Econ 203C: Systems Models

Problem Set 4

Michael Powell
Department of Economics, UCLA
May 4th, 2006

Question 1:

Consider the asymptotic distribution of the ML estimator derived in Lecture Note 8. Explain in detail whether or not

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)$$

is a consistent estimator for $I\left(\theta_0\right)$.

Solution: Assuming (y_i, x_i) are i.i.d., we have that by the weak law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\xrightarrow{p}E\left[\frac{\partial \ln f\left(Y_i,X_i;\theta_0\right)}{\partial\theta}\right]$$

And

$$\begin{aligned}
E\left[\frac{\partial \ln f\left(Y_i,X_i;\theta_0\right)}{\partial\theta}\right] &= \int\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}f\left(y_i,x_i;\theta_0\right)dx_idy_i\\
&= \int\frac{1}{f\left(y_i,x_i;\theta_0\right)}\frac{\partial f\left(y_i,x_i;\theta_0\right)}{\partial\theta}f\left(y_i,x_i;\theta_0\right)dx_idy_i\\
&= \int\frac{\partial f\left(y_i,x_i;\theta_0\right)}{\partial\theta}dx_idy_i\\
&= \frac{\partial}{\partial\theta}\int f\left(y_i,x_i;\theta_0\right)dx_idy_i\\
&= \frac{\partial}{\partial\theta}1=0
\end{aligned}$$

Where in the third to last step, I appealed to the dominated convergence theorem, which requires some assumptions about uniform upper bounds on the derivative of f (or equivalently, uniform continuity of f).

By Slutsky's theorem, we have that

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)\xrightarrow{p}0\cdot 0=0$$

And therefore,

$$\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(y_i,x_i;\theta_0\right)}{\partial\theta}\right)$$

Is not a consistent estimator for $I\left(\theta_0\right)=E\left[s\left(Y,X;\theta_0\right)s\left(Y,X;\theta_0\right)'\right]$. (Where we define $s\left(Y,X;\theta_0\right)=E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial \ln f\left(Y_i,X_i;\theta_0\right)}{\partial\theta}\right]$.

Question 2: (Question 3 in Greene, page 522)
Suppose that the joint distribution of the two random variables x and y is

$$f(x,y) = \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!},\ \beta > 0,\ \theta > 0,\ y \geq 0,\ x = 0, 1, 2, \ldots$$

1. Find the MLE for β and θ and provide their asymptotic joint distribution.
Solution: The likelihood function for a random sample from this distribution is

$$L(\beta, \theta; X, Y) = \prod_{i=1}^{n} \frac{\theta e^{-(\beta+\theta)Y_i}(\beta Y_i)^{X_i}}{X_i!}$$

Taking logs,

$$\begin{aligned}
\log L(\beta, \theta; X, Y) &= \sum_{i=1}^{n} \left[\log\theta - (\beta+\theta)Y_i + X_i \log(\beta Y_i) - \log X_i!\right] \\
&= n\log\theta - (\beta+\theta)\sum_{i=1}^{n} Y_i + \log\beta \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} X_i \log Y_i + \sum_{i=1}^{n} \log X_i!
\end{aligned}$$

The first order conditions are thus

$$\begin{aligned}
(\theta) &: \frac{n}{\hat{\theta}_{ML}} - \sum_{i=1}^{n} Y_i = 0 \\
(\beta) &: -\sum_{i=1}^{n} Y_i + \frac{1}{\hat{\beta}_{ML}} \sum_{i=1}^{n} X_i = 0
\end{aligned}$$

Which gives us

$$\begin{aligned}
\frac{n}{\hat{\theta}_{ML}} &= \sum_{i=1}^{n} Y_i \\
\hat{\theta}_{ML} &= \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^{-1}
\end{aligned}$$

And

$$\begin{aligned}
\frac{1}{\hat{\beta}_{ML}} \sum_{i=1}^{n} X_i &= \sum_{i=1}^{n} Y_i \\
\hat{\beta}_{ML} &= \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} Y_i}
\end{aligned}$$

Since we know that, under regularity conditions,

$$\sqrt{n}\left(\hat{\delta}_{ML} - \delta_0\right) \xrightarrow{d} N\left(0, [I_1(\delta_0)]^{-1}\right)$$

We need only calculate

$$I_1(\delta_0) = \frac{1}{n} I_n(\delta_0) = -\frac{1}{n} E\left[\frac{\partial \log L(\delta; X, Y)}{\partial\delta\partial\delta'}\right]$$

Where $\delta_0 = \begin{bmatrix} \theta_0 \\ \beta_0 \end{bmatrix}$. The first derivatives of the log likelihood function are

$$\begin{aligned}
\frac{\partial \log L}{\partial\theta} &= \frac{n}{\theta} - \sum_{i=1}^{n} Y_i = 0 \\
\frac{\partial \log L}{\partial\beta} &: -\sum_{i=1}^{n} Y_i + \frac{1}{\beta} \sum_{i=1}^{n} X_i = 0
\end{aligned}$$

The second derivatives are

$$\begin{aligned}
\frac{\partial^2 \log L}{\partial\theta\partial\theta} &= -\frac{n}{\theta^2} \\
\frac{\partial^2 \log L}{\partial\theta\partial\beta} &= \frac{\partial^2 \log L}{\partial\beta\partial\theta} = 0 \\
\frac{\partial^2 \log L}{\partial\beta\partial\beta} &= -\frac{1}{\beta^2}\sum_{i=1}^{n} X_i = 0
\end{aligned}$$

This gives us

$$\begin{aligned}
\frac{\partial^2 \log L}{\partial\delta\partial\delta'} &= \begin{bmatrix} \frac{\partial^2 \log L}{\partial\theta\partial\theta} & \frac{\partial^2 \log L}{\partial\theta\partial\beta} \\ \frac{\partial^2 \log L}{\partial\beta\partial\theta} & \frac{\partial^2 \log L}{\partial\beta\partial\beta} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{n}{\theta^2} & 0 \\ 0 & -\frac{1}{\beta^2}\sum_{i=1}^{n} X_i \end{bmatrix}
\end{aligned}$$

And

$$I_n(\delta_0) = -E\left[\left.\frac{\partial^2 \log L}{\partial\delta\partial\delta'}\right|_{\delta_0} \right] = \begin{bmatrix} \frac{n}{\theta_0^2} & 0 \\ 0 & \frac{n}{\beta_0^2} E[X_i] \end{bmatrix}$$

