



## TESTS FOR DISCRIMINATION BETWEEN TWO GENERALIZED RAYLEIGH DISTRIBUTIONS

Ion VLADIMIRESCU\*, Radu TUNARU\*\*

\*Faculty of Mathematics & Informatics, University of Craiova

\*\*Department of Economics, Finance & International Business, London Metropolitan University

Corresponding author: Ion Vladimirescu, Str. A.I. Cuza No 13, Craiova 1100 Romania, [vladi@central.ucv.ro](mailto:vladi@central.ucv.ro)

The generalized Rayleigh distribution is a two-parameter family of distributions having Rayleigh, Maxwell and chi-square distributions as particular cases. We propose uniformly most powerful unbiased tests for discriminating between the parameters  $\theta$  of two generalized Rayleigh distributions, conditional on the value of parameters  $k$ .

*Key words:* generalized Rayleigh distribution, population discrimination, uniformly most powerful unbiased tests.

### 1. THE GENERALIZED RAYLEIGH DISTRIBUTION

Let  $k$  and  $\theta$  be positive real numbers. The function  $\tilde{p}(\cdot; \theta, k) : (0, \infty) \rightarrow (0, \infty)$  defined by

$$\tilde{p}(x; \theta, k) = \frac{2\theta^{k+1}}{\Gamma(k+1)} x^{2k+1} e^{-x^2} \quad (1)$$

is a probability density function with respect to the Lebesgue measure restricted to  $(0, \infty)$ .

The probability measure defined over the domain  $(0, \infty)$  having the probability density function  $\tilde{p}(\cdot; \theta, k)$  is called the generalized Rayleigh distribution. This distribution has two parameters  $\theta$  and  $k$  and it is denoted in the following by  $RG_{\theta, k}$ . It was introduced by Voda in [1],[14] and [15] and independently by Bury ([2]) in another similar form. This distribution is different from a noncentral chi-square distribution advocated by Miller and others in [8], [9], [4] and [12], that is used in thermodynamics and signal processing and is also referred to as generalized Rayleigh distribution or Ricean distribution ([6]). The nomination of generalized Rayleigh distribution is somehow unfortunate since Miller generated a lot of research on the direction of norms of Gaussian variates, see [1] and [10].

A wide range of probability distributions with positive support are particular cases of the distribution given in (1). The Rayleigh distribution is obtained when  $k=0$  and  $\theta = \frac{1}{\omega^2}$ , the Maxwell distribution is a particular case for  $k = 1/2$  and  $\theta = \frac{1}{2\omega^2}$ , and when  $k = -1/2$  and  $\theta = \frac{1}{\omega^2}$  then one gets the half normal distribution, see [5]. The  $\chi^2(m)$  is also a special case of the generalized Rayleigh distribution and corresponds to  $k = \frac{m}{2} - 1$  and  $\theta = \frac{1}{2\tau^2}$ .

For statistical modelling and especially for reliability research and applications, employing a model based on the generalized Rayleigh distribution family seems to be more useful than the widely used Weibull model since the latter includes only the Rayleigh distribution while the former encompasses in addition the Maxwell distribution.

The cumulative distribution function of the generalized Rayleigh distribution can be calculated (e.g. [16]) as

$$\Lambda(x; \theta, k) = \frac{\Gamma_{\theta x^2}(k+1)}{\Gamma(k+1)} \quad (2)$$

where  $\Gamma_a(\cdot)$  is the incomplete Gamma function. Voda ([16]) also showed that if a random variable  $X$  is distributed with a generalized Rayleigh distribution then  $X^2$  has a Gamma distribution.

## 2. STATISTICAL MODELLING WHEN $k$ IS KNOWN

When the parameter  $k$  is known then  $\tilde{\rho}(x; \theta, k) = c(\theta)h(x)e^{Q(\theta)T(x)}$  where  $c : (0, \infty) \rightarrow (0, \infty)$ ,  $c(\theta) = \frac{2\theta^{k+1}}{\Gamma(k+1)}$ ,  $h : (0, \infty) \rightarrow (0, \infty)$ ,  $h(x) = x^{2k+1}$ , a  $(B_{(0,\infty)}, B_{(0,\infty)})$ -measurable function,  $Q : (0, \infty) \rightarrow R$ ,  $Q(\theta) = \theta$  and  $T : (0, \infty) \rightarrow R$ ,  $T(x) = -x^2$ , a  $(B_{(0,\infty)}, B_R)$ -measurable function. Hence, when  $k$  is known the statistical model

$$\left( (0, \infty), B_{(0,\infty)}, \{RG_{\theta,k} \mid \theta > 0\} \right) \quad (3)$$

is of exponential type ([3]) and the statistic  $T$  is sufficient for inference on the unknown parameter  $\theta$ .

