Supplementary Materials for "Efficient Likelihood-Based Inference for the Generalized Pareto Distribution"

Hideki Nagatsuka*[1] and N. Balakrishnan[2]

[1]Department of Industrial and Systems Engineering, Chuo University, 1-13-27 Kasuga,Bunkyo-ku,Tokyo 112-8551, Japan
[2]Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1

Contents

1 Computation for the Proposed Method

1.1 Computation of the proposed likelihood function and its derivatives

Computation of the proposed likelihood function and its derivatives require a numerical integration. For their computation in our programs, we use the Matlab function *integral()*, which works satisfactorily for computing integrals in the likelihood functions. Besides the Matlab function, we recommend to use the double exponential formulas provided by Takahasi and Mori (1974). For the implementation of the double exponential formulas in C program, the readers may find the C functions *intde ()* and *intdei ()* in Ooura's Mathematical Software Packages, provided by Prof. Takuya Ooura on his web page (http://www.kurims.kyoto-u.ac.jp/ ooura/index.html). We strongly recommend not to use the R function *integrate()* since the obtained results using this function may not be satisfactory, as we have witnessed.

Although, in general, the values of the likelihood function increase exponentially with sample size n, it is possible to prevent the values from getting too large by taking the logarithm of the likelihood function, and thus may overflow rarely happens in computation. However, it is not effective to take the logarithm of the proposed likelihood function $l(k;\boldsymbol{s}_n^{(j)})$ for preventing $l(k;\boldsymbol{s}_n^{(j)})$ from overflow since $l(k;\boldsymbol{s}_n^{(j)})$ has an integral with the integrand including factors that increase exponentially with n that needs to be evaluated before taking the logarithm. To remedy this difficulty, we recommend the use of log sum exponential method to compute the proposed log-likelihood function, it is as follows.

For $k \neq 0$, the logarithm of $l(k;\boldsymbol{s}_n^{(j)})$ can be expressed as

$$\log l(k;\boldsymbol{s}_n^{(j)}) \quad = \quad \log n! + D + \log \int_{\chi_k} \frac{1}{|k|}\left(\frac{u}{k}\right)^{n-1} \prod_{i=1}^{n} (1-us_i)^{1/k-1} \exp\left(-D\right)\, du, \tag{1}$$

*Corresponding Author: Hideki Nagatsuka, email:hideki@indsys.chuo-u.ac.jp where

$$
\begin{aligned}
D &= \sup_{u\in\chi_k} \log \frac{1}{|k|}\left(\frac{u}{k}\right)^{n-1}\prod_{i=1}^{n}\left(1-us_i\right)^{1/k-1} \\
&= \sup_{u\in\chi_k} (n-1)\log|u| + \left(\frac{1}{k}-1\right)\sum_{i=1}^{n}\log\left(1-us_i\right).
\end{aligned}
$$

Note that the integrand of the third term in Eq.(1) is less than or equal to 1. Thus, an overflow in computation is not expected to occur. This method can be used in the computation of the derivatives of the likelihood function as well. For $k=0$, $l(k;\boldsymbol{s}_n^{(j)})$ and its derivatives do not have integrals and the log sum exponential method is not needed in this case.

1.2 Computation of Fisher Information:

The Fisher information,

$$
I_{j,n}(k) = -E\left(\frac{\partial^2}{\partial k^2}\log l\left(k;\,\boldsymbol{S}_n^{(j)}\right)\right),
$$

requires n-fold integrals to compute the expectation of $\log l\left(k;\,\boldsymbol{S}_n^{(j)}\right)$. One possibility is to use Monte Carlo integration; but, instead, we use the observed Fisher information, instead of $I_{j,n}(k)$, which is given by

$$
\hat{I}_{j,n}(k) = -\frac{\partial^2}{\partial k^2}\log l\left(k;\,\boldsymbol{s}_n^{(j)}\right),
$$

where $\boldsymbol{s}_n^{(j)} = (s_{1:n}^{(j)},\ldots,s_{j-1:n}^{(j)},s_{j+1:n}^{(j)},\ldots,s_{n:n}^{(j)})$, $s_{i:n}^{(j)} = x_{i:n}/x_{j:n}$, $i\neq j$, $1\le i\le n$, and $x_{i:n}$ is the observed value corresponding to $X_{i:n}$. We carried out extensive Monte Carlo simulations and observed that the asymptotic properties of the interval estimation and hypothesis testing procedures based on the observed Fisher information were quite similar to those based on the expected Fisher information. In Sections 4 and 5, we therefore use the observed Fisher information for the inferential procedures carried out there.

