# §5.4大数定律

引子：在第一章统计概率一节中，我们从大量例子中知道：**事件发生的频率具有稳定性**。事实上，第二章引入的重复独立E中**事件A发生的频率稳定性**即**大量随机现象的统计规律性的典型表现**。这里应当注意两个事实:

(1)**频率稳定性表征 了事件概率的一种客观存在**;

( 2)**频率性质表征了随机事件概率性质**;

这**是概率公理化定义的实际背景！直观背景！**(n充分大时；二者之间有个近似)

但是**事件的频率稳定性是一种不确切且仅为一种直观的说法**(原因：频率具有不确定性)，怎样对频率的稳定性这一直观背景作一理论阐释呢？**初学者常把它理解为微积分中的变量与极限的关系，这是不正确的**。

由于**事件的概率只是某种平均结果**，于是我们亦可把**事件的频率稳定性处理为大量随机现象之平均结果。在概率论中，这类平均结果的稳定性Theorem，统称为大数定律**。

例1. 在分析天平上称量一质量为 $\mu$ 的物品，用$X_1,X_2,\ldots,X_n$表示n次重复测量值，

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\qquad X_1,\cdots X_n\text{为}n\text{个}r.v$$

当n愈来愈大时，$\overline{X}$ 对 μ 偏差愈小，实际上：

当 $n\to+\infty$ 时，在**一定收敛意义**下

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\to\mu$$

例2.频率稳定性：在n次重复贝努里E中$X_i$表示第i 次贝努里E成功次数

$$(i=\overline{1,n}),$$

$$\mu_n(A),\ f_n(A)$$

分别表示事件A出现的频数与频率

$$\mu_n(A)=\sum_{i=1}^{n}X_i,$$

$$f_n(A)=\frac{1}{n}\mu_n(A)=\frac{1}{n}\sum_{i=1}^{n}X_i$$

当 $n \to +\infty$ 时，在一定收敛意义下

$$f_n(A) = \frac{1}{n}\sum_{i=1}^{n} X_i \to P(A)$$

原因：单个随机现象的行为对大量随机现象共同产生的总平均效果几乎不发生影响。(许多单个随机现象具体实现—偏差相互作用—随机偏差互相抵消，补偿和拉平，致使总平均效果趋于稳定)

## § 5.4.1Tchebysheff lnequality

Theoren5.1对 $\forall$ r·vX，若方差DX存在，则对 $\forall \varepsilon > 0$ 有：

$$P(|X - EX| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} DX$$

Proof:(一)r·vX为连续型，其pdf为 p(x), dfF (x)

$$P(|X-EX|\geq \varepsilon)=\int_{|X-EX|\geq \varepsilon} dF(x)$$

$$\leq \int_{|X-EX|\geq \varepsilon} \frac{(x-EX)^2}{\varepsilon^2} dF(x)$$

$$\leq \frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-EX)^2 dF(x)=\frac{1}{\varepsilon^2}DX$$

(二)r·vX为离散型：证明时把积分号换成$\sum$

（1）等价形式：

由于$P(|X-EX|\geq\varepsilon)=1-P(|X-EX|<\varepsilon)$

$$P(|X-EX|<\varepsilon)\geq 1-\frac{1}{\varepsilon^{2}}DX$$

（2）本质：在 $r\cdot vX$ 分布未知情况下，利用 $EX$ 和 $DX$，对$r\cdot vX$ 的概率分布进行估计的一种方法。

例 对 $\forall \varepsilon = k\sqrt{DX} > 0,$

$$P\left(|X - EX| < k\sqrt{DX}\right) \geq 1 - \frac{1}{k^2 DX} \times DX = 1 - \frac{1}{k^2}$$

当$k$=3时有：

$$P\left(|X - EX| < 3\sqrt{DX}\right) \geq 1 - \frac{1}{3^2} \approx 0.8889$$

若 $X \sim N\left(\mu, \sigma^2\right)$

$$P\left(|X - \mu| < 3\sigma\right) = 0.9973$$

两式比较：切比晓夫不等式比较粗略，只利用 $r \cdot vX$的$EX, DX$信息!

