Statistical Issues in Applied Health Physics

Daniel J. Strom
Pacific Northwest National Laboratory
Richland, Washington USA
strom@pnl.gov +1 (509) 375-2626

Health Physics Society Professional Development School
Corvallis, Oregon
Saturday, July 14, 2007

Battelle operates Pacific Northwest National Laboratory for
the U.S. Department of Energy under Contract No. DE-AC05-76RLO 1830

Battelle

Pacific Northwest National Laboratory

Outline (First 8 Hours)

- HPS 2000 Summer School was devoted entirely to applications of probability and statistics in health physics (Borak 2000a)
- Focus only on a few aspects of the problem
  - reliance on earlier proceedings
  - links to much excellent material available for free over the internet
- Review of general statistical concepts from ISO GUM and other sources
  - meanings of probability
  - statistical distributions
  - population versus sample statistics
  - variability, uncertainty, bias, error, and blunder
  - correlation and "shared" uncertainty
  - the source of uncertainty
  - expressions of uncertainty
- Numerical thresholds for decisions
  - e.g., whether radioactive material is present above background levels
  - missed dose
  - unmonitored dose
- Case studies
  - reconstructing x-ray doses
  - reconstructing worker doses
  - uncertainty in internal dosimetry

1995 ISO Guide to the Expression of Uncertainty in Measurement (GUM)

- Extensive, well-thought-out framework for dealing with uncertainty in measurement
  - Clearly-defined concepts and terms
  - Practical approach
  - Some new words may not exist in Russian
- Doesn't cover
  - the use of measurements in models that have uncertain
    - assumptions
    - parameters
    - form
  - representativeness (e.g., of a breathing-zone air sample)

Meanings of Probability: 1 – ***Classical***

- The possible outcomes of an event are analyzed into a set of elementary, equally probable events
- The probability of a particular outcome is then the number of elementary events producing the outcome divided by the total number of elementary events" (Little 2000)
- Also called mathematical, formal, or axiomatic probability (Kaplan 1997)

Meanings of Probability: 2 – ***Frequentist***

- The probability of an outcome is its frequency in the limit of an infinite number of trails (Little 2000)
- This frequency, or fraction, definition has at least nine other incarnations (Kaplan 1997):

  – randomness
  – variability
  – aleatory probability
  – objective probability
  – in the world probability
  – stochastic ontological probability
  – reliability
  – chance
  – risk

Meanings of Probability: 3 – ***Subjective***

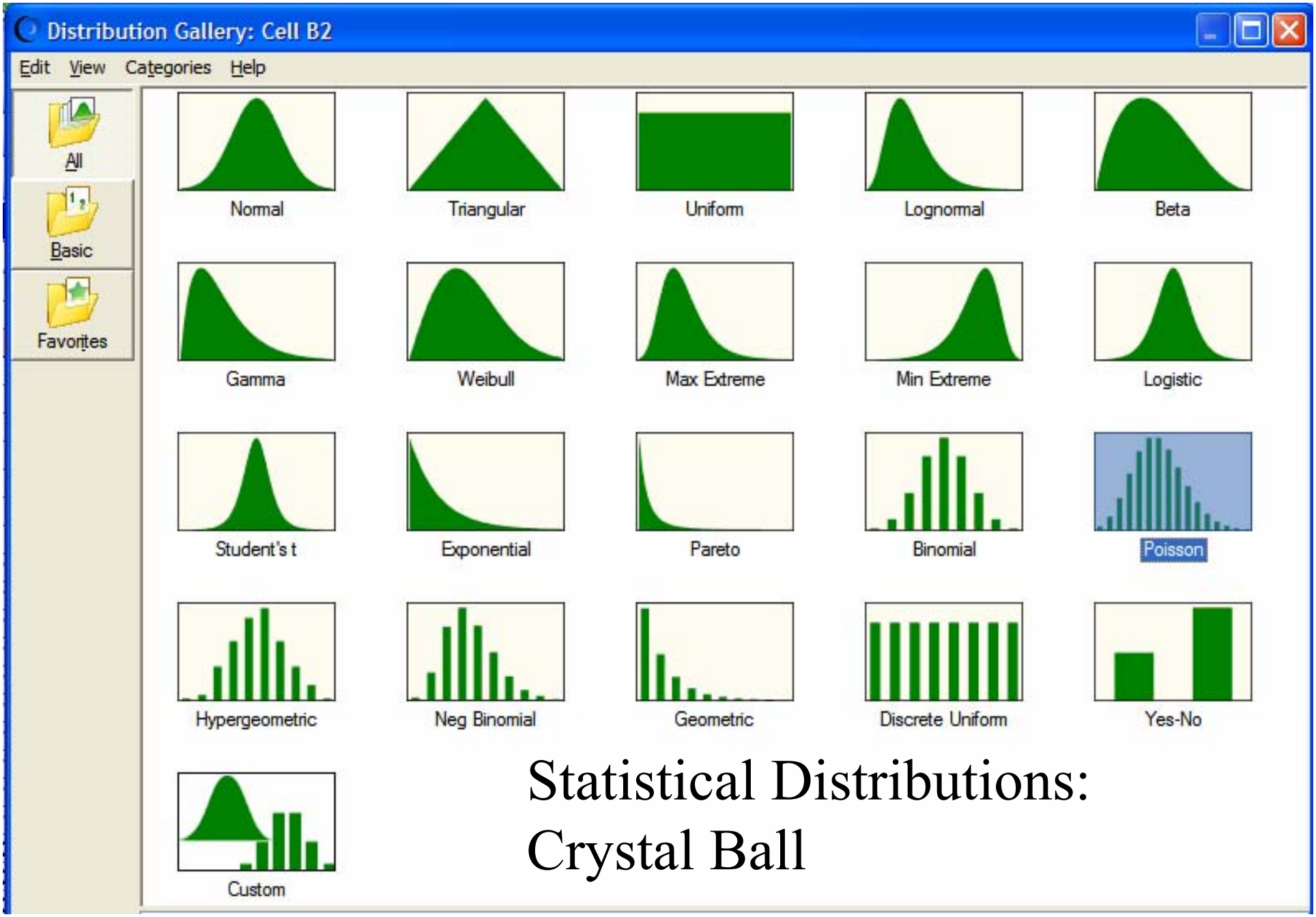
- The probability of an event is based on the observer's assessment of a coherent bet
- Different observers need not agree on the probability of an outcome (Little 2000)
- Synonyms (Kaplan 1997)
  - Bayesian probability
  - belief
  - personal probability
  - uncertainty
  - confidence
  - epistemic probability
  - plausibility
  - credibility
  - evidence-based probability
- "Probability is that degree of credibility or confidence dictated by the evidence through Bayes' theorem"

Meanings of Probability: ISO GUM (1995)

- [Probability is a] real number in the scale of 0 to 1 attached to a random event
- Note – it can be related to a long-run relative frequency of occurrence or to a degree of belief that an event will occur
- For a high degree of belief, the probability is near 1
- *This definition avoids controversy*

Statistical Distributions

- Borak (2000)
- Crystal Ball (2007)



Statistical Distributions: Crystal Ball

Population Parameters versus Sample Statistics

- A critical distinction
- Greek letters denote population parameters
- Roman letters to denote samples from populations
- Standard deviation
  - σ: population SD
  - s: SD of sample from that population
- when people talk about "2 sigma," for example, using "sigma" as a synonym for standard deviation, the distinction between a population parameter and a sample statistic is lost

Population Mean μ and Sample Mean $\bar{x}$

- arithmetic mean height of a population consisting of all n of the participants in this PDS:

$$\mu = \frac{1}{n}\sum_{1}^{n} x_i ,$$

where x_i is the height of the ith individual

- μ can be estimated by measuring the height of k randomly selected people from the population:

$$\bar{x} = \frac{1}{k}\sum_{1}^{k} x_i .$$

1995 ISO GUM General Metrological Terms - 1

ISO-GUM Term	Meaning
(measurable) quantity	attribute of a phenomenon, body, or substance that may be distinguished qualitatively and determined quantitatively
value (of a quantity)	magnitude of a particular quantity generally expressed as a unit of measurement multiplied by a number
true value (of a quantity)	value consistent with the definition of a given particular quantity. *Not used in the GUM; "value of a measurand" is used.*
conventional true value (of a quantity)	value attributed to a particular quantity and accepted, sometimes by convention, as having an uncertainty appropriate for a given purpose
measurement	set of operations having the object of determining a value of a quantity

1995 ISO GUM General Metrological Terms - 2

ISO-GUM Term	Meaning
principle of measurement	scientific basis of a measurement
method of measurement	logical sequence of operations, described generically, used in the performance of measurements
measurement procedure	set of operations, described specifically, used in the performance of particular measurements according to a given method
measurand	particular quantity subject to measurement. [the unknown value of a physical quantity representing the "true state of Nature"]
influence quantity	quantity that is not the measurand but that affect the result of the measurement

1995 ISO GUM General Metrological Terms - 3

ISO-GUM Term	Meaning
result of a measurement	value attributed to a measurand, obtained by measurement
uncorrected result	result of a measurement before correction for systematic error *(i.e., bias)*
corrected result	result of a measurement after correction for systematic error *(i.e., bias)*
accuracy of measurement	closeness of the agreement between the result of a measurement and a true value of the measurand
repeatability (of results of measurements)	closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement

1995 ISO GUM General Metrological Terms - 4

ISO-GUM Term	Meaning
reproducibility (of results of measurements)	closeness of agreement between the results of measurements of the same measurand carried out under changed conditions of measurement
experimental standard deviation	for a series of n measurements of the same measurand, the quantity $s(q_k)$ characterizing the dispersion of the results and given by the formula $$s(q_k) = \sqrt{\frac{\sum_{k=1}^{n}(q_k - \bar{q})^2}{n-1}}$$, q_k being the result of the kth measurement and being the arithmetic mean of the n results considered. Note that is an estimate of the standard deviation of the distribution of and is called the experimental standard deviation of the mean (not the standard error of the mean)

1995 ISO GUM General Metrological Terms - 5

ISO-GUM Term	Meaning
uncertainty (of measurement)	parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. *It is a bound for the likely size of the measurement error.*
error (of measurement)	result of a measurement minus a true value of the measurand *(i.e., the [unknowable] difference between a measured result the actual value of the measurand.)* "Error is an idealized concept and errors cannot be known exactly" (Note 3.2.1)
relative error	error of measurement divided by a true value of the measurand
random error	result of a measurement minus the mean that would result from an infinite number of measurements of the measurand carried out under repeatability conditions

1995 ISO GUM General Metrological Terms - 6

ISO-GUM Term	Meaning
systematic error	mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand
correction	value added algebraically to the uncorrected result of a measurement to compensate for systematic error
correction factor	Numerical factor by which the uncorrected result of a measurement is multiplied to compensate for systematic error

1995 ISO GUM Basic Statistical Terms & Concepts - 1

ISO-GUM Term	Meaning
probability	A real number in the scale of 0 to 1 attached to a random event. Note – it can be related to a long-run relative frequency of occurrence or to a degree of belief that an event will occur. For a high degree of belief, the probability is near 1.
random variable; variate	A variable that may take any of the values of a specified set of values and with which is associated a *probability distribution*
probability distribution (of a random variable)	A function giving the probability that a random variable takes any given value or belongs to a given set of values
distribution function	A function giving, for every value x, the probability that the random variable X will be less than or equal to x: $F(x) = \Pr(X \leq x)$

1995 ISO GUM Basic Statistical Terms & Concepts - 2

ISO-GUM Term	Meaning
probability mass function	A function giving, for each value x_i of a discrete random variable X, the probability p_i that the random variable equals x_i: $p_i = \Pr(X = x_i)$
parameter	A quantity used in describing the probability distribution of a random variable
correlation	The relationship between two or several random variables within a distribution of two or more random variables
expectation (of a random variable or of a probability distribution); expected value; mean	For a discrete random variable X taking the values x_i with the probabilities p_i, the expectation, if it exists, is $\mu = E(X) = \sum p_i x_i$. For a continuous random variable X having the probability density function $f(x)$, The expectation, if it exists, is $\mu = E(X) = \int x\, f(x)\, \mathrm{d}x$, the integral being extended over the interval(s) of variation of X.

1995 ISO GUM Basic Statistical Terms & Concepts - 3

ISO-GUM Term	Meaning
centered random variable	A random variable the expectation of which equals zero
variance (of a random variable or of a probability distribution)	The expectation of the square of the centered random variable: $\sigma^2 = V(X) = E\{[X - E(X)]^2\}$
standard deviation (of a random variable or of a probability distribution)	The positive square root of the variance: $\sigma = \sqrt{V(X)}$
central moment of order q	In a univariate distribution, the expectation of the qth power of the centered random variable (X - μ): $E[(X - \mu)^q]$

1995 ISO GUM Basic Statistical Terms & Concepts - 4

ISO-GUM Term	Meaning
normal distribution; Laplace-Gauss distribution	The probability distribution of a continuous random variable X, the probability density function of which is $$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right)$$ for $-\infty < x < \infty$.
characteristic	A property which helps to identify or differentiate between items of a given population
population	The totality of items under consideration
frequency	The number of occurrences of a given type of event or the number of observations falling into a specified class
frequency distribution	The empirical relationship between the values of a characteristic and their frequencies or relative frequencies

1995 ISO GUM Basic Statistical Terms & Concepts - 5

ISO-GUM Term	Meaning
arithmetic mean; average	the sum of values divided by the number of values: $$\bar{x} = \frac{1}{n}\sum x_i$$

Non- ISO-GUM Term	Meaning
geometric mean	the *n*th root of the product of *n* values: $$x_{gm} = \exp\left(\frac{1}{n}\sum \ln x_i\right)$$ For 2 values, $x_{gm} = \sqrt{x_1 x_2}$
median	the value in the middle of a distribution, such that there is an equal number of values above and below the median. Also known as the 50th percentile, x_{50}
mode	the most frequently occurring value

1995 ISO GUM Basic Statistical Terms & Concepts - 6

ISO-GUM Term	Meaning
variance; sample estimate of the variance	A measure of dispersion, which is the sum of the squared deviations of observations from their average divided by one less than the number of observations: $$s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2 .$$ Note – the sample variance is an unbiased estimator of the population variance.
(sample) standard deviation	The positive square root of the variance. Note – the sample standard deviation is a biased estimator of the population standard deviation.
central moment of order q	In a distribution of a single characteristic, the arithmetic mean of the qth power of the difference between the observed values and their average $\bar{x}$: $\frac{1}{n}\sum(x_i - \bar{x})^q$ where n is the number of observations.

