深度之眼
deepshare.net

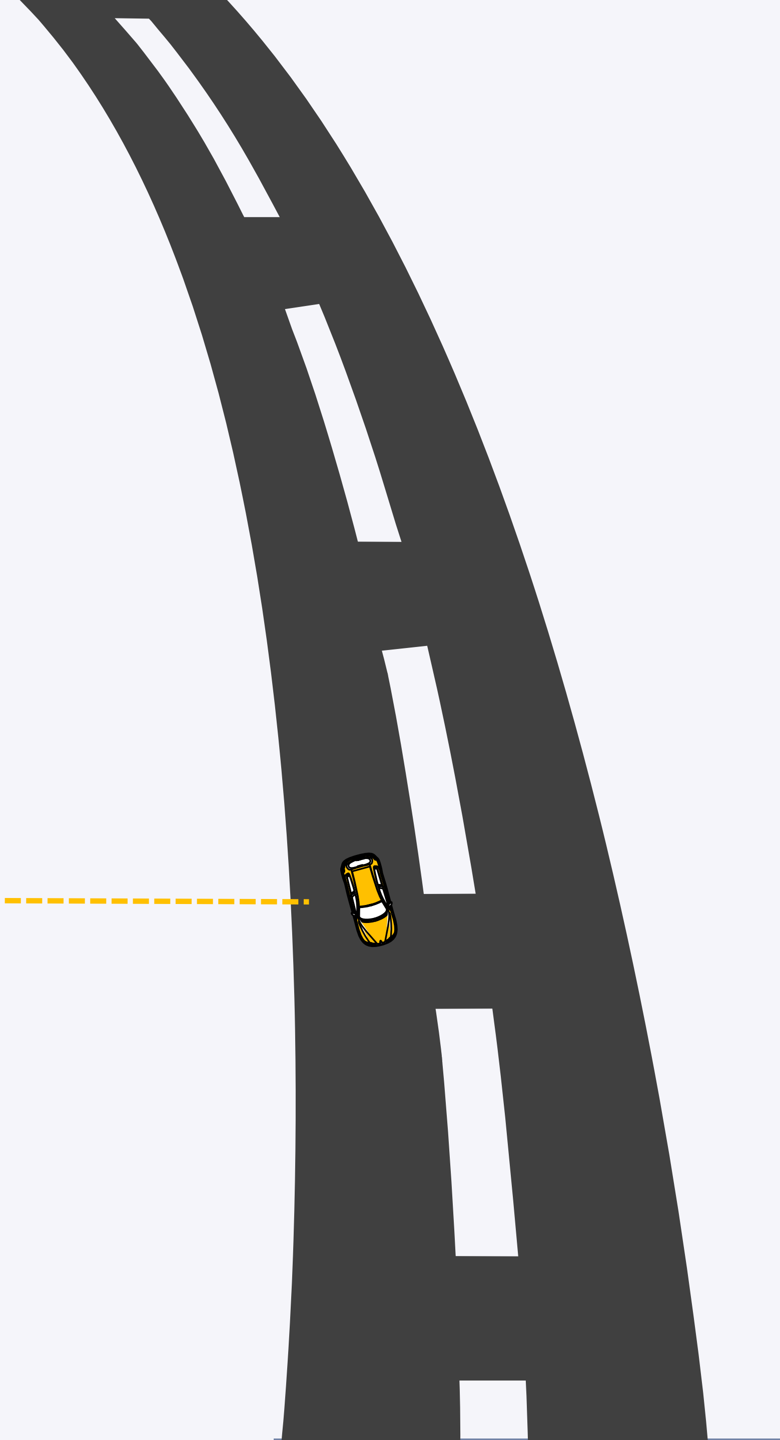
数学基础—概率论

导师：Johnson

主要内容

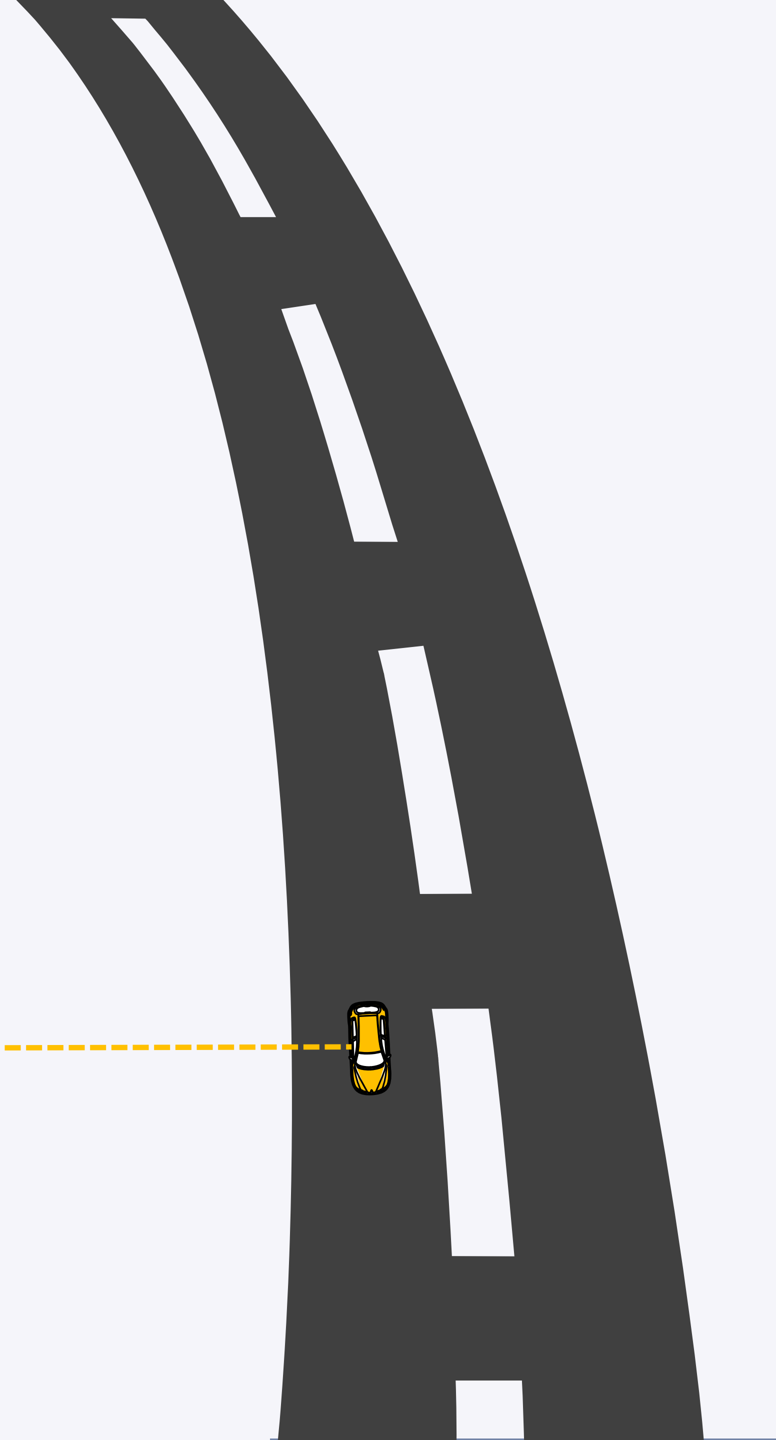
概率论

- **概率基础**
  - 1.随机试验样本空间和随机事件
  - 2.概率的定义
  - 3.条件概率与乘法公式
  - 4.全概率公式与贝叶斯公式以及应用
  - 5.独立性
- **随机变量与多维随机变量**
  - 1.离散随机变量
  - 2.连续随机变量
  - 3.多维随机变量
  - 4.边缘分布和条件分布
  - 5.独立性
- **期望与方差**
  - 1.期望与方差的定义
  - 2.常用随机变量期望与方差的计算
  - 3.期望与方差的性质
  - 4.协方差与协方差矩阵
- **参数的估计**
  - 1.矩估计
  - 2.极大似然估计
  - 3.最大后验估计



