1.2 随机变量的数字特征

- 一　数学期望的概念
- 二　随机变量函数的数学期望
- 三　方差、协方差与相关系数
- 四　条件数学期望

BACK

一　数学期望的概念

1. 离散型随机变量的数学期望

定义2.1　设离散型随机变量X的分布律为

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

若级数$\sum\limits_{k-1}^{\infty} x_k p_k$绝对收敛,则称

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

为X的数学期望,亦称为概率均值,简称均值或期望

2 连续型随机变量的数学期望

定义2.2 设连续型随机变量X的概率密度为$f(x)$,若积分$\int_{-\infty}^{+\infty} f(x)\,dx$ 绝对收敛,则称

$$E(X)=\int_{-\infty}^{+\infty} xf(x)\,dx$$

为X的数学期望,简称期望或均值

从力学的角度看,设在OX轴上分布着质点,其线密度为$f(x)$,因$\int_{-\infty}^{+\infty} f(x)dx=1$,故有

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

即数学期望值表示质量中心的坐标。

二 随机变量函数的数学期望

1 离散型随机变量的函数的期望

若X的分布律为$P\{X=x_k\}=p_k, k=1,2,\cdots, g(x)$为连续函数

且$\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛,则函数$Y=g(X)$的期望为

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

2 连续型随机变量的函数的期望

若连续型随机变量X的概率密度为$f(x)$,

且$\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛,则函数$Y=g(X)$的期望为

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

3 多维随机变量函数的数学期望

(1) 若已知(X,Y)的分布律为$P\{X=x_i,Y=y_j\}=p_{ij},i,j=1,2,\cdots$

则函数$Z=g(X,Y)$的数学期望为

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

(2) 若已知(X,Y)的概率密度为$f(x,y)$,

则函数$Z=g(X,Y)$的数学期望为

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

注1　上式中级数与积分均要求绝对收敛。

注2　对二维以上的函数的期望公式类似于上式。

三 随机变量的方差、协方差与相关系数

1 方差的定义

定义2.3 设X是一个随机变量，若$E\left\{[X-E(X)]^2\right\}$存在，则称$E\left\{[X-E(X)]^2\right\}$为$X$的方差，记为$D(X)$或$Var(X)$,$\sigma_X^2$,即

$$\boxed{D(X)=Var(X)=E\left\{\left[X-E(X)\right]^2\right\}}$$

显然，$D(X)\geq 0$，可将$D(X)$开平方，此时称$\sigma_X=\sigma(X)=\sqrt{D(X)}$为$X$的标准差或均方差。

方差是反映数据疏散程度特征的量。方差大，说明数据疏散；方差小，说明数据集中。 由方差的定义式易得

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

$$E(X^2)=D(X)+[E(X)]^2$$

若X为离散型随机变量，其概率分布为

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

则可得X的方差为

$$D(X)=\sum_{k=1}^{\infty}(x_k-E(X))^2 p_k$$

$$D(X)=\sum_{k=1}^{\infty}x_k^{\ 2} p_k-(\sum_{k=1}^{\infty}x_k p_k)^2$$

若X为连续型随机变量，其概率密度为$f(x)$,
则由可得X的方差为

$$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2 f(x)dx$$

$$D(X)=\int_{-\infty}^{+\infty}x^2 f(x)dx-\left[\int_{-\infty}^{+\infty}xf(x)dx\right]^2$$

2、协方差定义

定义2.4 设(X,Y)为二维随机变量，称

Cov(X , Y)=E[(X-E(X))(Y-E(Y)]

