ST 762, HOMEWORK 5 EXTRA PROBLEMS SOLUTIONS, FALL 2009

1. (a) Statistician A has specified a model for the "typical response profile." Thus, in her model, β_0 represents the value such that $H(\beta_0)$ is the probability that a randomly chosen individual from the population will have a flare-up at time 0 (response at time 0). Note this is different from the situation in B's model; in that model, the probability that a randomly chosen individual from the population will have a flare-up at time 0, depends on who that individual is and is equal to $H(\beta_0 + b_i)$ for the ith individual. Similarly, β_1 is the parameter that dictates how the probability that a randomly chosen individual will have a flare-up at time x_{ij} changes with time. That is, β_0 and β_1 dictate the average response profile (probabilities of flare-up) over time in the population.

   (b) In Statistician B's model, if $\pi_{ij} = H(\beta_0 + \beta_1 x_{ij})$, π_{ij} is the individual-specific probability of having a flare-up at time x_{ij} for subject i, and we have $H^{-1}(\pi_{ij}) = \beta_0 + \beta_1 x_{ij} + b_i$ ("linear predictor" for i), the log odds. Here, $H(\beta_0 + b_i)$ represents i's personal probability of a flare-up at baseline (time 0) at the start of treatment, and β_1 dictates the change in i's probability over time. There is no random effect associated with this change parameter, so B's model makes the assumption that this individual-specific dependence on time is the same for all subjects. Thus, β_0 represents the "typical" value of the parameter $\beta_0 + b_i$ in the population that determines a *specific individual's* baseline probability of flare-up. β_1 represents the "typical value" in the population of how the change in probability takes place for individual subjects (which just so happens to be the same for all subjects in this model).

   That is, β_0 and β_1 have to do with the population of individual-specific probabilities of toenail fungus flare-up.

   (c) First, it is straightforward to derive by a change of variables that if $F(u) = 1/(1 + e^{-u})$, then $X = (U - a)/d$ has CDF $G(x) = 1/[1 + \exp\{-(a + dx)\}$. This looks like A's model, identifying $a = \beta_0$ and $d = \beta_1$. Now from the hint we have that $F(u) = 1/(1+e^{-u}) \approx \Phi(u/c)$, where $\Phi(\cdot)$ is the standard normal CDF. Thus, writing $u = a + dx$, we may approximate $G(x)$

   $$G(x) \approx \Phi\left(\frac{a}{c} + \frac{dx}{c}\right), \tag{1}$$

   and we may approximate A's model as

   $$E(Y_{ij}|x_{ij}) \approx \Phi\left(\frac{\beta_0}{c} + \frac{\beta_1 x_{ij}}{c}\right).$$

   (d) To find an approximation for B's model, first note that, from the results in (c), identifying $a = \beta_0 + b_i$ and $d = \beta_1$, we may approximate his model for the conditional expectation $E(Y_{ij}|x_{ij}, b_i)$ by

   $$\Phi\left(\frac{\beta_0}{c} + \frac{\beta_1 x_{ij}}{c} + \frac{b_i}{c}\right).$$

   Thus, if we want to approximate $E(Y_{ij}|x_{ij})$ for B's model, we need to evaluate the right hand side of

   $$E(Y_{ij}|x_{ij}) = E\{E(Y_{ij}|x_{ij}, b_i)\} \approx E\left\{\Phi\left(\frac{\beta_0}{c} + \frac{\beta_1 x_{ij}}{c} + \frac{b_i}{c}\right)\right\},$$

   where the outer expectation is with respect to the distribution of b_i This may be written as

   $$\frac{1}{D^{1/2}\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi\left(\frac{\beta_0}{c} + \frac{\beta_1 x_{ij}}{c} + \frac{b_i}{c}\right) \exp\left(-\frac{b_i^2}{2D}\right) db_i.$$

Write $\beta_0^c = \beta_0/c$ and $\beta_1^c = \beta_1/c$. Making the transformation $b = b_i/c$ in the integral, and letting $D_c = D/c^2$, we may write the integral as

$$\frac{1}{D_c^{1/2}\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi\left(\beta_0^c + \beta_1^c x_{ij} + b\right) \exp\left(-\frac{b^2}{2D_c}\right) db.$$

The integral is of the form

$$\frac{1}{D_c^{1/2}} \int_{-\infty}^{\infty} \Phi(a+b)\varphi(b^2/D_c^{1/2}) db, \tag{2}$$

where for our problem $a = \beta_0^c + \beta_1^c x_{ij}$ and $\varphi(\cdot)$ is the standard normal density. We thus need to compute an integral of this form. Here are two solutions.

