Comparative statistics of Garman-Klass, Parkinson, Roger-Satchell and bridge estimators

S. Lapinova\(^1\)* and A. Saichev\(^2\)

Abstract: Comparative statistical properties of Parkinson, Garman-Klass, Roger-Satchell and bridge oscillation estimators are discussed. Point and interval estimations, related with mentioned estimators are considered.

Subjects: Environment & Economics; Statistics; Statistical Physics

Keywords: volatility estimator; Garman-Klass estimator; Bridge estimator; Parkinson estimator; comparative statistics of estimators

1. Examples of volatility estimators
Consider dependence on time \( t \) of the price \( P(t) \) of some financial instrument. As a rule, at discussing of volatility, one consider its logarithm

\[ X(t) = \ln P(t). \]

Let us point out one of the conventional volatility \( V(T) \) definition, which we are using in this paper: It is the variance

\[ V(T) = \operatorname{Var}[Y(t, T)] = E\left[Y^2(t, T)\right] - E[Y(t, T)]^2 \]  

\( (1) \)

of the log-price increment \( Y(t, T) = X(t + T) - X(t) \) within given time interval duration \( T \).

Recall, Garman and Klass, (G&K) (1980), Parkinson, (PARK) (1980) and Rogers and Satchell, (R&S) (1991) volatility estimators are resting on the high and low values:

\[ H = \sup_{t \in (0, T)} Y(t, t'), \quad L = \inf_{t \in (0, T)} Y(t, t') \]  

\( (2) \)

Accordingly, PARK estimator is equal to

\[ \hat{V}_p = \frac{(h - L)^2}{\ln 16}, \]  

\( (3) \)

ABOUT THE AUTHORS
S. Lapinova is an associate professor of the Department of Economics National research University “Higher school of economics”, Russia. Lapinova’s research interests include stochastic processes in finance, volatility estimates, stochastic net structures, multidimensional models of volatility, time series.

A. Saichev (1946–2013) is a professor at ETH Zurich – Department of Management, Technology and Economics, Switzerland. Saichev’s research interests include nonlinear mediums, stochastic processes in physics and in finance, volatility estimates, diffusion processes time series.

PUBLIC INTEREST STATEMENT
The volatility estimation is an important problem in finance. The level of volatility is a signal to trade for speculators. It is more important factor than a direction of trend. A cause of existence of some estimators of volatility is the principal differences of calculation of parameters, efficient of the estimation, what is depended from market situation. A goal of that paper is to compare parameters of main volatility estimators.
while G&K estimator given by expression

\[
\hat{V}_g = k_1(H - L)^2 - k_2(C(H - L) - 2HL) - k_3C^2, \tag{4}
\]

\[k_1 = 0.511, \quad k_2 = 0.0109, \quad k_3 = 0.383.\]

Here \(C := Y(t, T)\) is the close value of the log-price increment. Recall else R&S estimator, equal to

\[
\hat{V}_r = H(H - C) + L(L - C). \tag{5}
\]

Besides mentioned well-known estimators, we discuss bridge oscillation estimator. Below, we call it shortly by bridge estimator. Before to define it, recall bridge \(Z(t, t')\) stochastic process definition. It is equal to

\[
Z(t, t') = Y(t, t') - \frac{t'}{T}Y(t, T), \quad t' \in (0, T). \tag{6}
\]

Let high and low of the bridge be introduced:

\[
H = \max_{t' \in (0, T)} Z(t, t'), \quad L = \min_{t' \in (0, T)} Z(t, t'). \tag{7}
\]

Accordingly, mentioned above bridge volatility estimator given by

\[
\hat{V}_b = \kappa(H - L)^2 \tag{8}
\]

The value of the factor \(\kappa\) will be calculated later.

