴儿死亡率，这种方法是属于？-->A.统计描述

193、用某年全市的出生数和婴儿死亡数计算婴儿死亡率，这种方法属于（）。B.统计描述

194、用频数表计算平均数时，各组的组中值应为（）。-->D. (本组段变量值的下限值 + 本组段变量值的上限值) / 2

195、欲了解某市某年所有三级甲医院的病床数，该市每个三级甲医院就是一个（）。-->B.观察单位

196、欲描述末端元确定数据的资料平均水平时，宜采用的指标是（）。-->C.中位数

197、原始数据分布不明时，表示其集中趋势宜采用（）。-->A.中位数

198、在分析相关系数 r 时可（）。-->算出 r 值后，还学要做显著性检验才能确定两变量的有无线性相关

199、在寿命表编制中，假定的一代人为（）。-->B.100000 人

200、在同一总体中随机抽取样本含量相同的若干样本时，样本指标之间的差异以及样本指标与总体指标的差异。-->A.抽样误差

201、在下列何种情况下，可以认为判断矩阵具有较满意的一致性（）。-->A.当 CR0.10 时

202、在一个假设的总体（总体率 $\pi=30\%$ ）中，随机抽取 $n=100$ 的样本，得样本率 $p=32\%$ ，则产生样本率与总体率不同的原因是（）。-->C.抽样误差

203、在掌握有关历史资料的条件下，对于可数量化的指标，常用的筛选评价指标的方法是（）。-->B.多元回归法

204、正常男子的血铅含量系偏态分布资料，对数变换后的呈正态分布。欲描述血铅含量的平均水平宜用以下哪种指标？-->原始数据的几何均数

205、正态近似法估计总体率的 95 置信区间用以下哪个公式（）。--> $A. p \pm 1.96sp$

206、直条图适用于（）。-->C.各自独立的分类资料

207、直线回归的前提条件是（）。-->A.线性、独立、正态、等方差

208、直线回归系数假设检验，其自由度为（）。-->C.n-2

209、直线相关分析可以应用于以下哪项研究？（）。-->B.儿童的身高与肺活量

210、总体率置信区间的估计符合下列哪种情况时，可以借用正态近似法处理（）。-->A.np 和 n(1-p) 大于 5 时

211、作多个样本率的假设检验，检验假设是（）。-->C. $\pi_1=\pi_2=\dots=\pi_n$

判断(81)-

1、 $r=0.9$ ，就可以认为两变量为正相关，且相关密切。-->错

2、t 检验可用于同一批对象的身高和体重均数差异的比较。-->错

3、t 检验是对两个样本均数的差别进行假设检验的方法。-->对

4、t 检验是对两个样本均数的差别进行假设检验的方法之一。-->对

5、 χ^2 检验，若 $n \leq 40$ ，或 $1 < T \leq 5$ 时，需用确切概率算法。-->错

6、比较 8 岁男童与 18 岁男青年的身高的变异程度宜用指标标准差。-->错

7、编制频数表时，第一组段要包括最小观察值，最后一个组段要包括最大观察值。-->

8、表示全部观察值的变异度的统计指标是 MS 组间。-->错

9、参数估计是以样本统计量来直接代表总体参数的方法。-->错

10、参数检验的统计分析的效率较高，但是对样本所对应的总体分布有比较严格的要求。-->对

11、参数是指样本指标。-->错

12、抽样研究是对从所研究的总体中随机抽取有代表性的一部分个体构成的样本进行的研究称为抽样研究。-->对

13、单侧 Z 检验，检验水准 $\alpha=0.05$ ，对应的 z 界值是 1.96，检验水准 $\alpha=0.01$ ，对应的 Z 界值为 2.58。答案：错

答案：错

14、点估计是指用样本指标（称为统计量）估计总体指标（称为参数）。-->错

15、对变量进行测量或观察的值称为变量。-->错

16、对于有序行列表，在比较各处理组的效应有无差别时，用 χ^2 检验不是最好的方法。-->对

17、多个样本均数间的两两比较，在处理组数大于 2 时，若用 t 检验对任意两均数之间进行多重比较，则会增大犯 I 型错误的概率。-->对

18、反映等比资料集中趋势的指标宜用中位数。-->错

19、反映偏态分布资料的平均水平可用几何均数。-->错

20、方差分析的目的是分析各组总体方差是否不同。-->错

21、方差分析时，方差齐性是指各组样本的方差相等。-->错

22、方差分析时，要求各组方差齐性。-->错

23、方差分析中，F 值不会是负数。-->对

24、方差分析中自由度的计算公式为： $v_{总}=N-1$ ， $v_{组间}=K-1$ ， $v_{组内}=N-K$ ，N 是样本量，K 是处理组数。-->对

25、非参数检验的方法只有秩和检验。-->错

26、非参数统计方法不对待定分布的参数作统计推断，但仍要求数据服从正态分布。-->错

27、分类变量资料是指将全体观察单位按照某种性质的不同程度分为若干组，分别清点各组中观察单位的个数所得的资料，这种资料也称有序分类变量资料。-->错

28、分类变量资料中的各类别必须互相排斥，不能相互包含。-->对

29、根据样本进行相关分析，相关分析的假设检验 $P < 0.01$ ，说明两样本的来自高度相关的总体。-->错

30、观察单位（或个体）的某种属性或标志称为变量保留值。-->对

31、环比发展速度可以反映事物在一定时期的发展速度。-->错

32、假设检验结果为 $P < 0.01$ ，则可以认为两样本均数之间确实存在很大的差别。-->错

33、进行假设检验的前提之一是两组资料具有可比性。-->错

34、卡方检验，是一种用途较广的计数资料的假设检验方法，属于非参数检验的范畴。-->对

35、两次 t 检验都是对两个不同样本均数的差别进行假设检验，一次 $P < 0.01$ ，一次 $0.01 < P < 0.05$ ，就表明前者两样本均数差别大，后者两样本均数差别小。-->错

36、率的标准化的条件是不同人群各年龄组率未知，选择标准人群各年龄组人数或构成比，即可用直接法进行标化。-->错

37、描述对称分布或正态分布资料的平均水平可用均数。-->错

38、描述末端无确定值资料的离散程度宜用四分位数间距。-->对

39、某地某年交通事故死亡 200 人，其中摩托车造成的死亡 50 人，汽车造成死亡 150 人，因此，可以认为摩托车比汽车更安全。-->错

40、某地省级医院脑溢血患者的病死率高于县、乡级医院，故可认为县、乡级医院水平高于省级医院。-->错

41、某地省级医院脑溢血患者的病死率高于县、乡级医院，故可认为县乡级医院水平高于省级医院。-->错

42、某医生治疗一例癌症患者后痊愈，因此该医生的癌症治愈率为 100%。-->错

43、某医师比较甲乙两种治疗方法的疗效，作假设检验，若结果 $p < 0.05$ ，说明其中某一疗法优于另一疗法；若 $P < 0.01$ ，则说明其中某一疗法非常优于另一疗法。-->错

