实践六 随机变量的数字特征

一、 实践问题

1. 问题背景

随机变量的数字特征能够描述随机变量的重要性质和人们最关心的特征，如数学期望、方差、 协方差、 相关系数等. 它们在应用和理论上都非常重要. 但随机变量的数字特征的计算比较繁杂. MATLAB 软件提供了计算随机变量的数字特征的相关函数，计算起来方便、快捷、准确.

2. 实践目的与要求

(1) 熟练掌握随机变量数字特征的有关操作命令;

(2) 利用软件求边缘概率密度，求随机变量函数的数字特征;

(3) 利用软件处理简单的概率问题.

二、 实践操作过程

1. 样本均值

格式:

• Y=mean(X)　　　% X 是向量，返回的 Y 是 X 中元素的算术平均值。

例 6-1 随机抽取 6 个滚珠，测得直径(单位:mm)如下:

14.70, 15.21, 14.90, 14.91, 15.32, 15.32 .

试求该样本均值.

解 在命令窗口中输入:

```
X=[14.70  15.21  14.90  14.91  15.32  15.32];
mean(X)        %计算样本均值.
```

回车后结果如下:

```
ans =
    15.0600
```

2. 随机变量的数学期望

(1) 离散型随机变量的数学期望

基本数学原理: 设随机变量的分布律为

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

则

$$E(X)=\sum_{k=1}^{+\infty}x_kp_k$$

随机变量函数的数学期望:

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

在 MATLAB 中，离散型随机变量的数学期望函数是求和函数 sum.

调用格式

• EX=sum(X.*P)　　%计算随机向量 X 与对应概率向量 P 的乘积之和.

例 6-2 设随机变量 X 的分布律为

X	-2	-1	0	1	2
p	0.3	0.1	0.2	0.1	0.3

求 $E(X)$, $E(X^2-1)$.

解 在命令窗口中输入:

```
X=[-2 -1 0 1 2];
p=[0.3 0.1 0.2 0.1 0.3];
EX=sum(X.*p)        %数组 X 与 p 对应元素相乘，不可缺运算点“.”.
Y=X.^2-1            %数组 X 自身元素平方，不可缺运算点“.”.
EY=sum(Y.*p)
```

回车后显示:

```
EX =
    0
Y =
    3     0    -1     0     3
EY =
    1.6000
```

结果表明：随机变量 X 的数学期望为 0，随机变量函数 $Y=X^2-1$ 的数学期望为 1.6000.

（2）**连续型随机变量的数学期望**

基本数学原理: 设随机变量的概率密度为 f(x)，则

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

随机变量函数的数学期望:

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

在 MATLAB 中，连续型随机变量的数学期望函数是积分函数 int.

调用格式

•EX=int(x*fx, x, a, b)　　%用积分计算期望, fx 是随机变量的概率密度函数, [a, b]是积分区间

例 6-3 设连续型随机变量 X 的概率密度

$$f(x)=\begin{cases}x^2 & 0\le x\le 1,\\ 0 & others.\end{cases}$$

求随机变量 X 的数学期望 E(X).

解 在命令窗口输入:

```
syms x;
fx=x^2;
EX=int(x*fx,x,0,1)
```

回车后显示:

```
EX=
    1/4
```

结果表明：连续型随机变量 X 的数学期望是 1/4 .

例 6-4 求解概率密度函数为分段函数时的数学期望:

设随机变量 X 的概率密度为

$$f(x)=\begin{cases}0.5e^{x} & x\le 0,\\ 0.5e^{-x} & others.\end{cases}$$

求随机变量 Y=|X|的期望.

解 在命令窗口输入:

```
syms x;
fx1=0.5*exp(x); fx2=0.5*exp(-x);
EY=int(-x*fx1, x, -inf, 0)+int(x*fx2, x, 0, inf)    %用一行命令计算分段 函数的数学期望，可以分开计算，然后相加.
```

回车后显示:

```
EY=
    1
```

结果说明：随机变量函数 Y=|X|的期望为 1.

