

随机变量及其分布 01

1、随机变量的概念：（随机事件的结果展现出来的数学形式）

定义在样本空间 Ω 上的实值函数 $X = X(\omega)$ 称为随机变量，常用大写字母 X, Y, Z 等表示随机变量，其取值用小写字母 x, y, z 等表示。

假如一个随机变量仅可能取有限个或可列个值，则称其为**离散型随机变量**。假如一个随机变量的可能取值充满数轴上的一个区间 (a, b) ，则称其为**连续型随机变量**，其中 a 可以是 $-\infty$ ， b 也可以是 $+\infty$

2、随机变量的分布函数：设 X 为一个随机变量，对任意实数 x ， $F(x) = P(X \leq x)$ 为 X 的分布函数，且称 X 服从 $F(x)$ ，记为 $X \sim F(x)$ 或 $F_X(x)$

分布函数的三个性质：**单调性，有界性，右连续**，这三个基本性质成为判别某个函数是否能成为分布函数的充要条件。

对任意实数 $a < b$ ，有下列公式：

$$(1) P(a < x \leq b) = F(b) - F(a)$$

推导： $P(a < x \leq b) = P((x \leq b) - (x \leq a)) = P(x \leq b) - P(x \leq a) = F(b) - F(a)$

$$(2) P(X = a) = F(a) - F(a - 0)$$

推导： $P(X = a) = P(X \leq a) - P(X < a) = F(a) - F(a - 0)$

$$(3) P(X \geq b) = 1 - F(b - 0)$$

推导： $P(X \geq b) = 1 - P(X < b) = 1 - F(b - 0)$

$$(4) P(X > b) = 1 - F(b)$$

推导: $P(X > b) = 1 - P(X \leq b) = 1 - F(b)$

$$(5) P(a < x < b) = F(b - 0) - F(a)$$

推导: $P(a < x < b) = P(x < b) - p(x \leq a) = F(b - 0) - F(a)$

$$(6) P(a \leq x \leq b) = F(b) - F(a - 0)$$

推导: $P(a \leq x \leq b) = P(x \leq b) - P(x < a) = F(b) - F(a - 0)$

$$(7) P(a \leq x < b) = F(b - 0) - F(a - 0)$$
 特别的, 如果函数在 a, b 处

连续时, 有 $F(a - 0) = F(a), F(b - 0) = F(b)$

3、离散型随机变量的概率分布列

设离散型随机变量 X 的可能取值为 $x_1, x_2, \dots, x_n, \dots$ 加入 X 取 x_i 的概率为 $p_i = P(X = x_i), i = 1, 2, \dots$ 为 X 的概率分布列或简称为分布列, 记为 $X \sim \{P_i\}$, 分布列为:

X	x_1	x_2	\dots	x_n	\dots
P	p_1	p_2	\dots	p_n	\dots

4、分布列的性质: (1) 非负性 $p(x_i) \geq 0$ (2) 正则性 $\sum_{i=1}^n p(x_i) = 1$

由离散随机变量 X 的分布列可以得出 X 的分布函数为 $F(x) =$

$$\sum_{x_i \leq x} p(x_i)$$

比如现在有1, 2, 3, 4, 5这些点, 分别对应概率 p_1, p_2, p_3, p_4, p_5 ,

$$\text{则 } F(3) = p_1 + p_2 + p_3$$

5、求离散型随机变量的分布列应该注意:

(1) 确定随机变量的所有可能取值

(2) 计算每个取值点的概率

求离散型随机变量的分布列函数应该注意:

- (1) $F(x)$ 是递增的阶梯函数
- (2) 其间断点均为右连续的
- (3) 其间断点即为 X 的可能取值点
- (4) 其间断点的跳跃高度是对应的概率值

例如：
$$F(x) = \begin{cases} 0, & x < 0 \\ 0.4, & 0 \leq x < 1 \\ 0.8, & 1 \leq x < 2 \\ 1, & 2 \leq x \end{cases}$$

X	0	1	2
P	0.4	0.4	0.2

6、连续型随机变量的概率分布列

定义：设随机变量 X 的分布函数为 $F(x)$ ，如果存在实数轴上的一个非负可积函数 $p(x)$ ，使得对任意实数 x 有： $F(x) = \int_{-\infty}^x p(t) dt$ ，则称 $p(x)$ 为 X 的**概率密度函数**，简称密度函数或称密度。

