概率论与数理统计

Probability and Mathematical Statistics

主讲人:

第二章 随机变量及其分布

2.1 随机变量及其分布函数

2.2 离散型随机变量

2.3 连续型随机变量

2.4 随机变量函数的分布

注：第三章是此章维数的提升，注意联系与区别。

2.1 随机变量及其分布

E_1：观察掷骰子出现的点数X.

$\omega=\{$出现的点数为$i\}$ $\xleftrightarrow{\text{对应}}$ $X = i$

建立映射

E_2：抛硬币。观察出现的结果.

$\omega=\{$出现正面$\}$ $\xleftrightarrow{\text{赋值}}$ $X(\omega)=1$

$\omega=\{$出现反面$\}$ $\xleftrightarrow{\text{赋值}}$ $X(\omega)=0$

建立映射

$\{\omega: X(\omega)=0\}=\{X=0\}$
$=\{$反面$\}$

$X:\Omega \xrightarrow{\text{映射}} \mathbf{R}$

$\omega \longrightarrow X(\omega)$

一 随机变量(*Random Variable*)

1. 定义 (Ω,F)为可测空间, 称Ω上的实值函数 $X(\omega)$ 为随机变量，若对任意实数x, $\{\omega: X(\omega) \le x\} \in F$. 简记为 X .

注1. 引入R.V.后，事件就通过R.V.的取值表示。

例：E_1:$\{X=3\},\{X\le 3\}$等.

注2 若F为最大σ域，则Ω上的实值函数 $X(\omega)$都是随机变量.

2. R.V. X 的分类

R.V.

- 离散型: X的所有可能取值为有限个或可数个 (discrete r.v.)
- 非离散型 $\xrightarrow{\text{之一}}$ 连续型 (continuous r.v.)

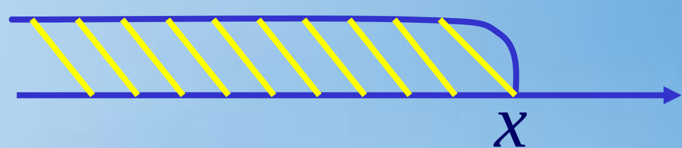
二 分布函数(distribution function)

1. 定义　设有R.V.X. 称

$$F(x)=P\{X\le x\}\ ,\ x\in\mathbb{R}$$

为R.V.X 的分布函数。

注1



2 设$x_1, x_2\in\mathbb{R}, x_1\le x_2$,则

$$P(x_1<X\le x_2)=P(X\le x_2)-P(X\le x_1)=F(x_2)-F(x_1)$$

2 分布函数的性质

（1）单调不减性：若 $x_1<x_2$，则 $F(x_1)\le F(x_2)$；

（2）有界性： $0\le F(x)\le 1$，且

$$F(+\infty):=\lim_{x\to+\infty}F(x)=1,$$

$$F(-\infty):=\lim_{x\to-\infty}F(x)=0.$$

（3）右连续性：对任意的实数 x_0，有

$$\lim_{x\to x_0^+}F(x)=F(x_0).$$

注：上述性质也是鉴别一个函数是否是某个随机变量的分布函数的充分条件.

2.2 离散型随机变量（D.R.V.）

一 D.R.V.及其分布列

1. 分布列定义　　2. 分布列的性质

二 常见D.R.

应用场合

在一定时间间隔内：
电话总机接到的电话次数；
一匹布上的疵点个数；
大卖场的顾客数...

1. 两点分布

3. 泊松分布　　4. 几何分布

5. 超几何分布

一 离散型随机变量及其分布列

1.定义 设D.R.V. X 的所有可能取值为$x_1, x_2, \cdots,$ $x_k, \cdots$，p_k 是X 取x_k的概率，称

$$P(X=x_k)=p_k, \quad k=1,2,3,\cdots$$

为R.V. X 的**概率分布列**，简称**分布列(律)**.

注1 分布列的表格形式：

X	x_1	x_2	$\cdots$	x_k	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

2 概率分布图：横坐标X的取值，纵坐标为对应概率

3 **概率分布**(简称**分布**)指R.V.的分布列或分布函数.

思考题：D.R.V.的分布列和分布函数能互相决定吗？

此时，$F(x)=P(X\le x)=\sum\limits_{x_k\le x} p_k$

2 分布列的性质

（1）非负性： $p_k \geq 0$， $k = 1, 2, \ldots$

（2）规范性： $\sum_{k=1}^{\infty} p_k = 1.$

注：上述性质也是鉴别一串非负实数是否是某随机变量的概率分布的充分条件.

例 将4个球随机地投入三个盒子中，X 表示有球的盒子的数目，试求

(1) X 的分布列;　(2) X 的分布函数, 并作图.

(3) $P(2\le X<5/2)$, $P(2<X<5/2)$, $P(1\le X\le 5/2)$.

解 (1)显然 X 可能的取值为 1, 2, 3.

$$P(X=1)=3/3^4=1/27;$$

$$P(X=2)=\frac{C_3^2(2^4-2)}{3^4}=14/27;=\frac{C_3^2(C_4^1+C_4^3+C_4^2)}{3^4}$$

$$P(X=3)=1-P(X=1)-P(X=2)=12/27;=\frac{C_4^2\,3!}{3^4}$$

X	1	2	3
p_k	1/27	14/27	12/27

X	1	2	3
p_k	1/27	14/27	12/27

(2) $F(x) = P(X \le x)$

$$F(x) = \sum_{x_k \le x} p_k$$

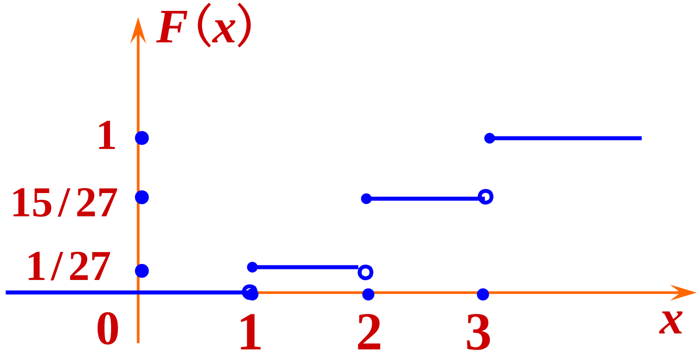
当 $x<1$ 时，$F(x) = P(\Phi) = 0$;

当 $1 \le x < 2$ 时，$F(x) = P(X=1) = 1/27$;

当 $2 \le x < 3$ 时，$F(x) = P(X=1) + P(X=2) = 15/27$;

当 $x \ge 3$ 时，$F(x) = P(X=1) + P(X=2) + P(X=3) = 1$,

$$\therefore F(x) = \begin{cases} 0, & x<1; \\ 1/27, & 1 \le x < 2; \\ 15/27, & 2 \le x < 3; \\ 1, & 3 \le x. \end{cases}$$



注：在点 x_k 处有跳跃，跃度为 p_k

X	1	2	3
p_k	1/27	14/27	12/27

(3) $P(2 \le X \le 5/2) = P(X=2) = 14/27$;

$P(2 < X < 5/2) = P(\Phi) = 0$;

$P(1 \le X \le 5/2) = P(X=1) + P(X=2) = 15/27$.

