Q1(a) Y = tellle free or not

X = sunny or not

$Y = \{0, 1\}$ & $X = \{0, 1\}$

Here we are using the Bernoulli distributions

Well we need to find

$P(Y|X=0)$ & $P(Y|X=1)$.

where the parameters are α_1 & α_2 for each of the bernoulli distributions

let $\alpha = (\alpha_1, \alpha_2)$.

$\therefore$ we have ~~pt~~ when $X = 0$

for a sample $x_i\, y_i$

$$-\ln p(y_i | x_i = 0, \alpha) = -\ln \alpha_1^{y_i} (1-\alpha_1)^{y_i}$$

& for $x_i y_1$ when $X = 1$

$$-\ln(p(y_i | x_i = 1, \alpha) = -\ln \alpha_2^{y_i} (1-\alpha_2)^{y_2}$$

since $x_i = 0$ or 1

for a set $I_1 = \{1, \dots n\} | x_i = 0$ & $I_2 = \{1, \dots n\} | x_i = 1$

These are mutually exclusive

$I_1 \cup I_2 = \{1 \dots n\}$ as $x_i = 0$ or 1

$\therefore$ Negative log likelihood

$$\sum_{i \in I_1} -\ln(p(y_i | x_i = 0, \alpha) + \sum_{i \in I_2} -\ln(p(y_i | x_i = 1, \alpha)$$

$$= \sum_{i \in I_0} -\ln(p_i | \alpha_1) + \sum_{i \in I_1} -\ln(p_i | \alpha_2).$$

(b) sps $P(Y=0|X=1) > P(Y=1|X=1)$

~~[illegible]~~ $\Leftrightarrow$ say $P(Y=0|X=1) > 0.5$

That implies that ~~[illegible]~~ given that it is sunny it most probable that the table is not free.

sps ~~$P(Y=1)$~~ $P(Y=1|X=1) > P(Y=0|X=1)$

$\Leftrightarrow$ say $P(Y=1|X=1) > 0.5$

That implies that given that it is sunny it's most probable that the table is free

(c) Well consider another varible $Z = \{0, 1, 2\}$

0 → morning
1 → afternoon
2 → evening

Well now they are 6 total possible outcome's

$P(Y \mid x = 0 \wedge z = 0)$
$P(Y \mid x = 0 \wedge z = 1)$
$P(Y \mid x = 0 \wedge z = 2)$
$P(Y \mid x = 1 \wedge z = 0)$
$P(Y \mid x = 1 \wedge z = 1)$
$P(Y \mid x = 1 \wedge z = 2)$

Well now the parameters are x q z

$$\therefore\ p(\{y_i\}_{i=1}^{n} \mid \underset{\alpha}{x}\ \&\ \underset{\beta}{z}) = p(\{y_i\}_{i=1}^{n} \mid \alpha, \beta)$$

$$= \prod_{i=1}^{n} \left[\alpha^{y_i} (1-\alpha)^{1-y_i} \prod_{j<i} \left(\frac{\alpha}{\alpha\ \beta} \right)^{y_j y_i} \right]$$

Well now these are six bernoulli distributions with parameter α_1 α_2 α_3 α_4 α_5 α_6

let $\alpha' = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$

Well the MLE becomes

$$p(y_i \mid x_i = \{0, 1\} \wedge z_i = \{0, 1, 2\}, \alpha)$$

Q2 X RV

$\mathcal{N}(\mu_1, \sigma_1^2) \quad \mathcal{N}(\mu_2, \sigma_2^2)$

$p(x) = \frac{1}{2}\left(\mathcal{N}(\mu_1, \sigma_1^2) + \mathcal{N}(\mu_2, \sigma_2^2)\right), \quad \mathcal{N}(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(w-\mu)^2} = p(w)$

