数理统计

1.3:随机变量的数字特征

所谓随机变量的数字特征, 就是用来表示随机变量某种特征的数字. 常用的数字特征包括: 数学期望, 方差, 协方差, 相关系数, 矩　等

1.离散随机变量的数学期望

定义：设离散型随机变量 X 的分布律为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
p_k	p_1	p_2	$\cdots$	p_i	$\cdots$

若级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛，则称级数 $\sum_{i=1}^{\infty} x_i p_i$ 的和为随机变量 X 的数学期望,记为 $E(X)$ 或 EX 。

即 $E(X)=\sum_{i=1}^{\infty} x_i p_i$

例 **1.**甲、乙两人进行打靶，所得分数分别记为 X_1 、X_2 它们的分布律分别为

X_1	**0**	**1**	2
p_k	**0**	**0.2**	0.8

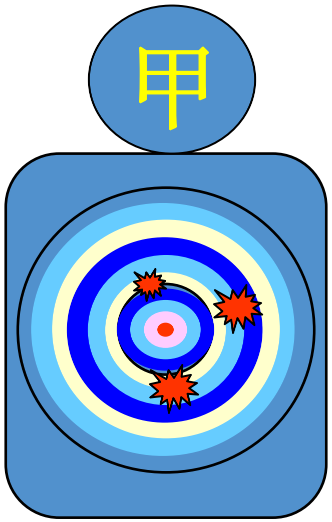
X_2	**0**	**1**	**2**
p_k	**0.6**	**0.3**	**0.1**

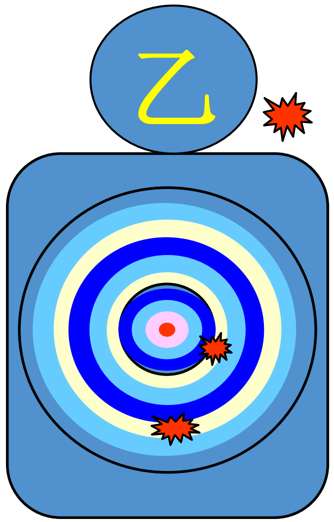
试评定他们的成绩的好坏。

解：$EX_1 = 0\times 0 + 1\times 0.2 + 2\times 0.8 = 1.8$

$EX_2 = 0\times 0.6 + 1\times 0.3 + 2\times 0.1 = 0.5$

即乙的成绩远不如甲的成绩。





2.连续随机变量的数学期望

定义：设连续型随机变量 X 的概率密度为 $\varphi(x)$，若积分 $\int_{-\infty}^{\infty} x\varphi(x)dx$ 绝对收敛（即 $\int_{-\infty}^{\infty} |x|\varphi(x)dx$ 存在），则称积分 $\int_{-\infty}^{\infty} x\varphi(x)dx$ 的值为随机变量 X 的数学期望。记为 $E(X)$ 或 EX。即

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

数学期望简称为期望，又称为均值。

例 2. 若随机变量 X 的概率密度函数为:

$$\varphi(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其它}\end{cases},$$

称X服从区间(a, b)上的均匀分布，记作$X\sim U(a,b)$,

求(a,b)上均匀分布的数学期望 EX.

解 $$EX=\int_a^b x\cdot\frac{1}{b-a}dx=\frac{1}{2(b-a)}x^2\Big|_a^b$$

$$=\frac{b^2-a^2}{2(b-a)}=\frac{a+b}{2}$$

例 **3.** 设随机变量 X 服从柯西分布，其密度函数为

$$\varphi(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < +\infty$$

求数学期望 EX 。

解：尽管 $\int_{-\infty}^{+\infty} x \cdot \frac{1}{\pi(1+x^2)} dx = 0$

因为$\int_{-\infty}^{+\infty} \frac{|x|}{1+x^2} dx$ 发散，即 $\int_{-\infty}^{+\infty} \frac{x}{1+x^2} dx$ 不绝对收敛。