1995 ISO GUM Basic Statistical Terms & Concepts - 7

ISO-GUM Term	Meaning
statistic	A function of the sample random variables. Note – a statistic, as a function of random variables, is also a random variable and as such it assumes different values from sample to sample. The value of the statistic obtained by using the observed values in this function may be used in a statistical test or as an estimate of the population parameter, such as a mean or a standard deviation.
estimation	The operation of assigning, from the observations in a sample, numerical values to the parameters of a distribution chosen as the statistical model of the population from which this sample is taken
estimator	A statistic used to estimate a population parameter
estimate	The value of an estimator obtained as a result of an estimation.

1995 ISO GUM Basic Statistical Terms & Concepts - 8

ISO-GUM Term	Meaning
two-sided confidence interval	When T_1 and T_2 are two functions of the observed values such that, θ being a population parameter to be estimated, the probability $\Pr(T_1 \leq \theta \leq T_2)$ is at least equal to $(1 - \alpha)$ [where $(1 - \alpha)$ is a fixed number, positive and less than 1], the interval between T_1 and T_2 is a two-sided $(1 - \alpha)$ confidence interval for θ.
one-sided confidence interval	When T is a function of the observed values such that, θ being a population parameter to be estimated, the probability $\Pr(T \geq \theta)$ [the probability $\Pr(T \leq \theta)$] is at least equal to $(1 - \alpha)$ [where $(1 - \alpha)$ is a fixed number, positive and less than 1], the interval from the smallest possible value of θ up to T (or the interval from T up to the largest possible value of θ) is a one-sided $(1 - \alpha)$ confidence interval for θ.

1995 ISO GUM Basic Statistical Terms & Concepts - 9

ISO-GUM Term	Meaning
confidence coefficient; confidence level	The value $(1-\alpha)$ of the probability associate with a confidence interval or a statistical coverage interval.
statistical coverage interval	An interval for which it can be stated with a given level of confidence that it contains at least a specified proportion of the population
degrees of freedom	In general, the number of terms in a sum minus the number of constraints on the terms of the sum
expectation	-
variance	-
standard deviation	-

1995 ISO GUM Basic Statistical Terms & Concepts - 10

ISO-GUM Term	Meaning
covariance	The covariance of two random variables is a measure of their mutual dependence. The covariance of random variables y and z is defined by $\mathrm{cov}(y,z) = \mathrm{cov}(z,y) = E\{[y - E(y)][z - E(z)]\}$, which leads to $\mathrm{cov}(y,z) = \mathrm{cov}(z,y)$ $= \iint (y-\mu_y)(z-\mu_z)p(y,z)\,\mathrm{d}\,y\,\mathrm{d}\,z$ $= \iint yzp(y,z)\,\mathrm{d}\,y\,\mathrm{d}\,z - \mu_y\mu_z$ where $p(y, z)$ is the joint probability density function of the two variables y and z. The covariance $\mathrm{cov}(y, z)$ [also denoted by $v(y,z)$] may be estimated by $s(y_i,z_i)$ obtained from n independent pairs of simultaneous observations y_i and z_i of y and z, $s(y_i, z_i) = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z})$

1995 ISO GUM Basic Statistical Terms & Concepts - 11

ISO-GUM Term	Meaning
covariance matrix	-
correlation coefficient	The correlation coefficient is a measure of the relative mutual dependence of two variables, equal to the ratio of their covariances to the positive square root of the product of their variances. Thus, $\rho(y,z) = \rho(z,y) = \frac{\upsilon(y,z)}{\sqrt{\upsilon(y,y)\upsilon(z,z)}} = \frac{\upsilon(y,z)}{\sigma(y)\sigma(z)}$ with estimates $r(y_i,z_i) = r(z_i,y_i) = \frac{s(y_i,z_i)}{\sqrt{s(y_i,y_i)s(z_i,z_i)}} = \frac{s(y_i,z_i)}{s(y_i)s(z_i)}$ The correlation coefficient is a pure number such that $-1 \le \rho \le +1$ or $-1 \le r \le +1$.
independence	Two random variables are statistically independent if their joint probability distribution is the product of their individual probability distributions.

1995 ISO GUM Basic Statistical Terms & Concepts - 12

ISO-GUM Term	Meaning
the *t*-distribution; Student's distribution	The *t*-distribution or Student's distribution is the probability distribution of a continuous random variable *t* whose probability density function is $$\rho(t,\nu) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]} \left[1+\frac{t^2}{\nu}\right]^{-(\nu+1)/2},$$ $$-\infty < t < +\infty$$ where Γ is the gamma function and $\nu > 0$. The expectation of the *t*-distribution is zero and its variance is $\nu/(\nu-2)$ for $\nu > 2$. As $\nu \to \infty$, the *t*-distribution approaches a normal distribution with $\mu = 0$ and $\sigma = 1$. The probability distribution of the variable is the *t*-distribution if the random variable *z* is normally distributed with expectation μ_z, where $(\bar{z} - \mu_z)/s(\bar{z})$ is the arithmetic mean of *n* independent observations z_i of *z*, $s(z_i)$ is the experimental standard deviation of the *n* observations, and $s(\bar{z}) = s(z_i)/\sqrt{n}$ is the experimental standard deviation of the mean with $\nu = n - 1$ degrees of freedom

1995 ISO GUM Additional Terms & Concepts - 1

ISO-GUM Term	Meaning
blunder	"Blunders in recording or analyzing data can introduce a significant unknown error in the result of a measurement. Large blunders can usually be identified by a proper review of all the data; small ones could be masked by, or even appear as, random variations. Measures of uncertainty are not intended to account for such mistakes." (3.4.7) Other terms include mistake and spurious error.
"Type A" uncertainty evaluation	uncertainty that is evaluated by the statistical analysis of series of observations
"Type B" uncertainty evaluation	uncertainty that is evaluated by means other than the statistical analysis of a series of observations

1995 ISO GUM Additional Terms & Concepts - 2

ISO-GUM Term	Meaning
combined standard uncertainty	standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances of covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities.
expanded uncertainty	quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand
coverage factor	numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

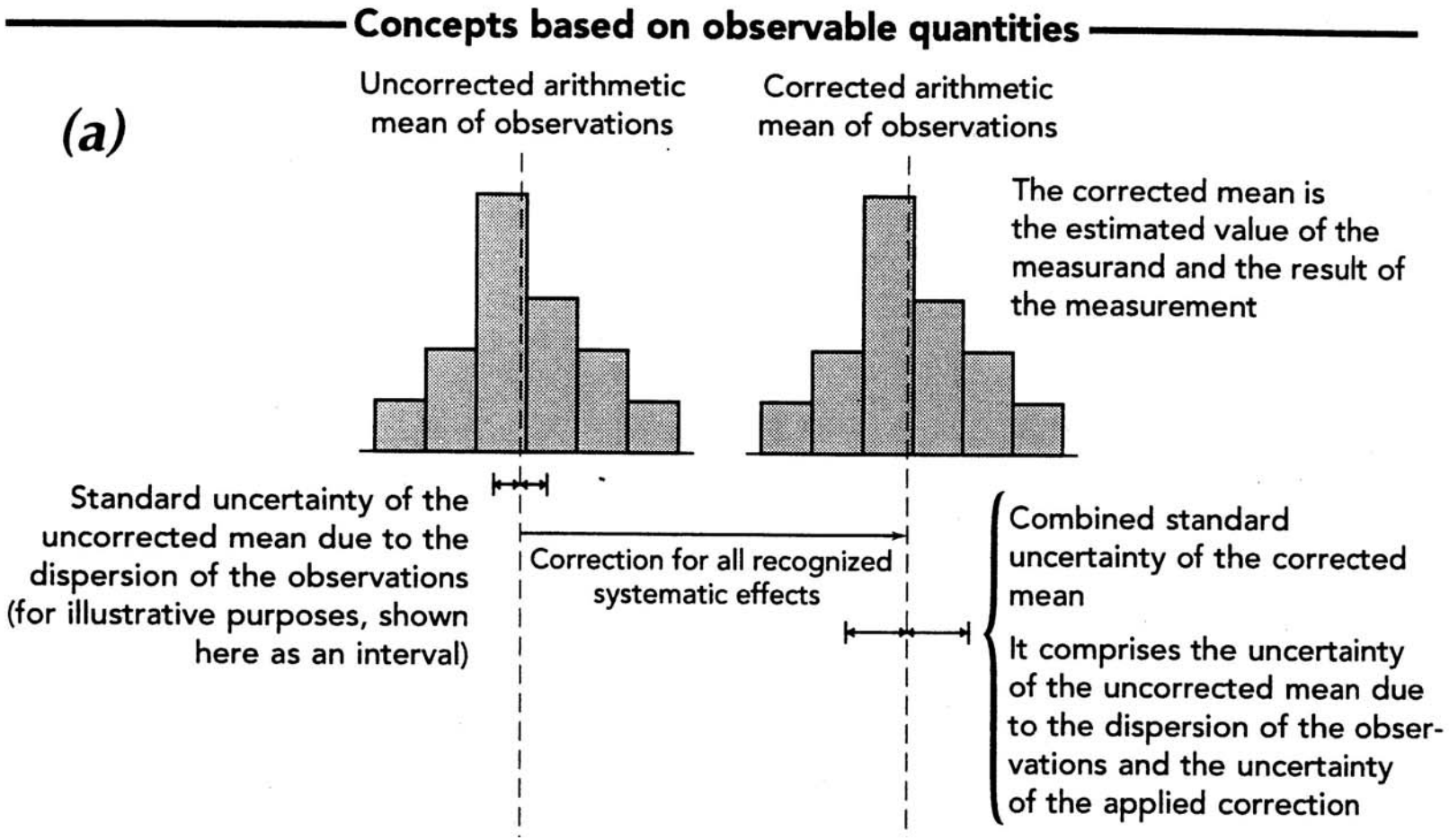
Type A and Type B Uncertainty

- Uncertainty that is evaluated by the statistical analysis of series of observations is called a "**Type A**" uncertainty evaluation.
- Uncertainty that is evaluated by means *other* than the statistical analysis of a series of observations is called a "**Type B**" uncertainty evaluation.

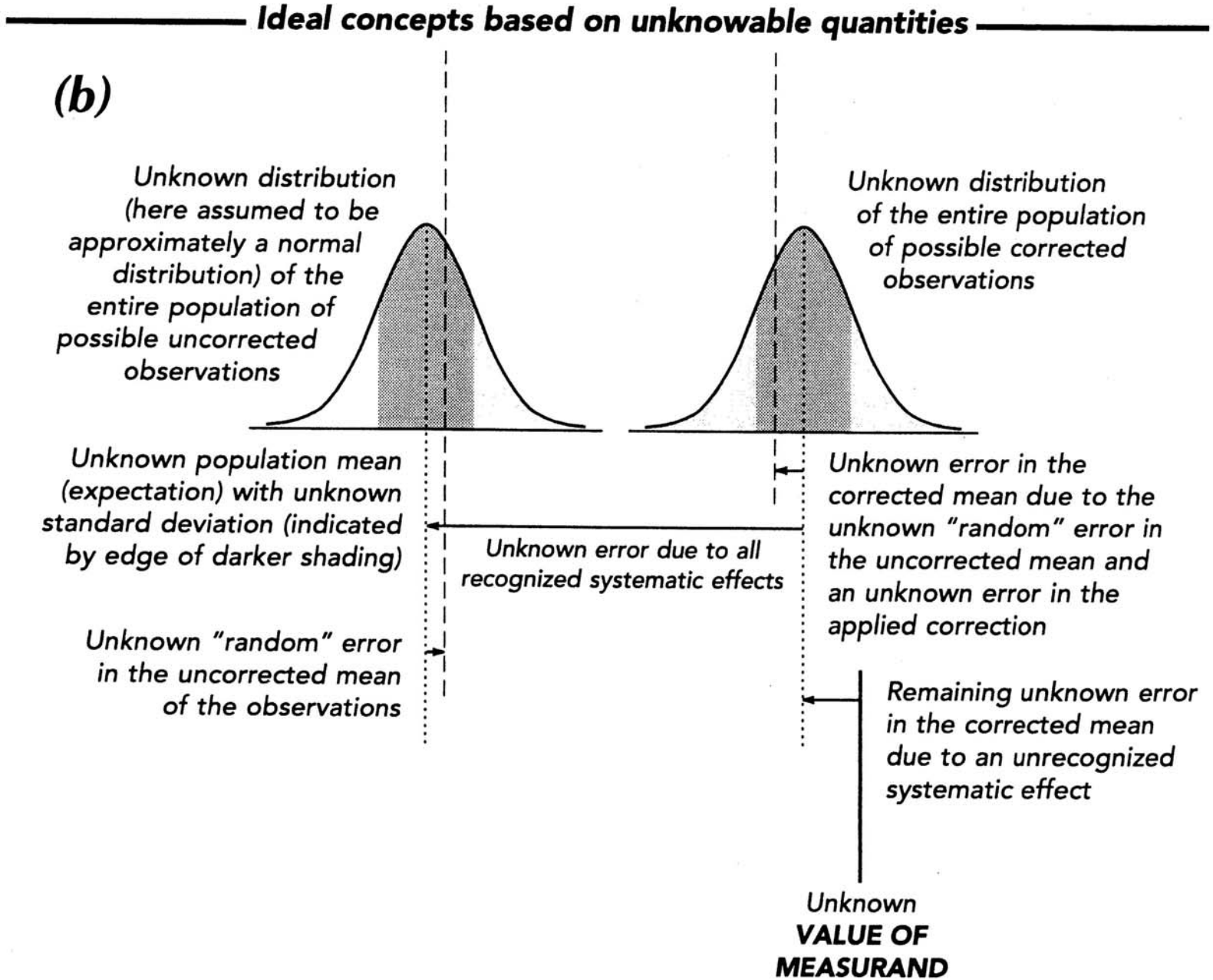
Terms: Error, Uncertainty, Variability

- "The difference between error and uncertainty should always be borne in mind."
- "For example, the result of a measurement after correction can unknowably be very close to the unknown value of the measurand, and thus have negligible error, even though it may have a large uncertainty."
- "Error bars?" No! "Uncertainty bars" is what we should say
- Variability is the range of values for different individuals in a population
  - e.g., height, weight, metabolism

