概率论:Week 8

1 事件的独立性

条件概率定义.

定义 2.1.1 设$(\Omega,\mathscr{F},P)$是一个概率空间,$B\in\mathscr{F}$,而且$P(B)>0$,则对任意$A\in\mathscr{F}$,记

$$P(A\mid B)=\frac{P(AB)}{P(B)} \tag{2.1.1}$$

并称$P(A|B)$为**在事件 B 发生的条件下事件 A 发生的条件概率**(conditional probability).

若未经特别指出,今后出现条件概率$P(A|B)$时,都假定$P(B)>0$.

条件概率也是概率.

非负性、规范性、可列可加性.

(i) $P(A|B)\geqslant 0$;

(ii) $P(\Omega|B)=1$;

(iii) $P\left(\sum_{i=1}^{\infty}A_i\mid B\right)=\sum_{i=1}^{\infty}P(A_i\mid B)$.

性质 (概率乘法公式)**.**

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1\cdots A_{n-1})$$

例 1 (Polya urn scheme)**.** *每次抽出后放回, 并多放入t个相同颜色的球.*

证明. 假设有r个红球, b个黑球. 记B_m是第m次抽中黑球. 断言: $P(B_m)=\frac{b}{b+r}$.

由数学归纳法, 当$m=1$时, 显然. 若$m=k$时成立, 则当$m=k+1$时, 关于事件B_1取条件,

$$P(B_{k+1}) = P(B_{k+1} \mid B_1)P(B_1) + P(B_{k+1} \mid B_1^c)P(B_1^c).$$

由于,

$$\begin{aligned} P(B_{k+1} \mid B_1) &= P(B_k) \quad \text{starting with b+c black and r red balls} \\ &= \frac{b+c}{b+r+c} \quad \text{by inductive assumption.} \end{aligned}$$

同理

$$\begin{aligned} P(B_{k+1} \mid B_1^c) &= P(B_k) \\ &= \frac{b}{b+r+c} \end{aligned}$$

因此,

$$\begin{aligned} P(B_{k+1}) &= \frac{b+c}{b+r+c}\frac{b}{b+r} + \frac{b}{b+r+c}\frac{r}{b+r} \\ &= \frac{b}{b+r}. \end{aligned}$$

□

注记. *抽签不但与顺序无关, 抽签还与抽签方法无关. 有放回抽样就是$t=0$, 无放回抽样就是$t=-1$, polya urn scheme为一般情况.*

性质 (全概率公式)**.**

$$P(B) = \sum_{i=1}^{\infty} P(A_i)P(B|A_i)$$

- $P(A_i)$: 先验概率
- $P(A_i|B)$: 后验概率

例 2 (Noise Channel Model)**.**

$$\begin{aligned} P(A_0 \mid B) &= \frac{P(A_0)P(B \mid A_0)}{P(A_0)P(B \mid A_0) + P(A_1)P(B \mid A_1)} \\ &\propto P(A_0)P(B \mid A_0) \end{aligned}$$

应用:spell-checking, machine translation...

$$\hat{E} = \operatorname*{argmax}_{E \in \mathit{English}} \overbrace{P(F \mid E)}^{\text{translation model}} \quad \overbrace{P(E)}^{\text{language model}}$$

2 事件的独立性

定义 (事件A与B相互独立)**.**

$$P(AB) = P(A)P(B)$$

注记. *若$P(B)>0$, 还有$P(A|B)=P(A)$.*

注记. 若事件 A 与 B 独立,则下列各对事件也相互独立:

$$\{\bar{A}, B\}, \{A, \bar{B}\}, \{\bar{A}, \bar{B}\}$$

例 3 (两两独立不蕴含相互独立)**.** 考虑一个均匀的四面骰子.

样本空间 $\Omega = \{1, 2, 3, 4\}$, 对每个 $\omega \in \Omega$ 发生的概率是相等的:

$$\forall \omega \in \Omega : Pr(\omega) = 1/4$$

考虑以下的事件: *S={A,B,C}* 其中: *A={1,2},B={1,3},C={1,4}* 我们有:

$$Pr(A) = Pr(B) = Pr(C) = 12$$

还有: $Pr(A \cap B)=Pr(A \cap C)=Pr(B \cap C)=Pr(\{1\})=1/4$ 因此

$$Pr(A)Pr(B) = Pr(A \cap B)Pr(A)Pr(C) = Pr(A \cap C)Pr(B)Pr(C) = Pr(B \cap C)$$

因此事件*A,B,C*是两两独立的. 再考虑: $Pr(A \cap B \cap C)=Pr(\{1\})=1/4$ 但是: $Pr(A)Pr(B)Pr(C)=1/8 \neq Pr(A \cap B \cap C)$

3 试验的独立性

要用到乘积测度空间.

