2012 Qualifying Exam Section 2

February 21, 2019

Question 1

1.a

Note that

$$P(Y_{i1}=1,Y_{i2}=1)=P(Y_{i2}=1|Y_{i1}=1)P(Y_{i1}=1)=\beta\alpha=\alpha\beta$$

$$P(Y_{i1}=1,Y_{i2}=0)=P(Y_{i2}=0|Y_{i1}=1)P(Y_{i1}=1)=(1-\beta)\alpha=\alpha(1-\beta)$$

$$P(Y_{i1}=0,Y_{i2}=1)=0$$
$$P(Y_{i1}=0,Y_{i2}=0)=P(Y_{i1}=0)=1-\alpha$$

The last equality follows because someone can only get the secondary illness if they've received a primary illness. We put these probabilities in a table below.

		Y_{i2}		
		0	1	
Y_{i1}	0	$1-\alpha$	0	$1-\alpha$
	1	$\alpha(1-\beta)$	$\alpha\beta$	α
		$1-\alpha\beta$	$\alpha\beta$	1

Now, the likelihood function can be written as

$$\begin{aligned}
L(\alpha,\beta)&=(\alpha\beta)^{x_1}[\alpha(1-\beta)]^{x_2}(1-\alpha)^{x_3}\\
&=\exp\{x_1\log(\alpha\beta)+x_2\log(\alpha[1-\beta])+x_3\log(1-\alpha)\}\\
&=\exp\{x_1\log(\alpha\beta)+x_2\log(\alpha[1-\beta])+(n-x_1-x_2)\log(1-\alpha)\}\\
&=\exp\left\{x_1\log\frac{\alpha\beta}{1-\alpha}+x_2\log\frac{\alpha(1-\beta)}{1-\alpha}+n\log(1-\alpha)\right\}\\
&=\exp\{\mathbf{T}(\mathbf{x})^T\boldsymbol{\theta}-b(\boldsymbol{\theta})\}
\end{aligned}$$

where $\boldsymbol{T}(\boldsymbol{x}) = (x_1, x_2)^T$, $\boldsymbol{\theta} = (\theta_1, \theta_2)^T, \theta_1 = \log \frac{\alpha\beta}{1-\alpha}, \theta_2 = \log \frac{\alpha(1-\beta)}{1-\alpha}$, and $b(\boldsymbol{\theta}) = n \log(1-\alpha)$.
We recognize $L(\alpha, \beta)$ as a multiparameter exponential family.

1.b.

The maximum likelihood estimators of α and β are found by setting the derivative of the log likelihood to 0. We have

$$\ell(\alpha, \beta) = (x_1 + x_2)\log \alpha + x_3 \log(1-\alpha) + x_1 \log \beta + x_2 \log(1-\beta)$$

Thus,

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} = \frac{x_1 + x_2}{\alpha} - \frac{x_3}{1-\alpha} \overset{\text{SET}}{=} 0 &\implies (1-\hat{\alpha})(x_1 + x_2) - \hat{\alpha} x_3 = 0 \\
&\implies x_1 + x_2 - \hat{\alpha}(x_1 + x_2 + x_3) = 0 \\
&\implies \hat{\alpha} = \frac{x_1 + x_2}{x_2 + x_2 + x_3} = \frac{x_1 + x_2}{n}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta} = \frac{x_1}{\beta} - \frac{x_2}{1-\beta} \overset{\text{SET}}{=} 0 &\implies (1-\hat{\beta})x_1 - \hat{\beta} x_2 = 0 \\
&\implies x_1 - \hat{\beta}(x_1 + x_2) = 0 \\
&\implies \hat{\beta} = \frac{x_1}{x_1 + x_2}
\end{aligned}$$

We have

$$\frac{\partial^2 \ell}{\partial \alpha^2} = -\frac{x_1 + x_2}{\alpha^2} - \frac{x_3}{(1-\alpha)^2} < 0$$

$$\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{x_1}{\beta^2} - \frac{x_2}{(1-\beta)^2} < 0$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \beta} = 0$$

So the negative Hessian is positive definite. Hence, $(\hat{\alpha}, \hat{\beta})$ is a maximum.

1.c.