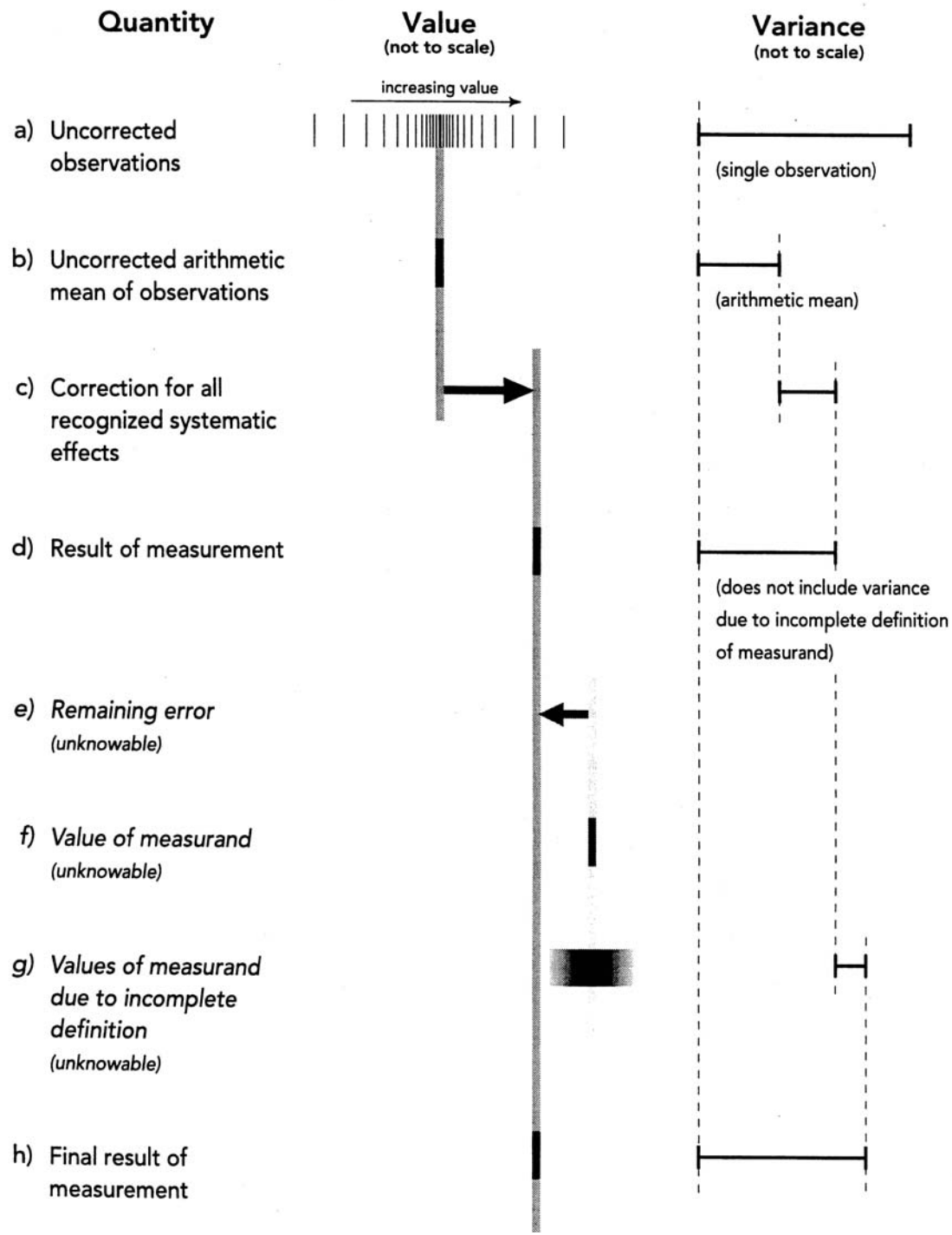
Graphical Illustration of Value, Error, and Uncertainty



Graphical Illustration of Value, Error, and Uncertainty



Graphical Illustration of Value, Error, and Uncertainty



Random and Systematic "Errors"

ISO-GUM Term	Meaning
random error	result of a measurement minus the mean that would result from an infinite number of measurements of the measurand carried out under repeatability conditions
systematic error	mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand

Random and Systematic Uncertainty versus Type A and Type B Uncertainty Evaluation

- GUM: There is not always a simple correspondence between the classification of uncertainty components into categories A and B and the commonly used classification of uncertainty components as “random” and “systematic.”
- The nature of an uncertainty component is conditioned by the use made of the corresponding quantity, that is, on how that quantity appears in the mathematical model that describes the measurement process.
- When the corresponding quantity is used in a different way, a “random” component may become a “systematic” component and vice versa.

Random and Systematic Uncertainty

- Thus the terms “random uncertainty” and “systematic uncertainty” can be misleading when generally applied.
- An alternative nomenclature that might be used is “component of uncertainty arising from a random effect,” “component of uncertainty arising from a systematic effect,” where a random effect is one that gives rise to a possible random error in the current measurement process and a systematic effect is one that gives rise to a possible systematic error in the current measurement process. In principle, an uncertainty component arising from a systematic effect may in some cases be evaluated by method A while in other cases by method B (see subsection 2.2), as may be an uncertainty component arising from a random effect.

Type A Uncertainty Evaluation

- represented by a statistically estimated standard deviation

$$s_i = \left| \sqrt{s_i^2} \right|$$

- associated number of degrees of freedom = v_i.
- the standard uncertainty is $u_i = s_i$.

Type B Uncertainty Evaluation

- represented by a quantity u_j

$$u_j \approx \text{corresponding standard deviation;}$$

$$u_j = \left|\sqrt{u_j^2}\right|;$$

$$u_j^2 \approx \text{corresponding variance obtained from}$$

an assumed probability distribution

based on all the available information

- Since the quantity u_j^2 is treated like a variance and u_j like a standard deviation, for such a component the standard uncertainty is simply u_j.

Combined Standard Uncertainty

The First Step

- Must know what y depends on, and how:

$$y = G(x_1, x_2, \ldots, x_n)$$

Uncertainty Propagation Formula

- Combined standard uncertainty

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j)$$

$$= \text{Sum of variance terms and covariance terms}$$

- Derived from first-order Taylor series expansion
- Not accurate for large uncertainties (e.g., broad lognormal distributions)
- Covariances usually unknown and ignored

Uncertainty Propagation Formula – 2

- Formulation using correlation coefficient $r(x_i,x_j)$

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i)$$

$$+ 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} r(x_i, x_j) s(x_i) s(x_j)$$

- See Rolf Michel's wipe test example:

  http://www.kernchemie.uni-mainz.de/downloads/saagas21/michel_2.pdf

Numerical Methods

- Monte Carlo simulations, with covariances, may be needed to explore uncertainty
- Crystal Ball® does this easily

Inferring Dose from Measurements, Models, Assumptions, or None of the Above

Dosimetry, Dosinference, and Doswaggery

Radiation Protection Dosimetry
Vol. 98, No. 4, pp. 363-366 (2002)
Nuclear Technology Publishing

Guest Editorial

Dosimetry, Dosinference and Doswaggery

For ionising radiation, absorbed dose is defined as the energy imparted per unit mass of an infinitesimal volume. While straightforward to define, absorbed dose is somewhat more difficult to measure or infer. There is growing attention to the uncertainty in values of absorbed dose and its derivative quantities such as equivalent dose and effective dose. The process of arriving at a value of dose to record is variously called measuring a dose, computing a dose, calculating a dose, assessing a dose, evaluating a dose, estimating a dose, reconstructing a dose, inferring a dose, or doing dosimetry. All values for the quantity dose are based on some amount of inference, but the effect of the inference on uncertainty may range from trivial to enormous. The more inferential steps required to arrive at a dose value, the greater the uncertainty associated with the result. It is suggested here that the word dosimetry is being stretched too far when it is applied to inferential processes whose total uncertainty is little affected by measurement uncertainty, but is dominated by uncertainty in models, their parameters, and imputed data that have little or no basis in measurement.

Word Choice Based on Uncertainty

Term	Typical Dominant Uncertainty				Ratio of 97.5%^{ile} to 2.5%^{ile} of Inferred Dose
	Measure-ments	Models	Model Parameters	Imputed Data	
Dosimetry	✓✓	✓	~	~	1.01 to 2
Dosinference	✓	✓✓	✓✓	~	2 to 20
Doswaggery	~	✓	✓	✓✓	>20

✓ denotes important; ✓✓ denotes very important; ~ relatively trivial

Calling a Spade a Spade…

- maybe it's time to choose different words when the dose in question is measured, inferred, or essentially assumed
- *dosimetry* when measurement uncertainty predominates
- *dosinference* when model parameter uncertainty predominates
- *doswaggery* when assumption or imputed value uncertainty predominates

Reporting and Recording of Measurement Results

“Censoring” of Data

- Censoring data means changing measured results from numbers to some other form that cannot be added or averaged or analyzed numerically
- Examples of data censoring
  - Left-censoring
    - changing results that are less than some value to zero
    - changing results that are less than some value to “less than” some value
  - Right-censoring
    - changing values from the measured result to “greater than” some value
  - Rounding

Why should censoring of data be avoided?

- Censoring means changing the numbers
- In a sense, it is dishonest
- If results are ever
  - summed,
  - averaged, or
  - used for some other aggregate analysis such as fitting a distribution,

  censoring makes this
  - difficult,
  - impossible, or
  - simply biased.

Censoring Examples

- Five results for discharge from a pipe taken over 1 year
  - uncensored results: –2, –1, 0, 1, and 2
  - sum = 0 (total discharge for the year is 0)
  - average = 0 (average discharge for the year is 0)
- Example 1: Set negative values to zero
  - censored results: 0, 0, 0, 1, and 2
  - sum = 3 (i.e., total discharge for the year is 3; this is not true)
  - average = 0.6 (i.e., average discharge for the year is 0.6; false)
- Example 2: Suppose $L_C = 2$. Set all values < 2 to "<"
  - censored results: <, <, <, <, and 2
  - sum = ? (total discharge for the year cannot be determined)
  - average = ? (average discharge for the year cannot be determined)

But Negative Activity Is Meaningless…

- No, it's not meaningless
- Just like money, subtracting a big number from a small number gives a negative value
  - You have 100€, you charge 200€, you owe 100€
  - 100€ – 200€ = –100€ (your net value)
  - this doesn't mean you can find a bank note for –100€
  - stocks go up and down; the end of the year value includes all changes, positive and negative
- Negative activity only means that random statistical fluctuations resulted in a negative number
- If negative, zero, or less-than values are suppressed, the sum is biased.

More Reasons Not to Censor

- Upper confidence limits of negative, zero, or less-than values
  - may be small positive numbers
  - needed for some applications (e.g., probability of causation)
- Censoring is prohibited by many standards and regulations
  - ANSI N13.30-1996: "Results obtained by the service laboratory shall be reported to the customer and shall include the following items …quantification using appropriate blank values of radionuclides whether positive, negative, or zero"
  - Many U.S. Department of Energy regulations require reporting raw data, calculated results (positive, negative, or zero), and total propagated uncertainties
  - Decision on actions can be made with uncensored data

Rounding Is Censoring

- Rounding a number is
  - changing its value
  - biasing the value
  - censoring
- Rounding often "justified" by claiming uncertainty
  - Uncertainty does not justify changing the answer
  - Explicitly state the uncertainty
- Beware of converting units of a rounded number and then rounding again!
- Intermediate results and laboratory records should never be rounded
- The only time to round is in presentations or communications

Report and Record All Measurements with No Censoring and Minimal Rounding

Expressing What We Know About a Quantity

What We Really Need to Know About a Number

1. the quantity
2. the unit
3. how the number was obtained (measurement, calculation using 1 or more measurements, model, estimate, …)
4. the value
5. the uncertainty
6. the kind of uncertainty (standard deviation, geometric standard deviation, range)
7. how the uncertainty was obtained (e.g., repeated measurements, calculations, models, estimate…)

Example of 7-Vector for 2 Numbers

Component of 7-Vector	External Irradiation	Intake
Quantity	Deep Dose Equivalent	Committed Effective Dose
Unit	mSv	mSv
How obtained	OSL Dosimeter	Serial Pu/Am Urine & Fecal Bioassay; ICRP 68 Models; IMBA
Value	0.13	42
Uncertainty	0.03	3.5
Type of Uncertainty	S.D.	G.S.D.
How uncertainty obtained	ISO T.P.U.	Multi-tracer method; Resampling Statistics

The Two Aspects of the Counting Problem

The Two Counting Problems

- Radioactive decay is a Bernoulli process described by a binomial or Poisson distribution
  - A Bernoulli process is one concerned with the count of the total number of independent events, each with the same probability, occurring in a specified number of trials
- The "forward problem"
  - from properties of the process, we predict the distribution of counting results (mean, standard deviation (SD))
  - measurand → distribution of possible observations
- The "reverse problem"
  - measure a counting result
  - from the counting result, we infer the parameters of the underlying binomial or Poisson distribution (mean, SD)

    see, e.g., Rainwater and Wu (1947)
  - this is the problem we're really interested in!

