随机变量及其分布01

1、随机变量的概念：(随机事件的结果展现出来的数学形式)

定义在样本空间Ω上的实值函数$X = X(\omega)$称为随机变量，常用大写字母X,Y,Z等表示随机变量，其取值用小写字母x,y,z等表示。

假如一个随机变量仅可能取有限个或可列个值，则称其为**离散型随机变量**。假如一个随机变量的可能取值充满数轴上的一个区间(a,b)，则称其为**连续型随机变量**，其中a可以是$-\infty$，b也可以是$+\infty$

2、随机变量的分布函数：设X为一个随机变量，对任意实数x，$F(x) = P(X \le x)$为X的分布函数，且称X服从$F(x)$, 记为$X \sim F(x)$或$F_X(x)$

分布函数的三个性质：单调性，有界性，右连续，这三个基本性质成为判别某个函数是否能成为分布函数的充要条件。

对任意实数$a<b$，有下列公式：

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

推导：$P(a < x \le b) = P\big((x \le b) - (x \le a)\big) = P(x \le b) - P(x \le a) = F(b) - F(a)$

(2) $P(X = a) = F(a) - F(a-0)$

推导：$P(X = a) = P(X \le a) - P(X < a) = F(a) - F(a-0)$

(3) $P(X \ge b) = 1 - F(b-0)$

推导：$P(X \ge b) = 1 - P(X < b) = 1 - F(b-0)$

(4) $P(X > b) = 1 - F(b)$

推导：$P(X>b)=1-P(X\leq b)=1-F(b)$

(5) $P(a<x<b)=F(b-0)-F(a)$

推导：$P(a<x<b)=P(x<b)-p(x\leq a)=F(b-0)-F(a)$

(6) $P(a\leq x\leq b)=F(b)-F(a-0)$

推导：$P(a\leq x\leq b)=P(x\leq b)-P(x<a)=F(b)-F(a-0)$

(7) $P(a\leq x<b)=F(b-0)-F(a-0)$ 特别的，如果函数在a,b处连续时，有$F(a-0)=F(a),F(b-0)=F(b)$

3、离散型随机变量的概率分布列

设离散型随机变量X的可能取值为$x_1,x_2,\ldots,x_n,\ldots$加入$X$取$x_i$的概率为$\boldsymbol{p_i=P(X=x_i),i=1,2,\ldots}$**为$X$的概率分布列或简称为分布列**，记为$X\sim\{P_i\}$，分布列为：

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

4、分布列的性质：（1）非负性$p(x_i)\geq 0$（2）正则性$\sum_{i=1}^{n}p(x_i)=1$

由离散随机变量X的分布列可以得出X的分布函数为$F(x)=\sum_{x_i\leq x}^{\infty}p(x_i)$

比如现在有1，2，3，4，5这些点，分别对应概率p_1,p_2,p_3,p_4,p_5，则$F(3)=p_1+p_2+p_3$

5、求离散型随机变量的分布列应该注意：

（1）确定随机变量的所有可能取值

（2）计算每个取值点的概率

求离散型随机变量的分布列函数应该注意：

（1）$F(x)$是递增的阶梯函数

（2）其间断点均为右连续的

（3）其间断点即为X的可能取值点

（4）其间断点的跳跃高度是对应的概率值

例如：$F(x)=\begin{cases}0, & x<0\\ 0.4, & 0\le x<1\\ 0.8, & 1\le x<2\\ 1, & 2\le x\end{cases}$

X	0	1	2
P	0. 4	0. 4	0. 2

6、连续型随机变量的概率分布列

定义：设随机变量X的分布函数为$F(x)$，如果存在实数轴上的一个非负可积函数$p(x)$，使得对任意实数x有：$F(x)=\int_{-\infty}^{x} p(t)\,dt$，则称$p(x)$为$X$的概率密度函数，简称密度函数或称密度。