The Fisher information matrix is the negative of the expectation of the Hessian matrix. Note that

$$\begin{aligned}
\mathbb{E}X_1 &= nP(Y_{i1}=1, Y_{i2}=1) = n\alpha\beta \\
\mathbb{E}X_2 &= nP(Y_{i1}=1, Y_{i2}=0) = n\alpha(1-\beta) \\
\mathbb{E}X_3 &= nP(Y_{i1}=0, Y_{i2}=0) = n(1-\alpha)
\end{aligned}$$

Hence,

$$\begin{aligned}
I_{\alpha\alpha} := -\mathbb{E}\frac{\partial^2 \ell}{\partial \alpha^2} &= \mathbb{E}\left\{\frac{X_1+X_2}{\alpha^2} + \frac{X_3}{(1-\alpha)^2}\right\} = \frac{n\alpha\beta + n\alpha(1-\beta)}{\alpha^2} + \frac{n(1-\alpha)}{(1-\alpha)^2} \\
&= \frac{n}{\alpha} + \frac{n}{1-\alpha} = \frac{n}{\alpha(1-\alpha)}
\end{aligned}$$

$$I_{\beta\beta} := -\mathbb{E}\frac{\partial^2 \ell}{\partial \beta^2} = \mathbb{E}\frac{X_1}{\beta^2} + \mathbb{E}\frac{X_2}{(1-\beta)^2} = \frac{n\alpha\beta}{\beta^2} + \frac{n\alpha(1-\beta)}{(1-\beta)^2} = \frac{n\alpha}{\beta} + \frac{n\alpha}{1-\beta} = \frac{n\alpha}{\beta(1-\beta)}$$

$$I_{\alpha\beta} = 0$$

The Fisher information matrix is

$$I(\alpha,\beta) = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} \\ I_{\beta\alpha} & I_{\beta\beta} \end{pmatrix} = \begin{pmatrix} \frac{n}{1-\alpha} & 0 \\ 0 & \frac{n\alpha}{\beta(1-\beta)} \end{pmatrix}$$

The asymptotic covariance matrix, which we will denote A-Cov, is the inverse of the fisher information matrix. Since the samples are i.i.d., we ignore the n in the Fisher information matrix.

$$\text{A-Cov}(\alpha,\beta) = [I(\alpha,\beta)]^{-1} = \begin{pmatrix} \alpha(1-\alpha) & 0 \\ 0 & \frac{\beta(1-\beta)}{\alpha} \end{pmatrix}$$

Note, this matrix is approximately the covariance matrix of $\sqrt{n}(\hat{\alpha},\hat{\beta})^T$. Putting the n's back in will yield the approximation for the covariance matrix of $(\hat{\alpha},\hat{\beta})^T$.

1.d

Suppose a UMP test ϕ exists for $H_0 : \beta = 0.5$ vs. $H_1 : \beta > 0.5$ Let $\beta_1 > \beta_2 > 0.5$. Consider testing

$$H_0^i : \beta = 0.5 \text{ vs. } H_1^i : \beta = \beta_i, \quad i = 1, 2$$

By the Neyman-Pearson Lemma, a UMP test exists for testing H_0^i vs. H_1^i, say, ϕ_i where ϕ_i takes the form

$$\phi_i(x) = \begin{cases} 1, & \text{if } \frac{p(x|\alpha,\beta_i)}{p(x|\alpha,0.5)} > k_i \\ \gamma_i, & \text{if } \frac{p(x|\alpha,\beta_i)}{p(x|\alpha,0.5)} = k_i \\ 0, & \text{if } \frac{p(x|\alpha,\beta_i)}{p(x|\alpha,0.5)} < k_i \end{cases}$$

where k_i and γ_i are chosen so that $E_{0.5}(\phi_i(x)) = \alpha$, $i = 1, 2$. Note that

$$\begin{aligned} \frac{p(x|\alpha,\beta_i)}{p(x|\alpha,0.5)} &= \frac{\alpha^{x_1+x_2}\beta_i^{x_1}(1-\beta_i)^{x_2}(1-\alpha)^{x_3}}{\alpha^{x_1+x_2}0.5^{x_1}0.5^{x_2}(1-\alpha)^{x_3}} \\ &= 2^{x_1+x_2}\beta_i^{x_1}(1-\beta_i)^{x_2} \end{aligned}$$

The rejection region of ϕ_i depends on β_i, and thus k_i is different for each i. Hence, ϕ_1 and ϕ_2 are UMP tests with different rejection regions. Thus, ϕ cannot be UMP for testing H_0 vs. H_1 since it has a different rejection region than the UMP test that tests H_0^i vs. H_1^i, $i = 1, 2$ where $\beta_1 > \beta_2 > 0.5$ were chosen arbitrarily.

