HW 2

Name: Yanan Liu

NetID: yl2248.

P1

1.1

$$P(y|\bar{x}) = \frac{P(\bar{x}|y)\cdot P(y)}{P(\bar{x})} = \frac{P(\bar{x}, y)}{P(\bar{x})} = \frac{P(\bar{x}, y)}{\sum_{i=0}^{n} P(\bar{x}, y_n)} = \frac{P(\bar{x}, y)}{P(\bar{x}, 1) + P(\bar{x}, -1)}$$

1.2.

goal: $1 - P(y=1|\bar{x}) = P(y=-1|\bar{x})$

$$P(y=+1|\bar{x}) = \frac{P(\bar{x}|y=+1)\cdot P(\bar{x})}{P(\bar{x})} = \frac{P(\bar{x}, +1)}{P(\bar{x})} = \frac{P(\bar{x}, +1)}{P(\bar{x}, +1) + P(\bar{x}, -1)}$$

$$1 - \frac{P(\bar{x}, +1)}{P(\bar{x}, +1) + P(\bar{x}, -1)} = \frac{P(\bar{x}, -1)}{P(\bar{x}, +1) + P(\bar{x}, -1)} = \frac{P(\bar{x}, -1)}{P(\bar{x})} = \frac{P(y=-1|\bar{x})\cdot P(\bar{x})}{P(\bar{x})}$$

$$= \frac{P(y=-1|\bar{x})}{P(\bar{x})}$$

∴ the error probability is $P(y=-1|\bar{x})$

1.3.

As proved in the 1.2

$P(y=-1|\bar{x}) = 1 - P(y=1|\bar{x})$

since the max value of the P_{max} can be Calculate MAP.

∴ the min value can by find by

$1 - P_{max}$

1.4:

• map estimator:

$$\arg\max P(y|\bar{x}) = \max\left[\frac{P(\bar{x}|y)\cdot 0.5}{P(\bar{x})}, \frac{P(\bar{x}|y')\cdot 0.5}{P(\bar{x})}\right]$$

$$= \max\left[\frac{P(\bar{x}|y)\cdot 0.5}{0.5\,[P(\bar{x}|-1) + P(\bar{x}|+1)]}, \frac{P(\bar{x}|y')\cdot 0.5}{0.5\,[P(\bar{x}|-1) + P(\bar{x}|+1)]}\right]$$

$$= \max\left[\frac{P(\bar{x}|y)}{P(\bar{x}|-1) + P(\bar{x}|+1)}, \frac{P(\bar{x}|y')}{P(\bar{x}|-1) + P(\bar{x}|+1)}\right]$$

$$= \max\left[\frac{P(\bar{x}|+1)}{P(\bar{x}|-1) + P(\bar{x}|+1)}, \frac{P(\bar{x}|-1)}{P(\bar{x}|-1) + P(\bar{x}|+1)}\right] \cdots ①$$

ML:

$$\arg\max P(\bar{x}|y) = \max\left[P(\bar{x}|-1), P(\bar{x}|+1)\right] \cdots ②$$

For ①, the denomitor does not affect the result and the numerator for ① & ② is the same.

∴ ML = Map.

1.5

Assume: $P(y=+1) = 0.6$ $P(\bar{x}|-1) = 0.8$

$P(y=-1) = 0.4$ $P(\bar{x}|+1) = 0.2$

Map:

3.

3.1

DAY	OUTLOOK	TEMPER.	HUMIDITY	WIND	PLAY? (LABEL)
1	Sunny	Hot	High	Weak	NO
2	Sunny	Hot	High	Strong	NO
3	Overcast	Hot	High	Weak	YES
4	Rain	Mild	High	Weak	YES
5	Rain	Cold	Normal	Weak	YES
6	Rain	Cold	Normal	Strong	NO
7	Overcast	Cold	Normal	Strong	YES
8	Sunny	Mild	High	Weak	NO
9	Sunny	Cold	Normal	Weak	YES
10	Rain	Mild	Normal	Weak	YES
11	Sunny	Mild	Normal	Strong	YES
12	Overcast	Mild	High	Strong	YES
13	Overcast	Hot	Normal	Weak	YES
14	Rain	Mild	High	Strong	NO

	Yes/+	No/+
Sunny	2/2+1	3/3+1
Overcast	4/5	0/1
Rain	3/4	1/2
Hot	2/3	2/3
mild	4/5	2/3
Cold	3/4	1/2
Normal	6/7	1/2
High	3/4	4/5
weak	6/7	2/3
storng	3/4	3/4

