Stat 513
HW 2 Solutions
January 30, 2004

- 1.
Let $Y_1, ..., Y_n \overset{iid}{\sim} \mathcal{B}(1, p_i)$ with

$$logit\ (p_i) \quad = \quad log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_i$$

for known covariate x_i.
Use a Newton-Raphson algorithm to find the MLE's for β_0 and β_1, providing the value of the loglikelihood, score, Fisher's information, and updated estimates at each iteration using initial estimates of $\hat{\beta}_{0(0)} = logit(\overline{Y})$ and $\hat{\beta}_{1(0)} = 0$.

For simplicity, let $x_i\beta = \beta_0 + \beta_1 x_i$

Recall that:

$$\begin{aligned}
\text{Likelihood} \quad &\Rightarrow \quad \prod_{i=1}^{n} \frac{e^{(x_i\beta)y_i}}{1+e^{x_i\beta}} \\
\text{Log-likelihood} \quad &\Rightarrow \quad \sum_{i=1}^{n} \left[(x_i\beta)y_i - ln(1+e^{x_i\beta})\right] \\
\text{Score: } \beta_0 \quad &\Rightarrow \quad \sum_{i=1}^{n} \left[y_i - \frac{1}{(1+e^{x_i\beta})} e^{x_i\beta}\right] \\
\beta 1 \quad &\Rightarrow \quad \sum_{i=1}^{n} \left[x_i y_i - \frac{x_i}{(1+e^{x_i\beta})} e^{x_i\beta}\right]
\end{aligned}$$

- 2.
Let $Y_1, ..., Y_n$ be independent, Normally distributed with $Y_i \sim N(\mu_i, \sigma^2)$, where $\mu_i = \beta_0 + \beta_1 x_i$.

  (a) Find formulas for the MLE's and covariance matix of those estimates. What is the distribution of the estimates? Are they unbiased? Consistent? Efficient?
  MLE's:

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

$$
\begin{aligned}
L(\beta,\sigma^2) &= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)} \\
\mathcal{L}(\beta,\sigma^2) &= -\frac{n}{2}\log(\sigma^2) - \frac{(Y-X\beta)'(Y-X\beta)}{2\sigma^2} \\
\mathcal{U}(\beta) &= \frac{X'(Y-X\beta)}{\sigma^2} \\
&= \frac{X'X}{\sigma^2}((X'X)^{-1}X'Y-\beta) \\
\mathcal{U}(\sigma^2) &= -\frac{n}{2\sigma^2} + \frac{RSS}{2\sigma^4} \\
&= -\frac{n}{2\sigma^4}(\frac{RSS}{n} - \sigma^2) \\
I(\beta) &= \frac{X'X}{\sigma^2} \\
I(\sigma^2) &= \frac{n}{2\sigma^4} \\
I(\beta,\sigma^2) &= \frac{1}{\sigma^2}\begin{pmatrix} X'X & 0 \\ 0 & \frac{n}{2\sigma^2} \end{pmatrix} \\
&\ \text{Since second partials have expectation zero}
\end{aligned}
$$

Formulas for MLE's:

$$
\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
\hat{\sigma}^2 &= \frac{RSS}{n}
\end{aligned}
$$

Now the normal is regular, and we can put both scores in nice form, so we know we have consistency, and asymptotic distribution based on information. We now check unbiasedness and show that $\hat{\beta}$ is efficient because the variance is equal to the inverse information:

$$
\begin{aligned}
E(\hat{\beta}) &= E((X'X)^{-1}X'Y) \\
&= (X'X)^{-1}X'X\beta \\
&= \beta \\
Var(\hat{\beta}) &= Var((X'X)^{-1}X'Y) \\
&= (X'X)^{-1}X'Var(Y)X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1} \\
E(\hat{\sigma}^2) &= E(\frac{RSS}{n}) \\
&= \frac{1}{n}E[(Y-X\hat{\beta})'(Y-X\hat{\beta})] \\
&= \frac{(n-2)\sigma^2}{n} \\
&\ \text{So } \hat{\sigma}^2 \text{ is not unbiased}
\end{aligned}
$$

(b) Now suppose that all is known is $E[Y_i] = \mu_i = \beta_0 + \beta_1 x_i$ and $Var(Y_i) = \sigma^2$. Discuss issues in finding the asymptotic distribution of β. Our previous claims of unbiasedness had nothing to do with the specific distribution of Y, just expectation and variance. Therefore, our inference is still true if we can apply a Linderberg-Feller condition to use the CLT and asymptotic distribution holds.

