AMW

2. n ⫫ and identical trials

Observe freq. of events $A \cap B$, $A \cap B^c$, $A^c \cap B$, and $A^c \cap B^c$

	A	A^c	Total
B	X_{11}	X_{12}	n_1
B^c	X_{21}	X_{22}	n_2
Total	m_1	m_2	n

a) Let $p_{ij} = E(X_{ij})/n$, $i=1,2$, $j=1,2$, $\sum_{ij} p_{ij} = 1$

Then, $X = (X_{11}, X_{12}, X_{21}, X_{22})$ is multinomial.

$$f(x_{11}, x_{12}, x_{21}, x_{22}) = \frac{n!}{\prod_{ij} x_{ij}!} \prod_{ij} p_{ij}^{x_{ij}}$$

Verify that this distn. is a member of an exponential family & write in canonical form.

$$f(\underset{\sim}{x} | \underset{\sim}{p}) = \frac{n!}{\prod_{ij} x_{ij}!} \prod_{ij} p_{ij}^{x_{ij}}$$

$$= \exp\left\{ \log(n!) - \sum_{ij} \log(x_{ij}!) + x_{ij} \log(p_{ij}) \right\}$$

$$= \exp\{\log(n!)\} \exp\{x_{ij} \log(p_{ij})\} \exp\{-\sum_{ij} \log(x_{ij}!)\}$$

$$= \exp\{\log(n!)\} \exp\{x_{11}\log(p_{11}) + x_{12}\log(p_{12}) + x_{21}\log(p_{21}) + \underbrace{x_{22}}_{n - x_{11} - x_{12} - x_{21}}\log(p_{22})\} \exp\{-\sum_{ij}\log(x_{ij}!)\}$$

$$= \exp\{\log(n!)\} \exp\{x_{11}\log(p_{11}) + x_{12}\log(p_{12}) + x_{21}\log(p_{21}) + (n - x_{11} - x_{12} - x_{21})\log(p_{22})\} \exp\{-\sum_{ij}\log(x_{ij}!)\}$$

$$= \exp\{\log(n!)\} \exp\{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22}) + n\log(p_{22})\} \exp\{-\sum_{ij}\log(x_{ij}!)\}$$

$$= \exp\{\log(n!) + n\log(p_{22})\} \exp\{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22})\} \exp\{-\sum_{ij}\log(x_{ij}!)\}$$

$$= \underbrace{n!\, p_{22}^{\,n}}_{c(\underset{\sim}{p})} \exp\{\underbrace{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22})}_{T(\underset{\sim}{x})Q(\underset{\sim}{p})}\} \underbrace{\exp\{-\sum_{ij}\log(x_{ij}!)\}}_{h(\underset{\sim}{x})}$$

which is a member of an exponential family in the canonical form,

$$f(\underset{\sim}{x} | \underset{\sim}{\theta}) = c(\underset{\sim}{p}) \exp\{T(\underset{\sim}{x}) Q(\underset{\sim}{p})\} h(\underset{\sim}{x})$$

AMW

2b. Show that $A \perp\!\!\!\perp B$ iff $\log\left(\frac{p_{11}}{p_{22}}\right) = \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{22}}\right)$.

$\ulcorner$ $A \perp\!\!\!\perp B \Leftrightarrow P(A \cap B) = P(A)P(B) \Leftrightarrow p_{11} = (p_{11} + p_{21})(p_{11} + p_{12})$

$\Leftrightarrow p_{11} = p_{11}^2 + p_{11}p_{12} + p_{11}p_{21} + p_{12}p_{21}$

$\Leftrightarrow p_{11} - p_{11}^2 - p_{11}p_{12} - p_{11}p_{21} - p_{12}p_{21} = 0$

$\Leftrightarrow p_{11}\underbrace{(1 - p_{11} - p_{12} - p_{21})}_{p_{22}} - p_{12}p_{21} = 0$

$\Leftrightarrow p_{11}p_{22} - p_{12}p_{21} = 0 \Leftrightarrow p_{11}p_{22} = p_{12}p_{21} \Leftrightarrow \frac{p_{11}\cancel{p_{22}}}{p_{22}^{\cancel{2}}} = \frac{p_{12}p_{21}}{p_{22}^2}$

