第二章 随机变量及其分布

刘 春 光

暨南大学数学系

2018年2月

目录

目录

随机变量

在实际问题中，随机试验的结果虽有各种不同的表现形式，但是都可以用数量来表示，由此就产生了随机变量的概念。

随机变量

有些试验的结果本身就是数值，如

- 掷一颗骰子所得的点数；
- 抽查样品时的废品个数；
- 广州每日的平均气温；
- 某电子管的使用寿命；
- ···························

随机变量

有些试验的结果看起来与数值无关，但我们可以引进一个映射使每个结果与一个数值对应，把试验结果数值化。

随机变量

Definition (随机变量(random variable))

设$(\Omega, \mathscr{F}, P)$为概率空间，$X = X(\omega)$为Ω上定义的实值函数。如果对于每个$x \in \mathbb{R}$，都有

$$\{\omega |\ X(\omega) < x\} \in \mathscr{F},$$

则称$X(\omega)$为随机变量。

随机变量一般用大写的英文字母或小写的希腊字母来表示。

随机变量

随机变量的分类

1. 离散随机变量(discrete r.v.)
2. 连续随机变量(continuous r.v.)
3. 奇异随机变量(singular r.v.)
4. 混合型(mixed type)

目录

离散随机变量

Definition (离散随机变量(discrete random variable))

只能取有限个或者可列个值的随机变量称为离散随机变量。

例：骰子的点数、抽取的次品数、收到的呼叫次数等

离散随机变量

Definition (离散随机变量的概率分布)

若离散随机变量X的所有可能值为$\{x_k\}$，分别对应概率$\{p_k\}$，则称

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots$$

为X的概率分布或者分布律。

离散随机变量

概率分布也常用下面表格的形式给出：

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

$\{p_k\}$的性质：$\begin{cases} p_k \geq 0, \quad k = 1, 2, \cdots \\ \sum_k p_k = 1 \end{cases}$

离散随机变量

练习：已知随机变量X有如下分布律

X	1	2	3	4	5
P	c	$4c$	$6c$	$4c$	c

求常数c并计算

$$P\{X=2|X<5\},\ P\{X^2+4=5X\}.$$

离散随机变量

例1 袋中装有两个白球与三个黑球，每次从袋中任取一个球，直至取得白球为止，求以下两种取法下取球次数的概率分布：

1. 每次取出的球不再放回；
2. 每次取球观察颜色后放回袋中。

离散随机变量

Definition (几何分布(geometric distribution))

如下概率分布

$$P\{X=k\}=pq^{k-1}, \quad k=1,2,\cdots$$

（其中$p,q>0,p+q=1$）称为几何分布。

注：几何分布描述反复进行的Bernoulli试验中第一次成功所需试验次数的概率。

目录
随机变量
离散型
常用离散分布
连续型
分布函数
概率密度
常用连续分布
函数的分布
联合分布
边缘分布
条件分布
随机变量的独立性
二维变量函数的分布

目录

0-1分布

Definition (两点分布)

只有两个可能取值的分布称为两点分布，其分布律可表示为：

$$P\{X=x_1\}=p, \quad P\{X=x_2\}=1-p,$$

其中$0<p<1$。

两点分布的对应分布表为

X	x_1	x_2
P	p	$1-p$

0-1分布

Definition (0-1分布)

若随机变量X只能取0或1，其概率分布为：

$$P\{X=1\}=p,\ P\{X=0\}=1-p \quad (0<p<1)$$

则称X服从参数为p的0-1分布，记为

$$X \sim B(1,p).$$

0-1分布的分布表为

X	0	1
P	$1-p$	p

二项分布

Definition (二项分布(binomial distribution))

如果随机变量X服从以下分布律

$$P\{X=k\}=b(k;n,p)=C_n^k\,p^k(1-p)^{n-k},$$

其中$p\in(0,1),k\in\overline{0,n}$，则称$X$服从参数为$n,p$的二项分布。简记为

$$X\sim B(n,p).$$

注：二项分布描述n重Bernoulli试验中成功次数的概率。

二项分布

二项分布与0-1分布的关系：

1. 0-1分布$B(1,p)$是二项分布$B(n,p)$在$n=1$时的特殊情况。
2. 设在某试验中事件A的概率为p，将该试验独立地进行n次。记X为n次试验中事件A发生的总次数，X_i为第i次试验中事件A发生的次数，则有

$$X = X_1 + X_2 + \cdots + X_n.$$

超几何分布

例1 某班有学生20名，其中有5名女生，今从班上任选4名学生去参观展览，被选到的女生数X是一个随机变量，求X的分布。

例2 某班有学生20名，其中有3名女生，今从班上任选4名学生去参观展览，被选到的女生数X是一个随机变量，求X的分布。

超几何分布

Definition (超几何分布(Hypergeometric distribution))

设N个元素分为两类，第一类有M个元素，从中不重复抽取n个，以X表示抽到的第一类元素的个数，则X的分布律为

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

称X服从超几何分布，记为$X\sim H(n,M,N)$。

超几何分布

Theorem

固定正整数n，若超几何分布的参数列$\{(M,N)\}$满足

$$\lim_{N\to\infty}\frac{M}{N}=p\in(0,1),$$

则对任意的$k=0,1,\cdots,n$，有

$$\lim_{N\to\infty}\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k}.$$

超几何分布

前述定理的意义为：对超几何分布$H(n,M,N)$，当N,M非常大，而n相对很小的时候，可将其近似视为二项分布$B(n,p)$，其中$p=\dfrac{M}{N}$。

泊松(Poisson)分布

Definition (泊松分布(Poisson distribution))

如果随机变量X服从以下分布律

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,\cdots$$

其中$\lambda>0$，则称X服从参数为λ的泊松分布，简记为

$$X\sim P(\lambda).$$

泊松(Poisson)分布

Example

以下随机变量常以泊松分布作为其概率模型：

1. 服务系统中对服务的呼唤数；
2. 产品的缺陷数；
3. 一定时期内出现的稀有事件（事故、灾害等）个数；
4. 放射性物质发射出的粒子数。

泊松(Poisson)分布

历史上，泊松分布是作为二项分布的近似，于1837年由法国数学家泊松引入的。

Theorem

对一列二项分布$\{B(n,p_n)\}$，若其参数列满足$\lim\limits_{n\to\infty} np_n = \lambda > 0$，则对任意的非负整数$k$有

$$\lim_{n\to\infty} b(k;n,p_n) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