主要内容

- 概率论
  - 概率基础
    - 1.随机试验样本空间和随机事件
    - 2.概率的定义
    - 3.条件概率与乘法公式
    - 4.全概率公式与贝叶斯公式以及应用
    - 5.独立性
  - 随机变量与多维随机变量
    - 1.离散随机变量
    - 2.连续随机变量
    - 3.多维随机变量
    - 4.边缘分布和条件分布
    - 5.独立性
  - 期望与方差
    - 1.期望与方差的定义
    - 2.常用随机变量期望与方差的计算
    - 3.期望与方差的性质
    - 4.协方差与协方差矩阵
  - 参数的估计
    - 1.矩估计
    - 2.极大似然估计
    - 3.最大后验估计



期望与方差的定义

期望

连续型： $E(X)=\int_{-\infty}^{\infty} x f(x) d x$

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

数学期望简称期望，又称为均值.

期望与方差的定义

期望

例1 某医院当新生儿诞生时，医生要根据婴儿的皮肤颜色、肌肉弹性、反应的敏感性、心脏的搏动等方面的情况进行评分，新生儿的得分X是一个随机变量。据以往的资料表明X的分布律为

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
$$\begin{aligned} E(X) &= 0 \times 0.002 + 1 \times 0.001 + 2 \times 0.002 + 3 \times 0.005 + 4 \times 0.02 \\ &\quad + 5 \times 0.04 + 6 \times 0.18 + 7 \times 0.37 + 8 \times 0.25 + 9 \times 0.12 + 10 \times 0.01 \\ &= 7.15(\text{分}) \end{aligned}$$

这意味着，若考察医院出生的很多新生儿，例如1000个，那么一个新生儿的平均得分约7.15分，1000个新生儿共得分约7150分.

期望与方差的定义

方差

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

来度量随机变量X与其均值$E(X)$的偏离程度.

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

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

在应用上还引入量$\sqrt{D(X)}$，记为$\sigma(X)$，称为**标准差**或**均方差**.

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

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

$$\begin{aligned}D(X)&=E\{[X-E(X)]^2\}=E\{X^2-2XE(X)+[E(X)]^2\}\\&=E(X^2)-2E(X)E(X)+[E(X)]^2\\&=E(X^2)-[E(X)]^2.\end{aligned}$$

常用随机变量期望与方差的计算

深度之眼
deepshare.net

0-1分布

X	0	1
p_k	1-p	p

$$E(X) = 0 \times (1-p) + 1 \times p = p$$

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

$$= 0^2 \times (1-p) + 1^2 \times p - p^2 = p(1-p)$$

常用随机变量期望与方差的计算

深度之眼 deepshare.net

二项分布

$$P\{X=k\}=C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n.$$

$$kC_n^k=k\frac{n!}{k!\,(n-k)!}=\frac{n(n-1)!}{(k-1)!\,(n-k)!}=nC_{n-1}^{k-1}\quad 当k\geq 1$$

$$E(X)=\sum_{k=0}^{n}kC_n^k p^k q^{n-k}=\sum_{k=1}^{n}nC_{n-1}^{k-1}p^k q^{n-k}$$

令$i=k-1$

$$=\quad n\sum_{i=0}^{n-1}C_{n-1}^{i}p^{i+1}q^{n-1-i}=np\sum_{i=0}^{n-1}C_{n-1}^{i}p^{i}q^{n-1-i}$$

$$=np(p+q)^{n-1}=np$$

常用随机变量期望与方差的计算

深度之眼 deepshare.net

二项分布

$P\{X=k\}=C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n.\quad E(X)=np$

$$k(k-1)C_n^k=k(k-1)\frac{n!}{k!\,(n-k)!}=\frac{n(n-1)(n-2)!}{(k-2)!\,(n-k)!}=n(n-1)C_{n-2}^{k-2}\quad \text{当} k\geq 2$$

$$E(X^2)=\sum_{k=0}^{n}k^2C_n^k p^k q^{n-k}=\sum_{k=0}^{n}k(k-1)C_n^k p^k q^{n-k}+\sum_{k=0}^{n}kC_n^k p^k q^{n-k}$$

$$=\sum_{k=2}^{n}n(n-1)\,C_{n-2}^{k-2}p^k q^{n-k}+np\qquad \text{令} i=k-2$$

$$=\sum_{i=0}^{n-2}n(n-1)\,C_{n-2}^{i}p^{i+2}q^{n-2-i}+np$$

$$=n(n-1)p^2\sum_{i=0}^{n-2}C_{n-2}^{i}p^i q^{n-2-i}+np=n(n-1)p^2+np$$

常用随机变量期望与方差的计算

二项分布

$P\{X=k\}=C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n.$

$E(X)=np$

$E(X^2)=n(n-1)p^2+np$

$$\begin{aligned} D(X)&=E(X^2)-E^2(X)=n(n-1)p^2+np-n^2p^2\\ &=np-np^2=np(1-p)\end{aligned}$$

