Lower Limit of Detection: Definition and Elaboration of a Proposed Position for Radiological Effluent and Environmental Measurements

The concept of Lower Limit of Detection (LLD) is used routinely in the NRC Radiological Effluent Technical Specifications (RETS) for measurement of radiological effluent concentrations within a nuclear power plant and of radiological environmental samples outside of the plant. The definition of LLD is subject to different interpretations by various groups. Consequently, difficulties arose when the NRC attempted to apply uniformly requirements on licensees. At present, NRC relies on documentation on LLDs that has been developed by other agencies for their own purposes. The material is for the most part difficult to obtain, and is only partially relatable to Technical Specifications requirements.

There was clearly a need to evaluate the various concepts and interpretations of LLD presented in the literature and to determine the current use and application of these concepts in practice in Technical Specifications for operating nuclear plants. This would then lead to a NUREG/CR document that could assist the NRC Nuclear Reactor Regulation staff in defining and elaborating its position relative to LLDs, as well as providing a technically sound basic document on detection capability for effluent and environmental monitoring.

Dr. Lloyd A. Currie of the National Bureau of Standards, a nationally recognized expert in statistics, was asked to undertake this task. At the start Dr. Currie performed an extensive literature search in the area of detection limits. He discussed concepts and problems of LLD with a number of individuals from licensed nuclear power plants, from contracting measurement laboratories,

and from NRC Headquarters and Regional Offices. He then integrated these nuclear-power oriented questions and concepts into his extensive experience in low-level measurement to develop a comprehensive document covering the problems of LLD in radiological effluent and environmental measurements.

It should be emphasized that this document represents Dr. Currie's interpretation of the situations he encountered and his recommendations to the NRC staff relative to these problems. It cannot of itself represent NRC policy. It will, however, be used by NRC staff in development of potential modifications in the definitions and bases sections of the model RETS relative to LLD. And of most immediate importance, it will provide a sound basis to licensees and NRC staff alike for use in clarifying thoughts and writings in the area of detection capability of radiological measurement systems.

> Frank J. Congel, Chief
> Radiological Assessment Branch
>
> Charles A. Willis, Leader
> Effluent Treatment Section
>
> NRC Division of Systems Integration

ABSTRACT

A manual is provided to define and illustrate a proposed use of the Lower Limit of Detection (LLD) for Radiological Effluent and Environmental Measurements. The manual contains a review of information regarding LLD practices gained from site visits; a review of the literature and a summary of basic principles underlying the concept of detection in Nuclear and Analytical Chemistry; a detailed presentation of the application of LLD principles to a range of problem categories (simple counting to multinuclide spectroscopy), including derivations, equations, and numerical examples; and a brief examination of related issues such as reference samples, numerical quality control, and instrumental limitations. An appendix contains a summary of notation and terminology, a bibliography, and worked-out examples.

EXECUTIVE SUMMARY

This document defines and illustrates a proposed use of the concept of
Lower Limit of Detection (LLD) for Radiological Effluent and Environmental
Measurements. It contains a review of information regarding LLD practices
gained from nuclear plant site visits, a review of the literature and a
summary of basic principles underlying the concept of detection in Nuclear
and Analytical Chemistry, and a detailed presentation of the application of
LLD principles to a range of problem categories (simple counting to multi-
nuclide spectroscopy), including derivations, equations, and numerical
examples. It also contains a brief examination of related issues such as
reference samples, numerical quality control, and instrumental limitations.
An appendix contains a summary of notation and terminology, a bibliography,
and worked-out examples.

The detection capability of any measurement process (MP) is one of
its most important performance characteristics. When one is concerned with
pressing an MP to its lower limit or with designing an MP to meet an extreme
measurement requirement, an objective measure of this capability is just as
important for characterizing the MP as is the more commonly understood
characteristics "precision" and "accuracy." As with these other characteristics,
the detection capability cannot be specified quantitatively unless the MP is
rigorously defined and in a state of control. In the monitoring environment,
for low levels of effluent and environmental radioactivity associated with
the operation of·nuclear power reactors, MPs must b capable of detecting the
relevant radionuclides at levels well below those of concern to the public
health and safety.

Much confusion surrounds the nomenclature, formulation, and assumptions associated with this important measurement process characteristic. For the purposes of this document the term "Lower Limit of Detection" (LLD) is used to describe the MP characteristic, and the same terminology, with appropriate adjustments for scale and dimensions is applied to amounts of radioactivity, concentrations, release rates, etc. In short, the same notation, LLD, is used as a universal descriptor for all of the MPs in question. The assumptions and mathematical and numerical formulations underlying LLDs are treated explicitly, and the practical usage (and limitations thereof) is illustrated with appropriate numerical examples. In particular, the special opportunities and pitfalls associated with "Poisson counting statistics" are duly noted.

Section I of the report provides an introduction that sets the stage for the technical sections that follow. Considerations that enter into an NRC Technical Position on LLD are recorded, including theoretical background, technical issues, policy issues, and implementation and documentation. High-lights from site visits are next presented, providing perspective on the problems and actual practices regarding LLD from the viewpoints of: the NRC (regional offices and inspectors), a trade association, nuclear utility labo-ratories, the EPA cross-check laboratory, and contracting laboratories.

The primary historical and theoretical background on detection decisions and detection limits is presented in Section II. The lack of and need for uniform practice, which was ascertained during the site visits, is underlined in the historical review of the literature. The basis for the approach to

LLD adopted here, hypothesis testing, is outlined in some detail. This is
followed by an examination of several crucial issues of general concern such
as the role of detection decisions, the meaning of a priori in the case of
interference, the treatment of systematic error, and the calibration func-
tion. The basic concepts are next applied to radioactivity, and to specific
issues related to the blank, counting technique, measurement process design
(to meet the requisite LLD), quality in communication and monitoring (control),
and the increase required in LLD to meet the demands of multiple detection
decisions.

Section III builds on the theory developed in Section II. Basic and
simplified formulations are presented in "stand-alone" form, with sufficient
notes, that they might be adapted for use in Radiological Efluent Technical
Specifications (RETS). The heart of Section III comprises detailed algebraic
reductions of the general equations for a variety of radioactivity measure-
ment situations, ranging from "simple counting" to multicomponent spectroscopy.
The treatment of extreme low-level counting is illustrated, as well as ordinary
Poisson error treatment and systematic error treatment in relation to the LLD.

The Appendix includes a condensed summary of notation, an index to the
tutorial notes in Section III, a more extended literature survey and biblio-
graphy, and worked-out numerical examples.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

A. Introductory Remark[1]

The detection capability of any measurement process (MP) is one of its most important performance characteristics. When one is concerned with pressing an MP to its lower limit or with designing an MP to meet an extreme measurement requirement, an objective measure of this capability is just as important for characterizing the MP as is the more commonly understood characteristics "precision" and "accuracy." As with these other characteristics, the detection capability cannot be specified quantitatively unless the MP is rigorously defined and in a state of control. (Thus, a secondary issue of major importance is the quality control of the measurement procedure.) In the monitoring environment -- in the present case, for low levels of effluent and environmental radioactivity associated with the operation of nuclear power reactors -- MPs must be capable of detecting the relevant radionuclides at levels well below those of concern to the public health and safety. (This need may be contrasted with others where, for example, adequate detection capability may be required to monitor biological conditions, natural hazards, industrial processes and materials properties, international agreements, etc.)

Much confusion surrounds the nomenclature, formulation, and assumptions associated with this important measurement process characteristic. For the purposes of this document, we shall somewhat arbitrarily select the term "Lower Limit of Detection" (LLD) to describe the MP characteristic, and we shall apply the same terminology, with appropriate adjustments for scale and dimensions, to amounts of radioactivity, concentrations, release rates, etc. -- in short, we shall use the same notation, LLD, as a universal descriptor

---

[1] In this report reference numbers are placed in parentheses and special numbered notes (preceded by series letter A or B), in brackets.

for all of the MPs in question. The assumptions and mathematical and numerical formulations underlying LLD's will be treated explicitly, and the practical usage (and limitations thereof) will be illustrated with appropriate numerical examples. In particular, the special opportunities and pitfalls associated with "Poisson counting statistics" will be duly noted.

B. Plan for the Report

The objective and background for an NRC Technical position (following section) sets the stage for this report-manual on LLD. Next, perspective is given on the problems and actual practices from the viewpoints of: the NRC (regional offices and inspectors), a trade association, nuclear utility laboratories, the EPA cross-check laboratory, and contracting laboratories.

The primary historical and theoretical background on detection decisions and detection limits is presented in section II. The lack of and need for uniform practice, which was ascertained during the site visits, is underlined in the historical review of the literature. The basis for the approach to LLD adopted here, hypothesis testing, is outlined in some detail. This is followed by an examination of several crucial issues of general concern such as the role of detection decisions, the meaning of a priori in the case of interference, the treatment of systematic error, and the calibration func-
tion. The basic concepts are next applied to radioactivity, and to specific issues related to the blank, counting technique, measurement process design (to meet the requisite LLD), quality in communication and monitoring (control), and the increase required in LLD to meet the demands of multiple detection decisions.

Section III builds on the theory developed in section II. Basic and simplified formulations are presented in "stand-alone" form, with sufficient notes, that they might be adapted for use in Radiological Effluent Technical Specifications (RETS). (This led to some necessary redundancy with ideas presented in section II.) The heart of section III comprises detailed algebraic reductions of the general equations for a variety of radioactivity measurement situations, ranging from "simple counting" to multicomponent spectroscopy. The treatment of extreme low-level counting is illustrated, as well as ordinary Poisson error treatment and systematic error treatment in relation to the LLD.

The Appendix includes a condensed summary of notation, an index to the tutorial notes in section III, a more extended literature survey and bibliography, and worked-out numerical examples.

C. Considerations for an NRC Technical Position

1. Objective of the NRC Position

Adequate measurement capabilities for effluent and environmental radioactivity are required to assure the safety of the public, as put forth in 10 CFR Parts 20 and 50 which mandate appropriate radiological effluent and environmental monitoring programs. In order to assure adequate detection capability for radionuclides to meet these requirements, the NRC has established numerical levels for Lower Limits of Detection (LLD) which are consistent with a sufficient capacity for detecting effluent and environmental radionuclides well below levels of concern for the public health and safety. For such LLDs to be meaningful and useful, they must (a) be soundly based in terms of measurement science, and (b) they must be accepted, understood, and applied in a uniform manner by the community responsible for

performing and evaluating the respective measurements. These limiting values as LLDs become part of the Operating License of a Nuclear Power Plant through the Radiological Effluent Technical Specifications (RETS) of the operating license.

2. Theoretical Background

A firm basis for evaluating LLDs is given by the statistical theory of hypothesis testing, which recognizes that the issue of detection involves a decision ("detected," "not detected") made on the basis of an experimental observation and an appropriate test statistic. Once the decision algorithm has been defined, one can evaluate the underlying detection capability (LLD) of the measurement process under consideration. Arbitrary rules for defining LLD's which do not have a sound base (such as hypothesis testing) yield LLD's with little meaning and needless incomparability among laboratories. The system for computing and evaluating LLDs to be recommended for effluent and environmental radioactivity measurement processes, is based on exactly the same principles which underlie more commonly used and understood confidence intervals. Key quantities which arise in the approach to LLDs are the probabilities of false positives (α) and false negatives (β) - both generally taken to be 5%.

3. Technical Issues

⊙ The adopted terminology (notation) to reflect the measurement (detection) capability shall be "LLD," and it shall refer to the intrinsic detection capability of the entire measurement process - sampling through data reduction and reporting.

An LLD for simply one stage of the measurement process, such as γ-ray spectroscopy or β-counting, may in some instances be far smaller than the overall LLD; as a result, the presumed capability to detect important levels of (e.g.) environmental contamination may be much too optimistic.

⊙   The LLD shall be defined according to the statistical hypothesis testing theory, using 5% for both "risks" (errors of the first and second kind), taking into consideration possible bounds for systematic error. This means that the detection <u>decision</u> (based on an experimental outcome) and its comparison with a critical or decision level must be clearly and consciously distinguished from the detection <u>limit</u>, which is an inherent performance characteristic of the measurement process. (Note that physical non-negativity implies the use of 1-sided significance tests.)

⊙   Both the critical level and the LLD depend upon the <u>precision</u> of the measurement process (MP) which must be evaluated with some care at and below the LLD in order for the critical level and LLD to be reliable quantities. Information concerning the nature and variability of the blank is crucial in this regard. (For $\alpha = \beta$, and symmetric distribution functions, LLD ≈ twice the critical level, numerically.)

⊙   Given the above statistical (random error) bases it is clear that the <u>overall random error (σ) of the MP must be evaluated</u> -- <u>via</u> propagation, replication, or "scientific judgment" -- to compute a meaningful LLD. "Meaningful," as used here, refers to an LLD which in fact reflects the desired α, β error rates or risks.

⊙   A great many assumptions must be recognized and satisfied for the LLD to be meaningful (or valid). These include: knowledge of the error distribution function(s) (they may not simply be Poisson or Normal); consid-

eration of all sources of random error; reliable estimation of random errors and appropriate use of Student's-t and careful attention to sources of systematic error.

⊙ Systematic error derives from non-repeated calibration, incorrect models or parameters (as in γ-ray spectroscopy), incorrect yields, efficiencies, sampling, and "blunders." Bounds for systematic error should always be estimated and made small compared to the imprecision (σ), if possible. Systematic calibration and estimation error may become a very serious problem for measurements of "gross" (α, β) activity where the response depends on the relative mix of half-lives and particle energies.

⊙ Control of the MP also is essential, and should therefore be guaranteed by both internal and external "cross-check" programs. External cross-checks should represent the same type (sample matrix, nuclide mixture) and level of activity as the "real" effluent and environmental samples including blanks for the "principal radionuclides", and the cross-checks should be available "blind" to the measuring laboratory. Note that without adequate control or without negligible systematic error, LLD loses meaning in the purely probabilistic sense. The issues of setting bounds for residual systematic error and bounds for possibly undetected activity under these circumstances both deserve careful consideration, however.

⊙ Radionuclide interference (and increased Compton baseline) necessarily inflates the LLD, and must be taken into consideration quantitatively. The use of "a priori" and "a posteriori" to refer to this issue is strongly discouraged, because of needless confusion thereby introduced involving another usage of these terms (related to detection decisions and LLD).

6

○   Reporting practices are crucial to the communication and
understanding of data (as well as the validity of the respective LLD). This
is a special problem for levels at or below the LLD, where sometimes even
negative experimental estimates obtain. Full data reporting is recommended,
from a technical point of view, to alleviate information-loss and the
possibility of introducing bias when periodic averages are required. (Also,
policy on uncertainty estimates and significant figures is in order.)

4.   Related Policy Issues

○   Once defined and agreed upon, a uniform approach to LLD, statement
of uncertainty, QA assessment (external), and data reporting should be
established.

○   Issues involving interference (and LLD relaxation) and reliance only
on Poisson counting statistics (vs adequate replication and full error propa-
gation) must be settled. Other factors such as branching ratios/γ-abundance
should be considered in setting practically-achievable nuclide LLDs.

○   Significant distortions which could arise from: a) "gross" (α,β)
activity measurements, b) sampling systematic errors, and c) concealed
software and bad nuclear parameters must be highlighted and controlled.
(Institution of an external data "cross-check" QA program, as the IAEA γ-ray
intercomparison spectra, may be one fruitful approach to the last problem.)

○   Difficulties between scientific vs public (political) perceptions
connected with "detected" vs "non-detected" radionuclides especially in
reporting contexts need to be addressed.

o   Means for dealing with situations where the purely statistical assumptions underlying LLD may not be satisfied must be defined. (That is one purpose of the present report. See section II for a catalog of assumption difficulties.)

Implementation and Documentation

A potential basis for the NRC position for effluent and environmental radioactivity measurement process LLD's is developed and illustrated in this technical manual (NUREG/CR document). This document is designed to provide explicit information on: a) the history and principles of LLD's; b) practices actually encountered in the field at the time of this study; c) simple, clear yet accurate exposition and numerical illustrations of detection decisions and LLD use, as applied to effluent and environmental radioactivity measurements; and d) special technical issues, data, and bibliographic material (in the Appendix).

D. Highlights from Site Visits

The highlights developed from a series of site visits are presented as a synthesis of information gained rather than as a report concerning individual discussions or specific organizations. The information represents _my understanding_ from numerous discussions; the more critical issues may need to be appropriately verified. Also, it should be understood that the contents in this section constitute a record of my observations, not necessarily an indication that all parts are directly applicable to the Radiological .Effluent Technical Specifications (RETS). (e.g., parts 12 and 13).

Organizations and Individuals Visited (besides NRC-Headquarters)

4 November 1982   Dave Harward, Atomic Industrial Forum, Bethesda, MD

19 November 1982   Dave McCurdy, Yankee Atomic Electric Company,
Framingham, MA, (Environmental Lab)

5 July 1983   Jerry Hamada (Inspector), NRC Region V Office,
Walnut Creek, CA

6 July 1983   Roger Miller, Rancho Seco Power Plant, CA (accompanied
by J. Hamada)

7 July 1983   Rod Melgard, EAL, Inc. (Contracting Lab.), Richmond, CA

11 July 1983   Art Jarvis and Gene Easterly,
EPA - Las Vegas (cross-check program)

12 July 1983   Jim Johnson, Colorado State University, Ft. Collins
(measurements for Ft. St. Vrain plant)

9 August 1983   Mary Birch and Bob Sorber, Duke Power Co., Charlotte, NC
(HQ, and Lab at Oconee site)

21 November 1983   Carl Paperiello, (Marty Schumacher, Steve Rozak,
Al Januska) NRC Region III Office, Glen Ellyn, IL

22 November 1983   Leonid Huebner, Teledyne Isotopes Midwest Lab (formerly
Hazelton), Northbrook, IL

9 February 1984   Tom Jentz, John Campisi, Joan Grover,
Charlie Marcinkiewicz, NUS (Contractor Lab.), Gaithersburg,
MD

1. <u>Need and approach for the planned LLD manual</u>. With one exception, I came away from the several meetings with strong support for the aim of producing a manual. Most of those I visited (especially in the West) were quite anxious to receive a copy of the manual as soon as possible. Valuable suggestions included requests to treat the basic concepts in a unified and complete, yet easy-to-grasp manner (<u>e.g.</u>, hypothesis testing). One approach would be to include mathematics and appropriate reprints in an appendix, but worked-through examples in the text.

2. _Diversity of training and experience_. This was evident in speaking to personnel ranging from lab technicians to lab managers to company officials. This diversity underlines the approach called for in item 1. (It was noteworthy that some of the younger and least professionally trained personnel raised some of the most penetrating questions about assumptions, alternative approaches to data presentation and evaluation, etc.)

3. _Diversity of terminology, usage, etc._ Despite the definition and references provided by the NRC for LLD (_e.g._, throughout NUREG-0472), there exist a number of popular terms (LLD, MDA, MDC, ...) and formulations (2σ, S/N, hypothesis testing risks, ...) to the detection limit, and an even wider diversity of assumptions recognized (or ignored!) in practice. Some of the more pertinent practices (_re:_ assumptions) will be noted below.

4. _Policy Issues_. I found many opportunities to become enmeshed in policy. Despite my advance letter (and copy of the "manual" - work statement), certain of my hosts seemed to believe I could speak to policy -- _i.e._, what numerical values should be established for LLD's to be met. I explained that this was not my charge, though in certain special cases -- _e.g._, the effects of severe radionuclide interference on detection capabilities -- it might be useful to consider the impact of policy on practical operations (see below).

In certain cases, I was advised that the "process environment" mandated special approaches to the evaluation and reporting of data, because of large sample loads and the need for rapid decisions. Under some circumstances this could imply (statistically) conservatively biased reporting of data, and non-specific radionuclide measurements (_e.g._, β^- counting of separated iodine

isotopes, and treating the result as though it were all I-131). The issue I perceive is whether it is appropriate to recommend different LLD and/or reporting schemes depending on how busy a laboratory is.

5. <u>Detection decisions</u>. I found the full range of criteria: from decisions based on the critical level (such that α and β risks each equal 5%) to those based on LLD (such that "false positives" are infinitesimal, but "false negatives" are 50%!). I have the impression that the decision-making aspect of detection -- <u>i.e.</u>, the actual <u>testing</u> of the null hypothesis -- is not fully appreciated by all workers.

6. <u>Reporting (when "not detected")</u>. Such results are equated to zero, some upper limit, LLD, LLD/2, etc. All of those I spoke to recognized that averaging (<u>e.g.</u>, over a quarter) of such reported results is either imposs-ible, or positively or negatively biased. I sensed some resistance to reporting the <u>observed</u> value (especially when it is negative), though one group preserves such information for unbiased averaging; but then reports the <u>same</u> data in two different (biased) ways according to the policies mandated by different users of the data! Also, during one visit, I learned that company (?) policy leads to <u>different</u> ways of reporting "non-detected" results between environmental and effluent measurements.

7. <u>Radionuclide interference</u>. A significant issue. It is (universally) recognized that interference increases detection limits (all else being equal). The same example (Ce-144 with very large amounts of Co-58, -60) was raised during two visits, but with somewhat different (policy) perspectives. In the one, it was suggested that prescribed LLD's be relaxed (or possibly remain "pure solution" or interference-free LLD's) when excessive

interference is present because the relative contribution of Ce-144 (here) is trivial by comparison. In the other, caution was suggested, because even a small amount of Ce-144 could be an important indicator for transuranics.

8. _Blank, background, baseline_. Some ambiguity was noted in the current proposed NRC definition for LLD. Also, the question of real background variability and number of degrees of freedom (and Student's-t) were raised. One laboratory always assumes Poisson-background variability, or, if this seems exceeded, it shuts down until a problem is identified or expected behavior resumes.

9. _Non-counting errors_. Almost universally it was recognized that actual probabilities of detection (and LLD) depend upon _all_ sources of error, yet nearly all workers are using Poisson statistics only (for the blank and sample, and ignoring errors for efficiency or chemical yield estimates) to calculate LLD. Since the Relative Standard Deviation ≈30% at the detection limit ($\alpha = \beta = 0.05$), this approximation is partly justified. Severe errors, however, in blank estimates, detection efficiency (_e.g._, for cartridge filters and for gross-α deposits), and sampling[2] can seriously invalidate this (Poisson) approximation. Several of the groups are working very hard to estimate (and minimize) non-counting error, but there is little movement toward considering its (necessary) effects on the LLD.

One interesting suggestion (mutually developed) was to distribute blind cross-check samples having radionuclide concentrations slightly (_e.g._, 50%) higher than the intended (NRC) LLD's to assess the actual significance of non-Poisson error on detection capabilities. (This might also include blanks of "principal radionuclides" to test α-risk performance.)

----

[2]Sampling Errors -- _e.g._, involving soil particles, coolant containing sediment, single ion exchange beads, -- were in some cases shown to be overwhelming, reducing all other errors to insignificance.

10. <u>Modeling rather than direct measurement</u>. Knowing (at least approximately) relative dilution factors (laboratory, atmosphere, coolant systems) in many cases allows more accurate inferences to be drawn from relatively high level measurements followed by calculation -- as opposed to direct measurements of the diluted (dispersed) material. (This is followed, for example, in preparation of the EPA cross-check samples.)

11. <u>QA and cross-check samples</u>. I found some excellent intralab QA, but at the same time I found extremely strong support for external cross-check programs -- especially because of the wide range of (e.g.) contractor or technician capabilities. The EPA sample program is valuable (essential, since there is no other) for this purpose, but several useful extensions were suggested: increased frequency (perhaps suited to QA performance), truly "blind" samples (EPA's are clearly recognizable, and often given special attention), and samples which are closer in composition and level to those encountered in the various programs (environmental, effluent, waste). (Splits, especially with mobile laboratories serve effluent QA well, but availability of "known" samples would be valuable.)

12. <u>"De minimis" reporting</u>. Media other than air and water are in many cases not covered by specified LLD's (<u>e.g.</u>, oil, charcoal, ...), so that any detected activity must be reported. Apparently, the situation is analogous to that arising from one interpretation of the Delaney Amendment, where non-detection is taken equivalent to absence; so that reporting requirements (and public perceptions) are strongly affected as measurement techniques improve.

13. <u>Uncertainties, reporting levels, litigation</u>. In view of measurement uncertainty, one often meets the question of whether an experimental observation implies that the true value exceeds or is less than a specified regulatory limit. The issue is perhaps compounded when one considers a summation,

$$\sum_{1}^{n} \left(\frac{\text{concentration}}{\text{reporting level}} \right)_{i} \geq 1$$

as on page 5 of the NRC Radiological Assessment Branch Technical Position (November 1979). Both the magnitude of the total errors and the number of terms (n) impact this matter. Actions and legal defense can be rather complex as a result; so cautious attention must be given to matters of relative "costs", experienced judgment on the part of inspectors, burden of proof, etc.

14. <u>Continuous and continual monitoring; averaging.</u> A difficult area: varied equipment age or quality can make continuous monitors difficult to integrate reliably, and errors in estimated time constants and flow rates can be substantial. Continual monitoring (for period averaging), on the other hand, must be done with care to avoid missing non-monotonic behavior (excursions, ...). Random variations may be approximately normal (gaussian) close to the emission site, but log-normal when mixed in the environmental system. Averaging procedures (arithmetic <u>vs</u>. geometric mean) may differ accordingly. (Weighted averaging is yet another topic.)

15. <u>Multiple detection decisions</u>. Basing all decisions on $\alpha = 5\%$ (single observation false positive risk) means that on the average 1 in 20 blanks <u>will</u> be reported as detected. Adjustment so that, <u>e.g.</u> in a multi-component Y-ray spectrum, there is only a 5% change of <u>any</u> false positive, was a seemingly esoteric matter noted by very few of those I visited.

Also, not widely appreciated was the too liberal nature of an outlier rule (Chauvenet's criterion) being sometimes employed.

16. <u>Hidden algorithms, bad parameters</u>. A widespread, but not too widely appreciated problem is the nature and lack of access to computer programs used for γ-ray spectrum evaluation. A number of parameters (<u>e.g.</u>, branching ratios) both in certain nuclear data compilations and in some "canned" software routines are wrong. The absence of adequate software documentation and the inaccessibility of source code has caused moderate difficulties in several laboratories -- problems which may be exacerbated for small activities ($\gtrsim$ LLD), for high levels of interference (base-line shape, pile-up, ...), and for multiplets. One interesting test that was described, revealed software artifacts (algorithm switching) when computer output was examined for a series of sequential (known) dilutions of a given radionuclide sample. (Note the similarity to the classic, Standard Addition Method to reveal or compensate <u>chemical</u> interference.)

II. BASIC CONCEPTS[1]

In order to meet the underlying objective of defining LLD for use in Radiological Effluent Technical Specifications (RETS) it is necessary first to adopt a uniform and reasonable conceptual approach to the specification of detection capability for an MP, and it is then necessary to set forth a carefully-constructed and consistent scheme of nomenclature and mathematical statistical relations for specific application to the range of problems encountered in measurements of effluent and environmental radioactivity. Our goal in this section is to outline the preferred conceptual approach together with a reasonably complete catalogue of assumptions and means for putting it

---

[1]See Appendix A for selected nomenclature and terminology.

into practice. Detailed reduction of the basic formulas presented in this section will take place in the next section, for the several common categories of nuclear and radiochemical measurement; and explicit numerical examples will be given in the Appendix. Let us begin with a glance at the past.

A. Overview and Historical Perspective

Some appreciation for the evolution of methods for expressing detection capability may be gained from Table 1. In this table, which refers only to detection capability (not detection decision levels), we observe that the development of detection terminology and formulations for Nuclear and Analytical Chemistry covers an extended period of time and that it has been characterized by diverse and non-consistent approaches. (Besides alternative terms for the same concept, one occasionally finds the same term applied to different concepts -- viz., Kaiser's "Nachweisgrenze", which refers to the test or detection decision level, is commonly translated "detection limit"; yet, in english "detection limit" generally relates to the inherent detection capability of the Chemical Measurement Process (CMP).) For information concerning the detailed assumptions and formulations associated with the terms presented in Table 1 the reader is referred to the original literature. The principal approaches, however, are represented by: (a) Feigl -- selecting a more or less arbitrary concentration (or amount), based on expert judgment of the current state of the art; (b) Kaiser and Altshuler -- grounding detection theory on the principles of hypothesis testing; (c) St. John -- using signal/noise (assumed "white") and considering only the error

of the first kind; (d) Nicholson -- considering detection from the perspective of a specific assumed probability distribution (Poisson); (e) Liteanu -- treating detection in terms of the directly observed frequency distribution, and (f) Grinzaid -- applying the weaker, but more robust approaches of non-parametric statistics to the problem. The widespread practice of ignoring the error of the second kind is epitomized by Ingle in his inference that it is too complex for ordinary chemists to use and comprehend! Treatment of detection in the presence of possible systematic and/or model error is considered briefly in Ref. [33].