泊松(Poisson)分布

前述定理的意义为：对二项分布$B(n,p)$，当n充分大、p很小时，对任意固定的非负整数k，有近似式

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

其中$\lambda = np$。

习惯设定：$n \geq 100, p \leq 0.01$，且$np \leq 20$。

泊松(Poisson)分布

例 设一批产品共2000个，其中有40个次品。随机抽取100个样品，求样品中次品数的概率分布，如果抽样方式是

1. 不放回抽样；
2. 放回抽样。

目录

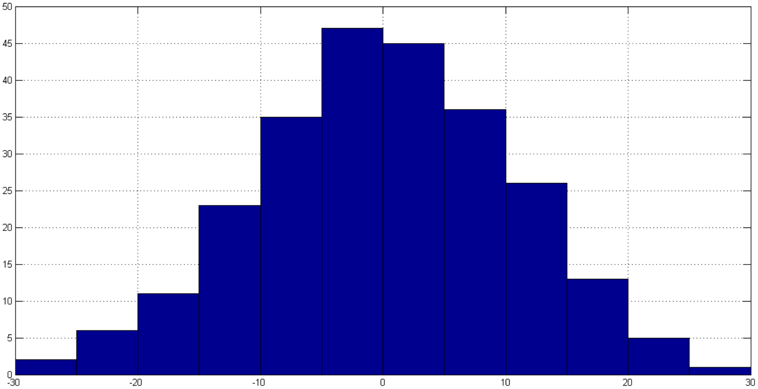
连续随机变量

例 某机床加工零件的长度与标准尺寸的误差范围为$[-30, 30]$（单位：微米）。抽取250个零件，以$5\mu m$为单位分组：

$$[-30, -25], (-25, -20], \cdots, (25, 30].$$

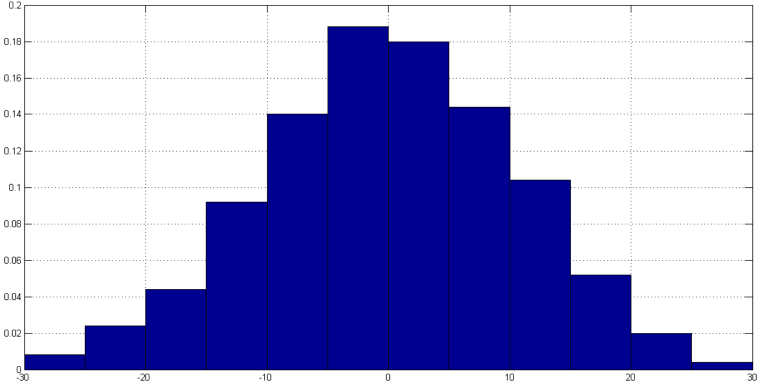
统计各组的样本频数，并做直方图说明样本的分布。

连续随机变量



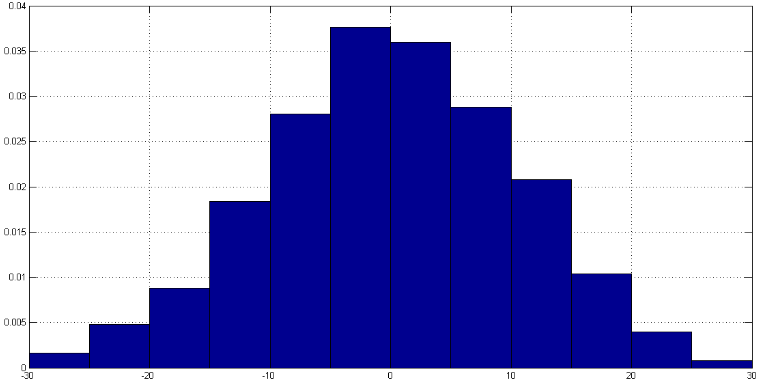
250个样本，频数直方图

连续随机变量



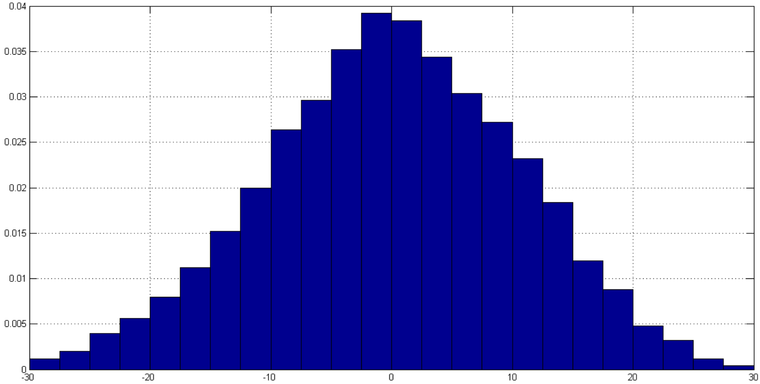
250个样本，频率直方图

连续随机变量



250个样本，密度直方图

连续随机变量



1000个样本，密度直方图

目录

分布函数

Definition (分布函数(cumulative distribution function (c.d.f.)))

