北京大学数学科学学院期末试题

2015-2016 学年第一学期

考试科目　数理统计　　　　考试时间　2016 年 1 月 12 日

姓　　名　[illegible]　　　　学　　号　[illegible]

本试题共 8 道大题，满分 100 分

一.（18 分）设 X 服从指数分布，分布函数为

$$F(x)=\begin{cases}1-e^{-\frac{x}{\theta}} & \text{当} x>0\\ 0 & \text{否则}\end{cases}$$

其中 θ 是未知参数。若 $x_1,\ldots,x_n$ 为来自 X 的样本

1. 找出 θ 的最大似然估计，说明它在 $n\to\infty$ 时是否有相合性；
2. 求 θ 的矩估计；
3. 求 Fisher 信息量 $I(\theta)$；
4. 求 θ 的无偏估计的方差的下界；
5. 1 中的最大似然估计是否是 θ 的最小方差无偏估计？
6. 试找出 θ^2 的一个无偏估计。

二.（12 分）若随机变量 X 的分布密度可取下面的 $f_0(x)$ 或 $f_1(x)$：

$$f_0=\begin{cases}2x & \text{当} 0<x<1\\ 0 & \text{其他}\end{cases}\quad;\quad f_1=\begin{cases}2(1-x) & \text{当} 0<x<1\\ 0 & \text{其他}\end{cases}$$

基于 X 的一个观测值，对检验问题 H_0：$f(x)=f_0(x)\leftrightarrow H_1$：$f(x)=f_1(x)$，利用 N-P 引理求检验水平为 $\alpha=0.2$ 的 UMP 检验 φ，并求其第 II 类错误的概率。

三.（12 分）设 $m+n$ 组观察值满足线性回归模型

$$y_i=a+b+\varepsilon_i,\qquad (i=1,\ldots,m);$$

$$y_j=a-b+\varepsilon_j,\qquad (j=m+1,\ldots,m+n),$$

其中 $\varepsilon_i,\varepsilon_j$ 是独立同分布，均值为 0，方差为 σ^2 的随机变量列。试求出 a,b 的最小二乘估计 $\hat{a},\hat{b}$，并证明 $\hat{a},\hat{b}$ 不相关的充分必要条件为 $m=n$。

四.（12 分）在单因素试验中，因素 A 有 $s=2$ 个水平 A_1,A_2。设水平 A_i 下重复进行了 r 次

试验（$r \ge 2$），数据为$Y_{i1}, \cdots, Y_{ir}$（$i = 1,2$）。假定模型为$Y_{ij} = \mu_i + \varepsilon_{ij}$，其中随机误差$\varepsilon_{ij}$相互独立，共同的分布为$N(0, \sigma^2)$，$\sigma^2$为未知参数。对假设检验问题$H_0$：$\mu_1 = \mu_2$，方差分析中给出的$F$检验与假设检验中两个正态总体的$t$检验是否一致？试证明你的结论。

五. (12 分) 称Y服从参数为μ、σ^2的对数正态分布，若$Y = e^X$，而X服从$N(\mu, \sigma^2)$。对于Y的简单随机样本$y_1, \cdots, y_n$（$n \ge 2$），试求μ和σ^2的最大似然估计。

六. (12 分) 设$X_1, \ldots, X_n$是来自正态总体$N(\mu_1, \sigma_1^2)$的简单随机样本，$Y_1, \ldots, Y_n$是来自正态总体$N(\mu_2, \sigma_2^2)$的简单随机样本，两总体相互独立，μ_1, μ_2已知，试求假设检验问题：H_0：$\sigma_1^2 \le \sigma_2^2 \leftrightarrow H_1$：$\sigma_1^2 > \sigma_2^2$ 的水平为α（$0 < \alpha < 1$）的广义似然比检验法（需给出确定临界值的方法）。

七. (12 分) 针对数据$(x_1, y_1), \cdots, (x_n, y_n)$，设一元齐次线性模型为$Y_i = bx_i + \varepsilon_i$，其中$\varepsilon_i$相互独立，共同的分布为$N(0, \sigma^2)$。

1. 求b的最小二乘估计$\hat{b}$；
2. 试利用线性模型的检验理论推导对$H_0: b = 0$的检验统计量，并指出其分布（包括自由度）。
3. 试给出相应的平方和分解公式。

八. (10 分) 设$X_1, \cdots, X_n$为来自$B(1, p)$的简单随机样本，试证明：1、$X_1 + \cdots + X_n$是参数p的充分统计量；2、$X_1 + \cdots + X_k$ $(1 \le k < n)$是参数p的完全统计量，但不是充分统计量；3、$(X_1 + \cdots + X_k, X_{k+1} + \cdots X_n)$不是参数$p$的完全统计量。