第六章总结 联合分布

一. Discrete Multivariate R.V.s

· Def: Joint CDF of r.v. X & Y is $F_{X,Y}$ given by:

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

则可知: $P_{X,Y}(x,y) = P(X=x, Y=y)$

且: $$\begin{cases} \sum_x \sum_y P(X=x, Y=y) = 1 \\ P(X=x) = \sum_y P(X=x, Y=y) \end{cases}$$ ← 边缘概率质量函数 (LOTP)

· 条件PMF: 对于离散 r.v. X. Y, X的条件PMF (在Y=y下) 是:

$$P_{X|Y}(x|y) = P(X=x|Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)}.$$

· r.v. X Y 称为独立 if for all x. y:

$$F_{X,Y}(x,y) = F_X(x) F_Y(y)$$

若X. Y 离散, 则 $\Leftrightarrow$ $P(X=x, Y=y) = P(X=x) P(Y=y)$

$\Leftrightarrow$ $P(Y=y|X=x) = P(Y=y).$

Eg: Chicken-egg 问题. 鸡随机产生N枚鸡蛋, $N \sim Pois(\lambda)$。每枚蛋 w.p 独立 hatch, $q=1-p$ 失败。: X: 成功hatch数量, Y: 未hatch数量, $X+Y=N$。求X. Y 联合概率质量函数

Solution: $P(X=i, Y=j) = \sum_{n=0}^{\infty} P(X=i, Y=j | N=n) P(N=n)$

$= P(X=i, Y=j | N=i+j) P(N=i+j).$

$$= \binom{i+j}{i} p^i q^j \cdot \frac{e^{-\lambda}\lambda^{i+j}}{(i+j)!} = \frac{e^{-\lambda p}(\lambda p)^i}{i!} \cdot \frac{e^{-\lambda q}(\lambda q)^j}{j!}$$

由上例可衍生出 Theorem:

If $X \sim Pois(\lambda p)$, $Y \sim Pois(\lambda q)$, X. Y独立, then $N = X+Y \sim Pois(\lambda)$ and $X|N=n \sim Bin(n,p)$

No.
Date

i.e.: If $N \sim Pois(\lambda)$ $X|N=n \sim Bin(n,p)$, then $X \sim Pois(\lambda p)$
$Y = N - X \sim Pois(\lambda q)$, X、Y独立

二. Continuous Multivariate R.V.s

Joint CDF: $F_{X,Y}(x,y) = P(X \le x, Y \le y)$

则联合PDF: $f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$

Validity: $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx\,dy = 1$

可利用之求更广的 range 的 "Collective Density":

$$P((X,Y)\in A) = \iint_A f_{X,Y}(x,y)\,dx\,dy$$

该 Scenario 的 Marginal PDF:

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

连续场景下的贝叶斯与LOTP:

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)\,f_Y(y)}{f_X(x)}$$

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

那 X Y 任意 DT or CT , Baye's Rule 如下:

	Y DT	Y CT
X DT	$P(Y=y\mid X=x) = \frac{P(X=x\mid Y=y)P(Y=y)}{P(X=x)}$	$f_X(y\mid X=x) = \frac{P(X=x\mid Y=y)f_Y(y)}{P(X=x)}$
X CT	$P(Y=y\mid X=x) = \frac{f_X(x\mid Y=y)P(Y=y)}{f_X(x)}$	$f_{Y\mid X}(y\mid x) = \frac{f_{X\mid Y}(x\mid y)f_Y(y)}{f_X(x)}$

谁DT, 谁用 P; 谁CT, 谁用 f

Campus

LOTP也如此：

	Y DT	Y CT
X DT	$P(X=x)=\sum_y P(X=x\mid Y=y)P(Y=y)$;	$P(X=x)=\int_{-\infty}^{\infty} P(X=x\mid Y=y)f_Y(y)dy$
Y CT	$f_X(x)=\sum_y f_X(x\mid Y=y)P(Y=y)$;	$f_X(x)=\int_{-\infty}^{\infty} f_{X\mid Y}(x\mid y)f_Y(y)dy$

