哈爾濱工業大學

概率论与数理统计课程论文

题目：实际生活中泊松分布的应用

班　号	
学　号	
姓　名	
日　期	**2022.11.10**
成　绩	

摘 要

泊松分布在概率论中是很重要的分布，现实中很多情况都满足了这一分布。本文对泊松分布进行了简单的说明，并重点介绍了泊松分布在一些实际情况中的应用，揭示了泊松分布在解决实际问题中的重要性。有助于理解和把握分布。

关键词：泊松分布；二项分布；实际应用。

1 提出背景

泊松分布最早是由贝努利家族的一位成员描述的。1837 年，法国数学家泊松(Simeon-Denis Poisson, 1781~1840)在《论判断的概率》一文中提出了这种描述随机现象的常见分布，因此在概率论中，这种分布被称为泊松分布。这种分布在公用事业、放射性现象等方面有许多应用。

2 泊松分布

若随机变量 X 的分布列是 $P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$,λ >0 (k=0,1,2,⋯),则称 X 服从参数为 λ 的泊松分布，并用记号 X ~ P(λ)表示[1]。泊松分布的参数 λ 是单位时间（或单位面积）内随机事件的平均发生次数。泊松分布适合于描述单位时间内随机事件发生的次数。泊松分布的数学期望和方差均为 λ。

3 泊松分布与二项分布的关系

当二项分布的 n 很大而 p 很小时，泊松分布可作为二项分布的近似，其中 λ 为 np。通常当 n≥20,p≤0.05 时，就可以用泊松公式近似地计算。

4 研究意义

泊松分布被广泛应用。在实际应用中，许多随机现象服从泊松分布。例如，电话交换机接到的电话数量、去商店的顾客数量、到达机场的飞机数量、穿过天空某一区域的流星数量、纺纱机线路上断头的数量、放射性物质释放出的粒子数量等等，都服从泊松分布[1]

接下来，我们将结合一个具体的例子进一步理解泊松分布的研究意义表 1。例如，有一家包子铺，一周的日营业额如表 1 所示。每天准备几个馒头才能供应充足不浪费?

日期	周一	周二	周三	周四	周五
销售	3	7	4	6	5

首先可以计算出均值为 5，但显然 5 作为答案是不太合适的（40%的时间不够卖），此时我们需要一种更加可靠的估计方法。设馒头店每天的营业时间为 T，将 T 均分为足够大的 n 份，使得对于每一个时间段内，卖出一个馒头和没有卖出馒头的概率之和为 1，这样的话，T 时间内卖出 k 个馒头的概率通过二项分布来计算就是 $C_n^k p^k(1-p)^{n-k}$，为了得到更好的效果，我们对 n 取极限，得到 $\lim_{n\to\infty} C_n^k p^k(1-p)^{n-k}$。接下来是对 p 的求解，在上面的假设中，问题已经被转化为了二项分布。二项分布的期望为：E(X)=np=μ，则 $p=\frac{\mu}{n}$，故有：$\lim_{n\to\infty} C_n^k p^k(1-p)^{n-k}=\lim_{n\to\infty} C_n^k(\frac{\mu}{n})^k(1-\frac{\mu}{n})^{n-k}$，计算极限可得 $\lim_{n\to\infty} C_n^k(\frac{\mu}{n})^k\left(1-\frac{\mu}{n}\right)^{n-k}=\frac{\mu^k}{k!}e^{-\mu}$，

令 μ=λ，可得在 T 时间内卖出 k 个馒头的概率为：P（X=k)=$\frac{\lambda^k e^{-\lambda}}{k!}$，这就是泊松分布。其中 λ 可以用样本均值 5 来代替，故 P（X=k)=$\frac{5^k e^{-5}}{k!}$，画出概率密度函数曲线就是：

