1. 二项分布解题及局限性

假设你是一个交通工程师，想知道任意时刻通过街上某一点的车辆数。例如，想确定某一小时内100辆车或5辆车通过的概率。

最好的方法就是先定义一个随机变量：X = # of cars pass in a hour

然后求出该变量的概率分布，这样就能求一小时内100辆车或5辆车通过的概率了。

在进行之前，有两个假设要说：

- 街上的某一地点任意时刻的情况没有差异。这是一种简化假设
- 一段时间的车流量对另一段时间没有影响，也就是说具有独立性。

有了这两个假设（Given that），我们就可以对分布进行建模了。

（1）首先，守在观测点，测得一个合理的随机变量的期望值。

对于任何分布，我们可以首先估计均值。可以通过连续的观察多个小时的车流量然后平均起来，这也许就是对总体均值的很好估计值了。X是一个随机变量，所以也就是期望值。假设期望值的最后估计值是λ:

$$E(x) = \lambda$$

(2) 用二项分布建模

我们知道二项分布的期望值等于试验的次数n乘以每一次成功的概率p。则概率等于期望值除以试验的次数：

$$p = \frac{\lambda}{n}$$

如果每分钟试验一次，那么一小时试验60次，每分钟通过的概率就是$\frac{\lambda}{60}$。则每个小时通过k辆车的概率：

$$P(X=k) = \binom{60}{k} (\frac{\lambda}{60})^k (1-\frac{\lambda}{60})^{60-k}$$

但是，如果每分钟不止一辆车通过怎么办？之前我们把有一辆车通过叫成功。但没有考虑到1分钟内头同时通过多辆车的情况。解决方法就是划分更多的区间，例如划分成秒，则每小时通过k辆车的概率：

$P(X=k)=\binom{3600}{k}(\frac{\lambda}{3600})^k(1-\frac{\lambda}{3600})^{3600-k}$ 这是更好的近似，但是也有可能一秒钟开过多辆车，我们可以继续进行区间分割，让这个数字越来越大。这种直观很对，一直下去就能得到泊松分布。

2.泊松分布

泊松分布来自于二项分布。二项分布来自于投硬币，这是一切的源头。当区间分布趋近于无穷大时，这就是泊松分布。

> 我们需要两个知识点：
> 首先：$\lim\limits_{x\to\infty}(1+\dfrac{a}{x})^x=e^a$。令 $\frac{1}{n}=\frac{a}{x}$,则$x=na$,当x趋近于无穷时，n也趋近于无穷。换元后，极限式等价于：
> $\lim\limits_{n\to\infty}(1+\dfrac{1}{n})^{na}=\lim\limits_{n\to\infty}((1+\dfrac{1}{n})^n)^a=(\lim\limits_{n\to\infty}((1+\dfrac{1}{n})^n)^a=e^a$
> 此外：$\dfrac{x!}{(x-k)!}=(x)(x-1)...(x-k+1)$

二项分布的区间趋近于无穷时，公式：

$$
\begin{aligned}
P(X=k) &= \lim_{n\to\infty}\binom{n}{k}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^{n-k} \\
&= \lim_{n\to\infty}\frac{n!}{(n-k)!k!}\frac{\lambda^k}{n^k}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k} \\
&= \lim_{n\to\infty}\frac{(n(n-1)...(n-k+1))}{n^k}\frac{\lambda^k}{k!}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k} \\
&= \lim_{n\to\infty}\frac{n^k+....}{n^k} \qquad (1) \\
&+ \frac{\lambda^k}{k!} \\
&+ \lim_{n\to\infty}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k} \qquad (2)
\end{aligned}
$$

(1)中只要上下两项最高次数相同，系数相同，n趋于无穷时，肯定趋于1.

(2)中n趋于无穷时，$(1-\frac{\lambda}{n})^n$的极限为$e^{-\lambda}$。而$(1-\frac{\lambda}{n})^{-k}$中，当n趋近于无穷时，$\frac{\lambda}{n}$为0，而1的任何次方都是1，所以极限为1.

最后：

$$P(X=k)=1\cdot\frac{\lambda^k}{k!}e^{-\lambda}\cdot 1=\frac{\lambda^k}{k!}e^{-\lambda}$$

例如，你是一个交通工程师，测出平均每小时9辆车通过。然后，你想知道某小时正好有2辆车经过的概率：

$$P(X=k)=\frac{9^2}{2!}e^{-9}\approx 0.5\%$$