随机过程及应用

Stochastic Processes with its Applications

电子科技大学——数学科学学院

武德安

附录预备知识

§1 Riemann-Stieltjes $(R-S)$积分简介

§2 随机变量的数字特征

§3 特征函数

一. 数学期望与方差

定义1.4.2： 随机变量X的分布函数为$F(x)$,若$\int_{-\infty}^{+\infty}|x|dF(x)<+\infty$,则

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

注： 若X是**连续型随机变量**,则

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

若X是**离散型随机变量**,则

$$E(X)=\sum_{i=1}^{\infty}x_iP\{X=x_i\}$$

定理1.4.2： 设$F(x)$是随机变量X的分布函数，$g(x)$在R上连续，

若 $\int_{-\infty}^{+\infty}|g(x)|dF(x)<\infty$，则$Y=g(X)$的数学期望存在，且

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

当随机变量X是**连续型随机变量**，则

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

当随机变量X是**离散型随机变量**，则

$$E[g(X)]=\int_{-\infty}^{+\infty}g(x)dF(x)=\sum_{k=-\infty}^{\infty}g(x_k)p_k$$

定义1.4.3：若随机变量X的分布函数为$F(x)$,以下积分存在

$$D(X) \triangleq \int_{-\infty}^{+\infty} [x - E(X)]^2 dF(x)$$

称为随机变量X的**方差**.

二、条件数学期望

1.条件数学期望概念

定义1.4.4: 设(X,Y)是二维随机变量,条件分布函数$F_{Y|X}(y|x)$存在，又若

$\int_{-\infty}^{+\infty} y|y|dF_{Y|X}(y|x) < \infty$称

$$E(Y|x) = E(Y|X=x) \triangleq \int_{-\infty}^{+\infty} y\,dF_{Y|X}(y|x)$$

为在$X=x$的条件下,随机变量Y的**条件数学期望**.

若(X,Y)是连续型随机变量,则(直观意义)

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

Eg.1: 若$(X,Y)\sim N(0,1;0,1;\rho),(-1<\rho<1)$,在$X=x$的条件下,$Y\sim N\left(\rho x,\sqrt{1-\rho^2}\right)$;

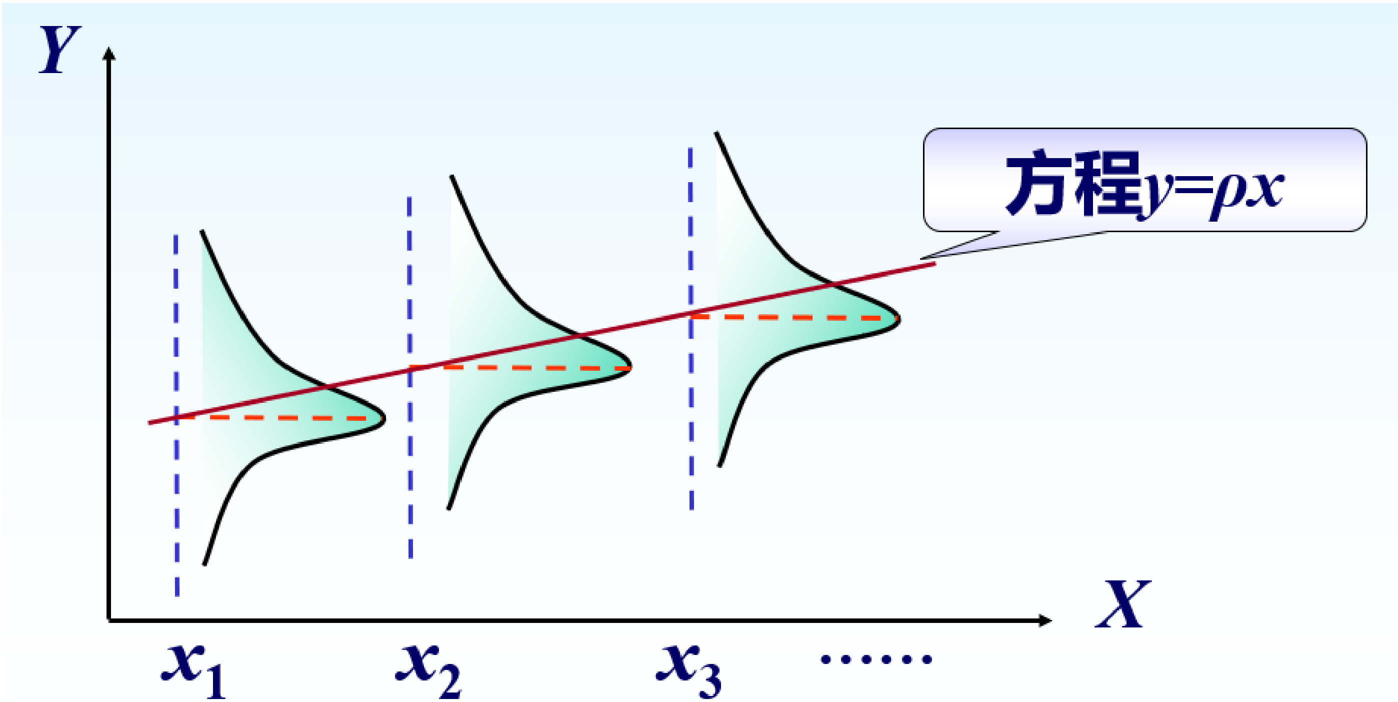
因$f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\dfrac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}e^{-\frac{1}{2}\left[\frac{y-\rho x}{\sqrt{1-\rho^2}}\right]^2}\ y\in R.$