$\therefore EX$ 不存在。

3.随机变量函数的期望

随机变量 Y 是 X 的函数， $Y = f(X)$

则 $$EY = \begin{cases} \sum_{k=1}^{n} f(x_k) p_k & \text{离散型} \\ \int_{-\infty}^{+\infty} f(x) p(x) dx & \text{连续型} \end{cases}$$

对多维随机变量,如（X，Y），$Z=f(X,Y)$ 也成立类似结论:

$$EZ=\begin{cases}\sum\limits_{i,j} f(x_i,y_j)p_{ij} & \text{离散型}\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)p(x,y)dxdy & \text{连续型}\end{cases}$$

数学期望的性质

性质1. $E(aX+b)=aEX+b$

性质2. $E(f(X)+g(X))=Ef(X)+Eg(X)$

性质3. 设$f(x)\le g(x)$，连续或分段连续，则

$$Ef(X)\le Eg(X)$$

现只对离散型的情形证明性质 **1**

$$E(aX+b)=\sum_i (ax_i+b)p_i=a\sum_i x_i p_i+b\sum_i p_i$$

$$=aEX+b$$

性质4. $E(a\xi+b\eta)=aE\xi+bE\eta \qquad (a,b\text{是常数})$

证明:不妨设为二维连续型随机变量

$$
\begin{aligned}
& E(a\xi+b\eta) \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(ax+by)\,p(x,y)dxdy \\
&= a\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)dxdy+b\int_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}yp(x,y)dxdy \\
&= aE\xi+bE\eta
\end{aligned}
$$

推论. $E(\sum_{i=1}^{n}a_i\xi_i)=\sum_{i=1}^{n}a_iE\xi_i$

性质5.设ξ,η相互独立，则$E\xi\eta = E\xi E\eta$

证明:仅对连续型证明,离散型同理可证

$$E\xi\eta = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp(x,y)dxdy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp_{\xi}(x)p_{\eta}(y)dxdy$$

$$= \int_{-\infty}^{+\infty} xp_{\xi}(x)dx\int_{-\infty}^{+\infty} yp_{\eta}(y)dy = E\xi\cdot E\eta$$

随机变量的数学期望表示随机变量取值的”中心”
那么如何来表示随机变量的分散程度?

4.方差的定义：称 $E(X-EX)^2$ 为随机变量 X 的方差，记为 $D(X)$，或 $\mathrm{var}(X)$。

离散：$D(X)=\sum_i (x_i-EX)^2 p_i$

连续：$D(X)=\int_{-\infty}^{+\infty}(x-EX)^2\varphi(x)dx$

其中 $\varphi(x)$ 是 X 的密度函数

说明：当随机变量的可能值密集在数学期望的附近时，方差较小；反之，方差较大。

方差的性质

性质 1)　$DX = EX^2 - (EX)^2$

证明：$DX = E(X - EX)^2$

$$= E(X^2 - 2X \cdot EX + (EX)^2)$$

$$= EX^2 - 2EX \cdot EX + (EX)^2$$

$$= EX^2 - (EX)^2$$

性质2）$D(aX+b)=a^2DX$

证明: $D(aX+b)=E[(aX+b)-E(aX+b)]^2$

$$=E[aX+b-aEX-b]^2$$

$$=E[a^2(X-EX)^2]=a^2DX$$

性质 3) $E(X-C)^2 \geq DX$,等号当且仅当 $C = EX$ 时成立.

例 4.称随机变量: $Y = \dfrac{X - EX}{\sqrt{DX}}$ 为随机变量 X 标准化.

证明: $EY = 0$, $DY = 1$.

注意到 EX, DX 都是数，根据期望和方差的性质易得

例5　设X服从两点分布，即

X	0	1
P	1-p	p

求方差 DX

解：$EX=0\times(1-p)+1\times p=p$

$$DX=EX^2-(EX)^2$$

$$=[0^2\cdot(1-p)+1^2\cdot p]-p^2$$

$$=p(1-p)$$

例 6：已知随机变量 $X \sim U(a,b)$,即密度函数为:

$$\varphi(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b \\ 0, & 其它\end{cases},$$

求随机变量 X 的方差.

