2010 Theory II #1

1a) $$P(Y_i = y_i \mid X_i) = \frac{p(x_i, y_i)}{\sum_{y=0}^{1} p(x_i, y_i)}$$

$$= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}\left[\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2\right\}\right]^{y_i}\left[\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_0)^2\right\}\right]^{1-y_i} \pi_1^{y_i}(1-\pi_1)^{1-y_i}}{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_0)^2\right\} + \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2\right\}}$$

Then

$$\operatorname{logit}\left(P(Y_i = 1 \mid X_i)\right) = \log\left\{\frac{P(Y_i=1 \mid X_i)}{P(Y_i=0 \mid X_i)}\right\}$$

$$= \log\left\{\frac{\pi_1 \exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2\right\}}{(1-\pi_1)\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_0)^2\right\}}\right\}$$

$$= \operatorname{logit}\pi_1 + \frac{1}{2\sigma^2}\left[\cancel{x_i^2} - 2\mu_0 x_i + \mu_0^2 - \left(\cancel{x_i^2} - 2\mu_1 x_i + \mu_1^2\right)\right]$$

$$= \underbrace{\operatorname{logit}\pi_1 + \frac{\mu_0^2-\mu_1^2}{2\sigma^2}}_{\alpha_0} + \underbrace{\frac{(\mu_1-\mu_0)}{\sigma^2}}_{\alpha_1} x_i$$

1b) See 4.9 in the GLM solutions for this.

No need to calculate θ yet

1c) $p(x_i, y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \left[\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_1)^2\right\}\right]^{y_i} \left[\exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu_0)^2\right\}\right]^{1-y_i} \pi_1^{y_i}(1-\pi_1)^{1-y_i}$

$$\ell_n(\theta) = \sum_{i=1}^{n} \left\{ -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{y_i}{2\sigma^2}(x_i-\mu_1)^2 - \frac{1-y_i}{2\sigma^2}(x_i-\mu_0)^2 + y_i \log\left(\frac{\pi_1}{1-\pi_1}\right) + \log(1-\pi_1) \right\}$$

Let $A = \{i : y_i = 0\}$, $B = \{i : y_i = 1\}$

$$\frac{\partial}{\partial \mu_0}\ell_n(\theta) = \frac{1}{\sigma^2}\sum_{i \in A}(x_i - \mu_0) \overset{set}{=} 0 \Rightarrow \hat{\mu}_0 = \bar{x}_A$$

$$\frac{\partial}{\partial \mu_1}\ell_n(\theta) = \frac{1}{\sigma^2}\sum_{i \in B}(x_i - \mu_1) \overset{set}{=} 0 \Rightarrow \hat{\mu}_1 = \bar{x}_B$$

$$\frac{\partial}{\partial \sigma^2}\ell_n(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i \in A}(x_i-\mu_0)^2 + \frac{1}{2(\sigma^2)^2}\sum_{i \in B}(x_i-\mu_1)^2 \overset{set}{=} 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i \in A}(x_i - \mu_0) + \sum_{i \in B}(x_i - \mu_1)}{n}$$

$$\frac{\partial}{\partial \pi}\ell_n(\theta) = \sum_{i=1}^{n}\frac{\partial}{\partial \pi}\left\{y_i \log \pi_1 + (1-y_i)\log(1-\pi)\right\}$$

$$= \sum_{i=1}^{n}\left\{\frac{y_i}{\pi_1} - \frac{1-y_i}{1-\pi}\right\} = \frac{n\bar{y}_n}{\pi_1} - \frac{n(1-\bar{y}_n)}{1-\pi_1} \overset{set}{=} 0$$

$$\Rightarrow (1-\hat{\pi}_1)\bar{y}_n = \hat{\pi}_1(1-\bar{y}_n) \Rightarrow \bar{y}_n - \cancel{\hat{\pi}_1\bar{y}_n} = \hat{\pi}_1 - \cancel{\hat{\pi}_1\bar{y}_n}$$