则$E(Y|X=x)=\displaystyle\int_{-\infty}^{+\infty}y\,f_{Y|X}(y|x)\,dy=\rho x\Longleftarrow$ $\boxed{\text{实值函数}}$

同理 $E(X|Y=y)=\rho y\Longleftarrow$ $\boxed{\text{实值函数}}$

在$X=x$的条件下,$Y\sim N\left(\rho x,\sqrt{1-\rho^2}\right)$;$E(Y\mid X=x)=\int_{-\infty}^{+\infty} y f_{Y\mid X}(y\mid x)\,dy=\rho x$

对于X的不同取值$x_1,x_2,\ldots,x_n$



Eg.2 设随机变量(X,Y)的联合概率密度为$f(x,y)=\begin{cases}\frac{1}{2}e^{-y}, & y>|x|, -\infty<x<\infty;\\ 0, & \text{其它}.\end{cases}$

试求$E(Y|X=x)$.

解: $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\frac{1}{2}e^{-|x|}, \quad x\in R$;在"$X=x$"的条件下,有条件概率密度

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}e^{-y+|x|}, & y>|x|;\\ 0, & \text{其它}.\end{cases}$$

$$E(Y|X=x)=\int_{-\infty}^{+\infty} y f_{Y|X}(y|x)dy=\int_{|x|}^{\infty} y\mathrm{e}^{-y+|x|}dy=1+|x| \Longleftarrow \boxed{\text{实值函数}}$$

注: 一般有$E(Y|X=x)=\mu(x)\Leftarrow$ $\boxed{Y\text{关于}X\text{的回归函数}}$, $E(X|Y=y)=\delta(y)\Leftarrow$ $\boxed{X\text{关于}Y\text{的回归函数}}$

定理0.4.4: 设函数$g(x)$在R上连续,若$\int_{-\infty}^{+\infty}|g(x)|dF_{X|Y}(x|y)<+\infty$则随机变量$g(X)$在"$Y=y$"条件下的条件数期望为

$$E[g(X)|Y=y]=\int_{-\infty}^{+\infty}g(x)dF_{X|Y}(x|y)$$

定义0.4.4: 称$D(X|Y=y)=E_X[X-E(X|Y=y)]^2=\int_{-\infty}^{+\infty}[x-E(X|Y=y)]^2dF_{X|Y}(x|y)$

为"$Y=y$"的条件下,随机变量X的**条件方差**.

注: 为随机变量X相对于条件数学期望$E(X|Y=y)$的偏离程度的衡量指标.

定理0.4.2: 设(X,Y)是二维随机变量,若函数$g(y)$在R上连续,若$\int_{-\infty}^{+\infty}|g(y)|dF_{Y|X}(y|x)<\infty$

$$E(g(Y)|x)=E(g(Y)|X=x)=\int_{-\infty}^{+\infty}g(y)dF_{Y|X}(y|x)$$

为在$X=x$的条件下,随机变量$g(Y)$的**条件数学期望**.

