Section 8.4. Example

Let $\{(R_i, X_i) : i = 1, \ldots, n\}$ be an i.i.d. sample of n random vectors (R, X). Here R is a response indicator and X is a covariate. We assume that

$$logit P[R = 1|X] = \alpha + \beta X$$

and assume that X is normally distributed with mean μ and variance σ^2. So, our probability model has four parameters, $\theta = (\alpha, \beta, \mu, \sigma^2)$. Let $\theta_0 = (\alpha_0, \beta_0, \mu_0, \sigma_0^2)$ denote the true value of θ.

a. For a given realization of the data, write out the likelihood function of θ.

$$L(\theta; \boldsymbol{x}, \boldsymbol{r}) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2\sigma^2}(x_i - \mu)^2) \frac{\exp(r_i(\alpha + \beta x_i))}{1 + \exp(\alpha + \beta x_i)}$$

b. Find the maximum likelihood estimator of θ, $\hat{\theta}_n$ (If there exists a closed form solution, then present it. If not, indicate how a solution can be found).

To find the MLE, solve the score equations. The log-likelihood for an individual is

$$l(\theta; x, r) \propto -\log(\sigma) - \frac{1}{2\sigma^2}(x-\mu)^2 + r(\alpha+\beta x) - \log(1+\exp(\alpha+\beta x))$$

The score vector for an individual is

$$\psi(x, r; \theta) = \begin{bmatrix} \frac{\partial l(\theta;x,r)}{\partial \alpha} \\ \frac{\partial l(\theta;x,r)}{\partial \beta} \\ \frac{\partial l(\theta;x,r)}{\partial \mu} \\ \frac{\partial l(\theta;x,r)}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} r - \frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)} \\ x(r - \frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}) \\ \frac{x-\mu}{\sigma^2} \\ -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^4} \end{bmatrix}$$

The score equations for the full sample are:

$$\sum_{i=1}^{n} r_i - \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)} = 0 \tag{1}$$

$$\sum_{i=1}^{n} x_i(r_i - \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)}) = 0 \tag{2}$$

$$\sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} = 0 \tag{3}$$

$$\sum_{i=1}^{n} -\frac{1}{2\sigma^2} + \frac{(x_i - \mu)^2}{2\sigma^4} = 0 \tag{4}$$

Note that Equations (3) and (4) can be solved explicitly to get solutions for $\hat{\mu}$ and $\hat{\sigma}^2$, i.e.,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ and } \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$$

The solutions for $\hat{\alpha}$ and $\hat{\beta}$ are obtained by solving Equations (1) and (2). This does not yield simple closed form solutions. Therefore, we use the Newton-Raphson algorithm. This entails computing the observed information matrix. Some of these computations will be needed later so let's compute the entire matrix of second partial derivatives. Let $p_i(\alpha, \beta) = \frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}$ and $q_i(\alpha, \beta) = 1 - p_i(\alpha, \beta)$, Now,

$$nJ_n(\theta) = \begin{bmatrix} J_{n1}(\alpha, \beta) & 0 \\ 0 & J_{n2}(\mu, \sigma^2) \end{bmatrix}$$

where

$$J_{n1}(\alpha, \beta) = \begin{bmatrix} \sum_{i=1}^n p_i(\alpha, \beta) q_i(\alpha, \beta) & \sum_{i=1}^n x_i p_i(\alpha, \beta) q_i(\alpha, \beta) \\ \sum_{i=1}^n x_i p_i(\alpha, \beta) q_i(\alpha, \beta) & \sum_{i=1}^n x_i^2 p_i(\alpha, \beta) q_i(\alpha, \beta) \end{bmatrix}$$

and

$$J_{n2}(\mu, \sigma^2) = \begin{bmatrix} \frac{n}{\sigma^2} & \frac{1}{\sigma^4}\sum_{i=1}^n (x_i - \mu) \\ \frac{1}{\sigma^4}\sum_{i=1}^n (x_i - \mu) & -\frac{n}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{i=1}^n (x_i - \mu)^2 \end{bmatrix}$$

To compute the the solution to (1) and (2), we start with initial values for α and β, say $\alpha^{(0)}$ and $\beta^{(0)}$, then update as follows:

$$\begin{pmatrix} \alpha^{(j+1)} \\ \beta^{(j+1)} \end{pmatrix} = \begin{pmatrix} \alpha^{(j)} \\ \beta^{(j)} \end{pmatrix} + \{J_{n1}(\alpha^{(j)}, \beta^{(j)})\}^{-1} \cdot \begin{pmatrix} \sum_{i=1}^n (r_i - p_i(\alpha^{(j)}, \beta^{(j)}) \\ \sum_{i=1}^n x_i(r_i - p_i(\alpha^{(j)}, \beta^{(j)}) \end{pmatrix}$$

We iterate until convergence.