$$\Rightarrow \hat{\pi}_1 = \bar{y}_n$$

$$\frac{\partial^2}{\partial\mu_0^2} l_n(\theta) = -\frac{1}{\sigma^2}\sum_{i\in A} 1 = -\frac{1}{\sigma^2}\sum(1-y_i) \qquad \frac{\partial^2}{\partial\mu_1^2} l_n(\theta) = -\frac{1}{\sigma^2}\sum_{i\in B} = -\frac{1}{\sigma^2}\sum y_i$$

$$\frac{\partial^2}{\partial(\sigma^2)^2} l_n(\theta) = \frac{n}{2(\sigma^2)^2} - \frac{1}{(\sigma^2)^3}\sum(1-y_i)(x_i-\mu_0)^2 - \frac{1}{(\sigma^2)^3}\sum y_i(x_i-\mu_1)^2$$

$$\frac{\partial^2}{\partial\sigma^2\partial\mu_0} l_n(\theta) = -\frac{1}{(\sigma^2)^2}\sum_{i\in A}(x_i-\mu_0) = -\frac{1}{(\sigma^2)^2}\sum(1-y_i)(x_i-\mu_0)$$

$$\frac{\partial^2}{\partial\sigma^2\partial\mu_1} l_n(\theta) = -\frac{1}{(\sigma^2)^2}\sum_{i\in B}(x_i-\mu_1) = -\frac{1}{(\sigma^2)^2}\sum y_i(x_i-\mu_1)$$

$$\frac{\partial^2}{\partial\pi^2} l_n(\theta) = -\frac{n\bar{y}_n}{\pi_1^2} - \frac{n(1-\bar{y}_n)}{(1-\pi_1)^2} = -n\left[\frac{\bar{y}_n(1-\pi_1)^2 + \pi_1^2(1-\bar{y}_n)}{\pi_1^2(1-\pi_1)^2}\right]$$

$$= -n\left[\frac{\bar{y}_n(1-2\pi_1+\cancel{\pi_1^2}) + \pi_1^2 - \cancel{\pi_1^2\bar{y}_n}}{\pi_1^2(1-\pi_1)^2}\right]$$

$$= -n\left[\frac{\bar{y}_n - 2\pi_1\bar{y}_n + \pi_1^2}{\pi_1^2(1-\pi_1)^2}\right]$$

0 for all other partial derivatives

Notice that

and similarly for $E[(1-Y_i)(X_i-\mu_0)]$

$$E[Y_i(X_i-\mu_1)] = E\left(Y_i\, E[X_i-\mu_1 \mid Y_i]\right)$$

$$\sum_{y_i=0}^{1} y_i\, E[X_i-\mu_1 \mid Y_i=y_i]\,\pi_1^{y_i}(1-\pi_1)^{1-y_i} = E[X_i-\mu_1 \mid Y_i=1]\,\pi_1 = 0$$

Then

$$I_n(\mu_0, \mu_1, \sigma^2, \pi_1) = \begin{bmatrix} \frac{n(1-\pi_1)}{\sigma^2} & & & \\ & \frac{n\pi_1}{\sigma^2} & & \\ & & -\left(\frac{n}{2(\sigma^2)^2} - \frac{n\pi_1}{(\sigma^2)^2} - \frac{n(1-\pi_1)}{(\sigma^2)^2}\right) = \frac{n}{2(\sigma^2)^2} & \\ & & & \frac{n}{\pi_1(1-\pi_1)} \end{bmatrix}$$

since

$$E\left[Y_i(X_i-\mu_1)^2\right] = E\left(Y_i\, E\left[(X_i-\mu_1)^2 \mid Y_i\right]\right)$$

$$= \sum_{y_i=0}^{1} y_i\, E\left[(X_i-\mu_1)^2 \mid Y_i = y_i\right] \pi_1^{y_i} (1-\pi_1)^{1-y_i}$$