Easy way: Let $X \sim \mathcal{N}(0, D)$ and $Z \sim \mathcal{N}(0, 1)$, where X and Z are independent. Then note that

$$\begin{aligned} P(X+Z<a) &= E\{I(X+Z<a)\} \\ &= E[E\{I(X+Z<a)|X\}] \\ &= E[P(X+Z<a|X)\} \\ &= \frac{1}{D_c^{1/2}} \int_{-\infty}^{\infty} P(Z<a-x|X=x)\varphi(x^2/D_c^{1/2})\, dx \\ &= \frac{1}{D_c^{1/2}} \int_{-\infty}^{\infty} \Phi(a-x)\varphi(x^2/D_c^{1/2})\, dx, \end{aligned}$$

where the last equality follows by the independence of X and Z. Upon the change of variables $b = -x$, this may be written as

$$\frac{1}{D_c^{1/2}} \int_{-\infty}^{\infty} \Phi(a+b)\varphi(b^2/D_c^{1/2})\, db,$$

which is exactly the integral (2) we seek. But note that, because X and Z are independent and normal, we know immediately that $X + Z \sim \mathcal{N}\{0, (1 + D_c)^{1/2}\}$, from whence it follows that

$$P(X+Z<a) = P\left\{\frac{X+Z}{(1+D_c)^{1/2}} < \frac{a}{(1+D)^{1/2}}\right\} = \Phi\{a(1+D_c)^{-1/2}\}.$$

Thus, we find that the integral we are interested in is equal to $\Phi\{a(1 + D_c)^{-1/2}\}$.

Less easy way: The integral in (2) may be written as

$$\frac{1}{2\pi D_c^{1/2}} \int_{-\infty}^{\infty} \left\{\int_{-\infty}^{a+b} \frac{1}{2\pi} \exp\left(-\frac{u^2}{2}\right) du\right\} \exp\left(-\frac{b^2}{2D_c}\right) db.$$

We may rewrite the inner integral to obtain

$$\frac{1}{(2\pi D_c)^{1/2}} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{a} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{(u-b)^2}{2}\right\} du\right] \exp\left(-\frac{b^2}{2D_c}\right) db.$$

Switching the order of integration, we have

$$\frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{a} \left(\int_{-\infty}^{\infty} \frac{1}{(2\pi D_c)^{1/2}} \exp\left[-\frac{1}{2D_c}\{D_c(u-b)^2 + b^2\}\right] db\right) du.$$

By some algebra (complete the square), we may show that

$$D_c(u-b)^2+b^2=(1+D_c)\left(b-\frac{uD_c}{1+D_c}\right)^2+\frac{D_cu^2}{1+D_c}.$$

Substituting this into the inner integral, we obtain

$$\frac{1}{(2\pi)^{1/2}(1+D_c)^{-1/2}}\int_{-\infty}^{a}\left[\int_{-\infty}^{\infty}\frac{1}{(2\pi D_c)^{1/2}(1+D_c)^{-1/2}}\exp\left\{-\frac{\left(b-\frac{uD_c}{1+D_c}\right)^2}{2D_c(1+D_c)^{-1}}\right\}db\right]$$

$$\times\exp\left\{-\frac{u^2}{2(1+D_c)}\right\}du.$$

Clearly, the inner integral is equal to 1, so that the expression reduces to

$$\frac{1}{(2\pi)^{1/2}(1+D_c)^{-1/2}}\int_{-\infty}^{a}\exp\left\{-\frac{u^2}{2(1+D_c)}\right\}du.$$

Clearly, this is equal to $\Phi\{a(1+D_c)^{-1/2}\}$.

Thus, we have two different arguments showing that (2) is equal to $\Phi\{a(1+D_c)^{-1/2}\}$. Applying this to the problem at hand, then, we have that

$$E(Y_{ij}|x_{ij})=\Phi\left\{\frac{\beta_0^c+\beta_1^cx_{ij}}{(1+D_c)^{1/2}}\right\}=\Phi(\beta_0^*+\beta_1^*x_{ij}),$$

where $\beta_0^*=\beta_0/(c^2+D)^{1/2}$ and $\beta_1^*=\beta_1/(c^2+D)^{1/2}$.

