Volatility Estimators and the Inverse Range Process in a Random Volatility Random Walk and Wiener Processes

Pierre P. Vallois, Charles S. Tapiero

Revised November 2007

Volatility Estimators and the Inverse Range Process in a Random Volatility Random Walk and Wiener Processes

Pierre Vallois
Département de Mathématiques, Université de Nancy I
B.P. 239 - 54506 Vandoeuvre les Nancy Cedex - France

Charles S. Tapiero,
Department of Risk and Financial Engineering,
Polytechnic University of New York,
6 Metro Tech Center, Brooklyn, N.Y. 11201
And ESSEC, France

1. Introduction

Empirical evidence has shown that financial time series are not always "well behaved". They may have an unpredictable variance, underscoring departures from the "random walk hypothesis". These effects have been recognized and have been the subject of considerable research under the heading of ARCH and GARCH related models (Engle, 1987, Bollerslev, 1986, Bollerslev, Engle and Nelson 1994) focusing on the estimation of an underlying process variance. A related effort based on samples range of a constant volatility random walk was pointed out by Parkinson (1980) however. Explicitly, using Feller's result (1957) Parkinson's volatility estimator is given by:

$$\hat{\sigma}^2(p) = \frac{1}{(4\ln 2)Tn}\sum_{i=1}^{n} R_i$$

where $R_i = \max_{i\in I_i} X(t) - \min_{i\in I_i} X(t)$ is the *ith* sample range, *n* is the number of intervals I_i, *i=1,2...,n* over which the range is estimated, *T* is the length of the interval while *X(t)* is assumed to be a normal process with volatility σ. An adjustment to this estimate, based on the transformation of Parkinson's samples is suggested by Kunitomo (1992), providing thereby an estimate which is equivalent to the estimates based on samples variances. Range based estimates in econometric have been studied and applied further by Martens and van Dijk (2007) as well as Alizadeh et al. (2002) and Brandt and Diebold (2006) who claim that the daily range in financial data is a more robust estimator against the effects of microstructure noise than the realized variance. By contrast, in the study of data with long memory, the Hurst index based on a range to standard deviation statistic provide the means to estimate a latent long run memory in long time series (for example, Tapiero and Vallois, 1996).

The purpose of this paper is to provide volatility estimators in stationary random volatility models using instead an inverse range process statistic which will be defined here. The intent of the paper is to indicate the importance of such a (range) process to estimate directly the functional parameters of an unknown underlying volatility and provide statistical estimates for this volatility. Unlike the ARCH-GARCH approach which is used to counter the effects of heteroscedasticity in linear regressions (where volatility is random), and uses specific models of volatility to negate their statistical effects, our approach is based on the information that a range process can provide to

estimate the underlying process volatility. Our approach therefore is not a substitute to ARCH and GARCH estimation techniques but complements it when the volatility is unknown or is indirectly observable. Of course, if the process can be directly observed and measured, such an estimate can be used in addition to the standard maximum likelihood techniques applied in estimating directly the volatility. For example, random (stationary) volatility (and variance) distributions such as the exponential, the lognormal, Cauchy (fat tail) and other distributions can be used as hypotheses regarding an underlying volatility process and the parameters estimated by an observation of the range.

Unlike papers by Feller 1951, Vallois 1996, Vallois and Tapiero 1995, 1997a, 1997b, we assume that the random walk process volatility is characterized by a known probability distribution and calculate its essential statistics (although approximations can be reached when such a distribution is not specified). Both discrete and continuous time observation of a given range may be used in our calculations. As a result, both random walk and random Wiener processes are considered here.

Problems related to volatility (variance) estimation are particularly important in finance. For example, an investor—a buyer of options, may have only a probabilistic assessment of the underlying stock volatility, which is essential for option's pricing. Random volatility tests are also an indicator of markets' incompleteness and therefore important to detect arbitrage opportunities. Similarly, volatility estimates are used in control charts to determine control limits (often based as well on sample ranges which assume that samples are iid). In such cases, process range statistics might be used to test the validity of a given set of control limits. We begin by developing estimators for a random volatility random walk and obtain an explicit expression to estimate the parameters of the volatility distribution. Subsequently, a random volatility Wiener process is considered and specific cases (volatility distributions and estimators) are resolved.