Finally, we have that

$$[I_1(\delta_0)]^{-1} = \left[\frac{1}{n} I_n(\delta_0)\right]^{-1} = \begin{bmatrix} \frac{1}{\theta_0^2} & 0 \\ 0 & \frac{1}{\beta_0^2} E[X_i] \end{bmatrix}^{-1} = \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix}$$

Thus,

$$\sqrt{n}\left(\hat{\delta}_{ML} - \delta_0\right) \xrightarrow{d} N\left(0, \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix}\right)$$

2. Find the MLE for $\frac{\theta}{\beta+\theta}$ and its asymptotic distribution.

Solution: By the invariance principle of MLEs, $\widehat{h(\delta)}_{ML} = h\left(\hat{\delta}_{ML}\right)$. In this particular case,

$$\begin{aligned}
\widehat{\frac{\theta}{\beta+\theta}}_{ML} &= \frac{\hat{\theta}_{ML}}{\hat{\beta}_{ML} + \hat{\theta}_{ML}} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^{-1}}{\frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} Y_i} + \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^{-1}} \\
&= \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1} = \frac{n}{\sum_{i=1}^{n} X_i + n}
\end{aligned}$$

Since from above, we know that $\sqrt{n}\left(\hat{\delta}_{ML} - \delta_0\right) \xrightarrow{d} N\left(0, \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix}\right)$, by the delta method, we have that

$$\sqrt{n}\left(h\left(\hat{\delta}_{ML}\right) - h(\delta_0)\right) \xrightarrow{d} N\left(0, \left.\begin{bmatrix} \frac{\partial h}{\partial\theta} & \frac{\partial h}{\partial\beta} \end{bmatrix}\right|_{\delta_0} \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix} \left.\begin{bmatrix} \frac{\partial h}{\partial\theta} \\ \frac{\partial h}{\partial\beta} \end{bmatrix}\right|_{\delta_0}\right)$$

Calculating the derivatives,

$$\begin{aligned}
\frac{\partial h}{\partial\theta} &= \frac{\partial}{\partial\theta}\left[\frac{\theta}{\beta+\theta}\right] = \frac{(\beta+\theta)-\theta}{(\beta+\theta)^2} = \frac{\beta}{(\beta+\theta)^2} \\
\frac{\partial h}{\partial\beta} &= \frac{\partial}{\partial\beta}\left[\frac{\theta}{\beta+\theta}\right] = \frac{-\theta}{(\beta+\theta)^2}
\end{aligned}$$

Thus, the asymptotic covariance matrix is

$$\begin{aligned}
& \begin{bmatrix} \frac{\beta_0}{(\beta_0+\theta_0)^2} & -\frac{\theta_0}{(\beta_0+\theta_0)^2} \end{bmatrix} \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix} \begin{bmatrix} \frac{\beta_0}{(\beta_0+\theta_0)^2} \\ -\frac{\theta_0}{(\beta_0+\theta_0)^2} \end{bmatrix} \\
= & \frac{\beta_0^2}{(\beta_0+\theta_0)^4}\theta_0^2 + \frac{\theta_0^2}{(\beta_0+\theta_0)^4}\frac{\beta_0^2}{E[X_i]} \\
= & \frac{\beta_0^2\theta_0^2(1+E[X_i])}{(\beta_0+\theta_0)^4 E[X_i]}
\end{aligned}$$

3. Prove that $f(x)$ is of the form

$$f(x) = \gamma(1-\gamma)^x, \; x = 0, 1, 2, \ldots$$

and find the ML estimator for γ and its asymptotic distribution.

Solution: Recall that $f(x,y) = \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}$. Integrating with respect to y,

$$\int_0^\infty \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}dy = \theta\frac{\beta^x}{x!}\int_0^\infty e^{-(\beta+\theta)y}y^x dy$$

Let

$$\begin{aligned}
u &= y^x \\
du &= xy^{x-1}dy \\
dv &= e^{-(\beta+\theta)y}dy \\
v &= -\frac{1}{\beta+\theta}e^{-(\beta+\theta)y}
\end{aligned}$$

By integration by parts,

$$\begin{aligned}
\int_0^\infty e^{-(\beta+\theta)y}y^x dy &= -\frac{1}{\beta+\theta}e^{-(\beta+\theta)y}y^x\Big|_0^\infty + \frac{x}{\beta+\theta}\int_0^\infty e^{-(\beta+\theta)y}y^{x-1}dy \\
&= \frac{x}{\beta+\theta}\int_0^\infty e^{-(\beta+\theta)y}y^{x-1}dy
\end{aligned}$$

Proceeding iteratively,

$$\begin{aligned}
\int_0^\infty e^{-(\beta+\theta)y}y^x dy &= \frac{x\cdot(x-1)\cdot\cdots\cdot 2\cdot 1}{(\beta+\theta)^x}\int_0^\infty e^{-(\beta+\theta)\cdot y}dy \\
&= \frac{x\cdot(x-1)\cdot\cdots\cdot 2\cdot 1}{(\beta+\theta)^x}\cdot\frac{1}{\beta+\theta}
\end{aligned}$$

This gives us

$$\begin{aligned}
f(x) &= \int_0^\infty \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}dy \\
&= \theta\frac{\beta^x}{x!}\int_0^\infty e^{-(\beta+\theta)y}y^x dy \\
&= \frac{\theta}{\beta+\theta}\left(\frac{\beta}{\beta+\theta}\right)^x\frac{x!}{x!} \\
&= \frac{\theta}{\beta+\theta}\left[1-\frac{\theta}{\beta+\theta}\right]^x
\end{aligned}$$

Define $\gamma = \frac{\theta}{\beta+\theta}$. Then $f(x) = \gamma(1-\gamma)^x$.

