Chapter 6

Inverse Gaussian distribution for Modelling Conditional Durations in Finance

6.1 Introduction

In traditional time series analysis investigators are concerned with the sequence of observations collected at equally spaced intervals. This was also our objective in the last five Chapters of this thesis. That is, in this case, the time process is considered as being non-stochastic. The general time series theory of Autoregressive Moving Average (see Box and Jenkins (1976)) or some of its modifications (see Brockwell and Davis (1991)) can be used in the modelling and forecasting of such situations. Although many financial data may be treated as time series, the standard techniques
of time series analysis cannot be employed here directly due to the rapid variation of the time intervals. Since many finance problems involve the arrival of events such as prices or trades in irregular time intervals, a new direction of modelling is necessary to explain the properties of such data.

In order to model the time durations between two successive events, Engle and Russell (1998) introduced the ACD model. Similar to the GARCH model for volatility, the ACD model catches duration clustering and is widely used for calculating expected duration. As mentioned by Hautsch (2012), the model can be directly applied to any other positive valued (continuous) process, such as trading volumes (Manganelli (2005)), market depth, bid-ask spreads or the number of trades (if they are sufficiently continuous). The basic idea is to (dynamically) parameterize the conditional duration mean rather than the intensity function itself.

Our objective in this chapter is to propose some conditional duration models based on inverse Gaussian distribution and study their properties. The motivation for this approach is: (i) inverse Gaussian distribution is a member of the natural exponential family of distributions and can be considered an alternative to exponential, log-Normal, log-logistic, Frechet and Weibull distributions, among others. Moreover, the inverse Gaussian has a hazard function which is non-monotonic; (ii) it is also likely to prove useful in statistical applications as a flexible and tractable model for fitting duration data, right-skewed unimodal data; (iii) it is a flexible closed form distribution that can be applied to model heavy-tailed processes (for example, it has been applied in many applications in studies of life times, reaction times, reliability and number of event occurrences in fields such as economics, and agricultural science).
Next section contains a brief review of the available ACD models in the literature. In section 6.3 we introduce the IG-ACD model and discuss the properties of the proposed model. The maximum likelihood method of estimation of IG-ACD model is discussed in Section 6.4. Sections 6.5 generalize the IG-ACD model to extended generalized inverse Gaussian (EGIG) ACD model and list the special cases of EGIG-ACD model. Section 6.6 discusses IG-SCD model and its properties. Section 6.7 briefly illustrates the efficient importance sampling method for maximum likelihood estimation of IG-SCD model and Section 6.8 contains the results of the simulation study. Finally Section 6.9 deals with a data analysis for illustrating the methods discussed in the previous Sections.

6.2 Review of ACD Models

Engle and Russell (1998) introduced the most popular ACD model that assumes that the error term follows the standard exponential distribution. The simplest and often very successful member of ACD family is the exponential ACD model. The exponential ACD model denoted by EACD (1,1), may be presented as

\[ X_i = \psi_i \varepsilon_i, \quad \psi_i = \omega + \alpha X_{i-1} + \beta \psi_{i-1}, \quad (6.1) \]

where \( \varepsilon_i \) follows the standard exponential distribution. We have \( E(\varepsilon_i) = 1, \ Var(\varepsilon_i) = 1, \) and \( E(\varepsilon_i^2) = 2. \)
Taking the expectation of the model, we obtain

\[ E(X_i) = E(\psi_i \varepsilon_i) = E(\psi_i) E(\varepsilon_i) = E(\psi_i), \]

\[ E(\psi_i) = \omega + \alpha E(X_{i-1}) + \beta E(\psi_{i-1}). \]

Under the weak stationarity assumption, \( E(X_i) = E(X_{i-1}) \), so that

\[ E(X_i) = E(\psi_i) = \frac{\omega}{1 - \alpha - \beta} = \mu_x. \quad (6.2) \]

Consequently, \( 0 \leq \alpha + \beta < 1 \) for a weakly stationary process \( \{X_i\} \).

We have \( E(X_i^2) = 2E(\psi_i^2) \).

Again, under weak stationarity,

\[ E(\psi_i^2) = \frac{\mu_x^2 [1 - (\alpha + \beta)^2]}{1 - 2\alpha^2 - \beta^2 - 2\alpha\beta}, \]

\[ Var(X_i) = \frac{\mu_x^2 [1 - \beta^2 - 2\alpha\beta]}{1 - 2\alpha^2 - \beta^2 - 2\alpha\beta}. \quad (6.3) \]

From these results, for the EACD(1,1) model to have a finite variance, we need \( 1 > 2\alpha^2 + \beta^2 + 2\alpha\beta \). Similar results can be obtained for the general EACD \((p,q)\) model, but the algebra involved becomes tedious. One may refer Engle and Russell (1998) for details.

The EACD model has several nice features. For instance, it is simple in theory and in ease of estimation. But the model also encounters some weaknesses. For example, the use of the exponential distribution implies that the model has a constant hazard
function. As stated in Tsay (2009) transaction duration in finance is inversely related to trading intensity, which in turn depends on the arrival of new information, making it hard to justify that the hazard function of duration is constant over time. To overcome this weakness, alternative innovation distributions have been proposed in the literature. Engle and Russell (1998) entertain the Weibull distribution for $\varepsilon_i$. A feature of Engle and Russell’s linear ACD specification with exponential or Weibull errors is that the implied conditional hazard functions are restricted to being constant, increasing or decreasing. Zhang et al. (2001), Hamilton and Jorda (2002) and Bauwens and Veredas (2004) questioned whether this assumption is an adequate one. As an alternative to the Weibull distribution used in the original ACD model, Lunde (1999) employs a formulation based on the generalized Gamma (GG) distribution, while Grammig and Maurer (2000) and Hautsch (2001) utilize the Burr and generalized F (GF) distributions respectively. Bhatti (2010) introduced Birnbaum-Saunders ACD model as an alternative to the existing ACD models which allow a unimodal hazard function. A recent review of the literature on the ACD models and their applications to finance can be found in Pacurar (2008).

Now, let us describe the specification of the Weibull ACD, GG-ACD, Burr-ACD and BS-ACD models. We begin with the WACD model. The Weibull probability density function with shape parameter $\theta$ and scale parameter $\sigma$ is given by

$$f(x; \theta, \sigma) = \frac{\theta}{\sigma} \left(\frac{x}{\sigma}\right)^{\theta-1} \exp\left[-\left(\frac{x}{\sigma}\right)^\theta\right], \quad x > 0; \quad \theta, \sigma > 0. \quad (6.4)$$

The mean of a Weibull($\theta, \sigma$) random variable is $E(X) = \sigma \Gamma(1 + \theta^{-1})$. Applying the change of variable $\varepsilon_i = X_i/\sigma \Gamma(1 + \theta^{-1})$ to (6.4); we obtain the unit mean
Weibull probability density function

\[ f_\varepsilon(\varepsilon_i) = \frac{\theta}{\Gamma(1 + \theta^{-1})} \left( \frac{\varepsilon_i}{\Gamma(1 + \theta^{-1})} \right)^{\theta-1} \exp \left[ - \left( \frac{\varepsilon_i}{\Gamma(1 + \theta^{-1})} \right)^{\theta} \right], \varepsilon_i > 0, \theta > 0 \]  

(6.5)

A final transformation must be applied to obtain the distribution of \( X_i \) parametrised in terms of the conditional mean \( \psi_i \). Applying the transformation \( \varepsilon_i = X_i/\psi_i \) yields the conditional probability density function

\[ f_{X_i|\psi_i}(X_i) = \frac{\theta}{\psi_i/\Gamma(1 + \theta^{-1})} \left( \frac{x_i}{\psi_i/\Gamma(1 + \theta^{-1})} \right)^{\theta-1} \exp \left[ - \left( \frac{x_i}{\psi_i/\Gamma(1 + \theta^{-1})} \right)^{\theta} \right]. \]  

(6.6)

The specification of the ACD model is completed by specifying the dynamic structure for the conditional mean \( \psi_i \).