2. A Random Volatility Random Walk

Consider the symmetric random walk, initiated at zero:

$$(1)\quad X_0 = 0, X_n = \sum_{i=1}^{n} \varepsilon_i,\ n \geq 1$$

where $(\varepsilon_i)_{i\geq n}$ is a series of independently distributed random variables with $P(\varepsilon_i = \pm 1) = 1/2$. We define as well the random (volatility) parameter, $\alpha > 0$ independent of the random walk $(X_n)_{n\geq 0}$ and consider the following process:

(2) $X_n^{(\alpha)} = \alpha X_n;\ n \geq 0.$

Such a process implies that the underlying random walk (a price for example) increases or decreases by increments of size α. When α is unknown, we then have a constant volatility process—albeit its parameters are unknown and therefore presumed to be defined by a probability distribution. When α is a stochastic process, say a Bernoulli process assuming values *a* with probability *p* and zero otherwise, the underlying process (2) becomes a trinomial random walk with unknown parameters (*a,p*) that may be determined in the same manner that α is estimated in this paper. However, such extensions and some of their difficulties are discussed in the appendix and provide an area for further research. In particular, while the trinomial random walk is the sum of independent random variables, the random walk with a Bernoulli mixture results in a mixture distribution where the random increments are dependent. For our current purpose, we rewrite (2) as follows:

(3) $X_n^{(\alpha)} = X_{n-1}^{(\alpha)} + \alpha\varepsilon_n$ or (3') $X_n^{(\alpha)} = X_{n-1}^{(\alpha)} + z_n$ where $z_n = \alpha\varepsilon_n$

Where $\alpha > 0$ is assumed to be a square integrable random variable. A priori, the process $(X_n^{(\alpha)}; n \geq 1)$ is not a random walk however, except when $\alpha > 0$ is constant, in which case it assumes the discrete values: $..., -3\alpha, -2\alpha, -\alpha, 0, \alpha, 2\alpha, 3\alpha, ...$ as stated above. Due to the independence of $\alpha > 0$ and $(X_n)_{n\geq 0}$, it is easy to show that $E(X_n^{(\alpha)}) = 0$ and $Var(X_n^{(\alpha)}) = nE(\alpha^2);\ n \geq 1$.

Thus, an unknown volatility, or non-perfectly observable process over a given period of time, can lead to uncertainty regarding a process evolution. Of course, if $\alpha > 0$ is defined by a known probability distribution, the range derived statistics can then be used to test hypotheses regarding the underlying process and estimate the process volatility. The intent of this paper is to provide some results that can be used to such ends.

Consider first the inverse range process $\theta_\alpha(a)$ which denotes the first time the range for a random walk of the type (3) is greater than a given level "a". The volatility being a-priori unknown. Both the first two moments and the probability distributions are determined. Let $\theta(a)$ be the first time the amplitude of process $(X_n)_{n\geq 0}$ crosses the value *a>0*;

$$(4)\quad \theta(a)=\inf\left\{n\geq 0;\max_{0\leq i\leq n}X_i-\min_{0\leq i\leq n}X_i\geq a\right\},\ a\geq 0.$$

In an analogous manner we define:

$$(5)\quad \theta_\alpha(a)=\inf\left\{n\geq 0;\max_{0\leq i\leq n}X_i^{(\alpha)}-\min_{0\leq i\leq n}X_i^{(\alpha)}\geq a\right\},\ a\geq 0.$$

α being positive, then

$$\max_{0\leq i\leq n}X_i^{(\alpha)}=\alpha\max_{0\leq i\leq n}X_i \text{ and } \min_{0\leq i\leq n}X_i^{(\alpha)}=\alpha\min_{0\leq i\leq n}X_i\ .$$

Since $(X_i)_{i\geq 0}$ is a process with values in $\mathbb{Z}$, we have:

$$(6)\quad \theta_\alpha(a)=\begin{cases}\inf\left\{n\geq 0;\max_{0\leq i\leq n}X_i-\min_{0\leq i\leq n}X_i\geq\dfrac{a}{\alpha}\right\}, \text{if } a/\alpha\in\mathbb{N}\\ \inf\left\{n\geq 0;\max_{0\leq i\leq n}X_i-\min_{0\leq i\leq n}X_i\geq\left[\dfrac{a}{\alpha}\right]+1\right\},\ \text{if } a/\alpha\notin\mathbb{N}\end{cases}$$

where [..] denotes the entry whole number. Thus,

$$(7)\quad \theta_\alpha(a)=\begin{cases}\theta\left(\dfrac{a}{\alpha}\right),\ \text{if } \dfrac{a}{\alpha}\in\mathbb{N}\\ \theta\left(\left[\dfrac{a}{\alpha}\right]+1\right),\ \textit{otherwise}\end{cases}$$

