Skills

一. 多变量问题，则只抓两个！

题欲证明："$f(p_1, p_2, \cdots, p_n)$ 当且仅当 $p_1 = p_2 = \cdots = p_n = ?$ 有min"

这种结论时，只抓 p_1 & p_2，证明在 $p_1' = p_2' = \frac{p_1+p_2}{2}$ 下有 f 较小值

i.e.， $f\left(\frac{p_1+p_2}{2}, \frac{p_1+p_2}{2}, p_3 \cdots\right) \le f(p_1, p_2, p_3 \cdots)$

其中 $p_3, \cdots, p_n$ 为常数，p_1, p_2 为变量

这样，欲min，则 $p_1 = p_3, p_2 = p_3, \cdots,$ $p_1 = \cdots = p_n$ 下有min

二. 用概率与期望证明存在性

关键思想：

好分数原则：设 X 为随机选择的某个对象的得分

如果 $E(X) \ge C$，则存在一个得分不小于 C 的对象

可能性原则：设 A 为一个事件表示在集合中随机选择的某个对象具有某种特定的性质。若 $P(A) > 0$，则存在这种性质的对象

例：100个人，分到15个规模为20人的组中，每个人服务于3个组。

证明：存在两个组，它们至少有3个公共的人。

解：X 表示为任取两个组，组内公共人的数量

令 $X = I_1 + I_2 + \cdots + I_{100}$，$I_j$ 代表第 j 个人同时在两个组里

$\therefore E(X) = 100E(I_j)$

而 $E(I_j) = P$(第 j 个人同时在两个组中) $= \frac{C_3^2}{C_{15}^2} = \frac{1}{35}$

$\therefore E(X) = \frac{100}{35} = \frac{20}{7}$

$\therefore$ 至少有两个组，有3个公共人

三. 灵活利用 Adam's Law

$E[E(Y|X)] = E(Y)$，可以让 Y 与 X 缠上关系，且可以凑题目欲求的项！

No.
Date

Eg.1 2022 Fall T5 (a).

欲证 $E[(\theta-E[\theta|X])h(X)]=0$ for any function $h(\cdot)$

Proof: $\theta-E[\theta|X]$ 与 $h(X)$ 不确定独立，但在X条件下便独立

$E(E[(\theta-E[\theta|X]h(X)|X]).$

$= h(X)\ E[(\theta-E[\theta|X])|X].$

$= h(X)\cdot[E(\theta|X)-E[E[\theta|X]|X]].=0.$

Eg. 2022 Fall T7 (a). $[P(T_1=T_2)<1]$

$T_2=E(T_1|R)$. 欲证 $E[(T_2-\theta)^2]<E[(T_1-\theta)^2]$.

右式如何凑 $T_1|R$ 呢？Conditioning on R!

$LHS = E[(E[T_1|R]-\theta)^2]$

$= E[(E(T_1-\theta)|R)^2] = E[E[(T_1-\theta)]^2|R]$

$RHS \overset{Adam}{=} E(E[(T_1-\theta)^2]|R).$

Jensen with Condition: $E(g(x)|R)\le g(E(X|R))$ 凹

$E[g(x)|R]\ge g(E[X|R])$，凸 ⇒ 即可证毕

四. MMSE ⇒ 后验"∝"找分布

$f_{Y|X}=\dfrac{f_{X|Y}\cdot f_Y}{\sum\leftarrow 定}\propto f_{X|Y}(x|y)f_Y(y)$，往往它好求！

找分布，然后 $E(f_{Y|X}(y|x))$ 就是分布的mean（公式）.

Eg! 2022 Fall T5 (b) 求 MMSE $E[\theta|X]$.

Solution: $f_{\theta|X}(\theta|x)\propto f_\theta(\theta)f_{X|\theta}(x|\theta)=\theta^x(1-\theta)^{n-x}$

则 $\theta|X=x \sim Beta(x+1,n-x+1)$ 故 $E[\theta|X=x]=\dfrac{x+1}{n+2}$.

