# 2018东北大学统计学研究生招生考试

**一、简答题**

**1、简述假设检验的步骤，并说明假设检验与区间估计的关系。**

**答：假设检验的步骤。**

**①确定原假设和备择假设。**

**②确定检验统计量和显著性水平。**

**③确定检验统计量和拒绝域。**

**④确定样本。**

**⑤计算检验统计量的值和值。**

**⑥根据拒绝域或值作出决策。**

**假设检验与区间估计的关系：**

**①二者使用同一样本。**

**②二者使用同一统计量。**

**③同一显著性水平下，区间估计的置信区间与假设检验的接受域相同。**

**④假设检验中参数的值是已知的，而在区间估计中是未知的。**

**2、简述回归分析中变量选择的方法及其优缺点。**

**答：变量选择有三种方法：向前选择、向后选择、逐步选择。**

**向前选择是从只含有一个变量的模型开始，选择最显著的变量然后引入模型，直到所有变量对模型的影响都不显著为止。**

**向后选择是从全变量模型开始，剔除对模型影响最不显著的变量，直到所有变量对模型的影响都是显著的。**

**逐步选择是在引入变量过程中同时考虑剔除变量，即前面引入的变量在后面可能被剔除。**

**向前和向后选择中，变量一旦被引入（或被剔除），便不会再次被剔除（或被引入），忽略了变量间存在相互作用的影响。逐步选择考虑了这个问题，所以，变量在被引入（或被剔除）之后，有可能再次被剔除（或引入）。**

**3、简述主成分分析的思想与步骤。**

**答：主成分分析是利用降维的思想，在损失很少信息的前提下，把多个指标转化为几个综合指标的多元统计方法。**

**其思想是在保留原始变量尽可能多的信息的前提下达到降维的目的，从而简化问题的复杂性并抓住问题的主要矛盾。**

**主成分分析的步骤：**

**① 根据研究问题选取初始分析变量；**

**② 根据初始变量特性判断由协方差阵求主成分还是由相关阵求主成分；**

**③ 求协方差阵或相关阵的特征根与相应标准特征向量；**

**④ 判断是否存在明显的多重共线性，若存在，则回到第①步；**

**⑤ 得到主成分的表达式，并确定主成分个数，选取主成分；**

**⑥ 结合主成分对研究问题进行分析并深入研究。**

**4、简述分布、分布的定义，并说明它们之间的关系。**

**答：分布定义：若随机变量，，且相互独立，则称**

**是服从分布的，记为。**

**分布定义：若随机变量，，且相互独立，则称**

**是服从分布的，记为。**

**二者关系：当时，。**

**5、简述充分统计量的定义及其统计意义。**

**答：设总体的分布函数为为未知参数，，为样本，为不带有未知参数的统计量，若在给定统计量时，的条件分布与无关，则称为的充分统计量。**

**统计意义：充分统计量已经将样本中所含的关于参数的信息充分提取出来。**

**二、证明题**

**1、设总体，为其样本，为样本均值，为样本方差，试证明：**

**(1)**

**证：由题意知独立同分布于，故**

**从而**

**(2)**

**证：令，则且独立同分布，记**

**对做如下正交线性变换**

**即**

**显然依然是正态随机变量，且**

**又**

**从而独立同分布于。**

**由于，故，又**

**为正交阵，因此，故**

**即**

**而**

**其中**

**由卡方分布的独立可加性可得**

**(3)**

**证：由(2)可得，的分布与无关，因此与相互独立，从而与独立。**

**由t分布的定义可知**

**2、一元线性回归模型，其中，相互独立，估计的回归方程为，。试证明：**

**(1)**

**证：由题意知，记**

**上式对、分别求偏导可得**

**由于、是的最小二乘估计，因此满足**

**从而**

**(2)**

**证：由(1)知**

**即**

**三、计算题**

**1、总体，为其样本，试求参数的极大似然估计。**

**解：由题意知，的概率密度函数为**

**参数的极大似然函数为**

**显然为一常数函数，因此考虑从定义出发**

**要使得不为零，则应满足**

**从而的极大似然估计为**

**2、列联表检验。**

**见课本**