概率统计第九讲：条件分布

史灵生　清华数学系

1、本讲提要

1. 条件分布
   - 离散型
   - 连续型
2. 条件特征数
   - 条件数学期望
   - 重期望公式

2、条件分布

设 二维离散型随机变量(X,Y)的联合分布列为：

$$P(X=x_i,Y=y_j)=p_{i,j}.$$

则 $P(X=x_i|Y=y_j)=\dfrac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\dfrac{p_{i,j}}{p_j^Y}=\dfrac{p_{i,j}}{\sum\limits_k p_{k,j}}$。

定义

称

X	x_1	$\cdots$	x_i	$\cdots$
P	$\frac{p_{1j}}{p_j^Y}$	$\cdots$	$\frac{p_{ij}}{p_j^Y}$	$\cdots$

为给定$Y=y_j$条件下X的条件分布列（*conditional distribution law*），记为$P_{X|Y}(x_i|y_j)=P(X=x_i|Y=y_j)$。

3、二项分布和Poisson分布

例

若$X \sim P(\lambda)$与$Y \sim P(\mu)$独立,则$X|X+Y=n \sim b(n, \lambda/(\lambda+\mu))$.

证明:

由Poisson分布的可加性得，$X+Y \sim P(\lambda+\mu)$。

$$
\begin{aligned}
P_{X|X+Y}(k|n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)} \\
&= \frac{P(X=k)P(Y=n-k)}{P(X+Y=n)} \\
&= \frac{\lambda^k}{k!}e^{-\lambda}\frac{\mu^{n-k}}{(n-k)!}e^{-\mu} \Big/ \frac{(\lambda+\mu)^n}{n!}e^{-(\lambda+\mu)} \\
&= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{n-k}.
\end{aligned}
$$

4、Poisson分布在随机选择下的不变性

例

设某商店的顾客数$X \sim P(\lambda)$，每位顾客是男性的概率为p，则男顾客数$Y \sim P(\lambda p)$。

$$
\begin{aligned}
P(Y=k) &= \sum_{n=0}^{\infty} P(Y=k \mid X=n) P(X=n) \\
&= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \lambda^n e^{-\lambda}/n! \quad \left(Y|X=n \sim b(n,p)\right) \\
&= \frac{(\lambda p)^k e^{-\lambda}}{k!} \sum_{n=k}^{\infty} [\lambda(1-p)]^{n-k}/(n-k)! \\
&= \frac{(\lambda p)^k}{k!} e^{-\lambda} e^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!} e^{-\lambda p}.
\end{aligned}
$$

5、条件分布的例

例

掷一硬币出现正面的概率为p，独立抛掷n次。求n次抛掷中仅出现了一次正面的条件下，首次正面出现在第k次的条件概率。

解：

- 记X为正面首次出现时抛掷的次数，
- 记Y为n次抛掷中正面出现的次数，则$Y \sim b(n,p)$。
- $$\begin{aligned} P_{X|Y}(k|1) &= P(X=k|Y=1) \\ &= P(X=k, Y=1)/P(Y=1) \\ &= p(1-p)^{n-1}/[np(1-p)^{n-1}] \\ &= 1/n。\end{aligned}$$

问 $P_{X|Y}(k|m) = ?$

6、条件分布的例

例

袋中装有a个白球，b个黑球，每次取出一球后，总是放入一个白球。问在第n次取白球的情况下，第$n+1$次取的还是白球的概率与第n次取白球的概率相同吗？

解：

- 记X_n为第n次取的是白球的示性函数，
- 记Y为第$n-1$次后袋中白球数。
- $P_{X_n|Y}(1|k) = k/(a+b)$。
- $$P(X_n=1) = \sum_k P(Y=k)P_{X_n|Y}(1|k) = \sum_k P(Y=k)k/(a+b) = EY/(a+b)。$$
- $EY = (a+b)P(X_n=1)$。

7、条件分布的例

已经得到$P_{X_n|Y}(1|k)=k/(a+b)$和$EY=(a+b)P(X_n=1)$;

$$\begin{aligned}
P_{X_{n+1}|X_n}(1|1) &= P(X_{n+1}=1\mid X_n=1)\\
\text{(全概公式)} &= \sum_k P_{Y|X_n}(k|1)P(X_{n+1}=1\mid X_n=1,Y=k)\\
\text{(Bayes公式)} &= \sum_k [P_{X_n|Y}(1|k)P(Y=k)/P(X_n=1)]k/(a+b)\\
&= \sum_k k^2P(Y=k)/[(a+b)^2P(X_n=1)]\\
&= EY^2/[(a+b)^2P(X_n=1)]\\
\text{(方差公式)} &= [(EY)^2+Var(Y)]/[(a+b)^2P(X_n=1)]\\
&= P(X_n=1)+Var(Y)/[(a+b)^2P(X_n=1)]\\
&\neq P(X_n=1),\quad \text{当}n>1\text{时。}
\end{aligned}$$

