



第四章 随机变量的数字特征

目录

1. 数学期望
2. 方差
3. 协方差与相关系数
4. 矩、协方差矩阵





1. 数学期望



数学期望

例：一射手进行打靶练习，规定射入区域 e_2 得2分；射入区域 e_1 得1分；脱靶(射入区域 e_0 得0分). 射手一次射击所得分数 X 是随机变量。 X 的分布律为：

$$P\{X = k\} = p_k, \quad k = 0, 1, 2$$

现射击 N 次，其中得0分 a_0 次，得1分 a_1 次，得2分 a_2 次， $a_0 + a_1 + a_2 = N$. 射击 N 次得分总和为 $a_0 \times 0 + a_1 \times 1 + a_2 \times 2$. 平均射击得分为：

$$\frac{a_0 \times 0 + a_1 \times 1 + a_2 \times 2}{N} = \sum_{k=0}^2 k \frac{a_k}{N}$$

这里 $\frac{a_k}{N}$ 为事件 $\{X = k\}$ 的频率. 当 N 很大时， $\frac{a_k}{N}$ 在一定意义下接近于事件 $\{X = k\}$ 的概率 p_k . 即随机变量 X 的观察值的算数平均 $\sum_{k=0}^2 k \frac{a_k}{N}$ 在一定意义下接近于 $\sum_{k=0}^2 k p_k$ ，我们称 $\sum_{k=0}^2 k p_k$ 为随机变量 X 的**数学期望**或**均值**。



数学期望

◆ **定义**：设离散型随机变量 X 的分布律为：

$$P\{X = x_k\} = p_k, \quad k = 1, 2, \dots$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称级数 $\sum_{k=1}^{\infty} x_k p_k$ 的和为随机变量 X 的**数学期望**，记为 $E(X)$ 。即

$$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

设连续型随机变量 X 的概率密度为 $f(x)$ ，若积分 $\int_{-\infty}^{\infty} x f(x) dx$ 绝对收敛，则积分 $\int_{-\infty}^{\infty} x f(x) dx$ 的值为随机变量 X 的**数学期望**，记为 $E(X)$ 。即

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

数学期望简称**期望**，又称为**均值**。



数学期望

例：某医院当新生儿诞生时，医生需要对婴儿的各方面情况进行评分，设新生儿的得分 X 是一个随机变量， X 的分布律如下表：

X	0	1	2	3	4	5	6	7	8	9	10
p_k	0.002	0.001	0.002	0.005	0.02	0.04	0.18	0.37	0.25	0.12	0.01

试求 X 的数学期望 $E(x)$.

解： $E(x) = 0 \times 0.002 + 1 \times 0.001 + 2 \times 0.002 + 3 \times 0.005 + 4 \times 0.02 + 5 \times 0.04 + 6 \times 0.18 + 7 \times 0.37 + 8 \times 0.25 + 9 \times 0.12 + 10 \times 0.01 = 7.15$ （分）

即若考察1000个新生儿，则一个新生儿的平均得分为7.15，1000个新生儿共得分7150分。



数学期望

例：有两个相互独立的电子装置，他们的寿命 X_k ($k = 1, 2$)服从同一指数分布，其概率密度为：

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x > 0, \theta > 0 \\ 0, & x \leq 0 \end{cases}$$

若将这两个电子装置串联成整机，求整机寿命 N 的数学期望。

解： X_k 的分布函数为：

$$F(x) = \begin{cases} 1 - e^{-x/\theta}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

而 $N = \min\{X_1, X_2\}$ 的分布函数为

$$F_{\min}(x) = 1 - [1 - F(x)]^2 = \begin{cases} 1 - e^{-\frac{2x}{\theta}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



数学期望

因而 N 的概率密度为:

$$f_{min}(x) = \begin{cases} \frac{2}{\theta} e^{-\frac{2x}{\theta}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

