

February 20, 2006

Dr W Edwin Sharp, Editor-in-Chief
Mathematical Geology
Department of Geological Sciences
University of South Carolina
Columbia, SC 29208

Subject: *Manuscript Number 05-76*

Dear Dr Sharp,

Your letter of January 11, 2006, made as mesmerizing a read as your thoughts on degrees of freedom when you were *MG*'s Book Review Editor and contacted me on March 16, 1998. I was surprised but pleased when you phoned on March 28, 2005, and expressed your interest in a paper on testing for spatial dependence and charting sampling variograms. I transmitted the first draft on June 13, 2005, took into account most of your commentary and criticism, and submitted three copies of the final paper on October 20, 2005.

In retrospect I regret that I left it up to you to decide whether Bre-X's Busang salting scam, Hecla's Grouse Creek failure or Cominco's Cerratepe bulk sampling program would make the most compelling case to show how to verify spatial dependence by applying Fisher's F-test. I would have selected Bre-X's infamous bogus grades because the first variance term of the ordered set played a crucial role in proving the intrinsic variance of bogus gold in Busang's barren rock to be statistically identical to zero, as it ought to be in a phantom gold resource.

Reviewer A concludes that $df_o=2(n-1)$, the number of degrees of freedom for the first variance term of the ordered set, is '*highly incorrect even if the data are a sample of the Normal distribution*'. So he aborts his review, takes off on a tangent with some simulation study and applies Kolmogorov's goodness-of-fit test to prove some point. If the formula for the first variance term were indeed '*highly incorrect*', it would invalidate the application of Fisher's F-test to verify spatial dependence and chart sampling variograms as described in a number of ISO standards. It was Halden who applied the formula for the first variance term of the ordered set in 1952, but your *very good and very fair-minded reviewer*' dismissed the same formula in 2005. Would your Reviewer A consider Halden's formula '*highly correct*' if the degrees of freedom for the *in situ* ordered gold and silver grades were $df_r=n-1=25$, and the tabulated value $F_{0.05;25;50}=2.17$ at 95% probability were $F_{0.05;25;25}=1.95$?

Your reviewer ought to prove that ISO Technical Committee 69–*Applications of Statistical Methods* approved this ‘*highly incorrect*’ formula, an invalid formula that is not even applicable if a data set were normally distributed. He should submit his findings in the format of a paper for review and publication in *Mathematical Geology*, preferably before your retire as its Editor-in-Chief.

Reviewer B suggests I need a ghostwriter because ‘*(T)he statistical notation and terminology is unclear and outdated*’. For example, he prefers the verb ‘*lag*’ but I consider it an ambiguous measure for length when variances are charted against $var_j(x)$ –symbols in a sampling variogram to derive a distance or length in m. His opinion becomes suspect when scrutinizing his $\hat{\mu}_i^2$ –symbol in Paragraph 7.1 of the review. In ISO Standards, the μ –symbol refers to the unknown true population mean, an elusive and esoteric statistic that plays a role in defining fundamental concepts such as bias, accuracy and trueness. In contrast, the \bar{x} –symbol refers to a central value (arithmetic mean or some weighted average) of a set of measured values with equal or variable weights determined in samples selected at positions with different coordinates in a sample space.

Reviewer B deems ‘*outdated*’ Youden (1951) and Davis (1986). However, Halden’s 1952 *Statistical Theory with Engineering Applications*, in Formula (13.5.10) on page 358, shows that, even though this author studied ratios between low and high variances, his inference agreed with Fisher and Snedecor’s who studied ratios between high and low variances. *MG*’s reviewers, in particular, ought not to dismiss the works of early statisticians because they should have looked beyond symbols and terms and grasped the essence of their thoughts before embracing Matheronian geostatistics.

Even though you reviewed and revised the first draft of my paper, you concurred with Reviewer B that the final paper required revisions on January 11, 2006, but changed your mind again on February 14, 2006. You allege not to know this reviewer but I know that most *MG*’s reviewers deem clear and up-to-date a wide range of puzzling neologisms and a cult like lingo of krige-inspired eponyms, the strangest of which by far is the counterintuitive ‘*zero kriging variance*’.

On March 26, 1998, I brought to your attention that a randomly distributed set of n measured values gives $df_r=n-1$ degrees of freedom, and that the ordered set gives $df_o=2(n-1)$. On April 8, 1998, you wrote, ‘*I do not dispute that the divisor for the regular case is $(n-1)$ and for the ordered case it is $2(n-1)$. However the word divisor is not synonymous [sic] with degrees of freedom*’. And you pointed out that I am, ‘*making the same mistake as the geostatisticians in having more degrees of freedom than you have observations*’, and that your factor corrects for this mistake. Paradoxically, the geostatisticians’s $2n$ is approximately right but your $n-1$ is exactly wrong!

I wonder if your bias correction factor was inspired by Journel's who pontificated on October 15, 1992, '*(T)he reason for the denominator (n-1) in the classical expression of the variance estimate...is indeed correction of the bias introduced by the prior estimation of the population mean \bar{z} by the same n data z_i* '. On April 29, 1998, I reiterated, '*Von Neumann's "bias correction factor" simply reflects that all but the first and last datum are used twice*'. It is beyond comprehension that *MG's* Editor-in-Chief cannot count degrees of freedom for ordered sets.

Meanwhile, your bias correction factor mutated into a scaling factor of sorts, which, when multiplied with $df_r=n-1$, the number of degrees of freedom for the randomized set, gives $df_o=2(n-1)$. Your hypothesis that this multiplier is a scaling or bias correction factor crumbles under scientific scrutiny. Mathematical proof is posted not only on my own website but also on *ai-geostats.org* under *Documents*.

On March 26, 2003, you announced, '*With respect to geostat, I am probably the best ally you will ever have as I have always had a certain amount of scepticism.*' How often need I tell an ally and a skeptic about geostatistics that ordered sets have $df_o=2(n-1)$ degrees of freedom simply because all data but the first and last are used twice? You were told at least as many times as *IAMG's Board* that it is a scientific fraud to assume spatial dependence without proof, interpolate by kriging, smooth pseudo kriging variances of subsets of infinite sets of kriged estimates to perfection, and rig the rules of mathematical statistics.

On October 10, 2002, I transmitted the names of several statistics scholars who represent the USA in ISO/TC69--*Applications of Statistical Methods*. Because I wrote the paper at your request, you should have had the courtesy to have one of those statisticians review my paper. The relentless vagaries of *MG's* peer review leave me no choice but to rescind my consent to have *Mathematical Geology* publish *Manuscript No 05-76* in whole or in part.

Scientific integrity demands that Reviewer A be requested to study spatial dependence and sampling variograms as defined and accepted by ISO Technical Committee 69, and that *Mathematical Geology* publishes his findings. Reviewer A should also be asked to explain whether or not it makes statistical sense to replace the variance of a distance-weighted average with the pseudo kriging variance of a subset of some infinite set of kriged estimates. The readers of *Mathematical Geology* deserve no less.

Yours truly,

J W Merks