2009 Theory I #1

1a) Obvious

1b) $\log\left(\frac{p_{11}}{p_{22}}\right) = \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{22}}\right) \iff 1 = \frac{p_{11}p_{22}}{p_{12}p_{21}}$

$p_{11} = p_{1\cdot}p_{\cdot 1} \iff p_{11} = (p_{11}+p_{12})(p_{11}+p_{21}) = p_{11}^2 + p_{11}p_{12} + p_{11}p_{21} + p_{12}p_{21}$

$\iff p_{11}(1-p_{11}) = p_{11}p_{12} + p_{11}p_{21} + p_{12}p_{21}$

$\iff p_{11}(p_{12}+p_{21}+p_{22}) = p_{11}p_{12} + p_{11}p_{21} + p_{12}p_{21}$

$\iff p_{11}p_{22} = p_{12}p_{21} \iff \frac{p_{11}p_{22}}{p_{12}p_{21}} = 1$

similarly obtain same result for p_{12}, p_{21}, p_{22}

1c) $p(x; p) = \exp\left\{\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij}\log p_{ij} + c(x)\right\}$

$= \exp\left\{x_{11}\log\left(\frac{p_{11}}{p_{22}}\right) + x_{12}\log\left(\frac{p_{12}}{p_{22}}\right) + x_{21}\log\left(\frac{p_{21}}{p_{22}}\right) + n\log p_{22} + c(x)\right\}$

$= \exp\left\{x_{11}\left[\log\left(\frac{p_{11}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) - \log\left(\frac{p_{21}}{p_{22}}\right)\right]\right.$

$\left. + \underbrace{(x_{11}+x_{12})}_{= x_{1\cdot}}\log\left(\frac{p_{12}}{p_{22}}\right) + \underbrace{(x_{11}+x_{21})}_{= x_{\cdot 1}}\log\left(\frac{p_{21}}{p_{22}}\right) + n\log p_{22} + c(x)\right\}$

(pg 328-330)

The UMPU test is of the form

$$\phi(\underline{x}) = \begin{cases} 1, & x_{11} < c_1(x_{1\cdot}, x_{\cdot 1}) \text{ or } x_{11} > c_2(x_{1\cdot}, x_{\cdot 1}) \\ \gamma_1^{(x_{1\cdot}, x_{\cdot 1})} & x_{11} = c_1(x_{1\cdot}, x_{\cdot 1}) \\ \gamma_2^{(x_{1\cdot}, x_{\cdot 1})} & x_{11} = c_2(x_{1\cdot}, x_{\cdot 1}) \\ 0 & \text{else} \end{cases}$$

where
$$\begin{cases} E_{\theta_0}[\phi(X_{11}) \mid X_{1\cdot}, X_{\cdot 1}] = \alpha \\ E_{\theta_0}[X_{11}\,\phi(X_{11}) \mid X_{1\cdot}, X_{\cdot 1}] = \alpha E_{\theta_0}[X_{11} \mid X_{1\cdot}, X_{\cdot 1}] \end{cases}$$

Now $X_{11} \mid (X_{1\cdot}, X_{\cdot 1})$ follows a noncentral hypergeometric distribution which under H_0 reduces to a central hypergeometric distribution. (see pages 156-157 of class text). Then the conditions become

$HG(n, X_{1\cdot}, X_{\cdot 1})$

$$\begin{cases} E_{\theta_0}[\phi(X_{11}) \mid X_{1\cdot}, X_{\cdot 1}] = \alpha \\ E_{\theta_0}[X_{11}\,\phi(X_{11}) \mid X_{1\cdot}, X_{\cdot 1}] = \alpha \dfrac{X_{1\cdot} X_{\cdot 1}}{n} \end{cases}$$

A little more detail here?

The power function is given by

$$\beta(\theta) \equiv E_\theta[\phi(X_{11}) \mid X_{1\cdot}, X_{\cdot 1}]$$

where $X_{11} \mid (X_{1\cdot}, X_{\cdot 1}) \sim HG(n, X_{1\cdot}, X_{\cdot 1}, e^{\theta})$

d) The hypothesis is equivalent to $H_0: p_{21} \geq p_{12}$ vs. $H_1: p_{21} < p_{12}$

$$p(x;p) = \exp\left\{x_{11}\log\left(\frac{p_{11}}{p_{22}}\right) + x_{12}\log\left(\frac{p_{12}}{p_{22}}\right) + x_{21}\log\left(\frac{p_{21}}{p_{22}}\right) + n\log p_{22} + c(x)\right\}$$

$$= \exp\left\{x_{21}\log\left(\frac{p_{21}}{p_{12}}\right) + (x_{12}+x_{21})\log\left(\frac{p_{12}}{p_{22}}\right) + x_{11}\log\left(\frac{p_{11}}{p_{22}}\right) + n\log p_{22} + c(x)\right\}$$

why?

at the boundary

Now $X_{21} \mid (X_{12}+X_{21}, X_{11}) \overset{d}{=} X_{21} \mid (X_{12}+X_{21}) \sim \text{bin}(X_{12}+X_{21}, 0.5)$

Then a UMPU test is of the form

$$\phi(x_{21}) = \begin{cases} 1, & x_{21} < c(x_{12}+x_{21}) \\ \gamma(t), & x_{21} = c(x_{12}+x_{21}) \\ 0, & \text{else} \end{cases}$$

$\boxed{\text{A little more detail here?}}$

where $E_{\theta_0}[\phi(X_{21}) \mid X_{12}+X_{21}] = \alpha$

1e i) Let $\tilde{p} = (\tilde{p}_{11}, \tilde{p}_{12}, \tilde{p}_{21}, \tilde{p}_{22})$ be the restricted MLEs.

