Overlap Coefficients Based on Kullback-Leibler of Two Normal Densities: Equal Means Case

Hamza Dhaker[1], Papa Ngom[2], Boubakari Ibrahimouf[3] & Malick Mbodj[4]

[1] Département de mathématiques et statistique,Faculté des Sciences, Université de Moncton, NB, Canada

[2] LMA,Université Cheikh Anta Diop, Dakar, Senegal

[3] Florida International University, Miami, FL, USA

[4] Bowie State University, Bowie, MD, USA

Correspondence: Hamza Dhaker Département de mathématiques et statistique,Université de Moncton, NB., Canada., E1A 3E9, Canada. E-mail: hamza.dhaker@umoncton.ca

Received: January 20, 2019 Accepted: February 23, 2019 Online Published: March 21, 2019

doi:10.5539/jmr.v11n2p114 URL: https://doi.org/10.5539/jmr.v11n2p114

Abstract

Overlap coefficient (OVL) represents the proportion of overlap between two probability distributions, as a measure of the similarity between them. In this paper, we define a new overlap coefficient Λ based on Kullback-Leibler divergence and compare its performance to three known overlap coefficients, namely Matusia's Measure ρ, Morisita's Measure λ, Weitzman's Measure δ. We study their properties, relations between them, and give approximate expressions for the biases and the variances.

Keywords: Kullback-Leibler, overlap coefficients, normal density

1. Introduction

Overlap measure are commonly used in reliability analysis to estimate the proportion of machines or electronic devices that have similar range of failure time. The machines may come from two different sources or may be under different stress, which implies different probability densities of failure time. This proportion can be measured by the overlap coefficients (OVL) of the two densities. For applications of OVLs in ecology refer to Pianka (Pianka, 1973), Hurlbert (Hurlbert, 1978), Horn (Horn, 1966), and in geology refer to Sneath (Sneath, 1977). However, due to the unknown nature of sampling distributions of these measures, decisions are often made using only point estimates.

In the literature, overlap coefficients are mostly used in ecology. Other applications include the lowest bound for the probability of failure in the stress-strength models of reliability analysis (Ichikawa 1993), an estimate of the proportion of genetic deviates in segregating populations (Federer et al. 1963), and a measure of disjunction (Sneath 1977). For more details and applications of OVL coefficients including application on income differentials, please see Mulekar and Mishra (Mulekar and Mishra, 1994, 2000) , Inmanand Bradley (Inmanand Bradley, 1989) and Gastwirth (Gastwirth, 1975).

The overlap coefficient have been used for two exponential populations with different means (Al-Saleh and Samawi 2007 and Dhaker et al. 2017), and Mulekar and Andrade (Mulekar and Andrade, 2017) established a necessary condition to obtain valid values of Weitzman's Measure for normal densities and extend the result to lognormal, exponential, Weibull, and Pareto densities. Sibil and Seemon (Sibil and Seemon, 2019) constructed of the confidence interval of the overlap coefficient under one way random models.

Let X be a random variable defined on the real line for two different populations and $f_1(x)$ and $f_2(x)$ their respective probability density functions. The overlapping coefficients are the common areas under the two functions, defined as follows:

- Matusia's Measure (Matusia, 1955)

$$\rho = \int \sqrt{f_1(x) f_2(x)} dx$$

- Morisita's Measure (Morisita, 1959)

$$\lambda = \frac{2 f_1(x) f_2(x)}{\int [f_1(x)]^2 dx + \int [f_2(x)]^2 dx}$$

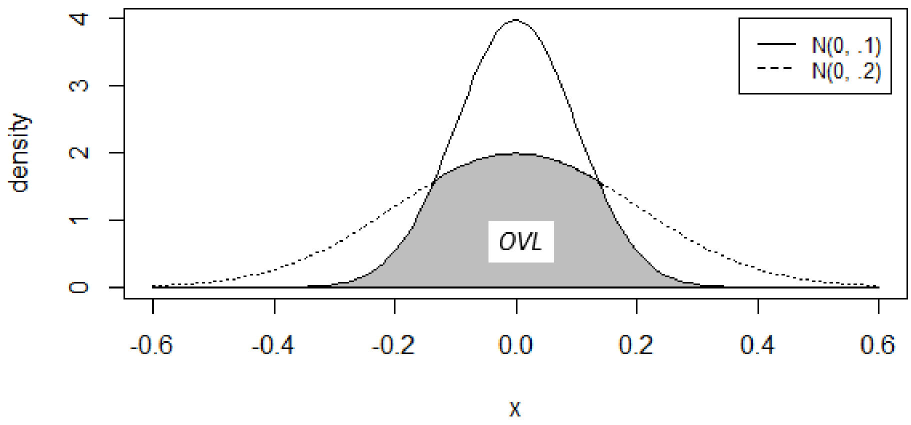

Figure 1. The overlap of two normal densities

- Weitzman's Measure (Weitzman, 1970)