$\Leftrightarrow \frac{p_{11}}{p_{22}} = \frac{p_{12}}{p_{22}}\frac{p_{21}}{p_{22}} \Leftrightarrow \log\left(\frac{p_{11}}{p_{22}}\right) = \log\left(\frac{p_{12}}{p_{22}}\frac{p_{21}}{p_{22}}\right)$ ✓

$\Leftrightarrow \log\left(\frac{p_{11}}{p_{22}}\right) = \log\left(\frac{p_{12}}{p_{22}}\right) + \log\left(\frac{p_{21}}{p_{22}}\right)$ $\lrcorner$

2c) Let $\theta = a_0\log\left(\frac{p_{11}}{p_{22}}\right) + a_1\log\left(\frac{p_{12}}{p_{22}}\right) + a_2\log\left(\frac{p_{21}}{p_{22}}\right)$ where (a_0, a_1, a_2) are constants.

Let $a_0 = 1$ and $a_1 = a_2 = -1$ derive a UMPU size α test

for testing $H_0: \theta = 0$ vs. $H_1: \theta \neq 0$, (ii)

(i) Also, derive power function.

(Hint: use a theorem for multiparameter exp families to construct the UMPU test).

$\ulcorner$ ① Write in exponential family form.

Told to let $\theta = a_0\log\left(\frac{p_{11}}{p_{22}}\right) + a_1\log\left(\frac{p_{12}}{p_{22}}\right) + a_2\log\left(\frac{p_{21}}{p_{22}}\right)$

Want to test $H_0: a_0\log\left(\frac{p_{11}}{p_{22}}\right) + a_1\log\left(\frac{p_{12}}{p_{22}}\right) + a_2\log\left(\frac{p_{21}}{p_{22}}\right) = 0$

vs. $H_1: a_0\log\left(\frac{p_{11}}{p_{22}}\right) + a_1\log\left(\frac{p_{12}}{p_{22}}\right) + a_2\log\left(\frac{p_{21}}{p_{22}}\right) \neq 0$

For $(a_0, a_1, a_2) = (1, -1, -1)$, these hypotheses become:

$H_0: \log\left(\frac{p_{11}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) - \log\left(\frac{p_{21}}{p_{22}}\right) = 0$

vs. $H_1: \log\left(\frac{p_{11}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) - \log\left(\frac{p_{21}}{p_{22}}\right) \neq 0$

Cont'd. →

2 c. cont'd

From part a), $f(\underset{\sim}{x} \mid \underset{\sim}{p}) = n! \, p_{22}^{n} \exp\{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22})\}$

$\cdot \exp\{-\sum_{ij}\log(x_{ij}!)\}$

$$= n! \, p_{22}^{n} \exp\Big\{ x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22})$$
$$+ x_{11}\Big(\log\Big(\frac{p_{11}}{p_{22}}\Big) - \log\Big(\frac{p_{12}}{p_{22}}\Big) - \log\Big(\frac{p_{21}}{p_{22}}\Big)\Big)$$
$$- x_{11}\Big(\log\Big(\frac{p_{11}}{p_{22}}\Big) - \log\Big(\frac{p_{12}}{p_{22}}\Big) - \log\Big(\frac{p_{21}}{p_{22}}\Big)\Big)\Big\} \exp\{-\textstyle\sum_{ij}\log(x_{ij}!)\}$$

$$= n! \, p_{22}^{n} \exp\Big\{(x_{11}+x_{12})\log(p_{12}/p_{22}) + (x_{11}+x_{21})\log(p_{21}/p_{22})$$
$$+ x_{11}\big(\log(p_{11}/p_{22}) - \log(p_{12}/p_{22}) - \log(p_{21}/p_{22})\big)\Big\} \exp\{-\textstyle\sum_{ij}\log(x_{ij}!)\}$$

where $\theta = \log\Big(\frac{p_{11}}{p_{22}}\Big) - \log\Big(\frac{p_{12}}{p_{22}}\Big) - \log\Big(\frac{p_{21}}{p_{22}}\Big)$

$U = x_{11}$

$\xi_1 = \log(p_{12}/p_{22})$, $\xi_2 = \log(p_{21}/p_{22})$

$T_1 = X_{11} + X_{12}$, $T_2 = X_{11} + X_{21}$

(2) Now, write the form of the UMPU level-α test.

