Hae Sung Jeong
005146648

# Homework 1

1. $A$ = event that an urn is selected

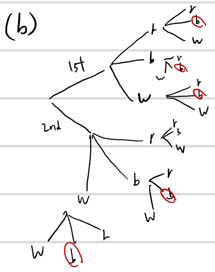
(a) Two balls are red

If the selected urn is 1st, $\left(\frac{4}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right)$

If the selected urn is 2nd, $\left(\frac{6}{10}\right)\left(\frac{2}{10}\right)\left(\frac{1}{9}\right)$

$\Rightarrow P(\text{two balls are red}) = \left(\frac{4}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right) + \left(\frac{6}{10}\right)\left(\frac{2}{10}\right)\left(\frac{1}{9}\right) = 0.0667$

(b)




$P(\text{the second ball is blue}) = \left(\frac{4}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right) + \left(\frac{4}{10}\right)\left(\frac{3}{10}\right)\left(\frac{2}{9}\right) + \left(\frac{4}{10}\right)\left(\frac{3}{10}\right)\left(\frac{3}{9}\right)$
$+ \left(\frac{6}{10}\right)\left(\frac{2}{10}\right)\left(\frac{4}{9}\right) + \left(\frac{6}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right) + \left(\frac{6}{10}\right)\left(\frac{4}{10}\right)\left(\frac{4}{9}\right) = 0.36$

(c) $P(\text{second ball is blue} \mid \text{first ball is red}) = \dfrac{P(\text{second ball is blue} \cap \text{first ball is red})}{P(\text{first ball is red})} = \dfrac{\left(\frac{4}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right) + \left(\frac{6}{10}\right)\left(\frac{2}{10}\right)\left(\frac{4}{9}\right)}{\left(\frac{4}{10}\right)\left(\frac{4}{10}\right) + \left(\frac{6}{10}\right)\left(\frac{2}{10}\right)} = 0.381$

2. Since the coin toss event is Binomial

(a) The expected value of $Z = np$
$= (4)\left(\frac{1}{2}\right) = 2$

(b) The variance of $Z = np(1-p) = (4)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$
$= 1$

3. (a) Let $X$ be the event that a student studied

Let $Y$ be the event that a student gets an A

$P(X \mid Y) = \dfrac{P(X \cap Y)}{P(Y)} \Rightarrow \dfrac{P(Y \mid X)P(X)}{P(Y)}$

$\Rightarrow P(Y) = \left(\frac{9}{10}\right)\left(\frac{85}{100}\right) + \left(\frac{1}{10}\right)\left(\frac{50}{100}\right)$

$\Rightarrow P(X \mid Y) = \dfrac{\left(\frac{9}{10}\right)\left(\frac{85}{100}\right)}{\left(\frac{9}{10}\right)\left(\frac{85}{100}\right) + \left(\frac{1}{10}\right)\left(\frac{50}{100}\right)} = 0.9387$

(b) Let $A$ be the event that a student didn't study

Let $B$ be the event that a student receive a B or lower

$P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B)}$

$\Rightarrow P(B) = \left(\frac{9}{10}\right)\left(\frac{15}{100}\right) + \left(\frac{1}{10}\right)\left(\frac{50}{100}\right)$

$P(A \mid B) = \dfrac{\left(\frac{1}{10}\right)\left(\frac{50}{100}\right)}{\left(\frac{9}{10}\right)\left(\frac{15}{100}\right) + \left(\frac{1}{10}\right)\left(\frac{50}{100}\right)} = 0.270$

4. (a) $E[X+Y] = E[X] + E[Y]$

$E[X+Y] = \sum_x \sum_y (x+y) P(x,y)$

$= \sum_x \sum_y x P(x,y) + \sum_x \sum_y y P(x,y)$

$= \sum_x x \sum_y P(x,y) + \sum_y y \sum_x P(x,y)$

$\Rightarrow \sum_y P(x,y) = P(x)$ , $\sum_x P(x,y) = P(y)$ by Total Probability Thm.