Table 1. Historical Perspective -- Detection Limit Terminology

---

Feigl ('23)	- Limit of Identification [Ref. 1]
Altshuler ('63)	- Minimum Detectable True Activity [Ref. 4]
Kaiser ('65-'68)	- Limit of Guarantee for Purity [Ref. 2]
St. John ('67)	- Limiting Detectable Concentration (S/N_{rms}) [Ref. 3]
Currie ('68)	- Detection Limit [Ref. 5]
Nicholson ('68)	- Detectability [Ref. 36]
IUPAC ('72)	- Sensitivity; Limit of Detection...[Ref. 22, 23]
Ingle ('74)	- ("[too] complex...not common") [Ref. 51]
Lochamy ('76)	- Minimum Detectable Activity [Ref. 7]
Grinzaid ('77)	- Nonparametric...Detection Limit [Ref. 44]
Liteanu ('80)	- Frequentometric Detection [Ref. 31]

A condensed summary of the principal approaches to signal detection is presented in Table 2. The hypothesis testing approach, which this author favors, serves also as the basis for the more familiar construction of confidence intervals for signals which are detected [83]. For more information on the relationship between the power of an hypothesis test and the

significance levels and number of replicates (for normally-distributed data)
the reader may refer to OC (Operating Characteristic) curves as compiled by
Natrella [84]. There it is seen, for example, that 5 replicates are neces-
sary if one wishes to establish a detection limit which is no greater than
2σ, taking [α] and [β] risks at 5% each. (Note the inequality statement;
this arises because of the discrete nature of replication.) Once we leave
the domain of simple detection of signals, and face the question of analyte
or radioactivity concentration detection, we encounter numerous added

---

Table 2. Detection Limits: Approaches, Difficulties

Signal/Noise (S/N) [Ref's 3,29,30,86]

Detection Limit $\equiv 2N_{p-p}$, $2N_{rms}$, $3s$ (n=16-20)

$$[N_{rms} \approx N_{p-p}/5.]$$

DC: white noise assumed, β-error ignored

AC: must consider noise power spectrum, non-stationarity,
digitization noise

Simple Hypothesis Testing [Ref's 2,5,26,56,83]

$$\hat{S} = y - \hat{B}$$

H_O: significance test (α-error) - 1-sided confidence interval

H_A: power of test (β-error) - Operating Characteristic Curve

Determination of S_D requires accurate knowledge of the distribution
function for $\hat{S}$

If $\hat{S} \sim N(S, \sigma^2)$, and α, β=0.05, then $S_D = 2S_C = 3.29~\sigma$

Other Approaches [Ref's 28,85,87,88]

Decision Analysis (uniformly best, Bayes, minimax), Information and Fuzzy
set theories.

---

problems or difficulties with assumption validity. That is, assumptions concerning the calibration function or functions -- i.e., the full analytic model -- and the "propagation" of errors (and distributional characteristics) become crucial. A catalog of some of these issues is given in Table 3; further discussion will be found in the following subsection. Finally, for more detailed summary of the relevant literature, the reader is referred to the review and bibliography in Appendix C.

Table 3. Concentration Detection Limits - Some Problems

⊙ σ^2 only estimated; H_o-test ok $(ts/\sqrt{n})$, but x_D is uncertain

⊙ Calibration function estimated, so normality not exactly preserved: $\hat{x} = (y-\hat{B})/\hat{A} \neq$ linear Fcn (observations)

⊙ B-distribution (or even magnitude) may not be directly observed

⊙ Effects of non-linear regression; effects of "errors in x- and y" (calibration)

⊙ Systematic error, blunders -- e.g., in the shape, parameters of A [$\delta \rightarrow \Delta$, without continual re-calibration]

⊙ Uncertain number of components (and identity) [Lack of fit tests lose power under multicollinearity]

⊙ Multiple detection decisions: $(1-\alpha) \rightarrow (1-\alpha)^n$

B. Signal Detection (principles)

1. Alternative Approaches

A necessary, first step in treating signal detection is to consider what magnitude observed (a posteriori) response (gross signal) constitutes a statistically significant deviation (increment, or net signal) from the zero-level (blank or background or baseline in radioactivity measurement). This increment, which really represents a critical or decision level (S_C)

with which the observed signal is compared, is derived from the distribution function for the noise. If the noise can be considered normal ("Gaussian") with parameter-σ (standard deviation), S_C is given by a fixed multiplier times σ, and the detection process becomes simply a significance test based on comparison of the observed with the critical signal to noise ratio. Certain non-trivial problems arise if the noise power spectrum is not "white" (Gaussian) and when the signal is continuous (in time) but is sampled periodically. These issues are treated in some depth in References indicated in Table 2.

The test, however, is incomplete (though widely practiced!) for our purposes. It speaks only to the question of signal detection (a posteriori) -- i.e., the detection decision given the noise probability density function (pdf) and an observed signal. It is important to us in that the significance level of the test $\underline{\alpha}$ is equivalent to the false positive probability or "error of the first kind." (That is, α equals the probability that one would, by chance, falsely conclude that a blank contained excess radioactivity.) This is insufficient, per se, for us to specify the detection capability or LLD, which is an a priori performance characteristic of the Measurement Process (MP).

A solution is found in the theory of Hypothesis Testing, wherein we use an experimental outcome $\hat{S}$ not simply to test for the presence of a signal but actually to discriminate between two possible states of the system: H_O and H_D. H_O and H_D are, respectively, the "null hypothesis" and the "alternative hypothesis" and the critical level S_C is set in such a way that an optimal decision (in the long run) is made between the two hypotheses. As the

subscripts imply, H_O refers to samples containing no net radioactivity, and H_D, to samples containing radioactivity at the LLD. In terms of the net signal, $\underline{H_O}$: $S=0$ and $\underline{H_D}$: $S=S_D$ (S being the true, but unknown net signal.)

Two of the basic forms of Hypothesis testing require information or assumptions that are not generally available for simple chemical or physical measurements. The first involves the use of the "Bayes Criterion" which requires prior probabilities for H_O and H_D, as well as the assignment of costs for making incorrect decisions. In this case S_C would be set to minimize the average (long-run) cost. The second approach, which is related to game theory, does not require prior probabilities. Rather, it is designed to minimize the maximum cost over the entire set of possible prior probabilities. Appropriately, this is termed the "Minimax" decision strategy. Lacking either costs or prior probabilities, we prefer to define detection capability (LLD) on the basis of simple hypothesis testing ("Neyman-Pearson criterion") which considers H_O, H_D and S_C simply in terms of the probabilities of drawing false conclusions when $\hat{S}$ is compared to S_C. Lucid expositions of all three decision strategies are given in Ref's 28, 29 and 79. A more complete development of simple hypothesis testing for direct application to LLD follows.

2. Simple Hypothesis Testing and the LLD

[adapted from Ref. 38]

The basic issue we wish to address is whether one primary hypothesis [the "null hypothesis", H_O] describes the state of the system at the point (or time) of sampling or whether the "alternative hypothesis" [H_D] describes it. The actual test is one of consistency – i.e., given the experimental sample, are the data consistent with H_O, at the specified level of signifi-

cance, α? That is the first question, and if we draw (unknowingly) the wrong conclusion, it is called an error of the first kind. This is equivalent to a false positive in the case of trace analysis - i.e., although the (unknown) true analyte signal S equals zero (state H_0), the analyst reports, "detected".

The second question relates to discrimination. That is, given a decision- (or critical-) level S_C used for deciding upon consistency of the experimental sample with H_0, what true signal level S_D can be distinguished from S_C at a level of significance β? If the state of the system corresponds to H_D ($S=S_D$) and we falsely conclude that it is in state H_0, that is called an error of the second kind, and it corresponds in trace analysis to a false negative. The probabilities of making correct decisions are therefore $1-\alpha$ (given H_0) and $1-\beta$ (given H_D); $1-\beta$ is also known as the "power" of the test, and it is fixed by $1-\alpha$ (or S_C) and S_D. One major objective in selecting a particular MP is thus to achieve adequate detection power $(1-\beta)$ at the signal level of interest (S_D), while minimizing the risk (α) of false positives. Given α and β (commonly taken to be 5% each), there are clearly two derived quantities of interest; S_C for making the detection decision, and S_D the detection limit. (If, for RETS, our concern were strictly with the net signal rather than radioactivity concentration, LLD would be taken equal to S_D.) Figure 1 illustrates the interrelation of α, β, S_C and the detection limit.

An assumption underlying the above test procedure is that the estimated net signal $\hat{S}$ is an independent random variable having a known distribution. (This is identical to the prerequisite for specifying confidence intervals.) Thus, knowing (or having a statistical estimate for) the standard deviation of the estimated net signal $\hat{S}$, one can calculate S_C and S_D, given the form of
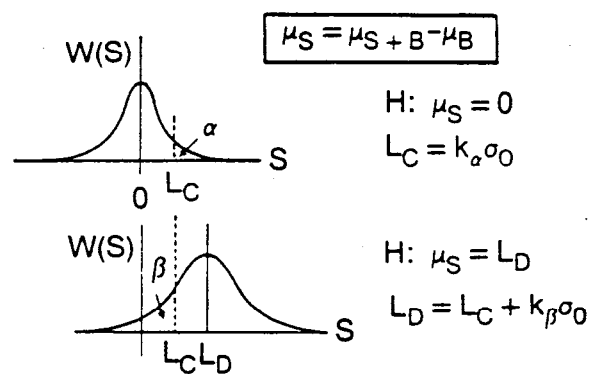
Fig. 1. Hypothesis testing; errors of the first and second kinds

the distribution and α and β. If the distribution is Normal with constant σ, and $\alpha = \beta = 0.05$, $S_D = 3.29 \, \sigma_S$ and $S_C = S_D/2$. Thus, the relative standard deviation of the estimated net signal equals 30% at the detection limit (5). Incidentally, the theory of differential detection follows exactly that of detection, except that ΔS_{JND} (the "just noticeable difference") takes the place of S_D, and for H_O reference is made to the base level S_O of the analyte rather than the zero level (blank). A small fractional change $(\Delta S/S)_D$ thus requires even smaller imprecision.

Obviously, the smallest detection limits obtain for interference-free measurements and in the absence of systematic error. Allowance for these factors not only increases S_D, but (at least in the case of systematic error) distorts the probabilistic setting, just as it does with confidence intervals. Special treatments for these questions and for non-normal distributions will be given as appropriate. Not so obvious perhaps is the fact that S_D depends on the specific algorithm selected for data reduction. As with interference effects on S_D, this dependence comes about because of the effect on σ_S, the standard deviation of the estimated net signal. More explicit coverage of these matters will be given below and detailed derivations and numerical examples will be found in section III and the Appendix of this report, respectively, (see also Ref. 33.).

Hypothesis testing is extremely important for other phases of chemical and radiochemical analysis, in addition to the question of analyte detection limits. Through the use of appropriate test statistics, one may test data sets for bias, for unexpected random error components, for outliers, and even for erroneous evaluation (data reduction) models (33). Because of statistical limitations of such tests, especially when there are relatively few degrees of freedom, they are somewhat insensitive (lack power) except for

quite large effects. <u>For this reason it is worth considerable effort on the part of the analyst to construct his MP so that it is as free from or resistant to bias, blunders, and imperfect models as possible.</u>

Figure 2 gives an illustration of the difficulties of detecting both systematic error and excess random error. There we see that just to detect systematic error when it is comparable to the random error (σ) requires about 15 observations; and to detect an extra random error component having a comparable σ requires 47 observations (89). In a simple case involving model error it has been shown that analyte components omitted from a least-squares multicomponent spectrum fitting exercise must be significantly above their detection limits (given the correct model) before misfit statistics signal the error (33). This limitation in "statistical power" to prevent significant model error bias, especialy in the fitting of multicomponent spectra, is one of the most important reasons for developing multidimensional chemical or instrumental procedures and improved detectors of high specificity or resolution.

C. <u>General Formulation of LLD - Major Assumptions and Limitations</u>

The foregoing discussion provides the basis for deriving specific expressions for the LLD for <u>signals</u>, given α and β, and σ_S as a function of concentration. Before treating concentration detection limits generally, and radioactivity concentration detection limits specifically, however, it is necessary to examine a number of basic assumptions connected with the concept and with the MP.
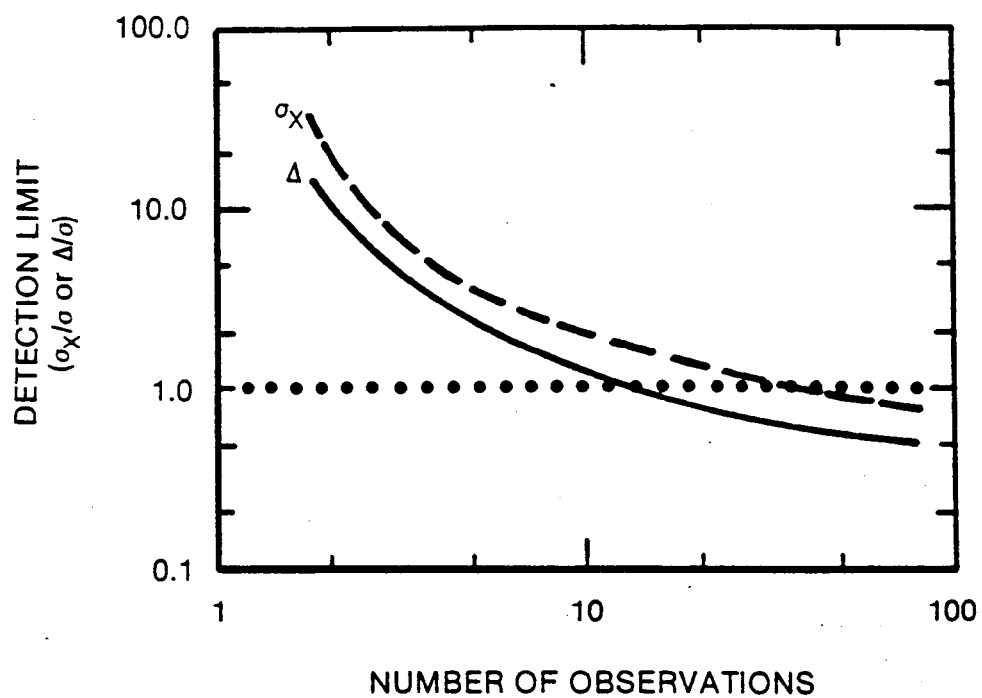
Fig. 2. Detection limits vs. number of observations for extraneous random
error (σ_X, dashed curve) and systematic error (Δ, solid curve).

1) Detection Decisions vs Detection Limits

The signal detection limit S_D is undefined unless α or S_C is defined <u>and</u> <u>applied</u>. That is, detection decisions are mandatory if detection limits (in the hypothesis testing sense) are to be meaningful. The relatively common practice of equating these two levels ($S_C = S_D$) is equivalent to setting the false negative risk at 50%. That is, a detection limit so defined will in fact be missed half the time! The recommended practice therefore is to take $\alpha = \beta = 0.05$, in which case,

$$S_C = z_{1-\alpha}\sigma_0 = 1.645\ \sigma_0 \tag{1}$$

$$S_D = S_C + z_{1-\beta}\sigma_0 = 2S_C = 3.29\sigma_0 \tag{2}$$

provided the standard deviation of the net signal σ_S is known and constant (at least up to the detection limit) and it is normally-distributed (z refers to the indicated percentile of the standard normal variate.) In Eq's (1) and (2), $\sigma_0 = \sigma_S$ (at $S=0$); this in turn equals σ_B if the average value of the blank is well-known (Ref. 5). (For "paired observations", $\sigma_0 = \sigma_B\sqrt{2}$.) S_C is used for testing whether an observed signal $\hat{S}$ is (statistically) distinguishable from the blank -- i.e. "detected"; S_D represents the corresponding MP performance characteristic, i.e., the detection limit. Although $S_D/S_C \approx 2$ generally, this is <u>not</u> universally true. A number of exceptional cases which do occur, especially in extreme low-level counting and in nuclear spectroscopy, are treated in section III of this manual.

2) A Priori vs A Posteriori; Changes in the MP (Interference, ...)

Some confusion exists in the usage of these terms which mean "before the fact" and "after the fact." The "fact" referred to is the experimental outcome -- i.e., the observation of a (random) signal $\hat{S}$, associated with the measurement of a particular sample. The MP, which <u>necessarily</u> includes the

influence of the sample on the characteristics of the measurement system is not the "fact", from the perspective of hypothesis testing. In order to make intelligent decisions regarding S we need therefore information concerning the MP characteristics, notably σ_S at S=0 and the variation of σ_S with concentration. This in turn is influenced by the level and nature of any interfering species in the sample in question. Also, as soon as we consider the real quantity of interest, the concentration detection limit (x_D), we require information concerning the overall calibration factor for the particular sample; this includes the (radio)chemical yield or recovery, detection efficiency (as perturbed by sample matrix effects: absorption and scattering), volume or mass of the sample, etc.

Thus prior knowledge concerning the sample in question is required in order to compute S_C which one needs for the a posteriori test of $\hat{S}$; it is needed also to compute the signal and concentration detection limits (S_D, x_D) for that sample. Such prior information may be obtained in a preliminary or screening experiment; it may be estimated from data resulting from the experiment, itself; or it may be assumed (not recommended) independent of the experiment. The last approach might be taken if one were interested in "pure solution" or ideal sample detection limits, where there is no interference, no matrix effects and perfect or unvarying recoveries. A slightly less disastrous alternative, to assume average values for such quantities or effects, results in needless information loss. To caricature the situation, it's equivalent to permitting the counting time to vary in a haphazard fashion from sample to sample and guessing an average time for calculating individual counting rates. The point is: the critical (decision) level and detection limit really do vary with the nature of the sample. So proper

assessment of these quantities demands relevant information on each sample, unless the variations among samples (e.g., interference levels) are quite trivial.

Some perspective and a suggested approach to this matter are given in Fig. 3. Here, we consider three possible outcomes for an experiment ("experiment-a") which is designed (sample size, expected interference level or background activities, counting time, etc.) according to our prior knowledge of the MP. This prior "knowledge", which here includes the assumption of zero interference ($I=0$), we designate "prior(a)"; it leads to a concentration detection limit x_{D_0} based on a background equivalent activity B_O. We consider the experiment adequately designed if this estimated detection limit x_D (actual LLD) does not exceed the specified maximum level x_R (prescribed LLD).

As soon as the (first) experiment is performed, we gain two kinds of information: new data on the MP-characteristics for the sample at hand, and an experimental result $\hat{x}_a$. The three possible outcomes (MP characteristics) depicted in Fig. 3 show progressively greater background- (or baseline-) equivalent activities ($B_3 > B_2 > B_1$) and therefore similarly increasing detection limits (x_D's). For outcome-1, the posterior MP characteristics ["post(a)"] are equivalent to our assumed prior MP-characteristics ["prior(a)"], -- i.e, $B_1 = B_0$ -- so of course the detection limit is as calculated ($x_{D_1} = x_{D_0}$) and the experiment is adequate ($x_D \le x_R$). For outcomes-2 and -3 the posterior characteristics differ from the prior; there <u>is</u> interference (B_2 and $B_3 > B_o$), so the detection limit is greater. Outcome-2 still shows an adequate detection limit ($x_{D_2} \le x_R$), so our task is complete --the initial design was sufficiently conservative ($x_{D_0} < x_R$) that some interference could be tolerated.
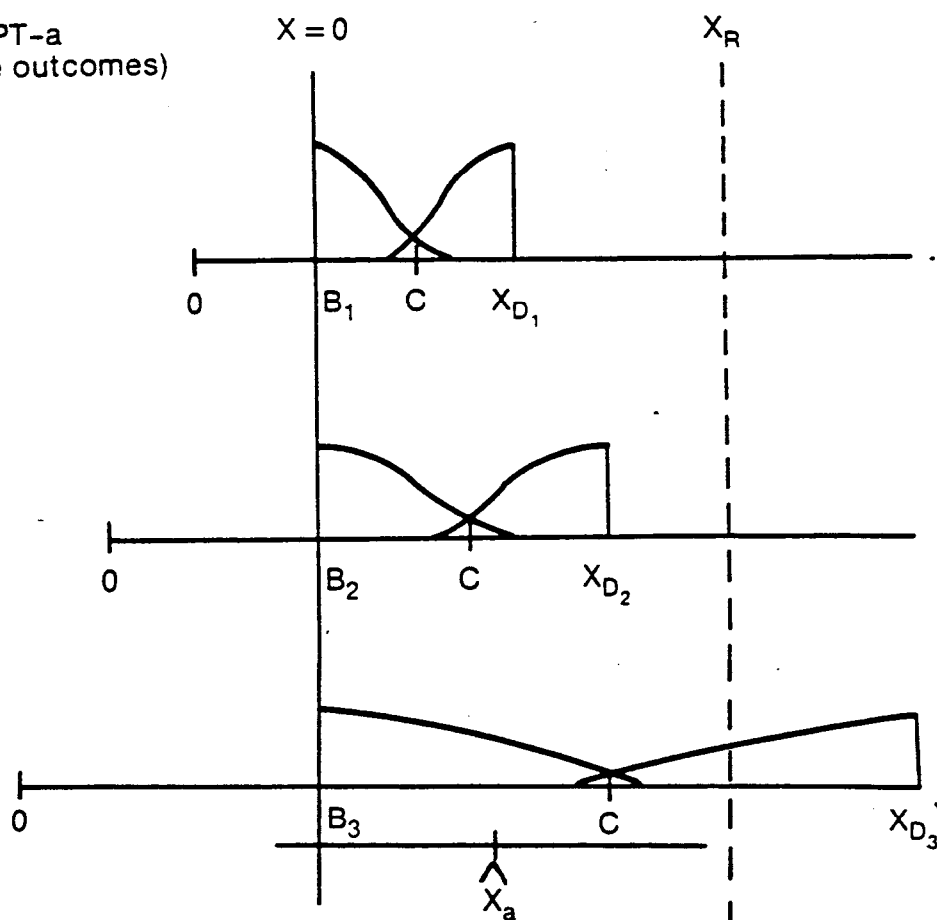
Fig. 3. A Priori vs A Posteriori, Sequential Experiments and the Effect
of Interference

The third set of MP-characteristics (outcome-3) correspond to a sample having so high a level of interference that the initial design was inadequate $(x_{D_3} > x_R)$. We therefore must use this posterior information ("post$^{(a)}$") as our new prior information ("prior$^{(b)}$") to re-design the MP to yield adequate characteristics $(x'_{D_3} \leq x_R)$, in preparation for a second (final) experiment. (This is still properly considered "a priori" in the technical sense of hypothesis testing until the second experimental result $\hat{x}_b$ ["fact" or observation] has been obtained.) Such re-design can be based on any of the MP-variables under our control, such as sample size, radiochemical separation or concentration, or counting time. (In Fig. 3 we indicate re-design simply as an extension of counting time for relatively long-lived radioactivity.) A 1-line summary of these comments regarding sequential experiments would be simply to state that one's posterior becomes another's prior.

3. Continuity of Hypotheses; Unprovability

Hypothesis-testing as outlined above was dichotomous -- that is, we referred to the null hypothesis (H_0: S=0) and the detection limit hypothesis (H_D: S=S_D) only. In fact, S is a continuous quantity which may take on any value from zero and some large, reasonable upper limit.[1] What takes place when we compare $\hat{S}$ with S_C and make the detection decision is to conclude that one or the other of our two hypotheses (H_0, H_D) is quite unlikely, or more correctly that such a result $\hat{S}$ is quite unlikely (here, $\leq 5\%$ chance of

----

[1] A logician might object to this statement on the basis that atoms are discrete; and such an argument might even seem relevant if we had, say, 100 atoms of a short-lived radionuclide and a perfect (100% efficient) detector. We could count them all. Even here, however, the "S" that as scientists we're interested in is not the number of atoms in that particular sample, but its expected value -- such as the long-run average that would arise from repeated, identical activation analyses. The underlying issue relates to compound probability distributions; a treatment for the case of radioactivity is given in Ref. 63.

occurring) given H_O or H_D. The other hypothesis (H_O if $S \leq S_C$, H_D if $S > S_C$) is said to be consistent with the observation, but it is by no means proved. An infinite number of intermediate values of S are also consistent! (The most likely is $S = \hat{S}$.) This bit of logic may seem trivial and obvious to some, and subtle and irrelevant to others, but there is one curious and important consequence. The habit of "accepting" the hypothesis that is not rejected, sometimes leads to biased reporting of data. For example, if $\hat{S} \leq S_C$, the value reported may be zero; the other extreme is reporting it as being at the detection limit, if $S > S_C$. A further comment on this matter is given in the subsection on Reporting of Results (section II.D.4). (See also note A13.)

4. The Calibration Function and LLD.

Since our concern is with the detection limit for radioactivity concentration -- i.e., the "lower limit of detection" (LLD) -- we must go beyond the above exposition on signal detection. If the calibration function, relating response $\underline{y}$ to concentration $\underline{x}$ is linear,

$$y = B + Ax + e_y \qquad (3)$$

where B represents the blank; A, the calibration constant or factor; and e_y, the error in the observation y.

The estimated net signal is

$$\hat{S} = y - \hat{B} \tag{4}$$

$\hat{B}$ being an independent estimate for B; and the estimated concentration is

$$\hat{x} = (y - \hat{B})/\hat{A} \tag{5}$$

$\hat{A}$ being an independent estimate for A. (Here, "independent" means independent of the observation y . Interdependence [correlation] of $\hat{B}$ and $\hat{A}$ always results, of course, when they are both estimated from the fitting of a single set of calibration data.)

Ideally we would next determine σ_x as a function of x either _via_ replication, or by error-propagation. Complete replication of the entire calibration and sample measurement process for the full range of sample matrixes and interfering activities to yield and adequate number (n) of replicates: $\hat{x}_i$ for i = 1 to n spanning the full concentration range of concern (from zero to - LLD) would be a _very large_ task. (For the estimated standard deviation to have a relative uncertainty (95% CI) of ±10% for example would require about n = 200 replicates at each concentration!) We favor therefore error propagation, reserving occasional full replication for control of quality and blunder identification.

Error-propagation is straightforward for linear functions of normally-distributed random variables. Thus,

$$V_S = V_y + V_B = \sigma_S^2 \tag{6}$$

where V represents the variance of the subscripted quantity. Since E(y) (the expected value of y) equals S + B,

$$V_O = V_S(S=0) = V_B + V_{\hat{B}} \tag{7}$$

33

so, if the observations leading to $\hat{B}$ and y are equivalent, $V_O = 2V_B$ or $o_O = o_B\sqrt{2}$ as noted earlier. Calculation of S_C and S_D follow immediately (assuming still Normality).

With the introduction of a random variable $\hat{A}$ in the denominator of Eq. 5, complications set in because we now have a non-linear function (ratio) of random variables. If ϕ_A (relative standard deviation or RSD of A) is quite small, the distribution of $\hat{x}$ is only slightly skew; however, the appropriate error propagation formula (not shown), which itself is an approximation, contains the unknown quantity A. The consequence is that both x_C and x_D are themselves uncertain. (Or, if we <u>choose</u> values for x_C and x_D, the hypothesis testing errors α and β are uncertain.) Full treatment of this matter is beyond the scope of this document, but further details may be found in Ref. 76.

The approach adopted for LLD purposes, which we label "S-based" is simpler in concept and straightforward in application. That is, we treat the detection decision strictly in the signal domain, using $\hat{S}$ and S_C. The corresponding signal detection limit S_D is then transformed into the "true" concentration detection limit x_D using the true calibration factor A, which we do not know.

$$x_D = S_D/A = (z_{1-\alpha}o_O + z_{1-\beta}o_D)/A \qquad (8)$$

Using bounds for A; $\hat{A} \pm z_{1-\gamma/2}o_A$, we can then calculate a confidence interval for x_D. Taking a conservative viewpoint, we go one step further; namely $A_m = \hat{A}-z_{1-\gamma/2}o_A$ is inserted in the denominator of Eq. (8). The resulting quantity is an upper limit for x_D for $\beta = 0.05$. (A dual interpretation, which will not be discussed here, defines x_D in conjunction with an upper limit for β. α, of course, remains at 0.05; and neither S_C nor S_D suffer from the A-uncertainty, because they are strictly signal-based. When A is

not randomly sampled, the uncertainty in x_D no longer represents a "confidence" interval. It must be viewed as a systematic error interval. Finally, if this conservative estimate (upper limit) for x_D is less than the prescribed regulatory limit (x_R), the objective of RETS will have been met.

