Data Science H.W1 b08502141 EE4 石旻翰

1.

(a).

①. Let X be a set. Let Σ be a σ-algebra over X.

②. Non-negativity: $\mu(E) \geq 0,\ \forall E \in \Sigma$

③. Null empty set: $\mu(\emptyset) = 0$.

④. Countable additivity: $\forall$ countable collections $\{E_k\}_{k=1}^{\infty}$ of pairwise disjoint sets in Σ: $\mu(\bigsqcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$.

(b).

$\forall x, y, z \in X$. $d(a,b)$ means the distance between a, b.

①. $d(x,y) \geq 0$, "=" holds if and only if $x=y$.

②. $d(x,y) = d(y,x)$.

③. $d(x,y) \leq d(x,z) + d(y,z)$

④. $|d(x,y) - d(x,z)| \leq d(y,z)$.

2. $F(x)$ is CDF. $F(x) - \int_0^x f(u)du = -e^{-u}\Big|_0^x = 1-e^{-x},\ x \geq 0$

(1). $F_Z(z) = P(Z \leq z) = P(\max(X,Y) \leq z) = P(X \leq z)P(Y \leq z)$

$= F_X(z) \cdot F_Y(z)$

$$\therefore f_Z(z) = \frac{dF_Z(z)}{dz} = f_X(z)F_Y(z) + F_X(z)f_Y(z) = \begin{cases} 2e^{-z}(1-e^{-z}), & x \geq 0 \\ 0, & \text{o.w.} \end{cases} \#$$

(2). $F_W(w) = P(W \leq w) = P(\min(X,Y) \leq w) = P(1 - P(\min(X,Y) > w))$

$$\therefore f_W(w) = -\frac{d}{dw}\left(P(X>w)P(Y>w)\right) = f_X(w)(1-F_Y(w)) + (1-F_X(w))f_y(w)$$

$$= \begin{cases} 2e^{-2w}, & x \geq 0 \\ 0, & \text{o.w} \end{cases} \#$$

3.

(1). If X is Gamma Distribution, then CDF is:

$$F(x) = \int_0^x \frac{u^{\alpha-1} e^{-\frac{u}{\theta}}}{\theta^k \Gamma(\alpha)} du \text{, now } \alpha=1 \text{.. } \theta=1.$$

$$\therefore F(x) = \int_0^x e^{-u} du = -e^{-u}\Big|_0^x = 1-e^{-x}$$

$$\because P(2X \leq x) = P(X \leq \tfrac{x}{2}) = F(\tfrac{x}{2}) = 1-e^{-\frac{x}{2}}$$

$\therefore$ $2X$'s distribution is $\Gamma(\alpha=1, \theta=2)$ ※

(2).

①.
$$f(x) = \begin{cases} \frac{1}{\Gamma(1) 2^1} x^0 \cdot e^{-\frac{x}{2}} = \frac{1}{2} e^{-\frac{x}{2}}, & \text{if } x>0 \\ 0, & \text{o.w.} \end{cases}$$

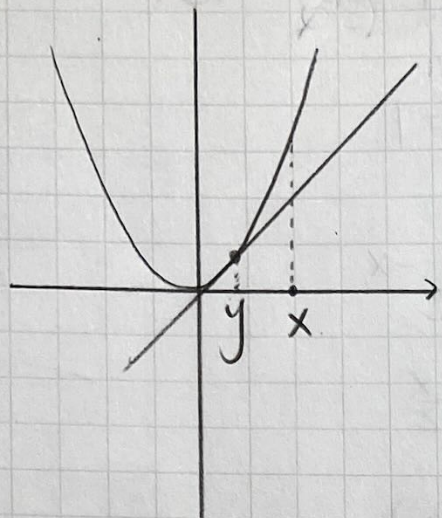
② PDF of X is $f(x) = \frac{1}{2} e^{-\frac{x}{2}}$ for $x>0$.

$\Rightarrow$ X follows a chi² distribution with 2 degrees of freedom, denoted $\chi^2(2)$.

$$\therefore f(x) = \begin{cases} \frac{x^0 e^{-\frac{x}{2}}}{2^1 \Gamma(1)} = \frac{1}{2} e^{-\frac{x}{2}}, & x>0. \\ 0, & \text{o.w.} \end{cases}$$

4.

(1).



Let $F: \Omega \to R$, which is a differentiable, strictly convex function on a closed convex set Ω. ($\Omega \subseteq R^d$), and x, y are two points in Ω

Take $F(x) = \|x\|^2$. $D_F(x,y) = F(y) - F(x) - \langle \nabla F(x), y - x \rangle$

$$= \|y\|^2 + \|x\|^2 - 2xy = \|y - x\|^2$$

$\therefore$ Squared Euclidean distance is Bregman divergence. #

(2). Entropy denotes as H.

$$H(x,y) = H(x|y) + H(y) = H(x) + \xrightarrow{\;0\;}\cancel{H(F(x)|x)}$$ $\because$ Knowing X, automatically knowing $f(x)$.

$$= H(F(x)) + F(x|F(x))$$

$\therefore H(x) = H(F(x)) + F(x|F(x)) \Rightarrow H(F(x)) \le H(x)$. #

5.

(1). $f(x)=\frac{1}{\theta}$ $M_1=\frac{1}{n}\sum_{i=1}^{n}X_i=\mu=E(x)=\frac{\theta}{2}\Rightarrow\hat{\theta}=2\bar{X}$

$M_2=\mu^2+\sigma^2=\frac{\theta^2}{3}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

$\Rightarrow\hat{\theta}=\sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$ #

(2). $L(\theta)=\prod_{i=1}^{n}f(x|\theta)$, $\hat{\theta}=\arg\max_{\theta\in\Omega}L(\theta)$. 又 $f(x|\theta)=\frac{1}{\theta}$

$\Rightarrow \ln(L(\theta))=\ln(\prod_{i=1}^{n}\frac{1}{\theta})=-n\ln\theta.$

$\frac{d}{d\theta}\ln(L(\theta))=-\frac{n}{\theta}$, $\forall x>0$.

Thus, max. of $L(\theta)$ is at $\theta=X_n\Rightarrow\hat{\theta}=X_n$ #

(3). MAP estimator is: $\arg\max_{\theta}f(x|\theta)f(\theta)=\arg\max_{\theta}\frac{1}{\theta}$

$\Rightarrow\hat{\theta}_{MAP}=\arg\max_{\theta}f(x|\theta)=\hat{\theta}_{MLE}=X_n$ #

(4). $L^2(\hat{\theta},\theta)=(\hat{\theta}-\theta)^2$ $\begin{cases}f(x|\theta)=\frac{1}{\theta}, \text{ when } 0<x<\theta\\ h(\theta)=1, \text{ when } 0<\theta<1\end{cases}$

$\Rightarrow u(x,\theta)=h(\theta)f(x|\theta)=\frac{1}{\theta}$, $0<x<\theta<1$.

$g(x)=\int_x^1 u(x,\theta)d\theta=\int_x^1\frac{1}{\theta}d\theta=-\ln x$, $0<x<1$

$k(\theta|x)=\frac{u(x,\theta)}{g(x)}=\frac{1}{-\theta\ln x}$, $0<x<\theta<1$

$\therefore\hat{\theta}=E(\theta|x)=\int_x^1\theta k(\theta|x)d\theta=-\frac{1}{\ln x}\int_x^1 d\theta=\frac{x-1}{\ln x}$ #