Assume that $|\alpha| \leq K_1$, $|\beta| \leq K_2$, $|\mu| \leq K_3$, and $K_4 \leq \sigma^2 \leq K_5$, where $K_1, \ldots, K_5$ are finite positive quantities. Is the parameter space, Θ, compact? YES! Assume that θ_0 (truth) belongs to the interior of Θ.

c. Verify all of the regularity conditions that are necessary to show that the MLE is both consistent and asymptotically normal.

Refer to Section 8c. Condition i) is satisfied by assumption. Condition ii) is satisfied since $\log p(x, r; \theta)$ is continuous at each θ for all $-\infty < x < \infty$ and $r = 0, 1$. Now, we need to show condition iii) that $|\log p(x, r; \theta)| \leq d(x, r)$ for all $\theta \in \Theta$ and $E_{\theta_0}[d(X, R)] < \infty$.

$$\log p(x, r; \theta) = -\frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{1}{2\sigma^2}(x-\mu)^2 + r(\alpha+\beta x) - \log(1+\exp(\alpha+\beta x))$$

By the triangle inequality, we know that

$$|\log p(x, r; \theta)| \leq \frac{1}{2}\log(2\pi) + |\log(\sigma)| + |\frac{1}{2\sigma^2}(x-\mu)^2| + |r(\alpha+\beta x)| + |\log(1+\exp(\alpha+\beta x))|$$

Some inequalities that we will need.

$$|x| \leq 1 + x^2$$

$$\log(1 + \exp(\alpha + \beta x)) \geq 0$$

$$\begin{aligned}
\log(1 + \exp(\alpha + \beta x)) &\leq \log(1 + \exp(|\alpha| + |\beta||x|)) \\
&\leq 1 + \log(\exp(|\alpha| + |\beta||x|)) \\
&= 1 + |\alpha| + |\beta||x| \\
&\leq 1 + K_1 + K_2(1 + x^2)
\end{aligned}$$

$$|\log(\sigma)| \leq \frac{1}{2}\max(|\log(K_4)|, |\log(K_5)|)$$

$$\begin{aligned} \left|\frac{1}{2\sigma^2}(x-\mu)^2\right| &\leq \frac{1}{2\sigma^2}x^2 + \frac{1}{\sigma^2}|\mu||x| + \frac{\mu^2}{2\sigma^2} \\ &\leq \frac{x^2}{2K_4} + \frac{K_3}{K_4}(1+x^2) + \frac{K_3^2}{2K_4} \end{aligned}$$

$$|r(\alpha+\beta x)| \leq |\alpha| + |\beta||x| \leq K_1 + K_2(1+x^2)$$

$$|\log(1+\exp(\alpha+\beta x))| \leq 1 + K_1 + K_2(1+x^2)$$

Let

$$\begin{aligned} d(x,r) = \; & \frac{1}{2}\log(2\pi) + \frac{1}{2}\max(|\log(K_4)|, |\log(K_5)|) + \\ & \frac{x^2}{2K_4} + \frac{K_3}{K_4}(1+x^2) + \frac{K_3^2}{2K_4} + K_1 + K_2(1+x^2) + \\ & 1 + K_1 + K_2(1+x^2) \end{aligned}$$

Now, $E_{\theta_0}[X^2] = \mu_0^2 + \sigma_0^2 < \infty$. This implies that $E_{\theta_0}[d(X,R)] < \infty$.

For condition iv), we have to show that $p(x,r;\theta)$ is twice-continuously differentiable and $p(x, r; \theta) > 0$ for θ in a neighborhood of θ_0. This follows simply from the fact that the density is made up of polynomial and exponentials for which derivatives exists. The density is clearly positive over the entire parameter space.