2. Geometric Brownian motion

One of the conventional models of price stochastic behavior is geometric Brownian motion (see Cont & Tankov, 2004; Jeanblanc, Yor, & Chesney, 2009; Saichev, Malevergne, & Sornette, 2010). In particular, it is used in theoretical justification of G&K, PARK and R&S estimators. Below we discuss statistics of mentioned volatility estimators in frame of geometric Brownian motion model. Namely, we assume that increment of the log-price is of the form

\[
Y(t, T) = \mu T + \sigma B(T). \tag{9}
\]

Here \(\mu\) is the drift of the price, while \(B(t)\) is the standard Brownian motion \(B(t) \sim N(0, t)\). Factor \(\sigma^2\) is the intensity of the Brownian motion.

Recall, Brownian motion posses by self-similar property

\[
B(\tau) \sim \sqrt{T}B\left(\frac{\tau}{T}\right), \quad \forall T > 0, \tag{10}
\]

where and below sign \(\sim\) means identity in law.

Using pointed out self-similar property, one can ensure that

\[
Y(t, t') \sim \sigma\sqrt{T}x(\tau, \gamma), \tag{11}
\]

\[x(\tau, \gamma) = \gamma \tau + B(\tau), \quad \gamma = \frac{\mu}{\sigma}\sqrt{T}, \quad \tau = \frac{t'}{T} \in (0, 1).\]

Henceforth we call process \(x(\tau, \gamma)\) by canonical Brownian motion, while factor \(\gamma\) by canonical drift.

Using relations (3), (4), (8) and (11), one find that
We have used above canonical estimators:

\[ \hat{\varepsilon}_p(\gamma), \hat{\varepsilon}_g(\gamma), \hat{\varepsilon}_r, \hat{\varepsilon}_b \]

containing high, low and close values

\[ \hat{\varepsilon}_p(\gamma) = \frac{d^2}{\ln 16}, \quad \hat{\varepsilon}_b = \kappa s^2, \quad d = h - l, \quad s = \xi - \zeta, \]

\[ \hat{\varepsilon}_g(\gamma) = k_1 d^2 - k_2 (cd - 2hc) - k_3 c^2, \quad h = \sup_{r \in [0,1]} x(r,\gamma), \quad l = \inf_{r \in [0,1]} x(r,\gamma), \quad c = x(1,\gamma), \]

of canonical Brownian motion, and high and low values

\[ \hat{\varepsilon}_r = h(h - c) + l(l - c), \quad \xi = \sup_{r \in [0,1]} z(r), \quad \zeta = \inf_{r \in [0,1]} z(r), \]

of the canonical bridge

\[ z(\tau) = x(\tau,\gamma) - \tau x(1,\gamma) = B(\tau) - \tau \cdot B(1), \quad \tau \in (0,1). \]

Plots of the typical paths of the canonical Brownian motion \( x(\tau,\gamma) \) (11) for \( \gamma = 1 \) and corresponding canonical bridge \( z(\tau) \) (15) are given in Figure 1.

It is worthwhile to note that the closer expected values of canonical estimators \( \hat{\varepsilon}_g(\gamma), \hat{\varepsilon}_p(\gamma), \hat{\varepsilon}_r, \hat{\varepsilon}_b \) to unity, the less biased corresponding original volatility estimators. Analogously, the smaller variances of canonical estimators the more efficient original volatility estimators \( \hat{V}_p, \hat{V}_g, \hat{V}_r, \hat{V}_b \).

Notice additionally that canonical drift \( \gamma \) of the canonical Brownian motion \( x(\tau,\gamma) \) (11) is, as a rule, unknown. Nevertheless, to get some idea about dependence on drift \( \mu \) of bias and efficiency of
volatility estimators, we will discuss below in details dependence of canonical estimators statistical properties on possible values of the factor $\gamma$.

3. Comparative efficiency of PARK and bridge estimators

Resting on, given at Appendix A, analytical formulas for probability density functions (pdfs) of random variables (13) and (14), we explore in this section some statistical properties of canonical PARK estimator $\hat{\sigma}_p(\gamma)$ and bridge one $\hat{\sigma}_b(\gamma)$.