$\Rightarrow \sum_x x P(x) + \sum_y y P(y)$

$= E[X] + E[Y]$

(b) $var(X+Y) = var(X) + var(Y)$ if $X$ and $Y$ are independent

$var(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$

$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$

$= E[X^2] + 2E[XY] + E[Y^2] - [(E[X])^2 + 2E[X]E[Y] + (E[Y])^2]$

$= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 + 2E[XY] - 2E[X]E[Y]$

$= var(X) + var(Y) + 2E[XY] - 2E[X]E[Y]$

$\Rightarrow 2E[XY] = 2E[X]E[Y]$ by Independence

$var(X+Y) = var(X) + var(Y)$.

5. (a) Let A be the event that the bus will not arrive within d more seconds

Let B be the event that you have waited r seconds

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

$\Rightarrow P(A \cap B) = 1 - P(T \le r+d)$

$= 1 - (1 - e^{-\lambda(r+d)})$

$= e^{-\lambda(r+d)}$

$P(B) = 1 - P(T \le r) = e^{-\lambda r}$

$\Rightarrow P(A|B) = \frac{e^{-\lambda(r+d)}}{e^{-\lambda(r)}} = e^{-\lambda d}$

(b) $P(T \le t) = F_T(t)$

$\Rightarrow f_T(t) = \frac{dF_T}{dt}(t) = \frac{d}{dt}(1 - e^{-\lambda t})$

$= \lambda e^{-\lambda t}$

$E[T] = \int_0^\infty t \lambda e^{-\lambda t} dt$

$u = t \qquad dv = \lambda e^{-\lambda t} dt$

$du = dt \qquad v = -e^{-\lambda t}$

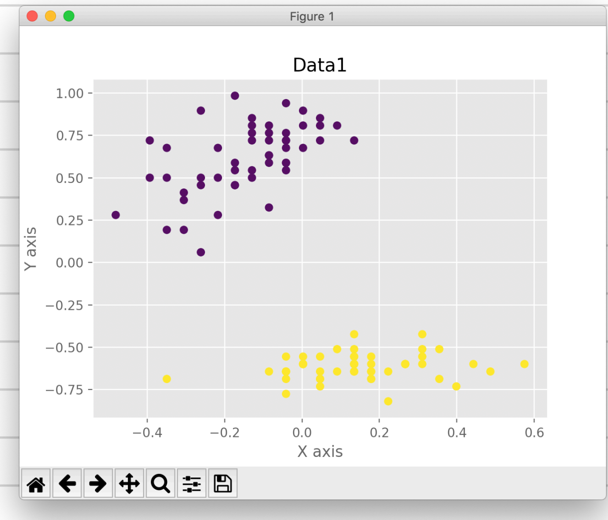
$= -t e^{-\lambda t} \Big|_0^\infty - \int_0^\infty -e^{-\lambda t} dt$

$= 0 - \frac{1}{\lambda} e^{-\lambda t} \Big|_0^\infty$

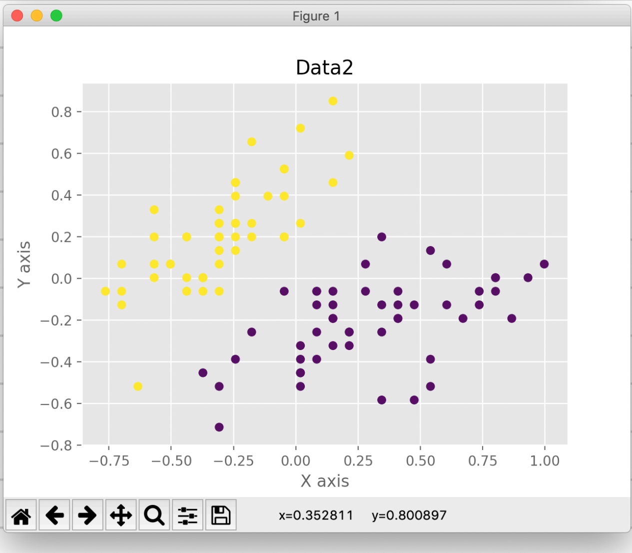
$= -\frac{1}{\lambda} e^{-\infty} - (-\frac{1}{\lambda} e^0)$

