概率论与数理统计
第二章 随机变量及其概率分布

金玲飞

复旦大学计算机学院
Email: lfjin@fudan.edu.cn

2018.10.16

例子 (2.2.4)

以往的病史资料表明，在患感冒的人群中因感冒而最终死亡的比例为*0.2%*。试求，目前正在患感冒的*1000*个病人中：

- 最终恰有*4*个人因感冒而死亡的概率
- 最终因感冒而死亡的人数不超过*50*人的概率

例子 (2.2.4)

以往的病史资料表明，在患感冒的人群中因感冒而最终死亡的比例为*0.2%*。试求，目前正在患感冒的*1000*个病人中：

- 最终恰有*4*个人因感冒而死亡的概率
- 最终因感冒而死亡的人数不超过*50*人的概率

解： X : 1000个感冒中因感冒死亡的人数，
则$X \sim B(1000, 0.002)$。

- $P(X = 4) = \binom{1000}{4} 0.002^4 0.998^{996}$
- $P(X \leq 50) = \sum_{k=0}^{50} \binom{1000}{k} 0.002^k 0.998^{1000-k}$

泊松定理

二项分布的近似估计？

定义 (2.2.1（Poisson定理）)

设对每个自然数n，$0 < p_n < 1$。若存在极限 $\lim_{n\to\infty} np_n = \lambda$，则对每个非负整数$k$，有

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

记 $p(k;\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \ldots$

泊松定理

二项分布的近似估计？

定义 (2.2.1（Poisson定理）)

设对每个自然数n，$0 < p_n < 1$。若存在极限$\lim_{n\to\infty} np_n = \lambda$，则对每个非负整数$k$，有

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

记 $p(k;\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, 2, \ldots$

- 适用于n较大，p很小，np 大小适中。

二项分布的泊松逼近

- 当 $p(p \leq 0.05)$ 很小，n 较大，

$$b(k;n,p) \approx p(k;np) = \frac{(np)^k}{k!} e^{-np}$$

二项分布的泊松逼近

- 当 $p(p \leq 0.05)$ 很小，n 较大，

$$b(k; n, p) \approx p(k; np) = \frac{(np)^k}{k!} e^{-np}$$

- 当 p 很大，n 较大，

二项分布的泊松逼近

- 当 $p(p \le 0.05)$ 很小，n 较大，

$$b(k;n,p) \approx p(k;np) = \frac{(np)^k}{k!}e^{-np}$$

- 当 p 很大，n 较大，

$$b(k;n,p) = b(n-k;n,1-p) \approx p(n-k,n(1-p))$$
$$= \frac{(n(1-p))^{n-k}}{(n-k)!}e^{-n(1-p)}$$

若p不大不小怎么办？ 正态逼近！

表 2.4.3 二项分布与泊松近似的比较

k	二项分布 按$\binom{n}{k}p^k(1-p)^{n-k}$计算				泊松近似 按$\frac{(np)^k}{k!}e^{-np}$计算
	$n=10$ $p=0.1$	$n=20$ $p=0.05$	$n=40$ $p=0.025$	$n=100$ $p=0.01$	$\lambda=np=1$
0	0. 349	0. 358	0. 363	0. 366	0. 368
1	0. 385	0. 377	0. 372	0. 370	0. 368
2	0. 194	0. 189	0. 186	0. 185	0. 184
3	0. 057	0. 060	0. 060	0. 061	0. 061
4	0. 011	0. 013	0. 014	0. 015	0. 015
>4	0. 004	0. 003	0. 005	0. 003	0. 004

例子 (2.2.6)

设某种数字传输系统每秒传送512×10^3个*0*或*1*，由于干扰，传送中会出现误码，即将*0*误传为*1*，或将*1*误传为*0*。设误码率为10^{-7}，求在*10s* 内出现*1*个误码的概率。

例子 (2.2.6)

设某种数字传输系统每秒传送512×10^3个*0*或*1*，由于干扰，传送中会出现误码，即将*0*误传为*1*，或将*1*误传为*0*。设误码率为10^{-7}， 求在*10s* 内出现*1*个误码的概率。

