2016年攻读博士学位研究生入学考试初试试题卷

考试科目：水利数理统计　　　　适用专业：水利水电工程　水工结构工程

（不用抄题，答案写在答题纸上，写明题号，答案写在试题上无效）

一. 填空题 (每题3分，共30分)

1.若 $X_1, X_2, \cdots, X_n$ 为来自正态分布 $X \sim N(\mu, \sigma^2)$ 的一个样本，S^2 为样本方差，则 $E(S^2) =$____________

2.设 $X_1, X_2, \cdots, X_n$ 是正态总体 $X \sim N(\mu, \sigma^2)$ 的样本，$\bar{X}$ 是其样本均值，其中 σ^2 已知,则 μ 的置信度为 $1-\alpha$ 的置信区间的长度为___________

3.设总体 X 服从正态分布 $N(0,\ 4)$，而 $(X_1, X_2 \cdots, X_{15})$ 是来自 X 的样本，则 $U = \dfrac{X_1^2 + \cdots + X_{10}^2}{2(X_{11}^2 + \cdots + X_{15}^2)}$ 服从的分布是_______

4.方差分析的目的是_______

5.设 (X_1, X_2, X_3, X_4) 是取自正态总体 $X \sim N(0, 2^2)$ 的简单随机样本且 $Y = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$，则 $a =$_______，$b =$________时，统计量 Y 服从 $\chi^2(2)$ 分布.

6.设随机变量 X, Y 相互独立，均服从 $N(0, 3^2)$ 分布且 $X_1, \cdots, X_9$ 与 $Y_1, \cdots, Y_9$ 分别是来自总体 X, Y 的简单随机样本，则统计量 $U = \dfrac{X_1 + \cdots + X_9}{\sqrt{Y_1^2 + \cdots + Y_9^2}}$ 服从的分布是____

7.已知正态总体 $X \sim N(\mu, \sigma^2)$，其中参数 σ^2 未知，那么假设检验 $H_0: \mu = \mu_0$ 应该选取的统计量为____________

8.若 $X_1, X_2, ..., X_n$ 为来自泊松分布 $P(\lambda)$ 的一个样本，$\overline{X}, S^2$ 分别为样本均值和样本方差，则 $E(\overline{X}) =$____________，$E(S^2) =$____________

9.设 $T_n = T_n(X_1, X_2, \cdots, X_n)$ 为 $g(\vartheta)$ 的估计量，称 T_n 为 $g(\vartheta)$ 的相合估计是指__

10.设有 10 个样本观测值：1.6,　0.5,　3.5, 4.2, **3.6**, 2.1, 6.1, 2.5, 1.5, 3.0　则其中 **3.6** 的秩等于__________

二．（10 分）设 X_1, X_2, X_3, X_4 为取自总体 $X \sim N(\mu, 4^2)$ 的样本，对假设检验问题 $H_0: \mu = 5, H_1: \mu \neq 5$，（1）在显著性水平 0.05 下求拒绝域；（2）若 μ=6，求上述检验所犯的第二类错误的概率 β.

$(\Phi(1.96) = 0.975; \Phi(1.46) = 0.92785; \Phi(4.92) = 0.99999)$

三．（15 分）设总体 X 的概率密度为 $f(x;\theta) = \begin{cases} \dfrac{1}{2\theta}, & 0 < x < \theta, \\ \dfrac{1}{2(1-\theta)}, & \theta \le x < 1, \\ 0, & \text{其他}, \end{cases}$ 其中参数 $\theta\,(0 < \theta < 1)$ 未知，$(X_1, X_2, \cdots, X_n)$ 是来自总体的一个样本，$\bar{X}$ 是样本均值.

（1）求参数 θ 的矩估计量 $\hat{\theta}$；（2）证明 $4\bar{X}^2$ 不是 θ^2 的无偏估计量.

四．（15 分）设总体 X 的概率密度函数为 $f(x) = \begin{cases} 2e^{-2(x-\vartheta)} & x > \vartheta \\ 0 & x \le \vartheta \end{cases}$，其中 $\vartheta > 0$ 是未知参数，从总体 X 中抽取简单随机样本 $X_1, X_2, \cdots, X_n$，记 $\hat{\vartheta} = \min(X_1, X_2, \cdots, X_n)$.

（1）. 求总体 X 的分布函数 $F(x)$；（2）求统计量 $\hat{\vartheta}$ 的分布函数 $F_{\hat{\vartheta}}(x)$；

（3）如果用 $\hat{\vartheta}$ 作为 ϑ 的估计量，讨论它是否具有无偏性.

五．（15 分）设有多元线性回归模型 $\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2), \text{且}\varepsilon_i\text{相互独立，}\varepsilon_i\text{=1,2,}\cdots\text{,n} \end{cases}$

（1）如何由样本 $(x_{i1}, x_{i2}, \cdots, x_{ip}, y_i)$ $(i = 1, 2, \cdots n)$ 求出 $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_p$，试简述建立回归方程 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p$ 的过程.

（2）所谓对建立的回归方程进行回归效果的显著性检验，是指要检验什么假设？拒绝该假设意味着什么？接受该假设又意味着什么？

（3）当所求得的回归方程效果显著时，可以用该回归方程做什么工作？

六.（15 分）在芫花叶总黄酮提取工艺的研究中，考察的因素水平如下表,试解决下列问题：

（1）选取下列合适正交表，并作表头设计；

$L_4(2^3)$；$L_8(2^7)$；$L_9(3^4)$；$L_{18}(3^7)$；$L_{27}(3^{13})$

（2）如把 A、B、C 放在 L_9（3^4）表的第 1，2，3 列上，所得总黄酮收率（%）依次为 0.55，0.95，0.96，0.48，0.58，0.79，0.75，1.02，1.65，试对结果作直观分析，并确定最佳工艺条件.