$= \frac{1}{\lambda}$

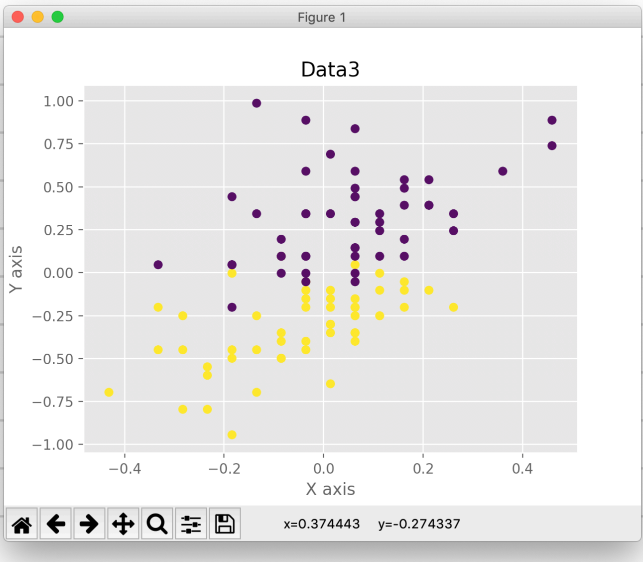
6. (a)




Data 1 is linearly separable

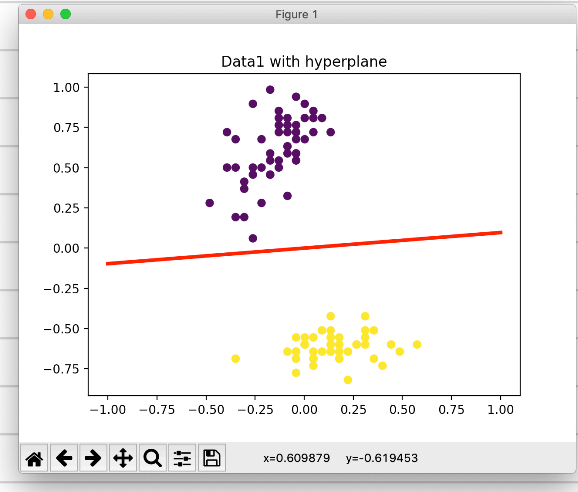



Data 2 is linearly separable




Data 3 is not linearly separable

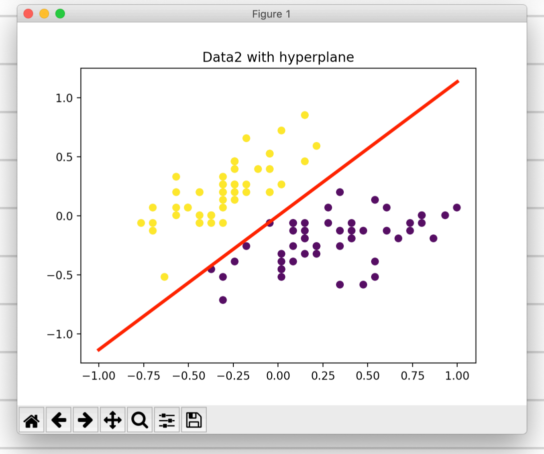
(b)




$w = [0.264028 \quad -2.72838]$

$b = 0$

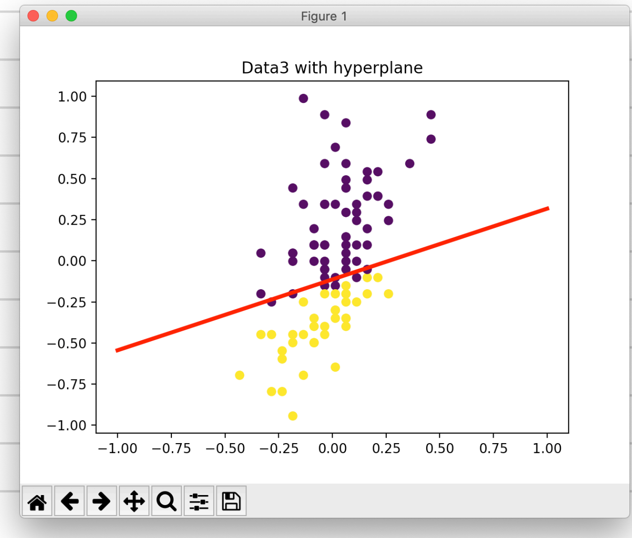
$u = 2$




$W = [-7.698916 \quad 6.785506]$

$b = 0$

$u = 32$




$W = [7.587248 \quad -17.6200492]$

$b = -2$

$u = 5909$

(c) For each linearly separable dataset, the number of updates is less than the upper bound ($\frac{1}{\gamma_{w,b}^2}$)