可靠性数学基础——泊松分布

光波
学习使人充实快乐；学无止境，其乐无穷！

在历史上拍松分布是作为二项分布的近似，于1837 年由法国数学家泊松(Poisson S. D. 1781 ~ 1840)首次提出，以后发现，很多取非负整数的离散随机变量都服从泊松分布，这里仍按历史发展次序来介绍泊松公布。

由二项分布演变成泊松分布

在二项分布b(n, p)中，若相对地说，n 大， p小，而乘积λ= np大小适中时，二项分布中诸概率有一个很好的近似公式。这就是著名的泊松定理。

定理 2. 2. 1(泊松定理) 在 n 重贝努里试验中，以 p_n 表示在一次试验中成功发生的概率。且随着 n 增大，p_n 在减小。若 $n\to\infty$时有 $\lambda_n=np_n\to\lambda$(正数)，则出现 x 次成功的概率

$$\binom{n}{x}p_n^x(1-p_n)^{n-x}\to\frac{\lambda^x}{x!}e^{-\lambda}\quad(n\to\infty)$$

证：由 $p_n=\lambda_n/n$，可得

$$\begin{aligned}&\binom{n}{x}p_n^x(1-p_n)^{n-x}\\&=\frac{n(n-1)\cdots(n-x+1)}{x!}\left(\frac{\lambda_n}{n}\right)^x\left(1-\frac{\lambda_n}{n}\right)^{n-x}\\&=\frac{\lambda_n^x}{x!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-x}\end{aligned}$$

对固定的 x 有

$$\lim_{n\to\infty}\lambda_n=\lambda$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{n-x}=e^{-\lambda}$$

即得

$$\lim_{n\to\infty}\binom{n}{x}p_n^x(1-p_n)^{n-x}=\frac{\lambda^x}{x!}e^{-\lambda}$$

就证明了本定理。

泊松分布的概率密度函数关键推导

$$\lim_{\infty}\lambda_n=n*p_n=\lambda$$

$$\lim_{n\to\infty}(1-\frac{\lambda_n}{n})^{(n-x)}=\lim_{n\to\infty}[1+\frac{-\lambda_n}{n}]^{\left(\frac{n}{-\lambda_n}\right)*\left(\frac{-\lambda_n}{n}\right)*(n-x)}$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{(n-x)}=\lim_{n\to\infty}[e]^{\left(\frac{-\lambda_n}{n}\right)*(n-x)}=\lim_{n\to\infty}[e]^{(-\lambda_n)*\left(\frac{n-x}{n}\right)}$$

$$=e^{-\lambda}$$

其中，应用了e的极限定义

自然对数的底e是由一个重要极限给出的。我们定义：当n趋于无穷大时，$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=e$.

e是一个无限不循环小数，其值约等于2.718281828459…，它是一个超越数。

$$\binom{n}{x}p_n^x(1-p_n)^{n-x} \doteq \frac{\lambda^x}{x!}e^{-\lambda}$$

其中 λ 就取 np_n。

例 2.2.11 在 500 人组成的团体中，恰有 k 个人的生日是在元旦的概率是多少？

解：在该团体中每个人的生日恰好在元旦的概率为 $p=1/365$，则该团体中生日为元旦的人数 $X \sim b(500,1/365)$。即

$$P(X=k)=\binom{500}{k}\left(\frac{1}{365}\right)^k\left(1-\frac{1}{365}\right)^{500-k}$$

$$k=0,1,\cdots,500$$

这个概率计算是复杂的，但为了比较，仍对 $k=0,1,\cdots,6$ 进行计算，然后再用近似公式(2.2.7)计算，其中 $\lambda=500/365=1.3699$。两者结果都列在表 2.2.3 上。

从表 2.2.3 可见，两者的差别都在第四位小数上才显示出来，其近似程度是相当好的。

表 2.2.3 二项分布与泊松近似的比较

k	$\binom{500}{k}\left(\frac{1}{365}\right)^k\left(1-\frac{1}{365}\right)^{500-k}$	$\frac{(1.3699)^k}{k!}e^{-1.3699}$
0	0.2537	0.2541
1	0.3484	0.3481
2	0.2388	0.2385
3	0.1089	0.1089
4	0.0372	0.0373
5	0.0101	0.0102
6	0.0023	0.0023
≥7	0.0006	0.0006