3. 样本方差与随机变量的方差

(1) 样本方差

调用格式:

• D=var(x)

• D=var(x,1)

当 x 是向量时，返回的是 x 的方差. 当 x 是矩阵时，返回的是一个向量，对应于矩阵每一列的方差值.

命令 D=var(x)计算的是

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

命令 D=var(x,1)计算的是

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(2) 样本标准差

调用格式:

• S=std(x,1)

(3) 随机变量的方差

基本数学原理: 随机变量 X 的方差:

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

离散型随机变量的方差:

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

其中$P\{X = x_k\} = p_k, \quad k = 1, 2, ...$为 X 的分布律.

连续随机变量的方差:

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

例 6-5 设随机变量 X 的概率密度为

$$f(x) = \begin{cases} \frac{2}{\pi}\cos 2x & |x| \le \frac{\pi}{2}, \\ 0 & others. \end{cases}$$

求随机变量 X 的期望和方差.

解法一 在命令窗口输入:

```
syms x;
fx=cos(2*x)*2/pi;
EX=int(x*fx, x, -pi/2, pi/2);
DX=int((x-EX)^2*fx,x,-pi/2,pi/2) %D(X) = ∫_{-∞}^{+∞}[x − E(X)]^2 f(x)dx
```

回车后显示:

```
EX=
    0
DX=
    -1
```

结果表明: 随机变量 X 的数学期望是 0，方差为-1.

解法二 在命令窗口输入:

```
syms x;
fx=cos(2*x)*2/pi;
EX=int(x*fx,x,-pi/2,pi/2)
E2X=int(x^2*fx,x,-pi/2,pi/2)   % 计算E(X^2).
DX=E2X-EX^2                    %用公式D(X) = E(X^2) − [E(X)]^2计算.
```

回车后显示:

```
EX=
    0
E2X=
    -1
DX=
    -1
```

结果表明: 随机变量 X 的数学期望是 0，方差为-1.

4. 常见分布的期望和方差

(1) 二项分布:

基本数学原理: 设随机变量 X～B(n, p)，则 E(X)=np，D(X)=npq，这里 q=1-p. 具体概率分布见实践二.

在 MATLAB 中,用函数 binostat 来计算 X 的期望和方差,其基本调用格式 为:

- [E,D]=binostat(n, p)　　% 返回二项分布(参数为 n 和 p)的均值 E 和方差 D. 输入的向量或矩阵 n, p 必须形式相同，输出 E 和 D 也与输入形式相同. 标量输入将被扩展成和其它输入具有相同维数的常数矩阵.

(2) 超几何分布: [E,D]=hygestat(M, N, K)

功能同(1). 具体概率分布见实践二.

(3) 泊松分布: [E,D]=poissstat(lambda)

功能同(1). 对于参数为 λ 的泊松分布，期望与方差都是 λ.

(4) 均匀分布: [E, D]=unifstat(a, b)

功能同(1). 对于参数为 a, b 的均匀分布，期望与方差分别为期望: $\frac{a+b}{2}$，方差: $\frac{(b-a)^2}{12}$

(5) 指数分布: [E, D]=expstat(lambda)

功能同(1). 对于参数为 λ 的指数分布，均值与方差分别为: λ，λ^2.

(6) 正态分布: [E, D]=normstat(mu, sigma)

功能同(1). 对于参数为μ与σ^2的正态分布，期望与方差分别为:μ, σ^2.

其它分布: gamstat( ), tstat(), fstat(),chi2stat( )等等. 具体概率分布见实践二.

例 6-6 求服从自由度为 25 的χ^2分布的期望和方差.

解 在命令窗口输入:

```
[E,D]=chi2stat(25)
```

回车后显示:

```
E=
    25
D=
    50
```

结果表明: 自由度为 25 的χ^2分布的期望为 25, 方差为 50.

5. 协方差和相关系数

(1) 统计数据的协方差

基本数学原理:

$$Cov(X,Y)=E\{[X-E(X)][(Y-E(Y)]\}=E(XY)-E(X)E(Y).$$

调用格式:

• cov(X) % 返回向量 X 的协方差, 即向量 X 的方差, 也可以用命令 var(X).