常用随机变量期望与方差的计算

深度之眼
deepshare.net

泊松分布

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},k=0,1,2,\cdots,n.$$

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

$$E(X^2)=E[X(X-1)+X]=E[X(X-1)]+E(X)$$

$$=\sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k e^{-\lambda}}{k!}+\lambda=\lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda$$

$$=\lambda^2 e^{-\lambda}e^{\lambda}+\lambda=\lambda^2+\lambda,$$

所以方差

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

常用随机变量期望与方差的计算

深度之眼
deepshare.net

均匀分布

$$f(x)=\begin{cases}\dfrac{1}{b-a}, a<x<b, \\ 0, \qquad\qquad \text{其他},\end{cases}$$

$$E(X)=\int_{-\infty}^{\infty} xf(x)\,dx=\int_{a}^{b}\frac{x}{b-a}\,dx=\frac{a+b}{2}$$

$$\begin{aligned}D(X)&=E(X^2)-[E(X)]^2\\&=\int_{a}^{b}x^2\frac{1}{b-a}\,dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}\end{aligned}$$

常用随机变量期望与方差的计算

深度之眼
deepshare.net

指数分布

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

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_{0}^{\infty}x\frac{1}{\theta}e^{-x/\theta}dx$$

$$=-xe^{-x/\theta}\Big|_{0}^{\infty}+\int_{0}^{\infty}e^{-x/\theta}dx=\theta,$$

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)dx=\int_{0}^{\infty}x^2\frac{1}{\theta}e^{-x/\theta}dx$$

$$=-x^2e^{-x/\theta}\Big|_{0}^{\infty}+\int_{0}^{\infty}2xe^{-x/\theta}dx=2\theta^2,$$

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

常用随机变量期望与方差的计算

深度之眼
deepshare.net

正态分布

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

令$t = \frac{x-\mu}{\sigma}$

$\Rightarrow dx = \sigma dt$

$$= \int_{-\infty}^{\infty} (\sigma t + \mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} \sigma\, dt$$

$$= \int_{-\infty}^{\infty} \sigma t \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} \sigma dt + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= 0 + \mu = \mu$$

$$D(X) = \sigma^2$$

期望与方差的性质

期望的性质

1. 设C是常数，则有$E(C)=C$.

2. 设X是一个随机变量，C是常数，则有

$$E(CX)=CE(X).$$

3. 设X,Y是两个随机变量，则有

$$E(X+Y)=E(X)+E(Y).$$

这一性质可以推广到任意有限个随机变量之和的情况.

4. 设X,Y是相互独立的随机变量，则有

$$E(XY)=E(X)E(Y).$$

这一性质可以推广到任意有限个相互独立的随机变量之积的情况.

期望与方差的性质

方差的性质

1. 设C是常数，则$D(C)=0$.

2. 设X是随机变量，C是常量，则有

$$D(CX)=C^2D(X), D(X+C)=D(X).$$

3. 设X,Y是两个随机变量，则有

$$D(X+Y)=D(X)+D(Y)+2E\{(X-E(X))(Y-E(Y))\}.$$

特别，若X,Y相互独立，则有

$$D(X+Y)=D(X)+D(Y).$$

这一性质可以推广到任意有限多个相互独立的随机变量之和的情况.

4. $D(X)=0$的充要条件是X以概率1取常数$E(X)$，即

$$P\{X=E(X)\}=1.$$

期望与方差的性质

方差的性质

1. 设C是常数，则$D(C)=0$.

证　$D(C)=E\{[C-E(C)]^2\}=0$.

2. 设X是随机变量，C是常量，则有

$$D(CX)=C^2D(X), D(X+C)=D(X).$$

证　$D(CX)=E\{[CX-E(CX)]^2\}=C^2E\{[X-E(X)]^2\}=C^2D(X)$.