泊松分布的概率之和为1及期望为λ

泊松分布 泊松定理中的泊松概率 $\lambda^x e^{-\lambda}/x!$ 对一切非负整数 x 都是非负的，且其和恰好为 1，因为

$$\boxed{\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}=e^{-\lambda}\cdot e^{\lambda}=1}$$

这样一来，泊松概率的全体组成的一个概率分布，称为泊松分布，记为 $P(\lambda)$，若随机变量 X 服从泊松分布，即 $X\sim P(\lambda)$，这意味着，X 仅取 0,1,2,…等一切非负整数，且取这些值的概率为

这样一来，泊松概率的全体组成的一个概率分布，称为泊松分布，记为P(λ)，若随机变量X 服从泊松分布，即X~ P(λ)，这意味着， X 仅取0 ， 1 ， 2 ，…等一切非负整数，且取这些值的概率为

$$P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}, \quad x=0,1,\cdots$$

其中参数 $\lambda>0$，它的数学期望容易算得，

$$E(X)=\sum_{x=0}^{\infty}x\cdot\frac{\lambda^x}{x!}e^{-\lambda}$$

$$=\lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}=\lambda$$

这表明：**泊松分布 $P(\lambda)$ 的数学期望就是参数 λ。**

参数λ是分布的均值。

泰勒公式及e^x的无穷级数展开

说明为什么泊松公式中含有e，e是一个非常非常重要的数。

补充说明泊松分布概率之和为1中无穷级数之和为e^x的说明

泰勒公式：将一个在 x=x0 处具有 n 阶导数的函数 f (x) 利用关于 (x-x0) 的 n 次多项式来逼近函数的方法。

若函数 f(x)在包含 x0 的某个闭区间[a,b]上具有 n 阶导数，且在开区间(a,b) 上具有 (n+1) 阶导数，则对闭区间[a,b]上任意一点 x，成立下式：

$$f(x)=\frac{f(x_0)}{0!}+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+R_n(x)$$

其中，表示 f (x) 的 n 阶导数，等号后的多项式称为函数 f (x) 在 x0 处的泰勒展开式，剩余的 Rn (x) 是泰勒公式的余项，是 (x-x0) n 的高阶无穷小。

e^x在x=0展开得

f(x) = e^x = f(0) + f′(0)*x + f″(0)*x²/2! + ... + fⁿ(0)*x^n/n! + Rn(x)

= 1 + x + x^2/2! + x^3/3! + ... + x^n/n! + Rn(x),

$_{n=0}$ $n!$ 2 6

泊松分布应用

泊松分布是常用的离散分布之一，现实世界中有很多随机变量都可直接用泊松分布描述，它们之间的差别表现在不同的 λ 上。下面是国内外文献上认可的服从或近似服从泊松分布的随机变量的一些例子：

(1)在一定时间内，电话总站接错电话的次数；
(2)在一定时间内，在超级市场排队等候付款的顾客人数；
(3)在一定时间内，来到车站等候公共汽车的人数；
(4)在一定时间内，某操作系统发生故障的次数；
(5)在一个稳定的团体内，活到 100 岁的人数；
(6)一匹布上，疵点的个数；
(7)100 页书上，错别字的个数；
(8)一个面包上，葡萄干的个数。

从这些例子可以看出，泊松分布总与计数过程相关联，并且计数是在一定时间内、或一定区域内、或一特定单位内的前提下进行的。下面我们详细地研究第 1 个例子。

例 2.2.12 一项研究表明：电话总站一天内接错电话号码的次数 X 是一个服从泊松分布 $P(\lambda)$ 的随机变量(简称泊松变量)。对某电话总站连续观察 100 天，共发现 320 只电话接错号码，具体数据按接错次数多少整理在表 2.2.4的前两列上。其中 x 表示一天内接错次数，n_x 表示接错次数为 x 的天数。由此可算得接错次数为 x 的频率为 $P^*(x)=n_x/100$，它放在第 3 列上。最后一列是按 $\lambda=3.2$ 算得的泊松概率 $p(x)=3.2^x e^{-3.2}/x!$。比较表上的最后两列可以看出：这组泊松概率 $p(x)$对这组观察频率 $p^*(x)$的拟合程度是很好的(在统计部分将进一步说明此种拟合程度)。