The likelihood function for a random sample is

$$\begin{aligned} L(\gamma; X) &= \prod_{i=1}^{n} \gamma (1-\gamma)^{X_i} \\ &= \gamma^n (1-\gamma)^{\sum_{i=1}^{n} X_i} \end{aligned}$$

Taking logs

$$\log L(\gamma; X) = n \log \gamma + \sum_{i=1}^{n} X_i \log (1-\gamma)$$

Taking first order conditions,

$$(\gamma) : \frac{n}{\hat{\gamma}_{ML}} - \frac{1}{1-\hat{\gamma}_{ML}} \sum_{i=1}^{n} X_i = 0$$

Or

$$\begin{aligned} \frac{n}{\hat{\gamma}_{ML}} &= \frac{\sum_{i=1}^{n} X_i}{1-\hat{\gamma}_{ML}} \\ n - n\hat{\gamma}_{ML} &= \hat{\gamma}_{ML} \sum_{i=1}^{n} X_i \\ \hat{\gamma}_{ML} \left(\sum_{i=1}^{n} X_i + n \right) &= n \\ \hat{\gamma}_{ML} &= \frac{n}{\sum_{i=1}^{n} X_i + n} \end{aligned}$$

Which is exactly what we would have expected from our result in part (2). Also, as in part (2), we know that

$$\sqrt{n} (\hat{\gamma}_{ML} - \gamma_0) \xrightarrow{d} N \left(0, \frac{\beta_0^2 \theta_0^2 (1 + E[X_i])}{(\beta_0 + \theta_0)^4 E[X_i]} \right)$$

But what is $E[X_i]$?

$$\begin{aligned} E[X_i] &= \sum_{k=0}^{\infty} k \cdot \gamma (1-\gamma)^k \\ &= \frac{1-\gamma}{\gamma} \quad \text{(Using Mathematica)} \\ &= \frac{1 - \frac{\theta}{\beta+\theta}}{\frac{\theta}{\beta+\theta}} = \frac{\beta}{\theta} \end{aligned}$$

Thus, the asymptotic variance is

$$\begin{aligned} \frac{\beta_0^2 \theta_0^2 \left(1 + \frac{\beta_0}{\theta_0} \right)}{(\beta_0 + \theta_0)^4 \frac{\beta_0}{\theta_0}} &= \frac{\beta_0^2 \theta_0^2 \left(\frac{\beta_0 + \theta_0}{\theta_0} \right)}{(\beta_0 + \theta_0)^4 \frac{\beta_0}{\theta_0}} \\ &= \frac{\beta_0 \theta_0^2}{(\beta_0 + \theta_0)^3} \end{aligned}$$

4. Prove that

$$f(y|x) = \frac{\lambda e^{-\lambda y} (\lambda y)^x}{x!}, \; y \geq 0, \; \lambda > 0.$$

Show that $f(y|x)$ integrates to 1. Find the MLE for λ and its asymptotic distribution.

Solution: By the definition of conditional probability,

$$\begin{aligned} f\left(y|\,x\right) &= \frac{f\left(x,y\right)}{f\left(x\right)} = \frac{\frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}}{\frac{\theta}{\beta+\theta}\left[\frac{\beta}{\beta+\theta}\right]^x} \\ &= \frac{\frac{e^{-(\beta+\theta)y}(y)^x}{x!}}{\frac{1}{\beta+\theta}\left[\frac{1}{\beta+\theta}\right]^x} \\ &= \frac{\left(\beta+\theta\right)e^{-(\beta+\theta)y}\left(\left(\beta+\theta\right)y\right)^x}{x!} \end{aligned}$$

Define $\lambda = \beta + \theta$. Then,

$$f\left(y|\,x\right) = \frac{\lambda e^{-\lambda y}\left(\lambda y\right)^x}{x!}$$

Integrating,

$$\int_0^\infty \frac{\lambda e^{-\lambda y}\left(\lambda y\right)^x}{x!}dy = \frac{\lambda^{x+1}}{x!}\int_0^\infty e^{-\lambda y}y^x dy$$

Let

$$\begin{aligned} u &= y^x \\ du &= xy^{x-1}dy \\ dv &= e^{-\lambda y}dy \\ v &= -\frac{1}{\lambda}e^{-\lambda y}dy \end{aligned}$$

Then, by integration by parts,

$$\begin{aligned} \int_0^\infty e^{-\lambda y}y^x dy &= \left[-\frac{1}{\lambda}e^{-\lambda y}y^x\right]_0^\infty + \frac{x}{\lambda}\int_0^\infty e^{-\lambda y}y^{x-1}dy \\ &= \frac{x}{\lambda}\int_0^\infty e^{-\lambda y}y^{x-1}dy \end{aligned}$$

Proceeding iteratively,

$$\begin{aligned} \int_0^\infty e^{-\lambda y}y^x dy &= \frac{x\cdot(x-1)\cdot\cdots\cdot 2\cdot 1}{\lambda^x}\int_0^\infty e^{-\lambda y}dy \\ &= \frac{x!}{\lambda^{x+1}} \end{aligned}$$

Thus,

$$\int_0^\infty \frac{\lambda e^{-\lambda y}\left(\lambda y\right)^x}{x!}dy = \frac{\lambda^{x+1}}{x!}\int_0^\infty e^{-\lambda y}y^x dy = \frac{\lambda^{x+1}}{x!}\frac{x!}{\lambda^{x+1}} = 1$$

In order to find the ML estimator and its asymptotic distribution, I will make use the invariance principle and delta method. Let $\lambda = h\left(\theta,\beta\right) = \beta + \theta$. Then, by the invariance principle,

$$\begin{aligned} \hat{\lambda}_{ML} &= \hat{\beta}_{ML} + \hat{\theta}_{ML} \\ &= \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n Y_i} + \frac{n}{\sum_{i=1}^n Y_i} = \frac{\sum_{i=1}^n X_i + n}{\sum_{i=1}^n Y_i} \end{aligned}$$

The asymptotic distribution will be given by

$$\sqrt{n}\left(\hat{\lambda}_{ML} - \lambda_0\right) \xrightarrow{n} N\left(0, \left[\begin{array}{cc} \frac{\partial h}{\partial\theta} & \frac{\partial h}{\partial\beta}\end{array}\right]\left[\begin{array}{cc} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]}\end{array}\right]\left[\begin{array}{c} \frac{\partial h}{\partial\theta} \\ \frac{\partial h}{\partial\beta}\end{array}\right]\right)$$

The asymptotic covariance is thus

$$\begin{aligned}
\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \theta_0^2 & 0 \\ 0 & \frac{\beta_0^2}{E[X_i]} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \theta_0^2 + \frac{\beta_0^2}{E[X_i]} \\
&= \theta_0^2 + \frac{\beta_0^2}{\frac{\beta_0}{\theta_0}} \\
&= \theta_0^2 + \theta_0 \beta_0 \\
&= \theta_0 (\theta_0 + \beta_0).
\end{aligned}$$

This gives us

$$\sqrt{n}\left(\hat{\lambda}_{ML} - \lambda_0\right) \xrightarrow{d} N(0, \theta_0(\theta_0 + \beta_0)).$$

5. Prove that

$$f(y) = \theta e^{-\theta y}, \ y \geq 0, \ \theta > 0.$$

Find the MLE for θ and its asymptotic distribution.

