T.1.) $p(y_2, y_1, y_0 \mid \alpha) = p(y_2 \mid y_1, y_0; \alpha)\, p(y_1 \mid y_0; \alpha)$

เนื่องจาก Markov process หาก y_2 มีค่า given y_1 แล้ว ไม่จำเป็นต้อง given y_0 แล้ว

จะได้ $p(y_2 \mid y_1; \alpha)\, p(y_1 \mid y_0; \alpha) = p(w_1 = y_2 - \alpha y_1)\, p(w_0 = y_1 - \alpha y_0)$

กำหนดให้ y is fixed แล้ว $p(y \mid \alpha) = L(\alpha)$

$$L(\alpha) = p(w_1 = y_2 - \alpha y_1)\, p(w_0 = y_1 - \alpha y_0) \quad ; \ w_1, w_0 \sim N(0, \sigma^2)$$

$$L(\alpha) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{((y_2 - \alpha y_1) - 0)^2}{2\sigma^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{((y_1 - \alpha y_0) - 0)^2}{2\sigma^2}}\right)$$

$$= \frac{1}{2\pi\sigma^2} e^{\frac{-(y_2 - \alpha y_1)^2 - (y_1 - \alpha y_0)^2}{2\sigma^2}}$$

$$\ln L(\alpha) = -\ln(2\pi) - 2\ln\sigma - \left[\frac{(y_2 - \alpha y_1)^2 + (y_1 - \alpha y_0)^2}{2\sigma^2}\right]$$

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{-1}{\cancel{2}\sigma^2}\left[\cancel{2}(y_2 - \alpha y_1)(-y_1) + \cancel{2}(y_1 - \alpha y_0)(-y_0)\right]$$

$$= \frac{1}{\sigma^2}\left[(y_2 - \alpha y_1)(y_1) + (y_1 - \alpha y_0) y_0\right] = 0$$

$$(y_1)(y_2 - \alpha y_1) = y_0(\alpha y_0 - y_1)$$

$$y_1 y_2 - \alpha y_1^2 = \alpha y_0^2 - y_1 y_0$$

$$y_1 y_2 + y_1 y_0 = \alpha(y_0^2 + y_1^2)$$

$$\therefore \boxed{\alpha = \frac{y_2 y_1 + y_1 y_0}{(y_0^2 + y_1)^2}}$$

OT 1.) $p(y_n, y_{n-1}, \ldots, y_0 \mid \alpha) = p(y_n \mid y_{n-1}, y_{n-2}, \ldots, y_0; \alpha)\, p(y_{n-1} \mid y_{n-2}, y_{n-3}, \ldots, y_0; \alpha) \ldots p(y_1 \mid y_0; \alpha)$

ตาม Markov Process จะได้ $p(y_n \mid y_{n-1}; \alpha)\, p(y_{n-1} \mid y_{n-2}; \alpha) \ldots p(y_1 \mid y_0; \alpha)$

$$= \prod_{k=1}^{n} p(y_k \mid y_{k-1}; \alpha) = \prod_{k=1}^{n} p(w_k = y_k - \alpha y_{k-1})$$

กำหนดให้ y is fixed แล้ว $p(y \mid \alpha) = L(\alpha)$

$$L(\alpha) = \prod_{k=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{((y_k - \alpha y_{k-1}) - 0)^2}{2\sigma^2}}\right)$$

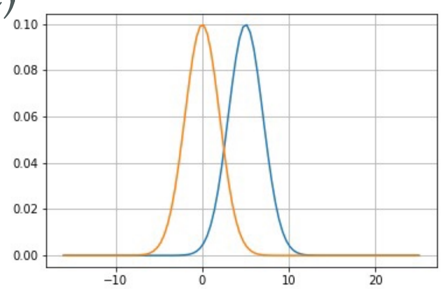
n

$$\ln L(\alpha) = \sum_{k=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \sum_{k=1} \frac{(y_k - \alpha y_{k-1})^2}{2\sigma^2}$$

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = \cancel{-}\frac{1}{\cancel{2\sigma^2}} \sum_{k=1}^{n} \cancel{2}\,(y_k - \alpha y_{k-1})(\cancel{-}y_{k-1}) = 0$$

$$\sum_{k=1}^{n} y_k y_{k-1} = \sum_{k=1}^{n} \alpha (y_{k-1})^2 \Rightarrow \therefore \boxed{\alpha = \frac{\sum_{k=1}^{n} y_k y_{k-1}}{\sum_{k=1}^{n} (y_{k-1})^2}}$$

T2)



$\dfrac{P(x|w_1)}{P(x|w_2)} \overset{?}{} \; 1$ เนื่องจาก prior ของทั้งคู่มีค่าเท่ากัน

