2.6 重要的离散分布

2.6.1 几何分布

2.6.1.1 几何分布的定义

对于n重伯努利实验，如果每次得到“是”的概率为p，设随机变量：

$$X = \text{首次得到“是”时进行的试验次数}$$

则称：

$$p(k) = P(X = k) = (1-p)^{k-1}p, \quad k = 1, 2, \cdots$$

为随机变量X的几何分布，也可以记作：

$$X \sim Ge(p)$$

其数学期望和方差为：

$$E(X) = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}$$

2.6.1.2 几何分布的无记忆性

假设梅累骑士和尼古拉斯抛掷骰子，谁先抛出6点谁获胜。由于谁先抛出6谁就获胜，所以这个问题很适合用几何分布。假设抛出6点的概率为p，那么有随机变量X：

$$X \sim Ge(p)$$

梅累骑士连抛m把都没有抛出6点，所以对应概率为：

$$P(X > m) = \sum_{k=m+1}^{\infty}(1-p)^{k-1}p = \frac{p(1-p)^m}{1-(1-p)} = (1-p)^m$$

在梅累骑士连抛m把都没有抛出6点的条件下，尼古拉斯接着抛了n把，就抛出6点的概率为（这是一个条件概率）：

$$P(X = m+n | X > m) = \frac{P(X = m+n)}{P(X > m)} = \frac{(1-p)^{m+n-1}p}{(1-p)^m}$$

$$= (1-p)^{n-1}p = P(X = n)$$

可以看到，梅累骑士连抛m把对尼古拉斯没有影响，计算结果与尼古拉斯直接抛掷n次的概率是一样的，所以游戏是公平的。

数学家把这种特性形象地称作几何分布的无记忆性，只要期待的“是”没有出现，那么几何分布就仿佛不记得之前发生的事件。值得注意的是在离散分布中只有几何分布具有此特性。

2.6.2 负二项分布

每次飞行之后，都要对飞机进行检查，其中一非关键的易损件，每次飞行后都有p的概率损坏。本着节约成本的原则，规定可以修理，但损坏r次后必须报废更换。问，第k次飞行后此零件报废的概率为多少?

假设随机变量为:

$$X = \text{第}k\text{次飞行，零件正好损坏}r\text{次}$$

那根据题意就是要求$P(X=k)$。所求概率可这么分析，之前$k-1$次飞行后，已经发生了$r-1$次损坏（所以k必须大于等于r），这里可通过二项分布来计算:

$$\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}, \quad k=r,r+1,\cdots$$

然后第k次飞行，此零件必须要损坏，所以所求概率为:

$$P(X=k)=\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}p$$

$$=\binom{k-1}{r-1}p^{r}(1-p)^{k-r}$$

2.6.2.1 负二项分布的定义

对于n重伯努利实验，如果每次得到“是”的概率为p，设随机变量:

$$X = \text{第}r\text{次“是”发生时的实验次数}$$

则称:

$$p(k)=P(X=k)=\binom{k-1}{r-1}p^{r}(1-p)^{k-r}, k=r,r+1,\cdots$$

为随机变量X的负二项分布，也称为帕斯卡分布，也可以记作:

$$X \sim Nb(r,p)$$

其数学期望为:

$$E(X)=\frac{r}{p}, \quad Var(X)=\frac{r(1-p)}{p^2}$$

2.6.2.2 负二项分布与几何分布的关系

1 几何是负二项的特例

负二项分布是这样的:

$$p(k)=P(X=k)=\binom{k-1}{r-1}p^{r}(1-p)^{k-r}, k=r,r+1,\cdots$$

$r=1$的时候，就得到了几何分布:

$$p(k)=P(X=k)=(1-p)^{k-1}p, \quad k=1,2,\cdots$$

2 负二项是几何的和

参数为r、p的负二项分布可以表示为如下事件序列：

$$\underbrace{\text{否 否}\cdots\text{是}}_{X_1}\quad\underbrace{\text{否 否}\cdots\text{是}}_{X_2}\quad\cdots\quad\underbrace{\text{否 否}\cdots\text{是}}_{X_r}$$

