ESTIMATES OF [illegible]

BY

[illegible] J. [illegible]

[illegible] REPORT NO. [illegible]

[illegible]EMBER 5, 1985

PREPARED UNDER CONTRACT

[illegible]-76-C-0475 (NR-042-267)

[illegible] OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
[illegible] UNIVERSITY
[illegible], CALIFORNIA

ESTIMATES OF CIRCULAR ERROR PROBABILITIES

BY

M. EVANS, Z. GOVINDARAJULU, and J. BARTHOULOT

TECHNICAL REPORT NO. 367
DECEMBER 5, 1985

Prepared Under Contract
N00014-76-C-0475 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ESTIMATES OF CIRCULAR ERROR PROBABILITIES

by

M. Evans, Z. Govindarajulu, and J. Barthoulot

1. Introduction

Consider the statistical model $M=(X,A,\{P_\theta:\theta\in\Omega\})$ and event $C\in A$ and suppose we wish to estimate $P_\theta(C)$ based on a sample $X=(x_1,\ldots,x_n)$ from M. The typical approach to this problem is to select a probability measure Q^X on A, dependent on the observed data, and then quote $Q^X(C)$ as the estimate.

For example the nonparametric estimator of $P_\theta(C)$ is

$$Q_E^X(C)=\frac{1}{n}\sum_{i=1}^{n}I_C(x_i) \tag{1}$$

where I_C is the indicator function of C; i.e. the empirical probability content of C. For a sufficiently broad class $\{P_\theta:\theta\in\Omega\}$ this estimator is known to be UMVU.

If m(X) is a complete minimal sufficient statistic for $\{P_\theta:\theta\in\Omega\}$ then

$$Q_U^X(C)=E[Q_E^X(C):m(X)] \tag{2}$$

is UMVU. Clearly Q_U^X is a probability measure on A as it is formed by mixing probability distributions.

Perhaps the most commonly used method of obtaining an estimate is to choose some estimator $\hat{\theta}(X)$ of θ; e.g. the MLE, and then quote

$$Q_S^X(C)=P_{\hat{\theta}(X)}(C) \tag{3}$$

as the estimate.

In a Bayesian context, or perhaps just as a method of generating a plausible estimate, a prior for θ gives rise to a posterior P_X for θ which in turn induces a distribution for $P_\theta(C)$. Then the minimum Bayes risk estimator with respect to squared error loss is given by

$$Q_B^X(C)=\int_\Omega P_\theta(C)\,dP_X(\theta) \tag{4}$$

the expected value of P_θ with respect to the posterior. Again Q_B^X is a probability measure on A as it is a mixture of probability measures.

Other strategies could also be devised for obtaining estimates but we will restrict our discussion to those presented above. In all of these approaches we note that the choice of Q_X does not depend on C. As such it seems more appropriate to say we are estimating P_θ rather than $P_\theta(C)$.

One way of inducing at least some dependence on C is via the joint invariant group of the model M and the event C; namely the class G of those 1-1, bimeasurable $g:X \to X$ satisfying gC= C and both of $P_\theta g$ and $P_\theta g^{-1}$ are in $\{P_\theta:\theta\in\Omega\}$ for all $\theta\in\Omega$. If X also satisfies topological requirements then it makes sense to require that g also preserve this structure; e.g. if $X=R^p$ then we require g to be a diffeomorphism (1-1, onto and infinitely differentiable both ways). If $x\sim P_\theta$ then $gx\sim P_\theta g^{-1}$ and $P_\theta(C)=P_\theta(g^{-1}C)=P_\theta g^{-1}(C)$. Thus the estimate of $P_\theta(C)$ should satisfy $Q^X(C)=Q^{gX}(C)$; i.e. our estimate should be the same whether we observe X or gX. As we shall see, this criterion leads to some restriction in the class of possible estimators for the problem we consider in the succeeding sections.

2. Circular Error Probabilities for the Bivariate Normal

Suppose that $x\sim N_2(\mu,\Sigma)$ and $C_k=\{x: x'x\leq k^2\}$. Thus x could give the coordinates of the hitting point for some projectile aimed at bullseye 0 and we wish to estimate the probability of coming within k of 0 as an assessment of the accuracy of the targeting procedure.

Even when (μ,Σ) is known the problem of calculating $P_{(\mu,\Sigma)}(C_k)$ is significant. For various tabulations and results related to this problem see, for example, Grad and Solomon(1955), Harter(1960), Lowe(1960), Groenewoud et al.(1967) and Govindarajulu(1983). For further problems involving probability calculations related to targeting problems see, for example, Solomon(1953) and Guenther and Terragno(1964).

The problem we are concerned with here is to estimate $P_{(\mu,\Sigma)}(C_k)$ based on a sample $X=(x_1,\ldots,x_n)$ from the $N_2(\mu,\Sigma)$ distribution where $\mu\in R^2$ and $\Sigma\in R^{2\times 2}$ positive definite are unknown. The relevant invariant group for this problem is O(2), the group of orthogonal transformations on R^2. In particular we discuss algorithms for the evaluation of estimators of the form (2), (3) and (4) which satisfy the invariance requirement. Further we present some Monte Carlo results which give some indication as to the behaviour of the estimates in repeated sampling and as such present additional information for the investigator who might be faced with choosing amongst them.

