第 4 讲*

数学期望（矩）

麻省理工学院 14.30 2006 年春季

Herman Bennett

7 数学期望

7.1 单变量模型

设 X 随机变量的概率质量/密度函数为 $f(x)$。X 的数学期望值和均值记作 $E(x)$ 和 μ_X，定义为：

$$E(X)=\mu_X=\sum_{x\in X}xf(x) \qquad \text{（离散型）}$$

$$E(X)=\mu_X=\int_{-\infty}^{\infty}xf(x)\,\mathrm{dx} \qquad \text{（连续型）} \qquad \text{（18）}$$

- 直观认识：集中趋势（一种分布的“平均”）
- 计算方法：加权平均

如果 $Z=z(X)$ 是一个新的随机变量，定义为随机变量 X 的函数（变换），则：

$$E[Z]=E[z(X)]=\mu_z=\sum_{x\in X}z(x)f(x) \qquad \text{（离散型）}$$

$$E[Z]=E[z(X)]=\mu_z=\int_{-\infty}^{\infty}z(x)f(x)dx \qquad \text{（连续型）} \qquad \text{（19）}$$

注意：这些讲义不一定是自封的。它们只是对讲座的一种补充而不是替代。

例 7.1　a）求 $E(X)$ 和 $E(X^2)$，这里 X 是随机变量，表示掷骰子的结果。b ）计算 $E(Z)$ 和 $E(X)$，这里，如果 $0<x<1$，随机变量 X 的概率密度函数为 $f(x)=2x$，否则，$Z=\sqrt{X}$ 。

- 均值与中位数

7．2 二维模型

设 (X,Y) 是联合密度函数为 $f_{XY}(x,y)$ 的随机向量。随机变量 $Z=z(X,Y)$ 的数学期望或均值为：

$$E(Z)=\sum_{(x,y)\in(X,Y)} z(x,y)\,f(x,y) \qquad E(Z)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z(x,y)\,f(x,y)\,dxdy \qquad (20)$$

- 二维以上的随机变量相关定义类似（见讲义第 3 讲最后的多维分布部分）

例 7.2 计算 $E(Z)$，其中，如果 $0<x,y<1$，则 $f(x,y)=1$，否则为 0，且 $Z=X^2+Y^2$。

7．3 数学期望的性质

设 $Y, X_1, X_2, \ldots X_n$ 为随机变量，a，b，c，d 为常数。那么，

a. $E[aX+b]=aE(X)+b$ 且 $E[az(X)+b]=aE[z(X)]+b$。

b. $E(aX_1+bX_2+\ldots cX_n+d)=aE(X_1)+bE(X_2)+\ldots cE(X_n)+d$

c. X 和 Y 是相互独立的随机变量→ $E(XY)=E(X)E(Y)$

例 7.3 证明 a 和 c。（作业：证明 b）

- $E[z(X)] \stackrel{?}{=} z(E[X])$ （詹生不等式）

8 方差

随机变量 X 的方差记作$Var(X)$或σ_X^2，定义为：

$$Var(X)=\sigma_X^2=E\left[\left(X-\mu_X\right)^2\right],\ \mu_X=E(X)。\tag{21}$$

- 标准差，${}_X=\sqrt{\sigma_X^2}$

8．1 方差的性质

设$X_1,X_2,\ldots X_n$为随机变量，a，b，c，d 为常数。那么，

a. $Var(X)=0\leftrightarrow\exists c.s.t\ P(X=c)=1$(退化分布)

b. $Var(X)=E(X^2)-\left[E(X)\right]^2$。

c. $Var(aX+b)=a^2Var(X)$。

d. 如果$X_1,X_2,\ldots X_n$是相互独立的随机变量，那么

$$Var(aX_1+bX_2+\ldots cX_n+d)=a^2Var(X_1)+b^2Var(X_2)+\ldots c^2Var(X_n)$$

例 8.1 证明 b 和 c。(作业：证明在二维随机变量条件下 d 结论也成立。)

例 8.2 计算$Var(X)$和$Var(Y)$，其中，若$x=-2,0,1,3,4$，$f(x)=1/5$；否则，为 0，且$Y=4X-7$。

例 8.3 如果 Y～二项分布 bin(n,p)，计算 $Var(Y)$。

9 协方差与相关系数

设X和Y为两个随机变量。X和Y的协方差记作 $Cov(X,Y)$，定义为

$$Cov(X,Y)=E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right] \tag{22}$$