A Fundamental Problem of Statistical Inference

- Aristotle's syllogisms
  - All humans are mortal
  - Socrates is human
  - Therefore, Socrates is mortal
- Necessary prerequisite: the premises must be absolutely certain
- This reasoning becomes invalid when applied to probabilistic premises*

**Beck-Bornholdt, H.P. Is the Pope an Alien?, Nature 381(27 June), 730 (1996).*

Reasoning becomes invalid when applied to probabilistic premises

- Is the Pope an Alien?
  - Probabilistic premise: If we randomly pick a human being, chances are 1 in 6 billion it's the pope ($p \approx 0.000\ 000\ 000\ 17$)
  - If an individual is human, it's probably not the pope ($p \approx 0.000\ 000\ 000\ 17$)
  - Benedict XVI is the Pope
  - Therefore, he is not a human being ($p \approx 0.000\ 000\ 000\ 17$)
  - Which is obviously not sensible!
- Change from absolute certainty to probability makes syllogistic reasoning false

Statistical Hypothesis Testing

- Usual line of reasoning
  - If the null hypothesis is true, these data are unlikely ($p < 0.05$)
  - the data have occurred
  - therefore the null hypothesis is wrong ($p < 0.05$)
- Aristotle noticed over 2,000 years ago that this type of inference is wrong

Two Kinds of Statistics

- Classical statistics
  - does the forward problem well
  - does not really do the reverse problem
- Bayesian statistics does the reverse problem using
  - a prior probability distribution
  - the observed results
  - a likelihood function (a classical expression of the forward problem)

The Forward Problem

The Forward Problem

- Use Poisson statistics to predict the distribution of observations from a given value of the measurand
- The measurand is best thought of as a count rate ρ
  - otherwise it is difficult to deal with different counting times
- The observable is a number of counts, N, sampled
  - from a Poisson distribution
  - during time t
  - with mean ρt
- $\mathrm{Var}(\mathrm{Poi}(N \mid \rho t)) = \rho t$

The Forward Problem: A Bernoulli Process

1. It consists of *M* trials (i.e., *M* atoms each having a chance to transition)
2. Each trial has a binary outcome: success of failure (transition or not)
3. The probability of success (transition) is constant from trial to trail (all atoms have an equal chance to transition)
4. The trials are independent

Turner JE. 1995. *Atoms, Radioactivity, and Radiation Protection,* 2nd ed., p. 290

Binomial Distribution - 1

- Consider

M radioactive atoms, each with probability

$p = 1 - e^{-\lambda t}$ of decaying in time t

- Probability that exactly N will decay in time t is

$$Bi(N \mid M, p) = \binom{M}{N} p^N (1-p)^{M-N}$$

where the binomial coefficient $\binom{M}{N} \equiv \dfrac{M!}{N!(M-N)!}$

Binomial Distribution - 2

$$Bi \text{ is normalized}: \quad \sum_{N=0}^{M} Bi(N \mid M, p) = 1$$

$$\text{Mean}: \quad \mu = Mp$$

$$\text{Standard deviation}: \quad \sigma = \sqrt{Mp(1-p)}$$

- Excellent discussion in Turner (1995)

Poisson Distribution - 1

- If

  $M >> 1$ (lots of radioactive atoms)

  $M >> N$ (not too many of them decay), and

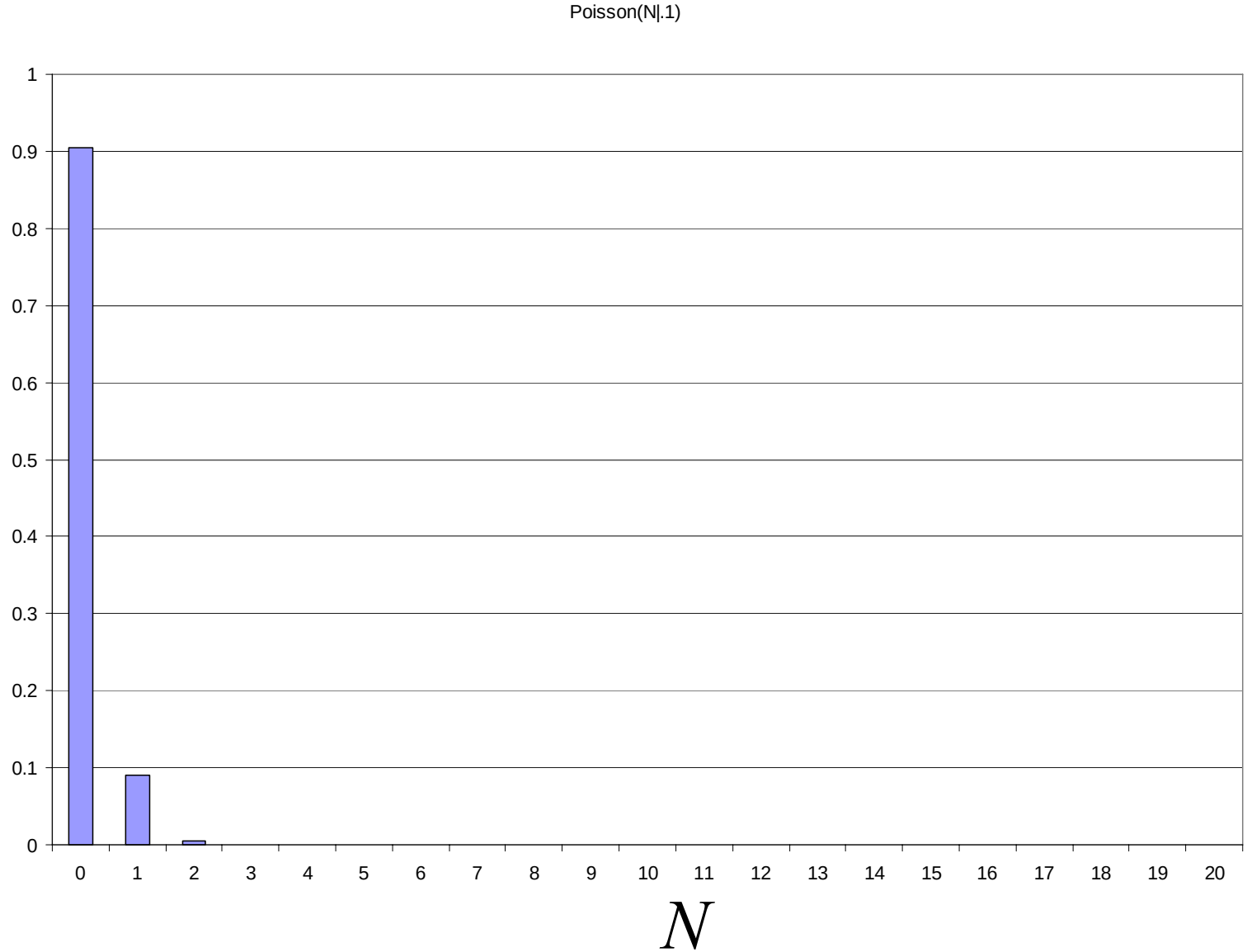
  $p << 1$ (decay probability during counting is low)

  then the Poisson distribution is an excellent approximation to the binomial distribution

$$Poi(N \mid \mu) = \frac{e^{-\mu}\mu^{N}}{N!}$$

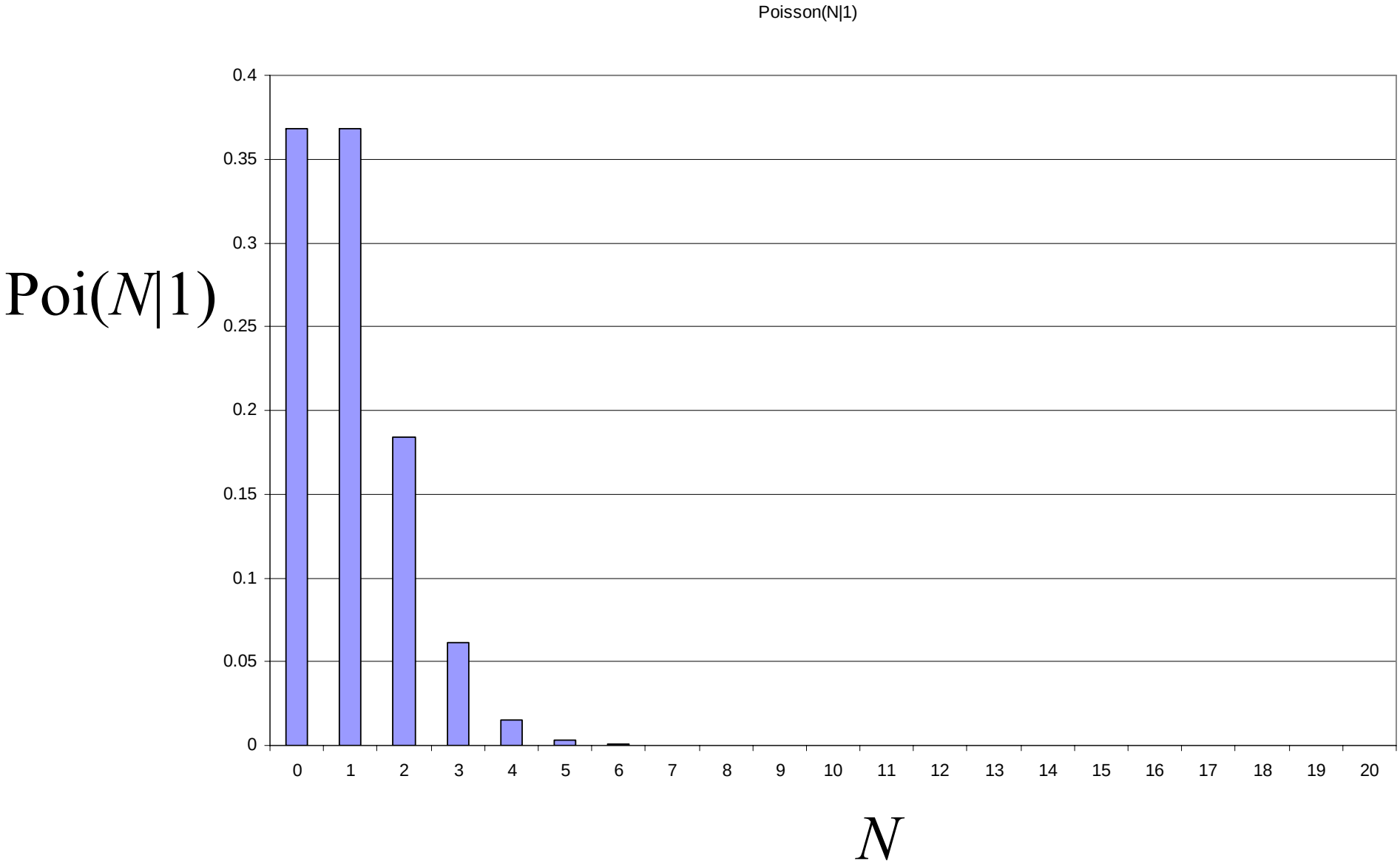
- Expected number of counts N given Poisson mean μ

