Ann. Inst. Statist. Math.
Vol. 50, No. 2, 361–377 (1998)

# PREDICTION OF THE MAXIMUM SIZE IN WICKSELL'S CORPUSCLE PROBLEM

RINYA TAKAHASHI[1] AND MASAAKI SIBUYA[2*]

[1] *Kobe University of Mercantile Marine, 5-1-1, Fukae-Minami, Higashi-Nada-ku, Kobe 658-0022, Japan*
[2] *Department of Mathematics, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan*

(Received November 29, 1995; revised May 6, 1997)

**Abstract.** In the Wicksell corpuscle problem, the maximum size of random spheres in a volume part is to be predicted from the sectional circular distribution of spheres cut by a plane. The size of the spheres is assumed to follow the generalized gamma distribution. Some prediction methods according to measurement methods on the sectional plane are proposed, and their performances are evaluated by simulation. The prediction method based on the $r$ largest sizes and the total number of the sectional circles is recommended, because of its satisfactory performance.

*Key words and phrases*: Extreme value theory, generalized gamma distribution, Gumbel distribution, metal fatigue, stereology.

## 1. Introduction

Spherical particles of random size are randomly scattered in a space, and sectional circles of the spheres cut by a plane are observed. To estimate the size distribution and the spatial density of the random spheres from those of circles on the sectional plane is Wicksell's corpuscle problem, Wicksell (1925). Our problem in this paper is to predict the maximum size of the spheres in a given volume and that of those intersecting with a given plane area.

For controlling the fatigue strength of steel, Murakami (1993, 1994) developed a prediction method using the Gumbel QQ-plot of the sectional maximum data. A feature of his method is to use only the maximum circle in each of some parts of the sectional plane. Takahashi and Sibuya (1996), assuming the size distribution of random spheres to be generalized gamma, proposed an extended Murakami's method for prediction. Simulation results show that the performance of this method is unsatisfactory.

* Now at Takachiho University, 2-19-1 Ohmiya, Suginami-ku, Tokyo 168-8508, Japan.

In this paper, we propose alternatives by adding more works to the measurement and using corresponding prediction methods, and evaluate their performance by simulation. Some prediction methods are preferable to the extended Murakami's one.

In Section 2, Wicksell transform in terms of the areas is summarized, and the prediction method in Takahashi and Sibuya (1996) is reviewed. In Section 3, we propose some measurement methods on the sectional plane and corresponding prediction methods, and in Section 4 evaluate their performance by simulation. The final Section 5 is supplementary and technical. In the Subsection 5.1, we prove a theorem on the maximum of Wicksell transforms of areas. In the Subsection 5.2, a theorem on the extreme value in a random sample of random size is stated, and in the Subsection 5.3, the Fisher informations of the estimates in Section 3 are evaluated.

## 2. A parametric model of Wicksell's corpuscle problem

### 2.1 *Wicksell's corpuscle problem*

Spherical particles of random size are randomly scattered in space. The centers of the spheres constitute a Poisson process of intensity $\lambda_V$, and the size $S_V$ of the sphere is independent of its position. The spheres are cut by a sectional plane, and the centers of the sectional circles constitute a Poisson process of intensity $\lambda_A$, and the size $S_A$ of the circle is independent of its position. We need to consider the spheres (with size $S_C$) crossing the sectional plane. Assume $E(S_V)$ is finite and small enough compared with $\lambda_V$, and the spheres are actually disjoint.

Usually, the size $S_\omega$, $\omega = A$, $C$ and $V$, is the diameter. In this paper, $S_A$ is the area of sectional circle and $S_C$ and $S_V$ are the areas of the great circles of spheres. Their p.d.f.'s (probability density functions), d.f.'s (distribution functions) and survival functions are denoted by $f_\omega(s)$, $F_\omega(s)$ and $\bar{F}_\omega(s) = 1 - F_\omega(s)$ ($\omega = V$, $C$ and $A$), respectively.

It is known that

$$(2.1)\qquad \lambda_V = \sqrt{\pi}\lambda_A/(2\mu_0) \quad \text{where} \quad \mu_0 = E(\sqrt{S_V}),$$

$$(2.2)\qquad f_C(s) = \sqrt{s} f_V(s)/\mu_0, \quad 0 < s < \infty, \quad S_A = S_C(1 - U^2),$$

$$(2.3)\qquad f_A(s) = \frac{1}{2\mu_0}\int_s^\infty \frac{1}{\sqrt{v-s}} f_V(v)dv, \quad \bar{F}_A(s) = \frac{1}{2\mu_0}\int_s^\infty \frac{1}{\sqrt{v-s}}\bar{F}_V(v)dv,$$

and

$$(2.4)\qquad \bar{F}_V(s) = \frac{1}{E(1/\sqrt{S_A})}\int_s^\infty \frac{1}{\sqrt{t-s}} f_A(t)dt,$$

where $U$ is the uniform random variable on $(0, 1)$ and independent of $S_C$. The probability distribution of $S_A$ is called Wicksell transform of that of $S_V$.

### 2.2 *Extreme value theory*

Let $(X_j)_{j=1}^{\infty}$ be a sequence of i.i.d. (independent and identically distributed) random variables with a d.f. $H$ and let $W_n = \max(X_1, \dots, X_n)$. If there exist coefficients $a_n$ ($> 0$) and $b_n$ such that $(W_n - b_n)/a_n$ converges in distribution, i.e. if $\lim_{n\to\infty} H^n(a_n x + b_n) = L(x)$ for some nondegenerate d.f. $L$, then $H$ belongs to the domain of attraction of $L$ (or $H \in \mathcal{D}(L)$) and $L$ is limited to the following d.f.'s.

$$L_{ic}(x) = \begin{cases} \exp(-x^{-c}), & x \geq 0, \quad c > 0,\ i = 1, \\ \exp(-(-x)^c), & x \leq 0, \quad c > 0,\ i = 2, \\ \exp(-\exp(-x)), & -\infty < x < \infty,\ c = 0,\ i = 3. \end{cases}$$

For the maximum of Wicksell transforms of the diameters, a general result was obtained by Drees and Reiss (1992). In terms of the d.f. $\Phi_\omega$ of $T_\omega = S_\omega^\beta$, $\beta > 0$, $\omega = V$ and $A$, which are used in later, their result is as follows.

PROPOSITION 2.1.

$$\begin{aligned} &\Phi_A \in \mathcal{D}(L_{1,c-(2\beta)^{-1}}) \quad \textit{if} \quad \Phi_V \in \mathcal{D}(L_{1c}),\ c > (2\beta)^{-1}, \\ &\Phi_A \in \mathcal{D}(L_{2,c+1/2}) \quad \textit{if} \quad \Phi_V \in \mathcal{D}(L_{2c}),\ c > 0, \\ &\Phi_A \in \mathcal{D}(L_{30}) \quad \textit{if} \quad \Phi_V \in \mathcal{D}(L_{30}). \end{aligned}$$

A simple proof of this proposition is given in Subsection 5.1. In the following, we shall discuss only the case of $L_{30}$, the Gumbel distribution, which is denoted by

$$\Lambda(x) = \exp(-\exp(-x)), \qquad -\infty < x < \infty.$$

In general, let $N$ be the Poisson variable with mean $\theta$ and independent of $(X_j)_{j=1}^{\infty}$. Under the condition $H \in \mathcal{D}(\Lambda)$ and $N > 0$, for the d.f. $H^*$ of $W_N = \max(X_1, \dots, X_N)$,

$$H^*(a_\theta y + b_\theta) \to \Lambda(y), \qquad \text{as} \quad \theta \to \infty,$$

where $b_\theta = b(\theta) = \bar{H}^{-1}(1/\theta)$, $a_\theta = a(\theta) = b(\theta e) - b(\theta)$ and $\bar{H}^{-1}(x) = \inf\{y \mid 1 - H(y) \leq x\}$. A more general result is shown in Subsection 5.2.