หา decision boundary ได้โดย $\dfrac{P(x|w_1)}{P(x|w_2)} = 1$

$\therefore P(x|w_1) = P(x|w_2)$ ซึ่งก็คือจุดตัดกันของ normal distribution ดังภาพ.

$$f(x_1) = f(x_2)$$

$$\frac{1}{\cancel{\sqrt{2}}\cancel{\sqrt{2\pi}}} e^{-\frac{(x-5)^2}{2(2)}} = \frac{1}{\cancel{(\sqrt{2})}\cancel{\sqrt{2\pi}}} e^{-\frac{(x-0)^2}{2(2)}}$$

$$e^{-\frac{(x-5)^2}{\cancel{4}}} = e^{-\frac{x^2}{\cancel{4}}}$$

$$\cancel{-}\frac{(x-5)^2}{\cancel{4}} = \cancel{-}\frac{x^2}{\cancel{4}}$$

$$(x-5)^2 - x^2 = 0$$

$$-5(2x-5) = 0$$

$$2x - 5 = 0$$

$$\therefore \boxed{x = \frac{5}{2}}$$

T3.) find decision boundary

happy cat, sad cat

$$p(x|w_1)\,p(w_1) = p(x|w_2)\,p(w_2)$$

$$p(x|w_1)\,(0.8) = p(x|w_2)\,(0.2) \quad ; \; p(w_1)+p(w_2) = 1.0$$

$$\frac{p(x|w_1)}{p(x|w_2)} = \frac{1}{4} \Rightarrow p(x|w_1) = 4\,p(x|w_2)$$

$$(4)\frac{1}{\cancel{(\sqrt{2})}\cancel{\sqrt{2\pi}}} e^{-\frac{(x-5)^2}{2(2)}} = \frac{1}{\cancel{(\sqrt{8})}\cancel{\sqrt{2\pi}}} e^{-\frac{(x-0)^2}{2(2)}}$$

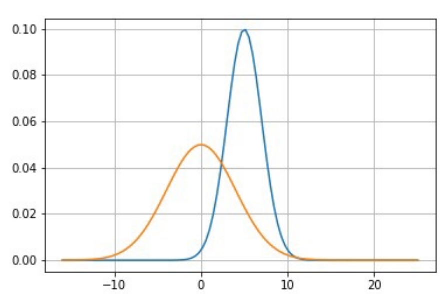
$$2\log 2 - \frac{(x-5)^2}{4} = -\frac{x^2}{4}$$

$$2\ln 2 = \frac{\cancel{x^2} - 10x + 25 - \cancel{x^2}}{4}$$

$$8\ln 2 = -10x + 25$$

$$10x = 25 - 8\ln 2$$

$$\boxed{\therefore x = 2.5 - 0.8\ln 2}$$



OT2) Determine $p(x|w_1) = N(\mu_1, \sigma^2)$, $P(x|w_2) = N(\mu_2, \sigma^2)$, $p(w_1) = p(w_2) = 0.5$, Prove that the decision boundary is a $x = \mu_1 + \mu_2$

And decision boundary

$$p(x|w_1)p(w_1) = p(x|w_2)p(w_2) \quad ; \quad p(w_1) = p(w_2) = 0.5$$

$$\frac{p(x|w_1)}{p(x|w_2)} = 1$$

$$p(x|w_1) = p(x|w_2) \; ; \; p(x|w_1) = N(\mu_1, \sigma^2), \; p(x|w_2) = N(\mu_2, \sigma^2)$$

$$\cancel{\frac{1}{\sigma\sqrt{2\pi}}} e^{-\frac{(x-\mu_1)^2}{2\sigma^2}} = \cancel{\frac{1}{\sigma\sqrt{2\pi}}} e^{-\frac{(x-\mu_2)^2}{2\sigma^2}}$$

$$\cancel{-}\frac{(x-\mu_1)^2}{\cancel{2\sigma^2}} = \cancel{-}\frac{(x-\mu_2)^2}{\cancel{2\sigma^2}}$$

$$(x-\mu_1)^2 - (x-\mu_2)^2 = 0$$

$$(\mu_2-\mu_1)(2x - (\mu_1+\mu_2)) = 0$$

$$[2x - (\mu_1+\mu_2)] = 0$$

$$\therefore \boxed{x = \frac{\mu_1+\mu_2}{2}}$$

- Determine $P(x|w_1) = N(5, 2)$

$\qquad\qquad P(x|w_2) = N(0, 4)$

find a decision boundary

$$p(x|w_1)p(w_1) = p(x|w_2)p(w_2)$$

$$p(x|w_1)\cancel{(0.5)} = p(x|w_2)\cancel{(0.5)}$$

$$\frac{1}{\sqrt{2}\,\cancel{\sqrt{2\pi}}} e^{-\frac{(x-5)^2}{2(2)}} = \frac{1}{\cancel{2}\,\cancel{\sqrt{2\pi}}} e^{-\frac{(x-0)^2}{2(4)}}$$