在实际中，人们常把在一次试验中出现概率很小(如小于 0.05)的事件称为**稀有事件**。由二项分布的泊松近似可以得到：n 重贝努里试验中稀有事件出现次数近似服从泊松分布。下面的例子说明这个现象。

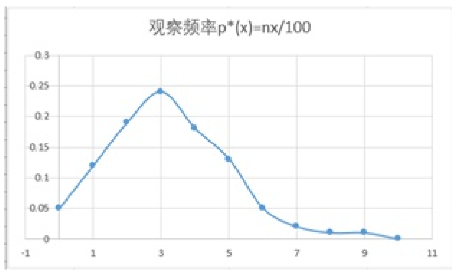
表 2.2.4 接错次数的观察频率与泊松概率

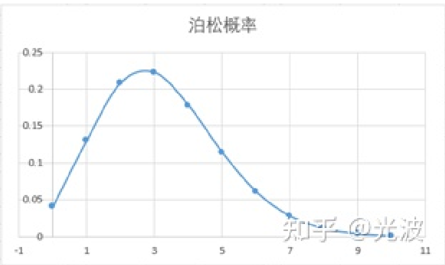
接错次数 x	观察天数 n_x	观察频率 $p^*(x)=n_x/100$	泊松概率 $p(x)=3.2^x e^{-3.2}/x!$
0	5	0.05	0.041
1	12	0.12	0.130
2	19	0.19	0.209
3	24	0.24	0.223
4	18	0.18	0.178
5	13	0.13	0.114
6	5	0.05	0.060
7	2	0.02	0.028
8	1	0.01	0.011
9	1	0.01	0.004
10	0	0.00	0.002
	100	1.00	1.000

泊松概率公式中唯一参数λ为加权平均值，先算出总错误次数为320，然后除以天数100，得到3.2。

	A	B	C	D	E	F	G	H
1	接错次数x	观察天数nx	观察频率p*(x)=nx/100	泊松概率		加权平均值λ	错误次数	
2	0	5	0.05	0.041		3.2	0	
3	1	12	0.12	0.130			12	
4	2	19	0.19	0.209			38	
5	3	24	0.24	0.223			72	
6	4	18	0.18	0.178			72	
7	5	13	0.13	0.114			65	
8	6	5	0.05	0.061			30	
9	7	2	0.02	0.028			14	
10	8	1	0.01	0.011			8	
11	9	1	0.01	0.004			9	
12	10	0	0	0.001			0	
13		100	1				320	总和

从下图可以可以形象得看出，观测频率与泊松分布的概率图，几乎一样，所以按错次数的天数分布，服从泊松分布。





n 台设备中同时发生故障的台数 X 服从二项分布 $b(n, 0.01)$，由于 $p=0.01$ 很小，故"每台设备发生故障"可看作稀有事件，故 X 又可近似看作服从泊松分布 $P(\lambda)$，其中 $\lambda=np=n\times 0.01$。下面用此看法来讨论几个问题。

(1)P 若用一名维修工负责维修 20 台设备，求设备发生故障而不能及时维修的概率是多少？

设 X_1 为 20 台设备中同时发生故障的台数，则 $X_1 \sim b(20, 0.01)$。由于稀有事件之故，可认为 $X_1 \dot{\sim} P(\lambda_1)$，其中 $\lambda_1=20\times 0.01=0.2$，这里符号 $\dot{\sim}$ 表示"近似服从"。于是 20 台设备中因故障得不到及时维修只在同时有 2 台和 2 台以上发生故障时才会出现。故所求概率

$$
\begin{aligned}
P(X_1 \geqslant 2) &= \sum_{x=2}^{\infty} \frac{0.2^x}{x!} e^{-0.2} \\
&= 1-e^{-0.2}-0.2e^{-0.2} \\
&= 0.0175
\end{aligned}
$$