例子 (2.2.7)

某电话总机有*300*个用户，但只有*8*条线路可供打进电话，在每个时刻各用户通话与否相互独立，各用户通话概率均为$\frac{1}{60}$。求在某给定 时刻有用户打不进电话的概率。

例子 (2.2.6)

设某种数字传输系统每秒传送512×10^3个*0*或*1*，由于干扰，传送中会出现误码，即将*0*误传为*1*，或将*1*误传为*0*。设误码率为10^{-7}， 求在*10s* 内出现*1*个误码的概率。

例子 (2.2.7)

某电话总机有*300*个用户，但只有*8*条线路可供打进电话，在每个时刻各用户通话与否相互独立，各用户通话概率均为$\frac{1}{60}$。求在某给定 时刻有用户打不进电话的概率。

解： X表示该时刻打电话的用户数，$X \sim B(300, \frac{1}{60})$。

$$P(X > 8) = \sum_{k=9}^{300} P(X = k) = \sum_{k=9}^{300} b(k; 300, \frac{1}{60}) \approx \sum_{k=9}^{\infty} p(k; 5)$$

例子 (2.2.8)

设*200*台同类型设备的工作是相互独立的，每台发生故障的概率均为*0.01*，一台设备的故障可由一人处理。
分如下两种情况计算设备发生故障得不到及时处理的概率。

- 由*5*人来维护这些设备，每人负责*40*台固定的设备。
- 由*4*人共同维护这*200*台设备。

例子 (2.2.8)

设*200*台同类型设备的工作是相互独立的，每台发生故障的概率均为*0.01*，一台设备的故障可由一人处理。
分如下两种情况计算设备发生故障得不到及时处理的概率。

- 由*5*人来维护这些设备，每人负责*40*台固定的设备。
- 由*4*人共同维护这*200*台设备。

答案见课本Page 50。

- 概率分析有助于有效的调节人力和物力。

泊松分布

Poisson分布（译名有泊松分布、普阿松分布、卜瓦松分布、布瓦松分布、布阿松分布等），是一种统计与概率学里常见到的离散概率分布，由法国数学家(Siméon-Denis Poisson)在1838年时发表。

定义 (Poisson distribution)

设X是概率空间$(\Omega, \mathscr{F}, P)$上的随机变量，如果X的取值范围是$0, 1, 2, \ldots$，且其分布律为

$$p(k; \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, \ldots$$

称X服从以λ为参数的泊松分布，记为$X \sim \mathscr{P}(\lambda)$。

例子 (2.2.9)

由一家商店过去的销售记录知，某种商品每月销售数X可以用参数$\lambda = 10$的泊松分布来描述。为了以*95%*以上的把握保证下月该商品不脱 销，问商店在月底至少应进该种商品多少件？

例子 (2.2.9)

由一家商店过去的销售记录知，某种商品每月销售数X可以用参数$\lambda = 10$的泊松分布来描述。为了以*95%*以上的把握保证下月该商品不脱 销，问商店在月底至少应进该种商品多少件？

解：设进货a件，则要求$P(X \leq a) \geq 0.95$。
因$X \sim \mathscr{P}(10)$，

$$\sum_{k=0}^{a} \frac{10^k}{k!} e^{-10} \geq 0.95$$

求得$a = 15$。

例子 (2.2.10)

实验室器皿产生甲、乙两类细菌的机会是相等的，且产生的细菌总数服从参数为λ的泊松分布，试求

- 产生了甲类细菌但没有乙类细菌的概率
- 在已知产生了细菌而且没有甲类细菌的条件下，有两个乙类细菌的概率。

例子 (2.2.10)

实验室器皿产生甲、乙两类细菌的机会是相等的，且产生的细菌总数服从参数为λ的泊松分布，试求

- 产生了甲类细菌但没有乙类细菌的概率
- 在已知产生了细菌而且没有甲类细菌的条件下，有两个乙类细菌的概率。

见课本Page 53。

超几何分布

设对某批N件产品进行不放回抽样检查，若这批产品中有M件次品，现从这批产品中随机地抽取n件，以X表示这n件中的次品数，则X服从超几何分布。

超几何分布

设对某批N件产品进行不放回抽样检查，若这批产品中有M 件次品，现从这批产品中随机地抽取n 件，以X表示这n 件中的次品数，则X 服从超几何分布。

定义 (2.2.4 Hypergeometric distribution)