$$\Delta = \int \min\{f_1(x), f_2(x)\}dx$$

- OVL based Kullback-Leibler (Kullback and Leibler, 1951)

$$\Lambda = \frac{1}{1 + KL(f_1 \| f_2)} \tag{1}$$

with $KL(f_1\|f_2) = \int (f_1 - f_2)\log\left(\frac{f_1}{f_2}\right)dx$

Our goal in this paper is to compare the Kullback-Leibler Measure Λ's performance to Matusia's Measure ρ, Morisita's Measure λ, and Weitzman's Measure δ. We study their properties, their relations, in addition to approximating expressions for their biases and variances. The organisation of this paper is as follows. In Section 2, we derive the expressions of the measures described above and study their properties along the lines of Mulekar and Mishra (Mulekar and Mishra, 1994). In Section 3, we provide their maximum likelihood estimators along with approximate bias and variances of $\widehat{OVL}$. In Section 4, a simulation study is perform to evaluate and compare biases and mean square errors of OVL measures estimates. In Section 5 we give an example using a real dataset. Finally, the conclusion is presented in Section 6.

2. Properties of Different Overlap Measures

Let $f_1(x)$ and $f_2(x)$ represent the two populations normal densities with common expectation parameter μ and variances σ_i^2 $(i = 1, 2)$ respectively. We define $C = \sigma_1/\sigma_2$ $(C \geq 0)$ as the ratio of standard deviations. Under the equal means condition, the four similarity measures of interest are given by:

$$\rho = \sqrt{\frac{2C}{1 + C^2}} \tag{2}$$

$$\lambda = \frac{2\sqrt{2}}{\sqrt{1 + C^2}}\left(\frac{C}{1 + C}\right) \tag{3}$$

$$\Delta = \begin{cases} 1 - 2\Phi(b) + 2\Phi(Cb) & if \quad 0 < C < 1 \\ 1 + 2\Phi(b) - 2\Phi(Cb) & if \quad C \geq 1 \end{cases} \tag{4}$$

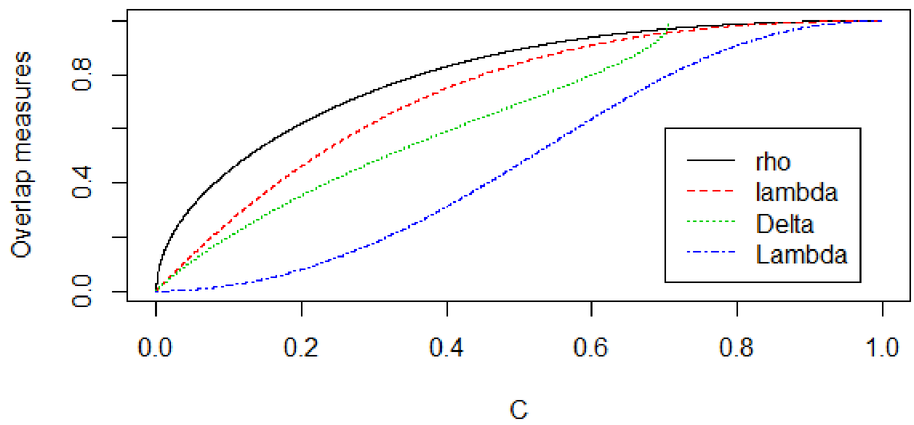

Figure 2. Measures of similarity as functions of C

$$\Lambda = \frac{2C^2}{C^4 + 1} \tag{5}$$

where $b = \sqrt{-\ln C^2/(1 - C^2)}$ and $\Phi(.)$, the cumulative distribution function of a standard normal deviate. Figure 1 shows overlap of $N(0, 0.1)$ and $N(0, 0.2)$.