· Independence of Continuous R.V.s

R.V. X, Y 独立 if $\forall x, y$: $F_{X,Y}(x,y)=F_X(x)F_Y(y)$

$\Leftrightarrow f_{X,Y}(x,y)=f_X(x)f_Y(y) \Leftrightarrow f_{Y\mid X}(y\mid x)=f_Y(y)$

Theorem: 若 $f_{X,Y}(x,y)=g(x)h(y)$ (factored as)
则 X Y独立。且若g. h中一个中Valid PDF，另一个也是Valid PDF，且g. h为X. Y的marginal PDFs.

· 2D LOTUS: $g: R^2\to R$，则：

$E(g(X,Y))=\sum_x\sum_y g(x,y)P(X=x,Y=y)$. 若X. Y连续，则：

$E(g(X,Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dxdy$

下见一些例子

Eg1: (Conditional PDF) $f(x,y)=\begin{cases}\frac{12x(2-x-y)}{5}, & \text{if } x\in(0,1),\ y\in(0,1)\\ 0, & \text{otherwise}\end{cases}$

求Y=y下，X的条件PDF

Solution: $f_{X\mid Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dx}$

$=\frac{f(x,y)}{\int_0^1 f(x,y)dx}=\frac{6x(2-x-y)}{4-3y}$

Eg2. (Baye's Rule):

$Y \sim Expo(\lambda)$，但 $\lambda \sim Unif\ [1, \frac{3}{2}]$。已知 $Y=y$，求 λ 分布:

Solution: Λ 代表 λ 的分布; Unif → Expo

$$f_{\Lambda|Y}(\lambda|y) = \frac{f_\Lambda(\lambda) \cdot f_{Y|\lambda}(y|\lambda)}{\sum f_\Lambda(\lambda) f_{Y|\lambda}(y|\lambda)} = \frac{(\frac{3}{2}-1)\cdot \lambda e^{-\lambda y}}{\int_1^{3/2} (\frac{3}{2}-1)\lambda e^{-\lambda y} d\lambda}$$

Eg3. (LOTP) $T_1 \sim Expo(\lambda_1)$ $T_2 \sim Expo(\lambda_2)$ 求 $P(T_1<T_2)$.

有: $P(T_1<T_2) = \int_0^{+\infty} P(T_1<T_2|T_2=t)\ P(T_2=t)\, dt$

$= \int_0^{+\infty} (1-\lambda_1 e^{-\lambda_1 t})\ \lambda_2 e^{-\lambda_2 t}\, dt$

Eg4. For $X, Y \overset{i.i.d}{\sim} Unif(0,1)$, find $E(|X-Y|)$, $E(\max(X,Y))$ $E(\min(X,Y))$.

Solution: $E|X-Y| = \int_0^1\int_0^1 |x-y| f_X(x) f_Y(y)\, dx dy$

$= \int_0^1\int_0^1 |x-y|\, dx dy = \int_0^1\int_y^1 (x-1)\, dx dy + \int_0^1\int_0^y (y-x)\, dx dy$

$= \int_0^1 (\frac{1}{2}x^2 - x)\Big|_y^1 dy + \int_0^1 (yx - \frac{1}{2}x^2)\Big|_0^y dy = 1/3$

而: $M = \max(X,Y)$ $N = \min(X,Y)$，则 $M+N = X+Y$

$\therefore E(M)+E(N) = E(M+N) = E(X+Y) = 1$

又: $E(M)-E(N)$: $M-N = \begin{cases} x-y, & \text{if } x>y \\ y-x, & \text{if } x<y \end{cases} = |X-Y|$

$\therefore E(M)-E(N) = E(M-N) = E(|X-Y|) = 1/3$

$\therefore$ 解出: $E(M) = \frac{2}{3}$ $E(N) = \frac{1}{3}$

No.

