第2章 随机变量及其分布

Chpt. 2 Random Variables & Probability Distributions

概率研究中的Case Study方法是否能解决所有问题？

前已介绍一些概率模型、很多例子，都有代表性；

但是可以找出更多的例子进行研究，永无止境！

这种方法属于Case Study。

存在问题：不同的case之间有无本质、统一的东西？

进一步，不同的概率现象之间有什么样的联系？

2.1 随机变量

2.1.1 引言

那么，概率函数是否可用?

概率函数定义在样本空间S上， $\forall A \in S$有$P(A)$与之对应；但由于样本空间的多样性，导致样本描述的多样性，使抽取其共性变得困难。

对于数值函数我们已经有足够多的研究成果

自然希望把一般的概率问题 $\longrightarrow$ 数值函数问题。

事件空间 $\longleftrightarrow$ 实数集合

研究事件 $\longrightarrow$ 研究数值

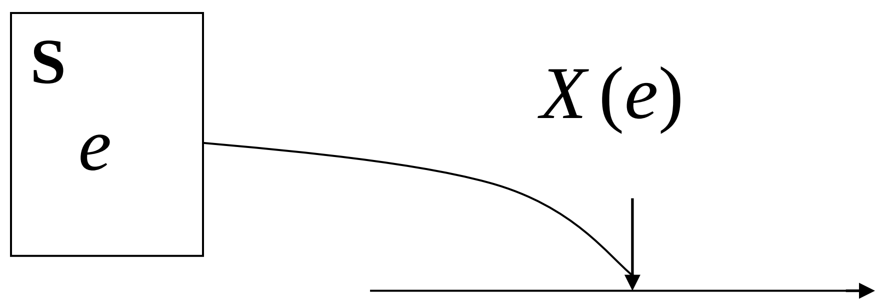
2.1 随机变量

2.1.1 随机变量的定义

随机试验E，其样本空间为S={e}，如果对每个e∈S ，有一个实数X(e)与之对应，就得到一个定义在S上的单值实函数X=X(e)，它是一个因变量，且因e的随机性而具有随机性，称为随机变量。

$$S \xrightarrow{X(\cdot)} R_X$$

$$X\text{：}\ e \in S \to X(e) \in R_X$$



2.1.1 随机变量的定义

通过随机变量可以描述事件，随机变量本身取值有一定的概率，区别于其他普通函数。对于一个实数集合L，随机变量X，$\{X\in L\}$表示的是一个事件A：

$$\{X \in L\} \Leftrightarrow \{e \mid X(e) \in L\} = A$$

因此，

$$P\{X \in L\} = P\{e \mid X(e) \in L\}$$

Example 2.1 某射手每次射击命中率为0.02，现在不断地射击，直到命中目标为止。

可以定义“ X＝命中目标所需的射击次数”

$$R_x = \{1,2,\cdots\}$$

$$L_1 = \{100\}$$

$$\{X \in L_1\} \Leftrightarrow \{e \mid \text{射击到第100次命中}\}$$

$$L_2 = \{1,2,\cdots 100\}$$

$$\{X \in L_2\} \Leftrightarrow \{e \mid \text{不多于100次命中目标}\}$$

有些试验结果本身就是数量描述的，其结果直接构成了随机变量。

例如：高炮在一定的条件下射击，用ρ表示弹着点与目标之间的距离，ρ就是随机变量；如以目标点为原点，建立直角坐标系，则弹着点(x,y)中的x、y均是随机变量，(x,y)本身是一个二维随机变量。

2.1.1 随机变量的定义

Example 2.2 盒子中有N个白球、M个黑球，从中抽取k个($k \le M+N$)，则以下变量为随机变量

X＝取得的白球数

Y＝取得的黑球数

Z＝两种球个数的差

Example 2.3 某公共汽车站每5分钟有一辆汽车通过，如果某人到达该车站的时刻是随机的。那么，一般来说，他等车的时间X是一个随机变量。

注意：此处随机变量X的取值是一个连续区间[0，5]

2.1.1 随机变量的定义

Example 2.2 盒子中有N个白球、M个黑球，从中抽取k个($k \le M+N$)，则以下变量为随机变量

X＝取得的白球数

Y＝取得的黑球数

Z＝两种球个数的差

$$P(X \in \{1\})$$

$$P(Y \in \{1,2,3,4,5\})$$

Example 2.3 某公共汽车站每5分钟有一辆汽车通过，如果某人到达该车站的时刻是随机的。那么，一般来说，他等车的时间X是一个随机变量。

$$P(2 < X \le 4)$$

注意：此处随机变量X的取值是一个连续区间[0，5]

2.2 离散型随机变量及其分布

2.2.1 离散型随机变量的定义

若随机变量X可能取的值为有限个或可列无限个， 则称X为离散型随机变量(discrete random variable)。

Example 2.4

[1] 上面例子中的射击命中次数X；

[2] 在[0，1]区间上方随机抛球，观察落到有理点X上的情形。

有理点为：

0, 1

1/2

1/3, 2/3

1/4, 2/4, 3/4

………

2.2 离散型随机变量及其分布

2.2.2 概率分布

对离散型随机变量涉及到两点：[1] 随机变量所有可能的取值；[2] 取每个值得概率。

如果X所有可能的取值为x_k (k=1,2,…)，X取 x_k (即事件 A={X=x_k}) 的概率记为

$$P(X=x_k)=p_k(k=1,2,\cdots)$$

称之为离散随机变量X的<u>概率分布或分布列</u>。

其直观的表示是列表

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p(x_1) & p(x_2) & \cdots & p(x_n) & \cdots \end{pmatrix}$$

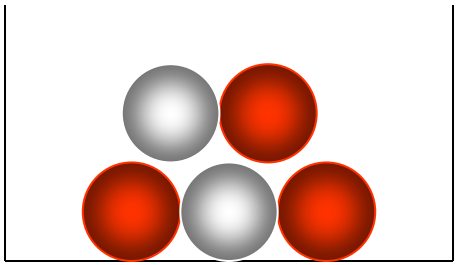
2.2.2 概率分布

Example 2.5 如右图所示，从盒中任取3个球。取到的白球数X是一个随机变量。*X*可能取的值是0,1,2。取每个值的概率为:

$$P(X=0)=\frac{C_3^3}{C_5^3}=\frac{1}{10}$$

$$P(X=1)=\frac{C_3^2C_2^1}{C_5^3}=\frac{6}{10}$$

$$P(X=2)=\frac{C_3^1C_2^2}{C_5^3}=\frac{3}{10}$$



其分布列为

X	0	1	2
p_i	0.1	0.6	0.3

Example 2.6 设随机变量X的分布列为

$$\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \dfrac{a-1}{4} & \dfrac{a+1}{4} & 0.1 & 0.2 & 0.1 \end{pmatrix}$$

求$P\{-1\le X\le 2\}$

解 (1) 由 $\dfrac{a-1}{4}+\dfrac{a+1}{4}+0.1+0.2+0.1=1$ 解得 $a=1.2$

故分布列为 $\begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 0.05 & 0.55 & 0.1 & 0.2 & 0.1 \end{pmatrix}$

(2) $P\{-1\le X\le 2\}=\sum\limits_{-1\le X\le 2} p(x_i)$

$= 0.55+0.1+0.2+0.1$

$= 0.95$

2.2.2 概率分布

Example 2.7 巴斯卡分布 在伯努里试验中，每次成功的概率为p。 记直至得到第r次成功时的试验次数为X，求X的分布列。

解 P{X＝k}＝P{前k-1次试验中有r-1次成功，有k-r次不成功，且第k次成功}

$$= C_{k-1}^{r-1} p^{r-1} (1-p)^{k-r} p$$

$$= C_{k-1}^{r-1} p^{r} (1-p)^{k-r}$$

$$(k = r, r+1, r+2, \cdots\cdots)$$

2.2.3 常见的概率分布

（一）离散均匀分布

设随机变量取值共有n个，且每个取值的概率相同，即

$$P(X=k)=\frac{1}{n}, k=1,2,\cdots,n$$

则称随机变量X服从离散均匀分布。(古典概型)

特殊地，设随机变量X只取一个常数值c， 即 P(X＝c)＝1，称它为退化(degenerate)分布，又称为单点分布。

事实上，{X＝c}是一个概率为1的事件，X可以看作一个常数，但有时我们宁愿把它看作（退化的）随机变量。

（二）0-1分布（两点分布/Bernoulli分布）

随机变量X只取两个值0和1，满足分布：

$$P(X=0)=1-p \qquad 0<p<1$$

$$P(X=1)=p$$

称X服从参数为p的 0-1分布（也称伯努里分布）。

2.2.3 常见的概率分布

Example 2.8 2000件产品，有1990件合格品，10 件次品。从中随机抽一件，规定e_1＝{取得的是合格品}，e_2＝{取得的是次品}

定义随机变量 $$X(e_1)=0, X(e_2)=1$$

$$p=10/2000=0.05$$

X服从参数为0.05的0-1分布，如定义。

[1] 伯努里试验具有广泛性：电路‘断’与‘不断’，产品‘合格’与‘不合格’，种子‘发芽’与‘不发芽’，掷硬币得‘正面’与‘反面’，…

[2] 任一伯努里试验的结果都可用伯努里分布描述

2.2.3 常见的概率分布

（三）二项分布

对于n重伯努里（Bernoulli）试验，假设每次成功的概率是p，定义随机变量X描述n次试验中事件A可能发生的次数k

$$P\{X=k\}=b(k,n,p)$$

$$=C_n^k p^k(1-p)^{n-k} \qquad (k=0,1,2,\cdots n)$$

称X服从参数为n和p的二项分布(Binomial distribution)，记为X~b(n,p)或X~B(n,p).