△: $Beta(a,b)\propto x^{a-1}(1-x)^{b-1}$，$E(x)=\dfrac{a}{a+b}$.

Eg2. Fall 2022 T6. (a) 令 $X=\sum_{i=1}^{n} x_i$

$f_{p|X}(p|X=k) \propto f_{X|p}(X=k|p)\cdot f_p(p) \stackrel{'}{=} p^k(1-p)^{n-k}$

$\therefore \sim Beta(k+1, n-k+1)$，其分布仅依赖于 $k.n.$

(b) $P(X_{n+1}=1 \mid X=k)$ 实质：$X=k \Rightarrow p \Rightarrow X_{n+1}$，$p\in[0,1]$.

$= \int_0^1 P(X_{n+1}|p)\, f_{p|X}(p|X=k)\,dp$

$= \int_0^1 p\cdot f_{p|X}(p|X=k)\,dp = E(Beta(k+1,n-k+1)) = \frac{k+1}{n+2}$

为什么可以不管不顾，只需Bayes找好算的，然后发现分布上"类似的部分"就断言它是XX分布？

这实际上是归一化的手法；Bayes Net 中使用了相同的结

五. Conjugate Prior

二项分布的PMF可视为Beta分布，这是一种巧合吗？

其实这样一类的分布可称为一个distribution family

若对p从Beta先验分布，而data是p下的二项分布，则从先验推到后验，我们没有离开Beta分布family。We say that Beta is the conjugate prior of the Binomial.

Conjugate (Likelihood) Model: Discrete

Sample Space	Sampling Dist	Conjugate Prior	Posterior
$\mathcal{X}=\{0,1\}$	$Bern(\theta)$	$Beta(\alpha,\beta)$	$Beta(\alpha+n\bar{x}, \beta+n-n\bar{x})$
$\mathcal{X}=Z_+$	$Pois(\lambda)$	$Gamma(\alpha,\beta)$	$Gamma(\alpha+n\bar{x}, \beta+n)$
$\mathcal{X}=Z_{++}$	$Geom(\theta)$	$Gamma(\alpha,\beta)$	$Gamma(\alpha+n, \beta+n\bar{x})$

Continuous:

	$Expo(\theta)$	$Gamma(\alpha,\beta)$	$Gamma(\alpha+n, \beta+n\bar{x})$

KOKUYO

No.

Date

六、MAP. MMSE & LLSE

A coin, heads up: θ. 令先验: PDF $f_\theta \sim Unif(0,1)$

n tosses, k heads up. (Denote X: num of heads up)

① MAP: $f_{\theta|X=k}(\theta) = \dfrac{f_{X=k|\theta=\theta}(k) \cdot \boxed{f_\theta(\theta)} \to 1}{\Sigma \to \text{constant}}$

$\propto f_{X=k|\theta=\theta}(k) = \binom{n}{k}\theta^k(1-\theta)^{n-k} \propto \theta^k(1-\theta)^{n-k} = g(\theta)$

欲 $g(\theta)$ max $\Rightarrow$ $\lg(\theta)$ max $\Rightarrow$ $k\ln\theta + n-k\ln(1-\theta) = \varphi(\theta)$ max

$\varphi'(\theta) = \dfrac{k}{\theta} - \dfrac{n-k}{1-\theta} = 0$, $\theta = \dfrac{k}{n}$

i.e., $\hat{\theta} = \underset{\theta}{\mathrm{argmax}}\, f(X|\theta) = \dfrac{X}{n}$ (X: r.v., $\theta = \frac{k}{n}$ when $X=k$)

② MMSE: $f_{\theta|k} \propto \theta^k(1-\theta)^{n-k} \Rightarrow f_{\theta|k} \sim Beta(k+1, n-k+1)$

$\therefore E[f_{\theta|k}] = \dfrac{k+1}{n+2}$

③ LLSE:

$L[\theta|X] = E[\theta] + \dfrac{Cov(\theta, X)}{Var(X)}[X - E(X)]$

$E[\theta] = \frac{1}{2}$, $Cov(\theta, X) = E[\theta X] - E[\theta]E[X]$

$E[\theta X] = E[E[\theta X|\theta]] = E[\theta \cdot E[X|\theta]] = nE[\theta^2] = n\cdot\int_0^1 1\cdot\theta^2 d\theta = \dfrac{n}{3}$

$E[X] = E[E(X|\theta)] = E[n\theta] = \frac{1}{2}n$

$\therefore Cov(\theta, X) = \dfrac{n}{3} - \dfrac{1}{2}\cdot\dfrac{n}{2} = \dfrac{n}{12}$

又 $Var(X) = E[Var(X|\theta)] + Var[E[X|\theta]]$ (Eve's Law)

$= E(n\theta(1-\theta)) + Var(n\theta) = n\cdot\left(\frac{1}{2} - \frac{1}{3}\right) + n^2\cdot\frac{1}{12}$

$= \frac{n}{12}(n+2)$

$\therefore L[\theta|X] = \dfrac{1}{2} + \dfrac{\frac{n}{12}}{\frac{n}{12}(n+2)}\left(X - \dfrac{n}{2}\right) = \dfrac{X+1}{n+2}$

No.

Date

七. 矩母用途

① 切比诺夫: $P(X\geq a)\leq \dfrac{E(e^{tX})}{e^{ta}}$

$E(e^{tX})$ 就是矩母; t灵活取值可以用来凑欲证不等式 (2023 T9, 2022 T8)

② 应对 $E(\sin X)$ $E(\cos X)$

Eg: $X, Y \overset{iid}{\sim} N(0,1)$, $Z_1=\sin(X+Y)$ $Z_2=\cos(X+Y)$ (2021 T9)

求 $E(Z_1)$ $E(Z_2)$ $Var(Z_1)$ $Var(Z_2)$

Solution: (X,Y) 是 BVN, 则 $X+Y\sim(0,2)=Z$.

△: 对于 $aW+b$, $W\sim(\mu,\sigma^2)$, 欲求 $E[\cos(aW+b)]$ $E[\sin(aW+b)]$

$\cos\theta=\dfrac{e^{i\theta}+e^{-i\theta}}{2}$, $E[\cos(aW+b)]=\frac{1}{2}(E[e^{i(aW+b)}]+E[e^{-i(aW+b)}])$

$=\frac{1}{2}\cdot[e^{ib}E[e^{iaW}]+e^{-ib}E[e^{-iaW}]]$ 矩母代t = ...

$E(e^{tW})=e^{\mu t+\frac{1}{2}\sigma^2t^2}$, 则 $\begin{cases}E[e^{iaW}]=e^{i\mu a-\frac{1}{2}\sigma^2a^2}\\E[e^{-iaW}]=e^{-i\mu a-\frac{1}{2}\sigma^2a^2}\end{cases}$

$\therefore E=\frac{1}{2}\cdot e^{-\frac{1}{2}\sigma^2a^2}\cdot(e^{i(\mu a+b)}+e^{-i(\mu a+b)})$

$=e^{-\frac{1}{2}\sigma^2a^2}\cdot\cos(a\mu+b)$

同理: $E[\sin(aW+b)]=e^{-\frac{1}{2}\sigma^2a^2}\sin(aW+b)$

$\therefore E(Z_1)=0$ $E(Z_2)=e^{-1}$.

$E(Z_1^2)=\frac{1}{2}-\frac{1}{2}E[\cos(2W)]=\frac{1}{2}-\frac{1}{2}e^{-4}$

$E(Z_2^2)=\frac{1}{2}+\frac{1}{2}E[\cos(2W)]=\frac{1}{2}+\frac{1}{2}e^{-4}$

$\therefore Var(Z_1)=\frac{1}{2}-\frac{1}{2}e^{-4}$

$Var(Z_2)=\frac{1}{2}+\frac{1}{2}e^{-4}-e^{-2}$

八. min. max 换元

$\min\{X,Y\}=M$, $\max\{X,Y\}=N$, 则 $X+Y=M+N$, $-M+N=|X-Y|$

由其: $P(N\leq t)=P(X\leq t)P(Y\leq t)=F_X(t)F_Y(t)$

则 $P(N=t)_{(pdf)}=(F_X(t)F_Y(t))'$

KOKUYO

No.