解： $$EX^2=\int_a^b x^2\cdot\frac{1}{b-a}dx=\frac{1}{3(b-a)}x^3\Big|_a^b$$

$$=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+ab+b^2}{3},$$

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

所以，$DX = EX^2 - (EX)^2$

$$= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{a^2 - 2ab + b^2}{12} = \frac{(a-b)^2}{12}$$

例 7.　设 $X \sim N(\mu,\sigma^2)$, 即密度函数

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

试证明: $EX=\mu$, $DX=\sigma^2$

证明: $EX=\int_{-\infty}^{+\infty}x\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

令 $\frac{x-\mu}{\sigma}=t$，上式写成

$$EX=\int_{-\infty}^{+\infty}(\mu+\sigma t)\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\mu\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt=\mu$$

$$\text{因为}EX^2=\int_{-\infty}^{+\infty} x^2\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$\text{令}\frac{x-\mu}{\sigma}=t$$

$$EX^2=\int_{-\infty}^{+\infty}(\mu+\sigma t)^2\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\mu^2\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt+\frac{2\mu\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt+\sigma^2\int_{-\infty}^{+\infty}\frac{t^2}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$= \mu^2 - \frac{\sigma^2}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} \Bigg|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{\sigma^2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$\therefore DX = EX^2 - (EX)^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

5.切贝雪夫不等式

设随机变量ξ的$E\xi$及$D\xi$存在，则对于任何$\varepsilon > 0$，有

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}$$

证明：不妨设ξ是连续型,密度函数为$\varphi(x)$

$$P(|\xi - E\xi| \geq \varepsilon) \leq \int_{|x-E\xi|\geq\varepsilon} \frac{(x-E\xi)^2}{\varepsilon^2}\varphi(x)dx$$

$$= \frac{1}{\varepsilon^2}\int_{|x-E\xi|\geq\varepsilon}(x-E\xi)^2\varphi(x)dx$$

$$\leq \frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-E\xi)^2\varphi(x)dx$$

$$= \frac{D\xi}{\varepsilon^2}$$

例 8: 随机变量ξ方差为**0,**当且仅当ξ服从单点分布，

即 $P\{\xi = E\xi\} = 1.$

证：单点分布方差为 **0** 显然.

反之, 对任$\varepsilon > 0$，有

$$P\{|\xi - E\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2} = 0.$$

即$P\{\xi = E\xi\} = 1.$

6. 协方差与相关系数

称矩阵$\begin{pmatrix} E(\xi-E\xi)^2 & E(\xi-E\xi)(\eta-E\eta) \\ E(\eta-E\eta)(\xi-E\xi) & E(\eta-E\eta)^2 \end{pmatrix}$为随机向量$(\xi,\eta)$的协方差矩阵

称$E(\xi-E\xi)(\eta-E\eta)$为(ξ,η)的协方差，记为 $\mathrm{cov}(\xi,\eta)$

称$\dfrac{\mathrm{cov}(\xi,\eta)}{\sqrt{\mathrm{D}\xi}\sqrt{\mathrm{D}\eta}}$为 ξ与η 的相关系数,记为 ρ(或 r)

即: $\mathrm{cov}(\xi,\eta) = E(\xi-E\xi)(\eta-E\eta)$

$$\rho = \frac{\mathrm{cov}(\xi,\eta)}{\sqrt{\mathrm{D}\xi}\sqrt{\mathrm{D}\eta}}$$

协方差与相关系数有如下性质

性质 1 $\operatorname{cov}(\xi,\eta)=E\xi\eta-(E\xi)\cdot(E\eta)$.

性质 2 $\operatorname{cov}(\xi,\eta)=\operatorname{cov}(\eta,\xi)$.

性质 3 $\operatorname{cov}(a\xi+b\eta,\ \zeta)=a\operatorname{cov}(\xi,\zeta)+b\operatorname{cov}(\eta,\zeta)$.

性质 4 若ξ,η独立，则 $\operatorname{cov}(\xi,\eta)=0$.

性质 5 $D(\xi\pm\eta)=D\xi+D\eta\pm2\operatorname{cov}(\xi,\eta)$.