44、某医师用针灸疗法治疗 3 例失眠患者，其中 2 例有效，针灸法的有效率为 66.6%。-->错

45、配对计数资料的卡方检验中， $b+c < 40b$ 时，可以不用校正的配对计数资料的卡方检验。-->错

46、配对设计差值符号的秩和检验的查表法对样本量没有要求。-->错

47、配对资料的差别的卡方检验，

如果 $\chi^2 > \chi^2_{0.01,1}$ ，可认为两样本来自的总体率不同。（）

答案：对

48、配对资料的差别的卡方检验，如果 $\chi^2 > \chi^2_{0.01,1}$ ，1，可认为两样本率不同。-->错

49、频数表常作为小样本资料的陈述形式。-->错

50、平均发展速度说明一定时期的平均发展速度。-->对

51、区间估计是按一定的概率估计总体参数所在的可能范围的方法。-->对

52、如果个体间有变异，则他们一定不是来自同一总体。-->错

53、数值变量和分类变量可以互相转换。-->错

54、数值变量可以转换成分类变量。-->对

55、数值变量资料是指用度量衡的方法测定每个观察单位的某项研究指标量的大小，所得到的数据（即测量值），就称为数值变量资料。-->对

56、双侧 Z 检验，检验水准 $\alpha=0.05$ ，对应的 z 界值是 1.96，检验水准 $\alpha=0.1$ ，对应的 Z 界值为 2.58。-->对

57、四格表资料校正 χ^2 检验，如 $\chi^2 > \chi^2_{0.05,1}$ ，可认为两样本率不同。-->对

58、通常 σ_x 是均数的标准差，它反映了均数之间的离散程度；也反映了均数抽样误差的大小。-->对

59、同一资料，回归系数 b 与相关系数 r 的符号相同。-->

60、同质的个体间不存在变异。-->错

61、统计表中某处无数字，对应表中应填写“.....”。-->错

62、统计量是指总体指标。-->错

63、统计量是指总体指标。-->错

64、统计推断是用统计图表或计算统计指标的方法表达一个特定群体（这个群体可以是总体也可以是样本）的某种现象或特征。-->错

65、统计资料要求样本量越大越好。-->错

66、无效假设不成立，则实际频数与理论频数之差一般不会很大，即出现大的 χ^2 值的概率是小的。-->错

67、系统误差是指样本指标与样本指标或样本指标与总体参数间存在的差别，是由个体变异造成的，是抽样机遇所致的误差。-->错

68、线图的纵坐标必须从零开始。-->错

69、相关系数的 t 检验，自由度 $v=n-2$ 。-->对

70、一个好的统计表可以说明多个中心内容，这是因为它能代替繁杂的文字。-->错

71、有序分类变量资料：是将全体观察单位按照某种性质或类别进行分组，然后分别清点各组中的例数，这样的资料也称计数资料。-->错

72、在 y 和 x 的回归分析中，若 $t_b < t_{0.05, v}$ ，可认为两变量存在直线关系。

在 y 和 x 的回归分析中，若 $t_b < t_{0.05, v}$ ，可认为两变量存在直线关系。（）

73、在同一总体中随机抽取样本含量相同的若干样本时，样本指标之间的差异以及样本指标与总体指标的差异称为抽样误差。-->对

74、直方图是用等宽直条的长短来表示相互独立的各指标的数值大小，直方图的要求变量分组后组距相等。-->错

75、直条图的纵轴应从零开始，中间不要折断。-->对

76、直线回归分析中，b 表示变量 X 每增加（减）一个单位，Y 平均改变 b 个单位。-->错

77、只要增加样本例数就可以避免抽样误差。-->错

78、秩和检验中编秩时，遇不同组相同数据取平均秩次。-->对

79、总体是根据研究目的确定的同质观察单位的全体，更确切地说，是同质的所有观察单位某种变量值的集合。-->对

80、总体研究是对无限总体中的每个个体都作观察。-->错

81、总体研究是对无限总体中的每个个体都作观察就称总体研究。-->错

填空(25)-

1、8 统计资料的两个主要来源可概括为（）、（）。-->经常性资料；一次性资料

2、参数估计是指用样本指标估计总体指标，参数估计有两种方法，一种方法是（），另一种方法是（）。-->点估计；区间估计

3、参数检验方法，如 u 检验，通常是在假设（）的基础上对总体参数进行的检验。-->样本来自总体分布型是已知的(如正态分布)。

4、常用的相对数有（），（），（）。-->率构成比；相对比

5、非参数检验适用于（）分布的资料，如（），又如（）的资料。-->任何严重偏态分布的资料；分布形状不明

6、非参数检验适用于（）分布的资料，如（）又如（）的资料。-->任何严重；偏态分布的资料；分布形状不明

7、描述计量资料集中趋势的常用指标有（）、（）、（）。-->均数；几何均数；中位数

8、频率是对（）而言，概率是对（）而言。-->样本；总体

9、评价指标的选取是否合适，直接影响到（）的结论。指标是不是选取的越多越好呢？太多了，事实上是（），会相互干扰；太少了，可能所选的指标缺乏（），会产生比较片面的结果。-->综合评价重复性的指标；足够的代表性

10、设计统计表的横纵标目时，基本要求是（），（）即横标目在表中作主语，纵标目作谓语，连贯在一起阅读，可以组成一句完整而通顺的话。-->符合逻辑；主谓分明

11、死亡率可间接反映（）的发病水平。-->病死率高且病程短的疾病

12、统计分析主要包括（）和（）。-->统计描述；统计推断

13、统计分析主要包括：（）、（）、（）。-->统计描述；统计推断

14、统计工作的四个步骤（）、（）、（）、（）。-->计划与收集资料；整理资料；分析资料

15、统计资料的两个主要来源（）、（）。-->经常性资料；一次性资料

16、为了达到合理选择统计图的目的，首先要遵循的制图要求是（）。-->根据资料的性质以及分析目的，选择合适的图形

17、为了达到合理选择统计图的目的，首先要遵循的制图要求是根据资料的性质以及分析目的，（）。-->选择合适的图形

18、卫生服务调查的分析指标包括（）、（）和（）。-->卫生服务需要指标；卫生服务利用指标；卫生服务资源指标

19、卫生统计资料一般分为三大类，主要有（）、（）和（）。-->计量资料；计数资料；等级资料

20、一般应用层次分析法时，注意在计算（）后，应当检查所求出的（）是否符合逻辑。通常用（）检验该目标层子目标的相对重要程度有无（）。-->归一化权重系数；权重系数；一致性指标；逻辑混乱

21、医学科学研究中的误差通常指测量值与真值之差，其中包括系统误差和（），以及样本指标与总体指标之差，即（）。-->随机测量误差；抽样误差

22、医学人口统计资料通常分为（）和（）两类。-->静态人口统计；动态人口统计

23、增长速度与发展速度的关系为（）。-->增长速度=发展速度-1

24、正态分布由位置参数即（）和变异度参数即（）所决定。--> μ 或（总体均数；总体标准差）。

25、制定调查表时应注意调查指标要精选，调查项目要具体、明确。调查表的结构顺序有时间顺序、（）和类别顺序，设计者应当灵活应用。-->内容顺序

名词解释(62)-

1、X2 分布:

2、I 型错误 (type I error) -->拒绝了实际上成立的 H_0 。这类“弃真”的错误称为 I 型错误。

3、II 型错误 (type II error) -->接受了实际上是不成立的 H_0 ，这类“存伪”的错误称为 II 型错误。

4、变量-->观察单位（或个体）的某种属性或标志称为变量。

5、变量值-->对变量进行测量或观察的值称为变量值（或测量值、观察值）。

6、变异-->同一性质的事物，其个体观察值（变量值）之间的差异，在统计学上称为变异。

7、变异系数-->用于观察指标单位不同或均数相差较大时两组资料变异程度的比较。用 CV 表示。

8、标准差-->是方差的正平方根，使用的量纲与原量纲相同，适用于近似正态分布的资料，大样本、小样本均可，最为常用。

9、参数-->是总体指标。

10、参数估计-->是用样本指标（称为统计量）估计总体指标（称为参数）。

11、参数检验-->样本来自的总体分布类型是已知的，在这种假设的基础上，对总体参数进行的检验，称为参数检验。

12、抽样误差-->在同一总体中随机抽取的若干样本，样本指标之间的差异以及样本指标与总体指标的差异，称为抽样误差。

13、抽样误差-->在抽样研究中，即使消除了系统误差，控制了随机测量误差，样本统计指标和总体参数间仍会存在差别。这是由于个体变异造成的，是抽样机遇所致，是客观存在，不可避免的。

14、抽样研究-->抽样研究：从所研究的总体中随机抽取一部分有代表性的样本进行研究，抽样研究的目的是通过用样本资料计算的指标去推论总体。

15、等级相关 (rank correlation) -->又称为秩相关，是用等级数据作相关分析，属于非参数统计方法的一种。

16、等级相关腔坪习-->等级相关是对等级数据作相关分析，它又称为秩相关，是一种非参数统计方法。

17、等级资料-->将观察单位按测量结果的某种属性的不同程度分组，所得各组的观察单位等级资料，等级资料又称有序资料。

18、动态数列 (dynamic series) -->是一系列按时间顺序排列起来的统计指标，包括绝对数、相对数或平均数，用以说明事物在时间上的变化和发展趋势。

19、方差分析-->多个样本均数的比较用方差分析 (analysis of variance, 简称为 ANOVA)，其目的是检验多个总体均数是否相等。

20、非参数检验-->非参数检验是一种不依赖总体分布类型，也不对总体参数（如总体均数）进行统计推断的假设检验。

21、分类变量资料-->将全体观察单位按照某种性质或类别进行分组，然后分别清点各组中的例数，这样得到的数据称为分类变量资料，也称计数资料。

22、概率-->概率：是指某随机事件发生的可能性大小的数值，常用符号 P 表示。

23、构成比-->构成比又称构成指标，它表示事物内部各组成部分所占比重或分布。

24、计量资料-->对每个观察单位用定量的方法测定某项指标量的大小，所得的资料称为计量资料。

25、假设检验-->是先对总体做出某种假定（检验假设），然后根据样本信息推断其是否成立的一类统计方法。

26、检验水准-->也称为显著性水准 (significance level)，符号为 α 。它是预先规定的概率值，它是“是否拒绝 H_0 的界限”。

27、检验效能-->又称检验功效 (1- β)，它的含义是：当两总体确实有差别时，按规定的检验水准 α 能够发现两总体间差别的能力。

28、可比性-->可比性：指除对处理因素外，其它可能影响结果的非处理因素在各组间应该尽可能相同或相近，即“齐同”。

29、离散趋势-->指从中间部份向两侧频数分布逐渐减少的趋势。

30、普通线图-->普通线图是用线段的升降来表示某事物随时间推移的发展变化情况，或某现象另一现象变迁的情况。

31、区间估计-->按一定的概率估计总体参数所在的可能范围的方法称为区间估计。

32、散点图-->散点图是用点的密集程度和趋势表示两种现象间的相关关系。

33、时点患病率-->指在某时点检查时，接受检查的人群中现患病例所占的比例，分子为新老病人数，又称现患率或流行率。

34、数值变量资料-->是指用度量衡的方法测定每个观察单位的某项研究指标量的大小，所得到的数据（即测量值），就称为数值变量资料。

35、四格表资料-->由四个基本数据构成是统计表，其它数据均可由这四个基本数据推算出来，这种形式的资料称为四格表资料。

36、随机测量误差-->这种误差是偶然机遇所致，故无方向性，对同一样品多次测定，结果有高低，不完全一致。

37、随机抽样-->所谓随机抽样，就是总体中每个个体都有均等机会被抽取，抽到谁具有一定的偶然性。

38、同质-->研究对象具有相同的背景、条件、属性称同质。

39、统计量-->是指样本指标。

40、统计描述-->统计描述：用统计图表或计算统计指标的方法表达一个特定群体（这个群体可以是总体也可以是样本）的某种现象或特征。

41、统计图-->统计图：是以各种几何图形（如点、线、面或立体）显示数据的大小、升降、分布以及关系等，它也是对资料进行统计描述时的一种常用手段。

42、统计推断-->根据样本资料的特性对总体的特性作估计或推论的方法称统计推断，常用方法是参数估计和假设检验。

43、卫生服务

卫生服务需要是人们因疾病影响健康，引起人体正常活动的障碍，实际应当接受各种卫生服务的需要。

44、卫生服务调查-->是对卫生服务状况、人群健康的危险因素、人群卫生服务的需求和利用、卫生服务资源的分配和利用所进行的一种社会调查。

45、卫生服务调查统计-->卫生服务调查统计：它是卫生统计的重要内容之一，卫生服务调查统计是从卫生服务资料的设计、收集、整理、分析的角度，来阐述卫生服务研究的特点、研究方法和注意事项，以便使卫生服务研究工作更具有科学性。

46、卫生服务需要

卫生服务需要是人们因疾病影响健康，引起人体正常活动的障碍，实际应当接受各种卫生服务的需要。

47、误差-->指测量值与真值之差。

48、系统误差-->系统误差：这种误差不是偶然机遇造成，而是某种必然因素所致，具有一定的倾向性。

49、小概率事件-->一般常将 $P \leq$

50、样本-->从总体中随机抽取有代表性的一部分个体，其测量值（或观察值）的集合称为样本。

51、优序法-->优序法：为对比较某几个事物或方案的优劣，在选定各项评价指标后，将待评的对象或方案就各项评价指标的测量值大小分别排序，并分别对各序号（等级）以相应的评分值即优序数，然后综合诸评价指标，分别计算评价对象的总赋优序数，并按总赋优序数大小评定其优劣顺序的方法。

52、有序分类变量资料-->指将全体观察单位按照某种性质的不同程度分为若干组，分别清点各组中观察单位的个数所得的资料，这种资料也称为等级资料。

53、正态分布-->正态分布是应用最广泛的连续性概率分布，其特征是倒置的“钟”形曲线。正态分布是 Z 检验的理论基础，t 检验的理论基础 t 分布等（f、x² 分布）都是在正态分布的基础上推导出来的。