Figure 2 shows curves of the three overlap measures according to C All three measures are monotone for all $C > 0$. Similar to Mulekar and Mishra (Mulekar and Mishra, 2000), ρ, λ, Δ and Λ have nice properties, such as, symmetry in C, i.e. $OVL(C) = OVL(1/C)$ and invariance under linear transformation, $Y = aX + b, a \neq 0$. They all attain the maximum value of 1 at $R = 1$.

Lemma 1. *For ρ, λ, Δ and Λ defined in equations 2-5, we have*

(i) $\lambda \leq \rho$ *and* $\Delta \leq \rho$.

(ii) $\rho \leq \Lambda$

For all $C > 0$, *equality holding if* $C = 1$.

Proof. (i) Lemma 2 of Mulekar and Mishra (Mulekar and Mishra, 1994).
(ii) we define $h(C)$ by

$$h(C) = \Lambda/\rho = \frac{\sqrt{2}\sqrt{C^3 + C^5}}{1 + C^4}$$

and the derivative with respect to C

$$h^{'}(C) = \frac{3C(1 - C)(1 + C)(C^2 + \frac{4-\sqrt{7}}{3})(C^2 + \frac{4+\sqrt{7}}{3})}{\sqrt{2}\sqrt{C + C^3}(1 + C^4)^2}$$

for $C > 0$ $h^{'}(C) = 0 \Longrightarrow C = 1$. Then $h(C)$ is an increasing function of C for $0 < C < 1$ and a decreasing function of C for $C > 1$ (See Figure 3). Also the sup $h(C) = 1$ is attained at $C = 1$. Thus, $0 < h(C) < 1$ for all $(0 < C < 1)$, which gives the desired result.

3. Estimation of OVL Measures

As in Folks and Chhikara (Folks and Chhikara, 1978), parallel results to those of the two inverse Gaussian populations can be established for the normal populations with common mean. This is based on the following results (Lemma 2 below)

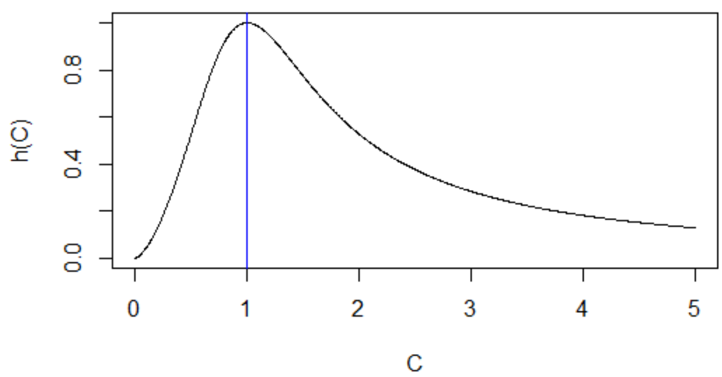

Figure 3. $h(C) = \Lambda/\rho$ as functions of C

from Mulekar and Mishra (Mulekar and Mishra, 1994).
Suppose $(X_{ij}; j = 1, \cdots, n_i; i = 1, 2)$ denote independent observations from two independent ormal random samples draw from $f_1(x)$ and $f_2(x)$ respectively, where:

$$f_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma_1})^2}$$

$$f_2(x) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma_2})^2}$$

Let $C = \sigma_1/\sigma_2$ be the ratio of standard deviations as above. An unbiased estimate of σ_i^2 is given by

$$\sigma_i^2 = S_i^2 = \frac{1}{n-1}\sum_i^{n_i}(X_i - \overline{X})$$

Lemma 2.

$$\mathbb{E}(\widehat{C^2}) = \gamma_1 C^2 \quad Var(\widehat{C^2}) = \gamma_2 C^4$$

where γ_1 and γ_2 are constants in $\mathbb{R}^+$, then γ_1 and γ_2 can be determined as functions of n_1 and n_2 only as,

$$\gamma_1 = \frac{n_2-1}{n_2-3}, \qquad \gamma_2 = \frac{(n_2-1)^2(n_1+1)}{(n_1-1)(n_2-3)(n_2-5)} - \gamma_1^2$$

provided $n_1 > 1$, and $n_2 > 5$.