小结

- 了解随机变量的概念及随机变量的分类
- 深入理解分布函数的概念和性质
- 深入理解D.R.V.及分布列的概念和性质，会熟练求解相关参数，事件的概率和分布函数

二 常见的D.R.V及其分布

1. **两点分布** 若R.V. X 的分布列为

X	0	1
P	$1-p$	p

$(0 < p < 1)$

则称X 服从两点分布，记为 $X \sim (0—1)$ 分布

或 $X \sim B(1, p)$.

2. 二项分布 (binomial distribution)

若R.V. X 的分布列为

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

$$(0<p<1)$$

称 X 服从参数为 n , p 的二项分布,

记为 $X\sim B(n,p)$.

注: $\sum_{k=0}^{n} C_n^k p^k q^{n-k}=(p+q)^n=1$

(1) n 重(独立)伯努利 (Bernouli) 试验

- 在相同条件下进行 n 次重复试验
- 各次试验的结果互不影响
- 每次试验只有两个种的结果 A 与 $\overline{A}$

(2) 问题

在n重伯努利试验中，用 X 表示事件 A 发生的次数，则 X 是一个随机变量，且其可能的取值为 0, 1, 2, ⋯ n，

分布列： $P(X=k)$ =？（$k=0, 1, 2, \cdots n$）

分析：记$P(A)=p$， $P(\bar{A})=1-p=q$,

A_i 表示“第 i 次试验中 A 发生了”则 $P(A_i)=p$,

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

(3) 二项分布的注释

1. n=1时，为两点分布.

2. 几何模型：n重伯努利试验中, 记事件A发生的次数为X，则$P(X=k)= C_n^k p^k q^{n-k}$ (即A恰好发生k次的概率).

3. $P(X=k)=:f(k)$的单调性：随k先增后减.

当$(n+1)p$是整数, 在$k= (n+1)p$及$(n+1)p-1$处达最大;

当$(n+1)p$不是整数, 在$k= [(n+1)p]$ 处达到最大.

中心项

例 **有 9 位工人间歇地使用电力**，假设在任一时刻每位工人都以 0. 2 的概率需要一个单位的电力，并且各位工人工作相互独立.问最大可能有多少位工人同时需要供应一个单位的电力?

解 设 X 为任一时刻同时需要供应一个电力的工人数，

则 $X \sim B(9, 0.2)$.

n 重(独立)伯努利 (Bernouli) 试验

在相同条件下进行 n 次重复试验，各次试验的结果互不影响，每次试验只有两个种的结果 A 与 $\overline{A}$.

判断方法：n 次试验，满足

① 每次试验只有两种结果 A 与 $\overline{A}$

② 试验可重复，$P(A)=p$

③每次试验独立进行

设 X 为 A 发生的次数，则 $X \sim B(n, p)$

n 次试验 1. A与$\overline{A}$
2. $P(A)=p$
3. 独立

设X为A发生的次数，则$X\sim B(n,P(A))$

例 有 9 位工人间歇地使用电力，假设在任一时刻每位工人都以 0. 2 的概率需要一个单位的电力，并且各位工人工作相互独立.问最大可能有多少位工人同时需要供应一个单位的电力?

解 设X为任一时刻同时需要供应电力的工人数，

则 $X \sim B(9, 0.2)$.

$(n+1)p = (9+1)\times 0.2 = 2$,

为整数

故最大可能有1位工人或2位工人同时需要供应一个单位的电力。

且最大可能性为$P(X=2)=C_9^2(0.2)^2(0.8)^7=0.3020$.

n 次试验　1. A与$\overline{A}$

2. $P(A)=p$

3. 独立

设X为A发生的次数，则$X \sim B(n,P(A))$

例　某车间里有12台车床，每台车床由于工艺需要时常需要停车.设每台车床的停(开)是相互独立的,每台车床在任一时刻处于停车状态的概率为1/3，求任一时刻

（1）车间里恰有 2 台车床处于停车状态的概率；

（2）处于停车状态的车床不少于 2 台的概率.

解　设 X 为在任一时刻车床停车的台数，则 $X \sim B(12, 1/3)$.

(1)所求为　$P(X=2)=C_{12}^{2}\cdot(\frac{1}{3})^{2}\cdot(1-\frac{1}{3})^{12-2}=0.1272$；

(2)所求为　$P(X\geq 2)=1-P(X<2)$

$$=1-P(X=0)-P(X=1)$$

$$=1-C_{12}^{0}\cdot(\frac{1}{3})^{0}\cdot(\frac{2}{3})^{12-0}-C_{12}^{1}\cdot(\frac{1}{3})^{1}\cdot(\frac{2}{3})^{12-1}$$

$$=0.9461.$$

例 一门大炮对目标进行轰击，假定此目标必须被击中r次才能被摧毁。若每次击中目标的概率为p $(0 < p < 1)$，且各次轰击相互独立，一次一次地轰击直到摧毁目标为止。求所需轰击次数X的概率分布.

解 $P(X=k)=P$(前$k-1$次击中$r-1$次, 第k次击中目标)

$=P$(前$k-1$次击中$r-1$次)P(第k次击中目标)

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

$$=C_{k-1}^{r-1}p^{r}(1-p)^{k-r},\qquad k=r,r+1,\cdots$$

负二项分布

例 某人射击，设每次射击的命中率为0. 001，独立射击5000次，至少命中两次的概率是多少？

解 设X为命中的次数，则$X \sim B(5000, 0.001)$.

