第三章 随机变量的数字特征

刘 春 光

暨南大学数学系

2018年4月

目录

主要内容

1. 数学期望：随机变量的平均取值
2. 方差：随机变量取值偏离平均值的程度
3. 协方差、相关系数：两个随机变量之间的线性相关程度
4. 大数定律：大量随机现象平均结果的稳定性

目录
数学期望
函数的期望
期望的性质
方差与标准差
常用分布的数字特征
矩
协方差与相关系数
大数定律

目录

离散随机变量的数学期望

概念引入：某服装公司生产两种套装，一种是大众装，每件价格200元，每月生产1万件；另一种是高档装，每件1800元，每月生产100件。现在问该公司生产的套装平均价格是多少？

离散随机变量的数学期望

Definition (数学期望(mathematical expectation))

设离散型随机变量X的分布律为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots .$$

若级数

$$\sum_k x_k p_k$$

绝对收敛，则称其和为随机变量X的数学期望，记为EX。

离散随机变量的数学期望

例1 0-1分布

X	0	1
P	$1-p$	p

的期望为

$$EX = p.$$

离散随机变量的数学期望

例2 袋中装有两个白球与三个黑球，每次从袋中任取一个球，直至取得白球为止，求以下两种取法下取球次数的期望：

1. 每次取出的球不再放回；
2. 每次取球观察颜色后放回袋中。

离散随机变量的数学期望

总结：几何分布

$$P\{X=k\}=pq^{k-1},\quad k=1,2,\cdots$$

（其中$p,q>0,p+q=1$）的数学期望为

$$EX=\frac{1}{p}.$$

离散随机变量的数学期望

练习：随机变量X的概率分布为

X	-1	0	1	2
P	0.2	0.1	0.3	0.4

求X的期望。

连续随机变量的数学期望

Definition

设连续型随机变量X的概率密度为$f(x)$，若积分

$$\int_{-\infty}^{+\infty} xf(x)dx$$

绝对收敛，则称此积分为随机变量X的数学期望，记为EX。

连续随机变量的数学期望

例3 向半径为R的圆形靶射击，击中点M落在以靶心O为中心、r为半径的圆内的概率与该圆的面积成正比，并且不会出现脱靶的情况。用X表示击中点M与靶心O的距离，求X的数学期望。

连续随机变量的数学期望

例4 柯西分布的概率密度为

$$f(x) = \frac{1}{\pi(1+x^2)},$$

该分布的数学期望不存在。

连续随机变量的数学期望

练习：已知随机变量X的概率密度为

$$\varphi(x) := \begin{cases} a + bx^2 & 0 < x < 1 \\ 0 & \text{else.} \end{cases}$$

且$EX = \dfrac{3}{5}$，求未知系数a, b。

补充：数学期望的一般定义

Definition

设$(\Omega, \mathscr{F}, P)$为概率空间，$X = X(\omega)$为Ω上随机变量。如果$X(\omega)$在Ω上Lebesgue可积，则积分

$$\int_{\Omega} X(\omega) P(d\omega)$$

称为X的数学期望，记为EX。

补充：数学期望的一般定义

Theorem (积分变换定理)

设$(\Omega, \mathscr{F}, P)$为概率空间，$X = X(\omega)$为Ω上随机变量，有分布函数$F(x)$。则对任意可测函数$g(x)$，有

$$\int_{\mathbb{R}} g(x)dF(x) = \int_{\Omega} g[X(\omega)]P(d\omega).$$

上式在如下意义下成立：如果等式一端积分有限，则另一端也有限，且二者相等。

补充：数学期望的一般定义

说明：前述定理中的积分

$$\int_{\mathbb{R}} g(x)dF(x)$$

为g对F的Lebesgue-Stieltjes积分。当g为连续函数时，该积分等价于g对F的Riemann-Stieltjes积分。进一步若F绝对连续，上述积分等价于：

$$\int_{\mathbb{R}} g(x)F'(x)dx.$$

二维随机变量的数学期望

设二维离散型随机变量(X,Y)的分布为

$$P\{X=x_i,Y=y_j\}=p_{ij},$$

$$i=1,2,\cdots,j=1,2,\cdots,$$

则

$$EX=\sum_i\sum_j x_i p_{ij},$$

$$EY=\sum_i\sum_j y_j p_{ij}.$$

二维随机变量的数学期望

设二维连续型随机变量(X, Y)的联合概率密度为$f(x, y)$，则

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

二维随机变量的数学期望

练习：求以下联合分布中随机变量X, Y的期望：

$X \backslash Y$	0	1	2	3	4
0	0	0	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{1}{42}$
1	0	$\frac{1}{14}$	$\frac{2}{7}$	$\frac{1}{7}$	0
2	$\frac{1}{70}$	$\frac{1}{7}$	$\frac{1}{7}$	0	0
3	$\frac{1}{105}$	$\frac{1}{42}$	0	0	0