Solution: We are given that $f(x,y) = \frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}$. Recall that the Taylor series expansion of $e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!}$. Here, let $a = \beta y$. Then $\sum_{k=0}^{n} \frac{(\beta y)^k}{k!} = e^{\beta y}$, which gives us

$$\begin{aligned}
f(y) &= \sum_{k=0}^{\infty} \frac{\theta e^{-(\beta+\theta)y} (\beta y)^x}{x!} \\
&= \theta e^{-(\beta+\theta)y} \sum_{k=0}^{\infty} \frac{(\beta y)^k}{k!} \\
&= \theta e^{-(\beta+\theta)y} e^{\beta y} \\
&= \theta e^{-\theta y}
\end{aligned}$$

Which is the desired result. We have already computed the MLE for θ, but I will do so again in this context. Let the likelihood function for a random sample of Y_i be

$$\begin{aligned}
L(\theta; Y) &= \prod_{i=1}^{n} \theta e^{-\theta Y_i} \\
&= \theta^n e^{-\theta \sum_{i=1}^{n} Y_i}
\end{aligned}$$

Taking logs,

$$\log L(\theta; Y) = n \log \theta - \theta \sum_{i=1}^{n} Y_i$$

Taking first order conditions,

$$(\theta): \frac{n}{\hat{\theta}_{ML}} - \sum_{i=1}^{n} Y_i = 0$$

Or

$$\hat{\theta}_{ML} = \left(\frac{1}{n} \sum_{i=1}^{n} Y_i\right)^{-1}$$

I will go about computing the asymptotic distribution slightly differently than above. (Though we could easily apply the delta method again with $h(\theta, \beta) = \theta$.) Computing the Fisher information for $\hat{\theta}_{ML}$,

$$\begin{aligned}
\frac{\partial \log L}{\partial \theta} &= \frac{n}{\theta} - \sum_{i=1}^{n} Y_i \\
\frac{\partial^2 \log L}{\partial \theta \partial \theta} &= -\frac{n}{\theta^2}
\end{aligned}$$

Thus,

$$\begin{aligned} I_n(\theta_0) &= -E\left[\frac{\partial^2 \log L}{\partial\theta\partial\theta}\right] = \frac{n}{\theta_0^2} \\ I_1(\theta_0) &= \frac{1}{\theta_0^2} \end{aligned}$$

Therefore,

$$\sqrt{n}\left(\hat{\theta}_{ML} - \theta_0\right) \xrightarrow{d} N\left(0, \theta_0^2\right)$$

6. Prove that

$$f(x|\, y) = \frac{e^{-\beta y}(\beta y)^x}{x!}, \;\; \beta > 0, \;\; x = 0, 1, 2, \ldots$$

Find the MLE for β and its asymptotic distribution.

Solution: Here, we have

$$\begin{aligned} f(x|\, y) &= \frac{f(x,y)}{f(y)} = \frac{\frac{\theta e^{-(\beta+\theta)y}(\beta y)^x}{x!}}{\theta e^{-\theta y}} \\ &= \frac{e^{-\beta y}(\beta y)^x}{x!} \end{aligned}$$

Which is the desired result. Computing the unconditional MLE for β, which we have already done, construct the likelihood function for a random sample of X_i, Y_i.

$$L(\beta; X, Y) = \prod_{i=1}^{n} \frac{e^{-\beta Y_i}(\beta Y_i)^{X_i}}{X_i!}$$

Taking logs

$$\begin{aligned} \log L(\beta; X, Y) &= \sum_{i=1}^{n} \left[-\beta Y_i + X_i \log(\beta Y_i) - \log X_i!\right] \\ &= -\beta \sum_{i=1}^{n} Y_i + \log\beta \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} X_i \log Y_i - \sum_{i=1}^{n} \log X_i! \end{aligned}$$

The FOCs are:

$$(\beta): -\sum_{i=1}^{n} Y_i + \frac{1}{\hat{\beta}_{ML}} \sum_{i=1}^{n} X_i = 0$$

Which gives us

$$\begin{aligned} \frac{1}{\hat{\beta}_{ML}} \sum_{i=1}^{n} X_i &= \sum_{i=1}^{n} Y_i \\ \hat{\beta}_{ML} &= \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} Y_i} \end{aligned}$$

In order to compute the asymptotic variance, we will need to compute the derivatives

$$\begin{aligned} \frac{\partial \log L}{\partial\beta} &= -\sum_{i=1}^{n} Y_i + \frac{1}{\beta}\sum_{i=1}^{n} X_i \\ \frac{\partial^2 \log L}{\partial\beta\partial\beta} &= -\frac{1}{\beta^2}\sum_{i=1}^{n} X_i \end{aligned}$$

Which gives us

$$\begin{aligned} I_n(\beta_0) &= -E\left[\frac{\partial^2 \log L}{\partial\beta\partial\beta}\right] \\ &= \frac{E[X_i]}{\beta_0^2} = \frac{\frac{\beta_0}{\theta_0}}{\beta_0^2} = \frac{1}{\beta_0\theta_0} \end{aligned}$$

And

$$I_1(\beta_0) = \frac{1}{n\beta_0\theta_0}$$

And therefore,

$$\sqrt{n}\left(\hat{\beta}_{ML} - \beta_0\right) \xrightarrow{d} N(0, n\beta_0\theta_0)$$

Alternatively, in order to compute the conditional MLE for $\beta\left(\tilde{Y}\right)$, assume that we have a random sample of $\{X_i\}$ holding $\tilde{Y}$ fixed. The likelihood function is then

$$L\left(\beta|\{X_i\},\tilde{Y}\right) = \prod_{i=1}^{n} f\left(X_i|\tilde{Y}\right) = \prod_{i=1}^{n} \frac{e^{-\beta\tilde{Y}}\left(\beta\tilde{Y}\right)^{X_i}}{X_i!}$$

Taking logs

$$\begin{aligned} \log L &= \sum_{i=1}^{n}\left[-\beta\tilde{Y} + X_i \log\left(\beta\tilde{Y}\right) - \log X_i!\right] \\ &= -n\beta\tilde{Y} + \log\beta\sum_{i=1}^{n} X_i + \sum_{i=1}^{n} X_i \log\tilde{Y} - \sum_{i=1}^{n}\log X_i! \end{aligned}$$