Let unbiasedness of canonical PARK estimator be checked. To make it, let mean square of oscillation $d^2 = h - l$ of the canonical Brownian motion $x(\tau, \gamma)$ at the zero canonical drift ($\gamma = 0$) be calculated with the help of pdf $q_x(\delta)$ (A7). After simple calculations obtain

$$E[d^2] = 2 + \sum_{m=1}^{\infty} \frac{2}{m(4m^2 - 1)} = \ln 16.$$  \hfill (16)

From here and from expression (12) of canonical PARK estimator $\hat{\sigma}_p(\gamma)$ one can see that the following expression is true

$$E[\hat{\sigma}_p(\gamma = 0)] = 1$$

Let us find now the factor $\kappa$ at expressions (8) and (12). To make it, calculate first of all the mean square of the bridge oscillation. Due to expression (A9)

$$E[s^2] = \sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{6}.$$  

Accordingly, unbiased canonical bridge estimator has the form

$$E[\hat{\sigma}_b] = 1 \Rightarrow \kappa = \frac{1}{E[s^2]} \Rightarrow \hat{\sigma}_b = \frac{6s^2}{\pi^2}.$$  \hfill (17)

The great advantage of the bridge estimator is its unbiasedness for any drift. This remarkable property of the pointed out estimator is the consequence of the fact that bridge $Z(t, t')$ (6) and its
canonical counterpart \( z(\tau) \) don’t depend on the drift \( \mu \) (canonical drift \( \gamma \)) at all. On the contrary, PARK estimator becomes essentially biased at nonzero drift. In Figure 2 depicted dependence on \( \gamma \) of canonical PARK estimator expected value, illustrates bias of PARK estimator at nonzero drift. Corresponding curve is obtained with the help of analytical expression (A6) for canonical bridge oscillation pdf.

Figure 3. Plots of dependence on \( \gamma \) of canonical PARK estimator variance. Straight line is the variance of canonical bridge estimator.

Figure 4. Plot of relative bias (20) of canonical PARK estimator as function of canonical drift \( \gamma \).

Let variances of canonical PARK and bridge estimators be calculated. After substitution into the rhs of expression

\[
E \left[ \hat{\nu}_P^2(\gamma = 0) \right] = \frac{1}{\ln^2 16} \int_0^\infty \delta^4 q_\gamma(\delta) d\delta
\]

the sum (A7) for the canonical Brownian motion oscillation pdf \( q_\gamma(\delta) \), and after summation obtain for \( \gamma = 0 \):
Accordingly, variance of canonical PARK estimator $\hat{\alpha}_p$ is

$$E\left[\hat{\alpha}_p^2(\gamma = 0)\right] = \frac{9\zeta(3)}{\ln^2 16} \approx 1.40733.$$  

As the next step, we calculate variance of canonical bridge estimator $\hat{\alpha}_b(17)$. Sought variance is equal to

$$\text{Var}[\hat{\alpha}_b] = \frac{36\pi^2}{\pi^2} E[s^4] - 1.$$  

After substitution here, following from (A9), relation

$$E[s^4] = \int_0^2 \delta^4 q_b(\delta) d\delta = 3 \sum_{m=1}^{\infty} \frac{1}{m^4} = \pi^4 / 30,$$

obtain

$$\text{Var}[\hat{\alpha}_b] = \frac{5}{6} - 1 = 0.2.$$  

Comparing equalities (18) and (19), one can see that variance of bridge estimator approximately twice smaller than variance of PARK estimator.

Recall, variance of bridge estimator does not depend on drift. On the contrary, variance of PARK estimator essentially depends on the drift. One can see it in Figure 3, where depicted plot of dependence, on canonical drift $\gamma$, of canonical PARK estimator variance.

Notice that bias of some estimator is insignificant only if it is much smaller than rms of corresponding estimator, i.e. the relative bias is small:

$$\rho = \frac{E[\hat{\alpha}(\gamma)] - 1}{\sqrt{\text{Var}[\hat{\alpha}(\gamma)]}}.$$  

Plot of canonical PARK estimator relative bias, as function of canonical drift $\gamma$ depicted in Figure 4.

