概率论与数理统计第五次习题课题目

题1 将编号为 1 至 n 的 n 个球随机投入编号为 1 至 n 的 n 个盒子中，并限制每一个盒子中只能放入一个球，设球与盒子的号码一致的个数为 S_n ，求证：

$$\frac{S_n - ES_n}{n} \xrightarrow{P} 0, \quad n \to \infty.$$

题2 一报贩发现每个路过他的报摊的行人向他买报的概率为 $\frac{1}{3}$ 。求在他出售了 100 份报纸时的过路人的数目在 280 人到 320 人之间的概率。（用两种不同的估计方法，并比较它们的优劣）

题3 设总体分布为$U[\theta-1,\theta+1]$，其中θ是未知参数，$X_1,\ldots,X_n$是来自该总体的简单随机样本。

1. 求θ的矩估计量$\hat{\theta}$，判断它的相合性和无偏性，计算均方误差$\mathrm{MSE}(\hat{\theta})$；
2. 证明对任何$0 \le t \le 1$，$\hat{\theta}_t := tX_{(n)} + (1-t)X_{(1)} + 1 - 2t$都是$\theta$的极大似然估计量；
3. 求$X_{(1)}$和$X_{(n)}$的概率分布以及数学期望$EX_{(1)}$、$EX_{(n)}$；
4. 问$\hat{\theta}_t$是否为θ的相合估计和无偏估计？
5. 求$X_{(1)},X_{(n)}$的联合分布，以及$X_{(1)}+X_{(n)}$的概率分布，并计算方差$\mathrm{Var}(\hat{\theta}_{1/2})$；对比第1问的结果，你有何结论？

题4 设总体分布为$U[\theta,2\theta]$，其中$\theta>0$是未知参数，$X_1,\ldots,X_n$是来自该总体的简单随机样本。

1. 利用矩估计方法求θ的无偏估计量$\hat{\theta}_1$，计算其方差；
2. 求θ的极大似然估计量$\hat{\theta}_{\mathrm{MLE}}$，并由它构造$\theta$的一个无偏估计$\hat{\theta}_2$，并计算$\hat{\theta}_2$的方差；
3. 把$X_{(1)}$当作θ的一个点估计，由它构造θ的一个无偏估计$\hat{\theta}_3$，并计算$\hat{\theta}_3$的方差；
4. 试比较上述无偏估计的有效性；
5. 求θ的置信水平为$1-\alpha$的置信区间。

题5 设某城市有 N 辆机动车，牌号依次是 $1,2,\cdots,N$ 。一个人将他一天内看到的所有机动车牌号（包括重复出现的牌号）都记录下来，得到 $X_1,X_2,\cdots,X_n$ 。如果用最大牌号 $X_{(n)}$ 作为对N的一个估计（即近似值），我们采取以下方式来评价这个估计：

1. 当n充分大时，$X_{(n)}$是否近似等于N？并且试证明 $X_{(n)}$ 是 N 的极大似然估计
2. 试给出 N 的一个矩估计，并与其极大似然估计$X_{(n)}$进行比较。
3. 如果这样的观察方式被多次重复进行，每次得到$X_{(n)}$的一个观测值，那么根据大数定律，$X_{(n)}$观测值的算术平均值将以$EX_{(n)}$为极限，求$EX_{(n)}-N$（称为这种近似方式的“偏”，即系统误差）的值。
4. 如果$X_{(n)}$存在系统误差（有偏，即$EX_{(n)}-N\neq 0$），那么你有什么办法可以消除这个系统误差？

如果不重复记录的话，如何用观测值$X_1,X_2,\cdots,X_n$给出 N 的一个估计？分析你给出的估计的性质，并与重复情况下的估计进行比较。

题6 甲乙两位编辑独立地对同一段文字进行校对，甲发现了n_1处错误，乙发现了n_2处错误，并且其中有n_3处错误是甲乙共同发现的。试用矩估计法和极大似然估计法估计这段文字的错误个数。

题7 设$X_1,X_2,\ldots,X_n$是来自总体$N(\mu,1)$的简单随机样本，其中μ是未知常数。

1. 求μ的置信水平为99%的置信区间；
2. 为使上述置信区间的长度不超过0.1，问样本容量n至少需要多大？