

## 实习二 数值变量资料统计描述

### 一、最佳选择题

- 血清滴度资料最常计算\_\_\_\_\_以表示其平均水平。  
A. 算术均数      B. 中位数      C. 几何均数      D. 全距
- 原始数据减去同一个不等于 0 的常数后\_\_\_\_\_。  
A.  $\bar{x}$  变,  $s$  变      B.  $\bar{x}$  变,  $s$  不变      C.  $\bar{x}$ 、 $s$  都不变      D.  $\bar{x}$ 、 $s$  都变
- 开展一项科学研究, 何时开始运用统计学知识\_\_\_\_\_。  
A. 有了实验数据之后      B. 有了实验设计之后  
C. 从最初实验设计开始      D. 从计算机算出结果开始
- 利用频数分布表及如下公式  $M=L+i/f_x * (n/2 - \sum f_i)$  计算中位数时\_\_\_\_\_。  
A. 要求数据对称分布      B. 要求数据偏态分布      C. 不要求组距相等      D. 要求组距相等
- 如一组观察值的标准差为 0, 则\_\_\_\_\_。  
A. 样本例数为 0      B. 抽样误差为 0      C. 平均数为 0      D. 以上都不对
- 以下指标中\_\_\_\_\_可用来描述数值变量资料离散程度。  
A. 算术均数      B. 几何均数      C. 中位数      D. 极差
- 偏态分布资料宜用\_\_\_\_\_描述其分布的集中趋势。  
A. 算术均数      B. 标准差      C. 中位数      D. 四分位数间距
- 用均数和标准差可全面描述\_\_\_\_\_资料的分布特征。  
A. 正态分布      B. 正偏态分布      C. 负偏态分布      D. 任何分布
- \_\_\_\_\_可用于比较身高与体重的变异度。  
A. 标准差      B. 变异系数      C. 全距      D. 四分位数间距
- 正态曲线下横轴上, 从均数  $\mu$  到  $+\infty$  的面积为\_\_\_\_\_。  
A. 97.5%      B. 95%      C. 50%      D. 5%
- 标准正态分布的均数与标准差分别为\_\_\_\_\_。  
A. 0 和 1      B. 1 和 0      C. 0 和 0      D. 1.96 和 2.58
- 若正常人的血铅含量  $X$  近似对数正态分布, 则可用公式制定 95% 正常值范围\_\_\_\_\_。  
(其中:  $Y = \log X$ )  
A.  $\bar{x} + 1.96S$       B.  $\bar{x} + 1.64S$       C.  $\log^{-1}(Y + 1.96S_Y)$       D.  $\log^{-1}(Y + 1.64S_Y)$
- 正态分布有两个参数  $\mu$  与  $\sigma$ , \_\_\_\_\_曲线的形状越扁平。  
A.  $\mu$  越大      B.  $\mu$  越小      C.  $\sigma$  越大      D.  $\sigma$  越小
- 正态分布曲线下, 横轴上从均数  $\mu$  到  $\mu + 1.96\sigma$  的面积为\_\_\_\_\_。  
A. 95%      B. 45%      C. 97.5%      D. 47.5%
- 关于标准差, 错误的是\_\_\_\_\_。  
A. 反映全部观察值的离散程度      B. 度量了一组数据偏离平均数的大小  
C. 反映了均数代表性的好坏      D. 不会大于算术均数
- 一组正态分布曲线, 经过  $u = (x - \mu) / \sigma$  变换后, 变成\_\_\_\_\_。  
A. 一组曲线,  $\mu = 0$ ,  $\sigma$  不确定      B. 一组曲线,  $\mu$ ,  $\sigma$  都不确定  
C. 一条曲线,  $\mu = 0$ ,  $\sigma = 1$       D. 一条曲线,  $\mu = 1$ ,  $\sigma = 0$
- 现有 5 人的血清滴度为 <1:2, 1:4, 1:8, 1:16, 1:32, 描述平均滴度以\_\_\_\_\_指标较好。  
A. 算术均数      B. 几何均数      C. 中位数      D. 百分位数

18. 原始数据分布不明时, 表示集中趋势的指标应是\_\_\_\_\_
- A. 均数      B. 中位数      C. 几何均数      D. 均数和中位数
19. 各观察值同乘以一个不等于 0 的常数后, \_\_\_\_\_ 不变。
- A. 均数      B. 标准差      C. 中位数      D. 变异系数
20. 若 X 服从以  $(\mu, \sigma)$  为中心的正态分布, 则 X 的第 95 百分位数即\_\_\_\_\_
- A.  $\mu - 1.645\sigma$       B.  $\mu - 1.96\sigma$       C.  $\mu + 1.645\sigma$       D.  $\mu + 1.96\sigma$

