Aplikace matematiky

Viorel Gh. Vodă

Inferential procedures on a generalized Rayleigh variate. II

Aplikace matematiky, Vol. 21 (1976), No. 6, 413--419

Persistent URL: `http://dml.cz/dmlcz/103664`

INFERENTIAL PROCEDURES ON A GENERALIZED RAYLEIGH VARIATE (II)

V. GH. VODĂ

(Received March 28, 1974)

1. CONFIDENCE INTERVALS

The GRV introduced in the first part has the property that if X is the underlying variable then X^2 is a Gamma variate with certain parameters.

In this way, if $x_1, x_2, \ldots, x_n$ is an independent sample on X then it is easy to prove that the statistic $2\theta\xi$ where $\xi = \sum_{i=1}^{n} x_i^2$ is distributed as a chi-square variable with $2(k + 1)\, n$ degrees of freedom, k being the shape parameter of the GRV and θ the scale parameter.

Therefore, we can determine two numbers l_i and l_s such that for a given confidence – say $(1 - \gamma)$ – we have

$$\text{Prob}\,\{l_i < 2\theta\xi < l_s\} = 1 - \gamma\,. \tag{1}$$

The length of the interval for θ is

$$Q = \frac{1}{2\xi}(l_s - l_i) \tag{2}$$

and if we look for Q-minimum, we obtain after some tedious algebra:

$$\int_{l_i}^{l_s} x^{n(k+1)-1} e^{-x/2}\, \mathrm{d}x = (1 - \gamma)\, 2^{nk+n}\, \Gamma(nk + n)\,, \quad \left(\frac{l_s}{l_i}\right)^{nk+n-1} = \exp\{\tfrac{1}{2}(l_s - l_i)\} \tag{3}$$

(see also Vodă [3]) which may provide values for l_i and l_s. In this situation seems to be more convenient to look for confidence intervals for $1/\theta$. We have

$$\text{Prob}\left\{\frac{2\xi}{l_s} < \frac{1}{\theta} < \frac{2\xi}{l_i}\right\} = 1 - \gamma\,. \tag{4}$$

The length is now $\tilde{Q} = 2\xi(1/l_i - 1/l_s)$ and the minimum condition yields finally

$$(5) \qquad \int_{l_i}^{l_s} x^{nk+n-1} e^{-x/2}\, dx = (1-\gamma)\, 2^{nk+n}\, \Gamma(nk+n)\,, \quad \left(\frac{l_s}{l_i}\right)^{nk+n+1} = \\ = \exp\{\tfrac{1}{2}(l_s - l_i)\}$$

which can be used for concrete solutions with the aid of Tate-Klett tables [2] but entering in the cell corresponding to $2(k+1)\,n$ degrees of freedom.

From (4) we obtain easily

$$(6) \qquad \text{Prob}\left\{\frac{\Gamma(k+\frac{3}{2})}{\Gamma(k+1)}\sqrt{\frac{2\xi}{l_s}} < E(X) < \frac{\Gamma(k+\frac{3}{2})}{\Gamma(k+1)}\sqrt{\frac{2\xi}{l_i}}\right\} = 1 - \gamma$$

or

$$(7) \qquad \text{Prob}\,\{\delta(l_s)^{-\frac{1}{2}} < E(X) < \delta(l_i)^{-\frac{1}{2}}\} = 1 - \gamma$$

where

$$(8) \qquad \delta = \frac{\Gamma(k+\frac{3}{2})}{\Gamma(k+1)}\,.\,(2\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}\,.$$

The above relation may be interpreted as a confidence interval with minimum length for the expected – life in a GR model. In the table below we give the values of the constant

$$(9) \qquad \omega = \frac{2^{\frac{1}{2}}\Gamma(k+\frac{3}{2})}{\Gamma(k+1)} \quad \text{where} \quad \delta = \omega(\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}\,.$$

N	Density function	ω	Degrees of freedom
1	Rayleigh ($k = 0$)	1·2533141373	$2n$
2	Maxwell ($k = \frac{1}{2}$)	1·595769121	$3n$

Table 1. Useful constants for computing confidence intervals.

For an application of the method we must take into account that Tate-Klett tables [2] are computed for $n = 2(1)\,30$. Therefore the sample sizes must be limited to stay within the range of the tables (in Rayleigh case: $n \leqq 15$ and in Maxwell case: $n \leqq 10$).

