2010 Theory II #3

3a) $P(\beta' X \geq \beta' Y) = P(\beta'(X-Y) \geq 0)$

$$= P\left(\frac{\beta'(X-Y) - \beta'(\mu_1-\mu_2)}{\sqrt{2\beta'\Sigma\beta}} \geq -\frac{\beta'(\mu_1-\mu_2)}{\sqrt{2\beta'\Sigma\beta}} \right)$$

$$= P\left(N(0,1) \geq -\frac{\beta'(\mu_1-\mu_2)}{\sqrt{2\beta'\Sigma\beta}} \right) = P\left(N(0,1) \leq \frac{\beta'(\mu_1-\mu_2)}{\sqrt{2\beta'\Sigma\beta}} \right)$$

3b) Recall the Cauchy-Schwartz inequality: for a, b vectors of an inner product space

$\|a'b\| \leq \|a\| \|b\| \Rightarrow (a'b)^2 \leq (a'a)(b'b)$

Then

$$\frac{\beta'(\mu_1-\mu_2)}{\sqrt{2\beta'\Sigma\beta}} = \left\{ \frac{[\beta'(\mu_1-\mu_2)]^2}{2\beta'\Sigma\beta} \right\}^{1/2}$$

$$\overset{\Sigma \text{ is pd}}{=} \left\{ \frac{[(\Sigma^{1/2}\beta)' \, \Sigma^{-1/2}(\mu_1-\mu_2)]^2}{2\beta'\Sigma\beta} \right\}^{1/2}$$

$$\overset{C\text{-}S}{\leq} \left\{ \frac{[(\Sigma^{1/2}\beta)'\Sigma^{1/2}\beta] \, [(\Sigma^{-1/2}(\mu_1-\mu_2))'\Sigma^{-1/2}(\mu_1-\mu_2)]}{2\beta'\Sigma\beta} \right\}^{1/2}$$

$$= \sqrt{\tfrac{1}{2}(\mu_1-\mu_2)\Sigma^{-1}(\mu_1-\mu_2)}$$

But choosing $\beta = \Sigma^{-1}(\mu_1-\mu_2)$ yields $\dfrac{\beta'(\mu_1-\mu_2)}{2\sqrt{\beta'\Sigma^{-1}\beta}} = \sqrt{\dfrac{(\mu_1-\mu_2)\Sigma^{-1}(\mu_1-\mu_2)}{2}}$

and since Φ is monotone this choice of β maximizes $AUC(\beta)$

$$p(\underline{x},\underline{y};\ \mu_1,\mu_2,\Sigma) = \left(\prod_{i=1}^{n} (2\pi)^{-1}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(x_i-\mu_1)'\Sigma^{-1}(x_i-\mu_1)\right\}\right)$$

$$\times\left(\prod_{j=1}^{m} (2\pi)^{-1}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(y_j-\mu_2)'\Sigma^{-1}(y_j-\mu_2)\right\}\right)$$

$$\ell_n(\mu_1,\mu_2,\Sigma) = -(n+m)\log(2\pi) - \frac{n+m}{2}\log(|\Sigma|)$$

$$-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_1)'\Sigma^{-1}(x_i-\mu_1) - \frac{1}{2}\sum_{j=1}^{m}(y_i-\mu_2)'\Sigma^{-1}(y_i-\mu_2)$$

$$\frac{\partial}{\partial\mu_1}\ell_n(\mu_1,\mu_2,\Sigma) = \frac{\partial}{\partial\mu_1}\left\{ n\mu_1'\Sigma^{-1}\bar{x}_n - \frac{n}{2}\mu_1'\Sigma^{-1}\mu_1\right\}$$

$$= n\Sigma^{-1}\bar{x}_n - n\Sigma^{-1}\mu_1 \stackrel{set}{=} 0 \Rightarrow \hat{\mu}_1 = \bar{x}_n$$

Similarly we obtain

$$\hat{\mu}_2 = \bar{y}_m$$

Next,

$$\frac{\partial}{\partial\Sigma^{-1}}\ell_n(\mu_1,\mu_2,\Sigma) = \frac{\partial}{\partial\Sigma^{-1}}\left\{\frac{n+m}{2}\log(|\Sigma^{-1}|) - \frac{1}{2}\sum_{i=1}^{n} tr\left[\Sigma^{-1}(x_i-\mu_1)(x_i-\mu_1)'\right]\right.$$

$$\left.-\frac{1}{2}\sum_{j=1}^{m} tr\left[\Sigma^{-1}(y_j-\mu_2)(y_j-\mu_2)'\right]\right\}$$

$$= \frac{n+m}{2}\Sigma - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_1)(x_i-\mu_1)' - \frac{1}{2}\sum_{j=1}^{m}(y_j-\mu_2)(y_j-\mu_2)' \stackrel{set}{=} 0$$

$$\Rightarrow \hat{\Sigma} = \frac{1}{n+m}\sum_{i=1}^{n}(x_i-\hat{\mu}_1)(x_i-\hat{\mu}_1)' + \frac{1}{n+m}\sum_{j=1}^{m}(y_j-\hat{\mu}_2)(y_j-\hat{\mu}_2)'$$

