Chpt.4 Digital Features of Random Variables

第四章 随机变量的数字特征

上节回顾

- 数学期望：
  - 离散型随机变量$E(X)=\sum x_k p_k$ （$\sum\limits_k |x_k|\, p_k < \infty$ ）
  - 连续型随机变量$E(X)=\int_{-\infty}^{\infty} xf(x)dx$（$\int_{-\infty}^{\infty} |x|\, f(x)dx < \infty$ ）
  - 随机变量函数 $E(Y)=E(g(X))=\int_{-\infty}^{\infty} g(x)f(x)dx$

$$E(Y)=\iint\limits_{R^n} g(x_1,\cdots,x_n)f(x_1,\cdots,x_n)dx_1\cdots dx_n$$

$$E(X_i)=\iint\limits_{R^n} x_i f(x_1,\cdots x_n)dx_1\cdots dx_n$$

思考题： 国际市场上每年对我国某种商品的需求量X服从 [2000, 4000]上的均匀分布.每售出一吨该商品可获利3万美元; 但若积压于仓库，每吨将损失1万美元. 问应组织多少货源才能使收益最大?

国际市场上每年对我国某种商品的需求量X服从 [2000, 4000]上的均匀分布.每售出一吨该商品可获利3万美元; 但若积压于仓库，每吨将损失1万美元. 问应组织多少货源才能使收益最大?

[解] 收益 a_s 与需求量X有关，也与组织的货源S有关.

a_s 是随机变量，故"收益最大"的含义是指"平均收益最大".

抽象得到需求量 $X \sim U[S_1, S_2]$;

季内售出一吨获利 b =3万美元，积压一吨损失 l =1万美元;

进货量为 S $(S_1 \le S \le S_2)$

考察收益:

$$a_s(X) = \begin{cases} bX-(S-X)l & S_1 \le X \le S \\ bS & S \le X \le S_2 \end{cases}$$

$$a_S(X)=\begin{cases}(b+l)X-Sl & S_1\le X\le S\\ bS & S\le X\le S_2\end{cases}$$

X的概率密度是 $f(x)=\begin{cases}\dfrac{1}{S_2-S_1} & S_1\le X\le S_2\\ 0 & others\end{cases}$

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

$$=\int_{S_1}^{S}\frac{1}{S_2-S_1}\left[(b+l)x-Sl\right]dx+\int_{S}^{S_2}\frac{bS}{S_2-S_1}dx$$

$$=\frac{1}{S_2-S_1}\left[\frac{1}{2}(b+l)(S^2-S_1^2)-Sl(S-S_1)\right]+\frac{bS(S_2-S)}{S_2-S_1}$$

$$=\frac{1}{S_2-S_1}\left[\frac{1}{2}(b+l)(S^2-S_1^2)-Sl(S-S_1)+bS(S_2-S)\right]$$

$$=\frac{1}{S_2-S_1}\left[-\frac{1}{2}(b+l)S^2-\frac{1}{2}(b+l)S_1^2+S(lS_1+bS_2)\right]$$

$E(a_s(X))$ 是S的函数，为了使其达到最大，S满足 $\frac{d}{ds}E(a_s(X))=0$
也就是：

$$-(b+l)S+(lS_1+bS_2)=0$$

$$S=\frac{lS_1}{b+l}+\frac{bS_2}{b+l}$$

容易知道：

(1) $S_1 \le S \le S_2$

(2）S是S_1和S_2的加权和，如果售出一件商品的利润远远高于售不出而亏损的价值，即 $b>>l$ ，则 $S\to S_2$，尽量多进货；反之，一件亏损额较大，则 $S\to S_1$ 。

上述例子中：当S＝3500时，$E(\alpha_s)$ 达到最大值825万.

4.1.3 数学期望的基本性质

尝试推导,会证明

(1) 设 C 为常数，则有 $E(C)=C$

(2) 设 X 是一个随机变量，C 为常数，则有 $E(CX)=CE(X)$

(3) 设X,Y是两个随机变量，则有 E(X+Y)=E(X)+E(Y)

设 $X1$，$X2$，…，X_n 是 n 个随机变量，$c_1,c_2,\cdots,c_n$ 为实数，则

$$E(c_1X_1+c_2X_2+\cdots+c_nX_n)=c_1E(X_1)+c_2E(X_2)+\cdots+c_nE(X_n)$$

(4) 设 X，Y 是相互独立的随机变量，则有

数学期望的线性性

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

可推广到任意有限个相互独立的随机变量的情况

$$E(\prod_{i=1}^{n}X_i)=\prod_{i=1}^{n}E(X_i),$$

其中 $X_i, i=1,2,\cdots,n,$ 相互独立.