Following [11] it is possible to choose a  $\sigma$ -finite measure  $\nu$  on  $((0, \infty), B_{(0,\infty)})$  that dominates the statistical model (3). The probability density function of the probability distribution  $RG_{\theta,k}$  with respect to  $\nu$  is

$$\rho(x; \theta) = c(\theta)e^{-\theta x^2} \quad (4)$$

for all positive  $x$ .

## 3. TESTS FOR DISCRIMINATING BETWEEN TWO POPULATIONS

The novelty of this paper consists in testing the discrepancy between two distributions from the generalized Rayleigh family. The tests proposed here refer to the  $\theta$ 's parameters when the  $k$ 's are known. Consider the statistical model

$$\otimes_{i=1}^2 \left( (0, \infty), B_{(0,\infty)}; \{RG_{\theta_i, k_i} \mid \theta_i > 0\} \right)^{n_i} \quad (5)$$

that is dominated by the measure  $\nu^{n_1} \otimes \nu^{n_2}$ , where  $n_1, n_2 \in N^*$  are given.

The probability density of the probability distribution  $(RG_{\theta_1, k_1})^{n_1} \otimes (RG_{\theta_2, k_2})^{n_2}$  with respect to  $\nu^{n_1} \otimes \nu^{n_2}$  is

$$L(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}; \theta_1, \theta_2) = c(\theta_1, \theta_2) \exp \left[ -\theta_1 \sum_{i=1}^{n_1} x_i^2 - \theta_2 \sum_{i=1}^{n_2} y_i^2 \right] \quad (6)$$

for all positive  $x$ 's and  $y$ 's, with  $c(\theta_1, \theta_2) = \frac{2\theta_1^{k_1+1}}{\Gamma(k_1+1)} \cdot \frac{2\theta_2^{k_2+1}}{\Gamma(k_2+1)}$ . The function in (6) can be rewritten equivalently as

$$c(\theta_1, \theta_2) \exp \left[ (\theta_2 - \theta_1) \sum_{i=1}^{n_1} x_i^2 - \theta_2 \left( \sum_{i=1}^{n_1} x_i^2 + \sum_{i=1}^{n_2} y_i^2 \right) \right].$$

Denoting

$$\theta^* = \theta_2 - \theta_1, \omega = -\theta_2, U(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \sum_{i=1}^{n_1} x_i^2, T(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \sum_{i=1}^{n_1} x_i^2 + \sum_{i=1}^{n_2} y_i^2$$

it follows that

$$L(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}; \theta_1, \theta_2) = \tilde{c}(\theta^*, \omega) \exp[\theta^* U + \omega T].$$

Let  $\Omega_0 = \{(\theta_1, \theta_2) \in (0, \infty)^2 \mid \theta_1 = \theta_2\}$ . The function  $h : (0, \infty) \times (0, \infty) \rightarrow R$ ,  $h(u, t) = u/t$  is measurable and for any  $t > 0$  the section application  $h(\cdot, t)$  is increasing.

**Lemma 1.** The statistic  $V = h(U, T)$  has the beta probability distribution  $b_{(k_1+1)n_1; (k_2+1)n_2}$  with respect to the measure  $(RG_{\theta_1, k_1})^{n_1} \otimes (RG_{\theta_2, k_2})^{n_2}$  when  $\theta_1 = \theta_2$ . In other words, the statistic  $V$  is free of parameters on the domain  $\Omega_0$ .