Explicitly we can also write:

$$(8)\quad m\alpha<a\leq(m+1)\alpha;\quad \theta_\alpha(a)=\theta(m+1),\ m=0,1,2,3,...$$

With these definitions on hand, we calculate some of the properties of the perturbed random walk inverse range process. We consider first the mean and the variance.

Proposition 1:

Let the amplitude $a>0$, $\alpha>0$ *be the volatility (independent of* $(X_n)_{n\geq 1}$*), and* $F(.)$ *be the function* $F(x)=P(\alpha<x);\ x\geq 0$. *Then:*

$$(9)\quad E\left[\theta_\alpha(a)\right]=\frac{1}{2}\sum_{n\geq 1}n(n+1)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

$$(10)\quad E\left[\theta_\alpha^2(a)\right]=\frac{1}{6}\sum_{n\geq 1}n(n+1)(2n^2+2n-1)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

with the convention that $a/0=\infty$ *and* $F(+\infty)=1$. *Further:*

$$(11)\quad P\left(\theta_\alpha(a)=k\right)=\sum_{n\geq 1}P\left(\theta(n)=k\right)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right],$$

$$(12)\quad P\left(\theta(n)=n+k\right)=\Lambda_{n+1,k-1}-\Lambda_{n,k}$$

with:

$$(13)\quad \begin{aligned}&\Lambda_{n,k}=\left(1+(-1)^{n+k+1}\right)\Gamma_{n,k-1}+\left(1+(-1)^{n+k}\right)\Gamma_{n,k}+\left(1+(-1)^{k+1}\right)\Gamma^{-}_{n,k-1}+\left(1+(-1)^{k}\right)\Gamma^{-}_{n,k}\\&\Gamma_{n,k}=\frac{1}{n}\sum_{0<m<n/2}(\cos(\xi_{n,m}))^{n+k},\ \Gamma^{-}_{n,k}=\frac{1}{n}\sum_{0<m<n/2}(-1)^{m+1}(\cos(\xi_{n,m}))^{n+k},\ \xi_{n,m}=\frac{m\pi}{n}\end{aligned}$$

Proof: (see Proposition 13, Vallois, 1996, p. 1026)

This proposition clearly defines the random (unknown) volatility distribution $P\left(\theta_\alpha(a)=k\right)$ in terms of the non-random volatility distribution $P\left(\theta(n)=k\right)$ calculated by Vallois and Tapiero in several papers (Vallois, 1996, Vallois and Tapiero 1995, 1997a, 1997b) and the a-priori defined probability distribution of the volatility $F(*)$. First, note that calculations of the first two moments can be simplified if we consider instead the distribution of the reciprocal volatility $\beta=\frac{1}{\alpha}$ with distribution function $G(x)=P\left(\beta\leq x\right)$. In this case, we have instead of (9) and (10):

$$(14)\quad E\left[\theta_\alpha(a)\right]=\frac{1}{2}\sum_{n\geq 1}n(n+1)\left[G\left(\frac{n}{a}\right)-G\left(\frac{n-1}{a}\right)\right]$$

$$(15)\quad E\left[\theta_\alpha^2(a)\right]=\frac{1}{6}\sum_{n\geq 1}n(n+1)(2n^2+2n-1)\left[G\left(\frac{n}{a}\right)-G\left(\frac{n-1}{a}\right)\right].$$

And using Proposition 1, we can write as well:

$$(16)\quad \begin{aligned}P\left(\theta_\alpha(a)=k\right)&=\sum_{n\geq 1}P\left(\theta(n)=k\right)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]\\&=\sum_{n\geq 1}P\left(\theta(n)=k\right)\left[G\left(\frac{n}{a}\right)-G\left(\frac{n-1}{a}\right)\right]\end{aligned}$$