（三）二项分布

二项分布是概率论中最重要的分布之一，应用很广

(1)考察某地n个人是否患某种非流行性疾病，患病人数X服从二项分布.
检查一人是否患某种非流行性疾病是一次伯努里试验，各人是否生这病可认为相互独立，并可近似认为患病的概率p相等。

(2) 保险公司对某种灾害 (自行车被盗，火灾，...) 保险
各人发生此种灾害与否可认为相互独立，并假定概率相等。设一年间一人发生此种灾害的概率为p， 则在参加此种保险的n人中发生此种灾害的人数X服从二项分布。

（三）二项分布

(3) 碰运气能否通过英语四级考试

$$P\{X = k\} = b(k, 85, 0.25)$$

$$P\{X \geq 51\} = b(51, 85, 0.25) + b(52, 85, 0.25) + \cdots + b(85, 85, 0.25)$$

$$\approx 8.74 \times 10^{-12}$$

(4) 机房内n台同类计算机，在一年内每台损坏的概率为p，则在一年时间内损坏的机器数X服从二项分布。

如此，我们要问这样的分布具有什么样的性质呢？

比如：我们看看一年内最多会坏多少台机器？

为保证工作，最多备多少台机器？

平均会坏多少？一般备多少台机器（平均）？

2.2.3 常见的概率分布

二项分布的重要性质

[1] $$b(k,n,p)=b(n-k,n,1-p)$$

这从二项分布的概率以及 $C_n^k=C_n^{n-k}$ 立即可得。

也可以这样理解：n次试验中，事件{k次成功}与事件{n-k次不成功}是同样的，而{n-k次不成功}的概率即为

$$b(n-k,n,1-p)$$

很多情况下二项分布的计算很复杂，有时备有相应的计算表格，但只限于$p\le 0.5$的情况，当$p>0.5$时就可利用上式来计算。

[2] 增减性以及最可能成功（发生）次数

对固定的n、p， 由于

$$\frac{b(k,n,p)}{b(k-1,n,p)}=\frac{C_n^k p^k (1-p)^{n-k}}{C_n^{k-1} p^{k-1} (1-p)^{n-k+1}}$$

$$=\frac{(n-k+1)p}{k(1-p)}$$

$$=1+\frac{(n+1)p-k}{k(1-p)}$$

上式是否大于1，主要看(n+1)p－k的正负，或者说看(n+1)p与k的大小问题。

当 $k<(n+1)p$ 时，$\dfrac{b(k,n,p)}{b(k-1,n,p)}>1$，$b(k,n,p)$单调递增；

当 $k>(n+1)p$ 时，$\dfrac{b(k,n,p)}{b(k-1,n,p)}<1$，$b(k,n,p)$单调递减；

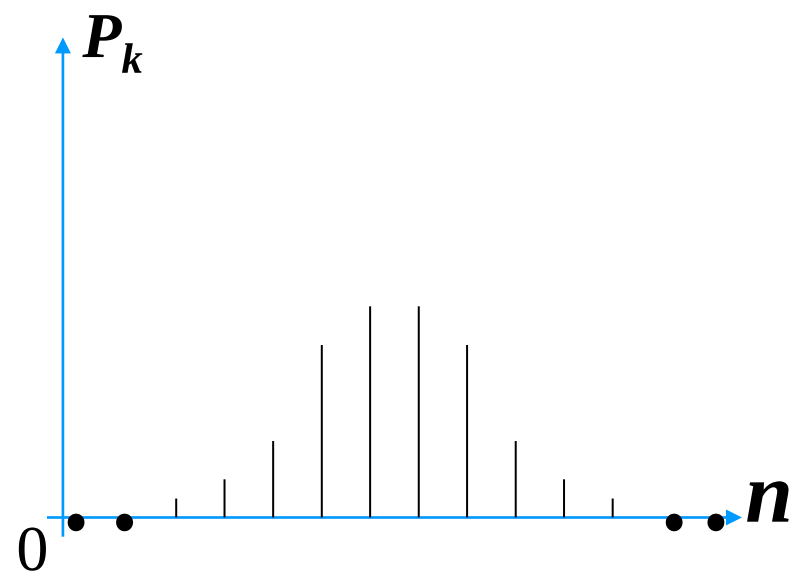
[2] 增减性以及最可能成功（发生）次数

(n＋1)p是整数且k＝(n＋1)p时, $b(k,n,p)=b(k-1,n,p)$ 达最大值;

我们称m＝(n＋1)p或(n＋1)p－1为最可能成功次数；

当(n＋1)p不是整数时，最可能成功次数 $m=[(n+1)p]$

直观推断：概率p是统计得到的，现在做n次试验，最可能的成功次数应该是np附近。



n=13, p=0.5

Example 2.9 邮局开设多少服务窗口合理

某居民区有n人，拟设一个邮局，开设m个服务窗口，每个窗口都办理所有业务。窗口数m太小，则经常派长队；窗口数m太大不经济。

假定在每一个指定时刻，小区n个人中每人是否去邮局是独立的，每人去邮局办理业务的概率都是p。

问题： "在营业中任意时刻保证每个窗口的排队人数（包括正在被服务的那个人）不超过s"这个事件的概率不小于α（一般取α＝0.80, 0.85, 0.90）, 则至少需开设多少个窗口？

分析： 假设在邮局办事的人数为X，有k个人在邮局的概率为

$$P\{X=k\}=C_n^k p^k (1-p)^{n-k} \qquad k=0,1,2,\cdots,n$$

Example 2.9 邮局开设多少服务窗口合理

“在营业中任意时刻每个窗口的排队人数（包括正在被服务的那个人）不超过s” 这件事相当于m个窗口总的人数X：{ X≤sm } 这个事件。其发生的概率为

$$P\{X \le sm\} = \sum_{k=0}^{sm} P\{X = k\} = \sum_{k=0}^{sm} C_n^k p^k (1-p)^{n-k}$$

题目要求此概率大于α，即

$$\sum_{k=0}^{sm} C_n^k p^k (1-p)^{n-k} \ge \alpha$$

找一个自然数m，使得上述不等式成立，此m即为问题的答案。

二项分布的重要性质

[3] n→∞ 时的渐近性质

假定p与n有关，记作p_n。 考虑 n →∞的情况，有下面的定理：

[泊松(Poisson)定理] 如果存在正常数 λ， 当 n→∞ 时有 $np_n \to \lambda$，则

$$\lim_{n\to\infty} b(k,n,p) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad k = 0,1,2,\cdots$$

Nankai University

[3] $n \to \infty$ 时的渐近性质

Proof: 记$\lambda_n = np_n$

$$b(k,n,p) = C_n^k (p_n)^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{(n-k)}$$

$$= \frac{\lambda_n^k}{k!}\,\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda_n}{n}\right)^n \Big/ \left(1-\frac{\lambda_n}{n}\right)^k$$

$$\frac{\lambda_n^k}{k!} \to \boxed{\frac{\lambda^k}{k!}}, \quad \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \to \boxed{1}, \quad \left(1-\frac{\lambda_n}{n}\right)^n \to \boxed{e^{-\lambda}}, \quad \left(1-\frac{\lambda_n}{n}\right)^k \to \boxed{1}$$

$$\to \frac{\lambda^k}{k!}e^{-\lambda} \qquad (n \to \infty)$$

[3] $n\to\infty$ 时的渐近性质

$$\lim_{n\to\infty} b(k,n,p)=\frac{\lambda^k}{k!}e^{-\lambda} \qquad k=0,1,2,\cdots$$

定理的条件 $np_n\to\lambda$ 意味着当 n很大时，p_n 必定很小.

因此，利用泊松定理，对于二项分布，当 n 很大，p 很小时有以下近似式：

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}$$

其中 $\boldsymbol{\lambda = np}$

实际计算中，$n\ge 30$, $np \le 5$ 时近似效果就很好.

Example 2.10 为保证设备正常工作，要配备适量的维修人员 .

设共有300台设备，每台的工作相互独立，发生故障的概率都是0.01. 在通常情况下，一台设备的故障可由一人来处理 .

问：至少应配备多少维修人员，才能保证当设备发生故障时不能及时维修的概率小于0.01?

分析：

设X为300台设备同时发生故障的台数，300台设备独立工作，每台出故障概率p=0.01 . 可看作n=300的伯努里概型.

可见， **X~B(n,p)，n＝300, p＝0.01**

设需配备N个维修人员, 所求问题就是满足$P(X>N) < 0.01$ 或 $P(X\le N)\ge 0.99$ 的最小的N.

Nankai University

$$P\{X > N\} = \sum_{k=N+1}^{300} C_{300}^{k} (0.01)^k (0.99)^{300-k}$$

$$\approx \sum_{k=N+1}^{300} \frac{3^k e^{-3}}{k!}$$

$$\approx \sum_{k=N+1}^{\infty} \frac{3^k e^{-3}}{k!}$$

我们求满足 $\sum_{k=N+1}^{\infty} \frac{3^k e^{-3}}{k!} < 0.01$ 的最小的N.

查泊松分布表（教材p.383）得

$$\sum_{k=9}^{\infty} \frac{e^{-3} 3^k}{k!} \approx 0.0038, \qquad \sum_{k=8}^{\infty} \frac{e^{-3} 3^k}{k!} \approx 0.012$$

$K = N+1 \geq 9$即可，即$N \geq 8$.　即至少需配备8个维修人员.

（四）泊松分布(Poisson)

从上面的泊松定理可引入另一类重要的分布。

设随机变量X可取一切非负整数值，取这些值的概率为：

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda} \qquad k=0,1,2,\cdots\cdots$$

其中λ>0是它的参数，称X服从参数为λ的泊松分布，简记 $X\sim\pi(\lambda)$

容易知道：

$$P\{X=k\}>0 \qquad (k=0,1,2\cdots)$$

$$\sum_{k=0}^{\infty}P\{X=k\}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=1$$

（四）泊松分布(Poisson)

泊松分布常与单位时间（或单位面积、单位产品等）上的计数过程相联系，比如下列计数都服从泊松分布：

- 在一天内，来到某商城的顾客数；
- 在一定时间内，某十字路口出现交通事故的次数；
- 在单位时间内，一电路收到外界电磁波的冲击次数；
- 一平方米玻璃上的气泡数；
- 一个铸件上的砂眼数；
- 在一定时间内，某种放射性物质放射出来的 α -粒子数。

