随机变量

汪超男

C_WANG@JNU.EDU.CN

2 随机变量

- 随机变量
- 离散随机变量
- 连续随机变量
- 随机变量的期望
- 联合分布的随机变量
- 矩母函数

3 随机变量

- 在做试验时，相对于试验结果本身，常常对结果的某些函数感兴趣
  - 掷两颗骰子，关心的是点数和
- 定义在样本空间上的实值函数称为随机变量（random variable, *r.v.*）
  - 随机变量将样本空间上每个可能的结果映射（map）到一个实数上
- 因为随机变量的值是由试验的结果决定的，可以给随机变量的可能值指定概率

- 随机变量不是变量，而是一个函数
- 映射不是随机的，而是一定的

4 例：以X记随机变量，它定义为两颗均匀的骰子的点数和

$$P\{X=2\}=P\{(1,1)\}=\frac{1}{36}$$
$$P\{X=3\}=P\{(1,2),(2,1)\}=\frac{2}{36}$$
$$P\{X=4\}=P\{(1,3),(2,2),(3,1)\}=\frac{3}{36}$$
$$P\{X=5\}=P\{(1,4),(2,3),(3,2),(4,1)\}=\frac{4}{36}$$
$$P\{X=6\}=P\{(1,5),(2,4),(3,3),(4,2),(5,1)\}=\frac{5}{36}$$
$$P\{X=7\}=P\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}=\frac{6}{36}$$
$$P\{X=8\}=P\{(2,6),(3,5),(4,4),(5,3),(6,2)\}=\frac{5}{36}$$
$$P\{X=9\}=P\{(3,6),(4,5),(5,4),(6,3)\}=\frac{4}{36}$$
$$P\{X=10\}=P\{(4,6),(5,5),(6,4)\}=\frac{3}{36}$$
$$P\{X=11\}=P\{(5,6),(6,5)\}=\frac{2}{36}$$
$$P\{X=12\}=P\{(6,6)\}=\frac{1}{36}$$

$$1=P\left(\bigcup_{n=2}^{12}\{X=n\}\right)=\sum_{n=2}^{12}P\{X=n\}$$

5 例：抛掷两枚均匀的硬币组成，以Y记出现正面的次数

$$P\{Y=0\}=P\{(T,T)\}=\frac{1}{4}$$
$$P\{Y=1\}=P\{(T,H),(H,T)\}=\frac{1}{2}$$
$$P\{Y=2\}=P\{(H,H)\}=\frac{1}{4}$$

$$P\{Y=0\}+P\{Y=1\}+P\{Y=2\}=1$$

6 示性随机变量

- 例：假定试验是观察电池在损耗前能用多久，并假定我们主要并不关心电池的实际寿命，而只是关心电池是否至少能用两年，在这种情形下，我们可以定义随机变量I为

$$I=\begin{cases}1, & \text{若电池的寿命是两年或更长}\\ 0, & \text{其他情形}\end{cases}$$

- 如果以E记电池能使用两年或更长，那么随机变量I称为事件E的示性（indicator）随机变量（I的取值依赖于E是否发生）

7 累积分布函数

- 累积分布函数（cumulative distribution function， cdf）,简称分布函数
- 对于任意实数$b, -\infty < b < \infty$，随机变量X的cdf $F(\cdot)$ 定义为

$$F(b) = P\{X \leq b\}$$

- $F(b)$为随机变量X取一个小于或者等于b的值的概率
  - $F(b)$是b的非减函数
  - $\lim_{b\to\infty} F(b) = F(\infty) = 1$
  - $\lim_{b\to-\infty} F(b) = F(-\infty) = 0$

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

注意：$F(b) \neq P\{X < b\}$

8 离散与连续随机变量

- 一个最多取可数个可能值的随机变量，称为离散（discrete）随机变量
  - 例：掷骰子

- 一个取值为某一范围实数（不可数的）的随机变量，称为连续（Continuous）随机变量
  - 例：顾客的到达时间（一般为某一事件发生的时间）

9 随机变量

✓ 随机变量
- 离散随机变量
- 连续随机变量
- 随机变量的期望
- 联合分布的随机变量
- 矩母函数
- 极限定理

10 概率质量函数

- 一个最多取可数个可能值的随机变量，称为离散（discrete）随机变量
- 概率质量函数（ probability mass function， pmf），记为$p(a)$

$$p(a) = P\{X = a\}$$

  - 如果X必须是值$x_1, x_2, \cdots$之一，那么
    - $p(x_i) > 0, i = 1, 2, \ldots$
    - $p(x) = 0,$ 所有其他x值
    - $\sum_{i=0}^{\infty} p(x_i) = 1$