Next we consider the construction of the GG-ACD model. The probability density function of the GG distribution is very similar to the probability density function of the Weibull distribution

\[ f(x; \kappa, \sigma, \theta) = \frac{\theta}{\sigma \Gamma(\kappa)} \left( \frac{x}{\sigma} \right)^{\kappa\theta-1} \exp \left[ - \left( \frac{x}{\sigma} \right)^{\theta} \right], x > 0; \kappa, \sigma, \theta > 0, \]  

(6.7)

where \( \Gamma(\kappa) \) is the usual gamma function defined by

\[ \Gamma(\kappa) = \int_0^\infty x^{\kappa-1} \exp(-x) \, dx. \]

Following the same steps as we did for the WACD model, we apply the transformation \( \varepsilon_i = X_i\phi(\kappa, \theta)/\sigma \) where \( \phi(\kappa, \theta) = \Gamma(\kappa)/\Gamma(\kappa + \theta^{-1}) \) to the probability density
function \( f(\varepsilon_i) = \frac{\theta}{\phi(\kappa, \theta) \Gamma(\kappa)} \left( \frac{\varepsilon_i}{\phi(\kappa, \theta)} \right)^{\theta-1} \exp \left[ - \left( \frac{\varepsilon_i}{\phi(\kappa, \theta)} \right)^{\theta} \right] \). \hspace{1cm} (6.8)

Applying the second transformation \( \varepsilon_i = X_i / \psi_i \) yields the conditional likelihood function for \( X_i \) given \( \psi_i \)

\[ f_{X_i|\psi_i}(X_i) = \frac{\theta}{\phi(\kappa, \theta) \psi_i \Gamma(\kappa)} \left( \frac{x_i}{\phi(\kappa, \theta) \psi_i} \right)^{\theta-1} \exp \left[ - \left( \frac{x_i}{\phi(\kappa, \theta) \psi_i} \right)^{\theta} \right]. \hspace{1cm} (6.9) \]

Grammig and Maurer (2000) proposed a more flexible specification based on the Burr distribution with probability density function

\[ f(x; \mu, \kappa, \sigma^2) = \frac{\mu \kappa x^{\kappa-1}}{(1 + \sigma^2 \mu x^\kappa)^{\frac{3}{2}} + 1}, \quad x > 0; \quad \mu, \kappa, \sigma^2 > 0. \hspace{1cm} (6.10) \]

Lancaster (1992) shows that the Burr distribution can be derived as a Gamma mixture of Weibull distributions. Exponential, Weibull and Log-Logistic are limiting cases. Unlike Weibull and Exponential, the Burr distribution is less frequently used in duration analysis.

Bhatti (2010) introduced BS-ACD model by specifying the time-varying model dynamics in terms of the conditional median duration, instead of the conditional mean duration. The probability density function of the BS(\( \kappa, \sigma \)) distribution is given by

\[ f(x; \kappa, \sigma) = \frac{1}{2\kappa \sigma \sqrt{2\pi}} \left[ \left( \frac{\sigma}{x} \right)^{\frac{3}{2}} + \left( \frac{x}{\sigma} \right)^{\frac{3}{2}} \right] \exp \left( -\frac{1}{2\kappa^2} \left[ \frac{x}{\sigma} + \frac{\sigma}{x} - 2 \right] \right), \quad x > 0; \quad \kappa, \sigma > 0. \hspace{1cm} (6.11) \]
Chapter 6. IG Conditional Duration Models

The conditional probability density function of $X_i$ given $\sigma_i$ is given by

$$ f_{X_i|\sigma_i}(X_i) = \frac{1}{2\kappa\sigma_i\sqrt{2\pi}} \left[ \left( \frac{\sigma_i}{x_i} \right)^{\frac{1}{2}} + \left( \frac{\sigma_i}{x_i} \right)^{\frac{3}{2}} \right] \exp \left( -\frac{1}{2\kappa^2} \left[ \frac{x_i}{\sigma_i} + \frac{\sigma_i}{x_i} - 2 \right] \right) \tag{6.12} $$

where $\sigma_i$ is the time-varying conditional median duration.

The other classes of ACD models are defined by different choices of functional form of conditional mean $\psi_i$. Bauwens et al. (2000) propose a logarithmic ACD (LACD) model that allows the introduction of additional variables without sign restrictions on their coefficients, as the LACD ensures the non-negativity of durations. Fernandes and Grammig (2006) develop a family of augmented ACD (AACD) models that encompasses the standard ACD model, the Log-ACD model and other ACD models inspired by the GARCH literature. Some extended ACD models allow for regime-dependence of the conditional mean function. Zhang et al. (2001) propose a threshold ACD (TACD) model to allow the expected duration to depend nonlinearly on past information variables. Unlike the TACD model, where the transition between states follows a jump process, Meitz and Teräsvirta (2006) introduce a smooth transition ACD (STACD) model. Based on the strong persistence of the trading duration, some long memory ACD models have been introduced. Based on the Ding and Granger (1996) two-component model for volatility, Engle (2000) applies the two-component model for duration. This allows for a slower decay autocorrelation function compared to the corresponding standard model. Jasiak (1998) introduces a fractionally integrated ACD (FIACD) model which is based on a fractionally integrated process for the expected duration. The FIACD model is closely linked with the fractionally integrated GARCH model proposed by Baillie et al. (1996). The
FIACD model is not covariance stationary and implies infinite first and second unconditional moments of the duration. Karanasos (2001) provides an alternative long memory ACD model which is analogous to the long-memory GARCH introduced by Robinson and Henry (1999). Drost and Werker (2004) develop a semiparametric ACD model that can relax the assumption of independently, identically distributed innovations of the standard ACD model. Like the similarity between the ACD and GARCH models, and based on the idea of SV model, Bauwens and Veredas (2004) propose the stochastic conditional duration model for duration. The SCD model is based on the assumption that the durations are generated by a dynamic stochastic latent variable.

Many physical phenomena exhibit hazard functions that are non-monotonic. Grammig and Maurer (2000) provide the motivation to deal with non-monotonic hazard functions when modelling financial duration processes. In the following section we propose a more flexible model for conditional durations based on the inverse Gaussian distribution. One of the important distributions studied in the context of modelling the sequence of durations is the inverse Gaussian distribution. For example, Lancaster (1972) used this distribution to model the intervals between events, such as duration of strikes. In the case of financial series, the sequences of log-returns are assumed to be realizations of Brownian motion or Gaussian process. It is well known that inverse Gaussian distribution arises as a first passage time distribution in any Gaussian process. Duration between events can be compared with the life times of units in renewal/reliability related studies. The distributions having non-monotonic failure rate are important in such studies and the inverse Gaussian distribution possesses such a property (see Chhikara and Folks (1977)).
6.3 Inverse Gaussian ACD Model

Now let us consider the construction of IG-ACD model. A random variable $X$ is said to have an inverse Gaussian distribution with parameters $\mu$ and $\lambda$ and is denoted by $IG(\mu, \lambda)$ if its probability density function is given by

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp \left\{ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right\}, \quad x > 0,$$

where $\mu$ and $\lambda$ are assumed to be positive. $\mu$ is the mean of the distribution and $\lambda$ is a shape parameter. This density is unimodal and skewed. The variance for the distribution is $\mu^3/\lambda$, implying $\mu$ is not a location parameter in the usual sense.

Assuming that $\varepsilon_i$ follows a unit mean IG distribution with probability density function

$$f_{\varepsilon}(\varepsilon_i) = \sqrt{\frac{\lambda}{2\pi \varepsilon_i^3}} \exp \left\{ -\frac{\lambda(\varepsilon_i - 1)^2}{2\varepsilon_i} \right\}, \quad \varepsilon_i > 0,$$

and IG-ACD (1,1) model can be written as

$$X_i = \psi_i \varepsilon_i, \quad \psi_i = \omega + \alpha X_{i-1} + \beta \psi_{i-1}.$$  (6.15)

Conditional on $F_{i-1}$, the probability density function of $X_i$ can be expressed as

$$f(x_i|F_{i-1}) = \frac{1}{\psi_i} f_{\varepsilon_i} \left( \frac{x_i}{\psi_i} \right) = \sqrt{\frac{\lambda \psi_i}{2\pi x_i^3}} \exp \left\{ -\frac{\lambda \psi_i(x_i - \psi_i)^2}{2\psi_i^2 x_i} \right\}.$$  (6.16)

That is, the conditional distribution of $X_i$ given the past information is $IG(\psi_i, \lambda \psi_i)$. 