于是 N 的数学期望为:

$$E(N) = \int_{-\infty}^{\infty} x f_{min}(x) dx = \int_0^{\infty} \frac{2x}{\theta} e^{-\frac{2x}{\theta}} dx = \frac{\theta}{2}$$



数学期望

例：某车站每天8:00~9:00，9:00~10:00都恰有一辆客车到站，但到站时刻是随机的，且两者到站时间相互独立，其规律为：

到站时刻	8:10	8:30	8:50
	9:10	9:30	9:50
概率	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

一旅客8:20到车站，求他候车时间的数学期望。

解：设旅客的候车时间为 X ， X 的分布律为：

X	10	30	50	70	90
p_k	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6} \times \frac{1}{6}$	$\frac{1}{6} \times \frac{3}{6}$	$\frac{1}{6} \times \frac{2}{6}$

则候车时间的数学期望为：

$$E(X) = 10 \times \frac{3}{6} + 30 \times \frac{2}{6} + 50 \times \frac{1}{36} + 70 \times \frac{3}{36} + 90 \times \frac{2}{36} = 27.22$$



数学期望

例：某商店对某种家电销售采取先使用后付款的方式，记使用寿命为 X ，规定：

$X \leq 1$ ，一台付款1500元； $1 < X \leq 2$ ，一台付款2000元；

$2 < X \leq 3$ ，一台付款2500元； $X > 3$ ，一台付款3000元；

设寿命 X 服从指数分布，其概率密度如下：

$$f(x) = \begin{cases} \frac{1}{10} e^{-x/10}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

试求该商店一台这种家电收费 Y 的数学期望。

解：先求寿命 X 落在各个时间区间的概率，有

$$P\{X \leq 1\} = \int_0^1 \frac{1}{10} e^{-x/10} dx = 1 - e^{-0.1} = 0.0952$$



数学期望

$$P\{1 < X \leq 2\} = \int_1^2 \frac{1}{10} e^{-x/10} dx = e^{-0.1} - e^{-0.2} = 0.0861$$

$$P\{2 < X \leq 3\} = \int_2^3 \frac{1}{10} e^{-x/10} dx = e^{-0.2} - e^{-0.3} = 0.0779$$

$$P\{X > 3\} = \int_3^{\infty} \frac{1}{10} e^{-x/10} dx = e^{-0.3} = 0.7408$$

一台家用电器收费 Y 的分布律为;

Y	1 500	2 000	2 500	3 000
p_k	0.095 2	0.086 1	0.077 9	0.740 8

则 $E(Y) = 2732.15$, 即平均一台收费2732.15元。



数学期望

例：在一个人数很多的团体中普查某种疾病，为此要抽验 N 个人的血，可以用两种方法进行. (i) 将每个人的血分别去验，这就需验 N 次. (ii) 按 k 个人一组进行分组，把从 k 个人抽来的血混合在一起进行检验. 如果这混合血液呈阴性反应，就说明 k 个人的血都呈阴性反应，这样，这 k 个人的血就只需验一次；若呈阳性，则再对这 k 个人的血液分别进行化验，这样， k 个人的血总共要化验 $k+1$ 次. 假设每个人化验呈阳性的概率为 p ，且这些人的试验反应是相互独立的. 试说明当 p 较小时，选取适当的 k ，按第二种方法可以减少化验的次数. 并说明 k 取什么值时最适宜.

解：各人的血呈阴性反应的概率为 $q = 1 - p$. 因而 k 个人的混合血呈阴性反应的概率为 q^k ， k 个人的混合血呈阳性反应的概率为 $1 - q^k$.



数学期望

设以 k 个人为一组时，组内每人化验次数为 X ，则 X 是一个随机变量，其分布律为：

X	$\frac{1}{k}$	$\frac{k+1}{k}$
p_k	q^k	$1 - q^k$

X 的数学期望为 $E(x) = \frac{1}{k}q^k + \left(1 + \frac{1}{k}\right)(1 - q^k) = 1 - q^k + \frac{1}{k}$.