The unrestricted MLEs are given by $\hat{p} = \left(\frac{x_{11}}{n}, \frac{x_{12}}{n}, \frac{x_{21}}{n}, \frac{x_{22}}{n}\right)$.

Then

$$LRT_n = 2\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij}\log\left(\frac{\hat{p}_{ij}}{\tilde{p}_{ij}}\right) = -2\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij}\log\left(\frac{\tilde{p}_{ij}}{\hat{p}_{ij}}\right)$$

Now a Taylor expansion of $\log(x)$ about a yields

$$\log(x) = \log(a) + \frac{x-a}{a} + \frac{(x-a)^2}{2a} + o\left((x-a)^2\right)$$

since in our case $\frac{\tilde{p}_{ij}}{\hat{p}_{ij}} \xrightarrow{a.s} 1$ by SLLN we choose $a = 1$ H_0?

and obtain

$$\log\left(\frac{\tilde{p}_{ij}}{\hat{p}_{ij}}\right) = \log(1) + \left(\frac{\tilde{p}_{ij}}{\hat{p}_{ij}} - 1\right) - \frac{1}{2}\left(\frac{\tilde{p}_{ij}}{\hat{p}_{ij}} - 1\right)^2 + o_p(1)$$

$$= \underbrace{\frac{\tilde{p}_{ij} - \hat{p}_{ij}}{\hat{p}_{ij}}}_{\in o_p(1)} - \frac{1}{2}\left(\frac{\tilde{p}_{ij} - \hat{p}_{ij}}{\hat{p}_{ij}}\right)^2 + o_p(1)$$

$$= -\frac{1}{2}\left(\frac{\tilde{p}_{ij} - \hat{p}_{ij}}{\hat{p}_{ij}}\right)^2 + o_p(1)$$

since $o_p(1) + o_p(1) = o_p(1)$. Thus

$$LRT_n = -2\sum_{i=1}^{2}\sum_{j=1}^{2} x_{ij} \log\left(\frac{\tilde{p}_{ij}}{\hat{p}_{ij}}\right)$$

$$= -2\sum_{i=1}^{2}\sum_{j=1}^{2} \frac{x_{ij}}{-2}\left(\frac{\tilde{p}_{ij} - \hat{p}_{ij}}{\hat{p}_{ij}}\right)^2 + o_p(1)$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{x_{ij}}{n^2} \frac{(x_{ij} - n\tilde{p}_{ij})^2}{\frac{x_{ij}}{n}\hat{p}_{ij}} + o_p(1)$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(\tilde{p}_{ij} - \hat{p}_{ij})^2}{n\hat{p}_{ij}} + o_p(1)$$

$$= \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(\tilde{p}_{ij} - \hat{p}_{ij})^2}{n\tilde{p}_{ij}} + o_p(1)$$

since $\hat{p} \to \tilde{p}$ a.s. by SLLN so that $\sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(\tilde{p}_{ij} - \hat{p}_{ij})}{n\tilde{p}_{ij}} - \sum_{i=1}^{1}\sum_{j=1}^{1} \frac{(\tilde{p}_{ij} - \hat{p}_{ij})}{n\hat{p}_{ij}} \xrightarrow{a.s} 0.$

and hence in probability, then two applications of Slutsky's thm yields desired result.

1e.ii) We know that under H_0 $LRT_n \xrightarrow{L} \chi^2_r$ where

$r = \underbrace{\partial im(\Theta)}_{\text{3 parameters}} - \underbrace{\partial im(\Theta_0)}_{\text{2 parameters}} = 1$

2 parameters since we need (e.g.) $\pi_{1\cdot}, \pi_{\cdot 1}$ only to obtain $\pi_{1\cdot}, \pi_{2\cdot}, \pi_{\cdot 1}, \pi_{\cdot 2}$ and hence $\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22}$

Under H_1, (see Agresti pg 592)

$$LRT_n \xrightarrow{L} \chi^2_1(\delta)$$

where $\delta = n \sum_{i=1}^{n} \frac{(p_{ni} - f(\theta_0))^2}{f(\theta_0)^2}$

sample values for n observations

f is (false) model

so that

$$\delta = n \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(x_{ij}/n - \pi_{i\cdot}\pi_{\cdot j})^2}{p_{i\cdot}\, p_{\cdot j}}$$