For approximation purposes, equations (14) and (15) provide an initial estimate for confidence intervals. For example, using Tchebycheff's inequality, we have:

$$(17)\quad P\left(\left|\theta_\alpha(a)-E\left[\theta_\alpha(a)\right]\right|\geq T\right)\leq\frac{\text{var}\left[\theta_\alpha(a)\right]}{T^2}$$

providing an approximate test periods for the unknown volatility random walk. Further, let the maximum likelihood is defined by the product probabilities of the amount of time that an amplitude a has been reached for the first time, at time $\theta_{\alpha,j}(a)$ which equals the observed sample k_j , with j, denoting the sample number. These probabilities are then given by the product of $P\left(\theta(n)=k_j\right)$ (where a non-volatile random walk will attain the amplitude n for the first time at time k_j, multiplied by the probabilities $\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$. In other words, the likelihood of a random volatility set of samples observed $k_1,k_2,k_3,...,k_M$ is given as follows:

$$(18)\quad \prod_{j=1}^{M}P\left(\theta_{\alpha,j}(a)=k_j\right)=\prod_{j=1}^{M}\sum_{n\geq 1}P\left(\theta(n)=k_j\right)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

Given, $k_1,k_2,k_3,...,k_M$ as well as the theoretically derived probabilities $P\left(\theta(n)=k_j\right)$, the parameters implied in the volatility distribution can then be derived by maximum likelihood. In other words, maximizing the ML below provides alternative estimators to the volatility of the random walk.

$$(19)\quad \underset{F(.)}{Max}\prod_{j=1}^{M}\sum_{n\geq 1}P\left(\theta(n)=k_j\right)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

Of course, this is most likely a difficult expression to optimize. Simple cases can be handled easily while approximations based on moments estimators can also be reached. Using equation (16), we have by have:

$$(20)\quad P\left(\theta_\alpha(a)=k_j\right)=\sum_{n\geq 1}P\left(\theta(n)=k_j\right)\left[G\left(\frac{n}{a}\right)-G\left(\frac{n-1}{a}\right)\right]$$

Where $P\left(\theta(n)=k_j\right)$ is given by (12) where $k_j\geq n$:

$$(21)\quad P\left(\theta(n)=k_j\right)=\Lambda_{n+1,k_j-n-1}-\Lambda_{n,k_j-n}$$

Example 1:

Let β assume an exponential distribution with parameter μ. Then,

$$G(x) = P(\beta \le x) = 1 - e^{-\mu x}$$

After some tedious manipulations (verified as well by Maple), we have:

$$E\left[\theta_\alpha(a)\right] = \frac{1}{\left(1-e^{-\mu/a}\right)^2}; \;\; Var\left[\theta_\alpha(a)\right] = \frac{e^{-\mu/a}\left(6+e^{-\mu/a}\right)}{\left(1-e^{-\mu/a}\right)^4}.$$

Of course, if $\theta_{\alpha,1}(a), \theta_{\alpha,2}(a), ..., \theta_{\alpha,n}(a)$ is a sample of $\theta_\alpha(a)$ then the estimator $\hat{\mu}_n$ for the parameter μ is a solution of:

$$\frac{\theta_{\alpha,1}(a) + \theta_{\alpha,2}(a) + ... + \theta_{\alpha,n}(a)}{n} = \left(1 - e^{-\hat{\mu}_n/a}\right)^{-2}$$

When the distribution of $\theta_\alpha(a)$ is calculated explicitly, we can calculate as well the time rate at which range growth occurs. This rate is given by:

$$\lambda_{a,k}^{(\alpha)} = \frac{P\left(\theta_\alpha(a) = k\right)}{P\left(\theta_\alpha(a) \ge k\right)},$$

where

$$\begin{aligned} P\left(\theta_\alpha(a) = k\right) &= \sum_{n\ge 1} P\left(\theta(n) = k\right)\left(F\left(\frac{a}{n-1}\right) - F\left(\frac{a}{n}\right)\right) \\ &= \sum_{n\ge 1} P\left(\theta(n) = k\right)\left(G\left(\frac{n}{a}\right) - G\left(\frac{n-1}{a}\right)\right) \\ &= \sum_{n\ge 1} P\left(\theta(n) = k\right)\left(e^{-\mu\frac{n-1}{a}} - e^{-\mu\frac{n}{a}}\right) \\ &= \left(e^{-\frac{\mu}{a}} - 1\right)\sum_{n\ge 1} P\left(\theta(n) = k\right) e^{-\mu\frac{n}{a}} \end{aligned}$$