Proof. For the Proof, please see Mulekar and Mishra (Mulekar and Mishra, 1994)

Theorem 1. *Let $\widehat{\rho}$, $\widehat{\lambda}$, $\widehat{\Delta}$ and $\widehat{\Lambda}$ be the estimates of ρ, λ, Δ and Λ respectively, by substituting $\widehat{C^2}$ for C^2, then for $n_1 > 1$ and $n_2 > 5$ we have the approximate expressions for bias and variance of $\widehat{\rho}$, $\widehat{\lambda}$, $\widehat{\Delta}$ and $\widehat{\Lambda}$ given above.*

where $H_{\widehat{OVL}} = \gamma_2(\gamma_1 - 1)Bias^2(\widehat{OVL})$, $\phi(.)$ is the density function of standard normal variate and

$$I_C = \begin{cases} 1 & if \quad 0 < C < 1 \\ -1 & if \quad C \geq 1 \end{cases}$$

Proof. Let $g(\theta)$ a one parameter function of θ and let $\widehat{\theta}$ be an almost sure consistent estimate of θ. Then the mean and variance of $g(\widehat{\theta})$ may be obtained from the linear Taylor approximation around θ.

OVL	$Bias(\widehat{OVL})$	$Var(\widehat{OVL})$
ρ	$\frac{(\gamma_1-1)\rho}{4}\frac{(1-C^2)}{1+C^2}$	$H_{\widehat{\rho}}$
λ	$\frac{\lambda(\gamma_1-1)}{2}\left[\frac{(1-C)(1+C+C^2)}{(1+C)(1+C^2)}\right]$	$H_{\widehat{\lambda}}$
Δ	$(\gamma_1-1)\left\{(C\phi(Cb)-\phi(b))\left(\frac{C^2b^2-1}{b(1-C^2)}\right)+Cb\phi(Cb)\right\}I_C$	$H_{\widehat{\Delta}}$
Λ	$(\gamma_1-1)\Lambda\frac{1-C^4}{1+C^4}$	$H_{\widehat{\Lambda}}$

$$g_1(\widehat{\theta}) = g_1(\theta) + (\widehat{\theta}-\theta)g_1^{'}(\theta) \tag{6}$$

For example, letting $\theta = C^2$, the estimator of $\widehat{\Lambda}$:

$$\widehat{\Lambda} = g_1(\widehat{\theta}), \quad g_1(\theta) = \frac{2\theta}{1+\theta^2}.$$

Since, in this case,

$$g_1^{'}(\theta) = \frac{2(1-\theta^2)}{(1+\theta^2)^2}$$

from (6)

$$\begin{aligned} Bias(\widehat{\Lambda}) = \mathbb{E}(\widehat{\Lambda}) - \Lambda = \mathbb{E}(\widehat{\theta}-\theta)g_1^{'}(\theta) &= (\gamma_1-1)\frac{2\theta}{1+\theta^2}\frac{1-\theta^2}{1+\theta^2} \\ &= (\gamma_1-1)\Lambda\frac{1-\theta^2}{1+\theta^2} \quad (7) \\ &= (\gamma_1-1)\Lambda\frac{1-C^4}{1+C^4} \quad (8) \end{aligned}$$

for the estimator $\widehat{\lambda}$:

$$\widehat{\lambda} = g_2(\widehat{\theta}) = \sqrt{\frac{2\sqrt{\theta}}{1+\theta}}$$

Since, in this case,

$$g_2^{'}(\theta) = \sqrt{2}\frac{\frac{1}{4}\theta^{-3/4} - (1+\theta)^{-1/2}\theta^{1/4}}{1+\theta}$$

$$\begin{aligned} Bias(\widehat{\rho}) = \mathbb{E}(\widehat{\rho}) - \rho = \mathbb{E}(\widehat{\theta}-\theta)g_2^{'}(\theta) &= (\gamma_1-1)\theta\sqrt{2}\frac{\frac{1}{4}\theta^{-3/4} - (1+\theta)^{-1/2}\theta^{1/4}}{1+\theta} \\ &= (\gamma_1-1)\rho\frac{\frac{1}{4}\sqrt{1+\theta} - \frac{1}{2}\frac{\theta}{\sqrt{1+\theta}}}{\sqrt{1+\theta}} \quad (9) \\ &= \frac{(\gamma_1-1)}{4}\rho\frac{1-\theta}{1+\theta} \quad (10) \\ &= \frac{(\gamma_1-1)}{4}\rho\frac{1-C^2}{1+C^2} \quad (11) \end{aligned}$$

Similar arguments can be used for the other overlap coefficients.