Then by the invariance property of MLEs,

$$\hat{A} = \Phi\left(\sqrt{\tfrac{1}{2}(\hat{\mu}_1-\hat{\mu}_2)\hat{\Sigma}^{-1}(\hat{\mu}_1-\hat{\mu}_2)}\right)$$

3d) By CLT and since $X_n \xrightarrow{L} X$, $Y_n \xrightarrow{L} Y$ all independent

$$\Rightarrow \begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{L} N\left(0, \begin{bmatrix} \Sigma_X & 0 \\ 0 & \Sigma_Y \end{bmatrix}\right)$$

$$\sqrt{n}\left(\begin{bmatrix} \bar{X}_{n1} \\ \bar{X}_{n2} \\ \overline{X^2_{n1}} \\ \overline{X^2_{n2}} \\ \overline{X_{n1}X_{n2}} \\ \bar{Y}_{n1} \\ \bar{Y}_{n2} \\ \overline{Y^2_{n1}} \\ \overline{Y^2_{n2}} \\ \overline{Y_{n1}Y_{n2}} \end{bmatrix} - \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \sigma_{11}+\mu_{11}^2 \\ \sigma_{22}+\mu_{21}^2 \\ \sigma_{12}+\mu_{11}\mu_{12} \\ \mu_{12} \\ \mu_{22} \\ \sigma_{11}+\mu_{12} \\ \sigma_{22}+\mu_{22}^2 \\ \sigma_{12}+\mu_{12}\mu_{22} \end{bmatrix}\right) \xrightarrow{L} N\left(0, \begin{bmatrix} \tilde{\Gamma}_X & 0 \\ 0 & \Gamma_Y \end{bmatrix}\right)$$

where $\tilde{\Gamma}_X$ and Γ_Y can be calculated using mgf's

Now $\Sigma^{-1} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1} = \dfrac{1}{\sigma_{11}\sigma_{12}-\sigma_{12}^2}\begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix}$

so that

$$(\mu_1-\mu_2)'\Sigma^{-1}(\mu_1-\mu_2) = \frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\Big\{(\mu_{11}-\mu_{12})^2\sigma_{22} - 2(\mu_{11}-\mu_{12})(\mu_{21}-\mu_{22})\sigma_{12} + (\mu_{21}-\mu_{22})^2\sigma_{11}\Big\}$$

Now we can construct a function $g : \mathbb{R}^{10} \mapsto \mathbb{R}$ s.t.

$$g(\bar{x}_{n1}, \bar{x}_{n2}, \overline{x^2_{n1}}, \overline{x^2_{n2}}, \overline{x_{n1}x_{n2}}, \overline{y_{n1}}, \overline{y_{n2}}, \overline{y^2_{n1}}, \overline{y^2_{n2}}, \overline{y_{n1}y_{n2}}) = \hat{A}$$

and it can be easily verified that

$$g(\mu_{11}, \mu_{21}, \sigma_{11}+\mu_{11}^2, \sigma_{22}+\mu_{21}^2, \sigma_{12}+\mu_{11}\mu_{12}, \mu_{12}, \mu_{22}, \sigma_{11}+\mu_{12}^2, \sigma_{22}+\mu_{22}^2, \sigma_{12}+\mu_{12}\mu_{22})$$
$$= A^{optimal}$$

Then ~~so~~ by the delta method, the asymptotic variance is given by $\dot{g}(\xi)\begin{bmatrix} \Gamma_X & 0 \\ 0 & \Gamma_Y \end{bmatrix}[\dot{g}(\xi)]'$

*Alternatively, we could verify the regularity conditions for MLE estimates. Then

$$\sqrt{n}\left(\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\Sigma} \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \\ \Sigma \end{bmatrix}\right) \xrightarrow{L} N\left(0, \left[\lim_{n\to\infty} \frac{1}{n} I_n(\mu_1, \mu_2, \Sigma)\right]^{-1}\right)$$

and we could again use the delta method to obtain the desired quantity

Need to do 3e!!