二维随机变量的数学期望

例8 设二维随机变量(X,Y)服从G上的均匀分布，其中G为抛物线$y=x^2$与直线$y=x+2$所围成的区域，求EX及EY。

目录

随机变量函数的期望

由积分变换定理立刻得到如下性质：

Theorem

若随机变量X有分布函数$F(x)$。则对任意可测函数$g(x)$，随机变量$Y=g(X)$的期望为

$$EY=E[g(X)]=\int_{\mathbb{R}}g(x)dF(x).$$

离散随机变量的函数的期望

若X为离散型随机变量，分布律为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots .$$

则随机变量$Y = g(X)$的期望为

$$EY = \sum_k g(x_k) p_k.$$

离散随机变量的函数的期望

例1 设随机变量X的分布为

X	-2	-1	0	1	2	3
P	0.1	0.2	0.25	0.2	0.15	0.1

求随机变量$Y=X^2$的数学期望。

连续随机变量的函数的期望

若X为连续型随机变量，概率密度为$f(x)$，则随机变量$Y=g(X)$的期望为

$$EY=\int_{-\infty}^{+\infty}g(x)f(x)dx.$$

例2 设随机变量X在区间$[0,\pi]$上服从均匀分布，求随机变量$Y=\sin X$的期望。

二维随机变量的函数的期望

设(X,Y)为随机向量，$Z=g(X,Y)$，则

1. 若(X,Y)为离散型随机变量，分布律为

$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots.$$

则

$$EZ=\sum_i\sum_j g(x_i,y_j)p_{ij}.$$

二维随机变量的函数的期望

2. 若(X,Y)为连续型随机变量，概率密度为$f(x,y)$，则

$$EZ=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy.$$

二维随机变量的函数的期望

练习：设二维离散型随机向量(X, Y)的概率分布如下表所示，求$Z = X^2 + Y$的期望。

$X \backslash Y$	1	2
1	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{8}$

二维随机变量的函数的期望

练习：假设(ξ,η)的联合分布为

(ξ,η)	$(0,0)$	$(0,2)$	$(1,1)$	$(1,2)$	$(2,1)$
P	0.15	0.25	0.1	0.2	0.3

求$2\xi-\eta$的期望。

二维随机变量的函数的期望

例3 设二维随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\frac{8}{\pi(x^2+y^2+1)^3}, & x\geq 0, y\geq 0,\\ 0, & \text{else},\end{cases}$$

求随机变量$Z=X^2+Y^2$的期望。

二维随机变量的函数的期望

练习：设随机变量X和Y相互独立，概率密度分别为

$$f_X(x) = \begin{cases} 4e^{-4x}, & x \geq 0, \\ 0, & \text{else}, \end{cases}$$

$$f_Y(y) = \begin{cases} 2e^{-2y}, & y \geq 0, \\ 0, & \text{else}, \end{cases}$$

求$E(XY)$。

目录

数学期望的性质

设$X, X_1, X_2, \cdots, X_n$为随机变量，c, k 为常数，则有

1. $E(c) = c$;
2. $E(kX) = kE(X)$;
3. $E(X_1 + X_2) = E(X_1) + E(X_2)$;
   推论：$E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$