6.3.1 Properties of IG-ACD Model

Conditional on \((X_{i-1}, X_{i-2}, \ldots)\) the mean and variance of \(X_i\) are given by \(E(X_i|F_{i-1}) = \psi_i\) and \(Var(X_i|F_{i-1}) = \psi_i^2 Var(\varepsilon_i) = \psi_i^2/\lambda\).

Further the model (6.15) implies that the unconditional mean and variance of the stationary distribution of \(\{X_i\}\) can be respectively obtained as

\[
\mu_x \equiv E(X_i) = E(\psi_i) = \frac{\omega}{1-\alpha - \beta}, \quad Var(X_i) = \frac{\mu^2(1 - \beta^2 - 2\alpha\beta)}{\lambda \left[1 - (1 + \frac{1}{\lambda})\alpha^2 - \beta^2 - 2\alpha\beta\right]}.
\]

Consequently, for weak stationarity of \(\{X_i\}\) we need the condition \(0 \leq \alpha + \beta < 1\).

**Autocorrelation function:**

The \(k^{th}\) order auto-covariance function of \(X_i\) is defined as

\[
\gamma_k = Cov(X_i, X_{i-k}) = Cov(\psi_i, X_{i-k})
\]

\[
= Cov(\omega + \alpha X_{i-1} + \beta \psi_{i-1}, X_{i-k})
\]

\[
= \alpha Cov(X_{i-1}, X_{i-k}) + \beta Cov(\psi_{i-1}, X_{i-k})
\]

\[
\gamma_k = (\alpha + \beta)\gamma_{k-1}
\]  

(6.17)

The first order auto-covariance function of \(X_i\) is

\[
\gamma_1 = Cov(X_i, X_{i-1}) = Cov(\psi_i, X_{i-1})
\]

\[
= Cov(\omega + \alpha X_{i-1} + \beta \psi_{i-1}, X_{i-1})
\]
\[ \gamma_1 = \alpha \gamma_0 + \beta \text{Var}(\psi_{i-1}) \]

where \( \text{Var}(\psi_{i-1}) = \frac{\alpha^2 \mu^2}{1 - 2\alpha^2 - \beta^2 - 2\alpha \beta} \).

Finally, the \( k^{th} \) order ACF of \( \{X_i\} \) is derived as

\[ \rho_k = (\alpha + \beta) \rho_{k-1}, \quad k > 1 \tag{6.18} \]

with

\[ \rho_1 = \frac{\alpha(1 - \beta^2 - \alpha \beta)}{1 - \beta^2 - 2\alpha \beta}. \]

Forecasts from an IG-ACD model can be obtained using a procedure similar to that of a GARCH model (cf. Pacurar (2008)).

**Intensity function or Hazard function:**

Let us denote by \( T \) the duration of stay in the state of interest and recall the definition of the hazard function as the instantaneous rate of leaving the interval between \( T = t \) and \( T = t + \Delta t \), given that it stayed up to time \( t \),

\[ h(t) = \lim_{\Delta t \to 0} \frac{p(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \]

Then the hazard function implied by the IG-ACD model may now be written as

\[ h(x_i) = \sqrt{\frac{\lambda \psi_i}{2\pi x_i}} \exp \left\{ -\frac{(\lambda \psi_i (x_i - \psi_i))^2}{2\mu^2 x_i} \right\} \Phi \left( \sqrt{\frac{\lambda \psi_i}{x_i}} \left( 1 - \frac{x_i}{\psi_i} \right) \right) - e^{2\lambda} \Phi \left( -\sqrt{\frac{\lambda \psi_i}{x_i}} \left( 1 + \frac{x_i}{\psi_i} \right) \right), \quad x_i > 0, \tag{6.19} \]

where \( \Phi(.) \) is the standard normal distribution function. The expression for \( h(x_i) \) is
rather complicated but it is not difficult to compute for any given values of parameters. Several typical hazard function curves are given in Figure 6.1. Inspection of these curves reveals that the hazard function is non-monotonic for all $\mu$ and $\lambda$.

![Figure 6.1: Hazard rate of inverse Gaussian distribution when $\mu = 1$.](image)

### 6.4 Estimation of IG-ACD Model

Let $X = (X_1, X_2, ..., X_n)$ be a realization from an IG-ACD(1,1) model and the parameter vector to be estimated be $\Theta = (\lambda, \omega, \alpha, \beta)$. The likelihood function of $\Theta$ based on $X$ may be expressed as

$$L(\Theta | X) = f(X_1 | \Theta) \prod_{i=2}^{n} f(X_i | F_{i-1}; \Theta),$$

(6.20)
where \( f(X_1; \Theta) \) is the density function of the initial random variable and it does not have a closed form expression. Further its influence on the overall likelihood function diminishes as the sample size \( n \) increases and hence we adopt the conditional likelihood method by ignoring the term \( f(X_1; \Theta) \). Now using (6.16) in (6.20) the conditional log-likelihood function is given by

\[
\log L = \frac{n}{2} \log \lambda + \frac{1}{2} \sum_{i=2}^{n} \log \psi_i - \frac{n}{2} \log(2\pi) - \frac{3}{2} \sum_{i=2}^{n} \log X_i - \frac{n}{2} \sum_{i=2}^{n} \frac{\psi_i(X_i - \psi_i)^2}{\psi_i^2 X_i}.
\]

(6.21)

The ML estimator of \( \lambda \) is given by

\[
\hat{\lambda} = \left[ \frac{1}{n-1} \sum_{i=2}^{n} \frac{\psi_i(X_i - \psi_i)^2}{\psi_i^2 X_i} \right]^{-1}.
\]

(6.22)

We obtain the ML estimates of the remaining parameters \((\omega, \alpha, \beta)\) by Newton-Raphson iteration method. In order to develop this method let us denote \( \Theta = (\omega, \alpha, \beta) = (\theta_1, \theta_2, \theta_3) \).

Taking the first and second order partial derivative of (6.21) with respect to \( \theta_j \), we get

\[
L'(\theta_j) = \frac{\partial \log L}{\partial \theta_j} = \sum_{i=2}^{n} \left( \frac{1}{2\psi_i} - \frac{\lambda}{X_i} + \frac{\lambda X_i}{2\psi_i^2} + \frac{\lambda}{2X_i} \right) \frac{\partial \psi_i}{\partial \theta_j},
\]

\[
L''(\theta_j) = \frac{\partial^2 \log L}{\partial \theta_j^2} = -\sum_{i=2}^{n} \left( \frac{1}{2\psi_i^2} + \frac{\lambda X_i}{\psi_i^3} \right) \left( \frac{\partial \psi_i}{\partial \theta_j} \right)^2, \quad j = 1, 2, 3.
\]

Now the iteration formula for estimating \( \theta_j \) is given by

\[
\hat{\theta}_j^{(m+1)} = \hat{\theta}_j^{(m)} - \left[ \frac{L'(\hat{\theta}_j^{(m)})}{L''(\hat{\theta}_j^{(m)})} \right], \quad m = 1, 2, \ldots,
\]

(6.23)
where $\hat{\theta}_j^{(m)}$ is the estimate of $\theta_j$ obtained at $m$th iteration. The computation of these estimates based on a simulated sample is illustrated in Section 6.8.

In the next section, we extend the IG-ACD framework to a more flexible specification based on the EGIG distribution. The major advantage of this distribution over the exponential and Weibull distribution, the most common distributions utilized in ACD models, is that these have non-monotonic hazard functions taking bathtub shaped or inverted bathtub shaped forms. In some cases, there is more than one turning point of the hazard rate for EGIG distribution. The shape properties of EGIG hazard are derived in Gupta and Viles (2011).