### 4. Interval estimations on the basis of PARK and bridge estimators

Given at Appendix analytical expressions (A6), (A7) and (A9) for canonical Brownian motion and canonical bridge random oscillations pdfs

$$W_b(x; \gamma) = \sqrt{\frac{a}{4x}} q_b\left(\sqrt{ax}; \gamma\right), \quad a = \ln 16.$$  

Similarly, pdf of canonical bridge estimator is equal to

$$W_b(x) = \sqrt{\frac{a}{4x}} q_b\left(\sqrt{ax}\right), \quad a = \frac{\pi^2}{6}.$$  

Here $q_b(\delta)$ (A9) is the pdf of canonical bridge oscillation. Plots of canonical PARK estimator pdf, for $\gamma = 0$, and pdf of canonical bridge estimator are depicted in Figure 5. In Figure 6 pdfs of canonical PARK estimator, for $\gamma = 1$, and pdf of canonical bridge estimator are compared. It is seen in both figures that pdf of canonical bridge estimator is better concentrated around its expected value $E\left[\hat{\alpha}_b\right] = 1$ than canonical PARK estimator pdf.
Knowing estimators pdfs, one can produce interval estimations of possible volatility values. Consider typical interval estimation: Let $\hat{V}$ is some volatility estimator, equal to

$$\hat{V} = V(T) \cdot \hat{\delta}.$$  

(23)

Here $\hat{\delta}$ is corresponding canonical estimator, while $V(T)$ is the measured volatility. One needs to find probability

$$F(N) = \Pr\{V(T) < N \cdot \hat{V}\}$$

that unknown (random) volatility $T$ is not more than $N$ times exceeds known (measured) volatility estimated value $\hat{V}$. It follows from (23) that following inequalities are equivalent:
Last means in turn that sought probability $F(N)$ is expressed through pdf of canonical estimator $\hat{\alpha}$ by the following way:

$$F(N) = Pr\left\{ \hat{\alpha} > 1/N \right\} = \int_{1/N}^{\infty} W(x) dx$$

(24)

here $W(x)$ is the pdf of canonical estimator $\hat{\alpha}$.

Calculations, resting on relations (21), (22), (24) give probability $F_b(2) \approx 0.918$ that true volatility is less than twice of given bridge volatility estimator value $\hat{V}_b$. It is substantially larger than analogous probability in the case of PARK estimator: $F_p(2, \gamma = 0) \approx 0.813$.

Plots of probabilities $F(N)$ (24) dependence on the level $N$, for PARK estimator (in the case of zero drift $\mu = 0$) and for bridge volatility estimator are given in Figure 7.

5. Comparative statistics of canonical estimators

Above, we explored in detail statistical properties of two, PARK and bridge estimators. Here, we compare their statistics and statistics of other well-known volatility estimators: G&K and R&S one. Despite the previous chapters, where we have used known analytical expressions for pdfs of canonical PARK and the bridge estimators, below we use predominantly results of numerical simulations.

Namely, we produce $M \gg 1$ numerical simulations of random sequences

$$x_n(\gamma) = \gamma \frac{n}{N} + \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \varepsilon_n, \quad n = 0, 1, \ldots, N, \quad x_0(\gamma) = 0,$$

(25)

where $(\varepsilon_n)$ are iid Gaussian variables $\sim N(0,1)$. Notice that stochastic process $x_n(\gamma)$ of discrete argument $n$ rather accurately approximates, for large $N \gg 1$, paths of canonical Brownian motion $x(\tau, \gamma)$ (11).
Knowing $M$ iid sequences $\{x_n(\gamma)\}$ one can find corresponding iid samples of pointed out above canonical estimators. Everywhere below we take number of iid samples $M$ and discretization number $N$ equal to $5 \cdot 10^4$, $M = 5 \cdot 10^5$.

Plots in Figure 8 demonstrate rather convincingly accuracy of numerical simulations. In Figure 9 are given two hundred samples of canonical G&K and bridge estimators, ensuring “by naked eye” that canonical bridge estimator is more efficient than G&K one.