为X与Y的协方差。

协方差是反映两随机变量X与Y相关关系的特征量。

由方差定义与协方差定义可知

$$\boxed{D(X+Y)=D(X)+D(Y)+2Cov(X,Y)}$$

$$\boxed{Cov(X,Y)=E(XY)-E(X)E(Y)}$$

协方差的基本性质有

1^0 $Cov(X,Y)=Cov(Y,X)$

2^0 $Cov(aX,bY)=abCov(X,Y)$

3^0 $Cov(X_1 \pm X_2,Y)=Cov(X_1,Y)\pm Cov(X_2,Y)$

4^0 $|Cov(X,Y)| \le \sqrt{D(X)}\sqrt{D(Y)}$

5^0 $Cov(X,X)=D(X)$

6^0 若X与Y相互独立，则$Cov(X,Y)=0$

7^0 $Cov(aX \pm b, cY \pm d)=acCov(X,Y)$

3、相关系数的定义

定义2.5 设(X,Y)为二维随机变量，$D(X),D(Y),Cov(X,Y)$分别为X,Y的方差与协方差，称

$$\boxed{\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}}$$

为随机变量X与Y的相关系数。

相关系数是反映X与Y相关关系的一个无量纲 的特征量。

相关系数具有以下两个性质：

1^0 $|\rho_{XY}|\le 1$ $|\rho_{XY}|=1$的充要条件是存在常数a,b,使

$$\boxed{P\{Y=aX+b\}=1}$$

2^0 若X,Y相互独立，且$D(X),D(Y)>0$,则 $\boxed{\rho_{XY}=0}$

若X,Y的相关系数$\rho_{XY}=0$，则称X与Y为不相关

四　条件数学期望

由概率论知识可知，若(X, Y)为二维随机变量，则在一定条件下可求条件分布函数与条件概率分布或条件概率密度，由此我们可用类似数学期望的定义去定义条件数学期望。

1　离散型随机变量的条件期望

设(X, Y)为二维离散型随机变量，若其概率分布为

$$\boxed{P\{X = x_i, Y = y_i\} = p_{ij},\ i, j = 1, 2, \cdots}$$

边缘概率分布为

$$p_{i\bullet} = P(X = x_i\} = \sum_{j=1}^{\infty} p_{ij}, \quad p_{\bullet j} = P\{Y = y_j\} = \sum_{i=1}^{\infty} p_{ij}$$

则条件数学期望如下定义：

定义2.6 若级数 $\sum_{j=1}^{\infty} y_j P\{Y = y_j \mid X = x_i\} = \sum_{j=1}^{\infty} y_j \frac{p_{ij}}{p_{i\bullet}}$ 绝对收敛，则称此级数为在 $\{X = x_i\}$ 条件下，Y的条件数学期望。记作

$$E(Y \mid X = x_i) \quad i = 1, 2, \cdots$$

即

$$E\{Y \mid X = x_i) = \sum_{j=1}^{\infty} y_j \frac{p_{ij}}{p_{i\bullet}}$$

类似地，在{ Y=y_j }条件下X的条件数学期望为

$$E\{X \mid Y = y_j) = \sum_{i=1}^{\infty} x_i \frac{p_{ij}}{p_{\bullet j}}$$

例2.1 设(X, Y)的概率分布如下

X \ Y	0	1	2
0	0.2	0.1	0.3
1	0.1	0.2	0.1

试求出其全部的条件数学期望。

解：由已知概率分布可得关于X与Y的边缘概率分布为：

X	0	1
$p_{i.}$	0.6	0.4

Y	0	1	2
$p_{.j}$	0.3	0.3	0.4

则条件概率分布与条件数学期望为

$$P\{Y=0 \mid X=0) = \frac{0.2}{0.6} = \frac{1}{3} \qquad P\{Y=1 \mid X=0) = \frac{0.1}{0.6} = \frac{1}{6}$$

$$P\{Y=2 \mid X=0) = \frac{0.3}{0.6} = \frac{1}{2}$$

$$E(Y \mid X=0) = \sum_{j=0}^{2} jP\{Y=j \mid X=0\}$$

$$= 0 \times \frac{1}{3} + 1 \times \frac{1}{6} + 2 \times \frac{1}{3} = \frac{5}{6}$$

$$P(X=0 \mid Y=0) = \frac{0.2}{0.3} = \frac{2}{3} \qquad P(X=1 \mid Y=0) = \frac{0.1}{0.3} = \frac{1}{3}$$

$$E(X \mid Y=0) = \sum_{i=0}^{1} iP\{X=i \mid Y=0\} = 0 \times \frac{2}{3} + 1 \times \frac{1}{3} = \frac{1}{3}$$

类似可得 $E(X \mid Y=1) = \dfrac{2}{3}$ $\qquad E(X \mid Y=2) = \dfrac{1}{4}$

例2.2 一射手进行射击，击中目标的概率为$p(0<p<1)$，射击到击中两次目标为止。设X表示首次击中目标所进行的射击次数，以Y表示总共进行的射击次数，试求X和Y的条件数学期望。

