统计量计算与分析

统计量分析

- 随机变量的均值与方差
- 随机变量的矩
- 多变量随机数的协方差分析
- 多变量正态分布的联合概率密度及分布函数

随机变量的均值与方差

- 连续随机变量 x 的概率密度函数为 $p(x)$
- 数学期望 $\mathrm{E}[x]$：

$$\mathrm{E}[x] = \int_{-\infty}^{\infty} x p(x) \, \mathrm{d}x$$

- 数学方差 $\mathrm{D}[x]$：

$$\mathrm{D}[x] = \int_{-\infty}^{\infty} (x - \mathrm{E}[x])^2 \, p(x) \, \mathrm{d}x$$

例9-18 统计量计算

- 用积分方法求取Gamma分布(a>0,λ>0)的均值与方差
- Gamma分布的概率密度

$$p(x)=\begin{cases}\dfrac{\lambda^a x^{a-1}}{\Gamma(a)}\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

- MATLAB求解语句：

```
>> syms x; syms a lam positive
   p=lam^a*x^(a-1)/gamma(a)*...
       exp(-lam*x); m=int(x*p,x,0,inf)
   s=simplify(int((x-1/lam*a)^2*p,x,0,inf))
```

- 结果：
  - $m=\dfrac{a}{\lambda}$ 和 $s=\dfrac{a}{\lambda^2}$

统计量数值计算

- 在实际中测出一组样本数据 $\boldsymbol{x} = [x_1, x_2, x_3, \cdots, x_n]^{\mathrm{T}}$
- 它们的均值和方差分别为——样本均值方差

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \hat{s}_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

m=mean($\boldsymbol{x}$)

s2=var($\boldsymbol{x}$)

- 无偏的方差

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

s=std($\boldsymbol{x}$)

- 称 $s_x \geqslant 0$ 为“标准差”

例9-22 伪随机数的统计量

- 生成一组30000个正态分布随机数
  - 均值为 0.5，标准差为 1.5
  - 分析数据实际的均值、方差和标准差
  - 减小随机变量个数有何影响
- MATLAB求解语句：

```
>> p=normrnd(0.5,1.5,30000,1);
   mean(p), var(p), std(p)
```

- 使用300个随机数:

```
>> p=normrnd(0.5,1.5,300,1);
   mean(p), var(p), std(p)
```

随机变量的矩

- 假设 x 为连续随机变量，且 $p(x)$ 为其概率密度函数
  - 该变量的 r 阶原点矩

$$\nu_r = \int_{-\infty}^{\infty} x^r p(x)\,\mathrm{d}x$$

  - r 阶中心矩

$$\mu_r = \int_{-\infty}^{\infty} (x-\mu)^r p(x)\,\mathrm{d}x$$

- 可见， $\nu_1 = \mathrm{E}[x],\ \mu_2 = \mathrm{D}[x]$

例9-21 Gamma分布的矩量计算

- 考虑Gamma分布(a>0，λ>0)的原点矩和中心矩
  - 由前几项结果总结一般规律
- MATLAB求解命令

```
>> syms x; syms a lam positive;
p=lam^a*x^(a-1)/gamma(a)*exp(-lam*x);
for n=1:5, m=int(x^n*p,x,0,inf), end
```

- 原点矩通项表达式

$$\nu_k = \frac{1}{\lambda^k}\prod_{i=0}^{k-1}(a+i)$$

矩量的解析推导

- 直接求出任意阶次的原点矩

```
>> syms k;
   m=simplify(int((x)^k*p,x,0,inf))
```

- 计算原问题的中心矩

```
>> for n=1:7,
      s=simplify(int((x-1/lam*a)^n*p,x,0,inf)),
   end
```

- 中心矩没有显而易见的通项公式

矩的数值计算

- 给定的随机数为一些样本点 $x_1, x_2, \cdots, x_n$
- 该随机变量的 r 阶原点矩

$$A_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r$$

A_r=sum(x.^r)/length(x)

- 该随机变量的 r 阶中心矩

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

B_r=moment(x,r)

例9-22 矩量的数值计算

- 给生成一组30000个正态分布随机数
  - 均值为 0.5，标准差为 1.5，试求出随机数的各阶矩
- MATLAB求解命令：

```
>> A=[]; B=[]; n=1:5; p=normrnd(0.5,1.5,30000,1);
   for r=n,
      A=[A, sum(p.^r)/length(p)]; B=[B,moment(p,r)];
   end, A, B
```

```
>> syms x; A1=[]; B1=[]; a=-inf; b=inf;
   p=1/(sqrt(2*sym(pi))*3/2)*exp(-(x-1/2)^2/(2*(3/2)^2));
   for i=1:5,
      A1=[A1,int(x^i*p,x,a,b)]; B1=[B1,int((x-1/2)^i*p,x,a,b)];
   end, A1, B1
```