2 Proofs

2.1 Proof of Proposition 2

For $k\neq 0$, $l'(k;\boldsymbol{s}_n^{(j)})$ is immediately obtained by deriving $l(k;\boldsymbol{s}_n^{(j)})$. In the case when $k=0$, for $\Delta k<0$, using the fact that $l(0;\boldsymbol{s}_n^{(j)}) = n!\int_0^\infty u^{n-1}\exp\left(-u\sum_{i=1}^n s_i\right)du$,

$$
\begin{aligned}
&\frac{l(\Delta k;\boldsymbol{s}_n^{(j)}) - l(0;\boldsymbol{s}_n^{(j)})}{\Delta k} \\
&= n!\left[\int_0^\infty u^{n-1}\exp\left\{\left(\frac{1}{\Delta k}-1\right)\sum_{i=1}^{n}\log\left(1-\Delta k u s_i\right)\right\} - u^{n-1}\exp\left(-u\sum_{i=1}^{n}s_i\right)du\right] \\
&\quad /\Delta k \\
&= n!\left[\int_0^\infty u^{n-1}\exp\left[\sum_{i=1}^{n}\left\{-us_i+\left(1-\frac{us_i}{2}\right)\Delta k u s_i\right\}+O\left(\Delta k\right)^2\right] - u^{n-1}\exp\left(-u\sum_{i=1}^{n}s_i\right)du\right] \\
&\quad /\Delta k \\
&\to n!\int_0^\infty u^{n-1}\left\{\sum_{i=1}^{n}\left(1-\frac{us_i}{2}\right)us_i\right\}\exp\left(-u\sum_{i=1}^{n}s_i\right)du, \quad \text{as } \Delta k\uparrow 0, \\
&= \left\{1-\frac{(n+1)\sum_{i=1}^n s_i^2}{2\left(\sum_{i=1}^n s_i\right)^2}\right\}\frac{(n!)^2}{\left(\sum_{i=1}^n s_i\right)^n}. \qquad (2)
\end{aligned}
$$

For $\Delta k > 0$, we see that

$$
\begin{aligned}
& \frac{l(\Delta k; \boldsymbol{s}_n^{(j)}) - l(0; \boldsymbol{s}_n^{(j)})}{\Delta k} \\
&= n! \left[\int_0^{1/\Delta k} u^{n-1} \exp\left\{ \left(\frac{1}{\Delta k} - 1 \right) \sum_{i=1}^n \log\left(1 - \Delta k u s_i\right) \right\} du - \int_0^\infty u^{n-1} \exp\left(-u \sum_{i=1}^n s_i \right) du \right] \\
&\quad / \Delta k \\
&\to n! \int_0^\infty u^{n-1} \exp\left\{ \left(\frac{1}{\Delta k} - 1 \right) \sum_{i=1}^n \log\left(1 - \Delta k u s_i\right) \right\} du - u^{n-1} \exp\left(-u \sum_{i=1}^n s_i \right) du \\
&\quad / \Delta k \\
&\to n! \int_0^\infty u^{n-1} \left\{ \sum_{i=1}^n \left(1 - \frac{u s_i}{2} \right) u s_i \right\} \exp\left(-u \sum_{i=1}^n s_i \right) du, \quad \text{as } \Delta k \downarrow 0, \\
&= \left\{ 1 - \frac{(n+1) \sum_{i=1}^n s_i^2}{2 \left(\sum_{i=1}^n s_i \right)^2} \right\} \frac{(n!)^2}{\left(\sum_{i=1}^n s_i \right)^n}. \qquad (3)
\end{aligned}
$$

It follows from (2) and (3) that

$$
l'(0; \boldsymbol{s}_n^{(j)}) = \lim_{\Delta k \to 0} \frac{l(\Delta k; \boldsymbol{s}_n^{(j)}) - l(0; \boldsymbol{s}_n^{(j)})}{\Delta k} = \left\{ 1 - \frac{(n+1) \sum_{i=1}^n s_i^2}{2 \left(\sum_{i=1}^n s_i \right)^2} \right\} \frac{(n!)^2}{\left(\sum_{i=1}^n s_i \right)^n}.
$$