2. PARAMETER ESTIMATION IN THE CASE OF A MIXTURE OF TWO GR VARIABLES

Consider now a random variable X_{mix} characterized by the following density:

$$(10) \qquad X_{\text{mix}} : f_{\text{mix}}(x;\theta_1,\theta_2,p,k) = p\,f(x;\theta_1,k) + (1-p)\,f(x;\theta_2,k)$$

where $x > 0$, $\theta_1, \theta_2 > 0$, $0 < p < 1$ and $k \geqq 0$ are assumed to be known and $f(x; \theta, k)$ is the density of a GRV.

Let our task be to estimate the parameters θ_1, θ_2 and p.

In this way, we shall generalize a former work of Krysicki [13] which concerns the mixture of two simple Rayleigh laws.

We shall apply the same method – namely the method of moments.

It is interesting also to investigate the behaviour of the density (10) with respect to the modal value.

We have

$$(11) \qquad f'_{\text{mix}}(x) = \frac{2(1-p)\,\theta_2^{k+2} x^{2k}(2x^2 - \theta_1^{-1}(2k+1))}{\Gamma(k+1)\exp(\theta_1 x^2)} \cdot \\ \cdot \left[\frac{2x^2 - \theta_2^{-1}(2k+1)}{\theta_1^{-1}(2k+1) - 2x^2} \exp\{(\theta_1 - \theta_2)\,x^2\} - \frac{p}{1-p}\left(\frac{\theta_1}{\theta_2}\right)^{k+2}\right].$$

To find the modal value, we must impose

$$(12) \qquad f'_{\text{mix}}(x; \theta_1, \theta_2, p, k) = 0\,.$$

It is clear that the product (11) vanishes if

$$(13) \qquad x = \left(\frac{2k+1}{2\theta_1}\right)^{\frac{1}{2}}.$$

But this value is not a solution of (12), therefore we have in fact to solve the equations:

$$(14) \qquad \frac{2x^2 - \theta_2^{-1}(2k+1)}{\theta_1^{-1}(2k+1) - 2x^2} \exp\{(\theta_1 - \theta_2)\,x^2\} = \frac{p}{1-p}\left(\frac{\theta_1}{\theta_2}\right)^{k+2}$$

which is a transcendental equation.

Since $(\theta_1/\theta_2)^{k+2} > 0$ for every $\theta_1, \theta_2 > 0$ and $k \geqq 0$ and as $0 < p < 1$, the right-hand side is an increasing function of p, due to the factor $p/(1-p)$.

The left-hand side becomes infinite for x given by (13) and vanishes for

$$(15) \qquad x = \left(\frac{2k+1}{2\theta_2}\right)^{\frac{1}{2}}.$$

Therefore, for x lying in the interval

$$(16) \qquad \mathscr{L} \equiv \left(\left(\frac{2k+1}{2\theta_1}\right)^{\frac{1}{2}},\ \left(\frac{2k+1}{2\theta_2}\right)^{\frac{1}{2}}\right)$$

the left-hand side is positive if $\theta_1/\theta_2 > 1$.

Let us denote for brevity the left-hand side by $g(x)$.

If follows that to study the behaviour of $g(x) = (p/(1-p))\,(\theta_1/\theta_2)^{k+2}$ in the interval $\mathscr{L}$, we have to take the derivative of $g(x)$. We have

$$g'(x) = \frac{8(\theta_1 - \theta_2)\,x}{[(2k+1)\,\theta_1^{-1} - 2x^2]^2} \exp[(\theta_1 - \theta_2)\,x^2] \,. \tag{17}$$
$$. \left\{ -x^4 + \tfrac{1}{2}(2k+1)\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right) x^2 - \frac{(2k+1)(2k+3)}{4\theta_1\theta_2} \right\}.$$

The sign of $g'(x)$ is determined by the expression in parentheses.

The discriminant of the equation in parantheses is

$$\delta = \tfrac{1}{4}(2k+1)^2 \left(\frac{1}{\theta_2^2} - 2\,\frac{2k+5}{2k+1} \cdot \frac{1}{\theta_1\theta_2} + \frac{1}{\theta_1^2} \right). \tag{18}$$

Therefore

$$\delta \geqq 0 \quad \text{if} \quad \frac{1}{\theta_2} \geqq \frac{2k + 5 + 2(4k+6)^{\frac{1}{2}}}{(2k+1)\,\theta_1} \,. \tag{19}$$

Under this condition we obtain for $g'(x)$ two points $x^{(1)}$ and $x^{(2)}$ for which $g'(x^{(i)}) = 0$, $i = 1, 2, \ldots$.