• cov(A) % 求矩阵 A 的协方差矩阵, 该协方差矩阵的对角线元素是 A 的各列的方差, 即: var(A)=diag(cov(A)). 在 i 行 j 列的元素为第 i 列与第 j 列向量的协方差.

• cov(X,Y) % X, Y 为等长列向量, 等同于 cov([X Y]).

例 6-7 求向量、矩阵的协方差及协方差矩阵.

解 在命令窗口中输入:

```
X=[0-1 1]'; Y=[1 2 2]';
C1=cov(X);     % 求列向量 X 的协方差 cov(X)=1.
C2=cov(X,Y)    % 求列向量 X、Y 的协方差矩阵, 对角线元素为各列向量的方差.
```

回车后显示:

```
C2 =
    1.0000         0
         0    0.3333
```

```
A=[1 2 3;4 0 -1;1 7 3];
C1=cov(A)      % 求矩阵 A 的协方差矩阵.
```

回车后显示:

```
C1 =
     3.0000   -4.5000   -4.0000
    -4.5000   13.0000    6.0000
    -4.0000    6.0000    5.3333
```

结果表明: 主对角线元素 3.0000,13.0000,5.3333 分别为 A 的第一列(1 4 1) ', 第二列(2 0 7) '和第三列(3 -1 3) '的方差, 其余(i, j)元素表示第 i 列与第 j 列向量的协方差.

(2) 统计数据的相关系数

基本数学原理:

$$\rho = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

调用格式:

- corrcoef(X,Y)　　%返回列向量 X,Y 的相关系数,等同于 corrcoef([X　Y]).
- corrcoef (A)　　%返回矩阵 A 的列向量的相关系数矩阵.

例 6-8 求二维随机变量(X,Y)的边缘概率密度、各自的数学期望与方差、X 与 Y 的协方差和相关系数:

设随机变量(X,Y)具有概率密度

$$f(x)=\begin{cases}\frac{1}{8}(x+y) & 0\le x\le 2,\quad 0\le y\le 2,\\ 0 & others.\end{cases}$$

解 在命令窗口输入:

```
syms x y m;
f1=(x+y)/8;
fx=int(f1,y,0,2);          % 计算关于 X 的边缘概率密度.
fy=int(f1,x,0,2);          % 计算关于 Y 的边缘概率密度.
f2=int(x*y*f1,y,0,2);      %计算 XY 的数学期望, 先对 y 积分一次.
EXY=int(f2,x,0,2);         %再对 x 积分一次,得到 XY 的数学期望.
EX=int(x*fx,x,0,2);        %计算 X 的数学期望.
EX1=int(x^2*fx,x,0,2);     %计算 X2 的数学期望.
DX=EX1-EX^2;               %计算 X 的方差.
EY=int(y*fy,y,0,2);        %计算 Y 的数学期望.
EY1=int(y^2*fy,y,0,2);     %计算 Y2 的数学期望.
DY=EY1-EY^2;               %计算 Y 的方差.
y1=EXY-EX*EY               %计算 X 与 Y 的协方差.
y2=sqrt(DX*DY);            %计算 X 与 Y 的相关系数的分母.
m=y1/y2                    %计算 X 与 Y 的相关系数.
```

回车后显示:

```
y1=
    -1/36
m=
    -1/11
```

结果表明: 随机变量(X, Y)的协方差为 Cov(X,Y)=-1/36, 相关系数 ρ =-1/11.

三、实践结论与总结

随机变量和样本的数字特征是由随机变量的分布确定的, 是描述随机变量和样本某一个方面的特征的常数. 本实践主要用 MATLAB 工具箱提供的计算随机变量和样本的数字特征命令, 如: EX=sum(X.*p), DX=var(X,1), cov(X,Y), corrcoef(X,Y)等计算随机变量和样本的数学期望、方差、协方差、相关系数, 计算起来方便、快捷和准确.