Poisson Distribution, $\mu = \rho t = 0.1$

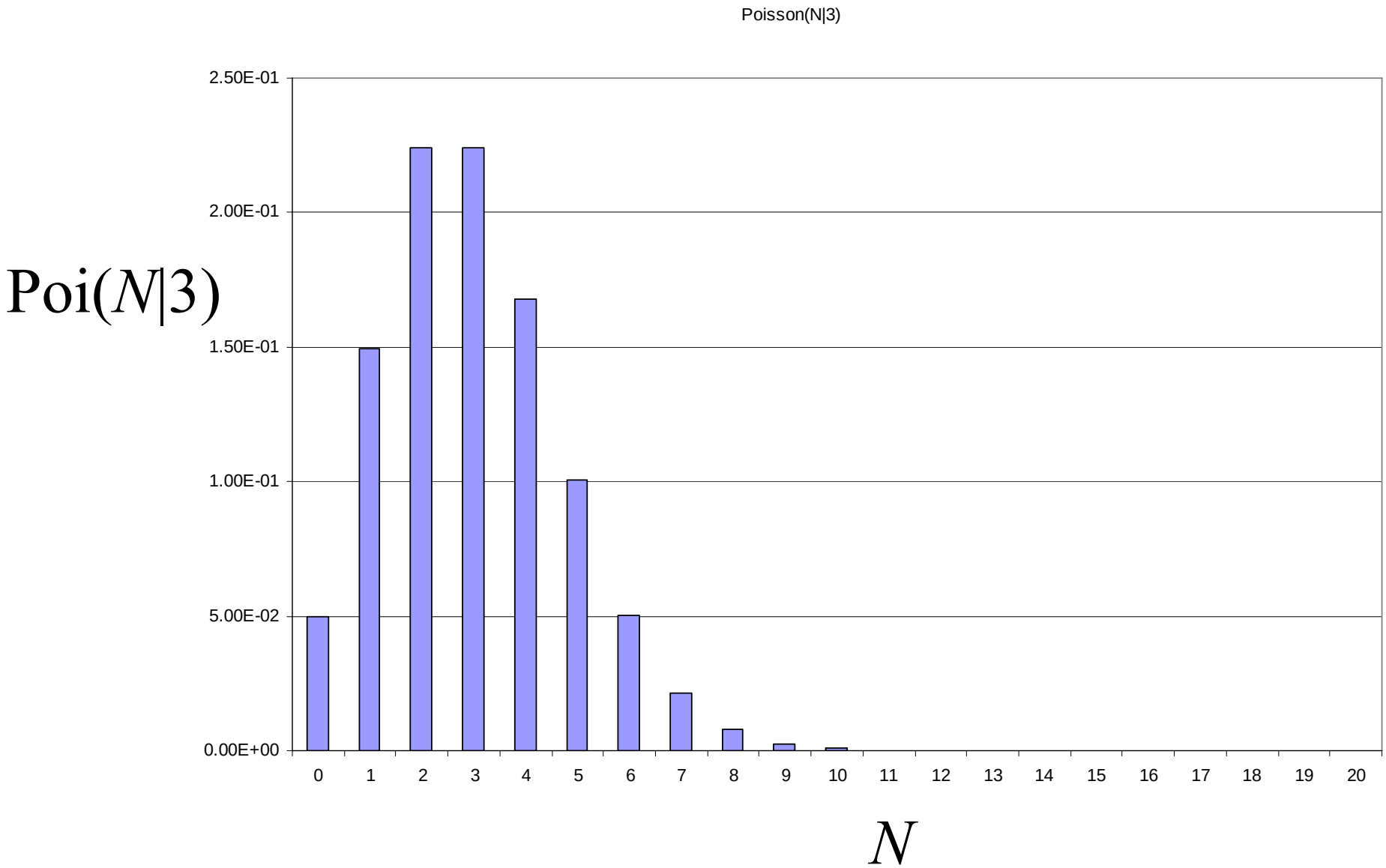


$\mathrm{Poi}(N|0.1)$

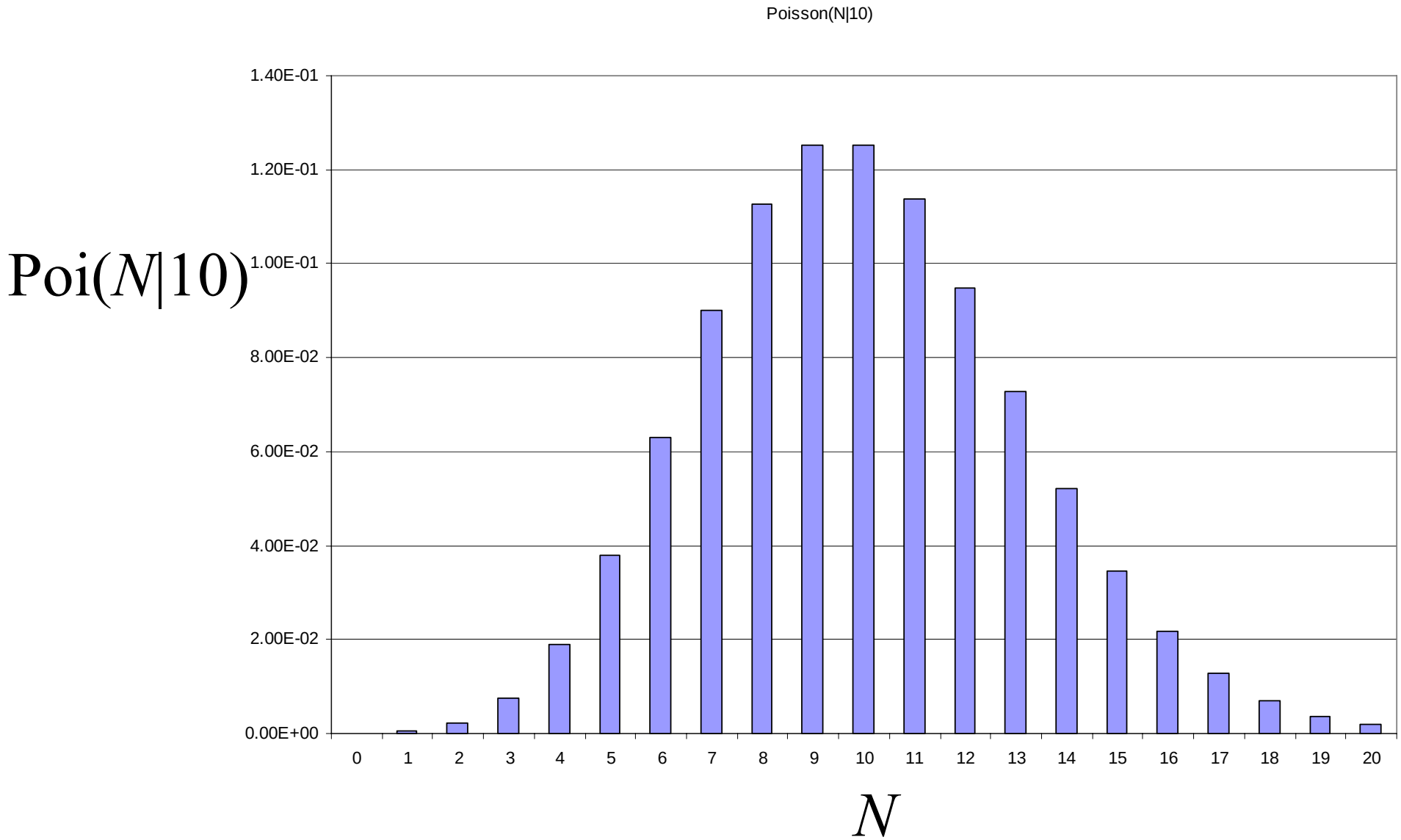
Poisson Distribution, $\mu = \rho t = 1$



Poisson Distribution, $\mu = \rho t = 3$



Poisson Distribution, $\mu = \rho t = 10$



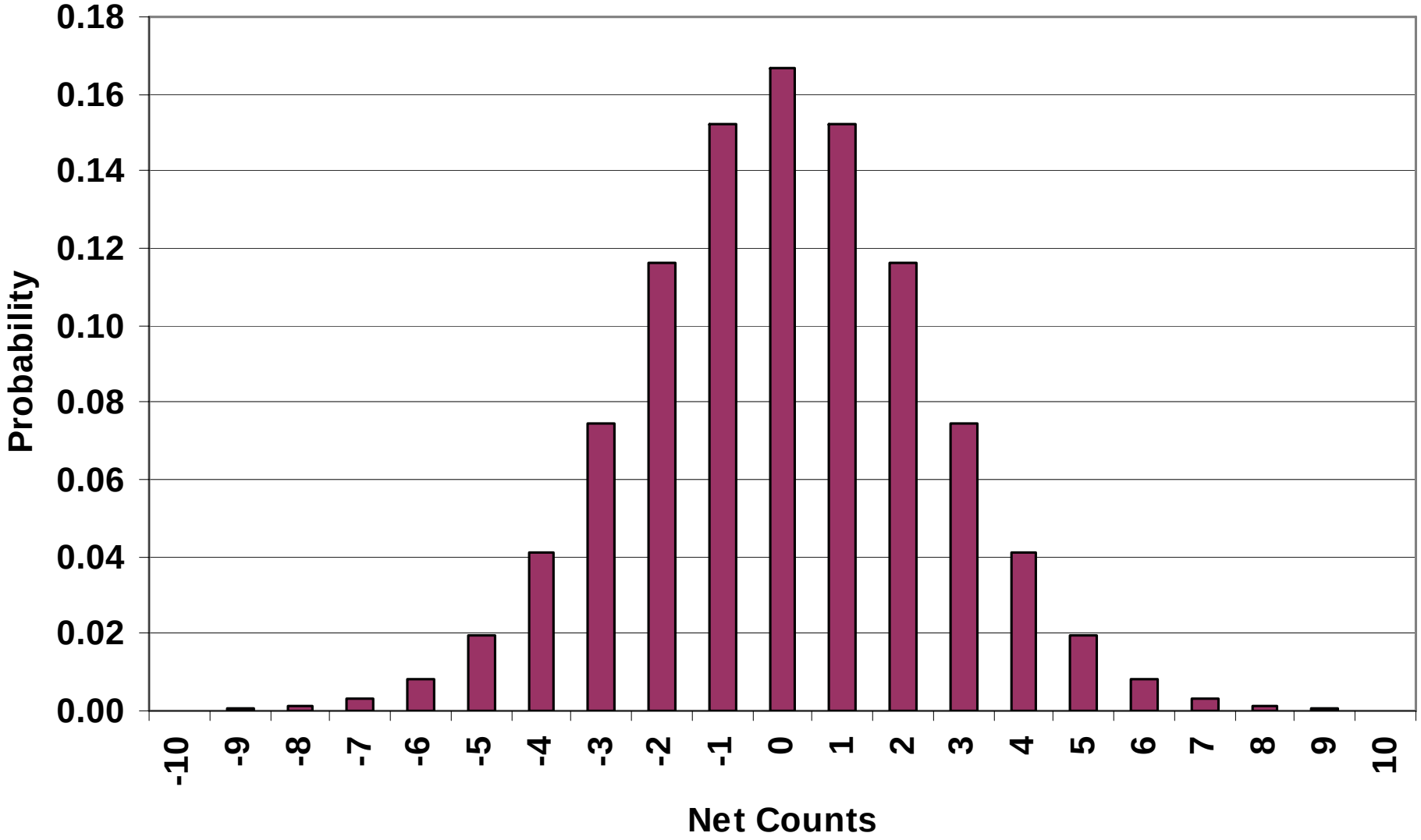
Normal Approximation to the Poisson

- No one tries to approximate a Poisson distribution with a Normal distribution in counting problems
- The normal approximation is applied to the *difference* of two Poisson distributions
  - typically much more symmetric

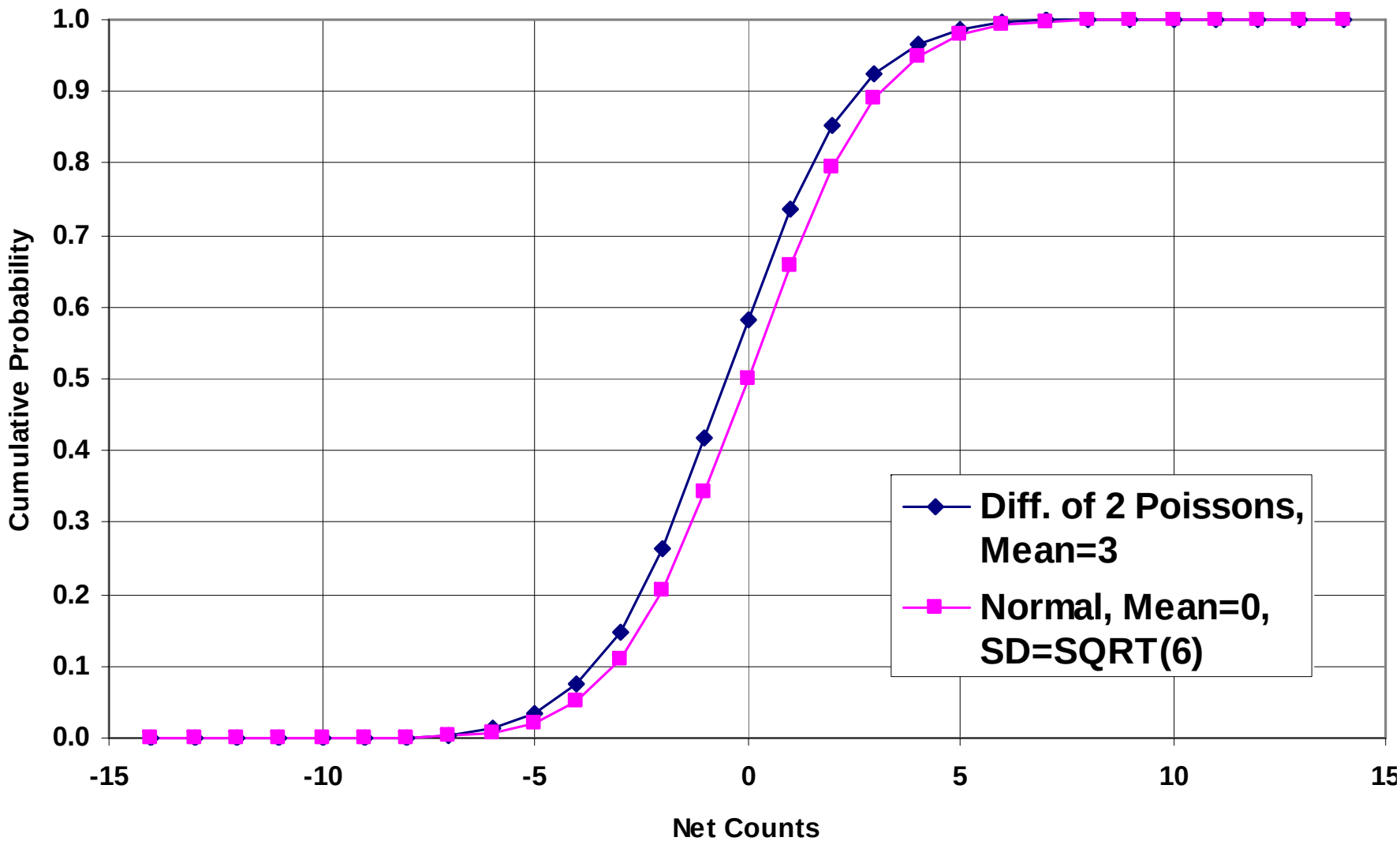
Difference of 2 Poisson Distributions

- When means are equal (e.g., blanks), are symmetric
- Discrete, not continuous
  - For $\mu_b = 3$, $P(N<0) = 0.42$, $P(N \leq 0) = 0.58$
  - For Normal, $P(N<0) = 0.500$, $P(N \leq 0) = 0.500$
- Probabilities on upper tails aren't too much different from Normal