对随机变量X，称函数

$$F(x) := P\{X \leq x\}, \quad x \in \mathbb{R}$$

为X的分布函数。

由上述定义得：对任意实数$a < b$，都有

$$P\{a < X \leq b\} = F(b) - F(a).$$

分布函数

Theorem (分布函数的性质)

设$F(x)$为某随机变量的分布函数，则其有以下性质：

1. 广义单增性：$a < b \Rightarrow F(a) \leq F(b)$;
2. 右连续性：$F(x) = \lim\limits_{t \to x^+} F(t)$;
3. 规范性：值域为$[0, 1]$，且

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

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

分布函数

广义单增性、右连续性与规范性是分布函数的本征性质，即$\mathbb{R}$上的函数$F(x)$是某个随机变量的分布函数当且仅当$F(x)$满足以上三条性质。

离散随机变量的分布函数

设离散型随机变量X的分布律为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots.$$

则

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

这里的和式是对所有满足$x_k \le x$的k求和。

离散随机变量的分布函数

例 袋中装有两个白球与三个黑球，不放回地依次从袋中取球，直至取得白球为止，用X表示取球次数，则其分布律为

X	1	2	3	4
P	0.4	0.3	0.2	0.1

则X的分布函数为

离散随机变量的分布函数

$$F(x)=\begin{cases}0, & x<1,\\ 0.4, & 1\le x<2,\\ 0.7, & 2\le x<3,\\ 0.9, & 3\le x<4,\\ 1, & x\ge 4.\end{cases}$$

容易看出，离散随机变量的分布函数是一个纯跳跃函数：它在X的每个可能值x_k上有跃度p_k，在不含任何可能值的区间上恒取常值。

连续随机变量的分布函数

例1 向半径为R的圆形靶射击，击中点M落在以靶心O为中心、r为半径的圆内的概率与该圆的面积成正比，并且不会出现脱靶的情况。用X表示击中点M与靶心O的距离，

1. 求X的分布函数；
2. 如果击中点M落在以靶心O为中心、内外半径分别为$\frac{i}{10}R$及$\frac{i+1}{10}R$ $(i=0,1,\cdots,9)$的圆环域内，则计为$10-i$环。求一次射击得到$10-i$环的概率。

连续随机变量的分布函数

例2 使用了t小时的电子元件在以后的Δt小时内损坏的概率等于$\lambda\Delta t + o(\Delta t)$，其中$\lambda > 0$为常数。求电子元件寿命的分布函数。

目录

连续随机变量的概率密度

Definition (continuous random variable, probability density function (p.d.f.))

如果存在函数$f(x)$，满足

$$(1)\ f(x) \geq 0, \quad (2)\ \int_{-\infty}^{+\infty} f(x)dx = 1,$$

且随机变量X的分布函数$F(x)$可以写成

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

则称X为连续随机变量，称$f(x)$为X的概率密度函数。

连续随机变量的概率密度

Definition (续)

此时，我们称X服从以$f(x)$为密度的概率分布，简记为

$$X \sim f(x).$$

连续随机变量的概率密度

练习：函数$f(x) = -\sin x, x \in I$可以做某随机变量的密度函数，若区间I为

(A) $\left[-\frac{\pi}{2}, 0\right]$　　(B) $\left[0, \frac{\pi}{2}\right]$

(C) $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$　　(D) $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$

★补充：绝对连续性

Definition (绝对连续性(absolute continuity))

设I为$\mathbb{R}$上的区间，F为定义在I上的函数。如果对任意的$\varepsilon > 0$，都存在$\delta > 0$，使得对任意有限个两两不交的区间$\{(x_k, z_k)\}_{1\leq k\leq n}$，只要

$$\sum_{k=1}^{n}(z_k - x_k) < \delta,$$

就有

$$\sum_{k=1}^{n}|F(z_k) - F(x_k)| < \varepsilon,$$

则称f在I上绝对连续。

★补充：绝对连续性

由定义可知：

$$\text{绝对连续} \Longrightarrow \text{一致连续} \Longrightarrow \text{连续。}$$

★补充：绝对连续性

Theorem (绝对连续性的等价条件)

设F为定义在$[a,b]$上的函数。以下论断相互等价：

1. F在$[a,b]$上绝对连续；
2. F在$[a,b]$上几乎处处可导，$F'(x)$ *Lebesgue*可积，且
$$F(x)=F(a)+\int_a^x F'(t)dt,\quad \forall x\in[a,b].$$
3. 存在*Lebesgue*可积的函数f，使得
$$F(x)=F(a)+\int_a^x f(t)dt,\quad \forall x\in[a,b].$$

★补充：绝对连续性

由等价定理可知：连续随机变量的分布函数是绝对连续的。

连续随机变量的概率密度

细节辨析：

1. 连续型随机变量的分布函数是连续的，但密度函数不一定连续；
2. 如果密度函数$f(x)$在点x_0连续，则

$$F'(x_0) = f(x_0);$$

且对足够小的$\Delta x > 0$,

$$P\{x_0 - \Delta x < X \le x_0 + \Delta x\} \approx f(x_0) \cdot 2\Delta x,$$

即$f(x_0)$可近似视为x_0附近小区间上的平均概率。

连续随机变量的概率密度

任意区间上概率的计算：由密度函数的定义可知，

$$P\{X \in (a,b]\} = F(b) - F(a) = \int_a^b f(x)dx.$$

上式中的区间$(a,b]$改为(a,b), $[a,b)$或$[a,b]$后等式仍成立。

特别地，对每个实数a，有

$$P\{X = a\} = 0.$$

连续随机变量的概率密度

由定积分的几何意义，$P\{X \in [a,b]\}$等于以下曲线

$$x = a,\ x = b;\ y = 0,\ y = f(x)$$

所围成的曲边梯形的面积。

连续随机变量的概率密度

细节辨析：

- 对连续型随机变量而言，概率为零的事件与不可能事件是不同的概念：
  不可能事件的概率一定为零；但概率为零的事件也可能发生。
- 同样道理，必然事件概率一定为1，但概率为1的事件也不一定是必然事件。

连续随机变量的概率密度

Definition

若事件A的概率为1，则称该事件几乎必然发生(happens almost surely (a.s.))。

连续随机变量的概率密度

例1 已知连续型随机变量X的分布函数为

$$F(r)=\begin{cases}0, & r<0,\\ \frac{r^2}{R^2}, & r\in[0,R],\\ 1, & r>R,\end{cases}$$

求其概率密度函数。

连续随机变量的概率密度

例2 设连续随机变量X的概率密度为

$$f(x) = \frac{A}{1+x^2}, \quad x \in \mathbb{R}.$$

求

1. 常数A的值；
2. 随机变量X落在区间$[0,1]$内的概率；
3. 随机变量X的分布函数。

连续随机变量的概率密度

Definition (柯西分布(Cauchy distribution))