密度函数的基本性质：（1）非负性$p(x)\ge 0$　（2）正则性$\int_{-\infty}^{+\infty} p(x)\,dx=1$（含有$p(x)$的可积性）

以上两条基本性质是确定或判别某个函数是否成为密度函数的充要条件。

(1)　$P(a\le X\le b)=\int_{a}^{b} p(x)\,dx$

(2)　$F(x)$是$(-\infty,+\infty)$上的连续函数

(3)　$P(X=x)=F(x)-F(x-0)=0$

(4)　$F(x+\triangle x)-F(x)=\int_{x}^{x+\triangle x} p(x)\,dx\to 0\ \ (\triangle x\to 0)$

(5)　$P(X=a)=0$，连续型随机变量在某一点的概率为 0

(6)　$P(a\le X\le b)=P(a<X\le b)=P(a\le X<b)=P(a<X<b)$

(7)　密度函数$X\sim p(x)$不唯一

7、数学期望的定义

(离散型) 设离散型随机变量X的分布列为：$p(x_i)=P(X=x_i),i=1,2,\dots,n,\dots$ 如果$\sum_{i=1}^{\infty}|x_i|\,p(x_i)<\infty$，则称$E(X)=\sum_{i=1}^{\infty}x_i\,p(x_i)$为随机变量$X$的数学期望，或称为该分布的数学期望，简称期望或均值，若$\sum_{i=1}^{\infty}|x_i|\,p(x_i)$不收敛，则$X$的数学期望不存在

(连续性) 设连续型随机变量X的密度函数为$p(x)$，如果$\int_{-\infty}^{+\infty}|x|p(x)\,dx<\infty$，则称$E(X)=\int_{-\infty}^{+\infty}xp(x)\,dx$为随机变量 X 的数学期望，或称为该分布$p(x)$的数学期望，简称期望或均值，若$\int_{-\infty}^{+\infty}|x|p(x)\,dx$不收敛，则$X$的数学期望不存在

8、数学期望的性质

设$Y=g(x)$是随机变量X的函数，若$E(g(X))$存在，则

$$E\big(g(x)\big)=\begin{cases}\sum_{i=1}^{\infty}g(x_i)P(X=x_i),\text{在离散场合}\\ \int_{-\infty}^{+\infty}g(x)p(x)dx,\text{在连续场合}\end{cases}$$

性质 1：若C是常数，则$E(c)=c$

性质 2：对任意数a，有$E(aX)=aE(X)$

性质 3：对任意两个函数$g_1(x)$和$g_2(x)$，有$E[g_1(x)\pm g_2(x)]=E(g_1(x))\pm E(g_2(x))$

9、方差与标准差的定义（方差反应的稳定程度）

定义：若随机变量X^2的数学期望$E(X^2)$存在，则称偏差平方$(X-EX)^2$的数学期望为随机变量X(或相应分布)的方差，记为

$var\ (X)=E(X-E(X))^2$ 称方差的正平方根$\sqrt{Var(X)}$为随机变量X(或相应分布)的标准差，记为$\sigma(X)$或σ_X

10、方差的性质

性质 1：$var\ (X)=E(X-E(X))^2=E(X^2-2*X*E(X)+E(X)^2)=E(X^2)-2E(X)*E(X)+E(X)^2=E(X^2)-[E(X)]^2$

求方差常用此性质，先求出$E(X)$, 再求$E(X^2)$

性质 2：常数 c 的方差为 0

性质 3：对任意常数a与b和随机变量X，有$Var(aX+b)=a^2Var(X)$

随机变量的标准化：设$Var(X)>0$, 令$Y=\frac{X-E(X)}{\sqrt{Var(X)}}$，则$E(Y)=0, Var(Y)=1$, 称$Y$为$X$的标准化

11、切比雪夫不等式

定理 1：对任一随机变量X，若X的数学期望和方差都存在，则对任一正数ε，恒有$P(|X-E(X)|\geq\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}$或$P(|X-E(X)|<\varepsilon\geq 1-\frac{Var(X)}{\varepsilon^2})$