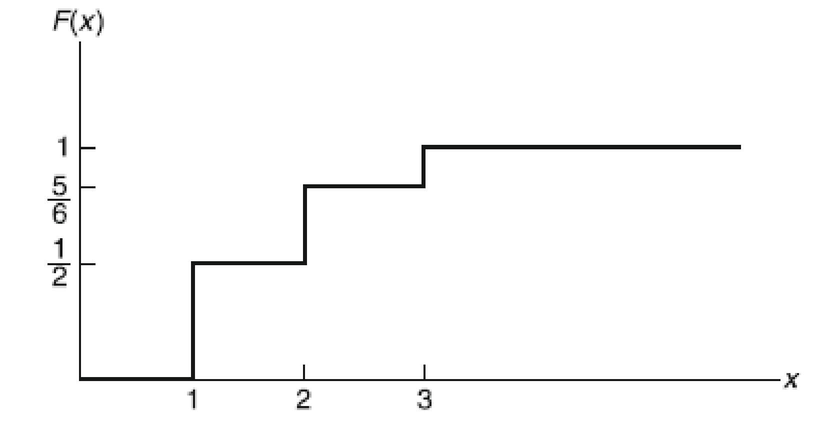
$$F(a) = \sum_{all\ x_i \leq a} p(x_i)$$

例：假定X的概率质量函数为

$$p(1)=\frac{1}{2}, p(2)=\frac{1}{3}, p(3)=\frac{1}{6}$$

X的累积分布函数为

$$F(a)=\begin{cases}0, & a<1\\ \frac{1}{2}, & 1\leq a<2\\ \frac{5}{6}, & 2\leq a<3\\ 1, & 3\leq a\end{cases}$$



12 伯努利随机变量

- 假定一个试验，其结果可以分为成功或者失败，如果试验的结果是成功时令X等于1，而在试验的结果是失败时令X等于0，那么该随机变量X称为伯努利（Bernoulli）随机变量，X的概率质量函数pmf为

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

  - 其中$p, 0 \leq p \leq 1$是试验的结果为成功的概率

二项随机变量

- 假定做了n次独立试验，其中每次结果为成功的概率为p，结果为失败的概率为$1-p$，如果以X代表出现在n次试验中的成功的次数，那么X称为具有参数(n,p)的二项（binomial）随机变量，概率质量函数为

$$p(i)=\binom{n}{i}p^i(1-p)^{n-i}, i=0,1,\ldots,n$$

$$\sum_{i=0}^{\infty}p(i)=\sum_{i=0}^{\infty}\binom{n}{i}p^i(1-p)^{n-i}=\big(p+(1-p)\big)^n=1$$

二项式定理

例子

- 抛掷4枚均匀的硬币，问得到2个正面、2个反面的概率是多少？（例2.6）
- 已知某机器生产的一个产品是废品的概率是0.1，在三个产品的样本中，至多有一个废品的概率是多少？（例2.7）
- 假定在飞行中，飞机发动机失效的概率为 $1-p$ ，如果至少 50% 的发动机保持运行，那么飞机就能完成一次成功的飞行，问p取什么样的值，4个发动机的飞机比2个发动机的飞机更可靠？（例2.8）

