Second Order Imhof Approximations to General Distribution Functions.

by M.Arellano and J.D.Sargan. January 1987

1. Introduction.

In a previous paper [ ] we suggested that an alternative to the Edgeworth approximation for the distribution of a general econometric estimator of a linear model with normally distributed errors might be used which derives from the Imhof method of calculating the distribution of the first order expansion of the estimator as a function of the sample data second moments . In the notation of the previous article suppose that ϕ, the econometric estimator, can be written

$$\phi = e(p),$$

where $p = m_2 - \mu_2$, and m_2 is a vector of sample data second moments, and $\mu_2 = E(m_2)$, and where $e(.)$ is a scalar function which has continuous first derivatives at the origin and $e(0) = 0$.

The simple Imhof approximation is then obtained by considering the distribution of $\phi^* = e_0'p$, where

$$e_0 = (\delta e/\delta p)_{p=0}.$$

Since p is a quadratic function of the original data and e_0 is a constant vector, ϕ^* is a scalar quadratic in a set of normally distributed variables. The distribution of ϕ^* can then be calculated using the Imhof algorithm. If F^* is the cumulative distribution function of ϕ^* then we consider approximations of the form $F^*(\psi(r))$, where $\psi(r)$ is an increasing function of r. More exactly since ϕ and ϕ^* are both stochastically of order $T^{-1/2}$, where T is the sample size, we consider

$$\Pr(T^{1/2}\phi < r) \simeq F(\psi(r)) \text{ , where } F = \Pr(T^{1/2}\phi^* < r),$$

and where for finite r , $\psi(r)-r = O(T^{-1/2})$.

If we take the Taylor series expansion of $\psi(r)$ at the origin so that its first three derivatives have values which make the first four moments equal to the corresponding moments of the exact distribution to $O(1/T)$ then both the exact and the approximating distribution will differ from the second order Edgeworth approximating distribution by $o(1/T)$. It then follows that for finite r the Imhof approximation has errors of order $o(1/T)$, and so is a second order approximation.

The argument is similar to that of our previous paper where a similar first order approximation was used. The new approximation can be expected to be of similar accuracy to the second order Edgeworth expansion at least for large T, and can be expected to be better in some models, particularly those where ϕ^* has a far from normal distribution and $e(p)$ is well approximated by a linear function of p near the origin.

2.General Theory.

Suppose as in the last section that we wish to approximate $\Pr(T^{1/2}\phi \leq r)$ by $F(h^*(r/\sigma))$, where

$$h^*(x) = x + (h_0^* + h_2^* x^2) T^{-1/2} + x(h_1^* + h_3^* x^2)T^{-1} ,$$

and σ is the asymptotic standard deviation defined by $\sigma^2 = \psi_{ij} e_i e_j$, where the notation is similar to that of [ ], $\psi_{ab\ldots d}$ representing the cumulant,defined as the derivative of the cumulant generating function differentiated at the origin with respect to $p_a,p_b,\ldots,p_d$ and rescaled by an appropriate power of T so as to be $O(1)$, and $e_{ab\ldots d}$ representing the derivative of $e(p)$ at the origin with respect to the same set of variables. It is possible to approximate $F(x)$ by its second order Edgeworth expansion with errors of $o(T^{-1})$, so that, defining $x = r/\sigma$, we can write

$$F(x) = I(\gamma(x)) + o(T^{-1}),$$

2nd order Edgeworth expansion of $T^{1/2}\phi^*$

where I(.) is the cumulative distribution function for the N(0,1) normally distributed variable, and

$$\gamma(x) \;=\; x \;+(\gamma_0 + \gamma_2\, x^2)T^{-1/2} + x(\gamma_1 \;+ \gamma_3 x^2)T^{-1},$$

and so we can write the Imhof second order approximation as

$$(1) \quad F(h*(x)) \;=\; I(\gamma(h*(x))) \;+\; o(T^{-1})\ .$$

Expanding the two cubic functions appropriately

$$(2) \qquad \gamma(h*(x)) \;=\; h(x) \quad + o(T^{-1})\ ,$$

where we can write

$$h(x) \;= x \;+ (h_0 \;+ h_2 x^2)T^{-1/2} \;+ x(h_1 \;+ h_3 x^2)T^{-1}.$$ 2nd order Edgeworth expansion of $T^{1/2}\phi$

If we choose h*(x) so that I(h(x)) is the second order Edgeworth approximation to the distribution function of $T^{1/2}\phi$ r, which has errors of order $o(T^{-1})$, then from (1) and (2) the second order Imhof approximation has errors of the same order.

