The choice of a mixing distribution in duration models

Sanjiv Jaggia

The Hougaard mixing distribution is considered for a Weibull duration model. This distribution is flexible and also encompasses the gamma and the inverse Gaussian distributions making it useful in discriminating between alternate distributions.

1. Introduction

In duration models, there has been a gradual shift of emphasis from the estimation and interpretation of econometric results to model testing and the evaluation of the consequences of misspecification. The focus of many studies has been the unobserved heterogeneity. [1] If test results indicate unobserved heterogeneity, then the crucial issue is to incorporate the mixing distribution of the heterogeneity term. Heckman and Singer (1984a) show that parameter estimates are very sensitive to the choice of this mixing distribution. In this paper, a fairly general Hougaard (1986) mixing distribution is considered for a Weibull model. The use of this distribution can be considered as a parametric alternative to the non-parameter alternative to the non-parametric method of Heckman and Singer. Further, a test for a commonly used gamma distribution is motivated within this general distribution.

The rest of the paper is organized as follows: Section 2 contains a discussion of the gamma and the inverse gaussian mixing distributions for a Weibull model. In section 3, the Hougaard distribution, that encompasses many attractive mixing distributions, is introduced. A test for a gamma distribution is motivated and the problem of identification is highlighted. Section 4 concludes.

2. Choosing a mixing distribution

Given the presence of unobserved heterogeneity, represented by V, the Weibull survivor function conditional on V is:

$$S(t; X, V) = \exp(-V\mu t^{\alpha}),\tag{2.1}$$

---

[1] See, for example, Burdett et al. (1985), Heckman and Singer (1984a, 1984b), Lancaster (1979, 1985).

where $\mu = \exp(\beta_0 = X_1\beta_1)$. As V is not observable, the unconditional distribution is derived as:

$$S(t;X) = \int \exp(-V\mu t^\alpha) p(V)\, dV, \tag{2.2}$$

where $p(V)$ represents the density function of V. The density and the hazard functions are derived similarly.

2.1. Gamma distribution

Several authors, including Lancaster (1979) and Vaupel et al. (1979), have used the gamma mixing distribution due to the flexibility of the family of gamma distributions. Furthermore, a closed form expression for $S(t;X)$ is obtainable that avoids resorting to numerical integration. The gamma density function with $E(V) = \delta/\theta$, $\mathrm{Var}(V) = \delta/\delta^2$ is:

$$p(V) = [1/\Gamma(\delta)]\theta^\delta V^{\delta-1} \exp(-V\theta). \tag{2.3}$$

If V has a finite mean, $E(V)$ can be set to equal 1 without loss of generality given that X includes a constant term. The unconditional survivor function, with $\delta = \theta$, is:

$$S(t;X) = [\theta/(\theta + \mu t^\alpha)]^\theta. \tag{2.4}$$

2.2. Inverse Gaussian distribution

The attractive features of a gamma mixing distribution are not unique and are shared by the entire class of non-negative exponential families [Hougaard (1984)]. The inverse gaussian distribution, for example, is as desirable as the gamma both analytically and computationally though it is not as widely used. [2] The density function of the inverse gaussian distribution with $E(V) = \delta\theta^{-1/2}$ and $\mathrm{Var}(V) = (\delta/2)\theta^{-3/2}$ is:

$$p(V) = \delta\pi^{-1/2} \exp(2\delta\theta^{1/2})V^{-3/2} \exp(-\theta V - \delta^2/V). \tag{2.5}$$

Again, the following are derived by setting $\delta = \theta^{1/2}$:

$$S(t;X) = \exp\left[-2\theta\left[((\theta + \mu t^\alpha)/\theta)^{1/2} - 1\right]\right]. \tag{2.6}$$

Since the estimates of a model are sensitive to the choice of mixing distribution [Heckman and Singer (1984a)], care must be taken in selecting any particular distribution. Further, as economic theory provides almost no basis on which to discriminate between competing models, sample information must be used for this purpose.

2.3. Graphical test

Informal graphical tests can be used in mixture models as well. These tests are based on:

$$\epsilon = -\ln(S(t;X)), \tag{2.7}$$

[2] Dean et al. (1989) use the inverse gaussian mixing distribution in a Poisson regression model.

where ϵ has a unit exponential distribution under the null and is a generalized error in the sense of Cox and Snell. An estimate of the survivor function of the residuals, $S(\hat{\epsilon})$, is computed and $-\ln S(\hat{\epsilon})$ is plotted against $\hat{\epsilon}$. If the model is correctly specified, the scatter plot clusters around a $45°$ line through the origin [see Lancaster and Chesher (1985)].

