2010 Theory I #1

1a) $p((x_1,y_1),\ldots,(x_n,y_n)) = \prod_{i=1}^{n} (2\pi)^{-1} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}\left(\begin{bmatrix}x_i\\y_i\end{bmatrix}-\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}\right)' \Sigma^{-1} \left(\begin{bmatrix}x_i\\y_i\end{bmatrix}-\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}\right)\right\}$

$$= \prod_{i=1}^{n} (2\pi)^{-1} (\sigma_1^2\sigma_2^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma_1^2}(x_i-\mu_1)^2 - \frac{1}{2\sigma_2^2}(y_i-\mu_2)^2\right\}$$

$$= (2\pi)^{-n}(\sigma_1^2\sigma_2^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\mu_1)^2 - \frac{1}{2\sigma_2^2}\sum_{i=1}^{n}(y_i-\mu_2)^2\right\}$$

$$= (2\pi)^{-n}(\sigma_1^2\sigma_2^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i^2-2x_i\mu_1+\mu_1^2) - \frac{1}{2\sigma_2^2}\sum_{i=1}^{n}(y_i^2-2y_i\mu_2-\mu_2^2)\right\}$$

$$= (2\pi)^{-n}(\sigma_1^2\sigma_2^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_1^2}\sum x_i^2 + \frac{n\mu_1}{\sigma_1^2}\bar{x}_n - \frac{n\mu_1^2}{2\sigma_1^2} - \frac{1}{2\sigma_2^2}\sum y_i^2 + \frac{n\mu_2}{\sigma_2^2}\bar{y}_n - \frac{n\mu_2^2}{2\sigma_2^2}\right\}$$

$$= (2\pi)^{-n}(\sigma_1^2\sigma_2^2)^{-n/2} \exp\left\{\left(\frac{1}{2\sigma_2^2}-\frac{1}{2\sigma_1^2}\right)\sum x_i^2 + \frac{n\mu_1}{\sigma_1^2}\bar{x}_n - \frac{n\mu_1^2}{2\sigma_1^2} - \frac{1}{2\sigma_2^2}\left(\sum x_i^2+\sum y_i^2\right) + \frac{n\mu_2}{\sigma_2^2}\bar{y}_n - \frac{n\mu_2^2}{2\sigma_2^2}\right\}$$

which is exponential family with

$T_1 = \sum X_i^2$ $\qquad \theta_1 = \frac{1}{2\sigma_2^2} - \frac{1}{2\sigma_1^2}$

$T_2 = \sum X_i^2 + \sum Y_i^2$ $\qquad \theta_2 = -\frac{1}{2\sigma_2^2}$

$T_3 = \bar{X}_n$ $\qquad \theta_3 = \frac{n\mu_1}{\sigma_1^2}$

$T_4 = \bar{Y}_n$ $\qquad \theta_4 = \frac{n\mu_2}{\sigma_2^2}$

It is clear that we can find an open set ~~[illegible]~~ in the neighborhood of, say, $(0, -1, 0, 0)$ so that the density is full-rank. ~~[illegible]~~ Thus $\underline{T}$ is sufficient for $\underline{\theta}$.

~~[illegible]~~ The hypothesis is equivalent to $H_0: \theta_1 = 0$ vs. $H_1: \theta_1 \neq 0$.
By Thm 2.7 ~~[illegible]~~ a UMPU test is of the form

$$\phi(t_1) = \begin{cases} 1, & t_1 < c_1(t_2, t_3, t_4) \text{ or } t_1 > c_2(t_2, t_3, t_4) \\ 0, & \text{else} \end{cases}$$

*Note assume wlog $\Delta_0 = 1$

where $c_1(\cdot)$, $c_2(\cdot)$ satisfy

$$\begin{cases} E_{\theta_0}[\phi(T_1) \mid T_2, T_3, T_4] = \alpha \\ E_{\theta_0}[T_1 \phi(T_1) \mid T_2, T_3, T_4] = \alpha E_{\theta_0}[T_1 \mid T_2, T_3, T_4] \end{cases}$$

Consider

want to find an increasing fcn of T which satisfies Basu's thm and that we can work toward an F-test

$$V = \frac{T_1 - T_3^2}{T_2 - T_3 - T_4} = \frac{\sum X_i^2 - n\bar{X}_n^2}{\sum X_i^2 + \sum Y_i^2 - n\bar{X}_n^2 - n\bar{Y}_n^2} = \frac{\sum (X_i - \bar{X}_n)^2}{\sum (X_i - \bar{X}_n)^2 + \sum (Y_i - \bar{Y}_n)^2}$$

$a(t)U + b(t)$ for $a(t) > 0$?

