期末总复习 (至 inequality；不等式. Markov. Inference见前)

Part 1: 连续型随机变量

· CDF: $f(x)$　　PDF: $F(x)=\int_{-\infty}^{x} f(t)dt$，$f(x)$取值不是概率！

△: 连续型 r.v. 取端点概率恒为0

· Def: The support of X, and of its distribution: $\{x \mid f(x)>0\}$

· 验证 PDF f 是否 valid 的流程：

① $f(x)\geq 0$　② $\int_{-\infty}^{\infty} f(x)dx=1$

support

· 介绍两个分布：Logistic：$F(x)=\dfrac{e^x}{1+e^x}$，$x\in\mathbb{R}$；$f(x)=\dfrac{e^x}{(1+e^x)^2}$，$x\in\mathbb{R}$

Rayleigh: $F(x)=1-e^{-x^2/2}$, $x>0$；$f(x)=xe^{-x^2/2}$, $x>0$

· 验证 CDF F 是否 valid 的流程：

① Increasing　② Right-Continuous　③ Convergence to 0 & 1 in the limits

· $E(X)=\int_{-\infty}^{+\infty} xf(x)dx$　这里 $f(x)$ 的使用很像概率，但实则不是！

$E(g(x))=\int_{-\infty}^{+\infty} g(x)f(x)dx$；令 $G(x)=1-F(x)$, $E(X)=\int_{0}^{+\infty} G(x)dx$

$=P(X>x)$　　Via Survival function

最后介绍了 Symmetric Property：

Let $X_1,\cdots,X_n$ be i.i.d. from a [Continuous Distribution] (Any!!):

$P(X_{a_1}<\cdots<X_{a_n})=1/n!$ for any permutation $a_1,\cdots,a_n$

之后是三大著名 Con. r.v.：

① 均匀：$f(x)=\begin{cases}1/(b-a), & x\in(a,b)\\ 0, & \text{otherwise}\end{cases}$　　$U\sim \text{Unif}(a,b)$

② 指数：$f(x)=\lambda e^{-\lambda x}$, $x>0$；$F(x)=1-e^{-\lambda x}$, $x>0$；$X\sim \text{Expo}(\lambda)$

它最重要的性质为：Memoryless：

$$P(X\geq s+t \mid X\geq s)=P(X\geq t)$$

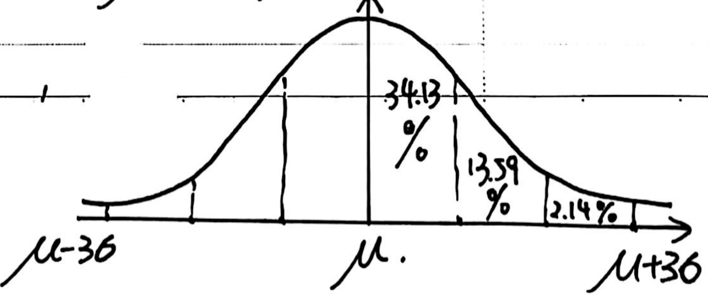
Def: Failure Rate Function：$r(t)=\dfrac{f(t)}{1-F(t)}$，而代入Expo，$r(t)=\lambda$

△: 而且一个 Con. r.v 若 memoryless，它必是指数分布

③ 正态：标准 $\varphi(z)=\dfrac{1}{\sqrt{2\pi}}e^{-z^2/2}$，$-\infty<z<\infty$，$Z\sim \text{Norm}(0,1)$　均　方差

而若：$X=\mu+\sigma Z$，则 $X\sim N(\mu,\sigma^2)$，均μ 方差σ^2

$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，



· 均匀分布的普遍性：（F为一个CDF函数）

1. 设 $U\sim Unif(0,1)$，$X=F^{-1}(U)$，则X是一个CDF为F的随机变量

2. X是CDF为F的r.v.，则 $F(X)\sim Unif(0,1)$

Eg: $F(x)=\frac{e^x}{1+e^x}$，$F^{-1}(U)=\ln\left(\frac{U}{1-U}\right)$，$U\sim Unif(0,1)$.