Date

三. Covariance & Correlation

· Def: r.v.s X, Y 有: $Cov(X,Y) = E((X-EX)(Y-EY))$

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

则协方差有一些性质:

$Cov(X,X) = Var(X)$ $\quad Cov(X,Y) = Cov(Y,X)$

$Cov(X,c) = 0$ $\quad Cov(a\cdot X, Y) = a\,Cov(X,Y)$

$Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$

$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)^{*}$

* 之间提过, X Y间独立, 则 $Var(X+Y) = Var(X) + Var(Y)$

$Var(X_1 + \cdots + X_n) = \sum_i Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$

· Def: r.v.s X. Y间 correlation 为:

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Given r.v.s X. Y, if $Cov(X,Y)=0$ or $Corr(X,Y)=0$,

则 X. Y不相关 (注意, 不是独立!)

Theorem: if X. Y不相关 (uncorrelated), X. Y 不一定独立

Eg: $X \sim N(0,1)$ $\quad Y = X^2$, $E(X)=0$, $E(XY) = E(X^3) = 0$

本质: 这里的 Correlation 实际仅指 线性 correlation

Theorem: if X. Y 独立, 则 X. Y必不相关

Theorem: $-1 \le Corr(X,Y) \le 1$

四. Multinomial Distribution

Story: n objects → k类，每个object进入第j类的概率为p_j，$\sum_{j=1}^{k} p_j = 1$. 令X_i是i类中objects个数，则$\sum_i X_i = n$.

则：$X = (X_1, \cdots, X_n)$ is said to have the Multinomial Distribution with parameters n and $\vec{p} = (p_1, \cdots, p_k)$

$$X \sim Mult_k(n, \vec{p})$$

可见：PMF：$P(X_1 = n_1, X_2 = n_2 \cdots) = \frac{n!}{n_1! \, n_2! \cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$

Proof: $\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n_k}{n_k} = \frac{n!}{n_1! \, n_2! \cdots n_k!}$

而每个 $\binom{n-\sum_{i=1}^{j-1} n_i}{n_j}$ 发生概率：$p_j^{n_j}$

Theorem: If $X \sim Mult_k(n, \vec{p})$，则 $X_j \sim Bin(n, p_j)$.

单独挑X_j出来，1个object进j类：p_j，不进则为$1-p_j$.

则 equivalent to Bern(p_j)

Theorem: Multinomial Lumping：$X \sim Mult_k(n, p)$

则 $(X_1 + X_2, X_3, \cdots) \sim Mult_{k-1}(n, (p_1+p_2, p_3, \cdots))$

Theorem: Multinomial Conditioning：

If $X \sim Mult_k(n, p)$，则：

$$(X_2, \cdots, X_k) \mid X_1 = n_1 \sim Mult_{k-1}(n-n_1, (p_2', \cdots, p_k'))$$

其中 $p_j' = p_j / (p_2 + \cdots + p_k)$

该条定理在直觉上很容易认同

Date . .

Theorem: Covariance in a Multinomial:

Let $(X_1, \cdots, X_k) \sim Mult_k(n, \vec{p})$, $\vec{p} = (p_1, \cdots, p_k)$, 则 $i \neq j$,
有: $Cov(X_i, X_j) = -np_ip_j$

Proof: $X_1 \sim Bin(n, p_1)$ $\quad X_2 \sim Bin(n, p_2)$

则 $X_1 + X_2 \sim Bin(n, p_1+p_2)$

且: for $Z \sim Bin(n, p)$, $Var(Z) = np(1-p)$

则 $Var(X_1+X_2) - Var(X_1) - Var(X_2) = 2Cov(X_1, X_2)$

$\Rightarrow \quad Cov(X_1, X_2) = -np_1p_2$

五 Multivariate Normal

Def: $X = (X_1, \cdots, X_k)$ is said to be Multivariate Normal (MVN) distribution if every Linear combination of X_j has a normal distribution.