Recognizing the distinction between x_R -- the maximum permissible LLD, or "regulatory limit", and x_D -- the actual LLD or "concentration detection limit" for a particular sample and measurement technique, and the RETS requirement:

$$x_D \leq x_R \qquad\qquad (9)$$

it becomes interesting to consider inequality approaches. One such inequality, forced on us because of the non-linear relation Eq. 5, has already been useful in conjuncion with Eq. 9. The crucial point is that Eq. 9 removes the necessity that x_D be known exactly or with a fixed small relative uncertainty. As long as a reasonably chosen upper limit for x_D satisfies this relation the problem is solved.

A second type of inequality involving x_D, of great practical importance, derives from upper bounds which can be derived <u>immediately from the experimental result</u> ($\hat{x}$, σ_x) which is necessarily produced for <u>every analysis</u>. The resulting upper bound for x, if $\hat{x} > x_C$, can be shown <u>always</u> to exceed x_D. Therefore, if for a given sample that bound satisfies Eq. 9, there is no need to re-determine the actual detection limit or to re-design the experiment. (See the comments on sequential experiments, accompanying Fig. 3 [section II, C.2], and the note [B4] in Section III for a slightly extended discussion of the use of inequalities for rapid estimation of bounds for the detection limit.)

A purposely controversial, "non-detected" result $(\hat{x}_a)$ has been shown in Fig. 3, so that we may address the matter of an inadequate MP $(x_D > x_R)$ for which a seemingly adequate result (x_a-upper limit $< x_R$) has been obtained. We advise caution. That is, if $x_{\bar{D}} > x_R$, the uncertainty associated with any given measurement is apt to yield rather gradually changing significance levels (and false negative errors, β). It is advisable in cases such as this to estimate directly the probability β which would obtain taking $\underline{x_R}$ as the upper limit. That is, assuming normality

$$z_{1-\beta'} = \frac{x_R - \hat{x}}{\sigma_x} = \frac{x_R - x_U}{\sigma_x} + z_{.95} \tag{10}$$

If the 90% CI upper limit ($x_u = \hat{x} + 1.645\ \sigma_x$) is smaller than x_R, then β is necessarily less than 5%. However, as is obvious from Eq 10, the statistical significance of a given difference (x_R-x_U) decreases with increasing σ_x, which is to say it decreases with increasing LLD (x_D). Taking the result in Fig. 3, $x_0 = x_a + 1.645\ \sigma_{xU} = 0.9\ x_R$ (where $x_{D_3} = 1.5\ x_R$), we find that $(x_R$-$x_U)/\sigma_x = 0.219$ [assumes $\sigma_0 = \sigma_x = $ const.], so $z_{1-\beta'} = 1.864$ or $\beta' = 0.031$. This is not so much smaller than the base value $\beta = 0.05$ or, put differently, the upper limit from a 95% CI would <u>exceed</u> x_R. Contrast this with outcome-1 in Fig. 3, where x_{D_1} (and therefore σ_0) is smaller by a factor of 3. There, if an x_u were $0.9\ x_R$, $z_{1-\beta'}$ would be $1.645 + 3\ (0.219) = 2.302$, so $\beta' = 0.01$, and a 98% CI would be required for the upper limit to reach x_R.[3]

A final set of precautionary notes regarding the calibration function are in order:

⊙   The presumed straight-line model (Eq. 3) is generally adequate over a small concentration range ("locally linear"), such as between

----

[3]The numerology in this paragraph takes an added impact when one faces the issue of multiple detection decisions, where still more stringent requirements are placed on α and β. (See section II.D.4.)

36

$x = 0$ and $x = x_D$. If there is any doubt, however, such a presumption should be checked; and, above all, the slope or "calibration constant" <u>A</u> in the region of the detection limit should not be derived from remote data ($x \gg x_D$) where the curve may exhibit non-linearity (Ref. 76).

⊙ Imposed (instrumental, software) thresholds, in place of S_C, will not only alter α but may change the relevent "local" slope -- unless the calibration curve is perfectly straight (Ref. 76).

⊙ The calibration factor A, and any of the factors that comprise it -- Y (yield), E (efficiency), V (sample mass or volume), T (counting time function) -- may show interactions with B (background, baseline, blank, interference). Such further distortions (of Eq. 3) are discussed briefly in section III.

⊙ If non-linear estimation techniques, such as non-linear least squares, are employed for nuclide identification or for estimation of calibration curve parameters, values of α and β and the distribution of x can be perturbed. (Ref. 90).

⊙ Obvious, but worth stating, is the fact that ϕ_A (RSD of A) for use in connection with Eq. (8) is

$$\phi_A = [\phi_Y^2 + \phi_E^2 + \phi_V^2 + \phi_T^2]^{1/2} \qquad (11)$$

provided that all the constituent ϕ's are small. (Sampling errors, which could be manifest in the factors Y, E, or V may not always satisfy this requirement. ϕ_T, on the other hand, is effectively zero in most counting situations -- though uncertain (temporal) sampling input functions, or uncertain half-lives or radionuclide mixes could affect even this quantity.)

5. Bounds for Systematic Error

It would be marvelous if all our errors were random and of known distribution (with known parameters), and even more so if we could rely on their being Poisson. Such is never the case, so it is inappropriate to apply the foregoing random-error based hypothesis testing framework for x_D-calculation, except as an asymptotic component. With carefully controlled experimental work, however, that asymptotic component fortunately can be the principal component.

A basis for the treatment of detection decisions and detection limits in the presence of possible (uncorrected) systematic error is given in Ref. 33 for the case of signal detection. We extend that here to include the case of "S-based" concentration detection, through the introduction of a second systematic error bound parameter. Building on Eq. (8) for the random-error-based concentration detection limit, we get

$$S_C = \Delta + z_{1-\alpha}\sigma_O \tag{12}$$

$$x_D = f(2\Delta + z_{1-\alpha}\sigma_O + z_{1-\beta}\sigma_D)/A \tag{13}$$

where the quantity in the numerator in parentheses in Eq. (13) is S_D (incorporating blank systematic error bounds), and $\underline{f}$ is a proportionate amplification factor to provide a conservative bound for possible systematic error in A. Thus, if A = YEVT (ignoring the 2.22 pCi conversion factor) were based on a one-time calibration such that random calibration errors became systematic,

$$f = 1 + z_{1-\gamma/2} \, \phi_A \tag{14}$$

where ϕ_A is given by Eq. (11). Δ represents the a bound for possible blank or interference systematic error. It can be further decomposed into $\Delta_B B$ where Δ_B denotes the relative systematic error bound in the blank (or interference) and B denotes the magnitude of this quantity. (See Eq. 4.)

38

If we re-cast Eq. (13) in terms of radioactivity, assuming $o_O = o_D$ and taking $z_{1-\alpha} = z_{1-\beta} = 1.645$

$$x_D = \frac{f(2\Delta_B B + 3.29 o_O)}{2.22 \text{ (YEVT)}} \qquad (15)$$

Here, the numerator is in units of counts, and x_D, in units of pCi per unit mass or volume.

Following our Δ-notation for the relative systematic error bound we obtain from Eq. (14)

$$f = 1 + \Delta_A \qquad (16)$$

Clearly, the best experimental practice would include exhaustive theoretical and/or experimental studies to obtain reliable values for Δ_B and Δ_A. That empirical evaluation of such quantities is not trivial is shown in Fig. 2, where we see that just to <u>detect</u> a systematic error equal in magnitude to the random error of the MP requires more than ten observations (for standard error reduction).

In lieu of this, and for the sake of providing explicit, reasonable limits for the Δ's, we suggest the following [see notes A11 and B3]:

$$\Delta_{Bk} = 0.05, \quad \Delta_I = 0.01, \quad \Delta_A = 0.10$$

where "Bk" refers to both the blank and background and "I" refers to baseline or interfering activity effects on B. Systematic error of still another type, systematic <u>model</u> error is beyond the scope of our discussion though it is treated briefly in section III. C and in some detail in Ref. 72.

Equations (12) and (15) thus reduce to

$$S_C = (0.05)B + 1.645 \, o_O \quad \text{[counts]} \qquad (17)$$

$$x_D = (0.11)BEA + (0.50)\left(\frac{3.29 \, o_O}{YEVT}\right) \quad \text{[pCi/g or L]} \qquad (18)$$

39

for the case of Blank (Bk) predominance. If I >> Bk, then the coefficients
of the first terms in Eq's (17 and 18) become 0.01 and 0.022. B, in Eq. (17)
represents the Blank counts; and BEA, in Eq. (18) is the Blank Equivalent
Activity. As we shall see in subsequent discussions, this is a very impor-
tant quantity both for the calculation of the systematic error bound (term-1,
Eq. (18)) and for derivation of the random error-based term-2 (through σ_0).
σ_0 is the standard deviation of the estimated net signal (counts) when its
true value is zero. Its magnitude depends on the specific counting (measure-
ment) process, and it is the subject of the second following subsection.

Equation (18) is the expression for the LLD (actual [x_D], not
prescribed [x_R]). It is valid only when used in conjunction with Eq. (17).
Also, it carries the assumption of normality, and it should therefore be used
only when the "blank experiment" yields B $\gtrsim$ 70 counts. (See section III for
the treatment of very low-level counting and other special situations.)

D. Special Topics Concerning the LLD and Radioactivity

1. The Blank, Blank Equivalent Activity (BEA), and Regions of Validity

The ultimate limit of detection for any nuclear or chemical measurement
process is governed by the systematic and random uncertainty in B. (For B,
read: background, blank, interference, model error bias, etc.) For this
reason BEA should be recognized as an important benchmark in considerations of
detection capabilities. Some useful perspective on the nature and importance
of B-variations is offered in the following three paragraphs (adapted from
Ref. 38.)

"Unfortunately, there is no alternative to extreme vigilence when
treating the limitations imposed by the blank. In the best of circumstances
the mean value of the blank might be expected to be constant and its

fluctuations ("noise") normally distributed. Given an adequate number of
observations, one could estimate the standard deviation of this noise and
therefore set detection limits and precisions for trace signals. In situa-
tions where the chemical (analyte) blank remains small compared to the
instrumental noise blank this procedure may be valid, as in many low-level
counting experiments. Even here, however, to assume that the noise is nor-
mally or Poisson distributed, or to estimate the background from one or two
observations is to invite deception. As indicated in Table 4, there is a
significant chance (5% for normally-distributed blanks) that the expected
value of the noise (blank standard deviation) will exceed the observed dif-
ference between two blanks by a factor of 16! Subtle perturbations arise
even in the instrumental blank situation. For example, if the analyte detec-
tion efficiency changes discretely or even fluctuates, it is quite possible
that the instrumental blank will suffer a disproportionate change (77).

Certain special cases occur where the blank can be reliably estimated,
and therefore adjusted, indirectly. This is the situation: for "on-line"
coincidence cancellation of the cosmic-ray mu-meson component of the back-
ground in low-level radioactivity measurement (where there is not even a
stochastic residue from the adjustment process); for the adjustment of the
baseline (due generally to multiple interfering processes) in the fitting of
spectra or chromatograms; and for correction for isonuclidic contamination
(due to interfering nuclear reactions) in high sensitivity nuclear activation
analysis.

When the blank is due to contamination (as opposed to interferences or
instrumental background), high quality trace analysis is at its greatest
risk. Assumptions of constancy, normality or even randomness are not to be
trusted. An apparent analyte signal may be almost entirely due to

contamination (78); and blank correction must take into account its point(s) of introduction and subsequent analyte recoveries. The randomness assumption may be inappropriate because the blank may depend upon the specific history of the sample, container or reagents (35). Also when procedures are applied to real sample matrices as opposed to pure solutions blank problems abound, as was observed, for example, in the analysis of Pb (at a concentration of 30 ng/g) in porcine blood in contrast to aqueous solutions (93). (Reference 93 is also commended to the reader for a more complete treatment of the blank in trace analysis.) The most severe test of this sort comes when "blind" blanks together with samples at or near the detection limit, all in actual sample matrices, are submitted for analysis. Horwitz, for example, referring to collaborative tests of "unknowns" for 2378-TCDD in pure solutions, beef fat, and human milk, noted that significant numbers of false negatives began to appear when concentrations were less than 9×10^{-12} (μg/g), and that false positives increased from 19% for blank "standards" to over 90% for human milk samples (94)!"

Table 4. The Blank

- Direct Observation – Crucial for Detection Limit

- Adequate No. of Measurements Needed: With but two, σ_B may be 16 times the difference

- Efficiency Correction May Differ Between Blank and Analyte (Scales, 1963) [Ref. 77]

- Yield Corrections Must Recognize Point(s) of Introduction of Blank (Patterson, 1976) [Ref. 78]

- Multisource Blanks Generate Strange Probability Distributions – Shape and Parameters Important (Kingston, 1983) [Ref. 95]

- Poisson Hypothesis Must be Tested for Counting Background and Blank

---

In relatively controlled environments, especially if B is not an excessive number of counts, the Poisson assumption ($\sigma^2_B \approx B$) may be reasonably valid. The possibility of additional systematic and random error components should never be dismissed, however; and it is recommended that both types of non-Poisson B-error be monitored _via_ internal as well as external quality control procedures. It has already been shown that such control is not easy -- _i.e._, in Fig. 2 (and Ref. 38) it was shown that more than 10 and nearly 50 observations are required just to detect systematic or additional random error, respectively equal in magnitude to the Poisson component. The alternative of <u>substituting</u> s^2_B for the Poisson estimate for the assessment S_C and x_D has some merit; but, for a number of reasons we recommend using it (s^2_B) rather as a measure of control. [See notes A1 and A2.] What has been recommended (preceding section) to cover the possibility of non-Poisson error is provision of a relative systematic-error bound $\not\!{\Delta}_B$.

In less-controlled environments, rather severe excursions in B and in its variability may take place. If B comes from contamination in sampling and analysis (reagent), its distribution function -- which is crucial for estimating detection limits -- may be derived from both normal (or approximately constant "offset") and log-normal components (Ref. 95), in which case a large suite of genuine blanks is a prerequisite to x_D estimation. In the worst of circumstances B fluctuations may be wild and non-random. In this case there is no substitute for experienced, "expert judgment" as to maximum non-significant excursions. (Modern statistical tools, such as Exploratory Data Analysis (Ref. 96) would make superb partners for "expert judgment" in these cases.) Formally, this could correspond to substitution of a site-specific, realistic value of $\not\!{\Delta}_B$, in

place of our suggested default value (0.05). One situation in which such relatively severe fluctuations might be expected would be continuous monitors (count rate meters - analog or digital) for effluent noble gases.

Model error, such as deviations of baselines from single functional shapes (linear, quadratic, ...) or incorrect components or peak shapes when fitting complex multiplets or spectra, constitutes another source of B-error. Here, the "B" involved actually is interference, and the problem is that high levels of interfer·ing activities can cause serious deviations from our assumed B (e.g., baseline) uncertainties and, hence, estimated detection limits. Our default value $\delta_I = 0.01$ is intended to provide some protection. Some discussion and illustration of this potentially complex issue is given in section III and Ref. 72.

Before leaving the topic of the Blank, let us consider some regions of validity in relation to 3 types of effects on the detection limit. Two of these have been noted already: systematic error (<u>via</u> δ_B) and normally-distributed random error (<u>via</u> σ_0). (See Eq. 15.) The third, of major concern in extreme low-level counting is Poisson effect, <u>viz.</u> Poisson deviations from Normality. For "simple counting" (gross signal minus background) this (Poisson effect) adds a term $z^2 = 2.71$ to the parenthetical quantity in the numerator of Eq. 15. (For the lowest level counting, where B = 0, Eq. 15 must be replaced with an exact Poisson treatment. (See section III.C.1.) Taking σ_0 equal to $\sigma_B = \sqrt{B}$ for the "well-known" blank case, and $\delta_B = 0.05$, we can directly compare the three terms which delimit the detection of net signals (units:counts):

[systematic] term-1: $2\delta_B B = 0.10\ B$ (counts)

[conventional] term-2: $3.29\ \sigma_0 = 3.29\ \sqrt{B}$ (counts)

[Poisson] term-3: $z^2 = 2.71$ (counts)

Two types of question interest us: (1) the cross-over points where each term becomes predominant, and (2) the points (B-magnitudes) by which the unconventional" terms-1 and -3 are negligible. For question (1), we set adjacent terms equal and solve for B; for question (2) we define negligible as 10% relative. The results:

   term-1 < term-2 for B < 1082 counts

   term-3 < term-2 for B > 0.68 counts[1]

Thus, the conventional, approximately Normal Poisson expression (term-2) predominates for roughly 1 to 1000 background counts observed. (For interference, substituting ϕ_I - 0.01 for ϕ_B, the upper limit is increased to about 27,000 counts.

Terms-1 and -3 are not so easily ignored, however. The systematic error term-1 exceeds 10% of term-2, for B > 10.8 counts; and the extra Poisson term-3 exceeds 10% of term-2 for B < 67.6 counts. Thus, Eq's (15) and (18) were recommended for use when B $\gtrsim$ 70 counts. (The above regions of validity apply strictly to the very common simple-counting, well-known blank case. Somewhat altered values come about when x is estimated from single or multicomponent least squares deconvolution.) (See also note B9 for a discussion of the approximation $\sigma_B \approx \sqrt{B}$.)

2. Deduction of Signal Detection Limits for Specific Counting Techniques

The concentration detection limit x_D or LLD can be expressed as (see Eq's (13) and (15)

$$x_D = const. \; BEA + const' \cdot S_D^o/(YEVT) \qquad (19)$$

----
[1] It is interesting to consider the exact Poisson treatment in this case. Using Table 7 in section III.C.1 we calculate a detection limit (S_D) of 5.63 counts, whereas the sum of terms-2 and -3 gives 5.42 counts.

where the first term relates purely to systematic uncertainty (error bounds) and both constants include the calibration systematic error factor f. S_D^o is the signal detection limit taking into account <u>random</u> error only. Apart from BEA, the LLD is controlled by the nature of the counting process (including the data reduction algorithm) as reflected in the random error-controlled quantity S_D^o and the calibration factors Y,E,V,T. In this subsection we shall consider the dependence of the all-important quantity S_D^o on the nature of the counting process. The calibration factors will be discussed in the following subsection on design.

Signal decision (critical) levels and detection limits were given in Eq's (1) and (2).

$$S_C^o = z_{1-\alpha}\, \sigma_o = 1.645\, \sigma_o \qquad\qquad (1)$$

$$S_D^o = S_C^o + z_{1-\beta}\, \sigma_D = 1.645\,(\sigma_o + \sigma_D) \qquad\qquad (2')$$

(A prime has been placed on Eq. (2) because we do not wish to restrict ourselves to the assumption that $\sigma_o = \sigma_D$ at this point.) The crucial quantities governing the signal detection limit are thus σ_o and σ_D -- the standard deviations of the estimated net signal ($\hat{S}$) when its true value is $S = 0$ and $S = S_D^o$. These are what we shall relate to the counting system. What follows is simply a concise summary for different systems of importance. Derivations and detailed expositions are to be found in section III.C. (Note that in the remainder of this section, since we shall refer strictly to the random error component, we shall omit the superscript - zero on S_C and S_D -- for ease of presentation. Also, $\alpha = \beta = 0.05$, so $z = 1.645$.)

a) <u>Meaning of</u> σ_o and σ_D. These quantities are central to the entire discussion. Let us therefore consider their definitions in terms of the observations (gross counts) y_1 and y_2, for "simple counting."

$$y_1 = S + B + e_1 \quad \text{(gross signal)} \tag{20}$$

$$y_2 = bB + e_2 \quad \text{(blank)} \tag{21}$$

(In Eq 21, one can envisage y_2 as the sum of b - measurements of the blank, so y_2/b equals the average observed blank.)

The estimated net signal is

$$\hat{S} = c_1 y_1 + c_2 y_2 = (1)y_1 - \left(\frac{1}{B}\right) y_2 \qquad . \tag{22}$$

(The coefficients c_i are introduced for later generalization.) Following the rules of error propagation, and using $V = \sigma^2$,

$$V_S = \Sigma c_i^2 \, V_{y_i} = V_1 + \left(\frac{1}{B}\right)^2 V_2 \tag{23}$$

When $S = 0$, $V_1 = V_B$; and $V_2 = bV_B$. Thus,

$$V_O = V_B + \left(\frac{1}{b}\right)^2 (bV_B) = V_B \left(1 + \frac{1}{b}\right) = V_B \eta \tag{24}$$

(Equation (24) <u>defines</u> the coefficient η.)

when $S = S_D$, $V_1 = V_{S_D} + B$ which may or may not differ from V_B in the most general case (e.g., <u>non</u>-counting systems, or systems where non-Poisson variations dominate). Thus, for variance which is relatively independent of signal amplitude, $V_1 = \text{const} = V_B$, so $V_D = V_O$. It follows, in this case that

$$S_C = 1.645 \, \sigma_O = 1.645 \, \sigma_B \sqrt{\eta} \tag{25}$$

$$S_D = S_C + 1.645 \, \sigma_D = 2 \, S_C \tag{26}$$

(Thus far we have said nothing about Poisson counting statistics. That will follow shortly.)

First, an important generalization: If we consider a rather more complicated measurement scheme (e.g., decay curve and/or γ-spectrum fitting by _linear_ least squares),

$$y_i = \Sigma \, a_{ij} S_j + B_i + e_i \qquad (27)$$

the solution to Eq. (27) is of the form (see section III.C.3),

$$S_j = \Sigma \, c_{ji} \, y_i \qquad (22')$$

or, denoting the component of interest as S_1 (or simply S) and the respective coefficients as c_{1i} (or simply c_i) we write

$$\hat{S} = \Sigma \, c_i \, y_i$$

just like Eq. (22). Therefore,

$$V_S = \Sigma \, c_i^2 \, V_i \qquad (23')$$

just like Eq. (23). Knowing the least squares coefficients (c_i) and the variances (V_i) of the observations (y_i), we can proceed according to exactly the sample principles developed for "simple counting." (Admittedly, non-trivial issues must be dealt with concerning Poisson statistics, identity and amplitudes of interfering components (S_j for $j \neq 1$), and possible semi-empirical shape functions for fitting the baseline b_i. Such complications will be treated in part below and in part in section III.C.)

In any case, Eq's (25) and (26) are the most important results of this introductory section. The signal detection limit is seen to be directly proportional to the standard deviation of the blank, where the constant of proportionality (for simple counting) is $3.29 \sqrt{\eta}$. The dimensionless quantity η depends on the relative amount of effort (replicate measurements, counting time) involved in estimating the mean value of the blank. The bounds for η are clearly 1 and 2 (taking $b \geq 1$).

b) Use of replication (s^2) and Student's-t. We have an enormous advantage but a subtle trap as a result of Poisson counting statistics. σ_B and σ_D can be estimated directly from the respective number of gross counts. The trap is that other sources of random error may be operating [Ref. 20].

One solution to this problem is to substitute $t_\nu s_B$ for 1.645 σ_B in Eq. (23), where t_ν is Student's-t (also at the $1-\alpha = 0.95$ significance level) for ν-degrees of freedom. ($\nu=b-1$ according to the convention of Eq. 21) s_B is the square root of the estimated blank variance, i.e.,

$$s_B^2 = \sum_1^n \frac{(B_i - B)^2}{n-1} \qquad (28)$$

where, for our example, $n = b$.

We strongly recommend the routine calculation of s_B as a control for the anticipated Poisson value, $\sqrt{B}$. If non-Poisson Normal, random error predominates and is well understood and in control, then it is appropriate to adopt $t_\nu s_B$ in place of 1.645 $\sqrt{B}$. Unless this is assured, blithe application of $t_\nu s_B$ could be foolhardy, for Eq. (28) will give a numerical value even if the blank is non-Normal or not in control. Further, information which can be deduced using Poisson statistics (e.g., from Eq's (22′) and (23′)) is generally far more general and more precise than what can be derived from a reasonable number of replicates. [For more on this topic, including the analogue of Eq. (26) under replication, see notes A1, A2, and B2.]

c) Simple Counting -- Poisson Statistics. If there are at least several blank counts expected ($B \gtrsim 5$), substitution of the Poisson variances for V_1 and V_2 at $S = 0$ and $S = S_D$ give a valid solutions:

$$S_C = 1.645 \sqrt{B\eta} = 1.645 \sqrt{B\left(\frac{b+1}{b}\right)} \qquad (25′)$$

$$S_D = z^2 + 2 S_C = 2.71 + 2 S_C \qquad (26′)$$

The constant z^2 in Eq. (24´) comes directly from Poisson statistics and the fact that $\sigma_D > \sigma_O$ [Ref. 5]. Thus, it is evident that the detection limit remains finite even with a zero blank.

d) <u>Multicomponent Counting</u>. When there are two or more mutually interfering species, σ_O and σ_D are not so easily expressed. More detail on these topics will be found in section III.C, but two of the results will be highlighted here.

For two mutually intefering components, where a solution is given by simultaneous equations or linear least squares, it can be shown that

$$S_C = 1.645 \sqrt{B\eta} \quad [\eta > 1] \tag{25´´}$$

$$S_D = z^2 \mu + 2 S_C \quad [\mu > 1] \tag{26´´}$$

where, now, B, η, and μ depend on the specific set of equations defining the observations in relation to the net signal of interest. ("S" and "B" remain useful and even meaningful labels for the components when there are only two.) These more general relations show that a universal consequence of Poisson statistics is the inequality: $S_D/S_C > 2$. Equality is approached, however, for simple counting when $B \geq 70$ counts.

For multiple interference, a closed (analytic) solution for S_D cannot be given. One must return to the original defintions, Eq's (1) and (2´), and tentatively estimate the corresponding σ's from the appropriate diagonal elements of the inverse least-squares (variance-covariance) matrix. (Non-~~linear~~ fitting introduces some rather peculiar problems. See section III.)

Fortunately, a limiting calculation for x_D, which derives from non-negatively ($S \geq 0$), can be made for any specific result ($\hat{x}$, σ_x) of multi-component analysis. Through the use of <u>Inequality Relations</u> ($\sigma_x \geq \sigma_O$, etc.) upper bounds for the critical level and detection limit can be immediately derived. (See note B4.)

50

A very significant point with respect to these more complicated, multicomponent cases is <u>algorithm dependence</u>. (See section III.) That is the particular data reduction algorithm (model and channels used for peak and baseline estimation; assumed number and type of interfering species, etc.) determines σ_0 and σ_D, and therefore the detection limit.

e) <u>Continuous Monitors</u>. Both analog and digital monitors are used for continuous monitoring in nuclear plants. As noted already in section II.D.1, one must be cautious in applying Poisson statistics in uncontrolled environments. Some basic information on the statistics of such count rate meters is given, however, in Evans (Ref. 74) and more recent publications such as Ref. 73. Some of this has been covered also in section III of this report. Two basic limiting relations, for example, are:

$$\sigma_R^2 = R/t \quad \text{if } t \gg \tau \tag{29}$$

$$\sigma_R^2 = R/2\tau \quad \text{if } t \ll \tau \tag{30}$$

where R refers to count rate, <u>t</u> to the averaging time, and <u>τ</u> to the time constant for an analog circuit. Applications of the relations for long-term (Eq. 29) and instantaneous (Eq. 30) measurements are treated in section III. (See also note B7.)

f) <u>Extreme Low-Level Counting</u>. When the expected number of blank counts for a sample measurement is less than about 5 it is advisable to use the exact Poisson distribution for making detection decisions and setting detection limits. (So long as the constant term z^2 is kept in the expression for simple counting [Eq. 26´], this gives a reasonable approximation even down to $E(y_1) = 1.$ count -- see section III.D.1.) Although treatments have been given where both gross signal (y_1) and blank (y_2) observations contain few if any counts (Ref. 36, 75), we recommend the MP be designed so that a

51

reasonably precise estimate be available for B. The expected number of blank counts in the 'blank' experiment ($y_2 = bB$), for example, should exceed 100, if possible.

In that case, a simple reduced activity diagram (Fig. 7) can be used to instantly determine S_C and the detection limit (in units of BEA) [Ref. 19]. A complete treatment of this subject is given in section III.C.1.

3. Design and Optimization

We consider briefly the question of experiment (i.e., MP) design because this is the very question one faces when attempting to alter the adjustable experimental variables in order to meet RETS requirements. The task is to bring about the condition,

$$x_D \leq x_R \qquad (9)$$

Optimization differs from design (in general) in that we adjust the variables to minimize x_D rather than simply to satisfy the inequality, Eq. (9). Design and optimization are literally multidimensional operations when one treats a multicomponent system with interfering spectra and/or decay curves and the possibility of different schemes of multiple time and multiple energy band observation. It is well beyond the scope of this manual.