数学期望的性质

例：设在某试验中事件A的概率为p，将该试验独立地进行n次。记X为n次试验中事件A发生的总次数，X_i为第i次试验中事件A发生的次数，则

$$X \sim B(n,p),\ X_i \sim B(1,p), i = 1,2,\cdots,n, \text{且}$$

$$X = X_1 + X_2 + \cdots + X_n$$

故

$$EX = EX_1 + EX_2 + \cdots + EX_n = np.$$

数学期望的性质

练习：从一副标准扑克牌（52张，不包含大小王）中任意抽取10张，求其中红桃张数的期望、脸牌（J,Q,K）张数的期望。（提示：利用抽签的公平性）

数学期望的性质

4. 若X_1, X_2相互独立，则有

$$E(X_1X_2) = E(X_1)E(X_2).$$

推论：若$X_1, X_2, \cdots, X_n$相互独立，则有

$$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k)$$

注：如果没有相互独立这一条件，上式一般不成立！！！

数学期望的性质

练习：设二维离散型随机向量(X,Y)的概率分布如下表所示：

$X \backslash Y$	1	2
1	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{8}$

求$E(X), E(Y)$和$E(XY)$。

目录

方差

概念引入：设甲、乙两台仪器测量某零件长度（单位：cm）的分布律如下：

甲：

X	9.8	9.9	10	10.1	10.2
P	0.2	0.2	0.2	0.2	0.2

乙：

Y	9.6	9.8	10	10.2	10.4
P	0.2	0.2	0.2	0.2	0.2

试判断两台仪器的测量精度孰优孰劣？

方差

概念引入：设有随机变量

$$X \sim U[-1,1], \quad Y \sim U[-100,100],$$

则

$$EX = EY = 0,$$

但两个随机变量取值的集中程度却有显著的不同。

方差

在实际问题中，仅靠期望值不能完善地说明随机变量的分布特征。我们常常需要知道分布相对于期望值的离散程度。

定义：若随机变量X的期望存在，称

$$X - EX$$

为随机变量X的离差。

方差

Definition (方差(variance)、标准差(standard deviation))

随机变量X的离差的平方的期望称为X的方差，记为$D(X)$（或$Var(X)$），即

$$D(X) := E[X - E(X)]^2.$$

称

$$\sigma(X) := \sqrt{D(X)}$$

为X的标准差。

方差

方差刻划了随机变量的取值对于其数学期望的偏离程度。

1. 若X 的取值比较分散，则方差较大；
2. 若X 的取值比较集中，则方差较小；特别地，$D(X)=0$当且仅当X取某个常数的概率为1。

方差

方差的常用计算公式：$D(X) = E(X^2) - [EX]^2$。

例1 0-1分布

X	0	1
P	$1-p$	p

的期望和方差分别为

$$EX = p, \qquad D(X) = p(1-p).$$

方差

例2 袋中装有两个白球与三个黑球，每次从袋中任取一个球，直至取得白球为止，求以下两种取法下取球次数的方差：

1. 每次取出的球不再放回；
2. 每次取球观察颜色后放回袋中。

方差

总结：几何分布

$$P\{X=k\}=pq^{k-1},\quad k=1,2,\cdots$$

（其中$p,q>0,p+q=1$）的数学期望和方差分别为

$$EX=\frac{1}{p},\quad DX=\frac{1}{p^2}-\frac{1}{p}.$$

方差

练习：随机变量X的概率分布为

X	-1	0	1	2
P	0.2	0.1	0.3	0.4

求$D(X)$。

方差

例3 向半径为R的圆形靶射击，击中点M落在以靶心O为中心、r为半径的圆内的概率与该圆的面积成正比，并且不会出现脱靶的情况。用X表示击中点M与靶心O的距离，求X的方差。

方差

练习：设连续型随机变量X的密度为

$$f(x) = \begin{cases} 2x, & x \in [0, 1], \\ 0, & \text{otherwise}, \end{cases}$$

求$D(X)$。

方差

练习：设(X, Y)的联合概率密度为

$$f(x, y) = \begin{cases} 3x, & x \in [0, 1], y \in [0, x), \\ 0, & \text{otherwise} \end{cases},$$

求$E(Y)$和$D(Y)$。

方差

练习：假设随机变量ξ的概率密度为

$$\varphi(x) = \begin{cases} ax^2 + bx + c, & x \in (0,1), \\ 0, & \text{otherwise} \end{cases},$$

已知$E\xi = 0.5$, $D(\xi) = 0.15$，求a, b, c。

方差的性质

设$X, X_1, X_2, \cdots, X_n$为随机变量，c, k 为常数，则有

1. $D(c) = 0$， $D(X + c) = D(X)$；
2. $D(X) \geq 0$，且等式成立当且仅当X 几乎必然为常数；
3. $D(kX) = k^2 D(X)$；

