

南京财经大学

2011 年攻读硕士学位研究生入学考试（初试）试卷 A

考试科目：432 统计学 适用专业：应用统计硕士 满分 150 分

考试时间：2011 年 1 月 16 日下午 2:00—5:00

注意事项：所有答案必须写在答题纸上，做在试卷或草稿纸上无效；

请认真阅读答题纸上的注意事项，试题随答卷一起装入试题袋中交回。

注：答题过程中可能要用到的数据： $u_{0.975}=1.96, u_{0.95}=1.65, t_{0.95}(9)=1.833, t_{0.95}(10)=1.812,$

$t_{0.975}(9)=2.262, t_{0.975}(10)=2.228$

一、单项选择题（本题包括 1-30 题共 30 个小题，每小题 1 分，共 30 分。在每小题给出的四个选项中，只有一个符合题目要求，把所选项前的字母填在答题纸相应的序号内）。

1. 从含有 N 个元素的总体中，抽取 n 个元素作为样本，使得总体中的每一个元素都有相同的机会（概率）被抽中，这样的抽样方式称为（ ）。

- A. 简单随机抽样 B. 分层抽样 C. 系统抽样 D. 整群抽样

2. 一家公司的人力资源部主管需要研究公司雇员的饮食习惯，改善公司餐厅的现状。他将问卷发给就餐者，填上后再收上来。他的收集数据的方法属于（ ）。

- A. 自填式问卷调查 B. 面访式问卷调查 C. 实验调查 D. 观察式调查

3. 将全部变量值依次划分为若干个区间，并将这一区间的变量值作为一组，这样的分组方法称为（ ）。

- A. 单变量值分组 B. 组距分组 C. 等距分组 D. 连续分组

4. 将某企业职工的月收入依次分为 2000 元以下、2000 元-3000 元、3000 元-4000 元、4000 元-5000 元、5000 元以上几个组。第一组的组中值近似为（ ）。

- A. 2000 B. 1000 C. 1500 D. 2500

5. 经验法则表明，当一组数据对称分布时，在平均数加减 1 个标准差的范围内大约有（ ）。

- A. 68% 的数据 B. 95% 的数据 C. 99% 的数据 D. 89% 的数据

6. 在比较两组数据的离散程度时，不能直接比较它们的标准差，因为两组数据的（ ）。

- A. 标准差不同 B. 方差不同 C. 数据个数不同 D. 计量单位不同

7. 在一家宾馆门口等待出租车的时间是左偏的，均值为 12 分钟，标准差为 3 分钟。如果从宾馆门口随机抽取 100 名顾客并记录他们等待出租车的时间，则该样本均值的分布服从（ ）。

- A. 正态分布，均值为 12 分钟，标准差为 0.3 分钟
B. 正态分布，均值为 12 分钟，标准差为 3 分钟
C. 左偏分布，均值为 12 分钟，标准差为 3 分钟
D. 左偏分布，均值为 12 分钟，标准差为 0.3 分钟

8. 一个 95% 的置信区间是指（ ）。

- A. 总体参数有 95% 的概率落在这一区间内
B. 总体参数有 5% 的概率未落在这一区间内

- C.在用同样方法构造的总体参数的多个区间中,有 95%的区间包含该总体参数
D.在用同样方法构造的总体参数的多个区间中,有 95%的区间不包含该总体参数
- 9.从一个正态总体中随机抽取 $n=20$ 的一个样本,样本均值为 17.25,样本标准差为 3.3。则总体均值的 95%的置信区间为 ()。
- A. (15.97, 18.53) B. (15.71, 18.79) C. (15.14, 19.36) D. (14.89, 20.45)
- 10.一项研究表明,司机驾车时因接打手机而发生事故的的比例超过 20%,用来检验这一结论的原假设和备择假设应为 ()。
- A. $H_0: \pi = 20\%, H_1: \pi \neq 20\%$ B. $H_0: \pi \neq 20\%, H_1: \pi = 20\%$
C. $H_0: \pi \geq 20\%, H_1: \pi < 20\%$ D. $H_0: \pi \leq 20\%, H_1: \pi > 20\%$
- 11.一项调查表明,5 年前每个家庭每天看电视的平均时间为 6.7 小时。而最近对 200 个家庭的调查结果是:每个家庭每天看电视的平均时间为 7.25 小时,标准差为 2.5 小时。在 $\alpha = 0.05$ 的显著性水平下,检验假设 $H_0: \mu \leq 6.7, H_1: \mu > 6.7$,得到的结论为 ()。
- A.拒绝 H_0 B.不拒绝 H_0
C.可以拒绝也可以不拒绝 H_0 D.可能拒绝也可能不拒绝 H_0
- 12.如果变量之间的关系近似地表现为一条直线,则称两个变量之间为 ()。
- A.正线性相关关系 B.负线性相关关系 C.线性相关关系 D.非线性相关关系
- 13.在一元线性回归方程中,回归系数 β_1 的实际意义是 ()。
- A.当 $x=0$ 时, y 的期望值
B.当 x 变动一个单位时, y 的平均变动数量
C.当 x 变动一个单位时, y 增加的总数量
D.当 y 变动一个单位时, x 的平均变动数量
- 14.残差平方和 SSE 反映了 y 的总变差中 ()。
- A.由于 x 和 y 之间的线性关系引起的 y 的变化部分
B.除了 x 对 y 的线性影响之外的其他因素对 y 变差的影响
C.由于 x 和 y 之间的非线性关系引起的 y 的变化部分
D.由于 x 和 y 之间的函数关系引起的 y 的变化部分
- 15.在多元线性回归分析中,如果 t 检验表明回归系数 β_i 不显著,则意味着 ()。
- A.整个回归方程的线性关系不显著
B.整个回归方程的线性关系显著
C.自变量 x_i 与因变量之间的线性关系不显著
D.自变量 x_i 与因变量之间的线性关系显著
- 16.一家出租车公司为确定合理的管理费用,需要研究出租车司机每天的收入(元)与他的行驶时间(小时)、行驶的里程(公里)之间的关系,为此随机调查了 20 位出租车司机,根据每天的收入(y)、行驶时间(x_1)和行驶的里程(x_2)的有关数据进行回归,得到下