所求概率为$P(X \geq 2) = 1 - P(X=0) - P(X=1)$

$$= 1 - \sum_{k=0}^{1} C_{5000}^{k} (0.001)^k (0.999)^{5000-k} = 0.9574.$$

注 $P(X=k) \approx \dfrac{\lambda^k e^{-\lambda}}{k!}, \lambda = np.$

3.泊松分布(*Poisson distribution*)

(1)泊松定理 在n重伯努利试验中,设事件A出现的概率为p_n

如果 $$\lim_{n\to\infty} np_n = \lambda \ (\lambda > 0),$$

则 $$\lim_{n\to\infty} C_n^k \, p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, \quad k = 0, 1, 2, \cdots$$

应用: $$C_n^k \, p^k (1-p)^{n-k} \approx \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, \quad (\lambda = np).$$

思考: $\lambda > 0$, $\boxed{\dfrac{\lambda^k e^{-\lambda}}{k!}, \ k = 0, 1, 2, \cdots}$

是不是某个R.V.的分布列?

1^0 $$\frac{\lambda^k e^{-\lambda}}{k!} \geq 0$$

2^0 $$\boxed{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}} e^{-\lambda} = 1$$

(2) 泊松分布 (*Poisson distribution*)

若R.V. X 的分布列为

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

称 X 服从参数为 λ 的泊松分布, 记为 $X\sim P(\lambda)$.

注　泊松分布也有中心项.

当λ是整数，在$k=\lambda$ 及 $\lambda-1$处达最大;

当λ不是整数，在$k=[\lambda]$ 处达到最大.

x	……	λ=4.5	λ=5.0
⋮	$\sum_{r=x}^{\infty}\frac{\lambda^r}{r!}e^{-\lambda}$	⋮	⋮
10	……	0.017093	0.031828
11	……	0.006669	0.013695
12	……	0.002404	0.005453
13	……	0.000805	0.002019
⋮		⋮	⋮

例 由某商店过去的销售记录可知，某种商品每

：件）可用参数为λ=5的泊松

9%以上的把握保证不脱销，问

进货多少件？

销售量(件)，则X~P(5).

则

9 $\sum_{k=0}^{N}\frac{5^k}{k!}e^{-5}\geq 0.99$

$$\sum_{k=N+1}^{\infty}\frac{5^k}{k!}e^{-5}\leq 0.01$$

查表得N+1≥12，即N≥11.

4. 超几何分布(*hypergeometric distribution*)

(1)问题 N件产品中有M件次品，从中任取n件(无放回)。

记X为这n件产品中的次品数，求X的分布。

解 $$\boxed{P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \quad k=0,1,\cdots,\min\{n,M\}.}$$

(2)超几何分布： 若R.V.X的分布列如上，其中$N \geq M$，$N \geq n$.

注 N很大，n很小时. 无放回近似有放回

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \overset{p=M/N}{\approx} C_n^k p^k (1-p)^{n-k}$$

5. 几何分布(*geometric distribution*)

(1)问题 设X是一个无穷次独立伯努利试验序列中事件A首次发生的次数，求X的分布。

解

$$P(X=k)=(1-p)^{k-1}p\,,\qquad k=1,2,3,\cdots.$$

(2)几何分布：若R.V.X的分布列如上, 其中$0<p<1$.

(3) 特征(无记忆性)：R.V.X取值自然数. 则

X服从几何分布$\Longrightarrow P(X>m+n\mid X>m)=P(X>n)$,

其中m,n为任意自然数.

$P(X>n)=q^n$

综合例子

例 设某高速公路一年内发生恶性交通事故的次数 $X\sim P(\lambda)$, 而每次交通事故造成一位司机死亡的概率为 p，且一次事故至多造成一位司机死亡。求一年内恰有 r 位司机因车祸丧生的概率。

解：A_n={一年内恰有 n 次恶性事故}={$X=n$}, n=0, 1, 2, ⋯

B={一年内恰有 r 位司机丧生}

$$P(B)=\sum_{n=0}^{\infty}P(B|A_n)P(A_n)$$

$$=\sum_{n=r}^{\infty}C_n^r p^r(1-p)^{n-r}\frac{\lambda^n e^{-\lambda}}{n!}=\cdots \qquad =\frac{(\lambda p)^r e^{-\lambda p}}{r!}$$

注：用 Y 表示一年内因车祸丧生的司机人数，则 $Y\sim P(\lambda p)$

综合例子

例 周末到达商场的顾客数$N\sim P(\lambda)$，λ=500(人)，其中女性顾客占70%(=p). 求到达商场的女性顾客N_F的分布列。假设每位顾客到达商场相互独立。

解：$P(N_F=k)=\sum_{n=0}^{\infty}P(N_F=k\mid N=n)P(N=n)$ $\quad k=0, 1, 2, \cdots$

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

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

$$=\frac{(\lambda p)^k e^{-\lambda p}}{k!}$$

注：$N_F\sim P(\lambda p)$

小结

- 熟练掌握五种常见D.R.V.的分布列或分布函数
- 熟练掌握常见D.R.V.解决的问题模型，尤其会判断n重伯努利试验，并设出相应的随机变量
- 注意与全概率公式结合的应用

泊松定理的证明

欣赏

若 $\lim\limits_{n\to\infty} np_n = \lambda$（$\lambda > 0$），则

$$\lim_{n\to\infty} C_n^k p_n^{\,k}(1-p_n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0,1,2,\cdots$$

证: $C_n^k p_n^{\,k}(1-p_n)^{n-k} = n(n-1)\ldots(n-k+1)\dfrac{(np_n/n)^k}{k!}(1-\dfrac{np_n}{n})^{n-k}$

$$= 1\cdot(1-\frac{1}{n})\ldots(1-\frac{k-1}{n})\ \frac{(np_n)^k}{k!}[(1-\frac{np_n}{n})^{-\frac{n}{np_n}}]^{-np_n}(1-p_n)^{-k}$$

$$\xrightarrow{n\to\infty} \frac{\lambda^k}{k!}e^{-\lambda}$$

实际推断原理

欣赏

例 **设鸡群中感染某无传染疾病的概率为20%**，新发现的某血清可能对预防此疾病有效，**对 25 只健康鸡注射了这种血清，注射后发现只有一只鸡感染此病**，问这种血清是否有用？

解 若血清无作用，则每只鸡染病的概率为20%，且25只鸡中感染此病的鸡数 X 是一个服从二项分布的随机变量，

即 $X \sim B(25, 0.20)$，

我们在无作用条件下，通过求事件“至多有一只鸡感染此病”的概率来判定血清是否有用.