We have :

$$\sum_{n\ge 1} P\left(\theta(n) = k\right) e^{-\mu\frac{n}{a}} = E\left[\sum_{n\ge 1} 1_{\{n = R_k, R_{k-1} < n\}} e^{-\mu\frac{n}{a}}\right] = E\left[1_{\{R_k > R_{k-1}\}} e^{-\frac{\mu R_k}{a}}\right]$$

Thus,

$$P\left(\theta_\alpha(a) = k\right) = E\left[1_{\{R_k > R_{k-1}\}} e^{-\mu\frac{R_k}{a}}\right]$$

Furthermore,

$$\sum_{n\ge 1} P\left(\theta_\alpha(a) \ge k\right) = \sum_{i\ge k} P\left(\theta_\alpha(a) = i\right) = \sum_{i\ge k} E\left[1_{\{R_i > R_{i-1}\}} e^{-\mu\frac{R_i}{a}}\right]$$

In these expressions, note that (see Figure 1 as well):

$$\left\{R_k > R_{k-1}\right\} = \left\{X_{k-1} = S_{k-1} \text{ and } X_k = 1 + X_{k-1}\right\} \cup \left\{X_{k-1} = I_{k-1} \text{ and } X_k = X_{k-1} - 1\right\}$$

With

$$S_k = \max_{i \le k} X_i, I_k = \min_{i \le k} X_i$$

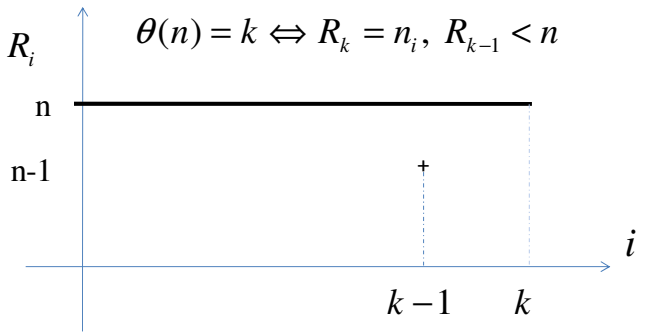

Figure 1

Example 2:

Say that the volatility is distributed between $[1/b, \infty]$; $b \ge 1$ and β assumes a uniform distribution on $[0, b]$; being an integer, $b \ge 1$. Then:

$$If \quad \frac{n}{a} \le b, \quad G\left(\frac{n}{a}\right) = \frac{n}{ab}$$

$$If \quad \frac{n}{a} > b, \quad G\left(\frac{n}{a}\right) = 1$$

Note that $G\left(\frac{n}{a}\right) - G\left(\frac{n-1}{a}\right) = \frac{1}{ab};\ 1 \le n \le ab$. It can be verified that:

$$E\left[\theta_\alpha(a)\right] = \frac{1}{2}\sum_{n=1}^{ab} n(n+1)\left[\frac{n}{ab} - \frac{n-1}{ab}\right] = \frac{(ab+1)(ab+2)}{6}$$

Similarly for the variance, we have (verified by Maple as well):

$$Var\left[\theta_\alpha(a)\right] = \frac{(ab-1)(ab+1)(ab+2)(7ab+16)}{180}.$$

Reasoning as in Example 1, it can be proved that:

$$P\left(\theta_\alpha(a) = k\right) = \frac{1}{ab} P\left(R_{k-1} \le R_k \le ab\right)$$

Which can be used to calculate as shown previously, $\lambda_{n,k}^{(\alpha)} = \frac{P\left(\theta_\alpha(n) = k\right)}{P\left(\theta_\alpha(n) \ge k\right)}$. Again, as in Example 1, note that:

$$\frac{\theta_{\alpha,1}(a) + \theta_{\alpha,2}(a) + \ldots + \theta_{\alpha,n}(a)}{n} = \frac{(1 + a\hat{b}_n)(2 + a\hat{b}_n)}{6}$$

where $\hat{b}_n$ is an estimator of *b*. If we consider the case $\alpha = 1$ or 2 we can also verify that :

$$E\left[\theta_\alpha(2a)\right] = E\left[1_{\{\alpha=1\}}\theta(2a)\right] + E\left[1_{\{\alpha=2\}}\theta(a)\right] =$$
$$= E\left[\theta(2a)\right]P(\alpha=1) + E\left[\theta(a)\right]P(\alpha=2)$$
$$= a(2a+1)P(\alpha=1) + \frac{a(a+1)}{2}P(\alpha=2)$$