1.e.

Under H_0, $\alpha - \beta = 0$, so $\alpha = \beta = \theta$ for some θ. Thus, the restricted likelihood is

$$L(\theta, \theta) = \theta^{x_1+x_2}\theta^{x_1}(1-\theta)^{x_2}(1-\theta)^{x_3} = \theta^{2x_1+x_2}(1-\theta)^{x_2+x_3}$$

Hence, the restricted log-likelihood is

$$\ell_0 := \ell(\theta, \theta) = (2x_1 + x_2)\log(\theta) + (x_2 + x_3)\log(1-\theta)$$

We have

$$\begin{aligned}
\frac{\mathrm{d}\ell_0}{\mathrm{d}\theta} &= \frac{2x_1 + x_2}{\theta} - \frac{x_2 + x_3}{1-\theta} \overset{\text{SET}}{=} 0 \\
&\implies (2x_1 + x_2) - \theta(2x_1 + x_2) - \theta(x_2 + x_3) = 0 \\
&\implies 2x_1 + x_2 = \theta(2x_1 + 2x_2 + x_3) \\
&\implies \hat{\theta} = \frac{2x_1 + x_2}{2x_1 + 2x_2 + x_3} = \frac{2x_1 + x_2}{n + x_1 + x_2}
\end{aligned}$$

where $\hat{\theta}$ is the MLE under H_0.

From here, simply take the ratio of the likelihood plugging in $\hat{\theta}$ in the restricted likelihood and plugging in $(\hat{\alpha}, \hat{\beta})$ in the full likelihood. It does not simplify well.

1.f.

Note that

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha} &= \frac{x_1+x_2}{\alpha} - \frac{x_3}{1-\alpha} = \frac{(x_1+x_2)-\alpha(x_1+x_2)-\alpha x_3}{\alpha(1-\alpha)} \\ &= \frac{(x_1+x_2)-\alpha(x_1+x_2+x_3)}{\alpha(1-\alpha)} \\ &= \frac{x_1+x_2-n\alpha}{\alpha(1-\alpha)}\end{aligned}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{x_1}{\beta} - \frac{x_2}{1-\beta} = \frac{(1-\beta)x_1-\beta x_2}{\beta(1-\beta)} = \frac{x_1-(x_1+x_2)\beta}{\beta(1-\beta)}$$

We can write the score equation as

$$S_n(\alpha,\beta) = \left(\frac{\partial \ell}{\partial \alpha}, \frac{\partial \ell}{\partial \beta}\right)^T = \left(\frac{x_1+x_2-n\alpha}{\alpha(1-\alpha)},\ \frac{x_1-(x_1+x_2)\beta}{\beta(1-\beta)}\right)^T$$

Hence,

$$S_n(\hat{\theta},\hat{\theta}) = \left(\frac{x_1+x_2-n\hat{\theta}}{\hat{\theta}(1-\hat{\theta})},\ \frac{x_1-(x_1+x_2)\hat{\theta}}{\theta(1-\theta)}\right)$$

The inverse fisher information evaluated at the restricted MLE is

$$[I_n(\hat{\theta},\hat{\theta})]^{-1} = \begin{pmatrix} \frac{1-\hat{\theta}}{n} & 0 \\ 0 & \frac{\hat{\theta}(1-\hat{\theta})}{n\hat{\theta}} \end{pmatrix} = \begin{pmatrix} \frac{1-\hat{\theta}}{n} & 0 \\ 0 & \frac{(1-\hat{\theta})}{n} \end{pmatrix}$$

The score statistic is given by

$$SC_n = S_n(\hat{\theta},\hat{\theta})^T [I_n(\hat{\theta},\hat{\theta})]^{-1} S_n(\hat{\theta},\hat{\theta})$$

which is asymptotically distributed as a $\chi^2(1)$ random variable under H_0.