密度函数的基本性质：(1) 非负性 $p(x) \geq 0$ (2) 正则性

$$\int_{-\infty}^{+\infty} p(x) dx = 1 \text{ (含有 } p(x) \text{ 的可积性)}$$

以上两条基本性质是确定或判别某个函数是否成为密度函数的充要条件。

- (1) $P(a \leq X \leq b) = \int_a^b p(x) dx$
- (2) $F(x)$ 是 $(-\infty, +\infty)$ 上的连续函数
- (3) $P(X = x) = F(x) - F(x - 0) = 0$
- (4) $F(x + \Delta x) - F(x) = \int_x^{x+\Delta x} p(x) dx \rightarrow 0 \quad (\Delta x \rightarrow 0)$
- (5) $P(X = a) = 0$ ，连续型随机变量在某一点的概率为 0
- (6) $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$

(7) 密度函数 $X \sim p(x)$ 不唯一

7、数学期望的定义

(离散型) 设离散型随机变量 X 的分布列为: $p(x_i) = P(X = x_i), i = 1, 2, \dots, n, \dots$ 如果 $\sum_{i=1}^{\infty} |x_i| p(x_i) < \infty$, 则称 $E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$ 为随机变量 X 的数学期望, 或称为该分布的数学期望, 简称期望或均值, 若 $\sum_{i=1}^{\infty} |x_i| p(x_i)$ 不收敛, 则 X 的数学期望不存在

(连续性) 设连续型随机变量 X 的密度函数为 $p(x)$, 如果 $\int_{-\infty}^{+\infty} |x| p(x) dx < \infty$, 则称 $E(X) = \int_{-\infty}^{+\infty} x p(x) dx$ 为随机变量 X 的数学期望, 或称为该分布 $p(x)$ 的数学期望, 简称期望或均值, 若 $\int_{-\infty}^{+\infty} |x| p(x) dx$ 不收敛, 则 X 的数学期望不存在

8、数学期望的性质

设 $Y = g(x)$ 是随机变量 X 的函数, 若 $E(g(X))$ 存在, 则

$$E(g(x)) = \begin{cases} \sum_{i=1}^{\infty} g(x_i) P(X = x_i), & \text{在离散场合} \\ \int_{-\infty}^{+\infty} g(x) p(x) dx, & \text{在连续场合} \end{cases}$$

性质 1: 若 C 是常数, 则 $E(c) = c$

性质 2: 对任意数 a , 有 $E(aX) = aE(X)$

性质 3: 对任意两个函数 $g_1(x)$ 和 $g_2(x)$, 有 $E[g_1(x) \pm g_2(x)] = E(g_1(x)) \pm E(g_2(x))$

9、方差与标准差的定义 (方差反应的稳定程度)

定义: 若随机变量 X^2 的数学期望 $E(X^2)$ 存在, 则称偏差平方

$(X - EX)^2$ 的数学期望为随机变量 X (或相应分布)的方差, 记为

$var(X) = E(X - E(X))^2$ 称方差的正平方根 $\sqrt{Var(X)}$ 为随机变量 X (或相应分布)的标准差, 记为 $\sigma(X)$ 或 σ_X

10、方差的性质

性质 1: $var(X) = E(X - E(X))^2 = E(X^2 - 2 * X * E(X) + E(X)^2) = E(X^2) - 2E(X) * E(X) + E(X)^2 = E(X^2) - [E(X)]^2$

求方差常用此性质, 先求出 $E(X)$, 再求 $E(X^2)$

性质 2: 常数 c 的方差为 0

性质 3: 对任意常数 a 与 b 和随机变量 X , 有 $Var(aX + b) = a^2 Var(X)$

随机变量的标准化: 设 $Var(X) > 0$, 令 $Y = \frac{X - E(X)}{\sqrt{Var(X)}}$, 则 $E(Y) =$

$0, Var(Y) = 1$, 称 Y 为 X 的标准化

11、切比雪夫不等式

定理 1: 对任一随机变量 X , 若 X 的数学期望和方差都存在, 则对任一正数 ε , 恒有 $P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$ 或 $P(|X - E(X)| < \varepsilon) \geq 1 - \frac{Var(X)}{\varepsilon^2}$

定理 2: 方差为 0 的随机变量 X 必几乎处处为常数。这个常数就是其期望 $E(X)$, 这个定理亦可表示为:

若 $Var(X) = 0$, 则 $P(X = E(X)) = 1$ 或 $P(X = a) = 1$

定理 2 的意思就是如果方差为 0, 那么随机变量 X 为常数列

12、常用的离散分布

(一) 二项分布

(1) 伯努利试验: 如果试验只有两个结果成功(A)和失败(\bar{A}), 且每次试验成功的概率是不变的, 则称这种试验为伯努利实验, 将一个伯

努利试验独立重复的进行 n 次，称为 n 重伯努利试验。

在每次试验中， $P(A) = P(0 < p < 1)$ ，则在 n 重伯努利试验中事件 A 出现 k 次的概率为

$$P_n(k) = C_n^k p^k (1-p)^{n-k} (k = 0, 1, 2, \dots, n)$$

若随机变量 X 的概率分布为 $P(X = k) = C_n^k p^k (1-p)^{n-k} (k = 0, 1, 2, \dots, n)$ ，则称随机变量 X 服从参数为 (n, p) 的二项分布，并记
 $X \sim B(n, p)$

二项分布的数学期望 $E(X) = np$ ， $Var(X) = np(1-p)$

(2) 两点分布（二项分布的特殊形式）

当 $n=1$ 时，二项分布为 $B(1, p)$ 称为两点分布，或称0-1分布，或称伯努利分布，其分布列为 $P(X=x) = p^x (1-p)^{1-x}, x = 0$ 或 1

二点分布的数学期望 $E(X) = p$ ， $Var(X) = p(1-p)$

(二) 泊松分布

若随机变量 X 的概率分布为 $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda} (k = 0, 1, 2, \dots)$ ，其中参数 $\lambda > 0$ ，则称随机变量 X 服从参数为 λ 的泊松分布，并记 $X \sim P(\lambda)$ ，泊松分布常与单位时间(或单位面积、单位产品等)的计算过程相联系。

泊松分布的数学期望和方差均为 λ

(三) 二项分布的泊松近似

在二项分布 $B(n, p)$ 中，当 n 较大时，计算量是令人烦恼的，而 p 较小时使用泊松定理，可以减少二项分布中的计算量。

定理(泊松定理)在 n 重伯努利实验中，记事件 A 在一次试验中发生的概率为 P_n ，(与试验次数 n 有关)，如果当 $n \rightarrow \infty$ 时，有 $np_n \rightarrow \lambda$ ，则

$$\lim_{n \rightarrow \infty} C_n^k p_n^k (1-p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

(四) 超几何分布

从一个有限总体中进行不放回抽样常会遇到超几何分布

设有 N 件产品，其中有 M 件不合格品，若从中不放回的随机抽取 n

件，则其中含有的不合格品的件数 X 服从超几何分布，记为 $X \sim h$

(n, N, M) 超几何分布的概率分布列为：

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0, 1, \dots, r$$

其中 $r = \min\{M, n\}$ ，且 $M \leq N, n \leq N, n, N, M$ 均为正整数

超几何分布的数学期望 $E(X) = n \frac{M}{N}$

方差 $Var(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)}$

(五) 超几何分布的二项近似

当 $n \ll N$ 时，即抽取个数 n 远小于产品个数 N 时，每次抽取后，总体中的不合格品率 $p = M/N$ 改变甚微，所以不放回抽样可近似地看成放回抽样，这时超几何分布可用二项分布近似。

$$\frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = C_n^k p^k (1-p)^{n-k}, \text{ 其中 } p = \frac{M}{N}$$

(六) 几何分布

在伯努利试验序列中，记每次试验中事件 A 发生的概率为 p ，如果 X 为事件 A 首次出现时的试验次数，则 X 的可能取值为 $1, 2, \dots$ ，称 X 服从几何分布，记为 $X \sim Ge(p)$ ，其分布列为 $p(X = k) = (1-p)^{k-1}p, k = 1, 2, \dots$ 几何分布的数学期望和方差 $E(X) = \frac{1}{p}, Var(X) = \frac{1-p}{p^2}$

几何分布的无记忆性：设 $X \sim Ge(p)$ ，则对任意正整数 m 与 n 有

$p(X > m + n | X > m) = p(X > n)$, 这个定理表明: 在前 m 次实验中 A 没有出现的条件下, 则在接下来的 n 次实验中 A 仍未出现的概率只与 n 有关, 而与以前的 m 次试验无关, 似乎忘记了之前 m 次的实验结果, 这就是无记忆性。

(七) 负二项分布

在伯努利实验序列中, 记每次试验中事件 A 发生的概率为 P , 如果 X 为事件 A 第 r 次出现时的试验次数, 则 X 的可能取值为 $r, r + 1, \dots, r + m, \dots$ 称 X 服从负二项分布或帕斯卡分布, 其分布列为

$$P(X = r) = C_{k-1}^{r-1} p^r (1-p)^{k-r}, k = r, r + 1, \dots$$

记为 $X \sim Nb(r, p)$, 当 $r = 1$ 时, 记为几何分布

负二项分布的数学期望和方差为 $E(X) = \frac{r}{p}; Var(X) = \frac{r(1-p)}{p^2}$