The generalized error for a Weibull-gamma model is:

$$\epsilon = \theta \ln[(\theta + \mu t^\alpha)/\theta],$$ (2.8)

and for a Weibull-inverse Gaussian model, it is:

$$\epsilon = 2\theta\left[((\theta + \mu t^\alpha)/\theta)^{1/2} - 1\right].$$ (2.9)

These errors can be used to visually infer the adequacy of the model. Sometimes it is difficult to determine whether the data sufficiently supports the theoretical specification of a model by simply examining a graphical plot. However, a comparison of graphs under two different specifications will at least determine which one is superior.

3. Hougaard distribution

Hougaard (1986) introduces a very general distribution that possesses all the properties of, as well as encompasses, the gamma and the inverse Gaussian distributions in it. The density function of this distribution is quite complex. However, only a simple Laplace transform of a mixing distribution is necessary for modelling purposes. The unconditional survivor function is:

$$S(t;X) = \int \exp(-V\mu t^\alpha)p(V)\ \mathrm{d}v = L(\mu t^\alpha),$$ (3.1)

where $L(s)$ is the Laplace transform [3] of the mixing distribution. The Laplace transform of the Hougaard distribution is:

$$L(s) = \exp\left[(-\delta/\gamma)\left[(\theta + s)^\gamma - \theta^\gamma\right]\right].$$ (3.2)

For $\theta > 0$ or $\gamma = 1$,

$$E(V) = \delta\theta^{\gamma-1} \text{ and } \mathrm{Var}(V) = \delta(1 - \gamma)\theta^{\gamma-2}.$$ (3.3)

Constraining $E(V)$ to 1 by setting $\delta = \theta^{1-\gamma}$ implies:

$$L(s) = \exp\left[(-\theta/\gamma)\left[((\theta + s)/\theta)^\gamma - 1\right]\right].$$

Therefore:

$$S(t;X) = \exp\left[(-\theta/\gamma)\left[((\theta + \mu t^\alpha)/\theta)^\gamma - 1\right]\right],$$ (3.4)

[3] If $f(y)$ is a real valued function that is defined for $y \geq 0$, then the Laplace transform of $f(y)$ is given by $\int \exp(-sy)f(y)\,\mathrm{d}y$.

and:

$$f(t;X) = \mu \alpha t^{\alpha-1} \left[\frac{\theta + \mu t^\alpha}{\theta} \right]^{\gamma-1} \exp\left[\frac{-\theta}{\gamma} \left[((\theta + \mu t^\alpha)/\theta)^\gamma - 1 \right] \right]. \tag{3.5}$$

The generality of the Hougaard distribution is desirable since economic theory does not provide much basis on which to choose a mixing distribution. Furthermore, this distribution can also be used to determine if the mixing distribution is gamma ($\gamma \to 0$), inverse Gaussian ($\gamma = 1/2$) or degenerate ($\gamma = 1$) since it encompasses all these distributions.

3.1. Score test for a gamma distribution

If all observations are complete, the log-likelihood function, using (3.5), is:

$$\mathscr{L} = \Sigma \left[\ln(\mu) = \ln(\alpha) + (\alpha - 1) \ln(t) \right.$$

$$\left. + (\gamma - 1) \ln((\theta + \mu t^\alpha)/\theta) - \frac{\theta}{\gamma} \left[((\theta + \mu t^\alpha)/\theta)^\gamma - 1 \right] \right]. \tag{3.6}$$

The likelihood function can be maximized in the usual way to estimate the parameters of the model. This function can also be used to test for a specific mixing distribution using a score test where the model needs to be estimated only under the null. The mean score $1/N(\partial\mathscr{L}/\partial\gamma)$ is:

$$\frac{1}{N} \Sigma \left[\ln((\theta + \mu t^\alpha)/\theta) + \frac{\theta}{\gamma^2} \left[((\theta + \mu t^\alpha)/\theta)^\gamma - 1 \right] \right.$$

$$\left. - \frac{\theta}{\gamma} ((\theta + \mu t^\alpha)/\theta)^\gamma \ln((\theta + \mu t^\alpha)/\theta) \right].$$