(e) Comparing the approximations to the models, we see that they are both of the same form. For Statistician A, we have

$$E(Y_{ij}|x_{ij})\approx\Phi\left(\frac{\beta_0}{c}+\frac{\beta_1x_{ij}}{c}\right).$$

For B, we have

$$E(Y_{ij}|x_{ij})\approx\Phi\left(\frac{\beta_0}{(c^2+D)^{1/2}}+\frac{\beta_1x_{ij}}{(c^2+D)^{1/2}}\right).$$

Thus, the approximation to the model for marginal mean is *the same* for each statistician – both may be approximated by probit models with a simple linear predictor in time. The difference is that the "intercept" and "slope" of the linear predictors are not the same. Those for Statistician B are *attenuated* relative to those of Statistician A; that is, $\beta_k/c\geq\beta_k/(c^2+D)^{1/2}$ for $k=0,1$. This shows explicitly that β_0 and β_1 in each model cannot have the same interpretation in general. In the probit approximation to Statistician A's model, the parameters producing the "typical response profile" are β_0/c and β_1/c, while for statistician B the parameters with this interpretation in the approximate marginal model are $\beta_0(c^2+D)^{-1/2}$ and $\beta_1(c^2+D)^{-1/2}$. This shows that, at least approximately, the marginal probabilities of fungus flare-up under B's model will always be smaller than those under A's model for the same values of β_0 and β_1. Note that when $D=0$, so that baseline probabilities do not vary across subjects in B's model, the models become the same.

(f) Statistician A thinks she is estimating the parameters producing the typical response profile. Because, at least approximately, her model and B's model have exactly the same form, she will indeed be estimating the parameters she thinks she is estimating. It so happens that these parameters have true values of $\beta_0(c^2+D)^{-1/2}$ and $\beta_1(c^2+D)^{-1/2}$ in B's model, so these are the values she is consistently estimating, but these are the values that produce the typical response profile under B's model.

2. We use obvious shorthand notation here.

(a) We have from page 384, combining (i) and (ii), that

$$\boldsymbol{Z}_i = \begin{pmatrix} 2v_{i11}^2 & 0 & \cdots & 0 \\ 0 & v_{i11}v_{i12} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 02v_{in_in_i} & \end{pmatrix}.$$

(b) The PL equation depends on the matrix $\boldsymbol{V}_i$ $(n_i \times n_i)$. Under the conditions in (a), $\boldsymbol{V}_i$ is diagonal and $\boldsymbol{V}_i^{-1} = \mathrm{diag}(v_{i11}^{-1}, \cdots, v_{in_in_i}^{-1})$. Thus, the matrix $\partial \boldsymbol{V}_i/\partial \xi_k$ is also diagonal $(n_i \times n_i)$, i.e.,

$$\partial \boldsymbol{V}_i/\partial \xi_k = \mathrm{diag}(v_{\xi i11}, \ldots, v_{\xi in_in_i}),$$

where $v_{\xi ijj} = \partial v_{ijj}/\partial \xi_k$. Thus, the quadratic form

$$(\boldsymbol{Y}_i - \boldsymbol{f}_i)^T \boldsymbol{V}_i^{-1}\{\partial \boldsymbol{V}_i/\partial \xi_k\}\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{f}_i) = \sum_{j=1}^{n_i} \frac{(Y_{ij}-f_{ij})^2}{v_{ijj}} \frac{v_{\xi ijj}}{v_{ijj}}$$

and

$$\mathrm{tr}\left(\boldsymbol{V}_i^{-1}\partial \boldsymbol{V}_i/\partial \xi_k\right) = \sum_{j=1}^{n_i} \frac{v_{\xi ijj}}{v_{ijj}}.$$

Thus, the PL equation has the usual form

$$(1/2)\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left\{\frac{(Y_{ij}-f_{ij})^2}{v_{ijj}}\frac{v_{\xi ijj}}{v_{ijj}}\right\} = 0.$$

Now consider the equation in (4). It is straightforward to see that the kth column of $\boldsymbol{E}_i$ is $\partial \boldsymbol{v}_i/\partial \xi_k$ $\{n_i(n_i+1)/2\}$. Because $\boldsymbol{V}_i$ is a diagonal matrix, this column vector will have elements $v_{\xi ijj}$ in the positions corresponding to the diagonal elements of $\boldsymbol{V}_i$ and zeroes in all other positions. Thus, the kth row will be of the form

$$\sum_{i=1}^{m}(v_{\xi i11}, 0, 0, \ldots, v_{\xi i22}, 0, \ldots, v_{\xi in_in_i}) \begin{pmatrix} (2v_{i11}^2)^{-1} & 0 & \cdots \\ 0 & (v_{i11}v_{i22})^{-1} & \cdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & (2v_{in_in_i}^2)^{-1} \end{pmatrix} (\boldsymbol{u}_i - \boldsymbol{v}_i).$$

