随机过程

林爱红
ahlin@ecust.edu.cn

教材：随机过程

刘次华编，华中科技大学出版社

机器学习的主要应用之一是对随机过程建模

“The only simple truth is that there is nothing simple in this complex universe. Everything relates. Everything connects”

— Johnny Rich, The Human Script

泊松过程：用于处理等待时间以及队列。

随机漫步和布朗运动过程：用于交易算法。

马尔可夫决策过程：常用于计算生物学和强化学习。

高斯过程：用于回归和优化问题(如,超参数调优和自动机器学习)。

自回归和移动平均过程：用于时间序列分析(如,ARIMA 模型)。

第一个随机过程：爱因斯坦的布朗运动

如果给随机过程打个比方，它就像是一个充满交叉小径的花园。你站在现在的点上，看未来的变化，未来有千万种变化的方式，　每一种可能又不断分叉变化出其它可能。

引例1 （排队模型）顾客来到服务站要求服务. 当服务站中的服务员都忙碌，即服务员都在为别的顾客服务时，来到的顾客就要排队等候. 顾客的到来、每个顾客所需的服务时间都是随机的，所以如果用$X(t)$表示t时刻的队长，用 $Y(t)$表示t时刻到来的顾客所需的等待时间，则$\{X(t),t\in T\},\{Y(t),t\in T\}$都是随机过程.

第一章 预备知识

一. 特征函数
二. 多元正态分布
三. 条件概率

常用分布的期望和方差

1.两点分布	$P\{X=1\}=p$	$EX=p$	$DX=pq,(p+q=1)$
2.二项分布	$B(n,p)$	$EX=np$	$DX=npq,(p+q=1)$
3.超几何分布	$H(n,M,N)$	$EX=\frac{nM}{N}$	$DX=\frac{nM(N-M)(N-n)}{N^2(N-1)}$
4.泊松分布	$P(\lambda)$	$EX=\lambda$	$DX=\lambda$
5.几何分布	$G(p)$	$EX=1/p$	$DX=q/p^2$
6.均匀分布	$U(a,b)$	$EX=\frac{a+b}{2}$	$DX=(b-a)^2/12$
7.正态分布	$N(\mu,\sigma^2)$	$EX=\mu$	$DX=\sigma^2$
8.指数分布	$E(\lambda)$	$EX=1/\lambda$	$DX=1/\lambda^2$
9. χ^2分布	$\chi^2(k)$	$EX=k$	$DX=2k$

随机向量的定义：

设 $\{X_i(\omega)\}\quad i=1,2,\cdots,n$ 是定义在同一个概率空间 $(\Omega,\mathcal{F},P)$ 上的 n 个随机变量，则称 $X(\omega)=(X_1(\omega),X_2(\omega),\cdots,X_n(\omega))$ 为 n 维随机向量或 n 维随机变量.

分布函数的定义

对任意实数$x_1, x_2, \cdots, x_n$，称函数

$$F(x_1, x_2, \cdots, x_n) = P\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\}$$

为随机向量ξ的（联合）分布函数.

若$(\boldsymbol{X_1}, \boldsymbol{X_2}, \cdots, \boldsymbol{X_n})$ 是 离散型随机向量，则联合概率分布列为

$$\boldsymbol{P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}}$$

若$(\boldsymbol{X_1}, \boldsymbol{X_2}, \cdots, \boldsymbol{X_n})$ 是 连续型随机向量，则联合概率密度为

$$p\left(x_1, x_2, \cdots, x_n\right)$$

分布函数

$$\boldsymbol{F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(t_1, t_2, \cdots, t_n) dt_1 \cdots dt_n}$$

独立性

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\} = \prod_{i=1}^{n} P\{X_i = x_i\}$$

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i)$$

$$P\{X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n\} = \prod_{i=1}^{n} P\{X_i \le x_i\}$$

$$E(X_1 X_2 \cdots X_n) = EX_1 EX_2 \cdots EX_n$$

$$D(X_1 \pm X_2 \cdots \mp X_n) = DX_1 + DX_2 \cdots + DX_n$$

概率母函数

讨论取非负整数值的离散型随机变量ξ,

ξ	0	1	$\cdots$	n	$\cdots$
p_k	p_0	p_1	$\cdots$	p_n	$\cdots$

定义: $$g(s)=E(s^{\xi})=\sum_{j=0}^{\infty}s^{j}p_{k},\qquad s\in[-1,1]$$

为随机变量ξ的概率母函数,简称母函数.