N 个人平均需化验次数为 $N(1 - q^k + \frac{1}{k})$

因此只要选择 k 使 $1 - q^k + \frac{1}{k} < 1$

则 N 个人平均需化验次数 $< N$. 当 p 固定时，选取 k 使得 $L = 1 - q^k + \frac{1}{k}$ 小于1且取到最小值，此时为最好分组法。

如： $p = 0.1; q = 0.9; k = 4$ 时， $L = 1 - q^k + \frac{1}{k}$ 取到最小值。

若 $N = 1000$ ，以 $k = 4$ 分组，按第二种方法化验只需
 $1000 \left(1 - 0.9^4 + \frac{1}{4}\right) = 594$ (次)。

减少了40%工作量



数学期望

例： 设随机变量 $X \sim \pi(\lambda)$, 求 $E(X)$ 。

X 的分布律为 $P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots, \lambda > 0$

X 的数学期望为：

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$$

例： 设随机变量 $X \sim U(a, b)$, 求 $E(X)$ 。

X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{others} \end{cases}$$

X 的数学期望为： $E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$



数学期望

◆ **定理**: 设 Y 是随机变量 X 的函数: $Y = g(X)$, g 是连续函数

➤ 如果 X 是离散型随机变量, 分布律为 $P\{X = x_k\} = p_k, k = 1, 2, \dots$, 若 $\sum_{k=1}^{\infty} g(x_k) p_k$ 绝对收敛, 则有

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k$$

➤ 如果 X 是连续型随机变量, 密度函数为 $f(x)$, 若 $\int_{-\infty}^{\infty} g(x) f(x) dx$ 绝对收敛, 则有

$$E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

注: 定理意义在于我们求 $E(Y)$ 时, 不要求 Y 的分布律或密度函数, 只需要知道 X 的分布律或密度函数即可



数学期望

证明: (只对下述特殊情况加以证明)

X 是连续型随机变量,且 $y = g(x)$ 满足第二章第五节中的定理条件, 则随机变量 $Y = g(X)$ 的概率密度为:

$$f_Y(y) = \begin{cases} f_X[h(y)]|h'(y)|, & \alpha < y < \beta \\ 0, & \text{others} \end{cases}$$

于是 $E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{\alpha}^{\beta} y f_X[h(y)]|h'(y)| dy$.

当 $h'(y) > 0$ 时,

$$E(Y) = \int_{\alpha}^{\beta} y f_X[h(y)]h'(y) dy = \int_{-\infty}^{\infty} g(x) f(x) dx$$

当 $h'(y) < 0$ 时,

$$E(Y) = - \int_{\alpha}^{\beta} y f_X[h(y)]h'(y) dy = \int_{-\infty}^{\infty} g(x) f(x) dx$$

综上所述, 证毕。



数学期望

推广:

设 Z 是随机变量 X, Y 的函数 $Z = g(X, Y)$ (g 是连续函数), 则 Z 是一个一维随机变量。若二维随机变量 (X, Y) 的概率密度为 $f(x, y)$ 则有

$$E(Z) = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

这里设上式右边的积分绝对收敛。

若 (X, Y) 为离散型随机变量, 其分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}, i, j = 1, 2, \dots$, 则有

$$E(Z) = E[g(X, Y)] = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} g(x_i, y_j) p_{ij}$$

这里设右边的级数绝对收敛。



数学期望

例： 设风速 V 在 $(0, a)$ 上服从均匀分布，即具有概率密度

$$f(v) = \begin{cases} \frac{1}{a}, & 0 < v < a \\ 0, & \text{others} \end{cases}$$

设飞机机翼受到的正压力 W 是 V 的函数： $W = kV^2$ ($k > 0$, 常数)，求 W 的数学期望。

解： $E(W) = \int_{-\infty}^{\infty} kv^2 f(v) dv = \int_0^a kv^2 \frac{1}{a} dv = \frac{1}{3} ka^2$