Difference of 2 Poissons with $\mu = \rho t = 3$



Difference of 2 Poissons with $\mu = \rho t = 3$



Difference of 2 Poissons with $\mu = \rho t = 3$

N_net	Diff. of 2 Poissons, Mean=3	Normal, Mean=0, SD= SQRT(6)	N_net	Diff. of 2 Poissons, Mean=3	Normal, Mean=0, SD= SQRT(6)
-14	1.4E-7	5.5E-9	0	0.583	0.500
-13	1.2E-6	5.6E-8	1	0.735	0.658
-12	6.4E-6	4.8E-7	2	0.851	0.793
-11	2.9E-5	3.6E-6	3	0.926	0.890
-10	0.00012	0.00002	4	0.967	0.949
-9	0.00044	0.00012	5	0.987	0.979
-8	0.0015	0.0005	6	0.995	0.993
-7	0.0046	0.0021	7	0.9985	0.9979
-6	0.0129	0.0072	8	0.9996	0.9995
-5	0.033	0.021	9	0.99988	0.99988
-4	0.074	0.051	10	0.99997	0.999978
-3	0.149	0.110	11	0.999992	0.999996
-2	0.265	0.207	12	0.999998	1
-1	0.417	0.342	13	0.999999	1
0	0.583	0.500	14	0.999999	1

The Observables and the Measurands

The Observables

- Same apparatus for blank and sample
- Assume count times known ("time preselection" in ISO parlance)
- Assume no non-Poisson variance
  - sometimes not valid in the real world
  - see, e.g., Kathren 2001, ISO 1995
- *Assume observed count is maximum likelihood estimate and estimate of its variance* ("the Great Leap of Inference")

Notation - 1: Observed Quantities

- Convention: Roman letters denote observed quantities
- N_b: number of blank counts observed
- N_g: number of gross counts observed
- t_b: blank count time (s)
- t_g: gross count time (s)
- R_b: blank count rate (s^{-1})
- R_g: gross count rate (s^{-1})
- R_n: net count rate (s^{-1})
- $s(R_n)$: standard deviation of net count rate (s^{-1})

Classical Statistics: Traditional Relationships Among Observed Quantities

$$R_{\mathrm{b}} = \frac{N_{\mathrm{b}}}{t_{\mathrm{b}}}\,;\, R_{\mathrm{g}} = \frac{N_{\mathrm{g}}}{t_{\mathrm{g}}}$$

$$R_{\mathrm{n}} = R_{\mathrm{g}} - R_{\mathrm{b}} = \frac{N_{\mathrm{g}}}{t_{\mathrm{g}}} - \frac{N_{\mathrm{b}}}{t_{\mathrm{b}}}$$

$$s(R_{\mathrm{n}}) = \sqrt{\frac{\mathrm{Var}(N_{\mathrm{g}})}{t_{\mathrm{g}}^2} + \frac{\mathrm{Var}(N_{\mathrm{b}})}{t_{\mathrm{b}}^2}} \approx \sqrt{\frac{N_{\mathrm{g}}}{t_{\mathrm{g}}^2} + \frac{N_{\mathrm{b}}}{t_{\mathrm{b}}^2}}$$

A Poor Assumption: "N is a good estimate of Var(N)"

- A better assumption may be

$$s(R_{\mathrm{n}}) = \sqrt{\frac{N_{\mathrm{b}}+1}{t_{\mathrm{b}}^2} + \frac{N_{\mathrm{g}}+1}{t_{\mathrm{g}}^2}}$$

Notation 2: The Measurands - [Unknown] Population Parameters

- By convention, Greek letters denote population parameters
- These reflect the *measurand*, the "true state of Nature" that we are trying to infer
- ρ_b: long-term blank count rate (s^{-1})
- ρ_n: long-term net count rate (s^{-1}) (due to analyte in unknown)
- ρ_g: long-term gross count rate (s^{-1})

Notation 3: The Measurands - [Unknown] Population Parameters

- Parameters are needed for sampling from population distributions
- μ_b: number of blank counts expected during t_b
- μ_g: number of gross counts expected t_g
- $\sigma(\rho_n)$: standard deviation of long-term net count rate (s^{-1})

Classical Statistics:
Relationships Among Population Parameters

$$\text{number of counts} = \text{rate} \times \text{time}$$

$$\mu_{\mathrm{b}} = \rho_{\mathrm{b}} t_{\mathrm{b}}$$

$$\mu_{\mathrm{g}} = (\rho_{\mathrm{b}} + \rho_{\mathrm{n}}) t_{\mathrm{g}}$$

$$\sigma(\rho_{\mathrm{n}}) = \sqrt{\frac{\rho_{\mathrm{b}}}{t_{\mathrm{b}}} + \frac{\rho_{\mathrm{b}} + \rho_{\mathrm{n}}}{t_{\mathrm{g}}}}$$

The Reverse Problem

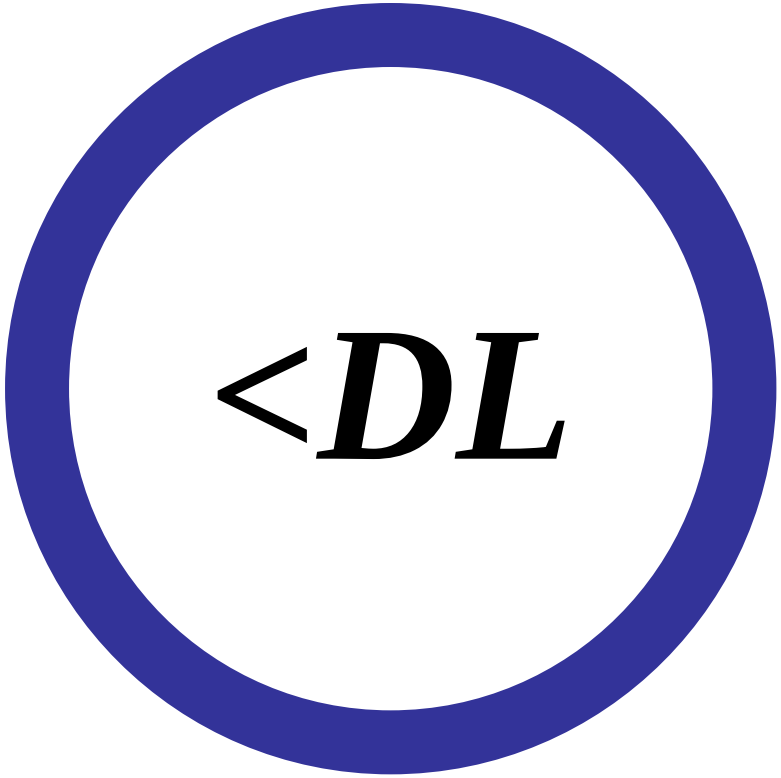
The Reverse Problem: Using Observed Quantities to Estimate Population Parameters (Measurands)

- Classical statisticians
  - use R_n to *estimate* ρ_n
  - use $s(R_n)$ to *estimate* $\sigma(\rho_n)$
    - *ofter a poor assumption for low numbers of counts*
  - every time you make another measurement, you get a new R_n and $s(R_n)$, that is, a new estimate of ρ_n and $\sigma(\rho_n)$
- Bayesian approach shown later

Decision Rules

Terminology Is a Mess! and This Is Just in English!

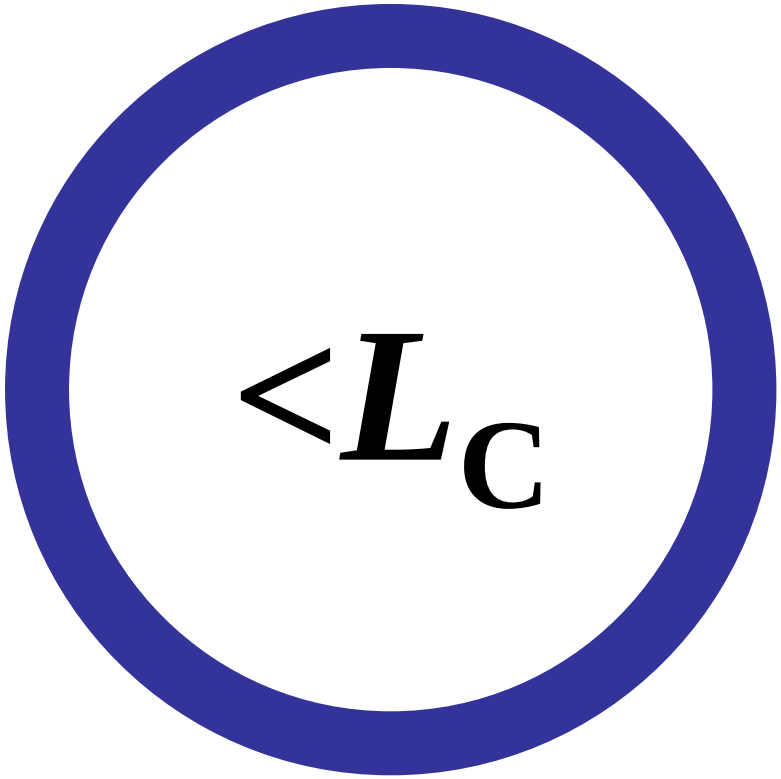
	"DL"	***"MDA"***
Name	decision level	minimum detectable amount
What?	the lowest useable action level	NOT an action level!
Use:	compare measurements to *DL*	Use in planning, advertising or in a statement of work for a contractor: "How much will you charge to provide counting services with this *MDA*?"
When?	*a posteriori*: *after* the measurement is made	*a priori*: *before* the measurement is made (but it does "vary with the nature of the sample" – NUREG-4007)
Defined in	HPS/ANSI N13.30	HPS/ANSI N13.30
Currie's Name	critical level, L_C	detection level, L_D
Ill-defined Names		lower limit of detection, *LLD*; also, un-fortunately, "lower level discriminator," detection limit, limit of detection ("LOD")
Turner's name	"minimum significant measured activity"	"minimum detectable true activity"
ISO 11929 name	"decision threshold"	"detection limit"
Spanish name	umbral de decision	limite de deteccion
MARLAP name	"critical value of []"	"minimum detectable amount" or "minimum detectable concentration"
Strom's name	"false alarm level"	"advertising level" "expected detection capability"



Always compare a result with *DL*

Never compare a result with MDA!

Translation:



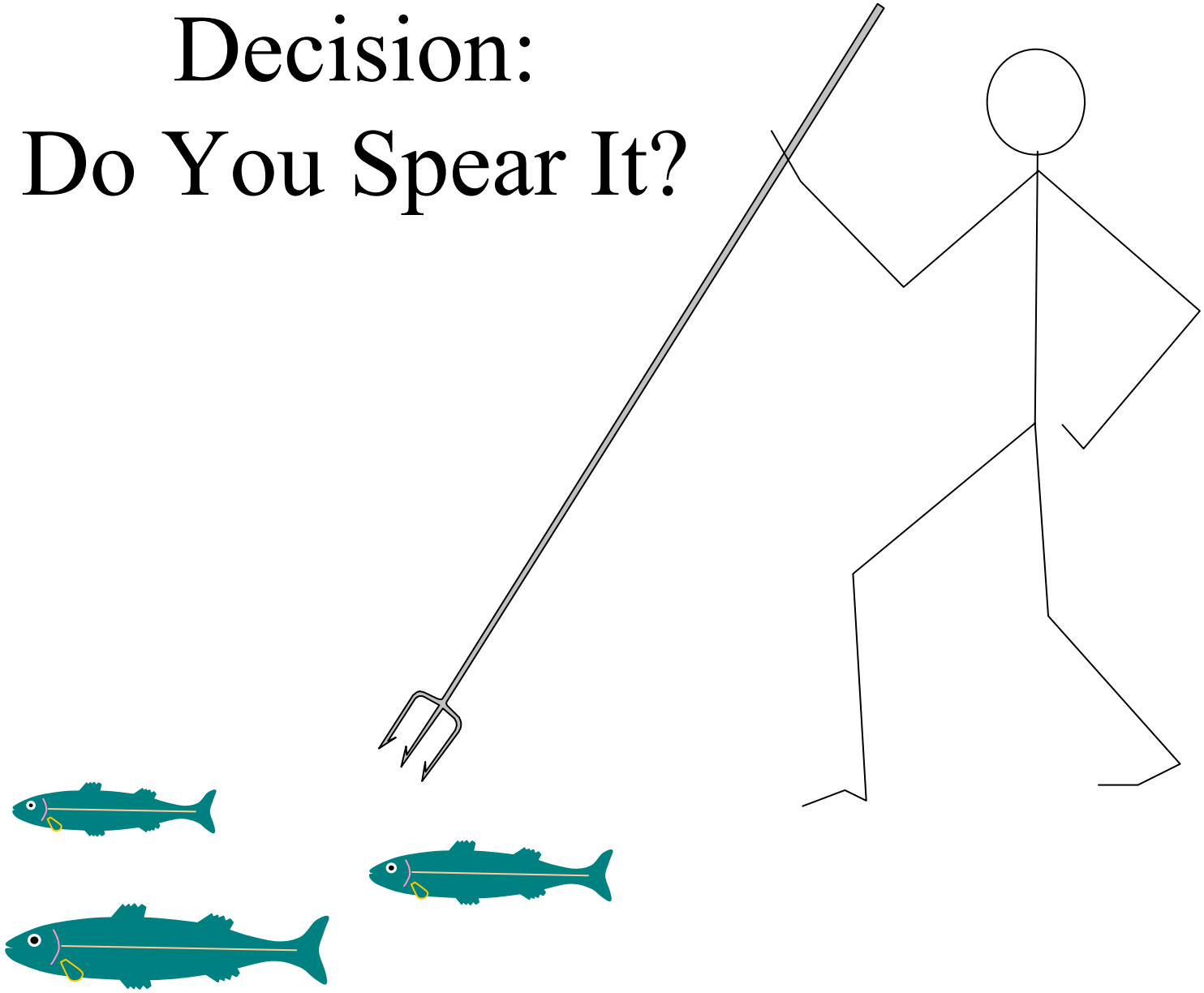
Always compare a result with L_C

Never compare a result with L_D!