*Proof:* Following [16], page 61, the statistic defined by  $2\theta_1 \sum_{i=1}^{n_1} x_i^2$  has a  $\chi^2(2(k_1+1)n_1)$  probability density function with respect to  $(RG_{\theta_1, k_1})^{n_1}$  while the statistic  $2\theta_2 \sum_{i=1}^{n_2} y_i^2$  has a  $\chi^2(2(k_2+1)n_2)$  probability density function with respect to  $(RG_{\theta_2, k_2})^{n_2}$ . Moreover, they are independent with respect to  $(RG_{\theta_1, k_1})^{n_1} \otimes (RG_{\theta_2, k_2})^{n_2}$ . Since

$$V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \frac{\sum_{i=1}^{n_1} x_i^2}{\sum_{i=1}^{n_1} x_i^2 + \sum_{i=1}^{n_2} y_i^2}$$

then we can say that  $V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) = \frac{\sum_{i=1}^{n_1} x_i^2}{\sum_{i=1}^{n_1} x_i^2 + \sum_{i=1}^{n_2} y_i^2}$  for  $\theta_1 = \theta_2$ . It is obvious now that  $V$  is

Beta distributed with the parameters stated above. This concludes the proof.

Let  $\beta_{p, q; \alpha}$  be the quantile of order  $\alpha$ ,  $F_{p, q}$  be the cumulative distribution function, and  $f_{p, q}$  be the probability distribution function of the Beta distribution with parameters  $p$  and  $q$ . For simplicity we denote  $(RG_{\theta_1, k_1})^{n_1} \otimes (RG_{\theta_2, k_2})^{n_2}$  by  $\mu_{\theta_1}$  when  $\theta_1 = \theta_2$ .

**Theorem 1.** Let  $\alpha \in (0, 1)$  and  $k_1, k_2$  be given. Then

- i. The test  $1_{C_1}$ , provided by the critical region

$$C_1 = \left\{ (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \mid V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) > \beta_{(k_1+1)n_1, (k_2+1)n_2; 1-\alpha} \right\} \quad (7)$$

is uniformly most powerful unbiased at the level of significance  $\alpha$  for testing the null hypothesis  $H_0^{(1)} : \theta_1 \geq \theta_2$  versus the alternative  $H_1^{(1)} : \theta_1 < \theta_2$ .

- ii. The test  $1_{C_2}$ , provided by the critical region

$$C_2 = \left\{ (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \mid V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) < \beta_{(k_1+1)n_1, (k_2+1)n_2; \alpha} \right\} \quad (8)$$

is uniformly most powerful unbiased at the level of significance  $\alpha$  for testing the null hypothesis  $H_0^{(1)} : \theta_1 \leq \theta_2$  versus the alternative  $H_1^{(1)} : \theta_1 > \theta_2$ .

- iii. For testing the null hypothesis  $H_0^{(3)} : \theta_1 = \theta_2$  versus the alternative  $H_1^{(3)} : \theta_1 \neq \theta_2$  the test  $1_{C_3}$  provided by the critical region

$$C_3 = \left\{ (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \mid V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) < c_1 \text{ or } V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) > c_2 \right\} \quad (9)$$

is uniformly most powerful unbiased test at the level of significance  $\alpha$ . The scalars  $c_1, c_2$  satisfy the conditions

$$\int_{c_1}^{c_2} f_{(k_1+1)n_1, (k_2+1)n_2}(x)dx = 1 - \alpha = \int_{c_1}^{c_2} f_{(k_1+1)n_1+1, (k_2+1)n_2}(x)dx \quad (10)$$

*Proof:* Since  $V$  is a free parameter statistic over  $\Omega_0$  then using theorem 1, chapter 5, from [7] it follows that the test given by the critical region  $C_1 = \{(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \mid V(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) > c\}$  where  $c$  is determined from

$$\mu_{\theta_1}(C_1) = \alpha \quad (11)$$

is uniformly most powerful unbiased test for testing the null hypothesis  $\tilde{H}_0^{(1)} : \theta^* \leq 0$  versus the alternative  $\tilde{H}_1^{(1)} : \theta^* > 0$  at the level of significance  $\alpha$ . From (11)

$$\alpha = \mu_{\theta_1}(C_1) = (\mu_{\theta_1} \circ V^{-1})[(c, \infty)] = b_{(k_1+1)n_1, (k_2+1)n_2}(0, \infty) = 1 - F_{(k_1+1)n_1, (k_2+1)n_2}$$

and thus  $c = \beta_{(k_1+1)n_1, (k_2+1)n_2; 1-\alpha}$ , proving i. The proof for ii. is similar with i.