$D = \{x_i\}_{i=1}^n$

(a) $p(D|f) = p(\underbrace{\{x_i\}_{i=1}^n}_{= w_i} | f)$

$f \in F \leftarrow$ family of Gaussians

$= \prod_{i=1}^n 0.5\, \mathcal{N}(\mu_1, \sigma_1^2) + 0.5\, \mathcal{N}(\mu_2, \sigma_2^2)$

$= \prod_{i=1}^n \frac{1}{2}\left(\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2} + \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2}(w_i - \mu_2)^2}\right)$

$= \prod_{i=1}^n \frac{1}{2\sqrt{2\pi}}\left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i - \mu_2)^2}}{\sigma_2}\right)$

Taking log:

$= \sum_{i=1}^n \log\left(\frac{1}{2\sqrt{2\pi}}\left(\underbrace{\frac{e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i - \mu_2)^2}}{\sigma_2}}\right)\right)$

$= \sum_{i=1}^n \left(-\log(2\sqrt{2\pi}) + \log\left(\quad \right)\right)$

$$= -\log(2\sqrt{2\pi})n + \cancel{\log} \sum_{i=1}^{n} \left(\log \left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i-\mu_2)^2}}{\sigma_2} \right) \right)$$

negative log likelihood $\Rightarrow$ $C(t)$

$$\Rightarrow -\sum_{i=1}^{n} \left(\log \left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_i} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i-\mu_2)^2}}{\sigma_2} \right) \right) + \log(2\sqrt{2\pi})n.$$

(b)

$$\frac{\partial}{\partial \sigma_1} \left(\left(-\sum_{i=1}^{n} \left(\log \left(\overbrace{\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i-\mu_2)^2}}{\sigma_2}}^{A} \right) \right) \right) \right)' + 0$$

$$= -\sum_{i=1}^{N} \frac{1}{A} \left(\left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1} \right)' + 0 \right)$$

NOW let $u = \frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1}$

$$\log u = -\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2 - \log(\sigma_1)$$

$$\frac{1}{u}\frac{du}{\partial \sigma_1} = \frac{(w_i-\mu_1)^2}{\sigma_1^3} - \frac{1}{\sigma_1}$$

$$\frac{du}{\partial \sigma_1} = \frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1} \left(\frac{(w_i-\mu_1)^2 \cancel{\sigma_1} - \sigma_1^2}{\sigma_1^3} \right)$$

$$\frac{du}{\partial \sigma_1} = e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2} \frac{((w_i-\mu_1)^2 \cancel{\sigma_1} - \sigma_1^2)}{\sigma_1^4}.$$

$$= -\sum_{i=1}^{N} \frac{1}{A}\left(e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2} \frac{((w_i-\mu_1)^2 - \sigma_1^2)}{\sigma_1^4}\right)$$

$$= \sum_{i=1}^{N} \frac{1}{A}\left(e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2} \frac{(\sigma_1^2-(w_i-\mu_1)^2)}{\sigma_1^4}\right)$$

For $\frac{\partial}{\partial\sigma_2}$ replace $\sigma_1 \to \sigma_2$ & $\mu_1 \to \mu_2$ in final expression

$$\frac{\partial}{\partial\mu_1}\left(-\sum_{i=1}^{N}\left(\log\left(\underbrace{\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i-\mu_1)^2}}{\sigma_2}}_{A}\right)\right)\right)$$

$$= -\sum_{i=1}^{N} \frac{1}{A}\left(\left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1}\right)' + 0\right)$$

let $v = \frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1}$

$$\log v = -\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2 - \log\sigma_1$$

$$\frac{1}{v}\frac{dv}{d\mu_1} = \left(\frac{-1}{\sigma_1^2}(w_i-\mu_1)\cdot -1 - 0\right)$$

$$\frac{dv}{d\mu_1} = \left(\frac{(w_i-\mu_1)}{\sigma_1^2}\right)\frac{e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}}{\sigma_1}$$

$$\frac{\partial}{\partial\mu_1}\cancel{\left(-\sum\right.}$$

$$\frac{\partial}{\partial\mu_1} = -\sum_{i=1}^{N} \cancel{\log}\ \frac{1}{A}\left(\frac{w_i-\mu_1}{\sigma_1^3} e^{-\frac{1}{2\sigma_1^2}(w_i-\mu_1)^2}\right)$$

for $\frac{\partial}{\partial\mu_2}$ replace $\sigma_1 \to \sigma_2$ & $\mu_1 \to \mu_2$ in