Set $p_1 = P(\alpha=1)$ for convenience. Since $P(\alpha=2) = 1 - P(\alpha=1) = 1 - p_1$, we have :

$$E[\theta_\alpha(2a)] = a(2a+1)p_1 + \frac{a(a+1)}{2}(1-p_1)$$
$$= \frac{a(a+1)}{2} + p_1\frac{a(3a+1)}{2}$$

And therefore an estimate $\hat{p}_1^n$ for the probability p_1 is given by:

$$\frac{\theta_{\alpha,1}(a) + \theta_{\alpha,2}(a) + ... + \theta_{\alpha,n}(a)}{n} = \frac{a(a+1)}{2} + \frac{a(3a+1)}{2}\hat{p}_1^n.$$

If the volatility parameter assumes k values, our analysis will be more complex, however if we let the amplitude a varies, we can obtain a linear system of equations by using the probabilities $\left(P(\alpha=\alpha_1), P(\alpha=\alpha_2), ..., P(\alpha=\alpha_k)\right)$ of the volatility.

3. The Random Volatility Wiener Process

Extension to a Wiener processes can be pursued in a similar manner. Imhof (1985) has calculated the probability generating function $E\left(e^{-\lambda\theta(a)}\right)$ of the range process explicitly. Thus, if we consider a random volatility Wiener process, the inverse range process is defined by $\theta_\alpha(a) = \theta(\beta a)$. Noting that in distribution (Vallois, 1996):

(21) $\theta(\beta a) \overset{(d)}{=} \beta^2\theta(a)$

We also note that the Laplace transform is (Imhof, 1985):

(22) $E\left(e^{-\lambda\theta(\beta a)}\right) = E\left(e^{-\lambda\beta^2\theta(a)}\right) = \frac{1}{\left[ch(a\beta\sqrt{\lambda/2})\right]^2}$, $\lambda \geq 0$,

and $E\left(\theta(a)\right) = \frac{a^2}{2}$; $\text{var}\left(\theta(a)\right) = \frac{a^4}{12}$. As a result, the mean and the variance of the time to attain a range *a*, is:

(23) $E\left(\theta_\alpha(a)\right) = \frac{a^2}{2}E\left(\beta^2\right)$.

Further, since $E\left(\theta_\alpha{}^2(a)\right)=E\left(\beta^4\theta^2(a)\right)$, additional manipulations lead to:

$$(24)\quad Var(\theta_\alpha(a))=\frac{a^4}{12}\left[4E\left(\beta^4\right)-3\left(E\left(\beta^2\right)\right)^2\right]$$

The cumulative distribution given by:

$$(25)\quad P\left(\theta_\alpha(a)\le T\right)=P\left(\theta\left(\frac{a}{\alpha}\right)\le T\right)=P\left(\frac{1}{\alpha^2}\theta(a)\le T\right)$$

can be determined explicitly. These results may of course be used to estimate the parameters of a random volatility in a Wiener process. Explicitly, we note that:

$$(26)\quad P\left(\frac{\theta(a)}{T}\le\alpha^2\right)=1-E\left[F_{\alpha^2}\left(\frac{\theta(a)}{T}\right)\right]$$

where $F_{\alpha^2}(.)$ is the cumulative probability distribution of α^2 which can be used to obtain maximum likelihood estimates for the random volatility distribution parameters. We shall consider first an example.