对于$s\in[-1,1]$, 有$|s^{\xi}|\le 1$, 规定$0^{0}=1$,

母函数的性质

（1）母函数和分布可唯一互相确定

(2) $P(X=k)=\frac{1}{k!}g^{(k)}(0),\quad k=0,1,2,\cdots;\quad EX=g'(1);$

（3） 如果 $EX<\infty$,则 $DX=g''(1)+g'(1)-(g'(1))^2;$

（4） 如果 $X_1,\ldots,X_n$ 相互独立，$g_i(s)=E(s^{X_i})$ 是 X_i 的母函数,

则 $Y=X_1+\ldots+X_n$ 有母函数

$$g_Y(s)=g_1(s)g_2(s)\cdots g_n(s),\qquad s[-1,1].$$

重要分布的母函数

例 1：退化分布 $I(x-c)$ 的母函数为 $g(s)=s^{c}$

例 2：二项分布 $B(n,p)$ 的母函数为 $g(s)=(ps+q)^{n}$

例 3：泊松分布 $P(\lambda)$ 的母函数为 $g(s)=e^{\lambda(s-1)}$

例 4：几何分布 $G(p)$ 的母函数为 $g(s)=\dfrac{sp}{1-s(1-p)}$

特征函数(Characteristic function)

若随机变量ξ的分布函数为$F_\xi(x)$，则称

$$f_\xi(t) = Ee^{it\xi} = \int_{-\infty}^{\infty} e^{itx} dF_\xi(x)$$

为ξ的特征函数. (分布函数的傅里叶变换)

1) **离散型随机变量**，分布列$\begin{bmatrix} x_1 & x_2 \cdots & x_n \cdots \\ p_1 & p_2 \cdots & p_n \cdots \end{bmatrix}$，则特

征函数为 $f(t) = \sum_{j=1}^{\infty} p_j e^{itx_j} = P(e^{it})$，母函数 $P(s)$

2) **连续型随机变量**，$f(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx$

特征函数方便计算r.v.的矩,以及求独立r.v.和的分布函数.

例 1．求离散型随机变量的特征函数 $\varphi_X(t)$，若随机变量 X 的分布为

（1）二项分布 $B(n,p)$；

（2）泊松分布 $P(\lambda)$．

解：（1）$\varphi_X(t)=\sum_{k=0}^{n} e^{itk} C_n^k p^k q^{n-k}=\sum_{k=0}^{n} C_n^k (pe^{it})^k q^{n-k}=(pe^{it}+q)^n$

当 $n=1$ 时，则 $X\sim B(1,p)$（两点分布），其特征函数为

$$\varphi_X(t)=pe^{it}+q$$

（2）$\varphi_X(t)=\sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^{it}\lambda)^k}{k!}=e^{-\lambda}e^{\lambda e^{it}}=e^{\lambda(e^{it}-1)}$

例 2 求正态分布 $N(\mu,\sigma^2)$ 的特征函数.

解: $$f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{itx}e^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos tx\cdot e^{-\frac{x^2}{2}}dx$$

由于正态分布一阶矩存在，求导，得

$$f'(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(-x)\sin txe^{-\frac{x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin txde^{-\frac{x^2}{2}}$$

$$=\left[\frac{1}{\sqrt{2\pi}}\sin tx\cdot e^{-\frac{x^2}{2}}\right]_{-\infty}^{\infty}-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t\cos tx\cdot e^{-\frac{x^2}{2}}dx=-tf(t)$$

$$\ln f(t)=-\frac{t^2}{2}+c,\qquad f(0)=1\Rightarrow c=0\Rightarrow f(t)=e^{-\frac{t^2}{2}}$$

2．注意：

1）存在性：$\varphi(t)$ 是复值函数，且必存在.

