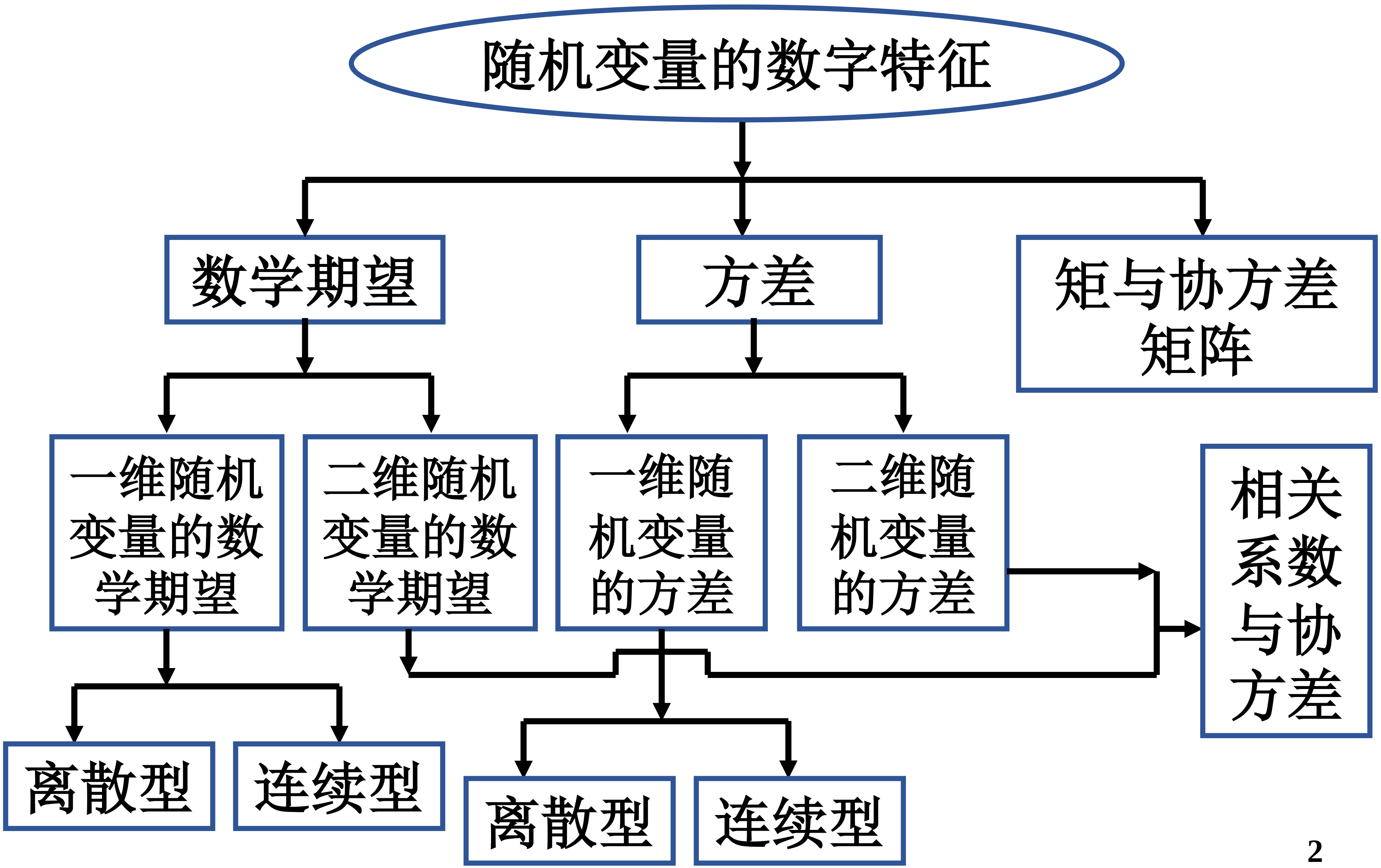
第四章
随机变量的数字特征
—刻画随机变量某一方面特征的常数

王笑尘

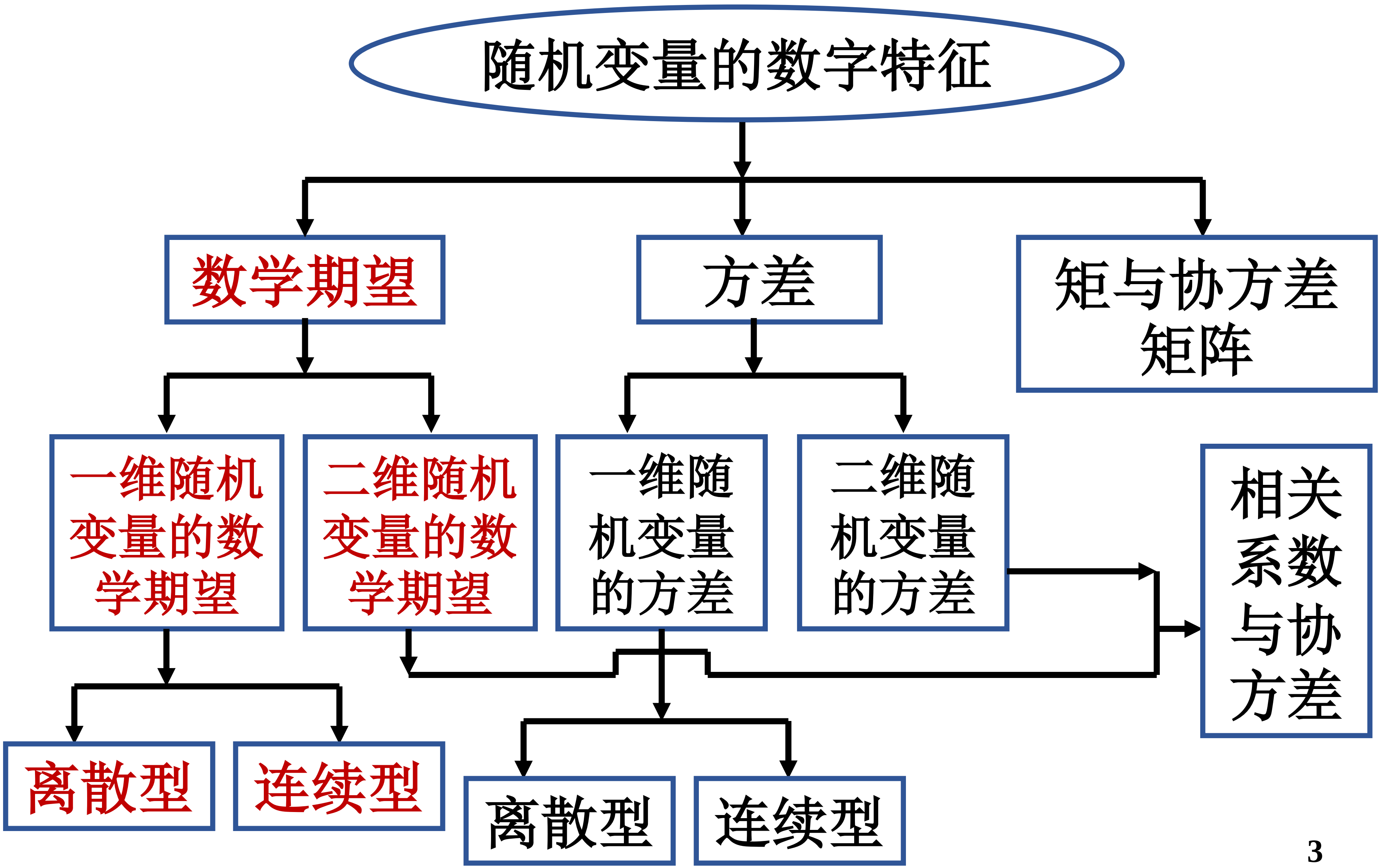
北京邮电大学网络空间安全学院

wxiaochen@bupt.edu.cn

第四章 知识结构图


随机变量的数字特征
数学期望
方差
矩与协方差矩阵
一维随机变量的数学期望
二维随机变量的数学期望
一维随机变量的方差
二维随机变量的方差
相关系数与协方差
离散型
连续型
离散型
连续型

第一节 数学期望



随机变量的数字特征
数学期望
方差
矩与协方差矩阵
一维随机变量的数学期望
二维随机变量的数学期望
一维随机变量的方差
二维随机变量的方差
相关系数与协方差
离散型
连续型
离散型
连续型