Now since V is increasing in T_1 for fixed (T_2, T_3, T_4), the hypothesis is unchanged if we replace t_1 with v.

Now V is a linear function of T_1. Want ~~[illegible]~~ to verify Basu's conditions, ~~[illegible]~~ Firstly T_1 ~~[illegible]~~ exponential family implies that it is complete sufficient. Notice that complete ~~[illegible]~~ is stronger than bounded complete as is needed for Basu's.

Next, under H_0, (write $\sigma^2 = \sigma_1^2 = \sigma_2^2$)

$$V = \frac{(n-1)\sum(X_i - \bar{X}_n)^2/\sigma^2}{(n-1)\sum(X_i-\bar{X}_n)^2/\sigma^2 + (n-1)\sum(Y_i-\bar{Y}_n)^2/\sigma^2} = \frac{\chi^2_{n-1}}{\chi^2_{n-1} + \chi^2_{n-1}}$$

so that V is ancillary of (T_2, T_3, T_4) and Basu's thm conditions are satisfied. Thus we may rewrite our UMPU test as

$$\phi(v) = \begin{cases} 1, & v < c_1 \text{ or } v > c_2 \\ 0, & \text{else} \end{cases}$$

where c_1, c_2 satisfy

$$\begin{cases} E_{\theta_0}[\phi(V)] = \alpha \\ E_{\theta_0}[V\phi(V)] = \alpha E_{\theta_0}[V] \end{cases}$$

Now

$$c_1 > V = \frac{\sum(X_i - \bar{X}_n)^2}{\sum(X_i-\bar{X}_n)^2 + \sum(Y_i - \bar{Y}_n)^2}$$

$$\Leftrightarrow c_1^* < 1 + \frac{\sum(Y_i - \bar{Y}_n)^2}{\sum(X_i - \bar{X}_n)^2}$$

$$\Leftrightarrow c_1^{**} < \frac{\sum(Y_i - \bar{Y}_n)^2/(n-1)}{\sum(X_i-\bar{X}_n)^2/(n-1)}$$

$$\Leftrightarrow c_1^{***} > F \equiv \frac{\sum(X_i - \bar{X}_n)^2/(n-1)}{\sum(Y_i - \bar{Y}_n)^2/(n-1)} \overset{H_0}{\sim} F_{n-1,n-1}$$

Similarly,

$c_2 < V \iff c_2 < F$. Thus we have our UMPU test of the form

~~It remains to find~~

$$\phi(F) = \begin{cases} 1, & F < c_1 \text{ or } F > c_2 \\ 0, & \text{else} \end{cases}$$

Note that $F \overset{d}{=} 1/F$ so that

$$P(F < c_1 \text{ or } F > c_2) = P(1/F < c_1 \text{ or } 1/F > c_2)$$

$$= P\left(F > \frac{1}{c_1} \text{ or } F < \frac{1}{c_2}\right)$$

But since the test is UMPU it must be that $c_1 = \frac{1}{c_2}$. Thus we can solve

$$\alpha = E_{\theta_0}[\phi(F)] = P_{\theta_0}[F < c_1] + P_{\theta_0}\left[F > \frac{1}{c_1}\right]$$

$$= 2 P_{\theta_0}[F < c_1] \implies c_1 = F_{n-1,n-1}(\alpha/2)$$

and similarly

$$\alpha = 2 P_{\theta_0}[F > c_2] \implies c_2 = F_{n-1,n-1}(1 - \alpha/2)$$

~~Thus the test is~~

~~$\phi(F) = \{1,$~~

1b) Again assume wlog $\Delta_0 = 1$. ~~Now~~ It is easy to show that under $H_0: \sigma^2 = \sigma_1^2 = \sigma_2^2$ we have