即求 $P(X \le 1) = P(X=0) + P(X=1)$

$$= C_{25}^{0}(0.20)^{0}(0.80)^{25} + C_{25}^{1}(0.20)^{1}(0.80)^{25-1}$$

$$= 0.0274 < 3\%$$

小概率事件

一个小概率事件现在却发生了，我们有理由对假设条件产生怀疑，即血清是有作用的.

2.3 连续型随机变量（C.R.V.）

一 C.R.V.及其概率

1. C.R.V.及概率

2. 概率密度的性质

应用场合 用指数分布描述的实例有：

随机服务系统中的服务时间；

电话问题中的通话时间；

无线电元件的寿命；

动物的寿命.

常作为各种“寿命”分布的近似.

二 常见C.R.

1. 均匀分布

应用场合 可用正态变量描述的实例非常多：

各种测量的误差； 人的生理特征；

工厂产品的尺寸； 农作物的收获量；

海洋波浪的高度； 学生们的考试成绩…

3. 正态分布

三 混合型R.V.

$$P(a-h<X\leq a)=F(a)-F(a-h)$$

若$F(x)$连续 $\longrightarrow 0 \quad (h\to 0)$

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

一 C.R.V.及其概率密度

1. C.R.V.及概率密度的定义

设R.V.X的分布函数为$F(x)$，若存在非负可积函数 $f(x)$，使得

$$\boxed{F(x)=\int_{-\infty}^{x} f(t)dt, \quad x\in \mathrm{R},}$$

则称X为**C.R.V.**，$f(x)$称为X的**(概率)密度函数**。

注1 C.R.V.的分布函数(绝对)连续.

注2 几何意义：$F(x)$为以$f(x)$为曲顶, $(-\infty, x]$为底的面积.

注3 任意$a\in\mathrm{R}$，$P(X=a)=0$.

$$0\leqslant P(X=a)\leqslant P(a-\Delta x<X\le a)=F(a)-F(a-\Delta x)\xrightarrow[\Delta x\to 0]{} 0$$

- 零概率事件不一定是不可能事件，同样，概率为1的事件也不一定是必然事件。
- $P(x_1<X\le x_2)=P(x_1<X<x_2)=P(x_1\le X<x_2)=P(x_1\le X\le x_2)$.

注4 $P(x_1<X\le x_2)=F(x_2)-F(x_1)=\int_{-\infty}^{x_2}f(t)dt-\int_{-\infty}^{x_1}f(t)dt$

$$=\int_{x_1}^{x_2}f(t)dt$$

一般，$B\subset \mathrm{R}$， $\boxed{P(X\in B)=\int_B f(x)dx.}$

注5 若$f(x)$在x处连续，则 $f(x)=F'(x)$.

2 概率密度的性质

（1）非负性：　$f(x) \geq 0$，　$x \in \mathrm{R}$

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

注：上述性质也是鉴别一个实函数是否是某C.R.V.的概率密度的充分条件.

例 设X的密度函数 $f(x)=\begin{cases} x, & 0\le x<1 \\ A-x, & 1\le x<2 \\ 0, & \text{其它}. \end{cases}$

试求(1)常数A；(2) $P(-1<X<3/2)$；(3) $F(x)$.

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

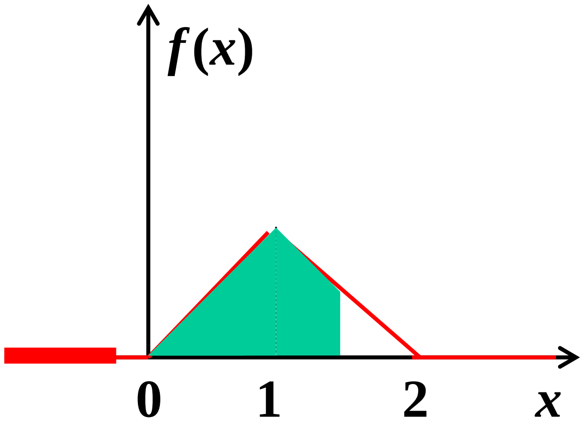
$$\int_0^1 xdx+\int_1^2 (A-x)dx = A-1 = 1$$

$$\Rightarrow \ A=2$$

（2）$P(-1<X<3/2)=\int_{-1}^{3/2} f(x)dx$

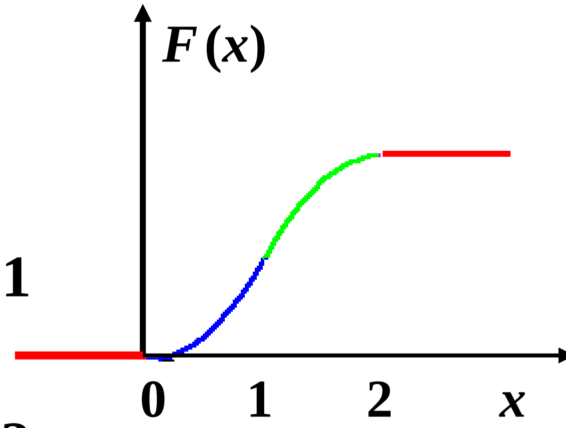
$$=\int_0^1 xdx+\int_1^{3/2}(2-x)dx=1/2+1-5/8=7/8.$$

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



（3）

$$F(x)=\begin{cases} \int_{-\infty}^{x} f(t)dt=0, & x<0 \\ \int_{-\infty}^{x} f(t)dt=\int_0^x tdt=\frac{x^2}{2}, & 0\le x<1 \\ \int_0^1 tdt+\int_1^x (2-t)dt=1-\frac{1}{2}(2-x)^2 & 1\le x<2 \\ 1 & 2\le x \end{cases}$$



例 C.R.V.X的分布函数 $F(x)=\begin{cases}0, & x<0\\ A\sin x, & 0\le x<\pi/2\\ 1, & x\ge \pi/2.\end{cases}$

试求(1)常数A；(2) $P(|X|<\pi/6)$；(3)概率密度 $f(x)$.