若随机变量X有概率密度

$$f(x) = \frac{1}{\pi(1+x^2)},$$

则称X服从柯西分布，其分布函数为

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x.$$

连续随机变量的概率密度

例3 已知连续型随机变量X的密度函数为

$$f(x)=\begin{cases} c+x, & x\in[-1,0), \\ c-x, & x\in[0,1], \\ 0 & \text{else} \end{cases},$$

求(1) 常数c；(2) 概率$P\{|X|\leq 0.5\}$；(3) X的分布函数$F(x)$。

目录

均匀分布

Definition (均匀分布(uniform distribution))

若随机变量X有概率密度

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

则称X服从区间$[a,b]$上的均匀分布，记为

$$X \sim U[a,b].$$

均匀分布

区间$[a,b]$上均匀分布的分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \frac{x-a}{b-a}, & x\in[a,b],\\ 1, & x>b.\end{cases}$$

均匀分布

例1 用电子表计时一般准确至0.01s，即如果以秒为时间的计量单位，则小数点后第二位数字是按“四舍五入”原则得到的，求使用电子表计时产生的随机误差X的概率密度，并计算误差的绝对值不超过0.002s的概率。

指数分布

Definition (指数分布(exponential distribution))

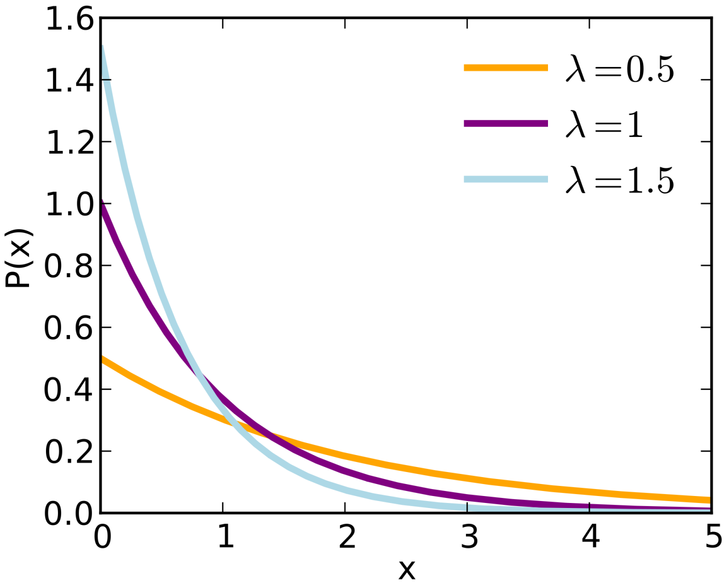
如果随机变量X有以下概率密度

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

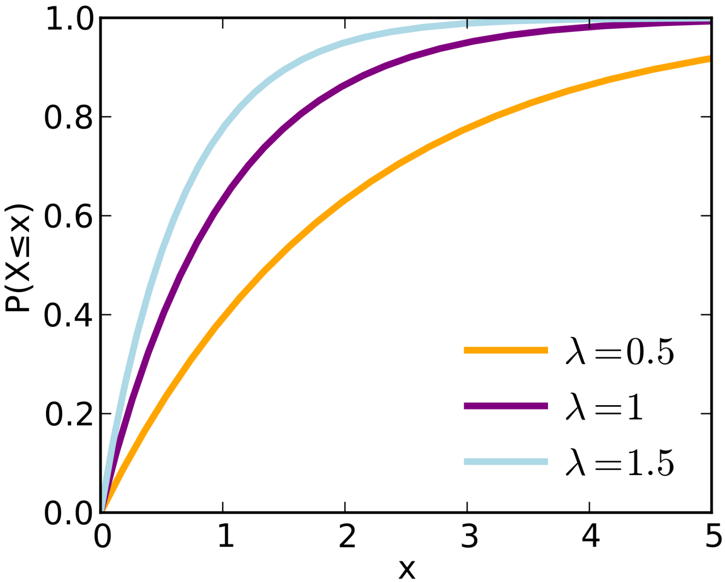
其中$\lambda > 0$，则称X服从参数为λ的指数分布。其分布函数为

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

指数分布的密度函数


1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
P(x)
0
1
2
3
4
5
x
$\lambda = 0.5$
$\lambda = 1$
$\lambda = 1.5$

指数分布的分布函数



指数分布

指数分布经常作为等待时间的分布，如产品的寿命，排队模型中的服务时间等。

例2 设某电子管的使用寿命X（单位：小时）服从参数为$\lambda = 0.001$的指数分布。求电子管使用寿命超过1000小时的概率。

指数分布

练习（指数分布的无记忆性）：随机变量ξ服从参数为λ的指数分布。设$s,t>0$，求以下条件概率

$$P\{\xi > s+t|\xi > s\}.$$

★补充：奇异型分布

Definition (奇异型分布(singular distribution))

若随机变量X的取值范围为零测度集，且X取每个可能值的概率均为零，则称X为奇异型随机变量，称其分布为奇异型分布。

奇异型分布的分布函数为连续函数，但不是绝对连续。如康托函数(Cantor function)。

★补充：分布函数的分解

Theorem (Lebesgue分解)

任意分布函数$F(x)$都可分解为如下形式：

$$F(x) = c_1F_1(x) + c_2F_2(x) + c_3F_3(x),$$

其中常数$c_1, c_2, c_3 \geq 0$, $c_1 + c_2 + c_3 = 1$,
$F_1(x), F_2(x), F_3(x)$都是分布函数，$F_1(x)$为纯跳跃函数，$F_2(x)$是绝对连续函数，$F_3(x)$为奇异函数。