## 5.4.2大数定律

Theoren5.2设在n重贝努里E中，总成功的次数为$Y_n$，而在每次中成功概率为

$p(0<p<1)$,则对$\forall \varepsilon>0$有

$$\lim_{n \to +\infty} P\left( \left|\frac{Y_n}{n} - p\right| \geq \varepsilon \right) = 0$$

or

$$\lim_{n \to +\infty} P\left( \left|\frac{Y_n}{n} - p\right| < \varepsilon \right) = 1$$

Proof: 由题设

$$Y_n \sim B(n,p) \qquad EY_n = np, DY_n = npq$$

故

$$E\frac{Y_n}{n} = p \qquad D\left(\frac{Y_n}{n}\right) = \frac{pq}{n}$$

$$\frac{Y_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad X_i \sim (0,1), i = \overline{1,n}$$

于是

$$P\left(\left|\frac{Y_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\cdot\frac{pq}{n} \xrightarrow{n\to+\infty} 0$$

故

$$\lim_{n\to+\infty} P\left(\left|\frac{Y_n}{n}-p\right|\geq\varepsilon\right)=0$$

而

故

$$P\left(\left|\frac{Y_n}{n}-p\right|\geq\varepsilon\right)=1-P\left(\left|\frac{Y_n}{n}-p\right|<\varepsilon\right)$$

$$\lim_{n\to+\infty} P\left(\left|\frac{Y_n}{n}-p\right|<\varepsilon\right)=1$$

Remark: $\frac{Y_n}{n}$ 为n重贝努里E中成功的频率 $f_n(A)$，$p$ 为每次贝努里E中成功的概率。

结论：当n充分大时，$P(|\frac{Y_n}{n}-p|\geq\varepsilon)$可小于任何预先指定的正数，即 $\frac{Y_n}{n}\xrightarrow{p}p$

频率稳定性的一种确切解释。n充分大时

$$f_n(A)\approx P$$

Theorem 5.3

设$\{X_n\}$是独立的r·v序列，若有常数c，使得 $DX_i \le C(i=1,2,\cdots)$ 对则 $\forall \varepsilon > 0$

有：

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| \ge \varepsilon\right) = 0 \quad (\mathrm{I})$$

or

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| < \varepsilon\right) = 1 \quad (\Pi)$$

Proof:
$$E(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}EX_i$$

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

利用切比晓夫等式：

$$P(|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}EX_i|\geq\varepsilon)\leq\frac{1}{\varepsilon^2}D(\frac{1}{n}\sum_{i=1}^{n}X_i)$$

$$=\frac{1}{\varepsilon^2}\times\frac{1}{n^2}\sum_{i=1}^{n} DX_i \le \frac{1}{\varepsilon^2}\cdot\frac{c}{n}\xrightarrow{n\to+\infty} 0$$

在概率论中r·v序列，满足(I)or(II)，称$\{X_n\}$服从大数定律。

推论：设$\{X_n\}$是独立同分布r·v序列，具有有限的数学期望和方差，

$$EX_i=\mu, EX_i=\sigma^2, i=1,2,\cdots$$

则对 $\forall \varepsilon > 0$ 有：

$$\lim_{n\to+\infty} P(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \varepsilon) = 0$$

or

$$\lim_{n\to+\infty} P(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| < \varepsilon) = 1$$

略

Note：贝努里大数定律是切比晓夫大数定律特例。

在推论中，DX存在，但实际上方差存在条件亦不是必要的

Theoren5.4

设$X_i$,…$X_n$,…是独立同分布的r·v序列，且具有有限的数学期望 $EX_i = \mu, (i = 1, 2, \cdots)$

则对 $\forall \varepsilon > 0$ 成立

$$\lim_{n \to +\infty} p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0$$

or

$$\lim_{n \to +\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right) = 1$$

