1. Basic Probability & Statistics

(1). $F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & x \leq 0 \\ x/4 & x \in [0,1) \\ 1/4 & x \in [1,3) \\ (3x-7)/8 & x \in [3,5] \\ 1 & x \in [5,+\infty) \end{cases}$

(2) $F_Y(y) = P\{Y \leq y\} = P\{\frac{1}{X} \leq y\} = P\{X \geq \frac{1}{y}\}. = F_X(\frac{1}{y}).$

$$f_Y(y) = f_X(\frac{1}{y})\cdot(\frac{1}{y})'$$

$$= \begin{cases} \frac{1}{4y^2} & y \in [1,+\infty) \\ \frac{3}{8y^2} & y \in (\frac{1}{5},\frac{1}{3}) \\ 0 & \text{elsewhere.} \end{cases}$$

(3). $P[Z \geq z] = \int_{z}^{\infty} f(t)\,dt = 1 - \int_{0}^{z} f(t)\,dt = 1 - F(z).$

$Z = \int_0^{Z} dt = \int_0^{\infty} I(t)\,dt. \Rightarrow I(t) := \begin{cases} 1 & t \leq Z \\ 0 & \text{elsewhere.} \end{cases}$

$E[Z] = E[\int_0^{\infty} I(t)\,dt] = \int_0^{\infty} E(I(t))\cdot dt = \int_0^{\infty} P(Z > t)\,dt = \int_0^{\infty} P(Z > z)\,dz$

验证:

$[X] = \int_0^{\infty} x f(x)\,dx = \int_0^1 \frac{1}{4} x\,dx + \int_3^5 \frac{3}{8} x\,dx$

$= \frac{1}{8}x^2\Big|_0^1 + \frac{3}{16}x^2\Big|_3^5 = \frac{25}{8}$

$[Y] = \int_0^{\infty} y\cdot f(y)\,dy = \int_1^{\infty} \frac{1}{4y}\,dy + \int_{\frac{1}{5}}^{\frac{1}{3}} \frac{3}{8y}\,dy$

$= +\infty$ (not exist).

2. Strong Convexity.

2.1 → 2.2 :

$$(2.1) \Rightarrow f((1-\alpha)\cdot x + \alpha y) \le (1-\alpha) f(x) + \alpha\cdot f(y) - \frac{\lambda}{2}\alpha(1-\alpha)\|x-y\|^2$$

$$\Rightarrow \frac{f(x+\alpha(y-x)) - f(x)}{\alpha} \le f(y) - f(x) - \frac{\lambda}{2}(1-\alpha)\|x-y\|^2$$

when $\alpha \to 0$

$$\Rightarrow \nabla f(x)^T (y-x) \le f(y) - f(x) - \frac{\lambda}{2}\|x-y\|^2$$

$$\Rightarrow f(y) \ge f(x) + \nabla f(x)^T (y-x) + \frac{\lambda}{2}\|y-x\|^2$$

2.2 → 2.1 :

$$2.2 \rightarrow f(y) \ge f(x) + \nabla f(x)^T (y-x) + \frac{\lambda}{2}\|y-x\|^2$$

set $y = \alpha(x_1) + (1-\alpha)(x_2)$

$$\Rightarrow \begin{cases} f(x_1) \ge f(y) + \nabla f(y)^T (x_1 - y) + \frac{\lambda}{2}\|x_1 - y\|^2 & (2.2.a) \\ f(x_2) \ge f(y) + \nabla f(y)^T (x_2 - y) + \frac{\lambda}{2}\|x_2 - y\|^2 & (2.2.b) \end{cases}$$

$(2.2.a) \times \alpha + (2.2.b)\cdot(1-\alpha)$

$$\Rightarrow \alpha\cdot f(x_1) + (1-\alpha) f(x_2) \ge f(\alpha\cdot x_1 + (1-\alpha) x_2) + \frac{\lambda}{2}\alpha\cdot\|(1-\alpha)(x_1 - x_2)\|^2 + \frac{\lambda}{2}(1-\alpha)\|\alpha(x_2 - x_1)\|^2$$

set $\cancel{\lambda' = 2\lambda}$ $x_2 = x$ $x_1 = y$

$$\Rightarrow f((1-\alpha)x + \alpha y) \le (1-\alpha) f(x) + \alpha f(y) - \frac{\lambda}{2}\alpha(1-\alpha)\|x-y\|^2$$