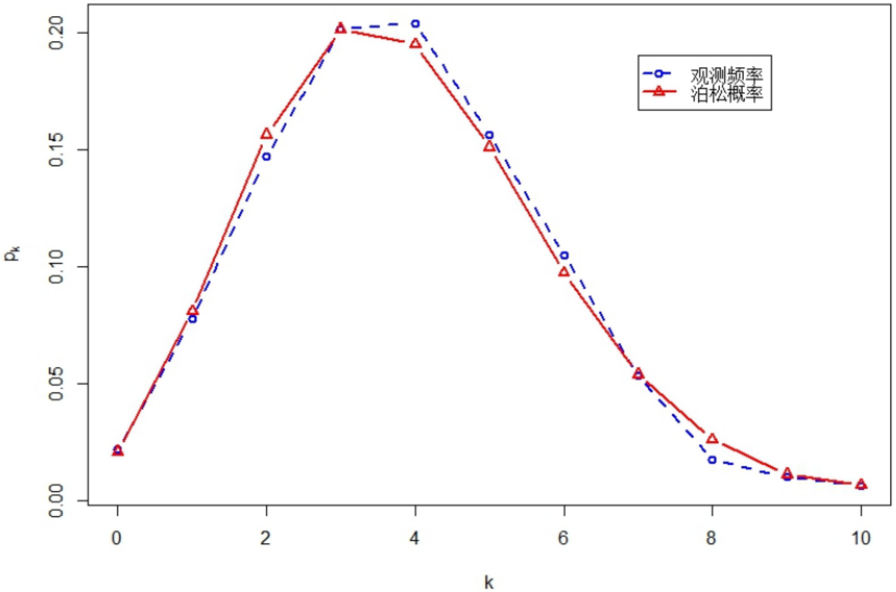
（四）泊松分布(Poisson)

例：1910 年，著名科学家 Rutherford(罗瑟福) 和 Geiger(盖克) 观察了放射性物质钋(polonium) 放射 α 粒子的情况。

- 他们进行了 $N = 2608$ 次观测, 每次观测 7.5 秒, 一共观测到 10094 个 α 粒子放出, 下面的表 2.2.1 是观测记录. 其中的 Y 是服从 Poisson(3.87) 分布的随机变量, 3.87=10094/2608 是 7.5 秒中放射出 α 粒子的平均数。
- 用 Y 表示这块放射性钋在7.5 秒内放射出的 α 粒子数, 表的最后两列表明事件 $\{Y = k\}$ 在 $N = 2608$ 次重复观测中发生的频率和 $P(Y = k)$ 基本相同。

（四）泊松分布(Poisson)

观测到的 α粒子数k	观测到k个粒子 的次数m_k	发生的 频率m_k/N	$P(Y=k)$ $Y\sim\text{Poisson}(3.87)$
0	57	0.022	0.021
1	203	0.078	0.081
2	383	0.147	0.156
3	525	0.201	0.201
4	532	0.204	0.195
5	408	0.156	0.151
6	273	0.105	0.097
7	139	0.053	0.054
8	45	0.017	0.026
9	27	0.010	0.011
10+	16	0.006	0.007
总计	2608	0.999	1.00



考虑在白天一分钟内某电话交换台接到的呼叫数X的分布，X可能取值范围为0,1,2…。那么，X服从何种分布？

我们采用分割-求和-取极限的方法来剖析。

[1] 首先把一分钟分割成n个足够小的区间（对等划分），假设每个区间只可能呼叫一次，对应的呼叫概率均为 p_n ，可以视n个小区间的呼叫为n重Bernoulli试验

[2] 在其中有k次呼叫的概率为 $b(k,n,p)=C_n^k(p_n)^k(1-p_n)^{n-k}$

假设在一分钟内的呼叫次数的统计平均数为λ，则 $np_n=\lambda$

[3] 取n足够的大，$n>>\lambda$ ，由泊松定理，我们知道

$$b(k,n,p)\to\frac{\lambda^k}{k!}e^{-\lambda}\quad(n\to\infty)$$

2.2.3 常见的概率分布

（四）泊松分布(Poisson)

Remark 1 这说明电话交换台在一分钟内呼叫的次数X满足参数为λ的泊松分布$X \sim \pi(\lambda)$。 λ的含义恰恰是在一分钟内的呼叫次数的统计平均数

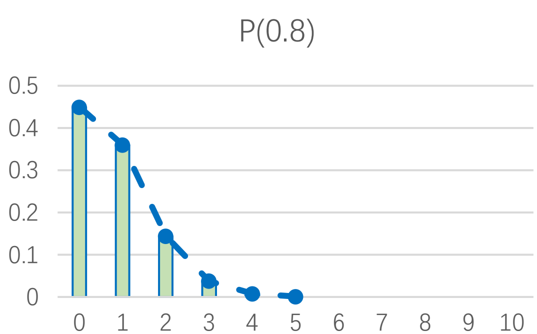
Remark 2 这也说明泊松分布是n重Bernoulli分布的极限。当n足够大时, p较小，可以近似有

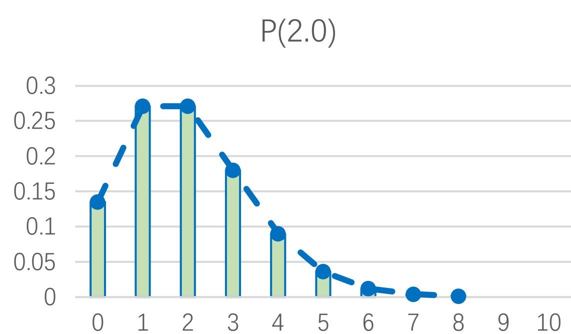
$$b(k,n,p) \sim \frac{\lambda^k}{k!}e^{-\lambda}$$

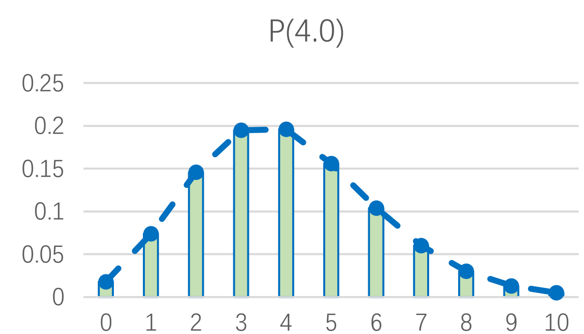
2.2.3 常见的概率分布

（四）泊松分布(Poisson)

k	0	1	2	3	4	5	6	7	8	9	10
P(0.8)	0.449	0.36	0.144	0.038	0.008	0.001					
P(2.0)	0.135	0.271	0.271	0.18	0.09	0.036	0.012	0.004	0.001		
P(4.0)	0.018	0.074	0.146	0.195	0.196	0.156	0.104	0.06	0.03	0.013	0.005







(1) $\lambda \leq 1$， $P(k)$随k增加而非增

(2) $\lambda > 1$， $P(k)$随k增加先增后降

2.2.3 常见的概率分布

（五）几何分布

单次试验中事件A 发生概率为p，现重复多次，直到A出现为止。定义X为事件A首次发生时所进行的试验次数，则X为一随机变量，其分布为

$$P(X=k)=(1-p)^{k-1}p \qquad (k=1,2,\cdots,)$$

称X服从参数为 p的几何分布，记为$X\sim Ge(p)$ 。

e.g., chpt 0生育政策、chpt1射击

（五）几何分布

几何分布的无记忆性

设$X \sim Ge(p)$，则对任意正整数m与n有

$$P(X > m + n | X > m) = P(X > n)$$

解释：在一列伯努利试验中，若事件A首次成功的试验次数X服从几何分布，则事件$\{X > m\}$表示前m次试验中A没有发生，假设在接下来的n次试验中A仍未出现，这个事件记为$\{X > m + n\}$.

这个定理表明：在前m次试验中A没有成功的条件下，则在接下来n次试验中A仍未出现的概率只与n相关，与前m次试验无关，似乎忘记了前m次试验的结果，故称为无记忆性。

（五）几何分布

几何分布的无记忆性

设$X \sim Ge(p)$，则对任意正整数m与n有

$$P(X > m + n | X > m) = P(X > n)$$

证明：因为

$$P(X > n) = \sum_{k=n+1}^{\infty} (1-p)^{k-1} p = \frac{p(1-p)^n}{1-(1-p)} = (1-p)^n$$

所以对任意的正整数m和n，条件概率

$$P(X > \mathrm{m} + n | X > m) = \frac{P(X > m + n)}{P(X > m)} = \frac{(1-p)^{m+n}}{(1-p)^m}$$

$$= (1-p)^n = P(X > n)$$

有一个赌徒在赌大小，他一直在押“大”，可是台上连续出了10把“小”，让他输了很多钱。你认为赌徒第11把该如何选择？

A 押“小”，找准了赌大小的规律

B 押“大”，前面出了很多把“小”，再出“小”的可能性很小

C 停止赌博，及时止损

提交

（五）几何分布

<u>几何分布的无记忆性</u>

赌徒心理：

有一个赌徒在赌大小，他一直在押“大”，可是台上连续出了10把“小”，让他输了很多钱。

赌徒认为，前面出了很多把“小”，再出“小”的可能性很小，于是他把全部身家押“大”，想搏一把翻本。这完全是赌徒心理，最合理的做法是马上停止赌博止损。

令X表示首次出现“大”的次数

$$P(X=11|X>10)=P(X=1)$$

2.3 随机变量的分布函数

2.3.1 WHY and HOW

WHY：分布列不能表示很多的随机变量的分布，比如机器的寿命，无法一一罗列这些值及其概率，有必要引入概率分布的新的表示法。

HOW：定义区间$(-\infty, x]$上的概率

- 事实上，我们经常关心的是一个随机变量落入一个区间、空间的概率，如机器寿命大于 T 的概率，产品中次品数少于 k 件的概率。
- 注意到随机变量取值为实数，可以说分布在实数轴上的任意的区间$(x_1, x_2]$都可以用$(-\infty, x]$的（有限或可列）并、（有限或可列）交、逆产生的集合，故此可以定义在区间$(-\infty, x]$上概率。

2.3 随机变量的分布函数

2.3.2 定义

设随机变量X，对任意的实数x，随机变量X取值落入区间$(-\infty, x]$内的概率为F(x)

$$F(x) = P(X \le x) \qquad -\infty < x < +\infty$$

称为随机变量X的分布函数 (distribution function)。

[1] 对确定的随机变量X, 其分布函数F(x)是唯一确定的，它是实变量x的函数，因此我们可以利用实变函数论这一有力工具来研究随机变量。

[2] 有了分布函数，则对任一波雷尔集A，概率P(A)可以用分布函数来表示。

$$\begin{aligned} P(a < X \le b) &= P\{(-\infty, b]\} - P\{(-\infty, a]\} \\ &= F(b) - F(a) \end{aligned}$$

2.3.3 分布函数的特性

由于分布函数已经是一个普通函数，就可以用数学分析、实变函数来研究。至此已经达到目的：

随机事件	$\Rightarrow$	随机变量(统一描述)
概率	$\Rightarrow$	分布函数(函数工具)