### 6.5 Extended Generalized Inverse Gaussian ACD Model

Now let us consider the construction of the EGIG-ACD model. The probability density function of the EGIG distribution is given by

$$f(x; a, b, \lambda, \delta) = \frac{1}{(2/\delta) (b/a)^{\lambda/2 \delta}} K_{\lambda} \left( \frac{2\sqrt{ab}}{2} \right) x^{\lambda-1} \exp \left(-ax^\delta - bx^{-\delta} \right), \ x > 0,$$

where $K_{\nu}(z)$ is a modified Bessel function of the third kind with index $\nu$ and is defined by

$$K_{\nu}(z) = \frac{1}{2} \int_0^\infty x^{\nu-1} \exp \left\{ -\frac{1}{2} z(x + x^{-1}) \right\} \ dx$$
Chapter 6. IG Conditional Duration Models

The domain of variation for the parameters is

\[ \lambda \in \mathbb{R}, \quad (a, b, \delta) \in \Omega_\lambda, \]

Where

\[ \Omega_\lambda = \begin{cases} 
(a, b, \delta) : a > 0, b \geq 0, \delta > 0 & \text{if } \lambda > 0 \\
(a, b, \delta) : a > 0, b > 0, \delta > 0 & \text{if } \lambda = 0 \\
(a, b, \delta) : a \geq 0, b > 0, \delta > 0 & \text{if } \lambda < 0.
\]

For more details on this distribution see Jørgensen (1982). This model includes as special cases the generalized inverse Gaussian distribution for \( \delta = 1 \), the inverse Gaussian for \( \delta = 1, \lambda = -1/2 \), and the generalized gamma distribution for \( \lambda > 0, b \to 0 \). Now it is straightforward that Exponential, Weibull, gamma distributions are particular cases of generalized gamma distribution for \( \lambda = \delta = 1, \lambda = \delta \) and \( \delta = 1 \) respectively.

The mean of the EGIG random variable is

\[ E(X) = \left( \sqrt{\frac{b}{a}} \right)^{1/\delta} K_{(\lambda+1)/\delta} \left( \frac{2\sqrt{ab}}{K_{\lambda/\delta} \left( 2\sqrt{ab} \right)} \right) = \varphi, \ (\text{say}). \]

Applying the change of variable \( \varepsilon_i = \frac{x_i}{\varphi} \), we obtain the unit mean EGIG probability density function

\[ f_{\varepsilon}(\varepsilon_i) = \frac{1}{\left( \frac{2}{\delta} \right) \left( \frac{b\varphi^{-\delta}}{a\varphi^\delta} \right)^{\lambda/2\delta} K_{\lambda/\delta} \left( \frac{2\sqrt{a\varphi^\delta b\varphi^{-\delta}}}{2\sqrt{ab}} \right)^{\varepsilon_i^{\lambda-1} e^{-(a\varphi^\delta)\varepsilon_i^\delta - (b\varphi^{-\delta})\varepsilon_i^{-\delta}}} } \] (6.25)
Then the conditional probability density function of \( x_i \) given \( \psi_i \) is

\[
f_{X_i|\psi_i}(X_i) = \left[ \left( \frac{2}{\delta} \right) \left( \frac{b \left( \frac{\varphi}{\psi_i} \right)^{-\frac{\delta}{\lambda}}} {a \left( \frac{\varphi}{\psi_i} \right)^{\frac{\delta}{\lambda}}} \right)^{\lambda/2\delta} K_{\frac{\lambda}{\delta}} \left( 2 \sqrt{a \left( \frac{\varphi}{\psi_i} \right)^{\delta} b \left( \frac{\varphi}{\psi_i} \right)^{-\delta}} \right) \right]^{-1}
\times x_i^{\lambda-1} e^{- \left[ a \left( \frac{\varphi}{\psi_i} \right)^{\delta} x_i^\delta - b \left( \frac{\varphi}{\psi_i} \right)^{-\delta} \right]} x_i^{-\delta}.
\]

That is conditional density of \( x_i \) given \( \psi_i \) follows \( \text{EGIG} \left( a \left( \frac{\varphi}{\psi_i} \right)^{\delta}, b \left( \frac{\varphi}{\psi_i} \right)^{-\delta}, \lambda, \delta \right) \).

### 6.5.1 Special Cases

Accordingly, EGIG distribution includes as special or limiting cases many distributions considered in econometrics and finance. This generalization consists of all the standard ACD models including EACD, WACD, GG-ACD models and IG-ACD model which we described in this chapter. For \( \lambda > 0 \) and \( b \to 0 \), EGIG-ACD model reduces to the GG-ACD model proposed by Lunde (1999). The probability density function of a generalized Gamma random variable is given by

\[
f(x; a, \lambda, \delta) = \frac{\delta}{\Gamma \left( \frac{\lambda}{\delta} \right)} a^{\frac{\lambda}{\delta}} x^{\lambda-1} e^{-ax^\delta} ; \ x > 0.
\]

Then the conditional density of \( x_i \) given \( \psi_i \) of GG-ACD is given by

\[
f_{X_i|\psi_i}(X_i) = \frac{\delta}{\Gamma \left( \frac{\lambda}{\delta} \right)} \left( a \left( \frac{\varphi}{\psi_i} \right)^{\delta} \right)^{\frac{\lambda}{\delta}} x_i^{\lambda-1} e^{-a \left( \frac{\varphi}{\psi_i} \right)^{\delta} x_i^\delta} ; \ x_i > 0.
\]

(6.27)
The generalized gamma family of density functions nests the Weibull distribution and exponential distribution. Both types of distributions have been already successfully applied in ACD framework by Engle and Russell (1998). For $\lambda = \delta = 1$, it reduces to EACD model and for $\lambda = \delta$, it reduces to WACD model. The problem of a flat conditional intensity of EACD was already raised by Engle and Russel as not having a good fit with some semiparametric estimate of the baseline hazard of the data and they therefore propose to extend the EACD model by generalizing the exponential density of the standardized durations to a Weibull density.

For $\delta = 1$, $\lambda = -1/2$, the EGIG-ACD model reduces to IG-ACD model which we have discussed in Section 6.3.

So far we have discussed about the inverse Gaussian ACD model for analysing the financial transaction durations. Next section will discuss about inverse Gaussian SCD model in which the durations are generated by a dynamic stochastic latent variable.

### 6.6 Inverse Gaussian SCD Model and Properties

Recall the SCD model of order one which we discussed in Chapter 2 as

$$X_i = e^{\psi_i} \varepsilon_i, \quad \psi_i = \omega + \beta \psi_{i-1} + u_i,$$  \hspace{1cm} (6.28)

In this section we discuss the model (6.28) when $\varepsilon_i$ follows a unit mean IG distribution and $\{u_i\}$ is an independent and identically distributed sequence of $N(0, \sigma^2)$
random variables. From the definition of the model it follows that \( \{ \psi_i \} \) is a Gaussian sequence and hence \( \{ e^{\psi_i} \} \) is a stationary log-normal Markov sequence. Now using the property of log-normal distribution, all the moments of \( X_i \) can be computed. In particular the mean and variance are respectively given by

\[
E(X_i) = \exp \left\{ \frac{\omega}{1 - \beta} + \frac{\sigma^2}{2(1 - \beta^2)} \right\}, \quad (6.29)
\]

and

\[
Var(X_i) = \exp \left\{ \frac{2\omega}{1 - \beta} + \frac{\sigma^2}{1 - \beta^2} \right\} \left[ \left( 1 + \frac{1}{\lambda} \right) \exp \left( \frac{\sigma^2}{1 - \beta^2} \right) - 1 \right]. \quad (6.30)
\]

**Autocorrelation function:**

The \( k^{th} \) autocovariance function of \( X_i \) is defined as

\[
\gamma_k = Cov(X_i, X_{i-k}) = E(X_i X_{i-k}) - E(X_i) E(X_{i-k})
\]

Now we need to compute the expectation of \( X_i X_{i-k} \), which is equal to

\[
E(X_i X_{i-k}) = E(e^{\psi_i} \epsilon_i e^{\psi_{i-k}} \epsilon_{i-k})
\]

\[
= E(e^{\psi_i + \psi_{i-k}}) E(\epsilon_i \epsilon_{i-k}).
\]

From the autoregressive equation of \( \psi_i \), we get

\[
\psi_i + \psi_{i-k} = \lambda_{i,k} = 2\omega + \beta \lambda_{i-1,k} + u_i + u_{i-k}
\]
which is a Gaussian ARMA(1, k) process (with restrictions in the MA polynomial). Unconditionally,

\[ e^{\lambda_{i,k}} \sim LN \left( \mu_k, \sigma_k^2 \right), \]

where

\[ \mu_k = \frac{2\omega}{1 - \beta} \]
\[ \sigma_k^2 = \frac{2\sigma^2 (1 + \beta^k)}{1 - \beta^2}. \]

Hence

\[ E(X_iX_{i-k}) = \exp \left\{ \frac{2\omega}{1 - \beta} + \frac{\sigma^2 (1 + \beta^k)}{1 - \beta^2} \right\}. \tag{6.31} \]