Now $\frac{d}{d\sigma_1} = 0$

$$\sum_{i=1}^{N} \frac{1}{A} \left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2} (\sigma_1^2 - (w_i - \mu_1)^2)}{\sigma_1^4} \right) = 0$$

$$= \sum_{i=1}^{N} e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2} (\sigma_1^2 - (w_i - \mu_1)^2) = 0$$

Well $e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2} \neq 0$ unless $w_i = \infty$

$$\Rightarrow \sum_{i=1}^{N} (\sigma_1^2 - (w_i - \mu_1)^2) = 0$$

$$n\sigma_1^2 - \sum_{i=1}^{N} (w_i - \mu_1)^2 = 0$$

$$\sigma_1 = \sqrt{\sum_{i=1}^{N} (w_i - \mu_1)^2 \cdot \frac{1}{n}}$$

Similarly $\sigma_2 = \sqrt{\sum_{i=1}^{N} (w_i - \mu_2)^2 \cdot \frac{1}{n}}$

$\frac{d}{d\mu_1} = 0$

$$\sum_{i=1}^{N} \frac{1}{A} \left(\frac{\mu_1 - w_i}{\sigma_1^3} e^{\frac{-1}{2\sigma_1^2}(w_i - \mu_1)^2} \right) = 0$$

$$\sum_{i=1}^{N} (\mu_1 - w_i) = 0$$

$$\sum_{i=1}^{N} \mu_1 - \sum_{i=1}^{N} w_i = 0$$

$$n\mu_1 = \sum_{i=1}^{N} w_i \Rightarrow \mu_1 = \frac{1}{n}\sum_{i=1}^{N} w_i$$

Similarly $\mu_2 = \frac{1}{n}\sum_{i=1}^{N} w_i$

Well $w_{t+1} \leftarrow w_t - \eta \frac{\partial}{\partial \sigma_1}$

$w_{t+1} \leftarrow w_t - \eta \frac{\partial}{\partial \sigma_2}$

$w_{t+1} \leftarrow w_t - \eta \frac{\partial}{\partial \mu_1}$

$w_{t+1} \leftarrow w_t - \eta \frac{\partial}{\partial \mu_2}$

$$\therefore \nabla C_w(t) = \left(\frac{\partial}{\partial \sigma_1}, \underbrace{\frac{\partial}{\partial \sigma_2}, \frac{\partial}{\partial \mu_1}, \frac{\partial}{\partial \mu_2}}_{\text{computed already}} \right)$$

$$w_{t+1} = w_t - \eta \nabla C_w(t).$$

$$, C_w(t) = \log(\sqrt{2\pi}\,2)n - \sum_{i=1}^{N} \log\left(\frac{e^{-\frac{1}{2\sigma_1^2}(w_i - \mu_1)^2}}{\sigma_1} + \frac{e^{-\frac{1}{2\sigma_2^2}(w_i - \mu_2)^2}}{\sigma_2} \right)$$

Q3			Mean	std
	(b)	~~Mean Regressor~~		
		Mean Regressor	31.3451	0.4715
		Random "	33.548	19.7075
		Range Regressor	65.6259	1.9842
		Stochastic "	11.9776	0.2316
	(c)	Batch "	24.5310	0.3728
		MiniBatch "	11.9908	0.2625
	(f)	Stochastic Heuristic "	14.9138	3.2064
		Batch Heuristic "	11.9512	0.2621
		MiniBatch Heuristic "	12.0006	0.2177