15 几何随机变量

- 假定进行独立试验直到出现一个结果为成功，其中每一个试验成功的概率都是p，X为直到出现首次成功所需要做的试验次数，那么称X为具有参数p的几何（geometric）随机变量，它的概率质量函数为

$$p(n) = P\{X = n\} = (1-p)^{n-1}p, n = 1,2,\ldots$$

$$\sum_{n=1}^{\infty} p(n) = p\sum_{n=1}^{\infty}(1-p)^{n-1} = 1$$

无穷递降等比数列求和

16 例：抛掷一枚出现正面的概率为p的硬币直至正面首次出现，以N记需要抛掷的次数

$$
\begin{aligned}
P\{N=1\} &= P\{H\} = p, \\
P\{N=2\} &= P\{(T,H)\} = (1-p)p, \\
P\{N=3\} &= P\{(T,T,H)\} = (1-p)^2 p, \\
&\vdots \\
P\{N=n\} &= P\{(\underbrace{T,T,\ldots,T}_{n-1},H)\} = (1-p)^{n-1}p,\ n \geq 1
\end{aligned}
$$

17 泊松随机变量

- 对于取值于$0,1,2,\ldots$的随机变量X，如果对于某个$\lambda > 0$，有概率质量函数

$$p(i) = P\{X = i\} = e^{-\lambda}\frac{\lambda^i}{i!}, i = 0,1,\ldots$$

则称X为具有参数λ的泊松（Poisson）随机变量

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1$$

- 泊松分布的重要性质：当n很大，p很小，参数为(n,p)的二项分布的概率值可以由参数为$\lambda = np$的泊松分布的概率值近似

实例

- 假定在书的一页上的印刷错误的个数是一个具有参数$\lambda = 1$的泊松随机变量，计算在此页上至少有一个错误的概率。（例2.10）

- 计算机硬件公司制造某种特殊型号的微型芯片，次品率达0.1%，各芯片成为次品相互独立，求在1000只产品中至少有2只次品的概率。

19 随机变量

- ✓ 随机变量
- ✓ 离散随机变量
- 连续随机变量
- 随机变量的期望
- 联合分布的随机变量
- 矩母函数
- 极限定理

20 # 连续随机变量

- 一个可能值是不可数的随机变量，称为连续（Continuous）随机变量
- 概率密度函数（probability density function，pdf）$f(x)$：对于连续随机变量X，如果存在一个定义在所有实数$x \in (-\infty, \infty)$上的非负函数$f(x)$，使得对于任意实数集合B有

$$P\{X \in B\} = \int_B f(x)dx$$

  - X在B中的概率可以由概率密度函数在集合B上求积分得到

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

概率密度函数

- 设$B = [a,b]$

$$P\{a \le X \le b\} = \int_a^b f(x)dx$$

- 如果在上面设$a = b$

$$P\{X = a\} = \int_a^a f(x)dx = 0$$

  - 说明了连续随机变量在假定为某个特殊值时的概率为零

22 累积分布函数

- 累积分布函数$F(\cdot)$与概率密度函数$f(\cdot)$的关系表示为

$$F(a) = P\{X \in (-\infty, a]\} = \int_{-\infty}^{a} f(x)dx$$

  - 对上式两边求微分

$$\frac{d}{da}F(a) = f(a)$$

- 概率密度函数是累积分布函数的导数

当h很小时，
$$P\left\{a - \frac{h}{2} \le X \le a + \frac{h}{2}\right\} = \int_{a-\frac{h}{2}}^{a+\frac{h}{2}} f(x)dx \approx hf(a)$$

- $f(a)$是随机变量在a附近可能性大小的量度

均匀随机变量

- 一个随机变量称为均匀（uniform）分布在区间(α,β)上，如果其概率密度函数为

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

$$\int_{-\infty}^{\infty}f(x)dx=\int_{\alpha}^{\beta}\frac{1}{\beta-\alpha}dx=1$$

- 对于(α,β)内任一长度为l的子区间，$\alpha\le c<c+l\le\beta$，有

$$P\{c<X\le c+l\}=\int_{c}^{c+l}f(x)dx=\int_{c}^{c+l}\frac{1}{\beta-\alpha}dx=\frac{l}{\beta-\alpha}$$