Figure 8. Upper panel: Histogram of $M$ samples of canonical bridge estimator $\hat{\sigma}_b$. Solid line is the plot of canonical bridge estimator’s pdf, given by analytical expression (22), (A9). Dashed line is the pdf of canonical PARK estimator for $\gamma = 0$. Lower panel: Histogram of $M$ samples of canonical G&K estimator $\hat{\sigma}_g$ for $\gamma = 0$. Solid line is the plot of the canonical bridge estimator pdf. Dashed line is the canonical PARK estimator pdf for $\gamma = 0$.

Figure 9. Plots of two hundreds samples of canonical estimators. Up to down are samples of G&K, R&S, bridge and PARK estimators. It is seen even by “naked eye” that bridge estimator estimates volatility more accurately than any other mentioned estimators.
In Figure 10 are given, obtained by numerical simulations, plots of canonical G&K, PARK, R&S and bridge estimators mean values, illustrating bias of G&K and PARK estimators for nonzero canonical drift $\gamma = 0.6$, and actual absence of bias for bridge and R&S estimators Figure 11.
Eventually, in Figure 12 are given plots of probabilities that true volatility $V(T)$ is larger than half of corresponding estimator value and less than twice of it:

$$P_\Delta = \Pr\left\{ \frac{V}{2} < V(T) < 2\hat{V} \right\} = \int_{1/2}^{2} W(x)dx. \quad (26)$$

It is seen that for any $\gamma$ mentioned probability is essentially larger for bridge estimator, than for G&K, R&S and PARK estimators.

Acknowledgments

We are grateful for scientific and financial help of Higher school of economics (Russia, Nizhny Novgorod).

Funding

The authors received no direct funding for this research.

Author details

S. Lapinova
E-mail: lapinovas@gmail.com
ORCID ID: http://orcid.org/0000-0002-4127-8135

A. Saichev
E-mail: saichev@hotmail.com

1 National Research University “Higher School of Economics”, Nizhny Novgorod, Russia.

2 Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland.

Citation information

Cite this article as: Comparative statistics of Garman-Klass, Parkinson, Roger-Satchell and bridge estimators, S. Lapinova & A. Saichev, Cogent Physics (2017), 4: 1303931.

References

Borodin, A. N., & Salminen, P. (2002). Handbook of Brownian motion – Facts and formulae (2nd ed.). Basel: Birkh äuser Verlag. http://dx.doi.org/10.1007/978-3-0348-s8163-0

Cont, R., & Tankov, P. (2004). Financial modelling with jump processes. London: CRC Press.

Garman, M., & Klass, M. J. (1980). On the Estimation of security price volatilities from historical data. The Journal of Business, 53, 67–78. http://dx.doi.org/10.1086/jb.1980.53.issue-1

Jeanblanc, M., Yor, M., & Chesney, M. (2009). Mathematical methods for financial markets. London: Springer Verlag. http://dx.doi.org/10.1007/978-1-84628-737-4

Park, M. (1980). The extreme value method for estimating the variance of the rate of return. The Journal of Business, 53, 61-65.

Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. The Annals of Applied Probability, 1, 504–512.

http://dx.doi.org/10.1214/aoap/1054649063

Saichev, A., Malevergne, Y., & Sornette, D. (2010). Theory of Zipf’s law and beyond. Heidelberg: Springer Verlag.

http://dx.doi.org/10.1007/978-3-642-02946-2

Saichev, A., Sornette, D. (2011 August 12). Time-bridge estimators of integrated variance. arXiv:1108.2611v1 [q-fin.ST].
Appendix A

A Probabilistic properties of high, low and close values

Here are given pdfs of random variables \(h,l,c\) (13) and variables \((\zeta,\xi)\) (14), which one needs for canonical estimators (12) statistical analysis. Let us begin with random variable \(c = x(1,\gamma)\). Obviously, its pdf is

\[
f(x;\gamma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\gamma)^2}{2}\right), \quad x \in (-\infty, \infty).
\]

It is easy to show, additionally, that joint pdf \(q_h(n,x;\gamma)\) of high values of canonical Brownian motion \(x(n,\gamma)\) and the close value \(c = x(1,\gamma)\) is equal to

\[
q_h(n, x; \gamma) = \sqrt{\frac{2}{\pi}}(2\eta - x) e^{2\eta \gamma} \exp\left(-\frac{1}{2}(2\eta - x + \gamma)^2\right),
\]

where \(\gamma > \chi > \eta\) (13). Using formulas, given at the monograph (Borodin & Salminen, 2002) and in the article (Saichev & Sornette, 2011), one might show that joint pdf given by:

\[
q_h(n, l, x; \gamma) = f(x; \gamma)S(n, l | x),
\]

where \(\chi \in (l, \eta), \quad h > \chi^1(x), \quad l > \chi^1(-x)\).