For rather simpler systems, however, we can consider design from the perspective of Equations 15, 18, and 26'.

$$x_D = \left(\frac{f}{2.22 \ YEV} \right) \left(\frac{2.71 + 3.29 \ \sqrt{R_B nt} + 0.10 \ R_B t}{\tau \ [1 - e^{-t/\tau}]} \right) \qquad (31)$$

That is,

$$x_D \propto \left(\frac{1}{YEV} \right) \left(\frac{C_1 + C_2 \ \sqrt{R_B t n} + C_3 [R_B t]}{1 - e^{-t/\tau}} \right) \qquad (32)$$

Eq. (32) has been cast, of course, to highlight the controllable variables: Y, E, V and t. (Note that $\tau = 1/\lambda$ = mean life.) Since the effects of these variables fall in two categories we shall treat each of the two main factors in Eq. 32 separately.

a) Proportionate Factors, $\underline{YEV}$. x_D decreases directly with each of these factors, so a requisite proportionate decrease to meet the prescribed LLD (i.e., x_R) can be achieved (in principle) by a corresponding reduction in any one of them or in their product.

The factor most readily available is $\underline{V}$, for this is a measure of the sample size taken. In certain situations, it may have reached an upper limit for various practical reasons, the most common of which is the size that the nuclear detector can accomodate. If the amount of sample (or disappearance through rapid decay) is not limiting, V may be effectively increased further through concentration and/or radiochemical separation. If such steps are too labor intensive, alternative approaches may be preferred. In general, however, because of its controllability and the inverse proportionality between x_D and V, this quantity provides the greatest leverage.

Y cannot exceed unity. In the absence of sample preparation steps, it is not even a relevant variable. The most important circumstances arise when Y is quite small; major improvements in procedures having poor recovery could have some impact.

The detection efficiency $\underline{E}$ is a complex factor. Changes possibly at our disposal include geometry, external or self-absorption (or quenching in the case of liquid scintillation counting), and the selection of nuclear particle or Y-ray to be measured.

Some effects are dictated by Nature, however. Most noteworthy is the decay scheme, especially branching ratios (or γ-abundances, etc.). Other things being equal, the LLD achievable -- i.e., x_D -- will vary inversely with the particle or γ-abundance of the radiation being measured. If nuclides having low γ-abundances are to achieve the same LLD's as those with high abundances, other factors will have to be accordingly adjusted.

Note that the effective detection efficiency may depend also on the data reduction algorithm -- e.g., fraction of a γ-spectrum used for radionuclide estimation. More efficient numerical information extraction schemes may thus be beneficial.

b) <u>Background (Blank) Rate; Counting Time</u>. It is clear from the numerator of the second factor in Eq. 32 that decreasing the background rate will decrease LLD up to a point. If t is fixed (say, the maximum feasible) then once the first (extreme Poisson) term C_1 predominates, further reduction in the background (or blank or interference) will have little effect. In contrast if B is so large that the third (systematic-error) term C_3B predominates, then B - reductions will have as large an effect as proportionate increases in V and E. In section II.D.1, we saw (for typical MP parameter values) that the B - transition points occurred at about 1 count and 1000 counts. Perhaps the most important opportunity for B-reduction occurs when it is due to large amounts of interfering nuclides which can be eliminated by decay or radiochemical separation.

A second quantity at our disposal is η. This depends on the amount of time or channels (for a simple peak) used for estimating B for simple counting. In more complex (multicomponent) situations, the data reduction algorithm (as embodied in η) will have some effect on x_D.

The major and most commonly considered variable is counting time. It is interesting here to consider two extremes for the factor in the denominator, $(1-e^{-t/\tau})$. [τ represents the mean life, $t_{1/2}/\ln 2$]. If $t \ll \tau$ this factor $\propto t$. At the other extreme ($t \gg \tau$) it approaches a constant (one). We can represent the situation in two dimensions as follows:

Table 5. LLD (x_D) Variations with B and t[a]

	$t \ll \tau$	$t \gg \tau$
$B \lesssim 1$	$x_D \propto t^{-1}$	$x_D = const$
$1 \lesssim B \lesssim 1000$[b]	$x_D \propto t^{1/2}$	$x_D \propto t^{1/2}$
$B \gtrsim 1000$[b]	$x_D = const$	$x_D \propto t$

a) Units for B are counts. τ equals the mean life ($t_{1/2}/\ln 2$).
b) For $\Delta_I = 0.01$, the upper crossing point changes from ~1000 to ~27000 counts.

Depending on which domain of B and t we are in, it is clear that increases in counting time may decrease x_D, have no effect, or at worst increase x_D. Also, it is interesting that in the region of extremely small B, all increases in t will be beneficial; in fact, the initial variation (if $t \ll \tau$) will be proportionate. (Admittedly, for fixed R_B, increased t will tend to move B out of the extreme Poisson region. However, if the expected value of B is significantly smaller than 1 count, increases in t can be of major advantage if one is measuring long-lived activity.)

When B is already quite large, increase in t can only make matters worse. The intermediate region is intriguing. Here ($1 \leq B \leq 1000$ counts) "conventional" counting statistics predominate; and for fixed R_D, x_D decreases with increased counting time for long-lived activity but reverses itself for

short-lived activity. Obviously there must be an optimum. Differentiating the appropriate term in Eq. (32) shows that optimum to be the solution of the transcendental equation.

$$t = \left(\frac{1}{2}\right) \; (1-e^{-t})/e^{-t} \qquad\qquad (33)$$

where t is in units of the mean life τ. The solution to equation (33) gives the optimum counting time as ~1.8 times the half-life.

It is worthy of re-statement that (Eq. 32, Table 5):

o   Knowing the time and B-domains, one can quickly scale x_D according to the expected variation with time.

o   Diminishing returns for background reduction set in when the term C_1 begins to dominate.

o   Diminished returns for LLD (x_D) reduction by extended counting set in once (a) t > 1.8 $t_{1/2}$ or (b) B > η $(z/\Delta_B)^2$ which equals 1082 and 27060 counts for the default values taken for blank and baseline relative systematic error bounds. (This latter statement is equivalent to indicating ~2% and ~11% of the BEA as minimum achievable bounds for x_D.)

o   A rapid graphical approach for experiment planning, for all 3 B-domains can be given in the form of the "Reduced Activity Diagram." Space does not permit an exposition on this topic, but one such diagram (for extreme low-level counting) is included as Fig. 7. Other diagrams for higher activity levels and including the effects of non-Poisson error may be found in Ref's 62 and 80.

4. Quality

a) <u>Communication</u>. Free and accurate exchange of information is one crucial link for assuring the quality of an MP and the evaluation of the consequent data. A few highlights in this area, relevant to LLD and RETS follow.

[i] <u>Mixed Nuclide Measurements</u>. Interpretation of non-specific radionuclide measurements is seldom possible unless the average temporal and detector responses are fixed. Calibrations and measurements of gross nuclide mixtures require controls on the relative amounts of nuclides having different half-lives and different detector responses for meaningful interpretation.

[ii] <u>"Black boxes" and Automatic Data Reduction</u>. One of the dis-benefits of automated data acquisition and evaluation is lack of information on source code or detailed algorithms employed, specific nuclear parameter values stored, and artificial thresholds and internal "data massaging" routines. A number of surprises and blunders could be prevented if there were adequate open communication in this area. One problem of hidden algorithms which can be especially troublesome for detection decisions and limits (as well as for quantification) is intentional (but unknown to the user) algorithm switching. A potential means of control for these kinds of problems is the use of artificial (known) reference data sets as distributed by the IAEA [Ref 81]. (Further comments on this are given below.)

[iii] <u>Reporting Without Loss of Information</u>. The following paragraphs and Figures are adapted from [Ref. 38].

"Quality data, poorly reported, leads to needless information loss. This is especially true at the trace level, where results frequently hover about the limit of detection. In particular, reports of upper limits or "not

57

detected" can mask important information, make intercomparison impossible, and even produce bias in an overall data set. An example is given in Fig. 4 which relates to a very difficult radioanalytic problem involving fission products in seawater (97). In this example, only six of the fifteen results could be compared and only eight could be used to calculate a mean. Since negative estimates were concealed by "ND" and "<", the mean was necessarily positively biased. (The true value τ in this exercise was, in fact, essentially zero; and the use of a robust estimator, the median [m] does give a consistent estimate.) Although upper limits convey more information than "ND", authors choose conventions ranging from the (possibly negative) estimated mean $(\hat{x})$ plus one standard error to some sort of fixed "detection limit." Such differences are manifest when one finds variable upper limits from one laboratory but constant upper limits from another (98).

The solution to the trace level reporting dilemma is to record all relevant information, including as a minimum: the number of observations (when pertinent), the estimated value $\hat{x}$ (even if it is negative!) and its standard deviation, and meaningful bounds for systematic error. More thorough treatments of this issue may be found in Eisenhart (99) and Fennel and West (100)."

When information is not fully preserved for a set of marginally detected results, distributional information and parameters may be recovered by statistical techniques (probability plots; maximum-likelihood estimates) which have been developed for "censored" data. [Ref. 48,69,82,91]. By "censored" we mean that although numerical results of some of the data may not be preserved, the number of such results is recorded. Though such techniques permit the partial recovery of information from censored data sets, they cannot fully compensate for such information loss.

$$\chi/\sqrt{\nu}$$

P = .975

P = .025

DEGREES OF FREEDOM (ν)
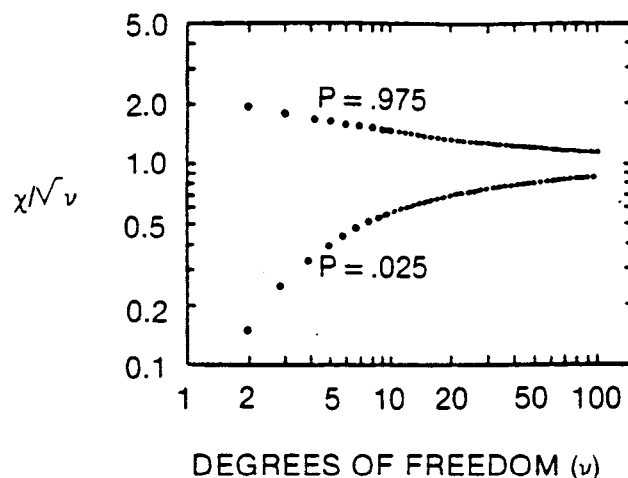
DATA		X ± SE	HISTOGRAM
<.10	ND		80|4
<7.30	ND		70|7
<40.00	9.50		60
<20.90 ±	<26.80		50 b^2
2.20 ±	0.30		40
9.20 ±	8.60		30
77.00 ±	11.00		20 |⊚
0.38 ±	0.23		10 |0]4
84.00 ±	7.00	—τ——m	0 |2 9 ⬚ 0
	14.1		ND|UNU[NUUU
0.44 ±	0.06		

Fig. 4a. $\chi/\sqrt{\nu}$ vs degrees of freedom. The curves enclose the 95% confidence interval for $\chi/\sqrt{\nu}$. They may be used for assessing the fit of single or multiple parameter models, and they give a direct indication of the precision of standard deviation estimates.

4b. Reporting deficiencies. International comparison of ^{95}Zr-^{95}Nb in sample SW-1-1 of seawater (pCi/kg). The symbols have the following meanings: τ = true value, X̄ = arithmetic mean (positive results), m = median (all results), and b^2 = a "double blunder" - i.e., inconsistent result 77 ± 11 was originally reported as 24. N and U indicate not detected, and upper limits, respectively.

So long as the full initial data are recorded and accessible, however, it may of course be reasonable to provide summary reports for special purposes which exclude tabulations of non-significant $\hat{x}$'s. But to set them all to either zero or to LLD guarantees confusion and biased averaging. The question of automated instrumentation and data reduction may again be involved here, if the "black box" does the censoring rather than the user.

b) <u>Monitoring (control)</u>. Three classes of control are considered important for reliable detection decisions and measurements in the region of the LLD. At the internal level it is crucial that blank variability be monitored by periodic measurements of replicates; similarly, complex fitting and/or interference (baseline) routines need to be regularly monitored by goodness-of-fit tests and residual analysis. If such tests do not indicate consistency with Poisson counting statistics, the simple substitution of s^2 or mutliplication by X^2/ν in place of the Poisson standard error is not generally recommended. It could mask assumption or model error unless that possibility has been carefully ruled out [Ref. 63]. Resulting LLD estimates could thereby be quite in error.

Reference samples, internal and external, blind and known, are crucial for maintaining accuracy and exposing unsuspected MP problems. "Blind splits" and the EPA Cross-Check samples thus serve a very important need. The utility of external quality control samples is highest, of course, when such samples resemble "real" samples as closely as possible in their nuclear and chemical properties, when their true values are known (to the distributors), and when they are really "blind" from the perspective of the laboratory wishing to maintain its quality. In connection with the LLD it might really be valuable to purposely monitor (internally and/or externally)

performance at this level -- i.e., to provide blind samples containing blanks and radionuclides in the neighborhood of the prescribed LLDs.

A third class of control relates to the data evaluation phase of the MP. The presumption that control is quite unnecessary for this step was belied by the IAEA Y-ray spectrum intercomparison study referred to earlier. A summary of the structure and outcome of that exercise (adapted from Ref. 38) follows.

"One of the most revealing tests of Y-ray peak evaluation algorithms was undertaken by the International Atomic Energy Agency (IAEA) in 1977. In this exercise, some 200 participants including this author were invited to apply their methods for peak estimation, detection and resolution to a simulated data set constructed by the IAEA. The basis for the data were actual Ge(Li) Y-ray observations made at high precision. Following this, the intercomparison organizers systematically altered peak positions and intensities, added known replicate Poisson random errors, created a set of marginally detectable peaks, and prepared one spectrum comprising nine doublets. The advantage was that the "truth was known" (to the IAEA), so the exercise provided an authentic test of precision and accuracy of the crucial evaluation step of the CMP.

"Standard, doublet and peak detection spectra (Fig. 5) were provided; Fig. 6 summarizes the results (81,92). While most participants were able to produce results for the six replicates of 22 easily detectable single peaks, less than half of them provided reliable uncertainty estimates. Two-thirds of the participants attacked the problem of doublet resolution, but only 23% were able to provide a result for the most difficult case. (Accuracy assessment for the doublet results was not even attempted by the IAEA because of the unreliability of participants' uncertainty estimates!) Of special

61
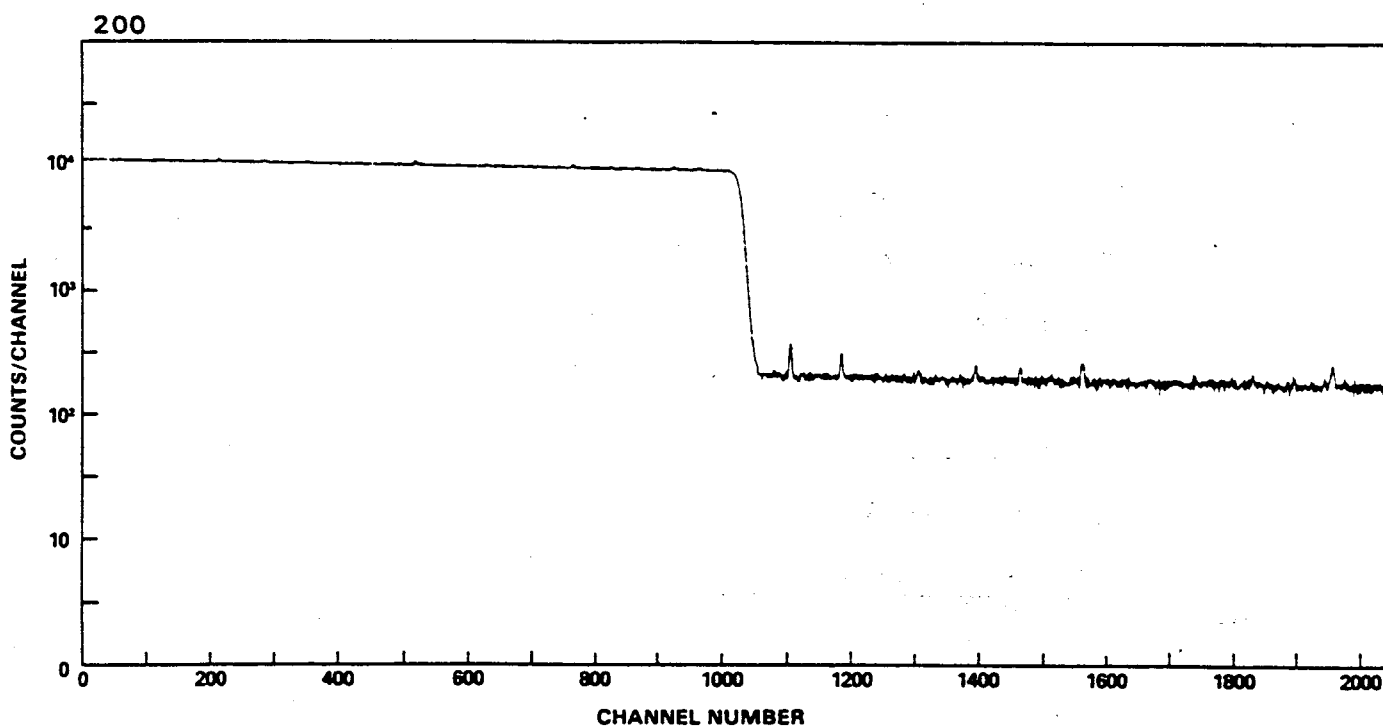
Fig. 5. IAEA test spectrum for peak detection

DATA EVALUATION--IAEA
γ-RAY INTERCOMPARISON

[Parr, Houtermans, Schaerf · 1979]

Peaks	Participants	Observations
22 · Singlets (m = 6)	205/212	• uncertainties: 41% (none), +17% (inaccurate)
9 · Doublets	144/212	• most difficult (1:10, 1 ch.) --49 results
22 · Subliminal	192/212	• correctly detected: 2 to 19 peaks • false positives: 0 to 23 peaks • best methods: visual (19), 2nd deriv. (18), cross correl. (17)

Fig. 6. Data evaluation - IAEA Y-ray intercomparison. Column two indicates the fraction of the participants reporting on the six replicates for 22 single peaks, 9 overlapping peaks, and 22 barely detectable peaks. Column three summarizes the results, showing (a) the percent of participants giving inadequate uncertainty estimates, (b) the number of results for the doublet having a 1:10 peak ratio with a 1 channel separation, and (c) the results of the peak detection exercise.

import from the point of view of trace analysis, however, was the outcome for the peak detection exercise. The results were surprising: of the 22 subliminal peaks, the number correctly detected ranged from 2 to 19. Most participants reported at most one spurious peak, but large numbers of false positives did occur, ranging up to 23! Considering the modeling and computational power available today, it was most interesting that the best peak detection performance was given by the 'trained eye' (visual method)."

5. Multiple Detection Decisions

It follows obviously that if all radionuclides measured are present either not at all (H_O) or at the LLD (H_D) and the errors α and β are each set at 5%, then 5% of the detection decisions will be wrong "in the long run." Thus, for example, in a γ-ray spectrum containing no radionuclides, if one were to examine say 200 locations for the possible presence of radionuclides, 10 false positives (on the average) would result. This carries some curious implications for any instructions to "report any activity detected" -- especially if one multiplies the 10 false positives by the number of spectra examined, for example in a Quarter. (One may find an apparently tighter constraint in a phrase such as "detected and identified," but this would require a second manual to struggle with a rigorous meaning for the term "identified" in such a context!)

If the number of nuclides sought is restricted purely to the "principal radionuclides," the situation is altered numerically but not qualitatively. If there were just one sample per month and 10 nuclides sought in each sample, we would expect after 1 year (or 12 samples) - 6 false positives (if there were in fact no activity).

Solutions to this dilemma are either to accept an error rate of 5% false positive or false negative results, or to redefine the goal such that there is only a 5% chance of getting a single false positive given the entire set of measurements. (This seems the only rational alternative when scanning a high resolution spectrum for the unsuspected tiny peaks.) The critical level must be correspondingly increased and with it, the detection limit. (If one were to assume some prior unequal apportionment of the samples to hypotheses H_O and H_D, the increases in S_C and S_D could differ substantially from one another, but we shall not treat this case.)

To address this matter explicitly, let us assume that N decisions (ergo, measurements) are made all at risk-level α'. The probability that none is incorrect can be given by the Binomial distribution:

$$\text{Prob } (0) = \binom{N}{0} (\alpha')^0 (1-\alpha')^N = (1-\alpha')^N \qquad (34)$$

The probability that no decision is incorrect is by definition $1-\alpha$, where α is the risk or probability that 1 or more is incorrect. Therefore, the α' we need to impose on each decision is

$$\alpha' = 1 - (1-\alpha)^{1/N} \approx \alpha/N \qquad (35)$$

for small α. If $N=100$, for example, and α remains 0.05, then

$$\alpha' = 1-(0.95)^{0.01} \approx 0.000513$$

If Normality could be assumed so far out on the tail of the distribution, $z_{1-\alpha'} = 3.27$. Treating β' in the same way, we would conclude that decision levels and detection limits would each need to be increased by about a factor of two (from 1.645).

A somewhat related issue involving the question of reporting non-principal radionuclides if detected is illustrated by result $\hat{x}_b$ in Fig. 3. Here an observation brings the decision "detected" and the <u>actual</u> LLD (x_D) is below the prescribed LLD (x_R). (Also, as shown, its upper limit as well lies

below x_R.) What follows is that unless there is truly zero activity in a set of samples examined, that the more sensitive MP's (lower x_D's) will "detect" more radionuclides even though they may be well below the prescribed LLD (x_R) if any.

III. PROPOSED APPLICATION TO RADIOLOGICAL EFFLUENT TECHNICAL SPECIFICATIONS

   (RETS)[1]

A. Lower Limit of Detection - Basic Formulation

1. Definition

The LLD is defined, for purposes of these specifications, as the smallest concentration of radioactive material in a sample that will yield a net count, above the measurement process (MP) blank, that will be detected with at least 95% probability with no greater than a 5% probability of falsely concluding that a blank observation represents a "real" signal. "Blank" in this context means (the effects of) everything apart from the signal sought -- i.e., background, contamination, and all interfering radionuclides.

For a particular measurement system, which may include radiochemical separation:

The Lower Limit of Detection is expressed in terms of radioactivity concentration (pCi per gram or liter [A3]); it refers to the a priori [A4] detection capability

$$LLD = \frac{f(2\Delta+[z_{1-\alpha}+z_{1-\beta}]\sigma_o)}{2.22 \ (YEV)T} = \frac{fS_D}{A} = x_D \qquad (1)$$

The detection decision is based on the observed net signal $\hat{S}$ (a posteriori [A4]) in comparison to the critical level (counts):

$$S_C = \Delta + z_{1-\alpha}\sigma_o \qquad (2)$$

where the "statistical" part of the definitions (when $f = 1$, $\Delta = 0$) sets the false positive and false negative risks at α and β, respectively [A5].

Meanings of the symbols follows. (See also App. A).

---

[1]Parts A and B of Section III represent proposed substitute RETS pages. Part A is the more comprehensive, and it is framed in a manner that should be applicable to most counting situations. Part B is offered as a simplified version, which will suffice for "simple" gross signal-minus-background measurements.

67

$\underline{A}$ is the overall calibration factor, transforming counts to pCi/g (or pCi/L).

$\underline{E}$ is the overall counting efficiency, as counts per disintegration; it comprises factors for solid angle, absorption and scattering, detector efficiency, branching ratios and even data reduction algorithm [A6, A7],

$\underline{V}$ is the sample size in units of mass or volume,

$\underline{2.22}$ is the number of disintegrations per minute per picocurie [A3],

$\underline{Y}$ is the fractional radiochemical yield, when applicable.

σ_0 is the Poisson standard deviation of the estimated net counts ($\hat{S}$) when the true value of S equals zero (i.e., a blank). (The relation of σ_0 to the background or baseline depends upon the exact mode of data reduction [see section III.C.3].)

$\underline{z_{1-\alpha}, z_{1-\beta}}$ = the critical values of the standard normal variate -- taking on the value 1.645 for 5% risks (underline one-sided) of false positives (α) and false negatives (β), when underline single detection decisions are made. (Multiple detection decisions require inflated values for $z_{1-\alpha}$ to prevent significant occurrence of spurious peaks -- as in high resolution γ-ray spectroscopy.) When α, β risks are equal, and systematic error negligible, the detection limit for net counts, S_D, is just twice S_C. (Assumes the Normal Limit for Poisson counts.) (When subscripts are omitted in the following text z will denote $z_{0.95}$ = 1.645).

$\underline{T}$ = the effective counting time, or decay function, to convert counts to initial counting rate (time "zero": end of sampling) [A9]. It is numerically equal to $(e^{-\lambda t_a} - e^{-\lambda t_b})/\lambda$, where t_a and t_b are the initial and final times (of the measurement interval) and λ, the decay constant. For $\lambda t \ll 1$, $T \to \Delta t = t_b - t_a$. [Multicomponent decay curve analysis yields a more complicated

expression for T -- and generally o_O/T, the standard deviation of the estimated initial <u>rate</u> is given directly.] (T must have units of minutes, for LLD to be expressed in pCi.] [A3,A6,A7,A9].

<u>f</u> and <u>Δ</u> are proportionate and additive parameters which represent bounds for systematic and non-Poisson random error. (The only totally acceptable alternative to this is <u>complete</u> replication of the entire measurement process (including recalibration, e.g., for <u>every</u> sample measured) and making several replicate measurements of the blank for each mixture of interfering nuclides and counting time under consideration [A10].)

<u>f</u> will be set equal to 1.1, to make allowance for up to a 10% systematic error in the denominator <u>A</u> of Eq. (1) --- viz., in the estimate of the product EVY [A11]. [If there are large <u>random</u> variations in A then full replication should be considered together with the use of $\hat{x}$ (radioactivity concentration) and o_X.] Note that <u>A</u> is equivalent to the slope of the calibration curve. If the curve deviates from linearity (e.g., -- due to saturation effects, algorithm deficiencies or changes with counting rate, signal amplitude, etc.) a more complex model and expression for LLD may be required.

<u>Δ</u> will be set equal to 5% of the blank counts plus 1% of the total interference counts (baseline minus blank) in order to give some protection against non-Poisson random or systematic error in the (assumed) magnitude of the blank (Ref's 20,72) [A11].

$\underline{S_D}$ is the detection limit expressed in terms of counts.

$\underline{S_C}$ (Eq. 2) is the critical number of counts for making the (a posteriori) Detection Decision, with false positive risk-α.

LLD (Eq. 1) is the lower limit of detection (radioactivity concentration), given the decision criterion of Eq. 2 (and risk-α), where the false negative risk (failing to detect a real signal) is β, [a priori]. The symbol x_D is used synonymously with LLD for later algebraic convenience.

S_C is applied to the observed net signal (units are counts) [A12]; whereas LLD refers to the smallest observable (detectable) concentration (units are disintegrations per unit time per unit volume or mass). LLD is without meaning unless the decision rule (S_C) is defined and applied [A13].

Bounds for systematic error in the blank (Δ, counts) and (relative) systematic error in the proportionate calibration factor (f) are included to prevent overly optimistic estimates of S_C or LLD based on extended counting times. Also, they take into account the possibility of systematic errors arising from the common practice of assuming simple models for peak baselines (linear or flat) and repeatedly using average values for blanks and calibration factors (Y,E). (Random calibration errors of course become systematic unless the system is recalibrated for each sample.) Inclusion of Δ and f in the equations for S_C and LLD converts the probability statements into inequalities. That is, $\alpha \leq 0.05$ and $\beta \leq 0.05$.

2. Tutorial Extensions and Notes

[A1]. Alternative Formulation in Terms of s_0. If the measurement process (including counting time, nature and levels of all interfering radionuclides, data reduction algorithm) is rigorously defined and under control, then it would be appropriate to replace $z_{1-\alpha}\sigma_0$ in Eq. (2) by ts_0, where t is Student's-t at the selected levels of significance (α, β) according to the number of degrees of freedom (df) accompanying the estimate s_0^2 of σ_0^2.

In this case, however, a small complication arises in calculating LLD, because S_D (detection limit in terms of counts) is approximately $2t\sigma_0$ (for $\alpha=\beta$) not $2ts_0$. A conservative approach would be to use the upper (95%) limit for σ_0 -- i.e. $s_0/\sqrt{F_L}$, where F_L is the lower (5%) limit for the distribution, χ^2/df. The recommended procedure is to use $z\sigma_0$ (Poisson) but to test the validity of the Poisson assumption through replication. (Ref. 20) [A2].