The MLEs for the two-parameter Normal distribution are asymptotically efficient and they are asymptotically normally distributed (see, Mulekar and Mishra 1994). However, the OVL measures are functions of the Normal distribution parameters. Therefore, by using the Delta-method, the OVL measures estimators are asymptotically normally distributed. Thus, the $100(1-\alpha)\%$ approximate confidence intervals are given by

$$\widehat{OVL} \pm Z_{1-\alpha/2}\sqrt{Var(\widehat{OVL})}$$

where $Z_{1-\alpha/2}$ is the $\alpha/2$ upper quantile of the standard normal distribution.

For large samples these confidence intervals work fairly well. However, for small sample sizes more refined versions of the above confidence intervals can be obtained by

$$\left\{\widehat{OVL} - Bias(\widehat{OVL}) - Z_{1-\alpha/2}\sqrt{Var(\widehat{OVL})}, \widehat{OVL} - Bias(\widehat{OVL}) + Z_{1-\alpha/2}\sqrt{Var(\widehat{OVL})}\right\}$$

4. Simulation Study

We performed a numerical study to examine the behavior of overlap coefficients and for comparing the approximate formula for biases and mean square errors derived in the previous section for different OVL measures. Samples of sizes $n = 10$, 25, 50, 100, 200, and 500 each were generated. From each pair of generated samples, the similarity measures ρ, λ, Δ and Λ were estimated and the amount of biases and the standard deviations of the estimates were determined. The mean squared error (MSE) and bias values for $C = 0.05, 0.25, 0.5, 0.75, 0.95$ are reported in Table 1. Tables 1 indicates that the bias of proposed OVL estimators is negligible and decreases as the sample size n increases. As expected, both bias and MSE decrease steadily as the sample size increases.

The bias estimates for $n = 25$ are plotted in Figure 4. Only one plot of bias values is presented because a similar pattern is observed for other sample sizes. For $C < 0.6$ the bias estimates of the measures $\widehat{\lambda}$, $\widehat{\Delta}$ and $\widehat{\Lambda}$ behave more similarly, but for the bias of $\widehat{\rho}$ shows a different pattern. For $C > 0.6$, the bias estimates of the measures $\widehat{\lambda}$ and $\widehat{\Delta}$ are still growing, but for that of $\widehat{\rho}$ and $\widehat{\Lambda}$ are decreasing and tends towards 0.

The standard deviation estimates for the overlap coefficients with sample size 25 are plotted in Figure 5. Again, only one figure for standard deviations is presented because similar pattern is observed for the other sample sizes. The standard deviation estimates for all four coefficients show the same behavior as the bias estimates of the measures.

The estimate of MSE are plotted in Figure 6 for all four overlap coefficients. For $C < 0.3$, the MSE estimates for the overlap coefficients have almost the same values. the MSE estimates of the measures $\widehat{\lambda}$ and $\widehat{\Delta}$ are still growing, but for that of $\widehat{\rho}$ is decreasing and tends towards 0, and for $\widehat{\Lambda}$ has a peak at $C = 0.6$ and declining steadily thereafter as C increases.

