**2016年东北大学研究生招生考试432统计学**

**一、简答题**

**1.直方图的作用，散点图的作用，盒子图的作用。**

**答：直方图：直方图是用于展示分组数据分布的一种图形，它是用矩形的宽度和高度（即面积）来表示频数分布的。绘制该图时，在平面直角坐标系中，用横轴表示数据分组，纵轴表示频数或频率。**

**散点图：散点图用于反映两个变量之间的相互关系，在直角坐标系中，以纵轴表示其中一个变量，横轴表示另一变量。**

**盒子图：盒子图是由最大值、最小值、中位数、上四分位数、下四分位数五个数字特征构成的用于反映原始数据分布特征的图形，也可用于多个总体分布特征的比较。**

**2.写出两个随机事件之间独立的公式与两个随机变量之间独立的公式，并说明如何从随机事件独立性的角度理解随机变量的独立性。**

**答：随机事件相互独立：记为两个随机事件，称随机事件相互独立，如果满足**

**显然，当或为零概率事件或必然事件时，一定相互独立。**

**随机变量相互独立：记为两个随机变量，称随机变量是相互独立的，如果对任意实数满足**

**若随机事件为“”，随机事件为“”，则随机事件与随机变量相互独立是等价的。**

**二、计算题**

**1.设且，相关系数，，求和。**

**解：由题意知**

**则有**

**2.设,,…,是来自均匀分布总体的随机样本，求的极大似然估计。**

**解：由题意知，总体的概率密度函数为**

**参数的极大似然函数为**

**显然为的单调递减函数，又若要使得，则应有**

**故当时取到最大值，即的极大似然估计为**

**3.法庭上，表示嫌疑人无罪，表示嫌疑人有罪，表示现场发现的证据，，，且，法官认为即认为嫌疑人有罪，问如果按照法官的标准，该嫌疑人是否有罪。**

**解：由题意可知，与为对立事件，故而**

**由全概率公式可知，嫌疑人留下证据F 的概率**

**则**

**因此在法官的标准下，该嫌疑人无罪。**

**4.设，，…，是来自参数为的泊松分布的随机样本。**

**（1）求参数的矩估计。**

**解：参数为的泊松分布的概率密度函数为**

**则**

**又、的矩估计分别为**

**故**

**即的矩估计为**

**（2）求参数为的极大似然估计。**

**解：参数的极大似然函数为**

**对上式取对数**

**对求偏导，并令其为0得**

**解得**

**由极大似然估计的不变性可知的极大然估计为**

**5.设，，是来自于总体的样本，设，。**

**（1）求显著性水平为下该假设检验的拒绝域。**

**解：由抽样分布定理知**

**其中**

**因此选择统计量作为检验统计量**

**由**

**得到拒绝域为**

**（2）若成立，求参数置信度为的置信区间。**

**解：当原假设成立时，，即总体服从分布**

**已知**

**故可由**

**得参数置信度为的置信区间为**

**三、论述题**

**1.简述表示平均的统计量，并说明他们的特点。**

**答：平均数、中位数、众数都是描述平均的统计量。**

**平均数是对所有变量值进行求和后，除以变量个数得到的值，是数据集中程度最主要的测度值，由于平均数利用了全部数据，因此它容易受到极端值的影响。平均数只能用于描述数值型数据。**

**中位数是对一组数据进行排序后，位于中间位置上的数据。中位数是一个位置代表值，因此不易受到极端值的影响。当数据拖尾严重时，宜使用中位数而非平均数。中位数可用于描述数值型数据和顺序型数据。**

**众数是一组数据中出现次数最多的变量值，它也是一个位置代表值。如果一组数据不存在明显的峰值，则众数可能不存在；若果有两个或以上的明显的峰值，则众数可能有多个。众数主要用于描述分类数据，也可用于描述数值型数据和顺序型数据。**

**2.简述方差分析的作用。**

**答：方差分析通过检验各总体的均值是否相等来判断分类型自变量对数值型因变量是否有显著影响。它是通过对数据误差来源的分析来判断不同总体的均值是否相等，进而分析自变量对因变量是否有显著性影响。**

**作用：**

**①检验各水平对因子是否有显著性影响。**

**②检验误差项是否服从正态分布。**

**③检验各总体间均值是否相等。**