- X落在(α,β)的子区间内的概率只依赖于子区间的长度而与子区间的位置无关

例：计算均匀分布在(α,β)上的随机变量的累积分布函数。（例2.13）

$$F(a)=\int_{-\infty}^{a} f(x)dx$$

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

例：如果X均匀分布在$(0,10)$上，计算概率$(a)\ X<3$，$(b)\ X>7$，$(c)\ 1<X<6$。（例2.14）

25 指数随机变量

- 具有参数λ的指数（Exponential）随机变量X的概率密度函数为

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

- 累积分布函数

$$F(a) = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}, a \geq 0$$

$$F(\infty) = \int_0^\infty \lambda e^{-\lambda x} dx = 1$$

26 伽马随机变量

- 具有参数λ和α的伽马（Gamma）随机变量X的概率密度函数为

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

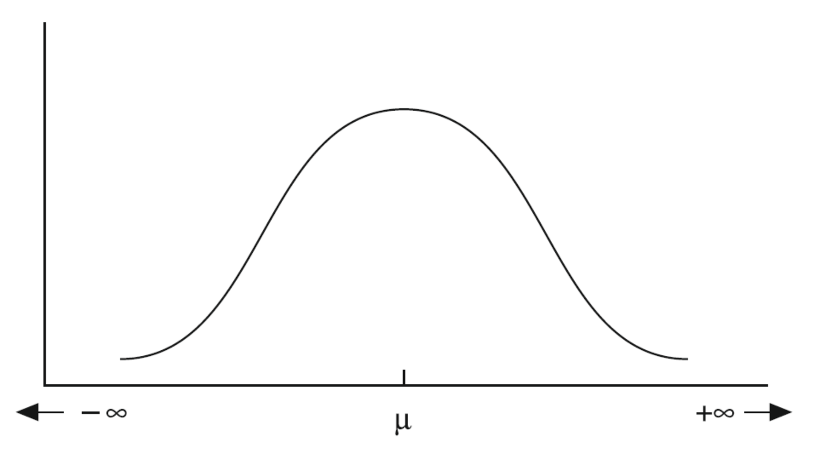
- 其中$\Gamma(\alpha)$称为伽马函数

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

27 正态随机变量

- 具有参数μ和σ^2的正态（Normal）随机变量X的概率密度分布为

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



- 密度函数是一条关于μ对称的钟形曲线
- 如果X以参数μ和σ^2正态地分布
  - $Y = \alpha X + \beta$以参数$\alpha\mu + \beta$和$(\alpha\sigma)^2$正态地分布
  - $Y = (X - \mu)/\sigma$以参数0和1正态地分布，该随机变量Y称为标准正态分布或单位正态分布

28 作业

- 课本4， 5， 9， 15， 32， 33

随机变量

- ✓ 随机变量
- ✓ 离散随机变量
- ✓ 连续随机变量
- 随机变量的期望
- 联合分布的随机变量
- 矩母函数
- 极限定理

离散随机变量的期望

- 对于离散随机变量X，其概率质量函数$p(x)$，期望值（expected value）为

$$E[X] = \sum_{x:p(x)>0} xp(x)$$

- X的期望值是X可能取的值的加权平均，每个值被X取此值的概率所加权
- 例：X等可能取值1或2，求期望值

$$E[X] = 1\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right) = \frac{3}{2}$$