The estimators appropriate to the situations when (μ,Σ) is restricted; e.g. requiring that $\Sigma=\sigma^2 I$ and $\sigma>0$ unknown, can typically be obtained by making obvious adjustments to our algorithms for the most general case. The computer programs for the evaluation of the estimators and the simulation were written in Fortran 77 and are available from the authors. All the calculations discussed in this paper were carried out on the PDP 11-70 in the Department of Statistics, University of Toronto.

The dual of the problem addressed here is to specify $p_0=P_{(\mu,\Sigma)}(C_k)$ and then based on the data X estimate k. When $p_0=.5$ the value of k is referred to as the circular probable error. This problem is discussed in Blischke and Halpin(1966).

3. The Standard Estimator

By the standard estimator we mean

$$Q_S^X(C_k)=P_{(\bar{x},(n-1)^{-1}S_X)}(C_k) \tag{5}$$

where $\bar{\mathbf{x}}=n^{-1}\sum_{i=1}^{n}\mathbf{x}_i$ and $S_X=(X-\bar{\mathbf{x}}1')(X-\bar{\mathbf{x}}1')'$; i.e. we have replaced μ and Σ by their UMVU estimators. This estimate is clearly invariant under O(2). The tabulations mentioned earlier are available for the calculation of (5). This approach, however, requires interpolation and is not appropriate for extensive Monte Carlo work. We discuss two approaches to the computer evaluation of this estimator.

First we write $S_X=S_\Delta S_\Delta'$ where $S_\Delta=(s_{ij})$ is the unique lower triangular matrix with positive diagonal elements satisfying this equation. Then if $\mathbf{z}\sim N_2(0,I)$ we can write (5) as

$$P_{(0,I)}(\ ||\bar{\mathbf{x}}+(n-1)^{-1/2}S_\Delta\mathbf{z}|| \le k\ ) = P_{(0,I)}(\ \sum_{i=1}^{2} d_i^2(z_i+b_i)^2 \le (n-1)k^2\ ) \tag{6}$$

where $\mathbf{b}=(n-1)^{1/2}QS_\Delta^{-1}\bar{\mathbf{x}}$ and QDQ is the spectral decomposition of $S_\Delta'S_\Delta$.

An algorithm for calculating (6) is obtained by using the fact that z_1, $z_2\sim N(0,1)$, statistically independent and thus we can write (6) as

$$\int_{-b_2-d_2^{-1}(n-1)^{1/2}k}^{-b_2+d_2^{-1}(n-1)^{1/2}k} [\Phi(u(\cdot)) - \Phi(l(z))]\phi(z)\,dz \tag{7}$$

where Φ is the distribution function for the N(0,1), ϕ is the density,

and u(z), l(z) equal

$$-b_1 \pm d_1^{-1}[(n-1)k^2 - d_2^2(z_2+b_2)^2]^{1/2} \tag{8}$$

respectively. An efficient algorithm is then obtained by using a packaged routine for the evaluation of Φ, e.g. IMSL, and performing the integration using a Gauss-Legendre rule.

A disadvantage of the above approach arises when we are interested in the higher dimensional analogs of this problem as the computation becomes progressively more complicated. A more efficient approach, and it is the one we have adopted, is based on an adaptation of an algorithm due to Sheil and O'Muircheartaigh(1977) which is in turn based on results due to Ruben(1962) concerning the evaluation of the distribution of (z+b)$'D$(z+b) where $z \sim N_p(0,I)$, b$\in R^p$ and $D \in R^{p \times p}$ is diagonal with nonnegative diagonal elements. The result gives a series representation for the distribution function of this quantity and thus also a series representation for (6). We controlled the accuracy in our calculation by stopping the summation when the contribution of the remaining terms was less than 10^{-11}. For p= 2, n= 20 an evaluation of the estimate takes approximately .1 seconds of CPU time.

4. The UMVU Estimator

We have that ($\bar{x}$, S_X) is a complete minimal sufficient statistic and thus the UMVU estimator is given by

$$Q_U^X(C_k) = E[Q_E^X(C_k) : \bar{x},\ S_X] = E[I_{C_k}(x_1) : \bar{x},\ S_X]. \tag{9}$$

This expectation can be evaluated using a result due to Laurent(1957); namely the conditional distribution of x_1 given ($\bar{x}$, S_X) has density proportional to

$$[1-(1-1/n)^{-1}(x_1-\bar{x})'S_X^{-1}(x_1-\bar{x})]^{(n-5)/2} \tag{10}$$

where x_1 is constrained so that the term in the brackets is positive. If we transform $X \to QX$ where Q$\in$O(2) we see that (10) is unchanged and thus this estimator is invariant under O(2).

Laurent proposed that a calculation such as (9) could be carried out by a double numerical integration. In fact we can simplify the calculation of (9) substantially. First we make the transformation $x_1 \to t$ where $t = \sqrt{1-1/n}\ S_\Delta(x_1 - \bar{x})$. The density of t is then proportional to

$$[1 - t't]^{(n-5)/2}. \tag{11}$$

Then in a particular quadrant of R^2 we make the transformation $t \to v$ where $v_i^{1/2} = d_i t_i$, $d_i = sgn(t_i)$. This transformation has Jacobian $(1/2)^2 v_1^{-1/2} v_2^{-1/2}$ and thus the density of v, conditional on a quadrant, is proportional to

$$v_1^{-1/2} v_2^{-1/2} (1 - v_1 - v_2)^{(n-5)/2}. \tag{12}$$

From this we conclude that $(v_1, v_2, 1 - v_1 - v_2)$ is distributed Dirichlet$(1/2, 1/2, (n-3)/2)$.