注: $D(X|Y=y)=E_X[X-E(X|Y=y)]^2=E\{[X-E(X|Y=y)]^2|Y=y\}$

$$D(Y|X=x)=\int_{-\infty}^{+\infty}[y-E(Y|X=x)]^2dF_{Y|X}(y|x)$$

2 条件数学期望性质

一般$E(Y|X=x)=\mu(x)$, $E(X|Y=y)=\delta(y)$ 是**实值函数**,

可以证明随机变量的函数$\mu(X)=E(Y|X)$, $\delta(Y)=E(X|Y)$仍是**随机变量**.

定理0.4.4(性质): 设X,Y,Z是(Ω,F,P)上的随机变量,$g(\cdot)$和$h(\cdot)$为R上连续函数,

且各数学期望存在

1) $E(c|Y)=c$是常数;

2) $E[aX+bY|Z]=aE(X|Z)+bE(Y|Z),a,b$ 是常数,

3) 如果X与Y相互独立,则$E(X|Y)=E(X)$.

4) $E[g(X)h(Y)|X]=g(X)E[h(Y)|X]$; $E[g(X)h(Y)|Y]=h(Y)E[g(X)|Y]$

5) $E\{E[g(X,Y)|Y]\}=E[g(X,Y)]$

6) $E[X-E(X|Y)]^2\leq E[X-g(Y)]^2$.

Eg.3: 已知二维随机变量(X,Y)在$\{(x,y):0\le y\le x\le 1\}$上服从均匀分布,求在$Y=y$条件下,随机变量$X$的条件数学期望$E(X\mid y)$,$E(X\mid Y)$以及其分布函数.

解: 因$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx=\int_y^1 2\,dx=\begin{cases}2(1-y), & 0\le y\le 1\\ 0, & \text{其他}.\end{cases}$

故当$0\le y<1$时

$$f_{X\mid Y}(x\mid y)=\frac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{1}{1-y}, & y<x<1\\ 0, & \text{其他}.\end{cases}$$

即$X\sim U(y,1)$,有$E(X\mid y)=\int_{-\infty}^{+\infty} x\,f_{X\mid Y}(x\mid y)\,dx=\dfrac{1+y}{2}$从而随机变量 $Z=\delta(Y)=E(X\mid Y)=\dfrac{1+Y}{2}$

定理0.4.4(性质) 设X,Y,Z是(Ω,F,P)上的随机变量,$g(\cdot)$和$h(\cdot)$为R上连续函数,且各数学期望存在,有

1) $E(c|Y)=c$, c是常数;

证: 1) 对$\forall y$, $E(c|Y=y)=\int_{-\infty}^{+\infty} c\,dF_{X|Y}(x|y)=c$, $\therefore E(c|Y)=c.$

2) $E[aX+bY|Z]=aE(X|Z)+bE(Y|Z)$, a,b是常数. 自证(有限的线性性质).

3) 如果X与Y相互独立,则$E(X|Y)=E(X)$.

证: X与Y独立, $\Rightarrow F_{X|Y}(x|y)=F_X(x)$, $\forall y\in R$. 对$\forall y\in R$,

$$E(X|y)=\int_{-\infty}^{+\infty} x\,dF_{X|Y}(x|y)=\int_{-\infty}^{+\infty} x\,dF_X(x)=E(X)\Rightarrow E(X|Y)=E(X)$$

4) $E[g(X)h(Y)|X]=g(X)E[h(Y)|X];E[g(X)h(Y)|Y]=h(Y)E[g(X)|Y]$

(参见教材P236)特别地:$g(X)=1$,则$E[h(Y)|Y]=h(Y)$.

5)* $E\{E[g(X,Y)|Y]\}=E[g(X,Y)]$

记$E[g(X,Y)|Y]=S(Y)$

则$S(y)=E[g(X,Y)|Y=y]=E[g(X,y)|y]$

$$E\{E[g(X,Y)|Y]\}=E[S(Y)]=\int_{-\infty}^{+\infty}S(y)\mathrm{d}F_Y(y)$$

$$=\int_{-\infty}^{+\infty}E[g(X,y)|y]dF_Y(y)=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}g(x,y)dF_{X|Y}(x|y)\right]dF_Y(y)$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)[dF_{X|Y}(x|y)dF_Y(y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)dF(x,y)=E[g(X,Y)]$$

6) $E[X-E(X|Y)]^2 \le E[X-g(Y)]^2$.或$E\{[X-E(X|Y)]^2|Y\} \le E\{[X-g(Y)]^2|Y\}$

类似方差性质: $E\{X-E(X)]^2\} \le E\{X-g(x)]^2\}$

证明: 函数$f(x)=E[(X-x)^2], x\in R$,当$x=E(X)$时达到最小.