定理 2：方差为 0 的随机变量X必几乎处处为常数。这个常数就是其期望$E(X)$，这个定理亦可表示为：

若$Var(X)=0$，则$P\big(X=E(X)\big)=1$或$P(X=a)=1$

定理 2 的意思就是如果方差为 0，那么随机变量X为常数列

12、常用的离散分布

（一）二项分布

(1) 伯努利试验：如果试验只有两个结果成功(A)和失败$(\bar{A})$，且每次试验成功的概率是不变的，则称这种试验为伯努利实验，将一个伯

努利试验独立重复的进行n次，称为n重伯努利试验。

在每次试验中，$P(A)=P(0<p<1)$，则在n重伯努利试验中事件A出现k次的概率为

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

若随机变量X的概率分布为$P(X=k)=C_n^k p^k(1-p)^{n-k}(k=0,1,2,\dots,n)$，则称随机变量$X$服从参数为$(n,p)$的二项分布，并记

$\boldsymbol{X \sim B(n,p)}$

二项分布的数学期望$E(X)=np$，$Var(X)=np(1-p)$

(2) 两点分布（二项分布的特殊形式）

当 n=1 时，二项分布为 B(1, p) 称为两点分布，或称 0-1 分布，或称伯努利分布，其分布列为 P(X=x)=$p^x(1-p)^{1-x}, x=0$ 或 1

二点分布的数学期望$E(X)=p$，$Var(X)=p(1-p)$

(二) 泊松分布

若随机变量X的概率分布为$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$ $(k=0,1,2,\dots)$，其中参数$\lambda>0$，则称随机变量X服从参数为λ的泊松分布，并记$X\sim P(\lambda)$，泊松分布常与单位时间(或单位面积、单位产品等)的计算过程相联系。

泊松分布的数学期望和方差均为λ

(三) 二项分布的泊松近似

在二项分布$B(n,p)$中，当 **n 较大**时，计算量是令人烦恼的，而 **P 较小**时使用泊松定理，可以减少二项分布中的计算量。

定理(泊松定理) 在n重伯努利实验中，记事件A在一次试验中发生的概率为P_n，（与试验次数n有关），如果当$n\to\infty$时，有$np_n\to\lambda$，则

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

(四)超几何分布

从一个有限总体中进行不放回抽样常会遇到超几何分布

设有N件产品，其中有M件不合格品，若从中不放回的随机抽取n件，则其中含有的不合格品的件数X服从超几何分布，记为$X \sim h(n, N, M)$超几何分布的概率分布列为：

$$P(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \quad ,k=0,1,\dots,r$$

其中$r=\min\{M,n\}$，且 $\mathrm{M} \le \mathrm{N}, \mathrm{n} \le \mathrm{N}, \mathrm{n}, \mathrm{N}, \mathrm{M}$ 均为正整数

超几何分布的数学期望　$E(X) = n\frac{M}{N}$

方差$Var(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)}$

(五)超几何分布的二项近似

当$n \ll N$时，即抽取个数n远小于产品个数N时，每次抽取后，总体中的不合格品率$p=M/N$改变甚微，所以不放回抽样可近似地看成放回抽样，这时超几何分布可用二项分布近似。

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

(六)几何分布

在伯努利试验序列中，记每次试验中事件A发生的概率为p，如果X为事件A首次出现时的试验次数，则X的可能取值为$1,2,\dots$，称X服从几何分布，记为$X \sim Ge(p)$，其分布列为$p(X=k) = (1-p)^{k-1}p, k = 1,2,\dots$　几何分布的数学期望和方差$E(X) = \frac{1}{p}$ ，$Var(X) = \frac{1-p}{p^2}$