Table 1. Bias and MSE of Estimates of OVLs

	$\widehat{\rho}$		$\widehat{\lambda}$		$\widehat{\Delta}$		$\widehat{\Lambda}$	
n	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*	*Bias*	*MSE*
c=0.05	$\rho = 0.316$		$\lambda = 0.134$		$\Delta = 0.112$		$\Lambda = 0.005$	
10	0.022	0.008	0.018	0.005	0.014	0.003	0.001	0.000*
25	0.007	0.001	0.006	0.0009	0.004	0.0006	0.0004	0.000*
50	0.003	0.0006	0.003	0.0004	0.002	0.0002	0.0002	0.000*
100	0.0016	0.0003	0.0013	0.0002	0.001	0.0001	0.0001	0.000*
200	0.0008	0.0001	0.0006	0.000*	0.0005	0.000*	0.000*	0.000*
500	0.0003	0.000*	0.0002	0.000*	0.0002	0.000*	0.000*	0.000*
c=0.25	$\rho = 0.686$		$\lambda = 0.549$		$\Delta = 0.418$		$\Lambda = 0.124$	
10	0.043	0.029	0.058	0.052	0.045	0.031	0.035	0.02
25	0.014	0.005	0.0185	0.01	0.0142	0.006	0.011	0.004
50	0.006	0.0022	0.009	0.004	0.007	0.002	0.005	0.001
100	0.0031	0.001	0.0042	0.0018	0.0032	0.001	0.0025	0.0007
200	0.0015	0.0005	0.0021	0.0009	0.0016	0.0005	0.0012	0.0003
500	0.0006	0.0002	0.0008	0.0003	0.0006	0.0002	0.0005	0.0001
c=0.5	$\rho = 0.894$		$\lambda = 0.843$		$\Delta = 0.677$		$\Lambda = 0.470$	
10	0.038	0.023	0.056	0.049	0.061	0.058	0.12	0.217
25	0.012	0.0042	0.018	0.0091	0.019	0.011	0.038	0.041
50	0.0057	0.0017	0.0084	0.0037	0.0092	0.0045	0.0177	0.017
100	0.0028	0.0008	0.004	0.0017	0.0044	0.0020	0.008	0.0076
200	0.0014	0.0004	0.002	0.0008	0.0022	0.001	0.0042	0.004
500	0.0005	0.0001	0.0008	0.0003	0.0009	0.0004	0.0017	0.001
c=0.75	$\rho = 0.988$		$\lambda = 0.970$		$\Delta = 0.862$		$\Lambda = 0.855$	
10	0.019	0.006	0.029	0.013	0.068	0.07	0.127	0.025
25	0.006	0.001	0.009	0.002	0.021	0.013	0.04	0.046
50	0.0029	0.0004	0.0044	0.001	0.01	0.005	0.019	0.019
100	0.0014	0.0002	0.0021	0.0005	005	0.002	0.009	0.009
200	0.0007	0.000*	0.001	0.0002	0.002	0.001	0.0045	0.001
500	0.0003	0.000*	0.0004	0.000*	0.0009	0.0005	0.002	0.001
c=0.95	$\rho = 0.999$		$\lambda = 0.999$		$\Delta = 0.975$		$\Lambda = 0.995$	
10	0.004	0.0002	0.005	0.0005	0.069	0.07	0.03	0.01
25	0.0012	0.000*	0.002	0.000*	0.022	0.01	0.009	0.002
50	0.0005	0.000*	0.0008	0.000*	0.01	0.006	0.004	0.0009
100	0.0003	0.000*	0.0004	0.000*	005	0.002	0.002	0.0004
200	0.0001	0.000*	0.0002	0.000*	0.002	0.001	0.001	0.0002
500	0.000*	0.000*	0.000*	0.000*	0.0009	0.0005	0.0004	0.000*