The estimate can then be expressed as

$$\frac{1}{4} \sum_{d_1, d_2 = \pm 1} P\{(\bar{x}_1 + \sqrt{1-1/n}\ s_{11} d_1 v_1^{1/2})^2 +$$

$$(\bar{x}_2 + \sqrt{1-1/n}\ s_{21} d_1 v_1^{1/2} + \sqrt{1-1/n}\ s_{22} d_2 v_2^{1/2})^2 \leq k^2\} \tag{13}$$

where $S_\Delta = (s_{ij})$ and P refers to the distribution of (v_1, v_2). We now make the transformation $(v_1, v_2) \to (u, v)$ where $u = v_1/(1 - v_2)$ and $v = v_2$. This transformation has Jacobian 1-v and thus the joint density of (u,v) is proportional to

$$u^{-1/2}(1-u)^{(n-5)/2} v^{-1/2} (1-v)^{(n-4)/2}. \tag{14}$$

Therefore $u \sim$ Beta$(1/2, (n-3)/2)$ is statistically independent of $v \sim$ Beta$(1/2, (n-2)/2)$. Then denoting the Beta(p,q) distribution and density functions by $B(\cdot \mid p,q)$ and $b(\cdot \mid p,q)$ respectively (13) can be written as

$$\frac{1}{4} \sum_{d_1, d_2 = \pm 1} \int_{s_1}^{s_2} [B(u_2(v) \mid 1/2, (n-3)/2) - B(u_1(v) \mid 1/2, (n-3)/2)]\, b(v \mid 1/2, (n-2)/2)\, dv \tag{15}$$

where putting r_2, r_1 equal to

$$[d_2(\bar{x}_1 s_{21} - \bar{x}_2 s_{11}) \pm (s_{11}^2 + s_{21}^2)^{1/2} k] / \sqrt{1-1/n}\ s_{11} s_{22} \tag{16}$$

respectively, we have that

$$v_i = [min(max(r_i, 0), 1)]^2 \tag{17}$$

and putting s_2, s_1 equal to

$$[-b(v)\pm(b^2(v)-4ac(v))^{1/2}]/2a\sqrt{1-v} \tag{18}$$

respectively where

$$a=(1-1/n)(s_{11}^2+s_{21}^2)$$

$$b(v)=2\sqrt{1-1/n}\ d_1(\bar{x}_1 s_{11}+\bar{x}_2 s_{21}+\sqrt{1-1/n}\ s_{21}s_{22}d_2 v^{1/2})$$

$$c(v)=(\bar{x}_2+\sqrt{1-1/n}\ s_{22}d_2 v^{1/2})^2+(\bar{x}_1^2-k^2) \tag{19}$$

then we have that

$$v_i(v)=[min(max(s_i,0),1)]^2. \tag{20}$$

To evaluate (15) we use an IMSL subroutine for $B(\cdot\ |1/2,(n-3)/2)$ and then use Gauss-Legendre integration. The efficiency of the integration is improved by using an IMSL routine for the inverse of the Beta$(1/2,(n-2)/2)$ distribution function to find the point vsup such that $B(\text{vsup}\,|1/2,(n-2)/2)=.999999$ and then using min(v_2 ,vsup) as the upper limit in the integration. This ensures that the Gauss points are concentrated where the probability lies. We further improved the efficiency of the integration by making the transformation $v\rightarrow w$ where $w=v^{1/2}$ so that w has density proportional to $(1-w^2)^{(n-4)/2}$. This transformation removes the singularity at 0 which $b(\cdot\ |1/2,(n-2)/2)$ possesses.

The accuracy of the calculation is controlled by dividing the interval of integration into subintervals of equal length and using a Gauss rule of the same order in each. The program allows the order of the rule to vary from 1 to 20 and for as many subdivisons as desired. Thus arbitrary accuracy can be achieved with the tradeoff being computing time. For n= 20, using 4 subdivisions and a Gauss(10) rule, stability was achieved in the fifth decimal place and took about 1.5 seconds of CPU time.

5. Bayes Estimate

There are of course many different Bayes estimates as there are many different choices for the prior distribution of θ and for the loss function. Here we will use mean-square error and choose the prior distribution for (μ,Σ) to be Jeffrey's prior; see for example Box and Tiao(1973).

An alternative approach to this problem, which also leads to the Bayes estimate associated with Jeffrey's prior, is to use the structural model for the multivariate normal model based on the affine group; i.e. the group $G=\{[a,C]: a\in R^2,\ C\in R^{2\times 2},\ det(C)\neq 0\}$, $[a_1,C_1][a_2,C_2]=[a_1+C_1a_2,C_1C_2]$ and we represent $x\sim N_2(\mu,\Sigma)$ as $x=[\mu,\Gamma]z=\mu+\Gamma z$ where Γ satisfies $\Sigma=\Gamma\Gamma'$ and $z\sim N_2(0,I)$. This approach is discussed in Fraser(1979) where the positive affine group is used; i.e. we also require $\det(C)>0$. Use of the full affine group requires only minor adjustment to the analysis presented in Fraser(1979) and it provides a convenient framework for obtaining our results.