- X 与 Y 的相关系数为：$Corr(X,Y)=\rho(X,Y)=\dfrac{Cov(X,Y)}{\sigma_X\sigma_Y}$（$Cov(X,Y)$ 的标准化定义版本）。

9．1 协方差与相关系数的性质

设 X 和 Y 为随机变量，a，b，c，d 为常数。那么

a. $Cov(X,Y)=Var(X)$。

b. $Cov(X,Y)=Cov(Y,X)$

c. $Cov(X,Y)=E(XY)-E(X)E(Y)$

d. X 和 Y 相互独立→$Cov(X,Y)=0$。

e. $Cov(aX+b)=acCov(X,Y)$。

f. $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$。

g. $\rho(X,Y)=\begin{cases} >0 \text{“正相关”} \\ =0 \text{“不相关”} \\ <0 \text{“负相关”} \end{cases}$

h. $\left|\rho(X,Y)\right|\leq 1$。

i. $\left|\rho(X,Y)\right|=1$ ，当且仅当$Y=aX+b$，$a\neq 0$。

例 9.1 证明 c，d 和 f。

例 9.2 计算 $Cov(X,Y)$ 和 $\rho(X,Y)$，其中：若 $0 \le x \le y \le 1$，$f(x,y) = 8xy$，否则，为 0。

10 条件数学期望和条件方差

设(X,Y)是条件概率密度函数为$f(y|x)$的随机向量。<u>给定$X = x$时Y的条件数学期望</u>记作$E(Y|X = x)$，定义为：

$$E(Y|X = x) = \sum_{y\in\Re} yf(y|x) \qquad E(Y|X = x) = \int_{-\infty}^{\infty} yf(y|x)dy \qquad (23)$$

（离散型） （连续型）

例 10.1 计算$E(Y|X = x)$，其中，对于$0 \le x \le y \le \infty$，$f(x,y) = e^{-y}$，否则，为 0。

<u>全数学期望法则</u>：设(X,Y)为随机向量。那么：

$$E\left[E(Y|X)\right] = E(Y) \qquad (24)$$

例 10.2 证明公式（24）

<u>条件方差的一致性</u>：对任意两个随机变量X和Y，X的方差可分解如下：

$$Var(X)=E\left[Var(X|Y)\right]+Var\left[E(Y|X)\right] \qquad (25)$$

例 10.3 某研发公司每年开发 N 项技术创新是一个随机过程，其中 $E(N)=2$，$Var(N)=1$。每一项创新成功应用于商业的概率为 0.2，且这个概率与过去的创新绩效相互独立。a）如果今年有 5 项创新，那么成功应用项目的数量的概率密度函数及其数学期望是多少？b）在知道开发的创新项目数量之前成功应用项目数量的数学期望是多少？c）在知道开发的创新项目数量之前，成功应用项目数量的方差是多少？

11 矩和矩母函数

11．1 矩

设 X 是一个连续随机变量，则矩 $E\left[g\left(X\right)\right]$ 为：

$$E\left[g\left(X\right)\right]=\int_{-\infty}^{\infty}g\left(x\right)f\left(x\right)dx \;\to\; \text{即 } g\left(X\right) \text{ 的数学期望。} \tag{26}$$

- 离散变量也类似。
- 例如，平均值反映矩的特征，这里 $g\left(X\right)=X$ 。
- X 的 n 阶矩记为 $E\left[X^{n}\right]$，它意味着 $g\left(X\right)=X^{n}$ 。

-偏态

-峰态

11．2 矩母函数

设 X 是一个随机变量。X 的矩母函数记作 $M_{X}\left(t\right)$，定义为

$$M_{X}\left(t\right)=E\left[e^{tX}\right] \tag{27}$$

满足如下性质：

$$M_{X}^{(n)}\left(t\right)=\frac{d^{n}M_{X}\left(t\right)}{dt^{n}}\Big|_{t=0}=E\left[X^{n}\right] \tag{28}$$

例 11.1 证明（28）式，并用矩母函数计算二项分布 $b\left(n,p\right)$ 的数学期望和方差。

12 不等式

12.1 马可夫不等式

设 X 为随机变量，且 $P(X \geq 0)=1$。那么，对于任意数 t>0，有

$$P\left(X \geq t\right) \leq \frac{E\left(X\right)}{t} \tag{29}$$

12.2 切比雪夫不等式

对任意随机变量 X，若它的方差 $Var\left(X\right)$ 存在，则对任意数 $t>0$，有

$$P\left(\left|X-E\left(X\right)\right| \geq t\right) \leq \frac{Var\left(X\right)}{t^2} \tag{30}$$