- 3.
Let $T_1, ..., T_n \overset{iid}{\sim} Weib(\rho, \lambda)$ where $Pr(T_i > t) = e^{-(\lambda t)^\rho}$ is the survivor function for $t, \rho, \lambda > 0$. Also let $C_i \sim \mathcal{U}(a, b)$ denote the censoring distribution where each C_i is independent of each other and of each T_i. Define $Y_i = min(T_i, C_i)$ as the obesrvation time for individual i, and $\delta_i = 1_{[Y_i = T_i]}$ is the indicator that Y_i is an observed failure time (as opposed to other censoring). Derive formulas for the MLE's of $\theta = (\rho, \lambda)$, and find their asymptotic distribution. Then find the MLE of the probability a subject surviving 5 years, and its asymptoic distribution.
This problem mirrors closely to the example using an exponential distribution in class.

$$\begin{aligned}
Weib(\lambda, \rho) &= \rho\lambda^\rho x^{\rho-1} e^{-(\lambda x)^\rho} \\
\Rightarrow L(\lambda, \rho) &= \prod_{i=1}^{n} \left[\rho\lambda^\rho t_i^{\rho-1}\right]^{\delta_i} \left[e^{-(\lambda t_i)^\rho}\right]^{1-\delta_i} \left[1 - F_{C_i}(t_i)\right]^{\delta_i} \left[f_{C_i}(t_i)\right]^{1-\delta_i} \\
\mathcal{L}(\lambda, \rho) &= \sum_{i=1}^{n} \delta_i \left[ln(\rho) + \rho ln(\lambda) + (\rho - 1) ln(t_i)\right] - (\lambda t_i)^\rho + \text{ schtuff} \\
\text{Formulas for MLE's:} \quad \mathcal{U}(\lambda) &= \sum_{i=1}^{n} \frac{\delta_i \rho - \rho(\lambda t_i)^\rho}{\lambda} \\
\mathcal{U}(\rho) &= \sum_{i=1}^{n} \frac{\delta_i}{\rho} + \delta_i \left[ln(\lambda) + ln(t_i)\right] - (\lambda t_i)^\rho ln(\lambda t_i)
\end{aligned}$$

Asymptotic distribution:

$$\sqrt{n}(\hat{\theta} - \theta) \underset{d}{\rightarrow} N(0, I^{-1}(\theta))$$

so we need to calculate $I(\theta)$. Note: Since θ is a vector of length 2, $I(\theta)$ will be a 2x2 matrix.

$$\begin{aligned}
I_{11}(\theta) &= E\left(\sum_{i=1}^{n} \frac{\delta_i}{\rho^2} + (\lambda t_i)^\rho (ln(\lambda t_i))^2\right) \\
I_{12}(\theta) = I_{21}(\theta) &= E\left(\sum_{i=1}^{n} \frac{\delta_i}{\lambda} + \lambda^{\rho-1} t_i^\rho - \rho\lambda^{\rho-1} t_i^\rho ln(\lambda t_i)\right) \\
I_{22}(\theta) &= E\left(\sum_{i=1}^{n} \frac{\delta_i \rho + \rho(\rho - 1)(\lambda t_i)^\rho}{\lambda^2}\right)
\end{aligned}$$

Now we need to find the MLE of the probability of surviving 5 years, and it's asymptotic distribution. We're given the survivor function, and since t is measured in years, we just need to plug-in $t = 5$ (and use the invariance property of MLE's).

$$Pr(T_i > 5) = e^{-(\lambda 5)^\rho}$$

$$\text{MLE} \;\Rightarrow\; e^{-(\hat{\lambda}5)^{\hat{\rho}}}$$

For the asymptotic distribution, use the delta method where $g(\theta) = e^{-(\lambda 5)^{\rho}}$.

$$\begin{aligned}
\nabla g_1(\theta) = \frac{dg(\theta)}{d\rho} &= -e^{-(\lambda 5)^{\rho}}(5\lambda)^{\rho} ln(5\lambda) \\
\nabla g_2(\theta) = \frac{dg(\theta)}{d\lambda} &= -e^{-(\lambda 5)^{\rho}}\rho 5^{\rho}\lambda^{\rho-1}
\end{aligned}$$

Resulting in the asymptotic distribution:

$$\sqrt{n}(e^{-(\hat{\lambda}5)^{\hat{\rho}}} - e^{-(\lambda 5)^{\rho}}) \underset{d}{\rightarrow} N(0, \nabla g'(\theta) I^{-1}(\theta)\nabla g(\theta))$$

- 4.
  Let $\hat{\theta}$ be the maximum likelihood estimate of θ from iid r.v.'s $X_1, ..., X_n$ satisfying regularity conditions. Let $I_1(\theta)$ denote the contribution to Fisher's information from a single observation. Thus we have
  $$\sqrt{n}(\hat{\theta} - \theta) \underset{d}{\longrightarrow} N(0, I_1^{-1}(\theta))$$
  and the MLE attains the CR-LB. Show that the MLE for $g(\theta)$ has an asymptotic distribution which acheives the CR-LB.
  First we need the MLE for $g(\theta)$, which is just $g(\hat{\theta})$ by the invariance property of MLE's.
  Now pile on the δ-method to that and we have:
  $$\begin{aligned}
  \sqrt{n}(g(\hat{\theta}) - g(\theta)) &\underset{d}{\rightarrow} g'(\theta) N(0, I_1^{-1}(\theta)) \\
  &\underset{d}{\rightarrow} N(0, \nabla g'(\theta) I_1^{-1}(\theta)\nabla g(\theta))
  \end{aligned}$$
  Note that the variance in the equation above is exactly that of the CR-LB for an unbiased function of $g(\theta)$. So while in small samples, $g(\hat{\theta})$ won't attain the lower bound, for large samples (asymptotically) it will converge to it. Yeehaw.
  (When θ is of dimension 1, we'd just have $g'(\theta)^2 I_1^{-1}$)