The structural model leads to the following relations

$$\begin{aligned}\mu &= \bar{x}-\Gamma\,\bar{z}\\ \Gamma &= C_X C^{-1}\end{aligned} \tag{21}$$

where $\Sigma=\Gamma\Gamma'$, $\Gamma\in R^{2\times 2}$ with $det(\Gamma)\neq 0$, $S_X=C_XC_X'$ with $C_X\in R^{2\times 2}$ calculated as described in Fraser(1979), $\bar{z}\sim N_2(0,n^{-1}I)$ statistically independent of C which has density as described in Fraser(1979) with the adjustment that the density is multiplied by 2^{-3} and is now a function on $\{C : C\in R^{2\times 2},\ det(C)\neq 0\}$.

As is well-known, the Bayes estimate with respect to mean-square error, is given by the mean-value, assuming it exists, of the marginal posterior distribution of the quantity to be estimated. Thus we wish to calculate

$$\int P_{(\mu,\Sigma)}(C_k)\,dP_X(\mu,\Sigma) \tag{22}$$

Denoting the joint posterior distribution of (μ,Γ) by P_X^* and using $\Sigma=\Gamma\Gamma'$ we can write (22) as

$$\int P_{(0,I)}(\,||\mu+\Gamma z||\leq k\,)\,dP_X^*(\mu,\Gamma) \tag{23}$$

This expectation can be evaluated by using the relations (21) and the joint distribution P^* for $(\bar{z},C)$ to obtain

$$\int P_{(0,I)}(\,||\bar{x}+\sqrt{1+1/n}\;\;C_X t||\leq k\,)\,dP^*(\bar{z},C) \tag{24}$$

where $t=\sqrt{1+1/n}\;\;C^{-1}(z-\bar{z})$. From Fraser(1979) we have that t has density function given by

$$A_{n-2}\,A_n^{-1}(1+t_1^2+t_2^2)^{-n/2} \tag{25}$$

where $A_f=2\pi^{f/2}/\Gamma(f/2)$; i.e. t has a canonical bivariate Student (n-1) distribution. Using the

Gram-Schmidt decomposition on the rows of C_X we obtain $C_X = S_\Delta Q$ where S_Δ is as before and $Q \in O(2)$. We note that the distribution of t is invariant under orthogonal transformations. Then by the theorem of total probability (24) can be written as

$$P((\bar{x}_1+\sqrt{1+1/n}\; s_{11}t_1)^2+(\bar{x}_2+\sqrt{1+1/n}\; s_{12}t_1+\sqrt{1+1/n}\; s_{22}t_2)^2 \leq k^2) \tag{26}$$

where P refers to the distribution of t.

We now make the transformation $(t_1,t_2)\rightarrow(u,v)$ where $u=\sqrt{n-1}\; t_1/\sqrt{1+t_2^2}$ and $v=\sqrt{n-2}\; t_2$. The density of (u,v) is proportional to

$$(1+\frac{u^2}{n-1})^{-n/2}(1+\frac{v^2}{n-2})^{-(n-1)/2} \tag{27}$$

and thus u∼Student(n-1) statistically independent of v∼Student(n-2). Therefore, denoting the Student(λ) distribution and density functions by $G(\cdot|\lambda)$, $g(\cdot|\lambda)$ respectively, we have that (27) can be written as

$$\int_{v_1}^{v_2} [G(u_2(v)\,|n-1)-G(u_1(v)\,|n-1)]\,g(v\,|n-2)\,dv \tag{28}$$

where v_2, v_1 equal

$$\sqrt{n-2}\;[\bar{x}_1 s_{21}-\bar{x}_2 s_{11} \pm k\sqrt{s_{11}^2+s_{21}^2}\;]\;/\sqrt{1+1/n}\; s_{11}s_{22} \tag{29}$$

respectively, $u_2(v), u_1(v)$ equal

$$\sqrt{n-1}\;[1+v^2/(n-2)]^{-1/2}[-b_1(v)\pm(b_1^2(v)-4a_1c_1(v))^{1/2}]/2a_1 \tag{30}$$

respectively and

$$a_1=(1+1/n)(s_{11}^2+s_{21}^2)$$

$$b_1(v)=2\sqrt{1+1/n}\;[(\bar{x}_1 s_{11}+\bar{x}_2 2\sqrt{1+1/n}\;[(\bar{x}_1 s_{11}+\bar{x}_2 s_{21})+\sqrt{1+1/n}\; s_{21}s_{22}v/\sqrt{n-2}\;]$$

$$c_1(v)=(\bar{x}_2+\sqrt{1+1/n}\; s_{22}v/\sqrt{n-2}\;)^2+\bar{x}_1^2-k^2 \tag{31}$$

The calculation is then carried out using an IMSL subroutine for the Student distribution function and Gauss-Legendre rules for the integration. The accuracy was controlled by subdividing the interval (v_1,v_2) into subintervals of equal length, carrying out the numerical integration within each subinterval using a Gauss rule of the same order and controlling the number of points in the rule. If $v_1<0<v_2$ then each of the intervals $(v_1,0)$ and $(0,v_2)$ was subdivided into the same number of subintervals of equal length. This is to ensure that the

mode 0 of the Student density serves as an endpoint for the integration as this improves the efficiency of the calculation. A further improvement was made by requiring that $|v_i|$ be no greater than $t_{.000001}(n-2)$ and this point was obtained from an IMSL subroutine for the inverse of a Student distribution function. For n= 20, using 4 subdivisions and a Gauss(10) rule, stability was obtained in the fifth decimal place. This calculation took approximately .8 seconds of CPU time.