令$(\Omega, \mathscr{F}, \mu)$, $(\tilde{\Omega}, \tilde{\mathscr{F}}, \nu)$是两个有限测度空间(概率空间). 令$\Omega \times \tilde{\Omega} := \{(\omega, \tilde{\omega}) : \omega \in \Omega, \tilde{\omega} \in \tilde{\Omega}\}$. $\mathscr{F} \times \tilde{\mathscr{F}}$: 以$\Omega \times \tilde{\Omega}$为全集的$\sigma$代数. 称可测矩体是: $E \subset \Omega \times \tilde{\Omega}$, 若$\exists A \in \mathscr{F}, B \in \tilde{\mathscr{F}}$使得$E = A \times B$.

性质. 任意的可测矩形E, $(\mu \times \nu)(E) := \mu(A)\nu(B)$, $\mathscr{F} \times \tilde{\mathscr{F}} = \sigma(\{$可测矩形的全体$\})$. 定义截口*(section)*:$\forall E \in \mathscr{F} \times \tilde{\mathscr{F}}, \forall \omega \in \Omega$, 令$E^{\omega} := \{\tilde{\omega} \in \tilde{\Omega}, (\omega, \tilde{\omega}) \in E\} \subset \tilde{\Omega}$, $E^{\tilde{\omega}} := \{\omega \in \Omega, (\omega, \tilde{\omega}) \in E\} \subset \Omega$, 则$E^{\tilde{\omega}} \in \mathscr{F}, E^{\omega} \in \tilde{\mathscr{F}}$

定义. 若对于任意的

$$A^{(1)} \in \mathscr{A}_1, A^{(2)} \in \mathscr{A}_2, \cdots, A^{(n)} \in \mathscr{A}_n$$

均成立

$$P(A^{(1)}A^{(2)} \cdots A^{(n)}) = P(A^{(1)})P(A^{(2)}) \cdots P(A^{(n)})$$

则称试验 $E_1, E_2, \cdots, E_n$是相互独立的.

注记. 所谓独立重复试验就是指$\Omega_1 = \cdots = \Omega_n$.

注记. *Bernoulli*试验就是指$\mathscr{F} = \{\phi, A, A^c, \Omega\}$.

- 每次试验至多出现两个可能结果之一 : A 或 $\bar{A}$;
- A 在每次试验中出现的概率 p 保持不变;
- 各次试验相互独立;
- 共进行 n 次试验.

n重Bernoulli试验段样本点是: $(\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_n)$. 显然这种样本点共$2^n$个, 是有限样本空间.

4 Bernoulli概型中的分布

4.1 Bernoulli分布(两点分布)

只进行一次Bernoulli试验.

4.2 二项分布

n重Bernoulli试验中事件A发生的次数

$$b\left(k;n,p\right)=\binom{n}{k}p^kq^{n-k},k=0,1,2,\cdots,n$$

4.3 几何分布

首次成功出现在第k次试验.

$$g\left(k;p\right)=q^{k-1}p,\quad k=1,2,\cdots$$

4.4 Pascal分布

第r次成功出现在第k次试验.

$$f(k;r,p)=\binom{k-1}{r-1}pq^{k-r},\quad k=r,r+1,\cdots$$

4.5 多项分布

n次Bernoulli试验出现k_1次事件A_1, ..., k_r次事件A_r. 发生的概率是

$$\frac{n!}{k_1!k_2!\cdots k_r!}p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}$$

例 4 (随机游走)**.** *假设一口井深10米, 一只青蛙从井底向上跳, 每次跳1米*

- *当它在井口时, 它就跳出去.*
- *当它在井底时, 有0.5的概率往上跳一次, 0.5的概率不动;*
- *当它既不在井口, 也不在井底时, 它有0.5的概率往上跳一次, 0.25的概率不动, 0.25的概率往下跳一次.*

求它从井底跳出去的概率.

4.6 Poisson极限定理

回顾P102, 教材上给出了二项分布的Poisson逼近, 事实上该定理是本书中较早提到的一种极限定理, 它往往被称为**Poisson极限定理**或**Law of Rare Events**.

定理 (Law of Rare Events)**.** 设p_n是$[0,1]$上的实数列, 且np_n收敛于有限的极限λ, 则有

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=e^{-\lambda}\frac{\lambda^k}{k!}$$

由于后续在证明De Moivre-Laplace定理时使用了Stirling公式, 这里给出一个使用Stirling公式证明上述定理的过程.

证明. 由Stirling公式,

$$n!\sim\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

以下简写$p=p_n$,

$$\begin{aligned}\binom{n}{k}p^k(1-p)^{n-k}&=\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}\\&\approx\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}k!}p^k(1-p)^{n-k}\\&=\sqrt{\frac{n}{n-k}}\frac{n^ne^{-k}}{(n-k)^{n-k}k!}p^k(1-p)^{n-k}\end{aligned}$$