例1.给定 $P(|X - EX| < \varepsilon) \geq 0.9, DX = 0.009$

利用切比晓夫不等式估计 ε

解：切比晓夫不等式，必须有：

$$p(|X-EX|<\varepsilon)\geq 1-\frac{DX}{\varepsilon^2}\geq 0.9$$

故 $\frac{1}{\varepsilon^2}DX\leq 0.1$ $\qquad DX=0.009$

$$\therefore \varepsilon\geq 0.3$$

例2.若DX=0.004，利用切比晓夫不等式估计概率 $P(|X-EX|<0.2)$

解；用切比晓夫不等式知

$$p(|X-EX|<0.2)\geq 1-\frac{1}{0.2^2}\cdot 0.004$$

$$=1-0.1=0.9$$

例2’ 例2”

§5.5中心极限定理

前言：在§5.1中我们讨论了独立r·v和的平均值

序列 $\frac{1}{n}\sum_{i=1}^{n}X_i$ 的**依概率收敛问题**。

下面我们将讨论**独立r·v和**$\sum_{i=1}^{n} X_i$ **的极限分布问题**。$(n \to +\infty)$ 我们知道：**正态分布在概率论和实际工程中占有特殊地位**，可否把 $\sum_{i=1}^{n} X_i$ (独立r·v和)的极限分布与之联系呢？

历史上，De Moivre-Laplace等人**首先讨论了n重贝努里E中成功次数** $Y_n = \sum_{i=1}^{n} X_i$ **的极限分布服从正态分布**——这是一个十分漂亮结果。直到19世纪**数学王子Gauss对(独立r·v和)** $\sum_{i=1}^{n} X_i$ 的**极限分布进行了全面**

**而系统的研究。**得到了**较一般结论即中心极限定理——在一定条件下断定独立r·v和的极限分布为正态分布的定理。条件：**被研究的r·v是大量独立r·v和，其中每一个别r·v对总和只起微小作用，则独立r·v和近似服从正态分布。上述问题从20世纪50年代上溯150年一直是概率学者研究的中心课题！

Theorem5.5(**独立同分布的中心极限定理**)若 r·v序列$\{X_n\}$是独立同分布的，且具有有限的数学期望与方差，

$$EX_i = \mu, DX_i = \sigma^2 (i = 1,2,\cdots)$$

则对切x有：

$$\lim_{n \to +\infty} P\left\{ \frac{1}{\sqrt{n}\sigma} (\sum_{i=1}^{n} X_i - n\mu) \le x \right\}$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

proof:略（Levy-Lindeberg定理）

当n充分大时

$$\frac{1}{\sqrt{n}\sigma}\left(\sum_{i=1}^{n} X_i - n\mu\right) = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \overset{\cdot}{\sim} N(0,1)$$

$$r \cdot v \sum_{i=1}^{n} X_i \overset{\cdot}{\sim} N(n\mu, n\sigma^2)$$

称 $\frac{1}{\sqrt{n}\sigma}(\sum_{i=1}^{n} X_i - n\mu)$ 渐近服从N(0,1)

Theorem5.6（De-Moirve-Laplace定理）

在n重贝努里E中，成功次数为$Y_n$，而在每次试验中，成功的概率为 $p(0<p<1), q=1-p$

则对一切x，有：

$$\lim_{n\to+\infty} P(\frac{Y_n - np}{\sqrt{npq}} \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

proof: $Y_n = \sum_{i=1}^{n} X_i \qquad EX_i = p, DX_i = pq(i = \overline{1,n})$

而$\{X_n\}$是独立同分布r·v序列，利用Th5.5

$$\lim_{n\to+\infty} P\left(\frac{Y_n - np}{\sqrt{npq}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