[A2]. <u>Uncertainty in the Detection Limit</u>. For reasonably well behaved systems, the critical level (S_C) which tests net <u>signals</u> for statistical significance can be fairly rigorously defined. (One needs a controlled MP and reliable functional and random error models.) The detection limit (radioactivity, trace element concentration, ...), however, requires knowledge of additional quantities which can only be estimated -- e.g., standard deviation of the blank, calibration factors, chemical recoveries, etc. Thus, although there exists a definite detection limit corresponding to the decision criterion (S_C or α) and the false negative error ($\beta = 0.05$), its exact magnitude may be unknown because of systematic and/or random error in these additional factors.

Two approaches may be taken to deal with this problem: (a) give an uncertainty interval for LLD, knowing that its true value (at $\beta = 0.05$) probably lies somewhere within (49) or (b) state the upper limit of the uncertainty interval as LLD, such that the false negative risk becomes an inequality -- i.e., $\beta \leq 0.05$. We prefer the latter procedure, because it provides a definite and conservative bound. Also, this is in keeping with the spirit of RETS, which simply requires that the actual LLD (x_D) not exceed the prescribed maximum (x_R).

One very important illustration of this matter arises in connection with signal detection limits based on replication. If the estimated net signal (when S=0) is normally distributed and sampled n-times (e.g., _via_ paired comparisons of appropriately selected blank pairs), the critical level is given by $t_{n-1}s/\sqrt{n}$, where _s_ is the square root of the estimated variance and t_{n-1} is Student's-t based on n-1 degrees of freedom. The minimum detectable signal is given by the non-central-t times the true (unknown) standard error. This is approximately $2t_{n-1}\sigma/\sqrt{n}$. Bounds for σ obtain from the χ^2 distribution: $(\chi^2/n-1)_{.05} \leq s^2/\sigma 2 \leq (\chi^2/n-1)_{.95}$. The upper bound for the signal detection limit ($\beta \leq 0.05$) would thus be .

$$[2t_{n-1}s/\sqrt{n}]\Big/[(\chi^2/n-1)_{.05}]^{1/2} \qquad (3)$$

For example, suppose that 10 replicate paired blank measurements were made, yielding a standard error $(s/\sqrt{10})$ _for the net signal_ (B_i-B_j) of 30 cpm. Then $t_9 = 1.83$ (for $\alpha = 0.05$) and $R_C = t_9 \cdot SE = 54.9$ cpm. Since $[\chi^2/9)_{.05}]^{1/2} = 0.607$, the upper bound for the detection limit would be higher by a factor of 2/0.607, or $R_D = 181.$ cpm. ($\beta \leq 0.05$). The total (90% CI) relative uncertainty for the standard error and hence R_D ($\beta = 0.05$) is given by the ratio of the upper and lower (.95, .05) bounds for $\sqrt{\chi^2}$, in this case (n = 10) equivalent to a factor of 2.26. To reduce the uncertainty in R_D to a factor of 2.00 (upper limit/lower limit) would require at least 13 replicates for the estimation of σ. [See Table 6 and Note B2.]

If, rather than paired replicates, a single sample measurement is to be compared with the estimated blank, and the latter is derived through replication,

$$\hat{\sigma}_0^2 = s_B^2 n = s_B^2 \left(\frac{n+1}{n}\right) \qquad (4)$$

Thus, the upper limit for S_D becomes

$$[2t_{n-1} s_B\sqrt{n}]/[(\chi^2/n-1)_{0.05}]^{1/2} = 2s_B\sqrt{n}[t(\sigma_{UL}/s_B)] \qquad (5)$$

72

[A3]. <u>S.I. Units.</u> The preferred (S.I.) unit for radioactivity is the Becquerel (Bq) which is defined as 1 disintegration/second (s^{-1}). To express LLD in units of Bq, the conversion factor 2.22 (dpm/pCi) in the denominator of Eq (1) would be replaced by 1. (dps/Bq) and the factor T would have units of seconds.

[A4]. <u>A priori (before the fact) and a posteriori (after the fact)</u> refer to the estimate $\hat{S}$ or $\hat{x}$ or decision process as the "fact." LLD is before the fact in that it does not depend on the specific (random) outcome of the MP. However, all parameters of the MP (including interference levels) must be known or estimated before "a priori" values for S_C or LLD (x_D) can be calculated. (Such parameters may be estimated from the results of the MP, itself, or they may be determined from a preliminary or "screening" experiment with the sample in question.)

[A5]. <u>Poisson Limit.</u> Equations (1) and (2) are valid only in the limit of large numbers of background or baseline counts. <u>If fewer than ~70 counts are obtained, special formulations are required to take into account deviations from normality.</u> (See section III.C.1 note B9, and Ref. 19). The simple sum in Eq (1) -- ($z_{1-\alpha} + z_{1-\beta}$) -- is an approximation; strictly valid only when $\sigma(\hat{S})$ is constant. This is a bad approximation for extreme low-level counting and for certain other measurement situations involving artificial thresholds (76).

[A6]. <u>Mixed Nuclides, Gross Counting.</u> For mixed, non-resolved radionuclides, where "gross" radiation measurements are made, the factors E and T are meaningful only if the particular mix (relative amounts and energies or half-lives) is specified. Common agreement on the radionuclides selected for efficiency calibration for "gross" counting is likewise mandatory.

[A7]. <u>For multicomponent spectroscopy and decay curve analysis</u>, the factors E and/or T are generally subsumed into the (computer-generated) expression for o_O, where o_O then has dimensions of disintegrations or (initial) counting rate or radioactivity (pCi or Bq). Both factors may thus depend upon the algorithm selected for data reduction -- i.e., the "information utilization efficiency" (see section III.C.3).

[A8]. <u>Formulation of the Basic Equations.</u> The expressions given for LLD and S_C are perfectly general, with one exception [A5], and intended to avert many pitfalls associated with errors in assumptions (non-Poisson random error, model error, systematic error, non-Normality from non-linear estimation) which can subvert the more familiar formulation. By formulating Eq's (1) and (2) in terms of o_O, we are able to apply them to all facets of radioactivity measurement, including the most intricate γ-ray spectrum deconvolution algorithms.

Use of $z_{1-\alpha}o_O$ in place of $t_{1-\alpha}s_O$ was a hard choice. I made it because LLD (as opposed to S_C) requires knowledge (or assumption) of o_O, as was noted in the discussion on replicate blanks [A2]; and γ-spectrum algorithms, for example, seldom are <u>really</u> applied to replicate baselines! Also, there is serious danger in s_O being estimated at one activity and interference level (<u>and counting time!</u>) and assumed equivalent [or $\propto 1/\sqrt{t}$] for changes according to Poisson statistics. The formally simple approach of adding the term Δ to Eq (2) limits both misuse and ignorance of a ts_O formulation. [To my knowledge, an all-encompassing rigorous solution to the problem (non-Poisson random and systematic error effects on detection capabilities) does not exist.]

[A9]. _Time Factor._ Obviously, T could be factored into an initial decay correction and decay during counting: $(e^{-\lambda t_a} - e^{-\lambda t_b})/\lambda = e^{-\lambda t_a}(1 - e^{-\lambda \Delta t})/\lambda$. Explicit expressions will not be given for decay during sampling or for multistep counting schemes, because they depend upon the exact design (and input function) for the sampling or counting process.

[A10]. _Excess (Non-Poisson) Random Error._ In place of a massive replication study (to replace $\Delta + z\sigma_0$ [Eq. 2] by ts_0) one could assume a two-component variance model and fit the non-Poisson parameter for approximate estimates of detection limit variations with counting time and interference level (20). This could become crucial when $B \gg 1$.

[A11]. _Systematic Error._ $\underline{\Delta}$ and $\underline{f}$ have been set at "reasonable" values to represent the routine state of the art. These may be subject to more careful evaluation by the NRC or specific estimation by the licensee. This is crucial for instruments in uncontrolled environments where these "reasonable" values may be too small; see footnote, p. 96. Similarly, _if demonstrated_ smaller bounds of_, say, 2% B limits could be substituted for the default bound of 5% B. A most important consequence of including reasonable bounds for systematic error is that LLD cannot be arbitrarily decreased by increasing T.

[A12]. _Multicomponent Analysis, A - Uncertainties._ In cases of multicomponent decay curve analysis or (α, β, γ) spectroscopy, S_C may be transformed to a critical level (decision level) for an initial rate or activity due to spectrum or decay curve _shape_ differences among the components. Common factor transformations (Y,E,V) applied, _with their uncertainties_, to S_C would simply needlessly increase the detection limit. As

75

shown in Eq. (1), such common (calibration) factors and their uncertainties must, however, be included to calculate the value of the a priori performance characteristic, LLD.

[A13]. <u>Decisions and Reporting of Data.</u> S_C (or LLD/2.20) is used for testing (a posteriori) each experimental result (S) for statistical significance. If $S > S_C$, the decision is "detected;" otherwise, not. Regardless of the outcome of this process, the experimental result and its estimated uncertainty should be recorded, even if it should be a negative number. (Proper averaging is otherwise impossible, except with certain techniques devised for lightly "censored" [but not "truncated"] data [Ref. 21, pp 7-16f].) The decision outcome, of course, should be noted and for non-significant results, the actual detection limit (for those particular samples) should be given. If desired, a second decision level of significance using $1.9 \cdot S_C$, may be noted, in view of the effects of multiple decisions on α and β. (See section II.D.5 on the treatment of multiple detection decisions and the origin of the coefficient 1.9.) <u>Obviously</u>, changes in S_C (<u>i.e.</u>, in $z_{1-\alpha}$) alter the detection limit, because of the sum, $(z_{1-\alpha} + z_{1-\beta})$, in Eq. (1).

[A14]. <u>Variance of the Blank.</u> Estimation of $\sigma_0{}^2$ by $s_0^2 = s_B^2 \eta$ is completely valid only if the entire rigorously defined, Measurement Process can be replicated. This is rarely achievable if there are significant levels of interference (B_I), for B_I will doubtless be unique for each sample. A suggested alternative, therefore, <u>if</u> the s_0^2 approach is to be applied, is to estimate s_{BK}^2 for the Blank (non-baseline) and to combine this (necessarily as an approximation) with the Poisson σ_0^2 from the spectrum fitting. One caution: χ^2 is appropriate to estimate bounds for non-Poisson variance (20) and lack-of-fit (model error), but it should <u>never</u> be used as an arbitrary correction factor for the Poisson variance (61,63).

[A15]. $\underline{S_D \text{ vs } x_D \text{ and Error Propagation}}$. The formulation given here is based on $\underline{signal}$ detection (S_C, S_D). Transformation to a $\underline{concentration}$ detection limit (x_D which $\underline{is}$ the LLD) involves uncertainties in the estimated denominator, A. In this report, we do not "propagate" such uncertainties directly, but rather use them to establish a corresponding uncertainty interval for x_D, given S_D. If ϕ_A (RSD) is small, and e_A random, then $x_D = S_D/A$ has the same RSD (ϕ_A). If $\phi_A \gtrsim 0.1$, then the uncertainty interval for x_D can be derived directly from the lower and upper bounds for $\underline{A}$. We take a conservative position, setting LLD equal to the upper bound for x_D. This can be further interpreted as a dualism: i.e., LLD [Eq. 1] is the upper (95%) limit for x_D, and $\beta = 0.05$; or, LLD [Eq. 1] $\underline{is}$ x_D, but $\beta < 0.05$ (upper 95% limit for β). (Eq. (1), where f = 1.1, takes the relative uncertainty in $\underline{A}$ to be ±10%.) S_C, of course, is unaffected by ϕ_A. An alternative treatment ("x-based" rather "S-based") is given Ref. 76, where x_D is estimated from full error propagation, but where one is left with uncertainty intervals for both α and β. The best solution clearly is to all but eliminate ϕ_A, but in any case it should be kept within the bounds given by the default value of factor $\underline{f}$ if at all possible.

[A16]. $\underline{\text{Calibration Factor (A) Variations}}$. If, for a given measurement process $\underline{A}$ actually varies -- e.g., if yields or efficiencies, etc., fluctuate about their mean values from sample to sample -- then the LLD itself varies. If this variation is significant (in a practical sense) and a mean value is used for $\underline{A}$, then x_D would best be described by a tolerance interval for the varying population sampled. Far better, in this case, is the use of direct or indirect measures for $\underline{A}$ (or its component factors -- Y,E,V) for $\underline{each}$ sample; such methods include isotope dilution (for Y) and internal and

77

external efficiency calibration (for E). Sampling errors, which can be very large indeed, come under this same topic; but further discussion is beyond the scope of this report.

B. Proposed Simplified RETS Page for "Simple" Counting[1]

(See footnote at beginning of section III.)

1. The LLD is Defined for purposes of these specifications, as the smallest concentration of radioactive material in a sample that will yield a net count, above system blank, that will be detected with at least 95% probability with no greater than a 5% probability of falsely concluding that a blank observation represents a "real" signal. "Blank" in this context means (the effects of) everything apart from the signal sought -- i.e., background, contamination, and all interfering radionuclides.[2]

For a particular measurement system, which may include radiochemical separation:

$$\text{LLD} \equiv (0.11)\ \text{BEA} + (0.50) \left(\frac{3.29\ \sigma_B \sqrt{n}}{\text{YEVT}} \right) \qquad (6)$$

The above equation gives a conservative estimate for LLD (in pCi per unit mass or volume (V)), including bounds for relative systematic error for the blank of 5%, for baseline (interference) of 1%, and for the calibration quantities (Y,E,V) of 10% [B5]. (A 5% blank systematic error bound [Δ_K] was used above; for baseline error, substitute Δ_I as indicated under 'BEA' below.) The "statistical" part -- numerator of the second term is based on

----

[1] "Simple", as used here, means that the net signal is estimated from just two observations (not necessarily of equal times or number of channels). One observation includes the signal + blank (or interference baseline); the other being a "pure" blank (or baseline) observation. Also, the "expected" (average) number of blank counts must exceed ~70 counts, for Eq. 6 to be adequately valid [B9].

[2] References to notes which follow (in section III.B.2) are indicated in brackets--e.g., [B5].

5% false positive and false negative risks and the standard deviation of the blank or baseline (interference) (σ_B) in units of counts, for the sample measurement time Δt. [See also Eq. (7), pg. 81.]

Meanings of the other quantities are:

BEA = Blank Equivalent Activity (pCi/mass or volume). If the baseline

(underneath an isolated Y-ray peak) is large compared to the blank, substitute "Baseline" for "Blank" in the first term of Eq. (6), and use a coefficient of 0.0220 in place of 0.11,

Y = Radiochemical recovery

E = Overall Counting efficiency (counts/disintegration [B6])

T = $e^{-\lambda t_a}(1-e^{-\lambda \Delta t})/\lambda$, the "effective" counting time (minutes); where λ is the decay constant, t_a is the time since sampling, and Δt is the length of the counting interval [For $\lambda t \ll 1$, $T = \Delta t$] [B7, A6, A9]

σ_B = $\sqrt{B}$ for Poisson counting statistics (B equals the expected number of Blank or Baseline counts). Do not use Eq. (6) unless B $\gtrsim$ 70 counts. Note that use of the observed number of blank counts, $\hat{B}$ in place of the unobservable true value B introduces a <u>relative</u> uncertainty (1σ) of $\leq 6\%$ (Poisson) in the estimated σ_B, if B > 70 counts [B9].

η = $1 + \left(\dfrac{\Delta t}{\Delta t_B}\right) g_A$ where Δt is the measurement time for the sample, and

Δt_B is the measurement time for the background. The dimensionless factor g_A takes into account possible influences of changes in the calibration factor <u>A</u> on the blank -- due to blank interactions/correlations with yield, efficiency or sample volume (mass). Generally, g_A will have the value, unity (77,78).

The <u>Detection Decision</u>: (a posteriori) is made using as the critical level LLD/2.20. Unless such a value is used in conjunction with Eq. (6), the probabilistic meaning (5% false-positive, negative-risks) is non-existent (5)!

2. Tutorial Extensions and Notes

[B1] <u>Simple Spectroscopy</u>: Eq. (6) may be used with isolated α- or γ-ray peaks by substituting: (a) baseline height (counts under the selected sample peak channels) for B in order to calculate BEA and σ_B; and (b) the expression $(1 + n_1/n_2)$ for η, where n_1 = number of peak channels taken and n_2 = total number of channels used to estimate the pure (linear or flat) baseline. (For a linear baseline, n_2 should be symmetrically distributed about the peak integration region.)

[B2]. <u>Replication</u>: The variability of the blank should always be tested by replication, using s^2 and χ^2. (See also notes <u>A2</u>, <u>A14</u>.) If the replication-estimated standard deviation significantly exceeds the Poisson value $(\sqrt{B})$, the cause should be determined.[1] If excess variability is random and stable the factors <u>3.29 σ_B</u> in Eq. (6) may be replaced by <u>2t σ_{UL}</u> as defined in note <u>A2</u>.

Some values of t and σ_{UL}/s (both at α = 0.05) follow:

Table 6. LLD Estimation by Replication: Student's-t and (σ/s) - Bounds
<u>vs</u> Number of Observations

<u>no. of replicates:</u>	5	10	13	20	120	∞
<u>Student's-t:</u>	2.13	1.83	1.78	1.73	1.66	1.645
<u>σ_{UL}/s:</u>	2.37	1.65	1.51	1.37	1.12	1.100

[B3]. <u>Systematic Error Bounds</u>. The presence of systematic error bounds limits unrealistic reduction of the LLD through extended counting. The values (1%, 5% and 10% for blank, baseline and calibration factors, resp.) are believed reasonable [Ref. 72], but if <u>demonstrated</u> lower bounds are achieved, they should be accordingly, substituted.

----

[1]At least 13 replicates are necessary to "assure" (90% confidence) that <u>s</u> be within ~50% of the true σ. [A2]

[B4]. Some Inequalities for Rapid Decision Making and LLD Estimation.
Equation (6) can be written:[1]

$$\text{LLD } x = x_D = 1.1 \ (2x_C) = 1.1 \ (2[\text{ΔBEA} + 1.645 \ \sigma_{xo}]) \qquad (7)$$

where x_C, BEA and σ_{xo} have dimensions of activity per unit mass or volume.
In the absence of systematic error bounds; $x_D = 2x_C$, Δ→0, and 1.1→1. The
standard deviation of the estimated concentration when its true value is
zero, is σ_{xo} which equals $\sqrt{B\eta}$ /[2.22 (YEVT)] for "simple" counting.

One result which is normally available following all radionuclide
measurements is the estimate of the radioactivity concentration, $\hat{x}$, and its
Poisson standard deviation σ_x. Since $\sigma_x \geq \sigma_{xo}$ necessarily (the equality
applying only when x=0 -- i.e., a blank),

$$x_C' = \text{ΔBEA} + 1.645 \ \sigma_x \geq x_C \qquad (8)$$

and

$$x_D' = 1.1 \ (2 \ x_C') \geq x_D \qquad (9)$$

with these two inequalities, using the result which is available with every
experiment (σ_x), we can instantly calculate quantities for conservative use
for Detection Decisions and for setting a bound for LLD.

Equation (8) should be considered as a new (quite legitimate) decision
threshold, for which $\alpha \leq 0.05$. Similarly, using x_C' for detection decisions,
x_D' (Eq. 9) may be considered a detection limit for which $\beta \leq 0.05$. (With a
little more work, one could calculate the (β = 0.05) LLD, which would be

----

[1]For convenience of algebraic statement, x_D will be used here to symbolize the
actual LLD. (See App. A.) Also, when units are concentration, "σ_o" will be
transformed accordingly: i.e., $\sigma_{xo} \equiv \sigma_o/A$, thus, σ_{xo} is σ_x for x=0.

<u>smaller</u> than x_D', using x_C'.) If, then, x_D' is less than the prescribed

regulatory value x_R for LLD, the requirements will have been met; and actual

calculation of σ_O, and LLD using Eq (1), would be unnecessary. Obviously,

this approach cannot be applied <u>completely</u> a priori, in the absence of any

experimental results. Operationally, however, it is straightforward,

conservative, and satisfies the goals of RETS.

Limits for the ratios of x_D'/x_D, which are necessarily the same for

x_C'/x_C, are readily given for simple counting. If the true value of sample

counts (S) is not zero, then the quantity $\sqrt{B\eta}$ is replaced with $\sqrt{B(\eta+r)}$ where

r = S/B, the ratio of sample to blank counts ("reduced activity" [Ref. 19]).

Thus, for S = B, for example, and $\eta=1$ (well-known blank), σ_O would be

increased by a factor of $\sqrt{1+r} = \sqrt{2}$, and this would be reflected in σ_x. The

ratio x_b'/x_D would likewise be $\sqrt{2}$, if there were no systematic error. When

systematic error dominates (ΔBEA in Eq. 8), then $x_D'/x_D \to 1$ showing no change.

[B5]. <u>Calibration Factor Variations</u>. If there are large <u>random</u> varia-

tions in Y, E, or V, the full replication of x (radioactivity concentration)

and σ_x should be considered in place of the f-systematic error bound

approach.

[B6]. <u>Branching Ratios</u> (or absolute radiation -- α, β, γ, e_K, --

fractions) may be shown explicitly by factoring the efficiency. Thus, for

example, $E = E_\gamma \cdot \xi_k$, where E_γ represents the counting efficiency for a γ-ray

of the energy in question, and ξ_k represents the branching ratio for that

energy γ-ray from radionuclide-k. All else being equal, then LLD $\propto 1/\xi_k$.

[B7]. <u>Continuous (Monitoring) Observations</u> [See also footnote: p.51].

When a <u>digital</u> count rate meter is employed (Ref. 73), or when a "long"

average estimate with an analog rate meter is made, the standard deviation of

the background rate is unchanged -- i.e., $\sigma_B/T = \sqrt{R_B/\Delta t}$ (for $\lambda t \ll 1$). When an

"instantaneous" analog reading is made, however, $T \to 2\tau$ (τ = resolving time of the circuit), so $\sigma_B/T \to \sqrt{\dfrac{R_B}{2\tau}}$ [Ref. 74]. Changes in analog ratemeter readings are governed by the instrumental time constant, just as they are in exponential radioactivity growth and decay, by the nuclear time constant.

[B8]. <u>Decisions and Reporting of Data</u>. S_C (or LLD/2.20) is used for testing each experimental (a posteriori) result ($\hat{S}$) for statistical significance. If $\hat{S} > S_C$, the decision is "detected"; otherwise, not. Regardless of the outcome of this process, the experimental result and its estimated uncertainty should be recorded, even if it should be a negative number. (Proper averaging is otherwise impossible, except with certain techniques devised for lightly "censored" [but not "truncated"] data [Ref. 21, pp 7-16f].) The decision outcome, of course, should be noted and for non-significant results, the actual detection limit (for those particular samples) should be given. If desired, a second level of significance, using 1.9 x S_C, may be noted, in view of the effects of multiple decisions on α and β. (See Section II.D.5 on the treatment of multiple detection decisions.)

[B9]. <u>Counts Required for Adequate Approximation of σ_B and S_D</u>. When B is large, the approximations

$$\text{(i)} \quad \underline{\sigma_B \approx \sqrt{B}} \quad \text{and} \quad \text{(II)} \quad \underline{S_D \approx 2S_C = 2z\sqrt{B\eta}}$$

become quite acceptable. They are, in fact, asymptotically correct, just as the Poisson distribution is asymptotically Normal. Regions of validity can be set by requiring, for example, that each approximate expression deviate no more than 10% from the correct expression.

<u>For Case (I)</u>, where the observed number of counts is used as an estimate for the Poisson parameter, we require:

$$0.90 \leq \sqrt{B} \,/\, \sqrt{\bar{B}} \leq 1.10$$

Taking the upper limit $\hat{B} = B + z_{1-\gamma} \sqrt{B}$, we have

$$(1.10)^2 \geq 1 + z_{1-\gamma}\sqrt{B}, \text{ or } B \geq (z_{1-\gamma}/0.21)^2 \text{ counts}$$

For '1σ' ($z=1$), this means $B \geq 22.7$ counts; for the '95% CI' ($z=1.96$), the limit is $B \geq 87.1$ counts. A most important point is that the B referred to is that associated with the <u>Blank</u> experiment, because that is the source of the estimate $\hat{B}$. Thus, if $b = \Delta t_B/\Delta t$ equals the ratio of counting times ["pure blank"/(signal + blank)], the RSD of $\hat{B}$ is given by $1/\sqrt{bB}$. The requisite number of counts bB is still $(z/0.21)^2$, but B itself is reduced to $(z_{1-\gamma}/0.21)^2/b$ $[b \geq 1]$. If, for example, the blank is measured twice as long as the sample, the '1σ' ($z=1$) limit for approximation (I) is $B \geq 11.3$ counts (expected).

For Case (II), we require that,

$$S_D/2S_C \leq 1.10$$

that is,

$$(z^2 + 2z \sqrt{B\eta})/(2z \sqrt{B\eta}) \leq 1.10$$

this reduces to (for $z_{1-\alpha} = z_{1-\beta} = 1.645$)

$$B \geq (5z)^2/\eta = (5 \cdot 1.645)^2/\eta = 67.6/\eta \text{ counts}$$

Taking the usual limits for η, we have

$$B \geq 67.6 \text{ counts } (\eta=1, \text{ "well-known" blank})$$

$$B \geq 33.8 \text{ counts } (\eta=2, \text{ "paired comparison"})$$

Since $\eta = 1 + 1/b$, this second approximation (II) is the more stringent.

C. <u>LLD for Specific Types of Counting</u>

1. <u>Extreme Low-Level Counting</u>

When fewer than ~70 background or baseline counts (B) are observed, the "simple" counting formula for S_D must have added the term $z^2 = 2.71$ (for

$\alpha=\beta=0.05$) to account for minor deviations of the Poisson distribution from Normality. [Ref. 5.] (Obviously, this term may be retained for B > 70, but its contribution is then relatively minor.)

When the mean (expected) number of background counts is fewer than about 5, such as may occur in low-level α-counting, further caution is necessary because of the rather large deviations from Normality. This issue has been treated in some detail in Ref's 19 and 75. The extreme case occurs, of course, when B≈0 where the asymptotic formula ($S_D \approx 3.29 \sqrt{B}$) would give a detection limit (counts) of zero, and the intermediate formula, 2.71. In fact, as will be shown below, the true detection limit ($\alpha=\beta=0.05$), in the case of negligible background, is 3.00 counts. Though the intermediate formula is not so bad in this case (within ~10% for S_D), the accuracy for S_C and S_D fluctuates as B increases from zero to ~5 counts; but above this point (B=5 counts) the deviations are generally within 10% relative. (Note that the symbol B refers to the true or expected value of the blank; $\hat{B}$ refers to an experimental estimate.)

