Maximum Likelihood Estimation of K Distribution: Application to the Environmental Data

Koji Kanefuji, The Institute of Statistical Mathematics, Japan
Kosei Iwase, Yokohama College of Pharmacy, Japan

ABSTRACT

We propose the K distribution related to the environmental statistics. This distribution has the same properties as the BS distribution in some aspects. For example, X/a and a/X are distributed as the same distribuion. In the statistical analysis for envionmental area studies, this type of property is very important. The K distribution has good character in the MLE. In this study, we also present the MLE of K distribution.

1. Definition of K distribution

The K distribution, $X \sim K(\mu, c^2)$ is defined by the following probability element

$$\frac{1}{2e^{1/c^2}K_0(1/c^2)}x^{-1}e^{-\frac{1}{2c^2}(\sqrt{x}-1/\sqrt{x})^2}dx, 0 < x < \infty.$$

This distribution is one of special cases of the GIG distribution(Barndorff-Nielsen and Halgreen; 1977).

Theorem 1

$$E\left[\left(\frac{X}{\mu}\right)^k\right] = \frac{K_k(1/c^2)}{K_0(1/c^2)}.$$

We can get the following equations.

$$E\left[\sqrt{\frac{X}{\mu}}\right] = E\left[\sqrt{\frac{\mu}{X}}\right] = \frac{c\sqrt{\pi/2}e^{-1/c^2}}{K_0(1/c^2)},$$

$$E\left[\frac{X}{\mu}\right] = E\left[\frac{\mu}{X}\right] = \frac{K_1(1/c^2)}{K_0(1/c^2)},$$

$$E\left[\left(\frac{X}{\mu}\right)^{3/2}\right] = E\left[\left(\frac{X}{\mu}\right)^{-3/2}\right] = \frac{c\sqrt{\pi/2}e^{-1/c^2}}{K_0(1/c^2)}(1+c^2)$$
$$= E\left[\sqrt{\frac{X}{\mu}}\right](1+c^2) = E\left[\sqrt{\frac{\mu}{X}}\right](1+c^2).$$

2. Properties of K distribution

Theorem 2

$$Population\ Arithmetic\ Mean: \mu_A = \mu\frac{K_1(1/c^2)}{K_0(1/c^2)},$$

$$Population\ Harmonic\ Mean: \mu_H = \mu\frac{K_0(1/c^2)}{K_1(1/c^2)}.$$

From the Theorem 2, we can have

$$\frac{\mu_A}{\mu_H} = \left(\frac{K_1(1/c^2)}{K_0(1/c^2)}\right)^2.$$

For fixed number of ν and $|z|$is large,

$$K_\nu(z) \simeq \sqrt{\frac{\pi}{2z}}e^{-z}\left(1 + \frac{4\nu^2-1}{8z} + \cdots\right).$$

$$\frac{K_1(1/c^2)}{K_0(1/c^2)} = 1 + \frac{1}{2}c^2 + \cdots$$

$$\frac{\mu_A}{\mu_H} \simeq 1 + c^2, c << 1$$

$$c^2 \simeq \frac{\mu_A}{\mu_H} - 1.$$

Theorem 3

$$\mu = \sqrt{\mu_A\mu_H},\ \ \mu = \frac{E\left[\sqrt{X}\right]}{E\left[\frac{1}{\sqrt{X}}\right]}.$$

If $X \sim K(\mu, c^2)$ and a is a constant then

$$X \sim K(\mu, c^2) \iff \frac{1}{X} \sim K\left(\frac{1}{\mu}, c^2\right).$$

$$X \sim K(\mu, c^2) \iff aX \sim K(a\mu, c^2).$$

Theorem 4

The parameter μ is population meadian.

$$\mu = \mu_{median}.$$

3. Estimation for K distribution

Theorem 5

$K(\mu, c^2)$ has the moment generating function $m(t)$ as follows

$$m(t) = \frac{K_0\left(\frac{1}{c^2}\sqrt{1-2c^2\mu t}\right)}{K_0\left(\frac{1}{c^2}\right)}, \frac{1}{2c^2} > \mu t.$$

We have the likelihood function $L(\mu, c^2)$ as follows

$$L(\mu, c^2) = \frac{1}{2^2 K_0^n(1/c^2)} \prod_{i=1}^{n} \frac{\mu}{x_i} e^{-\frac{1}{2c^2\mu}\sum_{i=1}^{n} x_i - \frac{\mu}{2c^2}\sum_{i=1}^{n}\frac{1}{x_i}} \frac{(dx)^n}{\mu^n}.$$

