Jane Downer
A20452471

Assignment 2

PROBLEM 1

(1)

$$P(X=1) = P(X=1 \mid Y=1)\cdot P(Y=1) + P(X=1 \mid Y=0)\cdot P(Y=0)$$

$$= 0.2 \times 0.1 = 0.02$$

$P(X=1 \mid Y=1)\cdot P(Y=1)$:

$$= P(X=1 \mid Y=1, Z=1)\cdot P(Y=1, Z=1) + P(X=1 \mid Y=1, Z=0)\cdot P(Y=1, Z=0)$$

$$= P(X=1 \mid Y=1, Z=1)\cdot P(Y=1)\cdot P(Z=1) + P(X=1 \mid Y=1, Z=0)\cdot P(Y=1)\cdot P(Z=0)$$

$$= 0.6 \times 0.9 \times 0.8 + 0.1 \times 0.9 \times 0.2 = 0.45$$

$$P(X=1) = 0.45 + 0.02 = \boxed{0.47}$$

(2) $$\mathbb{E}(Y) = 1 \times P(Y=1) + 0 \times P(Y=0)$$

$$= 1 \times 0.9 + 0 \times 0.1$$

$$= \boxed{0.9}$$

(3) $P(Y=115) = 0.9$

$P(Y=20) = 1 - 0.9 = 0.1$ $\longrightarrow$ $\mathbb{E}(Y) = 115 \times P(Y=115) + 20 \times P(Y=20)$

$$= 115 \times 0.9 + 20 \times 0.1$$

$$= \boxed{105.5}$$

PROBLEM 2

	$P(x)$	$P(\text{defect} \mid X)$
A	0.2	0.02
B	0.3	0.01
C	0.5	0.0005

$X \in (A, B, C)$

total = 10,000

(1) $P(\text{defective}) = 0.2 \times 0.02 + 0.3 \times 0.01 + 0.5 \times 0.0005 = \boxed{0.00725}$

(2) $P(A \mid \text{defective}) = \dfrac{P(\text{defective} \mid A) \cdot P(A)}{P(\text{defective})} = \dfrac{0.02 \times 0.2}{0.00725} \approx \boxed{0.5517}$

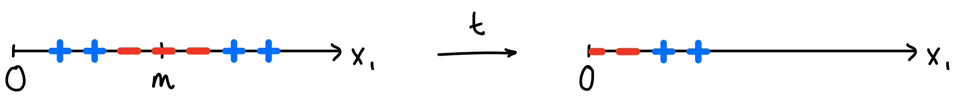
(3) $P(B \mid \text{defective}) = \dfrac{P(\text{defective} \mid B) \cdot P(B)}{P(\text{defective})} = \dfrac{0.01 \times 0.3}{0.00725} \approx \boxed{0.41379}$

(4) $P(C \mid \text{defective}) = \dfrac{P(\text{defective} \mid C) \cdot P(C)}{P(\text{defective})} = \dfrac{0.0005 \times 0.3}{0.00725} \approx \boxed{0.0207}$

PROBLEM 3

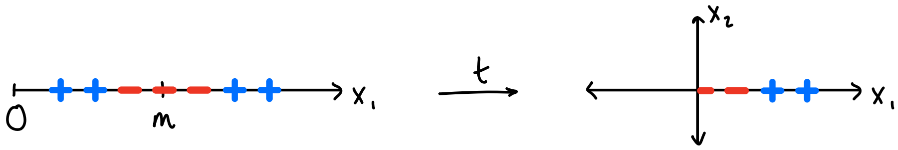
Designing Transformations

(1) m = median



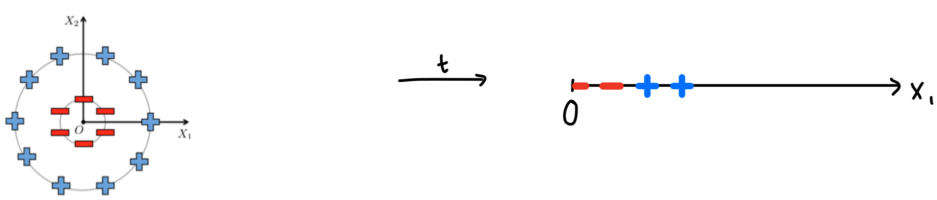
$$\boxed{t(x) = |m - x|}$$

(2)



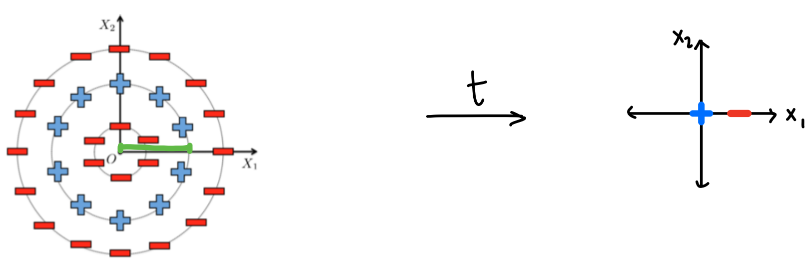
$$\boxed{t(x) = (|m-x|,\ 0)}$$

(3)



$$\boxed{t(x_1, x_2) = \sqrt{x_1^2 + x_2^2}}$$

(4)



d = shortest distance from origin to second ring

$$\boxed{t(x_1, x_2) = \left(\left|d - \sqrt{x_1^2 + x_2^2}\right|,\ 0\right)}$$

Kernel or Not

(1) $k(x,z) = (xz+1)^2 \rightarrow x = \begin{bmatrix} x_i \\ x_j \end{bmatrix} \quad z = \begin{bmatrix} z_i \\ z_j \end{bmatrix}$