54、正相关-->正相关：它是说明具有直线关系的两个变量间，存在有正的相关方向，即当对曾大时，对有相应增大的趋势，所算得的相关系数 r 为正值。

55、直线回归-->又称简单线性回归，用线性方程定量刻画两变量间的数量关系，是回归分析中最基本、最简单的一种。

56、直线相关分析-->是首先研究两变量间有无直线相关关系，如果有相关关系，则进一步回答相关的方向以及相关密切程度的统计方法。

57、直线相关系数（correlation coefficient）-->是说明具有直线关系的两个变量间，相关关系的密切程度与相关方向的统计指标。

58、中位数-->将一组观察值由小到大排列，n 为奇数时取位次居中的变量值；为偶数时，取位次居中的两个变量的平均值。

59、综合评价 i 度到-->综合评价指的是人们根据不同的评价目的，指标，并通过一定的数学模型，息。选择相应的评价形式，据此选择多个因素或将多个评价因素或指标转化为能反映评价对象总体特征的信息。

60、综合评价 I-->是指人们根据不同的评价目的，选择相应的评价形式，据此选择多个因素或指标，并通过一定的数学模型，将多个评价因素或指标转化为能反映评价对象总体特征的信息。

61、总体-->根据研究目的确定的同质观察单位的观察值全体所构成的集合。

62、总体研究-->对有限总体中的每个个体都作观察就称总体研究。

简答(15)-

1、比较直线相关分析与直线回归分析的研究目的...

2、非参数检验和参数检验的区别？各有何优缺点？...

3、绘制统计表的基本要求是什么？...

4、假设检验的结论为什么不能绝对化？...

5、简答四格表资料 x² 检验的适用条件。...

6、率的标准化法的基本思想是什么？直接标准化法...

7、请总结直线相关系数 r 与直线回归系数 b 的意义...

8、什么是根本死因？为什么死因统计要选择根本死...

9、是否说明两总体均数之间差别很大，为什么？...

10、四格表资料检验的适用条件？

11、完全随机设计的两样本均数比较的 t 检验与方...

12、为什么计算得到的相关系数 r 和回归系数 b 都要...

13、行 & 错列表资料一般包括哪些资料，它们的检验...

14、应用相对数时有哪些注意事项？...

15、应用相对数时有哪些注意事项？...

1、比较直线相关分析与直线回归分析的研究目的有何不同？

答案：直线相关分析的目的在于：一是首先研究两变量间有无直线相关关系，即 x 变化时，y 是否有相应改变；二是如果有相关，则进一步回答相关的方向，即当 x 增大，y 的变化趋势是增加还是减小，以及两者的相关密切程度有多大的问题。直线回归分析是定量地描述并分析两变量间的线性依存关系的一种统计方法。说明两变量间依存变化的数量关系用回归，说明变量间的相互关系用相关。

2、非参数检验和参数检验的区别？各有何优缺点？

答案：区别在于：参数检验要求样本来自正态分布总体，而非参数检验则不对总体分布有任何要求；参数检验是对总体参数进行的检验，而非参数检验考察的是总体的分布情况。参数检验的优点是能充分利用所提供的信息，检验效率较高。缺点是对样本所对应的总体分布有比较严格的要求，因此适用资料有限。非参数检验的优点是不受总体分布类型的限制，适用于任何分布的资料，其缺点是不直接对原始数据作检验，从而有可能会损失信息并降低其检验效率。

3、绘制统计表的基本要求是什么？...

答案：一是标题不可缺失，要求用一句简明扼要的话来说明表的内容，并写在表的上方。二是横纵标目的设计要符合逻辑，主谓分明，标目的文字应简明，有单位的要注明。三是线条同样要简洁，除基本线条外，应尽量减少不必要的线条。四是数字应准确无误，要求一律采用阿拉伯数字表示，同一指标的数值的位数应一致，位次对齐，表中不宜留空项。五是备注一般不列入表内，而写在表的下方。

4、假设检验的结论为什么不能绝对化？

答案：假设检验的结论不能绝对化，因为无论拒绝 H₀ 或不拒绝 H₀，假设检验的结论都有犯错误的可能。假设检验的结论通常可能发生两类的错误 I 型错误：拒绝实际上成立的 H₀。犯 I 型错误的概率是 α ， α 通常为 0.05。II 型错误：接受实际上不成立的 H₀。犯 II 型错误的概率是 β 。一般情况下 β 的大小是未知的。假设

检验的结论是建立在小概率反证法的基础上的，因此结论具有一定概率的性质，不是绝对的肯定或否定。

5、简答四格表资料 x² 检验的适用条件。

答案：四格表资料 x² 检验的适应条件：一是当 n>40，且所有 T>50 时，用 x² 检验的基本公式和四格表专用公式。二是当 n>40，但有 1<T<5 时，需用四格表 x² 检验的校正公式。三是若 n<40，或者 T<1，不能计算 x² 值，需要确切概率算法。

6、率的标准化法的基本思想是什么？直接标准化法的条件是什么？

答案：当不同的人群的总率进行比较时，若其人群的内部构成（如年龄、性别等）存在差异，而这些内部构成对率有影响，为消除构成不同的影响，要按照统一标准构成对两个人群进行校正。这种选择统一构成，然后计算标准化率的方法称为率的标准法。直接标准化法需要两个条件：一是资料条件，C 知实际人群的年龄别（组）率，且各年龄率无明显交叉，。二是选择标准，可选择标准人群的年龄组人口数或构成比。

7、请总结直线相关系数 r 与直线回归系数 b 的意义及特点？...

答案：直线相关系数 r 是说明具有直线关系的两个变量间，相关关系的密切程度与相关方向的统计指标。总体相关系数用 ρ 表示，样本相关系数用 r 表示，r 是 ρ 的估计值。相关系数没有单位，取值范围是 -1<r<1。r 值为正，表示两变量呈正相关，x 与 y 变化趋势是正向的。r 值为负，表示两变量呈负相关，x 与 y 呈反向变化，通常 r 的绝对值越大，表示两变量相关关系越密切。直线回归系数 b 即回归直线的斜率，b>0 表示直线从左下方向右上方，y 随 x 增大而增大；b<0 表示直线从左上方走向右下方，y 随 x 增大而减小；b=0 则直线与 x 轴平行，x 与 y 无直线关系。b 的统计学意义是 x 每增加（减）一个单位，y 平均改变 b 个单位。

8、什么是根本死因？为什么死因统计要选择根本死因？

答案：WHO 规定，根本死因是指：“（a）引起直接导致死亡的一系列病态事件的那些疾病或损伤，或者（b）造成致命损伤的事故或暴力的情况。”。制定根本死因的想法是从预防死亡的角度出发，去寻找带有根本性的、引起一系列疾病并最终导致死亡的那个原因，不管它发生在死前多长时间都应予以记录。根本死因可以是一个明确的疾病诊断，可以是一个元明确诊断的医学情况（如：症状、体征、临床表现），也可以是一个意外的损伤和中毒。

9、是否说明两总体均数之间差别很大，为什么？

答案：说明差别有统计学意义，可以认为两个总体均数不同；但是假设检验的结论不能直接回答差异的大小，“差别有统计学意义”并不意味着两个总体均数相差很大。差别的大小及差别有无实际意义只能进一步根据专业知识来确定。例如：当样本量足够大时，即使两个样本均数间相差很小也可能得出 P<0.05。此时差别虽具有统计学意义，但不一定有实际意义。