$K=2$, then X is called Bivariate Normal (BVN)

⚠ 不意味着: 只要 X_j 都为正态分布, 那么 X 就是BVN了!

Theorem: If (X_1, X_2, X_3) 是BVN. 则 (X_1, X_2) 也是

Theorem: If $X = (X_1, \cdots, X_n)$ 与 $Y = (Y_1, \cdots, Y_m)$ 是 MVN, 且 X.Y 间独立, 那么: $W = (X_1, \cdots, X_n, Y_1, \cdots, Y_m)$ 也是MVN

那么关于MVN还有许多性质与属性:

① mean vector $(\mu_1, \cdots, \mu_k)$. $E(X_j) = \mu_j$

② covariance matrix, $M[i][j]$ 为 $Cov(X_i, X_j)$

而对于 $f_{X_1,X_2}(x_1, x_2)$, i.e., X_1. X_2 两个正态分布的连合PDF:

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x_1^2+x_2^2-2\rho x_1x_2)\right\}$$

其中, $\rho \in (-1, 1)$ 为相关系数, 该式更多说明将在之后进行

$\rho = Corr(X_1, X_2)$

KOKUYO

扫描全能王 创建

No.

Date

· Def: 联合矩母函数: $X=(X_1,\cdots,X_k)$: 有 $t=(t_1,\cdots,t_k)$

$M(t)=E(e^{t'X})=E(e^{t_1X_1+\cdots+t_kX_k})$

而对于 正态 r.v. W: $E(e^W)=e^{E(W)+\frac{1}{2}Var(W)}$

$\therefore E(e^{t_1X_1+\cdots+t_kX_k})=exp\left(t_1E(X_1)+\cdots+t_kE(X_k)+\frac{1}{2}Var(t_1X_1+\cdots+t_kX_k)\right)$

Theorem: 在MVN向量内, uncorrelated ⇒ independent

That is: if $X\sim MVN$, $\vec{X}=(\vec{X_1},\vec{X_2})$, $\forall X_i\in\vec{X_i}$, $X_j\in\vec{X_j}$,

$Corr(X_i,X_j)=0$, 则 $\vec{X_1},\vec{X_2}$ 独立.

特别地: if (X,Y) is a BVN, $Corr(X,Y)=0$, 则 X,Y 独立

proof: Lemma: if X,Y 独立, 则: $M_{X,Y}(s,t)=M_X(s)\cdot M_Y(t)$

$M_{X,Y}(s,t)=e^{s\mu_1+t\mu_2+\frac{1}{2}s^2\sigma_1^2+\frac{1}{2}s^2\sigma_2^2+2\rho st\sigma_1\sigma_2}$

而 $(X,Y)\sim BVN$, $\rho=Corr(X,Y)=0$, 则 $M_{X,Y}(s,t)=e^{s\mu_1+\frac{1}{2}s^2\sigma_1^2}e^{s\mu_2+\frac{1}{2}s_2^2\sigma_2^2}$

$\therefore M_{X,Y}(s,t)=M_X(s)\cdot M_Y(t)$.

△: 原本, $Corr(X,Y)=0$ 推不出 X,Y 独立, 但由于 $(X,Y)\sim BVN$.

则 $M_{X,Y}(s,t)=E(e^{sX+tY})$

换而言之, 若 (X,Y) 非BVN, 则 $sX+tY$ 不一定是正态分布!

· BVN Generation: 设有 i.i.d. r.v.s $X,Y\sim N(0,1)$, 但欲生成

BVN (Z,W), $Corr(Z,W)=\rho$ 且 $Z,W\sim N(0,1)$.

则欲有 a.b.c.d: $Z=aX+bY$ $W=cX+dY$.