For condition v), we have to show that $\|\frac{\partial p(x,r;\theta)}{\partial\theta}\| \leq e(x,r)$, where $e(x,r)$ is integrable. Remember,

$$p(x,r;\theta) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\frac{\exp(r(\alpha+\beta x))}{1+\exp(\alpha+\beta x)}$$

and

$$\frac{\partial p(x,r;\theta)}{\partial\theta} = \begin{bmatrix} \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\frac{r\exp(r(\alpha+\beta x))+(r-1)\exp((r+1)(\alpha+\beta x)}{(1+\exp(\alpha+\beta x))^2} \\ \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\frac{xr\exp(r(\alpha+\beta x))+x(r-1)\exp((r+1)(\alpha+\beta x)}{(1+\exp(\alpha+\beta x))^2} \\ \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\frac{\exp(r(\alpha+\beta x))}{(1+\exp(\alpha+\beta x))}\{\frac{x-\mu}{\sigma^2}\} \\ \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\frac{\exp(r(\alpha+\beta x))}{(1+\exp(\alpha+\beta x))}\{-\frac{1}{2\sigma^2}+\frac{(x-\mu)^2}{2\sigma^4}\} \end{bmatrix}$$

Now,

$$\|\frac{\partial p(x,r;\theta)}{\partial\theta}\| \leq |\frac{\partial p(x,r;\theta)}{\partial\alpha}|+|\frac{\partial p(x,r;\theta)}{\partial\beta}|+|\frac{\partial p(x,r;\theta)}{\partial\mu}|+|\frac{\partial p(x,r;\theta)}{\partial\sigma^2}|$$

We want to bound each of the terms in the sum by a function of x and r which is integrable.

To do this, note that

$$\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2) \leq e^*(x)$$

where

$$e^*(x) = \begin{cases} \frac{1}{\sqrt{2K_4\pi}} & -K_3 \leq x \leq K_3 \\ \frac{1}{\sqrt{2K_4\pi}}\exp(-\frac{1}{2K_5}(x-K_3)^2) & x > K_3 \\ \frac{1}{\sqrt{2K_4\pi}}\exp(-\frac{1}{2K_5}(x+K_3)^2) & x < -K_3 \end{cases}$$

Now,

$$\begin{aligned}
|\frac{\partial p(x,r;\theta)}{\partial \alpha}| &\leq 2e^*(x) \equiv e_1(x) \\
|\frac{\partial p(x,r;\theta)}{\partial \beta}| &\leq 2|x|e^*(x) \leq 2(1+x^2)e^*(x) \equiv e_2(x) \\
|\frac{\partial p(x,r;\theta)}{\partial \mu}| &\leq e^*(x)(\frac{|x|+K_3}{K_4}) \leq e^*(x)\frac{1+x^2+K_3}{K_4} \equiv e_3(x) \\
|\frac{\partial p(x,r;\theta)}{\partial \sigma^2}| &\leq e^*(x)(\frac{1}{2K_4} + \frac{x^2(1+2K_3)+2K_3+K_3^2}{2K_4^2}) \equiv e_4(x)
\end{aligned}$$

We know that $\|\frac{\partial p(x,r;\theta)}{\partial \theta}\| \leq e(x) = e_1(x) + e_2(x) + e_3(x) + e_4(x)$. $e(x)$ is integrable since each $e_i(x)$ $(i = 1, \ldots, 4)$ is integrable.

For condition vi), we need to show that $I(\theta_0) = E_{\theta_0}[\psi(X, R; \theta_0)\psi(X, R; \theta_0)']$ is non-singular. It would suffice to show that $I(\theta_0)$ was positive definite. That is, $a'I(\theta_0)a > 0$ for all a such that $\|a\| > 0$. Alternatively, we can show that $Var_{\theta_0}[a'\psi(X, R; \theta_0)] > 0$.

Now,

$$\begin{aligned} a'\psi(X, R; \theta_0) \quad = \quad & a_1(R - \frac{\exp(\alpha_0 + \beta_0 X)}{1 + \exp(\alpha_0 + \beta_0 X)}) + \\ & a_2 X(R - \frac{\exp(\alpha_0 + \beta_0 X)}{1 + \exp(\alpha_0 + \beta_0 X)}) + \\ & a_3(\frac{X - \mu_0}{\sigma_0^2}) + a_4(-\frac{1}{2\sigma_0^2} + \frac{(X - \mu_0)^2}{2\sigma_0^4}) \end{aligned}$$

$$Var_{\theta_0}[a'\psi(X, R; \theta_0)] = E_{\theta_0}[Var_{\theta_0}[a'\psi(X, R; \theta_0)|X]] + Var_{\theta_0}[E_{\theta_0}[a'\psi(X, R; \theta_0)|X]]$$

$$
\begin{aligned}
Var_{\theta_0}[a'\psi(X,R;\theta_0)|X] &= (a_1 + a_2 X)^2 \frac{\exp(\alpha_0 + \beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2} \\
E_{\theta_0}[a'\psi(X,R;\theta_0)|X] &= a_3(\frac{X-\mu_0}{\sigma_0^2}) + a_4(-\frac{1}{2\sigma_0^2} + \frac{(X-\mu_0)^2}{2\sigma_0^4})
\end{aligned}
$$