(1) $F(x)$是一个不减的函数， $\forall x_1 \le x_2, F(x_1) \le F(x_2)$

(2) $0 \le F(x) \le 1$

(3) $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$ $\qquad F(\infty) = \lim\limits_{x \to \infty} F(x) = 1$

(4) 函数的右连续性 $F(x+0) = \lim\limits_{y \to x+0} F(y)$,

2.3 随机变量的分布函数

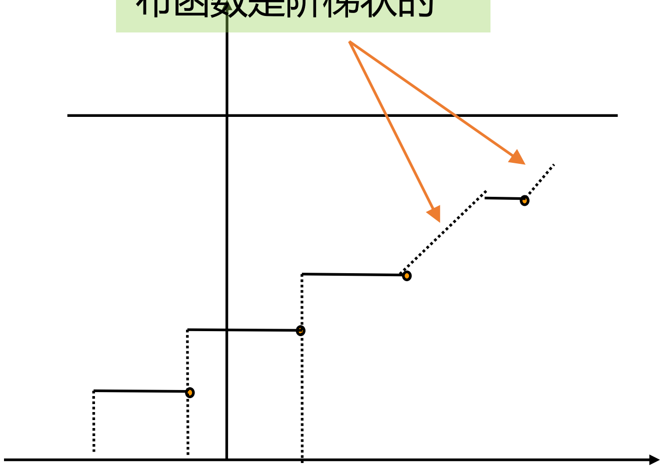
Example 2.11 离散随机变量的分布函数，随机变量X的分布列为

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p(x_1) & p(x_2) & \cdots & p(x_k) & \cdots \end{pmatrix}$$

且 $x_1 < x_2 < \cdots\cdots < x_k < \cdots\cdots$ ，则X的分布函数为 $F(x) = P(X \le x) = \sum_{x_k \le x} p_k$

省略号！
离散型随机变量的分布函数是阶梯状的

$$F(x) = \begin{cases} 0, & x < x_1, \\ p(x_1), & x_1 \le x < x_2, \\ \cdots & \cdots \\ \sum_{i \le k} p(x_i), & x_k \le x < x_{k+1}, \\ \cdots & \cdots \end{cases}$$



2.3 随机变量的分布函数

它是间断的分段函数，在x_k(k=1，2， …)各有一跳跃，跃度为$p(x_k)$，在每一段 $[x_k，x_{k+1})$中都是常数，呈阶梯形。

比如，随机变量X的分布律为

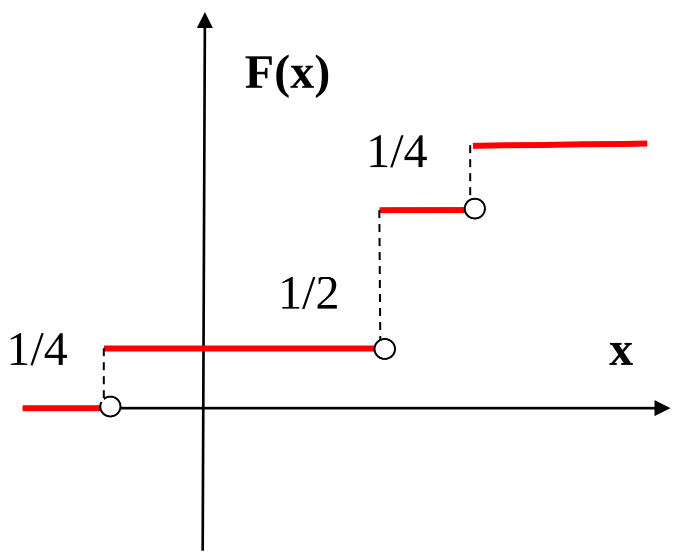
X	-1	2	3
p	1/4	1/2	1/4

X的分布函数为F(x)

$$F(x)=\begin{cases} 0 & x<-1 \\ P(X=-1) & -1\le x<2 \\ P\{X=-1\}+P\{X=2\} & 2\le x<3 \\ 1 & x\ge 3 \end{cases}$$

2.3 随机变量的分布函数

$$F(x)=\begin{cases}0 & x<-1\\ \dfrac{1}{4} & -1\le x<2\\ \dfrac{3}{4} & 2\le x<3\\ 1 & x\ge 3\end{cases}$$



$$P\left\{X\le\frac{1}{2}\right\}=F\left\{\frac{1}{2}\right\}=\frac{1}{4}$$

$$P\left\{\frac{3}{2}<X\le\frac{5}{2}\right\}=F\left(\frac{5}{2}\right)-F\left(\frac{1}{2}\right)=\frac{1}{2}$$

$$P\{2\le X\le 3\}=F(3)-F(2)+P(X=2)$$

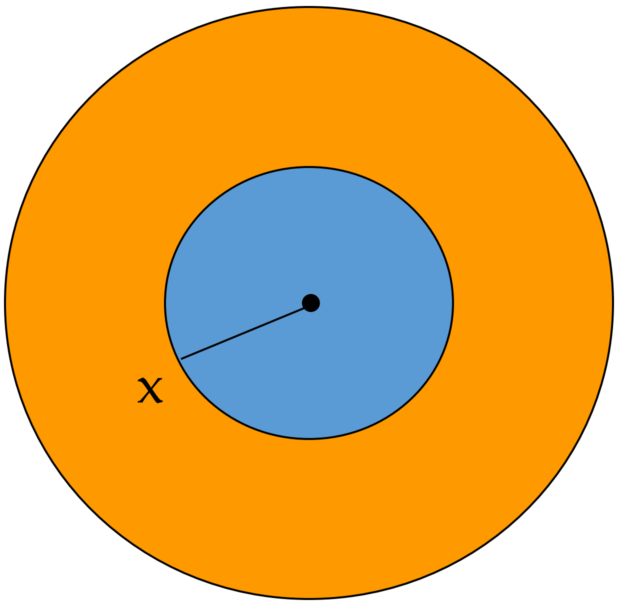
$$=\frac{3}{4}$$

Example 2.12

一个半径为2米的圆盘形靶子，设击中靶上任何一同心圆内的点的概率与该圆的面积成正比，并设射击都能中靶，以X表示弹着点与圆心的距离。

求：随机变量X的分布函数

解：如图，



[1] 若x<0，则{X≤x}是不可能事件

$$F(x)=P\{X\le x\}=0$$

[2] 若0≤ x≤2, 由假设知道{0≤ X≤x}的概率正比于圆的面积πx^2

$$P\{0\le X\le x\}=k\pi x^2$$

其中k为常数

当x＝2时，有 $P\{0\le X\le x\}=k\pi 4=1$

求得 $k\pi=1/4$

于是有 $P\{0\le X\le x\}=x^2/4$

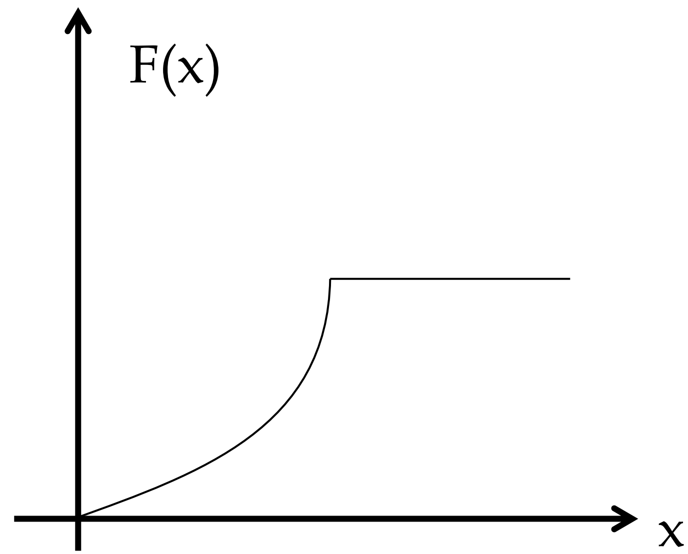
$F(x)=P\{X\le x\}=P(X<0)+P\{0\le X\le x\}=x^2/4$

[3] 若$x\ge 2$，由题意$\{X\le x\}$是必然事件

$F(x)=P\{X\le x\}=1$

综合以上知道：

$$F(x)=\begin{cases}0 & x<0\\ \dfrac{x^2}{4} & 0\le x<2\\ 1 & x\ge 2\end{cases}$$



同时分布函数可以写为积分形式

$$F(x)=\int_{-\infty}^{x} f(t)dt \qquad f(t)=\begin{cases}\dfrac{t}{2} & 0<t<2\\ 0 & \text{其他}\end{cases}$$

2.4 连续型随机变量及其分布

2.4.1 连续型随机变量及密度函数

猜测连续型随机变量的分布函数（依据）：

[1] 概率定义一再地引用几何面积的概念；

[2] 对于一个随机变量 X 取得某个区间的所有值时，离散随机变量的分布函数为概率的和，扩展到连续性随机变量上积分的形式：

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

Definition 1 如果对于随机变量 X 的分布函数$F(x)$，存在非负函数 $f(x)$，使对于任意实数 x 满足

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

则称 X 为连续型(continuous)随机变量，称 $f(x)$ 为 X 的概率密度函数，简称为**密度函数(density function)**。

2.4 连续型随机变量及其分布

2.4.2 连续型随机变量的特性

连续型随机变量的分布函数$F(x)$ 具有下列数学性质：

(1) 可以证明，$F(x)$ 必定处处连续；在 $f(x)$ 的连续点，$F(x)$可导，且 $F'(x) = f(x)$.