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

$p(X>m+n|X>m)=p(X>n)$，这个定理表明：在前m次实验中A没有出现的条件下，则在接下来的n次实验中A仍未出现的概率只与n有关，而与以前的m次试验无关，似乎忘记了之前m次的实验结果，这就是无记忆性。

(七) 负二项分布

在伯努利实验序列中，记每次试验中事件A发生的概率为P，如果X为事件A第r次出现时的试验次数，则X的可能取值为$r,r+1,\dots,r+m,\dots$称X服从负二项分布或帕斯卡分布，其分布列为

$$P(X=r)=C_{k-1}^{r-1}p^r(1-p)^{k-r},k=r,r+1,\dots$$

记为$X\sim Nb(r,p)$，当$r=1$时，记为几何分布

负二项分布的数学期望和方差为$E(X)=\frac{r}{p};Var(X)=\frac{r(1-p)}{p^2}$

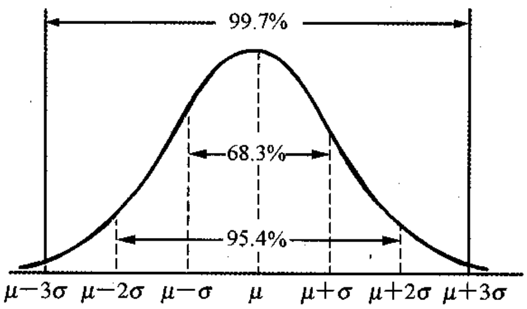
13、常用的连续分布

（一）正态分布

一个随机变量如果是由大量微小的、独立的随机因素的叠加结果，那么这个变量一般都可以认为服从正态分布。

若随机变量X的密度函数为

$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty$，则称$X$服从正态分布，称$X$为正态分布，记作$X\sim N(\mu,\sigma^2)$，其中参数$-\infty<\mu<+\infty$，$\sigma>0$



正态分布的3σ原则
尽管正态分布的取值范围是$(-\infty,+\infty)$，但他的99.73%的值在$(\mu-3\sigma,\mu+3\sigma)$内，这个性质被称为正态分布的$3\sigma$原则

正态分布$N(\mu,\sigma^2)$的分布函数为$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{\frac{(t-\mu)^2}{2\sigma^2}}dt$

正态分布的数学期望：$E(X)=\mu$　方差：$Var(X)=\sigma^2$

正态分布的性质：

（1）$p(x)$是一条钟形曲线，中间高，两边第、左右关于μ对称，μ是正态分布的中心，且在$x=\mu$附近取值的可能性大，在两侧取值的可能性小，$x\pm\sigma$是该曲线的拐点。

（2）μ为位置参数，σ为尺度参数

（3）若σ固定，μ改变，$p(x)$左右移动，形状保持不变

（4）若μ固定，σ改变，σ越大曲线越平坦(越分散)，σ越小曲线越陡峭(越集中)

标准正态分布

称$\mu=0$，$\sigma=1$是的正态分布$N(0,1)$为标准正态分布，记为U，因此密度函数与分布函数写成：$\varphi(u)=\frac{1}{\sqrt{2\pi}}e^{\frac{x^2}{2}},-\infty<u<+\infty$

$$\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{\frac{u^2}{2}}du,-\infty<u<+\infty$$

标准正态分布的性质：

（1）$\Phi(0)=\frac{1}{2}$

（2）$\Phi(-u)=1-\Phi(u)$

（3）$\mathrm{P}(U>u)=1-\Phi(u)$

（4）$\mathrm{P}(a<U<b)=\Phi(b)-\Phi(a)$

（5）$\mathrm{P}(|U|<c)=2\Phi(c)-1(c\geq 0)$

标准正态分布的数学期望：$E(X)=0$　方差：$Var(X)=1$

正态分布的标准化

对于一般正态分布都可以通过一个线性变换(标准化)化成标准正态分布。因此与正态分布变量有关的一切事件的概率都可以通过查正态分布函数表获得

定理 1：若随机变量$X \sim N(\mu,\sigma^2)$，则$U=(X-\mu)/\sigma \sim N(0,1)$，若$X \sim N(\mu,\sigma^2)$，则$P(X \le c)=\Phi(\frac{c-\mu}{\sigma})$，$P(a \le X \le b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi(\frac{a-\mu}{\sigma})$