证明: $f(x)=E[(X-x)^2]=E(X^2)-2xE(X)+x^2$

令 $f'(x)=0=-2E(X)+2x$ 得到 $x=E(X)$

又 $\because f''(x)|_{x=E(X)}=2>0$

故$f(x)$在$x=E(X)$时取到最小值.

随机变量关于其数学期望的偏离程度(方差)比它相对于其它任何值的偏离程度都小.

$D(X) \le E[(X-x)^2], x\in R$,等式在$x=E(X)$时成立.

证明1： $E\{[X-E(X|Y)]^2|Y\}\le E\{[X-g(Y)]^2|Y\}$

令 $h(t,y)\xrightarrow{t=g(y)}E\{[X-g(y)]^2|Y=y\}$

$=E\{X^2-2Xt+t^2|y\}=E\{X^2|y\}-2tE\{X|y\}+t^2$

$h(t,y)'_t=2t-2E[X|y]=0\Rightarrow t=E[X|y]$

$E\{[X-E(X|Y)]^2\}\le E\{[X-g(Y)]^2\}$

证明2： $E\{[X-g(Y)]^2\}=E\{[X-E(X|Y)+E(X|Y)-g(Y)]^2\}$

$=E[X-E(X|Y)]^2+E[(X|Y)-g(Y)]^2$

$+2E\{[X-E(X|Y)][E(X|Y)-g(Y)]\}\ge E[X-E(X|Y)]^2$

其中: $E\{[X-E(X|Y)][E(X|Y)-g(Y)]\}=E\{[X-E(X|Y)][E(X|Y)-g(Y)]|Y)\}$

$=[E(X|Y)-g(Y)]E(X-E(X|Y)|Y)=0$

3. 全数学期望公式

性质5)之特例 $\begin{cases} E(X)=E[E(X\mid Y)] \\ E[g(X)]=E\{E[g(X)\mid Y]\}=\int_{-\infty}^{+\infty} E[g(X)\mid y]dF_Y(y) \end{cases}$

思想: 计算X的期望可分两步走先算给定$Y=y$时X的条件期望，再对条件期望作加权平均就是总的平均或期望.

特别有**全概率公式**

对随机事件A,定义示性函数$I_A=\begin{cases} 1, & 若A发生; \\ 0, & 若A不发生. \end{cases}$

$\because \; E[I_A\mid Y=y]=P(A\mid Y=y)$

$\therefore P(A)=E(I_A)=E\{E[I_A\mid Y]\}$

$$P(A)=\begin{cases} \sum_y P(A\mid Y=y)P(Y=y), & Y \text{ 是离散型r.v.} \\ \int_{-\infty}^{+\infty} P(A\mid Y=y)f_Y(y)dy, & Y \text{ 是连续型r.v.} \end{cases}$$

Eg.4 常用全数学期望公式若$P\{Y=y_k\}=p_k,(k=1,2,\cdots),\sum\limits_{k=1}^{\infty}p_k=1,$

记$A_k=\{Y=y_k\}$, $\quad E(X\mid A_k)=E(X\mid Y=y_k)$,则有$E(X)=\sum\limits_{k=1}^{\infty}E(X\mid A_k)P(A_k).$

Eg.5 设某段时间内到达商场的顾客人数N服从参数为λ的泊松分布.每位顾客在该商场的消费额X服从$[a,b]$上的均匀分布.各位顾客之间消费是相互独立的且与N独立.求顾客在该商场总的消费额的期望值.