### 2.3 *Generalized gamma model and prediction problem*

As in Takahashi and Sibuya (1996), we study the problem within the framework of a parametric model. Let the generalized gamma distribution with the p.d.f.

$$\frac{1}{\Gamma(\alpha)} \cdot \frac{\gamma}{\xi^{\alpha\gamma}} x^{\alpha\gamma-1} e^{-(x/\xi)^\gamma} 1[0 < x < \infty], \qquad \alpha, \gamma, \xi > 0, \tag{2.5}$$

be denoted by $\mathsf{Ga}(\alpha, \gamma, \xi)$, and suppose the area $S_V$ of the great circle of the sphere to follow $\mathsf{Ga}(\alpha, \gamma, \xi^{1/\gamma})$ with known $\alpha, \gamma > 0$. The parameters $\xi$ and the intensity of the sphere $\lambda_V$ are unknown. This modelling was justified in Takahashi and Sibuya

(1996). We observe the circles in $k$ parts of identical area $A$ of the sectional plane. The number $N_A$ of the circles in a part of area $A$ is the Poisson variable with mean $\lambda_A A$. Let $W_A$ be the maximum area of the circles in a part of area $A$ if $N_A > 0$.

There are two problems (V) and (C). Since the latter is simple we discuss only the problem (V).

(V) Predict the maximum area $W_V$, of the great circles of the spheres in a part of volume $V$. The expected number of spheres in the part is $\lambda_V V$.

(C) Predict the maximum area $W_C$, of the great circles of the spheres which intersect with a part of area $A_C$. The expected number of spheres in the part is $\lambda_A A_C$.

2.4 *Base of prediction method*

Our prediction methods are based on the following facts on the scale and location parameters of $W_\omega^\gamma$, $\omega = V$ and $A$, and $S_V^\gamma$.

PROPOSITION 2.2. *Assume $T_V = S_V^\gamma$ to follow $\mathsf{Ga}(\alpha, 1, \xi)$. If $\lambda_A A$, $\lambda_V V \to \infty$, then the distribution of the power transformation of the maximum area $W_\omega^\gamma$ is approximated by the Gumbel distribution $\Lambda((t - \eta_\omega)/\xi)$, $\omega = V$ and $A$, where the scale parameter $\xi$ is common and equal to that of $T_V$, and the location is determined as follows:*

$$\eta_V/\xi = \tau_V + (\alpha - 1)\log\tau_V - \log\Gamma(\alpha), \qquad \tau_V = \log(\lambda_V V), \tag{2.6}$$

$$\eta_A/\xi = \tau_A + \left(\alpha + \frac{1}{2\gamma} - \frac{3}{2}\right)\log\tau_A - \log\left\{2\sqrt{\frac{\gamma}{\pi}}\,\Gamma\left(\alpha + \frac{1}{2\gamma}\right)\right\}, \tag{2.7}$$
$$\tau_A = \log(\lambda_A A).$$

*The relation between two intensities* (2.1) *is as follows:*

$$\lambda_V = \sqrt{\pi}\lambda_A/(2\mu_0), \qquad \mu_0 = \xi^{1/2\gamma}\Gamma(\alpha + 1/(2\gamma))/\Gamma(\alpha). \tag{2.8}$$

Remark that

$$\tau_V = \tau_A + \log\frac{V}{A} - \delta, \tag{2.9}$$
$$\delta = \log\left(\frac{2\mu_0}{\sqrt{\pi}}\right) = \frac{1}{2\gamma}\log\xi + \log\left\{\frac{2}{\sqrt{\pi}}\Gamma\left(\alpha + \frac{1}{2\gamma}\right)\Big/\Gamma(\alpha)\right\}.$$

For the better Gumbel approximation, we deal with $T_\omega = S_\omega^\gamma$, $\omega = V$ and $A$, instead of $S_\omega$.

Based on these facts, we predict $W_V^\gamma$ as follows. First, we estimate $\xi$ and $\tau_A$ by data. From the estimate $(\hat{\xi}, \widehat{\tau_A})$, we estimate further functions of $(\xi, \tau_A)$,

$$\hat{\delta} = \frac{1}{2\gamma}\log\hat{\xi} + \log\left\{\frac{2}{\sqrt{\pi}}\Gamma\left(\alpha + \frac{1}{2\gamma}\right)\Big/\Gamma(\alpha)\right\}, \tag{2.10}$$

$$\widehat{\tau_V} = \widehat{\tau_A} + \log\frac{V}{A} - \hat{\delta}, \quad \text{and} \quad \widehat{\left(\frac{\eta_V}{\xi}\right)} = \widehat{\tau_V} + (\alpha - 1)\log\widehat{\tau_V} - \log\Gamma(\alpha). \tag{2.11}$$

Finally, the mean and quantiles of $W_V^\gamma$ are estimated by linear expressions

$$\widehat{\eta_V} + c\hat{\xi} = \hat{\xi}(\widehat{\eta_V}/\hat{\xi} + c). \tag{2.12}$$

For the mean $c = \gamma_E$, Euler's constant, and for the quantile $c = \omega_p = -\log(-\log p)$.

The bias and variance of the above estimate $\widehat{\eta_V} + c\hat{\xi}$ are approximated by

$$\tau_V \operatorname{Bias}(\hat{\xi}) \quad \text{and} \quad (\tau_V)^2 \operatorname{Var}(\hat{\xi}), \tag{2.13}$$

respectively, for practically probable parameter values, $(\alpha, \gamma) \doteq (1, 1/2)$ and $\tau_A = 2 \sim 5$. See Takahashi and Sibuya (1996). Hence, an accurate estimate of $\xi$ is needed for good prediction.

## 3. Measurement methods and corresponding prediction methods

In this section, we investigate some measurement methods on the sectional plane and corresponding prediction methods. Recall that we can observe the circles in $k$ parts of common area $A$ of the sectional plane. The first two methods, the circles in each of $k$ parts of common area $A$ are measured. For the rest methods, the circles in a part of area $kA$ are measured, that is, the data in $k$ parts of common area $A$ are pooled. For convenience, we drop the subscript $A$ in $W_A$, $\eta_A$, $\tau_A$, etc.

### 3.1 *Maximum areas* (PM1)

Measure the maximum areas $W_j$ among circles in each of $k$ parts of common area $A$ on the sectional plane, $j = 1, \dots, k$. The prediction method, PM1, and simulation results are given in Takahashi and Sibuya (1996). The accuracy of PM1 is unsatisfactory.