当n充分大时 $Y_n \sim N(np, npq)$

推论1.对Th5.6中贝努里E，n充分大时

$$P\{|\frac{Y_n}{n} - p| < \varepsilon\} = P\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{Y_n - np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\}$$

$$\approx 2\Phi(\varepsilon\sqrt{\frac{n}{pq}}) - 1$$

n充分大时 $P\{|\frac{Y_n}{n} - p| < \varepsilon\} = 1$

推论2.对Th5.6中贝努里E，n充分大时

$$P(a< Y_n < b) \approx \Phi(\frac{b-np}{\sqrt{npq}}) - \Phi(\frac{a-np}{\sqrt{npq}})$$

$Proof$ :由于$n$很大时,利用Th5.6有:

$$Y_n \sim N(np, npq)$$

$$P(a < Y_n < b) \approx \Phi(\frac{b-np}{\sqrt{npq}}) - \Phi(\frac{a-np}{\sqrt{npq}})$$

例1. $X_i(i=\overline{1,50})$ 是独立的r·v且同分布

$X_i \sim P(0.03)$, $Z=\sum_{i=1}^{50} X_i$, 求 $P(Z \geq 3)$

解：

$n=50$充分大时

$$EX_i=\lambda=DX_i=0.03 \quad n=50$$

$$\frac{Z-n\lambda}{\sqrt{n\lambda}} \overset{\cdot}{\sim} N(0,1)$$

$$P(Z \geq 3) = P\left\{\frac{Z - 50 \times 0.03}{\sqrt{50 \times 0.03}} \geq \frac{3 - 50 \times 0.03}{\sqrt{50 \times 0.03}}\right\}$$

$$\approx 1 - \Phi\left(\frac{3 - 1.5}{\sqrt{50 \times 0.03}}\right) = 1 - \Phi(1.225) \doteq 0.1103$$

例2 重复投掷硬币100次，设每次出现正反概率均各为1/2，问“正面次数大于50而小于60”概率各为多少？

解： $n = 100$ 重贝努里 $E, p = 1/2$

$$np = 50$$

$$\sqrt{npq} = 5$$

$$P(50 < Y_n \le 60) = \Phi\left(\frac{60-50}{5}\right) - \Phi\left(\frac{50-50}{5}\right)$$

$$= 0.9772 - 0.5$$

$$= 0.4772$$

例3 大批种子中良种占1/6，我们有99%的把握断定，在6000粒种子中良种实际所占比例与1/6差是多少？这时相应良种数落在哪个范围内？

解： $n = 6000$重贝努里$E, p = 1/6(q = \frac{5}{6})$

$Y_{6000}$ 表示良种个数

于是

$$P\left(\left|\frac{Y_{6000}}{6000}-\frac{1}{6}\right|<\varepsilon\right)\approx 2\Phi\left(\varepsilon\sqrt{\frac{6000}{6\times\frac{5}{6}}}\right)-1$$

$$\doteq 2\Phi\left(120\sqrt{3}\varepsilon\right)-1=0.99$$

$$\therefore \Phi\left(120\sqrt{3}\varepsilon\right)=0.995$$

查表 $:120\sqrt{3}\varepsilon=2.58$

$$\varepsilon=0.124$$

将$\varepsilon = 0.0124$代入前式得:

$$P\left( \left| \frac{Y_{6000}}{6000} - \frac{1}{6} \right| < 0.0124 \right)$$

$$= P\left( \left| Y_{6000} - 1000 \right| < 74.4 \right)$$

$$= P\left( 925 < Y_{6000} < 1075 \right)$$

$$= 0.99$$

若实际检验结果，良种数不在此范围内，我们有理由诊断良种数占1/6假设不正确。

习题

1　设随机变量X具有分布列

$$P(X=K)=\frac{1}{5}\qquad K=1,2,3,4,5$$

求投掷四枚匀称的硬币，以X表示出现正面的个数，求$EX$，$DX$。

$$EX=\frac{1}{5}(1+2+3+4+5)=3$$

$$EX^2=\frac{1}{5}(1+4+9+16+25)=11$$

$$E(X+2)^2=\frac{1}{5}(9+16+25+36+49)=27$$

2 设$r\cdot vX$具有分布列

| X | -1 | 0 | 1/2 | 1 | 2 |
|---|---|---|---|---|---|
| P | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{12}$ | $\frac{1}{4}$ |