则 $E(Z)=E(W)=0$, $\begin{cases}Var(Z)=a^2+b^2=1\\Var(W)=c^2+d^2=1\end{cases}$

$\rho=\dfrac{Cov(Z,W)}{\sqrt{Var(Z)Var(W)}}=Cov(Z,W)=Cov(aX+bY,cX+dY)=\rho$

$\because X,Y$ 独立 $\therefore Cov(X,Y)=0$

$\therefore acCov(X,X)=ac$, $bdCov(Y,Y)=bd$.

$ac+bd=\rho$

有了 $\begin{cases} a^2+b^2=1 \\ c^2+d^2=1 \\ ac+bd=\rho \end{cases}$ 找一个解即可。$b=0 \Rightarrow a^2=1$. 令 $a=1$, $d=\sqrt{1-\rho^2}$

则: $\begin{cases} Z=X \\ W=\rho X+\sqrt{1-\rho^2}\,Y \end{cases}$ 这是很重要的结论!!

六. Change of Variables:

· Theorem: $Y=g(X)$, 则 $f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|$

Proof: w.l.o.g: g 单增. $x=g^{-1}(y)$, 则:

$F_Y(y)=P(Y\le y)=P(X\le g^{-1}(y))=F_X(g^{-1}(y))=F_X(x)$

$f_Y(y)=F_Y'(y)\overset{\text{链}}{=}\frac{dF_X(x)}{dx}\cdot\frac{dx}{dy}=f_X(x)\left|\frac{dx}{dy}\right|$

g 单减, 同理且正确 (有负号在 dx/dy, 但 abs 去掉了)

$\therefore$ $g(y)$ 可拆成一段段 $g_i(y_i)$, s.t. $y_i\in \underset{(\in R)}{\text{range}}$, $g_i(y_i)\uparrow$ or $\downarrow$

· Theorem: $\vec{X}=(X_1,\cdots,X_n)$ $\vec{Y}=g(X)$ $g: R^n\to R^n$.

$\frac{\partial x}{\partial y}=\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n}\end{bmatrix}$, 则 $f_{\vec{Y}}(y)=f_{\vec{X}}(x)\left|\frac{\partial x}{\partial y}\right|=f_{\vec{X}}(x)\frac{1}{\left|\frac{\partial y}{\partial x}\right|}$

△: Jacobian paradimn 仅对连续 r.v. 有效

例: Box-Muller: $U\sim Unif(0,2\pi)$. $T\sim Expo(1)$. T, U 独立

令 $X=\sqrt{2T}\cos U$, $Y=\sqrt{2T}\sin U$. 则 PDF of X&Y:

$f_{U,T}(u,t)=\frac{1}{2\pi}\cdot e^{-t}$, $\left|\frac{\partial(x,y)}{\partial(u,t)}\right|=-1$

$\therefore f_{X,Y}(x,y)=\frac{1}{2\pi}e^{-t}\cdot|-1|=\frac{1}{2\pi}\cdot e^{-\frac{1}{2}(x^2+y^2)}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$

$\therefore$ X, Y 独立且 X, Y ~ i.i.d. $N(0,1)$

No.

Date

t. Convolutions.

Theorem: 若 X.Y 为离散 r.v.s，则 $T=X+Y$:

$$P(T=t)=\sum_x P(Y=t-x)P(X=x)=\sum_y P(X=t-y)P(Y=y)$$

若连续: $f_T(t)=\int_{-\infty}^{+\infty} f_Y(t-x)f(x)dx=\int_{-\infty}^{+\infty} f_X(t-y)f_Y(y)dy$

△: 核心: LOTP

例: X. Y i.i.d Expo(λ). $T=X+Y$?

$$f_T(t)=\int_{-\infty}^{+\infty} f_Y(t-x)f_X(x)dx=\int_0^t \lambda e^{-\lambda(t-x)}\lambda e^{-\lambda x}dx$$

$$=\int_0^t \lambda^2 e^{-\lambda t}dx=\lambda^2 t e^{-\lambda t}.$$