As discussed above, the estimator (28) is obtained using Jeffrey's prior and the result is also obtained from the structural model using the affine group. In effect Jeffrey's prior results as the marginalization of the right Haar prior on this group. As is well-known other groups can be used to parametrize the multivariate normal model and their right Haar priors give rise to different priors for the full parameter (μ,Σ). For example, the affine lower triangular group leads to the estimate

$$P(\,||\bar{\mathbf{x}}+\sqrt{1+1/n}\;S_\Delta\mathbf{t}||\leq k\,) \tag{32}$$

where P refers to the distribution of $\mathbf{t}$ which has density proportional to

$$(1+t_1^2)^{-1}(1+t_1^2+t_2^2)^{-(n-1)/2} \tag{33}$$

by results in Fraser(1979). Thus we see we will obtain a different estimate of a form similar to (28) in this case.

The invariance considerations lead, however, to the choice of the estimator (28). For if we transform X to QX where $Q\in O(2)$ then $(\bar{\mathbf{x}},C_X)$ transforms to $(Q\bar{\mathbf{x}},QC_X)$ and from (24) we see that the estimator is invariant under O(2). This invariance property does not hold, however, for the estimator based on the affine lower triangular group. For example, if $S_X=$ I then S_Δ=I, $QS_XQ'=$ I and (33) is not invariant with respect to O(2) which proves the non-invariance of the estimate in this case.

6. The Monte Carlo Study

To study analytically the repeated sampling behaviour of the estimates we have discussed presents a difficult problem. Accordingly a simulation study was carried out to see how effective the estimators are and to assess their relative merits.

The performance of the estimators was considered for four sets of parameter values

(i) $\mu=0$, $\Sigma=I$

(ii) $\mu=1$, $\Sigma=I$

(iii) $\mu=0$, $\Sigma=.511'+.5I$

(iv) $\mu=1$, $\Sigma=.511'+.5I$.

For each parameter set we calculated k such that $P_{(\mu,\Sigma)}(C_k)=.5$ using the algorithm of Sheil and O'Muircheartaigh(1977). Our estimates were then always of the true value .5. For each parameter set we considered the estimation problem for sample sizes n= 10 and n= 20.

For a given parameter (μ,Σ) and sample size n we do not need to generate the full sample $X=(x_1, \ldots, x_n)$ from $N_2(\mu,\Sigma)$ to calculate the estimates. For we need only generate $(\bar{x},S_X)$ where $\bar{x}\sim N_2(\mu,n^{-1}\Sigma)$ statistically independent of $S_X\sim W_2(\Sigma,n-1)$. To do this we used the following relations

$$\bar{x}=\mu+n^{-1/2}\Sigma_\Delta z$$
$$S_X=\Sigma_\Delta S_\Delta S_\Delta{}'\Sigma_\Delta{}' \qquad (34)$$

where $z\sim N_2(0,I)$ statistically independent of $S_\Delta=(s_{ij})$ where $s_{11}^2\sim$Chi-square(n-1), $s_{22}^2\sim$Chi-square(n-2), $s_{21}\sim$N(0,1), $s_{12}=0$ and s_{11}, s_{22}, s_{21} are statistically independent. The N(0,1) variables were generated using the Box-Muller method; namely $z=(-2log(u_1))^{1/2}\cos(2\pi u_2)$ where u_1, u_2 are statistically independent and distributed U(0,1). The chi-square variables were generated using a method due to Cheng and Feast(1979) and in fact we used the program RGKM3 as it is listed in Bratley, Fox and Schrage(1983).

The uniform random variates needed for the generation of the normals and chi-squares were obtained using the routine due to Schrage(1979). To decrease possible effects due to not-quite-randomness we first filled a table with uniform values . Then each time a value was required we generated a random address ,the contents of which becomes the generated value, and replaced it in the table by a newly generated value.

For any given $(\bar{x},S_X)$, generated as above, we computed all three estimators; i.e we used common random numbers. Accordingly the estimators share equally in any effects due to deficiencies in the generators. Further, as we will see, this technique substantially improved the efficiency of the Monte Carlo study.

For a given sample size and parameter set we generated nrep values of $(\bar{x}, S_X)$ where nrep depended on the sample size n. Then for each estimator, denoted generically for convenience by x, we estimated

$$Av(x) = E(x)$$

$$MSE(x) = E(x - .5)^2 \tag{35}$$

by

$$\hat{A}v(x) = \frac{1}{nrep}\sum_{i=1}^{nrep} x_i$$

$$\hat{M}SE(x) = \frac{1}{nrep}\sum_{i=1}^{nrep}(x_i - .5)^2 \tag{36}$$

respectively. The standard errors of these estimates are given by

$$SD(\hat{A}v(x)) = [E(x^2) - (E(x))^2]^{1/2}/\sqrt{nrep}$$

$$SD(\hat{M}SE(x)) = [E(x-.5)^4 - (E(x-.5)^2)^2]^{1/2}/\sqrt{nrep} \tag{37}$$

respectively and they in turn are estimated by

$$\hat{S}D(\hat{A}v(x)) = [\sum_{i=1}^{nrep}(x_i - \hat{A}v(x))^2/nrep(nrep-1)]^{1/2}$$