2）唯一性：随机变量的特征函数和分布律一一对应.

3）操作性：特征函数易求，且具有良好的分析性质.

例 3．设随机变量 X 的特征函数为$\varphi_X(t)$， $Y=aX+b$，其中 a,b 为任意常数，

（1）证明Y 的特征函数$\varphi_Y(t)=e^{itb}\varphi_X(at)$；

（2）若 $Z\sim N(\mu,\sigma^2)$，求$\varphi_Z(t)$.

解:（1） $\varphi_Y(t)=E(e^{itY})=Ee^{it(aX+b)}=E(e^{itb}\cdot e^{i(at)X})=e^{itb}\varphi_X(at)$

（2）已知 $X\sim N(0,1)$， $\varphi_X(t)=e^{-\frac{t^2}{2}}$， $Z=\sigma X+\mu$，由（1）得

$$=e^{i\mu t-\frac{\sigma^2}{2}t^2}$$

3．性质：

1） $\varphi(0)=1, |\varphi(t)|\le 1, \varphi(-t)=\overline{\varphi(t)}$.

2） $\varphi(t)$ 在 R 上一致连续.

3）若随机变量的 n 阶原点矩存在，则有 $\varphi^{(k)}(0)=i^k EX^k$， $k\le n$

（即利用特征函数求各阶原点矩 $EX^k=\dfrac{\varphi^{(k)}(0)}{i^k}$）

4） $\varphi(t)$ 是非负定函数，即对任意正整数 n 及任意实数 $t_1,t_2,\cdots,t_n$

和复数 $z_1,z_2,\cdots,z_n$，有 $\sum\limits_{k,l=1}^{n}\varphi(t_k-t_l)z_k\bar{z}_l\ge 0$.

5）若 $X_1, X_2, \cdots, X_n$ 是相互独立的随机变量，则 $X = X_1 + X_2 + \cdots + X_n$ 的特征函数 $\varphi(t) = \varphi_1(t)\varphi_2(t)\cdots\varphi_n(t)$，其中 $\varphi_i(t)$ 是 X_i $(i = 1,2,\cdots,n)$ 的特征函数。(独立和的特征函数为各特征函数的乘积)

6）随机变量的分布函数由其特征函数唯一确定。

（分布与特征函数 1-1 对应）

例 4．由特征函数求数字特征：

（1）若 $X \sim B(n,p)$，求 EX, EX^2, DX；

（2）若 $X \sim N(0,\sigma^2)$，求 EX^n.

解：（1）若 $X \sim B(n,p)$，则 $\varphi_X(t)=(pe^{it}+q)^n$，

$$\varphi'_X(t)=npie^{it}(pe^{it}+q)^{n-1}，\quad \varphi'_X(0)=npi，\quad EX=\frac{\varphi'_X(0)}{i}=np$$

$$\varphi''_X(t)=npi^2e^{it}(pe^{it}+q)^{n-1}+n(n-1)p^2i^2e^{2it}(pe^{it}+q)^{n-2}$$

$$\varphi''_X(0)=-np-n(n-1)p^2，\quad EX^2=\frac{\varphi''_X(0)}{i^2}=np+n(n-1)p^2$$

$$DX=EX^2-(EX)^2=np(1-p)$$

（2）若 $X\sim N(0,\sigma^2)$，则 $\varphi_X(t)=e^{-\frac{\sigma^2}{2}t^2}$，展开成幂级数为

$$\varphi_X(t)=\sum_{k=0}^{\infty}\frac{(-\frac{\sigma^2}{2}t^2)^k}{k!}=\sum_{k=0}^{\infty}\frac{(-1)^k\frac{\sigma^{2k}}{2^k}t^{2k}}{k!},$$

