

二〇二〇~二〇二一学年 第一学期 《应用统计学》 考试试题

考试日期: 2021年1月15日

试卷类型: A

试卷代号: 090059

班号

学号

姓名

题号	一	二	三	四	五	六	七	八	九	十	总分
得分											

本题分数	10
得分	

一、选择题 (每题2分)

1. 设 A, B 为两个随机事件, 已知 $P(A) = 0.4, P(B) = 0.3, P(A|B) = 0.4$, 则有 ()

- A. A, B 相互独立 B. A, B 互斥 C. $B \subset A$ D. $P(A \cup B) = P(A) + P(B)$

2. 若函数 $F(x) = 2/(2+x^2)$ 为随机变量 X 的分布函数, 则 X 可能的取值范围为 ()

- A. $(-\infty, \infty)$ B. $(0, \infty)$ C. $(-\infty, 0)$ D. $(0, 1)$

3. 在假设检验中, 关于 P 值, 以下说法正确的是 ()

- A. P 值越小, 拒绝备择假设的证据越充分 B. P 值是能够拒绝原假设的最小的显著性水平
C. 显著性水平越大, P 值减小 D. 以上都不对

4. 关于最小二乘法得到的样本回归直线 $\hat{y}_i = \beta_0 + \beta_1 X_i$, 以下说法错误的是 ()。

- A. 残差求和一定为 0 B. 样本均值 (\bar{X}, \bar{Y}) 不一定在样本回归线上
C. 回归系数估计量是最佳线性无偏估计量 D. \hat{y}_i 的均值与 Y_i 的均值一定相等

5. 关于季节变动的测定, 下列说法错误的是 ()

- A. 目的在于掌握事物变动的季节周期性 B. 常用的方法是按月 (季) 平均法
C. 一般需要计算季节指数 D. 按月计算的季节指数之和应为 400%

本题分数	10
得分	

二、填空题 (每题2分)

1. 新邻居家有两个小孩, 已知至少有一个是男孩, 那么两个都是男孩的概率为 (假设生男生女概率相同) _____。

2. 考虑一个抛均匀硬币试验, 记 X 为出现首次正面时已抛掷的次数, 对任意的 $l \geq 1, P(X = l) = \underline{\hspace{2cm}}$ 。

3. 若 X_1, \dots, X_n 是来自均匀分布 $U(0, a)$ 的一组简单随机样本 ($a > 0$), 则参数 a 的极大似然估计量为 _____。

4. 设 X_1, X_2, \dots, X_n 为一组来自总体 $N(0, 1)$ 的简单随机样本, 则统计量 $(n-1)X_1^2 / \sum_{i=2}^n X_i^2$ 服从 _____。

分布。

5. 以 2000 年为基期, 国家 2002 年、2003 年的广义货币供应量的定基发展速度分别为 150%, 则 2003 年和 2002 年相比的环比发展速度为_____。

本题分数	8
得分	

三、假设一个单项选择题有四个选项, 一个班级里真正会做同学占 40%, 能够去掉一个错误答案的同学占 30%, 其余同学不会。真正会做的同学能够给出正确答案, 能够去掉一个错误答案的同学从剩余的三个选项中随机猜一个, 而完全不会的同学从四个选项中随机猜一个。

(1) 假设一个同学给出了正确答案, 那他真的会做这道题的概率是多少?

(2) 假设有 100 道这样的选择题, 每题 1 分。对于从该班级里随机抽取的一个学生, 得分服从什么分布? (明确指出参数)

本题分数	8
得分	

四、设随机变量 X 的密度函数为

$$f(x) = \begin{cases} a+x, & -1 \leq x \leq 0 \\ a-x, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$$

试求 (1) 常数 a ; (2) 概率 $P(|X| < \frac{1}{2})$ 。

本题分数	10
得分	

五、设随机变量 X, Y 的分布律分别为

X	0	1
P	1/3	2/3

Y	-1	0	1
P	1/3	1/3	1/3

且有 $P(X^2 = Y^2) = 1$ 。求 (1) (X, Y) 的联合分布; (2) $Z = XY$ 的分布。

本题分数	10
得分	

六、某生产线上产品的次品率为 p , 现随机抽取了 1000 个产品, 发现有 10 个次品。

(1) 求 p 的 95% 的置信区间 (给出计算过程并说明依据);

(2) 能否说 p 以 95% 的概率落入 (1) 中所得的置信区间? 为什么?

