

桂林电子科技大学《概率与统计》2020-2021 学年第一学期期末试卷

系别_____班级_____学号_____姓名_____

题目	一	二	三	四	五	总分
得分						
阅卷人签字						

一、选择题（每题 3 分，共 30 分）

1. 设随机变量 X 服从标准正态分布 $N(0,1)$ ，则 $P(|X| > 1.96) \approx$ _____。
A. 0.05
B. 0.025
C. 0.95
D. 0.975
2. 下列哪个不是概率分布的特性？
A. 所有可能取值的概率之和为 1
B. 每个可能取值的概率都非负
C. 概率分布函数是连续的
D. 对于离散型随机变量，每个可能取值对应一个概率
3. 设随机变量 X 和 Y 的相关系数为 0.8，则下列说法正确的是：
A. X 和 Y 完全相关
B. X 和 Y 不相关
C. X 和 Y 正相关且相关性较强
D. X 和 Y 负相关且相关性较弱
4. 在卡方检验中，自由度的计算公式是：
A. (行数 - 1) * (列数 - 1)
B. 行数 + 列数 - 1
C. 样本量 - 参数个数 - 1
D. 分类变量的类别数 - 1
5. 下列哪个不是假设检验的基本步骤？
A. 提出原假设和备择假设
B. 确定检验的统计量和拒绝域
C. 根据样本数据计算检验统计量的值
D. 根据先验概率判断原假设是否成立
6. 设随机变量 X 服从参数为 λ 的指数分布，则 $P(X > 1/\lambda) =$ _____。

- A. $1/e$
 - B. $1 - 1/e$
 - C. $e^{(-1)}$
 - D. $1 - e^{(-1)}$
7. 在多元线性回归分析中，如果某个解释变量的 t 检验不显著，则通常意味着：
- A. 该变量对被解释变量的解释作用很大
 - B. 该变量与被解释变量之间存在强烈的线性关系
 - C. 该变量对被解释变量的解释作用不大或没有解释作用
 - D. 该变量与被解释变量之间存在非线性关系
8. 设随机变量 X 服从自由度为 n 的卡方分布，则 $E(X^2) = \underline{\hspace{2cm}}$ 。
- A. n
 - B. $2n$
 - C. n^2
 - D. $2n^2$
9. 在时间序列分析中，如果数据呈现明显的季节性波动，则可以使用_____模型进行拟合和预测。
- A. ARIMA
 - B. SARIMA
 - C. MA
 - D. AR
10. 在贝叶斯网络中，节点之间的有向边表示的是：
- A. 条件概率
 - B. 联合概率
 - C. 因果关系或相关性
 - D. 边缘概率

二、填空题（每题 2 分，共 20 分）

1. 设随机变量 X 服从均匀分布 $U(a, b)$ ，则 $P(a < X < c) = \underline{\hspace{2cm}}$ （其中 $a < c < b$ ）。
2. 在二项分布 $B(n, p)$ 中，当试验次数 n 很大且每次试验成功的概率 p 很小时，可以使用_____分布来近似。
3. 设随机变量 X 和 Y 相互独立，且分别服从正态分布 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ ，则 $Z = X + Y$ 服从_____分布。
4. 在假设检验中， β 表示的是_____。
5. 设随机变量 X 服从参数为 λ 的泊松分布，则 $D(X) = \underline{\hspace{2cm}}$ 。
6. 在多元线性回归分析中，如果 R^2 的值接近 1，则意味着模型对数据的_____很强。
7. 设随机变量 X 服从自由度为 n 的 t 分布，则当 n 趋近于无穷大时， t 分布趋近于_____分布。
8. 在时间序列分析中，如果数据存在明显的趋势性，则可以使用_____模型进行拟合和预测（写出一种即可）。
9. 设随机变量 X 和 Y 的协方差为 $Cov(X, Y)$ ，则 $Corr(X, Y) = \underline{\hspace{2cm}}$ （用 $Cov(X, Y)$ 和其他相关量表示）。
10. 在贝叶斯定理中，如果事件 A 和 B 是相互独立的，则 $P(A|B) = \underline{\hspace{2cm}}$ 。

三、计算题（每题 10 分，共 20 分）

题目 1:

设随机变量 X 服从正态分布 $N(\mu, \sigma^2)$ ，已知 $P(\mu - 1.645\sigma < X < \mu + 1.645\sigma) = 0.9$ 。现从该分布中随机抽取一个样本 $X_1 = 5, X_2 = 6, \dots, X_n = 10$ ($n = 100$)，样本均值为 $\bar{X} = 7$ ，样本标准差为 $s = 2$ 。请利用这些信息估计 μ 的置信水平为 90% 的置信区间。

题目 2:

某公司为了评估其新产品的市场接受度，随机选取了 1000 名潜在消费者进行问卷调查。结果显示，有 600 名消费者表示愿意购买该产品。试在 $\alpha = 0.05$ 的显著性水平下，利用正态近似法检验该比例是否显著大于 50%。

四、证明题（10 分）

题目:

证明大数定律：设随机变量 $X_1, X_2, \dots$ 是两两不相关的，且它们具有相同的数学期望 $E(X_i) = \mu$ 和相同的方差 $D(X_i) = \sigma^2$ ($i = 1, 2, \dots$)。则对于任意的正数 $\varepsilon > 0$ ，有 $\lim_{n \rightarrow \infty} P(|1/n \cdot \sum(X_i - \mu)| < \varepsilon) = 1$ 。

五、应用题（20 分）

案例背景:

某电商平台为了优化其推荐系统，随机选取了 10000 名用户作为试验组，向他们推送了基于机器学习算法生成的个性化推荐商品列表；同时，另选 10000 名用户作为对照组，向他们推送了基于热门商品生成的推荐列表。经过一个月的观察，试验组的用户平均购买量为 5.2 件，对照组的用户平均购买量为 4.8 件。假设两组用户的购买量均服从正态分布，且方差相等。

题目:

1. 请在 $\alpha = 0.05$ 的显著性水平下，利用两独立样本均值 t 检验判断试验组的推荐系统是否显著提高了用户的购买量。
2. 如果试验组的推荐系统确实有效，请结合实际情况分析可能的原因，并提出进一步的优化建议。