Likelihood Theory and Extensions (BIOS:7110)
Breheny

Assignment 9
Due: Monday, November 10

1. *Score test with nuisance parameters.* This problem consists of proving the following theorem, which was stated in our lecture on score, Wald, and likelihood ratio approaches:

   Theorem: If (A)-(D) hold and $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_1^*$, then

   $$\mathbf{u}_1(\hat{\boldsymbol{\theta}}_0)^\top \mathscr{V}_{11}^n(\hat{\boldsymbol{\theta}}_0)\mathbf{u}_1(\hat{\boldsymbol{\theta}}_0) \xrightarrow{\text{d}} \chi_r^2.$$

   In this theorem, you may make use of the following lemma, which you do not need to prove (its proof is essentially identical to the case for the unrestricted MLE $\hat{\boldsymbol{\theta}}$).

   Lemma: If (A)-(D) hold and $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_1^*$, then

   $$\sqrt{n}(\hat{\boldsymbol{\theta}}_2(\boldsymbol{\theta}_0) - \boldsymbol{\theta}_2^*) = \mathscr{I}_{22}^{-1}(\boldsymbol{\theta}^*)\tfrac{1}{\sqrt{n}}\mathbf{u}_2(\boldsymbol{\theta}^*) + o_p(1),$$

   where $\mathbf{u} = (\mathbf{u}_1^\top \; \mathbf{u}_2^\top)^\top$ describes the partitioned score vector, $\mathscr{I}_{22}$ is the lower right block of the Fisher information for a single observation, and here the final $o_p(1)$ term is a $(d-r)\times 1$ vector, all of whose elements are converging in probability to zero.

   (a) Take a Taylor series expansion of $\frac{1}{\sqrt{n}}\mathbf{u}(\boldsymbol{\theta})$ about $\boldsymbol{\theta}^*$, evaluated at $\hat{\boldsymbol{\theta}}_0$.

   (b) Show that, under the null, $\frac{1}{\sqrt{n}}\mathbf{u}_1(\hat{\boldsymbol{\theta}}_0) = \mathbf{A}\frac{1}{\sqrt{n}}\mathbf{u}(\boldsymbol{\theta}^*) + o_p(1)$, where $\frac{1}{\sqrt{n}}\mathbf{u}_1(\hat{\boldsymbol{\theta}}_0)$ is the first component of the expansion in part (a) and $\mathbf{A}$ is a matrix for you to determine.

   (c) Using the result from (b), prove the above theorem.

2. *Finish Wilks's Theorem.* In our proof of Wilks' theorem, we saw that

   $$2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_0)\} = \frac{1}{\sqrt{n}}\mathbf{u}(\hat{\boldsymbol{\theta}}_0)^\top \mathscr{I}^{-1}\frac{1}{\sqrt{n}}\mathbf{u}(\hat{\boldsymbol{\theta}}_0) + o_p(1)$$

   and that

   $$\frac{1}{\sqrt{n}}\mathbf{u}(\hat{\boldsymbol{\theta}}_0) = (\mathbf{I} - \mathscr{I}\mathbf{H})\frac{1}{\sqrt{n}}\mathbf{u}(\boldsymbol{\theta}^*) + o_p(1),$$

   where $\mathscr{I}$ is evaluated at $\boldsymbol{\theta}^*$ and

   $$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathscr{I}_{22}^{-1} \end{bmatrix}.$$

   (a) Suppose that $\mathbf{z} \sim \text{N}(\mathbf{0}, \mathbf{I})$ and that $\mathbf{P}$ is a symmetric projection matrix (i.e., $\mathbf{PP} = \mathbf{P}$). Show that $\mathbf{z}^\top\mathbf{P}\mathbf{z} \sim \chi_r^2$, where r is the rank of $\mathbf{P}$ (you may wish to consult assignment 3 for some related results on projection matrices).

   (b) Show that $\mathbf{I} - \mathscr{I}^{1/2}\mathbf{H}\mathscr{I}^{1/2}$ is a projection matrix.

   (c) Given all of the above, show that $2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_0)\} \xrightarrow{\text{d}} \chi_r^2$, where r is the number of parameters we are testing.

3. *Score, Wald, and likelihood ratio tests for the rate parameter of the Gamma distribution.* This is a continuation of the problem, "Quadratic approximation for the Gamma distribution". In that problem, you derived the score and information, and simulated a random sample from the Gamma distribution. Now, implement score, Wald, and likelihood ratio approaches for carrying out inference regarding the rate parameter β. We discussed these approaches conceptually in class; this problem ensures that you understand all the specifics well enough to program them. Specifically, turn in a separate `.R` file that defines the following six functions (where `x` is the data, `level` is the confidence level, and `b0` is the hypothesized value of β:

   - `wald_ci(x, level=0.95)`
   - `wald_test(x, b0=1)`
   - `score_ci(x, level=0.95)`
   - `score_test(x, b0=1)`
   - `lr_ci(x, level=0.95)`
   - `lr_test(x, b0=1)`

   In other words, after sourcing the file you turn in, I should be able to run `score_ci(x)` and get a 95% confidence interval, but I should also be able to run `score_ci(x, 0.9)` and get a 90% confidence interval. The `_ci` functions should return a vector of length 2 (the lower and upper confidence limits), while the `_test` functions should return a vector of length 1 (the p-value).

   Your `.R` file can define other functions (for example, you will likely want an `mle()` function, since many of these methods require the MLE), but it should not execute anything or load any packages.

   Hint: For the score and likelihood ratio confidence interval functions, you may wish to look into the `extendInt` argument to `uniroot()`; note in particular that you can choose to extend the interval `"down"` or `"up"`.