Using again theorem 1, chapter 5, from [7], the test characterized by the critical region  $C_3 = (V < c_1) \cup (V > c_2)$  where  $c_1, c_2$  are determined from the conditions

$$E_{\mu_{\theta_1}}(1_{C_3}) = \alpha, \quad E_{\mu_{\theta_1}}(1_{C_3} \cdot V) = \alpha E_{\mu_{\theta_1}}(V) \quad (12)$$

is uniformly most powerful unbiased at the level of significance  $\alpha$  for testing the null hypothesis  $\tilde{H}_0^{(3)} : \theta^* = 0$  versus the alternative  $\tilde{H}_1^{(3)} : \theta^* \neq 0$ . The proof is finished by showing that the condition (12) is equivalent with (10). This is true because

$$E_{\mu_{\theta_1}}(V) = \frac{(k_1 + 1)n_1}{(k_1 + 1)n_1 + (k_2 + 1)n_2}$$

and

$$xf_{(k_1+1)n_1, (k_2+1)n_2}(x) = \frac{(k_1 + 1)n_1}{(k_1 + 1)n_1 + (k_2 + 1)n_2 - 1} f_{(k_1+1)n_1+1, (k_2+1)n_2-1}(x)$$

for any  $x > 0$ .

#### 4. CONCLUSION

The generalized Rayleigh distribution is a two-parameter class of distributions that includes Rayleigh, Maxwell and chi-square. Statistical models based on this relatively unknown distribution may provide a general platform for inference related to many areas of statistics, probability and engineering.

In this paper some uniformly most powerful unbiased tests were proposed to discriminate between two generalized Rayleigh populations. The tests are conditional on the value of the parameter  $k$ .

#### REFERENCES

1. BLUMENSON, L.E., Miller, K.S., *Properties of generalized Rayleigh distributions*, Annals of Mathematical Statistics, **34**, 3, pp. 903-910, 1963.
2. BURY, K.V., *Statistical Models in Applied Science*, New York, Wiley, 1975.
3. COX, D.R., HINKLEY, D., *Theoretical Statistics*, London, Chapman & Hall, 1974.
4. FORD, F., *A note on the paper of Miller, Bernstein and Blumenson*, Quarterly Applied Mathematics, **17**, pp. 446, 1959-1960.
5. JOHNSON, N.L., KOTZ, S., *Continuous Univariate Distributions*, vol. 1, Boston, Houghton Mifflin Company, 1970.
6. JOHNSON, N.L., KOTZ, S., BALAKRISHNAN, N., *Continuous Univariate Distributions*, vol. 2, second edition, Boston, Wiley, 1995.
7. LEHMANN, E. (1959), *Testing Statistical Hypotheses*, New York, Wiley, 1959.

8. MILLER, K., BERNSTEIN, R., BLUMENSON, L., *Generalized Rayleigh processes*, Quarterly of Applied Mathematics, **16**, pp. 137-145, 1958.
9. MILLER, K., BERNSTEIN, R., BLUMENSON, L., *A note on the paper "Generalized Rayleigh processes"*, Quarterly of Applied Mathematics, **20**, pp. 395, 1963.
10. MILLER, K., SACKROWITZ, H., *Relationships between biased and unbiased Rayleigh distributions*, SIAM Journal on Applied Mathematics, **15**, 6, pp. 1490-1495, 1967.
11. MONFORT, A., *Cours de Statistique Mathematique*, Paris, Economica, 1982.
12. PARK, J., *Moments of the generalized Rayleigh distribution*, Quarterly of Applied Mathematics, pp. 202-232, 1961.
13. VODA, D.G., *On the "inverse Rayleigh" distributed random variable*, Rep. Stat. Appl. Res. (JUSE), **19**, pp. 13-21, 1972.
14. VODA, D.G., *Inferential procedures on a generalized Rayleigh variate: part I*, Aplikace Matematiky, **21**, 6, pp. 395-412, 1976.
15. VODA, D.G., *Inferential procedures on a generalized Rayleigh variate: part II*, Aplikace Matematiky, **21**, 6, pp. 413-419, 1976.
16. VODA, D.G., *New Statistical Models for Product Durability Study*, in Romanian, Bucharest, Romanian Academy, 1980.

*Received March 4, 2003*