统计 笔记

任云玮

目录

1 常见分布

1.1 离散分布

1.1.1 Hypergeometric Distribution

含义 设有 N 个球，其中 M 个为红色，$N-M$ 个为绿色，它们除颜色以外无区别。从中选取 K 个，考虑恰有 x 个是红球的概率分布。

公式

$$
\begin{aligned}
&P(X=x|N,M,K)=\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}.\\
&\mathrm{E}\,X=\frac{KM}{N}.\\
&\mathrm{Var}\,X=\frac{KM}{N}\frac{(N-M)(N-K)}{N(N-1)}.\\
&x=0,1,\ldots,K.
\end{aligned}
$$

1.1.2 Binomial Distribution

含义 考虑 n 次相同的成功概率为 p 的 Bernoulli 试验，考虑其中恰有 y 次成功的概率分布。

公式

$$
\begin{aligned}
&P(Y=y|n,p)=\binom{n}{y}p^y(1-p)^{n-y}.\\
&\mathrm{E}\,X=np\\
&\mathrm{Var}\,X=np(1-p).\\
&M_X(t)=[pe^t+(1-p)]^n.
\end{aligned}
$$

1.1.3 Poisson Distribution

含义 Poisson 分布可以用来描述在一段时间内，某事件发生的次数的概率分布，例如类似于等车的行为。设 $t\geq 0$ 而 N_t 是关于 t 的整数值随机变量，设它满足如下性质：

1. 初始无到达：$N_0=0$；
2. 不相交时间区间内的达到次数相互独立：设 $s<t$，则 N_s 和 N_t-N_s 独立；
3. 到达次数近于区间长度有关：N_s 和 $N_{t+s}-N_t$ 是同分布的；
4. 当时间区间充分小时，到达可能性正比于区间长度：$\lim_{t\to 0}\dfrac{P(N_t=1)}{t}=\lambda$；

5. 不会出现同时到达的情况：$\lim_{t\to 0}\dfrac{P(N_t>1)}{t}=0$；

经常的，我们会进行归一化，同时只考虑 $t=1$ 的情况，即“仅等一单位时间情况下，等到车的数量”。

公式

$$
\begin{aligned}
&P(N_t=n|\lambda)=e^{-\lambda t}\frac{(\lambda t)^n}{n!}.\\
&\mathrm{E}\,N_1=\lambda.\\
&\mathrm{Var}\,N_1=\lambda.\\
&M_{N_1}(t)=e^{\lambda(e^t-1)}.
\end{aligned}
$$

利用 Poisson 分布近似 注意到 Poisson 描述的是连续的时间段中事件发生的情况，在某些离散的情况下，亦可以使用 Poisson 分布来近似。考虑二项分布的情况，我们可以把 Bernoulli 试验的成功理解成 Poisson 分布中的“到达”，同时把单个的 Bernoulli 试验理解成一个单位时间，则自然的 p 就对应着 λ。若 p 充分小，则可以认为这些离散的 Bernoulli 试验已经和连续的情况差不多乐，所以此时可以有

$$P(Y=y|n,p)\sim P(N_n=y|p)=P(N_1=y|np).$$

我们可以通过比较它们的递推式来证明这一结论。

1.1.4 Negative Binomial Distribution

含义 考虑 n 次独立的 Bernoulli(p) 试验，设随机变量 X 表示在第 X 轮发生了等 r 次成功，即描述一个成功了 r 次后才会停止的试验。或者考虑一种等价的形式，设随机变量 Y 表示在第 r 次成功前的失败的次数，显然有 $Y=X-r$。

公式

$$
\begin{aligned}
&P(X=x|r,p)=\binom{x-1}{r-1}p^r(1-p)^{x-r}\\
&P(Y=y|r,p)=\binom{r+y-1}{y}p^r(1-p)^y=(-1)^y\binom{-r}{y}p^r(1-p)^y.\\
&\mathrm{E}\,Y=r\frac{1-p}{p}.\\
&\mathrm{Var}\,Y=r\frac{1-p^2}{p}=\mu+\frac{1}{2}\mu^2.
\end{aligned}
$$

利用负二项分布近似 Poisson 分布 注意到如果我们令 $p \to 1$ 而 $r \to \infty$ 且同时满足 $r(1-p) \to \lambda$，则期望与方差都会趋于 λ，和 Poisson 分布的结果一致。这暗示了在这一条件下，负二项分布或许可以近似 Poisson 分布，而事实确实如此。

我们可以这样考虑这个事情。当 r 变大时候，一次试验在其中所占的份就变小了，从而变得更加的连续了，而在这里把一次失败理解成一次到达，则 $p \to 1$ 的条件同 Poisson 分布中关于概率与区间长度的要求是一致的，而 $r(1-p) \to \lambda$ 的要求可以从归一化之后的角度看待。

1.1.5 Geometric Distribution

含义 负二项分布中，取 $r = 1$ 的特殊情况，即一种成功就停止的试验。

公式

$$P(X = x|p) = p(1-p)^{x-1}.$$
$$\mathrm{E}\,X = \frac{1}{p}.$$
$$\mathrm{Var}\,X = \frac{1-p}{p^2}.$$

性质 之前的失败对于之后的概率没有影响，即有

$$P(X > s|X > t) = P(X > s - t).$$

这也意味着如果某种概率是随着试验次数/时间的增长而有变化时，几何分布是不适用的。