性质 6. 若ρ为随机变量ξ与η的相关系数,则 $|\rho| \leq 1$

且$|\rho|=1$的充要条件是ξ与η有线性关系:

即存在a, b使得 $P\{\eta = a\xi + b\} = 1$

只证明性质5:

$$D(\xi-\eta)=E[(\xi-\eta)-E(\xi-\eta)]^2$$

$$=E[(\xi-E\xi)\text{-}(\eta\text{-}E\eta)]$$

$$=E[(\xi-E\xi)^2\text{-}2(\xi-E\xi)(\eta\text{-}E\eta)+(\eta\text{-}E\eta)^2]$$

$$=E[(\xi-E\xi)^2\text{-}2\mathrm{E}(\xi-E\xi)(\eta\text{-}E\eta)+\mathrm{E}(\eta\text{-}E\eta)^2$$

$$=D\xi-2\operatorname{cov}(\xi,\eta)+D\eta$$

同理可证: $D(\xi+\eta)=D\xi+2\operatorname{cov}(\xi,\eta)+D\eta$

推论: 若ξ与η独立, 则 $D(\xi\pm\eta)=D\xi+D\eta$

更一般地: 若ξ_i独立, 则 $D(\xi_1+\xi_2+...+\xi_n)=D\xi_1+D\xi_2+...+D\xi_n$

例 9. 设n重贝努里试验(独立重复试验)中

每次试验事件A发生的概率为p,

X为n次试验中事件A发生的次数.试证明:

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

2) EX=np; DX=np(1-p)

证明: 1) X=k表示n次试验中事件A 恰好发生了k 次,

事件A可能发生在第1~k 次，也可能在第2~k+1 次,...，总共有C_n^k 种情况

对每一种情况,比如: 事件A 可能发生在第1~k 次而第2~n 次A 不发生

每种情况出现的概率显然都为: $p\cdots p\cdot(1-p)\cdots(1-p)=p^k(1-p)^{n-k}$

因此有: $P\{X=k\}=C_n^k p^k (1-p)^{n-k}$ (k=0,1,2,...,n)

2) 令 $X_i=\begin{cases}1 & \text{第i次试验A发生}\\0 & \text{否则}\end{cases}$, i=1,2,...,n

则X_i为两点分布. X_i相互独立. $\mathrm{E}X_i = p,\ DX_i = p(1-p)$

$$\mathrm{X} = X_1 + X_2 + \ldots + X_n$$

故: $E\,\mathrm{X} = E(X_1 + X_2 + \ldots + X_n) = EX_1 + EX_2 + \ldots + EX_n = np$

$$\mathrm{D\,X} = D(X_1 + X_2 + \ldots + X_n) = DX_1 + DX_2 + \ldots + DX_n = np(1-p)$$

本例中X的分布称为二项分布,记为 X~B(n,p)

7. 定义：若随机变量ξ与η的相关系数为0，

称ξ与η不相关.

定理：对随机变量ξ与η下列命题是等价的。

（1）ξ与η不相关

（2）$Cov(\xi,\eta)=0$

(3) $E\xi\eta=E\xi\, E\eta$

(4) $D(\xi+\eta)=D\xi+D\eta$

显然,若随机变量ξ与η相互独立,则ξ与η不相关.

若随机变量ξ与η相互独立,则ξ与η不相关.

反之一般不成立,即不相关与独立是不等价的

但是，当$(\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$时,可以证明:

$$\rho_{\xi\eta}=\frac{\operatorname{cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=\frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2}=\rho$$

而当 $\rho=0$时,有联合密度函数 $p(x,y)=p_\xi(x)p_\eta(y)$,即ξ与η独立

故：二维正态分布独立与不相关等价

总结我们前面曾经提到常见分布：

离散型：

1)两点分布　EX=p;　DX=np(1-p)

2)二项分布　$X \sim B(n, p)$　$\mathrm{EX} = \mathrm{np};$

$\mathrm{DX} = \mathrm{np}(1-\mathrm{p})$

3)泊松分布　$X \sim P(\lambda)$　$\mathrm{EX} = \lambda;$　$\mathrm{DX} = \lambda$

(根据期望和方差的定义证明)