### 3.2 *Maximum areas and the number of the circles* (PM2)

Measure the maximum area and the number of the circles $(W_j, N_j)_{j=1}^k$ in $k$ parts, each of area $A$, of the sectional plane. Because of Proposition 2.2, assume the data $(W_j^\gamma)_{j=1}^k$ to follow the Gumbel distribution $\Lambda(t/\xi - \zeta)$, where

$$\zeta = \frac{\eta}{\xi} = \tau + \left(\alpha + \frac{1}{2\gamma} - \frac{3}{2}\right)\log\tau - \log\left\{2\sqrt{\frac{\gamma}{\pi}}\,\Gamma\left(\alpha + \frac{1}{2\gamma}\right)\right\}, \tag{3.1}$$
$$\tau = \log(\lambda_A A).$$

The prediction method, PM2, is as follows. First, estimate $\tau$ by $\hat{\tau} = \log \bar{N}$, where $\bar{N} = \sum_{j=1}^k N_j/k$, and let

$$\widehat{\zeta_0} = \hat{\tau} + \left(\alpha + \frac{1}{2\gamma} - \frac{3}{2}\right)\log\hat{\tau} - \log\left\{2\sqrt{\frac{\gamma}{\pi}}\,\Gamma\left(\alpha + \frac{1}{2\gamma}\right)\right\}.$$

Second, fit $\Lambda(t/\xi - \widehat{\zeta_0})$ to $(W_j^\gamma)_{j=1}^k$ and estimate $\xi$ by the maximum likelihood method. The likelihood equation is solved by the Newton-Raphson method starting from $\widehat{\xi_C} = \bar{W}/(\widehat{\zeta_0} + \gamma_E)$, where $\bar{W} = \sum_{j=1}^k W_j^\gamma/k$, the moment estimate of $\xi$. From $(\hat{\tau}, \widehat{\xi_C})$, we can predict $W_V^\gamma$ by the same way as in Subsection 2.4.

The asymptotic conditional variance of the estimate $\widehat{\xi_C}$ is evaluated by

$$\mathrm{Var}(\widehat{\xi_C}) \doteq \frac{\xi^2}{k\{(\zeta_0+\gamma_E-1)^2+\pi^2/6\}} - \frac{1}{kI_1(\xi)},$$

where $I_1(\xi)$ is the Fisher information of $\xi$ in $\Lambda(t/\xi-\zeta_0)$, provided that $\tau$ is known and $\widehat{\zeta_0}=\zeta_0$ is a constant, and

$$\mathrm{Var}(\widehat{\xi_C}) \doteq \frac{1}{kI_1(\xi)} + \frac{\xi^2}{(\gamma_E+\zeta_0)^2}\mathrm{Var}(\widehat{\zeta_0}), \qquad \text{where} \quad \zeta_0 = E(\widehat{\zeta_0}), \tag{3.2}$$

for the estimate $\widehat{\zeta_0}$ (see Subsection 5.3). Hence we expect that the variance of $\widehat{\xi_C}$ is small if $\zeta_0$ is large. Simulation supports this conjecture.

3.3 *The r largest areas and the number of the circles* (PM3)

The data in $k$ parts of area are combined, and the prediction is based on the $r(\geq 1)$ largest areas on a part of area $kA$. The estimation method is due to Weissman (1978).

If the expected number $\lambda_A(kA)$ of the circles within the part of area $kA$ is larger, we may expect that the asymptotic joint distribution of the $r$ largest order statistics will well approximate the finite distribution, and hence expect that estimate of $\xi$ using the $r$ largest order statistics, $V_1 \geq V_2 \geq \cdots \geq V_r$, is good. However, this estimation is unsatisfactory, and we use the total number $N_W$ of circles within the part of area $kA$ as well as the $r$ largest areas.

The joint density of $X_j = V_j^\gamma$, $j=1,\dots,r$, is approximated by

$$f(x_1,x_2,\dots,x_r;\xi,\zeta) = \xi^{-r}\exp\left[-\exp\left(-\frac{x_r}{\xi}+\zeta\right) - \sum_{j=1}^{r}\left(\frac{x_j}{\xi}-\zeta\right)\right], \tag{3.3}$$
$$x_1 \geq x_2 \geq \cdots \geq x_r,$$

where $\zeta$ is defined by (3.1) with $\tau$ replaced by $\tau_W = \log(\lambda_A(kA))$.

Now, the prediction method is as follows. First, we estimate $\tau_W$ by $\widehat{\tau_W} = \log N_W$, and let

$$\widehat{\zeta_0} = \widehat{\tau_W} + \left(\alpha+\frac{1}{2\gamma}-\frac{3}{2}\right)\log\widehat{\tau_W} - \log\left\{2\sqrt{\frac{\gamma}{\pi}}\,\Gamma\left(\alpha+\frac{1}{2\gamma}\right)\right\}.$$

Second, using this estimate $\widehat{\zeta_0}$, we fit the joint density $f(\cdot;\xi,\widehat{\zeta_0})$ to $(X_j)_{j=1}^r$ and estimate $\xi$ by the maximum likelihood method. The likelihood equation is solved by the Newton-Raphson method starting from the estimate $\widehat{\xi_0} = \bar{X} - X_r$, which was proposed by Weissman (1978). Let $\widehat{\xi_W}$ denote the estimate. From $(\widehat{\tau_W},\widehat{\xi_W})$, we can predict $W_V^\gamma$ by the same way as in Subsection 2.4. In this case, $\hat{\xi}$, $\hat{\tau}$ and $A$ must be replaced by $\widehat{\xi_W}$, $\widehat{\tau_W}$ and $kA$, respectively.

The asymptotic conditional variance of the estimate $\widehat{\xi_W}$ is evaluated by

$$\mathrm{Var}(\widehat{\xi_W}) \doteq \frac{\xi^2}{C_r + \zeta_0\{\zeta_0 r - 2(1+r\psi(r))\}} - \frac{1}{I_r(\xi)},$$

where $I_r(\xi)$ is the Fisher information of $\xi$ in (3.4), provided that $\tau_W$ is known and $\widehat{\zeta_0} = \zeta_0$ is a constant, and

$$(3.4) \qquad \mathrm{Var}(\widehat{\xi_W}) \doteq \frac{1}{I_r(\xi)} + \frac{\xi^2}{(\zeta_0 - \log r)^2}\mathrm{Var}(\widehat{\zeta_0}), \qquad \text{where} \quad \zeta_0 = E(\widehat{\zeta_0}),$$

for the estimate $\widehat{\zeta_0}$, where $C_r = r\{\psi^2(r+1) + \psi'(r+1) + 1\}$, $\psi(x) = \Gamma'(x)/\Gamma(x)$. For the details see Subsection 5.3. Hence, we expect again that the variance of $\widehat{\xi_W}$ is small if $\zeta_0$ is large. Simulation supports this expectation.

Generally, it is hard to choose the optimal $r$ (see, for example, Hughey (1991)). A theoretical method of choosing $r$ is given by Smith (1987). Applying Smith's method, we need the exact distribution of $S_A^\gamma$ and the attraction coefficients which satisfy the condition in Cohen (1982). However, it is difficult to treat directly the exact distribution of $S_A^\gamma$. Moreover, we use the attraction coefficients determined by the asymptotically approximated distribution of $S_A^\gamma$. Hence, Smith's method is not applicable in this case.

Here, we consider the following $r$'s:

$$r_0 = [c_0 \times N_W + 0.5], \qquad r_1 = [c_1 \times \sqrt{N_W} + 0.5],$$
$$r_2 = [c_2 \times \log N_W + 0.5], \qquad r_3 = [c_3 \times (\log N_W)^2 + 0.5],$$

where $c_0$, $c_1$, $c_2$, $c_3$, are constants and $[y]$ is the largest integer not exceeding $y$. The prediction method with $r_i$ is denoted by PM3$i$, $i = 0, 1, 2, 3$.

Another method of choosing $r$ was proposed by Pickands (1975). However, the asymptotic tail of distribution $S_A^\gamma$ is not close to a tail of exponential distribution, Pickands method is not well.