$$\gamma(h*(x)) = x \;+\; (h_0* + h_2*x^2\ )T^{-1/2} \;+\; x(h_1* \;+\; h_3*x^2\ )T^{-1} \;+\; \gamma_0 T^{-1/2}$$
$$+\; \gamma_2[\ x^2 \;+\; 2x(h_0* \;+\; h_2 x^2)T^{-1/2}\ ]T^{-1/2} \;+\; x(\gamma_1 \;+\; \gamma_3 x^2\ )T^{-1}$$
$$+o(T^{-1}),$$ so that,

$$(3) \qquad h_0 \;=\; h_0* \quad +\gamma_0,$$

$$(3) \qquad h_2 \;=\; h_2* \quad +\gamma_2,$$

$$(3) \qquad h_1 \;=\; h_1* \quad +\gamma_1 \quad + 2\gamma_2 h_0*,$$

$$(3) \qquad h_3 \;=\; h_3* \quad +\gamma_3 \quad + 2\gamma_2 h_2*.$$

Then for given h_i and γ_i we can solve for the h_i* the equations

$$(4) \qquad h_0* \;=\; h_0 \quad -\gamma_0,$$

$$(4) \qquad h_2* \;=\; h_2 \quad -\gamma_2,$$

$$(4) \qquad h_1* \;=\; h_1 \quad -\gamma_1 \quad - 2\gamma_2(h_0 - \gamma_0),$$

$$(4) \qquad h_3* \;=\; h_3 \quad -\gamma_3 \quad - 2\gamma_2(h_2 - \gamma_2),$$

In the appendix we have extended the ideas of [ ] to allow E(p) to be nonzero but of O(1/T). The required h_i are defined in terms of the

cumulants of the p and the derivatives of $e(p)$ using the intermediate ψ ?
parameters defined in the Appendix as follows:- $\alpha_0 = \omega_i e_i$, $\alpha_1 = \omega_{jkm} e_j e_k e_m$, $\alpha_2 = \omega_{ijkm} e_i e_j e_k e_m$, $\alpha_3 = e_i \omega_{ia} e_{ab} \omega_{bj} e_j$, $\alpha_4 = e_{ab}\omega_{ab}$, $\alpha_5 = e_{ab}\omega_{abj} e_j$, $\alpha_6 = e_{abc}\omega_{ai} e_i \omega_{bj} e_j \omega_{ck} e_k$, $\alpha_7 = e_{abc}\omega_{ab}\omega_{cj} e_j$, $\alpha_8 = e_i \omega_{ia} e_{ab} \omega_{bc} e_{cd} \omega_{dj} e_j$, $\alpha_9 = \omega_{ab} e_{bc} \omega_{cd} e_{da}$, $\alpha_{10} = e_i \omega_{ia} e_{ab} \omega_{bjk} e_j e_k$, $\delta_1 = e_i \omega^*_{ij} e_j$, $\delta_2 = e_i \omega_{ij} e_{jk} \omega_k$, and $\delta = \delta_1 + \delta_2$, where in these definitions the repeated suffix summation convention is used.

Then
$$h_0 = [(\alpha_1 + 3\alpha_3)/\sigma^3 - 3(\alpha_4 + 2\alpha_0)/\sigma]/6,$$
$$h_2 = -(\alpha_1 + 3\alpha_3)/6\sigma^3,$$
$$72\sigma^6 h_1 = -14(\alpha_1 + 3\alpha_3)^2 + \sigma^2[9(\alpha_2 + 4\alpha_6 + 12\alpha_8 + 12\alpha_{10})$$
$$+12(\alpha_4 + 2\alpha_0)(\alpha_1 + 3\alpha_3)] - 18\sigma^4(2(\alpha_5 + \alpha_7 + \delta) + \alpha_9),$$
$$72\sigma^6 h_3 = 8(\alpha_1 + 3\alpha_3)^2 - 3\sigma^2(\alpha_2 + 4\alpha_6 + 12\alpha_8 + 12\alpha_{10}).$$

Since ϕ^* is a linear approximation to ϕ, it has the same first derivatives with respect to p at the origin, but all the higher order derivatives are zero so that in calculating the γ_i we put $e_{ab} = e_{abc} = 0$. This makes $\alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = \alpha_9 = \alpha_{10} = \delta_2 = 0$, and so
$$\gamma_0 = (\alpha_1 - 6\sigma^2\alpha_0)/6\sigma^3,$$
$$\gamma_2 = -\alpha_1/6\sigma^3,$$
$$72\sigma^6\gamma_1 = -14\alpha_1^2 + \sigma^2(9\alpha_2 + 24\alpha_0\alpha_1) - 18\sigma^4\delta_1,$$
$$72\sigma^6\gamma_3 = 8\alpha_1^2 - 3\sigma^2\alpha_2.$$

So, from equations (4),
$$h_0^* = (\alpha_3/\sigma^3 - \alpha_4/\sigma)/2,$$
$$h_2^* = -\alpha_3/2\sigma^3,$$
$$4\sigma^6 h_1^* = -\alpha_3(4\alpha_1 + 7\alpha_3) + 2\sigma^2((\alpha_6 + 3\alpha_8 + 3\alpha_{10}) + \alpha_3(\alpha_4 + 2\alpha_0))$$
$$- \sigma^4(2(\alpha_5 + \alpha_7 + \delta_2) + \alpha_9),$$
$$6\sigma^6 h_3^* = 3\alpha_3(\alpha_1 + 2\alpha_3) - \sigma^2(\alpha_6 + 3\alpha_9 + 3\alpha_{10}).$$

Note that these coefficients do not depend on α_2, and so do not depend on the fourth cumulants, but that they do depend on the third cumulants through α_5 and α_{10}. Note also that if e_{ab} and e_{abc} are all small then all the h_i* are proportionally small.