$$\phi(x_{11}) = \begin{cases} 1 & \text{if } U < c_1(t) \text{ or } U > c_2(t) \\ \gamma_i & \text{if } U = c_i(t) \\ 0 & \text{else} \end{cases} = \begin{cases} 1 & \text{if } X_{11} < c_1(t) \text{ or } X_{11} > c_2(t) \\ \gamma_1 & \text{if } X_{11} = c_1(t) \\ \gamma_2 & \text{if } X_{11} = c_2(t) \\ 0 & \text{else} \end{cases}$$

where $E_{\theta_0}[\phi(u) \mid T = t] = \alpha$

& $E_{\theta_0}[U\phi(u) \mid T = t] = \alpha E_{\theta_0}[U \mid T = t]$

(3) Need to find distribution of $U \mid T = t \Rightarrow X_{11} = x \mid (X_{11}+X_{12}) = n_1, (X_{11}+X_{21}) = m_1)$

First, need to show that $X_{11} \perp\!\!\!\perp X_{12}$ given $X_{11} + X_{21} = m_1$.

Take $P(X_{11} = x, X_{12} = y \mid X_{11} + X_{21} = m_1) = \dfrac{P(X_{11} = x, X_{12} = y, X_{11} + X_{21} = m_1)}{P(X_{11} + X_{21} = m_1)}$

$$= \frac{P(X_{11} = x, X_{12} = y, X_{21} = m_1 - x, X_{22} = n - m_1 - y)}{P(X_{11} + X_{21} = m_1)} = \frac{\dfrac{\cancel{n!}}{x!\,y!\,(m_1-x)!\,(n-m_1-y)!} p_{11}^{x} p_{12}^{y} p_{21}^{m_1-x} p_{22}^{n-m_1-y}}{\dfrac{\cancel{n!}}{m_1!\,(n-m_1)!}(p_{11}+p_{21})^{m_1} \underbrace{(1-p_{11}-p_{21})}_{p_{12}+p_{22}}{}^{n-m_1}}$$

$m_2 - y$

$m_2 - y$

m_2

$(n-m_1)! = m_2$

cont'd next pg.

AMW

2. c. cont'd.

$$\Rightarrow P(X_{11}=x, X_{12}=y \mid X_{11}+X_{21}=m_1) = \frac{m_1!}{x!(m_1-x)!}\cdot\frac{m_2!}{y!(m_2-y)!}\left(\frac{p_{11}}{p_{11}+p_{21}}\right)^{m_1}\left(\frac{p_{12}}{p_{12}+p_{22}}\right)^{m_2}$$

$$\cdot\left(\frac{p_{21}}{p_{11}+p_{21}}\right)^{m_1-x}\left(\frac{p_{22}}{p_{12}+p_{22}}\right)^{m_2-y}$$

$$= \underbrace{\binom{m_1}{x}\left(\frac{p_{11}}{p_{11}+p_{21}}\right)^{m_1}\left(1-\frac{p_{11}}{p_{11}+p_{21}}\right)^{m_1-x}}_{P(X_{11}=x \mid X_{11}+X_{21}=m_1)} \underbrace{\binom{m_2}{y}\left(\frac{p_{12}}{p_{12}+p_{22}}\right)^{m_2}\left(1-\frac{p_{12}}{p_{12}+p_{22}}\right)^{m_2-y}}_{P(X_{12}=y \mid X_{11}+X_{21}=m_1)}$$

$\Rightarrow$ $X_{11} \perp\!\!\!\perp X_{12}$ if conditioning on $X_{11}+X_{21}=m_1$

Then, $P(X_{11}=x \mid X_{11}+X_{12}=n_1, X_{11}+X_{21}=m_1)$

$$= \frac{P(X_{11}=x, X_{11}+X_{12}=n_1 \mid X_{11}+X_{21}=m_1)}{P(X_{11}+X_{12}=n_1 \mid X_{11}+X_{21}=m_1)}$$

$$= \frac{P(X_{11}=x, X_{12}=n_1-x \mid X_{11}+X_{21}=m_1)}{\sum_{z=0}^{\min(n_1,m_1)} P(X_{11}=z, X_{12}=n_1-z \mid X_{11}+X_{21}=m_1)} = \frac{P(X_{11}=x \mid X_{11}+X_{21}=m_1)\, P(X_{12}=n_1-x \mid X_{11}+X_{21}=m_1)}{\sum_{z=0}^{\min(n_1,m_1)} P(X_{11}=z \mid X_{11}+X_{21}=m_1)\, P(X_{12}=n_1-z \mid X_{11}+X_{21}=m_1)}$$