解：由题意得(X, Y)的联合概率分布为

$$P\{X=i,Y=j\}=p^2(1-p)^{j-2} \quad j=2,3,\cdots,\ i=1,2,\cdots,j-1$$

易计算得X与Y的边缘分布律为

$$P\{X=i\}=p(1-p)^{i-1},\ i=1,2,\cdots$$

$$P\{Y=j\}=(j-1)p^2(1-p)^{j-2},\ j=2,3,\cdots$$

当$Y=j=2,3,\cdots$时，X的条件分布律为

$$P\{X=i\mid Y=j\}=\frac{1}{j-1} \qquad i=1,2,\cdots,j-1$$

$$E(X \mid Y = j) = \sum_{i=1}^{j-1} \frac{1}{j-1} = 1, \quad j = 2, 3, \cdots$$

此条件数学期望与j无关。

又当$X = i = 1,2,\cdots$时，Y的条件分布律为

$$P\{Y = j \mid X = i\} = p(1-p)^{j-i-1}, \quad j = i+1, i+2, \cdots$$

$$E(Y \mid X = i) = \sum_{j=i+1}^{\infty} jp(1-p)^{j-i-1} = \sum_{k=0}^{\infty} (k+i+1)p(1-p)^k$$

$$= \sum_{k=0}^{\infty} (k+1)p(1-p)^k + i\sum_{k=0}^{\infty} p(1-p)^k = \frac{1}{p} + i$$

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

2、连续型随机变量的条件期望

设(X, Y)为连续型随机变量，其概率密度为$f(x, y)$，边缘概率密度为

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx$$

$$\boxed{f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)},\ f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(x)}}$$

则其相应的条件数学期望如下定义：

定义2.7 若广义积分 $\int_{-\infty}^{+\infty} y f_{Y|X}(y \mid x)dy$ 绝对收敛，则称此积分为在条件$X = x$下Y的条件数学期望，记为$E(Y|X=x)$，即

$$E(Y \mid X = x) = \int_{-\infty}^{+\infty} y f_{Y|X}(y \mid x)dy$$

类似地，在$Y = y$条件下X的条件数学期望为

$$E(X \mid Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y)dx$$

例2.3 设(X, Y)在圆域 $X^2+Y^2 \leqslant 1$上服从均匀分布，分别求$Y = y$与$X = x$条件下的条件数学期望。

解：由题意知，(X, Y)的概率密度为

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

$$f_Y(y)=\begin{cases}\frac{2}{\pi}\sqrt{1-y^2} & -1<y<1\\ 0 & \text{其它}\end{cases}\qquad f_X(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2} & -1<x<1\\ 0 & \text{其它}\end{cases}$$

当 $Y=y$ $(-1<y<1)$ 时X的条件概率密度为

$$f_{X|Y}(x\mid y)=\frac{1}{2\sqrt{1-y^2}}\qquad -\sqrt{1-y^2}<x<\sqrt{1-y^2}$$

$$E(X\mid Y=y)=\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}x\frac{1}{2\sqrt{1-y^2}}dx=0$$

类似可得$X=x$ $(-1<x<1)$时Y的条件数学期望为

$$E(Y\mid X=x)=0$$

例2.4 设在$Y=y$ $(0<y<1)$条件下X的条件概率密度为

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{3x^2}{y^3} & 0<x<y\\ 0 & \text{其它}\end{cases}$$

试求条件数学期望:

$$E(X|Y=y)$$

解：由定义可得

$$E(X|Y=y)=\int_0^y x\frac{3x^2}{y^3}dx=\frac{3}{y^3}\cdot\frac{1}{4}x^4\Big|_0^y=\frac{3y}{4}$$

3、条件数学期望的性质

设X, Y, Z为随机变量，$g(x)$在R上连续，且$E(X)$, $E(Y)$, $E(Z)$及$E(g(Y)X)$均存在，容易证明条件数学期望有如下性质，我们只在连续型随机变量情况下给出证明，离散型情况类似可得。

1° 当X与Y相互独立时，必有

$$\boxed{E(X \mid Y = y) = E(X)} \qquad \boxed{E(Y \mid X = x) = E(Y)}$$

即**X与Y独立时，条件期望与无条件期望相等**。

2° 全期望公式: $\boxed{E(X) = E[E(X \mid Y = y)]}$

证：$E[E(X \mid Y = y)] = \int_{-\infty}^{+\infty} E(X \mid Y = y) f_Y(y) dy$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y) f_Y(y) dy dx$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy$$

$$= \int_{-\infty}^{+\infty} x \left[\int_{-\infty}^{+\infty} f(x, y) dy \right] dx = \int_{-\infty}^{+\infty} x f_X(x) dx = E(X)$$