又根据泰勒展式 $\varphi_X(t)=\sum_{n=0}^{\infty}\frac{\varphi_X^{(n)}(0)}{n!}t^n$，即

当 $n=2k$ 时，$\varphi_X^{(2k)}(0)=\frac{(-1)^k(2k)!\sigma^{2k}}{2^k k!}=(-1)^k(2k-1)!!\sigma^{2k}$

$$EX^{2k}=\frac{\varphi_X^{(2k)}(0)}{i^{2k}}=(2k-1)!!\sigma^{2k}$$

当 $n=2k-1$ 时，$\varphi_X^{(2k-1)}(0)=0$，$EX^{2k-1}=\frac{\varphi_X^{(2k-1)}(0)}{i^{2k-1}}=0$

即 $EX^n=\begin{cases}(2k-1)!!\sigma^{2k}, & n=2k\\ 0, & n=2k-1\end{cases}\quad(k=1,2,\cdots)$

例 5．设随机变量 X 在 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 上服从均匀分布，$Y=\cos X$，利用特征函数求 Y 的密度函数。（由性质（6），从形式上得出）

解：X 的密度函数为 $f_X(x)=\begin{cases}1/\pi, & x\in(-\pi/2,\pi/2)\\ 0, & \text{其它}\end{cases}$

$$\varphi_Y(t)=Ee^{itY}=Ee^{it\cos X}=\int_{-\pi/2}^{\pi/2}e^{it\cos x}\cdot\frac{1}{\pi}dx=2\int_0^{\pi/2}e^{it\cos x}\cdot\frac{1}{\pi}dx$$

$$\underline{\underline{y=\cos x}}\quad 2\int_1^0 e^{ity}\cdot\frac{1}{\pi}\cdot\left(-\frac{1}{\sqrt{1-y^2}}\right)dy=\int_0^1 e^{ity}\cdot\frac{2}{\pi\sqrt{1-y^2}}dy$$

由唯一性得 $f_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1\\ 0, & \text{其它}\end{cases}$

4. 定义： 设 $X=(X_1,X_2,\cdots,X_n)$ 是 n 维随机变量， $t=(t_1,t_2,\cdots,t_n)\in R^n$ ，

则称

$$\varphi(t)=\varphi(t_1,t_2,\cdots,t_n)=Ee^{itX^T}=Ee^{i(t_1X_1+t_2X_2+\cdots+t_nX_n)}$$

为 $X=(X_1,X_2,\cdots,X_n)$ 的特征函数.

n 维随机变量的特征函数具有类似于一维随机变量的特征函数的性质.

二、多维正态分布

1. 定义： 若 n 维随机变量 $X=(X_1,X_2,\cdots,X_n)$ 的联合概率密度为

$$f(x)=f(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}}\exp\{-\frac{1}{2}(x-a)B^{-1}(x-a)^T\},$$

式中，$a=(a_1,a_2,\cdots,a_n)$ 是常向量，$B=(b_{ij})_{n\times n}$ 是正定矩阵，则称 X 为 n 维正态随机变量或服从 n 维正态分布，记做 $X\sim N(a,B)$.

2. 结论：若 n 维正态随机变量 $X \sim N(a,B)$，

1）则 X 的特征函数为 $\varphi(t)=\varphi(t_1,t_2,\cdots,t_n)=e^{iat^T-\frac{1}{2}tBt^T}$

2）则 $EX_k=a_k$， $\mathrm{cov}(X_k,X_l)=b_{kl}$， $k,l=1,2,\cdots,n$

(即 $EX=a$， B 为协方差矩阵)

3）则 $X_1,X_2,\cdots,X_n$ 两两相互独立 $\Leftrightarrow$ $X_1,X_2,\cdots,X_n$ 两两不相关. (正态时)

4）$Y=XA$， 且 A^TBA 正定， 则 $Y \sim N(aA,A^TBA)$.