(2) $\int_{-\infty}^{+\infty} f(x)dx = 1.$

(3) 随机变量X落在任意区间(a，b]的概率:

$$P\{a < x \le b\} = F(b) - F(a)$$

$$= \int_a^b f(x)dx.$$

良好的性质

(4) 特别，对任一常数c，$P\{X = c\} = 0.$

2.4 连续型随机变量及其分布

Remark 1：对连续型随机变量，计算在一点的概率是没有意义的，这也是不能用分布列描写连续型随机变量的理由之一。

Remark 2：对连续型随机变量$\{X=C\}$是一个可能发生的事件，因此一事件A的概率为0并不表明$A=\varnothing$（如在一个区间上方抛点，点可以落到任意一点 x，但是落到任意一点的概率为0）；同样若$P(A)=1$，也并不表明$A=S$（样本空间）。

Remark 3：除了离散型，连续型以外，随机变量还有其它类型吗？

$$F(x)=\begin{cases} 0, & x<0, \\ (1+x)/2, & 0\le x\le 1, \\ 1, & x>1. \end{cases}$$

是分布函数，它不是离散型的，也不是连续型的(因为 $F(x)$ 本身就不连续)。

2.4.3 常见的连续型随机变量

（一）均匀(Uniform)分布

考虑在一个区间[a,b]上方抛点，假设点落入任意一个等长度区间的概率（可能性）是相同的（或者说点落入某个区间的可能性大小只与区间长度有关，与区间端点无关），用随机变量 X 表示区间的端点 $(-\infty, x]$，则 X 的分布函数为

$$F(x)=P\{X\le x\}=\begin{cases}0, & x\le a\\ \dfrac{x-a}{b-a}, & a<x<b\\ 1, & x\ge b\end{cases}$$

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{otherwise}\end{cases}$$

称具有上述概率密度的连续型随机变量 X 为在区间 $[a, b]$ 上服从**均匀分布**，简记作 $X\sim U(a, b)$。

2.4.3 常见的连续型随机变量

（一）均匀(Uniform)分布

Example 2.13 考察一个数据，它在小数点n位后四舍五入，则其真值x与其近似值$\hat{x}$之间的误差$\varepsilon=\hat{x}-x$，一般假定服从$\left[-0.5\times10^{-n},0.5\times10^{-n}\right]$上的均匀分布。

利用均匀分布理论可以对大量运算后的数据进行误差分析。这对于使用计算机解题时是很重要的，因为计算机的字长总是有限的。

Example 2.14 列车从车站每40分钟开出一次，乘客到站时刻是机会均等的，问等候超过10分钟的概率是多少？

以T表示等候时间，它是一个随机变量。$P\{t_1<T\le t_1+\Delta t\}$指在区间$[t,t+\Delta t]$等到车的概率.

T取值范围为[0, 40]，其服从均匀分布$\boldsymbol{U}$[0, 40]。

$$P\{T\ge10\}=1-P\{T\le10\}=\int_{10}^{40}\frac{1}{40}dt=\frac{3}{4}$$

2.4.3 常见的连续型随机变量

（二）指数(Exponential)分布

随机变量 X 的密度函数为（当然，我们说它应该满足分布密度）

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}=\dfrac{1}{\theta}Exp\left(-\dfrac{x}{\theta}\right), & x>0\\ 0, & else\end{cases}$$

其中$\theta>0$为常数，则称X服从参数为θ的**指数分布**(Exponential)。
其分布函数为

$$F(x)=\begin{cases}1-e^{-x/\theta}, & x>0\\ 0, & else\end{cases}$$

指数分布常用来描述产品的寿命。

（二）指数分布

指数分布常用来描述某个事件发生的等待时间的分布，比如：

- 从现在开始，下一次地震发生的等待时间；
- 一场新的战争爆发的时间间隔；
- 从现在开始接到下一个误拨的电话的时间间隔；
- 产品失效的时间（产品寿命）。

（二）指数分布

Example 2.15 若是用了t小时的电子管，在以后的Δt小时内失效的概率为 $\lambda\Delta t+o(\Delta t)$，其中$\lambda$是不依赖于$t$的正常数，求电子管在 T 小时内失效的概率（假定电子管寿命为0的概率是0）。

解： 设 X 为电子管的寿命。注意，对于成批的电子管而言，X 是一个随机变量。按题意要找出 $P\{X\le T\}$，即要找出 X 的分布函数 $F(t)$，求 $F(t)$ 在 T 点的值。

从问题给出的条件入手，其可以表示为：

$$P\{t<X<t+\Delta t\,|\,t<X\}=\lambda\Delta t+o(\Delta t)$$

$$\frac{P\{(t<X<t+\Delta t)\bigcap(t<X)\}}{P\{t<X\}}=\lambda\Delta t+o(\Delta t)$$

$$\frac{P\{(t<X<t+\Delta t)\}}{P\{t<X\}}=\lambda\Delta t+o(\Delta t)$$

（二）指数分布

$$\frac{F(t+\Delta t)-F(t)}{1-F(t)}=\lambda\Delta t+o(\Delta t)$$

$$\frac{F(t+\Delta t)-F(t)}{\Delta t}=\lambda\big(1-F(t)\big)+o(\Delta t)/\Delta t$$

$$F'(t)=\lambda\big(1-F(t)\big)$$

解此线性微分方程组（注意到初始条件$F(0)=0$），解得

$$F(t)=1-e^{-\lambda t}$$

X 有密度函数 $$f(t)=\begin{cases}\lambda e^{-\lambda t} & t>0\\ 0 & else\end{cases}$$

电子管在T小时内失效的概率为 $1-e^{-\lambda T}.$

注意：寿命←→失效的对立性

寿命为T小时的概率为 $1-F(T)=e^{-\lambda T}.$

（二）指数分布

指数分布具有类似几何分布的“**无记忆性**”：使用s小时后元件的寿命不低于t的概率与初始时使用寿命为t的概率相等。

事实上，设随机变量ξ服从参数为λ的指数分布，则对于任意的$s>0$，$t>0$,

$$P\{(\xi>s+t)\cap(\xi>s)\}$$

$$P(\xi>s+t\mid\xi>s)=P(\xi>s+t,\xi>s)/P(\xi>s)$$

$$=e^{-\lambda(s+t)}/e^{-\lambda s}$$

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

$$=P(\xi>t)$$

（二）指数分布

Example 2.16 (泊松分布和指数分布的关系) 若某设备在长为t的时间$[0, t]$内发生故障的次数$N(t)$服从参数为λt的泊松分布，则相继两次故障之间的时间间隔T服从参数为λ的指数分布。

解： 因为$N(t)\sim\pi(\lambda t)$，即$P(N(t)=k)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}, k=0,1,2,\cdots$

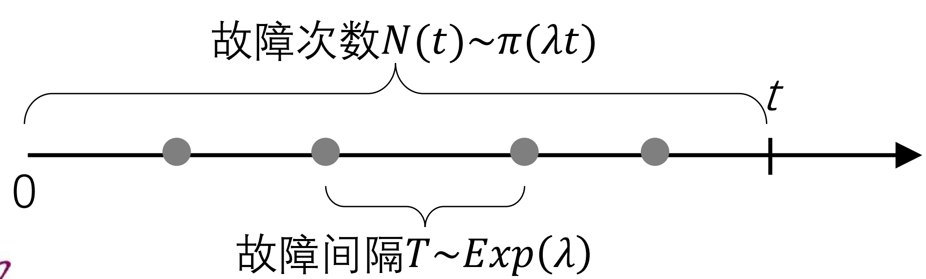
相继两次故障之间的时间间隔T是非负随机变量，**且事件$\{T>t\}$说明该事件在$[0,t]$内没有发生故障，即$\{T>t\}=\{N(t)=0\}$**，由此，

当$t<0$时，$F_T(t)=P(T\le t)=0$

当$t\ge 0$时，

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

所以$T\sim Exp(\lambda)$，即T服从参数为λ的指数分布



单选题 10分

设置

设随机变量X在区间(2, 5)上服从均匀分布，先对X进行三次独立观测，则至少有2次观测值大于3的概率为 (　　)

A 20/27

B 27/30

C 2/5

D 2/3

提交

填空题 20分

设置

设连续型随机变量X的分布函数为：

$$F(x)=\begin{cases}0, x<0,\\ A\sqrt{X}, 0\le x<1,\\ 1, x\ge 1\end{cases}$$

求：

(1) $A=$ [填空1]

(2) $P(X>0.36)=$ [填空2]

正常使用填空题需3.0以上版本雨课堂

南開大學 Nankai University

作答

（三）正态分布

若随机变量 X 的密度函数为　　　　　　　　　　　　　《正态分布的前世今生》

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty<x<\infty$$

就称 X 服从**正态(Normal)分布**，记作 $X\sim N(\mu,\sigma^2)$。其中, $\sigma>0$，$-\infty<\mu<\infty$。

我们来证明上述定义的 $f(x)$ 的确是密度函数。

显然 $f(x)>0$;

进一步证明 $\int_{-\infty}^{\infty}f(x)=1$，即证明

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=1 \xrightarrow{t=\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=1.$$

（三）正态分布

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=1 \xrightarrow{t=\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=1.$$

$$\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\right]^2=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}e^{-\frac{s^2}{2}}dtds$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{t^2+s^2}{2}}dtds$$

上述二重积分可用极坐标表示成

$$\frac{1}{2\pi}\int_{0}^{2\pi}d\theta\int_{0}^{\infty}e^{-\frac{1}{2}r^2}rdr=\int_{0}^{\infty}e^{-\frac{1}{2}r^2}rdr=1.$$

也即 $\int_{-\infty}^{\infty}f(x)=1.$

（三）正态分布

Remark 1： 正态分布是概率论中最重要的一种分布，与二项分布、泊松分布并称为**三大分布**。

正态分布是19世纪初高斯（Gauss）在研究测量误差时首次引进的，故正态分布又称**误差分布**或**高斯分布**。

Remark 2： 对正态分布应用很广，一般说来，若影响某一数量指标的随机因素很多，而每一因素所起的作用又不很大，则这个数量指标服从正态分布。例如进行测量时，由于仪器精度、人的视力、心理因素、外界干扰等多种因素影响，测量结果大致服从正态分布，测量误差也服从正态分布。

Remark 3： 正态分布具有良好的性质，一定条件下，很多分布可用正态分布来近似表达; 另一些分布又可以通过正态分布来导出，因此，正态分布在理论研究中也相当重要。