A test for a gamma distribution can be constructed by evaluating the above mean score at $\gamma \to 0$, which after some algebra is:

$$= \frac{1}{N\theta} \Sigma \left[\theta \ln((\theta + \mu t^\alpha)/\theta) - \tfrac{1}{2} [\theta \ln((\theta + \mu t^\alpha)/\theta)]^2 \right],$$

$$= \frac{1}{N\theta} \Sigma \left[\epsilon - \epsilon^2/2 \right]. \tag{3.7}$$

It is interesting to note that the above equation is just a function of the generalized error, ϵ as defined in (2.8). As $E(\epsilon^j) = j!$, $\partial\mathscr{L}/\partial\gamma$ clearly has expectation zero under the null hypothesis. The test can be implemented by substituting $\hat{\epsilon}$ in (3.7) and dividing the quantity by its estimated standard error. A score test indicator function can similarly be computed when the data are censored.

3.2. Problem of identification

The Hougaard distribution also includes the stable distribution as a special case ($\theta = 0$). Interestingly, a mix of the Weibull and the stable distributions results in a Weibull distribution with different parameters. This hightlights the problem of identification [4]. As finite moments do not

[4] See Elbers and Ridder (1982) and Heckman and Singer (1984b) for a general discussion.

exist for the stable distribution, the restriction $E(V) = 1$ cannot be imposed. The three parameter distribution, from (3.2), is:

$$L(s) = \exp\left[-\delta/\gamma\left((\theta + s)^\gamma - \theta^\gamma\right)\right],$$

$$= \exp\left[-\delta/\gamma(s^\gamma)\right] \quad \text{if} \quad \theta = 0.$$

Given $S(t; X, V) = \exp(-V\mu t^\alpha)$, the unconditional function is:

$$S(t; X) = \exp\left(-\delta/\gamma(\mu t^\alpha)^\gamma\right) = \exp(-\mu^* t^{\alpha*}), \tag{3.8}$$

where $\mu^* = \exp(\beta 0^* + \ln(\delta/\gamma) + X_1\beta_1^*)$; $\beta^* = \beta\gamma$ and $\alpha^* = \alpha\gamma$. If the Weibull model is estimated, the tests will indicate no misspecification. However, as $\gamma \in (0, 1]$, the structural parameters, β as well as α, will be underestimated.

4. Conclusion

In this paper, a general Hougaard mixing distribution is motivated for a Weibull duration model. This distribution encompasses many interesting distributions, thus making Hougaard a useful distribution to discriminate between such alternate models, in addition to its use as a flexible mixing distribution itself. Furthermore, the identification problem is highlighted when the mixing distribution is stable with no finite moments.

References

Burdett, K., N.M. Kiefer and S. Sharma, 1985, Layoffs and duration dependence in a model of turnover, Journal of Econometrics 28, 51–62.

Dean, C., J.F. Lawless and G.E. Willmot, 1989, A mixed Poisson-inverse Gaussian regression model, The Canadian Journal of Statistics 17, 171–181.

Elbers, C. and G. Ridder, 1982, True and spurious duration dependence: The identifiability of the proportional hazard model, Review of Economics Studies 49, 403–409.

Heckman, J. and B. Singer, 1984a, A method for minimizing the impact of distributional assumptions in econometric models for duration data, Econometrica 52, 271–320.

Heckman, J. and B. Singer, 1984b, The identifiability of the proportional hazard model, Review of Economic Studies 51, 231–241.

Hougaard, P., 1984, Life table methods for heterogeneous populations: Distributions describing the heterogeneity, Biometrica 71, 75–83.

Hougaard, P., 1986, Survival models for heterogeneous populations derived from stable distributions, Biometrica 73, 387–396.

Lancaster, T., 1979, Econometric methods for the duration of unemployment, Econometrica 47, 939–956.

Lancaster, T., 1985, Generalized residuals and heterogeneous duration models: With applications to the Weibull model, Journal of Econometrics 28, 155–169.

Lancaster, T. and A.D. Chesher, 1985, Residuals, tests and plots with a job Matching application, Annales de l'INSEE, 59/60, 47–70.

Vaupel, J.W., K.G. Manton and E. Stallard, 1979, The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography 16, 439–454.