$$\hat{S}D(\hat{M}SE(x) = [\sum_{i=1}^{nrep}((x_i - .5)^2 - \hat{M}SE(x))^2/nrep(nrep-1)]^{1/2} \tag{38}$$

respectively. We then tested the null hypothesis that x is unbiased for .5 via a z-test using the statistic

$$z = (\hat{A}v(x) - .5)/\hat{S}D(\hat{A}v(x)) \tag{39}$$

i.e we compare this value with the N(0,1) distribution by computing the observed level of significance $P(|Z| > |z|)$ where $Z \sim N(0,1)$.

The primary purpose of the study was to compare MSE(x) with MSE(y) for estimators x and y. When x is the nonparametric estimator we have that MSE(x) = .25/nrep and the test statistic takes the form

$$z=(\hat{MSE}(y)-MSE(x))/\hat{SD}(\hat{MSE}(y)) \quad (40)$$

When x and y are two of the estimators we discussed in the preceding sections the test statistic takes the form

$$z=(\hat{MSE}(y)-\hat{MSE}(x))/\hat{SD}(\hat{MSE}(y)-\hat{MSE}(x)) \quad (41)$$

where

$$\hat{SD}(\hat{MSE}(y)-\hat{MSE}(x))=[\hat{SD}^2(\hat{MSE}(y))+\hat{SD}^2(\hat{MSE}(x))-2\hat{Cov}(\hat{MSE}(y),\hat{MSE}(x))]^{1/2} \quad (42)$$

is the estimate of

$$SD(\hat{MSE}(y)-\hat{MSE}(x))=[(Var(y-.5)^2+Var(x-.5)^2-2Cov((y-.5)^2,(x-.5)^2))/nrep]^{1/2} \quad (43)$$

and where

$$\hat{Cov}(\hat{MSE}(y),\hat{MSE}(x))=\sum_{i=1}^{nrep}((y_i-.5)^2(x_i-.5)^2-\hat{MSE}(y)\hat{MSE}(x))/nrep(nrep-1). \quad (44)$$

The covariance term is required in (43) because we have used common random numbers. From this we see where the gain in efficiency was obtained as in all cases $\hat{MSE}(x)$ and $\hat{MSE}(Y)$ were positively correlated and this reduced $SD(\hat{MSE}(y)-\hat{MSE}(x))$ substantially. The estimated correlation between these two quantities ranged from .460 to .996 and in most cases was greater than .700.

The number of replications for each sample size and parameter set was determined by first performing a trial run of 100 for all cases and calculating the estimates we have just described. The primary determinant of sampling variability turned out to be the sample size. We then estimated an upper bound for $SD(\hat{MSE}(x))$ for all estimators over all parameter sets within a sample size. On the basis of this information we chose nrep so that when n= 10 the half-length of a .95-confidence interval for MSE(x) would be less than .001 and when n= 20 so that a .95-confidence interval would have half-length less than .00025. With some extra margin for safety this lead to choosing nrep= 15000 when n= 10 and nrep= 10000 when n= 20. The results we obtained tend to confirm our expectations. These choices gave conclusive results for the comparisons amongst the MSE(x) because of the use of common random numbers. Further these values of nrep gave that the half-length of a .995-confidence interval for Av(x) is less than .005.

For each estimator the accuracy of the calculation was controlled so that the error was less than 5×10^{-6} in each evaluation of the estimate; i.e. if $\hat{x}$ denotes the computed value of the estimate and x^* denotes the actual value of the estimate then

$$|\hat{x}-x^*|\leq 5\times10^{-6} \tag{45}$$

Thus the absolute error in $\hat{Av}(x)$ is less than 5×10^{-6} and since

$$|(x^*-.5)^2-(\hat{x}-.5)^2|\leq |(x^*)^2-\hat{x})^2|+2|x^*-\hat{x}|=|x^*-\hat{x}||x^*+\hat{x}|+2|x^*-\hat{x}|\leq 4|x^*-\hat{x}| \tag{46}$$

we have that the absolute error in $\hat{MSE}(x)$ is less than 2×10^{-5}. All calculations were done in double precision.

Table 1

Sample size= 10

Parameters	Estimator	$\hat{A}v(x)$	$\hat{S}D(\hat{A}v(x))$	α_A	$\hat{M}SE(x)$	$\hat{S}D(\hat{M}SE(x))$	α_M
(i)	Laurent	0.501630	0.00094879	0.085	0.013505	0.00016076	0.0
	Standard	0.507941	0.00088234	0.0	0.011740	0.00014092	0.0
	Bayes	0.422713	0.00077622	0.0	0.015011	0.00013103	0.0
(ii)	Laurent	0.499720	0.00102002	0.787	0.015606	0.00018432	0.0
	Standard	0.496583	0.00101844	0.001	0.015569	0.00018073	0.0
	Bayes	0.443320	0.00087711	0.0	0.014752	0.00015769	0.0
(iii)	Laurent	0.501258	0.00092129	0.174	0.127320	0.00015824	0.0
	Standard	0.505730	0.00086404	0.0	0.011231	0.00013910	0.0
	Bayes	0.422180	0.00076096	0.0	0.014879	0.00012915	0.0
(iv)	Laurent	0.499396	0.00100313	0.549	0.015093	0.00018800	0.0
	Standard	0.491079	0.00100830	0.0	0.015329	0.00018525	0.0
	Bayes	0.434217	0.00088902	0.0	0.016182	0.00017134	0.0