$$= \frac{\binom{m_1}{x}\Big(\overbrace{\frac{p_{11}}{p_{11}+p_{21}}}^{p_1}\Big)^{x}\Big(1-\overbrace{\frac{p_{11}}{p_{11}+p_{21}}}^{p_1}\Big)^{\cancel{m_1}-x}\binom{m_2}{n_1-x}\Big(\overbrace{\frac{p_{12}}{p_{12}+p_{22}}}^{p_2}\Big)^{\cancel{n_1}-x}\Big(1-\overbrace{\frac{p_{12}}{p_{12}+p_{22}}}^{p_2}\Big)^{\cancel{m_2}-\cancel{n_1}+x}}{\sum_{z=0}^{\min(n_1,m_1)}\binom{m_1}{z}\Big(\underbrace{\frac{p_{11}}{p_{11}+p_{21}}}_{p_1}\Big)^{z}\Big(1-\underbrace{\frac{p_{11}}{p_{11}+p_{21}}}_{p_1}\Big)^{\cancel{m_1}-z}\binom{m_2}{n_1-z}\Big(\underbrace{\frac{p_{12}}{p_{12}+p_{22}}}_{p_2}\Big)^{\cancel{n_1}-z}\Big(1-\underbrace{\frac{p_{12}}{p_{12}+p_{22}}}_{p_2}\Big)^{\cancel{m_2}-\cancel{n_1}+z}}$$

$$= \frac{\binom{m_1}{x}\binom{m_2}{n_1-x} p_1^{x}(1-p_1)^{-x} p_2^{-x}(1-p_2)^{x}}{\sum_{z=0}^{\min(n_1,m_1)}\binom{m_1}{z}\binom{m_2}{n_1-z} p_1^{z}(1-p_1)^{-z} p_2^{-z}(1-p_2)^{z}} = \frac{\binom{m_1}{x}\binom{m_2}{n_1-x}\left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right)^{x}}{\sum_{z=0}^{\min(n_1,m_1)}\binom{m_1}{z}\binom{m_2}{n_1-z}\left(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\right)^{z}}$$

$$= \frac{\binom{m_1}{x}\binom{m_2}{n_1-x}\omega^{x}}{P_0} \quad \text{where } \omega = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} \text{ and } P_0 = \sum_{z=0}^{\min(n_1,m_1)}\binom{m_1}{z}\binom{m_2}{n_1-z}\omega^{z}$$

$\Rightarrow$ $X_{11}=x \mid X_{11}+X_{12}=n_1, X_{11}+X_{21}=m_1 \sim$ Non-Central HyperGeom(m_1, m_2, n_1, ω)

2.c. cont'd AMW

Then, under $H_0: p_1 = p_2 \Rightarrow \quad \omega = \dfrac{p_1/(1-p_1)}{p_2/(1-p_2)} = 1$

$\Rightarrow$ Under H_0: $X_{11} = x \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1 \sim \underline{HyperGeom(m_1, m_2, n_1)}$

(4) Thus, the UMPU α-level test is,

$$\phi(x_{11}) = \begin{cases} 1 & \text{if } x_{11} < c_1(t) \text{ or } x_{11} > c_2(t) \\ \gamma_1 & \text{if } x_{11} = c_1(t) \\ \gamma_2 & \text{if } x_{11} = c_2(t) \\ 0 & \text{else} \end{cases}$$

where $\alpha = E_{\Theta_0}[\phi(u) \mid T = t] = 1 \cdot P_0(x_{11} < c_1(t) \text{ or } x_{11} > c_2(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$
$+ \gamma_1 P_0(x_{11} = c_1(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$
$+ \gamma_2 P_0(x_{11} = c_2(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$ (1)

where P_0 is the probability w.r.t. Hypergeom(m_1, m_2, n_1)

$\&$ $\alpha E_{\Theta_0}[U \mid T = t] = E_{\Theta_0}[U \phi(u) \mid T = t]$

$$\alpha \left(\frac{n_1 m_1}{n}\right) = E_{\Theta_0}[X_{11} \phi(x_{11}) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1] \quad (2)$$

where E_0 is the expectation w.r.t. Hypergeom(m_1, m_2, n_1)

Thus, we would find $c_1(t)$ and $c_2(t)$ from the above eqns. (1) and (2).