Then, the loglikelihood function $l(\mu, c^2)$ is

$$\begin{aligned} l(\mu, c^2) &= -n\log 2 - n\log K_0(1/c^2) + \sum_{i=1}^{n}\log\frac{\mu}{x_i} - \frac{1}{2c^2\mu}\sum_{i=1}^{n} x_i \\ &\quad -\frac{\mu}{2c^2}\sum_{i=1}^{n}\frac{1}{x_i} - n\log\frac{\mu}{dx}. \end{aligned}$$

$$\frac{\partial l}{\partial \mu} = \frac{1}{2c^2}\left(\frac{1}{\mu^2}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\frac{1}{x_i}\right).$$

$$\frac{\partial l}{\partial c^2} = -n\frac{d}{dc^2}\log K_0(1/c^2) + \frac{\sum_{i=1}^{n} x_i}{2c^4\mu} + \frac{\mu\sum_{i=1}^{n}\frac{1}{x_i}}{2c^4}.$$

Theorem 6

We have the MLEs for μ and c^2

$$\hat{\mu}_{MLE} = \sqrt{\bar{X}_A \bar{X}_H}.$$

$$\frac{K_1(1/\hat{c^2}_{MLE})}{K_0(1/\hat{c^2}_{MLE})} = \sqrt{\frac{\bar{X}_A}{\bar{X}_H}},$$

where, $\bar{X}_A = \frac{1}{n}\sum_{i=1}^{n} X_i \ and \ \bar{X}_H = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right)^{-1}$.

Theorem 7

From the probability element, we have the mode of $X \sim (\mu, c^2)$,

$$\mu_{mode} = \mu c^2 \left(\sqrt{1+\frac{1}{c^4}} - 1\right),$$

$$c^2 = \frac{1}{2}\left(\frac{\mu_{median}}{\mu_{mode}} - \frac{\mu_{mode}}{\mu_{median}}\right),$$

$$\texttt{Arcsinh} c^2 = \log\frac{\mu_{median}}{\mu_{mode}}.$$

Theorem 8

For $c << 1$,if $X \sim K(\mu, c^2)$,

$$U = \frac{1}{c}\left(\sqrt{\frac{X}{\mu}} - \sqrt{\frac{\mu}{X}}\right) \to N(0, 1^2).$$

Theorem 9

If $X \sim K(\mu, c^2)$, then we have

$$c^2 = \frac{E\left[\left(\frac{X}{\mu}\right)^{3/2}\right]}{E\left[\left(\frac{X}{\mu}\right)^{1/2}\right]} - 1 = \frac{E\left[\left(\frac{X}{\mu}\right)^{-3/2}\right]}{E\left[\left(\frac{X}{\mu}\right)^{-1/2}\right]} - 1,$$

$$\hat{c^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i/\tilde{X})^{3/2}}{\frac{1}{n}\sum_{i=1}^{n}(X_i/\tilde{X})^{1/2}} - 1,$$

$$\hat{c^2} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i/\tilde{X})^{-3/2}}{\frac{1}{n}\sum_{i=1}^{n}(X_i/\tilde{X})^{-1/2}} - 1.$$

Theorem 10

Let $X \sim K(\mu, c^2)$, then

$$c^2 = E\left[\frac{X}{\mu}\log\frac{X}{\mu}\right],$$

$$E\left[\log\frac{X}{\mu}\right] = 0 = \log 1,$$

$$E\left[\arctan\frac{X}{\mu}\right] = \frac{\pi}{4}.$$

4. Relation between the Lognormal and K distribution

If $X \sim LN(\mu, c^2)$, then its probability element is defined by the following

$$f(x)dx = \frac{1}{\sqrt{2\pi}c}\left(\frac{x}{\mu}\right)^{-1} e^{-\frac{1}{2}\left(\frac{1}{c}\log\frac{x}{\mu}\right)^2}\frac{dx}{\mu},$$

where $[X] = [\mu], [c] = [1], 0 < x < \infty, 0 < c < \infty$.

$LN(\mu, c^2)$	$K(\mu, c^2)$
$\mu = geometric\ mean$	$\mu = geometric\ mean$
$E\left[\log\frac{X}{\mu}\right] = 0$	$E\left[\log\frac{X}{\mu}\right] = 0$
$E\left[\frac{X}{\mu}\log\frac{X}{\mu}\right] = c^2$	$E\left[\frac{X}{\mu}\log\frac{X}{\mu}\right] = c^2$
$E\left[\frac{X/\mu-\mu/X}{2}\log\frac{X}{\mu}\right] = c^2$	$E\left[\frac{X/\mu-\mu/X}{2}\log\frac{X}{\mu}\right] = c^2$
$E\left[\left(\frac{X}{\mu}-1\right)\log\frac{X}{\mu}\right] = c^2$	$E\left[\left(\frac{X}{\mu}-1\right)\log\frac{X}{\mu}\right] = c^2$
$E\left[\left(\log\frac{X}{\mu}\right)^2\right] = c^2$	$E\left[\left(\log\frac{X}{\mu}\right)^2\right] \simeq c^2$

References

Barndorff-NielsenO. E., and Halgreen, C. (1977) Infinite divisibility of teh hyperbolic and generalized inverse Gauusian distribution. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*,**38**, 309-311.