连续型：

4) 均匀分布 $X \sim U(a,b)$　　$\mathrm{EX} = \tfrac{a+b}{2}$;　$\mathrm{DX} = \dfrac{(b-a)^2}{12}$

5) 指数分布 $X \sim E(\lambda)$　　$\mathrm{EX} = \tfrac{1}{\lambda}$;　$\mathrm{DX} = \dfrac{1}{\lambda^2}$ (证明略)

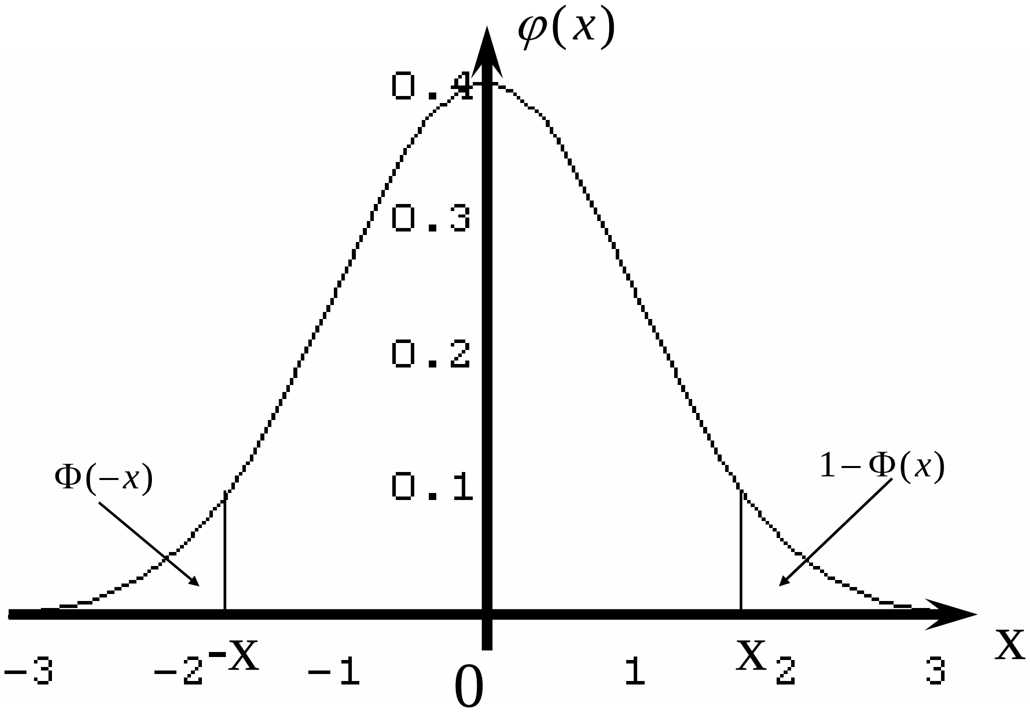
6) 正态分布 $X \sim N(\mu,\sigma^2)$　　$\mathrm{EX} = \mu$;　$\mathrm{DX} = \sigma^2$

$X \sim N(\mu,\sigma^2)$特别当$\mu=0,\sigma^2=1$时称X服从标准正态分布

$X \sim N(0,1)$ 的密度函数为:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

如图:



X 的分布函数记为$\Phi(x)$,

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

$\Phi(x)$ 的值可查表

$\Phi(x)$的性质:

(1) $\Phi(0)=0.5$, **(2)** $\Phi(+\infty)=1$, **(3)** $\Phi(-x)=1-\Phi(x)$.