$$= E\left[(X_i-\mu_1)^2 \mid Y_i = 1\right] \pi_1 = \sigma^2 \pi_1$$

and similarly

$$E\left[(1-Y_i)(X_i-\mu_0)^2\right] = \sigma^2(1-\pi_1)$$

Then

$$\left[\lim_{n\to\infty} \frac{1}{n} I_n(\theta)\right]^{-1} = \sigma^2 \begin{bmatrix} (1-\pi_1)^{-1} & & & \\ & \pi_1^{-1} & & \\ & & 2\sigma^2 & \\ & & & \frac{\pi_1(1-\pi_1)}{\sigma^2} \end{bmatrix}$$

1d) $$g(\mu_0, \mu_1, \sigma^2, \pi_1) = \left(\operatorname{logit}\pi + \frac{\mu_0^2 - \mu_1^2}{2\sigma^2},\ \frac{\mu_1 - \mu_0}{\sigma^2}\right)$$

$$\frac{\partial}{\partial \mu_0} g = \left(\frac{\mu_0}{\sigma^2},\ -\frac{1}{\sigma^2}\right) \qquad \frac{\partial}{\partial \mu_1} g = \left(-\frac{\mu_1}{\sigma^2},\ \frac{1}{\sigma^2}\right)$$

$$\frac{\partial}{\partial \sigma^2} g(\theta) = \left(-\frac{\mu_0^2 - \mu_1^2}{2(\sigma^2)^2} \, , \; -\frac{\mu_1 - \mu_0}{(\sigma^2)^2} \right)$$

$$\frac{\partial}{\partial \pi_1} g(\theta) = \left(\frac{\partial}{\partial \pi} \log \frac{\pi_1}{1-\pi_1} \, , \; 0 \right) = \left(\frac{(1-\pi)}{\pi} \frac{1-\pi_1+\pi_1}{(1-\pi_1)^2} \, , \; 0 \right) = \left(\frac{1}{\pi(1-\pi_1)} \, , \; 0 \right)$$

$$\dot{g}(\theta) \Sigma [\dot{g}(\theta)]'$$

$$= \frac{1}{\sigma^2} \begin{bmatrix} \mu_0 & -\mu_1 & -\frac{\mu_0^2-\mu_1^2}{2\sigma^2} & [\sigma^2 \pi (1-\pi)]^{-1} \\ -1 & 1 & -\frac{\mu_1-\mu_0}{\sigma^2} & 0 \end{bmatrix} \sigma^2 \begin{bmatrix} (1-\pi_1)^{-1} & & & \\ & \pi_1^{-1} & & \\ & & 2\sigma^2 & \\ & & & \frac{\pi_1(1-\pi_1)}{\sigma^2} \end{bmatrix} [\dot{g}(\theta)]'$$

$$= \begin{bmatrix} \frac{\mu_0}{1-\pi_1} & -\frac{\mu_1}{\pi_1} & -(\mu_0^2-\mu_1^2) & 1 \\ -\frac{1}{1-\pi_1} & \frac{1}{\pi_1} & -2(\mu_1-\mu_0) & 0 \end{bmatrix} \frac{1}{\sigma^2} \begin{bmatrix} \mu_0 & -1 \\ -\mu_1 & 1 \\ -\frac{\mu_0^2-\mu_1^2}{2\sigma^2} & -\frac{\mu_1-\mu_0}{\sigma^2} \\ \frac{\sigma^2}{\pi_1(1-\pi_1)} & 0 \end{bmatrix}$$

$$= \frac{1}{\sigma^2} \begin{bmatrix} \frac{\mu_0^2}{1-\pi_1} + \frac{\mu_1^2}{\pi_1} + \frac{(\mu_0^2-\mu_1^2)^2}{2\sigma^2} + \frac{\sigma^2}{\pi_1(1-\pi_1)} \, , & -\frac{\mu_0}{1-\pi_1} - \frac{\mu_1}{\pi_1} + \frac{(\mu_0^2-\mu_1^2)(\mu_0-\mu_1)}{\sigma^2} \\ -\frac{\mu_0}{1-\pi_1} - \frac{\mu_1}{\pi_1} + \frac{(\mu_0^2-\mu_1^2)(\mu_0-\mu_1)}{\sigma^2} \, , & \frac{1}{\pi_1(1-\pi_1)} + \frac{2(\mu_1-\mu_0)^2}{\sigma^2} \end{bmatrix}$$