（即正态随机变量的线性变换仍为正态随机变量.）

5）则必存在 n 阶正交矩阵 A， 使得 $Y=(Y_1,Y_2,\cdots,Y_n)=(X-a)A^T$ 是 n 维独立正态随机变量，即 $Y_1,Y_2,\cdots,Y_n$ 相互独立，且 $Y_k \sim N(0,d_k)$，其中 $d_k>0$ 是 B 的特征值， $k=1,2,\cdots,n$. (相当于实对称矩阵的对角化)

例 6. 设 X_1, X_2, X_3 独立同分布于标准正态分布，证明 $Y_1 = \frac{1}{\sqrt{2}}(X_1 - X_2)$，

$Y_2 = \frac{1}{\sqrt{6}}(X_1 + X_2 - 2X_3)$， $Y_3 = \frac{1}{\sqrt{3}}(X_1 + X_2 + X_3)$ 也独立同分布于标准正态分布.

解：因 X_1, X_2, X_3 独立同分布于标准正态分布，则 $X = (X_1, X_2, X_3)$ 的均值为 a=0，协方差矩阵为 $B = I$，即 $X \sim N(0, I)$.

又 $Y = (Y_1, Y_2, Y_3) = (X_1, X_2, X_3)\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$,

而 $A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$ 是正交矩阵，故 $A^T BA = I$ 正定，

则由 4）： $Y \sim N(0, I)$， 即 Y_1, Y_2, Y_3 也独立同分布于标准正态分布.

三、条件期望

1．定义：随机变量 X 和 Y，

1） X 在 $Y=y$ 时的条件期望：$E(X|Y=y)$

（是 y 的函数）

2） X 在 Y 下的条件期望：$E(X|Y)$

（是随机变量 Y 的函数）

2．性质：

1）X 和 Y 的期望都存在，则 $EX = E[E(X|Y)]$.

注意：$E[E(X|Y)] = \int E(X|y)dF_Y(y)$，具体地，

（1）若 Y 是离散型随机变量，则 $EX = \sum_y E(X|Y=y)P\{Y=y\}$

（2）若 Y 是连续型随机变量，其概率密度为 $f(y)$，则

$$EX = \int_{-\infty}^{+\infty} E(X|Y=y)f(y)dy$$

2）示性函数

事件 A 的示性函数 $I_A = I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$ 是一个二值随机变量

$$EI_A = P(A)，\quad E(I_A|Y=y) = P(A|Y=y)$$

$$\therefore P(A) = E(I_A) = E[E(I_A|Y)] = \int E(I_A|Y=y)dF_Y(y) = \int P(A|Y=y)dF_Y(y)$$

（相当于全概率公式）

例 7．设某段时间内到达某商场的顾客人数 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且与 N 独立。求顾客在该商场总的消费额的数学期望.

解：设 X_i 表示第 i 位顾客的消费额，

已知 $X_i \overset{i.i.d}{\sim} U[a,b]$， $EX_i = \dfrac{a+b}{2}$， $i = 1,2,\cdots$， 则

N 位顾客的总消费额为 $S = \sum_{i=1}^{N} X_i$ ，

$$ES = E(E(S \mid N)) = \sum_{n=0}^{\infty} E(S \mid N = n)P(N = n) \quad (\because E(S \mid N = 0) = 0)$$

$$= \sum_{n=1}^{\infty} E(\sum_{i=1}^{n} X_i)P(N = n) = \sum_{n=1}^{\infty} (\sum_{i=1}^{n} EX_i)P(N = n) = \sum_{n=1}^{\infty} nEX_1 P(N = n)$$

$$= EX_1 \sum_{n=1}^{\infty} nP(N = n) = EX_1 \sum_{n=0}^{\infty} nP(N = n) = EX_1 \cdot EN = \frac{a+b}{2}\lambda$$

若某商场日均顾客人数为 1000 人，人均消费额为 300 元，则该商场的平均日消费额为 30 万元 .