(二) 均匀分布

均匀分布的密度函数和分布函数

随机变量X的密度函数为$p(x)=\begin{cases}\frac{1}{b-a}, a<x<b \\ 0, 其他\end{cases}$

则称X服从区间(a,b)上均匀分布，记作$X \sim U(a,b)$

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

均匀分布的数学期望与方差：$E(X)=\frac{a+b}{2}$ $Var(X)=\frac{(b-a)^2}{12}$

(三) 指数分布

指数分布的密度函数与分布函数

若随机变量X的密度函数为$p(x)=\begin{cases}\lambda e^{-\lambda x}, x \ge 0 \\ 0, x<0\end{cases}$则称$X$服从指数分布，记作$X \sim Exp(\lambda)$或$X \sim E(\lambda)$，其中$\lambda>0$，指数分布的分布函数为

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

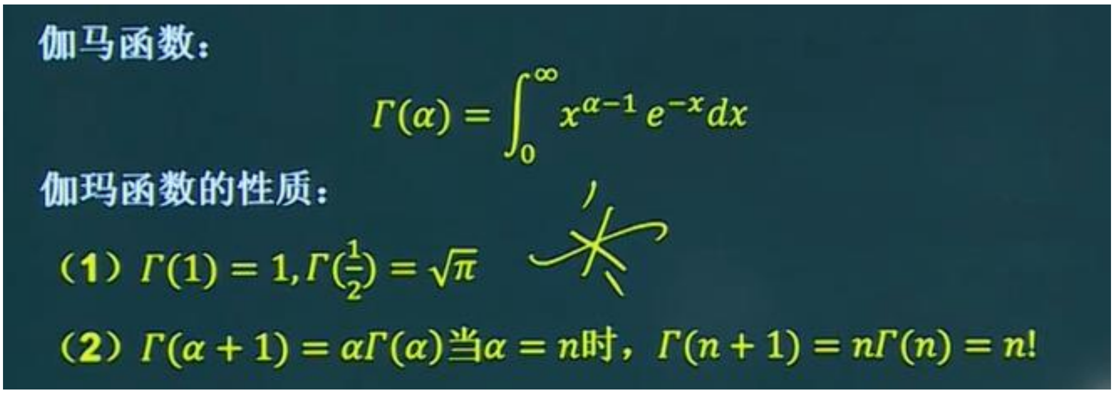
指数分布是一种偏态分布，由于指数分布随机变量只可能取非负实数，所以指数分布常被用作各种“寿命”分布。

指数分布的数学期望与方差$E(X)=\frac{1}{\lambda}\quad Var(X)=\frac{1}{\lambda^2}$

定理 1：（指数分布的无记忆性）如果随机变量$X\sim Exp(\lambda)$，则对任意$s>0,t>0$，有$P(X>s+t|X>s)=P(X>t)$

上式含义为：记X是种产品的使用寿命(h)，若X服从指数分布，那么已知此产品使用了$s(h)$没发生故障，则在能使用$t(h)$而不发生故障的概率与已使用的$s(h)$无关。只相当于重新开始使用$t(h)$的概率，即对已使用过的$s(h)$没有记忆

(四)伽马分布



若随机变量X的密度函数为$p(x)=\begin{cases}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},x\geq 0\\ 0,x<0\end{cases}$ 则称X服从伽马分布，记为$X\sim Ga(\alpha,\lambda)$，其中α为形状参数，λ为尺度参数