$xz = x_i z_i + x_j z_j$

$xz^2 = x_i^2 z_i^2 + 2x_i z_i x_j z_j + x_j^2 z_j^2$

$$\begin{aligned}(xz+1)^2 &= (xz)^2 + 2xz + 1 \\ &= x_i^2 z_i^2 + 2x_i z_i x_j z_j + x_j^2 z_j^2 + 2x_i z_i + 2x_j z_j + 1 \\ &= \begin{bmatrix} x_i^2 \\ \sqrt{2}x_i x_j \\ x_j^2 \\ \sqrt{2}x_i \\ \sqrt{2}x_j \\ 1 \end{bmatrix} \cdot \begin{bmatrix} z_i^2 \\ \sqrt{2}z_i z_j \\ z_j^2 \\ \sqrt{2}z_i \\ \sqrt{2}z_j \\ 1 \end{bmatrix} \\ &= \langle \phi(x), \phi(z) \rangle\end{aligned}$$

✓ $\boxed{\text{valid kernel}}$

(2) $k(x,z) = (xz-1)^3$

$$\begin{aligned}(a-1)^3 &= (a-1)^2(a-1) \\ &= (a^2 - 2a + 1)(a-1) \\ &= a^3 - 2a^2 + a - a^2 + 2a - 1 \\ &= a^3 - 3a^2 + 3a - 1\end{aligned}$$

$xz = x_i z_i + x_j z_j$

$xz^2 = x_i^2 z_i^2 + 2x_i z_i x_j z_j + x_j^2 z_j^2$

$xz^3 = x_i^3 z_i^3 + 2x_i^2 z_i^2 x_j z_j + x_i z_i x_j^2 z_j^2 + x_j z_j x_i^2 z_j^2 + 2x_i z_i x_j^2 z_j^2 + x_j^3 z_j^3$

$$\begin{aligned}(xz-1)^3 = \; & x_i^3 z_i^3 + 2x_i^2 z_i^2 x_j z_j + x_i z_i x_j^2 z_j^2 + x_j z_j x_i^2 z_j^2 + 2x_i z_i x_j^2 z_j^2 + x_j^3 z_j^3 \\ & - 3x_i^2 z_i^2 - 6x_i z_i x_j z_j - 3x_j^2 z_j^2 \\ & + x_i z_i + x_j z_j \\ & - 1\end{aligned}$$

→ cannot express this in terms of $\langle \phi(x), \phi(z) \rangle$

→ $\boxed{\text{not a valid kernel}}$

PROBLEM 4

(1) Exponential family distribution: $p(y;\eta) = b(y)\exp(\eta^T T(y)) - a(\eta))$

$$p(y;\phi) = (1-\phi)^{y-1}\phi$$

$$= \exp[(y-1)\cdot\log(1-\phi) + \log(\phi)]$$

$$= \exp[(y-1)\cdot\log(1-\phi) - \log(1-\exp[\log(1-\phi)])] \checkmark$$

$$\boxed{\begin{aligned} \eta &= \log(1-\phi) \\ a(\eta) &= -\log(1-\exp(\eta)) \\ T(y) &= y-1 \\ b(y) &= 1 \end{aligned}}$$

(2) From assumptions about GLMs:

$$\eta = w^T x$$

$$\text{log-likelihood} = \log\left[\exp[(y-1)\cdot\log(1-\phi) - \log(1-\exp[\log(1-\phi)])]\right]$$

$$= (y-1)\cdot\log(1-\phi) - \log(1-\exp[\log(1-\phi)])$$

$$= \eta^T(y-1) - \log(1-\exp(\eta))$$

$$= (w^T x)^T(y-1) - \log(1-\exp(w^T x))$$

$$= (x^T w)(y-1) - \log(1-\exp(w^T x))$$

$$\frac{\partial\,\text{log-likelihood}}{\partial w} = x^T(y-1) - (1-\exp(w^T x))^{-1}\cdot\exp(w^T x)\cdot x$$

$$= x^T(y-1) - \frac{\exp(w^T x)\cdot x}{1-\exp(w^T x)}$$

$$w \longleftarrow w + \alpha\left(\frac{\partial\,\text{log-likelihood}}{\partial w}\right)$$

$$\boxed{w \longleftarrow w + \alpha\left[x^T(y-1) - \frac{\exp(w^T x)\cdot x}{1-\exp(w^T x)}\right]}$$

PROBLEM 5 – attached (Python code)