2.2 Proof of Proposition 3

For $k \neq 0$, $l''(k; \boldsymbol{s}_n^{(j)})$ is immediately obtained by deriving $l'(k; \boldsymbol{s}_n^{(j)})$. In the following, we will derive $l''(0; \boldsymbol{s}_n^{(j)})$. We obtain for $\Delta k \neq 0$ such that $|\Delta k|$ is sufficiently small,

$$
\begin{aligned}
\Psi(\Delta k, u) &= (1/\Delta k - 1) \sum_{i=1}^n \log\left(1 - \Delta k u s_i\right) \\
&= -u \sum_{i=1}^n s_i + \Delta k u \sum_{i=1}^n \left(1 - \frac{u s_i}{2} \right) s_i + \Delta k^2 u^2 \sum_{i=1}^n \left(\frac{1}{2} - \frac{u s_i}{3} \right) s_i^2 \\
&\quad + \Delta k^3 u^3 \sum_{i=1}^n \left(\frac{1}{3} - \frac{u s_i}{4} \right) s_i^3 + O\left(\Delta k^4\right), \qquad (4)
\end{aligned}
$$

and

$$
\begin{aligned}
\Psi'(\Delta k, u) &= -\frac{n}{\Delta k} - \frac{\sum_{i=1}^n \log\left(1 - \Delta k u s_i\right)}{\Delta k^2} \\
&= -\left(n - u \sum_{i=1}^n s_i \right) / \Delta k + \frac{u^2}{2} \sum_{i=1}^n s_i^2 + \Delta k \frac{u^3}{3} \sum_{i=1}^n s_i^3 + O\left(\Delta k^2\right), \qquad (5)
\end{aligned}
$$

where $\Psi'(\Delta k, u) = \partial \Psi(\Delta k, u) / \partial \Delta$.

From (4) and (5), and a Maclaurin expansion for $\exp\left(\Psi\left(\Delta k, u\right)\right)$ around 0, we have

$$
\begin{aligned}
\lim_{\Delta k \to 0} l'(\Delta k; \boldsymbol{s}_n^{(j)}) &= \lim_{\Delta k \to 0} n! \int_{\mathcal{X}_{\Delta k,1}} u^{n-1} \Psi'\left(\Delta k, u\right) \exp\left(\Psi\left(\Delta k, u\right)\right) du \\
&= \lim_{\Delta k \to 0} \int_0^\infty \Bigg[-\left(n - u\sum_{i=1}^n s_i\right)/\Delta k \\
&\quad -un\sum_{i=1}^n s_i + u^2\left\{\frac{n+1}{2}\sum_{i=1}^n s_i + \left(\sum_{i=1}^n s_i\right)^2\right\} - u^3\frac{1}{2}\sum_{i=1}^n s_i \sum_{i=1}^n s_i^2 \\
&\quad + \Bigg[-\frac{n}{2}u^2\left\{\left(\sum_{i=1}^n s_i\right)^2 + \sum_{i=1}^n s_i^2\right\} \\
&\quad + u^3\left\{\frac{n+1}{3}\sum_{i=1}^n s_i^3 + \frac{n+2}{2}\sum_{i=1}^n s_i \sum_{i=1}^n s_i^2 + \frac{1}{2}\left(\sum_{i=1}^n s_i\right)^3\right\} \\
&\quad - u^4\left\{\frac{n+2}{8}\left(\sum_{i=1}^n s_i^2\right)^2 + \frac{1}{2}\left(\sum_{i=1}^n s_i\right)^2 \sum_{i=1}^n s_i^2 + \frac{1}{3}\sum_{i=1}^n s_i \sum_{i=1}^n s_i^3\right\} \\
&\quad + \frac{1}{3}u^5 \sum_{i=1}^n s_i \left(\sum_{i=1}^n s_i^2\right)^2 \Bigg] \Delta k + O\left(\Delta k^2\right)\Bigg] n!\, u^{n-1} \exp\left(-u\sum_{i=1}^n s_i\right) du \\
&= l'(0; \boldsymbol{s}_n^{(j)}) \\
&\quad + \left\{1 - (n+1)\frac{\sum_{i=1}^n s_i^2}{\left(\sum_{i=1}^n s_i\right)^2} + \frac{(n+2)(n+3)}{4}\frac{\left(\sum_{i=1}^n s_i^2\right)^2}{\left(\sum_{i=1}^n s_i\right)^4} - \frac{2(n+2)}{3}\frac{\sum_{i=1}^n s_i^3}{\left(\sum_{i=1}^n s_i\right)^3}\right\} \\
&\quad \times \frac{n!\,(n+1)!}{\left(\sum_{i=1}^n s_i\right)^n} \lim_{\Delta k \to 0} \Delta k + \lim_{\Delta k \to 0} O\left(\Delta k^2\right). \qquad (6)
\end{aligned}
$$