例： 设随机变量 (X, Y) 的概率密度

$$f(x, y) = \begin{cases} \frac{3}{2x^3y^2}, & \frac{1}{x} < y < x, x > 1 \\ 0, & \text{others} \end{cases}$$

求数学期望 $E(Y), E\left(\frac{1}{XY}\right)$



数学期望

解:

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dy dx = \int_1^{\infty} \int_{\frac{1}{x}}^x \frac{3}{2x^3 y} dy dx \\ &= \frac{3}{2} \int_1^{\infty} \frac{1}{x^3} [\ln y]_{\frac{1}{x}}^x dx = 3 \int_1^{\infty} \frac{\ln x}{x^3} dx = \left[-\frac{3}{2} \frac{\ln x}{x^2} \right]_1^{\infty} + \frac{3}{2} \int_1^{\infty} \frac{1}{x^3} dx \\ &= \frac{3}{4} \end{aligned}$$

$$E\left(\frac{1}{XY}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{xy} f(x, y) dy dx = \int_1^{\infty} dx \int_{\frac{1}{x}}^x \frac{3}{2x^4 y^3} dy = \frac{3}{5}$$



数学期望

例：某公司计划开发一种新产品市场，并试图确定该产品的产量。他们估计出售一件产品可获利 m 元，积压一件产品损失 n 元，他们预测销售量 Y 服从指数分布，其概率密度为：

$$f_Y(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta}, & y > 0, \\ 0, & y \leq 0 \end{cases}, \quad \theta > 0$$

问若要获利的数学期望最大，应生产多少件产品。

解： 设生产 x 件，则获利 Q 是 x 的函数

$$Q = Q(x) = \begin{cases} mY - n(x - Y), & Y < x \\ mx, & Y \geq x \end{cases}$$

Q 是随机变量，它是 Y 的函数，其数学期望为



数学期望

$$\begin{aligned} E(Q) &= \int_0^{\infty} Q f_Y(y) dy \\ &= \int_0^x [my - n(x - y)] \frac{1}{\theta} e^{-y/\theta} dy + \int_x^{\infty} mx \frac{1}{\theta} e^{-y/\theta} dy \end{aligned}$$

$$= (m + n)\theta - (m + n)\theta e^{-x/\theta} - nx$$

$$\text{令 } \frac{d}{dx} E(Q) = (m + n)e^{-x/\theta} - n = 0$$

$$\text{得 } x = -\theta \ln \frac{n}{m+n}$$

$$\text{而 } \frac{d^2}{dx^2} E(Q) = \frac{-(m+n)}{\theta} e^{-x/\theta} < 0$$

故知当 $x = -\theta \ln \frac{n}{m+n}$ 时 $E(Q)$ 取极大值，且可知也是最大值。



数学期望

例：甲与其他三人参与竞拍，价格高者获胜，若甲中标则将此项目以10千美金转让给他人，可以认为其他三人竞价相互独立，且都在7~11千美金之间均匀分布，问甲应该如何报价才能使获利期望最大。

解：设 X_1, X_2, X_3 是其他三人得报价，按题意 X_1, X_2, X_3 相互独立，且在区间(7,11)上服从均匀分布。其分布函数为

$$F(u) = \begin{cases} 0, & u < 7 \\ \frac{u-7}{4}, & 7 \leq u < 11 \\ 1, & u \geq 11 \end{cases}$$

以 Y 为三人得最高出价,即 $Y = \max\{X_1, X_2, X_3\}$, Y 的分布函数为

$$F_Y(u) = \begin{cases} 0, & u < 7 \\ \left(\frac{u-7}{4}\right)^3, & 7 \leq u < 11 \\ 1, & u \geq 11 \end{cases}$$



数学期望

若甲报价为 x , 按题意 $7 \leq x \leq 10$, 知甲能赢这一项目得概率为

$$p = P\{Y \leq x\} = F_Y(x) = \left(\frac{x-7}{4}\right)^3$$

以 $G(X)$ 为甲赚钱数, $G(X)$ 的分布律为

$G(x)$	$10-x$	0
概率	$\left(\frac{x-7}{4}\right)^3$	$1 - \left(\frac{x-7}{4}\right)^3$

