

名词

1. 总体: 总体是根据研究目的确定的所有同质观察单位的全体, 它包括所有定义范围内的个体变量值。
2. 样本: 是指从总体中选取的有代表性的一部分观察单位或个体, 通常使用随机选取方法得到。
3. 参数: 总体是根据研究目的确定的所有同质观察单位的全体, 它包括所有定义范围内的个体变量值, 那么描述总体特征的指标称为参数。
4. 统计量: 从研究总体中抽取部分有代表性的观察单位, 对变量进行观测构成一个样本, 描述样本特征的指标称为统计量
5. 医学参考值范围: 指“正常人”的解剖、生理、生化指标等数据大多数个体值的波动范围。
6. 置信区间: 区间估计是指按预先给定的概率, 计算出一个区间, 使他能够包含未知的总体参数, 事先给定的概率 $1-\alpha$ 称为可信度 (通常取 0.95 或 0.99), 计算得到的区间称为可信区间或置信区间, 其通常由下限与上限两个可信限构成。
7. 小概率事件: 习惯上将 $P \leq 0.05$ 的事件称为小概率事件, 表示在一次随机抽样中发生的可能性很小。
8. 回归方程: 是对变量之间统计关系进行定量描述的一种数学表达式, 指具有相关的随机变量和固定变量之间关系的方程。
9. 相关系数: 又称 Pearson 积差相关系数, 是说明具有直线相关关系的两个数值变量之间相关的方向和密切程度的统计量。
10. 定量数据: 也称计量资料, 变量的观测结果是数值型的, 用来说明研究对象的数量特征, 其特点是能够用数值大小来衡量观察单位不同特征水平的高低, 一般有计量单位。分为连续型定量数据与离散型定量数据。
11. 定性数据: 也称计数资料, 变量的观测值是定性的, 说明的是研究对象的品质特征, 表现为互不相容的类别或属性。
12. 假设检验: 亦称显著性检验, 是统计推断的另一重要内容, 其目的是定性比较总体参数之间有无差别或总体分布是否相同。

简答题

1、非配对四格表卡方检验的应用条件?
 $R \times C$ 列联表资料卡方检验应用条件---定性数据(1)专用公式: ①当各格子理论频数不小于 1, 并且 $1 \leq T < 5$ 的格子数不宜超过格子总数的 $1/5$ 时使用专用公式②若不满足①, 增加样本含量使理论频数增大仍可使用专用公式③若不满足①, 根据专业知识删去合并理论频数太小的行或列, 仍可使用专用公式。
(2)不满足①时直接用 $R \times C$ 表的 Fisher 确切概率法, 四格表资料的卡方检验应用条件---定性数据, 两个样本率比较①当 $n \geq 40$ 且所有的 $T \geq 5$ 时, 用 χ^2 检验的基本公式或四格表资料 χ^2 检验的专用公式。
②当 $n \geq 40$ 且 $5 > T \geq 1$ 时, 用四格表资料 χ^2 检验的校正公式。
③当 $T < 1$ 或 $n < 40$, 用四格表资料的 Fisher 确切概率法。

2、配对四格表卡方检验的应用条件?
配对四格表卡方检验---配对设计的定性数据, 样本含量不是很大(1) $\chi^2 = (b-c)^2 / b+c$, 自由度 = 1 条件: $b+c \geq 40$
(2) $\chi^2 = (|b-c|-1)^2 / b+c$, 自由度 = 1 条件: $b+c < 40$

3、数据的种类有哪些?
定量数据: 也称计量资料, 变量的观测值是定量的, 其特点是能够用数值大小衡量其水平的高低, 一般有计量单位, 根据变量的取值特征可分为连续型数据和离散型数据。
定性数据: 也称计数资料, 变量的观测值是定性的, 表现为互不相容的类别或属性。例如, 血型分为 A、B、O、AB 等。
有序数据: 也称半定量数据或等级资料, 变量的观测值是定性的, 但各类别 (属性) 之间有程度或顺序上的差别, 如尿糖的化验结果分为 -、+、++、+++。