1.g.

Note that this is a linear test. Let $\xi = (\alpha, \beta)^T$. Then $\alpha - \beta = R\xi$ where $R = (1, -1)$. The Wald Test statistic for a linear hypothesis is given by

$$(R\hat{\xi} - b_0)^T [R[I_n(\hat{\xi})]^{-1} R^T]^{-1} (R\hat{\xi} - b_0)$$

$R\hat{\xi} = \hat{\alpha} - \hat{\beta}$ and

$$[R[I_n(\hat{\xi})]^{-1} R^T]^{-1} = \left\{ \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{\hat{\alpha}(1-\hat{\alpha})}{n} & 0 \\ 0 & \frac{\hat{\beta}(1-\hat{\beta})}{n\alpha} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}^{-1}$$

$$= \left\{ \frac{1}{n} \left[\hat{\alpha}(1-\hat{\alpha}) + \frac{\hat{\beta}(1-\hat{\beta})}{\alpha} \right] \right\}^{-1}$$

$$= \frac{n\hat{\alpha}}{\hat{\alpha}^2(1-\alpha) + \hat{\beta}(1-\hat{\beta})}$$

Thus, the Wald Test is given by

$$W_n = \frac{n\hat{\alpha}(\hat{\alpha} - \hat{\beta})^2}{\hat{\alpha}^2(1-\hat{\alpha}) + \hat{\beta}(1-\hat{\beta})} \xrightarrow{\text{d}} \chi^2(1)$$

as $n \to \infty$ under H_0.

1.h.

Note that

$$\begin{aligned}
L(\alpha,\beta) &= \alpha^{x_1+x_2}\beta^{x_1}(1-\beta)^{x_2}(1-\alpha)^{x_3} \\
&= \exp\{(x_1+x_2)\log\alpha + x_1\log\beta + x_2\log(1-\beta) + x_3\log(1-\alpha)\} \\
&= \exp\{(n-x_3)\log\alpha + x_1\log\beta + (n-x_1-x_3)\log(1-\beta) + x_3\log(1-\alpha)\} \\
&= \exp\left\{x_1\log\frac{\beta}{1-\beta} + x_3\log\frac{1-\alpha}{\alpha} + n\log[\alpha(1-\beta)]\right\}
\end{aligned}$$

This is a full rank multivariate exponential family and since $\log\frac{\beta}{1-\beta}$ is a 1-1 function of β, we have that X_1 is a sufficient statistic for β and similarly X_3 is a sufficient statistic for α. Marginally, $X_3 \sim$Bin$(n, 1-\alpha)$. We have

$$P(x_1,x_2,x_3|x_3) = \frac{P(x_1,x_2,x_3)}{P(x_3)}$$

$$\begin{aligned}
&= \frac{\frac{n!}{x_1!x_2!x_3!}\alpha^{x_1+x_2}\beta^{x_1}(1-\beta)^{x_2}(1-\alpha)^{x_3}}{\frac{n!}{x_3!(n-x_3)!}(1-\alpha)^{x_3}\alpha^{n-x_3}} \\
&= \frac{\frac{n!}{x_1!x_3!(n-x_1-x_3)!}\alpha^{n-x_3}\beta^{x_1}(1-\beta)^{n-x_1-x_3}(1-\alpha)^{x_3}}{\frac{n!}{x_3!(n-x_3)!}(1-\alpha)^{x_3}\alpha^{n-x_3}}
\end{aligned}$$

$$= \frac{(n-x_3)!}{x_1!(n-x_3-x_1)!}\beta^{x_1}(1-\beta)^{n-x_1-x_3}$$

$$= \binom{n-x_3}{x_1}\beta^{x_1}(1-\beta)^{(n-x_3)-x_1}$$

This is the p.m.f. of a Bin$(n-x_3, \beta)$ random variable. From basic probability theory, we know that $\hat{\beta}_c$ is simply the average number of successes, i.e.,

$$\hat{\beta}_c = \frac{x_1}{n-x_3} = \frac{x_1}{x_1+x_2+x_3-x_3} = \frac{x_1}{x_1+x_2}$$

This estimate is intuitive since it is equal to the unconditional MLE, which was independent of α