$$\tilde{\mu}_1 = \bar{X}_n, \quad \tilde{\mu}_2 = \bar{y}_n, \quad \text{and} \quad \tilde{\sigma}^2 = \frac{1}{2n}\left[\sum(X_i - \bar{X}_n)^2 + \sum(y_i - \bar{y}_n)^2\right]$$

and that in the unrestricted case

$$\hat{\mu}_1 = \bar{x}_n, \quad \hat{\mu}_2 = \bar{y}_n, \quad \hat{\sigma}_1^2 = \frac{1}{n}\sum(x_i - \bar{x}_n)^2, \quad \hat{\sigma}_2^2 = \frac{1}{n}\sum(y_i - \bar{y}_n)^2$$

$$LRT_n = \frac{(2\pi)^{-n}(\tilde{\sigma}^2)^{-n}\exp\left\{-\frac{1}{2\tilde{\sigma}^2}\left[\sum(x_i - \tilde{\mu}_1)^2 + \sum(y_i - \tilde{\mu}_2)\right]\right\}}{(2\pi)^{-n}(\hat{\sigma}_1^2\hat{\sigma}_2^2)^{-n/2}\exp\left\{-\frac{1}{2\hat{\sigma}_1^2}\sum(x_i - \hat{\mu}_1)^2 - \frac{1}{2\hat{\sigma}_2^2}\sum(y_i - \hat{\mu}_2)^2\right\}}$$

$$= \left\{\frac{\hat{\sigma}_1^2\hat{\sigma}_2^2}{(\tilde{\sigma}^2)^2}\right\}^n \frac{\exp\{-n\}}{\exp\{-n\}} = \left\{\frac{\frac{1}{n^2}\left[\sum(x_i - \bar{x}_n)^2\right]\left[\sum(y_i - \bar{y}_n)^2\right]}{\frac{1}{4n^2}\left[\sum(x_i - \bar{x}_n)^2 + \sum(y_i - \bar{y}_n)^2\right]^2}\right\}^n$$

Thus we reject if

$$k < \left\{ \cdot \right\} \iff k^* < \frac{\left[\sum(x_i - \bar{x}_n)^2\right]\left[\sum(y_i - \bar{y}_n)^2\right]}{\left[\sum(x_i - \bar{x}_n)^2 + \sum(y_i - \bar{y}_n)^2\right]^2}$$

$$= \frac{\left[\sum(x_i - \bar{x}_n)^2\right]\left[\sum(y_i - \bar{y}_n)^2\right]}{\left[\sum(x_i - \bar{x}_n)^2 + \sum(y_i - \bar{y}_n)^2\right]^2} \cdot \frac{1/\left[\sum(y_i - \bar{y}_n)^2\right]^2}{1/\left[\sum(y_i - \bar{y}_n)^2\right]^2}$$

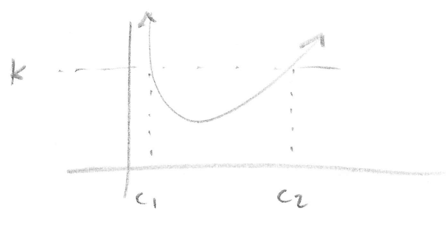
$$= \frac{\dfrac{\sum(x_i - \bar{x}_n)^2}{\sum(y_i - \bar{y}_n)^2}}{\left[\dfrac{\sum(x_i - \bar{x}_n)^2}{\sum(y_i - \bar{y}_n)^2} + \dfrac{\sum(y_i - \bar{y}_n)^2}{\sum(x_i - \bar{y}_n)^2}\right]^2} = \frac{F}{[1 + F]^2}$$

for F as before

$$\iff \quad k^{*} > \frac{(1+F)^2}{F}$$

Now $\frac{\partial^2}{\partial t^2} \frac{(1+t)^2}{t} = \frac{2}{t^3} > 0$ for $t > 0$ so we see

that the graph of $f(t) = \frac{(1+t)^2}{t}$ is given by



so we reject if $F < c_1$ or $F > c_2$. But ~~since~~ ~~k~~ then

$\cancel{F = \frac{\sum(X_i - \bar{X}_n)^2}{\sum(Y_i - \bar{Y}_n)^2}}$ the same argument as in (1a) ~~see~~ follows

so that we obtain the same test.