## 二、填空题

1. 医学统计学的用途\_\_\_\_\_。
2. 医学统计工作基本步骤\_\_\_\_\_。
3. 统计分析包括\_\_\_\_\_、\_\_\_\_\_两个部分。
4. 频数分布的两个重要特征是\_\_\_\_\_、\_\_\_\_\_。
5. 正态分布两个参数是\_\_\_\_\_、\_\_\_\_\_。
6. 统计资料的来源有\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
7. 编制频数表的步骤为\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。
8. 小概率事件是指\_\_\_\_\_。
9. 频数表的用途\_\_\_\_\_。
10. 常见的统计资料类型有\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_。

## 三、简答题

1. 数值变量资料频数表的组段数目是否越多越好? 组距和组段数目的关系是什么?
2. 描述数值变量资料集中趋势的指标有哪些? 其适用范围有何异同?
3. 描述数值变量资料离散趋势的指标有哪些? 其适用范围有何异同?
4. 中位数和百分位数在意义、计算和应用上有何区别与联系?
5. 同一资料的标准差是否一定小于均数?
6. 正态分布、标准正态分布在概念和应用上有何异同?
7. 医学正常值范围的涵义是什么? 确定的原则和方法是什么?
8. 标准差和标准误有何区别与联系?
9. 举例说明总体与样本的关系。
10. 预防医学教研室测得大气中SO<sub>2</sub>的浓度, 用两种计量单位表示: mg/m<sup>3</sup>, μg/m<sup>3</sup>, 分别计算几何均数G及标准差S<sub>1gx</sub>, 结果发现两种不同计量单位的S<sub>1gx</sub>相等, 试解释其原因。
11. 下表为儿童健康检查登记表的一部分, 试说出下列各指标哪些属数值变量? 哪些属无序分类变量? 哪些属有序分类变量? (姓名、性别、年龄、既往病史、身高、坐高、收缩压、舒张压、血型、肝大指数)

## 四、计算分析题

1. 随机测得 15 名健康大学生的脉搏数 (次/分) 为 75, 70, 74, 78, 65, 66, 77, 79, 71, 70, 73, 74, 68, 67, 69。试计算其平均数、标准差、方差、极差和变异系数。
2. 某地 101 例 30~50 岁健康男子血清总胆固醇值 (mmol/L) 测定结果如下:  
 4.77   3.37   6.14   3.95   3.56   4.23   4.31   4.71   5.69   4.12   4.56   4.37   5.39  
 6.39   5.21   **7.22**   5.54   3.93   5.21   6.51   5.18   5.77   4.79   5.12   5.20   5.10  
 4.70   4.74   3.50   4.69   4.38   4.89   6.25   5.32   4.50   4.63   3.61   4.44   4.43  
 4.25   4.03   5.85   4.09   3.35   4.08   4.79   5.30   4.97   3.18   3.97   5.16   5.10  
 5.86   4.79   5.34   4.24   4.32   4.77   6.36   6.38   4.88   5.55   3.04   4.55   3.35

4.87 4.17 5.85 5.16 5.09 4.52 4.38 4.31 4.58 5.72 6.55 4.76 4.61  
4.17 4.03 4.47 3.40 3.91 **2.70** 4.60 4.09 5.96 5.48 4.40 4.55 5.38  
3.89 4.60 4.47 3.64 4.34 5.18 6.14 3.24 4.90 3.05

- (1) 编制频数分布表并绘制直方图，简述其分布特征。
- (2) 计算均数、标准差、变异系数。
- (3) 计算该地 30~50 岁健康男子血清总胆固醇值 95% 的正常值范围。
- (4) 现测得一名 40 岁男子总胆固醇为 6.99mmol/L，问其血清总胆固醇值是否正常？

3. 某医院检测一批肝炎患者的相关抗原 (HAA)，滴度分布如下。求平均滴度。

|      |      |       |       |       |        |        |        |    |
|------|------|-------|-------|-------|--------|--------|--------|----|
| 抗体滴度 | 1: 8 | 1: 16 | 1: 32 | 1: 64 | 1: 128 | 1: 256 | 1: 512 | 合计 |
| 例数   | 2    | 4     | 6     | 10    | 5      | 2      | 1      | 30 |

4. 某年某地某医师调查 205 例伤寒患者的潜伏期 (天)，资料如下：

|         |    |    |    |    |     |     |     |     |     |     |     |
|---------|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|
| 潜伏期 (天) | 2~ | 4~ | 6~ | 8~ | 10~ | 12~ | 14~ | 16~ | 18~ | ≥20 | 合计  |
| 人数 (f)  | 26 | 29 | 42 | 50 | 48  | 4   | 2   | 2   | 1   | 1   | 205 |

- (1) 求平均潜伏期、四分位数间距。
- (2) 求伤寒患者潜伏期 95% 的正常值范围。

5. 某医生调查了某地 20 岁男子 100 人，其身高均数为 169.03cm，标准差为 4.15cm；体重均数为 53.72kg，标准差为 4.16kg。试比较身高和体重的变异大小。