解（1）$\because \lim\limits_{x\to\pi/2} F(x)=F(\pi/2)$

$\therefore A\sin(\pi/2)=1 \Rightarrow A=1.$

（2） $P(|X|<\pi/6)=F(\pi/6)-F(-\pi/6)=1/2.$

（3） $f(x)=F'(x)=\begin{cases}\cos x, & 0\le x<\pi/2,\\ 0, & \text{其它}.\end{cases}$

综合例子

例　有一批晶体管，已知每只的使用寿命 X (单位：小时) 为连续型随机变量，其概率密度函数为：

$$f(x)=\begin{cases} c/x^2, & x>1000 \\ 0, & \text{其它}. \end{cases} \quad （c\text{为常数}）$$

（1）求常数c.

(2）已知一只收音机上装有3只这样的晶体管，每只晶体管能否正常工作相互独立, 求在使用的最初1500小时只有一个损坏的概率.

解（1） $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1 \Rightarrow \int_{1000}^{+\infty} \frac{c}{x^2}\mathrm{d}x=1 \Rightarrow c=1000.$

(2) A ={一只晶体管的寿命小于1500小时}

$$P(A)=P(0\le X<1500)=\int_{1000}^{1500} 1000/x^2\mathrm{d}x=1/3.$$

设Y为三只晶体管中A发生的个数, 则$Y\sim B(3,1/3)$

$$P(Y=1)=C_3^1(1/3)(2/3)^2=4/9.$$

二 常见的C.R.V.及其分布

1. 均匀分布(*uniform distribution*)

若R.V. X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b],\\ 0, & 其它,\end{cases}$$

则称X 服从$[a,b]$上的均匀分布，记为 $X\sim U[a,b]$.

注 设$(c,c+l)\subset(a,b)$, $P(c<X<c+l)=\int_c^{c+l}\frac{1}{b-a}dx=\frac{l}{b-a}$

即X的取值在 $(a，b)$ 内任何长为 l的小区间的概率只与其长度成正比,与小区间的位置无关．这正是几何概型的情形.

例 数值计算中由于处理小数位数而四舍五入引起的舍入误差一般可看做是一个服从均匀分布的R.V.，现考虑一种要求保留整数的计算，问这一计算的舍入误差X在-0.1到+0.1之间的概率是多少？

解 由题意，$X \sim U[-0.5,+0.5]$. $X \sim U(-0.5,+0.5)$

所求概率为

$$P(-0.1 \le X \le 0.1) = \int_{-0.1}^{0.1} \frac{1}{0.5-(-0.5)} dx = 0.2.$$

2. 指数分布(*exponential distribution*)

(1) 定义 若R.V. X 的概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0,\\ 0, & \text{其它},\end{cases}\quad(\lambda>0)$$

则称X 服从参数为λ的指数分布，记为 $X \sim E(\lambda)$.

(2) 特征(无记忆性)：R.V.X取值非负实数. 则

X服从指数分布 $\Rightarrow P(X>x+y\mid X>x)=P(X>y)$,

对任意实数$x, y>0$.

$x>0, P(X>x)=e^{-\lambda x}$

例　设一次打电话所用的时间$X \sim E(\lambda)$,$\lambda=0.1$(分钟)。若某人刚好在你前面走进电话亭,求你等待10–20分钟之间的概率。

解

$$P(10 \le X \le 20) = \int_{10}^{20} 0.1 e^{-0.1x} \mathrm{d}x$$

$$= e^{-1} - e^{-2}.$$

例 设一大型设备在任何长为t时间内发生故障的次数$N(t)\sim P(\lambda t)$.

(1)求相继两次故障之间的时间间隔T的概率分布;

(2)求在设备无故障工作8小时的条件下，再无故障工作8小时的概率P.

解（1）$F(t)=P(T\le t)=1-P(T>t)=1-P(N(t)=0)$

$$=1-\frac{(\lambda t)^0}{0!}e^{-\lambda t}=1-e^{-\lambda t}\qquad t\geqslant 0$$

$t<0$时，$F(t)=P(\Phi)=0$ 即 $T\sim E(\lambda)$

（2）$P=P(T>16\,|\,T>8)=P(T>8)=e^{-8\lambda}$.

无记忆性

3. 正态分布(*normal distribution*)

(1) 定义 若R.V. X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},x\in R,\quad (\mu\in R,\sigma>0)$$

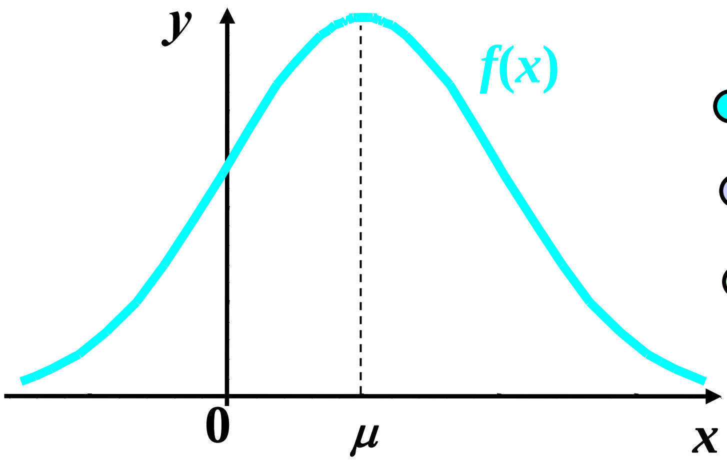
则称X 服从参数为μ,σ^2的正态分布，记 $X \sim N(\mu,\sigma^2)$.