$^*|value| < 0.00001$

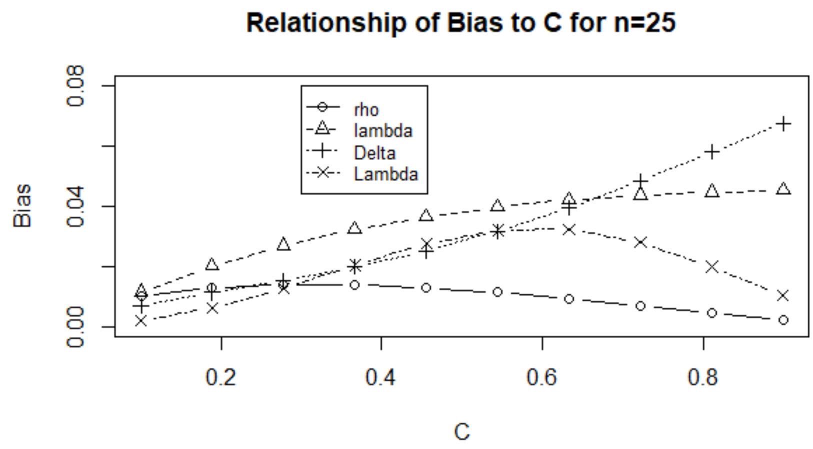

Figure 4. The bias estimates of overlap coefficients by *C*

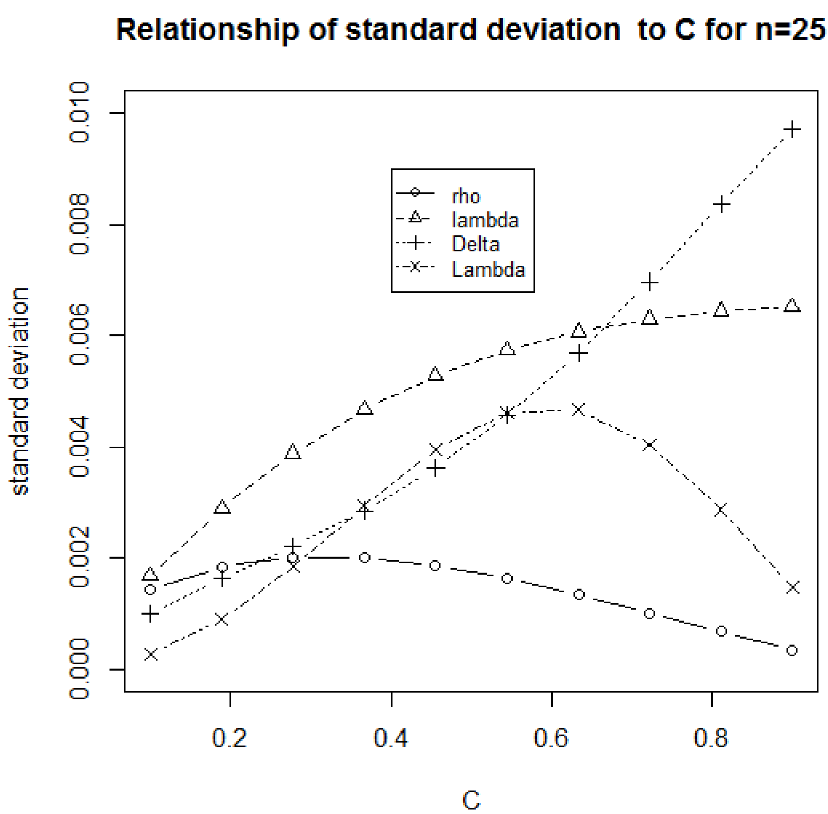

Figure 5. The standard deviation estimates of overlap coefficients by *C*

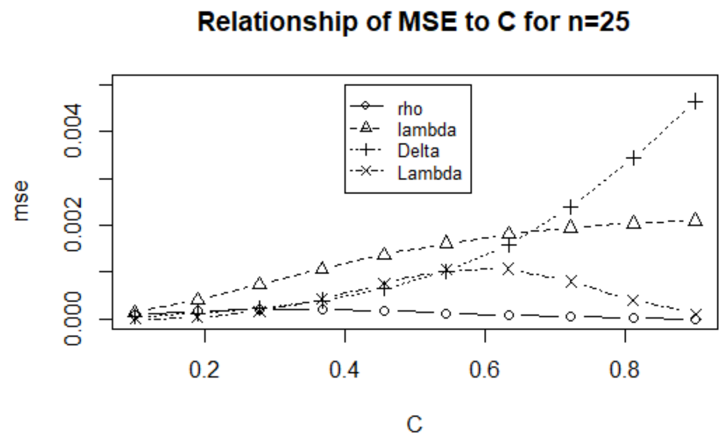

Figure 6. The MSE estimates for overlap coefficients by C

5. Example

As application of the new method, consider the dataset discussed by Federer et al. (Federer et al., 1963). The authors estimated the superior genetic deviates in segregating populations of sugar beets. The population is assumed to be normally distributed with mean μ (as in Mulekar and Mishra 1994). Using a sample of size 320, the estimated total and environmental deviations are $\widehat{\sigma}_t^2 = 0.911$ and $\widehat{\sigma}_e^2 = 0.509$, respectively (total variance= genetic variance + environmental variance). The estimate of the ratio $\widehat{C}$ is given as $\widehat{C} = \widehat{\sigma}_e/\widehat{\sigma}_t = 0.7475$. Applying Theorem 1, the estimates of bias, variance, and the 95%, confidence intervals for the OVLs are obtained and presented in Table 2.

From Table 2, all four confidence intervals does not include the value 1 the population distributions for the two groups should not be considered to be identical. However, the large lower bounds for the OVLs (For the four similarity coefficients) indicates substantial similarity between the two distributions.