The FOC is

$$(\beta): -n\tilde{Y} + \frac{1}{\hat{\beta}_{ML}\left(\tilde{Y}\right)}\sum_{i=1}^{n} X_i = 0$$

Or

$$\begin{aligned} \frac{1}{\hat{\beta}_{ML}\left(\tilde{Y}\right)}\sum_{i=1}^{n} X_i &= n\tilde{Y} \\ \hat{\beta}_{ML}\left(\tilde{Y}\right) &= \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\tilde{Y}} \end{aligned}$$

In order to compute the asymptotic variance, we will need to compute the derivatives

$$\begin{aligned} \frac{\partial \log L}{\partial\beta} &= -n\tilde{Y} + \frac{1}{\beta\left(\tilde{Y}\right)}\sum_{i=1}^{n} X_i \\ \frac{\partial^2 \log L}{\partial\beta\partial\beta} &= -\frac{1}{\left[\beta\left(\tilde{Y}\right)\right]^2}\sum_{i=1}^{n} X_i \end{aligned}$$

Which gives us

$$\begin{aligned} I_n\left(\beta_0\left(\tilde{Y}\right)\right) &= -E\left[\frac{\partial^2 \log L}{\partial\beta\partial\beta}\right] \\ &= \frac{nE\left[X_i|\tilde{Y}\right]}{\left[\beta_0\left(\tilde{Y}\right)\right]^2} \end{aligned}$$

But since $X_i|\tilde{Y} \sim Poisson\left(\beta_0\left(\tilde{Y}\right)\tilde{Y}\right)$

$$E\left[X_i|\tilde{Y}\right] = \beta_0\left(\tilde{Y}\right)\tilde{Y}$$

$$I_n\left(\beta_0\left(\tilde{Y}\right)\right) = \frac{n\beta_0\left(\tilde{Y}\right)\tilde{Y}}{\left[\beta_0\left(\tilde{Y}\right)\right]^2} = \frac{n\tilde{Y}}{\beta_0\left(\tilde{Y}\right)}$$

And

$$I_1\left(\beta_0\left(\tilde{Y}\right)\right) = \frac{\tilde{Y}}{\beta_0\left(\tilde{Y}\right)}$$

And therefore,

$$\sqrt{n}\left(\hat{\beta}_{ML}\left(\tilde{Y}\right) - \beta_0\left(\tilde{Y}\right)\right) \xrightarrow{d} N\left(0, \frac{\beta_0\left(\tilde{Y}\right)}{\tilde{Y}}\right)$$

Question 3:
Consider the binary probit model given by

$$\Pr\left(Y_i = 1 | X_i = x_i; \gamma\right) = \Phi\left(x_i'\gamma\right), \ i = 1, \ldots, n,$$

where $\Phi(\cdot)$ denotes the cdf of a standard normal random variable.
1. Define the population parameter vector γ_0.
Solution: The population parameter vector γ_0 is the vector which maximizes the normalized expected log likelihood function:

$$\gamma_0 = \arg\max E\left[\frac{1}{n}\log L(\gamma)\right]$$

This will become more apparent in part (2) where we derive the sample analog

$$\hat{\gamma}_n = \arg\max \frac{1}{n}\log L(\gamma)$$

By the laws of large numbers (under appropriate moment assumptions),

$$\frac{1}{n}\log L(\gamma) \xrightarrow{p} E\left[\frac{1}{n}\log L(\gamma)\right]$$

If we assume that our data is i.i.d., then

$$\frac{1}{n}\log L(\gamma) \xrightarrow{p} E\left[\log f_{Y_i}\left(y_i | X_i = x_i, \gamma\right)\right]$$

And we would have that

$$\gamma_0 = \arg\max_{\gamma} E\left[\log f_{Y_i}\left(y_i | X_i = x_i, \gamma\right)\right]$$

2. Define the sample log-likelihood function $L(\gamma)$ and the first-order conditions for the MLE, say $\hat{\gamma}_n$.
Solution: First, note that the conditional pdf for Y_i is

$$f_{Y_i}\left(y_i | X_i = x_i, \gamma\right) = \left[\Phi\left(x_i'\gamma\right)\right]^{y_i}\left[1 - \Phi\left(x_i'\gamma\right)\right]^{1-y_i}$$

Assuming that $\{Y_i | X_i = x_i\}$ are independent., then our likelihood function is

$$\begin{aligned} L\left(\gamma | \{Y_i\}, \{X_i\}\right) &= \prod_{i=1}^{n} f_{Y_i}\left(Y_i | X_i, \gamma\right) \\ &= \prod_{i=1}^{n} \left[\Phi\left(X_i'\gamma\right)\right]^{Y_i}\left[1 - \Phi\left(X_i'\gamma\right)\right]^{1-Y_i} \end{aligned}$$

Taking logs,

$$\log L = \sum_{i=1}^{n}\left[Y_i \log\left(\Phi\left(X_i'\gamma\right)\right) + \left(1 - Y_i\right)\log\left(1 - \Phi\left(X_i'\gamma\right)\right)\right]$$

We thus have that

$$\begin{aligned} \hat{\gamma}_n &= \arg\max \frac{1}{n}\log L \\ &= \arg\max \frac{1}{n}\sum_{i=1}^{n}\left[Y_i \log\left(\Phi\left(X_i'\gamma\right)\right) + \left(1 - Y_i\right)\log\left(1 - \Phi\left(X_i'\gamma\right)\right)\right] \end{aligned}$$

Taking FOCs

$$(\gamma) : \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i}{\Phi\left(X_i'\hat{\gamma}_n\right)}\phi\left(X_i'\hat{\gamma}_n\right)X_i + \frac{1}{n}\sum_{i=1}^{n}\frac{1 - Y_i}{1 - \Phi\left(X_i\hat{\gamma}_n\right)}\left(-\phi\left(X_i'\hat{\gamma}_n\right)\right)X_i = 0$$

Where ϕ is the pdf for a standard normal random variable.