设X是概率空间$(\Omega, \mathcal{F}, p)$上的随机变量，N, M, n是自然数，$M < N, n \leq N, \ell = \min(M, n)$。如果$X$的取值范围是$0, 1, \ldots, \ell$，且其分布律为

$$p_k = P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, k = 0, 1, 2, \ldots, \ell$$

则称X服从超几何分布。

描述了不放回的抽样的统计规律。

几何分布

在事件A发生的概率为p的多重伯努利试验中，若以X记A首次发生时的试验次数，则X服从参数为p的几何分布。

几何分布

在事件A发生的概率为p的多重伯努利试验中，若以X记A首次发生时的试验次数，则X服从参数为p的几何分布。

定义 (2.2.5 Geometric distribution)

设X是概率空间$(\Omega, \mathcal{F}, p)$上的随机变量，$0 < p < 1, q = 1 - p$。如果X的取值范围是$1, 2, 3, \ldots$，且其分布律为

$$p_k = P(X = k) = pq^{k-1}, k = 1, 2, 3, \ldots$$

则称X服从参数为p的几何分布。

几何分布具有无记忆性。

$$P(X=k+m|X>m)=P(X=k), k, m=1,2,\ldots$$

注：在离散型分布中，只有几何分布是无记忆的。

负二项分布,Pascal 分布

负二项分布：重复进行n次独立试验，成功的概率为p，第r次成功出现在第k次试验的概率

$$
\begin{aligned}
f(k;r,p) &= \binom{k-1}{r-1} p^{r-1}(1-p)^{k-1-(r-1)} \cdot p && (1)\\
&= \binom{k-1}{r-1} p^{r}(1-p)^{k-r} && (2)
\end{aligned}
$$

负二项分布

例子 (巴拿赫火柴问题–Banach's match problem)

某数学家有两盒火柴，每盒都有n根。每次使用，都随机的从中抽取一根。问他发现一盒空而另一盒还有$r(0 \leq r \leq n)$根的概率多少？

负二项分布

例子 (巴拿赫火柴问题–Banach's match problem)

某数学家有两盒火柴，每盒都有n根。每次使用，都随机的从中抽取一根。问他发现一盒空而另一盒还有$r(0 \le r \le n)$根的概率多少？

记 A：取左边口袋火柴；$\overline{A}$:取右边口袋火柴

$$P(A) = P(\overline{A}) = 1/2$$

负二项分布

例子 (巴拿赫火柴问题–Banach's match problem)

某数学家有两盒火柴，每盒都有n根。每次使用，都随机的从中抽取一根。问他发现一盒空而另一盒还有$r(0 \le r \le n)$根的概率多少？

记 A：取左边口袋火柴；$\overline{A}$:取右边口袋火柴

$$P(A) = P(\overline{A}) = 1/2$$

若巴拿赫首次发现他左衣袋中的一盒火柴变空，这时事件A已经是第$n+1$次发生，而此时他右边衣袋中火柴盒中恰剩r根火柴相当于他在此前已在右衣袋中取走了$n-r$ 根火柴。即在$2n-r+1$次试验中，第$2n-r+1$次A发生。

$$f(2n-r+1; n+1, \frac{1}{2}) = \binom{2n-r}{n}(\frac{1}{2})^{2n-r+1}$$

例子

设一个人一年内患感冒的次数服从参数$\lambda = 5$的泊松分布。现有某种预防感冒的药对75%的人有效（能将泊松分布的参数降低为$\lambda = 3$），对另外25%的人不起作用。如某人服用了此药，一年内患了两次感冒，那该药对此人有效的可能性多大？

例子

设一个人一年内患感冒的次数服从参数$\lambda = 5$的泊松分布。现有某种预防感冒的药对75%的人有效（能将泊松分布的参数降低为$\lambda = 3$），对另外25%的人不起作用。如某人服用了此药，一年内患了两次感冒，那该药对此人有效的可能性多大?