目录

随机变量的函数

在实际问题中，有时我们关心的随机变量Y不容易直接测量，而是要测量另外一个随机变量X，把Y表示为X的函数$Y = g(X)$。

由此引出的问题是：已知X的分布，如何得到Y的分布？例如：已知圆球直径D的分布，求圆球体积$V = \dfrac{\pi D^3}{6}$的分布。

离散随机变量的函数的分布

例1 设随机变量X的分布为

X	-2	-1	0	1	2	3
P	0.1	0.2	0.25	0.2	0.15	0.1

求$Y_1 = -2X, Y_2 = X^2$的概率分布。

离散随机变量的函数的分布

总结：设离散型随机变量X的概率分布为

$$P\{X = x_k\} = p_k, \quad k = 1, 2, \cdots$$

令$Y = g(X)$，则Y也是一个离散型随机变量，其分布可按如下步骤求得

1. 根据函数关系列出Y的所有可能值；
2. 对Y的每个可能值y，$P\{Y = y\}$等于所有满足$g(x_k) = y$的p_k之和。

离散随机变量的函数的分布

练习：假设随机变量ξ的分布为

ξ	-1	0	1	2
P	0.2	0.25	0.3	0.25

求$\eta_1 = \xi^2 + 1, \eta_2 = \xi^2 - \xi$的概率分布。

离散随机变量的函数的分布

例2 设随机变量X的分布为

$$p(k)=\frac{1}{2^k},\quad k=1,2,\cdots.$$

求随机变量$Y=\sin(\frac{\pi}{2}X)$的概率分布。

连续随机变量的函数的分布

对连续型随机变量X，求$Y=g(X)$的密度函数的基本方法是

1. 根据函数关系先求Y的分布函数

$$
\begin{aligned}
F_Y(y) &= P\{Y \le y\} = P\{g(X) \le y\} \\
&= P\{X \in I_y\} = \int_{I_y} f(x)dx,
\end{aligned}
$$

其中$I_y = \{x|\ g(x) \le y\}$;

2. 然后对$F_Y(y)$求导可得Y的概率密度。

连续随机变量的函数的分布

例3 设随机变量X有概率密度函数$f_X(x)$，求$Y=a+bX$的概率密度函数，其中a及$b\neq 0$都是常数。

例5 设随机变量X服从参数为λ的指数分布，求$Y=e^X$的概率密度函数。

连续随机变量的函数的分布

练习：假设随机变量ξ的概率密度为

$$\varphi_\xi(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

证明：$\eta = 1 - e^{-2\xi}$服从区间$(0,1)$上的均匀分布。

连续随机变量的函数的分布

Theorem

设随机变量X有密度函数$f_X(x)$，$x \in (\alpha, \beta)$。如果$g(x)$在(α, β)上是严格单调的连续函数，存在唯一的反函数$x = h(y)$，$y \in (a, b)$，并且$h'(y)$存在且连续，那么$Y = g(X)$也是连续型随机变量，具有密度函数

$$f_Y(y) = f_X[h(y)] \cdot |h'(y)|, \quad y \in (a, b).$$

连续随机变量的函数的分布

例4 设$X \sim U[0,\pi]$，求$Y = \sin X$的概率密度函数。

例6 设$X \sim U[-1,2]$，求$Y = X^2$的概率密度函数。

连续随机变量的函数的分布

例：某种电子元件的使用寿命X服从参数为0.2的指数分布（单位：年）。当$X > 0.5$时称元件为合格品，每个合格元件的利润为a元；不合格品则需要生产厂家免费更换新品，造成损失b元（即利润为$-b$元）。若用$Y = g(X)$ 表示零件的利润，则

$$Y = g(X) = \begin{cases} a, & X > 0.5, \\ -b, & X \leq 0.5. \end{cases}$$

求Y的概率分布。

目录

多维随机变量

很多随机现象只用一个随机变量来描述是不够的，需要用几个随机变量同时来描述。如：

- 平面上一点的位置需要用两个坐标来表示；
- 天气通常由最高、最低气温，相对湿度，风力，降水量等因素决定；
- 钢材的质量有含碳量、含硫量和硬度等基本指标。

多维随机变量

Definition (多维随机变量(multivariate random variable))

设Ω是某随机试验的样本空间，$X_1, X_2, \cdots, X_n$是该空间上的随机变量，称

$$\vec{X} = (X_1, X_2, \cdots, X_n)$$

为Ω上的随机向量或n维随机变量。

二维离散随机变量的联合分布

Definition

设二维离散型随机变量(X, Y)的所有可能取值为

$$(x_i, y_j), \quad i = 1, 2, \cdots, j = 1, 2, \cdots$$

称

$$P\{X = x_i, Y = y_j\} = p_{ij},$$

$$i = 1, 2, \cdots, j = 1, 2, \cdots$$

为该随机向量的联合概率分布。

二维离散随机变量的联合分布

二维离散型随机向量的联合分布也可以用表格表示：

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

二维离散随机变量的联合分布

例1 设有10件产品，其中3件一等品，5件二等品，2件三等品。现从中任取四件。求一等品、二等品件数的联合概率分布。

二维离散随机变量的联合分布

例3 随机掷三枚骰子，用X, Y分别表示三枚骰子中出现点数的最小值与最大值，求二维随机变量(X, Y)的联合分布。

二维离散随机变量的联合分布

练习：袋中有标上号码1、2、2的三个球，从中任取一个且不再放回，然后再从袋中任取一球，以X,Y分别表示第一、二次取到球上的号码数，求(X,Y)的分布律。

二维随机变量的联合分布函数

Definition

设(X, Y)为二维随机向量，称二元函数

$$F(x, y) = P\{X \leq x, Y \leq y\}$$

为(X, Y)的联合分布函数(joint distribution function)。

二维随机变量的联合分布函数

分布函数的性质：

1. $F(x,y)$对每个自变量都是广义单增的；
2. $0 \le F(x,y) \le 1$；
3. $F(x,-\infty) = F(-\infty,y) = 0$，$F(+\infty,+\infty) = 1$；
4. 随机向量(X,Y)落在矩形区域$(x_1,x_2] \times (y_1,y_2]$内的概率为

$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\}$$

$$=F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1).$$

二维连续随机变量的联合密度

Definition (joint density)

如果存在一个非负函数$f(x,y)$，使得二元随机变量(X,Y)的分布函数$F(x,y)$可以写成

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(s,t)dtds,$$

则称(X,Y)为二元连续性随机变量。函数$f(x,y)$称为(X,Y)的联合概率密度。

二维连续随机变量的联合密度

联合概率密度函数的基本性质:

1. $f(x,y) \geq 0, \forall x,y;$
2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1;$
3. 若函数f在点(x,y)处连续，则有