4、样本均数 t 检验有哪些常用方法?
(1)单样本 t 检验又称单样本均数 t 检验, 适用于来自正态分布的某个样本均数 \bar{x} 与已知总体均数 μ 的比较, 其比较目的是检验样本均数 \bar{x} 所代表的总体均数 μ 是否与已知总体均数 μ 有差别。
(2)配对样本均数 t 检验简称配对 t 检验, 又称非独立两样本均数 t 检验, 适用于配对设计计量资料均数的比较, 理论上假设配对差值服从正态分布, 其比较目的是检验两相关样本均数所代表的未知总体均数是否有差别。
(3)两独立样本 t 检验, 又称成组 t 检验, 适用于完全随机设计下两样本均数的比较, 其目的是检验两样本所来自总体的均数是否相等。完全随机设计是将受试对象随机地分配到两组中, 每组对象分别接受不同的处理, 分析比较两组的处理效应。

5. 假设检验时需要设立哪些检验, 都叫什么名字?
t 检验, z 检验, F 检验, 和卡方检验等。
6. 定性数据的统计描述方法都有哪些?
根据不同研究目的常用率, 构成比, 相对比等指标进行描述。
(1)率
(2)构成比
(3)相对比①两类别例数之比②相对危险度③比值比
(4)标准化率

7. 统计表的结构都有哪些?
从外形上看, 统计表可由标题、标目 (包括横标目、纵标目)、线条、数字和备注 5 部分构成。

8. 医学统计学的基本内容包括哪些?
医学统计学的基本内容包括统计设计、数据整理与核查、统计描述和统计推断。
①统计设计能够提高研究效率, 并使结果更加准确和可靠;
②数据整理与核查主要是对数据进行归类、检查数据质量, 以及是否符合特定的统计分析方法要求等;
③统计描述用来描述及总结数据的重要特征;
④统计推断指由样本数据的特征推断总体特征的方法, 包括参数估计和假设检验。

9. 常用的定量数据集中趋势描述方法有哪些?
(1)统计指标----均数
①算数均数--主要适用于对称分布或偏斜度不大的资料, 尤其适合正态分布资料。
②几何均数--观察值间接倍数关系变化或数值的对数成正态分布③中位数--变量值中出现个别特小或特大的数值, 资料的分布呈明显的偏态; 变量值分布一端或两端无确定数值, 只有小于或大于某个数值, 资料的分布不清④百分位数--任何频数分布资料
(2)统计表比如频数表
(3)统计图比如直方图, 线图, 箱式图, 误差条图等

10. 常用的定量数据离散趋势描述方法有哪些?
(1)统计指标
①极差--任何分布类型的资料②四分位数间距--适用于任何分布类型的资料③方差--主要是正态分布④标准差--正态分布⑤变异系数--用于对均数相差较大或单位不同的两组或多组观察值的变异程度进行比较
(2)统计表比如频数表
(3)统计图比如直方图, 线图, 箱式图, 误差条图等

11. 简述假设检验计算的基本过程。
①建立假设和确定检验水准 α
②选择检验方法与计算检验统计量
③根据 p 值做出统计推断

12. 统计推断有哪几类基本的方式?
包括参数估计和假设检验参数估计--用样本统计量估计总体参数包括点估计与区间估计假设检验--t 检验, z 检验, F 检验, 和卡方检验等。

13. 简述最小二乘法的基本原理。
该方法是保证各实测点到回归直线的纵向距离的平方和最小, 即使 $Q = \sum (Y - \hat{Y})^2$ 最小, 从而使计算出的回归直线最能代表实测数据所反映出的直线趋势。

14. 简述相关系数数值的含义。
相关系数是说明具有直线相关关系的两个数值变量之间相关的方向和密切程度的统计量, 相关系数 r 没有度量衡单位, 其取值范围是 $-1 \leq r \leq 1$, $r > 0$ 表示正相关, $r < 0$ 表示负相关, $r = 0$ 表示无线性相关, 当 $|r| = 1$ 时为完全相关, 相关系数的绝对值越接近 1, 相关越密切, 相关系数越接近 0, 相关越不密切。