A:服用此药后，一年患两次感冒； B:服用此药后有效

$$P(A) = P(B)P(A|B)+P(\overline{B})P(A|\overline{B}) = 0.75*\frac{3^2}{2}e^{-3}+0.25\frac{5^2}{2}e^{-5}$$

$$P(B|A) = 0.889$$

2.3 连续型随机变量

连续型随机变量

定义 (2.3.1 continuous random variable)

设X是概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，$F(x)$为其分布函数，如果存在定义在$(-\infty, \infty)$上的非负实值函数 $f(x)$，使得

$$F(x) = \int_{-\infty}^{x} f(y)dy, -\infty < x < \infty$$

则称X为连续型随机变量。称$F(x)$为连续型分布函数，称$f(x)$为X的概率密度函数，简称概率密度（或分布函数）。

概率密度必须满足的两个条件：

$$f(x) \geq 0, -\infty < x < \infty$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

概率密度必须满足的两个条件：

$$f(x) \geq 0, -\infty < x < \infty$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- $f(x)$表示X在一些地方取值机会大，另一些地方取值机会小。

概率密度必须满足的两个条件：

$$f(x) \geq 0, -\infty < x < \infty$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- $f(x)$表示X在一些地方取值机会大，另一些地方取值机会小。
- 连续型随机变量的分布函数是连续的。
- $F(x)$连续不代表它是连续型分布函数。

定理 (2.3.1)

设X为连续型随机变量，$F(x)$与$f(x)$分别为其分布函数和概率密度。

- 对任意常数$a < b$，有

$$P(a < X \leq b) = \int_a^b f(x)dx$$

- $F(x)$是连续函数，且当$f(x)$在$x = x_0$点连续时，$F'(x_0) = f(x_0)$；

- 对任意常数c，$P(X=c)=0$，从而对任何$a<b$，

$$\begin{aligned} & P(a<X<b)=P(a<X\le b) \\ =& P(a\le X<b)=P(a\le X\le b) \end{aligned}$$

- $P(X\in\Omega-\{c\})=1$.

推论

对实轴上任意“可求长”集合D

$$P(X \in D) = \int_D f(x)dx$$

注：掌握了概率密度，即掌握了它的统计规律。

推论

对实轴上任意“可求长”集合D

$$P(X \in D) = \int_D f(x)dx$$

注： 掌握了概率密度，即掌握了它的统计规律。

- 概率密度函数不唯一：可在任何有限多个点上任意改变概率密度函数的值（改变后非负），所得的函数仍为同一随机变量的概率密度函数。

怎么判断一个随机变量是连续型？

定理 (2.3.2)

设随机变量X的分布函数$F(x)$是连续的，且除有限个点$(c_1 < c_2 < \cdots < c_n)$外，导数$F'(x)$存在且连续，则$X$是连续型随机变 量，它具有如下的概率密度：

$$f(x) = \begin{cases} F'(x), & x \notin \{c_1, \ldots, c_n\} \\ 0, & x \in \{c_1, \ldots, c_n\} \end{cases}$$

离散型随机变量VS 连续性随机变量

例子 (2.3.1)

已知随机变量X的概率密度形如

$$f(x)=\begin{cases} ax+b, & 1<x<3 \\ 0, & \text{其他} \end{cases}$$

且$P(2<X<3)=2P(1<X<2)$。求常数a,b。

例子

已知随机变量X的概率密度形如

$$f(x) = \begin{cases} x, & 0 \le x < 1 \\ 2 - x, & 1 \le x < 2 \\ 0, & \text{其他} \end{cases}$$

求$F(x)$.

均匀分布

定义 (Uniform distribution)

设X为概率空间$(\Omega, \mathcal{F}, P)$上的连续型随机变量。若X的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其他} \end{cases}$$

称X服从区间$[a, b]$上的均匀分布， 记为$X \sim U[a, b]$。

均匀分布

定义 (Uniform distribution)

设X为概率空间$(\Omega, \mathcal{F}, P)$上的连续型随机变量。若X的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其他} \end{cases}$$

称X服从区间$[a, b]$上的均匀分布，记为$X \sim U[a, b]$。

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

均匀分布

下图给出了均匀分布的概率密度和分布函数的图像.