3. Find the asymptotic distribution of the MLE, that is, show that

$$\sqrt{n}\left(\hat{\gamma}_n - \gamma_0\right) \xrightarrow{d} N\left(0, \Lambda_0\right)$$

and provide the exact formula for Λ_0.
Solution: By the FOCs, we have that

$$0 = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \log f_{Y_i}\left(Y_i | X_i, \hat{\gamma}_n\right)}{\partial \gamma} = \frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \hat{\gamma}_n\right)$$

By the mean value theorem, (or more precisely, a first order Taylor approximation around the population parameter γ_0)

$$0 = \frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \hat{\gamma}_n\right) = \frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \gamma_0\right) + \frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right) \cdot \left(\hat{\gamma}_n - \gamma_0\right)$$

Where $H\left(W_i, \tilde{\gamma}\right) = \frac{\partial S(W_i, \tilde{\gamma})}{\partial \gamma'}$ and $\tilde{\gamma} \in |\hat{\gamma}_n, \gamma_0|$. Assuming $\frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right)$ is nonsingular,

$$\begin{aligned}
-\frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right) \cdot \left(\hat{\gamma}_n - \gamma_0\right) &= \frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \gamma_0\right) \\
\left(\hat{\gamma}_n - \gamma_0\right) &= \left[-\frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right)\right]^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \gamma_0\right)\right] \\
\sqrt{n}\left(\hat{\gamma}_n - \gamma_0\right) &= \left[-\frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right)\right]^{-1} \sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \gamma_0\right)\right]
\end{aligned}$$

If we assume the appropriate moment conditions, since $E\left[s\left(W_i, \gamma_0\right)\right] = 0$, we have that by the central limit theorem,

$$\sqrt{n}\left[\frac{1}{n}\sum_{i=1}^{n} s\left(W_i, \gamma_0\right)\right] \xrightarrow{d} N\left(0, V\left[s\left(W_i, \gamma_0\right)\right]\right) \stackrel{d}{=} N\left(0, E\left[s\left(W_i, \gamma_0\right) s\left(W_i, \gamma_0\right)'\right]\right)$$

Also, assuming appropriate moment conditions, we have that, by the uniform law of large numbers and the Mann-Wald theorem,

$$\left[-\frac{1}{n}\sum_{i=1}^{n} H\left(W_i, \tilde{\gamma}\right)\right]^{-1} \xrightarrow{p} E\left[-H\left(W_i, \gamma_0\right)\right]^{-1}$$

By Slutsky's theorem, then, we have that

$$\sqrt{n}\left(\hat{\gamma}_n - \gamma_0\right) \xrightarrow{d} N\left(0, \Lambda_0\right)$$

Where by symmetry of Hessian matrices,

$$\Lambda_0 = E\left[-H\left(W_i, \gamma_0\right)\right]^{-1} E\left[s\left(W_i, \gamma_0\right) s\left(W_i, \gamma_0\right)'\right] E\left[-H\left(W_i, \gamma_0\right)\right]^{-1}$$

Recognizing that, by the information equality, which I will establish in part (4),

$$E\left[-H\left(W_i, \gamma_0\right)\right] = E\left[s\left(W_i, \gamma_0\right) s\left(W_i, \gamma_0\right)'\right]$$

We then have that

$$\begin{aligned}
\Lambda_0 &= E\left[-H\left(W_i, \gamma_0\right)\right]^{-1} \\
&= E\left[s\left(W_i, \gamma_0\right) s\left(W_i, \gamma_0\right)'\right]^{-1}
\end{aligned}$$

4. Provide two alternative estimators for the asymptotic covariance matrix, one that is based on the Hessian matrix and one that is based on the outer product gradient (OPG). For each of these two estimators, show that it is a consistent estimator for the asymptotic covariance matrix derived in (3). State clearly all the assumptions you make in order to show consistency of the estimators.

Solution: First, in order to prove the information equality (we have to do this at least once per quarter, right?), recognize that

$$1 = \int f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) dy_i \tag{1}$$

Since

$$\begin{aligned}
\frac{\partial}{\partial \gamma} \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) &= \frac{1}{f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)} \frac{\partial}{\partial \gamma} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) \\
\frac{\partial}{\partial \gamma} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) &= \left[\frac{\partial}{\partial \gamma} \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)\right] f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)
\end{aligned}$$

If we differentiate both sides of (1) with respect to γ and assume that we have uniform continuity of $f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)$ (so that we may interchange the order of differentiation and integration by the dominated convergence theorem)

$$\begin{aligned}
0 &= \int \frac{\partial}{\partial \gamma} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) dy_i \\
&= \int \left[\frac{\partial}{\partial \gamma} \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)\right] f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) dy_i
\end{aligned} \tag{2}$$

Differentiating both sides of (2) with respect to γ' (and once again assuming that we can interchange the order of differentiation and integration), we have

$$0 = \int \left[\frac{\partial^2 \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma \partial \gamma'} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) + \frac{\partial \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma} \frac{\partial}{\partial \gamma'} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)\right] dy_i$$

Or

$$\begin{aligned}
\int -\frac{\partial^2 \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma \partial \gamma'} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) &= \int \frac{\partial \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma} \frac{\partial}{\partial \gamma'} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) dy_i \\
&= \int \frac{\partial \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma} \frac{\partial \log f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right)}{\partial \gamma'} f_{Y_i}\left(y_i \middle| X_i, \gamma_0\right) dy_i
\end{aligned}$$

Which gives us the information equality

$$E\left[-H\left(W_i, \gamma_0\right)\right] = E\left[s\left(W_i, \gamma_0\right) s\left(W_i, \gamma_0\right)'\right] \equiv \Lambda_0$$

I claim that the following two estimators are both consistent for Λ_0 :

$$\begin{aligned}
\hat{\Lambda}_n^{-1} &= \frac{1}{n} \sum_{i=1}^{n} \left[-\frac{\partial^2 \log f_{Y_i}\left(Y_i \middle| X_i, \hat{\gamma}_n\right)}{\partial \gamma \partial \gamma'}\right] \\
\tilde{\Lambda}_n^{-1} &= \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial \log f_{Y_i}\left(Y_i \middle| X_i, \hat{\gamma}_n\right)}{\partial \gamma} \frac{\partial \log f_{Y_i}\left(Y_i \middle| X_i, \hat{\gamma}_n\right)}{\partial \gamma'}\right]
\end{aligned}$$