It then follows from (6) that

$$
\begin{aligned}
l''\left(0; \boldsymbol{s}_n^{(j)}\right) &= \lim_{\Delta k \to 0} \frac{l'(\Delta k; \boldsymbol{s}_n^{(j)}) - l'(0; \boldsymbol{s}_n^{(j)})}{\Delta k} \\
&= \left\{1 - (n+1)\frac{\sum_{i=1}^n s_i^2}{\left(\sum_{i=1}^n s_i\right)^2} + \frac{(n+2)(n+3)}{4}\frac{\left(\sum_{i=1}^n s_i^2\right)^2}{\left(\sum_{i=1}^n s_i\right)^4} - \frac{2(n+2)}{3}\frac{\sum_{i=1}^n s_i^3}{\left(\sum_{i=1}^n s_i\right)^3}\right\} \\
&\quad \times \frac{n!\,(n+1)!}{\left(\sum_{i=1}^n s_i\right)^n}.
\end{aligned}
$$

2.3 Proof of Proposition 4

The proof is very similar to that of Proposition 3 and is therefore omitted for the sake of brevity.

2.4 Proof of Theorem 10

It suffices to prove the proposition only in the cases of $\hat{k} < 0$ and $0 < \hat{k} < 1$. We let

$$\Psi\left(\sigma\right) = n - \left(\frac{1}{\hat{k}} - 1\right)\sum_{i=1}^n \frac{\hat{k}\, x_i}{\sigma - \hat{k}\, x_i}.$$

Then, for $\hat{k} < 0$ and $0 < \hat{k} < 1$, we obtain

$$\frac{\partial \Psi\left(\sigma\right)}{\partial \sigma} = \left(\frac{1}{\hat{k}} - 1\right)\sum_{i=1}^n \frac{\hat{k}\, x_i}{\left(\sigma - \hat{k}\, x_i\right)^2} > 0. \qquad (7)$$

For $\hat{k} < 0$, we have $\lim_{\sigma \downarrow 0} \Psi\left(\sigma\right) = n/\hat{k} < 0$, $\lim_{\sigma \to \infty} \Psi\left(\sigma\right) = n > 0$ and $\Psi\left(\sigma\right)$ is continuous with respect to $\sigma > 0$. It follows from these facts and (7) that $\Psi\left(\sigma\right)$ changes sign only once with respect to σ. For $0 < \hat{k} < 1$, we must have $\sigma > \hat{k}\ \max_i x_i$ since the support of the pdf of the GPD is $(0, \sigma/k)$ when $k > 0$. Then, we observe that $\lim_{\sigma \downarrow \hat{k}\ \max_i x_i} \Psi(\sigma) = -\infty < 0$, $\lim_{\sigma \to \infty} \Psi(\sigma) = n > 0$, and $\Psi(\sigma)$ is continuous with respect to $\sigma \in (\hat{k}\ \max_i x_i, \infty)$, which also implies that $\Psi(\sigma)$ changes sign only once.

From the facts established above, $\Psi(\sigma) = 0$ always has a unique solution with respect to σ, which completes the proof of Theorem 10.

