第五次作业

冯海桐-522031910557

2024 年 12 月 25 日

1

引理（Jensen不等式）：设$f(x)$为区间I上的凸函数，则对于$\forall x_K \in I, \forall \lambda_k \in (0,1)(k=1,2,\cdots,n)$且, $\sum_{k=1}^{n} \lambda_k = 1$，有

$$f(\sum_{k=1}^{n} \lambda_k x_k) \leq \sum_{k=1}^{n} \lambda_k f(x_k) \ (n \geq 2)$$

证明：对于$n=2$的情况，设$x_1, x_2 \in I$，则有

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

这是凸函数的定义，所以对于$n=2$的情况，不等式成立。假设对于$n-1$的情况不等式成立，考察n的情况。

记$t = \sum_{k=2}^{n} \lambda_k \in (0,1)$，则有$\frac{1}{t}\sum_{k=2}^{n} \lambda_k = 1$以及$\lambda_1 + t = 1$。又因为

$$\min_{2 \leq k \leq n} \{x_k\} \leq \sum_{k=2}^{n} \frac{\lambda_k}{t} x_k \leq \max_{2 \leq k \leq n} \{x_k\}$$

所以有

$$\sum_{k=2}^{n} \frac{\lambda_k}{t} x_k \in I$$

于是有

$$f(\sum_{k=1}^{n} \lambda_k x_k) = f(\lambda_1 x_1 + t \sum_{k=2}^{n} \frac{\lambda_k}{t} x_k) \leq \lambda_1 f(x_1) + t f(\sum_{k=2}^{n} \frac{\lambda_k}{t} x_k)$$

$$\leq \lambda_1 f(x_1) + t \sum_{k=2}^{n} \frac{\lambda_k}{t} f(x_k) = \sum_{k=1}^{n} \lambda_k f(x_k)$$

故对于n的情况不等式成立。综上，对于任意$n \geq 2$的情况，不等式成立。

显然，对于原题，下式显然成立

$$\phi\left(E\left[x\right]\right) \leq E\left[\phi\left(x\right)\right]$$

2

$$D_{KL}(p(x)||q(x)) = \int p(x) \log \frac{p(x)}{q(x)} dx \geq \int p(x) \left(1 - \frac{q(x)}{p(x)}\right) dx$$

$$= \int p(x)dx - \int q(x)dx = 1 - 1 = 0$$

由于$\log x \geq 1 - 1/x$当且仅当$x = 1$时取等号，所以当$p(x) = q(x)$时取等号。

值得注意的是，对于任意一点单独的积分值为0，所以取等时允许有限个点不满足取等条件，即等号成立当且仅当$p(x)$与$q(x)$几乎处处相等。

3

$$I(X,Y) = D_{KL}(p(x,y)||p(x)p(y)) = \int_Y \int_X p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dxdy$$

$$= \int_Y \int_X p(x,y) \log p(x,y) dxdy - \int_X \int_Y p(x,y) \log p(x) dydx - \int_Y \int_X p(x,y) \log p(y) dxdy$$

$$= \int_Y \int_X p(x,y) \log p(x,y) dxdy - \int_X p(x) \log p(x) dx - \int_Y p(y) \log p(y) dy$$

$$= H(X,Y) - H(X) - H(Y)$$