例：X是参数为p的伯努利随机变量，计算$E[X]$ （例2.16）

$$E[X] = p$$

在一次试验中平均成功的次数正是成功的概率

例：X是参数为n和p二项随机变量，计算$E[X]$ （例2.17）

$$E[X] = np$$

n次独立试验的平均成功次数是n乘以一次试验成功的概率

例：X是参数为p的几何随机变量，计算$E[X]$ （例2.18）

$$E[X] = \frac{1}{p}$$

直至达到首次成功所需做的独立试验的期望数等于任意一次试验结果是成功的概率的倒数

例：X是参数为λ的泊松随机变量，计算$E[X]$ （例2.19）

$$E[X] = \lambda$$

32 连续随机变量的期望

- 对于连续随机变量X，其概率密度函数$f(x)$，期望值为

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

例：计算均匀分布在(α,β)上的随机变量的期望（例2.20）

$$E[X]=\frac{\alpha+\beta}{2}$$

在(α,β)上均匀分布的随机变量的期望值正是区间的中点

例：X是参数为λ的指数随机变量，计算$E[X]$（例2.21）

$$E[X]=\frac{1}{\lambda}$$

例：X是参数为μ和σ^2的正态随机变量，计算$E[X]$（例2.22）

$$E[X]=\mu$$

34 随机变量的函数的期望

- 假定X有如下的概率质量函数，计算$E[X^2]$（例2.23）

$$p(0) = 0.2, p(1) = 0.5, p(2) = 0.3$$

- 假定X在$(0,1)$上均匀分布，计算$E[X^3]$（例2.24）

35 随机变量的函数的期望

- 如果X是离散随机变量，有概率质量函数$p(x)$，那么对于任意实值函数g

$$E[g(x)] = \sum_{x:p(x)>0} g(x)p(x)$$

- 如果X是连续随机变量，有概率密度函数$f(x)$，那么对于任意实值函数g

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx$$

- 如果a和b都是常数，那么$E[aX + b] = aE[X] + b$

36 n阶矩与方差

- $E[X^n]$，$n \geq 1$称为X的n阶矩

$$E[X^n] = \begin{cases} \sum_{x:p(x)>0} x^n p(x)\,, & 若X是离散的 \\ \int_{-\infty}^{\infty} x^n f(x)dx\,, & 若X是连续的 \end{cases}$$

- 随机变量X的方差，记为$\mathrm{Var}(X)$，定义为

$$\mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

  - X的方差度量了X与其期望之间的偏差平方的期望

37 方差

- 设X是参数为μ和σ^2的正态随机变量，其方差（例2.27）

$$\mathrm{Var}(X)=\sigma^2$$

- 例：如果X代表掷一颗均匀的骰子的结果，求$\mathrm{Var}(X)$（例2.28）

随机变量

- ✓随机变量
- ✓离散随机变量
- ✓连续随机变量
- ✓随机变量的期望
- 联合分布的随机变量
- 矩母函数
- 极限定理

39 离散随机变量的联合分布

- 任意两个离散随机变量X和Y的联合累积概率分布函数（joint cumulative probability distribution function）为$F(a,b) = P\{X \le a, Y \le b\}$，$-\infty < a, b < \infty$
- 则X和Y的累积分布函数（边缘分布函数）为

$$F_X(a) = P\{X \le a\} = P\{X \le a, Y < \infty\} = F(a,\infty)$$

$$F_Y(b) = P\{Y \le b\} = P\{X < \infty, Y \le b\} = F(\infty, b)$$

- X和Y联合概率质量函数$p(x,y) = P\{X = x, Y = y\}$
- 则X和Y的概率质量函数（边缘分布律）为

$$p_X(x) = \sum_{y:p(x,y)>0} p(x,y) \qquad p_Y(y) = \sum_{x:p(x,y)>0} p(x,y)$$

连续随机变量的联合分布

- 我们说X和Y联合地连续（jointly continuous），如果存在一个对于所有的实数x和y定义的函数$f(x,y)$，对于所有的实数集合A和B满足

$$P\{X \in A, Y \in B\} = \int_B \int_A f(x,y)dxdy$$

  - $f(x,y)$为X和Y的联合概率密度函数

- X和Y的概率密度函数（边缘概率密度）为

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

- 概率分布函数与概率密度函数

$$F(a,b) = P(X \le a, Y \le b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)dydx \qquad \frac{d^2}{da\,db}F(a,b) = f(a,b)$$

练习

- Suppose two discrete *r.v.s* X and Y have the following joint p.m.f as shown in the table. Thus, X assumes values $1, 2,$ and 3; Y assumes values of $2, 3,$ and 4.

$X \backslash Y$	2	3	4
1	$1/6$	$1/6$	$1/6$
2	0	$1/6$	$1/6$
3	0	0	$1/6$

- Find the marginal p.m.f. p_X and p_Y

联合分布函数的期望

- 如果X和Y都是随机变量，g是一个双变量函数，那么

$$E[g(X,Y)] = \begin{cases} \sum_y \sum_x g(x,y)p(x,y) & \text{离散情形} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy & \text{连续情形} \end{cases}$$

- $g(X,Y) = X + Y$ $\qquad E[X+Y] = E[X] + E[Y]$
- 对于任意常数a和b $\qquad E[aX+bY] = aE[X] + bE[Y]$
- 若$X_1, \dots, X_n$是n个随机变量，那么对于n个常数$a_1, \dots, a_n$