Finally, the power function is,

$\beta = E_{\Theta_1}[\phi(u) \mid T = t] = E_{\Theta_1}[\phi(x_{11}) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1]$
$= P(x_{11} < c_1(t) \text{ or } x_{11} > c_2(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$
$+ \gamma_1 P(x_{11} = c_1(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$
$+ \gamma_2 P(x_{11} = c_2(t) \mid X_{11} + X_{12} = n_1, X_{11} + X_{21} = m_1)$

where P is the probability w.r.t. Noncentral HyperGeom(m_1, m_2, n_1, ω)

where $\underbrace{\omega}_{\text{non-centrality param.}} = \dfrac{p_1/(1-p_1)}{p_2/(1-p_2)}$

2.d) Derive a UMPU size α test for testing,

$H_0: P(A) \geq P(B)$

vs. $H_1: P(A) < P(B)$

(Hint: Use the techniques of part c) in setting up the hypotheses in terms of θ and then constructing the test.)

Γ $H_0: P(A) \geq P(B) \iff p_{11} + p_{21} \geq p_{11} + p_{12}$

$$\iff \frac{p_{21}}{p_{22}} \geq \frac{p_{12}}{p_{22}} \iff \underbrace{\log\left(\frac{p_{21}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right)}_{\theta} \geq 0$$

Test: $H_0: \theta \geq 0$

vs. $H_1: \theta < 0$

From a), $f(\underset{\sim}{x}|\underset{\sim}{p}) = n!\, p_{22}^n \exp\{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22})\}$

$\cdot \exp\{-\sum_{ij}\log(x_{ij}!)\}$

$= n!\, p_{22}^n \exp\{x_{11}\log(p_{11}/p_{22}) + x_{12}\log(p_{12}/p_{22}) + x_{21}\log(p_{21}/p_{22}) \underline{- x_{21}\log(p_{12}/p_{22})}$

$\underline{+ x_{21}\log(p_{12}/p_{22})}\} \exp\{-\sum_{ij}\log(x_{ij}!)\}$

$= n!\, p_{22}^n \exp\left\{x_{11}\log(p_{11}/p_{22}) + (x_{12}+x_{21})\log(p_{12}/p_{22}) + x_{21}\left[\log(p_{21}/p_{22}) - \log\left(\frac{p_{12}}{p_{22}}\right)\right]\right\}$

$\cdot \exp\{-\sum_{ij}\log(x_{ij}!)\}$

where $\theta = \log(p_{21}/p_{22}) - \log(p_{12}/p_{22})$

$U = X_{21}$

$\xi_1 = \log(p_{11}/p_{22})$, $\xi_2 = \log(p_{12}/p_{22})$

$T_1 = X_{11}$, $T_2 = X_{12} + X_{21}$

contd. $\longrightarrow$

AMW

2. d. cont'd

The UMPU α-level test associated w/ $H_0: \theta \geq 0$ vs. $H_1: \theta < 0$, where

$\theta = \log\left(\frac{p_{21}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) \geq 0$, is of the form,

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

where $\alpha = E_0[\phi(x_{21}) \mid x_{11} = x, x_{12} + x_{21} = y]$

Need to find distribution of $x_{21} = z \mid x_{11} = x, x_{12} + x_{21} = y$

$$P(X_{21} = z \mid X_{11} = x, X_{12} + X_{21} = y) = \frac{P(X_{11} = x, X_{12} = y - z, X_{21} = z, X_{22} = n-x-y)}{P(X_{11} = x, X_{12} + X_{21} = y, X_{22} = n-x-y)}$$

$$= \frac{\frac{n!}{x!(y-z)!\,z!(n-x-y)!} p_{11}^{x} p_{12}^{y-z} p_{21}^{z} p_{22}^{n-x-y}}{\sum_{w=0}^{y} P(X_{11} = x, X_{12} = y-w, X_{21} = w, X_{22} = n-x-y)}$$