3. Models to b Simulated.

e

Appendix A. e 1.

In [ ] it was assumed that the vector p had the property E(p)=0, and the function g(p) had the property g(0)=0. This had the advantage of leading to somewhat simpler formulae and was justified by noting that the first property could be ensured by merely changing the origin of p, and the second property by a trivial change in the definition of g. However the previous argumments are not perfect if E(p) is O(1/T), when the alternative treatment of this Appendix is considerably simpler to use. At the same time it was pointed out in [ ] that σ might depend upon T, and in the particular case where

$$\sigma^2 = \sigma_0^2 + O(1/T)$$

then again a more useful form of approximation can be developed. Similar formulae have already been given in [ ] for the χ^2 approximation. It is assumed throughout this paper that the function g(p) does not vary with T.

Suppose that $E(p) = p_o/T$, and define $p^* = p - p_o/T$, so that p^* has $E(p^*)=0$, and then define $g^*(p^*) = g(p) - g(p_o/T)$. Then the theorem of [ ] can be applied to p^* and $g^*(p^*)$. Define $\alpha_o = Tg(p_o/T)$, and then

$$\Pr(\, T^{1/2} g < r) = \Pr(\, T^{1/2} g^* < r - T^{-1/2}\alpha_o \,)$$

$$\text{(A1)} \qquad = I[\, (r - T^{-1/2}\alpha_o)/\sigma + h_o/T^{1/2} + h_1 r/T\sigma + h_2 (r - T^{-1/2}\alpha_o)^2/T^{1/2}\sigma^2 + h_3 r^3/T\sigma^3 \,]$$

Suppose also that $E(p_i p_j) = \psi_{ij} = \psi_{oij} + \psi^*_{ij}/T$, where ψ_{oij} is independent of T and ψ^*_{ij} is $O(1)$. then

$\sigma^2 = e^*_i \psi_{ij} e^*_j$, where e^*_i is the first derivative

$$(\delta g^*/\delta p^*_i)_{p^*=0} = (\delta g/\delta p_i)_{p=p_o/T}$$

$$= (\delta g/\delta p_i)_{p=0} + T^{-1}(\delta^2 g/\delta p_i \delta p_j) p_{oj} + o(1/T)$$

$$= g_i + g_{ij} p_{oj}/T \;,$$

from which it follows that

$$\sigma^2 = g_i \psi_{oij} g_j + g_i \psi^*_{ij} g_j/T + 2\, g_i \psi_{oij} g_{jk} p_{ok}/T + o(1/T).$$

Define $\delta_1 = g_i \psi^*_{ij} g_j$, $\delta_2 = 2(g_i \psi_{oij} g_{jk} p_{ok})$, $\delta = \delta_1 + \delta_2$, and $\sigma_o^2 = g_i \psi_{oij} g_j$.

Then $\qquad \sigma^2 = \sigma_o^2 + \delta/T + o(1/T)$,

and $\qquad \sigma^{-1} = \sigma_o^{-1} - \tfrac{1}{2}\sigma_o^{-3}\delta/T + o(1/T)$.

Then substituting in (A1) ,

$$\Pr(T^{1/2} g < r) = I[r/\sigma_o + h^*_o/T^{1/2} + h^*_1 r/T\sigma_o + h_2 r^2/T^{1/2}\sigma_o^2 + h_3 r^3/T\sigma_o^3] + o(1/T),$$

where

$$h^*_o = h_o - \alpha_o = -[3(\alpha_4 + 2\alpha_o) - (\alpha_1 + 3\alpha_3)/\sigma_o^2]/6\sigma_o,$$

$$h^*_1 = h_1 - \tfrac{1}{2}\delta/\sigma_o^2 - 2\,\alpha_o h_2/\sigma_o, \quad \text{or}$$

$$72\sigma_o^6 = -14(\alpha_1 + 3\alpha_3)^2 + [9(\alpha_2 + 4\alpha_6 + 12(\alpha_8 + \alpha_{10}))$$

$$+12(\alpha_4 + 2\alpha_o)(\alpha_1 + 3\alpha_3)]\sigma_o^2 - 18(2(\alpha_5 + \alpha_7 + \delta) + \alpha_9)\sigma_o^4.$$

and ,as in [ ],

$$h_2 = -(\alpha_1 + 3\alpha_3)/6\sigma_o^3, \text{ and}$$

$$72\sigma_o^6 h_3 = 8(\alpha_1 + 3\alpha_3)^2 - 3[\alpha_2 + 4\alpha_6 + 12(\alpha_8 + \alpha_{10})]\sigma_o^2.$$