They are given by

$$x^{(i)} = \left[\frac{(2k+1)(\theta_1 + \theta_2)}{4\theta_1\theta_2} \mp \tfrac{1}{2}\delta^{\frac{1}{2}} \right]^{\frac{1}{2}} . \tag{20}$$

The behaviour of g' is indicated in Table 2.

x	$\frac{2k+1}{2\theta_1}$		$x^{(1)}$		$x^{(2)}$		$\left(\frac{2k+1}{2\theta_2}\right)^{\frac{1}{2}}$
$g'(x)$	$+\infty$	$-$	$\searrow 0 \nearrow$ min	$+$	Max $\nearrow 0 \searrow$	$-$	0

Table 2. Behaviour of $g'(x)$ for $x \in \mathscr{L}$.

Therefore

$$g_{\min} = g(x^{(1)}) \quad \text{and} \quad g_{\max} = g(x^{(2)}) \,. \tag{21}$$

It follows also that the values of $(p/(1-p))/(\theta_1/\theta_2)^{k+2}$ vary between $g_{\min}$ and $g_{\max}$. Hence we can determine an interval for p as

$$\frac{g_{\min}}{g_{\min} + (\theta_1/\theta_2)^{k+2}} < p < \frac{g_{\mathrm{Max}}}{g_{\mathrm{Max}} + (\theta_1/\theta_2)^{k+2}} \,. \tag{22}$$

We shall show now that

$$g_{\min} = g(x^{(1)}) = g^*(\theta_1/\theta_2)\,, \tag{23}$$

$$g_{\max} = g(x^{(2)}) = g^{**}(\theta_1/\theta_2)\,. \tag{24}$$

This facts can be easily seen if we write (18) in the form

$$\delta = \frac{(2k+1)^2}{4\theta_1^2}\left(\frac{\theta_1^2}{\theta_2^2} - 2\,\frac{2k+5}{2k+1}\,\frac{\theta_1}{\theta_2} + 1\right), \tag{25}$$

$$g_{\min} = \frac{1 - r - \left(r^2 - 2\,\dfrac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}}{1 - r + \left(r^2 - 2\,\dfrac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}}\,. \tag{26}$$

$$.\exp\left\{\frac{2k+1}{4}\left(r - \frac{1}{r}\right) - \frac{2k+1}{4}\left(1 - \frac{1}{r}\right)\left(r^2 - 2\,\frac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}\right\}$$

where we have denoted $\theta_1/\theta_2 = r$.

We have by similar calculation

$$g_{\mathrm{Max}} = \frac{1 - r + \left(r^2 - 2\,\dfrac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}}{1 - r - \left(r^2 - 2\,\dfrac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}}\,. \tag{27}$$

$$.\exp\left\{\frac{2k+1}{4}\left(r - \frac{1}{r}\right) + \frac{2k+1}{4}\left(1 - \frac{1}{r}\right)\left(r^2 - 2\,\frac{2k+5}{2k+1}\,r + 1\right)^{\frac{1}{2}}\right\}.$$

Now it is clear that we must require

$$r \geqq \frac{2k + 5 + 2(4k+6)^{\frac{1}{2}}}{2k+1} \tag{28}$$

taking into account (19).

For instance, if we wish to tabulate limits for the values of p for different particular densities we must begin from a value of r given by (28) where we have insert the specific value of k.

Example. In the case of Rayleigh distribution we have $k = 0$; therefore

$$r \geqq 5 + 2\sqrt{6} \cong 9{\cdot}899 \tag{29}$$

and tabulation may begin from $r = 10$.

The values $g_{\min}$ and g_{Max} are respectively

(30) $$g_{\min}^{(k=0)} = \frac{1 - r - (r^2 - 10r + 1)^{\frac{1}{2}}}{1 - r + (r^2 - 10r + 1)^{\frac{1}{2}}} \exp\left\{\tfrac{1}{4}(r - \tfrac{1}{4})(r^2 - 10r + 1)^{\frac{1}{2}}\right\},$$

(31) $$g_{\max}^{(k=0)} = \frac{1 - r + (r^2 - 10r + 1)^{\frac{1}{2}}}{1 - r - (r^2 - 10r + 1)^{\frac{1}{2}}} .$$

$$. \exp\left\{\frac{1}{4}\left(r - \frac{1}{r}\right) + \frac{1}{4}\left(1 - \frac{1}{r}\right)(r^2 - 10r + 1)^{\frac{1}{2}}\right\}.$$

As concerns the estimation, let $x_1, x_2, \ldots, x_n$ be an independent sample form the underlying population. Therefore

(32) $$E(X_{\min}^j) = \frac{\Gamma(k + \frac{1}{2}j + 1)}{\Gamma(k + 1)}\left[p\theta_1^{-\frac{1}{2}j} + (1 - p)\,\theta_2^{-\frac{1}{2}j}\right].$$

Since three unknown parameters are involved we take for j successively the values 1, 2, 3.