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$$

例：掷三颗均匀的骰子，计算点数和的期望（例2.29）

例：n个伯努利随机变量和的期望（例2.30）

例：在一次聚会上，n个人将帽子扔到房间的中央，帽子混杂了以后，每个人随机的取一个，求取到自己的帽子的人的期望数（例2.31）

独立随机变量

- 如果对于一切a和b，$P\{X \le a, Y \le b\} = P\{X \le a\}P\{Y \le b\}$，称随机变量$X$和$Y$是独立的

$$F(a,b) = F_X(a)F_Y(b) \quad \text{对于一切}a,\ b\text{成立}$$

- 当X和Y都是离散的，独立的条件简化为$p(x,y) = p_X(x)p_Y(y)$
- 当X和Y联合地连续，独立的条件简化为$f(x,y) = f_X(x)f_Y(y)$
- 若X和Y是独立的，那么对于任意函数g和h，$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
- n个随机变量$X_1, X_2, \dots, X_n$称为独立的，如果对于所有的值$a_1, a_2, \dots, a_n$有

$$P\{X_1 \le a_1, X_2 \le a_2, \dots, X_n \le a_n\} = P\{X_1 \le a_1\}P\{X_2 \le a_2\} \cdots P\{X_n \le a_n\}$$

45 协方差

- 任意两个随机变量X和Y的协方差，记为$\mathrm{Cov}(X,Y)$，定义为

$$\mathrm{Cov}(X,Y)=E[(X-E[X])(Y-E[Y])]=E[XY]-E[X]E[Y]$$

  - $\mathrm{Cov}(X,Y)=0$，X和Y独立
  - $\mathrm{Cov}(X,Y)>0$，X增加时，Y倾向于增加
  - $\mathrm{Cov}(X,Y)<0$，X增加时，Y倾向于减少

- 对于任意随机变量X,Y,Z和常数c

(1) $\mathrm{Cov}(X,X)=\mathrm{Var}(X)$

(2) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$

(3) $\mathrm{Cov}(cX,Y)=c\mathrm{Cov}(X,Y)$

(4) $\mathrm{Cov}(X,Y+Z)=\mathrm{Cov}(X,Y)+\mathrm{Cov}(X,Z)$

例2.33：X, Y的联合密度函数是

$$f(x,y) = \frac{1}{y} e^{-(y+x/y)}, 0 < x, y < \infty$$

(a) 验证上述函数是联合密度函数

(b) 计算$\mathrm{Cov}(X,Y)$

方差的性质

- 对于任意随机变量X, Y和常数c
- $\mathrm{Var}(c) = 0$
- $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$
- 如果X, Y独立， $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

48 样本均值

- 若$X_1, \dots, X_n$是独立同分布的，则随机变量$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$称为样本均值(Sample mean)
- 若$X_1, \dots, X_n$是独立同分布的，具有期望值μ和方差σ^2，那么

  $(a)\ E[\bar{X}] = \mu$

  $(b)\ \mathrm{Var}(\bar{X}) = \sigma^2/n$

  $(c)\ \mathrm{Cov}(\bar{X}, X_i - \bar{X}) = 0, i = 1, \dots, n$

例子

- 计算参数为n和p的二项随机变量X的方差

- 考虑一个有N个人的总体，他们中一些人赞同某个提议，特别假定他们中的Np个人反对，这里假定p未知，我们关心的是通过随机地选取并确定总体中n个成员的态度，从而估计总体中赞同这个提议的人员的比率p，通常用被取样的总体赞同这个提议的比率作为p的估计，求该估计的均值和方差

50 超几何分布

- 一般地，设有N件产品，其中有$M(M \le N)$件次品。从中任取n件产品，用X表示取出的$n(n \le N)$件产品中次品的件数，X服从参数为N, M, N的超几何（hypergeometric）分布，其概率质量函数为

$$\mathrm{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

独立随机变量和的分布

- X和Y为相互独立的连续随机变量，X的概率密度函数为$f_X(x)$，Y的概率密度函数为$f_Y(y)$，$X+Y$的概率密度函数$f_{X+Y}(x)$为f_X和f_Y的卷积，记为$f_X * f_Y$