面的有关结果 ($\alpha = 0.05$):

方程的截距 $\hat{\beta}_0 = 42.38$ 截距的标准差 $S_{\hat{\beta}_0} = 36.59$ 回归平方和 $SSR = 29882$

回归系数 $\hat{\beta}_1 = 9.16$ 回归系数的标准差 $S_{\hat{\beta}_1} = 4.78$ 残差平方和 $SSE = 5205$

回归系数 $\hat{\beta}_2 = 0.46$ 回归系数的标准差 $S_{\hat{\beta}_2} = 0.14$

根据以上结果计算的判定系数为 ()。

A. 0.9229 B. 1.1483 C. 0.3852 D. 0.8516

17. 多重相关系数 R^2 的平方根度量了 ()。

A. k 个自变量之间的相关程度 B. 因变量同 k 个自变量之间的相关程度
C. 因变量之间的相关程度 D. 因变量同某个自变量之间的相关程度

18. 时间序列在长期内呈现出来的某种持续向上或持续下降的变动称为 ()。

A. 趋势 B. 季节性 C. 周期性 D. 随机性

19. 某地区农民家庭的年平均收入 2004 年为 1500 元, 2005 年增长了 8%, 那么 2005 年比 2004 年相比, 每增长 1 个百分点增加的收入额为 ()。

A. 7 元 B. 8 元 C. 15 元 D. 40 元

20. 拉氏指数方法是指在编制价格综合指数时 ()。

A. 用基期的销售量加权 B. 用报告期的销售量加权
C. 用固定某一时期的销售量加权 D. 选择有代表性时期的销售量加权

21. 指出下列指数公式中哪个是帕氏价格指数公式 ()。

A. $\frac{\sum p_1 q_1}{\sum p_0 q_0}$ B. $\frac{\sum p_1 q_1}{\sum p_0 q_1}$ C. $\frac{\sum p_1 q_0}{\sum p_0 q_0}$ D. $\frac{\sum p_1 q_1}{\sum p_1 q_0}$

22. 某地区 2009 年的零售价格指数为 108%, 这说明 ()。

A. 商品销售量增长了 8% B. 商品销售价格平均增长了 8%
C. 由于价格变动使销售量增长了 8% D. 由于销售量变动使价格增长了 8%

23. 对于正偏 (右偏) 分布, 平均数、中位数和众数之间的关系是 ()。

A. 平均数 > 中位数 > 众数 B. 中位数 > 平均数 > 众数
C. 众数 > 中位数 > 平均数 D. 众数 > 平均数 > 中位数

24. 指出下面的说法哪一个是正确的 ()。

A. 置信水平越大, 估计的可靠性越大 B. 置信水平越大, 估计的可靠性越小
C. 置信水平越小, 估计的可靠性越大 D. 置信水平的大小与估计的可靠性无关

25. 在多元回归分析中, 多重共线性是指模型中 ()。

A. 两个或两个以上的自变量彼此相关 B. 两个或两个以上的自变量彼此无关
C. 因变量与一个自变量相关 D. 因变量与两个或两个以上的自变量相关

26. 设 $0 < P(A) < 1$, $0 < P(B) < 1$, $P(A|B) + P(\bar{A}|\bar{B}) = 1$, 则 ()。

A. 事件 A 和 B 互不相容 B. 事件 A 和 B 对立
C. 事件 A 和 B 不独立 D. 事件 A 和 B 相互独立

27. 设 $X \sim N(\mu, 4^2)$, $Y \sim N(\mu, 5^2)$, 记 $P(X \leq \mu - 4) = p_1$, $P(Y \geq \mu + 5) = p_2$, 则