10、四格表资料检验的适用条件？

答案：四格表资料检验的适用条件：一是当 n>40，且所有 T>50 时，检验的基本公式和四格表专用公式。二是当 n>40，15、有 1<T<5 时，需用四格表检验的校正公式。三是若 n<40 法，或 T<1 时，不能计算值，需用确切概率算法。

11、完全随机设计的两样本均数比较的 t 检验与方差分析之间的关系如何？。

答案：两个样本均数比较可以看作是多个样本均数比较的特例，因此完全随机设计的两个样本均数比较的 t 检验，可以用完全随机设计的方差分析代替。两者的计算结果有如下关系：F~t。反之，则不成立，即多个样本均数比较的方法应该用方差分析，而不能用两个样本均数比较的 t 检验代替，否则会增大犯 1 型错误的概率。

12、为什么计算得到的相关系数 r 和回归系数 b 都要做显著性检验？

答案：由于 r 与 b 都是根据样本资料求得的，而在由变量 x 与 y 组成的总体中作随机抽样，获取样本的过程中，不可避免地会存在抽样误差。因此当算得的 r 与 b 不为零时，有可能实际情况是 x 与 y 之间是不相关或不存在直系回归关系的，只是由于抽样误差，r 与 b 才会不为零。作显著性检验的目的正是对这种情况作出判断，看 r 或 b 是否确实由总体相关系数 ρ 或总体回归系数 β 自不为零的总体中抽得。

13、行 & 错列表资料一般包括哪些资料，它们的检验目的有何不同？1

答案：行错列表资料一般包括多个样本率、多个构成比资料，其基本数据可整理成 R 行 C 列，称为 R 错 C 表，又称行错列表。多个样本率或构成比的错 2 检验的目的是推断其总体率或构成比是否不同。对同一样本资料按其两个无序分类变量（行变量和列变量）归纳成双向交叉排列的统计表，其行变量可分为 R 类，列变量可分为 C 类，这种表称为 R 错 C 列联表。列联表资料检验的目的是推断两变量（行变量、列变量）之间分布是否相互独立，用列联表的独立性检验。尽管这两种行列表检验目的和检验假设方面有所不同，但计算值和自由度的公式完全相同。

14、应用相对数时有哪些注意事项？

答案：一是构成比和率是含义不同的两个统计指标，应用时不能混淆构成比说明事务内部各组成部分所占的比重而率则说明某事务或现象的发生频率或强度。不能以构成比代替率来说明问题。（2 分）二是样本含量太小时，不宜计算相对数，最好用绝对数来表示。（2 分）三是对各组观察例数不等的几个率，不能直接相加求其总率，四是在比较相对数时应注意资料的可比性，。

15、应用相对数时有哪些注意事项？。

答案：一是构成比和率是意义不同的两个统计指标，应用时不能混淆。构成比说明事务内部各组成部分所占的比重，而率则说明某事务或现象的发生频率或强度，不能以构成比代替率来说明问题。二是样本含量太小时，不宜计算相对数，最好用绝对数来表示。三是对各组观察例数不等的几个率，不能直接相加求其总率。四是在比较相对数时应注意资料的可比性。

应用分析题(18)-

- 1、表中的资料是某市城乡女性原发性骨质疏松症...
- 2、测得 10 名铝作业工人和 10 名非铝作业工人的血...
- 3、某工厂 1995 年年底和 1996 年年底对该厂工人进...
- 4、某省从三个水中氟含量不同的地区随机抽取 20...
- 5、某研究单位观察 3 种药物驱虫的疗效，在服药 7 天...
- 6、某研究者分别用离子交换法与蒸馏法测定 8 名...
- 7、某医生为比较中药和西药治疗气管炎的疗效，随...
- 8、某医生为比较中药和西药治疗气管炎的疗效，随...
- 9、某医院比较几种疗法对慢性胃炎病人的疗效：单...
- 10、某医院对医院的医生和护士的年龄分布进行了...
- 11、某医院护理部抽测 10 名护士用水冲手 30 分钟前...
- 12、某医院内科治疗一般类型胃溃疡 80 例，治愈 60...
- 13、随机抽样调查 100 名上海市区男孩出生体重，均...
- 14、随机抽样调查 100 名上海市区男孩出生体重，均...
- 15、完成以下表格中的患病率的计算，表某市城乡女...
- 16、为比较两种方法对牛乳中脂肪含量测定结果是...
- 17、为了解某中药治疗原发性高血压的疗效，将 44 名...
- 18、用两种方法做实验，物理方法用了 30 只小白鼠，10...

1、表中的资料是某市城乡女性原发性骨质疏松症患病率情况，请进行比较和分析城乡总的骨质疏松症患病率情况。

表 1 某市城乡女性原发性骨质疏松症患病率情况

年龄组(岁)	城市			农村		
	调查人数	患病人数	患病率(%)	调查人数	患病人数	患病率(%)
50~	300	70		250	56	
60~	200	96		300	128	
70~	150	90		200	115	
80岁及以上	40	30		60	43	
合计						

参考答案：完成以下表格中的患病率的计算

表 1 某市城乡女性原发性骨质疏松症患病率情况

年龄组(岁)	城市			农村		
	调查人数	患病人数	患病率(%)	调查人数	患病人数	患病率(%)
50~	300	70	23.3	250	56	22.4
60~	200	96	48.0	300	128	42.7
70~	150	90	60.0	200	115	57.5
80岁及以上	40	30	75.0	60	43	71.7
合计	690	286	41.4	810	342	42.2

以城市的人口作为标准人口，则农村的总患病率为：

标化率 =
$$\frac{\sum \text{标准人群各年龄组人口数} \times \text{被标化人群各年龄组率}}{\sum \text{标准人群各年龄组人口数}}$$

=
$$(300 \times 22.4\% + 200 \times 42.7\% + 150 \times 57.5\% + 40 \times 71.7\%) / 690 \times 100\%$$

= 38.8%

农村的标化患病率为 38.8%，低于城市的总患病率 41.4%，因此消除城市和农村的人口构成的不同后，可以看出城市的女性骨质疏松症的患病率和分年龄段的该病患病率均高于农村的总的女性骨质疏松症的患病率和分年龄段的该病患病率。

2、测得 10 名铝作业工人和 10 名非铝作业工人的血铝（ $\mu\text{mol/L}$ ）

测得 10 名铝作业工人和 10 名非铝作业工人的血铝（ $\mu\text{mol/L}$ ）

，问：

- （1）这是什么资料？
- （2）适合用什么统计方法进行统计推断？
- （3）如果本组资料不符合正态分布，请分析两组工人的血铝含量是否不同？

表 2 两组工人的血铝含量比较

铝作业工人(1)	非铝作业工人(2)
0.96	0.34
0.97	0.22
1.35	0.25
0.89	0.54
>1.87	0.29
1.56	0.56
1.34	0.45
1.22	0.76
1.38	0.87
1.45	1.10
$n_1=10$	$n_2=10$