由序列的条件, $np\to\lambda$, $\sqrt{\dfrac{n}{n-k}}\to 1$, 当$n\to\infty$,

$$\begin{aligned}\binom{n}{k}p^k(1-p)^{n-k}&\approx\frac{n^np^k(1-p)^{n-k}e^{-k}}{(n-k)^{n-k}k!}\\&=\frac{n^n\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}e^{-k}}{n^{n-k}\left(1-\frac{k}{n}\right)^{n-k}k!}\\&=\frac{\lambda^k\left(1-\frac{\lambda}{n}\right)^{n-k}e^{-k}}{\left(1-\frac{k}{n}\right)^{n-k}k!}\\&\approx\frac{\lambda^k\left(1-\frac{\lambda}{n}\right)^{n}e^{-k}}{\left(1-\frac{k}{n}\right)^{n}k!}\to e^{-\lambda}\frac{\lambda^k}{k!}\end{aligned}$$

□

4.7 Le Cam定理(拓展)

对于n个独立伯努利随机变量X_i, 满足$P(X_i=1)=p_i$(ind: independent, non-identically distributed). 记$S_n=X_1+\cdots+X_n$, Le Cam 给出了如下的上界估计不等式:

$$\sum_{k=0}^{\infty}\left|P\left(S_n=k\right)-e^{-\lambda}\lambda^k/k!\right|<2\sum_{i=1}^{n}p_i^2$$

其中$\lambda=p_1+\cdots+p_n$. 特别地, 若有$p_i=\lambda/n$, 则

$$\sum_{k=0}^{\infty}\left|P\left(S_n=k\right)-e^{-\lambda}\lambda^k/k!\right|<2\sum_{i=1}^{n}(\lambda/n)^2=\frac{\lambda^2}{n}$$

5 Poisson过程

设一个随机过程$N(t)$满足:

- 计数过程:$N(0)=0$;
- 独立增量性: $\forall 0<t_1<t_2\cdots<t_n, N(t_1), N(t_2)-N(t_1),\cdots,N(t_n)-N(t_{n-1})$是相互独立的 (即$N(t)$ 为增量独立的);
- 平稳增量性: $N(t)$与$N(t+s)-N(s)$同分布;
- 普通性: 存在$\lambda>0$,使得对任意的$t>0$和充分小的$h>0$ 有$P(N(t+h)-N(t)=1)=\lambda h+o(h)$且$P(N(t+h)-N(t)\geq 2)=o(h)$.

定理 (poisson过程和指数分布的关系)**.** *计数过程*$\{N(t),t\geq 0\}$*是泊松过程的充分必要条件是*$\{X_n,n\geq 1\}$*是独立且参数同为λ的指数分布.*

6 符号测度*

在微积分中, 如果函数f在区间$[a,b]$上连续, 那么 $F(x)=\int_a^x f(y)\mathrm{d}y$称为不定积分, 而$f$是$F$的导数.

设f是测度空间$(X,\mathscr{F},\mu)$上积分存在的可测函数, 我们自然地把$\mathscr{F}$上定义的集函数

$$\varphi(A)=\int_A f\mathrm{d}\mu,\quad \forall A\in\mathscr{F}$$

也称为不定积分, f也叫φ对测度μ的导数. 由积分的定义, 有$\varphi(\phi)=0$, 再由Levi定理, φ是可列可加的.

其正式定义如下:

定义. *设*$X(,\mathscr{F})$*是一个可测空间, 从*$\mathscr{F}$*到*$\bar{\mathbb{R}}$*的集函数*φ*被称为符号测度, 若它满足*

1. $\varphi(\phi)=0$*;*
2. *可列可加性: 对于任何两两不交的*$\{A_n\}\subset\mathscr{F}$*, 有*

$$\varphi\Big(\bigcup_{n=1}^{\infty}A_n\Big)=\sum_{n=1}^{\infty}\varphi(A_n).$$

定义 (RN导数)**.** *设*φ*是测度空间上的符号测度. 如果存在*$a.e.$*意义下唯一的可测函数*f*使得*

$$\varphi(A)=\int_A f\mathrm{d}\mu,\quad \forall A\in\mathscr{F}$$

则称f*是*φ*对于*μ*的Radon-Nikodym导数. 记之为*$\dfrac{d\varphi}{d\mu}=f$*.*

这就是符号测度的公理化定义。满足非负性的符号测度是测度。

什么样的符号测度才有RN导数呢? 这有当φ对μ绝对连续时才有可能(φ is dominated by μ).

定义 (绝对连续(dominated by))**.** 设ϕ, μ是可测空间上的符号测度和测度. 若$\mu(A) = 0 \Longrightarrow \varphi(A) = 0, \forall A \in \mathscr{F}$.则称$\varphi$对于$\mu$绝对连续, 记之为$\varphi \ll \mu$. *(.. is absolutely continuous with respect to ... or dominated by ...)*

有这样的结论: 只要μ是σ有限的, 而且$\varphi \ll \mu$, 则φ对μ的RN导数一定存在.

性质. 若φ是可测空间上的符号测度, μ是同一个可测空间上的σ有限测度, 当$\varphi \ll \mu$, 存在$(X, \mathscr{F}, \mu)$上*a.e.*唯一的可测函数f使得不定积分式成立, 且若φ是σ有限的, 则f是*a.e.*有限的.