$\frac{1}{n}\sum(x_i-\bar{x})(x_i-x)^{*}$

1c) In the unrestricted model

$$\hat{\mu}_1 = \bar{x}_n, \quad \hat{\mu}_2 = \bar{y}_n, \quad \hat{\sigma}_1^2 = \frac{1}{n}\sum(x_i - \bar{x}_n)^2, \quad \hat{\sigma}_2^2 = \frac{1}{n}\sum(y_i - \bar{y}_n)^2$$

$$\hat{\sigma}_{12} = \frac{1}{n}\sum(x_i - \bar{x}_n)(y_i - \bar{y}_n)$$

Since $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_1^2 \sigma_2^2}}$ we have $R = \frac{\hat{\sigma}_{12}}{\sqrt{\hat{\sigma}_1 \hat{\sigma}_2}}$

then

$$L(\mu, \Sigma) = (2\pi)^{-n} |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\begin{bmatrix} x_i \\ y_i\end{bmatrix} - \begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}\right)' \Sigma^{-1}\left(\begin{bmatrix} x_i \\ y_i\end{bmatrix} - \begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}\right)\right\}$$

$$|\Sigma| = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2 - \sigma_1^2\sigma_2^2\rho^2 = \sigma_1^2\sigma_2^2(1-\rho^2)$$

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2\end{bmatrix}$$

$$(a\ b)\begin{bmatrix}x_{11} & x_{12} \\ x_{12} & x_{22}\end{bmatrix}\begin{pmatrix}a \\ b\end{pmatrix} = (a\ b)\begin{bmatrix}ax_{11} + bx_{12} \\ ax_{12} + bx_{22}\end{bmatrix} = a^2x_{11} + 2abx_{12} + b^2x_{22}$$

$$= (2\pi)^{-n}\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right]^{-n/2}\exp\left\{-\frac{1}{2\sigma_1^2\sigma_2^2(1-\rho^2)}\sum_{i=1}^{n}\left[(x_i-\mu_1)^2\sigma_2^2 - 2(x_i-\mu_1)(y_i-\mu_2)\sigma_{12} + (y_i-\mu_2)^2\right]\right\}$$

$$= (2\pi)^{-n}\left[\sigma_1^2\sigma_2^2(1-\rho^2)\right]^{-n/2}\exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{i=1}^{n}\left[\frac{(x_i-\mu_1)^2}{\sigma_1^2} - \frac{2(x_i-\mu_1)(y_i-\mu_2)\sigma_{12}}{\sigma_1^2\sigma_2^2} + \frac{(y_i-\mu_2)^2}{\sigma_2}\right]\right\}$$

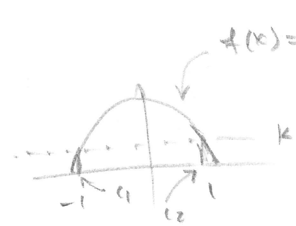
so that

$$-\frac{2n(1-R^2)}{2(1-R^2)} = -n$$

$$L(\hat{\mu}, \hat{\Sigma}) = (2\pi)^{-n}\left[\hat{\sigma}_1^2\hat{\sigma}_2^2(1-R^2)\right]^{-n/2}\exp\left\{-\frac{1}{2(1-R^2)}\left[n - 2nR^2 + n\right]\right\}$$

then

$$\Lambda = \frac{L(\tilde{\mu}, \tilde{\Sigma})}{L(\hat{\mu}, \hat{\Sigma})} = (1-R^2)^{n/2}$$



so we reject when

$$k < 1-R^2 \iff R^2 < k^* \iff |R| < k^{**}$$

~~Under Ho R is symmetrical about 0 so the test rejects when |R| > c~~ Under Ho R is symmetric about 0 so choosing k that satisfies