①建立检验假设。
H₀: 两组工人的血铝含量总体分布相同
H₁: 两组工人的血铝含量总体分布不同 a=0.05
②编秩（过程见上表）
③求秩和（过程见上表）（并确定统计量 T: 本例 n₁=n₂=10 较小，可选取检验统计量 T=T₂=58
3、某工厂 1995 年年底和 1996 年年底对该厂工人进行普查,人数均为 2000 人
某工厂 1995 年年底和 1996 年年底对该厂工人进行普查,人数均为 2000 人,1995 年查出高血压患者 197 人;1996 年查出高血压患者 208 人,208 人中有 181 人在 1995 年系高血压。
请计算（1）1995 年年底和 1996 年年底的高血压患病率。
（2）1996 年的高血压发病率。
答: 1995 年年底的高血压患病率=197/2000×100%=9. 85%
1996 年年底的高血压患病率=208/2000×100%=10. 4%
1996 年的高血压发病率=（208-181）/2000×100%=1. 35%
4、某省从三个水中氟含量不同的地区随机抽取 20 岁成人,进行患龋率的调查
某省从三个水中氟含量不同的地区随机抽取 20 岁成人,进行患龋率的调查,数据见表 1。研究者采用行列表的 X² 检验,X²=10. 66,P<0. 01。差异有统计学意义,故认为三个地区成人的患龋率都不同。

表 1 某省三个地区成人患龋率比较

地区	患龋人数	未患人数	合计	患龋率(%)
高氟区	50	20	70	71.4
中氟区	40	30	70	57.1
低氟区	50	10	60	83.3
合计	140	60	200	70.0

（1）该资料是什么资料？
答：该资料是 3 组二分类资料。
（2）该研究是什么设计？

答：研究设计为完全随机设计。
（3）统计分析的目的是什么？统计方法是否正确？
答：统计分析的目的是比较三个地区成人患龋率有无差别,统计方法是正确的。但是研究者的卡方检验结果,只能认为各总体率之间总的来说不同,但不能说明它们彼此之间都不同,或某两者之间有差别。
5、某研究单位观察 3 种药物驱虫的疗效，在服药 7 天后查粪中虫卵的阴转率如表，问三种药物的疗效是否相同？
某研究单位观察 3 种药物驱虫的疗效，在服药 7 天后查粪中虫卵的阴转率如表，问三种药物的疗效是否相同？
3 种药物驱虫的疗效比较

药物	服药人数	阴转人数	阴转率%
甲药	37	28	75.7
乙药	38	18	47.4
丙药	34	10	29.4

答：如下（1）建立检验假设，确定检验水准。
H₀: 三种药物阴转率相同,即 $\pi_1=\pi_2=\pi_3$ 。（1 分）
H₁: 三种药物阴转率不相同或不全相同。（1 分）
 $\alpha=0.05$ (1 分)
（2）计算统计量 χ^2 值。（1 分）
本例为 3×2 表资料,按下面的公式计算 χ^2 值。（1 分）
 $\chi^2=n(\sum\frac{A^2}{n_{gk}n_c}-1)=15.56$ (2 分)
（3）确定 P 值,作出推断结论。（1 分）
 $v=(3-1)(2-1)=2$ (1 分)
查附表 6(χ^2 界值表),本例 $x^2=15.56$, (2 分) $\chi^2=15.56>5.99$, (1 分)故 $P<0.05$,
按 $\alpha=0.05$ 水准,拒绝 H₀,接受 H₁,可认为三种药物阴转率不相同或不全相同

6、某研究者分别用离子交换法与蒸馏法测定 8 名健康人的尿汞，结果见表，问这两种测定方法检测结果有无差别？
某研究者分别用离子交换法与蒸馏法测定 8 名健康人的尿汞，结果见表，问这两种测定方法检测结果有无差别？

8 名健康人用离子交换法与蒸馏法测定尿汞值(ug/L)								
编号	1	2	3	4	5	6	7	8
离子交换法	0.4	2.1	0.1	1.9	0.9	4.5	1.6	2.2
蒸馏法	0.1	1.3	0.1	1.1	2.3	3.4	1.1	1.6

答：

两种方法测定尿汞值(μg/L)的结果比较				
编号	离子交换法	蒸馏法	差值	秩次
(1)	(2)	(3)	(4)=(2)-(3)	(5)
1	0.4	0.1	0.3	1
2	2.1	1.3	0.8	4
3	0.1	0.1	0.0	—
4	1.9	1.1	0.8	5
5	0.9	2.3	-1.4	-7
6	4.5	3.4	1.1	6
7	1.6	1.1	0.5	2
8	2.2	1.6	0.6	3

$T_+=21, \quad T_-=7$

检验步骤如下:

(1) 建立检验假设。

H_0 : 两法所得结果无差别。

H_1 : 两法所得结果有差别。

$\alpha=0.05$ 。

(2) 求差值:

如表所示,将第(2)和(3)栏相减得到第(4)栏差值。

(3) 编秩:

将表2第(4)栏的差值按绝对值从小到大编秩,如表第(5)栏所示,编好之后在秩次之前仍保持原差值的正负号。编秩时遇差值为零则舍去不予编秩,同时在总秩次中要相应减去零的个数。本例有8个差值,其中1个差值为零,故秩次为由1至7。

(4) 求秩和并确定检验统计量 T 。

分别求出正负秩和,正秩和是将所有正的秩次相加得到,以 T_+ 表示,负秩和是对所有负秩次的绝对值求和,以 T_- 表示。本例的 $T_+=21, T_-=7$ 。习惯上取秩和绝对值最小者为检验统计量 (T)。

本例取 $T=T_-=7$ 。

(5) 确定 P 值和作出统计结论。

当 $n \leq 50$ 时,可查附表 7, T 界值表来确定 P 值。本例有一个差值为零,未参与编秩。故例数 $n=7$, 双侧 $\alpha=0.05$ 检验水平时检验统计量 T 的可能取值范围(界值范围)是 5-40。由于 $T=T_-=7$, 在界值范围内, 故 $P>0.05$, 按 $\alpha=0.05$ 水平不拒绝 H_0 , 不能认为两法测得的尿汞值有差别。

7、某医生为比较中药和西药治疗气管炎的疗效, 随机抽取 140 例气管炎患者, 随机分成中药组和西药组, 结果西药组治疗 70 例, 有效 56 例, 中药组治疗 79 例, 有效 42 例。请比较两种方法有无差别?

某医生为比较中药和西药治疗气管炎的疗效, 随机抽取 140 例气管炎患者, 随机分成中药组和西药组, 结果西药组治疗 70 例, 有效 56 例, 中药组治疗 79 例, 有效 42 例。请比较两种方法有无差别?

中西药治疗气管炎的有效率不同。(2 分)

8、某医生为比较中药和西药治疗气管炎的疗效, 随机抽取 149 例气管炎患者, 随机分成中药组和西药组, 结果西药组治疗 70 例, 有效 56 例, 中药组治疗 79 例, 有效 42 例。问中西药治疗气管炎的疗效是否不同?