6

e) This implies that $X_i \perp\!\!\!\perp Y_i$ since $p(x_i, y_i) = p(x_i|y_i)p(y_i)$

$= p(x_i)p(y_i)$. so that $\text{logit}(P(Y_i=1|X_i)) = \text{logit}(P(Y_i=1))$

$= \text{logit}\,\pi$

Then

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

$$\frac{\partial^2}{\partial (\sigma^2)^2} l_n(\theta) = \frac{\partial}{\partial \sigma^2}\left\{-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

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

$$\frac{\partial^2}{\partial \pi_1^2} l_n(\theta) = -n\left[\frac{\bar{y}_n - 2\pi_1\bar{y}_n + \pi_1^2}{\pi_1^2(1-\pi_1)^2}\right] \qquad 0 \text{ otherwise}$$

so that

$$\left[\lim_{n\to\infty}\frac{1}{n} I_n(\theta)\right]^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} & & \\ & \frac{1}{2\sigma^4} & \\ & & \frac{1}{\pi_1(1-\pi_1)} \end{bmatrix}^{-1} = \begin{bmatrix} \sigma^2 & & \\ & 2\sigma^4 & \\ & & \pi_1(1-\pi_1) \end{bmatrix}$$

Let $h(\theta) = \text{logit}\,\pi$ then

$$\frac{\partial}{\partial \pi} h(\theta) = \frac{1-\pi_1}{\pi_1}\cdot\frac{1}{(1-\pi_1)^2} = \frac{1}{\pi_1(1-\pi_1)}$$

So that $\dot{h}(\theta) = (0, 0, [\pi_1(1-\pi_1)]^{-1})$ and hence

$$\sqrt{n}\,(\text{logit}\,\hat{\pi} - \text{logit}\,\pi) \xrightarrow{\mathcal{L}} N\left(0, \frac{1}{\pi(1-\pi_1)}\right)$$

The model is

$$\text{logit}(\pi_i) = \alpha_0 \Rightarrow \pi_i = \frac{e^{\alpha_0}}{1+e^{\alpha_0}}$$

Then

$$\frac{\partial}{\partial \alpha_0} l_n(\alpha_0) = \frac{\partial}{\partial \alpha_0} \sum_{i=1}^{n} \left\{ y_i \,\text{logit}\,\pi_i + \log(1-\pi_i) \right\}$$

$$= \frac{\partial}{\partial \alpha_0} \sum_{i=1}^{n} \left\{ y_i \alpha_0 - \log(1+e^{\alpha_0}) \right\} = \sum_{i=1}^{n} \left\{ y_i - \frac{e^{\alpha_0}}{1+e^{\alpha_0}} \right\}$$

and

$$\frac{\partial^2}{\partial \alpha_0^2} l_n(\alpha_0) = \frac{\partial}{\partial \alpha_0} \sum_{i=1}^{n} \left\{ y_i - \frac{e^{\alpha_0}}{1+e^{\alpha_0}} \right\} = -n \frac{e^{\alpha_0}}{1+e^{\alpha_0}} \cdot \frac{1}{1+e^{\alpha_0}} = -n\pi_1(1-\pi_1)$$

so that

$$\left[\lim_{n\to\infty} \frac{1}{n} I_n(\alpha_0)\right]^{-1} = \frac{1}{\pi_1(1-\pi_1)}$$