Therefore, the lag \( k \) auto-covariance function of \( X_i \) is

\[ \gamma_k = \exp \left\{ \frac{2\omega}{1 - \beta} + \frac{\sigma^2}{1 - \beta^2} \right\} \left[ \exp \left( \frac{\sigma^2 \beta^k}{1 - \beta^2} \right) - 1 \right]. \tag{6.32} \]

Finally, the \( k^{th} \) order autocorrelation function of \( X_i \) is \( \rho_k = \gamma_k/\gamma_0 \) and is given by

\[ \rho_k = \frac{\exp \left( \frac{\sigma^2 \beta^k}{1 - \beta^2} \right) - 1}{(1 + \frac{1}{\lambda}) \left\{ \exp \left( \frac{\sigma^2}{1 - \beta^2} \right) - 1 \right\}} \approx \frac{\sigma^2 \beta^k/(1 - \beta^2)}{(1 + \frac{1}{\lambda}) \left\{ \exp \left( \frac{\sigma^2}{1 - \beta^2} \right) - 1 \right\}} \approx \beta \rho_{k-1}, \tag{6.33} \]

The autocorrelation function \( \rho_k \) geometrically decreases at the rate \( \beta \) as the lag \( k \) increases.

**Hazard Function:**

The hazard function \( h(.) \) implied by the IG-SCD model can be computed by the formula

\[ h(x_i) = \frac{f(x_i)}{1 - \int_{0}^{x_i} f(u) \, du}, \tag{6.34} \]
where

\[ f(x_i) = \frac{\sqrt{\lambda}}{2\pi\sigma\sqrt{x_i^3}} \int_0^\infty \frac{1}{\sqrt{z}} \exp \left\{ -\frac{1}{2} \left( \frac{\lambda z (\frac{x_i}{z} - 1)^2}{x_i} + \frac{(\log z - \mu)^2}{\sigma^2} \right) \right\} dz. \]

One way of estimating the hazard function is to replace the parameters in the above expressions by their respective estimates. We do it for the simulated and the real data in Sections 6.8 and 6.9.

### 6.7 Estimation of IG-SCD Model

A relatively new method for computing the integral needed for evaluating the likelihood function of models with latent variables relies on the efficient importance sampling procedure, recently developed by Richard and Zhang (2007). This method is an extension of the well known importance sampling technique and seems to be particularly well suited for the computation of the multidimensional though relatively well behaved integral needed for evaluation of the SCD likelihood. Given a sequence \( X \) of \( n \) realizations of the process, with density \( g(X|\psi, \theta_1) \) indexed by the parameter vector \( \theta_1 \), conditional on a vector \( \psi \) of a latent variables of the same dimension as \( X \), and given the density \( h(\psi|\theta_2) \) indexed by the parameter \( \theta_2 \), the likelihood function of \( X \) can be written as:

\[
L(\theta; X) = L(\theta_1, \theta_2; X) = \int g(X|\psi, \theta_1) \ h(\psi|\theta_2) \ d\psi. \quad (6.35)
\]
Actually, the integrand in the previous equation is the joint density \( f(X, \psi | \theta) \). Given the assumptions we made, it can be sequentially decomposed as

\[
f(X, \psi | \theta) = \prod_{i=1}^{n} d(X_i, \psi_i|X_{i-1}, \psi_{i-1}, \theta) = \prod_{i=1}^{n} p(X_i|\psi_i, \theta_1) q(\psi_i|\psi_{i-1}, \theta_2),
\]

(6.36)

where \( p(X_i|\psi_i, \theta_1) \) is obtained from \( p(\varepsilon_i) \) (so that \( \theta_1 \) corresponds to the parameters of inverse Gaussian distribution), and \( q(\psi_i|\psi_{i-1}, \theta_2) \) is the Gaussian density \( N(\omega + \beta \psi_{i-1}, \sigma^2) \) (so that \( \theta_2 \) includes \( \omega, \beta \) and \( \sigma^2 \)).

A natural Monte Carlo (MC) estimate of the likelihood function in (6.36) is given by

\[
\tilde{L}(\theta; x) = \frac{1}{S} \sum_{j=1}^{S} \left[ \prod_{i=1}^{n} p(X_i|\tilde{\psi}_i^{(j)}, \theta_1) \right],
\]

(6.37)

where \( \tilde{\psi}_i^{(j)} \) denotes a draw from the density \( q(\psi_i|\psi_{i-1}, \theta_2) \). This approach bases itself only on the information provided by the distributional assumptions of the model and does not consider the information that comes from the observed sample. It turns out that this estimator is highly inefficient since its sampling variance rapidly increases with the sample size. In any practical case of a duration data set, where the sample size \( n \) lies between 500 and 50000 observations, the Monte Carlo sampling size \( S \) required to give precise enough estimates of \( L(\theta; x) \) would be too high to be affordable and it turns out that this estimator cannot be relied on practically.

EIS tries to make use of the information provided by the observed data in order to come to a reasonably fast and reliable numerical approximation. The principle of EIS is to replace the model-based sampler \( \{ q(\psi_i|\psi_{i-1}, \theta_2) \}_{i=1}^{n} \) with an optimal
auxiliary parametric importance sampler. Let \( \{ m(\psi_i|\psi_{i-1}, a_i) \}_{i=1}^n \) be a sequence of auxiliary samplers indexed by the set of auxiliary parameter vectors \( \{ a_i \}_{i=1}^n \). These densities can be defined as a parametric extension of the natural samplers \( \{ q(\psi_i|\psi_{i-1}, \theta_2) \}_{i=1}^n \). We rewrite the likelihood function as

\[
L(\theta; X) = \int \left[ \prod_{i=1}^n \frac{d(X_i, \psi_i|X_{i-1}, \psi_{i-1}, \theta)}{m(\psi_i|\psi_{i-1}, a_i)} \prod_{i=1}^n m(\psi_i|\psi_{i-1}, a_i) \right] d\psi. \tag{6.38}
\]

Then, its corresponding IS-MC estimator is given by

\[
\hat{L}(\theta; X, a) = \frac{1}{S} \sum_{j=1}^S \left[ \prod_{i=1}^n \frac{d(X_i, \tilde{\psi}_i^{(j)}(a_i)|X_{i-1}, \tilde{\psi}_{i-1}^{(j)}(a_{i-1}), \theta)}{m(\tilde{\psi}_i^{(j)}(a_i)|\tilde{\psi}_{i-1}^{(j)}(a_{i-1}), a_i)} \right], \tag{6.39}
\]

where \( \{ (\tilde{\psi}_i^{(j)}(a_i)) \}_{i=1}^n \) are trajectories drawn from the auxiliary samplers.

The optimality criterion for choosing the auxiliary samplers is the minimization of the MC variance of (6.39). Relying on the factorized expression of the likelihood, the MC variance minimization problem can be decomposed in a sequence of sub-problems for each element \( i \) of the sequence of observations, provided that the elements depending on the lagged values \( \psi_{i-1} \) are transferred back to the \((i - 1)^{th}\) minimization sub-problem. More precisely, if we decompose \( m \) in the product of a function of \( \psi_i \) and \( \psi_{i-1} \) and one of \( \psi_{i-1} \) only, such that

\[
m(\psi_i|\psi_{i-1}, a_i) = \frac{k(\psi_i, a_i)}{\chi(\psi_{i-1}, a_i)} = \frac{k(\psi_i, a_i)}{\int k(\psi_i, a_i) d\psi_i},
\]
we can set up the following minimization problem:

\[ \hat{a}_i(\theta) = \arg \min_{a_i} \sum_{j=1}^{S} \left\{ \ln \left[ d(X_i, \tilde{\psi}_i^{(j)} | \tilde{\psi}_{i-1}^{(j)}, X_{i-1}, \theta) \chi(\tilde{\psi}_i^{(j)}, \hat{a}_{i+1}) \right] - c_i - \ln(k(\tilde{\psi}_i^{(j)}, a_i)) \right\}^2 \]

(6.40)

where \( c_i \) is constant that must be estimated along with \( a_i \). If the density kernel \( k(\psi_i, a_i) \) belongs to the exponential family of distributions, the problem becomes linear in \( a_i \), and this greatly improves the speed of the algorithm, as a least squares formula can be employed instead of an iterative routine.