Note that premultiplication of $\boldsymbol{Z}_i^{-1}$ by the $\{1 \times n_i(n_i+1)/2\}$ vector of derivative with respect to ξ_k yields a $\{1 \times n_i(n_i+1)/2\}$ row vector with nonzero elements only in the positions that correspond to squared terms in $\boldsymbol{u}_i$. That is, we have

$$(1/2)\sum_{i=1}^{m}(v_{\xi i11}/v_{i11}, 0, \ldots, v_{\xi i22}/v_{i22}, 0, \ldots, v_{\xi in_in_i}/v_{in_in_i}) \left\{ \begin{array}{c} (Y_{i1}-f_{i1})^2 - v_{i11} \\ \vdots \\ (Y_{in_i}-f_{in_i})^2 - v_{in_in_i} \end{array} \right\}.$$

Upon multiplication, it is straightforward to show (you should have done it) that this is identical to the expression above.

(c) We now have $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2})^T$,

$$\boldsymbol{V}_i = \begin{pmatrix} v_{i11} & v_{i12} \\ v_{i12} & v_{i22} \end{pmatrix}, \quad \boldsymbol{u}_i = \begin{pmatrix} (Y_{i1} - f_{i1})^2 \\ (Y_{i1} - f_{i1})(Y_{i2} - f_{i2}) \\ (Y_{i2} - f_{i2}) \end{pmatrix}.$$

Thus, we have $\boldsymbol{Z}_i = \text{var}(\boldsymbol{u}_i | \boldsymbol{x}_i) =$

$$\begin{pmatrix} \text{var}(Y_{i1} - f_{i1})^2 & \text{cov}\{(Y_{i1} - f_{i1})^2, (Y_{i1} - f_{i1})(Y_{i2} - f_{i2})\} & \text{cov}\{(Y_{i1} - f_{i1})^2, (Y_{i2} - f_{i2})^2\} \\ & \text{var}\{(Y_{i1} - f_{i1})(Y_{i2} - f_{i2})\} & \text{cov}\{(Y_{i1} - f_{i1})(Y_{i2} - f_{i2}), (Y_{i2} - f_{i2})^2\} \\ & & \text{var}(Y_{i2} - f_{i2})^2 \end{pmatrix}.$$

From page 384 of the notes, this matrix becomes under the Gaussian working assumption

$$\boldsymbol{Z}_i = \begin{pmatrix} 2v_{i11}^2 & 2v_{i11}v_{i12} & 2v_{i12}^2 \\ & v_{i11}v_{i22} + v_{i12}^2 & 2v_{i12}v_{i22} \\ & & 2v_{i22}^2 \end{pmatrix}.$$

(d) First consider the PL equations. We have

$$\partial \boldsymbol{V}_i / \partial \xi_k = \begin{pmatrix} v_{\xi i11} & v_{\xi i12} \\ v_{\xi i12} & v_{\xi i22} \end{pmatrix}.$$

Note further that

$$\boldsymbol{V}_i^{-1} = \frac{1}{(v_{i11}v_{i22} - v_{i12}^2)} \begin{pmatrix} v_{i22} & -v_{i12} \\ -v_{i12} & v_{i11} \end{pmatrix}.$$

Thus, $\partial \boldsymbol{V}_i / \partial \xi_k \boldsymbol{V}_i^{-1}$ is equal to

$$\frac{1}{(v_{i11}v_{i22} - v_{i12}^2)} \begin{pmatrix} v_{i22}v_{xii11} - v_{i12}v_{\xi 12} & v_{i22}v_{\xi 12} - v_{i12}v_{\xi i22} \\ -v_{i12}v_{\xi i11} + v_{i11}v_{\xi i12} & -v_{i12}v_{xi12} + v_{i11}v_{xii22} \end{pmatrix}.$$

Also,

$$(1/2)\text{tr}\left(\boldsymbol{V}_i^{-1} \partial \boldsymbol{V}_i \partial \xi_k\right) = (1/2)(v_{i11}v_{i22} - v_{i12}^2)^{-1}(v_{i22}v_{\xi i11} - 2v_{i12}v_{\xi i12} + v_{i11}v_{\xi i22}). \qquad (3)$$