For accurate setting of critical levels (for detection decisions) and detection limits, when B < 5 counts, we therefore recommend using the exact Poisson distribution. In the following text we shall use the development given in Ref. 19 and make explicit use of Fig. 1 from that reference -- which appears here as Fig. 7. Before fully discussing the use of this figure, let us make some critical observations:

o   The mean number of background counts is assumed known. Such an assumption is both reasonable and necessary. It is reasonable in that, even for the lowest level counting arrangements, long-term background measurements should be made yielding, say, at least 100 counts. (An RSD of 10% is trivial in the present context.) The assumption is more or less

85

necessary, in that a rigorous detection limit cannot be stated for the difference between two estimated Poisson variables, although rigorous detection decisions and _relative_ limits can be given. (See references 19, 36, and 75 for further details.)

o   Fig. 7 gives the detection limits in units of BEA (background equivalent activity) as a function of B. For relatively small uncertainties in B, one can deduce limiting values from the curve.

o   The integers above the curve envelope indicate the critical number of gross counts ($y_C = S_C + B$). (Though B and S and y -- i.e., true or expected values are real numbers, the critical level for y (y_C) as well as all _observed_ gross counts are necessarily integers.)

o   The "sawtooth" structure of the envelope reflects the discrete (digital) nature of the Poisson distribution. A consequence is that the false positive risk becomes an inequality -- i.e., $\alpha \leq 0.05$. At each peak $\alpha = 0.05$, and then it is gradually decreases until the next integer satisfies the $\alpha = 0.05$ condition.

o   The dashed curve represents the locus of the intermediate expression ($S_D \approx 2.71 + 1.645 \sqrt{B}$).

o   It is seen that the extreme low-level situation generally applies to the case where the Poisson detection limit exceeds the BEA. In fact, this occurs once B is less than ~16 counts. It is recommended that Fig. 7 be used for detection decisions ($S_C + B$ = integers above the curve envelope) and estimated detection limits (ordinate = detection limit, in BEA units). In addition, the figure can be useful for designing (planning) the measurement process. For example, if the BEA for a particular nuclide is 1 pCi/L and one wishes to be able to detect 5 pCi/L, it is clear that the expected number of background counts must be at least
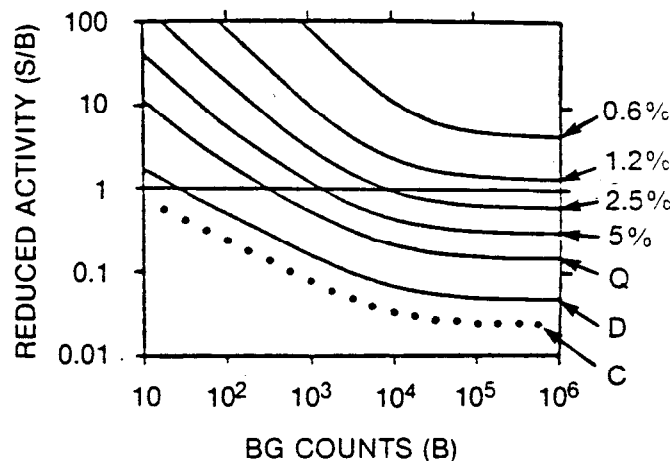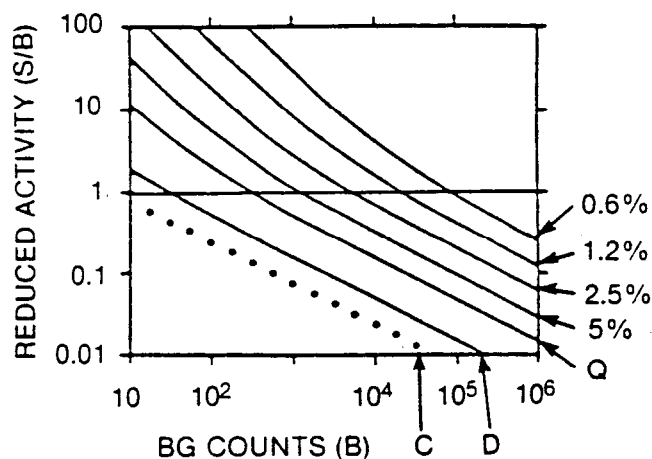
86
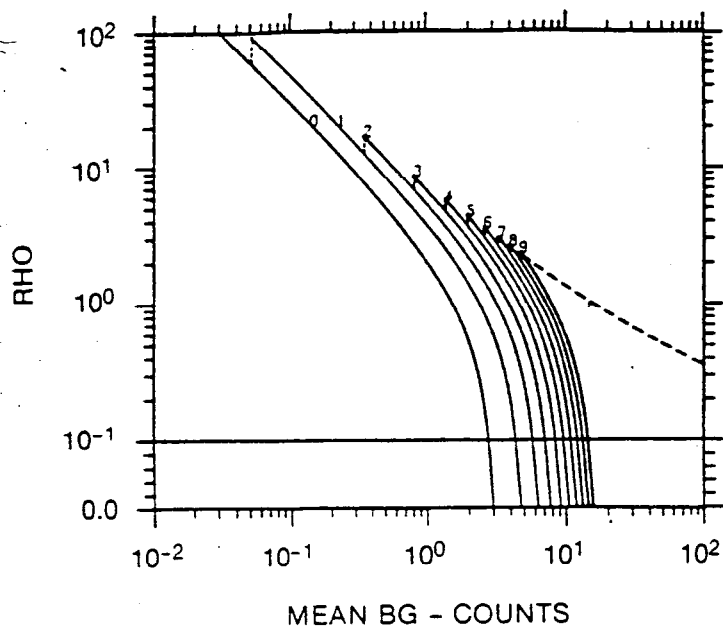
Figure 7a. Reduced Activity (ρ) vs Mean Background Counts (μ_B) and Observed
  Gross Counts (n). Each of the solid curves represents the upper
  limits for ρ vs μ_B, given n. The envelope of the curves,
  connected by a dotted line, represents the detection limit (ρ_D)
  and critical counts (n_C) as a function of μ_B. (α = β = 0.05)

7b. Reduced activity curves. Contour plots are presented for reduced
  activity (S/B = ρ) versus background counts (B) and counting
  precision (θ). Part (a) includes Poisson errors only; part (b)
  incorporates additional random error (0.50% for counting
  efficiency, 1.0% for background variability).

-1.3. If the background rate is, e.g., 3 counts/hour, this means a 26 min measurement is necessary (assuming the mean background rate to be reasonably well known).

o   A further use for Fig. 7 is the setting of the upper limits when $y <$ y_C. That is, the sequence of curves below the detection limit envelope, which have integers less than y_C, represent all possible outcomes when activity is not detected. For example, if B (expected value) = 1.0 count, y_C = 3 (so S_C = 2.0) and the normalized detection limit is 6.75 · BEA. If an experimental result were y = 1 count, the second curve below (labeled "1") intersects with B = 1.0 and the ordinate at the (5%) upper limit of 3.74 · BEA.

o   Table 7 is offered as an alternative to Fig. 7. Again, the mean background rate is assumed well-known, and $\alpha \leq 0.05$ while $\beta = 0.05$. For the case earlier discussed (B = 1.3 counts), we see that the net critical number of counts is 1.7 [i.e., 3 - 1.3] where y_C is necessarily an integer; and the detection limit is 7.75 - 1.30 = 6.45 counts, which is indeed ~ 5 · BEA. (Though β = 0.050, for this particular case it can be shown that α = 0.043.) The intermediate formula would have given 1.88 counts (1.645 $\sqrt{B}$) for S_C and 6.46 counts for S_D -- results that are fortuitously close to the correct values. (The fortuitousness becomes clear when one calculates S_C and S_D for B = 2.0, for example.)

Table 7. Critical Level and Detection Limits for Extreme Low-Level
          Counting (Assumes B, known)

Background Counts B - Range	Gross Counts	
	$y_C = S_C + B$ (integer)	$y_D = S_D + B$
0 - 0.051	.0	3.00
0.052 - 0.35	1	4.74
0.36 - 0.81	2	6.30
0.82 - 1.36	3	7.75
1.37 - 1.96	4	9.15
1.97 - 2.60	5	10.51
2.61 - 3.28	6	11.84
3.29 - 3.97	7	13.15
3.98 - 4.69	8	14.43
4.70 - 5.42	9	15.71

2. <u>Reductions of the General Equations.</u>

         For direct application of Eq's (1) and (2) we take the
following parameter values,

$f = 1.10$ (10% YEV - "calibration" systematic error bound)

$z_{1-\alpha} = z_{1-\beta} = 1.645$ (5% false positive and negative risks)

$\Delta = \Delta_K + \Delta_I = \delta_K B_K + \delta_I B_I = 0.05 B_K + 0.01 B_I$

where: Δ_K, Δ_I represent systematic error bounds (counts) from the blank

       and interference (e.g., non-blank component of a baseline),

       respectively.[1]

$\delta_{K,I}$ denote <u>relative</u> systematic error bounds of the Blank (counts,

B_K) and of the Interference (B_I). 5% and 1% values are taken as

reasonable for routine measurement, but these may be replaced by

laboratory-specific values (δ) which have <u>demonstrated</u> validity.

----

[1]Note that the Blank and Baseline (non-blank portion) are properly treated
   apart (a) because the Blank may contribute directly to a peak (α, Y-ray) due
   to contamination by the very nuclide sought, and (b) because of difference
   in both the origins of their systematic errors, and their (external)
   variability.

89

[Symbols without subscripts will denote summation, e.g.,

$\Delta B = \Sigma \Delta_i B_i = \|\Delta B\|$].

Thus, Eq. (1) takes the form,

$$\text{LLD} = \frac{1.1 \ (2 \ S_C)}{2.22 \ (\text{YEV})T} = x_D \qquad (10)$$

$$\text{LLD} = 0.11 \ (\text{BEA})_K + 0.022 \ (\text{BEA})_I + \frac{(1.1)3.29\sigma_0}{2.22 \ (\text{YEV})T} \qquad (11)$$

and

$$S_C = 0.05 \ B_K + 0.01 \ B_I + 1.645 \ \sigma_0 \qquad (12)$$

where: BEA = Blank (or Interference) Equivalent Activity

i.e., $\text{BEA} = B/[2.22 \ (\text{YEV})T] = B/A$ \qquad (13)

From the above equations it is clear also that the critical level,

expressed in the same units as LLD, is just LLD/2.2. Use of this is equiva-

lent to applying S_C to test net counts for significance; and the form of data

output available may make it (LLD/2.2) more convenient to use than S_C. In the

absence of systematic calibration error, of course, this equals LLD/2.

3.   Derivation and Application of Expressions for σ_0 -- The Poisson Standard
     Deviation of the Estimated Net Signal, Under the Null Hypothesis [Blank][1]

A.   "Simple" counting (gross signal minus blank)

i) Derivation

   When two (sets of) observations (y_1, y_2) are made, one of the sample and

one of the pure Blank (or Interference), we have

$$y_1 = S + B \quad + e_1 \text{ (counts)} \qquad \text{[observed]} \qquad (14)$$

----

[1]In the following text, Δ A, and B will be used without subscripts, in order
to simplify the presentation. The context will indicate whether the Blank
(B_K) or interference (B_I) predominates. As noted elsewhere, if the number
of background (or interference) counts exceeds ~70, the normal approximation
of (Poisson statistics) is adequate, and the relative uncertainty in
estimating σ_0 (or σ_B) will be less than 6%.

$$y_2 = \quad bB \quad + e_2 \text{ (counts)} \qquad \text{[observed]} \qquad (15)$$

(where e_1, e_2 are error terms)

Then, $\hat{S} = y_1 - y_2/b$.

$$\sigma_S^2 = \sigma_{y_1}^2 + \left(\frac{1}{b}\right)^2 \sigma_{y_2}^2 \approx y_1 + \left(\frac{1}{b}\right)^2 y_2 \quad {}^* \quad (16) \ .$$

---

* The approximation (to be used throughout this section) involves taking y (or $\hat{B}$), rather than the expected value $E(y)$ (or B), to estimate the Poisson variance [B9].

For the null hypothesis (S = 0),

$$\sigma_S^2 = \sigma_O^2 = B + \left(\frac{1}{b}\right)^2 (bB)$$

i.e. $\sigma_O = \sigma_B \sqrt{\eta} = \sqrt{B\eta}$, where $\eta = (1 + 1/b)$

The critical level S_C thus equals

$$S_C = z_{1-\alpha} \ \sigma_O = 1.645 \ \sqrt{B\eta} \qquad (17)$$

The detection limit (counts) is defined from the basic relation,

$$S_D = S_C + z_{1-\beta}\sigma_D = z_{1-\alpha}\sigma_O + z_{1-\beta}\sigma_D = z_{1-\alpha}\sigma_O + z_{1-\beta} \sqrt{\sigma_O^2 + S_D} \quad (18)$$

Taking $\alpha = \beta$, this leads to,

$$S_D = z^2 + 2z\sigma_O \approx 2S_C \qquad (19)$$

Since for $\alpha = \beta = 0.05$, $z^2 = (1.645)^2 = 2.71$,

$$S_D = 2.71 + 3.29 \ \sqrt{B\eta} \text{ (counts)} \qquad (20)$$

The first term is not completely negligible if B is small. For approximate normality, $B \gtrsim 9$ counts (Ref. 19); but to make the first term above (2.71) negligible -- i.e., less than 10% of S_D, we require at least 67 counts, since $\eta \geq 1$. [Below B $\approx$ 5 to 10 counts, the "extreme Poisson" techniques for detection limits, discussed in Ref. 19 and section III.C.1, should be employed; and for $5 \leq B \leq 70$ counts the full equation above should be used (See also Ref. 36.).]

ii) Two Special Cases

[1.] Gross Signal - blank

[RANDOM PART]

If the sample is measured for time t_1, yielding y_1 counts; and the blank for time t_2, yielding y_2 counts, then

$$b = t_2/t_1 \text{ and } \eta = 1 + \frac{1}{b} = \left(\frac{t_1+t_2}{t_2}\right)$$

$$\hat{B} = y_2/b$$

and $\hat{S} = y_1 - y_2/b = y_1 - y_2 (t_1/t_2)$

(Note that if $t_2 \geq t_1$, then the limits for η are obviously, 1 and 2.)

This is to be compared with the critical number of counts S_C,

$$S_C = 1.645 \, \sigma_0 \text{ where } \sigma_0 = \sqrt{B\eta} = \sqrt{\hat{B} \left(\frac{t_1+t_2}{t_2}\right)}$$

If $S \leq S_C$, we conclude ND; otherwise D

$$S_D = 2.71 + 3.29 \, \sigma_0 \tag{21}$$

and,

$$x_D = LLD = 2S_D/[(YEVT)(2.22)] \tag{22}$$

or, using Eq. (11) directly [last term divided by 1.1]

$$x_D = \frac{3.29 \, \sigma_0}{(2.22)(YEVT)} = \frac{3.29 \sqrt{\hat{B} \left(\frac{t_1+t_2}{t_2}\right)}}{2.22(YEVT)}$$

where the first approximation comes from dropping the term 2.71 in the numerator, and the second approximation comes from using $\hat{B}$ for the unobservable true value [B9]. (Both approximations are adequate so long as $B \gtrsim 70$ counts, and $t_2 \geq t_1$.)

If t_1 is small compared to the half-life, then $T \approx t_1$ (called Δt, earlier) Since $B = R_B \cdot t_1$, σ_0 and S_C scale as $t_1^{1/2}$, and $x_D \propto t_1^{-1/2}$. (For fixed t_2/t_1 or for $t_2 \gg t_1$.)

When decay during counting is not negligible then x_D decreases less rapidly with increasing t_1; and eventually ($t_1 \gg t_{1/2}$) T assumes the form $e^{-\lambda t_a}/\lambda$ which is independent of Δt (i.e. t_1), so x_D asymptotically increases as $t_1^{1/2}$. Obviously, there is an optimum (minimum LLD or x_D). [See section II.D.3 on Design.]

[+SYSTEMATIC PART]

Eq's (11) and (12) include terms for systematic error bounds for B (viz, ΔB_K and ΔB_I), where for the Blank (all that's being considered here), the relative error λ is taken as 0.05.

$$S_C = 0.05\ B_K + 1.645\ \sigma_0 \qquad (23)$$

$$\approx 0.05\ B + 1.645\ \sqrt{B\left(\frac{t_1+t_2}{t_2}\right)} \quad \text{[counts]}$$

and

$$x_D = 0.11\ (BEA)_K + \frac{(1.1)(3.29)\sigma_0}{2.22(YEVT)} \qquad (24)$$

$$\approx \frac{0.11\ B + 3.62\ \sqrt{\hat{B}\left(\frac{t_1+t_2}{t_2}\right)}}{2.22(YEVT)}$$

Since the first term in the numerator varies more rapidly with B than the second, the systematic error bound will predominate above a certain number of Blank counts;

$$B_{eq.}^{(K)} = \left(\frac{3.62}{0.11}\right)^2 \eta = 1082 \cdot \eta = 1082 \left(\frac{t_1+t_2}{t_2}\right) \text{ counts} \qquad (25)$$

Again, for long-lived radionuclides, ($t_1 \ll t_{1/2}$), $T = t_1$, and since $B = R_B t_1$,

$$x_D = \left[0.11\ R_B + 3.62 \sqrt{\frac{R_B \eta}{t_1}} \middle/ [2.22(YEV)] \right] \qquad (26)$$

The asymptotic constant value for x_D is determined therefore by the Blank rate, as indicated in the first term.

For $t_1 \gg t_{1/2}$, $T \to e^{-\lambda t_a}/\lambda = $ constant; so, from equation (24)

$$x_D = \text{const}\ (R_B t_1) + \text{const}'\ \sqrt{R_B t_1} \qquad (27)$$

thus, x_D asymptotically <u>increases</u> with t_1.

As stated elsewhere, the use of systematic error bounds converts the statistical risks into inequalities: $\alpha \le 0.05$, $\beta \le 0.05$.

[<u>REPLICATION</u>]

Let us suppose that 11-observations were made of the Blank; <u>all for the same time</u>, t_1. (Otherwise, the simple replication model is invalid.) Then, following the common estimation procedure,

$$\hat{S} = y_1 - \hat{B}, \text{ where } B = \sum_1^n B_i/n, \quad s_B^2 = \hat{\sigma}_B^2 = \frac{\Sigma(\overline{B}_i - B)^2}{n-1}$$

and

$$S\hat{E}(\hat{B}) = s/\sqrt{n} \qquad (28)$$
$$\hat{\sigma}_S \lesssim [y_1 + s_B^2 n]^{1/2}$$

Now, in place of $z\sigma_B$, we use ts_B, so $z\sigma_0 \to t\ s_B\sqrt{\eta}$, where now $\eta = (n+1)/n$ because t_2 has been replaced with $\sum_1^n t_1 = n \cdot t_1$.

In the absence of systematic error, the critical number of counts is given by

94

$$S_C = t_{1-\alpha,\nu}\, s_B \sqrt{n} \qquad\qquad (29)$$

where $t_{\alpha,\nu} = t_{.05,n-1}$ is Student's-t at the 5% significance level with n-1 degrees of freedom (ν),

$$x_D = \frac{2t_{1-\alpha,\nu}\cdot o_B\sqrt{n}}{2.22(\text{YEVT})} \leq \frac{2t_{1-\alpha,\nu}\cdot s_B\sqrt{n}/\sqrt{F_{1-\alpha,\nu}}}{2.22(\text{YEVT})} \qquad (30)$$

The inequality gives an upper limit for x_D, taking into account the uncertainty of o_B through the use of the χ^2. ($F_{1-\alpha,\nu,\infty}$ is equal to χ^2/ν for ν-degrees of freedom at the αth percentage point.)[1]

An alternative treatment, wherein a non-Poisson (or "extraneous") variance component is estimated and combined with the Poisson estimate, $\sqrt{B}$, is described in Ref. 20.

[REPORTING]

Recommendations for reporting the results following the above tests: the estimate $\hat{x} = \hat{S}/(2.22\ \text{YEVT})$, the estimated bound for systematic error $[f_\Delta(\text{BEA})]$, and the standard error $o_S/(2.22\ \text{YEVT})$, should all be recorded regardless of the outcome of the detection test for significance (whether $\hat{S} > S_C$ or not). This is vital both for unbiased averaging, and for the possibility of future tests at different levels of significance or with different estimates of systematic error. For "ND" results, the corresponding estimate of x_D should be provided. For the sake of uniform reporting practice and to avoid straining the distributional assumptions (Poisson $\approx$ Normal) one standard deviation (not a multiple thereof) should be reported.

----

[1] Because of the large uncertainty interval for s/o unless ν is very large, the use of an upper limit for x_D is preferred to the simple substitution of s for o_B in the previous equation. [A2]

[CONTINUOUS MEASUREMENT]

A long-term ($t_{1,2} \gg \tau$) measurement with an analog count-rate meter or a digital count rate meter measurement follow essentially the same statistics as above.

For an "instantaneous" measurement with an analog meter, however, the uncertainty in the rate is given by

$$\sigma = \sqrt{R/(2\tau)}$$

where τ is the RC time constant.

The product $\sigma_B \sqrt{\eta}/T$ in Eq (1B) is therefore replaced by $\sqrt{R_B \eta}/e^{-\lambda t_a}$, where now $\eta = \left(\dfrac{1}{2\tau} + \dfrac{1}{t_2} \right) \approx 1/(2\tau)$ (assuming $t_2 \gg \tau$ and $t_{1/2} \gg \tau$). For an "instantaneous" observation of a sample, we correspondingly find:

$$R_{net} = \frac{R_{gross} - R_B}{e^{-\lambda t_a}}; \qquad \sigma R_{net} = \frac{\sqrt{\dfrac{R_{gross}}{2\tau} - \dfrac{R_B}{t_2}}}{e^{-\lambda t_a}} \qquad (32)$$

The corresponding radioactivity concentrations are found by dividing the respective R's by (2.22)(YEV); and the factors 1.645 and 3.29 are used, respectively, to calculate critical levels and detection limits (LLD).[1]

A further complication with rate meters is the equilibration time (RC for analog instruments) which must be taken into consideration (74).

---

[1]The reader should be alerted to the fact that an instrument in a relatively uncontrolled environment, such as a count rate meter, may be subject to rather significant non-Poisson "background" variations. Therefore, it is urgent that the χ^2 test for background reproducibility be carried out, and if non-Poisson random variability is implied, s^2 should be used in place of the Poisson variance estimate. (See the earlier section on the use of Student's-t with replication procedures.)

Worse still, such fluctuations may be non-Normal or even non-random in character. In this case a system-specific estimate should be made for the relative uncertainty bounds -- i.e., ϕ_B. (One should not simply adopt the "reasonable" value of 5%, suggested for controlled environment [well-shielded] counting systems.)

[INSTRUMENTAL THRESHOLD]

On occasion, when there is "sensitivity to spare" a fixed, possibly arbitrary threshold (K) will be set in place of S_C. The minimum detectable number of counts is then given by:

$$S_D = K + z \sqrt{S_D + o_0^2} \qquad (33)$$

This equation has the approximate solution,

$$S_D = K + z^2 + z [K + o_0^2 + z^2/2]^{1/2} \qquad (34a)$$

or, if $K \gg o_0^2 = B\eta$,

$$S_D = K + 1.645 \sqrt{K} \qquad (34b)$$

For such a solution, $\alpha \ll 0.05$, but $\beta = 0.05$. Also, since K is a fixed number (like 103 counts, or in x-units 30 pCi/g for example), S_D is no longer much influenced by the statistical uncertainty in B. On the other hand, the detection limit is increased by an amount K or more.

[2] Simple Spectroscopy

[linear or flat baseline]

If a baseline underlying a spectral peak $(\alpha-, \gamma-)$ is estimated from a region well removed from that peak, then the decision and detection equations are formally identical to those presented above. One simply substitutes (for t_1 and t_2) n_1 and n_2, the respective number of channels used for estimating the peak and the baseline. The only other difference is that the full expression for Δ -- $(\Delta_K + \Delta_I)$ -- must be used, when one includes bounds for systematic error.

[RANDOM PART]

If two equivalent, pure baseline regions lie symmetrically about the peak, as shown in Fig. 8, each having $n_2/2$ channels, then
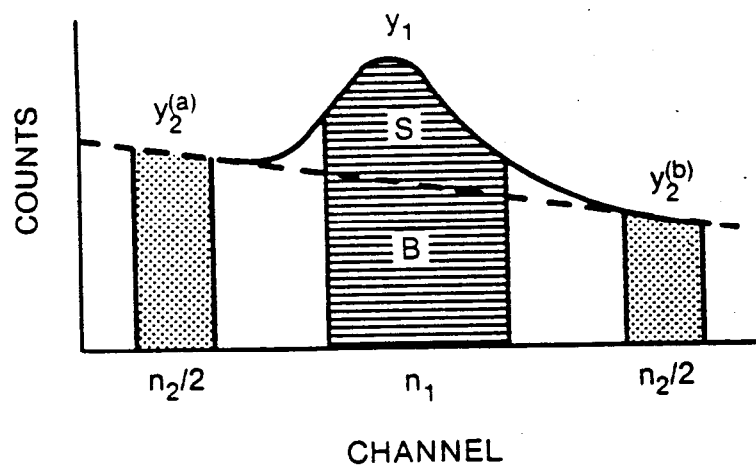
Fig. 8. Simple Counting: Detection Limit for a Spectrum Peak

$$y_1 = \sum_{n_1} y_i = S_1 + \sum_{n_1} B_i + e_1$$

$$y_2 = \sum_{n_2} y_i = \sum_{n_2} B_i + e_2 \quad \text{where}$$

$S_1 = \sum_{n_1} S_i$ equals the number of net sample counts in the peak region.

Under the assumption of linearity for B_i (baseline counts in channel-i),

$$y_1 = S_1 + n_1 \bar{B} + e_1 = S_1 + B_1 + e_1 \qquad (35)$$

$$y_2 = n_2 \bar{B} + e_2 = (n_2/n_1) B_1 + e_2 \qquad (36)$$

Thus,

$$S_1 = y_1 - y_2 (n_1/n_2)$$

$$V_{S_1} = y_1 + (n_1/n_2)^2 y_2$$

$$\sigma_0^2 = B_1 + \left(\frac{n_1}{n_2}\right)^2 (n_2 \bar{B})$$

$$= B_1 + \left(\frac{n_1}{n_2}\right)^2 \left(\frac{n_2}{n_1} B_1\right) = B_1 \eta \qquad (37)$$

$$\text{where } \eta = 1 + \frac{n_1}{n_2} = \frac{n_1 + n_2}{n_2}$$

Thus, the formulation is identical to the preceding one for gross signal minus blank, except that n_i's replace the t_i's.

[+SYSTEMATIC PART]

The formal structure again is unchanged. However, since we now treat baseline error bounds rather than blank systematic error bounds, $\Delta \to 0.01$ rather than 0.05. (The common, limiting case when one has baseline interference is assumed here: that $B_K \ll B_I$, so $\Delta \approx 0.01 \, B_I$, with B_I = baseline in region-1 (peak). This quantity is estimated as $y_2(n_1/n_2)$.

Thus,

$$S_C = 0.01\ B_I + 1.645\ \sigma_o$$

$$= 0.01\ B_1 + 1.645\ \sqrt{B_1 \eta} \qquad (23')$$

and,

$$x_D = 0.022\ (BEA)_I + \frac{(1.1)(3.29)\sigma_o}{2.22(YEVT)} \qquad (24')$$

$$\approx \frac{0.022\ B_1 + 3.62\sqrt{B_1\ \eta}}{2.22(YEVT)}$$

where $\eta = (n_1 + n_2)/n_2$

The point at which the systematic baseline error term dominates the expression for x_D is,

$$B_{eq}^I = \left(\frac{3.62}{0.022}\right)^2 \eta = (2.70 \times 10^4)\eta\ \text{counts} \qquad (25')$$

B. Mutual Interference (2 components)

i) Zero degrees of freedom – 2 observations

In both the evaluation of decay curves and simple spectroscopy, one often encounters the situation where there is "mutual interference" -- i.e., where radiations from two components contribute to each of the observations taken, or to each of the two classes of observations. If the relative contributions differ, the two components may be resolvable (depending upon statistics). [For the following discussion, refer to Fig. 9 for simple decay curve resolution, and Fig. 10 for simple spectrum peak analysis.]
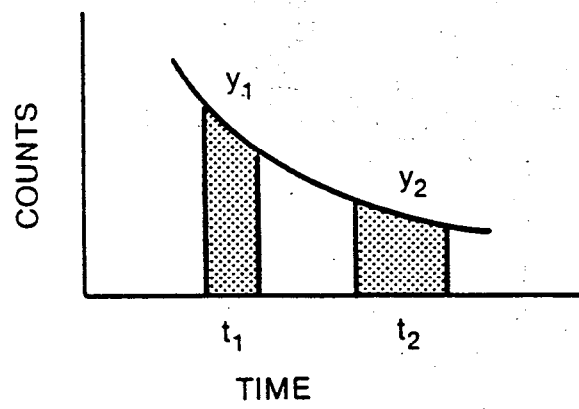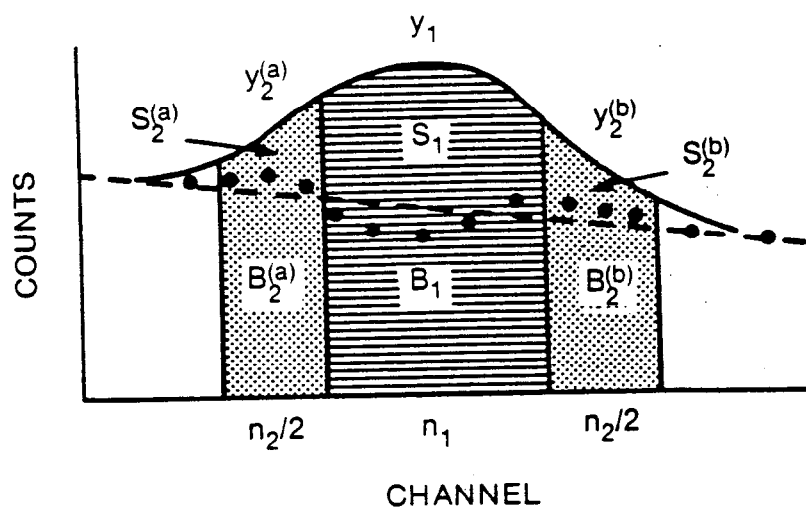
Figure 9.

Figure 10.

Here the signal dominates in region-1 (time or energy) and the blank, in region-2.