8、条件分布

定义

- 设X是概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，事件$A \in \mathcal{F}$满足$P(A) > 0$。则称

$$F_{X|A} : \mathbb{R} \to \mathbb{R}, \qquad F_{X|A}(x) = P(X \leq x|A) = \frac{P(X \leq x, A)}{P(A)}$$

  为在已知A发生的情况下X的条件分布函数。
- 如果存在非负可积函数$p_{X|A} : \mathbb{R} \to \mathbb{R}$使得条件分布函数

$$F_{X|A}(x) = \int_{-\infty}^{x} p_{X|A}(u)\mathrm{d}u,$$

  则称$p_{X|A}$为在已知A发生的情况下X的条件密度函数。

9、条件分布

- 对连续型随机变量(X,Y)，我们希望研究在已知$Y=y$的条件下X的概率分布。
- 但是由于Y是连续型随机变量，$P(Y=y)=0$，
- 所以这时的条件分布不能通过传统的条件概率直接得到。
- 为此我们先放宽条件事件的限制，考虑Y落在y附近一个小区间$[y,y+\Delta y)$ 内的情形。
- 如果p_Y在y连续,并且$p_Y(y)>0$,则$P(y\le Y<y+\Delta y)>0.$

$$P(X\le x\,|\,y\le Y<y+\Delta y)=\frac{P(X\le x,y\le Y<y+\Delta y)}{P(y\le Y<y+\Delta y)}$$

$$=\frac{F(x,y+\Delta y)-F(x,y)}{F_Y(y+\Delta y)-F_Y(y)}=\frac{[F(x,y+\Delta y)-F(x,y)]/\Delta y}{[F_Y(y+\Delta y)-F_Y(y)]/\Delta y}$$

$$\xrightarrow{\Delta y\to 0}\frac{\frac{\partial F(x,y)}{\partial y}}{F'_Y(y)}=\frac{\int_{-\infty}^{x}p_{X,Y}(u,y)\mathrm{d}u}{p_Y(y)}=\int_{-\infty}^{x}\frac{p_{X,Y}(u,y)}{p_Y(y)}\mathrm{d}u.$$

10、条件分布

定义

- 设(X,Y)是连续型随机变量，p_Y在y连续且$p_Y(y)>0$。则称
  $F_{X|Y}(\cdot|y):\mathbb{R}\to\mathbb{R}, \qquad F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{p_{X,Y}(u,y)}{p_Y(y)}\mathrm{d}u$
  为在已知$Y=y$发生的情况下X的条件分布函数。
- 在已知$Y=y$发生的情况下X的条件密度函数为：
  $$p_{X|Y}(x|y)=\frac{p_{X,Y}(x,y)}{p_Y(y)}。$$

1. 乘法公式的密度形式：$p_{X,Y}(x,y)=p_Y(y)p_{X|Y}(x|y)$。
2. 全概公式的密度形式：$p_X(x)=\int_{-\infty}^{\infty}p_Y(y)p_{X|Y}(x|y)\mathrm{d}y$。
3. Bayes公式的密度形式：$p_{X|Y}(x|y)=\frac{p_X(x)p_{Y|X}(y|x)}{\int_{-\infty}^{\infty}p_X(x)p_{Y|X}(y|x)\mathrm{d}x}$。
4. 如果$p_{X,Y}(x,y)=g(y)h(x,y)$，而且对任意给定的y，$h(\cdot,y)$非负可积且$\int_{-\infty}^{\infty}h(x,y)\mathrm{d}x=1$，则
   $$g(y)=p_Y(y),\quad h(x,y)=p_{X|Y}(x|y)$$

11、条件数学期望

定义

1. 如果(X, Y)是概率空间$(\Omega, \mathcal{F}, P)$上的离散型随机变量,$y \in \mathbb{R}$满足$P(Y = y) > 0$，则如果级数