则 $\ln\left(\frac{U}{1-U}\right)\sim$ Logistic

· 中心极限定理：$X_1,\cdots,X_n$ be i.i.d. r.v，均μ方σ^2，定义：

$\bar{X}_n=\frac{1}{n}\sum_{j=1}^{n}X_j$，则当 $n\to\infty$ 时：

$$\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right)\to N(0,1)$$

最后，介绍了矩生成函数：X：$M(t)=E(e^{tX})$ （与期中前 t^X 不同）

MGF encodes the moment of an r.v.，且有以下完美性质：

① $E(X^n)=M^{(n)}(0)$ ② 若X,Y独立，应：$M_{X+Y}(t)=M_X(t)M_Y(t)$

☆ 正态分布 MGF：$X\sim N(\mu,\sigma^2)$，$M_X(t)=e^{\mu t+\frac{1}{2}\sigma^2t^2}$

Part 2：联合分布

$\boxed{\text{Discrete}}$

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

Joint PMF：$P_{X,Y}(x,y)=P(X=x,Y=y)$

Marginal PMF：（用LOTP，只关注一个r.v.）$P(X=x)=\sum_y P(X=x,Y=y)$

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

从上述公式视角，能额外给出独立判断条件：

若X,Y独立：（离散）：$F_{X,Y}(x,y)=F_X(x)F_Y(y)$

$P(X=x,Y=y)=P(X=x)P(Y=y)$ $P(Y=y|X=x)=P(Y=y)$

例：鸡-蛋问题：

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

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

△：Poisson：PMF：$\frac{e^{-\lambda}\lambda^k}{k!}$，mean & variance 均为$\lambda$

TITLE Continuous

SUBTITLE

PAGE/NO

DATE

先补充：Conditional PDF Given an Event:

X: 连续 r.v.，Given Event A: $P(A)>0$:

· $P(X|A)=\int f_{X|A}(x)dx$

· $A_1,\dots,A_n$ 为 disjoint set，$P(A_i)>0$，则 $f_X(x)=\sum_{i=1}^{n}P(A_i)f_{X|A_i}(x)$

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

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

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

它们也有LOTP，Bayes' Rule:

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

而利用上述式子，也能帮助判断独立性：

若 X,Y（连续）独立：

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

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

$f_{Y|X}(y|x)=f_Y(y)$

除此之外，还有一条黄金性质！

$f_{X,Y}(x,y)=g(x)h(y)$

$\Rightarrow X,Y$ 独立

i.e.，解耦 $\Rightarrow$ 独立

LOTUS: $\quad g: \mathbb{R}^2\to\mathbb{R}$

$E(g(X,Y))=\sum_x\sum_y g(x,y)P(X=x,Y=y)$ （离散）

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

之后引入了协方差与相关系数概念：

协方差：$Cov(X,Y)=E((X-EX)(Y-EY))=E(XY)-E(X)E(Y)$

若 X,Y 独立，则 $Cov(X,Y)$ 显然为0

Cov有以下性质：

① $Cov(X,X)=Var(X)$ ② $Cov(X,Y)=Cov(Y,X)$

③ $Cov(aX,Y)=a\cdot Cov(X,Y)$ ④ $Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)$

⑤ $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$

	PAGE/NO	TITLE
DATE		SUBTITLE

相关系数：$Corr(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \in [-1,1]$

有以下的关系：

$Cov(X,Y)$ or $Corr(X,Y)=0 \Longrightarrow$ Uncorrelated $\underset{\checkmark}{\overset{\times}{\rightleftarrows}}$ 独立.

· Multinomial：$\sum p_i = 1$，$\sum n_i = n$，$P(X_i=n_i)=p_i$

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

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

$p_i' = \dfrac{p_i}{p_2+\cdots+p_k}$，$i\in[2,k]$

若 $(X_1,\cdots,X_k) \sim Mult_k(n,\vec{p})$，$\vec{p}=(p_1,\cdots,p_k)$. 则：

$\triangle$：$Cov(X_i,X_j) = -np_ip_j \Leftarrow$ 是用 $Cov(X_i,X_j) = \dfrac{Var(X_i+X_j)-Var(X_i)-Var(X_j)}{2}$

· 多元的形式不止 Multinomial 一种！还有：[Multivariate Normal]

$\vec{X}=(X_1,\cdots,X_k)$ 若 $\sum t_iX_i$ 均为正态分布，则 $\vec{X}$ 为MVN 分布

且对于MVN来说：$\vec{X}\sim MVN$，$\vec{X}=(\vec{X_1},\vec{X_2})$，则若 $\vec{X_1}$ 每个部分与 $\vec{X_2}$ 每个部分都不相关，则 $\vec{X_1}$ 与 $\vec{X_2}$ 独立