3. Doubly Stochastic Matrix

use second-order condition for convexity. $f(x) = -x \log x$

(1)

$$(x\log x)' = -(x \cdot \frac{1}{x}\frac{1}{\ln 10} + \log x) = (\frac{1}{\ln 10} + \log_{10} x)$$

$$(x\log x)'' = \frac{1}{x \ln 10} \geq 0 \quad (\text{for all } x \in [0,1])$$

$\Rightarrow$ $x\log x$ is a convex function, then use Jenson's Inequality.

$$f(\sum_{i=1}^{d} \theta_i x_i) \leq \sum_{i=1}^{d} \theta_i f(x_i)$$

here. $\theta_i = 1$ $f(\,) = x \log x$

$$\Rightarrow \quad -\sum_{i=1}^{d}\sum_{j=1}^{d} x_{ij}\log(x_{ij}) \leq -\log(\sum_{i=1}^{d}\sum_{j=1}^{d} x_{ij}) \cdot (\sum_{i=1}^{d}\sum_{j=1}^{d} x_{ij})$$

$$= + d \cdot \log(d).$$

(2) Using Perron-Frobenius eigenvalue inequality:

$$\min_i \sum_j a_{ij} \leq \rho(X) \leq \max_i \sum_j a_{ij}$$

while $\min_i \sum_j a_{ij} = \max_i \sum_j a_{ij} = 1$

$\Rightarrow$ $\rho(X) = 1$

Also

$$X \cdot (\underbrace{1, 1, \cdots 1}_{d个})^T = (\sum_j a_{1j}, \cdots \sum_j a_{dj}). = 1 \cdot (\underbrace{1, 1, \cdots 1}_{d个})^T$$

$\Rightarrow$ 1 is X's eigenvalue

(actually $\rho(X)$ is defined as the largest eigenvalue...)

4. Hypothesis Testing.

$$\tau_{\chi^2} = \frac{k-1}{k} \frac{12N}{k^2-1} \sum_{i=1}^{k} (r_i - \frac{k+1}{2})^2$$

N ~ dataset

k ~ algorithm.

$$= \frac{12N}{k(k+1)} \left(\sum_{i=1}^{k} r_i^2 - \frac{k(k+1)^2}{4} \right)$$

$$= \frac{12 \times 5}{5 \times 6} \cdot \left(3.8^2 + 4^2 + 3.2^2 + 1.2^2 + 2.8^2 - \frac{5 \times 6^2}{4} \right)$$

$$= 9.92$$

$$\tau_F = \frac{(N-1)\tau_{\chi^2}}{N(k-1) - \tau_{\chi^2}} = 3.9365 > \tau_F(5,5, \alpha=0.05) = 3.007$$

=> 这些算法性能不同.

Algorithm	A	B	C	D	E
Average performance	3.8	4	3.2	1.2	2.8

for Nemenyi Test:

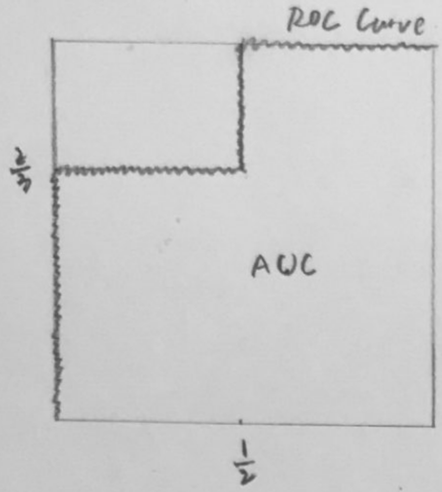
$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}}$$

$$= 2.728 \times \sqrt{\frac{5 \times 6}{6 \times 5}} = 2.728$$

$1.2 + 2.728 = 3.928$ => Algorithm B & D are different on performance.

5. ROC and AUC.

(1).



$m^+ = 3 \quad m^- = 2$

样本	x_1	x_4	x_5	x_2	x_3
输出	0.9	0.7	0.4	0.3	0.1
标记	+	+	-	+	-

$$AUC = \frac{1}{2}\left(\frac{2}{3}+1\right) = \frac{5}{6} \approx 0.83$$

$$\ell_{rank} = \frac{1}{6}(1) = \frac{1}{6}.$$

(2) 文字叙述即可.