概率论知识复习

二、依概率收敛及极限理论

依概率收敛 设 $X_1,\cdots,X_n,\cdots$ 是随机变量序列，如果存在常数 c，使得对 $\forall \varepsilon > 0$，总有 $\lim\limits_{n\to\infty} P(|X_n - c| < \varepsilon) = 1$，那么就称序列 $\{X_1,\cdots,X_n\cdots\}$ 依概率收敛于 c，记作 $X_n \xrightarrow{\text{Pr.}} c$。

或等价表示为 $\lim\limits_{n\to\infty} P(|X_n - c| \geq \varepsilon) = 0.$

大数定律　设$X_1,\cdots,X_n,\cdots$是独立同分布的随机变量序列,

并且 $E(X_1)=\mu, D(X_1)=\sigma^2$, 则 $\bar{X}\triangleq\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{\text{Pr.}} \mu.$

因为 $E(\bar{X})=\mu$, 所以上式也可写成 $\bar{X}\xrightarrow{\text{Pr.}} E(\bar{X}).$

或等价表示为 $\lim_{n\to\infty}P(|\bar{X}-\mu|\geq\varepsilon)=0.$

例　频率的稳定性　在 n 次重复独立试验中，记随机变量

$$X_i = \begin{cases} 1, & \text{事件}A\text{在第}\,i\,\text{次试验中发生}, \\ 0, & \text{事件}A\text{在第}\,i\,\text{次试验中不发生}. \end{cases} \quad i = 1, \cdots, n.$$

那么在 n 次重复独立试验中，随机事件 A 发生的频率为

$f_n(A) \triangleq \dfrac{N_A}{n} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$，其中 N_A 表示事件 A 发生的次数.

记 $P(A) = p$，则 $X_i \sim B(1, p)$ 且 $E(\dfrac{1}{n}\sum_{i=1}^{n} X_i) = E[f_n(A)] = p$

由大数定律即知 $\dfrac{N_A}{n} \xrightarrow{P} p$，即频率依概率收敛到概率.

频率的稳定性可以用下面的贝努利大数定律来表示:

贝努利大数定律

设 $X_1,\cdots,X_n,\cdots \overset{i.i.d}{\sim} B(1,p)$, 则 $\bar{X} \xrightarrow{\text{Pr.}} p$.

中心极限定理　设独立同分布随机变量序列 $X_1,\cdots,X_n,\cdots$ 满足 $E(X_i)=\mu, D(X_i)=\sigma^2\neq 0$, 则对 $\forall x\in(-\infty,+\infty)$, 有

$$\lim_{n\to\infty}P(\{\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n\sigma^2}}\leq x\})=\Phi(x).$$

定理表明当 n充分大时, $Y=\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n\sigma^2}}$ 近似服从标准正态分布 $N(0,1)$. **或者** $Z=\sum\limits_{i=1}^{n}X_i$ 近似服从 $N(n\mu,n\sigma^2)$.

或 $\sqrt{n}\dfrac{\bar{X}-\mu}{\sigma}\xrightarrow{n\to\infty}N(0,1).$

例　利用中心极限定理计算：当掷一枚均匀的铜币时，需要投掷多少次才能保证使得正面出现的频率在0.4至0.6之间的概率不小于90%?

解　记 $X_i=\begin{cases}1, & 第\, i\, 次抛掷出现正面,\\ 0, & 第\, i\, 次抛掷出现反面.\end{cases}\quad i=1,2,\cdots,n.$

则 $X_1,\cdots,X_n$ 独立同分布，且 $X_i\sim B(1,0.5)$.

而 $\bar{X}\triangleq\frac{1}{n}\sum_{i=1}^{n}X_i$ 即表示抛掷 n 次后正面向上发生的频率.

计算知 $E(\sum_{i=1}^{n}X_i)=0.5n, D(\sum_{i=1}^{n}X_i)=0.25n$,

由中心极限定理 $\sum_{i=1}^{n} X_i \dot{\sim} N(0.5n, 0.25n)$,

由题意知 $P(0.4 < \bar{X} < 0.6) \geq 0.9$,

即 $0.9 \leq P(0.4 < \dfrac{\sum_{i=1}^{n} X_i}{n} < 0.6) = P(0.4n < \sum_{i=1}^{n} X_i < 0.6n)$

$$\approx \Phi(\frac{0.6n - 0.5n}{\sqrt{0.25n}}) - \Phi(\frac{0.4n - 0.5n}{\sqrt{0.25n}})$$

$$= \Phi(0.2\sqrt{n}) - \Phi(-0.2\sqrt{n}) = 2\Phi(0.2\sqrt{n}) - 1,$$

故由 $2\Phi(0.2\sqrt{n}) - 1 \geq 0.9 \Rightarrow \Phi(0.2\sqrt{n}) \geq 0.95$

$\Rightarrow 0.2\sqrt{n} \geq u_{0.95} = 1.645$, 解得 $n \geq 67.65$, 故取 $n = 68$.