Label "Yes"

$P(\text{Sunny}|\text{Yes}) = 0.25$ $P(\text{overcast}|\text{Yes}) = 0.416$

$P(\text{Rain}|\text{Yes}) = 0.333$

$P(\text{Cold}|\text{Yes}) = 0.333$ $P(\text{Mild}|\text{Yes}) = 0.416$ $P(\text{Hot}|\text{Yes}) = 0.25$

$P(\text{High}|\text{Yes}) = 0.363$ $P(\text{Normal}|\text{Yes}) = 0.636$

$P(\text{weak}|\text{yes}) = 0.636$ $P(\text{strong}|\text{Yes}) = 0.363$

Label "No"

$P(\text{Sunny}|\text{no}) = 0.5$ $P(\text{overcast}|\text{no}) = 0.125$ $P(\text{Rain}|\text{no}) = 0.375$

$P(\text{Hot}|\text{no}) = 0.25$ $P(\text{Mild}|\text{No}) = 0.375$ $P(\text{Cold}|\text{no}) = 0.25$

$P(\text{High}|\text{no}) = 0.714$ $P(\text{normal}|\text{no}) = 0.285$

$P(\text{weak}|\text{No}) = 0.428$ $P(\text{strom}|\text{no}) = 0.571$

3.2.

Label Yes: $P(\text{Yes}) = \frac{9}{14}$

Label No: $P(\text{No}) = \frac{5}{14}$

3.3

$P(\text{overcast}, \text{Hot}, \text{High}, \text{Strong})$

when $\beta = 0$

$P(\text{Data}|\text{Yes}) = \frac{9}{14} \times \frac{4}{9} \times \frac{2}{9} \times \frac{3}{9} \times \frac{3}{9} = 0.007$

$P(\text{Data}|\text{no}) = 0$

when $\beta = 1$

$P(\text{Data}|\text{Yes}) = \frac{9}{14} \times \frac{5}{12} \times \frac{3}{12} \times \frac{4}{11} \times \frac{4}{11} = 0.008$

$P(\text{Data}|\text{no}) = \frac{5}{14} \times \frac{1}{8} \times \frac{3}{8} \times \frac{4}{7} \times \frac{5}{7} = 0.006$

when $\beta = \infty$

$P(\text{Data}|\text{Yes}) = \frac{9}{14} = 0.642$

$P(\text{Data}|\text{no}) = \frac{5}{14} = 0.357$

For all β above, the predict is all "Yes".

P4

4.1

the joint density: $P(x_1, x_2, \cdots x_n) = \prod_{i=0}^{n} P(x_i)$

since maximizing $f(x)$ is equl to maxining $\log(f(x))$.

∴ $\ln(P(x_1, x_2, \cdots x_n)) = \sum_{i=1}^{n} \ln P(x_i)$

$$\arg\max P(y|\bar{x}) = \left(\frac{P(y=+1)\cdot P(\bar{x}|y)}{P(\bar{x})}, \frac{P(y=-1)\cdot P(\bar{x}|y=-1)}{P(\bar{x})} \right)$$

$$= \left(\frac{0.6\times 0.2}{0.4\times 0.8 + 0.2\times 0.6}, \frac{0.4\times 0.8}{0.4\times 0.8 + 0.2\times 0.6} \right)$$

$$= \max(0.272, 0.727)$$

Label -1

ML:

$$\arg\max P(\bar{x}|y) = \max(P(\bar{x}|y=+1), P(\bar{x}|y=-1))$$

$$= \max(0.8, 0.2)$$

Label: $+1$

For map we pick $y=-1$, For ML we pick $y=+1$

P2

2.1

$A(n_H, n_T)$

the probability of n_H for A is

$$P = C_{H+T}^{H}\, \mu^{n_H} (1-\mu)^{n_T}$$

in order to find the max value.

$$P' = [C_{H+T}^{H}\, \mu^{n_H} (1-\mu)^{n_T}]'$$

as C_{H+T}^{H} is a constant, we can ignore it.

apply log function to both sides.

$$h(x) = \log P = \log \mu^{n_H} (1-\mu)^{n_T}$$

$$= \log \mu^{n_H} + \log (1-\mu)^{n_T}$$

$$= n_H \log \mu + n_T \log(1-\mu)$$

$$h'(x) = \frac{n_H}{\mu} + \frac{n_T}{1-\mu}$$

when $h'(x)=0$ the P is max.

$$\therefore \frac{n_H}{\mu} + \frac{n}{1-\mu} = 0$$

$$\frac{n_H}{\mu} = -\frac{n}{1-\mu}$$

$$n_H = \mu n + \mu n_H$$

Therefore $\mu = \frac{n_H}{n+n_H}$

2.2.

as the process is n independent toss.