求 $EX, E(1-X), EX^2$

解：
$$EX = -1\times\frac{1}{3}+0\times\frac{1}{6}+\frac{1}{2}\cdot\frac{1}{6}+\frac{1}{12}+\frac{2}{4}=\frac{1}{3}$$

$$E(1-X)=1-EX=\frac{2}{3}$$

$$EX^2=(-1)^2\times\frac{1}{3}+\frac{1}{4}\cdot\frac{1}{6}+1^2\times\frac{1}{12}+4\cdot\frac{1}{4}=\frac{35}{24}$$

3 同时投掷四枚匀称的硬币，以X表示出现正面的个数，求$EX$，$DX$.

解：（1）$X\sim B\left(4,\frac{1}{2}\right)$

故 $EX=4\times\frac{1}{2}=2, DX=4\times\frac{1}{2}\times\frac{1}{2}=1$

(2)设$X_i$表示投掷第i枚匀称的硬币出现正面的个数(i=1,2,3,4)

$$X=\sum_{i=1}^{4}X_i,\quad EX_i=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}$$

$$EX=4\times\frac{1}{2}=2$$

$\because X_1,X_2,X_3,X_4$相互独立

$$\therefore D\left(\sum_{i=1}^{4}X_i\right)=\sum_{i=1}^{4}D(X_i)=\sum_{i=1}^{4}\frac{1}{2}\times\frac{1}{2}=1$$

4 从分别标有数字0,1,2,…,8,9的十张卡片中，每次任取一张，然后放回，直到首次取到标有数字9的卡片为止，求所需抽取次数$X$的数学期望与方差。

解 $X$的分布列为

$$P(X=K)=\frac{1}{10}\cdot\left(\frac{9}{10}\right)^{K-1} \qquad K=1,2,\cdots$$

这是几何分布，其一般形式为

$$P(X=K)=p(1-p)^{K-1}, \quad K=1,2,\cdots$$

此时

$$EX=\sum_{K=1}^{\infty}Kp(1-p)^{K-1}=p\sum_{K=1}^{\infty}K(1-p)^{K-1}$$

$$=p\sum_{K=1}^{\infty}\left(x^{k}\right)'\Big|_{x=1-P}=p\left(\sum_{K=1}^{\infty}x^{k}\right)'\Big|_{x=1-p}$$

$$=p\left(\frac{x}{1-x}\right)'_{x=1-p}=p\cdot\frac{1}{(1-x)^{2}}\Bigg|_{x=1-p}=\frac{1}{p}$$

$$EX^2=\sum_{K=1}^{\infty}K^2p(1-p)^{K-1}=\sum_{K=1}^{\infty}K(K-1+1)p(1-P)^{K-1}$$

$$=\sum_{K=1}^{\infty}Kp(1-p)^{K-1}+\sum_{K=1}^{\infty}K(K-1)p(1-p)^{K-1}$$

$$=\frac{1}{p}+p(1-p)\left(\sum_{K=1}^{\infty}x^K\right)''\Bigg|_{x=1-p}$$

$$=\frac{1}{p}+p(1-p)\frac{2}{p^3}=\frac{1}{p}+2\frac{1-p}{p^2}$$

$$DX = \frac{1}{p} + \frac{2}{p^2} - \frac{2}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2}$$