Example 3:

If we let α^2 be an exponential probability distribution with mean μ. We have in this case:

$$P\left(\frac{\theta(a)}{T}\le\alpha^2\right)=1-E\left[F_{\alpha^2}\left(\frac{\theta(a)}{T}\right)\right]=\exp\left(-\mu\frac{\theta(a)}{T}\right)$$

And therefore, using the Laplace transform of the inverse range process $\theta(a)$, we obtain:

$$P\left(\theta_\alpha(a)\le T\right)=\left[ch\left(a\sqrt{\frac{\mu}{2T}}\right)\right]^{-2}$$

And as a result, the density function of the random variable $\theta_\alpha(a)$ is:

$$f_{\theta_\alpha(a)}(T)=\left[ch\left(a\sqrt{\frac{\mu}{2T}}\right)\right]^{-3}\left(\frac{a}{2}\sqrt{\frac{\mu}{2}}\right)T^{-3/2}$$

However, using the cumulative distribution calculated above, an estimator associated to the sample of size n, $\theta_{\alpha,1}(a),\theta_{\alpha,2}(a),...,\theta_{\alpha,n}(a)$ will provide for a fixed T, the following result :

$$\frac{1}{n}\sum_{i=1}^{n}1_{\{\theta_{\alpha,i}(a)\le T\}}=\left[ch\left(a\sqrt{\frac{\hat{\mu}_n}{2T}}\right)\right]^{-2}$$

We may assume of course other distributions. For example, volatility might be lognormally distributed, Weibull, Cauchy (with infinite variance) etc. The resulting problems will be technically more difficult but these problems problem can be handled numerically.

4. Discussion and Applications

In this paper we have calculated the statistics of an unknown volatility random walk, but with known probability distribution. The volatility albeit unknown is determined when the process can be observed. If the random volatility has a discrete probability distribution, this is equivalent to saying that the underling process may be a binomial or multinomial process of various orders. For example if $\alpha = (1, 2)$, each with an appropriate probability, (implying that the underlying process is a simple binomial or a quadrinomial model). Randomizing a simple binomial model by two random values leads of course, to a trinomial model. Explicitly, set $\alpha = (\sigma, 0)$ with probability p. Then our equation (3) is reduced to:

$$X_n^{(\alpha)} = X_{n-1}^{(\alpha)} + \sigma\zeta_n$$

where ζ_n is a trinomial random walk defined by :

$$\zeta_n = \begin{cases} +1 & wp \quad \frac{1}{2}p \\ 0 & wp \quad 1-p \\ -1 & wp \quad \frac{1}{2}p \end{cases}$$

Such a process and its inverse range process has been studied extensively by Vallois (1996) and Vallois and Tapiero (2001) and its results carry over to the case treated here (namely of unknown volatility). Such a case requires however further research due to the dependence which is introduced by the unknown volatility.

When α has a continuous probability distribution, the number of potential and underlying processes is very large. In this sense, the inverse-range process can be used as a statistic to reveal the process departure from an underlying process with a specified volatility.

Appendix 1: Proof of Proposition 1

α being positive, relation (7) implies:

$$E\left[\theta_\alpha(a)\right]=\sum_{n\geq 1}E\left[\theta(n)\,1_{\left\{\frac{a}{\alpha}=n\right\}}\right]+\sum_{n\geq 1}E\left[\theta(n)\,1_{\left\{n-1<\frac{a}{\alpha}<n\right\}}\right]$$

Since α is independent of $(X_n)_{n\geq 1}$, then:

$$E\left[\theta_\alpha(a)\right]=\sum_{n\geq 1}E\left[\theta(n)\right]P\left(\frac{a}{\alpha}=n\right)+\sum_{n\geq 1}E\left[\theta(n)\right]P\left(n-1<\frac{a}{\alpha}<n\right)$$

Or

$$E\left[\theta_\alpha(a)\right]=\sum_{n\geq 1}E\left[\theta(n)\right]P\left(n-1<\frac{a}{\alpha}\leq n\right)$$

But

$$\left\{n-1<\frac{a}{\alpha}\leq n\right\}=\left\{\frac{a}{n}\leq\alpha<\frac{a}{n-1}\right\};\ n\geq 1$$

$$\left\{0<\frac{a}{\alpha}\leq 1\right\}=\left\{\alpha\geq a\right\}$$

Let F be the function $F(x)=P(\alpha<x);x\geq 0.$ Recall that the distribution function of α is $x\to P(\alpha\leq x)$. It coincides with F if α has no atom. According to Vallois (1996):

$$E\left[\theta(a)\right]=\frac{a(a+1)}{2}$$

Consequently,

$$E\left[\theta_\alpha(a)\right]=\frac{1}{2}\sum_{n\geq 1}n(n+1)\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