Date (b). $X, Y \overset{i.i.d}{\sim} N(0,1)$，求 〃

例: (a). $X, Y \overset{i.i.d}{\sim} Unif(0,1)$，求 $[Corr(\max(X,Y), \min(X,Y))]$

Solution: $M = \max(X,Y)$, $N = \min(X,Y)$

$$Corr(M,N) = \frac{E(MN) - E(M)E(N)}{\sqrt{Var(M)Var(N)}}$$

$E(MN) = E(XY) = E(X)E(Y) = 1/4$

$E(M) = \int_0^1 t \cdot p(M=t)\,dt$， $p(M \le t) = t^2$ $\therefore p(M=t) = 2t$

$= \int_0^1 2t^2\,dt = \frac{2}{3}$ 同理 $E(N) = \frac{1}{3}$*

$E(M^2) = \int_0^1 t^2 \cdot 2t\,dt = \frac{1}{2}$ $\therefore Var(M) = E(M^2) - E(M)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$

*: $p(N \ge t) = (1-t)^2$, $p(N=t) = 2(1-t)$

$E(N) = \int_0^1 2t(1-t)\,dt = (t^2 - \frac{2}{3}t^3)\Big|_0^1 = \frac{1}{3}$

$E(N^2) = \int_0^1 2(1-t)t^2 = (\frac{2}{3}t^3 - \frac{1}{2}t^4)\Big|_0^1 = \frac{1}{6}$

$Var(N) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}$ $\therefore Corr(MN) = 18 \cdot (\frac{1}{4} - \frac{2}{9}) = 18 \cdot \frac{1}{36} = \frac{1}{2}$

(b). $M = \max(X,Y)$ $N = \min(X,Y)$, $E(MN) = E(XY) = 0$

但此处$E(M)$在用CDF ⇒ PDF不再好用！因为正态分布CDF未知！

$E(M) = E(X|X>Y) + E(Y|Y>X) = 2E(X|X-Y>0)$ (symmetric)

欣赏：令 $W = X+Y$ $Z = X-Y$，则 $X = \frac{W+Z}{2}$，$E(X|X-Y) = \frac{1}{2}E(W+Z|Z>0)$

$= \frac{1}{2}E(W|Z>0) + \frac{1}{2}E(Z|Z>0)$，那么$W$与$Z$独立吗？

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

$= Var(X) - Var(Y) = 0$，一般 $Cov = 0 \ne i.i.d$，但是有定理：

$BVN(X,Y)$, if $Corr(X,Y) = Cov(X,Y) = 0 \Rightarrow X, Y$ 独立

$\therefore W \sim (0,2)$ $Z \sim (0,2)$，$E(W|Z>0) = 0$

$E(Z|Z>0) = \int_0^\infty z f(z)\,dz = \int_0^\infty z \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{z^2}{2\sigma^2}}\,dz.$

令 $t = \frac{z}{\sigma}$，$z = \sigma t$，$dz = \sigma dt$，则：

$I = \int_0^\infty \sigma t \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2}\sigma\,dt = \frac{\sigma}{\sqrt{2\pi}}\int_0^\infty t e^{-\frac{t^2}{2}}\,dt = \sigma\sqrt{\frac{2}{\pi}}$

$\sigma = \sqrt{2}$，则 $\frac{1}{2}E(Z|Z>0) = \sqrt{1/\pi} = E(M)$，同理，$E(N) = -\frac{1}{\pi}$，

而：$M^2 + N^2 = X^2 + Y^2$，$E(M^2) = E(N^2) = 1$ $\therefore Var(M) = Var(N) = 1 - (\frac{1}{\sqrt{\pi}})^2 = 1 - \frac{1}{\pi}$

Campus

$\therefore Corr(M,N) = \frac{\frac{1}{\pi}}{1 - \frac{1}{\pi}} = \frac{1}{\pi - 1}$