对于本题$p = \frac{1}{10}$,故

$$EX = 10$$

$$DX = 90$$

5 设rv分别具有下列密度，求其期望与方差。

（1） $$p(X)=\frac{1}{2}e^{-|x|},$$

（2） $$p(X)=\begin{cases}1-|x|, & |x|\le 1\\ 0, & |x|>1\end{cases}$$

（3） $$p(x)=\begin{cases}\frac{15}{16}x^2(x-2)^2, & 0\le x\le 2\\ 0, & \text{其它}\end{cases}$$

（4）
$$p(X)=\begin{cases} x, & 0\le x<1 \\ 2-x, & 1\le x\le 2 \\ 0, & \text{其它} \end{cases}$$

解：（1） $EX=\int_{-\infty}^{+\infty}xp(x)dx=\frac{1}{2}\int_{-\infty}^{+\infty}xe^{-x}dx$

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

$$=\frac{1}{2}xe^{x}\Big|_{-\infty}^{0}-\frac{1}{2}e^{x}\Big|_{-\infty}^{0}+\frac{1}{2}x\left(-e^{-x}\right)\Big|_{0}^{+\infty}+\frac{1}{2}e^{-x}\Big|_{0}^{+\infty}$$

$$=-\frac{1}{2}+\frac{1}{2}=0$$

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

$$=\frac{1}{2}\left[x^2e^x\Big|_{-\infty}^{0}-2\int_{-\infty}^{0}xe^xdx-x^2e^{-x}\Big|_{0}^{+\infty}+2\int_{0}^{+\infty}xe^{-x}dx\right]$$

$$=-xe^x\Big|_{-\infty}^{0}+e^x\Big|_{-\infty}^{0}-xe^{-x}\Big|_{0}^{+\infty}+\left(-e^{-x}\right)\Big|_{0}^{+\infty}$$

$$=1+1$$

$$=2$$

$$\therefore DX=2$$

（2） $$EX=\int_{-\infty}^{+\infty} xp(x)dx=\int_{-1}^{1} x(1-|x|)dx=0 \qquad or$$

$$EX=\int_{-\infty}^{+\infty} xp(x)dx=\int_{-1}^{0} x(1+x)dx+\int_{0}^{1} x(1-x)dx$$

$$=\frac{x^2}{2}\Big|_{-1}^{0}+\frac{x^3}{3}\Big|_{-1}^{0}+\frac{x^2}{2}\Big|_{0}^{1}-\frac{x^3}{3}\Big|_{0}^{1}$$

$$=-\frac{1}{2}+\frac{1}{3}+\frac{1}{2}-\frac{1}{3}$$

$$=0$$

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

$$= \frac{x^3}{3}\Big|_{-1}^{0} + \frac{x^4}{4}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{0}^{1} - \frac{x^4}{4}\Big|_{0}^{1}$$

$$= \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4}$$

$$= \frac{1}{6}$$

$$\therefore DX = \frac{1}{6}$$

（3）

$$EX=\int_0^2 x\cdot\frac{15}{16}x^2(x-2)^2dx=\frac{15}{16}\int_0^2 x^3\left(x^2-4x+4\right)dx$$

$$=\frac{15}{16}\left[\frac{x^6}{6}-\frac{4}{5}x^5+\frac{4}{4}x^4\right]\Big|_0^2=\frac{15}{16}\left(\frac{64}{6}-\frac{4\times32}{5}+\frac{4\times16}{4}\right)$$

$$=\frac{15}{16}\times64\times\left(\frac{1}{6}-\frac{2}{5}+\frac{1}{4}\right)$$

$$=1$$

$$EX^2 = \int_0^2 x^2 \frac{15}{16} x^2 (x-2)^2 dx$$

$$= \frac{15}{16} \int_0^2 x^4 \left(x^2 - 4x + 4\right) dx$$

$$= \frac{8}{7}$$

$$\therefore DX = \frac{8}{7} - 1 = \frac{1}{7}$$

（4）

$$EX=\int_0^1 x^2dx+\int_1^2\left(2x-x^2\right)dx=\frac{x^3}{3}\Big|_0^1+x^2\Big|_1^2-\frac{x^3}{3}\Big|_1^2$$