By the same token,

$$E\left[\theta_\alpha(a)^2\right]=\sum_{n\geq 1}E\left[\theta(n)^2\right]\left[F\left(\frac{a}{n-1}\right)-F\left(\frac{a}{n}\right)\right]$$

But,

$$E\left\{\theta(n)^2\right\}=\frac{n(n+1)(2n^2+2n-1)}{6}$$

and therefore:

$$E\left[\theta_\alpha(a)^2\right] = \frac{1}{6}\sum_{n\geq 1} n(n+1)(2n^2+2n-1)\left[F\left(\frac{a}{n-1}\right) - F\left(\frac{a}{n}\right)\right]$$

The distribution of $\theta_\alpha(a)$ is similarly given by:

$$P\left(\theta_\alpha(a)=k\right) = \sum_{n\geq 1} P\left(\theta(n)=k\right)\left[F\left(\frac{a}{n-1}\right) - F\left(\frac{a}{n}\right)\right]$$

Recall (Vallois, 1996, Proposition 15 and Vallois and Tapiero (1997, p.331):

$$P\left(\theta(n)\geq k+1\right) = \Gamma_{n,k} - \Gamma_{n-1,k}$$

with:

$$\Gamma_{n,k} = \frac{1}{n}\sum_{\substack{0<m<n/2 \\ \varepsilon=\pm 1}} \varepsilon^{k+1-n}\frac{\varepsilon^n + (-1)^{m+1}}{(1-\varepsilon\cos(\xi_{n,m}))^2}\cos^{k-1}(\xi_{n,m})\sin^2(\xi_{n,m});\ \xi_{n,m} = m\pi/n$$

Proposition 1 follows immediately.

□

References

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002, Range based estimation of stochastic volatility models, *Journal of Finance*, 47, 1047-1092.

Bollerslev, T. (1986), `Generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., R. F. Engle and D. B. Nelson (1994), `ARCH models', In R. F. Engle and D. L. McFadden (eds), *Handbook of Econometrics* IV, Elsevier Science, Amsterdam, pp. 2961-3038.

Brandt, M.W. Diebold, F.X., 2006, A no-arbitrage approach to range-based estimation of return covariances and correlations, *Journal of Business*, 79, 61-74.

Engle R., 1987, Autoregressive Conditional Heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica*, 987-1008

Feller W., 1951, The asymptotic distribution of the range of sums of independent random variables, *Annals of Math. Stat*., 22, 427-432

Imhoff J.P. 1985, On the range of Brownian motion and its inverse process, *Ann Prob*.13, 3, 1011-1017

Kunitomo N., 1995, Improving the Parkinson method of estimating security price volatilities, *Journal of Business*, 65, no.2, 295-302.

Martens M. and D. van Dijk, 2007, Measuring volatility with the realized range, *Journal of Econometrics*, 138, 181-207

Parkinson M., 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business*, 53, no.1, 61-65

Tapiero C.S. and P. Vallois, 1996, Run length statistics and the Hurst exponent in random and birth-death random walks, *Chaos, Solitons and Fractals,* 7, no. 9, 1333-1341

Vallois P., 1996, The range of a simple random walk on Z . *Advances in Applied Probability*, 28, 1014-1033

Vallois P. and C.S. Tapiero, 1995, Moments of an amplitude process in a random walk, *Recherche Operationnelle/Operation Research (RAIRO)*, vol. 29, no.1, 1-17

Vallois P. and C.S. Tapiero, 1997a, Range Reliability in Random Walks, *Mathematical Methods of Operational Research*, 45, 325-345

Vallois P. and C.S. Tapiero, 1997b, The range process in random walks: Theoretical results and applications, in *Advances in Computational Economics*, Ed. H. Ammans, B. Rustem and A. Whinston, Kluwer Publications, 291-307

Vallois P. and C.S. Tapiero, 2001, The Range Inter-Event Process in a Symmetric Birth Death Random Walk, *Applied Stochastic Models in Business and Industry*.

ABSTRACT

The purpose of this paper is to study the mean, the variance, the probability distribution and the hazard rate of the inverse range process of an a-priori unknown volatility random walk. Motivation for this process arises when it is necessary to obtain statistics that pertain to a process volatility in addition to the usual variance statistics. As a result, range process statistics are indicated as an additional source of information in the study of processes' volatility. Examples and applications are considered.