注：若事件A的概率为1，则称该事件几乎必然成立(happens almost surely (abbreviated as a.s.))。

方差的性质

例4 设随机变量X的期望和方差分别为$E(X)$和$D(X)$，且$D(X) > 0$，求

$$Y = \frac{X - E(X)}{\sqrt{D(X)}}$$

的期望和方差。

称$Y = \dfrac{X - E(X)}{\sqrt{D(X)}}$为$X$的标准化的随机变量。

方差的性质

4. 若X_1, X_2相互独立，则有

$$D(X_1 + X_2) = D(X_1) + D(X_2).$$

推论：若$X_1, X_2, \cdots, X_n$相互独立，则有

$$D\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} D(X_k)$$

注：如果没有相互独立这一条件，上式一般不成立！！！

方差的性质

练习：一台设备由三个部件构成，在设备运转中各部件需要调整的概率分别为$0.01, 0.02, 0.03$。设各部件的状态相互独立，用X表示同时需要调整的部件数，求X的期望和方差。

方差的性质

练习：设二维离散型随机向量(X,Y)的概率分布如下表所示：

$X \backslash Y$	1	2
1	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{8}$

求$D(X),D(Y)$和$D(X+Y)$。

目录

超几何分布的期望与方差

如果随机变量$X \sim H(n, M, N)$，即

$$P\{X = k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n},$$

则

$$EX = \frac{nM}{N}, \quad D(X) = \frac{nM}{N} \cdot \frac{(N-M)(N-n)}{N(N-1)}.$$

二项分布的期望与方差

如果随机变量$X \sim B(n,p)$，即

$$P\{X=k\} = b(k;n,p) = C_n^k p^k (1-p)^{n-k},$$

则

$$EX = np, \quad D(X) = np(1-p).$$

二项分布的期望与方差

练习：已知随机变量X服从二项分布$B(n,p)$，

$$EX = 12, \quad D(X) = 8,$$

求n和p。

泊松分布的期望与方差

例：如果随机变量$X \sim P(\lambda)$, $\lambda > 0$，即

$$P\{X = m\} = \frac{\lambda^m}{m!}e^{-\lambda}, \quad m = 0, 1, \cdots,$$

则

$$EX = \lambda, \qquad D(X) = \lambda.$$

均匀分布的期望与方差

例：设随机变量$X \sim U[a,b]$，即其概率密度为

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b], \\ 0, & \text{otherwise}, \end{cases} \qquad (a < b)$$

则

$$EX = \frac{a+b}{2}, \qquad D(X) = \frac{1}{12}(b-a)^2.$$

指数分布的期望与方差

例：设随机变量X服从参数为λ的指数分布，即其概率密度为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

则

$$EX = \frac{1}{\lambda}, \qquad D(X) = \frac{1}{\lambda^2}.$$

目录

原点矩与中心矩

Definition (原点矩(raw moments)、中心矩(central monents))

对随机变量X与正整数k，

1. 称$\nu_k = E(X^k)$为X的k阶原点矩，
2. 称$\mu_k = E[(X - EX)^k]$为X的k阶中心矩。

原点矩与中心矩

Example

对任意的随机变量X,

1. X的一阶原点矩即其期望EX;
2. X的一阶中心矩总是为零;
3. X的二阶中心矩即其方差DX。

原点矩与中心矩

Example

中心矩可用原点矩表示，事实上，对任意$k \geq 2$

$$\mu_k = E[(X-\nu_1)^k] = \sum_{i=0}^{k} C_k^i (-\nu_1)^i \nu_{k-i}.$$

如

$$\mu_2 = \nu_2 - n_1^2,$$

$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3,$$

$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4.$$

原点矩与中心矩

例 设随机变量X服从参数为λ的指数分布，求X的任意阶原点矩及三、四阶中心矩。

目录

协方差与相关系数

对于二维随机向量(X, Y)，除了其分量X和Y的期望与方差外，还有一些数字特征，用以刻画X与Y之间的相关程度，其中最主要的就是下面要讨论的协方差和相关系数。

协方差

Definition (协方差(Covariance))

对于二元随机变量(X, Y)，称

$$\mathrm{cov}(X, Y) := E[(X - EX)(Y - EY)]$$

为X与Y的协方差。

由定义直接可得：任意随机变量与其自身的协方差就是该随机变量的方差，即

$$\mathrm{cov}(X, X) = D(X).$$

协方差的性质

设X, X_1, X_2, Y为随机变量，a, b, c, d为常数，则有

1. 对称性：$\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$；
2. $\mathrm{cov}(aX + b, cY + d) = ac \cdot \mathrm{cov}(X, Y)$；
3. $\mathrm{cov}(X_1 + X_2, Y) = \mathrm{cov}(X_1, Y) + \mathrm{cov}(X_2, Y)$；
4. $\mathrm{cov}(X, Y) = E(XY) - EX \cdot EY$；