### 3.4 *Threshold method* (PM4)

We measure all the areas of circles exceeding a threshold $u$ and the total number $N_W$ of circles within a part of area $kA$. The areas $U_1, U_2, \ldots, U_n$ ($> u$) and $N_W$ are the available data, where $n$ is a random variable. The variable $X_j = U_j^\gamma - u^\gamma$, $j = 1, \ldots, n$, follows asymptotically ($u \to \infty$) the exponential distribution with p.d.f.

$$f(x;\xi) = \xi^{-1}\exp(-x/\xi), \qquad x \geq 0.$$

See, for example, Pickands (1975). In this case, the prediction method, PM4, is as follows. First, we estimate $\xi$ by $\widehat{\xi_T} = (\sum_{j=1}^n X_j)/n$, and second, we estimate $\tau_W$ by $\widehat{\tau_W} = \log N_W$. Then, we predict $W_V^\gamma$ by the same way as in Subsection 2.4.

Generally, it is hard to choose the threshold $u$ balancing the bias and variance.

## 4. Simulation results

All prediction methods in Section 3 were compared by simulation using S-Plus. The parameter were set to the following practically probable values: the area $S_V$

Table 1a. $\lambda_A A = 5$.

| Parameters | $k = 40$ | $\xi$ | $\tau_V$ | $E(\sqrt{W_V})$ | $\omega^*_{0.95}$ | $\omega^*_{0.99}$ |
|---|---|---|---|---|---|---|
| Methods | True | 1.000 | 13.695 | 14.272 | 16.665 | 18.295 |
| PM1 | Bias | 0.149 | −0.094 | 1.978 | 2.335 | 2.578 |
| | S. D. | 0.141 | 0.327 | 1.679 | 2.016 | 2.246 |
| | M. S. E. | 0.042 | 0.116 | 6.733 | 9.519 | 11.693 |
| PM2 | Bias | 0.156 | −0.138 | 2.061 | 2.435 | 2.689 |
| | S. D. | 0.085 | 0.105 | 1.103 | 1.305 | 1.442 |
| | M. S. E. | 0.031 | 0.030 | 5.466 | 7.632 | 9.313 |
| PM31 | Bias | −0.015 | 0.023 | −0.202 | −0.239 | −0.264 |
| | S. D. | 0.079 | 0.105 | 1.045 | 1.233 | 1.361 |
| | M. S. E. | 0.006 | 0.012 | 1.134 | 1.578 | 1.922 |
| PM32 | Bias | −0.031 | 0.039 | −0.415 | −0.490 | −0.542 |
| | S. D. | 0.071 | 0.100 | 0.942 | 1.111 | 1.226 |
| | M. S. E. | 0.006 | 0.012 | 1.060 | 1.475 | 1.797 |
| PM33 | Bias | −0.015 | 0.023 | −0.194 | −0.229 | −0.253 |
| | S. D. | 0.079 | 0.105 | 1.049 | 1.237 | 1.365 |
| | M. S. E. | 0.006 | 0.011 | 1.137 | 1.582 | 1.927 |
| | $k = 80$ | | | | | |
| PM1 | Bias | 0.165 | −0.153 | 2.153 | 2.547 | 2.815 |
| | S. D. | 0.098 | 0.222 | 1.156 | 1.390 | 1.549 |
| | M. S. E. | 0.037 | 0.073 | 5.972 | 8.417 | 10.324 |
| PM2 | Bias | 0.153 | −0.135 | 2.031 | 2.398 | 2.648 |
| | S. D. | 0.058 | 0.074 | 0.762 | 0.901 | 0.996 |
| | M. S. E. | 0.027 | 0.024 | 4.704 | 6.561 | 8.003 |
| PM31 | Bias | −0.019 | 0.027 | −0.242 | −0.286 | −0.317 |
| | S. D. | 0.057 | 0.075 | 0.766 | 0.903 | 0.996 |
| | M. S. E. | 0.004 | 0.006 | 0.645 | 0.897 | 1.093 |
| PM32 | Bias | −0.025 | 0.033 | −0.327 | −0.387 | −0.428 |
| | S. D. | 0.055 | 0.073 | 0.739 | 0.871 | 0.961 |
| | M. S. E. | 0.004 | 0.007 | 0.653 | 0.908 | 1.106 |
| PM33 | Bias | −0.016 | 0.025 | −0.208 | −0.246 | −0.272 |
| | S. D. | 0.059 | 0.077 | 0.785 | 0.925 | 1.021 |
| | M. S. E. | 0.004 | 0.006 | 0.659 | 0.917 | 1.117 |

of the great circle of the sphere follows Ga$(1, 1/2, 1)$, $k = 40, 80$, $V/A = 200{,}000$, $\lambda_A A = 5, 10, 15$, and the prediction was repeated 1,000 times.

Simulation study was done for PM3$i$ with various values of $c_i$, $i = 0, 1, 2, 3$, and we found that the following three prediction methods are satisfactory:

$$\text{PM31} \quad \text{with} \quad r_1 = [\sqrt{N_W} + 0.5],$$
$$\text{PM32} \quad \text{with} \quad r_2 = [4 \times \log N_W + 0.5],$$

and

Table 1b. $\lambda_A A = 10$.

| Parameters | $k = 40$ | $\xi$ | $\tau_V$ | $E(\sqrt{W_V})$ | $\omega^*_{0.95}$ | $\omega^*_{0.99}$ |
|---|---|---|---|---|---|---|
| Methods | True | 1.000 | 14.388 | 14.965 | 17.358 | 18.988 |
| PM1 | Bias | 0.138 | −0.192 | 1.789 | 2.119 | 2.344 |
| | S. D. | 0.142 | 0.427 | 1.662 | 2.001 | 2.232 |
| | M. S. E. | 0.039 | 0.220 | 5.963 | 8.494 | 10.476 |
| PM2 | Bias | 0.091 | −0.085 | 1.260 | 1.476 | 1.624 |
| | S. D. | 0.060 | 0.076 | 0.829 | 0.971 | 1.069 |
| | M. S. E. | 0.012 | 0.013 | 2.273 | 3.123 | 3.779 |
| PM31 | Bias | −0.018 | 0.020 | −0.256 | −0.300 | −0.329 |
| | S. D. | 0.057 | 0.076 | 0.796 | 0.932 | 1.025 |
| | M. S. E. | 0.004 | 0.006 | 0.700 | 0.959 | 1.158 |
| PM32 | Bias | −0.025 | 0.027 | −0.349 | −0.408 | −0.449 |
| | S. D. | 0.055 | 0.075 | 0.769 | 0.899 | 0.989 |
| | M. S. E. | 0.004 | 0.006 | 0.712 | 0.975 | 1.178 |
| PM33 | Bias | −0.015 | 0.017 | −0.206 | −0.241 | −0.265 |
| | S. D. | 0.059 | 0.078 | 0.833 | 0.975 | 1.071 |
| | M. S. E. | 0.004 | 0.006 | 0.736 | 1.008 | 1.218 |
| | $k = 80$ | | | | | |
| PM1 | Bias | 0.142 | −0.230 | 1.828 | 2.168 | 2.398 |
| | S. D. | 0.103 | 0.299 | 1.209 | 1.455 | 1.623 |
| | M. S. E. | 0.031 | 0.143 | 4.805 | 6.816 | 8.387 |
| PM2 | Bias | 0.088 | −0.082 | 1.228 | 1.439 | 1.582 |
| | S. D. | 0.044 | 0.054 | 0.609 | 0.714 | 0.785 |
| | M. S. E. | 0.010 | 0.010 | 1.878 | 2.579 | 3.120 |
| PM31 | Bias | −0.022 | 0.026 | −0.311 | −0.365 | −0.401 |
| | S. D. | 0.045 | 0.059 | 0.634 | 0.742 | 0.815 |
| | M. S. E. | 0.003 | 0.004 | 0.498 | 0.683 | 0.826 |
| PM32 | Bias | −0.021 | 0.024 | −0.289 | −0.339 | −0.372 |
| | S. D. | 0.046 | 0.059 | 0.643 | 0.753 | 0.827 |
| | M. S. E. | 0.003 | 0.004 | 0.497 | 0.681 | 0.823 |
| PM33 | Bias | −0.016 | 0.019 | −0.219 | −0.256 | −0.282 |
| | S. D. | 0.048 | 0.060 | 0.665 | 0.779 | 0.856 |
| | M. S. E. | 0.003 | 0.004 | 0.490 | 0.672 | 0.813 |