Here \(1(\chi)\) is the unit step function, equal to unity for \(\chi > 0\) and zero otherwise. Besides, above there is function

\[
S(n, l | x) = \sum_{m=-\infty}^{\infty} m|mF(m(n - l), x) + (1 - m)F(m(n - l) + l, x)|,
\]

\[
F(n, x) = \left|(x - 2\eta)^2 - 1\right| e^{2\eta(x - \eta)}.
\]

We need, at exploring statistical properties of canonical G&K estimator, in joint pdf \(q_{\delta}(\xi,\gamma)\) of canonical Brownian motion \(x(\xi,\gamma)\) (11) oscillation \(d = h - l\) and the close value \(c = x(1,\gamma)\). As it follows from (A3), (A4), mentioned pdf is equal to

\[
q_{\delta}(\xi, x, \gamma) = 4f(x; \gamma)\sum_{m=-\infty}^{\infty} m\sqrt{\frac{1}{2}|m| + 2m^2} \exp\left((-m + 1)|x| + 2m\delta^2 - 1 - (m + 1)(|x| + 2m\delta)\right) e^{-2m|\delta| + 1},
\]

where \(\delta > |x|, \quad x \in (-\delta, \delta)\).

After integration above joint pdf over all \(x\) values obtain pdf \(q_{\delta}(\gamma)\) of oscillation \(d:\)
\[ q_{\delta}(\delta, \gamma) = \sum_{m=-\infty}^{\infty} m \left[ \frac{8}{\pi} \exp \left( - \frac{(1 + 2m)^2 \delta^2 + 2\delta \gamma + \gamma^2}{2} \right) \times \left( 2 \exp \left( \frac{\delta(\delta + 4m\delta + 2\gamma)}{2} \right)(2m^2 \delta^2 - 1 - m(2 + \gamma^2)) + (1 + e^{2m\gamma})(1 + m(2 + \gamma^2)) \right) - 2\gamma(a(\delta, \gamma, m) - a(\delta, -\gamma, m)) \right], \quad \delta < 0. \] (A6)

Here we have used auxiliary function

\[ a(\delta, \gamma, m) = e^{2m\gamma}[1 + m(3 + \gamma(\delta + 2m\delta + \gamma))], \quad \delta < 0. \]

In particular case of zero drift (\( \gamma = 0 \)), one get from (A6) following expression

\[ q_{\delta}(\delta) = \frac{8}{\pi} \sum_{m=-\infty}^{\infty} (2m^2 \delta^2 - 1 - m(2 + \gamma^2)) \exp \left( - \frac{(1 + 2m)^2 \delta^2}{2} \right) - 4m^2 e^{-2m^2 \delta^2}. \] (A7)

All statistical properties of high and low values (14) of canonical bridge (15) are defined by their twofold joint pdf \( q_{\delta, \gamma}(\delta, \gamma) \), given by relation

\[ q_{\delta, \gamma}(\delta, \gamma) = \sum_{m=-\infty}^{\infty} m [mF(m\eta - 1) + (1 - m)F(m\eta + 1)] \quad \text{with} \quad F(\eta) = 4(4\eta^2 - 1)e^{-\eta^2}. \] (A8)

Following from here pdf \( q_{\delta}(\delta) \) of canonical bridge oscillation \( s = \xi - \zeta \) given by equality

\[ q_{\delta}(\delta) = 8\delta \sum_{m=-\infty}^{\infty} m^2 (4m^2 \delta^2 - 3)e^{-2m^2 \delta^2}, \quad \delta < 0. \] (A9)