Web Appendix. Independence between Step 1 and Step 2 Statistics

We show that the test statistics in Steps 1 and 2 of our 2-step procedure are asymptotically independent. This will follow from a general result that a large sample test (likelihood ratio, Wald, or score test) based on a conditional likelihood $L(\theta)=\Pr(Y \mid X; \theta)$ is asymptotically uncorrelated to any statistic S(X) based on the conditioning variables only. The proof of the latter, which we sketch below, relies heavily on the following property of conditional expectation(1). For any random variable U and random vector $\mathbf{X}$,

$$E\left[(U - E[U \mid \mathbf{X}])S(\mathbf{X})\right] = 0 \tag{A1}$$

for any function $S(\mathbf{X})$ of $\mathbf{X}$. This fundamental property, which expresses that $E[U \mid \mathbf{X}]$ is the orthogonal projection of U onto the space generated by all functions of $\mathbf{X}$, can in fact be used as the defining property of conditional expectation(1).

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be an independent and identically distributed random sample and $L(\theta) = \prod_{i=1}^{n} \Pr(Y_i \mid X_i; \theta)$ be the conditional likelihood. For simplicity we assume that Y_i, and X_i are one-dimensional random variables and θ is a one-dimensional parameter, but the result generalizes readily to random vectors and a multidimensional parameter (which might include nuisance parameters). Let $U(\theta) = \partial \log L(\theta) / \partial \theta$ be the score function, $I(\theta) = -E\left[\partial^2 L / \partial \theta^2\right]$ the expected fisher information, $\hat{\theta}$ the maximum likelihood estimate of θ, $\mathbf{X}=(X_1,\ldots,X_n)$, and $\mathbf{Y}=(Y_1,\ldots,Y_n)$,

By likelihood theory(1, 2) assuming standard regularity conditions, the score has conditional expectation zero, i.e. $E[U(\theta) \mid \mathbf{X}] = 0$ (conditional on $\mathbf{X}$ in this case because the likelihood is L=Pr(**Y** | **X**)) from which it also follows that $E[U(\theta)]=0$ unconditionally. Additionally, the maximum likelihood estimate of θ is closely related to the score as follows(1, 2):

$$\sqrt{I(\hat{\theta})}(\hat{\theta}-\theta)=U(\theta)\big/\sqrt{I(\theta)}+o_P(1)\,. \tag{A2}$$

Let S(**X**) be a statistic based solely on **X**. By the fundamental property of conditional expectation stated above in (A1),

$$0=E\big[\big(U(\theta)-E_\theta[U(\theta)\,|\,\mathbf{X}]\big)S(\mathbf{X})\big]=E\big[U(\theta)S(\mathbf{X})\big], \tag{A3}$$

i.e. the score is orthogonal to S(**X**). Assume we are interested in testing the hypothesis $H_0{:}\theta=\theta_0$. using the Wald statistic, $W=\sqrt{I(\hat{\theta})}(\hat{\theta}-\theta_0)$. By (A2), the Wald statistic is asymptotically equivalent to $T=U(\theta)\big/\sqrt{I(\theta)}+\mu_0$, where $\mu_0=\sqrt{I(\theta)}(\theta-\theta_0)$. By (A3) and the fact that $E\big[T\big]=\mu_0$ because $E[U(\theta)]=0$, we have,

$$\operatorname{cov}\big(T,S(\mathbf{X})\big)=E[TS(\mathbf{X})]-E[T]E[S(\mathbf{X})]=E\big[U(\theta)S(\mathbf{X})\big]\big/\sqrt{I(\theta)}+\mu_0E[S(\mathbf{X})]-\mu_0E[S(\mathbf{X})]=0$$

It follows that the Wald statistic W (and the likelihood ratio and score statistics which are asymptotically equivalent to it) is asymptotically uncorrelated with S(**X**). Notice that this holds regardless of the true parameter, i.e. either under the null hypothesis $\theta=\theta_0$ or an alternative $\theta\neq\theta_0$.

Now, as shown in Equation 1, the likelihood for trios factors as

$$\Pr(\mathbf{G}_c, \mathbf{E}_c, \mathbf{G}_p \mid \mathbf{D}_c = 1) = \Pr(\mathbf{G}_c \mid \mathbf{E}_c, \mathbf{G}_p, \mathbf{D}_c = 1)\Pr(\mathbf{E}_c, \mathbf{G}_p \mid \mathbf{D}_c = 1)$$

Where $\mathbf{G}_c$ and $\mathbf{G}_p$, denote the genotypes of all the children and all the parents in the sample respectively, and $\mathbf{E}_c$ and $\mathbf{D}_c$ denote the environmental variable and disease status of all the children in the sample respectively. The Step 2 statistic is a likelihood ratio statistic (or Wald statistic) based on the conditional likelihood comprising the first factor above, $L_2 = \Pr(\mathbf{G}_c \mid \mathbf{G}_p, \mathbf{E}_c, \mathbf{D}_c = 1)$. The statistic in Step 1 is a likelihood ratio statistic (or Wald statistic) based on the second factor $L_1 = \Pr(\mathbf{E}_c, \mathbf{G}_p \mid \mathbf{D}_c = 1)$, which only involves the conditioning variables $\mathbf{G}_p$ and $\mathbf{E}_c$ of the conditional likelihood L_2. Thus, the result above applies with $\mathbf{Y} = \mathbf{G}_c$, and $\mathbf{X} = (\mathbf{E}_c, \mathbf{G}_p)$. The test statistic W is the Wald version of the Step 2 statistic, and S(**X**) is the Wald version of the Step 1 statistic. Furthermore, these two statistics are jointly asymptotically normal, and, because two jointly normal variables that are uncorrelated are also independent, the Step 1 and Step 2 statistics are asymptotically independent.

Web Appendix References

1. Vaart AWvd. Asymptotic statistics. Cambridge, UK; New York NY, 1998.
2. Cox D, Hinkley D. Theoretical Statistics. London: Chapman and Hall, 1974.