$$= \frac{\frac{\cancel{n!}}{\cancel{x!}(y-z)!\,z!\,\cancel{(n-x-y)!}} \cancel{p_{11}^{x}} p_{12}^{y-z} p_{21}^{z} \cancel{p_{22}^{n-x-y}}}{\sum_{w=0}^{y} \frac{\cancel{n!}}{\cancel{x!}(y-w)!\,w!\,\cancel{(n-x-y)!}} \cancel{p_{11}^{x}} p_{12}^{y-w} p_{21}^{w} \cancel{p_{22}^{n-x-y}}} = \frac{\frac{y!}{z!(y-z)!} p_{21}^{z} p_{12}^{y-z}}{\sum_{w=0}^{y} \frac{y!}{w!(y-w)!} p_{21}^{w} p_{12}^{y-w}}$$

$$= \frac{\frac{y!}{z!(y-z)!} p_{21}^{z} p_{12}^{y-z}}{(p_{21} + p_{12})^{y}} = \frac{y!}{z!(y-z)!} \left(\frac{p_{21}}{p_{21} + p_{12}}\right)^{z} \left(\frac{p_{12}}{p_{21} + p_{12}}\right)^{y-z}$$

$\Rightarrow X_{21} = z \mid X_{11} = x, X_{12} + X_{21} = y \sim \text{Binom}\left(y, p = \frac{p_{21}}{p_{21} + p_{12}}\right)$

Thus, the UMPU α-level test is,

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

where $\alpha = E_0[\phi(x_{21}) \mid X_{11} = x, X_{12} + X_{21} = y]$
$= 1 \cdot P_0(X_{21} < c(x,y) \mid X_{11} = x, X_{12} + X_{21} = y)$
$+ \gamma \cdot P_0(X_{21} = c(x,y) \mid X_{11} = x, X_{12} + X_{21} = y)$

where P_0 is the probability w.r.t. $\text{Binom}\left(y, p = \frac{p_{21}}{p_{21} + p_{12}}\right)$
under the constraint that $\theta = \log\left(\frac{p_{21}}{p_{22}}\right) - \log\left(\frac{p_{12}}{p_{22}}\right) \geq 0$

$\Rightarrow P_0$ is prob. w.r.t. $\text{Binom}(y, 1/2)$

AMW

2e) Derive the LRT statistic, denoted by Λ_n for the hypothesis in part c) and show that it is asymptotically equivalent to the Pearson chi-square statistic.

Specifically,

(i) Show that $-2\log(\Lambda_n) = \sum_{j=1}^{2}\sum_{i=1}^{2} \frac{(X_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}} + o_p(1)$

where $\hat{p}_{ij}$ denotes the MLE of p_{ij} under H_0.

① First, find the unrestricted MLE for p_{ij} (call it π_{ij} at end to distinguish from under the null)

$$\mathcal{L}(p_{ij} \mid x_{ij}) = \frac{n!}{X_{11}!\,X_{12}!\,X_{21}!\,X_{22}!} p_{11}^{X_{11}} p_{12}^{X_{12}} p_{21}^{X_{21}} (1-p_{11}-p_{12}-p_{21})^{X_{22}}$$

$$\Rightarrow \ell(p_{ij} \mid x_{ij}) \propto X_{11}\log(p_{11}) + X_{12}\log(p_{12}) + X_{21}\log(p_{21})$$

$$\Rightarrow \frac{\partial \ell}{\partial p_{ij}} = \frac{X_{ij}}{p_{ij}} - \frac{X_{22}}{1-p_{11}-p_{12}-p_{21}} \overset{set}{=} 0 \Rightarrow \frac{X_{ij}}{p_{ij}} = \frac{X_{22}}{1-p_{11}-p_{12}-p_{21}}$$

$$\Rightarrow \hat{\pi}_{ij} = \frac{X_{ij}}{\frac{X_{22}}{\frac{1-p_{11}-p_{12}-p_{21}}{p_{22}}}} = \frac{X_{ij}}{n}$$

b/c $p_{22} = X_{22}/n \Rightarrow n = \frac{X_{22}}{p_{22}}$

② Call the restricted MLE under H_0 as $\hat{p}_{ij}$

③ Write the LRT statistic using $\hat{\pi}_{ij}$ and $\hat{p}_{ij}$

$$\Lambda_n = \frac{\sup_{\theta \in \Theta_0} p(x \mid \theta)}{\sup_{\theta \in \Theta_0 \cup \Theta_1} p(x \mid \theta)} = \frac{p(x \mid \tilde{\theta})}{p(x \mid \hat{\theta})} = \frac{f(\underset{\sim}{x} \mid \underset{\sim}{\hat{p}})}{f(\underset{\sim}{x} \mid \underset{\sim}{\hat{\pi}})}$$