$$\text{PM33} \quad \text{with} \quad r_3 = [0.5 \times (\log N_W)^2 + 0.5].$$

From now on, PM3$i$ denotes the prediction methods with above $r_i$, $i = 1, 2, 3$. Under the considered simulation model, it holds

$$r_1 \approx r_2 \approx r_3.$$

The simulation results for prediction of the expectation $E(\sqrt{W_V})$, 95% quantile $\omega^*_{0.95}$ and 99% quantile $\omega^*_{0.99}$ of the distribution of $\sqrt{W_V}$ are summarized in

Table 1c. $\lambda_A A = 15$.

| Parameters | $k = 40$ | $\xi$ | $\tau_V$ | $E(\sqrt{W_V})$ | $\omega^*_{0.95}$ | $\omega^*_{0.99}$ |
|---|---|---|---|---|---|---|
| Methods | True | 1.000 | 14.793 | 15.371 | 17.764 | 19.303 |
| PM1 | Bias | 0.114 | −0.175 | 1.498 | 1.772 | 1.958 |
| | S. D. | 0.140 | 0.481 | 1.635 | 1.969 | 2.197 |
| | M. S. E. | 0.033 | 0.262 | 4.918 | 7.016 | 8.660 |
| PM2 | Bias | 0.069 | −0.066 | 0.993 | 1.159 | 1.272 |
| | S. D. | 0.052 | 0.065 | 0.743 | 0.867 | 0.951 |
| | M. S. E. | 0.008 | 0.009 | 1.537 | 2.095 | 2.523 |
| PM31 | Bias | −0.020 | 0.021 | −0.290 | −0.338 | −0.370 |
| | S. D. | 0.049 | 0.065 | 0.705 | 0.822 | 0.901 |
| | M. S. E. | 0.003 | 0.005 | 0.580 | 0.789 | 0.949 |
| PM32 | Bias | −0.022 | 0.023 | −0.313 | −0.365 | −0.400 |
| | S. D. | 0.048 | 0.064 | 0.687 | 0.801 | 0.878 |
| | M. S. E. | 0.003 | 0.005 | 0.570 | 0.774 | 0.932 |
| PM33 | Bias | −0.015 | 0.016 | −0.216 | −0.252 | −0.276 |
| | S. D. | 0.051 | 0.067 | 0.733 | 0.854 | 0.937 |
| | M. S. E. | 0.003 | 0.005 | 0.583 | 0.793 | 0.955 |
| | $k = 80$ | | | | | |
| PM1 | Bias | 0.118 | −0.220 | 1.540 | 1.822 | 2.014 |
| | S. D. | 0.095 | 0.315 | 1.120 | 1.347 | 1.502 |
| | M. S. E. | 0.023 | 0.148 | 3.624 | 5.135 | 6.315 |
| PM2 | Bias | 0.067 | −0.062 | 0.969 | 1.130 | 1.240 |
| | S. D. | 0.037 | 0.045 | 0.528 | 0.616 | 0.676 |
| | M. S. E. | 0.006 | 0.006 | 1.217 | 1.657 | 1.995 |
| PM31 | Bias | −0.023 | 0.027 | −0.332 | −0.388 | −0.425 |
| | S. D. | 0.038 | 0.048 | 0.548 | 0.639 | 0.702 |
| | M. S. E. | 0.002 | 0.003 | 0.411 | 0.559 | 0.673 |
| PM32 | Bias | −0.018 | 0.022 | −0.258 | −0.302 | −0.331 |
| | S. D. | 0.041 | 0.051 | 0.595 | 0.693 | 0.761 |
| | M. S. E. | 0.002 | 0.003 | 0.420 | 0.572 | 0.688 |
| PM33 | Bias | −0.015 | 0.019 | −0.217 | −0.253 | −0.278 |
| | S. D. | 0.042 | 0.051 | 0.608 | 0.709 | 0.777 |
| | M. S. E. | 0.002 | 0.003 | 0.416 | 0.566 | 0.682 |

Tables 1a, 1b and 1c. The performance of PM4 is not well, and its simulation results are omitted. Figure 1 shows boxplots of the five predictors in the case of $\lambda_A A = 10$, $k = 40, 80$ and the horizontal dotted line represents the true value of $E(\sqrt{W_V})$. All prediction methods have non-negligible biases, and the mean square errors (m.s.e.'s) of them are computed in Tables 1a, 1b and 1c.

By these tables and figure we evaluate the prediction methods as follows. PM1 and PM2 have positive bias, while PM31, PM32 and PM33 have negative bias. If $k$ increase, then all the prediction methods decrease the s.d.'s. The better

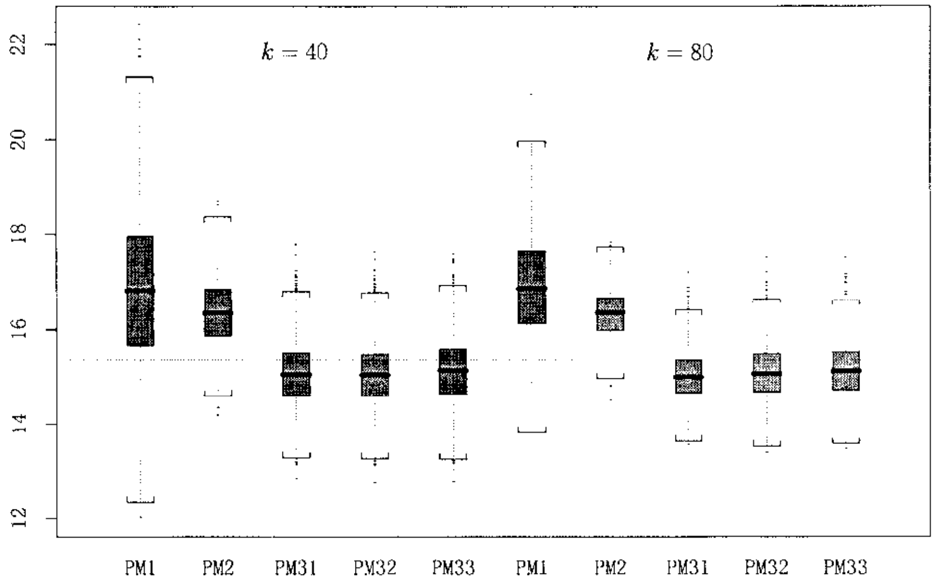


Fig. 1. Prediction of the mean $E(\sqrt{W_V})$, $\lambda_A A = 10$ and $k = 40, 80$.