例 8. 设迷宫中某处有三个路口。若选择路口 1，则 2 小时可走出迷宫；若选择路口 2，则 4 小时后又返回原处；若选择路口 3，则 6 小时后又返回原处；并设每次选择各路口的概率等可能，求走出迷宫所需时间的期望值 .

解：设 X 为选择路口的号码，且 $P(X=k)=\frac{1}{3},\ k=1,2,3$，

T 为走出迷宫所需要的时间，

T_1 为返回原处后再走出迷宫所需要的时间， $T=\begin{cases} 2, & X=1 \\ 4+T_1, & X=2 \\ 6+T_1, & X=3 \end{cases}$，

注意：$\{T_1 \mid X=2\}$、$\{T_1 \mid X=3\}$ 与 T 有相同分布，

则 $E(T_1 \mid X=2)=E(T_1 \mid X=3)=ET$， 因此

$$ET=E(E(T \mid X))=\sum_{k=1}^{3} E(T \mid X=k)P(X=k)$$

$$=E(T=2 \mid X=1)P(X=1)+E(T=4+T_1 \mid X=2)P(X=2)$$
$$+E(T=6+T_1 \mid X=3)P(X=3)$$

$$=(2+4+ET+6+ET)\cdot\frac{1}{3}=4+\frac{2}{3}ET$$，得 $ET=12$

例 9. 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） $E(Y \mid X=x), 0<x<a$；（2） EY .

解：（1）已知 $X=x$ 时， $Y \sim U[x,a]$，因此

$$E(Y \mid X=x)=\frac{x+a}{2}$$

$$(2)\ EY=E(E(Y \mid X))=\int_0^a E(Y \mid x)\,p_X\,dx$$

$$=\int_0^a \frac{a+x}{2}\frac{1}{a}dx=\frac{3}{4}a$$

例 10．为了研究抽烟和身体健康的关系，以 X 表示每人每天抽烟支数，分为三类：每天抽 0，10，20 支。以 Y 表示同一个人的身体状态，分为三等：好，中，坏，分别相应于 $Y=0$，1，2。在一个地方随机抽样，得到 X 和 Y 的概率分布及边际分布如下表。根据下表数据研究：（1）抽烟多与健康状态好坏有无关系；（2）每天抽 x 支烟的人的平均健康状态。

Y \ X	0	10	20	$p_Y(y_j)$
0	0．25	0．05	0．05	0．35
1	0．05	0．15	0．05	0．25
2	0．05	0．10	0．25	0．40
$p_X(x_i)$	0．35	0．30	0．35	1

解：（1）即求 $P(Y=k\mid X=20), k=0,1,2$ 和 $P(Y=2)$， 并比较

$$P(Y=0\mid X=20)=\frac{0.05}{0.35}=\frac{1}{7},\quad P(Y=1\mid X=20)=\frac{0.05}{0.35}=\frac{1}{7}$$

$$P(Y=2\mid X=20)=\frac{0.25}{0.35}=\frac{5}{7}\approx 0.71>P(Y=2)=0.4$$

即:每天抽 20 支烟的人健康不好的概率(约 0.71)远远超过一般健康不好(0.4)的概率.

结论：抽烟多的人健康不好的概率比较大。

（2）即求$E(Y|X)$的取值。

$$E(Y|X=0)=0\cdot\frac{0.25}{0.35}+1\cdot\frac{0.05}{0.35}+2\cdot\frac{0.05}{0.35}=\frac{3}{7}\approx 0.43$$

$$E(Y|X=10)=0\cdot\frac{0.05}{0.30}+1\cdot\frac{0.15}{0.30}+2\cdot\frac{0.10}{0.30}=\frac{7}{6}\approx 1.17$$

$$E(Y|X=20)=0\cdot\frac{0.05}{0.35}+1\cdot\frac{0.05}{0.35}+2\cdot\frac{0.25}{0.35}=\frac{11}{7}\approx 1.57$$

即： $E(Y|X=0)<E(Y|X=10)<E(Y|X=20)$

结论：抽烟少的人平均健康状况比较好，而抽烟多的人的平均健康状况最差