Table 2

Sample size= 10

Parameters	Comparison	z	$\hat{S}D(\hat{M}SE(y)-\hat{M}SE(z))$
(i)	Laurent vs Standard	74.6	0.00002363
	Laurent vs Bayes	-10.2	0.00014714
	Standard vs Bayes	-23.2	0.00014068
(ii)	Laurent vs Standard	23.2	0.00001572
	Laurent vs Bayes	73.5	0.00011612
	Standard vs Bayes	77.6	0.00010525
(iii)	Laurent vs Standard	66.5	0.00002257
	Laurent vs Bayes	-14.7	0.00014510
	Standard vs Bayes	-26.5	0.00013760
(iv)	Laurent vs Standard	11.4	0.00002064
	Laurent vs Bayes	-8.33	0.00013064
	Standard vs Bayes	-7.47	0.00011426

Table 3

Sample size = 20

Parameters	Estimator	$\hat{A}v(x)$	$\hat{S}D(\hat{A}v(x))$	α_A	$\hat{M}SE(x)$	$\hat{S}D(\hat{M}SE(x))$	α_M
(i)	Laurent	0.500797	0.00079951	0.682	0.006392	0.00009160	0.0
	Standard	0.504993	0.00077234	0.0	0.005989	0.00008667	0.0
	Bayes	0.460413	0.00072084	0.0	0.006763	0.00008111	0.0
(ii)	Laurent	0.500332	0.00087090	0.529	0.007584	0.00010944	0.0
	Standard	0.498784	0.00087101	0.161	0.007587	0.00010904	0.0
	Bayes	0.473130	0.00080648	0.0	0.007225	0.00009973	0.0
(iii)	Laurent	0.500907	0.00077780	0.246	0.006050	0.00009016	0.0
	Standard	0.503960	0.00075381	0.0	0.005697	0.00008533	0.0
	Bayes	0.460345	0.00070728	0.0	0.006575	0.00007921	0.0
(iv)	Laurent	0.500133	0.00085500	0.89	0.007309	0.00010730	0.0
	Standard	0.495674	0.00085777	0.0	0.007376	0.00010738	0.0
	Bayes	0.465767	0.00081321	0.0	0.007784	0.00010672	0.0

Table 4

Sample size= 20

Parameters	Comparison	z	$\hat{S}D(\hat{M}SE(y) - \hat{M}SE(x))$
(i)	Laurent vs Standard	52.6	0.00000765
	Laurent vs Bayes	-5.73	0.00006470
	Standard vs Bayes	-11.55	0.00006607
(ii)	Laurent vs Standard	-0.74	0.00000448
	Laurent vs Bayes	7.44	0.00004820
	Standard vs Bayes	8.06	0.00004490
(iii)	Laurent vs Standard	53.48	0.00000659
	Laurent vs Bayes	-8.22	0.00006380
	Standard vs Bayes	-13.56	0.00006470
(iv)	Laurent vs Standard	-8.32	0.00000800
	Laurent vs Bayes	-7.93	0.00005991
	Standard vs Bayes	-7.70	0.00005308

In every case it turns out that we have no evidence against the hypothesis that the Laurent estimator is unbiased for .5 and this is as theory predicts. We see that in every case except n= 20,(ii) we reject the hypothesis that the Standard estimator is unbiased for .5. We note, however, that the bias in this estimator is quite small in every case with the largest estimate of the bias being about .009 and the bias decreases as n increases. In every case we reject the hypothesis that the Bayes estimate is unbiased with the smallest estimate of its bias being about .027. The bias decreases as n increases and can be severe for small sample sizes.

In every case we reject the hypothesis that the mean-square error of the estimator included in the study was equal to that of the nonparametric estimator. The Laurent, Standard, and Bayes estimators would all appear to be substantial improvements over the nonparametric estimator. The smallest relative efficiency , as measured by the ratio of the mean-square errors, of an estimator to the nonparametric estimator was 154%.

We now compare the mean-square errors of the estimators included in the study. We note that in every case except for n= 20, (ii) Laurent versus Standard, we categorically reject the hypothesis that the mean-square errors are equal. For (i), in both sample sizes, we have that the Standard estimator is superior to Laurent's which is in turn superior to the Bayes estimator. For (ii), the Bayes estimator is superior to the other two while the Standard is superior to Laurent's when n= 10 and they are equivalent when n= 20. For (iii), we have the same ranking as in (i). For (iv), Laurent's estimator was best followed by the Standard which in turn was better than the Bayes estimator and this applied for both sample sizes.

We see from the above discussion that no estimator can be categorically accepted or rejected as the best or worst in the circumstances we considered. On the other hand, when taking account of both bias and mean-square error, it would seem that the Standard estimator would be the most practical choice. In fact the lowest relative efficiency of the Standard estimator to the best estimator, when it was not best, was about 94%. The lowest relative efficiency of Laurent's estimator to the best was about 87% and the corresponding value for the Bayes estimator was about 75%. Perhaps most surprising in our results was the good performance of the Standard estimator relative to Laurent's estimator given that the latter possesses an optimality property. A further point in favour of the Standard estimator is given by the fact that a much more efficient algorithm is available for its evaluation than for the other two.