这表明，一名维修工负责维修 20 台设备时，因同时发生故障得不到及时维修的概率不到 0.02。

(2)若用三名维修工负责维修 80 台设备，求设备发生故障而不能及时维修的概率是多少？

设 X_2 为 80 台设备中同时发生故障的台数，则 $X_2 \sim b(80, 0.01)$。类似地可认为，$X_2 \dot{\sim} P(\lambda_2)$，其中 $\lambda_2=80\times 0.01=0.8$，于是

$$P(X_2 \geqslant 4)=\sum_{k=4}^{\infty} \frac{0.8^k}{k!} e^{-0.8}=1-\sum_{k=0}^{3} \frac{0.8^k}{k!} e^{-0.8}$$

上述概率可以直接算得。这里可查泊松公布表获得，附表 2 对 x 为非负整数和 $\lambda=0.02(0.02)0.10(0.05)1.0(0.1)2.0(0.2)8(0.5)15(1)25$ 给出泊松分布函数值。

$$F(x)=P(X \leqslant x)=\sum_{k=0}^{x} \frac{\lambda^k}{k!} e^{-\lambda}$$

如 $\lambda=0.8, x=3$ 时，可查得

$$P(x_2 \leqslant 3)=\sum_{k=0}^{3} \frac{0.8^k}{k!} e^{-0.8}=0.991$$

利用这个值可算得

$$P(x_2 \geqslant 4)=1-P(X_2 \leqslant 3)=0.009$$

这表明，三名维修工负责维修 80 台设备时，因同时发生故障得不到及时维修的概率为 0.009，几乎为前面的 0.0175 的一半，提高了效率。

(3)若有 300 台设备，需要配多少名维修工，才能使得不到及时维修的概率不超过 0.01。

设 X_3 为 300 台设备中同时发生故障的台数，N 为所需配的维修工的人数，类似地可认为 $X_3 \dot{\sim} P(\lambda_3)$，$\lambda_3=300\times 0.01=3$。于是 N 应满足下列等式：

$$P(X_3 \geqslant N+1)=\sum_{k=N+1}^{\infty} \frac{3^k}{k!} e^{-3} \leqslant 0.01$$

或

$$\sum_{k=0}^{N} \frac{3^k}{k!} e^{-3} \geqslant 0.99$$

从附表 2 查得，当 $N=7$ 和 8 时，有

$$\sum_{k=0}^{7} \frac{3^k}{k!} e^{-3}=0.988$$

$$\sum_{k=0}^{8} \frac{3^k}{k!} e^{-3}=0.996$$

故 $N=8$ 时满足要求，即要用 8 名维修工才能使 300 台设备得不到及时维修的概率不超过 0.01。

可靠性工程中泊松分布应用

由于二项分布在实际计算中较为繁琐，因此希望能找到一个便于计算的近似公式。泊松分布被认为是当n为无限大时的二项分布的扩展，从上面的泊松公式概率公式的推导中可以看出，泊松公式是二项分布概率公式中n趋近于无穷大时的极限。事实上，在工程应用中，当n＞20，并且p≤0.05（发生概率5%，稀有事件）时，就可以用泊松分布近似表示二项分布。

泊松分布的表达式为：

$$P(X=r)=\frac{(n*p)^r}{r!} * e^{-(n*p)}=\frac{(\lambda)^r}{r!} * e^{-\lambda}$$

在与时间相关的可靠性度量函数中，需要计算成功的概率或可靠度时，常常把n换成时间t，p可以换成单位时间内的失效率Pλ，那么参数λ = t*Pλ。

例：控制台指示灯平均失效率为每小时0.001次。如果指示灯的失效数不能超过2个，该控制台指示灯工作500小时的可靠度是多少？

解：已知 $P=P_\lambda=0.001$，$n=t=500$ 小时，$n*p=0.5$，$r\leqslant 2$

发布于 2022-04-05 11:26

可靠性 概率论与数理统计 泊松分布

文章被以下专栏收录

可靠性数学
可靠性数据分析、寿命分布和数理统计等