()。

A. 对任意实数 μ 有 $p_1 = p_2$

B. $p_1 < p_2$

C. $p_1 > p_2$

D. 只对 μ 的个别值才有 $p_1 = p_2$

28. 设两个相互独立的随机变量 X 和 Y , 分别服从正态分布 $N(0,1)$ 和 $N(1,1)$, 则 ()。

A. $P(X+Y \leq 0) = 0.5$

B. $P(X+Y \leq 1) = 0.5$

C. $P(X-Y \leq 0) = 0.5$

D. $P(X-Y \leq 1) = 0.5$

29. 已知 $X \sim B(n, p)$, $EX = 2.4$, $DX = 1.44$, 则二项分布的参数为 ()。

A. $n=4, p=0.6$

B. $n=6, p=0.4$

C. $n=8, p=0.3$

D. $n=24, p=0.1$

30. 一家计算机软件开发公司的人事部门最近做了一项调查, 发现在最近两年内离职的公司员工中有 40% 是因为对工资不满意, 有 30% 是因为对工作不满意, 有 15% 是因为他们对工资和工作都不满意。则两年内离职的员工中, 离职原因是因为对工资不满意、或者对工作不满意、或者两者皆有的概率为 ()。

A. 0.40

B. 0.30

C. 0.15

D. 0.55

二、简要回答下列问题 (本题包括 1-6 题共 6 个小题, 每小题 5 分, 共 30 分)。

1. 举例说明总体、样本、参数、统计量这几个概念及他们之间的区别和联系。

2. 简述众数、中位数和平均数的特点和应用场合。

3. 简述影响抽样误差大小的因素有哪些。

4. 简述假设检验的基本思想。

5. 简述大数定理在统计研究中的方法论意义。

6. 正态分布的概率密度函数 $f(X)$ 有两个参数 μ 和 σ , 请结合函数 $f(X)$ 的几何形状说明 μ 和 σ 这两个参数的意义。

三、计算与分析题 (本题包括 1-6 题共 6 个小题, 每小题 15 分, 共 90 分)

1. 某公司所属三个企业生产同种产品, 2009 年实际产量、计划完成情况及产品优质品率资料如下:

企 业	实际产量 (万件)	完成计划 (%)	实际优质品率 (%)
甲	100	120	95
乙	150	110	96
丙	250	80	98

要求:

(1) 计算该公司产量计划完成百分比;

(2) 计算该公司实际的优质品率。

2. 现从某公司职工中随机抽取 60 人调查其工资收入情况, 得到有关资料在下表, 假定职工的月收入服从正态分布。

月收入	800	900	950	1000	1050	1100	1200	1500
工人数	6	7	9	10	9	8	7	4

(1) 以 95% 的置信度估计该公司工人的月平均工资所在范围;

(2) 以 95% 的置信度估计月收入在 1000 元及以上工人所占比重。

3. 为研究产品销售额和销售利润之间的关系, 某公司对所属 7 家企业进行调查, 设产品销售额为 X (万元), 销售利润为 Y (万元)。对调查资料进行整理和计算, 其结果如下:

$$\sum x = 795 \quad \sum x^2 = 72925 \quad \sum y = 1065 \quad \sum y^2 = 121475 \quad \sum xy = 93200$$

要求: (1) 计算销售额与销售利润之间的相关系数:

(2) 配合销售利润对销售额的直线回归方程 $y_c = a + bx$ 。

(3) 解释回归系数 b 的经济意义。

4. 某企业声明有 30% 以上的消费者对其产品质量满意。如果随机调查 600 名消费者, 表示对该企业产品满意的有 220 人。试在显著性水平 0.05 下, 检验调查结果是否支持企业的自我声明。

5. 三个地区同一种商品的价格和销售量资料如下:

地区	商品价格/元		销售量/万件	
	基期	报告期	基期	报告期
代表符号	P_0	P_1	q_0	q_1
甲	9	9	20	40
乙	10	9	30	40
丙	12	12	30	20
合计				

要求: (列表并写出计算公式及计算过程)

(1) 计算三个地区总的平均价格指数。

(2) 用相对数和绝对数分析说明三个地区总的平均价格变动中, 各地区价格变动和销售量结构的变动的影响。

6. 有一大批产品, 其验收方案如下, 先作第一次检验: 从中取 10 件, 经检验无次品接受这批产品, 次品数大于 2 拒收; 否则作第二次检验, 其做法是从中再任取 5 件, 仅当 5 件中无次品时接受这批产品。若产品的次品率为 10%, 求:

(1) 这批产品经第一次检验就能接受的概率。

(2) 需作第二次检验的概率。

(3) 这批产品按第二次检验的标准被接受的概率。

(4) 这批产品在第一次检验未能做决定且第二次检验时被通过的概率。

(5) 这批产品被接受的概率。