1. 正态分布的特性

正态分布的密度曲线是一条关于纵轴对称的钟形曲线。特点是“两头小，中间大，左右对称”。

[1] 曲线 $f(x)$ 关于 $x=\mu$对称。

$$f(\mu-x)=f(\mu+x) \qquad P\{\mu-h<X<\mu\}=P\{\mu<X<\mu+h\}$$

[2] 当$x=\mu$时，$f(x)$取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$

当$x<\mu$时，$f(x)$单调递增；

当$x>\mu$时，$f(x)$单调递减；

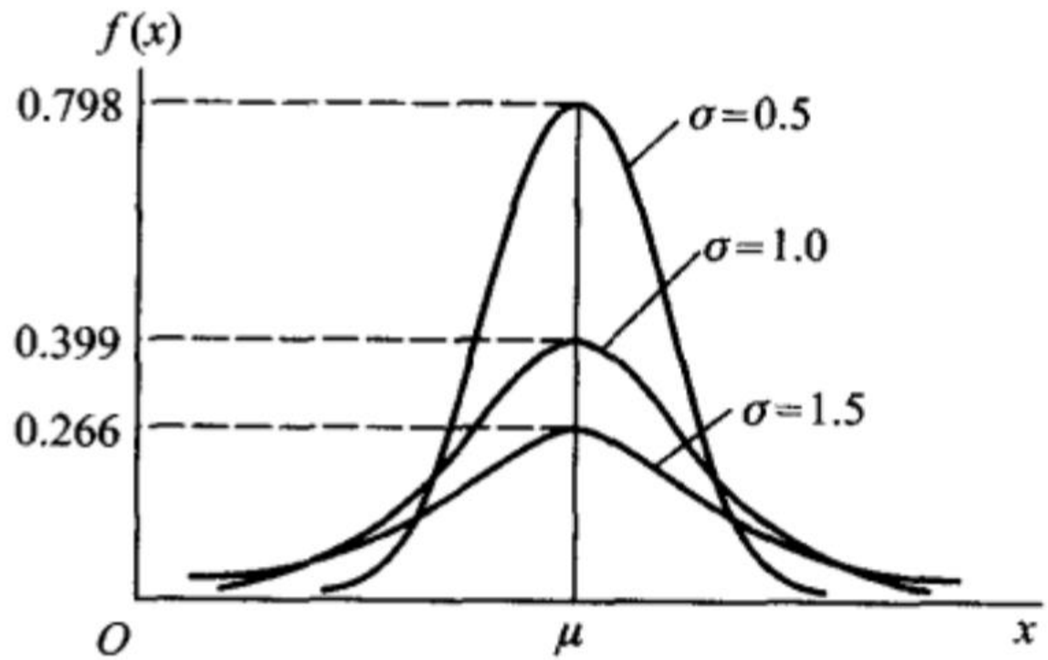
$x\to\infty$ 时，$f(x)\to 0.$

这些表明，x 离 μ 越远，$f(x)$的值越小。对于同样长度的区间，当区间离 μ 越远时，X 落在此区间的概率越小。

1. 正态分布的性质

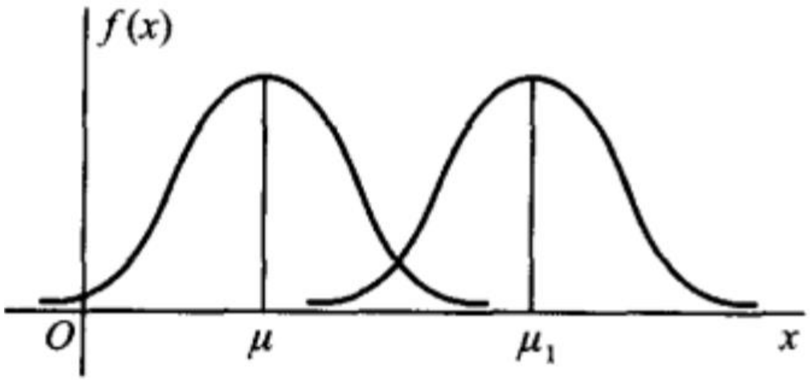
[3] σ^2（方差）越大，最高点越低，$f(x)$ 的图形越扁平，说明 X 分布比较分散，X 取值离开 μ 点远的概率也越大；

σ^2（方差）越小，最高点越高，说明 X 分布比较集中；$f(x)$的图形越陡峭，X 取值越集中在点 μ 附近。



1. 正态分布的性质

[4] 固定σ，改变 μ 的值，则图形沿着 x 轴平移，而不改变其形状。。



可见正态分布的概率密度曲线的位置完全由参数μ决定，μ 称为**位置参数**。

[1] μ、σ 分别称为**均值**与**均方差/标准差**；

[2] 高考：四科成绩分布，希望是均值尽量高（总分过线），方差也要尽量小（别偏科）。

甲：总分650，语文120，数学130，英语130，理综270

乙：总分670，语文130，数学130，英语130，理综280

丙：总分670，语文90，数学145，英语145，理综290

1. 正态分布的性质

[5] 当 $\mu=0, \sigma=1$时， 称为**标准正态分布**(standardized normal distribution)， 它的密度曲线关于纵轴对称， 其密度函数及分布函数特别记为 $\varphi(x)$ 和 $\Phi(x)$：

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \qquad -\infty < x < \infty$$

容易知道 $\Phi(-x)=1-\Phi(x)$.

2. 正态分布的计算

[1] 标准正态分布概率的计算

标准正态分布的计算不容易，有专门的表格供查阅。

当 $x \geq 0$ 时, 每隔一定数值，可以查到对应的分布函数 $\Phi(x)$的值;

在这些数值之间，用线性插值法求得相应的函数值;

(附录中这个纵向间隔是0.1，横向间隔是0.01）。

当 $x < 0$ 时, 注意到

$$\Phi(-x) = 1 - \Phi(x).$$

结合 $-x > 0$时的 $\Phi(-x)$ 表就可算出 $x < 0$时 $\Phi(x)$ 的值。

2. 正态分布的计算

Example 2.17 设 $X\sim N(0,1)$。

(1) 计算 $P\{-1<X<1\}$;

(2) 已知 $P\{X<\lambda\}=0.9755$，求λ。

解： (1) $P(-1<X<1)=P(X<1)-P(X<-1)$

$$
\begin{aligned}
&=\Phi(1)-\Phi(-1)\\
&=\Phi(1)-(1-\Phi(1))\\
&=2\Phi(1)-1\\
&=2\times 0.8413-1\\
&=0.6826
\end{aligned}
$$

$P(|X|<2)=0.9544,\qquad P(|X|<3)=0.9973$

2. 正态分布的计算

(2) $\Phi(\lambda)=0.9755$， 求λ。

我们倒查表格， 看看0.9755对应的值是多少。

但是， 表格没有这样的值, 只有相近的。

$$\Phi(1.96)=0.9750<\Phi(\lambda)<\Phi(1.97)=0.9756$$

由于$\Phi(x)$是单调不减的， 故λ在1.96与1.97之间

$$1.96<\lambda<1.97$$

在这个小的范围内， 我们把λ与$\Phi(\lambda)$的关系近似看作线性：

$$\frac{\lambda-1.96}{1.97-1.96}\approx\frac{\Phi(\lambda)-\Phi(1.96)}{\Phi(1.97)-\Phi(1.96)}$$

$$\lambda\approx1.96+\frac{\Phi(\lambda)-\Phi(1.96)}{\Phi(1.97)-\Phi(1.96)}(1.97-1.96)$$

$$\approx1.968$$

以上思路称为**线性插值法**。

[2] 一般正态分布的概率计算

对一般的 $X \sim N(\mu, \sigma^2)$，可以变换为标准正态分布加以计算。

记 $Y = (X - \mu) / \sigma$ (称为 X 的**标准化随机变量**)，则 Y 服从 $N(0, 1)$。

事实上 Y 的分布函数 $F_y(y)$

$$F_y(y) = P\{Y \le y\}$$

$$= P\{X \le \sigma y + \mu\}$$

$$= \int_{-\infty}^{\sigma y+\mu} \frac{1}{\sqrt{2\pi}\sigma} Exp\left(-(t-\mu)^2 / 2\sigma^2\right) dt$$

$$\Big\downarrow \; z = \frac{t-\mu}{\sigma}$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} Exp\left(-\frac{z^2}{2}\right) dz$$

[2] 一般正态分布的概率计算

由此可知：当 $X \sim N(\mu, \sigma^2)$ 时， $Y = (X-\mu)/\sigma \sim N(0,1)$

则

$$P\{X \le x\} = P\left\{\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right\}$$

$$= P\left\{Y \le \frac{x-\mu}{\sigma}\right\}$$

$$= \Phi\left\{\frac{x-\mu}{\sigma}\right\}$$

对于任意的区间 $(x_1, x_2]$

$$P\{x_1 < X \le x_2\} = P\left\{\frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{x_2-\mu}{\sigma}\right\}$$

$$= \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right)$$

[2] 一般正态分布的概率计算

Example 2.18 汽车设计手册中指出：人的身高服从正态分布 $X\sim N(\mu,\sigma^2)$。根据各国统计资料，可得各国、各民族男子身高的 μ 和 σ。对于中国人，$\mu=1.75$，$\sigma=0.05$。现要求上下车时要低头的人不超过0.5%，车门需要多高?

解： 设公交车车门高位h，X 为乘客的身高，则 $X\sim N(1.75,0.05^2)$，根据题意

$$P\{X\le h\}\ge 99.5\%$$

$$P\left\{\frac{X-1.75}{0.05}\le\frac{h-1.75}{0.05}\right\}\ge 99.5\%$$

$$\Phi\left(\frac{h-1.75}{0.05}\right)\ge 99.5\%$$

$$\frac{h-1.75}{0.05}\ge 2.58$$

$$h\ge 1.879$$

如此，知道车门应该高1.88米。

[2] 一般正态分布的概率计算

Example 2.19 从南郊某地乘车到北区火车站有两条路可走，第一条路较短，但交通拥挤，所需时间T_1服从 $N(50, 100)$ 分布；第二条路线略长，但意外阻塞较少，所需时间T_2服从 $N(60, 16)$。

(1) 若有70分钟可用，问应走哪一条路？

(2) 若只有65分钟可用，又应走哪一条路？

分析： 应该走在允许时间内有较大概率赶到火车站的路线。

解： (1) 若有70分钟可用

走第一条路线及时赶到的概率。

$$P\{T_1 \le 70\} = \Phi\left(\frac{70-50}{10}\right) = \Phi(2) = 0.9772$$

走第二条路线及时赶到的概率。

$$P\{T_2 \le 70\} = \Phi\left(\frac{70-60}{4}\right) = \Phi(2.5) = 0.9938$$

在这种场合，应走第二条路线。

[2] 一般正态分布的概率计算

(2) 若只有65分钟可用，又应走哪一条路?