The estimated \( \hat{a}_i \) are then substituted in (6.39) to obtain the EIS estimate of the likelihood. The EIS algorithm can be initialized by direct sampling, as in Eq. (6.37), to obtain a first series of \( \tilde{\psi}_i^{(j)} \) and then iterated to allow the convergence of the sequences of \( \{a_i\} \), which is usually obtained after 3-5 iterations. EIS-ML estimates are finally obtained by maximizing \( \tilde{L}(\theta; X, a) \) with respect to \( \theta \). Here we adopt a inverse Gaussian distribution for \( \varepsilon_i \) with parameter \( \lambda \) and a \( N(0, \sigma^2) \) for \( u_i \), we come up with the following expressions:

\[ p(X_i | \psi_{i-1}, \lambda) = \sqrt{\frac{\lambda e^{\psi_i}}{2\pi X_i^3}} \exp \left\{ -\frac{\lambda e^{\psi_i}(X_i - e^{\psi_i})^2}{2(e^{\psi_i})^2 X_i^3} \right\} \]

(6.41)

and

\[ q(\psi_i | \psi_{i-1}, \theta_2) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2}(\psi_i - \omega - \beta \psi_{i-1})^2 \right\} . \]

(6.42)

A convenient choice for the auxiliary sampler \( m(\psi_i, a_i) \) is a parametric extension of the natural sampler \( q(\psi_i | \psi_{i-1}, \theta_2) \), in order to obtain a good approximation of the integrand without too heavy a cost in terms of analytical complexity. Following Liesenfeld and Richard (2003), we can start by the following specification of the
function \( k(\psi_i, a_i) \):

\[
k(\psi_i, a_i) = q(\psi_i|\psi_{i-1}, \theta_2) \zeta(\psi_i, a_i),
\]

where \( \zeta(\psi_i, a_i) = \exp\{a_{1,i}\psi_i + a_{2,i}\psi_i^2\} \) and \( a_i = (a_{1,i} \ a_{2,i}) \). This specification is rather straightforward and has two advantages. Firstly, as \( q(\psi_i|\psi_{i-1}, \theta_2) \) is present in a multiplicative form, it cancels out in the objective function in (6.40), which becomes a least squares problem with \( \ln \zeta(\psi_i, a_i) \) that serves to approximate \( \ln p(x_i|\psi_i, \theta_1) + \ln \chi(\psi_i, a_i) \). Secondly, such a functional form for \( k \) leads to a distribution of the auxiliary sampler \( m(\psi_i, a_i) \) that remains Gaussian, as stated in the following theorem, whose proof is given in Bauwens and Galli (2009).

**Theorem 6.1.** If the functional form for \( q(\psi_i|\psi_{i-1}, \theta_2) \) and \( k(\psi_i, a_i) \) are as in equations (6.42) and (6.43) respectively, then the auxiliary \( m(\psi_i|\psi_{i-1}, a_i) = \frac{k(\psi_i, a_i)}{\chi(\psi_{i-1}, a_i)} \) is Gaussian, with conditional mean and variance respectively given by:

\[
\mu_i = v_i^2 \left( \frac{\omega + \beta \psi_{i-1}}{\sigma^2} + a_{1,i} \right) \quad \text{and} \quad v_i^2 = \frac{\sigma^2}{1 - 2\sigma^2 a_{2,i}},
\]

and the function \( \chi(\psi_{i-1}, a_i) \) is given by

\[
\frac{1}{\sqrt{1 - 2\sigma^2 a_{2,i}}} \exp \left\{ \frac{\sigma^2}{2(1 - 2\sigma^2 a_{2,i})} \left( \frac{\omega + \beta \psi_{i-1}}{\sigma^2} + a_{1,i} \right)^2 - \frac{1}{2} \left( \frac{\omega + \beta \psi_{i-1}}{\sigma} \right)^2 \right\}.
\]

By applying these results, it is possible to compute the likelihood function of the IG-SCD model for a given value of \( \theta \), based upon the following steps:

**Step 1:** Use the natural sampler \( q(\psi_i|\psi_{i-1}, \theta_2) \) to draw \( S \) trajectories of the latent variable \( \{\tilde{\psi}_i^{(j)}\}_{i=1}^n \) as in (6.37).
Step 2: The draws obtained in step 1 are used to solve for each $i$ (in the order from $n$ to 1) the least squares problems described in (6.40), which takes the form of the auxiliary linear regression:

$$
\frac{1}{2} \log \lambda - \log(2\pi) + \frac{1}{2} \log e^{\psi_i} - \frac{3}{2} \log X_i - \frac{\lambda e^{\psi_i}(X_i - e^{\psi_i})^2}{2(e^{\psi_i})^2 X_i} + \ln \chi(\tilde{\psi}_i^{(j)}, \hat{a}_{i+1})
$$

$$
= a_{0,i} + a_{1,i}\tilde{\psi}_i^{(j)} + a_{2,i}(\tilde{\psi}_i^{(j)})^2 + \varepsilon_i^{(i)}, \quad j = 1, 2, \ldots, S,
$$

where $\varepsilon_i^{(i)}$ is the error term, $a_{0,i}$ is the constant term, and $\chi(\tilde{\psi}_i^{(j)}, \hat{a}_{i+1})$ is set equal to 1 for $i = n$ and defined by (6.44) for $i < n$. The reverse ordering from $n$ to 1 is due to the fact that for determining $\hat{a}_i$, $\hat{a}_{i+1}$ is required, see (6.40).

Step 3: Use the estimated auxiliary parameters $\hat{a}_i$ to obtain $S$ trajectories $\{\tilde{\psi}_i^{(j)}(\hat{a}_i)\}_{i=1}^N$ from the auxiliary sampler $m(\psi_i|\psi_{i-1}, \hat{a}_i)$, applying the result of Theorem.

Step 4: Return to step 2, this time using the draws obtained with the auxiliary sampler. Steps 2, 3 and 4 are usually iterated a small number of times (from 3 to 5), until a reasonable convergence of the parameters $\hat{a}_i$ is obtained.

Once the auxiliary trajectories have attained a reasonable degree of convergence, the simulated samples can be plugged in formula (6.39) to obtain an EIS estimate of the likelihood. This procedure is embedded in a numerical maximization algorithm that converges to a maximum of the likelihood function. Throughout the EIS steps described above and their iterations, we employed a single set of simulated random numbers to obtain the draws from the auxiliary sampler. This technique, known as common random numbers, is motivated in Richard and Zhang (2007). The same
random numbers were also employed for each of the likelihood evaluations required by the maximization algorithm. The number of draws used (S in Eq. (6.39)) for all estimations in this article is equal to 100.

6.8 Simulation Study

A simulation study is carried out here in order to evaluate the performance of the estimation methods proposed for ACD and SCD models with inverse Gaussian innovations.

6.8.1 IG-ACD Model

For the IG-ACD (1,1) model (6.15), we performed the simulation experiment for different sample sizes and for different values of \((\omega, \alpha, \beta)\), fixing \(\lambda = 1\). Based on the simulated samples of size \(n= 1000, 2000, 3000\) and \(4000\), we obtained the ML estimates of \(\lambda, \omega, \alpha\) and \(\beta\). We repeated this computation 100 times and took the average value as the final estimate. These estimates are presented in Table 6.1 with corresponding mean square error in the parentheses.