特别地：BVN (X,Y)，若 $Corr(X,Y)=0 \Rightarrow X,Y$ 独立

☆ 最后一个核心：Change of Variable

$Y=g(X)$，则 $f_Y(y) = f_X(x)\cdot\left|\dfrac{dx}{dy}\right|$，$x=g^{-1}(y)$

$\Rightarrow \vec{X}=(X_1,\cdots,X_n)$ $\vec{Y}=g(\vec{X})$，$\dfrac{\partial\vec{x}}{\partial\vec{y}} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}$

$f_Y(y) = f_X(x)\left|\dfrac{dx}{dy}\right|$

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

$X=\sqrt{2T}\cos U$ $Y=\sqrt{2T}\sin U$

则 $f_{U,T}(u,t) = \dfrac{e^{-t}}{2\pi}$，$X^2+Y^2=2T$，$\dfrac{\partial(x,y)}{\partial(u,t)} = \begin{pmatrix} -\sqrt{2t}\sin u & \frac{1}{\sqrt{2t}}\cos u \\ \sqrt{2t}\cos u & \frac{1}{\sqrt{2t}}\sin u \end{pmatrix}$

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

*：这是十分惯用的手法！$\frac{\partial(u,t)}{\partial(x,y)}$ 难求，

就求 $\left|\frac{\partial(x,y)}{\partial(u,t)}\right|^{-1}$!

最后介绍了二元 r.v. 间的常见关系：和一定：

离散：$P(T=t)=\sum_x P(X=x)P(Y=T-t)$.

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

Part3：条件期望

'条件'可以是事件，也可以是随机变量

Event: $E(Y|A)=\sum_y y\,P(Y=y|A)$ (Dis)

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

全期望公式：设事件 $A_1,\cdots,A_n$ 是样本空间的一个分割，$\forall i$，$P(A_i)>0$

则：$E(Y)=\sum_{i=1}^{n} E(Y|A_i)P(A_i)$

LOTUS Rule: $\begin{cases} E(g(Y)|A)=\sum_y g(y)\,P_{Y|A}(y) \\ E(g(Y)|A)=\int_{-\infty}^{\infty} g(y)\cdot f_{Y|A}(y)dy \end{cases}$

Random Variable: Let $g(x)=E(Y|X=x)$，则 $E(Y|X)$ 定义为了 $g(X)$

重要的是：$E(Y|X)$ 是关于X的函数，且是一个r.v. 不是数！

因此能计算 $E(E(Y|X))$ & $Var(E(Y|X))$

例：$X,Y \overset{i.i.d}{\sim} Expo(\lambda)$，求 $E(max(X,Y)|min(X,Y))$.

Solution：令 $M=max(X,Y)$ $L=min(X,Y)$

$E(M|L=l)=E(L|L=l)+E(M-L|L=l)=l+E(M-L|L=l)$

指数分布无记忆性，故 $M-L$ 与 L 独立，且 $M-L\sim Expo(\lambda)$

$\therefore l+E(M-L|L=l)=l+E(M-L)=l+\frac{1}{\lambda}$

性质：若 X, Y 相互独立，则 $E(Y|X)=E(Y)$

- $E(h(X)Y|X)=h(X)E(Y|X)$
- 亚当定律：$E(E(Y|X))=E(Y)$

Adam's Law: Proof: $E(g(X)) = \sum_x g(x) \cdot P(X=x)$

$= \sum_x \sum_y y P(X=x) \; P(Y=y|X=x) = \sum_y y \sum_x P(X=x, Y=y)$

$= \sum_y y P(Y=y) = E(y)$ 且是对 $\forall$ r.v.s X & Y

其实 Adam's Law 很类似于全期望公式

且 Adam's Law 可进一步扩展:

$$E(E(Y|X,Z)|Z) = E(Y|Z)$$

而条件期望方差:

$$Var(Y|X) = E(Y^2|X) - [E(Y|X)]^2$$ 'EVVE'

Eve's Law: $Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$

Prediction: LLSE of Y given X:

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

事实上，它的形式为 $a+bx$，$L[Y|X]$ 可使得 $E[(Y-a-bX)^2]$ 取min

MMSE of Y given X: $g(X) = E[Y|X]$

投影解释: $\forall$ function h，$Y - E(Y|X)$ 与 $h(X)$ uncorrelated

i.e., $E((Y - E(Y|X))h(X)) = 0$

$E(Y|X)$ 是基于 X 对 Y 的预测，且是最佳预测

Theorem: 若 X, Y 为 jointly normal r.v. 则

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