走第一条路线及时赶到火车站的概率为

$$P\{T_1 \le 65\} = \Phi\left(\frac{65-50}{10}\right) = \Phi(1.5) = 0.9332$$

走第二条路线及时赶到的概率为

$$P\{T_2 \le 65\} = \Phi\left(\frac{65-60}{4}\right) = \Phi(1.25) = 0.8944$$

因此在这种场合，应走第一条路线更为保险。

Remark：此处的μ和σ的含义分别是走完这段路程所需要的平均时间和均方差。

[2] 一般正态分布的概率计算

Example 2.20 设 $X \sim N(\mu, \sigma^2)$,

求 $P\{|X-\mu|<\sigma\}, P\{|X-\mu|<2\sigma\}, P\{|X-\mu|<3\sigma\}$

解： $Y=(X-\mu)/\sigma$，则 $Y \sim N(0,1)$，故由标准正态分布函数 $\Phi(y)$，

可以求得
$$P\{|X-\mu|<k\sigma\} = P\left\{-k<\frac{X-\mu}{\sigma}<k\right\} = \Phi(k)-\Phi(-k) = 2\Phi(k)-1$$

$$P\{|X-\mu|<\sigma\} = 2\Phi(1)-1 \approx 0.6826$$

$$P\{|X-\mu|<2\sigma\} = 2\Phi(2)-1 \approx 0.9544$$

$$P\{|X-\mu|<3\sigma\} = 2\Phi(3)-1 \approx 0.9973$$

说明正态随机变量的99.73 %的值落在(μ-3σ, μ+3σ)之中，落在该区间之外的概率几乎为零，这种情况在实际应用中称为“**3σ原则**”。

（四）伽玛分布

伽玛函数 《神奇的伽玛函数》

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \ , \alpha > 0$$

伽玛函数具有如下性质:

（1） $\Gamma(1) = 1$， $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

（2）利用分部积分法，$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

$$\Gamma(\alpha+1) = \int_0^{\infty} x^{\alpha} e^{-x} dx \ = -e^{-x}x^{\alpha}|_0^{\infty} + \alpha \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

$$= \alpha\Gamma(\alpha)$$

Esp. $\Gamma(n) = (n-1)!\,\Gamma(1) = (n-1)!$

（四）伽玛分布

伽玛分布：若随机变量X的密度函数为

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x\geq 0\\[2ex] 0, x<0\end{cases}$$

则称X服从伽玛分布，记作$X\sim\Gamma(\alpha,\lambda)$，其中$\alpha>0$，为形状参数，$\lambda>0$为尺度参数

当α为一整数，如$\alpha=n$时，$\Gamma(\alpha,\lambda)$在实践中常用作某个事件总共要发生n次的等待时间的分布。

- 当$\alpha=1$时，该分布退化为指数分布
- 当$\alpha=\frac{n}{2},\lambda=\frac{1}{2}$时的伽玛分布成为自由度为$n$的$\chi^2$（卡方）分布

（四）伽玛分布

$$f(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x\geq 0\\ 0, x<0\end{cases}$$

Example 2.21 若在长为t的时间$[0, t]$内某事件发生的次数$N(t)$服从参数为λt的泊松分布，则等待该事件发生n次的时间服从参数为(n,λ)的伽玛分布$\Gamma(n,\lambda)$

解： 若令T_n表示事件第n次发生的时间，易知T_n是非负随机变量。则$T_n\leq t$等价于在时刻t之前该事件至少发生了n次，即$N(t)\geq n$，因此

$$P\{T_n\leq t\}=P\{N(t)\geq n\}=\sum_{j=n}^{\infty}P\{N(t)=j\}=\sum_{j=n}^{\infty}\frac{e^{-\lambda t}(\lambda t)^j}{j!}$$

$$f(t)=\sum_{j=n}^{\infty}\frac{e^{-\lambda t}j(\lambda t)^{j-1}\lambda}{j!}-\sum_{j=n}^{\infty}\frac{\lambda e^{-\lambda t}(\lambda t)^j}{j!}=\frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}$$

$$=\Gamma(n,\lambda)$$

（五）贝塔分布

贝塔分布：若随机变量X的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{B(a,b)}x^{a-1}(1-x)^{b-1},0<x<1\\ 0,\text{其他}\end{cases}$$

则称X服从贝塔分布，记作$X\sim\beta(a,b)$，其中

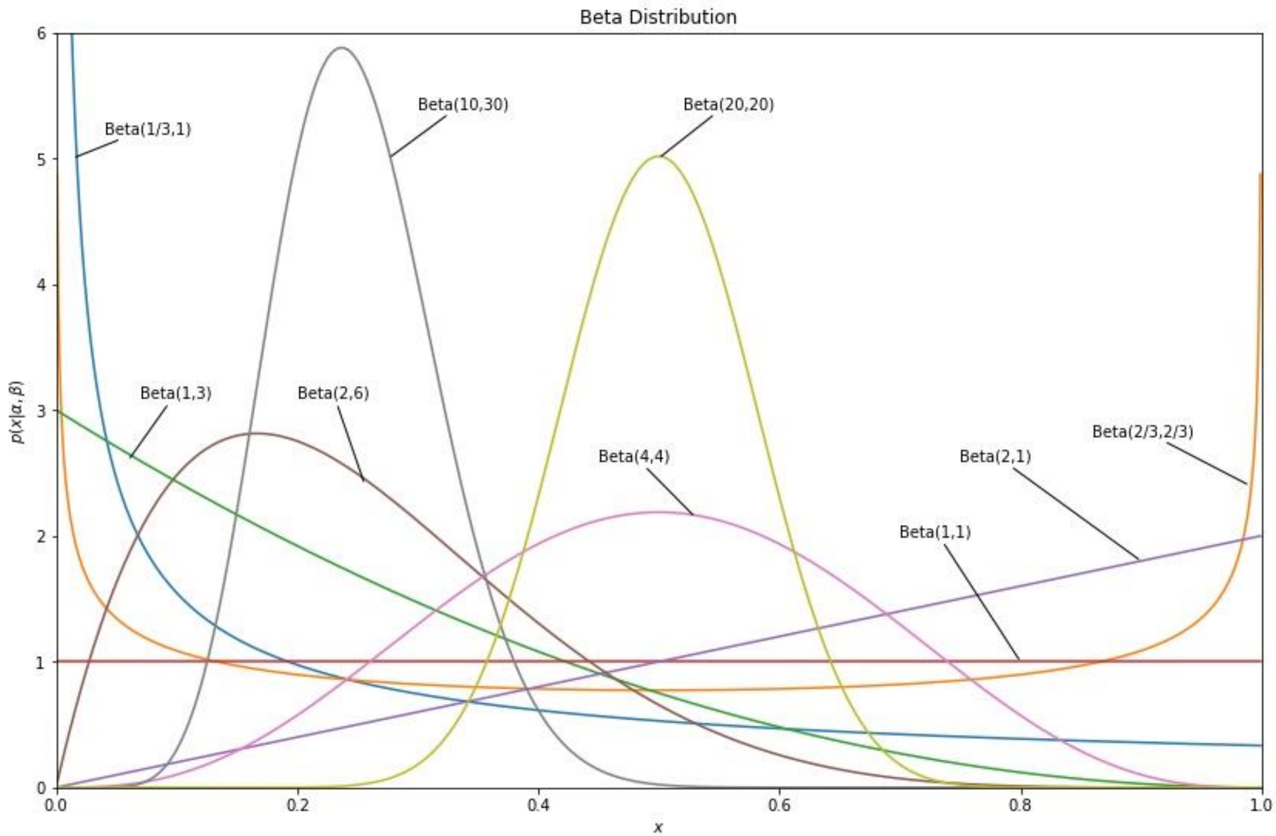
$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

可以证明 $B(a,b)=\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

https://www.zhihu.com/question/359358157

贝塔分布常用于为取值为[c, d]的有限区间随机变量建模（任意有限区间[c, d]可以转化为[0, 1]）

- ✓ 当$a=b$时，$p(x)$关于$x=1/2$对称
- ✓ a越大，取决于$x=1/2$附近的权重会越大
- ✓ 当$a=b=1$时，退化为$U(0,1)$
- ✓ 当$b>a$时，密度函数向左倾斜；反之，密度函数向右倾斜
- ✓ 当$a<1$且$b<1$时，$p(x)$是下凸的单峰函数
- ✓ 当$a>1$且$b>1$时，$p(x)$是上凸的单峰函数



2.5 随机变量的函数及其分布

人们经常碰到随机变量的函数。例如

(1) 分子运动动能 $T = mv^2/2$是分子运动速度—随机变量 v 的函数;

(2) 数理统计中经常用到 $\chi^2(n)$ 分布，相应的随机变量

$$\chi^2 = \xi_1^2 + \xi_2^2 + \cdots + \xi_n^2$$

其中各 ξ_i 相互独立，都服从 $N(0, 1)$。χ^2是$\xi_1, \ldots, \xi_n$的函数。

一般地，若 X 是随机变量， $y = g(x)$是普通的实函数，则 $Y = g(X)$ 是 X 的函数。

产生两个问题:

(1) Y 是随机变量吗?

(2) 如果是，Y 的分布是什么? 它与 X 的分布有什么关系?

2.5.1 离散型随机变量的函数

Example 2.22 假设随机变量 X 具有如下分布律

X	-1	0	1	2
p_n	0.2	0.3	0.1	0.4

求： $Y = (X-1)^2$ 的分布。

解： Y 的的可能取值为0,1,4, 是有限个，只须算出对应的概率。

$$P\{Y=0\} = P\{X=1\} = 0.3,$$

$$P\{Y=1\} = P\{X=-1\} + P\{X=1\} = 0.3$$

$$P\{Y=4\} = P\{X = 2\} = 0.4$$

类似可得其它概率。我们得到 Y 的分布:

Y	0	1	4
p_n	0.3	0.3	0.4

一般，设 X 有分布列 $P\{X = x_i\} = p\{x_i\}$，$i=1,2,\ldots$，则 $Y= f(X)$ 有分布列 $$\boxed{P\{Y=y_j\} = \sum_{f(x_i)=y_j} p(x_i), \qquad j=1,2,\cdots}$$

2.5.1 离散型随机变量的函数

Example 2.23 设 $X \sim B(n_1, p), Y \sim B(n_2, p)$， X、Y相互独立，

求： $Z = X + Y$ 的分布。

解： X, Y 各可取值 $0, 1, \dots, n_1$ 和 $0, 1, \dots, n_2$, 则 Z 可取值$0, 1, \dots ,$ n_1+n_2, 得

$$P\{Z = s\} = \sum_{k=0}^{s} P\{X = k, Y = s-k\}$$

$$= \sum_{k=0}^{s} P\{X = k\} P\{Y = s-k\}$$

$$= \sum_{k=0}^{s} C_{n_1}^{k} p^{k} (1-p)^{n_1-k} C_{n_2}^{s-k} p^{s-k} (1-p)^{n_2-(s-k)}$$

$$= p^{s} (1-p)^{n_1+n_2-s} \sum_{k=0}^{s} C_{n_1}^{k} C_{n_2}^{s-k}$$

$$= C_{n_1+n_2}^{s} p^{s} (1-p)^{n_1+n_2-s}$$

$$\boxed{X+Y \sim B(n_1+n_2, p)}$$

2.5.1 离散型随机变量的函数

Remark1： 上面公式的推导用到了组合数的性质。

$$\sum_{k=0}^{s} C_{n_1}^{k} C_{n_2}^{s-k} = C_{n_1+n_2}^{s}$$

Remark2： $X+Y \sim B(n_1+n_2, p)$ 这个事实显示了二项分布一个很重要的性质：**两个独立的二项分布，当它们的第二参数相同时，其和也服从二项分布，它的第一参数恰为这两个二项分布第一参数的和。**