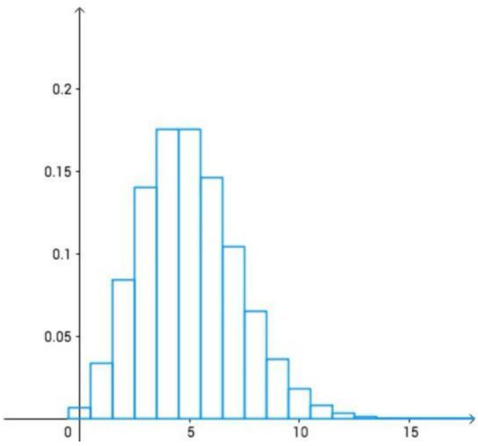

图 1

可以看到，如果每天准备 8 个馒头的话，那么足够卖的概率就是把前 8 个的概率加起来：

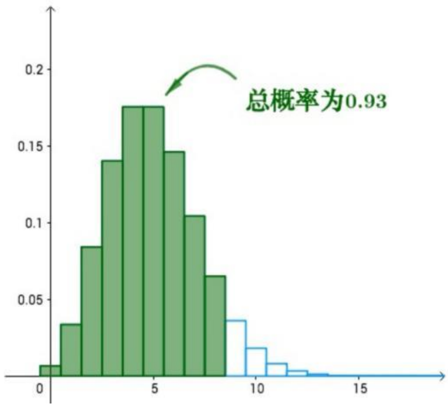

图 2

这样的话 93%的情况馒头够卖，因此把 8 作为这个问题的答案是比较合理的。

从这个例子中可以看出，泊松分布在解决许多实际问题中有很重要的研究意义。

5 实际应用

5.1 泊松分布在随机行为中的应用

学生公寓晚退行为是一种随机事件，具有较大的安全风险。为了便于管理，我们可以运用概率论的方法对学生晚归行为进行统计和预测。在大多数情况下，我们不能确切地预测我们的一些同学会晚归，但我们可以从 1 月或一个学期的迟交数据中预测，泊松分布更适合描述这些随机事件。学生公寓晚归行为具有以下特点：不会受之前晚回家学生的影响；除涉及假期的月份外，其他月份的晚归学生总数相对稳定，说明学生晚归的概率是稳定的。

从学生宿舍迟到回家是一种罕见的现象。这些特点使得学生公寓晚归行为满足以下三个条件：(1)随机事件具有独立性。(2)随机事件在不同时间段内发生的概率具有稳定性。(3)随机事件是个小概率事件[2]。因此可以将其视为泊松过程进行研究。

λ≈1.5，该大学共有 10000 名在校生。由于周末不单独发布晚归数据，每周五、周六和周日的晚归数据会下一周周一起发布，即每周一的公告实际包含了周五、周六和周日 3 天的数据。所以，为了补全公告未体现的数据，对这部分未体现的天数做了如下处理：根据 λ=1.5 得到泊松分布的概率分布，可认定某天晚上不发生晚归的概率为 P(0)=0.22，如果周末 3 天都未发生晚归（其概率为 0.223），则认定为这 3 天为无晚归。根据上述处理，补全了确实的数据，得到该学校 2013 年 9 月~2018 年 12 月以来的晚归次数的概率分布[3]。

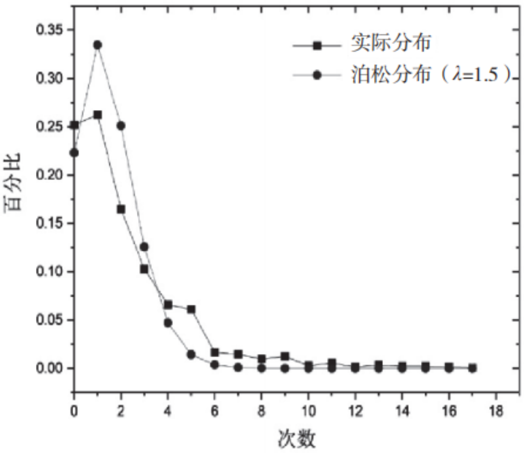

图 3

从图 3 中可以看出，泊松分布与实际分布趋势相似，且符合率较高，说明学生晚归次数的概率分布与泊松分布类似。因此，只要知道一段时间内晚归次数的平均值，我们就可以用对应的泊松分布来描述晚归次数的分布，将未知的概率分布转化为已知的概率分布。