If we assume that $\{Y_i\}$ are independent, Λ_0 is nonsingular, and

$$\begin{aligned}
E\left[\sup_{\gamma \in \Gamma} \left\| -\frac{\partial^2 \log f_{Y_i}\left(Y_i \middle| X_i, \gamma\right)}{\partial \gamma \partial \gamma'} \right\|\right] &< +\infty \\
E\left[\sup_{\gamma \in \Gamma} \left\| \frac{\partial \log f_{Y_i}\left(Y_i \middle| X_i, \gamma_0\right)}{\partial \gamma} \frac{\partial \log f_{Y_i}\left(Y_i \middle| X_i, \gamma_0\right)}{\partial \gamma'} \right\|\right] &< +\infty
\end{aligned}$$

then by the uniform law of large numbers, we will have that, since $\hat{\gamma}_n \xrightarrow{p} \gamma_0$ (from part (c)),

$$\hat{\Lambda}_n^{-1} \xrightarrow{p} E\left[-\frac{\partial^2 \log f_{Y_i}\left(y_i | X_i, \gamma_0\right)}{\partial\gamma\partial\gamma'}\right]^{-1} = \Lambda_0^{-1}$$

and

$$\tilde{\Lambda}_n^{-1} \xrightarrow{p} E\left[\frac{\partial \log f_{Y_i}\left(y_i | X_i, \gamma_0\right)}{\partial\gamma}\frac{\partial \log f_{Y_i}\left(y_i | X_i, \gamma_0\right)}{\partial\gamma'}\right]^{-1} = \Lambda_0^{-1}.$$

Therefore by the Mann-Wald theorem, $\hat{\Lambda}_n$ is a consistent estimator for Λ_0 based on the Hessian and $\tilde{\Lambda}_n$ is a consistent estimator for Λ_0 based on the outer product gradient. (OPG)

Question 4:

Consider the data provided in ps4q4.xls provided on the class website. In this data file you are provided with data on y, x_1, x_2, x_3, and x_4, where y_i, $(i = 1, \ldots, n)$ takes on two possible values 0 (failure) or 1 (success), according to the following model

$$y_i = I\left(\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i > 0\right),$$

where $\varepsilon_i | x_i \sim N(0, 1)$, $x_i = (x_{1i}, x_{2i}, x_{3i}, x_{4i})'$, and $I(\cdot)$ indicates the usual indicator function. The data are also stored in the matlab file ps4q4.mat which contains a 500×5 matrix with the variables y, x_1, x_2, x_3, and x_4 in the five columns, respectively.

1. Provide the MLE for $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$, say $\hat{\beta}_n$.

Solution: Here, we have that

$$\begin{aligned}
\Pr\left[Y_i = 1 | X_i\right] &= \Pr\left(\beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \varepsilon_i > 0 | X_i\right) \\
&= \Pr\left(\varepsilon_i > -\left(\beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i}\right) | X_i\right) \\
&= \Pr\left(\varepsilon_i \leq \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} | X_i\right) \\
&= \Phi\left(\beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i}\right) \\
&= \Phi\left(X_i'\beta\right)
\end{aligned}$$

Where

$$X_i = \begin{bmatrix} 1 \\ X_{2i} \\ X_{3i} \\ X_{4i} \end{bmatrix}, \text{ and } \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix}.$$

Thus, this is just a standard probit model. Using STATA to estimate $\hat{\beta}_{ML}$, we have

$$\begin{aligned}
\hat{\beta}_1^{ML} &= 0.9217 \\
\hat{\beta}_2^{ML} &= 1.0207 \\
\hat{\beta}_3^{ML} &= -0.9670 \\
\hat{\beta}_4^{ML} &= 0.4485
\end{aligned}$$

2. Provide a consistent estimate for the asymptotic covariance matrix of the estimator derived in (1). Provide also the sample standard errors for $\hat{\beta}_n$.

Solution: Since $\hat{\beta}_n$ is a MLE, its asymptotic covariance matrix will be given by Λ_0 where

$$\Lambda_0 = E\left[s\left(W_i, \beta_0\right) s\left(W_i, \beta_0\right)'\right]^{-1}$$

A consistent estimator for Λ_0 as shown in part (4) of question 3 is $\tilde{\Lambda}_n$ where

$$\tilde{\Lambda}_n^{-1} = \frac{1}{n} \sum_{i=1}^{n} \left[s\left(W_i, \hat{\beta}_n\right) s\left(W_i, \hat{\beta}_n\right)'\right]$$

And

$$\begin{aligned}
s\left(W_i, \hat{\beta}_n\right) &= \frac{\partial \log f_{Y_i}\left(Y_i | X_i, \hat{\beta}_n\right)}{\partial \beta} \\
&= \frac{Y_i}{\Phi\left(X_i'\hat{\beta}_n\right)} \phi\left(X_i'\hat{\beta}_n\right) X_i - \frac{1 - Y_i}{1 - \Phi\left(X_i\hat{\beta}_n\right)} \phi\left(X_i'\hat{\beta}_n\right) X_i
\end{aligned}$$

Estimating this using STATA, we have

$$\tilde{\Lambda}_n = \begin{bmatrix} 0.0488 & 0.0020 & -0.0130 & -0.0028 \\ 0.0020 & 0.0120 & -0.0082 & 0.0035 \\ -0.0130 & -0.0082 & 0.0107 & -0.0040 \\ -0.0028 & 0.0035 & -0.0040 & 0.0059 \end{bmatrix}$$

The standard errors are thus

$$\begin{aligned} se\left(\hat{\beta}_1^{ML}\right) &= \sqrt{0.0488} = 0.2209 \\ se\left(\hat{\beta}_2^{ML}\right) &= \sqrt{0.0120} = 0.1095 \\ se\left(\hat{\beta}_3^{ML}\right) &= \sqrt{0.0107} = 0.1034 \\ se\left(\hat{\beta}_4^{ML}\right) &= \sqrt{0.0059} = 0.0077 \end{aligned}$$

3. Test the null hypothesis $H_0 : \beta_2 + \beta_3 = 0$ against H_1 : not H_0.
Solution: For this hypothesis, we will use the following Wald statistic

$$W_0 = \left(\Gamma\hat{\beta}_{ML} - \gamma\right)' \left(\Gamma\tilde{\Lambda}_n\Gamma'\right)^{-1} \left(\Gamma\hat{\beta}_{ML} - \gamma\right) \overset{A}{\sim} \chi^2(r)$$

Where

$$\Gamma = \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix}, \gamma = 0, \text{ and } r = 1$$

Or

$$\begin{aligned} W_0 &= \left(\hat{\beta}_2 + \hat{\beta}_3\right)^2 \left(\begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \tilde{\Lambda}_n \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}\right)^{-1} \\ &= (0.0537)^2 \left(\begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.0488 & 0.0020 & -0.0130 & -0.0028 \\ 0.0020 & 0.0120 & -0.0082 & 0.0035 \\ -0.0130 & -0.0082 & 0.0107 & -0.0040 \\ -0.0028 & 0.0035 & -0.0040 & 0.0059 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}\right)^{-1} \\ &= 0.4577 \end{aligned}$$

Since $W_0 = 0.4577 \leq 3.8415 = c^*_{\chi^2(1),0.05}$, we fail to reject the null.