- 这性质称为二项分布的**再生性**(或**可加性**(additive property))
- 从X, Y的概率意义来看，这结果是非常明显的：X 和Y 分别是 n_1和 n_2重贝努里试验中成功的次数，两组试验合起来，$Z=X+Y$应该就是n_1+n_2重贝努里试验中成功的次数。

2.5.2 一维连续型随机变量的函数的分布

设 X 的密度函数为 $f_X(x)$，我们要求出$Y=g(X)$的分布函数 $F_Y(y)$ 。事实上

$$F_Y(y)=P(Y\le y)=P(g(X)\le y)$$

$$=P(X\in D)$$

而 $D=\{x\mid g(x)\le y\}$是一维的集合，故

$$F_Y(y)=P(X\in D)=\int_{x\in D} f_X(x)dx$$

至于Y是不是连续型随机变量，它的密度函数是什么，没有一般性的确定结论，但在某些特殊而又常见的场合，我们可以直接导出 Y 的密度函数 $f_Y(y)$。

2.5.2 一维连续型随机变量的函数的分布

某些情形下，我们可以直接导出 Y 的密度函数 $f_Y(y)$ 。

定理1 X 为连续型随机变量，有概率密度函数

$$f_X(x)(-\infty < x < \infty)$$

若$g(X)$严格单调且处处可导，则$Y = g(X)$也是连续型随机变量。若令其中 $h(y)$ 是 $g(x)$ 的反函数，则Y的密度函数为

$$f_Y(y) = \begin{cases} |h'(y)| f_X[h(y)] & y \in g(x)\text{的值域} \\ 0 & y \in \text{其它} \end{cases}$$

2.5.2 一维连续型随机变量的函数的分布

证： 不妨设 $g(x)$ 严格单调增加，且$-\infty < x < +\infty$时, $\alpha< g(x) <\beta$。显然若 $y \le \alpha$，则$F_Y(y) = 0$；若 $y > \beta$，则$F_Y(y) = 1$，两种情形下都有$f_Y(y) = 0$。当 $\alpha < y < \beta$时，$\{ Y \le y \} = \{ g(X) \le y \} = \{ X \le h(y) \}$，故

$$F_Y(y) = P\left(Y \le y\right) = P\left(X \le h(y)\right) = \int_{-\infty}^{h(y)} f_X(x)dx$$

$F_Y(y)$ 对 y 求导，得到 Y 的概率密度函数

$$f_Y(y) = \begin{cases} h'(y) f_X\left[h(y)\right] & y \in g(x)\text{的值域} \\ 0 & y \in \text{其他} \end{cases}$$

当 $y = f(x)$为严格单调减少时，类似可证：

$$f_Y(y) = \begin{cases} -h'(y) f_X\left[h(y)\right] & y \in g(x)\text{的值域} \\ 0 & y \in \text{其他} \end{cases}$$

二者结合，可得定理的结论。

2.5.2 一维连续型随机变量的函数的分布

Example 2.24 设 $X \sim N(\mu, \sigma^2)$，求 $Y = aX+b$ 的密度函数 ($a \neq 0$)。

解： $y = g(x) = ax+b$ 满足上述定理1中的条件，$h(y) = (y-b)/a$

Y 的密度

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$

$$f_Y(y) = f_X(h(y))|h'(y)|$$

结论：$Y \sim N\left(a\mu + b, (|a|\sigma)^2\right)$

$$= \frac{1}{\sqrt{2\pi}\sigma} Exp\left(-\left(\frac{y-b}{a} - \mu\right)^2 \Big/ 2\sigma^2\right) \times \frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi}|a|\sigma} Exp\left(-\left[y-(a\mu+b)\right]^2 \big/ 2(a\sigma)^2\right)$$

$$= \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\left[y-(a\mu+b)\right]^2 / 2(a\sigma)^2}$$

2.5.2 一维连续型随机变量的函数的分布

Remark1： 若$X\sim N\left(\mu,\sigma^2\right)$，则$Y=aX+b\sim N\left(a\mu+b,a^2\sigma^2\right)$
服从正态分布的随机变量经过线性变换后，得到的随机变量仍然服从正态分布

Remark2：

令 $a\mu+b=0,a^2\sigma^2=1$可以得到

$$a=\frac{1}{\sigma},\ \ b=-\frac{\mu}{\sigma}$$

即$Y=\frac{X-\mu}{\sigma}\sim N(0,1)$

任一正态分布可以转换成标准正态分布。

2.5.2 一维连续型随机变量的函数的分布

Example 2.25（对数正态分布） 设 $X \sim N(\mu, \sigma^2)$，那么随机变量$Y = e^X$就称为参数为μ, σ^2的对数正态随机变量（如果$\ln(Y)$服从正态分布，则随机变量Y服从对数正态分布）。

对数正态分布常用于描述一天结束时的证券价格和前一天结束时价格的比率的分布。即：设S_n为第n天结束时某证券的价格，那么通常假定$\frac{S_n}{S_{n-1}}$服从对数正态分布，因此$S_n = S_{n-1}e^X$，其中X服从正态分布。

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y}\exp\left\{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}\right\}, y > 0$$

2.5.2 一维连续型随机变量的函数的分布

Example 2.26 $X \sim N(0, 1)$, 求 $Y = X^2$ 的概率密度函数。

解： 先求分布函数 $F_Y(y)$，再微分求得 $f_Y(y)$。

$$F_Y(y) = P\left(Y \le y\right)$$

$$= P\left(X^2 \le y\right)$$

$$= \begin{cases} 0 & y \le 0 \\ P\left(-\sqrt{y} < X < \sqrt{y}\right) & y > 0 \end{cases}$$

$$= \begin{cases} 0 & y \le 0 \\ \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx & y > 0 \end{cases}$$

$$f_Y(y) = \begin{cases} 0 & y \le 0 \\ \dfrac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right] & y > 0 \end{cases}$$

当 $X \sim N(0,1)$ 时， $Y = X^2$

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}}\left(\dfrac{1}{\sqrt{2\pi}} e^{-y/2} + \dfrac{1}{\sqrt{2\pi}} e^{-y/2}\right) & y > 0 \\ 0 & y \le 0 \end{cases}$$

$$= \begin{cases} \dfrac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} & y > 0 \\ 0 & y \le 0 \end{cases}$$

此时称 Y **服从自由度为1的$\chi^2(1)$ 分布**， 它也是 **Γ-分布**的特例。

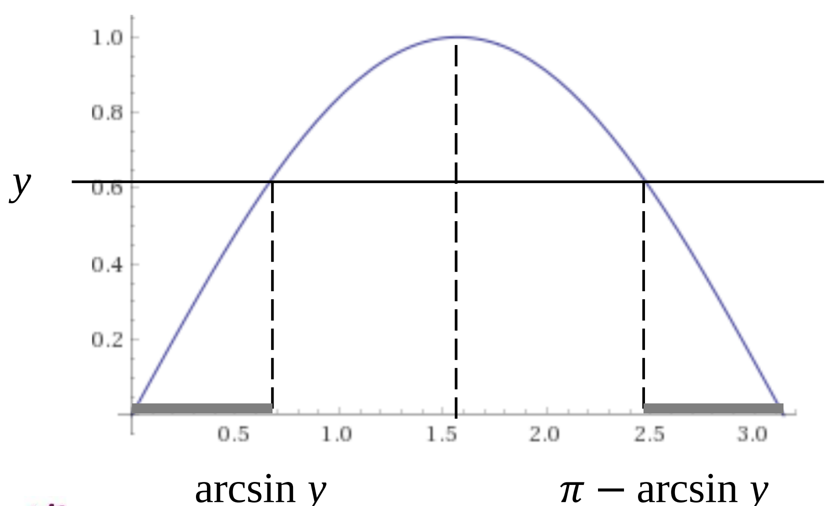
2.5.2 一维连续型随机变量的函数的分布

Example 2.26 设随机变量$X \sim p_X(x) = I_{0<x<\pi}\frac{2x}{\pi^2}$，求$Y$=sin$X$的密度函数。（$I_A$表示指示函数，indicator function）

解： $X \in (0,\pi) \Rightarrow Y \in (0,1)$，在此区间外， $p_Y(y) = 0$

当$Y \in (0,1)$时，使得$Y \le y$的X的取值分为两个区域

$$D_1 = [0, \arcsin y], D_2 = [\pi - \arcsin y, \pi]$$



2.5.2 一维连续型随机变量的函数的分布

Example 2.26 设随机变量$X \sim f_X(x) = I_{0<x<\pi}\frac{2x}{\pi^2}$，求$Y$=sin$X$的密度函数。（$I_A$表示指示函数，indicator function）

解： $X \in (0,\pi) \Rightarrow Y \in (0,1)$，在此区间外， $f_Y(y) = 0$

当$Y \in (0,1)$时，使得$Y \le y$的X的取值分为两个区域

$$D_1 = [0, \arcsin y], D_2 = [\pi - \arcsin y, \pi]$$

于是

$$F_Y(y) = P(Y \le y) = \int_0^{\arcsin y} f_X(x)dx + \int_{\pi-\arcsin y}^{\pi} f_X(x)dx$$

上式两边对y求导，得

$$f_Y(y) = \frac{2\arcsin y}{\pi^2\sqrt{1-y^2}} + \frac{2(\pi - \arcsin y)}{\pi^2\sqrt{1-y^2}} = \frac{2}{\pi\sqrt{1-y^2}}$$

Example 2.27 已知随机变量$X \sim U(0,1), Y = -\frac{\ln(1-X)}{\lambda}, \lambda > 0$，求$Y$的概率密度函数。

- https://www.wolframalpha.com

谢谢！

Nankai University