$$\sqrt{2}$$

$$(2)^{\frac{1}{2}} e^{\frac{-(x-5)^2}{4}} = e^{-\frac{x^2}{8}}$$

$$\frac{1}{2}\ln 2 - \frac{(x-5)^2}{4} = -\frac{x^2}{8}$$

$$4\ln 2 - 2(x-5)^2 = -x^2$$

$$4\ln 2 - 2(x^2 - 10x + 25) = -x^2$$

$$4\ln 2 - \cancel{2}x^2 + 20x - 50 = \cancel{-x^2}$$

$$x^2 - 20x + (50 - 4\ln 2) = 0$$

$$x = \frac{-(-20) \pm \sqrt{400 - 4(1)(50 - 4\ln 2)}}{2(1)} = 20 \pm \sqrt{400 - 200 + 16\ln 2}$$

$$= 10 \pm \sqrt{50 + 4\ln 2}$$

$\therefore \boxed{x = 10 - \sqrt{50 + 4\ln 2}}$ is a decision boundary

T4.) Age มี bin ที่เป็น 0 อยู่ 0

MonthlyIncome มี bin ที่เป็น 0 อยู่ 0

DistanceFromHome มี bin ที่เป็น 0 อยู่ 11

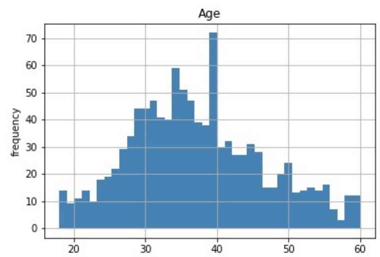
ที่ส่วนของ Age และ MonthlyIncome เป็นการ discretize ที่ดี เนื่องจาก ทุก bin มีค่า สามารถ generalize ปัญหาได้ ส่วน DistanceFromHome ยังไม่ดี เนื่องจากมี zero count bin ถึง 11 จาก 40 bin

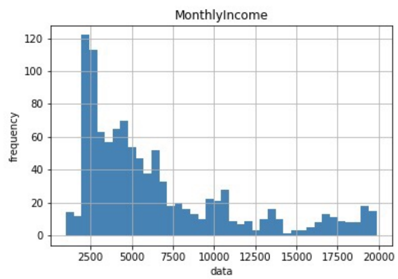
T5.) ไม่ได้ เนื่องจาก มีหลาย feature ที่ควรนำไป discretize ก่อน

ยกตัวอย่างเช่น Gender ถ้าเอาไม่เอาไป discretize ก่อน จะแทบไม่เห็นความต่างของ data ที่แบ่งเป็น Male(1) และ Female(0) โดย ∵ จะเกาะกลุ่มเป็น normal distribution เดียว แต่ discretize จะทำให้แยกเป็นคนละ normal distribution กัน

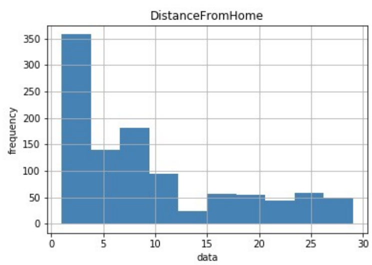
T6.)

bins = 40)





bins = 40



bins = 10

เนื่องจาก $\dfrac{\#\text{ zero count bin}}{\text{bins}}$ น้อยที่สุด

T2.7) feature ที่ควร discretize เช่น BusinessTravel, Department, EducationField เป็นต้น criteria ในการเลือก คือ เป็น feature ที่มีข้อมูล แตกต่างกันไม่มาก จำนวน bins ที่เหมาะสมที่นี้ คือ 10

T2.8) สังเกตจาก กราฟที่ plot ได้ใน ไฟล์ Employee_Attrition_Prediction
เหมาะใช้ "Binomial distribution" ∵ feature ที่สนใจส่วนใหญ่มีแค่สองค่า

T2.9) → prior distribution of two classes is "Bernoulli's distribution"

T10) ไม่ทำการคิดส่วนที่เป็น likelihood เลย หาก $P(x_i \mid attrition) = 0$

T11)

T12) สามารถดูคำตอบใน notebook ได้ในส่วนของการ implement และ Report the result ที่ไฟล์ Employee_Attrition_Prediction

T13)

T14)

T15)

T16) สำหรับวิธีการที่ implement ตรงๆ เลย threshold ที่ได้คือ 0.5
แต่สำหรับวิธีการที่นำไปทำ normal distribution ก่อน ได้ threshold เป็น 0.0