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y).$$

二维连续随机变量的联合密度

联合概率密度函数的基本性质：

4. 对任意的平面区域D，

$$P\{(X,Y)\in D\} = \iint\limits_{(x,y)\in D} f(x,y)dxdy.$$

特别地，

$$P\{a<X\leq b, c<Y\leq d\} = \int_a^b \int_c^d f(x,y)dydx.$$

二维连续随机变量的联合密度

Definition

设D是平面上的有界区域，其面积为d，若二维随机向量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}\frac{1}{d} & (x,y)\in D\\ 0 & \text{else}\end{cases}$$

则称(X,Y)服从D上的均匀分布。

二维连续随机变量的联合密度

若(X,Y)服从D上的均匀分布，则(X,Y)落在某一区域A内的概率

$$\begin{aligned}P\{(X,Y)\in A\} &= \iint_A f(x,y)dxdy \\ &= \iint_{A\cap D}\frac{1}{d}dxdy \\ &= \frac{S}{d}\end{aligned}$$

其中S为$A\cap D$的面积。

二维连续随机变量的联合密度

例2 设二维随机变量(X, Y)服从圆域

$$R = \{(x, y) \mid x^2 + y^2 \leq r^2\}$$

上的均匀分布，求其联合概率密度。

二维连续随机变量的联合密度

例4 设二维随机变量(X, Y)服从G上的均匀分布，其中G为抛物线$y = x^2$与直线$y = x + 2$所围成的区域，求

1. (X, Y)的联合概率密度；
2. 概率$P\{X + Y \geq 2\}$。

目录

边缘分布

Definition (边缘分布(Marginal distribution))

二维随机向量(X,Y)作为一个整体，有联合分布函数$F(x,y)$，其分量X与Y都是随机变量，有各自的分布函数，分别记成$F_X(x)$和$F_Y(y)$，称为X和Y的边缘分布函数。

边缘分布由联合分布完全确定：

$$F_X(x)=F(x,+\infty),\qquad F_Y(y)=F(+\infty,y).$$

二维离散随机变量的边缘分布

设二维离散型随机向量(X, Y)的分布律为

$$P\{X = x_i, Y = y_j\} = p_{ij},$$
$$i = 1, 2, \cdots, j = 1, 2, \cdots$$

则

$$P\{X = x_i\} = \sum_j p_{ij}, \qquad i = 1, 2, \cdots$$
$$P\{Y = y_j\} = \sum_i p_{ij}, \qquad j = 1, 2, \cdots$$

二维离散随机变量的边缘分布

通常记

$$p_{i\cdot} = P\{X = x_i\} = \sum_j p_{ij}, \qquad i = 1, 2, \cdots$$

$$p_{\cdot j} = P\{Y = y_j\} = \sum_i p_{ij}, \qquad j = 1, 2, \cdots$$

两个分布分别写在联合分布的右边和下边，
“边缘”一词由此而来。

二维离散随机变量的边缘分布

例1 求以下联合分布的边缘分布

$X \backslash Y$	0	1	2	3	4
0	0	0	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{1}{42}$
1	0	$\frac{1}{14}$	$\frac{2}{7}$	$\frac{1}{7}$	0
2	$\frac{1}{70}$	$\frac{1}{7}$	$\frac{1}{7}$	0	0
3	$\frac{1}{105}$	$\frac{1}{42}$	0	0	0

二维离散随机变量的边缘分布

练习：袋中有5个球，三个标数字0，两个标数字1。现依次从袋中取出两个球，分别以X, Y表示第一和第二个球上的数字。对有放回和不放回两种抽取方式，分别写出(X, Y)的联合分布与边缘分布。

总结：边缘分布由联合分布完全确定，但反之不真，仅由边缘分布一般不能得到联合分布。

二维离散随机变量的边缘分布

练习：设随机变量ξ与η各只有$-1,0,1$等三个可能值，且满足条件

$$P\{\xi=1\}=P\{\xi=-1\}=\frac{1}{4},$$
$$P\{\xi+\eta=0\}=1.$$

试求(ξ,η)的联合概率分布。

二维连续随机变量的边缘分布

设(X,Y)是二维连续型随机向量，有联合概率密度函数$f(x,y)$，则

$$F_X(x)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(s,t)dt\right]ds,$$

$$F_Y(y)=\int_{-\infty}^{y}\left[\int_{-\infty}^{+\infty}f(s,t)ds\right]dt.$$

故X,Y的边缘概率密度分别为

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

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx.$$

二维连续随机变量的边缘分布

例2 设(X,Y)服从圆域

$$R=\{(x,y)\,|\,x^2+y^2\le r^2\}$$

上的均匀分布。求X和Y的边缘概率密度。

例4 设二维随机变量(X,Y)服从G上的均匀分布，其中G为抛物线$y=x^2$与直线$y=x+2$所围成的区域，求X和Y的边缘概率密度。

二维连续随机变量的边缘分布

练习：设(X, Y)的联合概率密度为

$$f(x, y) = x + y, \qquad x \in [0, 1], y \in [0, 1].$$

求X与Y的边缘概率密度。

二维连续随机变量的边缘分布

练习：设(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y} & 0\le x\le y \\ 0 & \text{else} \end{cases}$$

求X与Y的边缘概率密度。

目录

条件分布

设X为样本空间Ω上的随机变量，固定事件B ($P(B) > 0$)，则每一事件A都有关于B的条件概率$P(A|B)$。

特别地，X关于这个条件概率的分布就是事件B发生条件下X的条件分布。如X关于B的条件分布函数为

$$F(x|B) = P\{X \leq x|B\} = \frac{P\{X \leq x, B\}}{P(B)}.$$

二维离散随机变量的条件分布

Definition

当$p_{\cdot j} > 0$时，称

$$P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{\cdot j}} \qquad i = 1, 2, \cdots$$

为$Y = y_j$时X的条件概率分布。

当$p_{i\cdot} > 0$时，称

$$P\{Y = y_j | X = x_i\} = \frac{p_{ij}}{p_{i\cdot}} \qquad j = 1, 2, \cdots$$

为$X = x_i$时Y的条件概率分布。

二维离散随机变量的条件分布

例1 求以下联合分布中，$Y = 2$时X的条件概率分布，及$X = 1$时Y的条件概率分布：

$X \backslash Y$	0	1	2	3	4
0	0	0	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{1}{42}$
1	0	$\frac{1}{14}$	$\frac{2}{7}$	$\frac{1}{7}$	0
2	$\frac{1}{70}$	$\frac{1}{7}$	$\frac{1}{7}$	0	0
3	$\frac{1}{105}$	$\frac{1}{42}$	0	0	0