L_C is a *threshold.*

Spearing Fish in a Mountain River: Statistical Decision Criteria

Decision:
Do You Spear It?



Spearing Fish in a Mountain River

- You are forbidden to keep fish shorter than 30 cm
- It's a lot of work to spear a fish, and you don't want to injure or kill a lot of fish that you can't keep
- The water is rough and foamy, the current is fast, and the fish are going in all directions, so you can't tell the size of a fish very well before you spear it
- If you spear every fish that *appears* to be at least 30 cm, you needlessly injure or kill lots of fish that are less than 30 cm; and of course you do spear some "keepers" that are longer than 30 cm

Fisher's Decision: Is It a Keeper?

- Question: How long must a fish *appear* to you so that only a few (say, about 1 in 20) will actually be less than 30 cm?
- You determine from experience that if you only spear fish that look like they're, say, 40 cm, then only a few (about 1 in 20) will actually be less than 30 cm
- 40 cm is your ***threshold for the decision to spear***
- However, you miss some fish that are somewhat longer than 40 cm because, due to water conditions, they appear to you like they're less than 40 cm

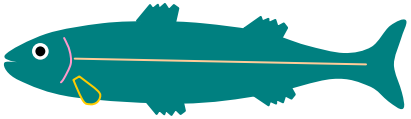
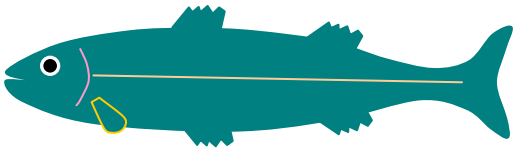
What Size Will You Almost Always Spear?

- Question: How long does a fish *really have to be* so that there's only a small chance (say, about 1 in 20) that it will *appear* to be less than 40 cm?
- You determine from experience that fish that are actually 50 cm have only a small chance (about 1 in 20) of appearing to be 40 cm or less
- 50 cm is the ***shortest fish that you will almost always catch*** when you decide to spear only those fish that *appear* to be at least 40 cm.

Results of Your Decision Rule

- You only spear fish whose true size is 40 cm about half the time, because
  - half the time they look like they're shorter than 40 cm (don't spear), and
  - half the time they look like they're longer than 40 cm
- You only spear fish whose true size is 30 cm about 1 time in 20 because they rarely look like they're 40 cm or more
- You almost always (19 times out of 20) spear fish that are 50 cm or longer because they almost always appear to be over 40 cm

Analogy with Counting Statistics

True Size	Apparent Size	Fishing	Counting	% Speared
> 30 cm		A keeper	above background	5%
	> 40 cm 	Good chance of *being* a keeper	decision threshold, "decision level" *(DL)* or "critical level" (L_C)	50%
> 50 cm 		Good chance of *appearing to be* a keeper	*"MDA"* or "detection level" (L_D)	95%

Measurand versus Measurement Result

- 2 “types” of errors (wrong decisions)

		Is anything there? (Is any activity present [above blank]?)	
		Yes	No
Did I detect anything? (Was the result above the decision level?)	Yes	• I made the correct decision (no error)	• False alarm • False positive • I’ve committed a Type I error
	No	• The alarm should have gone off, but didn’t • False negative • I’ve committed a Type II error	• I made the correct decision (no error)

Error Terminology

- A **Type I error** (wrong decision) is falsely concluding there's activity present when no activity is present
- A **Type II error** is falsely concluding there's no activity present when activity is present
- The **probability of a Type I** error is called α
- The **probability of a Type II** error is called β
- The number of standard deviations above zero on the standard normal distribution having a probability of α or β of being higher is known as the "**standard normal deviate**," k_α or k_β
  - these are $k_{1-\alpha}$ or $k_{1-\beta}$ in ISO notation
- For $\alpha = 0.05$ (a 5% chance of making a Type I error), $k_\alpha = 1.645$
- For $\beta = 0.05$ (a 5% chance of making a Type II error), $k_\beta = 1.645$

Characteristics of Many Decision Rules

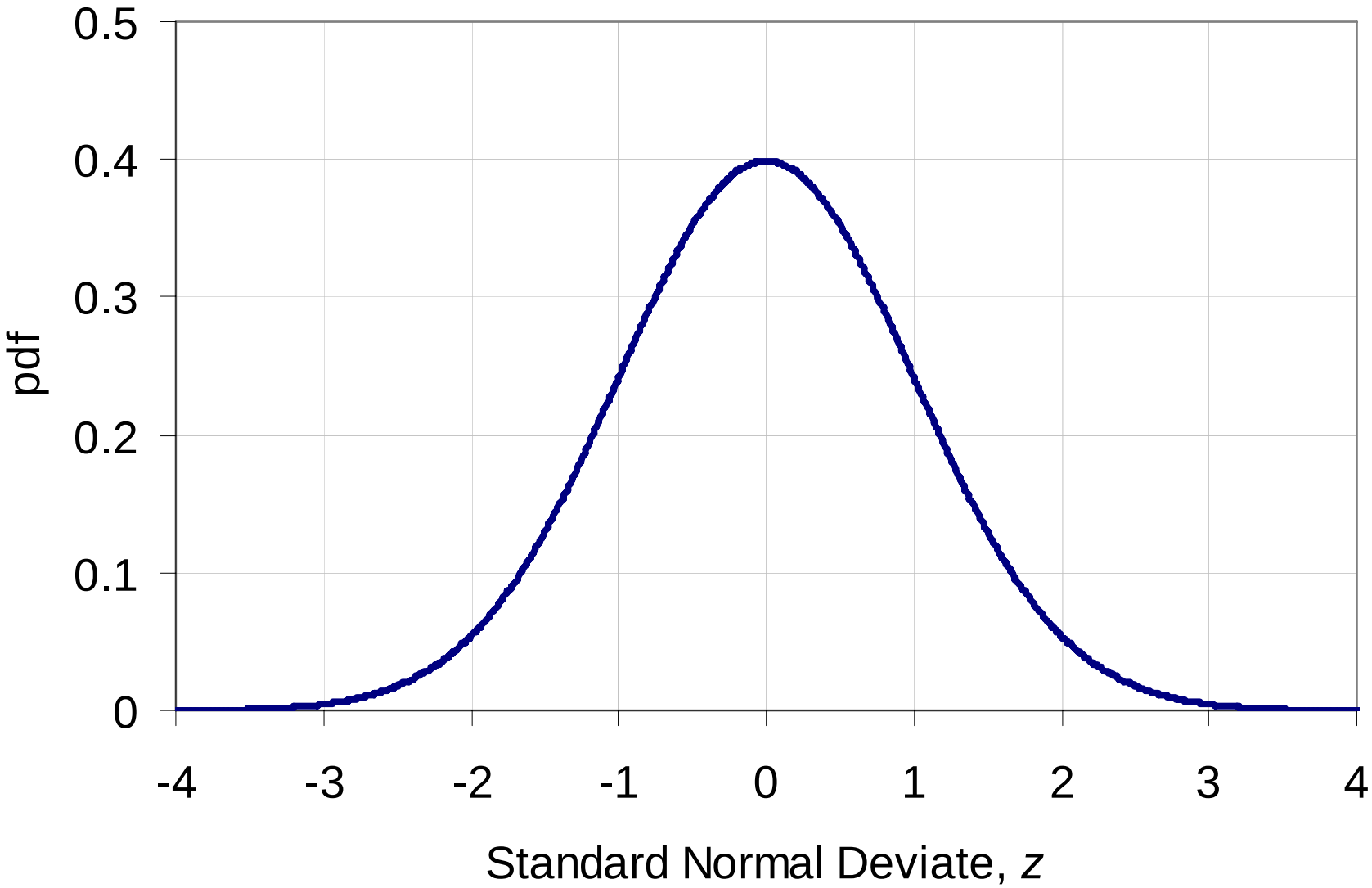
	Assumes Var(μ) = N	Assumes or permits Var(μ) > N	Exact Method or Binomial Distribution
Uses Blank (background) Counts Only	ISO 11929-1 2000 Currie (1968) ANSI N-13.30-1996 Altshuler & Pasternak (1963) Nicholson D_2 (1963)	DL_{N+1} (in Strom & MacLellan 2001)	
Uses Blank and Sample Counts	Nicholson D_1, D_3 (1963)	"Stapleton's decision criterion" (in Strom & MacLellan 2001)	Nicholson D_e (1963) Sumerling and Darby (1981)

Current Decision Level

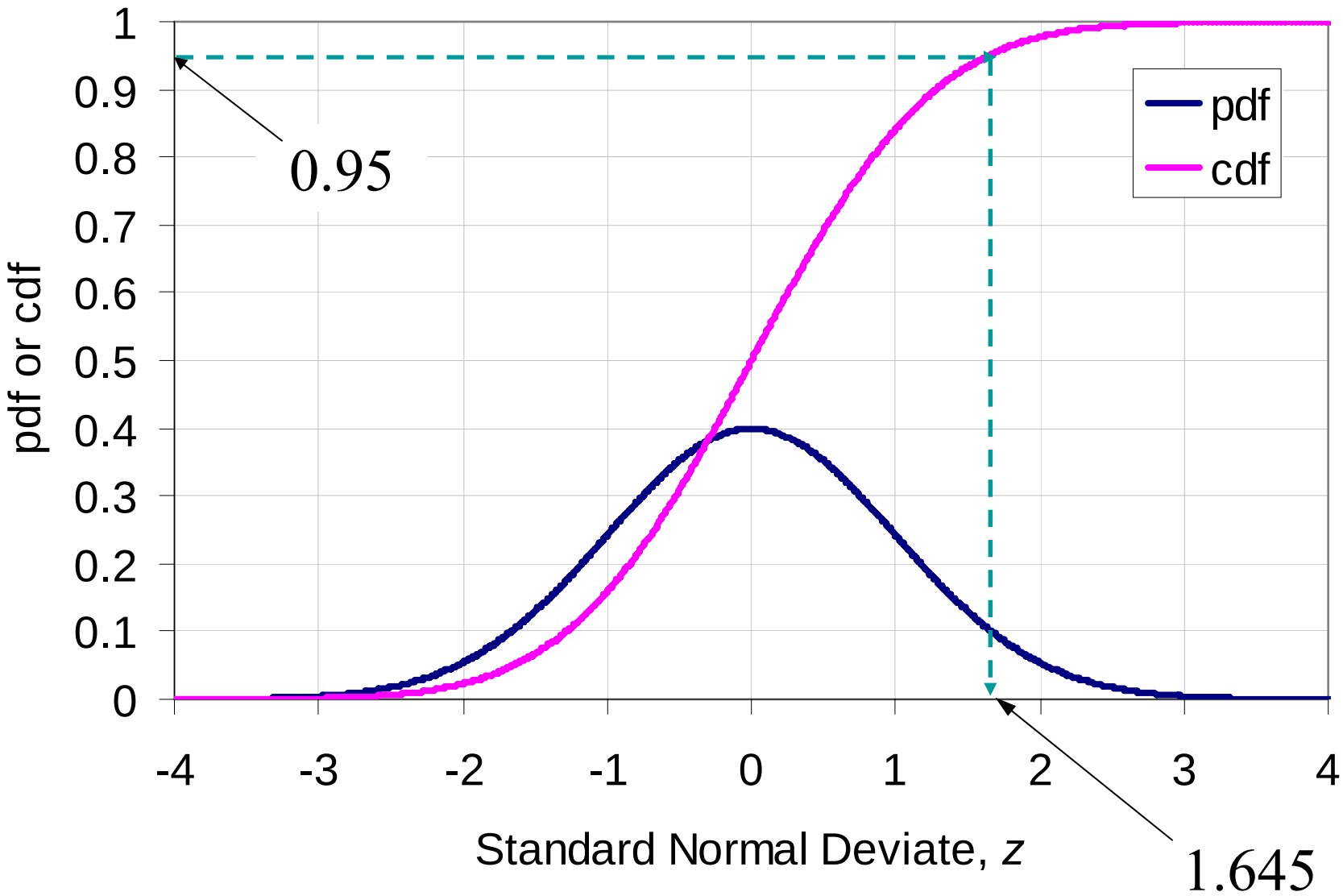
(a.k.a. Critical Level)