$$\sum_x x P_{X|Y}(x|y)$$

绝对收敛，则称它的值为X在已知$Y = y$的条件下的条件数学期望（*conditional expectation*），记为$E(X|Y = y)$。

2. 不难证明：若EX存在，则对满足$P(Y = y) > 0$的任意y，$E(X|Y = y)$存在。将它看成y的函数，记为

$$f(y) = E(X|Y = y),$$

则定义随机变量$f(Y)$为X对Y的条件数学期望（*conditional expectation*），记为$E(X|Y)$。

12、条件期望

定义

- 设(X, Y)是连续型随机变量，联合概率为p，$y \in \mathbb{R}$满足$p_Y(y) > 0$。如果

$$\int_{-\infty}^{+\infty} x p_{X|Y}(x|y)\mathrm{d}x$$

绝对收敛，则称之为X在给定条件$Y = y$下的条件数学期望，记为$E(X|Y = y)$。

- X关于Y的条件数学期望为：

$$E(X|Y) : \Omega \to \mathbb{R}, \quad E(X|Y)(\omega) = E[X|Y = Y(\omega)].$$

13、正态分布条件期望与方差

$(X, Y) \sim N(0, 0, 1, 1, \rho)$

- $p_{X|Y}(x|y) = p(x, y)/p_Y(y) = g(x, y \quad \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}}$；
- $X|Y = y \sim N(\rho y, 1 - \rho^2)$;
- $E(X|Y = y) = \rho y, \quad E(X|Y) = \rho Y$；
- $Var(X|Y) = 1 - \rho^2$。

$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

- $X^* = (X - \mu_1)/\sigma_1$，$Y^* = (Y - \mu_2)/\sigma_2$;
- $(X^*, Y^*) \sim N(0, 0, 1, 1, \rho)$，$X^*|Y^* = y \sim N(\rho y, 1 - \rho^2)$;
- $(X^*|Y = y) = (X^*|Y^* = \frac{y-\mu_2}{\sigma_2}) \sim N\left(\rho(y - \mu_2)/\sigma_2, 1 - \rho^2\right)$;
- $$(X|Y = y) = (\sigma_1 X^* + \mu_1 | Y = y) \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2), (1 - \rho^2)\sigma_1^2\right);$$
- $E(X|Y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(Y - \mu_2)$;
- $Var(X|Y) = (1 - \rho^2)\sigma_1^2$。

14、条件期望性质

由于条件数学期望是条件分布下的数学期望，故它也满足数学期望的常见性质如线性性质等，而外我们还有

定理（重期望公式）

若EX存在，则$EX = E[E(X|Y)]$。

证明：（离散型）

$$EX = \sum_{x,y} xP(X = x, Y = y) = \sum_{x,y} xP(Y = y)P(X = x|Y = y)$$
$$= \sum_{y} \left[\sum_{x} xP(X = x|Y = y) \right] P(Y = y)$$
$$= \sum_{y} E(X|Y = y)P(Y = y) = E[E(X|Y)]。$$

15、重期望公式

证明：（连续型）

$$
\begin{aligned}
EX &= \iint_{\mathbb{R}^2} xp(x,y)\mathrm{d}x\mathrm{d}y = \iint_{\mathbb{R}^2} xp_{X|Y}(x|y)p_Y(y)\mathrm{d}x\mathrm{d}y \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} xp_{X|Y}(x|y)\mathrm{d}x \right) p_Y(y)\mathrm{d}y \\
&= \int_{\mathbb{R}} E(X|Y=y)p_Y(y)\mathrm{d}y \\
&= \mathrm{E}[\mathrm{E}(X|Y)]。
\end{aligned}
$$

16、重期望公式推论

推论

重期望公式对任何随机变量都成立。例：设A为事件，

1. 则当X为连续型，$Y=I_A$为离散型时，得到全概率公式
$$P(A)=EI_A=E[E(I_A|X)]=\int_{-\infty}^{\infty}P(A|X=x)p_X(x)\mathrm{d}x.$$
2. 若$P(A)>0$，则$E(X|A)=E(XI_A)/EI_A=E(XI_A)/P(A)$。

证明：（2）

$$\begin{aligned}E(XI_A)&=E[E(XI_A|I_A)]=E(XI_A|A)P(A)+E(XI_A|\bar{A})P(\bar{A})\\&=E(X|A)P(A)。\end{aligned}$$

17、Bagdad窃贼问题

例

从A地出发等可能地选择三条道路：沿第一条路走t_1小时可到B地，沿第二条路走t_2小时回到A，沿第三条路走t_3小时回到A。问从A到B平均用多长时间？

解：

设 X为从A到B所用的时间，Y为选择的路的编号，则

$$\mathrm{EX} = \mathrm{E}[\mathrm{E}(X|Y)] = \sum_{k=1}^{3} E(X|Y=k)P(Y=k)$$

$$= \frac{1}{3}t_1 + \frac{1}{3}(t_2 + EX) + \frac{1}{3}(t_3 + EX),$$

故 $\mathrm{E}X = t_1 + t_2 + t_3$。

18、电力供应

例

一工厂的利润Z以如下方式决定于生产耗电量Y和供电量X，

$$Z = \begin{cases} 30Y, & 若 Y \le X; \\ 30X + 10(Y - X), & 若 Y > X。\end{cases}$$