3. $E(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)dxdy$

$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy$

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

4. $E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dxdy$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf_X(x)f_Y(y)dxdy$$

$$=\int_{-\infty}^{+\infty}xf_X(x)dx\int_{-\infty}^{+\infty}yf_Y(y)dy$$

$$=E(X)E(Y).$$

Example 设ξ是服从超几何分布的随机变量，N件产品中有M件次品，随机抽取n件， ξ为其中次品的数目

$$P(\xi = m) = C_M^m C_{N-M}^{n-m} / C_N^n \,(m = 0,1,2,\cdots,n)$$

求 Eξ.

解 不放回抽样. 令ξ_i (i=1,2,...,n)为第 i 次抽取时的废品数，则

$$\xi = \sum_{i=1}^{n} \xi_i$$

$$P(\xi_i = 1) = M / N, i = 1,2,\cdots,n$$

（由第一章古典概型知：不论是否放回，每次取得次品的概率都是M/N)

$$E\xi = \sum_{i=1}^{n} E\xi_i = nM / N$$

Example

一客车在有20位旅客从机场出发，沿途有 10 个车站可以下车。如到达一个车站没有旅客下车就不停车，以 X 表示停车次数。设每位旅客在各个车站下车时等可能的，并设旅客是否下车时相互独立的。

求：平均停车次数

解：引入随机变量X_i　　　(这样的0-1变量通常称为指示变量)

$$X_i = \begin{cases} 0, & \text{在第 } i \text{ 站没有人下车}, \\ 1, & \text{在第 } i \text{ 站有人下车}, \end{cases} \quad i = 1, 2, \cdots, 10.$$

则

$$X = X_1 + X_2 + \cdots + X_{10}.$$

$$P\{X_i = 0\} = \left(\frac{9}{10}\right)^{20}$$

$$(i = 1, 2, \cdots, 10)$$

$$P\{X_i = 1\} = 1 - \left(\frac{9}{10}\right)^{20}$$

$$E(X_i) = 1 - \left(\frac{9}{10}\right)^{20}, \quad i = 1, 2, \cdots.$$

$$E(X) = E(X_1 + X_2 + \cdots + X_{10})$$

$$= \boldsymbol{E(X_1) + E(X_2) + \cdots + E(X_{10})}$$

$$= \mathbf{10}\left[\mathbf{1} - \left(\frac{\mathbf{9}}{\mathbf{10}}\right)^{\mathbf{20}}\right]$$

$$= \mathbf{8.784}(\text{次}).$$

4.2 随机变量的方差（Variation）

4.2.1 Introduction/Definition

Example 比较两个随机试验，已知分布为：

甲X：

X	4	6	8	10	12	14	16
p	1/7	1/7	1/7	1/7	1/7	1/7	1/7

乙Y：

Y	4	5	7	10	13	15	16
p	1/7	1/7	1/7	1/7	1/7	1/7	1/7

二人的平均分：$E(X)=E(Y)=10$

从均值无法分辨优劣.还要考虑他们取值的离散程度：如同研究生入学考试，既要看总成绩，又要约定任一单科成绩.

4.2.1 Introduction/Definition

几种可能的比较方式：

（1）最大值－最小值＝**极差**

X：16－4=12，Y：16－4=12

不反映分散程度，其他值被忽略。

（2）**平均差** E(X－E(X))

E(X－E(X)) =0

X的离差正负相消.

（3）必须相加不能抵消. $E|X-E(X)| \Rightarrow E(X-EX)^2$

Definition X为随机变量，若 $E(X-E(X))^2$ 存在，就称它是随机变量X的**方差**(variance)，记作D(X)或Var(X). **反映随机变量波动性**

Remark1. 注意到D(X)的量纲与X不同

为了统一量纲，有时用$\sqrt{D(X)}$，称为X的**标准差**

(standard deviation)或均方差，记为 $\sigma(X)$.

Remark2. 前面例子中，区分两个成绩的优异

E(X)＝E(Y)＝10

D(X)＝16，D(Y)＝20

如此，我们知道甲比乙要好一些

方差的求法

$$D(X)=E(X-E(X))^2$$

取 $g(x)=[x-E(X)]^2$，则 $D(X)=E(g(X))$.

对于离散型随机变量X，其分布律为

$$P(X=x_i)=p_i,\ i=1,2,\cdots,$$

则
$$\boxed{D(X)=\sum_{i=1}^{\infty}[x_i-E(X)]^2 p_i;}$$

对于连续型随机变量X，其概率密度函数为 $f(x)$,

则
$$\boxed{D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x)dx.}$$

方差的求法

Important. **方差的等价计算公式**

$$D(X)=E(X-E(X))^2$$

$$=\int_{-\infty}^{\infty}(x-E(X))^2 f(x)dx$$

$$=\int_{-\infty}^{\infty}\left\{\, x^2-2xE(X)+\left[E(X)\right]^2\right\} f(x)dx$$

$$=\int_{-\infty}^{\infty}x^2 f(x)dx-2E(X)\int_{-\infty}^{\infty}xf(x)dx+\left[E(X)\right]^2$$

$$=E(X^2)-\left[E(X)\right]^2$$

Example 泊松分布X~π(λ)的方差.