   推论：两随机变量相互独立，则协方差等于零；反之未必成立。

协方差

例1 设二维离散型随机向量(X,Y)的概率分布如下表所示：

$X \backslash Y$	-1	0	1
0	0	$\frac{1}{3}$	0
1	$\frac{1}{3}$	0	$\frac{1}{3}$

求$\mathrm{cov}(X,Y)$。

协方差

练习：假设二维随机变量(X, Y)的联合分布为

$X \backslash Y$	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

求$P\{X = 2Y\}$及$\mathrm{cov}(X - Y, Y)$。

协方差的性质

5. $D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y)$;
6. Cauchy-Schwarz不等式：

$$[\text{cov}(X, Y)]^2 \leq D(X) \cdot D(Y),$$

且等号成立当且仅当X与Y之间有线性关系，即存在常数a, b使得

$$Y = aX + b \text{（或} X = aY + b\text{）}, \quad \text{a.s.}$$

相关系数

Definition (相关系数(Correlation))

对于二元随机变量(X, Y)，如果两个变量的方差都不为零，称

$$\frac{\text{cov}(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$$

为X与Y的相关系数，记为$R(X, Y)$，$\text{corr}(X, Y)$，或ρ_{XY}。

Definition (线性无关(Linear independence))

若$R(X, Y) = 0$，则称X与Y线性无关，或线性不相关，简称不相关。

相关系数

练习：已知随机变量X和Y的方差分别为1和4，相关系数为-0.5。
求$D(X+Y)$和$D(X-Y)$。

相关系数的性质

设X, Y为随机变量，则有

1. $|R(X,Y)| \leq 1$，且
   $R(X,Y) = 1$当且仅当存在常数$a > 0, b$使得$Y = aX + b$，a.s.;
   $R(X,Y) = -1$当且仅当存在常数$a < 0, b$使得$Y = aX + b$，a.s..
2. 若X与Y相互独立，则$R(X,Y) = 0$；反之未必成立。

相关系数

练习：设(X,Y)服从$D=\{(x,y):x^2+y^2\le 1\}$上的均匀分布，即概率密度函数为

$$f(x,y)=\begin{cases}\frac{1}{\pi} & x^2+y^2\le 1\\ 0 & \text{else}\end{cases}$$

求$R(X,Y)$。

相关系数

相关系数$R(X,Y)=0$意味着X与Y不存在任何线性关系，但它们仍然有可能存在其它函数关系，如从如下联合分布

$X \backslash Y$	-1	0	1
0	0	$\frac{1}{3}$	0
1	$\frac{1}{3}$	0	$\frac{1}{3}$

可以得到$R(X,Y)=0$及函数关系$X=Y^2$。

协方差矩阵

Definition (协方差矩阵(Covariance matrix))

对二维随机向量(X, Y)，称矩阵

$$\begin{pmatrix} \operatorname{cov}(X, X) & \operatorname{cov}(X, Y) \\ \\ \operatorname{cov}(Y, X) & \operatorname{cov}(Y, Y) \end{pmatrix}$$

为(X, Y)的协方差矩阵。

协方差矩阵

Definition (协方差矩阵(Covariance matrix))

设$\vec{X} = (X_1, X_2, \cdots, X_n)$为$n$维随机向量，称矩阵

$$\boldsymbol{B} = (\mathrm{cov}(X_i, X_j))_{n\times n}$$

为$\vec{X}$的协方差矩阵。

性质：协方差阵为对称的半正定矩阵。

目录

Chebyshev不等式

Theorem (Chebyshev不等式)

设随机变量X有期望和方差，则对于任给的$\varepsilon > 0$，有

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

即

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

Chebyshev不等式

练习：设随机变量X的期望和方差分别为μ和σ^2，估计以下概率

$$P(|X-\mu|<\sigma),$$

$$P(|X-\mu|<2\sigma),$$

$$P(|X-\mu|<3\sigma).$$

大数定律

Definition

设$\{X_\lambda\}_{\lambda\in\Lambda}$是随机变量的集合，如果该集合的任意有限子集中的随机变量相互独立，则称该集合$\{X_\lambda\}_{\lambda\in\Lambda}$相互独立。

大数定律

Theorem (切比雪夫大数定律(Chebyshev's law))

设随机变量序列$\{X_n\}_{n\geq 1}$相互独立，且方差有界，即存在常数$M>0$使得

$$D(X_n) \leq M, \quad n = 1, 2, \cdots .$$

则对任意正数ε，

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

大数定律

Definition (依概率收敛(Convergence in probability))