$E_{\theta_0}[\phi(R)] = \alpha$ yields rejection region symmetric about 0

1c.iii) We have $R = \dfrac{\sum (x_i - \bar{x}_n)^2 (y_i - \bar{y}_n)^2}{\sqrt{\sum (x_i - \bar{x}_n)^2 \sum (y_i - \bar{y}_n)^2}} =$

Want to find the distribution of

$$\frac{\sqrt{n-2}\, R}{\sqrt{1-R^2}} = \frac{s_{xy}}{\sqrt{s_x^2 s_y^2}} \Big/ \sqrt{1 - \frac{s_{xy}^2}{s_x^2 s_y^2}} = \frac{s_{xy}}{\sqrt{s_x^2 s_y^2}} \Big/ \sqrt{\frac{s_x^2 s_y^2 - s_{xy}^2}{s_x^2 s_y^2}}$$

can't get the coefficients to work out

$$= \frac{s_{xy}}{\sqrt{s_x^2 s_y^2 - s_{xy}^2}} = \frac{s_{xy}}{\sigma_{12}} \sqrt{\frac{\sigma_1 \sigma_2}{s_x^2 s_y^2 - s_{xy}^2}}$$

We observe that

$$s_{xy} = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum \{ x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x}\bar{y} \}$$

$$= \frac{1}{n} \left[\sum x_i y_i - n \bar{x} \bar{y} \right]$$

$$\sum_{i=1}^{n} X_i Y_i \mid Y_i \sim N\left(\mu_x \bar{Y},\ \sigma_x^2 \sum Y_i^2 \right)$$

$$\boxed{\begin{aligned} E\left[\sum X_i Y_i \mid Y_i\right] &= \sum Y_i E X_i = \mu_x \sum Y_i = \mu_x \bar{Y} \\ Var\left[\sum X_i Y_i \mid Y_i\right] &= \sum Y_i^2 Var[X_i] = \sigma_x^2 \sum Y_i^2 \end{aligned}}$$

$$n\bar{X}\bar{Y} \mid Y \overset{d}{=} n\bar{Y}\, N\left(\mu_x, \frac{\sigma_x^2}{n}\right) \Big| Y \sim N\left(n\mu_x \bar{Y},\ n\sigma_x^2 \bar{Y}^2\right)$$

$$\cancel{\left(\sum X_i Y_i - n\bar{X}\bar{Y}\right) \mid Y \sim N\left(0, \sigma_x^2\left(\sum Y_i^2 + n\bar{Y}^2\right)\right)}$$

$$\cancel{\overset{d}{=} N\left(0, \sigma_x^2 (n-1) s_Y^2\right)}$$

$$\mathrm{Cov}\left[\sum X_i Y_i,\ n\bar{X}\bar{Y} \mid Y\right] = \bar{Y}\, \mathrm{Cov}\left[\sum X_i Y_i,\ \sum X_i \mid Y\right]$$

$$= \bar{Y} \sum \mathrm{Cov}\left[X_i Y_i,\ X_i \mid Y\right] = \bar{Y} \sum Y_i \sigma_x^2 = n\sigma_x^2 \bar{Y}^2$$

Then

$$\mathrm{Var}\left[s_{XY} \mid Y\right] = \frac{1}{n^2} \mathrm{Var}\left[\sum X_i Y_i - n\bar{X}\bar{Y} \mid Y\right]$$

$$= \frac{1}{n^2}\left\{\mathrm{Var}\left[\sum X_i Y_i\right] - 2\,\mathrm{Cov}\left[\sum X_i Y_i,\ n\bar{X}\bar{Y} \mid Y\right] + \mathrm{Var}\left[n\bar{X}\bar{Y} \mid Y\right]\right\}$$

$$= \frac{1}{n^2}\left\{\sigma_x^2 \sum Y_i^2 - 2n\sigma_x^2 \bar{Y}^2 + n\sigma_x^2 \bar{Y}^2\right\} = \frac{1}{n^2}\left\{\sigma_x^2\left(\sum Y_i^2 - n\bar{Y}^2\right)\right\} = \frac{\sigma_x^2 s_Y^2}{n}$$

So that

$$s_{XY} \mid Y \sim N\left(0, \frac{\sigma_x^2 s_Y^2}{n}\right)$$