It is also straight]forward to show from these results that

$$\boldsymbol{V}_i^{-1}\{\partial \boldsymbol{V}_i \partial \xi_k\}\boldsymbol{V}_i^{-1} = (v_{i11}v_{i22} - v_{i12}^2)^{-2} \begin{pmatrix} A_{i11} & A_{i12} \\ A_{i12} & A_{i22} \end{pmatrix},$$

where

$$\begin{aligned} A_{i11} &= v_{i22}^2 v_{\xi i11} - 2v_{i12}v_{i22}v_{\xi i12} + v_{i12}^2 v_{\xi i22} \\ A_{i12} &= (v_{i12}^2 v_{i11}v_{i22})v_{\xi i12} - v_{i12}v_{i22}v_{\xi i11} - v_{i11}v_{i12}v_{\xi i22} \\ A_{i22} &= v_{i12}^2 v_{\xi i11} - 2v_{i11}v_{i12}v_{\xi i12} + v_{i11}^2 v_{i22}. \end{aligned}$$

Thus, the quadratic form

$$(\boldsymbol{Y}_i - \boldsymbol{f}_i)^T \boldsymbol{V}_i^{-1}\{\partial \boldsymbol{V}_i / \partial \xi_k\}\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{f}_i)$$

$$= (1/2)(v_{i11}v_{i22} - v_{i12}^2)^{-2}\{(Y_{i1} - f_{i1})^2 A_{i11} + 2(Y_{i1} - f_{i1})(Y_{i2} - f_{i2})A_{i12} + (Y_{i2} - f_{i2})A_{i22}\}. \tag{4}$$

To show the equivalence, we'd thus like to show that the kth row of

$$\sum_{i=1}^{m} \boldsymbol{E}_i^T(\boldsymbol{\beta}, \boldsymbol{\xi})\boldsymbol{Z}_i^{-1}(\boldsymbol{\beta}, \boldsymbol{\xi})\boldsymbol{v}_i(\boldsymbol{\beta}, \boldsymbol{\xi})$$

is equivalent to (3) and

$$\sum_{i=1}^{m} \boldsymbol{E}_i^T(\boldsymbol{\beta}, \boldsymbol{\xi})\boldsymbol{Z}_i^{-1}(\boldsymbol{\beta}, \boldsymbol{\xi})\boldsymbol{u}_i$$

is equivalent to (4), where $\boldsymbol{Z}_i$ is the inverse of the matrix found in (c). This turns out to be, using standard formulæ for the inverse of a (3×3) matrix,

$$\boldsymbol{Z}_i^{-1} = \frac{1}{4(v_{i11}v_{i22} - v_{i12}^2)^2} \begin{pmatrix} 2v_{i22}^2 & -4v_{i12}v_{i22} & 2v_{i12}^2 \\ -4v_{i12}v_{i22} & 4(v_{i11}v_{i22} + v_{i12}^2) & -4v_{i12}v_{i11} \\ 2v_{i12}^2 & -4v_{i12}v_{i11} & 2v_{i11}^2 \end{pmatrix}.$$

The kth row of $\boldsymbol{E}_i^T$ is $(v_{\xi i11}, v_{\xi i12}, v_{\xi i22})$, so that the kth row of $\boldsymbol{E}_i^T\boldsymbol{Z}^{-1} + i$ is the transpose of the (3×1) vector

$$\frac{1}{2(v_{i11}v_{i22} - v_{i12}^2)^2} \begin{pmatrix} v_{\xi i11}v_{i22}^2 - 2v_{\xi i12}v_{i12}v_{i22} + v_{\xi i22}v_{i12}^2 \\ -2v_{i12}v_{i22}v_{\xi i11} + 2v_{i11}v_{i22}v_{\xi i12} + 2v_{\xi i12}v_{i12}^2 - 2v_{i12}v_{i11} \\ v_{i12}^2 v_{\xi i11} - 2v_{i12}v_{i11}v_{\xi i12} + v_{i12}^2 v_{\xi i22} \end{pmatrix}.$$

We may now compute the two pieces of the estimating equation. After simplification, we end up with

$$\boldsymbol{E}_i^T\boldsymbol{Z}_i^{-1}\boldsymbol{v}_i = (1/2)(v_{i11}v_{i22} - v_{i12}^2)^{-1}(v_{i22}v_{\xi i11} - 2v_{i12}v_{\xi i12} + v_{\xi i22}v_{i11},$$

which is the ith element of (3). Similarly,

$$\boldsymbol{E}_i^T\boldsymbol{Z}_i^{-1} = (1/2)(v_{i11}v_{i22} - v_{i12}^2)^{-2}(A_{i11}, 2A_{i12}, A_{i22}).$$

Algebra can then be used to show that the the kth row of $\boldsymbol{E}_i^T\boldsymbol{Z}_i^{-1}\boldsymbol{u}_i$ is exactly equal to that of a summand in (4).