周期四 慕课作业

T1

> 基尼系数（Gini Index）是衡量经济不平等的一个主要指标。在这个问题中，我们考虑家庭收入 x 服从离散分布，其分布律为 $P(X=y_i)=\frac{1}{n}$，其中 $i=1,\ldots,m$，并且 $a\leq y_1\leq y_2\leq\ldots\leq y_n\leq b$。
>
> 定义 L_i 为收入不超过 y_i 的家庭的总收入占整体家庭总收入的比例。约定 $L_0=0$，在 Oxy 平面上连接点 $\left(\frac{i}{n},L_i\right)$，$i=0,1,\ldots,n$ 的分段线性函数 L 称为洛伦兹曲线（Lorenz Curve）。洛伦兹曲线 L 与线段 $y=x,0\leq x\leq 1$ 围成的区域记为 S。洛伦兹曲线 L 与 x 轴和线段 $x=1,0\leq y\leq 1$ 围成的区域记为 B。
>
> 基尼系数 G 定义为区域 S 的面积 A 的两倍。我们需要给出 G 的表达式，并简要说明 G 如何反映家庭收入差异程度。

首先，根据定义，我们有：

$$L_i=\frac{\sum_{j=1}^{i}y_j}{\sum_{j=1}^{n}y_j},\quad i=1,\ldots,n.$$

区域 B 的面积 S_B 可以表示为：

$$S_B=\sum_{i=1}^{n}\frac{L_{i-1}+L_i}{2}\cdot\frac{1}{n}=\frac{1}{2n}\left(2\sum_{i=1}^{n}L_i-L_n\right).$$

进一步简化得到：

$$S_B=\frac{1}{n}\sum_{i=1}^{n}L_i-\frac{1}{2n}=\frac{1}{n\sum_{j=1}^{n}y_j}\sum_{i=1}^{n}(n-i+1)y_i-\frac{1}{2n}.$$

因此，基尼系数 G 可以表示为：

$$G=2\left(\frac{1}{2}-S_B\right)=\frac{n+1}{n}-\frac{2}{n\sum_{j=1}^{n}y_j}\sum_{i=1}^{n}(n-i+1)y_i$$

由于 $y_1\leq y_2\leq\ldots\leq y_n$，故 $L_i\leq\frac{i}{n}$，即洛伦兹曲线 L 总在直线 $y=x$ 的下方。基尼系数 G 越大，说明洛伦兹曲线在横坐标较小时增长较慢，在横坐标较大时增长迅速，这表明少部分人占有收入比例很高，家庭收入差异大。

T2

> 小剧场一排有 n 个座位。由于各排之间空隙较小，如果某座位已有人入座，则入座者必须起身才能让后来者通过该座位。若一与会者进入会场时该排有若干个座位可供其选择，则他以相等的概率选择其中一个座位就坐。
>
> 1. 若该排座位只有左侧一个入口，所有与会者一旦就坐就不愿起身让后来者通过。记 E_n 为该排最终入座人数的期望，试写出 E_n 满足的递推关系，并求 E_n。
> 2. 若该排座位在左右两侧均有入口，所有与会者以 p 的概率起身让后来者通过，以 $1-p$ 的概率不让后来者通过。记 F_n 为该排最终入座人数的期望，试写出 F_n 满足的递推关系。

若进入会场时有 n 个座位可供选择，他选择每个座位的概率均为 $\frac{1}{n}$。若他选择了第 i 个座位，下一个进入会场者只有其左侧的 $i-1$ 个座位可供选择。因此，可得递推关系：

$E_n = 1 + \frac{1}{n}\sum_{i=1}^{n} E_{i-1}$

由上式可得:

$nE_n = n + \sum_{i=1}^{n} E_{i-1}$

$(n+1)E_{n+1} = n + 1 + \sum_{i=1}^{n+1} E_{i-1}$

两式相减可得 $(n+1)E_{n+1} - nE_n = 1 + E_n$，即:

$E_{n+1} - E_n = \frac{1}{n+1}$

由 $E_1 = 1$ 可得:

$E_n = \sum_{i=1}^{n} \frac{1}{i}$

记 G_n 为只有一侧有入口时最终入座人数的期望，则 G_n 满足递推关系:

$G_n = 1 + pG_{n-1} + (1-p) \cdot \frac{1}{n}\sum_{i=1}^{n} G_{i-1}.$

F_n 满足递推关系:

$F_n = 1 + pF_{n-1} + (1-p)\frac{1}{n}\sum_{i=1}^{n}(G_{i-1} + G_{n-i}).$