第一节 数学期望

问题的引出

引例. 某车间对工人的生产情况进行考察。车工小张每天生产的废品数 X 是一个随机变量。

问：如何定义X 的平均值呢？

一般来说，

若统计 n天, (假定小张每天至多出三件废品) 得：n_0天没有出废品; n_1天每天出一件废品;n_2天每天出两件废品； n_3天每天出三件废品。

可以得到 n天中每天的平均废品数为：

$$0\cdot\frac{n_0}{n}+1\cdot\frac{n_1}{n}+2\cdot\frac{n_2}{n}+3\cdot\frac{n_3}{n}$$

这是以频率为权的加权平均

现若统计100天得：

32天没有出废品; 30天每天出1件废品； 17天每天出2件废品；21天每天出3件废品;

于是，可以得到这100天中每天的平均废品数为：

$$0\cdot\frac{32}{100}+1\cdot\frac{30}{100}+2\cdot\frac{17}{100}+3\cdot\frac{21}{100}=1.27$$

可以想象：

若另外统计100天，车工小张不出废品，出一件、二件、三件废品的天数与前面的100天一般不会完全相同，这另外100天每天的平均废品数也不一定是1. 27。

这个数能否作为X的平均值吗？

由频率和概率的关系，不难想到，在求废品数 X 的平均值时，用概率代替频率，得平均值为：

$$0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3$$

这是以概率为权的加权平均

这样得到了一个确定的数。

现问：用这个数作为随机变量 X 的平均值是否合理呢？

注意到：对于一个随机变量，若它可能取的值是：

$x_1, x_2, \ldots,$ 相应的概率为 $p_1, p_2, \ldots,$

则对 X 作一系列观察(试验)，所得 X 的试验值的平均值也是随机的。

但是，如果试验次数很大，出现 x_k 的频率会接近于 p_k，于是试验值的平均值接近于：

$$\sum_{k=1}^{\infty} x_k p_k$$

由此：以概率为权的加权平均值作为随机变量 X 的平均值是合理的。

一. 离散型随机变量的数学期望

1. 定义1 设 X 是离散型随机变量，它的分布律为:

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

如果级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称此级数的和为随机变量 X 的**数学期望**，记为:

$$E(X)=\sum_{k=1}^{\infty} x_k p_k$$

注：▲ $E(X)$是个(实)数。它形式上是X的可能取值的加权平均值；本质上体现了X的真正的平均，故常称 $E(X)$为 X 的均值；物理上表示为一个质点系的重心坐标。

▲ $E(X)$ 的计算：当 X 的可能取值为有限时，则计算有穷和；当 X 的可能取值为无限时，则计算级数的和。

▲ 若 $\sum_{k=1}^{\infty} x_k p_k$ 不绝对收敛，则称 $E(X)$ 不存在。

▲ 推广到二维：
$$\begin{cases} E(X)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i\, p_{ij} \\ E(Y)=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} y_j\, p_{ij} \end{cases}$$

p_{ij} 为联合分布律

例1 某人的一串钥匙上有 n 把钥匙，其中只有一把能打开自己的家门，他随意地试用这串钥匙中的某一把去开门，若每把钥匙试开一次后除去。

求：打开门时试开次数的数学期望。

例1 某人的一串钥匙上有 n 把钥匙，其中只有一把能打开自己的家门，他随意地试用这串钥匙中的某一把去开门，若每把钥匙试开一次后除去。

求： 打开门时试开次数的数学期望。

解： 设试开次数为 X，则：

$$P(X=k)=1/n\ ,\quad k=1,2,\ldots,n$$

于是，由数学期望的定义得：

$$E(X)=\sum_{k=1}^{n}k\cdot\frac{1}{n}=\frac{1}{n}\cdot\frac{(1+n)n}{2}=\frac{n+1}{2}$$

2. 几种常见分布的数学期望

(1) (0−1)分布

若随机变量 X 只能取 0 与 1 两个值，它的分布律为:

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

则: $\mathrm{E(X)}=0\cdot q+1\cdot p=p \qquad (q=1-p)$

(2) 二项分布

设随机变量 X 服从参数为(n,p)的二项分布，

即 $X\sim B(n,p)$, 它的分布律为:

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

则:

$\because k=0$ 时 $k\,p_k=0$

$$E(X)=\sum_{k=0}^{n} k\cdot c_n^k p^k q^{n-k} = \sum_{k=1}^{n} k\cdot\frac{n!}{k!(n-k)!}p^k q^{n-k}$$

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

$$=np\boxed{\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}q^{\underline{(n-1)}-\underline{(k-1)}}}$$

令: $k-1=k'$

$$=np\sum_{k'=0}^{n-1}\frac{(n-1)!}{k'![(n-1)-k']!}\cdot p^{k'}q^{(n-1)-k'}$$

$$=np\sum_{k'=0}^{n-1}C_{n-1}^{k'}p^{k'}q^{(n-1)-k'} \quad =np(p+q)^{n-1}$$

$$=np[p+(1-p)]^{n-1}=np$$

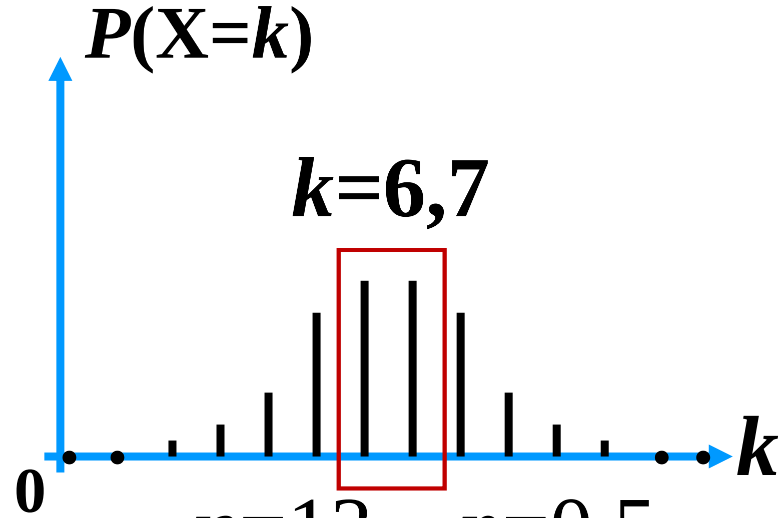
即: $E(X)=np$

二项分布

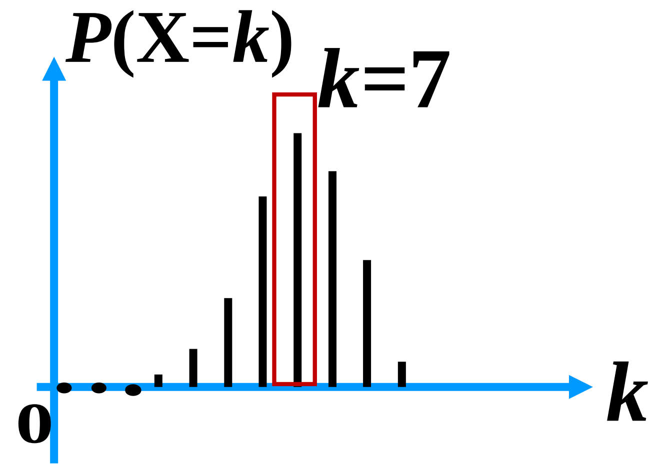
$$P_n(k)=C_n^k p^k(1-p)^{n-k}$$

$$k=0,1,2\cdots n$$

- 当$(n+1)p$为整数时，概率$P(X=k)$在$k=(n+1)p$ 和 $k=(n+1)p-1$处达到最大值；
- 当 $(n+1)p$ 不为整数时，概率 $P(X=k)$在$k=[(n+1)p]$ 达 到 最大值。



n=13， p=0.5



n=10， p=0.7

其中：$[x]$ 表示不超过 x 的最大整数

(3) 泊松分布

若随机变量X 的所有可能取值为: $0, 1, 2, \cdots\cdots$
而它的分布律(它所取值的各个概率)为:

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,2,\cdots \quad \text{即:} X \sim \pi(\lambda)$$

则: $$E(X)=\sum_{k=0}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}$$

令:
$k-1=k'$

$$=\lambda e^{-\lambda}\sum_{k'=0}^{\infty}\frac{\lambda^{k'}}{k'!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda$$

即: $E(X)=\lambda$

例2. 某银行信贷部门对前来申请贷款的两个企业进行调查，对其产品在市场上畅销、适销和滞销三种状况的盈利额和相应的概率作了如下估计：

甲企业:

产品	畅销	适销	滞销
盈利额 X_1（万元）	50	30	-20
概率	0.15	0.6	0. 25

乙企业:

产品	畅销	适销	滞销
盈利额 X_2（万元）	60	36	-40
概率	0.1	0.6	0. 3

问：当其它条件均相同时，信贷部门应先批准哪个企业的贷款更为稳妥？

解：当其它条件均相同时，应考查两个企业盈利额的平均值的情况。故分别求其数学期望：

$$E(X_1) = 50\times 0.15 + 30\times 0.6 + (-20)\times 0.25 = 20.5$$（万元）

$$E(X_2) = 60\times 0.1 + 36\times 0.6 + (-40)\times 0.3 = 15.6$$（万元）

由此可见，甲企业的经济效益高于乙企业，所以信贷部门应先批准甲企业的贷款更为稳妥。

二. 连续型随机变量的数学期望

1. 连续型随机变量数学期望的定义

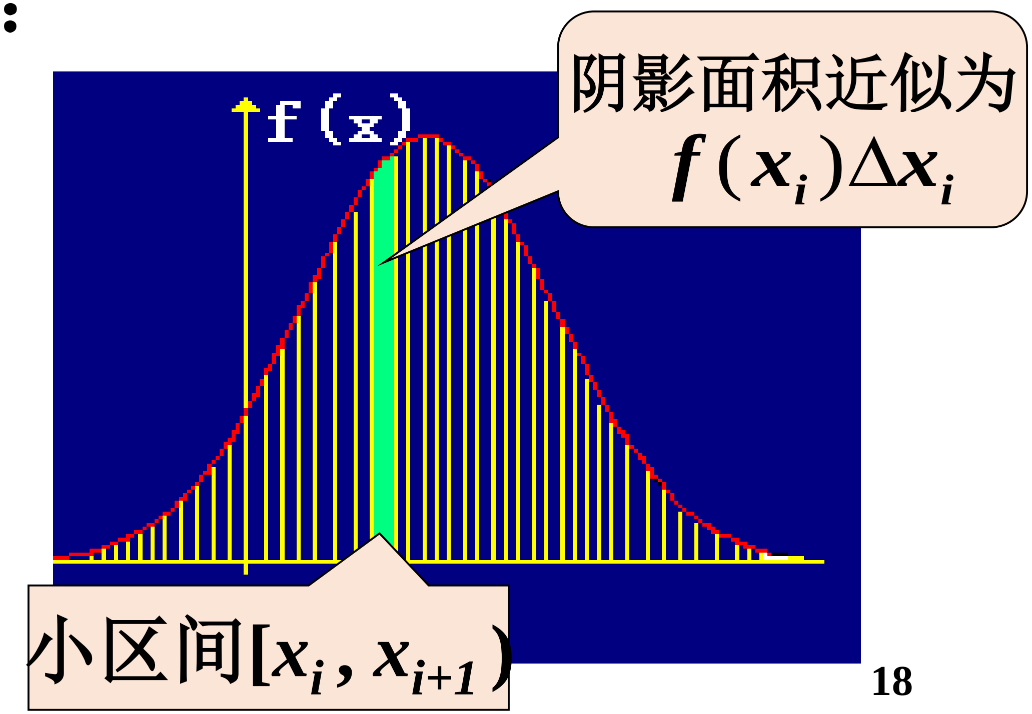
连续型随机变量的数学期望的引出

设X是连续型随机变量，其密度函数为 $f(x)$, 在数轴上取很密的分点 $x_0 < x_1 < x_2 < \dots$，则 X 落在小区间 $[x_i, x_{i+1})$ 的概率是:

$$\int_{x_i}^{x_{i+1}} f(x)dx$$

$$\approx f(x_i)(x_{i+1} - x_i)$$

$$= f(x_i)\Delta x_i$$



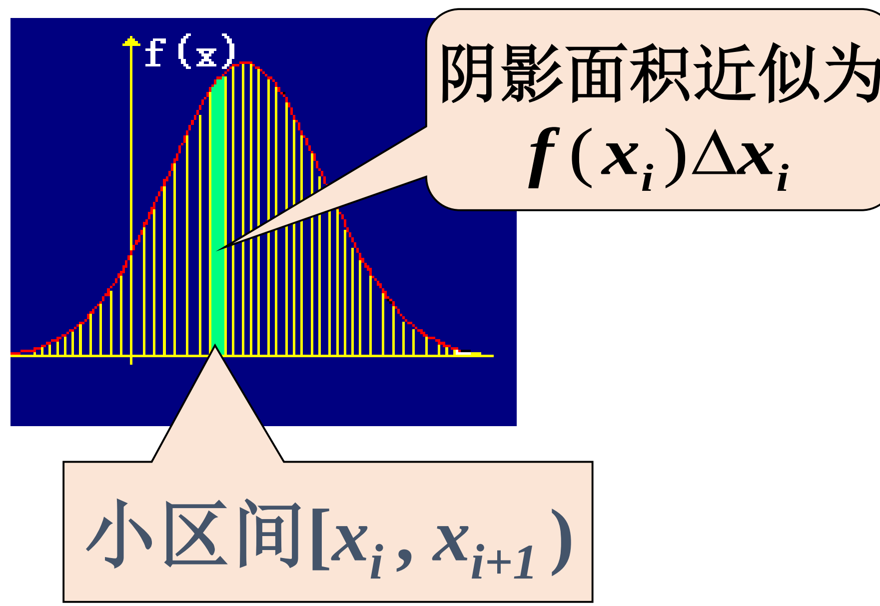
注意到：由于 x_i 与 x_{i+1} 很接近，所以区间 $[x_i, x_{i+1})$ 中的值可以用 x_i 来近似代替。

因此 X 与以概率 $f(x_i)\Delta x_i$ 取值 x_i 的离散型随机变量近似，该离散型随机变量的数学期望为：

$$\sum_i x_i f(x_i)\Delta x_i$$

这正是 $\int_{-\infty}^{\infty} x f(x)dx$

的渐近和式。



由此启发引进如下定义2：

定义2 设 X 是连续型随机变量，其概率密度函数为 $f(x)$，如果积分：

$\int_{-\infty}^{\infty} x \cdot f(x)dx$ 绝对收敛，则称此积分的值为**连续型随机变量 X 的数学期望**，记为：

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

也就是说，连续型随机变量的数学期望是一个绝对收敛的积分。

注：

▲ 推广到二维：

$$E(X)=\int_{-\infty}^{+\infty} x\cdot f_X(x)dx=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\cdot f(x,y)dydx$$

$$E(Y)=\int_{-\infty}^{+\infty} y\cdot f_Y(y)dy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y\cdot f(x,y)dxdy$$

$f(x,y)$为联合概率密度。

2. 几种常见分布的数学期望

(1). 均匀分布

若连续型随机变量 X 具有概率密度$f(x)$为：

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

即 $X \sim U[a,b]$

则： $E(X)=$

2. 几种常见分布的数学期望

(1). 均匀分布

若连续型随机变量 X 具有概率密度$f(x)$为:

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

即 $X \sim U[a,b]$

则: $E(X)=\int_{-\infty}^{+\infty} x\, f(x)dx = \int_a^b x \dfrac{1}{b-a}dx$

$$=\frac{a+b}{2}$$

即: $\boxed{E(X)=\dfrac{a+b}{2}}$

(2). 指数分布

若连续型随机变量 X 具有概率密度$f(x)$为:

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & 其它\end{cases}$$

其中$\theta > 0$为常数

则: $E(X)=\int_{-\infty}^{+\infty} x\, f(x)dx$

(2). 指数分布

若连续型随机变量 X 具有概率密度$f(x)$为:

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & 其它\end{cases}$$

其中$\theta>0$为常数

则: $$E(X)=\int_{-\infty}^{+\infty}x\,f(x)dx=\int_{0}^{+\infty}x\cdot\frac{1}{\theta}e^{-\frac{x}{\theta}}dx$$

$$=\frac{1}{\theta}\int_{0}^{+\infty}x\cdot e^{-\frac{x}{\theta}}dx=-\int_{0}^{+\infty}x\cdot d\,e^{-\frac{x}{\theta}}=\theta$$

由分部积分

即: $E(X)=\theta$

(3). 正态分布

若随机变量 X 的概率密度为:

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

即：$X \sim N(\mu,\sigma^2)$

则: $E(X)=\int_{-\infty}^{+\infty}x\,f(x)dx=\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