$$f_{X+Y}(z) = f_X * f_Y = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

例子

- 如果X和Y是独立的随机变量，两者都均匀地分布在$(0,1)$上，计算$X+Y$的概率密度（例2.36）

- 假定X和Y是独立的泊松随机变量，分别具有均值λ_1和λ_2，计算$X+Y$的分布（例2.37）

两个独立的泊松随机变量之和仍然服从泊松分布

随机变量

- ✓随机变量
- ✓离散随机变量
- ✓连续随机变量
- ✓随机变量的期望
- ✓联合分布的随机变量
- 矩母函数
- 极限定理

54 矩母函数

- 随机变量X的矩母函数（Moment Generating Functions）$\phi(t)$对所有值t定义为

$$\phi(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx}p(x), & 若X离散 \\ \int_{-\infty}^{\infty} e^{tx}f(x)dx, & 若X连续 \end{cases}$$

- X的所有的矩能有$\phi(t)$相继地求微分得到

$$\phi^{(n)}(0) = E[X^n], n \geq 1$$

$$\phi'(t) = \frac{d}{dt}E[e^{tX}] = E\left[\frac{d}{dt}(e^{tX})\right] = E[Xe^{tX}] \Rightarrow \phi'(0) = E[X]$$

$$\phi''(t) = \frac{d}{dt}\phi'(t) = \frac{d}{dt}E[Xe^{tX}] = E\left[\frac{d}{dt}(Xe^{tX})\right] = E[X^2e^{tX}] \Rightarrow \phi''(0) = E[X^2]$$

计算一些常见分布的$\phi(t)$

55

例2.40：参数为n和p的二项分布

例2.41：参数为λ的泊松分布

例2.42：参数为λ的指数分布

例2.43：参数为μ和σ^2的正态分布

Discrete probability distribution	**Probability mass function, $p(x)$**	**Moment generating function, $\phi(t)$**	**Mean**	**Variance**
Binomial with parameters n, p, $0 \le p \le 1$	$\binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$	$(pe^t + (1-p))^n$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda} \frac{\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$	λ	λ
Geometric with parameter $0 \le p \le 1$	$p(1-p)^{x-1}$, $x = 1, 2, \ldots$	$\frac{pe^t}{1-(1-p)e^t}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Continuous probability distribution	**Probability density function, $f(x)$**	**Moment generating function, $\phi(t)$**	**Mean**	**Variance**
Uniform over (a, b)	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x < 0 \end{cases}$	$\frac{\lambda}{\lambda - t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma with parameters (n, λ), $\lambda > 0$	$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}, & x \ge 0 \\ 0, & x < 0 \end{cases}$	$\left(\frac{\lambda}{\lambda - t}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \times \exp\{-(x-\mu)^2/2\sigma^2\}$, $-\infty < x < \infty$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$	μ	σ^2

56 矩母函数的性质

- 独立随机变量和的矩母函数正是单个矩母函数的乘积
  - 假设X和Y是独立的，矩母函数分别为$\phi_X(t)$和$\phi_Y(t)$，那么

$$\phi_{X+Y}(t) = E\left[e^{t(X+Y)}\right] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}] = \phi_X(t)\phi_Y(t)$$

- 矩母函数唯一地确定了分布
  - 在随机变量的矩母函数和分布函数之间存在一一对应

例2.44：独立二项随机变量的和
如果X和Y分别是以(n, p)和(m, p)为参数的独立二项随机变量，那么$X+Y$的分布是什么?

例2.45：独立泊松随机变量的和
如果X和Y分别是以λ_1和λ_2为参数的独立泊松随机变量，那么$X+Y$的分布是什么?

例2.46：独立正态随机变量的和
如果X和Y分别是以(μ_1, σ_1^2)和(μ_2, σ_2^2)为参数的独立正态随机变量，那么$X+Y$的分布是什么?