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

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

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

$$= e^{-\lambda} \sum_{j=0}^{\infty} j \frac{\lambda^{j+1}}{j!} + \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$$

Example 计算泊松分布π(λ)的方差.

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

$$= \lambda^{2} e^{-\lambda}\sum_{j=1}^{\infty}\frac{\lambda^{j-1}}{(j-1)!} + \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^{j}}{j!}$$

$$= \lambda^{2} + \lambda$$

所以 $D(X) = E\left(X^{2}\right) - (EX)^{2} = \lambda^{2} + \lambda - \lambda^{2} = \lambda$

Example 设X服从正态分布 $N(\mu,\sigma^2)$，求D(X).

解 E(X)＝μ

$$D(X)=E(X-\mu)^2$$

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

$$\Big\Downarrow\ t=\frac{x-\mu}{\sigma}$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-t^2/2}dt$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi}=\sigma^2$$

可见正态分布中参数σ²就是它的方差, σ就是标准差.

Example 指数分布的方差:设随机变量X服从指数分布，

其概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$

其中λ>0是常数. **求：D(X)**

$$E(X^2)=\int_{-\infty}^{\infty}x^2 f(x)dx$$

$$=\int_0^{\infty}x^2\lambda e^{-\lambda x}dx$$

$$=-\int_0^{\infty}x^2 de^{-\lambda x}$$

$$=-x^2e^{-\lambda x}\Big|_0^{\infty}+\int_0^{\infty}2xe^{-\lambda x}dx$$

$$=\frac{2}{\lambda}\int_0^{\infty}x(\lambda e^{-\lambda x})dx$$

$$=\frac{2}{\lambda}E(X)$$

$$=\frac{2}{\lambda^2}$$

$$D(x)=E(X^2)-(EX)^2$$

$$=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}$$

$$=\frac{1}{\lambda^2}$$

4.2.3 方差的性质

(1) $D(C)=0;\quad (C\text{为常数})$

（2） $D(X+b)=D(X);$

$$D(aX)=a^2D(X)$$

随机变量平移X+b不影响方差；

随机变量的比例（尺度）变换aX，会比例放大方差。

4.2.3 方差的性质

(3) 设X,Y是两个随机变量，则有

$$D(X+Y)=D(X)+D(Y)+2[E(XY)-E(X)E(Y)]$$

特别的，若X,Y独立，则有

$$D(X+Y)=D(X)+D(Y)$$

推广到一般情形：

$X_1,\cdots,X_n$为随机变量，如果$X_1,\cdots,X_n$相互独立，

$c_1,\cdots,c_n$为常数， $D(c_1X_1+\cdots+c_nX_n)=\sum_{j=1}^{n}c_i^2D(X_i)$

(4) $D(X)=0$ $\Longleftrightarrow$ **以概率1取得常数C，即** $P(X=C)=1$

Example 设随机变量X具有数学期望E(X)=μ，方差D(X)= σ²，则其标准化变量 $X^* = \dfrac{X-\mu}{\sigma}\ (\sigma > 0)$ 的均值为0，方差为1。

不一定是高斯分布

$$E(X^*) = E(\frac{X-\mu}{\sigma}) = \frac{1}{\sigma}[E(X)-\mu] = 0$$

$$E(X^{*2}) = E\left(\frac{X-\mu}{\sigma}\right)^2$$

$$= \frac{1}{\sigma^2}E[(X-\mu)^2]$$

$$= 1$$

X^* 称作X的标准化向量

$$D(X) = E(X^{*2}) - \left[E(X^*)\right]^2 = 1$$

Example 设相互独立同分布随机变量 $X_1, \cdots, X_n$

$$E(X_i) = \mu, \ \ D(X_i) = \sigma^2 \ (i = 1, 2, \cdots, n) \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

求 $E(\bar{X}), D(\bar{X})$

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$$

$$D(\bar{X}) = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{\sigma^2}{n}$$

Remark 这说明在独立同分布时，作为各X_i的算术平均 $\bar{X}$，它的数学期望与各X_i的数学期望相同，**但方差只有X_i的1/ n倍**. 这一事实在数理统计中有重要意义.