$$=\frac{1}{3}+4-1-\frac{8}{3}+\frac{1}{3}=1$$

$$EX^2=\int_0^1 x^3dx+\int_1^2\left(2x^2-x^3\right)dx$$

$$=\frac{x^4}{4}\Big|_0^1+\frac{2}{3}x^3\Big|_1^2-\frac{x^4}{4}\Big|_1^2$$

$$=\frac{1}{4}+\frac{16}{3}-\frac{2}{3}-4+\frac{1}{4}=\frac{1}{6}+1$$

$$\therefore DX=\frac{1}{6}$$

5 设二维离散型 $r\cdot v(X,Y)$在点$(1,1),\left(\frac{1}{2},\frac{1}{4}\right)$

$\left(-\frac{1}{2},-\frac{1}{4}\right),(-1,-1)$ 取值的概率均为 $\frac{1}{4}$,

求$EX,EY,DX,DY,EXY.$

解：

$$EX=\sum_i\sum_j x_i p_{ij}=\sum_{i=1}^{4}x_i p_{i\cdot}=1\times\frac{1}{4}+\left(-\frac{1}{2}\right)\times\frac{1}{4}+\frac{1}{2}\times\frac{1}{4}+(-1)\times\frac{1}{4}$$

$$=0$$

| Y \ X | $-1$ | $-1/4$ | $1/4$ | $1$ | $P_{i}\cdot$ |
|---|---|---|---|---|---|
| $-1$ | 1/4 | 0 | 0 | 0 | 1/4 |
| $-1/2$ | 0 | 1/4 | 0 | 0 | 1/4 |
| $1/2$ | 0 | 0 | 1/4 | 0 | 1/4 |
| $1$ | 0 | 0 | 0 | 1/4 | 1/4 |
| $P_{\cdot j}$ | 1/4 | 1/4 | 1/4 | 1/4 | 1 |

$$EX^2 = \frac{1}{4}\left[1 + \frac{1}{4} + \frac{1}{4} + 1\right] = \frac{5}{8}$$

$$DX = \frac{5}{8}$$

$$EY = \sum_i \sum_j y_j p_{ij} = \sum_{j=1}^{4} y_j p_{\cdot j} = \frac{1}{4}\left[1 + \frac{1}{4} - \frac{1}{4} - 1\right] = 0$$

$$EY^2 = \frac{1}{4}\left[1 + \frac{1}{16} + \frac{1}{16} + 1\right] = \frac{17}{32}$$

$$DY = \frac{17}{32}$$

$$EXY = \sum_i \sum_j x_i y_i p_{ij}$$

$$= \frac{1}{4}\left[1 \times 1 + (-\frac{1}{2}) \times (-\frac{1}{4}) + (-\frac{1}{2}) \times (-\frac{1}{4}) + (-1) \times (-1)\right]$$

$$= \frac{9}{16}$$

6 将$n$只球放入$M$只盒子中，设每只球落入各个盒子是等可能的，每一个盒子可放任意多个球，求有球的盒子数$X$的数学期望。

解：

$$设X_i=\begin{cases}1, & 第i个盒中有球\\ 0, & 第i个盒中无球\end{cases} \quad i=1,2,\cdots,n$$

则 $$X = \sum_{i=1}^{M} X_i$$

$$EX = \sum_{i=1}^{M} EX_i$$

$X_i$的概率分布为

| $X_i$ | 0 | 1 |
|---|---|---|
| P | $\left(1-\frac{1}{M}\right)^n$ | $1-\left(1-\frac{1}{M}\right)^n$ |