(3). 正态分布

若随机变量 X 的概率密度为：

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

即：$X\sim N(\mu,\sigma^2)$

则：$E(X)=\int_{-\infty}^{+\infty}x\,f(x)dx=\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

令 :y $=\dfrac{x-\mu}{\sigma}\Rightarrow$ $=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma y+\mu)e^{-\frac{y^2}{2}}dy$

$$=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}ye^{-\frac{y^2}{2}}dy+\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}dy$$

$$= 0 + \frac{\mu}{\sqrt{2\pi}}\sqrt{2\pi} = \mu$$

即: $E(X) = \mu$

结论：正态分布中密度函数的参数 μ 恰好就是随机变量X的数学期望。

三. 随机变量的函数的数学期望

1. 问题的提出:

设已知随机变量 X 的分布，且 $Y= g(X)$, 那么应该如何计算$Y= g(X)$的数学期望呢？也即如何计算随机变量的函数的数学期望?

一种方法是:

- 因为 $g(X)$也是随机变量，故应有概率分布，它的分布可以由已知的 X 的分布求出来。
- 一旦知道了$g(X)$的分布，就可以按照数学期望的定义把 $E[\,g(X)\,]$ 计算出来。

这种方法的缺点:

使用这种方法必须先求出随机变量函数 $g(X)$ 的分布，一般是比较复杂的。

问题:

是否可以不先求 $g(X)$ 的分布而只根据 X 的分布求得 $E[\,g(X)\,]$呢?

定理. 设Y是随机变量X的函数：$Y = g(X)$（g 是连续函数）则：

(1) X是**离散型随机变量**，它的分布律为：

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

若 $\sum_{k=1}^{\infty} g(x_k) p_k$ 绝对收敛，则有：

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k$$

(2) X是**连续型随机变量**，它的概率密度为 $f(x)$

若 $\int_{-\infty}^{\infty} g(x) f(x) dx$ 绝对收敛，则有：

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

证明： (略)，特殊情况的证明见教材（第五版）P96

注 ▲ 定理的意义：

给出了求随机变量函数的数学期望时，可以直接利用原来随机变量的分布，而不必先求随机变量函数的分布。

▲ 此定理可以推广到二个或二个以上随机变量的情形：

例如，设 Z 是随机变量 X，Y 的函数

$Z = g(X,Y)$, g 是连续的函数，则有：

例如，设 Z 是随机变量 X，Y 的函数

$Z = g(X, Y)$, g 是连续的函数，则有：

(1) 若(X,Y)为离散型随机变量，其联合分布律为：

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

则有：$E(Z) = E[g(X, Y)] = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i, y_j) p_{ij}$

这里设上式右边的级数绝对收敛。

(2) 若(X,Y)为连续型随机变量，其联合概率密度为

$f(x, y)$, 则有：

$$E(Z) = E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y) dxdy,$$

这里设右边的积分绝对收敛。

例4. 设二维连续型随机变量(X,Y)的联合概率密度为：

$$f(x,y)=\begin{cases}3x & 0<y<x<1\\ 0 & \text{其它}\end{cases}$$

求： $Z=XY$ 的数学期望

例4. 设二维连续型随机变量(X,Y)的联合概率密度为：

$$f(x,y)=\begin{cases}3x & 0<y<x<1\\ 0 & \text{其它}\end{cases}$$

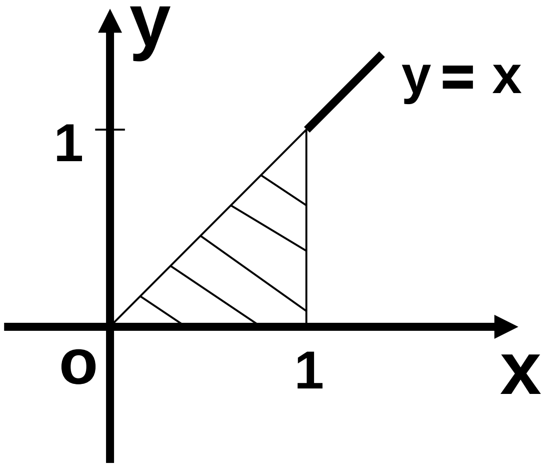
求： $Z = XY$ 的数学期望

解： $E(Z)=E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy\,f(x,y)\,dxdy$

显然，由题设 $f(x,y)\neq 0$ 的区域如图中的阴影部分

因此有：

$$E(Z)=\int_0^1 dx\int_0^x xy\cdot 3x\,dy=\frac{3}{10}$$



例5. 设国际市场对我国某出口商品的年需求量是一个随机变量 X（单位:吨），它在 [2000, 4000] 上服从均匀分布。设每售出这种商品一吨，可为国家挣外汇3万元，若售不出，则每吨需花费仓储费1万元。

问：需组织多少货源，才能使国家的收益最大?