7. Conclusions

This paper has been concerned with the problem of estimating circular error probabilities when we require that the estimator be invariant under the invariant group of the circle. Three competing estimators were proposed and we developed efficient methods for their evaluation. A Monte Carlo study was carried out to provide more information concerning the relative merits of the estimators. On the basis of this study and the relative efficiencies of their algorithms a recommendation can be made that the Standard estimator is perhaps the most practically useful for this problem. In all cases the estimators were substantially better than the nonparametric estimator when we are assuming bivariate normality.

8. Acknowledgements

The authors sincerely thank Professor Herbert Solomon for his advice and encouragement and for drawing to our attention the work of Professor Laurent. The authors also thank the Dept. of Statistics, Stanford University for its hospitality during the 1983/84 academic year.

The work of M. Evans was in part supported by the Natural Sciences and Engineering Council of Canada, Grant No. A3120. The work of Z. Govindarajulu was in part supported by a grant from the Office of Naval Research Grant No. N00014-84-K-D184. The work of J. Barthoulot was in part supported by a grant from the Swiss National Foundation for Scientific Research.

REFERENCES

1. Blischke, W.R. and Halpin, A.H. (1966). Asymptotic properties of some estimators of quantiles of circular error. JASA, 61, 618-632.

2. Box, G. and Tiao, G. (1973). Bayesian Inference in Statistical Analysis. Addison Wesley Publishing Co.

3. Bratley, P., Fox, B.L. and Schrage, L.E. (1983). A Guide to Simulation, Springer Verlag.

4. Cheng, R.C.H. and Feast, G.M. (1979). Some simple gamma variate generators. Appl. Statist., 28, 290-295.

5. Fraser, D.A.S. (1979). Inference and Linear Models. McGraw-Hill International Book Company.

6. Govindarajulu, Z. (1983). A Note on Circular Error Probabilities. University of Kentucky, Dept. of Statistics, Technical Report No. 212.

7. Groenewoud, C. Hoaglin, D.C. and Vitalis, J.A. (1967). Bivariate Normal Offset Circle Probability Tables. Cornell Aeronautical Laboratory, Inc. CAL No. XM-2464-G-1.

8. Guenther, W. and Terragno, P. (1964). A review of the literature on a class of coverage problems. Ann. Math. Statist., 64, 232-260.

9. Harter, H.L. (1960). Circular error probabilities. JASA, 55, 723-731.

10. IMSL Library (1982). Problem-solving software for mathematical and statistical Fortran programming. Revised June, 1982.

11. Laurent, A. (1957). Bombing problems - a statistical approach. Opns. Res., 10, 75-89.

12. Laurent, A. (1962). Bombing problems - a statistical approach II. Opns. Res., 10, 380-387.

13. Lowe, J.R. (1960). A table of the integral of the bivariate normal distribution over a offset circle. JRSS B, 22, 177-187.

14. Muirhead, R. (1982). Aspects of Multivariate Statistical Theory. John Wiley and Sons, Inc.

15. Ruben, H. (1962). Probability content of regions under spherical normal distributions, IV.

The distributions of homogeneous and nonhomogeneous quadratic functions of normal variables. Ann. Math. Statist., 33, 542-570.

16. Schrage, L.E. (1979). A more portable Fortran random number generator. ACM Trans. Math. Software, 5, 132-138.

17. Sheil, J. and O'Muircheartaigh, I. (1977). Algorithm AS 106. The distribution of non-negative quadratic forms in normal variables. Appl. Statist., 26, 92-98.

18. Solomon, H. (1953). Distribution of the measure of a random two-dimensional set. Ann. Math. Statist., 24, 650-656.

19. Solomon, H. and Grad, A. (1955). Distribution of quadratic forms and some applications. Ann. Math. Statist., 26, 464-477.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
367

2. GOVT ACCESSION NO.
AD-A163 257

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
Estimates Of Circular Error Probabilities

5. TYPE OF REPORT & PERIOD COVERED
TECHNICAL REPORT

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
M. Evans, Z. Govindarajulu, and J. Barthoulot

8. CONTRACT OR GRANT NUMBER(s)
N00014-76-C-0475

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Department of Statistics
Stanford University
Stanford, CA 94305

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
NR-042-267

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Statistics & Probability Program Code 411SP

12. REPORT DATE
December 5, 1985

13. NUMBER OF PAGES
25

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
Also supported in part by the Office of Naval Research under Contract No. N00014-84-K-D184 at University of Kentucky.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Circular probable error; bivariate normal; quadratic forms.

sub theta   by R   sub theta

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
The general problem we are concerned with here is the estimation of $P_\theta(C)$ where C is some fixed event and P_θ is unknown in some class. The various available estimation procedures seem to involve the choice of some random probability measure. In particular we consider this problem when C is a disk in $\mathbb{R}^2$ centered at o and P_θ is restricted to be bivariate normal. Details concerning the implementation of the estimation procedures and a Monte Carlo study are discussed for this case. This particular problem arises when we are concerned with assessing the accuracy of a targeting procedure.