26. (1)建立检验假设,确定检验水准(2 分)

H_0 : 中西药治疗气管炎的有效率相同, 即 $\pi_1 = \pi_2 \backslash$

H_1 : 中西药治疗气管炎的有效率不同, 即 $\pi_1 \neq \pi_2 \backslash$

$\alpha=0.05$ (2 分)

(2)选定检验方法,计算检验统计量(2 分)

$$\chi^2 = \frac{(ad - bc)^2 \cdot n}{(a + b)(c + d)(a + c)(b + d)} = 11.87,$$

332

(3)确定 P 值,作出推断结论、

本例 $|\chi^2| = 6.667 > 3.84$, 所以 $P < 0.05$ 。(2 分) | 按 $\alpha = 0.05$ 水准, 拒绝 H_0 , 接受 H_1 , 可认为中西药治疗气管炎的有效率不同(

9、某医院比较几种疗法对慢性胃炎病人的疗效: 单纯西药组治疗 79 例, 有效 63 例;

单纯中药组治疗 54 例, 有效 47 例; 中西医结合组治疗 68 例, 有效 65 例。问: (1) 该资料属何种资料? 实验设计属何种设计? (2) 欲比较 3 种疗法的疗效的差别, 宜选用何种假设检验方法?

(3) 写出该种检验方法的 H 。与 H_1 ; (4) 若求得的检验统计量为 8.143, 相应于 $\alpha=0.05$ 的检验统计量的临界值为 5.99, 你如何做出结论? (5) 根据你的结论, 你可能犯哪一类统计错误?

解: (1) 该资料属计数资料, 实验设计属完全随机设计;

(2) 宜选用 $R \times C$ 表的 χ^2 检验方法;

(3) H_0 : 3 种疗法的总体有效率相等; H_1 : 3 种疗法的总体有效率不全相等;

(4) 在 $\alpha=0.05$ 的水准上, 拒绝 H_0 , 接受 H_1 , 差别有统计学意义, 可认为 3 种疗法的总体有效率不全相等, 即 3 种疗法对慢性胃炎病人的疗效有差别;

(5) 可能犯 I 型错误。

10、某医院对医院的医生和护士的年龄分布进行了调查, 整理得到下表, 请根据统计表制表原则和注意事项指出问题所在并予以修改。

某医院对医院的医生和护士的年龄分布进行了调查, 整理得到下表, 请根据统计表制表原则和注意事项指出问题所在并予以修改。

某医院医生和护士年龄分布表		
	医生	护士
人次	200	150
30 岁以下	5.0	26.7
31~40	20.0	33.3
41~50	50.0	33.3
51~60	25.0	6.7

答案如下

表中的问题主要有

(1) 标题不确切。

(2) 主谓语设置不当。

(3) 表中数字不明。
(4) 线条过多。
建议修改的表如下表 (12 分)。

某医院医生和护士年龄构成				
年龄组	医生		护士	
	人数	构成比 (%)	人数	构成比 (%)
<30	10	5.0	40	26.7
30~	40	20.0	50	33.3
40~	100	50.0	50	33.3
≥51	50	25.0	10	6.7
合计	200	100.0	150	100.0

11、某医院护理部抽测 10 名护士用水冲手 30 分钟前后手上细菌情况 (个/cm²) 结果, 见下表, 试分析用水冲洗这种措施的效果。(提示自由度为 9 时, $\alpha=0.05$ 和 $\alpha=0.01$ 的双侧检验 t 界值分别为 2.26 和 3.25)
表 10 名护士用水清洗 30 分钟前后手上细菌情况(个/cm²)

编号	洗前	洗后
1	58	18
2	55	20
3	48	30
4	30	20
5	35	15
6	48	25
7	40	11
8	55	23
9	35	24
10	36	23

答: (1) 建立检验假设, 确定检验水准。(2 分)
 H_0 : 用水冲手 30 分钟前后手上细菌情况相同, $\mu_d=0$ (1 分)
 H_0 : 用水冲手 30 分钟前后手上细菌情况不相同, 即 $\mu_d \neq 0$ (1 分)
(2) 计算统计量 t 值。(1 分)
 $n=10$

$$t = \frac{\bar{d}-0}{s/\sqrt{n}} = 6.94 \text{ (2 分)}$$

(3) 确定 P 值, 作出推断结论。(2 分)
 $v=n-1=9$ (1 分)
查 t 界值表, 本例 $t=6.94>3.25$, 故 $P<0.01$, (2 分)
可以认为用水冲手 30 分钟前后手上细菌情况不同。(2 分)

12、某医院内科治疗一般类型胃溃疡 80 例, 治愈 60 例; 治疗特殊型胃溃疡患者 90 例, 治愈 30 例, 该医院内科对两种类型的胃溃疡病人治愈率是否不同?

附: χ^2 界值表, 双侧检验, 自由度为 1, $P=0.05$ 和 0.01 的 χ^2 界值分别是: 3.84 和 6.63。

解: (1) 建立检验假设, 确定检验水准。

H_0 : 该医院内科对两种类型的胃溃疡病人治愈率相同, 即 $\pi_1=\pi_2$
 H_1 : 该医院内科对两种类型的胃溃疡病人治愈率不相同, 即 $\pi_1 \neq \pi_2$, $\alpha=0.05$

(2) 计算统计量 χ^2 值。

本题中全部格子的理论频数大于 5, 同时 $n=170>40$, 故可用四格表专用公式计算 χ^2 值。
 $\chi^2=(60 \times 60 - 20 \times 30)^2 \times 170 / (90 \times 80 \times 80 \times 90) = 29.51$ (1 分)

(3) 确定 P 值, 作出推断结论。

$$v=(2-1)(2-1)=1$$

查 χ^2 界值表, 本例 $\chi^2=29.51>6.63$, 故 $P<0.01$, 可以认为该医院内科对两种类型的胃溃疡病人治愈率不相同。

13、随机抽样调查 100 名上海市区男孩出生体重, 均数为 3.29kg, 标准差为 0.44kg。郊区抽查男童 100 人的出生体重, 得均数 3.23%, 标准差 0.47kg, 问市区和郊区男孩出生体重均数是否不同?

答案: (1) 建立检验假设, 确定检验水准。(2 分)

$H_0: \mu_1=\mu_2$, 市区和郊区男孩出生体重均数相同。(1 分)

$H_1: \mu_1 \neq \mu_2$, 市区和郊区男孩出生体重均数不同。(1 分)

本例为双侧检验, 检验水准 $\alpha=0.05$ (1 分)

(2) 计算检验统计量。(2 分)

两样本含量较大, 故采用完全随机设计的两样本 Z 检验。

$$\chi^2 = \frac{(ad-bc)^2 \cdot n}{(a+b)(c+d)(a+c)(b+d)} = 6.6667, \text{ (2 分)}$$

(3) 确定 P 值, 作出推断结论。(2 分)

因为双侧 $Z_{0.05}=1.96$, 本例 $Z=1.45<1.96$, 所以 $P>0.05$ 。(2 分) 按 $\alpha=0.05$ 水准, 不拒绝 H_0 , 故认为市区和郊区男孩出生体重均数相同。(2 分)

14、随机抽样调查 100 名上海市区男孩出生体重, 均数为 3.29kg, 标准差为 0.44kg。郊区抽查男童 100 人的出生体重, 得均数 3.23 (kg), 标准差 0.47 (kg), 问市区和郊区男孩出生体重均数是否不同?