例5. 设国际市场对我国某出口商品的年**需求量**是一个随机变量 X（单位:吨），它在 [2000, 4000] 上服从**均匀分布**。设每售出这种商品一吨，可为国家**挣外汇3万元**，若售不出，则每吨需**花费仓储费1万元**。

问： 需组织多少货源，才能使国家的收益最大?

解： 设 y：准备某年出口此种商品的量

由题设可知： $x \in [2000, 4000]$

设收益为Y，则有：

$$Y = g(X) = \begin{cases} 3y & x \geq y \\ 3x - (y - x) \cdot 1 & x < y \end{cases}$$

而随机变量 X 的概率密度为：

$$f(x)=\begin{cases}\dfrac{1}{2000} & 2000\le x\le 4000\\ 0 & \text{其它}\end{cases}$$

所以：

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

$$=\frac{1}{2000}\int_{2000}^{4000}g(x)\ dx$$

$$=\frac{1}{2000}[\int_{2000}^{y}(4x-y)\,dx+\int_{y}^{4000}3y\,dx]$$

$$=\frac{1}{1000}[-y^2+7000y-4000000]$$

注意到：E(Y)是变量 y 的函数

所以，对 E(Y)关于 y 求极值，可得：

当 $y = 3500$ 时 $E(Y)$ 达最大

结论：应组织3500(吨)货源才能使国家的收益最大。

四. 数学期望的性质

1. 设 c 是常数，则：$E(c)=c$

2. 设 c 是常数，X 是随机变量，则：

$$E(cX)=cE(X)$$

线性性质

3. X,Y是两个随机变量，则：

$$E(X+Y)=E(X)+E(Y)$$

注：这一性质可推广到任意有限个随机变量之和的情形。

4. X,Y是两个相互独立的随机变量，则：

$$E(XY)=E(X)E(Y)$$

注：这一性质可推广到任意有限个相互独立的随机变量之积的情形。

性质3、性质4的证明见教材（第五版）P100

例6. 已知 X 的概率密度 $f_X(x)=\begin{cases}\dfrac{1}{2}e^{x} & x\le 0\\ \dfrac{1}{2}e^{-x} & x>0\end{cases}$

求：$Y=|X|+b$ 的数学期望

例6. 已知 X 的概率密度 $f_X(x)=\begin{cases}\frac{1}{2}e^{x} & x\le 0\\ \frac{1}{2}e^{-x} & x>0\end{cases}$

求: $Y=|X|+b$ 的数学期望

解: $E(Y)=\int_{-\infty}^{+\infty}(|x|+b)\cdot f(x)dx$

$$=\int_{-\infty}^{+\infty}|x|f(x)dx+\int_{-\infty}^{+\infty}b\,f(x)dx$$

$$=\int_{-\infty}^{0}(-x)\frac{1}{2}e^{x}dx+\int_{0}^{+\infty}x\frac{1}{2}e^{-x}dx$$

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

例7. 设 $X_1, X_2, \cdots, X_n$ 相互独立，且均服从 $(0-1)$ 分布

求： $E(X_1 + X_2 + \cdots + X_n)$

例7. 设 $X_1, X_2, \cdots, X_n$ 相互独立，且均服从 $(0-1)$ 分布

求： $E(X_1 + X_2 + \cdots + X_n)$

解： 设：$Y = X_1 + X_2 + \cdots + X_n$,

由题意，每一个随机变量均服从（0—1）分布

即每一个随机变量 X_i 服从：$\begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$.

则由数学期望的性质有：

$$E(Y) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

而：$E(X_i) = 0 \cdot q + 1 \cdot p = p$

所以得：$E(Y) = p + p + \cdots + p = np$