4. Provide the elasticities for x_2, x_3, and x_4 when the probabilities are evaluated at the sample average of x_2, x_3, and x_4.
Solution: Here, we calculate

$$\begin{aligned} e\left(Y_i, X_{2i}\right) &= \left.\frac{\partial E\left[Y_i | X_i, \beta\right]}{\partial X_{2i}} \frac{X_{2i}}{Y_i}\right|_{X_i = \bar{X}, Y_i = \bar{Y}} \\ &= \phi\left(\bar{X}'\hat{\beta}_{ML}\right) \hat{\beta}_2^{ML} \frac{\bar{X}_2}{\bar{Y}} \\ &= \phi\left(\begin{bmatrix} 1 & 2.0799 & 2.0651 & 1.9346 \end{bmatrix} \begin{bmatrix} 0.9217 \\ 1.0207 \\ -0.9670 \\ 0.4485 \end{bmatrix}\right) (1.0207) \frac{2.0799}{0.7440} \\ &= (2.8534)\,\phi\,(1.9154) \end{aligned}$$

Where

$$\begin{aligned}\phi(1.9154) &\approx \frac{1}{\sqrt{2*(3.1415)}}\exp\left\{-\frac{(1.9154)^2}{2}\right\}\\ &= 0.0637\end{aligned}$$

Thus,

$$e(Y_i, X_{2i}) = (2.8534)(0.0637) = 0.1818$$

Proceeding similarly for the other two elasticities,

$$\begin{aligned}e(Y_i, X_{3i}) &= \left.\frac{\partial E[Y_i|X_i,\beta]}{\partial X_{3i}}\frac{X_{3i}}{Y_i}\right|_{X_i=\bar{X},Y_i=\bar{Y}}\\ &= \phi\left(\bar{X}'\hat{\beta}_{ML}\right)\hat{\beta}_3^{ML}\frac{\bar{X}_3}{\bar{Y}}\\ &= \phi\left(\begin{bmatrix}1 & 2.0799 & 2.0651 & 1.9346\end{bmatrix}\begin{bmatrix}0.9217\\1.0207\\-0.9670\\0.4485\end{bmatrix}\right)(-0.9670)\frac{2.0651}{0.7440}\\ &= (-2.6841)\phi(1.9154)\\ &= (-2.6841)(0.0637)\\ &= -0.1710\end{aligned}$$

$$\begin{aligned}e(Y_i, X_{4i}) &= \left.\frac{\partial E[Y_i|X_i,\beta]}{\partial X_{4i}}\frac{X_{4i}}{Y_i}\right|_{X_i=\bar{X},Y_i=\bar{Y}}\\ &= \phi\left(\bar{X}'\hat{\beta}_{ML}\right)\hat{\beta}_4^{ML}\frac{\bar{X}_4}{\bar{Y}}\\ &= \phi\left(\begin{bmatrix}1 & 2.0799 & 2.0651 & 1.9346\end{bmatrix}\begin{bmatrix}0.9217\\1.0207\\-0.9670\\0.4485\end{bmatrix}\right)(0.4485)\frac{1.9346}{0.7440}\\ &= (1.1662)\phi(1.9154)\\ &= (1.1662)(0.0637)\\ &= 0.0743\end{aligned}$$

5. Consider now the function

$$h(\beta) = \frac{\beta_1\beta_2}{\beta_3^2}.$$

Provide an estimate for $h(\beta)$ and an estimate for its standard error using the delta method.
Solution: By the invariance principle,since h is continuous,

$$\widehat{h(\beta)} = \frac{\hat{\beta}_1^{ML}\hat{\beta}_2^{ML}}{\left(\hat{\beta}_3^{ML}\right)^2} = \frac{(0.9217)(1.0207)}{(-0.9670)^2} = 1.0061$$

By the delta method, we have that

$$\sqrt{n}\left(\frac{\hat{\beta}_1^{ML}\hat{\beta}_2^{ML}}{\left(\hat{\beta}_3^{ML}\right)^2} - \frac{\beta_1\beta_2}{\beta_3^2}\right) \xrightarrow{d} N(0, \Lambda_0)$$

Where

$$\Lambda_0 = \frac{\partial h(\beta)}{\partial \beta'} Var(\beta) \frac{\partial h(\beta)}{\partial \beta}$$

A consistent estimator for Λ_0 is

$$\hat{\Lambda} = \frac{\partial h\left(\hat{\beta}_{ML}\right)}{\partial \beta'} \widehat{Var(\beta)} \frac{\partial h\left(\hat{\beta}_{ML}\right)}{\partial \beta}$$

Calculating these values,

$$\left.\frac{\partial h}{\partial \beta'}\right|_{\hat{\beta}_{ML}} = \begin{bmatrix} \frac{\hat{\beta}_2^{ML}}{\left(\hat{\beta}_3^{ML}\right)^2} \\ \frac{\hat{\beta}_1^{ML}}{\left(\hat{\beta}_3^{ML}\right)^2} \\ -\frac{2\hat{\beta}_1^{ML}\hat{\beta}_2^{ML}}{\left(\hat{\beta}_3^{ML}\right)^3} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1.0207}{(-0.9670)^2} \\ \frac{0.9217}{(-0.9670)^2} \\ -\frac{2(0.9217)(1.0207)}{(-0.9670)^3} \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0916 \\ 0.9857 \\ 2.0808 \\ 0 \end{bmatrix}$$

we have that

$$\begin{aligned} \hat{\Lambda} &= \begin{bmatrix} 1.0916 & 0.9857 & 2.0808 & 0 \end{bmatrix} \begin{bmatrix} 0.0488 & 0.0020 & -0.0130 & -0.0028 \\ 0.0020 & 0.0120 & -0.0082 & 0.0035 \\ -0.0130 & -0.0082 & 0.0107 & -0.0040 \\ -0.0028 & 0.0035 & -0.0040 & 0.0059 \end{bmatrix} \begin{bmatrix} 1.0916 \\ 0.9857 \\ 2.0808 \\ 0 \end{bmatrix} \\ &= 0.0277 \end{aligned}$$

The standard error is thus $\sqrt{\hat{\Lambda}} = \sqrt{0.0277} = 0.1664$