November 23, 1990

The Honorable Kenneth M. Ca.
Chairman
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Dear Chairman Carr:

This is being written as an individual letter, since it deals with
a subject -- statistics - that holds more interest for me than it
does for the other members of ACRS, or indeed for the NRC staff.
I see no good reason why the NRC shouldn't do its statistics
correctly.

Since my "additional comment" on the ACRS diesel generator letter
of August 14, 1990 has been treated with benign neglect by the
staff, and has had no visible effect on the draft Regulatory Guide
1.9. Rev. 3, it may be useful to spell out the statistical point
a bit more expansively.

The matter is not trivial, since the purpose of the entire exercise
is to establish a level of reliability for a licensee's diesel
generators (either 0.95 or 0.975), and to provide a mechanism
through which the licensee can demonstrate maintenance of that
level.  In our meeting of November 8, 1990, Commissioner Curtiss
emphasized this point. The trigger value method given in the draft
Regulatory Guide provides no such demonstration.  Any statistician
could have made this point to the staff, and I am embarrassed to
have to do so.

I will use only one of the cases cited in the draft Regulatory
Guide, the trigger value of three failures in twenty attempts, for
a presumed reliability level of 0.95.   Presumably, failure to
trigger is to be construed as the necessary demonstration that the
desired reliability level of 0.95 is being maintained (otherwise,
why does the failure trigger any action?).   I will give these
numbers a standard elementary statistical analysis.

First notice that the statement "three or more failures in twenty
attempts" is ambiguous. There is a considerable difference between
the case in which one studies twenty attempts to see how many
failures have occurred and that in which one waits to count three
failures, and counts how many attempts it took to get there.  They
have different probabilities.  In this case, since the event is to
be used as a trigger for action, I will assume that the latter is
what was meant.

In that case, the precise statement of the situation is that at each failure the licensee should check how many have occurred in the previous nineteen attempts. If the number is zero or one, that will be construed as demonstration that the reliability objective of 0.95 is being maintained. If there are two or more the demonstration has failed. (We could have the same discussion about any of the other thresholds - this is just one case. In addition, the term "two or more" is redundant - if there are more, then the trigger has already been pressed.)
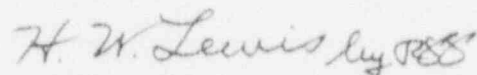
The fact that there has been zero or one failure to start in the previous nineteen attempts provides information that can be subjected to a standard and classical statistical analysis, assuming that the failures are independent. Such an analysis, which is trivial and can be performed by any statistician, leads to the conclusion that the 90% confidence bounds on the underlying diesel reliability are 0.774 and 0.961. Only such information can be drawn from those data. That falls rather short of the desired demonstration that the reliability is greater than 0.95. (If none of the three triggers - 3/20, 5/50, 8/100 - are pressed, the 90% confidence bounds on the underlying reliability become 0.892 and 0.982, again hardly a demonstration that the reliability has been maintained above 0.95, as required.)

I have no problem with an aspirational statement about diesel reliability, and none with arbitrary trigger levels for attention to the diesels. But these are not connected in the way stated in the draft Regulatory Guide, and it is simply error to say that they are.

Finally, I would think it useful for the NRC staff to include a few statisticians, and for the engineers to take statistics seriously. There is no sin in doing mathematics correctly. In fact, it would be easy to construct a monitoring and evaluation program that provides a running estimate of each diesel generator's underlying reliability, and of its uncertainty, and provides a means of detecting trends. Any statistician could devise such a program, and it could provide a sound basis for regulatory attention.

Additional comments by ACRS Member J. Ernest Wilkins, Jr., are presented below.

Sincerely,

H. W. Lewis by RSS

Harold W. Lewis, Member
Advisory Committee on Reactor
Safeguards

Additional Comments by ACRS Member J. Ernest Wilkins, Jr.

I concur with Professor Lewis' concerns in this specific case, as well as with his general observations about doing "statistics correctly."   I believe, however, that the NRC staff already includes at least one competent statistician.   The problem, therefore, may be that neither his professional services nor those of any other such individual are utilized for the purposes described in this letter.