Thus,

$$y_1 = \sum_{n_1} y_i = S_1 + B_1 + e_1 \qquad (38)$$

$$y_2 = \sum_{n_2} y_i = a S_1 + bB_1 + e_2 \qquad (39)$$

For the decay curve, the parameters $\underline{a}$ and $\underline{b}$ are uniquely determined by the $t_{1/2}$'s (or λ's) of the 2 components, the spacing (time) of the two observations, and the measurement intervals t_1 and t_2.[1]. [If $\lambda_2 = 0$, then the 2nd component is equivalent to a blank and/or long-lived interfering nuclide.] For the spectrum peak, n_1 and n_2 represent the respective numbers of channels as before; and the n_2's are symmetrically placed about the peak region (symmetric with respect to the mid-n_1-channel) for a baseline model which is linear or flat. The same formalism applies also for the case of two overlapping spectra (provided the blank is negligible or corrected), such as γ-ray doublets. (It should not be overlooked that, for the γ-peak, the effective detection efficiency [E] here depends upon the algorithm -- i.e., the locations, widths and separations of regions -1 and -2.)

Simply to solve these equations, we must assume that $\underline{a}$ and $\underline{b}$ -- i.e., the decay curve or spectrum shapes -- are known. When B (component-2) is a linear baseline or a constant blank or interference (decay curve), $\underline{b}$ is dictated by the model, then $a < b$, and

decay curve: $b = t_2/t_1$

spectrum: $b = n_2/n_1$

---

[1]Thus, $\underline{a}$ (and $\underline{b}$, if $\lambda_2 \neq 0$) subsumes the parameter T in Eq. (11) [Good approximation if t_o is set at the midpoint of the first interval (t_1)].

[RANDOM PART]

The solutions (Poisson statistics part) follow:

$$\hat{S}_1 = \left(\frac{by_1 - y_2}{b-a}\right) \quad \text{and} \quad \sigma^2_{S_1} \approx \left(\frac{b}{b-a}\right)^2 y_1 + \left(\frac{1}{b-a}\right)^2 y_2 \qquad (40)$$

and, replacing S_1 by zero,

$$\sigma^2_0 \approx \left(\frac{b}{b-a}\right)^2 B_1 + \left(\frac{1}{b-a}\right)^2 bB_1 = B_1 \eta \qquad (41)$$

where $\eta = \dfrac{b(b+1)}{(b-a)^2}$

[When $a \to 0$, as in "simple counting", we get the previous result, that $\eta \to (b+1)/b$]

As before,

$$S_C = z_{1-\alpha}\sigma_0 = 1.645 \sqrt{B_1 \eta} \qquad (17')$$

However, the minimum detectable S_1 - counts takes the form,

$$S_D = z^2 \mu + 2S_C = (2.71)\, \mu + 3.29 \sqrt{B_1 \eta} \qquad (20')$$

where $\mu = \dfrac{b^2 + a}{(b-a)^2} > 1$

Some generalizations follow:

(a) If $a \approx b$, both S_C and S_D diverge (S_D more rapidly)

(b) The term $z^2 \mu$ which comes about because of Poisson counting statistics has greater influence than the term z^2 which we find in "simple counting".

(c) In fact the previous approximation $S_D \approx 2S_C$ is poorer, especially when $\underline{a}$ approaches $\underline{b}$,

$$S_D/S_C = 2 + \kappa\, (z/\sqrt{B_1})$$

where $\kappa = \dfrac{(b^2+a)/\sqrt{b(b+1)}}{(b-a)} = \mu/\sqrt{\eta}$

Asymptotic forms for κ:

- ⊙ when $b \gg 1$ and $b \gg a$ (e.g., $n_2 \gg n_1$, or $t_2 \gg t_1$ or for barely overlapping peaks), $\kappa \to \approx (\frac{b}{b-a}) \approx 1$ [also, μ and $\eta \approx 1$]
- ⊙ when $b \approx 1$ (e.g., for blank or linear baseline), $\kappa \to \dfrac{(1+a)/\sqrt{2}}{(1-a)}$

For the first asymptote, $S_D \approx 2S_C$ (within 10%) when $B > 68$ counts, as before ("simple" counting). For the second asymptote, κ ranges from 0.707 [a=0] to ∞ [a=1]. Taking for example, a=1/2 [$\kappa = 2.12$], we find that $S_D \approx 2S_C$ once $B > 304$ counts. Thus, the extra Poisson term ($z^2\mu$) cannot be so readily ignored as in the case of "simple" counting.

Once again, $x_D = S_D/[2.22 \text{ (YEVT)}]$ where T will already have been included in the coefficients a and b for the decay curve example, and E will be influenced by the normalization of the coefficients for the spectrum peak example. (Here, $E=E_1$, the total efficiency corresponding to the fraction of the peak contained in region-1.)

That is, for the decay-curve mutual interference example, $x_D = \overset{o}{R_D}/[(\text{YEV})(2.22)]$ because $\overset{o}{R_D}$ (initial counting rate of the 'signal' radionuclide) depends on the equations including T:

$$y_1 = \overset{o}{R_S} T_{S1} + \overset{o}{R_B} T_{B1} + e_1 \qquad (42)$$

$$y_2 = \overset{o}{R_S} T_{S2} + \overset{o}{R_B} T_{B2} + e_2 \qquad (43)$$

where

$$T_{ji} = e^{-\lambda_j t_i} \left(\frac{1 - e^{-\lambda_j \Delta t_i}}{\lambda_j} \right) \qquad (44)$$

[SYSTEMATIC PART]

Let us next consider bounds for systematic error in B. At this point, a new problem presents itself: should we assume that the relative uncertainty ϕ_B applies to B_1 in y_1, or to $B_2 = bB$, in y_2, or both? In fact, the question as posed is inappropriate. The systematic error in fitting is due to model or <u>shape</u> error (in the baseline) rather than a discrete shift from a signal-free blank observation as in "simple counting."

In order to simply present the systematic (shape) error contribution to x_D, it will be convenient first to change the normalization basis from region-1 to the entire portion of the spectrum or decay curve involved in the fitting. We accomplish this by re-writing Eq's (38) and (39) to read,

$$y_1 = a_1 S + b_1 B + e_1 \qquad (45)$$

$$y_2 = a_2 S + b_2 B + e_2 \qquad (46)$$

where the a's and the b's are normalized to unity ($\Sigma a = 1$, $\Sigma b = 1$). Thus S and B represent the contributions of the net signal and blank to the total peak area that we analyze ($S + B = y_1 + y_2$).

The solution is formally identical to that obtained before,

$$\hat{S} = c_1 y_1 + c_2 y_2 \qquad (47)$$

$$\sigma_o^2 = c_1^2 (b_1 B) + c_2^2 (b_2 B) = Bn \qquad (48)$$

where now

$$c_1 = b_2/D, \ c_2 = -b_1/D, \ D = (a_1 b_2 - a_2 b_1) \qquad (49)$$

and

$$n = \Sigma c_i^2 b_i \qquad (50)$$

The relation,

$$S_D = 2S_C + z^2 \mu = 2z \ \sigma_o + z^2 \mu \qquad (51)$$

where

$$\mu = \Sigma\, c_i^2\, a_i \qquad\qquad (52)$$

is still valid, and it can be shown that $c = c^*/a_1$, $\sigma_o = \sigma_o^*/a_1$, and $\mu = \mu^*/a_1$ where the asterisk refers to the previous normalization (where $a_1, b_1 \to 1$). It follows that

$$x_D = \frac{S_D}{(2.22\ YVT)E} = \frac{S_D^*}{(2.22\ YVT)Ea_1} = \frac{S_D^*}{(2.22\ YVT)E^*} = x_D^*$$

Thus, the (Poisson part of) the detection limit does not depend on the a, b normalization.

With this re-normalization it becomes straightforward to treat systematic error. Substituting Δy_i for y_i in Eq. (47) we obtain

$$\Delta\hat{S} = c_1\,\Delta y_1 + c_2\,\Delta y_2 \qquad\qquad (53)$$

If the Δy_i's are due to systematic shape errors in the baseline, we have

$$\Delta\hat{S} = c_1 B\,\Delta b_1 + c_2 B\,\Delta b_2 = B\,\Sigma\, c_i\,\Delta b_i \qquad\qquad (54)$$

where the Δb_i's are the deviations of the actual baseline shape from the assumed shape and B represents the baseline area (counts) under the fitted region. Thus the quantity $\Sigma\, c_i\,\Delta b_i$ replaces the Δ_B which occurred in the expression for "simple counting" systematic error, so exactly the same equation may be used for calculating the detection limit. (Because of orthogonality between the $\{c_i\}$ and the true baseline $\{b_i\}$, Δ_B can be also calculated directly from the alternative baseline-shape b_i').

$$\Delta_B = \Sigma\, c_i\, b_i' \qquad\qquad (55)$$

A significant change in concept has entered, however, in that the Δb_i represent systematic baseline shape alternatives rather than simply a baseline level shift. (Thus, the Δb_i's represent generally a smooth transition in function -- as from a linear to a quadratic baseline, etc.)

The formalism developed here can be extended quite directly to the estimation of systematic model error even for multicomponent least squares fitting of spectra and decay curves. Some of the basic theory and details have been developed in Ref. 72 ("bias matrix").

(ii) Finite Degrees of Freedom - Least Squares

For just two components (as baseline and spectral peak, etc.) it is relatively simple to extend the above considerations to many observations such as one finds with multichannel spectrum analysis or multiobservation decay curve analysis. (This is because it is trivial to write down the expression for the inversion of the 2 x 2 "normal-equations" matrix.) The same basic matrix formulation applies, however, for any number of components.

[1] General WLS Formulation

In this case ($P \geq 2$, $n > P$), the observations (counts) y_i can be written:

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix} \cdot S + \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} \cdot B + \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix} \tag{56}$$

or, in matrix notation,

$$y = M \theta + e \tag{57}$$

where

$$M = \begin{bmatrix} a_1 & b_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ a_n & b_n \end{bmatrix} \quad \text{and} \quad \theta^T = (S \quad B)$$

The weighted least-squares (WLS) solution to Eq. (57) is,

$$\hat{S} = \hat{\theta}_1 = [(M^T wM)^{-1} M^T wy]_1 \tag{58}$$

and

$$V_{\hat{S}} = (M^T \ w \ M)^{-1}_{11} \tag{59}$$

where the weights $\underline{w}$ are,

$$w_{ii} = 1/V_{y_i} = 1/(M\theta)_i = 1/y_i \tag{60}$$

where the second equality applies for Poisson statistics. (If the observations are independent, $\underline{w}$ is a diagonal matrix -- i.e., $w_{ij} = 0$ for $i \neq j$.)

Defining,

$$c_i \equiv [(M^T \ w \ M)^{-1} \ M^T \ w]_{i1} \tag{61}$$

we can alternately express $V_{\hat{S}}$ by means of error propagation, that is,

$$\hat{S} = \hat{\theta}_1 = \Sigma \ c_i \ y_i \tag{58'}$$

$$V_{\hat{S}} = \Sigma \ c_i^2 V_{y_i} = \Sigma \ c_i^2 (a_i \ S + b_i \ B) \tag{59'}$$

Thus, for the case of Poisson counting statistics,

$$V_{\hat{S}} = S \ \mu + B\eta \tag{62}$$

where

$$\mu \equiv \Sigma \ c_i^2 \ a_i \qquad \qquad \eta \equiv \Sigma \ c_i^2 \ b_i$$

Beyond this, the development is identical to that given above for zero degrees of freedom (P=n). Thus,

$$S_C = z\sigma_o \qquad \text{where } \sigma_o = \sqrt{B\eta} \tag{63}$$

$$S_D = 2S_C + z^2 \ \mu \tag{64}$$

$$\Delta\hat{S} = B\hat{4} = B \ \Sigma \ c_i \ b_i' \tag{65}$$

where b_i' is an alternative baseline shape, used for estimating possible systematic error.

The development thus far has been perfectly general; that is, neither the number of components (P) nor the number of observations have been restricted. Components other than the one of interest ($S = \theta_1$) have, however, been coalesced to form a composite interference or baseline, B.

[2] Explicit Solution for P=2

If we treat the baseline (or any other single component) as a "pure" second component, having fixed shape, then the explicit solution for $\hat{S}$, $V_{\hat{S}}$ and the c_i may easily be stated. The results follow from the inversion of the 2 x 2 matrix

$$(M^T wM)^{-1} = \begin{pmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12} & \Sigma_2 \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_2 - \Sigma_{12} \\ -\Sigma_{12} & \Sigma_1 \end{pmatrix} \Big/ Det. \qquad (66)$$

where

$$Det = (\Sigma_1 \Sigma_2 - \Sigma_{12}^2)$$

and

$$\Sigma_1 = \Sigma w a^2 \qquad \Sigma_2 = \Sigma w b^2 \qquad \Sigma_{12} = \Sigma wab$$

Taking the null case ($S = 0$), the weights equal,

$$w_{ii} = 1/(Bb_i) \qquad (67)$$

using the above expression for w_{ii} and the previous definition for c_i, together with the explicit expression for the matrix $\underline{M}$ and its inverse, it can be shown that,

$$c_i = [a_i/b_i - 1)]/[\Sigma a_i^2/b_i - 1] \qquad (68)$$

All other quantities of interest -- μ, η, σ_0, S_C, S_D and λ (given b') -- follow directly as indicated above. (A reminder: we have normalized all "spectrum" components for the foregoing derivations. That is, $\Sigma a_i = \Sigma b_i = 1$.)

Equation (68) yields specific solutions once the peak (a_i) and baseline (b_i) shapes are given. For a flat baseline ($b_i = 1/n$), for example, Eq (68) reduces to

$$c_i = [a_i - 1/n] / [\Sigma a_i^2 - 1/n] \tag{69}$$

It follows that

$$\eta_0 = \Sigma c_i^2 \left(\frac{1}{n}\right) = 1 / [\Sigma a_i^2 - 1/n] \tag{70}$$

Gaussian Peak

If the peak shape (a_i) is symmetric, then a_i and c_i are even functions, which means that if alternative b_i' are odd (and share the same center of symmetry) then the systematic, baseline-model error vanishes.

$$\Delta = \Sigma c_i b_i' = \text{even vector} \cdot \text{odd vector} = 0$$

This suggests that for a symmetric isolated peak, one can treat the baseline as flat--even though it may be linear or otherwise odd (about the peak center) -- without introducing bias.

Passing beyond just the assumption of symmetry, and specifying the peak to be gaussian, we can calculate explicit values for the c_i once n is known.

It is interesting to examine this case as a function of channel density (number of channels per FWHM or per ± 3 standard deviations, (SD), etc.). The results of such a calculation are illustrated below.

Channel Density n = ch/peak peak $\equiv \pm$ 3 SD	$\mu = \Sigma\, c_i^2\, a_i$	$\eta = \Sigma\, c_i^2\, b_i = \Sigma\, c_i^2/n$
3	2.80	1.83
6	2.03	1.60
∞	1.92_4	1.44_4

The above values for μ and η may be used to estimate the several quantities of interest for the detection of an isolated peak. Note that if the observations are extended well beyond the peak (beyond ±3 SD), μ and η can be reduced substantially. The limiting values ($n = \infty$) become 1.16 and 0.591, respectively.

[3] Some Final Comments

The immediately preceding discussion was given from the perspective of γ-ray (or α-particle) spectra. The same formalism would follow (up to the specification of a gaussian or symmetric peak) for detection in decay curve analysis, or β-spectrum analysis, etc.

Except for the general matrix formulation and treatment as composite interference (baseline), the full multicomponent decay or spectrum analysis detection issue will not be treated here. Further discussion would require explicit assumed models (interfering radionuclides); but the basic principles and basic equations would be unchanged.

Regarding this more complicated situation, however, three procedural comments, and three notes of caution may be given:

[PROCEDURAL COMMENTS]:

○ peak searching efficiency and detection power depend on the exact
nature of the algorithm employed. For the IAEA test spectrum for peak
detection, for example, at least six independent principles were used
by 212 participants to detect peaks in the same digitized, synthetic
γ-ray spectrum [Fig. 6 and Ref. 81]. Yet, false positives ranged from
0 to 23 peaks, and false negatives ranged from 3 to 20 peaks. (The
number of actual peaks in the spectrum was 22.)

○ Eq. (64) is approximate only because of changing statistical weights as
S increases from zero to S_D. An exact solution may be obtained by
iteration (Ref. 61).

○ Systematic model error for the mutlicomponent situation may be derived
with the use of a "Bias Matrix," which can be derived from the least
squares solution for $\hat{S}$, --- together with alternative models (Ref.
72).

[CAUTIONS]:

○ Searches for multiple components often lead to multiple detection
decisions. The overall probability of a false positive (α) in
searching a single spectrum can thus be substantially more than the
single-decision risk. (See Ref. 53 and Section II.D.5 for more on
this topic.)

○ If non-linear searches (involving, for example, estimation of half-
lives and/or γ-energies as well as amplitudes) are made, the estimated
signal distribution ($\hat{S}$) is no longer normal. Again, substantial
deviations from presumed values of $\underline{\alpha}$ may be the result (Ref. 90).

⊙ Bad models and experimental blunders may inflate χ^2 because of poor fit. Multiplication of Poisson standard errors by mis-fit $\chi/\sqrt{df}$ will yield misleading random error estimates, and erode detection capability. (See note [A14] and Ref. 63.)

APPENDIX

Appendix A. Notation and Terminology

Response

$$E(y) = B + Ax = B + S \qquad \begin{cases} y = B + Ax + e_y = E(y) + e_y & \text{[observation]} \\ \hat{y} = \hat{B} + \hat{A}x & \text{[estimate]} \\ \hat{x} = (y - \hat{B})/\hat{A} \end{cases}$$

$E(y)$ = response or gross signal (counts), true or "expected" value [y_i denotes the ith sample or time period or energy bin, etc]

y = observed ("sampled") value of y, characterized by an error e_y

δ = random error

σ = standard deviation (SD); $\sigma/\sqrt{n}$ = SD of the mean (standard error, SE); RSD = relative standard deviation

Δ = systematic error (bound)

ϕ = relative-σ (RSD)

$\blacklozenge$ = relative-Δ

$\hat{y}$ = statistically estimated value for y (e.g., weighted mean, ...) (similarly for $\hat{S}$, $\hat{B}$, $\hat{A}$, $\hat{x}$)

$\tilde{y}$ = assumed or "scientifically" estimated value for y

S = true net signal (counts) ["expected value"]

B = true background or blank or baseline (counts) (B_K = blank; B_I = interference counts)

BEA = Background Equivalent Activity = B/A

x = true radioactivity concentration, per unit mass or volume [pCi or Bq/g or L]. To be referred to in the text simply as "concentration"

A = generalized calibration factor; for simple counting, with x in pCi/(g or L), $A = 2.22$ (YEVT), where

Y = (radio)chemical yield or recovery

115

E = detection efficiency (overall, including branching ratio)

V = volume or mass of sample

T = appropriate time factor or function (minutes)

V_S, V_X, etc = variance of the subscripted quantity

σ_S, σ_X, etc = SD of the subscripted quantity = $\sqrt{V}$

σ_O = SD of $\hat{S}$ (at S=0) [counts]; σ_{XO} = SD of $\hat{x}$ (at x=0) [concentration]

η = a multiplier which converts σ_B to σ_O: $\sigma_O = \sigma_B\sqrt{\eta}$. Its value depends on the design of the Measurement Process.

b = ratio of counting times (or channels) blank/(signal + blank); then

$\eta = 1 + 1/b$

S_C, x_C = critical or decision levels for judging whether radioactivity is present, with false positive risk-α

S_D, x_D = corresponding detection limits, with false negative risk -β

$z_{1-\alpha}$, $z_{1-\beta}$ = percentiles of the standardized Normal distribution, equal to 1.645 for α, β = 0.05

LLD = Lower Limit of Detection (for radioactivity concentration) = x_D

x_R = prescribed regulatory LLD -- i.e., limiting value which licensee is supposed to meet. This is in contrast to the actual LLD (x_D) which is achieved under specific experimental circumstances. (Thus, generally, $x_D \leq x_R$)

ν or df = degrees of freedom

Appendix B. Guide to Tutorial Extensions and Notes

Section III.A and III.B were prepared as proposed substitute RETS pages -- the former cast as a more or less comprehensive statement, and the latter, for "simple" gross signal-minus-blank counting. A-series and B-series notes, respectively, were appended to these sections, so that each

116

could be to a large extent self-contained. The following guide (or index) to these notes is given because of their possible general utility, and because the two sets of notes are not only redundant (as intended) but also complementary.

1. Basic Issues

   a) Use of S.I. units - Note A3

   b) General formulation - Note A8

   Eq. (1) was developed for application to most counting situations, through the introduction of parameters σ_0, f and Δ which can be evaluated for the specific counting method and data reduction algorithm in use. The equation must be modified, however, when small numbers of counts are involved. Normal variate percentiles ($z_{1-\alpha}$, $z_{1-\beta}$) are included as parameters which may be modified as appropriate (e.g., multiple detection decisions).

   c) A priori vs a posteriori - Note A4

   Measurement process characteristics must be known in advance before an "a priori" detection limit can be specified -- may call for a preliminary experiment.

   d) Decisions and reporting - Notes A13, B8 (identical)

   The critical level (S_C) may need to be increased in the case of multiple detection decisions; LLD then automatically increases. Non-detected and negative results should be recorded; related topics; averaging, truncation.

2. LLD Formulation -- Conventional (Poisson) Counting Statistics

   a) Rapid detection decisions, LLD bounds via inequalities - Note B4

   b) Extension of the simplified expression (Eq. 6) to isolated spectrum peaks. - Note B1

c) Continuous monitors - Note B7

d) Mixed nuclides, "gross" radioactivity - Note A6

e) Factors for detection efficiency (E), counting time (T). - Notes A7, A9, B6

Branching ratios, spectrum shapes, decay curves and sampling designs all affect LLD beyond just the matter of counting statistics. Interpretation of Eq. (1) (mixing of factors σ_0, E, T) varies accordingly.

3. Non Poisson (P) - Normal (N) Errors

a) Extreme low-level counting (P $\neq$ N); limits of validity for approximate expressions - Notes A5, B9

b) Replication, lack of fit, use and misuse of s^2, χ^2 - Notes A1, A2, A8, A10, A14, B2

c) Uncertainty in and variability of the LLD. Blank variations; multiplicative parameters: Y,E,V - Notes A2, A12, A15, A16, B5

d) Systematic error bounds - Notes A8, A11, B3

Additive and multiplicative components; default values; limiting effect on LLD reduction.

Appendix C.1

DETECTION CAPABILITIES OF CHEMICAL AND RADIOCHEMICAL MEASUREMENT SYSTEMS:

A Survey of the Literature (1923-1982+)

L. A. Currie
Center for Analytical Chemistry
National Bureau of Standards
Washington, DC 20234

Introduction

The twin issues of the detection capability of a Chemical Measurement Process (CMP) and the detection decision regarding the outcome of a specific measurement are fundamental in the practice of Nuclear and Analytical Chemistry, yet the literature on the topic is extremely diverse, and common understanding has yet to be achieved. Besides their importance to the fundamentals of chemical and radiochemical measurement these issues have great practical importance in application, ranging from the detection of impurities in industrial materials, to the detection of chemical signals of pathological conditions in humans, to the detection of hazardous chemical and radioactive species in the environment. It is in connection with this last area, as related to the regulation of nuclear effluents and environmental radioactive contamination, and at the request of the Nuclear Regulatory Commission (NRC), that this report has been prepared. Highlights from our extensive search of the literature are given in the following text.

Scope of the Survey

The focus of the literature survey was directed toward two points: (1) basic principles, terminology and formulations relating to detection in Analytical Chemistry; and (2) basic, but more detailed or specialized studies

relating to detection limits in the measurement of radionuclides, as well as important practical applications in this area. The search was conducted with the aid of five computer data bases, complemented by the examination of major reviews and books treating mathematical and statistical aspects of Analytical Chemistry.

Carefully constructed patterns of keywords led to a total of 1711 titles (1964-1982) which were scanned. From these, 700 were identified as important to our purpose, so abstracts were copied and studied. A final catalog of 387 articles from the computer literature search was prepared, and from this about 100 were marked as having special relevance. Discovering so extensive a literature on so esoteric a topic was somewhat surprising; also surprising, or at least noteworthy, is the fact that a very large fraction of the work on this topic has originated in foreign institutions with major contributions coming from Western and Eastern Europe, the Soviet Union, and Japan.

Basic References and Key Issues

For the purposes of this appendix-report our discussion of the literature must be highly selective; thus only a few of the most critical sources are discussed. We have given primary emphasis to the "archived" literature (e.g., journal articles as opposed to reports); and more general publications treating mathematics, statistics, radioactivity measurement, and quality assurance have been cited only if detection limits were given major focus. A slightly expanded, classified bibliography appears in appendix C.2.d.

The key issues which were addressed or cited in the literature included, as noted above, terminology and formulation (definitions) resulting from exposition of the basic principles of statistical estimation and hypothesis testing in chemical analysis. Special (but basic) topics treated by several

authors included: the effects of counting statistics, non-counting and non-normal random errors, random and systematic variations in the blank, reporting and averaging practices, multiple detection decisions, Bayesian approaches, the influence of the number of degrees of freedom, interlaboratory errors, control and stability, optimization of detection limits, interference effects, data truncation, and decisions vs detection vs determination vs identification limits. Major topics which related specifically to radioactivity measurements included the influence of alternative (γ-, β-) spectrum deconvolution techniques, comparison/selection of alternative instruments and radiochemical schemes of analysis (especially in the area of activation analysis), the treatment of very low-level activity and the treatment of very short-lived radionuclides. Titles in the highly selected bibliography reflect a number of these specific issues.

To conclude this summary report, I should like to cite just a few sources which I believe either set forth or review some of the more basic issues. The groundwork (within the present time frame) was laid by Kaiser (2), who adopted the basic statistical principles of hypothesis testing (and type-I, type-II errors) to detection in spectrographic analysis. Other frequently-cited works from the 60's are papers by St. John, McCarthy and Winefordner (3), Altshuler and Pasternack (4), and Currie (5), the latter two treating the question of radioactivity. Later important works which specifically treat radioactivity detection are given in references (6) - (21). (Further comments cannot be given in this brief report; see the titles for the focus of each paper.)

Finally, some of the most useful expositions and summaries of LLD treatments and principles and unsolved problems may be found in the books and reviews beginning with reference (22). Special attention should be directed

to the IUPAC statement (22,23), the papers by Wilson (34), the chapter by Currie (33), the review by Boumans (26), and the books by Winefordner (30), Kateman and Pijpers (29), and Massart, Dijkstra and Kaufman (28).

Appendix C.2

BIBLIOGRAPHY

(a) Basic References

1. Feigl, F. Tüpfel- und Farbreaktionen als mikrochemische Arbeits-
    methoden, Mikrochemie 1: 4-11; 1923. (See also Chapt. II in Feigl,
    F., Specific and Special Reactions for Use in Qualitative Analysis,
    New York: Elsevier; 1940.

2. Kaiser, H. Anal. Chem. 209: 1; 1965.
    Kaiser, H. Anal. Chem. 216: 80; 1966.
    Kaiser, H. Two papers on the Limit of Detection of a Complete Analyti-
    ical Procedure, English translation of [1] and [2], London: Hilger;
    1968.

3. St. John, P. A.; Winefordner, J. D. A statistical method for evaluation
    of limiting detectable sample concentrations. Anal. Chem. 39:
    1495-1497; 1967.

4. Altshuler, B.; Pasternack, B. Statistical measures of the lower limit
    of detection of a radioactivity counter. Health physics 9: 293-298;
    1963.

5. Currie, L. A. Limits for qualitative detection and quantitative
    determination. Anal. Chem. 40(3): 586-693; 1968.

(b) Radioactivity - Principal References

6. Donn, J. J.; Wolke, R. L. The statistical interpretation of counting data from measurements of low-level radioactivity. Health Physics, Vol. 32: 1-14; 1977.

7. Lochamy, J. D. The minimum-detectable-activity concept. Nat. Bur. Stand. (U.S.) Spec. Publ. 456: 1976, 169-172.

8. Nakaoka, A.; Fukushima, M.; Takagi, S. Determination of environmental radioactivity for dose assessment. Health Physics, Vol. 38: 743-748; 1980.

9. Pasternack, B. S.; Harley, N. H. Detection limits for radionuclides in the analysis of multi-component gamma-spectrometer data. Nucl. Instr. and Meth. 91: 533-540; 1971.

10. Guinn, V. P. Instrumental neutron activation analysis limits of detec tion in the presence of interferences. J. Radioanal. Chem. 15: 473-477; 1973.

11. Hartwell, J. K. Detection limits for radioanalytical counting techniques. Richland, Washington: Atlantic Richfield Hanford Co.; Report ARH-SA-215; 1975. 42p.

12. Tschurlovits, V. M. Zür festlegung der nachweisgrenze bei nichtselek tiver messung geringer aktivitäten. Atomkernenergie. 29: 266; 1977.