解: 设第i个顾客消费额为X_i,全体顾客在该商场总消费额为$S=\sum_{i=1}^{N} X_i$

根据全数学期望公式得

$$E(S)=E[E(S\mid N)]=E\left[E\left(\sum_{i=1}^{N} X_i \mid N\right)\right]$$

$$=\sum_{n=0}^{\infty} E\left[\sum_{i=1}^{n} X_i \mid N=n\right] P\{N=n\}=\sum_{n=0}^{\infty}\left[\sum_{i=1}^{n} E(X_i) P\{N=n\}\right]$$

$$=E(X)\sum_{n=0}^{\infty} n P\{N=n\}=E(X)E(N)=\frac{a+b}{2}\lambda$$

如果随机变量Y等于一串独立随机变量$X_1, X_2, \cdots, X_N$的和,且下标N也是一个整数值随机变量,则Y是随机变量的随机和.这在实际问题中常常会碰到.比如:

(1) N是在某一时段到达服务台的顾客数,而X_i代表第i个顾客所需要的服务时间,则$Y=\sum_{i=1}^{N} X_i$ 代表总服务时间.这是排队论问题中的重要变量.

(2) 在讨论人口和生物群体生长模型时,在某一地区某类生物的总数为N,X_i为第i个个体的后代数，则Y就是该群体的总数.下面讨论随机和Y的矩母函数与数字特征

$$Y=\sum_{i=1}^{N} X_i$$

Eg.6： 已知随机变量X服从$[0,a]$上的均匀分布,随机变量Y服从$[X,a]$上的均匀分布,试求

1) $E(Y|X=x), 0<x<a$;

2) $E(Y)$.

解1)：由条件知对任意$0<x<a$,有$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{a-x}, & x<y<a\\ 0, & \text{其它}\end{cases}$

$$E(Y|X=x)=\int_x^a \frac{y}{a-x}dy=\frac{a+x}{2},\ E(Y|X)=\frac{a+X}{2}$$

2) $$E(Y)=E[E(Y|X)]=E\left(\frac{a+X}{2}\right)=\frac{a+\frac{a}{2}}{2}=\frac{3}{4}a$$

$$=\int_{-\infty}^{+\infty}E(Y|x)f_X(x)dx=\int_0^a \frac{a+x}{2}\cdot\frac{1}{a}dx=\frac{3}{4}a$$

Eg.7(几何分布的期望): 连续抛掷一枚出现正面的概率为p的硬币直至出现正面为止.
问需要抛掷的次数的期望多少?

解: 以N记需要抛掷的次数,而令$Y=\begin{cases}1 & \text{如果第一次抛掷的结果是正面,}\\ 0 & \text{如果第一次抛掷的结果是反面.}\end{cases}$ 则

$$\begin{aligned}E(N)&=E(E(N\mid Y))=P(Y=1)\cdot E(N\mid Y=1)+P(Y=0)\cdot E(N\mid Y=0)\\ &=p\cdot E(N\mid Y=1)+(1-p)\cdot E(N\mid Y=0).\end{aligned}$$

我们有$E(N\mid Y=1)=1$, $E(N\mid Y=0)=1+E(N)$由于$Y=1$,我们知道第一次抛掷结果是正面,所以,需要抛掷的次数的期望是1.

另一方面,如果$Y=0$,第一次抛掷结果是反面.然而,由于假定相继的抛掷是独立的,这就推出在第一次出现反面直到正面首次出现时的附加抛掷次数的期望是$E(N)$.所以有

$E(N)=p+(1-p)\cdot(1+E(N))$,解方程,得 $E(N)=\dfrac{1}{p}$.

Eg.8: 某矿工身陷有三个门的矿井之中.经第1个门的通道行进2小时后,他将到达安全地;经第2个门的通道前进3小时后,他将回到矿井原地;经第3个门的通道前进5小时后,他又将回到矿井原地.假定这个矿工每次都等可能地任意一个门,问直到他到达安全地所需时间的期望是多少?

解: 令X记矿工到达安全地所需的时间,以Y记他最初选取的门.Y的分布律为$E(Y=i)=\dfrac{1}{3}, i=1,2,3$.现在

$$E(X)=E(E(X)|Y)=\sum_{i=1}^{3}P(Y=i)\cdot E(X|Y=i)=\sum_{i=1}^{3}\frac{1}{3}\times E(X|Y=i)$$

则有$E(X|Y=1)=2, E(X|Y=2)=3+E(X),\quad E(X|Y=3)=5+E(X)$。

以$E(X|Y=2)$为例,给出其如下推理.如果矿工选取第2个门,那么3小时后他将回到他的矿井.但是,一旦他回到矿井,问题就和以前一样了,而直到他到达安全地的附加时间的期望正是$E(X)$.

因此, $E(X)=\dfrac{1}{3}(2+3+E(X)+5+E(X))$,解方程,得$E(X)=10$.