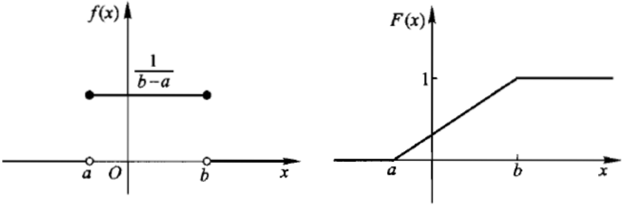

图 2.7 $[a, b]$ 上的均匀分布

例子 (2.3.4)

设随机变量$X \sim U[2,5]$。现在对X进行三次独立观测，求至少有两次观测值大于3的概率。

例子 (2.3.5)

设随机变量$\xi \sim U[1,6]$。求x的二次方程$x^2+\xi x+1=0$有实根的概率。

例子 (2.3.5)

设随机变量$\xi \sim U[1,6]$。求x的二次方程$x^2+\xi x+1=0$有实根的概率。

解. 由题设, ξ 的概率密度为

$$f(x)=\begin{cases}\frac{1}{5}, & 1\leqslant x\leqslant 6\\ 0, & \text{其他}\end{cases}$$

由于方程有实根的充分必要条件是判别式

$$\Delta=\xi^2-4\geqslant 0$$

而

$$\begin{aligned}P(\xi^2-4\geqslant 0) &= P(|\xi|\geqslant 2)\\ &= P(\xi\leqslant -2)+P(\xi\geqslant 2)\\ &= \int_{-\infty}^{-2} f(x)\mathrm{d}x+\int_{2}^{\infty} f(x)\mathrm{d}x\\ &= \int_{2}^{6}\frac{1}{5}\mathrm{d}x\end{aligned}$$

指数分布

定义 (Exponential distribution)

设X为概率空间$(\Omega, \mathcal{F}, P)$上的连续型随机变量。若X的概率密度函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, \text{其他} \end{cases}$$

（其中$\lambda > 0$为常数），则称X服从参数为λ的指数分布，记为$X \sim \mathscr{E}(\lambda)$。

指数分布

定义 (Exponential distribution)

设X为概率空间$(\Omega, \mathcal{F}, P)$上的连续型随机变量。若X的概率密度函数为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, \text{其他} \end{cases}$$

（其中$\lambda > 0$为常数），则称X服从参数为λ的指数分布，记为$X \sim \mathscr{E}(\lambda)$。

- 指数分布的确是分布。

指数分布函数

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

指数分布函数

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

应用：指数分布在排队论和可靠性理论中有着广泛的应用，常做各种寿命分布的近似。

指数分布函数

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

应用：指数分布在排队论和可靠性理论中有着广泛的应用，常做各种寿命分布的近似。

指数分布的无记忆性

$$P(X > s + t | X > s) = P(X > t)$$

例子 (2.3.7 指数分布与泊松分布的关系)

设某设备在任何长为t的时间内发生故障的次数$N(t)$服从参数为λt的泊松分布。

- 求相继两次故障之间的时间间隔T的概率分布
- 求在设备已经无故障工作$8h$的情况下，再无故障运行$8h$的概率Q

例子 (2.3.7 指数分布与泊松分布的关系)

设某设备在任何长为t的时间内发生故障的次数$N(t)$服从参数为λt的泊松分布。

- 求相继两次故障之间的时间间隔T的概率分布
- 求在设备已经无故障工作$8h$的情况下，再无故障运行$8h$的概率Q

解：(1)

- $t < 0$, $F(t) = P(T \leq t) = 0$.
- $t \geq 0$, $\{T > t\}$ 与 $\{N(t) = 0\}$ 相等，

$$F(t) = 1 - P(T > t) = 1 - P(N(t) = 0) = 1 - e^{-\lambda t}$$

(2) $Q = P(T \geq 8 + 8 | T \geq 8) = P(T \geq 8) = e^{-8\lambda}$.