5.2 流行性疾病预警

设某种流行性疾病在某一周期内的平均发病水平（均值）为 λ，则可以根据这一平均水平计算该种疾病发生 x 例病例时的概率，如果 P{X=k}的概率大于给定的 α（通常取 0.05）水准，并且 P(X=k+1)的概率小于给定的 α 水准时，由于此时出现 X=k+l 为小概率事件，认为其在通常情况下不会发生，此时 k 即为发生与不发生的临界值（也称预警值）。例如，某市 2012 年某流行性疾病的发病数如表 2 所示，求其发病预警线。

表 2

周次	例数	周次	例数	周次	例数	周次	例数	周次	例数	周次	例数
1	76	10	60	19	61	28	34	37	22	46	65
2	67	11	59	20	62	29	32	38	13	47	54
3	70	12	64	21	52	30	29	39	16	48	75
4	69	13	61	22	51	31	21	40	24	49	81
5	65	14	82	23	49	32	19	41	18	50	75
6	59	15	75	24	50	33	17	42	12	51	80
7	72	16	70	25	40	34	21	43	20	52	68
8	64	17	58	26	38	35	12	44	38	53	72
9	72	18	57	27	61	36	15	45	49	54	59

以 2012 年每个观察周的数据求样本均值，将其作为总体均值 λ，并把该数据作为之后时期的预测值，利用 Poisson 分布原理制作预警线。计算各周的均数 λ，经计算知，λ=53。然后，根据 λ 的值，利用泊松分布的计算公式计算 P{X=k}的值，其中 k=0,1,2,……计算后可得：P{X=0}=0，P{X=1}=0，P{X=2}=0，……P{X=55}=0.0509，P{X=56}=0.0483，……经计算后可得预警值为 k=55。根据此预警值 55 对下一年的该流行性疾病发病水平制作预警图，如图 4。

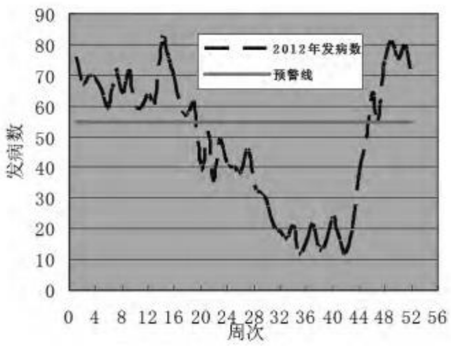

图 4

从图 4 可以看出，利用 Poisson 布原理制作的预警线为一条水平的直线，而实际生活中流行性疾病具有季节性，误差比较大。为了提高精确度，缩小误差，将 8 周作为一个周期，利用 Poisson 分布原理算出每个周期的预警值并且制作出分段警线，结果见图 5。

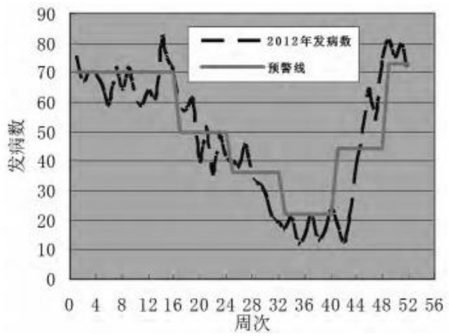

图 5

Poisson 分布预警原理，是基于疾病实际发病水平的预警，以往年的发病情况来制作预警线，可作为来年疾病水平的一个估计。这种预警方法比较简单，为此类流行性疾病的预警提供了一种可行的探讨模型[4]。

6 小结

泊松分布是一种在现实生活中应用非常广泛的分布，本文略述了很少一部分细节，简要介绍了泊松分布在几个方面的应用，体现了它在解决实际问题中的重要意义。如果有任何不当，请进行批评要求。

参考文献

[1]方茹,周永春,李朝艳,田波平,王勇.大学数学 概率论与数理统计 第二版[M].北京:高等教育出版社.2014.

[2]夏元睿,吴俊,叶冬青.泊松分布与概率论的发展——西蒙・丹尼尔・泊松[J].中华疾病控制杂志,2019,23(07):881-884.

[3]叶超.泊松分布在随机行为中的应用研究[J].中国新技术新产品,2020(13):125-126.

[4]蔡秋娥,段小华.基于泊松分布的流行性腮腺炎疾病预警的应用研究[J].科技创新与应用,2016(29):33-34.