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

期望与方差的性质

方差的性质

3. 设X,Y是两个随机变量，则有

$$D(X+Y)=D(X)+D(Y)+2E\{(X-E(X))(Y-E(Y))\}.$$

特别，若X,Y相互独立，则有

$$D(X+Y)=D(X)+D(Y).$$

这一性质可以推广到任意有限多个相互独立的随机变量之和的情况.

证 $D(X+Y)=E\{[(X+Y)-E(X+Y)]^2\}=E\left\{\left[\left(X-E(X)\right)+\left(Y-E(Y)\right)\right]^2\right\}$

$$=E\left\{\left(X-E(X)\right)^2\right\}+E\left\{\left(Y-E(Y)\right)^2\right\}+2E\{[X-E(X)][Y-E(Y)]\}$$

$$=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}.$$

$2E\{[X-E(X)][Y-E(Y)]\}$

$=2\mathrm{E}\{XY-XE(Y)-YE(X)+E(X)E(Y)\}$

$=2\{E(XY)-E(X)E(Y)-E(Y)E(X)+E(X)E(Y)\}$

$=2\{E(XY)-E(X)E(Y)\}.$

协方差与协方差矩阵

在本章方差性质3的证明中，我们已经看到，如果两个随机变量X和Y是相互独立的，则

$$E\{[X-E(X)][Y-E(Y)]\}=0.$$

这意味着当$E\{[X-E(X)][Y-E(Y)]\}\neq 0$时，$X$与$Y$不相互独立，而是存在着一定的关系的.

定义 量$E\{[X-E(X)][Y-E(Y)]\}$称为随机变量X与Y的**协方差**.记为$Cov(X,Y)$，即

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

而

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

称为随机变量X与Y的**相关系数**

协方差与协方差矩阵

$Cov(X, Y) = Cov(Y.X)$
$Cov(X, X) = D(X)$
$D(X + Y) = D(X) + D(Y) + 2Cov(X, Y)$
$Cov(X, Y) = E(XY) - E(X)E(Y)$

协方差具有下述性质：

1. $Cov(aX, bY) = abCov(X, Y), a, b$是常数.
2. $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$.

定理 1. $|\rho_{XY}| \leq 1$.

2. $|\rho_{XY}| = 1$的充要条件是，存在常数a, b使

$$P\{Y = a + bX\} = 1.$$

协方差与协方差矩阵

二维随机变量(X_1, X_2)有四个二阶中心矩（设它们都存在），分别记为

$$c_{11} = E\{[X_1 - E(X_1)]^2\},$$
$$c_{12} = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\},$$
$$c_{21} = E\{[X_2 - E(X_2)][X_1 - E(X_1)]\},$$
$$c_{22} = E\{[X_2 - E(X_2)]^2\}.$$

将它们排成矩形的形式

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

这个矩阵称为随机变量(X_1, X_2)的协方差矩阵.

协方差与协方差矩阵

设n维随机变量$(X_1, X_2, \cdots, X_n)$的二阶混合中心矩

$$c_{ij} = Cov(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}, i, j = 1,2, \cdots, n$$

都存在，则称矩阵

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为n维随机变量$(X_1, X_2, \cdots, X_n)$的**协方差矩阵**.由于$c_{ij} = c_{ji} (i \neq j; i, j = 1,2, \cdots, n)$,
因此上述矩阵是一个对称矩阵.

协方差与协方差矩阵

引入列矩阵

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\text{和}\mu=\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}=\begin{pmatrix}E(X_1)\\E(X_2)\\\vdots\\E(X_n)\end{pmatrix}.$$

n维正态随机变量$(X_1,X_2,\cdots,X_n)$的概念密度定义为

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{n/2}(\det C)^{1/2}}\exp\left\{-\frac{1}{2}(X-\mu)^TC^{-1}(X-\mu)\right\},$$

其中C是$(X_1,X_2,\cdots,X_n)$的协方差矩阵.

深度之眼
deepshare.net

联系我们：

电话：18001992849

邮箱：service@deepshare.net

Q Q： 2677693114

公众号

客服微信