Table 2. Results based on the real data of Federer et al (1963)

	ρ	λ	Δ	Λ
$\widehat{OVL}$	0.9793	0.9691	0.8701	0.8516
$Bias(\widehat{OVL})$	0.00044	0.00065	0.00216	0.00281
$Var(\widehat{OVL})$	0.00006	0.00014	0.0015	0.0025
95% confidence	(0.963, 0.994)	(0.945, 0.991)	(0.806, 0.911)	(0.772, 0.930)

6. Conclusion

In this paper we considered four measures of overlap, namely Matusia's measure ρ, Morisita's measure λ, Weitzman's measure Δ and Kullback-Leibler Λ. We used these measures in the case two Normal distributions having the same expectations and different standard deviations. The overall conclusion is that the biases and MSE of each of the OVL measures are close to zero and approximations are adequate for samples of size as small as 50. The values of the OVL measures are very similar, the coefficient based on Kullback-Leibler is always one of the best for having small values of Bias and MSE. It is clear, in general, that the approximations to bias and MSE presented here may require extremely large samples for example $n > 50$.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

References

Al-Saleh, M. F., & Samawi, H. M. (2007). Interference on overlapping coefficients in two exponential populations. *Journal of Modern Applied Statistical Methods, 6*(2), 15. https://doi.org/10.22237/jmasm/1193890440

Dhaker, H., Ngom, P., & Mbodj, M. (2017). Overlap coefficients based on Kullback-Leibler divergence: exponential populations case. *International Journal of Applied Mathematical Research, 6*(4), 135-140. https://doi.org/10.14419/ijamr.v6i4.8493

Federer, W. T., Powers, L., & Payne, M. G. (1963). Studies on statistical procedures applied to chemical genetic data from sugar beets. Technical Bulletin, Agricultural Experimentation Station, Colorado State University 77.

Folks, J. L., & Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application–a review. *Journal of the Royal Statistical Society. Series B (Methodological)*, 263-289. https://doi.org/10.1111/j.2517-6161.1978.tb01039.x

Gastwirth, J. L. (1975). Statistical measures of earnings differentials. *The American Statistician, 29*(1), 32-35.

Horn, H. S. (1966). Measurement of" overlap" in comparative ecological studies. *The American Naturalist, 100*(914), 419-424. https://doi.org/10.1086/282436

Hurlbert, S. H. (1978). The measurement of niche overlap and some relatives. *Ecology, 59*(1), 67-77. https://doi.org/10.2307/1936632

Ichikawa, M. (1993). A meaning of the overlapped area under probability density curves of stress and strength. *Reliability Engineering and System Safety, 41*(2), 203-204. https://doi.org/10.1016/0951-8320(93)90033-U

Inman, H. F., & Bradley Jr, E. L. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Communications in Statistics-Theory and Methods, 18*(10), 3851-3874. https://doi.org/10.1080/03610928908830127

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics, 22*(1), 79-86. https://doi.org/10.1214/aoms/1177729694

Matusita, K. (1955). Decision rules, based on the distance, for problems of fit, two samples, and estimation. *The Annals of Mathematical Statistics, 26*(4), 631-640. https://doi.org/10.1214/aoms/1177728422

Morisita, M. (1959). Measuring of interspeci?c association and similarity between communities. *Mem. Fac. Sci. Kyushu Univ. Series E, 3*, 65-80.

Mulekar, M. S., & Mishra, S. N. (1994). Overlap coefficients of two normal densities: equal means case. *Journal of the Japan Statistical Society, Japanese Issue, 24*(2), 169-180.

Mulekar, M. S., & Mishra, S. N. (2000). Confidence interval estimation of overlap: equal means case. *Computational statistics and data analysis, 34*(2), 121-137.

Mulekar, M. S., & Pereira De Andrade, N. (2017). A necessary condition for computation of the overlap measure. *American Journal of Mathematical and Management Sciences, 36*(2), 112-117.

Pianka, E. R. (1973). The structure of lizard communities. *Annual review of ecology and systematics, 4*(1), 53-74. https://doi.org/10.1146/annurev.es.04.110173.000413

Jose, S., & Thomas, S. (2019). Interval Estimation of the Overlapping Coefficient of Two Normal Distributions: One Way ANOVA with Random Effects. *Thailand Statistician, 17*(1), 84-92.

Sneath, P. H. A. (1977). A method for testing the distinctness of clusters: a test of the disjunction of two clusters in euclidean space as measured by their overlap. *Journal of the International Association for Mathematical Geology, 9*(2), 123-143. https://doi.org/10.1007/BF02312508

Weibull, W. (1939). A statistical theory of the strength of materials. *Ing. Vetenskaps Akad. Handl. 151*, 1–45.

Weitzman, M. S. (1970). Measures of overlap of income distributions of white and Negro families in the United States (Vol. 22). US Bureau of the Census.