随机变量

- ✓随机变量
- ✓离散随机变量
- ✓连续随机变量
- ✓随机变量的期望
- ✓联合分布的随机变量
- ✓矩母函数
- 极限定理

59 马尔可夫不等式

- 如果X是只取非负值的随机变量，那么对于任意$a > 0$，有

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

切比雪夫不等式

- 如果X是具有均值μ和方差σ^2的随机变量，那么对于任意$k > 0$

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

例子

- 假设一个工厂每星期生产的产品数均值为500
  - 估算这个星期的产品至少有1000的概率
  - 如果这星期生产的产品的方差已知等于100，估算这个星期的产品在400与600之间的概率。

62 强大数定律

- 假定$X_1, X_2, \ldots$是一列独立同分布的随机变量，令$E[X_i]=\mu$，那么，当$n \longrightarrow \infty$时以概率1有

$$\frac{X_1+X_2+\cdots+X_n}{n} \longrightarrow \mu$$

- 假设做一系列独立试验，令E为某一事件，而$P(E)$记在每次特定的试验中E出现的概率，令

$$X_i = \begin{cases} 1, & \text{若在第}i\text{次试验}E\text{发生} \\ 0, & \text{若在第}i\text{次试验}E\text{未发生} \end{cases}$$

则$X_1+X_2+\cdots+X_n$代表事件E在这n次试验中发生的次数，$\frac{X_1+X_2+\cdots+X_n}{n}$为事件$E$发生次数的极限比例以概率1是$P(E)$

中心极限定理

- 假定$X_1, X_2, \ldots$是一列独立同分布的随机变量，具有均值μ和方差σ^2，那么，当$n \to \infty$时，

$$\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

的分布趋于标准正态分布，也就是说，当$n \longrightarrow \infty$时，

$$P\left\{\frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}\, dx$$

以上结论对任意分布都成立！！！

例子

- 令X为抛掷一枚均匀的硬币40次出现正面的次数，求$X = 20$的概率，用正态近似，并将结果与精确解比较.（例2.50）

- 令$X_i (i = 1, \dots, 10)$是独立的随机变量，每个均匀地分布在$(0, 1)$上，估计$P\{\sum_1^{10} X_i > 7\}$.（例2.51）

- 一种特殊型号的电池的寿命是随机变量，具有均值40小时，标准差20小时，一个电池使用到失效，就换一个新的，假定库存这样的电池25个，它们的寿命是独立的，求能得到使用多于1100小时的近似概率。（例2.52）

小结（I）

- 随机变量
  - 定义在样本空间上的实值函数；示性随机变量
- 离散随机变量
  - 累积分布函数（cdf）；概率质量函数（pmf）
  - 伯努利随机变量；二项随机变量；几何随机变量；泊松随机变量
- 连续随机变量
  - 累积分布函数（cdf）；概率密度函数（pdf）
  - 均匀随机变量；指数随机变量；伽马随机变量；正态随机变量
- 随机变量的期望
  - 离散：$E[X] = \sum_{x:p(x)>0} xp(x)$；连续：$E[X] = \int_{-\infty}^{\infty} xf(x)dx$

小结（2）

- 随机变量的函数
  - 期望：$E[g(x)] = \sum_{x:p(x)>0} g(x)p(x)$； $E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
  - X的n阶矩： $E[X^n]$， $n \geq 1$
  - X的方差： $\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$
- 随机变量的联合分布
  - 离散：cdf，pmf；连续：cdf，pdf
  - 期望：$E[g(X,Y)] = \begin{cases} \sum_y \sum_x g(x,y)p(x,y) & \text{离散情形} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy & \text{连续情形} \end{cases}$
  - 独立： $P\{X_1 \leq a_1, X_2 \leq a_2, \ldots, X_n \leq a_n\} = P\{X_1 \leq a_1\}P\{X_2 \leq a_2\} \cdots P\{X_n \leq a_n\}$
  - 协方差：$\text{Cov}(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$
  - 样本均值的期望，方差，协方差

67 小结（3）

- 矩母函数：$\phi(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx}p(x), & 若X离散 \\ \int_{-\infty}^{\infty} e^{tx}f(x)dx, & 若X连续 \end{cases}$
  - $\phi^{(n)}(0) = E[X^n], n \geq 1$
  - 独立随机变量和的矩母函数$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$
- 极限定理
  - 马尔可夫不等式：$P\{X \geq a\} \leq \frac{E[X]}{a}$
  - 切比雪夫不等式：$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$
  - 强大数定律
  - 中心极限定理

68 作业

- 40，43，54，70，86