- 5.
  The fish fry problem.

  (a) Suppose a homozygous black male and homozygous red male are placed with a homozygous red female. Let N denote the total number of fry, and X denote the total number that were sired by the black male, and Y denote the total number of black fry (In this instance

$X = Y$, part (b) will be different though). Find an estimate for the probability p that the black male sires a fry and derive the asymptotic distribution for the estimate.

Since the only way for a fry to be black is to be sired by the black male, and every single fry sired by the black male would be black, we can just count the number of black fry and divide by the total number of fry. From the set-up, this would be exactly $\frac{X}{N}$. Now we also know the distribution of this estimate since the random variable is Binomial.

$$\frac{X}{N} \quad \dot{\sim} \quad N\left(p, \frac{p(1-p)}{N}\right)$$

(b) Now suppose the scenario of part (a) but with a hetrozygous black male (hence, $X \neq Y$). Assume that conditional upon the value of X, $Y|X \sim \mathcal{B}(X, 0.5)$. Again find an estimate for p and find the asymptotic distribution for this estimate.

Now again the only way for a fry to be black is to be sired by the black male, but only half of the fry sired by the black male are black. Thus, if we count the number of black fry, we'll still be missing half of the fry that the black male sired (since it's a 50/50 chance when the black male is heterozygous). But then all we need to do is double the number of black fry we see, so our estimate is $\frac{2Y}{N}$. Now we need to figure out what the distribution of Y is. We know by conditional probability that

$$p_Y(y) = p_{Y|X}(y|x) \cdot p_X(x)$$

Then to get rid of the X's, we just sum over all possibilities.

$$\begin{aligned}
p_Y(y) &= \sum_{x=0}^{N} \underbrace{Pr(Y = k|X = x)}_{\mathcal{B}(X,0.5)} \cdot \underbrace{P(X = x)}_{\mathcal{B}(N,\frac{X}{N})} \\
&= \sum_{x=0}^{N} \frac{x!}{(x-k)!x!} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{x-k} 1_{[0,x]}(k) \cdot \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x} \\
&= \sum_{x=0}^{N} \frac{x!}{x!(x-k)!} \left(\frac{1}{2}\right)^x 1_{[k,N]}(x) \cdot \frac{N!}{x!(N-x)!} p^x (1-p)^{N-x} \\
&= \sum_{x=k}^{N} \frac{N!}{(x-k)!k!(N-x)!} \left(\frac{p}{2}\right)^x (1-p)^{N-x}
\end{aligned}$$

Now this is looking sorta like the form of a binomial random variable, so let's try to make it look that way and set $u = x - k$.

$$= \sum_{u=0}^{N-k} \frac{N!}{(u+k-k)!k!(N-u-k)!} \left(\frac{p}{2}\right)^{u+k} (1-p)^{N-u-k}$$

$$= \frac{N!}{(N-k)!k!}\left(\frac{p}{2}\right)^k \sum_{u=0}^{N-k} \frac{N!}{(N-k-u)!u!}\left(\frac{p}{2}\right)^u (1-p)^{N-k-u}$$

Notice the second half of this expression is the Binomial theorem and thus is $(\frac{p}{2}+1-p)^{N-k} = (1-\frac{p}{2})^{N-k}$ and this gives the form of a $\mathcal{B}(N, \frac{p}{2})$ overall for the distribution for Y. Thus for our estimate:

$$\frac{2Y}{N} \;\dot{\sim}\; N\left(p, \frac{\frac{4p}{2}(1-\frac{p}{2})}{N}\right)$$

(c) Find the asymptotic relative efficiency (ARE) of the estimate derived in part (a) compared to that of part (b). Interpret the results in regard to the importance of knowing the genotype of the males.
Using the asymptotic distribution results from parts (a) and (b):

$$\text{ARE} = \frac{\frac{p(1-p)}{N}}{\frac{\frac{4p}{2}(1-\frac{p}{2})}{N}} = \frac{p(1-p)}{p(2-p)} = \frac{1-p}{2-p}$$

The 'importance' of knowing the genotype of the males will depend on p, although, since this ratio is monotone over all possible values of p (0 to 1), we know that the largest value is at p=0 where the ARE is 1/2.