二维离散随机变量的条件分布

练习：二维随机变量(X,Y)的联合分布及边缘分布为

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{7}{15}$	$\frac{7}{30}$	$\frac{7}{10}$
1	$\frac{7}{30}$	$\frac{1}{15}$	$\frac{3}{10}$
$p_{\cdot j}$	$\frac{7}{10}$	$\frac{3}{10}$	

求Y在X的各个取值下的条件分布。

二维连续随机变量的条件分布

Definition

给定y，设对任意$\Delta y > 0$,

$$P\{y - \Delta y < Y \leq y + \Delta y\} > 0.$$

若对任意x，极限

$$\lim_{\Delta y \to 0} P\{X \leq x | y - \Delta y < Y \leq y + \Delta y\}$$

存在，则称该极限为$Y = y$条件下，X的条件分布函数，记为$P\{X \leq x | Y = y\}$或$F_{X|Y}(x|y)$。

二维连续随机变量的条件分布

Theorem

若$f(\cdot,\cdot)$在点(x,y)处连续，$f_Y(\cdot)$在点y处连续，且$f_Y(y)>0$，则

$$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{f(s,y)}{f_Y(y)}ds.$$

二维连续随机变量的条件分布

Definition (conditional probability density function)

若$f_Y(y) > 0$，称

$$f_{X|Y}(x|y) := \frac{f(x,y)}{f_Y(y)}$$

为在$Y = y$条件下，X的条件概率密度；同样，

当$f_X(x) > 0$时，称

$$f_{Y|X}(y|x) := \frac{f(x,y)}{f_X(x)}$$

在$X = x$条件下，Y的条件概率密度。

二维连续随机变量的条件分布

例2 设(X,Y)服从圆域

$$R=\{(x,y)|\,x^2+y^2\le r^2\}$$

上的均匀分布。求X和Y的条件概率密度。

例4 设二维随机变量(X,Y)服从G上的均匀分布，其中G为抛物线$y=x^2$与直线$y=x+2$所围成的区域，求X和Y的条件概率密度。

二维连续随机变量的条件分布

练习：设(X, Y)的联合概率密度为

$$f(x,y) = x + y, \qquad x \in [0,1], y \in [0,1].$$

求条件概率密度。

二维连续随机变量的条件分布

练习：设(X,Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y} & 0\le x\le y \\ 0 & \text{else} \end{cases}$$

求条件概率密度。

目录
随机变量
离散型
常用离散分布
连续型
分布函数
概率密度
常用连续分布
函数的分布
联合分布
边缘分布
条件分布
随机变量的独立性
二维变量函数的分布

目录

随机变量的独立性

Definition

设二维随机向量(X,Y)的联合分布函数为$F(x,y)$，边缘分布分别为$F_X(x)$和$F_Y(y)$，若对任意实数x,y有

$$F(x,y)=F_X(x)\cdot F_Y(y),$$

则称X,Y相互独立。

X,Y相互独立即是指对任意实数x,y，事件$\{X\le x\}$与$\{Y\le y\}$相互独立。

离散随机变量的独立性

设二维离散型随机向量(X,Y)的分布律为

$$P\{X=x_i, Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots,$$

则X与Y的边缘分布分别为

$$p_{i\cdot}=P\{X=x_i\}=\sum_j p_{ij},\qquad i=1,2,\cdots;$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_i p_{ij},\qquad j=1,2,\cdots.$$

则X,Y相互独立的充要条件为

$$p_{ij}=p_{i\cdot}\cdot p_{\cdot j},\quad i,j=1,2,\cdots.$$

离散随机变量的独立性

练习：二维随机变量(X,Y)的服从以下分布律

$X \backslash Y$	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

判断X,Y的独立性。

离散随机变量的独立性

练习：袋中有标上号码1、2、2的三个球，从中任取一个且不再放回，然后再从袋中任取一球，以X,Y分别表示第一、二次取到球上的号码数，求(X,Y)的分布律，并判断两个随机变量的独立性。

离散随机变量的独立性

练习：袋中有5个球，三个标数字0，两个标数字1。现依次从袋中取出两个球，分别以X, Y表示第一和第二个球上的数字。对有放回和不放回两种抽取方式，分别讨论两个随机变量的独立性。