13. Robertson, R.; Spyrou, N. M.; Kennett, T. J. Low level Y-ray spectrometry; Nal(Tl) vs. Ge(Li). Anal. Chem. 47: 65; 1975.

14. Zimmer, W. H. Limits of detection of isotopic activity. ARHCO Analytical Method Standard, CODE EA001. Richland, Washington: Atlantic Richfield Hanford Co; 1972.

15. Zimmer, W. H. LLD versus MDA. EG&G Ortec Systems Application Studies. PSD No. 14; 1980.

16. Harley, J. H., ed. EML Procedures Manual. Dept. of Energy, Environmental Measurements Laboratory, HASL-300; 1972.

17. Nuclear Regulatory Commission, Radiological Assessment Branch technical position on regulatory guide 4.8; 1979.

18. Nuclear Regulatory Commission, Regulatory Guide 4.8 "Environmental Technical Specifications for Nuclear Power Plants". 1975 Dec.; Regulatory Guide 4.14 "Measuring, Evaluating, and Reporting Radioactivity in Releases of Radioactive Materials in Liquid and Airborne Effluents from Uranium Mills". 1977 June; Regulatory Guide 4.16 "Measuring, Evaluating, and Reporting Radioactivity in Releases of Radioactive Materials in Liquid and Airborne Effluents from Nuclear Fuel Processing and Fabrication Plants. 1978 March.

19. Currie, L. A. The Measurement of environmental levels of rare gas nuclides and the treatment of very low-level counting data. IEEE Trans. Nucl. Sci. NS-19: 119; 1972.

20. Currie, L. A. The limit of precision in nuclear and analytical chemistry. Nucl. Instr. Meth. 100: 387; 1972.

21. Health Physics Society, Upgrading environmental radiation data. Health Physics Society Committee Report HPSR-1: 1980. (See especially sections 4, 5, 6, and 7.)

(c)   Important Reviews and Texts

22. IUPAC Comm. on Spectrochem. and other Optical Procedures for Analysis. Nomenclature, symbols, units, and their usage in spectrochemical

analysis. II. Terms and symbols related to analytical functions and their figures of merit. Int. Bull. I.U.P.A.C., Append. Tentative Nomencl., Symb., Units, Stand.; 1972, 26; 24p.

23. IUPAC Commission on Spectrochemical and Other Optical Procedures, Nomenclature, Symbols, Units and Their Usage in Spectrochemical Analysis, II. Data Interpretation, Pure Appl. Chem. 45: 99; 1976; Spectrochim. Acta 33b: 241; 1978.

24. Nalimov, V. V. The application of mathematical statistics to chemical analysis. Oxford: Pergamon Press; 1963.

25. Svoboda, V; Gerbatsch, R. Definition of limiting values for the sensitivity of analytical methods. Fresenius' Z. Anal. Chem. 242(1): 1-13; 1968.

26. Boumans, P. W. J. M. A Tutorial review of some elementary concepts in the statistical evaluation of trace element measurements. Spectrochim. Acta 33B: 625; 1978.

27. Liteanu, C.; Rica, I. On the detection limit. Pure Appl. Chem. 44(3): 535-553; 1975.

28. Massart, D. L.; Dijkstra, A.; Kaufman, L. Evaluation and optimization of laboratory methods and analytical procedures. New York: Elsevier Scientific Publishing Co.; 1978.

29. Kateman, G.; Pijpers, F. W. Quality control in analytical chemistry. New York: John Wiley & Sons; 1981.

30. Winefordner, J. D., ed. Trace analysis. New York: Wiley; 1976. (See especially Chap. 2, Analytical Considerations by T. C. O'Haver.)

31. Liteanu, C.; Rica, I. Statistical theory and methodology of trace analysis. New York: John Wiley & Sons; 1980.

32. Eckschlager, K.; Stepanek, V. Information theory as applied to chemical analysis. New York: John Wiley & Sons; 1980.

33. Currie, L. A. Sources of error and the approach to accuracy in analytical chemistry, Chapter 4 in Treatise on Analytical Chemistry, Vol. 1, P. Elving and I. M. Kolthoff, eds. New York: J. Wiley & Sons; 1978.

34. Wilson, A. L. The performance-characteristics of analytical methods. Talanta 17: 21; 1970; 17; 31: 1970; 20: 725; 1973; and 21: 1109; 1974.

35. Hirschfeld, T. Limits of analysis. Anal. Chem. 48: 17A; 1976.

(d) Further Classified References

[i] Basic Principles

36. Nicholson, W. L. "What Can Be Detected," Developments in Applied Spectroscopy, v.6, Plenum Press, p. 101-113, 1968; Nicholson, W. L., Nucleonics 24 (1966) 118.

37. Rogers, L. B. Recommendations for improving the reliability and acceptability of analytical chemical data used for public purposes. Subcommittee dealing with the scientific aspects of Regulatory Measurements, American Chemical Society, 1982.

38. Currie, L. A. Quality of analytical results, with special reference to trace analysis and sociochemical problems. Pure & Appl. Chem. 54(4): 715-754; 1982.

39. Winefordner, J. D.; Vickers, T. J. Calculation of the limit of detectability in atomic absorption flame spectrometry. Anal. Chem. 36: 1947; 1964.

40. Buchanan, J. D. Sensitivity and detection limit. Health Physics Vol. 33: pp. 347-348, 1977.

41. Gabriels, R. A general method for calculating the detection limit in chemical analysis. Anal. Chem. 42: 1439; 1970.

42. Crummett, W. B. The problem of measurements near the limit of detection. Ann. N. Y. Acad. Sci. 320: 43-7; 1979.

43. Karasek, F. W. Detection limits in instrumental analysis. Res./Dev. 26(7); 20-24; 1975.

44. Grinzaid, E. L.; Zil'bershtein, Kh. I.; Nadezhina, L. S.; Yufa, B. Ya. Terms and methods of estimating detection limits in various analytical methods. J. Anal. Chem. - USSR 32: 1678; 1977.

45. Winefordner, J. D.; Ward, J. L. The reliability of detection limits in analytical chemistry. Analytical Letters 13(A14): 1293-1297; 1980.

46. Liteanu, C.; Rica, I.; Hopirtean, E. On the determination limit. Rev. Roumaine de Chimie 25: 735-743; 1980.

47. Blank, A. B. Lower limit of the concentration range and the detection limit. J. Anal. Chem. - USSR 34: 1; 1979.

48. Kushner, E. J. On determining the statistical parameters for pollution concentration from a truncated data set. Atmos. Environ. 10(11): 975-979; 1976.

49. Liteanu, C.; Rica, I. Frequentometric estimation of the detection limit. Mikrochim. Acta 2(3): 311-323; 1975.

50. Tanaka, N. Calculation for detection limit in routine chemical analyses. Kyoto-fu Eisei Kogai Kenkyusho Nempo 22: 121-122; 1978.

51. Ingle, J. D., Jr. Sensitivity and limit of detection in quantitative spectrometric methods. J. Chem. Educ. 51(2): 100-5; 1974.

52. Pantony, D. A.; Hurley, P. W. Statistical and practical considerations of limits of detection in x-ray spectrometry. Analyst 97: 497; 1972.

[ii] Applications to Radioactivity

53. Head, J. H. Minimum detectable photopeak areas in Ge(Li) spectra. Nucl. Instrum. & Meth. 98: 419; 1972.

54. Fisenne, I. M.; O'Tolle, A.; Cutler, R. Least squares analysis and minimum detection levels applied to multi-component alpha emitting samples. Radiochem. Radioanal. Letters 16(1): 5-16; 1973.

55. Obrusnik, I.; Kucera, J. The digital methods of peak area computation and detection limit in gamma-ray spectrometry. Radiochem. Radioanal. Lett. 32(3-4): 149-159; 1978.

56. Heydorn, K.; Wanscher, B. Application of statistical methods to activation analytical results near the limit of detection. Fresenius' Z. Anal. Chem. 292(1): 34-38; 1978.

57. Hearn, R. A.; McFarland, R. C.; McLain, M. E., Jr. High sensitivity sampling and analysis of gaseous radwaste effluents. IEEE Trans. Nucl. Sci. NS-23(1): 354; 1976.

58. Tschurlovits, M.; Niesner, R. On the optimization of liquid scintillation counting of aqueous solutions. Int. J. Appl. Rad. and Isotopes 30: 1-2; 1979.

59. Bowman, W. B. II; Swindle, D. L. Procedure optimization and error analysis for several laboratory routines in determining specific activity. Health Physics, Vol. 31: 355-361; 1976.

60. Mundschenk, H. Sensitivity of a low-level Ge(Li) spectrometer applied to environmental aquatic studies. Nucl. Instruments & Methods 177: 563-575; 1980.

61. Currie, L. A. The discovery of errors in the detection of trace components in gamma spectral analysis, in Modern trends in activation analysis, Vol. II. J. R. DeVoe; P. D. LaFleur, eds. Nat. Bur. Stand (U.S.) Spec. Publ. 312; p. 1215, 1968.

62. Currie, L. A. The evaluation of radiocarbon measurements and inherent statistical limitations in age resolution, in Proceedings of the Eighth International Conference on Radiocarbon Dating, Royal Society of New Zealand; p. 597, 1973.

63. Currie, L. A. On the interpretation of errors in counting experiments. Anal. Lett. 4: 873; 1971.

64. Pazdur, M. F. Counting statistics in low level radioactivity measurements with fluctuating counting efficiency. Int. J. Appl. Rad. & Isotopes 27: 179-184; 1976.

65. Rogers, V. C. Detection limits for gamma-ray spectral analysis. Anal. Chem. 42: 807; 1970.

66. Bosch, J.; Hartwig, G.; Taurit, R.; Wehinger, H.; Willenbecher, H. Studies on detection limits of radiation protection instruments. GIT Fachz. Lab 25(3): 202, 205-8; 1981.

67. Stevenson, P. C. Processing of counting data, National Academy of Sciences - National Research Council, Nuclear Science Series NAS-NS-3109: Springfield, Va.: Dept. of Commerce; 1965.

68. Sterlinski, S. The lower limit of detection for very short-lived radio-isotopes used in activation analysis. Nuclear Instruments and Methods 68: 341-343; 1969.

69. Gilbert, R. O.; Kinnison, R. R. Statistical methods for estimating the mean and variance from radionuclide data sets containing negative, unreported or less-than values. Health Phys. 40(3): 377-390; 1981.

70. Tschurlovits, M. Definition ot the detection limit for non-selective measurement of low activities. Atomkernenergie 29(4): 266-269; 1977.

71. Plato, P.; Oller, W. L. Minimum-detectable activity of beta-particle emitters by spectrum analysis. Nuclear Instruments and Methods 125: 177-181; 1975

(e) Extension of References (non-classified)

72. Currie, L. A. J. of Radioanalytical Chemistry 39, 223-237 (1977).

73. Swanston, E.; Jefford, J. D.; Krull, R.; and Malm, H. L. "A Liquid Effluent Beta Monitor," IEEE Trans. Nucl. Sci., NS-24 (1977) 629.

74. Evans, R. D. The Atomic Nucleus (McGraw-Hill, New York) (1955), pp. 803 ff.

75. Sumerling, T. J. and Darby, S. C. "Statistical Aspects of the Interpretation of Counting Experiments Designed to Detect Low Levels of Radioactivity," (National Radiological Protection Board, Harwell, Didcot, England), NRPB-R113 (1981) 29 pp.

76. Currie, L. A. "The Many Dimensions of Detection in Chemical Analysis," Chemometrics in Pesticide/Environmental Residue Analytical Determinations, ACS Sympos. Series (1984).

77. Scales, B. Anal. Biochem. 5, 489-496 (1963).

78. Patterson, C. C. and Settle, D. M. 7th Materials Res. Symposium, NBS Spec. Publ. 422, U.S. Government Printing Office, Washington, D.C., 321 (1976).

79. Helstrom, C. W. Statistical Theory of Signal Detection (MacMillan Co., New York) 1960.

80. Currie, L. A. "Accuracy and Merit in Liquid Scintillation Counting," Chapter 18, Liquid Scintillation Counting, Crook, Ed., (Heyden and Son, Ltd., London) p. 219-242 (1976).

81. Parr, R. M.; Houtermans, H.; and Schaerf, K. The IAEA Intercomparison of Methods for Processing Ge(Li) Gamma-ray Spectra, Computers in Activation Analysis and Gamma-ray Spectroscopy, CONF-780421, p. 544 (1979).

82. Tsay, J-Y; Chen, I-W; Maxon, H. R.; and Heminger, L. "A Statistical Method for Determining Normal Ranges from Laboratory Data Including Values below the Minimum Detectable Value," Clin. Chem. 25 (1979) 2011.

83. Ku, H. H. Edit., Precision Measurement and Calibration, NBS Spec. Public. 300 (1969). [See Chapt. 6.6 by M. G. Natrella, "The Relation Between Confidence Intervals and Tests of Significance."]

84. Natrella, M. G. Experimental Statistics, NBS Handbook 91, Chapt. 3 (1963).

85. Nalimov, V. V. Faces of Science, ISI Press, Philadelphia, 1981.

86. Lub, T. T. and Smit, H. C. Anal. Chem. Acta 112 (1979) 341.

87. Frank, I. E.; Pungor, E.; and Veress, G. E. Anal. Chim. Acta 133 (1981) 433.

88. Eckschlager, K. and Stepanek, V. Mikrochim. Acta II (1981) 143.

89. Currie, L. A. and DeVoe, J. R. Systematic Error in Chemical Analysis, Chapter 3, Validation of the Measurement Process, J. R. DeVoe, Ed., American Chemical Society, Washington, DC, 114-139 (1977).

90. Liggett, W. ASTM Conf. on Quality Assurance for Environmental Measurements, 1984, Boulder, CO, In Press.

91. Cohen, A. C. Jr. "Tables for Maximum Likelihood Estimates: Singly Truncated and Singly Censored Samples," _Technometrics_ 3 (1961) 535.

92. Ritter, G. L. and Currie, L. A. _Resolution of Spectral Peaks: Use of Empiricial Peak Shape_, Computers in Activation Analysis and Gamma-ray Spectroscopy. CONF-780421, p. 39 (1979).

93. Murphy, T. J. _The Role of the Analytical Blank in Accurate Trace Analysis_, NBS Spec. Publ. 422, Vol. II, U.S. Government Printing Office, Washington, DC, 509 (1976).

94. Horwitz, W. _Analytical Measurements: How Do You Know Your Results Are Right?_ Special Conference: The Pesticide Chemist and Modern Toxicology (June 1980)

95. Kingston, H. M.; Greenberg, R. R.; Beary, E. S.; Hardas, B. R.; Moody, J. R.; Rains, T. C.; and Liggett, W. S. "The Characterization of the Chesapeake Bay: A Systematic Analysis of Toxic Trace Elements," National Bureau of Standards, Washington, DC, 1983, NBSIR 83-2698.

96. Tukey, J. W. _Exploratory Data Analysis_, Addison-Wesley, Reading, MA (1977).

97. Fukai, R.; Ballestra, S.; and Murray, C. N. _Radioact. Sea_ 35 (1973).

98. Currie, L. A. _Detection and Quantitation in X-ray Fluorescence Spectrometry_, Chapter 25, X-ray Fluorescence Analysis of Environmental Samples, T. Dzubay, Ed., Ann Arbor Science Publishers, Inc., p. 289-305 (1977).

99. Eisenhart, _Science_ 160, 1201 (1968).

100. Fennel, R. W. and West, T. S. _Pure Appl. Chem._ 18, 437 (1969).

Appendix D. Numerical Examples[1]

1. Isolated γ-Ray Peak

Consider a Ge(Li) measurement of an isolated γ-ray, in which a 500 mL - H_2O sample is counted for 200 min, and for which the detection efficiency (cpm-peak/dpm) is 2% absolute. Let us assume that the expected blank rate for the peak region is 2.0 cpm, and that equal numbers of channels are used to estimate the baseline as are used to estimate the gross peak counts. This makes the net peak area estimation calculation exactly equivalent to the "simple" gross-signal-minus-background measurement, with equal counting times. Referring to Figure 8 and Eq's (35)-(37), we see that $n_1 = n_2$ (here 6 channels each), so $\eta = 2$ and $\sigma_0 = \sigma_B\sqrt{2}$.

a) Simplest Case

Ignoring possible systematic error components, the calculations are as follows:

$$Y = 1, \quad E = 0.02, \quad V = 0.5 \text{ L}, \quad T = 200 \text{ min}$$

$$R_B = 2.0 \text{ cpm}, \quad B = R_B T = 400 \text{ counts}$$

$$S_C = 1.645 \, \sigma_0 = 1.645 \, \sigma_B \sqrt{\eta} = 1.645 \sqrt{(400)(2)} = 46.5 \text{ counts}$$

Thus, if the net peak exceeded 46.5 counts one would conclude that a signal had been detected. (Obviously any observed net signal must be an integer, though S_C itself can be a real number.) The detection limit (in counts) is

$$S_D = 2.71 + 2S_C = 95.8 \text{ counts}$$

The concentration detection limit x_D is

$$LLD = x_D = \frac{S_D}{2.22(YEVT)} = \frac{95.8}{(2.22)(1)(0.02)(0.5)(200)} = 21.6 \text{ (pCi/L)}$$

----

[1]All equation numbers refer to Section III of this report, except for example 1g which refers to Section II.

If the mandated LLD [x_R] were a typical 30 pCi/L, the experimental "sensitivity" would be considered adequate.

b) Interference

The above calculation was "pure a priori." Let us suppose, however, that the actual sample being measured exhibited a Compton baseline of 30 cpm over the peak region (6 channels). Everything then becomes scaled by a factor of $\sqrt{30/2}$ (because $\sigma_o \propto \sqrt{B}$). Thus,

$B = R_B T = 6000$ counts

$S_C = 1.645 \sqrt{Bn} = 1.645 (109.5) = 180.2$ counts

$$x_D = \frac{2.71 + 2S_C}{2.22(YEVT)} = \frac{363.1}{4.44} = 81.1 \text{ pCi/L}$$

This exceeds the hypothetical mandated value (30 pCi/L), so we next face the issue of Design -- i.e., change of the Measurement Process, to attain the desired limit.

For long-lived activity in the absence of non-Poisson error, S_C and x_D both decrease as $(YEV)^{-1}$ and as $\sqrt{B}/T = \sqrt{R_B/T}$. A lowered LLD ($x_D$) could be achieved therefore by (1) decreasing the blank rate or increasing the counting time by a factor of $(81.8/30)^2 = 7.43$, or, (2) increasing the product (YEV) by $(81.8/30) = 2.73$. For the present example, neither Y nor R_B may be altered (unless radiochemical separation could be applied to remove the interfering activity); and we shall assume that E is fixed. Increase of the effective volume (possibly via concentration) would probably be the most efficient procedure, but, failing that, the counting time might be extended to 1487 min (~ 1 day).

c) Blank Variability (s_B^2)

To illustrate another point, let us assume that a series of 20 replicate blanks (200 min each) were obtained for which $s_B = 105$ counts, to be compared with the baseline Poisson estimate above,

$$\sigma_B = \sqrt{\bar{B}} = \sqrt{6000} = 77.4 \text{ counts}$$

Thus, $s_B^2/\sigma_B^2 = (105/77.4)^2 = 1.84$, which exceeds the 95 percentile of the χ^2/df distribution (just slightly - see Fig. 4A). We might conclude that this is due to bad luck (chance), or that there is non-random structure associated with the series of blanks, or that there is actually additional (non-Poisson) variability. For this last assumed case, we could use $ts_B\sqrt{n}$ and $2t\sigma_{UL}\sqrt{n}$ for S_C and S_D (bound), resp. (see equations 3-5, and note B2). That is

$$S_C = ts_B\sqrt{n} = 1.73(105)\sqrt{2} = 256.9 \text{ counts}$$

$$S_D = 2S_C (\sigma_{UL}/s) = 2(256.9)(1.37) = 703.9 \text{ counts}$$

and

$$x_D = S_D/(2.22 \text{ YEVT}) = 158.5 \text{ pCi/L}$$

[The factor σ_{UL}/s may be found in Table 6 accompanying note B2.] Thus, the critical level is inflated by roughly 40%, compared to the earlier (Poisson) estimate [S_C(Poisson) = 180.2 counts]; and the Detection Limit is nearly doubled. (Note that S_D and x_D are both upper limits.)

d) Rapid Estimation of LLD, Using Inequality Relations

Following Eq's (8) and (9) in note B4, we can set a limit for LLD directly from an experimental result -- for example, from a weighted least squares (WLS) spectrum deconvolution. Continuing the same example, let us suppose that the result from WLS fitting was

$$\hat{x} \pm \sigma_{\hat{x}} = 95.6 \pm 32.2 \text{ pCi/L}$$

135

Ignoring systematic error for the moment, we would take $\hat{\sigma}_x \geq \sigma_{xo}$.
Therefore,

$$x_C' = 1.645\ \sigma_x = 53.0\ pCi/L \geq x_C$$

$$x_D' = 2x_C' = 106\ pCi/L \geq x_D$$

The result $\hat{x}$ would thus be judged significant (detected), and 106 pCi/L
could be taken as an upper limit for LLD.

e) <u>Calibration and Systematic Blank Error</u>

Continuing with the same example, with interference: B = 6000 counts,
T = 200 min, η = 2, Y = 1, E = 0.02, and V = 0.5 L, we can use Eq. (6) for a
direct estimate of x_D.

$$LLD = x_D = (0.0220)\ BEA + (0.50)\ \frac{3.29\ \sigma_B\sqrt{\eta}}{YEVT}$$

(0.11 has been replaced with 0.0220 because we are treating a baseline rather
than a blank for the purpose of this illustration.) The baseline equivalent
activity is $R_B/(2.22\ YEV)$, or 30 cpm/0.0222 = 1351. pCi/L. Thus, the LLD,
taking a limit of 1% for baseline systematic error (e.g. -- deviation from
the assumed shape) and 10% for possible relative error in (YEV), we obtain

$$LLD = (0.0220)(1351.) + (0.50)\left(\frac{(3.29)\sqrt{(6000)(2)}}{(1)(0.02)(0.5)(200)}\right)$$

$$LLD = 29.7 + 90.1$$

Thus, the Poisson part (90.1/f = 90.1/1.1 = 81.9 pCi/L) is increased by 10%
to account for uncertainty in the multiplicative factors, plus a very
significant 33% (29.7/90.1) to account for possible B uncertainty -- using
Λ_I = 0.01.

f) Limits for LLD Reduction

A finite half-life (such as the 8.05 days for ^{131}I) and the systematic error bounds (f, $\AA_I$) both limit the amount of LLD reduction that can be accomplished through increased counting time. In the above example (x_D = 81.9 pCi/L for t = 200 min), taking $t_{1/2}$ = 8.05 d and $\AA_I$ = 0.01, f = 1.10, it can be shown that with the optimum counting interval (1.8 x $t_{1/2}$, or ~ 2 weeks), the Poisson component of LLD is reduced only to 13.9 pCi/L, and the added contribution from the systematic error bound ($\AA_I$) in the baseline then equals 47.3 pCi/L. (Setting f → 1.1 gives a further increase of 10%.) Thus, for this example, increasing the counting time by about a factor of 100 results in an overall LLD reduction of only ~ 25%!

g) Multiple Detection Decisions

If we wished to compensate for the number of nuclides sought but not found in a multicomponent spectrum search, we should increase S_C (and therefore necessarily LLD) from the above values. For example, if just 10 specific peaks were sought in a given spectrum, and we wished to maintain an overall 5% risk of a (single) false positive, we could employ Eq. 2-35 to calculate the needed adjustment in α and $z_{1-\alpha}$. That would be:

$$\alpha' = 1 - (1 - 0.05)^{0.1} = 0.00512$$

$z_{1-\alpha'}$ is thus 2.57. If we were to similarly decrease the false negative risk (β), both S_C and S_D (and therefore LLD) would be increased by the same factor 2.57/1.645 = 1.56. The resulting x_D for the peak under discussion would be,

$$x_D \to 1.56 \, (81.9) = 128 \text{ pCi/L}$$

2. Simple Beta Counting

Consider the measurement of ^{90}Sr, where $R_B = 0.50$ cpm, $Y = 0.85$, $E = 0.40$, and $t = 1000$ min. (V is irrelevant for this example.) We must consider decay during counting for the 64 hr $(t_{1/2})$ ^{90}y actually measured; and we shall take $\Delta_B = 0.05$, and $f = 1.10$ as before.

The LLD is given by Eq. (6):

$$LLD = (0.11)\ BEA + (0.50)\ \frac{3.29\ \sigma_B\sqrt{\eta}}{YEVT}$$

For this example we shall assume a very long averaged background $(\therefore \eta \approx 1)$, $BEA = (R_Bt)/(2.22\ YEVT)$, and $T = (1-e^{-\lambda t})/\lambda = 915$ min. Thus, $YEVT = (0.85)(0.40)(1)(915) = 311$ min, and

$$LLD = (0.11)\left[\frac{500}{(2.22)(311)}\right] + (0.50)\left[\frac{(3.29)\sqrt{500}}{311}\right]$$

$$= 0.080 + 0.118 = 0.198\ pCi,$$

where the systematic error bounds in the blank and multiplicative factors (5% and 10%, resp.) account for ~46% of the total. That is, with $f \rightarrow 1$ and $\Delta \rightarrow 0$, $LLD = 3.29\ \sqrt{500}/[(2.22)(311)] = 0.106$ pCi. The corresponding decision point x_C is $x_D/(2f)$ or $0.198/2.20 = 0.090$ pCi.

3. Low-Level α-Counting

Assume that a Measurement Process for ^{239}Pu had the following characteristics.

$$R_B = 0.01\ cpm, \quad E = 0.30, \quad Y = 0.80, \quad t = 1\ hr$$

Referring to Table 7, and taking $\bar{B} = 0.60$ counts, we find $y_C = 2$ counts and $y_D = 6.30$ counts. That is, if in a 60 min observation more than 2 counts (gross) were observed, the ^{239}Pu would be considered "detected". The LLD is given by

$$x_D = \frac{(y_D - B)}{2.22(YEVT)} = \frac{(6.30 - 0.60)}{2.22(0.80)(0.30)(60)} = 0.18 \text{ pCi}$$

If R_B were known to only 10% (<u>i.e.</u>, based on 100 counts observed), we could set limits: $B = 0.60 \pm 0.06$ counts, so y_C and y_D remain unchanged, but

$$x_D = \frac{6.30 - (0.60 \pm 0.06)}{2.22(0.80)(0.30)(60)} = 0.178 \pm 0.0019$$

The conservative (upper) limit for x_D thus equals 0.18_0 pCi.

The above estimates could, of course, have been obtained using Fig. 7A.

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned by DDC)	
NUREG/CR-4007	

4. TITLE AND SUBTITLE (Add Volume No., if appropriate)

Lower Limit of Detection: Definition and Elaboration of a Proposed Position for Radiological Effluent and Environmental Measurements

2. (Leave blank)

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)

L. A. Currie

5. DATE REPORT COMPLETED

MONTH	YEAR
August	1984

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

National Bureau of Standards
Washington, D.C. 20234

DATE REPORT ISSUED

MONTH	YEAR
September	1984

6. (Leave blank)

8. (Leave blank)

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

Division of Systems Integration
Office of Nuclear Reactor Regulation
U.S. Nuclear Regulatory Commission
Washington, D. C. 20555

10. PROJECT/TASK/WORK UNIT NO.

11. FIN NO.

B-8615

13. TYPE OF REPORT

Technical

PERIOD COVERED (Inclusive dates)

June 1982 - August 1984

15. SUPPLEMENTARY NOTES

14. (Leave blank)

16. ABSTRACT (200 words or less)

A manual is provided to define and illustrate a proposed use of the Lower Limit of Detection (LLD) for Radiological Effluent and Environmental Measurements. The manual contains a review of information regarding LLD practices gained from site visits; a review of the literature and a summary of basic principles underlying the concept of detection in Nuclear and Analytical Chemistry; a detailed presentation of the application of LLD principles to a range of problem categories (simple counting to multinuclide spectroscopy), including derivations, equations, and numerical examples; and a brief examination of related issues such as reference samples, numerical quality control, and instrumental limitations. An appendix contains a summary of notation and terminology, a bibliography, and worked-out examples.

17. KEY WORDS AND DOCUMENT ANALYSIS

17a. DESCRIPTORS

Lower Limit of Detection, LLD, Effluent Measurement, Environmental Measurement, Radiological Monitoring, RETS.

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT

Unlimited

19. SECURITY CLASS (This report)	21 NO OF PAGES
Unclassified	
20. SECURITY CLASS (This page)	22 PRICE
Unclassified	S