若对于任给的$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{|\xi_n - \xi| < \varepsilon\} = 1,$$

则称随机变量序列$\{\xi_n\}$依概率收敛于ξ，记作

$$\xi_n \xrightarrow{P} \xi.$$

大数定律

Corollary (辛钦大数定律(Khintchine's law))

设有独立同分布的随机变量序列$\{X_n\}_{n\geq 1}$，其数学期望及方差均存在：

$$EX_n = \mu, \ \ D(X_n) = \sigma^2, \quad n = 1, 2, \cdots .$$

则

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

实际意义：多次试验求平均值能够有效地逼近期望。

大数定律

Theorem (伯努利大数定律(Bernoulli's law))

在独立实验序列中，记事件A发生的概率为p。以$f_n(A)$表示前n次试验中事件A发生的次数，则

$$\frac{f_n(A)}{n} \xrightarrow{P} p.$$

实际意义：当重复试验次数充分大时，频率是概率的有效近似。

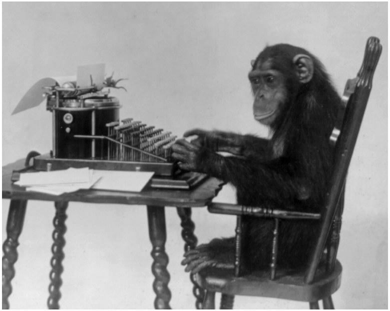
小概率原理

小概率原理(Little probability event principle)：一个事件的发生概率很小，那么

- 它在一次试验中是几乎不可能发生的，
- 但在无限次重复试验中是几乎必然发生的。

小概率原理

无限猴子定理(Infinite monkey theorem)：一只猴子随机在打字机键盘上按键，在无穷久的时间里几乎必然打出一套莎士比亚全集。



★补充：强大数定律

Definition (几乎必然收敛(Almost sure convergence))

如果有

$$P\left\{\omega \in \Omega \middle| \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right\} = 1,$$

则称随机变量序列$\{\xi_n\}$几乎必然收敛于ξ，记作

$$\xi_n \xrightarrow{\text{a.s.}} \xi.$$

★补充：强大数定律

Theorem

随机变量序列$\{\xi_n\}$几乎必然收敛于ξ的充分必要条件是对任意的$\varepsilon > 0$，都有

$$\lim_{n\to\infty} P\left\{\bigcup_{k=n}^{\infty}[|\xi_k - \xi| \geq \varepsilon]\right\} = 0,$$

Corollary

若$\xi_n \xrightarrow{a.s.} \xi$，则$\xi_n \xrightarrow{P} \xi$。

★补充：强大数定理

Example ($\xi_n \xrightarrow{P} \xi$时未必有$\xi_n \xrightarrow{\text{a.s.}} \xi$)

取区间$[0,1)$上的几何概率空间，即$\Omega=[0,1)$，

$\mathscr{F}$是$[0,1)$内的Borel集类，P为Lebsgue测度。

令

$$\eta_{m,k} = \begin{cases} 1, & \omega \in [\frac{k-1}{m}, \frac{k}{m}), \\ 0, & \text{else}, \end{cases} \quad \begin{pmatrix} k=1,\cdots,m, \\ m=1,2,\cdots. \end{pmatrix}$$

令$\xi_1=\eta_{1,1}$，$\xi_2=\eta_{2,1}$，$\xi_3=\eta_{2,2}$，$\xi_4=\eta_{3,1}$，$\cdots$，

则$\xi_n \xrightarrow{P} 0$，但$\xi_n \xrightarrow{\text{a.s.}} 0$ 不成立。

★补充：强大数定律

Definition (大数定律(Law of large numbers)、强大数定律(Strong law of large numbers))

设$\{\xi_n\}$为随机变量序列，每个ξ_n都有有限的数学期望。

如果

$$\frac{1}{n}\sum_{k=1}^{n}[\xi_k - E\xi_k] \xrightarrow{P} 0,$$

则称$\{\xi_n\}$满足大数定律。如果进一步有

$$\frac{1}{n}\sum_{k=1}^{n}[\xi_k - E\xi_k] \xrightarrow{\text{a.s.}} 0,$$

则称$\{\xi_n\}$满足强大数定律。

★补充：强大数定律

Theorem (Borel强大数定律)

在独立实验序列中，记事件A发生的概率为p。以$f_n(A)$表示前n次试验中事件A发生的次数，则

$$\frac{f_n(A)}{n} \xrightarrow{a.s.} p.$$