2.5 Proof of Theorem 11

When $k = 0$, $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathcal{P}} \sigma$, since $X_1, \ldots, X_n$ are i.i.d. random variables from the exponential distribution with mean σ, where $\xrightarrow{\mathcal{P}}$ denotes convergence in probability. When $k \geq 1$, for any $\epsilon > 0$, we have

$$P(|k\, X_{n:n} - \sigma| < \epsilon) \;=\; F^n\left(\frac{\sigma}{k} + \frac{\epsilon}{k}; k, \sigma\right) - F^n\left(\frac{\sigma}{k} - \frac{\epsilon}{k}; k, \sigma\right) \to 1, \qquad \text{as } n \to \infty,$$

which implies that $k\, X_{n:n} \xrightarrow{\mathcal{P}} \sigma$. When $k < 0$ or $0 < k < 1$, by Proposition 10 and Corollary 3.8 of Lehmann and Casella (1998, p.448), the proposed estimator $\hat{\sigma}$, in which $\hat{k}$ is replaced by the true value k, is consistent for σ.

From the above facts, Theorem 4 and Slutsky's theorem, we have $\hat{\sigma} \xrightarrow{\mathcal{P}} \sigma$, which completes the proof.

3 Illustrarive Example Based on Fatigue Data

In the second example, we fit the GPD to the fatigue data concerning the Kevlar/Epoxy strand lifetime (in hours) at 70% stress level, provided by Andrews and Herzberg (1985), and reanalyzed by Castillo and Hadi (1997) and Chen et al (2017). These data are presented in Table 1. The interest in this example is in the left tail of the corresponding lifetime distribution (see Castillo and Hadi (1997) for details). But, the GPD is applicable to upper extremes. So, as suggested by Castillo and Hadi (1997), we fit the GPD to negative values of these data so that the lower tail can be transformed to upper tail, and then the obtained results can be readily inverted back for the original data.

Table 1: The fatigue data: the Kevlar/Epoxy strand lifetime (in hours) at 70% stress level

1051	1337	1389	1921	1942	2322	3629	4006	4012	4063
4921	5445	5620	5817	5905	5956	6068	6121	6473	7501
7886	8108	8546	8666	8831	9106	9711	9806	10205	10396
10861	11026	11214	11362	11604	11608	11745	11762	11895	12044
13520	13670	14110	14496	15395	16179	17092	17568	17568	

In Tables 2 and 3, the estimates of k and σ by the ML, ZS, Zj, WMD and the proposed method (Proposed) are all presented for each threshold. The thresholds are taken as $u = -1.8, -1.6, -1.4, -1.2, -1.0$ and -0.8 $(\times 10^3)$, same as those chosen by Castillo and Hadi (1997). As several authors have reported, the ML estimates are not found for all the cases.

As in the preceding example, we computed ASAE for all the methods, the corresponding results for the ML, ZS, Zj, WMD and Proposed methods are all presented in Table 4.

Tables 2 and 3 also present the 95% the confidence intervals for k and σ by bootstrap-t technique, based on ZS, Zj, WMD and the proposed methods based on Wald type statistic (Proposed-Wald) and LR type statistic (Proposed-LR).

Table 2: Estimates of k for the fatigue data: u is the threshold and m is the number of exceedances

u $(\times 10^3)$	m	ZS	Zj	WMD	Proposed
-1.8	49	1.017	0.885	1.137	1.057
-1.6	45	0.968	0.841	1.079	1.014
-1.4	42	0.859	0.748	0.908	0.914
-1.2	39	0.672	0.599	0.441	0.760
-1.0	28	0.825	0.676	0.849	0.914
-0.8	21	0.853	0.640	0.894	0.955

Table 3: Estimates of σ for the fatigue data: u is the threshold and m is the number of exceedances

u $(\times 10^3)$	m	ZS	Zj	WMD	Proposed
-1.8	49	17627	16074	19354	18206
-1.6	45	14884	13605	16256	15472
-1.4	42	11627	10697	12170	12210
-1.2	39	8137	7665	7048	8814
-1.0	28	7922	7095	8205	8560
-0.8	21	6411	5501	6734	6995

Table 4: ASAE for the fatigue data: u is the threshold and m is the number of exceedances

u $(\times 10^3)$	m	ZS	Zj	WMD	Proposed
-1.8	49	0.0386	0.0511	0.0348	0.0361
-1.6	45	0.0366	0.0477	0.0350	0.0344
-1.4	42	0.0234	0.0302	0.0254	0.0257
-1.2	39	0.0357	0.0355	0.0378	0.0405
-1.0	28	0.0314	0.0358	0.0314	0.0337
-0.8	21	0.0475	0.0507	0.0479	0.0495