Table 2. The true value and m.s.e.'s (the upper and lower figures respectively) of prediction $E(W_V^{1/\gamma})$ by PM31, PM32 and PM33 for known $(\alpha, \gamma)$. $\lambda_A A = 10$ and $k = 40$.

| $\alpha$ / $\gamma$ | 0.5 | 1.0 | 1.5 | 2.0 |
|---|---|---|---|---|
| | 13.337 | 14.272 | 15.049 | 15.707 |
| | 0.346 | 0.514 | 1.400 | 3.582 |
| 0.25 | 0.336 | 0.418 | 1.109 | 3.154 |
| | 0.357 | 0.595 | 1.536 | 3.838 |
| | 13.729 | 14.965 | 15.999 | 16.889 |
| | 0.697 | 0.700 | 0.666 | 1.069 |
| 0.5 | 0.645 | 0.712 | 0.589 | 0.837 |
| | 0.741 | 0.736 | 0.722 | 1.217 |
| | 13.612 | 15.086 | 16.294 | 17.327 |
| | 2.591 | 1.019 | 0.967 | 0.957 |
| 1.0 | 2.869 | 0.900 | 0.923 | 0.865 |
| | 2.437 | 1.078 | 1.001 | 1.044 |
| | 13.416 | 15.063 | 16.381 | 17.498 |
| | 12.306 | 2.346 | 1.316 | 1.383 |
| 2.0 | 13.541 | 2.469 | 1.190 | 1.319 |
| | 11.616 | 2.353 | 1.425 | 1.465 |

Note: The case $(\alpha, \gamma) = (1, 1)$ is the exponential distribution that is the unique invariant d.f. of the Wicksell transformation. Thus, the ordinary estimation method is available.

Table 3. The true value, bias, s.d. and m.s.e. (the upper and lower figures respectively) of prediction $E(\sqrt{W_V})$ by PM31 misspecifying $(\alpha, \gamma) = (1, 0.5)$. $\lambda_A A = 10$ and $k = 40$.

| $\gamma$ \ $\alpha$ | 0.7 | 0.8 | 0.9 | 1.0 | 1.1 | 1.2 | 1.3 |
|---|---|---|---|---|---|---|---|
| | | | | 46.626 | | | |
| 0.35 | | | | −13.587 | | | |
| | | | | 2.330 | | | |
| | | | | 190.023 | | | |
| | | | 28.606 | 29.120 | 29.614 | | |
| 0.4 | | | −6.062 | −5.711 | −5.420 | | |
| | | | 1.509 | 1.557 | 1.563 | | |
| | | | 39.024 | 35.041 | 31.817 | | |
| | | 19.463 | 19.806 | 20.135 | 20.453 | 20.760 | |
| 0.45 | | −2.637 | −2.377 | −2.093 | −1.850 | −1.566 | |
| | | 1.041 | 1.071 | 1.060 | 1.052 | 1.154 | |
| | | 8.040 | 6.797 | 5.504 | 4.529 | 3.785 | |
| | 17.654 | 17.977 | 18.287 | 18.586 | 18.874 | 19.153 | 19.422 |
| 0.4625 | −2.313 | −2.054 | −1.791 | −1.525 | −1.279 | −0.955 | −0.670 |
| | 0.962 | 0.950 | 0.993 | 1.034 | 1.008 | 1.069 | 1.073 |
| | 6.273 | 5.121 | 4.194 | 3.396 | 2.651 | 2.054 | 1.601 |
| | 16.378 | 16.672 | 16.954 | 17.227 | 17.489 | 17.743 | 17.988 |
| 0.475 | −1.787 | −1.560 | −1.296 | −1.009 | −0.774 | −0.486 | −0.250 |
| | 0.898 | 0.906 | 0.922 | 0.925 | 0.920 | 0.929 | 0.957 |
| | 3.999 | 3.255 | 2.530 | 1.875 | 1.446 | 1.100 | 0.979 |
| | 15.252 | 15.521 | 15.779 | 16.028 | 16.268 | 16.500 | 16.725 |
| 0.4875 | −1.311 | −1.064 | −0.843 | −0.614 | −0.369 | −0.082 | 0.127 |
| | 0.847 | 0.821 | 0.866 | 0.839 | 0.859 | 0.881 | 0.907 |
| | 2.435 | 1.806 | 1.461 | 1.080 | 0.875 | 0.783 | 0.839 |
| | 14.254 | 14.500 | 14.737 | 14.965 | 15.186 | 15.399 | 15.605 |
| 0.5 | −0.928 | −0.693 | −0.436 | −0.256 | −0.009 | 0.235 | 0.522 |
| | 0.795 | 0.809 | 0.825 | 0.796 | 0.835 | 0.847 | 0.855 |
| | 1.493 | 1.135 | 0.871 | 0.700 | 0.697 | 0.773 | 1.003 |
| | 12.567 | 12.777 | 12.978 | 13.173 | 13.361 | 13.542 | 13.718 |
| 0.525 | −0.294 | −0.098 | 0.156 | 0.378 | 0.614 | 0.825 | 1.080 |
| | 0.690 | 0.683 | 0.710 | 0.720 | 0.732 | 0.731 | 0.730 |
| | 0.563 | 0.476 | 0.529 | 0.661 | 0.913 | 1.215 | 1.698 |
| | | 11.387 | 11.560 | 11.728 | 11.890 | 12.047 | |
| 0.55 | | 0.393 | 0.642 | 0.829 | 1.056 | 1.256 | |
| | | 0.646 | 0.638 | 0.644 | 0.662 | 0.655 | |
| | | 0.572 | 0.819 | 1.102 | 1.552 | 2.007 | |
| | | | 9.438 | 9.567 | 9.691 | | |
| 0.6 | | | 1.300 | 1.497 | 1.669 | | |
| | | | 0.530 | 0.504 | 0.537 | | |
| | | | 1.972 | 2.494 | 3.074 | | |
| | | | | 8.049 | | | |
| 0.65 | | | | 1.888 | | | |
| | | | | 0.442 | | | |
| | | | | 3.761 | | | |

is the estimate of $\xi$, the better is the prediction values. PM2 is better than PM1. PM31, PM32 and PM33 are preferable to PM1 and PM2.

Compared with PM1, PM31, PM32 and PM33 need more measurement work for combining the data in parts of area and counting the total number of circles. However, these prediction methods are accurate under the practically probable assumption ($S_V \sim \mathsf{Ga}(1, 1/2, \cdot)$, $\lambda_A A = 10$ and $k = 40$, see Murakami (1993)). So, we recommend PM31, PM32 and PM33.

Further, the performance of these prediction methods for other parameter values and for misspecified parameters are computed in Tables 2 and 3.

Table 2 shows the simulation result of prediction $E(W_V^{1/\gamma})$ by PM31, PM32 and PM33 for known $(\alpha, \gamma)$. It shows PM31, PM32 and PM33 are relatively good in the neighborhood of $(\alpha, \gamma) = (1, 1/2)$.

Table 3 shows the simulation result of the prediction $E(\sqrt{W_V})$ by PM31 misspecifying $(\alpha, \gamma) = (1, 1/2)$ when $(\alpha, \gamma) \neq (1, 1/2)$. It shows that PM31 is robust if $(\alpha, \gamma)$ is close to $(1, 1/2)$. However, the misspecification can be serious, especially that of $\gamma$ causes bias, it is important for using PM31 to specify rather exactly $\alpha$ and $\gamma$. The simulation results of PM32 and PM33 are similar to that of PM31, so we omit them.