其中$X \sim U(10, 30)$，$Y \sim U(10, 20)$。求EZ。

- 当$20 < x < 30$时，$E(Z|X = x) = \int_{10}^{20} 30y p_Y(y) \mathrm{d}y = 450$；
- 当$10 < x < 20$时，

$$E(Z|X = x) = \int_{10}^{x} 30y p_Y(y) \mathrm{d}y + \int_{x}^{20} [30x + 10(y - x)] p_Y(y) \mathrm{d}y$$
$$= 50 + 40x - x^2。$$

$$EZ = \mathrm{E}[\mathrm{E}(Z|X)] = \int_{10}^{30} E(Z|X = x) p_X(x) \mathrm{d}x$$
$$= \int_{10}^{20} (50 + 40x - x^2) \frac{1}{20} \mathrm{d}x + \int_{20}^{30} 450 \cdot \frac{1}{20} \mathrm{d}x \approx 433。$$

19、分支过程

例

一个家族第n代男性成员有X_n个人，$X_0 = 1$。假设这个家族中每个男性成员的儿子的个数是独立同分布的随机变量。求这个家族第n代的平均男性成员数。

- 记Y_k $(k = 1, \ldots, X_1)$是第一代的第k个男性成员的儿子数。
- 则$X_2 = \sum\limits_{k=1}^{X_1} Y_k$，其中$X_1, Y_1, Y_2, \ldots$独立同分布。
- 故$EX_1 = EY_k =: a \quad \forall k \in \mathbb{N}$。

$$E(X_2|X_1 = n) = E\left(\sum_{i=1}^{X_1} Y_i \middle| X_1 = n\right) = E\left(\sum_{i=1}^{n} Y_i\right) = nEY_1 = na,$$

$$EX_2 = E[E(X_2|X_1)] = \sum\nolimits_n naP(X_1 = n) = a\sum\nolimits_n nP(X_1 = n) = a^2.$$

由数学归纳可得$EX_n = a^n$。

20、灭绝概率

例（灭绝概率）

若$a<1$，则当$n\to\infty$时，灭绝概率$P(X_n=0)\to 1$。

证明：

- $P(X_n>0)\leq EX_n=a^n\to 0$ as $n\to\infty$，
- $P(X_n=0)=1-P(X_n>0)\geq 1-a^n\to 1$。

思考题

若$a\geq 1$，则当$n\to\infty$时，灭绝概率$P(X_n=0)\to$?

21、赌徒输光问题续

记T_i表示甲最初有i元赌博所需的时间，求$t_i = E(T_i)$。

- 记A表示第一次甲赢。则

$$\begin{aligned} t_i &= \mathrm{E}(T_i) = E[E(T_i|I_A)] = E(T_i|A)P(A) + E(T_i|\bar{A})P(\bar{A}) \\ &= p\big(1 + ET_{i+1}\big) + q\big(1 + ET_{i-1}\big) \\ &= 1 + pt_{i+1} + qt_{i-1}, \qquad 0 < i < a, \end{aligned}$$

- 从而：$t_i + \frac{i}{p-q} = p\big(t_{i+1} + \frac{i+1}{p-q}\big) + q\big(t_{i-1} + \frac{i-1}{p-q}\big)$。
- 注意到边界条件：$t_0 = 0, \quad t_a = 0,$
- 可得：$t_i = \begin{cases} \frac{1}{p-q}\left[a\frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^a} - i\right], & (p \neq q) \\ i(a-i), & (p = q = \frac{1}{2}). \end{cases}$
- 当$p = q = 1/2$时，上述结论表明：只要赌博是公平的，面对具有无穷财富的赌场，虽然赌客几乎注定要破产，但赌博平均用时为无穷，因此我们能看到有赌客从赌场赢钱。

22、Poisson分布和指数分布

例

一商店在t时刻的男顾客数是服从参数为λt的Poisson分布，女顾客数是服从参数为μt的Poisson分布，且相互独立。则第一位顾客是男性的概率为$\lambda/(\lambda+\mu)$。

设第一位男、女顾客的到达时刻分别为X, Y，则X, Y分别服从参数为λ, μ的指数分布，并且相互独立。

$$
\begin{aligned}
P(X<Y) &= \int_{-\infty}^{+\infty} P(X<Y|Y=y)p_Y(y)\mathrm{d}y \quad (\text{全概率公式}) \\
&= \int_{-\infty}^{+\infty} P(X<y)p_Y(y)\mathrm{d}y \quad (\text{独立性}) \\
&= \int_0^{+\infty} (1-e^{-\lambda y})\mu e^{-\mu y}\mathrm{d}y \\
&= \lambda/(\lambda+\mu).
\end{aligned}
$$