图中所示的每一段X_1、X_2、$\cdots$、X_r都是几何分布，所以有：

$$X = X_1 + X_2 + \cdots + X_r \sim Nb(r, p)$$

所以负二项分布的期望为：

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_r) = \frac{r}{p}$$

2.6.3 超几何分布

2.6.3.1 超几何分布的定义

设有N件产品，其中有M件不合格品，随机抽取n件产品，则其中含有m件不合格产品的概率为多少？

假设随机变量：

$$X = \text{随机抽取的}n\text{件中有}m\text{件不合格品}$$

这个随机变量的概率可以用古典概率来求，首先，样本空间就是从N件中随便抽取n件，所以：

$$|\Omega| = \binom{N}{n}$$

然后有m件从不合格品中抽取，剩下的在合格品中抽取，则有：

$$|X| = \binom{M}{m}\binom{N-M}{n-m}$$

所求概率即为：

$$P(X = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, m = 0, 1, \cdots, r$$

其中$r = min(M, n)$。此时称X服从超几何分布，可以记作：

$$X \sim h(n, N, M)$$

其数学期望和方差为：

$$E(X) = n\frac{M}{N},\quad Var(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\left(1 - \frac{n-1}{N-1}\right)$$

2.6.3.2 超几何分布与二项分布

超几何分布与二项分布类似，都是求抽取n次其中有m次“是”的概率，只是：

- 二项分布：相当于抽取之后放回
- 超几何分布：抽取之后不放回

所以在超几何分布中，如果被抽取的总数N特别大，那么放回不放回区别也就不大了，此时，那么超几何分布可以近似看作二项分布。

这点从两者的期望、方差也可以看出来：

	二项分布	超几何分布
μ	np	$n\frac{M}{N}$
σ	$np(1-p)$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)$

令$p=\dfrac{M}{N}$，超几何分布的期望和方差可以写作：

$$\mu=n\frac{M}{N}=np$$

$$\sigma=n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)=np(1-p)\left(1-\frac{n-1}{N-1}\right)$$

对超几何分布而言，当N足够大的时候，$\dfrac{M}{N}$可看作取出不合格产品的概率，那此时超几何分布可看作二项分布。

2.7 离散分布的总结

伯努利分布	抛硬币，二选一
二项分布	n重伯努利，出现k次“是”
泊松分布	二项分布的极限
几何分布	n重伯努利，第k次首次出现“是”
负二项分布	几何分布的和
超几何分布	不放回抽样的二项分布

2.8 连续分布的概率密度函数

2.8.1 概率密度函数严格定义

如果函数$p(x)$满足下列两个条件（对应了概率的三大公理）：

- **非负性**:

$$p(x) \geq 0$$

- **规范性**（暗含了可加性），因为是连续的，所以通过积分相加:

$$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1$$

则称其为概率密度函数（Probability Density Function，简写为PDF）。

2.8.2 连续的累积分布函数

2.8.2.1 连续的累积分布函数的定义

连续随机变量X的概率密度函数为$p(x)$，则:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(t)\mathrm{d}t$$

称为X的累积分布函数。

2.8.2.2 连续的累积分布函数的性质

根据微积分第一基本定理有:

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = p(x)$$

> 例题
>
> 设连续随机变量X的分布函数为:
>
> $$F(X) = \begin{cases} 0 & x < 0 \\ Ax^2 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$
>
> 求A的值
>
> 因$F(x)$为累计分布函数，由概率密度函数定义结合题意可知
>
> $$F(1) = \int_{-\infty}^{1} p(x) = 1 \quad F(0) = \int_{-\infty}^{0} p(x) = 0$$
>
> 由定积分的运算法则知

> $$\int_0^1 p(x)dx = F(1) - F(0) = 1$$
>
> 而根据题意有,当$0 < x < 1$时
>
> $$\int_0^x p(t)dt = Ax^2$$
>
> 根据基本积分表可得
>
> $$p(x) = 2Ax$$
>
> 则
>
> $$\int_0^1 2Ax = 1$$
>
> 因此
>
> $$A = 1$$

2.9 均匀分布

如果连续随机变量X的概率密度函数为:

$$p(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其它} \end{cases}$$

则称X服从区间(a,b)上的均匀分布，记作$X \sim U(a,b)$，其累积分布函数为:

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

期望和方差分别为:

$$E(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12}$$