When parameters $\alpha$ and $\gamma$ are unknown, we have to measure all the areas of circles on the sectional plane, but this is not always possible in practice. We shall discuss the prediction method for this case in a sequel paper.

## 5. Supplements

### 5.1 *Maximum of Wicksell transforms*

The limit behavior of the maximum of Wicksell transforms is as follows.

THEOREM 5.1. (1) *The following assertions hold:*

(A.1) $\quad F_A \in \mathcal{D}(L_{1,c-1/2}) \quad \textit{if} \quad F_V \in \mathcal{D}(L_{1c}), \quad c > 1/2.$

(A.2) $\quad F_A \in \mathcal{D}(L_{2,c+1/2}) \quad \textit{if} \quad F_V \in \mathcal{D}(L_{2c}), \quad c > 0.$

(A.3) $\quad F_A \in \mathcal{D}(L_{30}) \quad \textit{if} \quad F_V \in \mathcal{D}(L_{30}).$

(2) *Suppose the p.d.f.'s fulfill the von Mises conditions (see Drees and Reiss* (1992)).

(B.1) *The following three conditions are equivalent, if* $c > 1/2$.

i) $F_V \in \mathcal{D}(L_{1c})$, ii) $F_C \in \mathcal{D}(L_{1,c-1/2})$, iii) $F_A \in \mathcal{D}(L_{1,c-1/2})$.

(B.2) *The following three conditions are equivalent, if* $c > 0$.

i) $F_V \in \mathcal{D}(L_{2c})$, ii) $F_C \in \mathcal{D}(L_{2c})$, iii) $F_A \in \mathcal{D}(L_{2,c+1/2})$.

(B.3) *The following three conditions are equivalent.*

i) $F_V \in \mathcal{D}(L_{30})$, ii) $F_C \in \mathcal{D}(L_{30})$, iii) $F_A \in \mathcal{D}(L_{30})$.

The assertions ii) on $F_C$ are important to discuss the prediction in the problem (C) which is omitted in this paper. The following proof of this theorem is simpler than that of Theorem 1 in Drees and Reiss (1992). We prove only the case of Gumbel distribution ($L_{30}$). The other two cases are similarly proved.

We need the following lemma:

LEMMA 5.1. *Suppose that positive functions $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$ satisfy*

$$\int_0^\omega f(s,v)dv < \infty, \qquad \int_0^\omega g(s,v)dv < \infty$$

*for some $\omega$ $(0 < \omega \leq \infty)$ and for $s \leq v < \omega$*

$$\lim_{s\uparrow\omega} f(s,v)/g(s,v) = c, \qquad 0 \leq c \leq \infty.$$

*Then*

$$\lim_{s\uparrow\omega} \int_s^\omega f(s,v)dv \bigg/ \int_s^\omega g(s,v)dv = c.$$

PROOF OF THEOREM 5.1(A.3). The upper endpoints of the support of $F_V$ and $F_A$ are equal and positive (denoted by $\omega$). Since $F_V \in \mathcal{D}(L_{30})$, there exists the auxiliary function $h$ such that

$$\bar{F}_V(s + xh(s))/\bar{F}_V(s) \to e^{-x}, \qquad \text{as} \quad s \uparrow \omega \tag{5.1}$$

for all $x \in \boldsymbol{R}$. Suppose $h$ satisfies the same conditions in Drees and Reiss ((1992), p. 211), then

$$\begin{aligned}\bar{F}_A(s + xh(s)) &= c\int_{s+xh(s)}^\omega \frac{1}{\sqrt{v-(s+xh(s))}}\bar{F}_V(v)dv \\ &= c\int_s^\omega \frac{1+xh'(t)}{\sqrt{t-s+x(h(t)-h(s))}}\bar{F}_V(t+xh(t))dt,\end{aligned}$$

where $c = 1/(2E(\sqrt{S_V}))$. Hence, by Lemma 5.1, (5.1) and the properties of $h$

$$\bar{F}_A(s + xh(s))/\bar{F}_A(s) \to e^{-x}, \qquad \text{as} \quad s \uparrow \omega,$$

that is, $F_A \in \mathcal{D}(L_{30})$. □

PROOF OF THEOREM 5.1(B.3). (i) $\Longleftrightarrow$ (ii): Suppose $F_V \in \mathcal{D}(L_{30})$, then there exists the auxiliary function $h$ such that

$$f_V(s + xh(s))/f_V(s) \to e^{-x}, \qquad \text{as} \quad s \uparrow \omega,$$

for all $x \in \boldsymbol{R}$, where $h$ satisfies the same conditions in Drees and Reiss (1992). On the other hand, by (2.2), we have $f_C(s) = 2c_1\sqrt{s}f_V(s)$, $c_1 = 1/(2E(\sqrt{S_V}))$. Thus,

$$f_C(s + xh(s))/f_C(s) - \sqrt{s+xh(s)}f_V(s+xh(s))/\sqrt{s}f_V(s) \to e^{-x}, \qquad \text{as} \quad s \uparrow \omega.$$

Hence, $F_C \in \mathcal{D}(L_{30})$.

The converse is proved in an analogous way using the fact that

$$f_V(s) = f_C(s)/(2c_1\sqrt{s}).$$

(i) $\Longleftrightarrow$ (iii): Suppose $F_A \in \mathcal{D}(L_{30})$, then there exists the auxiliary function $h$ which satisfies the same conditions in Drees and Reiss (1992), such that

$$f_A(s + xh(s))/f_A(s) \to e^{-x}, \qquad \text{as} \quad s \uparrow \omega,$$

for all $x \in \boldsymbol{R}$. From (2.4), and replacing $\bar{F}_A$ by $\bar{F}_V$ and $\bar{F}_V$ by $f_A$ in the above proof of Theorem 5.1(A.3), we have $F_V \in \mathcal{D}(L_{30})$.

The converse is trivial from (A.3). □

From this Theorem and the results of de Haan (1970), we have Proposition 2.1.

5.2 *Extreme value in a random sample of random size*

Under the same condition in Subsection 2.2, let $N_\theta$ be a random variable taking non-negative integers (not necessarily Poisson variable) such that

$$p(n;\theta) = P(N_\theta = n), \qquad n = 0, 1, 2, \ldots$$

and

$$0 < \mu_\theta = E(N_\theta) < \infty, \qquad 0 < \sigma_\theta^2 = \mathrm{Var}(N_\theta) < \infty.$$

THEOREM 5.2. *Suppose $L$ is an extreme value distribution and $H \in \mathcal{D}(L)$, that is*

$$H^n(a(n)x + b(n)) \to L(x), \qquad \text{as} \quad n \to \infty.$$

*Suppose the random variable $N_\theta$ has the probability function $p(\cdot;\theta)$ such that $p(0;\theta) \to 0$, $\mu_\theta \to \infty$, and $\sigma_\theta^2 = O(\mu_\theta)$ $(\mu_\theta \to \infty)$, as $\theta \to \theta_1$. Then*

$$H^*(a(\mu_\theta)x + b(\mu_\theta)) \to L(x), \qquad \text{as} \quad \theta \to \theta_1.$$

Theorem 5.2 follows from Galambos (1987), Chapter 6.