切比雪夫(Chebyshev)不等式 若随机变量的方差存在，则

对任意给定的正数ε，恒有 $P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}$

Proof:

$$P\{|X-\mu|\geq\varepsilon\}=\int_{|x-u|\geq\varepsilon}f(x)dx$$

$$\leq\int_{|x-u|\geq\varepsilon}\frac{(x-\mu)^2}{\varepsilon^2}f(x)dx$$

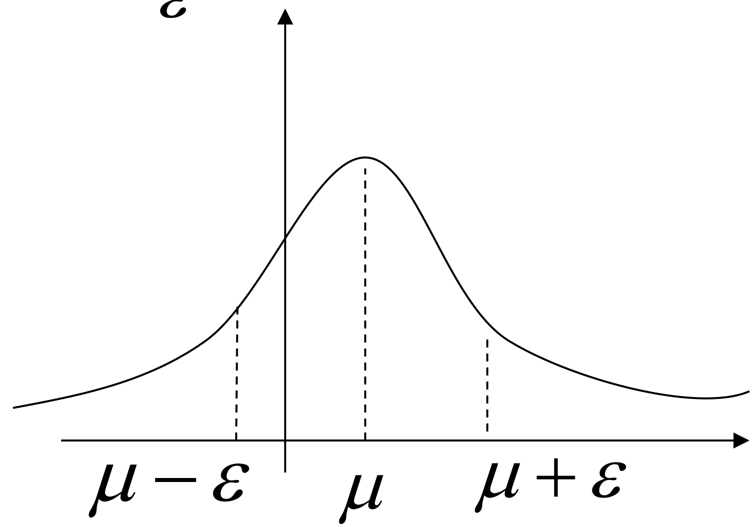
$$\leq\int_{-\infty}^{\infty}\frac{(x-\mu)^2}{\varepsilon^2}f(x)dx=\frac{\sigma^2}{\varepsilon^2}$$

上面的式子等价于

$$P\{|X-\mu|<\varepsilon\}=1-P\{|X-\mu|\geq\varepsilon\}\geq1-\frac{\sigma^2}{\varepsilon^2}$$

切贝雪夫不等式意义

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



X落在区间 $(E(X)-\varepsilon, E(X)+\varepsilon)$ 之外的概率小于 $D(X)/\varepsilon^2$，之内的概率大于 $1-D(X)/\varepsilon^2$，从而只用数学期望和方差就可对上述概率进行估计。

切贝雪夫不等式意义

这一定律可用来估计分布未知情况下事件的概率

$$\{|X-\mu|\geq\varepsilon\}\text{或}\{|X-\mu|\leq\varepsilon\}$$

车比雪夫下限：

$P\{|X-u|\geq\sigma\}\leq 1$!!

$P\{|X-u|<\sigma\}\geq 0$

$P\{|X-u|\geq 2\sigma\}\leq 0.25$

$P\{|X-u|<2\sigma\}\geq 0.75$

$P\{|X-u|\geq 3\sigma\}\leq 0.1112$

$P\{|X-u|<3\sigma\}\geq 0.8888$

切贝雪夫不等式意义

以正态分布$X \sim N(u,\sigma^2)$ 为例，已知其均值和方差μ,σ^2，

我们有3σ原则：

$$P\{|X-u|<\sigma\}=2\Phi(1)-1\approx 0.6826$$

$$P\{|X-u|<2\sigma\}=2\Phi(2)-1\approx 0.9544$$

$$P\{|X-u|<3\sigma\}=2\Phi(3)-1\approx 0.9973$$

$$P\{|X-u|<4\sigma\}=2\Phi(4)-1\approx 1$$

可与Chebyshev不等式比较；可见Chebyshev不等式的估计还是比较粗略。

车比雪夫不等式提供了一个loosen lower boundary（松弛的下限）

方差性质的证明：

(4) $D(X)=0 \Longleftrightarrow$ **以概率1取得常数C，即** $P\left(X=C\right)=1$

证明：（1）已知P(X＝C)＝1

那么，E(X)＝C

$$P\left(X=E(X)\right)=1$$

$$P\left(X^2=(E(X))^2\right)=1$$

$$E(X^2)=\left(E(X)\right)^2$$

$$D(X)=E(X^2)-\left(E(X)\right)^2=0$$

方差性质的证明：

（2）已知D(X)＝0，

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

(Chebyshev不等式)

$$P\{|X-\mu|\geq \frac{1}{n}\}\leq 0$$

$$P\{|X-\mu|\geq \frac{1}{n}\}=0$$

$$P\{|X-\mu|< \frac{1}{n}\}=1$$

拓展思考:为什么此处说P{X=C}=1，而不说X=C

$$\Downarrow$$

$$P\{|X-\mu|=0\}=1$$

令$C=\mu$, $\mathrm{P}\{\mathrm{X}=\mathrm{C}\}=1$

常见分布及其期望和方差列表

分布名称	数学期望 E(X)	方差 D(X)
0-1分布		
二项分布		
泊松分布		
均匀分布		
正态分布		
指数分布		