离散型随机变量情形类似。

$$3^\circ \quad E[g(Y)\cdot X \mid Y=y]=g(Y)E(X \mid Y=y)$$

实际上，对于任意固定的y，成立等式:

$$E[g(y)\cdot X \mid Y=y]=g(y)E(X \mid Y=y)$$

可知3°结论为真。

$$4^\circ \quad E[g(Y)X]=E[g(Y)\cdot E(X \mid Y=y)]$$

这由2°与3°结论可得。

5° 设C为常数,则 $E(C \mid Y=y)=C$

由条件数学期望定义立知其为真。

$$6^\circ \quad E[(aX+bY) \mid Z=z]=aE(X \mid Z=z)+bE(Y \mid Z=z)$$

由定义直接可得。

$8^\circ \quad E[X - E(X \mid Y = y)]^2 \le E[X - g(Y)]^2$

证：对任一固定的y,

$$E[X - g(Y)]^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - g(y))^2 f(x,y)dxdy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x - g(y))^2 f_{X|Y}(x \mid y) f_Y(y)dxdy$$

$$= \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}[x - g(y)]^2 f_{X|Y}(x \mid y)dx\right] f_Y(y)dy$$

由数学期望性质知，当$g(y) = E(X \mid Y = y)$时，积分

$\int_{-\infty}^{+\infty}(x - g(y))^2 f_{X|Y}(x \mid y)dx$ 达到最小，

因此$E[X - g(Y)]^2$当$g(Y) = E(X \mid Y = y)$时，达到最小。

例2.5 （巴格达窃贼问题）一窃贼被关在3个门的地牢中，其中第1个门通向自由，出这个门后3个小时便回到地面；第2个门通向一个地道，在此地道中走5个小时后将返回地牢；第3个门通向一个更长的地道，沿这个地道走7个小时也回到地牢，如果窃贼每次选择3个门的可能性总相等，试求他为获得自由而奔走的平均时间。

解：设窃贼需走X个小时到达地面，并设Y为窃贼每次对3个门的选择，则Y均以1/3的概率取值为1, 2, 3，可利用全期望公式得:

$$E(X)=E[E(X\mid Y=j)]=\sum_{i=1}^{3}E(X\mid Y=j)P(Y=j)$$

而有$E(X\mid Y=1]=3, E(X\mid Y=2]=5+E(X)$, 这是因为:

若窃贼选第1个门，则3个小时后肯定到达地面，故有E(X|Y=1)=3，而窃贼选第2个门时，他花5个小时重回地牢，此时处境与开始时完全一样，故有

$$E(X \mid Y=2)=5+E(X)$$

类似地，窃贼选第3个门时，有

$$E(X \mid Y=3)=7+E(X)$$

故得 $E(X)=\dfrac{1}{3}\left[3+5+E(X)+7+E(X)\right]$

解得: $E(X)=15$(小时）

即窃贼若从3个门中等可能地选择逃跑时，平均15个小时后获得自由。

例2.6　设随机变量X的概率密度为

$$f_X(x)=\begin{cases}\lambda^2 xe^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

随机变量Y在区间(0, X)上均匀分布，试求条件期望E(Y|X=x)与无条件期望E(Y)

解：当x > 0时:

$$E(Y\mid X=x)=\int_{-\infty}^{+\infty} yf_{Y|X}(y\mid x)dy=\int_0^x y\cdot\frac{1}{x}dy=\frac{x}{2}$$

再由全期望公式得

$$E(Y)=E\left[E(Y\mid X=x)\right]=\int_{-\infty}^{+\infty}E(Y\mid X=x)\cdot f_X(x)dx$$

$$=\int_0^{+\infty}\frac{x}{2}\cdot\lambda^2 xe^{-\lambda x}dx=\frac{1}{2}\int_0^{+\infty}(\lambda x)^2e^{-\lambda x}dx$$

$$\underline{\underline{u=\lambda x}}\frac{1}{2\lambda}\int_0^{+\infty}u^2e^{-u}du=\frac{1}{2\lambda}\Gamma(3)=\frac{1}{\lambda}$$

Forward

亦可通过无条件期望定义求E(Y)加以印证