5.3 *Evaluation of the expected information*

Following Smith (1986) we evaluate the Fisher informations of $\xi$ in Section 3. First, we assume that $\tau$ is known, $\widehat{\zeta_0} = \zeta_0$ is a constant, and the joint density of $(X_1, \ldots, X_r)$ is $f(\cdot;\xi,\zeta_0)$ in (3.3). The log likelihood is

$$l = -r\log\xi - \exp(-x_r/\xi + \zeta_0) - \sum_{j=1}^{r}(x_j/\xi - \zeta_0),$$

and

$$-(d/(d\xi))^2 l = -r\xi^{-2} + (x_r^2\xi^{-4} - 2x_r\xi^{-3})\exp(-x_r\xi^{-1} + \zeta_0) + 2\xi^{-3}\sum_{j=1}^{r} x_j.$$

Now $Z_j = X_j/\xi - \zeta_0$ has the p.d.f. $\exp(-jz - e^{-z})/\Gamma(j)$, $-\infty < z < \infty$, so for integer $m$ and $\alpha > -j$:

$$E(Z_j^m \exp(-\alpha Z_j)) = (-1)^m \Gamma^{(m)}(j + \alpha)/\Gamma(j),$$

where $\Gamma^{(m)}$ is the $m$-th derivative of the gamma function. Using the relation $X_j = \xi(Z_j + \zeta_0)$, we have

$$\begin{aligned}
&E(X_j) = \xi(\zeta_0 - \psi(j)),\\
&E(X_r e^{-Z_r}) = \xi\{\zeta_0 r - (1 + r\psi(r))\},\\
&E(X_r^2 e^{-Z_r}) = \xi^2\{2\psi(r) + r(\psi'(r) + \psi^2(r)) - 2\zeta_0(1 + r\psi(r)) + \zeta_0^2 r\}.
\end{aligned}$$

Thus, we have

$$I_r(\xi) = E(-(d/(d\xi))^2 l) = \{C_r + \zeta_0[\zeta_0 r - 2(1 + r\psi(r))]\}/\xi^2,$$

and, if $r = 1$,

$$I_1(\xi) = \{(\zeta_0 + \gamma_E - 1)^2 + \pi^2/6\}/\xi^2.$$

Next, to consider the case $\zeta_0$ is the estimate $\widehat{\zeta_0}$, recall that for a general bivariate random vector $(X, Y)$,

$$\operatorname{Var}(Y) = E^X(\operatorname{Var}^Y(Y \mid X)) + \operatorname{Var}^X(E^Y(Y \mid X)). \tag{5.2}$$

If $r = 1$, consider $(\widehat{\zeta_0}, \hat{\xi})$, $\hat{\xi} = \widehat{\xi_C}$, of the sample size $k$. The expectation of conditional variance is

$$E^{\widehat{\zeta_0}}\left(\frac{1}{kI_1(\xi, \widehat{\zeta_0})}\right) \doteq \frac{1}{kI_1(\xi)}.$$

To evaluate the variance of conditional mean, we need, for example, to consider the Edgeworth expansion of $E(\hat{\xi} \mid \widehat{\zeta_0})$, because $\hat{\xi}$ is given by the estimation equation. Assuming the MLE and the moment estimate are nearly equal,

$$\begin{aligned}
&E(\hat{\xi} \mid \widehat{\zeta_0}) \doteq \xi(\gamma_E + \zeta_0)/(\gamma_E + \widehat{\zeta_0}), \qquad \text{where} \quad \zeta_0 = E(\widehat{\zeta_0}),\\
&\frac{\partial}{\partial\zeta} E(\hat{\xi} \mid \zeta)\Big|_{\zeta_0} \doteq -\frac{\xi}{\gamma_E + \zeta_0},
\end{aligned}$$

and the variance of conditional mean is approximately equal to

$$\xi^2(\gamma_E + \zeta_0)^{-2}\operatorname{Var}(\widehat{\zeta_0}).$$

Thus we have

$$\operatorname{Var}(\hat{\xi}) \doteq \frac{1}{kI_1(\xi)} + \frac{\xi^2}{(\gamma_E + \zeta_0)^2}\operatorname{Var}(\widehat{\zeta_0}).$$

If $r$ is sufficiently large, the expectation of conditional variance is approximately equal to $I_r(\xi)^{-1}$. Suppose the maximum likelihood estimate $\hat{\xi} = \widehat{\xi_W}$ is nearly equal to $X_r/(\widehat{\zeta_0} - \log r)$, which is the maximum likelihood estimate proposed by Weissman (1978). Then

$$E(\hat{\xi} \mid \widehat{\zeta_0}) \doteq \frac{\xi(\zeta_0 - \psi(r))}{\widehat{\zeta_0} - \log r}, \qquad \frac{\partial}{\partial \zeta} E(\hat{\xi} \mid \zeta) \Big|_{\zeta_0} \doteq -\frac{\xi(\zeta_0 - \psi(r))}{(\zeta_0 - \log r)^2},$$

and the variance of conditional mean is approximately equal to

$$\frac{\xi^2(\zeta_0 - \psi(r))^2}{(\zeta_0 - \log r)^4} \operatorname{Var}(\widehat{\zeta_0}) \doteq \frac{\xi^2}{(\zeta_0 - \log r)^2} \operatorname{Var}(\widehat{\zeta_0}).$$

Thus, we have

$$\operatorname{Var}(\hat{\xi}) \doteq \frac{1}{I_r(\xi)} + \frac{\xi^2}{(\zeta_0 - \log r)^2} \operatorname{Var}(\widehat{\zeta_0}).$$

## Acknowledgements

This study was supported by Ministry of Education, Japan, Grant C 06680285. The authors thank to Mr. Wataru Sakamoto for the simulation work and the referees for their helpful comments and suggestions.

## References


Cohen, J. P. (1982). Convergence rates for the ultimate and penultimate approximations in extreme-value theory, *Adv. in Appl. Probab.*, **19**, 1811–1825.

de Haan, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tracts, **32**, Amsterdam.

Drees, H. and Reiss, R.-D. (1992). Tail behavior in Wicksell's corpuscle problem, *Probability Theory and Applications: Essays to the Memory of József Mogyoródi* (eds. J. Galambos and I. Kátai), 205–220, Kluwer, Dordrecht.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed., Krieger, Melbourne.

Hughey, R. L. (1991). A survey and comparison of methods for estimating extreme right tail-area quantiles, *Comm. Statist. Theory Methods*, **20**, 1463–1496.

Murakami, Y. (1993). *Metal Fatigue: Effects of Small Defects and Nonmetallic Inclusions*, Yokendo, Tokyo (in Japanese).

Murakami, Y. (1994). Inclusion rating by statistics of extreme values and its application to fatigue strength prediction and quality control of materials, *J. Res. Nat. Inst. Standards Tech.*, **99**, 345–351.

Pickands, J. (1975). Statistical inference using extreme order statistics, *Ann. Statist.*, **3**, 119–131.

Smith, R. L. (1986). Extreme value theory based on the $r$ largest annual events, *Journal of Hydrology*, **86**, 27–43.

Smith, R. L. (1987). Estimating tails of probability distributions, *Ann. Statist.*, **15**, 1174–1207.

Takahashi, R. and Sibuya, M. (1996). The maximum size of the planar sections of random spheres and its application to metallurgy, *Ann. Inst. Statist. Math.*, **48**, 127–144.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the $k$ largest observations, *J. Amer. Statist. Assoc.*, **73**, 812–815.

Wicksell, S. D. (1925). The corpuscle problem. A mathematical study of a biometric problem, *Biometrika*, **17**, 84–99.