So,

$$
\begin{aligned}
Var_{\theta_0}[a'\psi(X,R;\theta_0)] &= E_{\theta_0}[(a_1 + a_2 X)^2 \frac{\exp(\alpha_0 + \beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] + \\
& \quad Var_{\theta_0}[a_3(\frac{X-\mu_0}{\sigma_0^2}) + a_4(-\frac{1}{2\sigma_0^2} + \frac{(X-\mu_0)^2}{2\sigma_0^4})] \\
&= E_{\theta_0}[(a_1 + a_2 X)^2 \frac{\exp(\alpha_0 + \beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] + \\
& \quad \frac{a_3^2}{\sigma_0^2} + \frac{a_4^2}{4\sigma_0^8} Var_{\theta_0}[(X-\mu_0)^2]
\end{aligned}
$$

The variance can only equal zero if $a_1 = a_2 = a_3 = a_4 = 0$. So,

$I(\theta_0)$ is positive definite.

Conditions vii) and viii) involve showing that the norm of the matrix of second derivatives of the log likelihood and likelihood are bounded by functions of x and r which are integrable. The norm of a matrix $A = [a_{ij}]$ is equal to $(\sum_{i,j} a_{ij}^2)^{1/2}$. For each of these matrices, it is sufficient to show that the absolute value of each of entries is bounded by an integrable function of x and r. The bounding techniques are very similar to those used to show condition v).

d. Find the asymptotic variance of $\hat{\theta}_n$.

The asymptotic variance is found by taking the inverse of the information matrix.

$$I(\theta_0)^{-1} = \begin{bmatrix} E_{\theta_0}[\frac{\exp(\alpha_0+\beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] & E_{\theta_0}[X\frac{\exp(\alpha_0+\beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] & 0 & 0 \\ E_{\theta_0}[X\frac{\exp(\alpha_0+\beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] & E_{\theta_0}[X^2\frac{\exp(\alpha_0+\beta_0 X)}{(1+\exp(\alpha_0+\beta_0 X))^2}] & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_0^2} & 0 \\ 0 & 0 & 0 & \frac{1}{2\sigma_0^4} \end{bmatrix}^{-1}$$

e. Find an estimator for the probability of response for an individual with a covariate value which falls at the pth quantile of the covariate distribution. Find the asymptotic variance of this estimator.

The pth quantile of the covariate distribution is $\mu + \sigma\Phi^{-1}(p)$, where $\Phi(\cdot)$ is the c.d.f of a standard Normal distribution. The probability of response for an individual with this covariate value is

$$h(\theta) = \frac{\exp(\alpha + \beta(\mu + \sigma\Phi^{-1}(p)))}{1 + \exp(\alpha + \beta(\mu + \sigma\Phi^{-1}(p)))}$$

Thus, our estimator is

$$h(\hat{\theta}) = \frac{\exp(\hat{\alpha} + \hat{\beta}(\hat{\mu} + \sqrt{\hat{\sigma}^2}\Phi^{-1}(p)))}{1 + \exp(\hat{\alpha} + \hat{\beta}(\hat{\mu} + \sqrt{\hat{\sigma}^2}\Phi^{-1}(p)))}$$

The asymptotic variance is given by $(\frac{\partial h(\theta_0)}{\partial\theta})' I^{-1}(\theta_0)(\frac{\partial h(\theta_0)}{\partial\theta})$.

f. At what quantile of the covariate distribution would we expect the probability of response to be 0.5? Find an estimator for this quantile and find its asymptotic variance.

What value of x leads to a conditional response probability of 0.5? Solving

$$\frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} = 0.5$$

we find that $x = -\alpha/\beta$. The associated quantile is $p(\theta) = \Phi(\frac{-\alpha/\beta - \mu}{\sigma})$. Thus, our estimator is

$$p(\hat{\theta}) = \Phi(\frac{-\hat{\alpha}/\hat{\beta} - \hat{\mu}}{\sqrt{\hat{\sigma}^2}})$$

The asymptotic variance of $p(\hat{\theta})$ is given by $(\frac{\partial p(\theta_0)}{\partial \theta})' I^{-1}(\theta_0)(\frac{\partial p(\theta_0)}{\partial \theta})$.