Table 5: CIs of k for the fatigue data: u is the threshold and m is the number of exceedances

u $(\times 10^3)$	m	ZS	Zj	WMD	Proposed-Wald	Proposed-LR
-1.8	49	(0.776, 1.366)	(0.806, 1.240)	(0.725, 1.695)	(0.719, 1.394)	(0.755, 1.458)
-1.6	45	(0.705, 1.334)	(0.739, 1.200)	(0.668, 1.671)	(0.667, 1.360)	(0.700, 1.427)
-1.4	42	(0.579, 1.219)	(0.625, 1.104)	(0.524, 1.535)	(0.567, 1.261)	(0.583, 1.322)
-1.2	39	(0.394, 1.043)	(0.434, 0.949)	(0.103, 1.074)	(0.397, 1.123)	(0.359, 1.164)
-1.0	28	(0.495, 1.306)	(0.543, 1.132)	(0.331, 1.634)	(0.461, 1.367)	(0.466, 1.468)
-0.8	21	(0.435, 1.448)	(0.487, 1.164)	(0.208, 1.974)	(0.400, 1.510)	(0.408, 1.666)

Table 6: CIs of σ for the fatigue waves data: u is the threshold and m is the number of exceedances

u $(\times 10^3)$	m	ZS	Zj	WMD	Proposed-Wald	Proposed-LR
-1.8	49	(13434, 23097)	(13936, 21430)	(12648, 26536)	(12187, 23626)	(12803, 24716)
-1.6	45	(10971, 19787)	(11568, 18293)	(10254, 22402)	(9968, 20335)	(10466, 21338)
-1.4	42	(8142, 15689)	(8753, 14571)	(7480, 17096)	(7337, 16322)	(7546, 17123)
-1.2	39	(5417, 11232)	(5882, 10584)	(3957, 10213)	(4349, 12293)	(3935, 12749)
-1.0	28	(5133, 11412)	(5527, 10258)	(3905, 12250)	(4124, 12233)	(4171, 13136)
-0.8	21	(3611, 9648)	(4062, 8289)	(2414, 10542)	(2780, 10492)	(2832, 11577)

Table 7: P-values under $H_0 : k \leq 0.5$ for the proposed methods for the fatigue data: u is the threshold and m is the number of exceedances

$u \times 10^3$	m	Proposed-Wald	Proposed-Score	Proposed-LR
-1.8	49	0.001	0.000	0.000
-1.6	45	0.004	0.000	0.002
-1.4	42	0.019	0.001	0.016
-1.2	39	0.160	0.110	0.178
-1.0	28	0.073	0.014	0.067
-0.8	21	0.108	0.022	0.096

Furthermore, we also carry out the three proposed tests for k and σ. For this, we consider the hypothesis $H_0 : k \leq 0.5$ vs. $H_1 : k > 0.5$. Under the null hypothesis $H_0 : k \leq 0.5$, we carry out the three proposed tests, and the p-value at $k = 0.5$ is the largest among $H_0 : k \leq 0.5$, for all the cases. These p-values are presented in Table 7 for each threshold level. For $u = -18,000, -16,000$ and $-14,000$, all p-values are less than 0.05, providing evidence to k being larger than 0.5. Moreover, almost all estimates of k (see Table 2) are larger than 0.5, and the ML estimates are not found. This certainly provides a justification for the methods developed in this work.

Next, suppose we are interested in testing $H_0 : \sigma \geq 30,000$ vs. $H_1 : \sigma < 30,000$. The p-value at $\sigma = 30,000$ is the largest among $H_0 : \sigma \geq 30,000$, for all the cases. These p-values, computed for each threshold, all turned out to be 10^{-4}, but the results are not presented here for brevity.

References

Andrews DF, Herzberg AM (1985) Data: A Collection of Problems from Many Fields for the Student and Research Worker. Springer, New York

Castillo E, Hadi AS (1997) Fitting the generalized Pareto distribution to data. Journal of the American Statistical Association 92:1609–1620

Chen P, Ye Z, Zhao X (2017) Minimum distance estimation for the generalized Pareto distribution. Technometrics 59:528–541

Lehmann EL, Casella G (1998) Theory of Point Estimation, 2nd ed. Springer, New York

Takahasi H, Mori M (1974) Double exponential formulas for numerical integration. Publications of the Research Institute for Mathematical Sciences, Kyoto University 9:721–741