特别的，若$\mu=0,\sigma=1$，则称X 服从标准正态分布。并分别以$\varphi(x)$和$\Phi(x)$记其的概率密度和分布函数。

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\qquad \Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt.$$

可查表

(2) 密度函数$f(x)$的性质



- $x=\mu$为对称轴，此处取到最大值
- $x\to\pm\infty$，以x轴为渐近线
- σ固定，μ变动，图形沿x轴平移
  μ固定，σ变动，σ越小图形越陡

(3) 分布函数的性质

$\boxed{\Phi(-x)=1-\Phi(x)}$ 特别，$\boxed{\Phi(0)=1/2}$

$$\boxed{X\sim N(\mu,\sigma^2),\text{ 则 }F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right).}$$

例　公共汽车门的高度是按男子与车门顶碰头的机会在0.01以下来设计的。设男子身高服从μ=170cm，σ =6cm的正态分布，即$X\sim N(170,6^2)$. 问车门应如何确定？

解 设车门高度为$h$cm，

$$P(X>h)<0.01 \Rightarrow P(X\le h)>0.99$$

$$\Rightarrow F(h)=\Phi(\frac{h-170}{6})>0.99$$

查表 $$\Rightarrow \frac{h-170}{6}>2.33$$

$$\Rightarrow h>183.98$$

例（3σ原理） 设$X \sim N(\mu, \sigma^2)$, 求 $P(|X-\mu|<3\sigma)$

解 $P(|X-\mu|<3\sigma)=P(\mu-3\sigma<X<\mu+3\sigma)$

$$=F(\mu+3\sigma)-F(\mu-3\sigma)$$

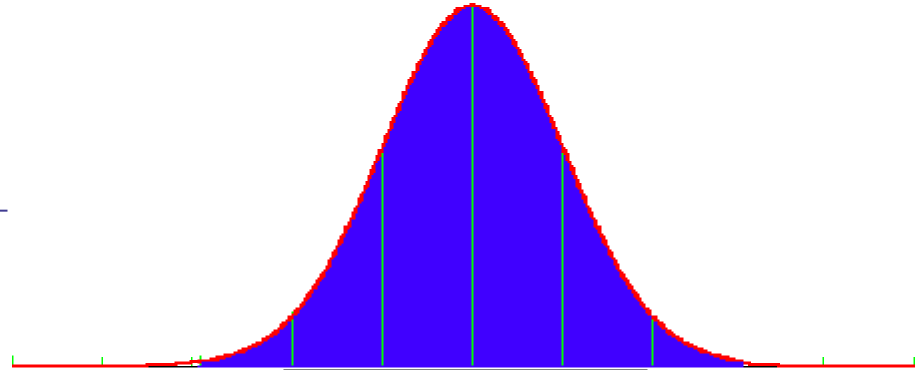
$$=\Phi\left(\frac{\mu+3\sigma-\mu}{\sigma}\right)-\Phi\left(\frac{\mu-3\sigma-\mu}{\sigma}\right)=\Phi(3)-\Phi(-3)$$

$$=\Phi(3)-(1-\Phi(3))$$

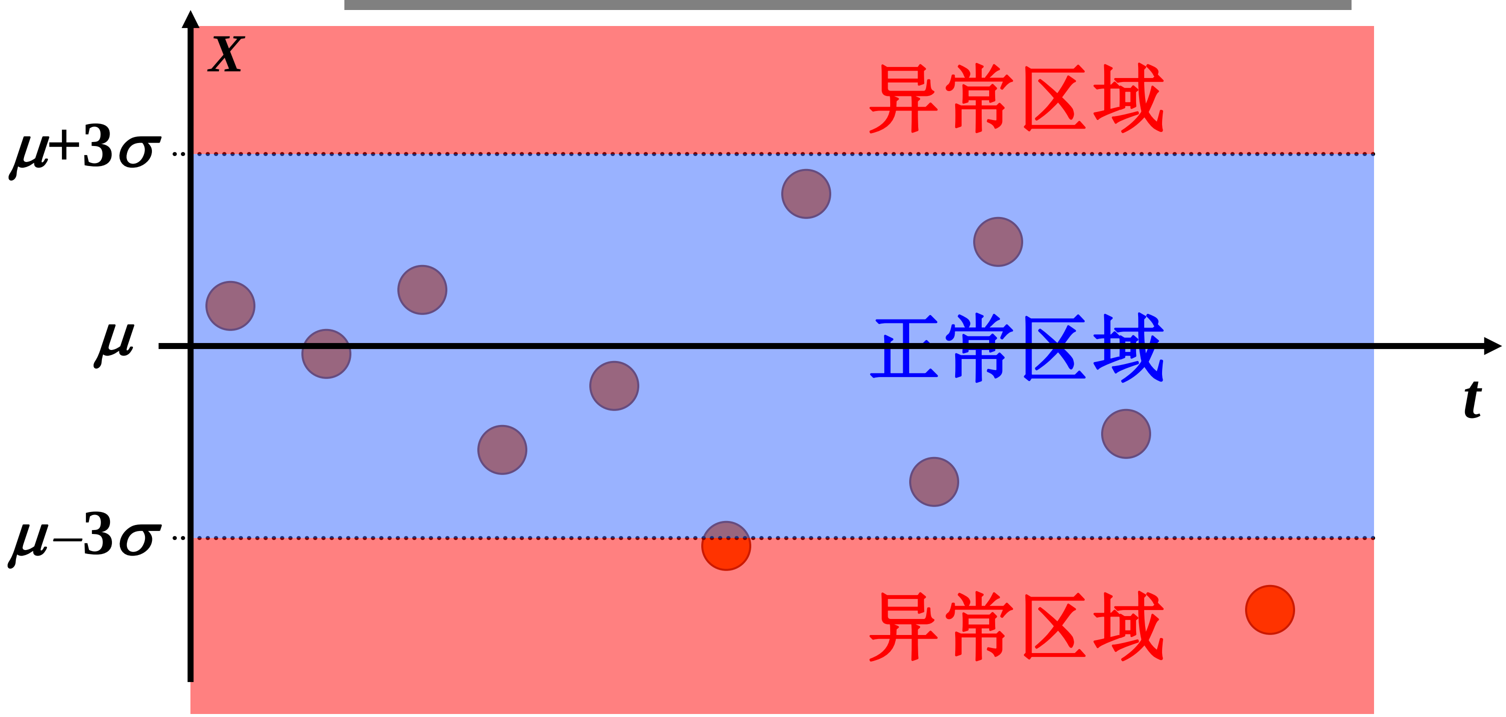
$$=2\Phi(3)-1$$

$$=2\times 0.9987-1 \;=0.9974$$

3σ原则



a	1	2	3
$P(\|X-\mu\|<k\sigma)$	0.6826	0.9545	0.9974



综合例子

例　设X~N(3,1). 现对X进行3次独立观测，试求至少有两次观测值大于3的概率P.