故

$$EX_i = 1 - \left(1 - \frac{1}{M}\right)^n \qquad i = 1, 2, \cdots, M$$

从而

$$EX = \sum_{i=1}^{M} EX_i = M\left[1 - \left(1 - \frac{1}{M}\right)^n\right]$$

7　袋中有$n$张卡片，分别记有号码1，2，…,$n$从中有放回地抽出$K$张卡片来，以X表示所得号码之和，求$EX$，$DX$。

解：设第i次抽得的号码为$X_i$，i=1,2,…K，则

$$X = \sum_{i=1}^{K} X_i$$

$X_i$的概率分布为

| $X_i$ | 1 | 2 | 3 | $\cdots$ | n |
| --- | --- | --- | --- | --- | --- |
| P | $\frac{1}{n}$ | $\frac{1}{n}$ | $\frac{1}{n}$ | $\cdots$ | $\frac{1}{n}$ |

$X_i$的数学期望为

$$EX_i = \frac{1}{n}(1+2+\cdots+n) = \frac{n+1}{2}$$

$X_i^2$ 的数学期望为

$$EX_i^2=\frac{1}{n}\left(1+4+\cdots+n^2\right)=\frac{1}{n}\times\frac{n(n+1)(2n+1)}{6}$$

$$=\frac{(n+1)(2n+1)}{6}$$

故 $$DX_i=\frac{(n+1)(2n+1)}{6}-\frac{(n+1)^2}{4}=\frac{n^2-1}{12}$$

于是

$$EX = \sum_{i=1}^{K} EX_i = K \cdot \frac{n+1}{2}$$

$$DX = \sum_{i=1}^{K} DX_i = K \cdot \frac{n^2-1}{12}$$

7 同时掷$n$只骰子，求出现点数之和的数学期望和方差。

解:设$X_i$=第i颗骰子现出的点数,i=1,…,n, $X$为点数之和，则 $X=\sum_{i=1}^{n}X_i$，且$X_i$相互独立。

$X_i$的概率分布为$P(X_i=K)$=1/6 K=1,2,…,6

$$EX_i = \sum_{K=1}^{6} K \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{6\times 7}{2} = 3.5$$

$$EX_i^2 = \sum_{K=1}^{6} K^2 \cdot \frac{1}{6} = \frac{1}{6} \cdot \frac{6\times 7\times 13}{6} = \frac{91}{6}$$

$$DX_i = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

于是

$$EX = \frac{7}{2}\times n = \frac{7}{2}n, DX = \frac{35}{12}n$$

8 已经$DX$=25，$DY$=36，$\rho_{XY}=0.4$求$D(X+Y)$及$D(X\text{-}Y)$

解：
$$\operatorname{cov}(X,Y)=\rho_{XY}\sqrt{DXDY}=12$$
$$D(X+Y)=DX+DY+2\operatorname{cov}(X,Y)$$
$$=25+36+24=85$$
$$D(X-Y)=DX+DY-2\operatorname{cov}(X,Y)$$
$$=25+36-24=37$$

9 设$X$，$Y$，$Z$为三个$rv$且$EX=EY=1$，$EZ=-1$，$DX=DY=DZ=1$，$\rho_{XY}=0$，$\rho_{yz}=-\frac{1}{2}$

$\rho_{XZ}=\frac{1}{2}$，若$W=X+Y+Z$，求$EW$，$DW$。

解：

$$EW = EX + EY + EZ = 1$$

$$DW = D(X+Y+Z) = DX + DY + DZ$$

$$+2\cos(X,Y) + 2\operatorname{cov}(X,Z) + 2\operatorname{cov}(Y,Z)$$

$$= 1+1+1+2\times\frac{1}{2}\sqrt{DXDZ} - 2\times\frac{1}{2}\sqrt{DXDZ}$$

$$= 3+1-1 = 3$$

# 第五章 数字特征及极限定理总结

1、期望定义、r.v函数期望及性质；

2、方差概念及性质；

3、协方差、相关系数概念及性质；

4、大数定律及中心极限定理内容及应用。

（1）、切比晓夫不等式和大数定律（三个）

（2）、中心极限定理（两个定理及推论）