本题分数	12
得分	

七、用传统工艺加工的某种果汁每瓶维生素 C 的含量平均为 19 毫克。现采用一种新的加工工艺, 试图减少在加工过程中对维生素 C 的破坏, 抽查了 16 瓶果汁, 测得维生素 C 的含量 (单位: 毫克) 如

22	23	20.5	21	20	22.5	19	20	23	20.5	18.8	20	19.5	22	18	23
----	----	------	----	----	------	----	----	----	------	------	----	------	----	----	----

已知该瓶装果汁中维生素含量服从正态分布, 问新工艺下维生素的含量是否比旧工艺有所提高(显著性水平 $\alpha = 0.05$)。

题分数	S
分	

八、某电器公司 3 种电器 2018 年 1 月份和 2 月份的销售额和销售单价的相关资料如下表所示, 试计算 3 种电器的帕氏价格指数和拉氏销售量指数, 并对销售额做因素分析。

商品名称	销售额 (万元)		2018 年 2 月份销售单价与 2018 年 1 月份销售单价之比 (%)
	2018. 1	2018. 2	
A	170	240	120
B	140	114	95
C	30	33	110

本题分数	12
得分	

九、某轮胎生产厂家设计了 5 种型号的轮胎, 要检验这些轮胎在平均刹车停止距离方面是否有显著差异, 以便对各型号轮胎定价。该厂家选择了 50 辆相同的小汽车, 为这 5 种轮胎各随机选取了 10 辆车, 以相同的速度进行试驾测试, 获得各型号轮胎对应的刹车距离数据, 运用 Excel 做方差分析后得方差分析表 (显著性水平 $\alpha = 0.05$):

方差分析	SS	df	MS	F	P-value	F crit
差异源						
组间	3622.6			4.0659108	0.006739724	
组内						
总计	13646					

青 (1) 基于已知信息完善上述方差分析表 (阴影部分), 并给出该方差分析的最终结论; (2) 写出 P 值计算公式。

本题分数	12
得分	

十、某公司在调查某产品推广费用对产品销售额的影响, 市场部收集了销售额 (单位: 万元) 和推广费用 (单位: 万元) 数据, 并运用线性回归进行了相关分析 (推广费用为解释变量 (自变量), 销售额为被解释变量 (因变量)), 利用 Excel 得到的回归结果如下 ($\alpha=0.05$)

回归统计	
Multiple R	0.95157
R Square	0.905486
Adjusted R Square	0.903338
标准误差	
观测值	

方差分析

df	SS	MS	F	Significance F
----	----	----	---	----------------

回归分析	1	1.17E+08	1.17E+08	421.5399
残差	44	12192771	277108.4	—
总计	45	1.29E+08	—	—

显著性检验

	Coefficients	标准误差	t Stat	P-value
Intercept	4361.486	416.7749	10.46485	1.61E-13
X Variable	1.198017	0.05835	20.53144	3.62E-24

- (1) 完善回归统计表 (阴影部分);
- (2) 写出总体回归模型, 估计回归方程, 并解释回归系数估计值的意义;
- (3) 检验回归系数的显著性;
- (4) 评价该回归模型对数据的拟合效果;
- (5) 如果推广费用为 3500 万元, 试预测产品平均销售额。

附录: 分位数

$$z_{0.05} = 1.645 \quad z_{0.025} = 1.96 \quad z_{0.01} = 2.325,$$

$$t_{0.05}(16) = 1.7459 \quad t_{0.025}(16) = 2.1199 \quad t_{0.05}(15) = 1.7531 \quad t_{0.025}(15) = 2.1315,$$

$$F_{0.95}(45,4) = 0.3878, F_{0.975}(4,45) = 0.1191, F_{0.05}(45,4) = 5.7073.$$