解　A={观测值大于3} ={X>3}

设Y为A发生的次数，则Y~B(3,P(A)=p)，

$$p = P(X > 3) = 1 - F(3) = 1 - \Phi(\frac{3-3}{1}) = 1 - \Phi(0) = 1/2.$$

$$\mathrm{P} = P(Y \geq 2) = P(Y = 2) + P(Y = 3)$$

$$= C_3^2 p^2 (1 - p) + C_3^3 p^3$$

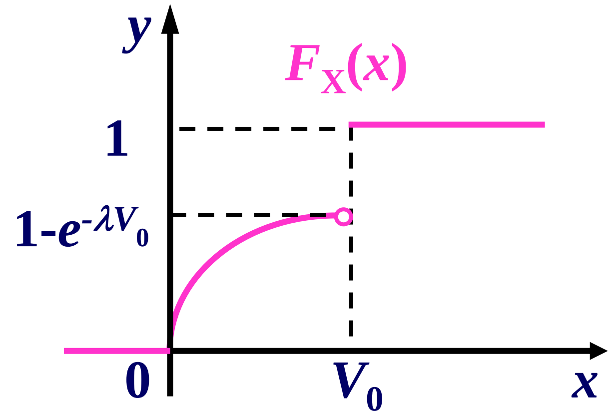
$$= C_3^2 (1/2)^2 1/2 + C_3^3 (1/2)^3 = 1/2.$$

三 混合型R.V.

例 某电路受外界刺激，电压V随机波动，且$V\sim E(\lambda)$. 先用电压表进行测量，电压表的最大读数为V_0. 以X记电压表的读数，求X的分布函数。

解 $X=\min\{V, V_0\}$.

$$F_X(x)=P(X\le x)=\begin{cases}0, & x<0,\\ 1-e^{-\lambda x}, & 0\le x<V_0,\\ 1, & V_0\le x.\end{cases}$$



$$F_X(x)=(1-e^{-\lambda V_0})F_1(x)+e^{-\lambda V_0}F_2(x),$$

$$F_1(x)=\begin{cases}0, & x<0,\\ \dfrac{1-e^{-\lambda x}}{1-e^{-\lambda V_0}}, & 0\le x<V_0,\\ 1 & V_0\le x.\end{cases}\qquad F_2(x)=\begin{cases}0, & x<V_0,\\ 1, & x\ge V_0.\end{cases}$$

小结

- 理解C.R.V.，概率密度的概念及性质
- 熟练掌握常见C.R.V.的概率密度或分布函数及其性质，并求解相关应用问题
- 注意C.R.V.与二项分布结合的应用
- 了解混合型R.V.

2.4 R.V.的函数的分布

设$y=g(x)$为一通常函数，令$Y=g(X)$, 其中X为R.V., 若Y也为R.V.，称 Y为R.V. X的函数。

问题：已知X的分布，如何求Y的分布？

一 D.R.V.函数的分布

二 C.R.V.函数的分布

1. 分布函数法　　　2. 定理法

一 D.R.V.函数的分布

例 已知X的分布列，求$Y_1=2X+1$，$Y_2=X^2$的分布。

X	-2	-1	0	1	2
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
$Y_1=2X+1$	-3	-1	1	3	5
$Y_2=X^2$	4	1	0	1	4

解

Y_1	-3	-1	1	3	5
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$

Y_2	0	1	4
P	$\frac{4}{10}$	$\frac{4}{10}$	$\frac{2}{10}$

二 C.R.V.函数的分布

1. 分布函数法

例 设X的概率密度为$f_X(x)$，求$Y=X^2$的概率密度。

解 $F_Y(y)=P(Y\le y)=P(X^2\le y)$

$y\le 0$时， $F_Y(y)=0;$

$y>0$时， $F_Y(y)=P(-\sqrt{y}\le X\le\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y})$

$f_Y(y)=F_Y'(y)$

$$=\begin{cases}\frac{1}{2}f_X(\sqrt{y})y^{-\frac{1}{2}}+\frac{1}{2}f_X(-\sqrt{y})y^{-\frac{1}{2}}, & y>0\\ 0, & y\le 0\end{cases}$$

2. 定理法

设C.R.V.X有密度函数$f_X(x)$. 设$y=g(x)$严格单调且反函数$x=g^{-1}(y)$有连续导数，则$Y=g(X)$也是C.R.V.，其密度函数为

$$f_Y(y)=\begin{cases} f_X[g^{-1}(y)]\,|[g^{-1}(y)]'|, & \alpha<y<\beta \\ 0\quad, & \text{其它} \end{cases}$$

其中(α, β)为$g(x)$的值域。

$$f_Y(y)=\begin{cases} f_X[g^{-1}(y)]\,|[g^{-1}(y)]'|, & \alpha<y<\beta \\ 0\quad, & \text{其他} \end{cases}$$

例 已知$X\sim N(\mu,\sigma^2)$, 求$Y=aX+b(a\neq 0)$的概率密度。

解: $y=g(x)=ax+b$严格单调,

$$x=g^{-1}(y)=\frac{y-b}{a},\quad [g^{-1}(y)]'=\frac{1}{a}.$$

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

即 $\boxed{X\sim N(\mu,\sigma^2)\Rightarrow aX+b\sim N(a\mu+b,\ (a\sigma)^2)}$

特别 $\dfrac{X-\mu}{\sigma}\sim N(0,\ 1)$

推论

设R.V.X有密度函数$f_X(x)$. 设$y=g(x)$在不重叠区间I_1, I_2 …, 逐段满足上述定理条件，以$g_i^{-1}(y)$, $(g_i^{-1}(y))'$记$y=g(x)$在I_i上的反函数及导数(i=1,2,…)，则$Y= g(X)$也是C.R.V.，且

$$f_Y(y)=\begin{cases}\sum\limits_{\{i:y\in g(I_i)\}} f_X[g_i^{-1}(y)]\,|[g_i^{-1}(y)]'|, & \alpha<y<\beta\\ 0\quad, & \text{其它}\end{cases}$$

其中(α, β)为$g(x)$的值域。

小结

- 熟练求解D.R.V.函数的分布列
- 熟练掌握C.R.V.函数的概率密度求解方法

  1. 分布函数法(万能方法)

  2.定理法(公式法)，注意定理条件

作 业 选 讲

1.(练习一5(3)) $A_1, A_2, \cdots, A_n$为n个事件，$A_1 \cup A_2 \cup \cdots \cup A_n$表示成个$n$两两不相容事件之和。

$$A_1 \cup A_2 \cup \cdots \cup A_n$$
$$= A_1 + (A_2 - A_1) + (A_3 - A_1 - A_2) + \cdots + (A_n - A_1 - \cdots - A_{n-1})$$