在 1.2 节中我们已经证明:

若随机变量 $X \sim N(\mu,\sigma^2)$,则 $\dfrac{X-\mu}{\sigma} \sim N(0.1)$

这里 $\dfrac{X-\mu}{\sigma}$ 就是随机变量X的标准化

有关正态分布的概率计算问题,一般是先标准化

例 10. $X \sim N(1,4)$，求 $P(0 < X \le 1.6)$

解： $P(0 < X \le 1.6) = \Phi(\dfrac{1.6-1}{2}) - \Phi(\dfrac{0-1}{2})$

$= \Phi(0.3) - \Phi(-0.5) \quad = \Phi(0.3) - [1 - \Phi(0.5)]$

$= 0.6179 - 1 + 0.6915$

$\mathbf{= 0.3094}$

1.4 随机变量序列的极限定理

一、中心极限定理

概率论中，有关论证一定条件下随机变量累加和的极限分布是正态分布的那些定理统称为 中心极限定理

最常用的条件是：　独立同分布

中心极限定理的结论：

$\sum_{i=1}^{n} \xi_i$ 近 似 服 从 正 态 分 布 N (期 望 , 方 差)

因涉及正态分布的概率计算往往需要标准化：

$$\frac{\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\mathrm{E}\xi_i}{\sqrt{\mathrm{D}(\sum_{i=1}^{n}\xi_i)}} \sim N(0,1)$$

二、林德贝格-勒维 中心极限定理

设独立随机变量 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$，服从相同分且

$E\xi_i = \mu \quad D\xi_i = \sigma^2 > 0 \quad i = 1,2,\cdots,n,\cdots$，

则对任意实数 x，有：

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n}\xi_i - n\mu}{\sqrt{n}\sigma} \le x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

即：当 n 充分大时，$\dfrac{\sum_{i=1}^{n}\xi_i - n\mu}{\sqrt{n}\sigma}$ 近似服从 $N(0,1)$

或：$\sum_{i=1}^{n}\xi_i \sim N(n\mu, n\sigma^2)$

例 1： 计算机进行加法计算时，把每个加数取为最接近于它的整数来计算。设所有的取整误差是相互独立的随机变量，并且都在区间**[-0.5,0.5]**上服从均匀分布，求 **300** 个数相加时误差总和的绝对值小于 **10** 的概率。

解： 设 ξ_i 表示第 i 个加数的取整误差，则 ξ_i 在区间**[-0.5,0.5]**上服从均匀分布，并且有

$$E\xi_i = \frac{-0.5+0.5}{2} = 0,$$

$$D\xi_i = \frac{[0.5-(-0.5)]^2}{12} = \frac{1}{12},$$

由林德贝格-列维中心极限定理得：

$$P\left(\left|\sum_{i=1}^{n}\xi_i\right|<10\right)=P\left(\frac{\left|\sum_{i=1}^{n}\xi_i-0\right|}{\sqrt{300\times\frac{1}{12}}}<\frac{10-0}{\sqrt{300\times\frac{1}{12}}}\right)$$

$$=P\left(\frac{\left|\sum_{i=1}^{n}\xi_i-0\right|}{\sqrt{300\times\frac{1}{12}}}<2\right)$$

$$=\Phi(2)-\Phi(-2)=2\Phi(2)-1$$

$$=2\times0.9772-1=0.9544$$

答：300个数相加时误差总和的绝对值小于10的概率为0.9544。

三、大数定律

概率论中所有论述一定条件下随机变量的平均结果具有稳定性的一系列定理都叫大数定理(大数定律).

"随机变量的平均结果"是指: $\dfrac{\sum_{i=1}^{n}\xi_i}{n}$

"稳定性" 是指 "随机变量的平均结果” 在某个数附近波动（有 ‘近似’ 的含义，严格描述稳定是“以概率收敛”）

问题：随机变量的平均结果 $\dfrac{\sum_{i=1}^{n}\xi_i}{n}$ 应该稳定在哪个数附近？

大数定理的一般表达式：

$$\lim_{n\to\infty} \mathrm{P}\{|\frac{\sum_{i=1}^{n}\xi_i}{n} - \frac{\sum_{i=1}^{n}E\xi_i}{n}| < \varepsilon\} = 1 \qquad (\forall \varepsilon > 0)$$

大数定理的条件：

- 马尔科夫条件
- 切比雪夫条件
- 辛钦条件（独立,同分布,且期望存在）
- 泊松条件

大数定理的应用：

- 概率的统计定义的理论依据
- 统计中参数矩法估计的理论依据