$\therefore P(H), P(T)$ are independent.

$$\therefore \arg\max_{\mu} p(\mu|A(n_H, n_T)) = \arg\max_{\mu} \left(\frac{P(\mu)\, P(A(n_H, n_T)|\mu)}{P(A(n_H, n_T))}\right)$$

since the denominator will not affect the final result

therefore

$$\arg\max_{\mu} p(\mu|A(n_H, n_T)) = \arg\max_{\mu} P(\mu)\, p(A(n_H, n_T)|\mu)$$

2.3

from 2.2. we can get:

$$\arg\max_{\mu} p(\mu|n_H, n_T) = \arg\max_{\mu} P(\mu)\, P(A(n_H, n_T)|\mu)$$

$$F(\mu) = P(\mu)\, P(A(n_H, n_T), n_T)|\mu) = \frac{\mu^{\alpha}(1-\mu)^{\beta}}{\int_0^1 x^{\alpha}(1-x)^{\beta}dx} \cdot \binom{n_H+n_T}{n_T} \mu^{n_H} (1-\mu)^{n_T}$$

In order to find max value, we should find $F'(\mu)=0$

$$F(\mu) = \frac{\binom{n_H+n_T}{n_T}}{\int_0^1 x^{\alpha}(1-x)^{\beta}dx} \cdot \mu^{n_H+\alpha} (1-\mu)^{\beta+n_T}$$

$$\log F(\mu) = \log \mu^{n_H+\alpha} + \log (1-\mu)^{\beta+n_T}$$

$$F'(\mu) = \frac{n_H+\alpha}{\mu} - \frac{\beta+n_T}{1-\mu} = 0.$$

$$\frac{n_H+\alpha}{\mu} = \frac{\beta+n_T}{1-\mu}$$

$$\mu = \frac{n_H+\alpha}{n_H+\alpha+n_T+\beta}$$

Therefore, $\arg\max_{\mu} P(\mu|n_H, n_T) = \frac{n_H+\alpha}{n_H+\alpha+n_T+\beta}$

$$\sum_{i=1}^{n} \ln P(x_i) = \sum_{i=1}^{n} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right)$$

$$= \sum_{i=1}^{n}\left(\ln \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(x-\mu)^2}{2\sigma^2}\right)$$

we can ignore $\ln(\sqrt{2\pi\sigma^2})^{-1}$ as it is a constant

$$\therefore \text{Fuction} = \sum_{i=1}^{n} -\frac{(x_i-\mu)^2}{2\sigma^2} = H(\mu)$$

in order to find the max μ.

$$H'(\mu) = 0 = \sum_{i=1}^{n} \frac{x_i-\mu}{\sigma^2} = 0$$

$$\sum_{i=1}^{n} x_i = n\mu \qquad \mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

Therefore $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$

4.2.

from 4.1. We can get

$$F(\sigma) = \sum_{i=1}^{n}\left(\ln \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \sum_{i=1}^{n} \left(-\ln\sigma - \ln\sqrt{2\pi} - \frac{(x-\mu)^2}{2\sigma^2}\right)$$

the same as P4.1, in order to get the max σ, the $F'(\sigma)=0$

$$F'(\sigma) = \frac{dF(\sigma)}{d\sigma} = \sum_{i=1}^{n}\left(-\frac{1}{\sigma} + \frac{(x_i-\mu)^2}{\sigma^3}\right) = 0$$

$$\therefore \frac{n}{\sigma} = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^3}$$

$$n\sigma^2 = \sum_{i=1}^{n} (x_i-\mu)^2$$

Therefore: $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

4.3.

$$E(\mu_{ML}) = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\cdot n\cdot \mu = \mu$$

$$V(\mu_{ML}) = V\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2\right] = \frac{1}{n^2} V\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} V[(x_i-\mu)^2]$$

$$= \frac{1}{n}\cdot\sigma^2$$

$$= \frac{\sigma^2}{n}$$