则甲的赚钱数的数学期望为 $E[G(X)] = \left(\frac{x-7}{4}\right)^3 (10-x)$

$$\text{令 } \frac{d}{dx} E[G(X)] = \frac{1}{4^3} [(x-7)^2(37-4x)] = 0$$

$$\text{得 } x = \frac{37}{4}, x = 7(\text{舍去}), \text{ 又 } \frac{d^2}{dx^2} E[G(X)]|_{x=37/4} < 0$$

故当甲报价为 $x = \frac{37}{4}$ 千美元时, 数学期望达到最大值。



数学期望

◆ 数学期望性质:

- 设 C 是常数, 则有 $E(C) = C$.
- 设 X 是一个随机变量, C 是常数, 则有 $E(CX) = CE(X)$.
- 设 X, Y 是两个随机变量, 则有 $E(X + Y) = E(X) + E(Y)$
(可推广到任意有限个随机变量之和的情况)
- 设 X, Y 是相互独立的随机变量, 则有 $E(XY) = E(X)E(Y)$
(可推广到任意有限个随机变量之积的情况)



数学期望

证3: 设二维随机变量 (X, Y) 的概率密度为 $f(x, y)$, 其边缘概率密度为 $f_X(x), f_Y(y)$.

$$\begin{aligned} E(X + Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy \\ &= E(X) + E(Y) \end{aligned}$$

证4: 若 X 和 Y 相互独立

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dx dy \\ &= \left[\int_{-\infty}^{\infty} x f_X(x) dx \right] \left[\int_{-\infty}^{\infty} y f_Y(y) dy \right] = E(X) E(Y) \end{aligned}$$



数学期望

例：一民航送客车载有20位旅客自机场开出，旅客有10个车站可以下车，如到达一个车站没有旅客下车就不停车，以 X 表示停车的次数，求 $E(X)$ (设旅客在各车站下车是等可能的、且相互独立)。

解：引入随机变量

$$X_i = \begin{cases} 0, & \text{在第} i \text{站没有人下车} \\ 1, & \text{在第} i \text{站有人下车} \end{cases}, \quad i = 1, 2, \dots, 10$$

易知 $X = X_1 + X_2 + \dots + X_{10}$ ，现求 $E(X)$ 。

任一旅客在第 i 站不下车的概率为 $9/10$ ，因此20位旅客都不在第 i 站下车的概率为 $(\frac{9}{10})^{20}$ ，在第 i 站有人下车的概率为 $1 - (\frac{9}{10})^{20}$

即 $P\{X_i = 0\} = (\frac{9}{10})^{20}$ ， $P\{X_i = 1\} = 1 - (\frac{9}{10})^{20}$

由此 $E(X_i) = 1 - (\frac{9}{10})^{20}$



数学期望

进而 $E(X) = E(X_1 + X_2 + \cdots + X_{10}) = E(X_1) + E(X_2) + \cdots + E(X_{10}) = 10 \left[1 - \left(\frac{9}{10} \right)^{20} \right] = 8.784(\text{次})$

例： 设一电路中电流 I 与电阻 R 是两个相互独立的随机变量，其概率密度分别为

$$g(i) = \begin{cases} 2i, & 0 \leq i \leq 1 \\ 0, & \text{others} \end{cases}, h(r) = \begin{cases} \frac{r^2}{9}, & 0 \leq r \leq 3 \\ 0, & \text{others} \end{cases}$$

试求电压 $V = IR$ 的均值。

解： $E(V) = E(IR) = E(I)E(R) =$
 $\left[\int_{-\infty}^{\infty} i g(i) di \right] \left[\int_{-\infty}^{\infty} r h(r) dr \right] = \left(\int_0^1 2i^2 di \right) \left(\int_0^3 \frac{r^3}{9} dr \right) = \frac{3}{2} (V)$