答案: 如下

(1) 建立检验假设, 确定检验水准。

$H_0: \mu_1 = \mu_2$, 市区和郊区男孩出生体重均数相同。

$H_1: \mu_1 \neq \mu_2$, 市区和郊区男孩出生体重均数不同。

本例为双侧检验, 检验水准 $\alpha=0.05$ 。

(2) 计算检验统计量。

两样本含量较大, 故采用完全随机设计的两样本 Z 检验。

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{3.29 - 3.23}{\sqrt{\frac{0.44^2}{100} + \frac{0.47^2}{100}}} = 1.45$$

(3) 确定 P 值, 作出推断结论。

因为双侧 $Z_{0.05} = 1.96$, 本例 $Z=1.45<1.96$, 所以 $P>0.05$ 。按 $\alpha=0.05$ 水准, 不拒绝 H_0 , 可认为市区和郊区男孩出生体重均数相同。

15、完成以下表格中的患病率的计算, 表某市城乡女性原发性骨质疏松症患病率情况 (以城市人口为标准人口)

完成以下表格中的患病率的计算（10 分）
解：

表 1 某市城乡女性原发性骨质疏松症患病率情况

年龄组(岁)	城市			农村		
	调查人数	患病人数	患病率(%)	调查人数	患病人数	患病率(%)
50~	300	70	23.3	250	56	22.4
60~	200	96	48.0	300	128	42.7
70~	150	90	60.0	200	115	57.5
80岁及以上	40	30	75.0	60	43	71.7
合计	690	268	41.4	810	342	42.2

以城市的人口作为标准人口，则农村的总的患病率为：

标准化 =
$$\frac{\sum \text{标准人群各年龄组人口数} \times \text{被标化人群各年龄组率}}{\sum \text{标准人群各年龄组人口数}}$$

= (300x22.4%+200x42.7%+150x57.5%+40x71.7%) /690x100%=38.8%
农村的标化患病率为 38.8%，低于城市的总的患病率 41.4%，因此消除城市和农村的人口构成的不同后，可以看出城市的女性骨质疏松症的患病率和分年龄段的该病患病率均高于农村的总的女性骨质疏松症的患病率和分年龄段的该病患病率。
16、为比较两种方法对牛乳中脂肪含量测定结果是否不同，某人随机抽取了 8 份牛乳制品，分别用甲、乙两种方法测定结果如表，请问两种方法测定结果是否不同？
两种方法测定牛乳脂肪含量结果

编号	甲法(μmol/L)	乙法(μmol/L)
1	0.642	0.573
2	0.613	0.582
3	0.753	0.618
4	0.542	0.473
5	0.713	0.682
6	0.653	0.578
7	1.082	0.834
8	0.786	0.695

(2)计算检验统计量 t 值(2 分)

本例为同一受试对象接受了两种方法的处理（n 较小），选用配对设计 t 检验，目的是检验两种方法是否不同。

已知 $n=8$ ，算得 $\sum d=0.749$ ， $\sum d^2=0.1051$ ，
 $\bar{d}=\sum d/n=0.749/8=0.0936(\mu\text{mol/L})$

$$S_d=\sqrt{\frac{\sum d^2-(\sum d)^2/n}{n-1}}$$

$$=\sqrt{\frac{0.1051-(0.749)^2/8}{8-1}}=0.0707(\mu\text{mol/L})$$

$$t=\frac{\bar{d}-0}{S_d/\sqrt{n}}=\frac{0.0936}{0.0707/\sqrt{8}}=3.74,(2\text{分})\quad v=n-1=8-1=7$$

(3) 确定 P 值，作出推断结论

查 t 界值表(附表 2)，当 $v=7$ 时，双侧 $t_{0.05/2,7}=2.365$ ，

本例 $t=3.74>2.365$ ，所以 $P<0.05$ 。

按 $\alpha=0.05$ 水准，拒绝 H_0 ，认为两种方法测得结果不同。

17、为了解某中药治疗原发性高血压的疗效，将 44 名高血压患者随机分为两组。
实验组用该药加辅助治疗，对照组用安慰剂加辅助治疗，观察结果如下表，问：该药治疗原发性高血压是否有效？
两种疗法治疗原发性高血压的疗效

分组	例数	有效	有效率/%
实验组	23	21	91.30
对照组	21	5	23.81

实验组

答:本题为完全随机设计两样本率的比较。整理表格见下表。

两种疗法治疗原发性高血压的疗效				
分 组	有效	无效	合计	有效率/%
实验组	21	2	23	91.30
对照组	5	16	21	23.81
合计	26	18	44	59.09

建立检验假设,确定检验水准

$H_0:\pi_1=\pi_2$,即该药治疗原发性高血压无效

$H_1:\pi_1\neq\pi_2$,即该药治疗原发性高血压有效

$\alpha=0.05(2\text{分})(2)$ 计算检验统计量

$T_{min}=21 \times 1844=8.59>5$
 $\chi^2=(ad-bc)^2/n$
 $(a+b)(c+d)(a+c)(b+d)=(21 \times 16-2 \times 5)^2 \times 4423 \times 21 \times 26 \times 18=20.69(2 \text{ 分})$
 $v=(2-1)(2-1)=1(2 \text{ 分})$ 确定 P 值,作出统计推断
查 χ^2 界值表得 $P<0.005$,按 $\alpha=0.05$ 水准拒绝 H_0 ,接受 H_1 ,差别有统计学意义,可以认为该药治疗原发性高血压有效。
18、用两种方法做实验,物理方法用了 30 只小白鼠,10 天内死亡 17 只;用化学方法处理 31 只小白鼠,同期内死亡 9 只,问两种方法对小白鼠的致死作用是否不同?
注: χ^2 值在自由度=1, $\alpha=0.05$ 和 $\alpha=0.01$ 分别为 3.84 和 6.63
解:该资料可以整理为如下表格的形式
两种方法处理小白鼠的死亡率的比较

方法	致死数	存活数	合计	致死率%
物理	17	13	30	56.27
化学	9	22	31	29.0
合计	26	35	61	42.6

(1)建立检验假设,确定检验水准(3 分)

H_0 :两种方法对小白鼠的致死率相同,即 $\pi_1=\pi_2(1 \text{ 分})$

H_1 :两种方法对小白鼠的致死率不相同,即 $\pi_1 \neq \pi_2(1 \text{ 分})$

$\alpha=0.05(1 \text{ 分})$

(2)计算统计量 χ^2 值。(2 分)

本题中全部格子的理论频数大于 5,同时 $n=61>40$,故可用四格表专用公式计算 χ^2 值。

$$\chi^2=(17 \times 22-13 \times 9)^2 \times 61 / (30 \times 31 \times 26 \times 35)=4.76$$

(3)确定 P 值,作出推断结论

$$v=(2-1)(2-1)=1$$

查 χ^2 界值表,本例 $\chi^2=4.76>3.84$,故 $P<0.05$,
可以认为两种方法对小白鼠的致死率不同。