（1）$0<\alpha<1$时，$p(x)$是严格下降函数，且在$x=0$处有奇异点

（2）$\alpha=1$时，$p(x)$是严格下降函数，且在$x=0$处$p(0)=\lambda$

（3）$1<\alpha\leq 2$时，$p(x)$是单峰函数，先上凸，后下凹

（4）$2<\alpha$时，$p(x)$是单峰函数，先下凸、中间上凸、后下凸。且α越大，$p(x)$越接近正态密度，但伽马分布总是偏态分布，α越小其偏斜程度越严重。

$X\sim Ga(\alpha,\lambda)$的数学期望和方差为 $E(X)=\frac{\alpha}{\lambda}$　$Var(X)=\frac{\alpha}{\lambda^2}$

注：（1）$\alpha = 1$时的伽马分布就是指数分布，即$Ga(1,\lambda) = Exp(\lambda)$

(2) 称$\alpha = \frac{n}{2}$, $\lambda = \frac{1}{2}$时的伽马分布是自由度为n的卡方分布，记为$\chi^2(n)$，即$Ga\left(\frac{n}{2},\frac{1}{2}\right) = \chi^2(n)$，这里$n$是卡方分布的唯一参数，称为自由度。卡方分布的数学期望和方差是$E(X) = n$，$Var(X) = 2n$

(五) 贝塔分布

贝塔函数：$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$ ，其中参数$a > 0, b > 0$，贝塔函数具有如下性质：

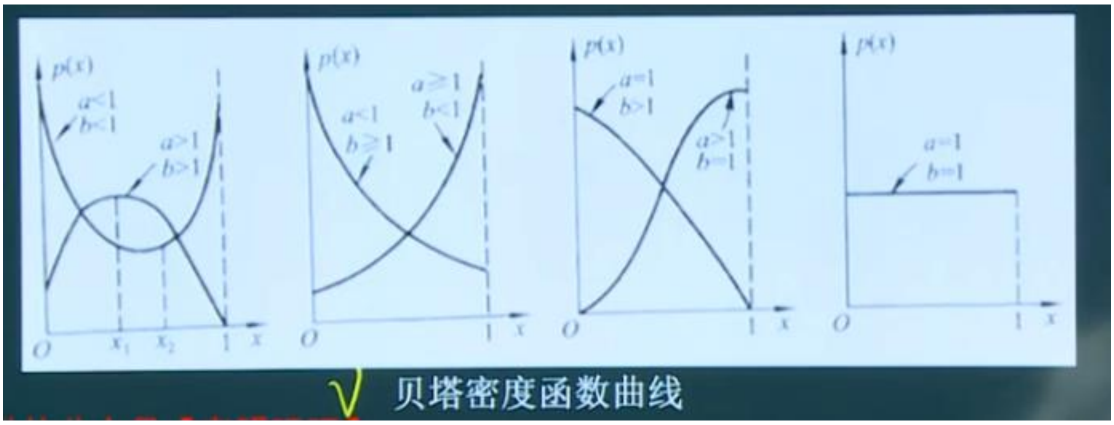
（1）$B(a,b) = B(b,a)$

（2）贝塔函数与伽马函数间有关系$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

贝塔分布：若随机变量X的密度函数为

$$p(x) = \begin{cases} \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} x^{a-1}(1-x)^{b-1}, 0 < x < 1 \\ 0, 其他 \end{cases}$$

则称X服从贝塔分布，记作$X \sim Be(a,b)$，其中参数$z > 0$，$b > 0$都是形状参数



贝塔密度函数曲线

（1）当$a < 1, b < 1$时，p(x)是下凸的U形函数

（2）当$a > 1, b > 1$时，p(x)是上凸的单峰函数

（3）当$a > 1$时，p(x)是上凸的单调增函数

（4）当 b>1 时 p(x) 是下凸的单调减函数

（5）当$a = 1, b = 1$时，$p(x)$是常数函数，且$Be(1,1) = U(0,1)$

贝塔的数学期望与方差$E(X) = \frac{a}{a+b}$　$Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$