6.8.2 IG-SCD Model

In this sub-section, we carry out a simulation study to evaluate the performance of the EIS-ML estimation method described in Section 6.7 for the IG-SCD model.
<table>
<thead>
<tr>
<th>n</th>
<th>True Values ((\lambda, \omega, \alpha, \beta))</th>
<th>(\hat{\lambda})</th>
<th>(\hat{\omega})</th>
<th>(\hat{\alpha})</th>
<th>(\hat{\beta})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>((1.00,0.10,0.10,0.80))</td>
<td>1.0096</td>
<td>0.1231</td>
<td>0.0931</td>
<td>0.7851</td>
</tr>
<tr>
<td></td>
<td>((1.00,0.10,0.20,0.60))</td>
<td>1.0046</td>
<td>0.1085</td>
<td>0.0988</td>
<td>0.7896</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.00,0.30,0.50))</td>
<td>1.0095</td>
<td>1.0348</td>
<td>0.2961</td>
<td>0.4930</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.50,0.50,0.40))</td>
<td>1.0058</td>
<td>1.5288</td>
<td>0.4930</td>
<td>0.4038</td>
</tr>
<tr>
<td></td>
<td>((1.00,2.00,0.70,0.10))</td>
<td>1.0119</td>
<td>2.0301</td>
<td>0.6905</td>
<td>0.0997</td>
</tr>
<tr>
<td>2000</td>
<td>((1.00,0.10,0.10,0.80))</td>
<td>1.0009</td>
<td>0.1016</td>
<td>0.0988</td>
<td>0.7896</td>
</tr>
<tr>
<td></td>
<td>((1.00,0.10,0.20,0.60))</td>
<td>1.0039</td>
<td>0.5223</td>
<td>0.1982</td>
<td>0.5924</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.00,0.30,0.50))</td>
<td>1.0023</td>
<td>1.0160</td>
<td>0.3018</td>
<td>0.4928</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.50,0.50,0.40))</td>
<td>1.0044</td>
<td>1.5169</td>
<td>0.5001</td>
<td>0.3973</td>
</tr>
<tr>
<td></td>
<td>((1.00,2.00,0.70,0.10))</td>
<td>1.0003</td>
<td>1.9837</td>
<td>0.6982</td>
<td>0.1018</td>
</tr>
<tr>
<td>3000</td>
<td>((1.00,0.10,0.10,0.80))</td>
<td>0.9993</td>
<td>0.1028</td>
<td>0.0978</td>
<td>0.7990</td>
</tr>
<tr>
<td></td>
<td>((1.00,0.10,0.20,0.60))</td>
<td>1.0059</td>
<td>0.5132</td>
<td>0.1990</td>
<td>0.5939</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.00,0.30,0.50))</td>
<td>1.0023</td>
<td>1.0006</td>
<td>0.2993</td>
<td>0.5004</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.50,0.50,0.40))</td>
<td>1.0018</td>
<td>1.5086</td>
<td>0.5019</td>
<td>0.3969</td>
</tr>
<tr>
<td></td>
<td>((1.00,2.00,0.70,0.10))</td>
<td>1.0039</td>
<td>2.0128</td>
<td>0.6986</td>
<td>0.0982</td>
</tr>
<tr>
<td>4000</td>
<td>((1.00,0.10,0.10,0.80))</td>
<td>1.0011</td>
<td>0.1066</td>
<td>0.1006</td>
<td>0.7934</td>
</tr>
<tr>
<td></td>
<td>((1.00,0.10,0.20,0.60))</td>
<td>1.0032</td>
<td>0.5150</td>
<td>0.1995</td>
<td>0.5925</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.00,0.30,0.50))</td>
<td>1.0038</td>
<td>1.0091</td>
<td>0.2987</td>
<td>0.4981</td>
</tr>
<tr>
<td></td>
<td>((1.00,1.50,0.50,0.40))</td>
<td>1.0024</td>
<td>1.5090</td>
<td>0.5018</td>
<td>0.3986</td>
</tr>
<tr>
<td></td>
<td>((1.00,2.00,0.70,0.10))</td>
<td>1.0003</td>
<td>2.0099</td>
<td>0.6997</td>
<td>0.0987</td>
</tr>
</tbody>
</table>

**Table 6.1:** The average ML estimates and the corresponding mean square error for IG-ACD model.
<table>
<thead>
<tr>
<th>$n$</th>
<th>True values $(\omega, \beta, \sigma, \lambda)$</th>
<th>$\hat{\omega}$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\sigma}$</th>
<th>$\hat{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>(0.00, 0.80, 0.50, 2.00)</td>
<td>0.0089 (0.0502)</td>
<td>0.7814 (0.0567)</td>
<td>0.4982 (0.0571)</td>
<td>2.0151 (0.0585)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.70, 0.30, 1.50)</td>
<td>0.0088 (0.0346)</td>
<td>0.6633 (0.0597)</td>
<td>0.2951 (0.0660)</td>
<td>1.5175 (0.0752)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.50, 0.20, 1.00)</td>
<td>0.0079 (0.0382)</td>
<td>0.4755 (0.0793)</td>
<td>0.1850 (0.0805)</td>
<td>1.0164 (0.0756)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.30, 0.05, 0.50)</td>
<td>0.0040 (0.0484)</td>
<td>0.2977 (0.0802)</td>
<td>0.0650 (0.0816)</td>
<td>0.5132 (0.0603)</td>
</tr>
<tr>
<td>2000</td>
<td>(0.00, 0.80, 0.50, 2.00)</td>
<td>0.0098 (0.0507)</td>
<td>0.7794 (0.0523)</td>
<td>0.5010 (0.0554)</td>
<td>2.0098 (0.0606)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.70, 0.30, 1.50)</td>
<td>0.0073 (0.0352)</td>
<td>0.6678 (0.0539)</td>
<td>0.2881 (0.0622)</td>
<td>1.4998 (0.0786)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.50, 0.20, 1.00)</td>
<td>0.0044 (0.0385)</td>
<td>0.4772 (0.0760)</td>
<td>0.1892 (0.0767)</td>
<td>1.0075 (0.0757)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.30, 0.05, 0.50)</td>
<td>0.0035 (0.0504)</td>
<td>0.3008 (0.0778)</td>
<td>0.0626 (0.0802)</td>
<td>0.5120 (0.0597)</td>
</tr>
<tr>
<td>3000</td>
<td>(0.00, 0.80, 0.50, 2.00)</td>
<td>0.0118 (0.0498)</td>
<td>0.7860 (0.0505)</td>
<td>0.5100 (0.0565)</td>
<td>2.0133 (0.0602)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.70, 0.30, 1.50)</td>
<td>0.0065 (0.0348)</td>
<td>0.6749 (0.0540)</td>
<td>0.2974 (0.0649)</td>
<td>1.5006 (0.0811)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.50, 0.20, 1.00)</td>
<td>0.0052 (0.0389)</td>
<td>0.4834 (0.0744)</td>
<td>0.1830 (0.0780)</td>
<td>0.9962 (0.0783)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.30, 0.05, 0.50)</td>
<td>0.0021 (0.0468)</td>
<td>0.3100 (0.0819)</td>
<td>0.0695 (0.0817)</td>
<td>0.5147 (0.0562)</td>
</tr>
<tr>
<td>4000</td>
<td>(0.00, 0.80, 0.50, 2.00)</td>
<td>0.0030 (0.0438)</td>
<td>0.7825 (0.0496)</td>
<td>0.4956 (0.0518)</td>
<td>2.0123 (0.0636)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.70, 0.30, 1.50)</td>
<td>0.0029 (0.0377)</td>
<td>0.6840 (0.0554)</td>
<td>0.2965 (0.0625)</td>
<td>1.5007 (0.0856)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.50, 0.20, 1.00)</td>
<td>0.0051 (0.0312)</td>
<td>0.4864 (0.0733)</td>
<td>0.1908 (0.0747)</td>
<td>0.9981 (0.0828)</td>
</tr>
<tr>
<td></td>
<td>(0.00, 0.30, 0.05, 0.50)</td>
<td>0.0031 (0.0476)</td>
<td>0.3034 (0.0818)</td>
<td>0.0695 (0.0780)</td>
<td>0.5054 (0.0564)</td>
</tr>
</tbody>
</table>

Table 6.2: The average estimates and the corresponding mean square error for the EIS ML estimates
We conducted several repeated simulation experiments with different values of $(\beta, \sigma, \lambda)$, fixing $\omega = 0$. The trajectories of 1000, 2000, 3000 and 4000 observations from a SCD data generating process were simulated 100 times and the model parameters were estimated. These estimates are presented in Table 6.2 with corresponding mean square error in parentheses.

From the table we can see that the EIS-ML method provides estimates which in mean closer to the true parameter values and mean square error of estimates are always remarkably small. The details of computation are given in Appendix E.