Hence we obtain the following equations:

(33) $$\hat{p}\hat{\theta}_1^{-\frac{1}{2}} + (1 - \hat{p})\,\hat{\theta}_2^{-\frac{1}{2}} = \frac{\Gamma(k + 1)}{n\,\Gamma(k + \frac{3}{2})}\sum_{i=1}^{n} x_i,$$

(34) $$\hat{p}\hat{\theta}_1^{-1} + (1 - \hat{p})\,\hat{\theta}_2^{-1} = \frac{1}{n(k + 1)}\sum_{i=1}^{n} x_i^2,$$

(35) $$\hat{p}\hat{\theta}_1^{-\frac{3}{2}} + (1 - \hat{p})\,\hat{\theta}_2^{-\frac{3}{2}} = \frac{\Gamma(k + 1)}{n\,\Gamma(k + \frac{5}{2})}\sum_{i=1}^{n} x_i^3$$

Let us denote by u and v the following expressions:

(36) $$u = \hat{\theta}_1^{-\frac{1}{2}}, \quad v = \hat{\theta}_2^{-\frac{1}{2}}.$$

We have after some calculations

(37) $$\hat{p}(u - v) = \frac{\Gamma(k + 1)}{n\,\Gamma(k + \frac{3}{2})}\sum_{i=1}^{n} x_i - v,$$

(38) $$\hat{p}(u^2 - v^2) = \frac{1}{n(k + 1)}\sum_{i=1}^{n} x_i^2 - v^2,$$

(39) $$\hat{p}(u^3 - v^3) = \frac{\Gamma(k + 1)}{n\,\Gamma(k + \frac{5}{2})}\sum_{i=1}^{n} x_i^3 - v^3.$$

Still other calculations yield

$$
(40) \qquad u + v = \frac{\frac{\Gamma(k+1)}{n\,\Gamma(k+\frac{5}{2})} \sum_{i=1}^{n} x_i^3 - \frac{\Gamma(k+1)}{n^2(k+1)\,\Gamma(k+\frac{3}{2})} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2}{\frac{1}{n(k+1)} \sum_{i=1}^{n} x_i^2 - \left[\frac{\Gamma(k+1)}{n\,\Gamma(k+\frac{3}{2})} \sum_{i=1}^{n} x_i\right]^2}
$$

$$
(41) \qquad uv = \frac{\frac{\Gamma(k+1)}{n^2\,\Gamma(k+\frac{3}{2})\,\Gamma(k+\frac{5}{2})} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \frac{1}{n^2(k+1)^2} (\sum_{i=1}^{n} x_i^2)^2}{\frac{1}{n(k+1)} \sum_{i=1}^{n} x_i^2 - \left[\frac{\Gamma(k+1)}{n\,\Gamma(k+\frac{3}{2})} \sum_{i=1}^{n} x_i\right]^2} .
$$

Supposing that the common denominator of the two ratios is not zero we have a second degree equation.

This equation will provide the moment estimators for $\hat{\theta}_1^{-\frac{1}{2}}$ and $\hat{\theta}_2^{-\frac{1}{2}}$. Then an estimate for p is easily established from (37).

References

[1] *W. Krysicki:* Application de la méthode des moments a l'estimation des paramètres d'un melange de deux distributions de Rayleigh. Rev. Stat. Appl. vol. XI (1963), 25–45.

[2] *R. F. Tate, G. W. Klett:* Optimum confidence intervals for the variance of a normal population. J. Amer. Stat. Asoc. 54 (1959), 674–682.

[3] *V. Gh. Vodă:* On the "inverse Rayleigh" distributed random variable. Rep. Stat. Appl. Res. JUSE (Japan), vol. 19 (1972), 13–21.

[4] *V. Gh. Vodă:* Inferential procedures on a generalized Rayleigh variate (I). Apl. mat. 21 (1976), 395–412.

Souhrn

PROCEDURY STATISTICKÉ INDUKCE PRO ZOBECNĚNOU RAYLEIGHOVU PROMĚNNOU (II)

V. Gh. Vodă

V této části se konstruují intervaly spolehlivosti minimální délky pro střední hodnotu zobecněné Rayleighovy proměnné. Dále se studují některé problémy týkající se odhadování ve směsi dvou zobecněných Rayleighových proměnných.

Author's address: Viorel Gh. Vodă, Center of Mathematical Statistics, Ştirbey Voda Street No 174, Bucharest 7000 Romania.