作业选讲

2.(练习二 5) n个座位依次从1号编到n号，将1至n的n个号码分给n个人，每人一个号码，这n个人随意坐到座位上，求至少有一个人手里的号码恰与座位号码相符的概率。

解 A_i={第i个人坐到第i号座位}

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)$$

$$=\sum_{i=1}^{n} P(A_i) - \sum_{1\le i\ne j\le n} P(A_i A_j) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n)$$

$$= n \cdot \frac{(n-1)!}{n!} - C_n^2 \frac{(n-2)!}{n!} + \cdots + (-1)^{n-1}\frac{1}{n!}$$

$$= 1 - \frac{1}{2!} + \cdots + (-1)^{n-1}\frac{1}{n!}.$$

作业选讲

3.(练习三 5) 设共有10张彩票，其中只有2张可获奖，甲乙丙三人依次抽取彩票一张，规则如下：每人抽出后不放回，但补入两张与所抽彩票不同的彩票。问：甲乙丙三人中谁中奖的概率大?

解 A={甲中奖}，B={乙中奖}，C={丙中奖}

$$P(A)=2/10,$$

$$P(B)=P(B\mid A)P(A)+P(B\mid \bar{A})P(\bar{A})=\frac{1}{11}\cdot\frac{2}{10}+\frac{4}{11}\cdot\frac{8}{10}=\frac{34}{110}.$$

$$P(C)=P(C\mid AB)P(AB)+P(C\mid \bar{A}B)P(\bar{A}B)$$

$$+P(C\mid A\bar{B})P(A\bar{B})+P(C\mid \bar{A}\bar{B})P(\bar{A}\bar{B})$$

$$=0+\frac{3}{12}P(B\mid \bar{A})P(\bar{A})+\frac{3}{12}P(\bar{B}\mid A)P(A)+\frac{6}{12}P(\bar{B}\mid \bar{A})P(\bar{A})$$

$$=\cdots=41/110.$$

一. 选择题

练

习

1.设随机变量$X\sim N(\mu,\sigma^2)$,则当σ增加时

$P(|X-\mu|\geq\sigma)$____C____

A 单调增大　　B 单调减小

C 保持不变　　D 增减不定

一. 选择题

练

习

2. 设X为连续型随机变量，$f(x), F(x)$分别为其密度函数和分布函数，对固定的a $(F(a)<1)$，

$$g(x)=\begin{cases} f(x)/1-F(a), & x\geq a; \\ 0, & x<a. \end{cases}$$

则$g(x)$为（ **A** ）。

A 密度函数 　　B 分布函数

C 连续函数 　　D 单调增函数

二. 填空题

练

习

1. 一整数等可能地在1到10中取值，以X记除得尽这一整数的正整数的个数，则EX =______， DX=______。

2. 已知随机变量X概率分布列为$P(X=1)=0.2$，$P(X=2)=0.3$， $P(X=3)=0.5$，则其分布函数$F(x)$=__________。

三. 解答题

练

习

1. 从学校乘汽车到火车站的途中有3个交通岗，假设在各个交通岗遇到红灯的事件是相互独立的，并且概率都是1/3。设X为途中遇到红灯的次数, 求随机变量X的分布列和分布函数.

(提示：二项分布)

三. 解答题

练

习

2. 设随机变量X在（0，2π）内服从均匀分布，求随机变量$Y=\cos X$的分布密度。

(提示：随机变量的函数的计算)

三. 解答题

练

习

3.（综合） 某交通路口设有电子监控系统。设该系统一天内监测到驾驶员违章的次数服从泊松分布，且各天违章次数相互独立。以往的资料表明一天内平均有6次违章记录。试求：

（1）该路口一天内至少有两次违章的概率；

（2）该路口一年（365天）内违章记录少于2190次的概率。

(提示：二项分布，大数定理)

三. 解答题

练

习

4. 某仪器有三个独立工作的同种型号电子元件，其寿命(单位：小时)都服从E(1/60).

(1)求在仪器使用的最初的200小时内至少有一只电子元件损坏的概率。

(2)若这三个电子元件都无故障时仪器正常，否则不正常,求仪器正常工作时间T的分布函数。

(提示：C.R.V.相关概率,二项分布及分布函数的计算)

三. 解答题

练

习

5. 设随机变量X的概率密度函数为：

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

以Y表示X的三次独立重复观察中事件$\{X\le 1/2\}$出现的次数，求Y的概率分布列。

(提示：C.R.V.与二项分布的结合)

第二章结束!

2020/11/19

THE END

例 某人射击，设每次射击的命中率为0.001，独立射击5000次，问

（1）最可能命中多少次？并求相应概率。

（2）至少命中两次的概率是多少？

解 设X为命中的次数，则$X \sim B(5000, 0.001)$.

（1）$\because (n+1)p=(5000+1)\times 0.001=5.001$

故最可能命中5次。相应的概率为

$$P(X=5)=C_{5000}^{5}0.001^{5}0.999^{4995}\approx 0.1756.$$

（2）所求概率为$P(X\geq 2)=1-P(X=0)-P(X=1)$

$$=1-\sum_{k=0}^{1}C_{5000}^{k}(0.001)^{k}(0.999)^{5000-k}=0.9574.$$

n 次试验　1. A与$\overline{A}$
2. $P(A)=p$
3. 独立

设X为A发生的次数，则$X\sim B(n,P(A))$

例 设有一决策系统，其中每个成员作出决策互不影响，且每个成员作出正确决策的概率均为$p(0<p<1)$。当占半数以上的成员作出正确决策时，系统作出正确决策。问p多大时，5个成员的决策系统比3个成员的决策系统更为可靠?

解 记X_n为n个成员的决策系统中, 作出正确决策的成员数.

$X_5 \sim B(5,p)$　$P(X_5 \geq 3) = C_5^3 p^3(1-p)^2 + C_5^4 p^4(1-p) + p^5$

$X_3 \sim B(3,p)$　$P(X_3 \geq 2) = C_3^2 p^2(1-p) + p^3$

$$C_5^3 p^3(1-p)^2 + C_5^4 p^4(1-p) + p^5 > C_3^2 p^2(1-p) + p^3$$

$$\Rightarrow \quad p > 1/2$$