离散随机变量的独立性

例2 二维随机变量(X, Y)的服从以下分布律

$X \backslash Y$	y_1	y_2	y_3
x_1	0.08	α	0.12
x_2	β	0.3	γ

当α, β, γ取何值时，X, Y相互独立？

连续随机变量的独立性

若(X,Y)是连续型随机向量，则X,Y相互独立的充要条件为：对几乎所有的实数x,y，有

$$f(x,y)=f_X(x)\cdot f_Y(y).$$

注：这里的“几乎所有”是指：例外的点(x,y)组成的集合在xOy平面上面积为零。

连续随机变量的独立性

例1 设(X,Y)的概率密度为

$$f(x,y)=\begin{cases} 2e^{-(x+2y)} & x,y>0 \\ 0 & \text{else}, \end{cases}$$

判断X与Y是否相互独立。

连续随机变量的独立性

练习：设(X,Y)的联合概率密度为

$$f(x,y)=x+y,\qquad x\in[0,1],y\in[0,1].$$

判断X与Y是否相互独立。

连续随机变量的独立性

练习：设(X,Y)的概率密度为

$$f(x,y)=\begin{cases} xe^{-(x+y)} & x,y>0 \\ 0 & \text{else} \end{cases}$$

判断X与Y是否相互独立。

连续随机变量的独立性

练习：设(X, Y)的概率密度为

$$f(x,y)=\begin{cases} e^{-y} & 0\le x\le y \\ 0 & \text{else} \end{cases}$$

判断X与Y是否相互独立。

连续随机变量的独立性

若两个随机变量相互独立，则可直接由边缘分布得到联合分布。

练习：设随机变量X与Y相互独立，且都服从$[0,1]$上的均匀分布，求(X,Y)的联合概率密度。

练习：设随机变量X与Y相互独立，分别服从$[2,4]$上的均匀分布与参数为3的指数分布，求(X,Y)的联合概率密度。

连续随机变量的独立性

若两个随机变量相互独立，则每个随机变量的条件分布都等于边缘分布。

目录

离散随机向量的函数的分布

总结：设离散型随机向量(X,Y)的联合分布为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots$$

令$Z=g(X,Y)$，则Z也是一个离散型随机变量，其分布可按如下步骤求得

1. 根据函数关系列出Z的所有可能值；
2. 对Z的每个可能值z，$P\{Z=z\}$等于所有满足$g(x_i,y_j)=z$的p_{ij}之和。

离散随机向量的函数的分布

例1 设随机变量X与Y相互独立，且都服从二项分布$B(2,\dfrac{1}{2})$，求它们的和$Z=X+Y$的分布。

推广：设随机变量X与Y相互独立，分别服从二项分布$B(n,p)$与$B(m,p)$，则它们的和$Z=X+Y$服从二项分布$B(m+n,p)$。

离散随机向量的函数的分布

例2 设随机变量X与Y相互独立，分别服从参数为λ_x与λ_y的泊松分布，求它们的和$Z = X + Y$的分布。

总结：服从泊松分布的相互独立的随机变量的和也服从泊松分布，且参数等于相加的随机变量参数的和。

离散随机向量的函数的分布

练习：二维随机变量(X,Y)服从以下分布律

$X \backslash Y$	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

1. 求$Z = X + Y$的分布；
2. 求$U = \max\{X,Y\}, V = \min\{X,Y\}$的分布。

离散随机向量的函数的分布

练习：二维随机变量(X,Y)服从以下分布律

$X \backslash Y$	-2	0	3
1	0.07	0.28	0.15
2	0.09	0.22	0.19

1. 求$Z=X+Y$的分布；
2. 求$U=\max\{X,Y\}, V=\min\{X,Y\}$的分布。

连续随机向量的函数的分布

对连续型随机变量(X,Y)，求$Z=g(X,Y)$的密度函数的基本方法是

1. 根据函数关系先求Z的分布函数

$$F_Z(z)=P\{Z\le z\}=P\{g(X,Y)\le z\}$$
$$=P\{(X,Y)\in D_z\}=\iint_{D_z} f(x,y)dxdy,$$

其中$D_z=\{(x,y)|\, g(x,y)\le z\}$;

2. 然后对$F_Z(z)$求导可得Z的概率密度。

连续随机变量的和的分布

设连续型随机变量(X,Y)的密度函数为$f(x,y)$，
则$Z=X+Y$的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)dx = \int_{-\infty}^{+\infty} f(z-y, y)dy.$$

连续随机变量的和的分布

如果X与Y相互独立，概率密度分别为$f_X(x)$和$f_Y(y)$，则$Z=X+Y$的概率密度为

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx \\ &= \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y)\,dy. \end{aligned}$$

上述公式称为卷积(convolution)公式。

连续随机变量的和的分布

例3 设随机变量X与Y相互独立，且都服从$[a,b]$上的均匀分布，求它们的和$Z=X+Y$的分布。

连续随机变量的和的分布

练习：设某种商品在一周内的需要量是一个随机变量，其概率密度函数为

$$f(x)=\begin{cases} xe^{-x} & x>0 \\ 0 & x\le 0 \end{cases}$$

如果各周的需要量相互独立，求两周需要量的概率密度函数。

连续随机变量的商的分布

例4 设随机变量X与Y相互独立，分别服从参数为λ和μ的指数分布，求随机变量$Z=\dfrac{X}{Y}$的概率密度。

连续随机变量的平方和的分布

例5 设二维随机变量(X, Y)的联合概率密度为

$$f(x,y)=\begin{cases}\frac{8}{\pi(x^2+y^2+1)^3}, & x\geq 0, y\geq 0,\\ 0, & \text{else},\end{cases}$$

求随机变量$Z=X^2+Y^2$的概率密度。

连续随机变量的最值的分布

设连续型随机变量(X, Y)的联合分布函数为$F(x, y)$，则$U = \max\{X, Y\}$的分布为

$$F_U(u) = F(u, u).$$

特别若X与Y相互独立，边缘分布分别为$F_X(x)$和$F_Y(y)$，则有

$$F_U(u) = F_X(u)F_Y(u).$$

连续随机变量的最值的分布

设(X,Y)的联合分布函数为$F(x,y)$，边缘分布分别为$F_X(x)$和$F_Y(y)$，则$V=\min\{X,Y\}$的分布为

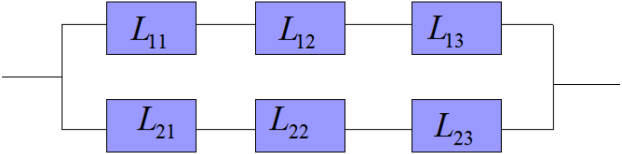
$$F_V(v)=F_X(v)+F_Y(v)-F(v,v).$$

特别若X与Y相互独立，则有

$$\begin{aligned}F_V(v)&=F_X(v)+F_Y(v)-F_X(v)F_Y(v)\\&=1-(1-F_X(v))(1-F_Y(v)).\end{aligned}$$

连续随机变量的最值的分布

例6 电子仪器由六个相互独立的部件L_{ij} $(i=1,2;\ j=1,2,3)$组成，联接方式如图所示。设各个部件的使用寿命X_{ij}均服从参数为λ的指数分布，求仪器使用寿命的概率密度。



混合型随机向量的函数的分布

例：设随机变量X, Y相互独立，X的概率分布为$P\{X=i\}=\dfrac{1}{3}\,(i=-1,0,1)$，$Y$的概率密度为

$$f_Y(y)=\begin{cases} 1 & 0\le y\le 1 \\ 0 & \text{其它} \end{cases},$$

记$Z=X+Y$，$U=\max\{X,Y\}$，

1. 求$P\{Z\le \frac{1}{2}|X=0\}$；
2. 求Z的概率密度；
3. 求U的分布函数。