$$= \frac{\cancel{\frac{n!}{X_{11}!\,X_{12}!\,X_{21}!\,X_{22}!}}\, \hat{p}_{11}^{X_{11}} \hat{p}_{12}^{X_{12}} \hat{p}_{21}^{X_{21}} \hat{p}_{22}^{X_{22}}}{\cancel{\frac{n!}{X_{11}!\,X_{12}!\,X_{21}!\,X_{22}!}}\, \hat{\pi}_{11}^{X_{11}} \hat{\pi}_{12}^{X_{12}} \hat{\pi}_{21}^{X_{21}} \hat{\pi}_{22}^{X_{22}}} = \frac{\hat{p}_{11}^{X_{11}} \hat{p}_{12}^{X_{12}} \hat{p}_{21}^{X_{21}} \hat{p}_{22}^{X_{22}}}{\hat{\pi}_{11}^{X_{11}} \hat{\pi}_{12}^{X_{12}} \hat{\pi}_{21}^{X_{21}} \hat{\pi}_{22}^{X_{22}}}$$

$$= \frac{\hat{p}_{11}^{X_{11}} \hat{p}_{12}^{X_{12}} \hat{p}_{21}^{X_{21}} \hat{p}_{22}^{X_{22}}}{\left(\frac{X_{11}}{n}\right)^{X_{11}} \left(\frac{X_{12}}{n}\right)^{X_{12}} \left(\frac{X_{21}}{n}\right)^{X_{21}} \left(\frac{X_{22}}{n}\right)^{X_{22}}} = \prod_{j=1}^{2}\prod_{i=1}^{2} \frac{\hat{p}_{ij}^{X_{ij}}}{\left(\frac{X_{ij}}{n}\right)^{X_{ij}}}$$

cont'd next pg. $\longrightarrow$

2c) cont'd.

Then, $-2\log(\Lambda_n) = -2\left\{ \sum_{j=1}^{2} \sum_{i=1}^{2} X_{ij} \log\left(\frac{\hat{p}_{ij}}{\left(\frac{X_{ij}}{n}\right)} \right) \right\}$

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

Taylor series expansion
$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \dots$

$$= -2\left\{ \sum_{j=1}^{2} \sum_{i=1}^{2} X_{ij} \left[-\left(1 - \frac{n\hat{p}_{ij}}{X_{ij}}\right) - \frac{1}{2}\left(1 - \frac{n\hat{p}_{ij}}{X_{ij}}\right)^2 - o_p(1) \right] \right\}$$

$$= \cancel{2 \sum_{j=1}^{2} \sum_{i=1}^{2} \cancel{X_{ij}} \left(\frac{X_{ij} - n\hat{p}_{ij}}{\cancel{X_{ij}}} \right)}^{0} + \sum_{j=1}^{2} \sum_{i=1}^{2} \overset{X_{ij}}{\overset{\shortparallel}{X_{ij}}} \left(\frac{X_{ij} - n\hat{p}_{ij}}{X_{ij}} \right)^2 + o_p(1)$$

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

(ii) Find the asymptotic distribution of $-2\log(\Lambda_n)$ under H_0 and H_1.

As by Theorem 2.7, assuming the "usual" regularity conditions hold (which, as Dr. Li noted, is equivalent to showing $f(\underline{x}|\underline{p})$ is a member of an exponential family, which we showed first in part a),

then $-2\log(\Lambda_n) = \sum_{j=1}^{2} \sum_{i=1}^{2} \frac{(X_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}} + o_p(1) \xrightarrow{H_0} \chi^2_1$ as $n \to \infty$

and $-2\log(\Lambda_n) \xrightarrow{H_1} \chi^2_1(\delta)$

↑

a non-central chi-squared distribution w/ non-centrality parameter

Know from 762 notes, practice 9.pdf that

$$\delta = \sum_{j=1}^{2} \sum_{i=1}^{2} \frac{(np_{ij} - np_{i+}p_{+j})^2}{np_{ij}}$$