The above estimates can be used to evaluate the estimated hazard function through the relation (6.34) to compare with the unconditional empirical hazard function. To obtain the unconditional empirical hazard function we use the relation $\hat{h}(t) = \frac{\hat{f}(t)}{1 - \hat{F}(t)}$, where $\hat{f}(t)$ is the kernel based estimator of the marginal probability density function of the IG-SCD sequence and $\hat{F}(t)$ is the empirical distribution function obtained using the relation $\hat{F}(t) = \int_{0}^{t} \hat{f}(u) \, du$. We have used formula $\hat{f}(t) = \frac{1}{n\Delta t} \sum_{i=1}^{n} \xi \left( \frac{t - T_i}{\Delta t} \right)$ to compute the density estimate in which $\xi(.)$ is the Epanechnikov kernel. See (Silverman (1986), pp 11-13) for details. We computed $\hat{h}(t)$ for the simulated series of IG-SCD sequences for different parameter combinations and compared with the estimated hazard function. Figure 6.2 gives one such graphical comparison, where the dotted line and dashed line represents the unconditional hazard function of the IG-SCD model for different parameter specifications $\omega = 0.10$, $\beta = 0.80$, $\lambda = 2.00$ and its estimated values $\hat{\omega} = 0.1003$, $\hat{\beta} = 0.7895$, $\hat{\lambda} = 2.0145$ respectively. The solid line is the kernel based empirical hazard function of one of the simulated series.
The figure clearly indicates that the empirical hazard function behaves similar to the true as well as the simulation based estimated hazard function.

### 6.9 Data Analysis

In the present section we apply the inverse Gaussian duration models for analysing the real data sets. We consider the data of intraday foreign exchange rates of Australian Dollar vs Canadian Dollar and US Dollar vs Singapore Dollar for the day 25\textsuperscript{th} April, 2012 and 2\textsuperscript{nd} May, 2012 respectively. The data sets are downloaded from the website of Swiss Forex bank. This is the trade book data (tick-by-tick) corresponding to exchange rate of different currencies traded in Swiss Forex bank.
From the traded data, we took the trade entered time (HH:MM:SS) and find the time duration between the consecutive trades in seconds. The zero durations are excluded. There is a strong seasonality in the durations and we adjusted the data to take care of this diurnal pattern of intraday durations using the method described in Tsay (2005).

Let \( f(t_i) \) be the mean value of the diurnal pattern at time \( t_i \). Then define

\[
X_i^* = \frac{X_i}{f(t_i)},
\]

be the adjusted duration and \( X_i \) be the observed duration \( i^{th} \) and \( i-1^{th} \) transactions. We construct \( f(t_i) \) using two simple time functions.

Define

\[
O(t_i) = \begin{cases} 
  t_i - 34200 & \text{if } t_i < 43200 \\
  0 & \text{otherwise},
\end{cases}
\]

and

\[
C(t_i) = \begin{cases} 
  57600 - t_i & \text{if } t_i \geq 43200 \\
  0 & \text{otherwise},
\end{cases}
\]

where \( t_i \) is the time of \( i^{th} \) transaction measured in seconds from midnight and 34200, 43200 and 57600 denote, respectively, market opening, noon and market closing times measured in seconds. Consider the multiple linear regression,

\[
\ln(X_i) = \beta_0 + \beta_1 O(t_i) + \beta_2 C(t_i) + e_i,
\]
Where \( o(t_i) = O(t_i)/10000 \) and \( c(t_i) = C(t_i)/10000 \). Let \( \hat{\beta}_i \) be the ordinary least squares estimates of above linear regression. The residual is then given by

\[
\hat{e}_i = \ln(X_i) - \hat{\beta}_0 - \hat{\beta}_1 o(t_i) - \hat{\beta}_2 c(t_i).
\]

Then \( f(t_i) = \exp\{\hat{e}_i\} \). Using \( f(t_i) \), we obtain the adjusted duration \( X^*_i \).

Figure 6.3 shows the time series plot of adjusted durations.

![Aus Dollar Vs Can Dollar](image)

**Figure 6.3:** Time series plot of adjusted durations.

In Table 6.3 we report the summary statistics for the data sets, where \( Q(10) \) denotes the Ljung-Box statistic of order 10.

The parameters estimated under IG-ACD(1,1) and IG-SCD models using the methods developed in Sections 6.4 and 6.7 respectively are summarized in Table 6.4.
### Table 6.3: Descriptive statistics for the data

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Australian Dollar vs Canadian Dollar</th>
<th>US Dollar vs Singapore Dollar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>2650</td>
<td>6933</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.1765</td>
<td>0.0546</td>
</tr>
<tr>
<td>Maximum</td>
<td>31.9400</td>
<td>175.50</td>
</tr>
<tr>
<td>Mean</td>
<td>1.7290</td>
<td>4.5300</td>
</tr>
<tr>
<td>Median</td>
<td>1.0210</td>
<td>0.7748</td>
</tr>
<tr>
<td>Q(10)</td>
<td>36.7616</td>
<td>70.2127</td>
</tr>
</tbody>
</table>

### Table 6.4: Parameter estimates for the data

<table>
<thead>
<tr>
<th>Parameters</th>
<th>IG-ACD</th>
<th>IG-SCD</th>
<th>IG-ACD</th>
<th>IG-SCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td>1.4734</td>
<td>0.1890</td>
<td>0.4960</td>
<td>0.2910</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.2520</td>
<td>—</td>
<td>0.2527</td>
<td>—</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1574</td>
<td>0.4102</td>
<td>0.2773</td>
<td>0.4507</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>—</td>
<td>0.4312</td>
<td>—</td>
<td>0.6846</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.6019</td>
<td>0.9011</td>
<td>0.2819</td>
<td>0.9801</td>
</tr>
</tbody>
</table>

We now perform a diagnostic check of the models based on the residuals. For IG-ACD model, the standardized residual is defined as $\hat{\varepsilon}_i = x_i / \hat{\psi}_i$. If the fitted ACD model is adequate, then $\{\hat{\varepsilon}_i\}$ should behave as an independent and identically distributed sequence of random variables with the assumed distribution. Since the model assumes that the residuals are independent; any dependence in either the standardized residuals or their squares indicates misspecification of the model. In particular, if the fitted model is adequate, both series $\{\hat{\varepsilon}_i\}$ and $\{\hat{\varepsilon}_i^2\}$ should have no autocorrelations.

Regarding the IG-SCD model, once the estimates of the parameters are obtained,
the unobservable component $\psi_i$ is estimated using Kalman filtering. We define the standardized residuals of the IG-SCD model as $\hat{\varepsilon}_i = x_i / e^{\hat{\psi}_i}$, where $\hat{\psi}_i$ is the estimator of $\psi_i$ provided by the Kalman filter at the SML estimate.

Figure 6.4 shows the time series plot of standardized innovations and Figure 6.5 gives the sample autocorrelation function of the standardized innovations for the fitted IG-ACD(1,1) and IG-SCD models respectively. From the figures, the innovations appear to be random and their ACFs fail to indicate any significant serial dependence.

![Figure 6.4: Time plot of standardized innovation series of IG-ACD(1,1) and IG-SCD models.](image)
Figure 6.5: ACF of the standardized residual series of IG-ACD(1,1) and IG-SCD models.

Figure 6.6 is the histogram of standardized residuals superimposed by the unit mean inverse Gaussian density curve for IG-ACD(1,1) and IG-SCD models. The figures show that the inverse Gaussian distribution is a good approximation for the standardized residuals. Yet, we have to check formally the serial correlations of the series $\{\hat{\varepsilon}_i\}$ and $\{\hat{\varepsilon}^2_i\}$. We adopt Ljung-Box statistics for checking the serial correlations of these two series. The Ljung-Box statistics for the standardized innovations ($Q(.)$) and for the squared innovations ($Q^*(.)$) are calculated and given in Table 6.5. The corresponding Chi-square table values are given in the parenthesis. From the Table 6.5 Ljung-Box statistics for the standardized innovations and the squared innovations are insignificant for both the data sets, so that the fitted models are adequate in describing the dynamic dependence of the adjusted durations.
Finally, we have demonstrated the significance of allowing non-monotonic hazard functions for modelling the conditional durations in financial time series. In Figure 6.7, we have plotted the empirical hazard function of the data and compared it with the unconditional hazard function of IG-SCD model plotted for SML estimates for...
the data. Thus the use of inverse Gaussian distribution is motivated by the fact that it allows for a non-monotonic hazard function which has been found empirically relevant.

**Figure 6.7:** Empirical and estimated hazard functions of adjusted durations of exchange rates data.

A part of this chapter is published in Balakrishna and Rahul (2014).