- α: acceptable probability of making wrong decision (Type I error): false alarm or false positive
  - α is often taken to be 0.05
- k_α: value of standard normal deviate for area 1-α
  - $k_{0.05}$ is 1.645
- ignore non-Poisson uncertainty for simplicity

Standard Normal Distribution, $\mu = 0$, $\sigma = 1$



Cumulative Standard Normal Distribution



Current "N13.30" Decision Rule

- Nicholson's (1963) D_2 rule; Currie's (1968) rule; ANSI/HPS N13.30-1996; MARSSIM; Equation 15a, Table 1 of ISO 11929-1:2000

$$DL_{\mathrm{N13.30}}(N_{\mathrm{b}},\alpha) = k_{\alpha} s_0 = k_{\alpha}\sqrt{2N_{\mathrm{b}}}$$

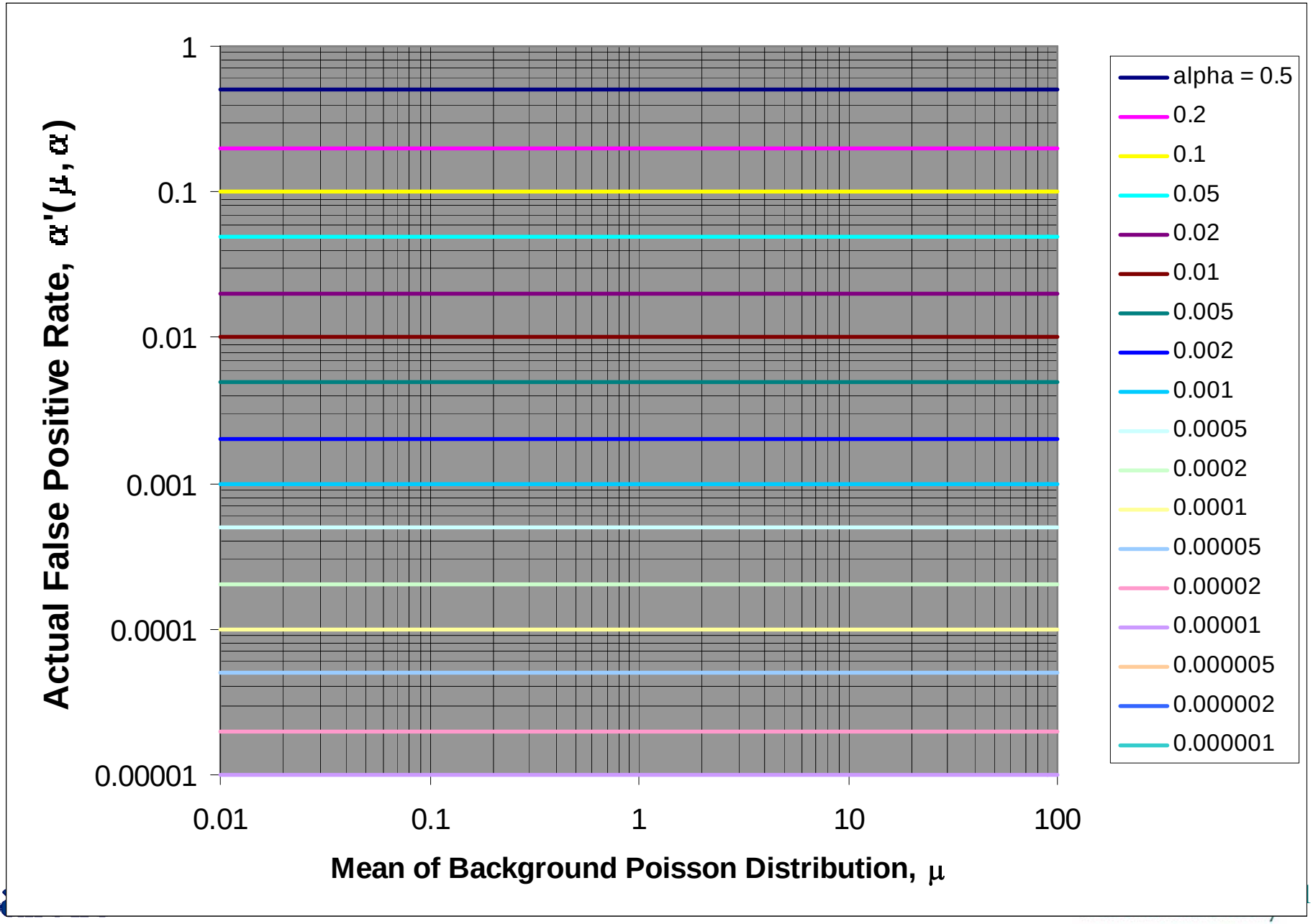
- For $\alpha = 0.05$

$$DL_{\mathrm{N13.30}}(N_{\mathrm{b}},0.05) = 1.645\sqrt{2N_{\mathrm{b}}} = 2.329\sqrt{N_{\mathrm{b}}}$$

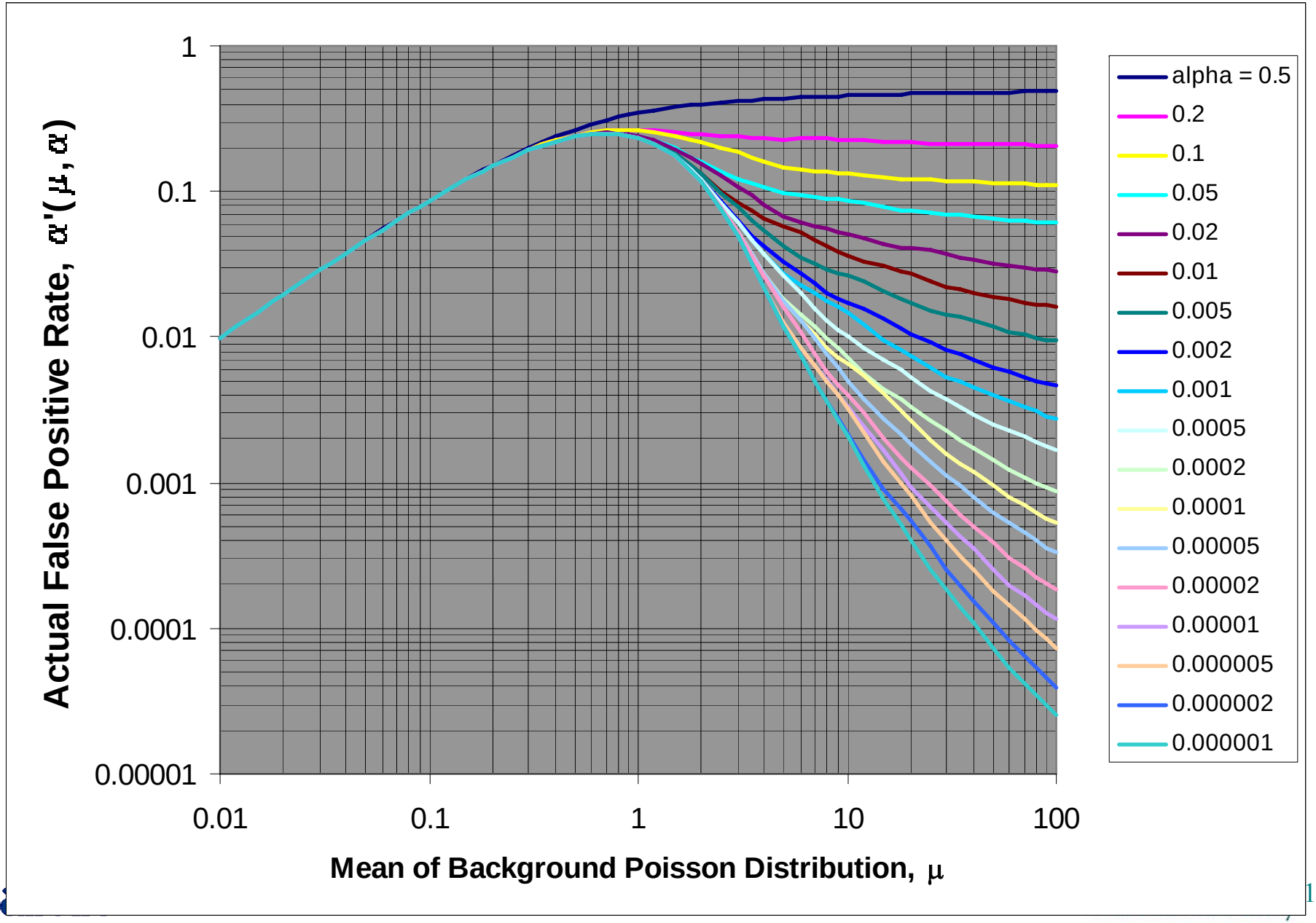
- Expressed as a rate, for non-paired blank:

$$DL_{\mathrm{N13.30}}(R_{\mathrm{n}},\alpha) = k_{\alpha}\sqrt{\frac{N_{\mathrm{b}}}{t_{\mathrm{b}}}\left(\frac{1}{t_{\mathrm{b}}}+\frac{1}{t_{\mathrm{g}}}\right)}$$

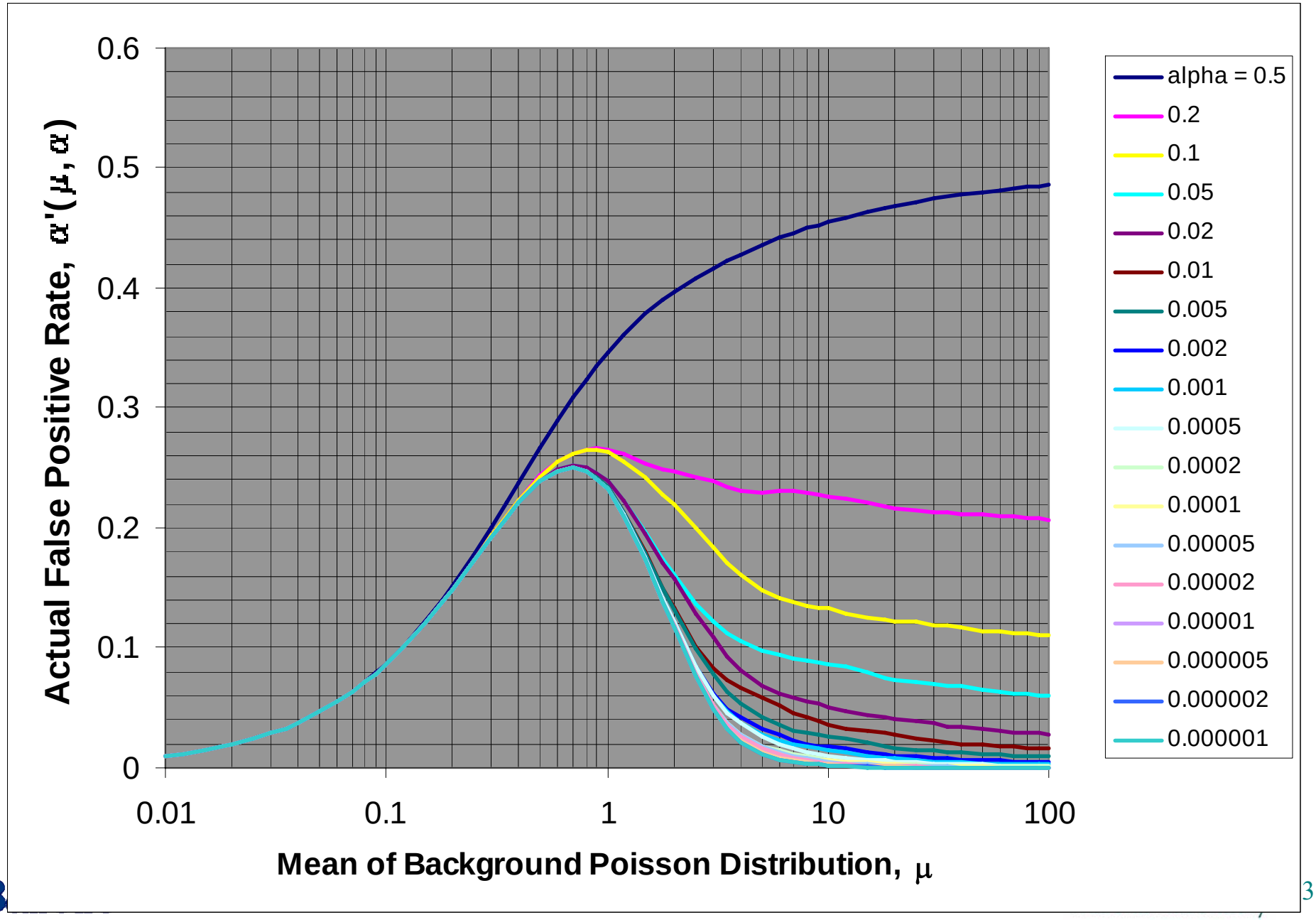
Ideal (Impossible) *DL*



$$DL(\alpha, N_b) = k_\alpha \sqrt{2N_b}$$



$$DL\,(\alpha, N_b) = k_\alpha \sqrt{2N_b}\ \text{(on a linear vertical scale)}$$



Problems with the "N13.30" Decision Rule

- Should be horizontal lines at $\alpha' = \alpha$
- 25% wrong decisions at $\mu_b \approx 0.7$ count, regardless of α
- Actual false positive rate α' is *independent* of α at very small numbers of counts

$$\mu_b = \rho_b t_b << 1$$

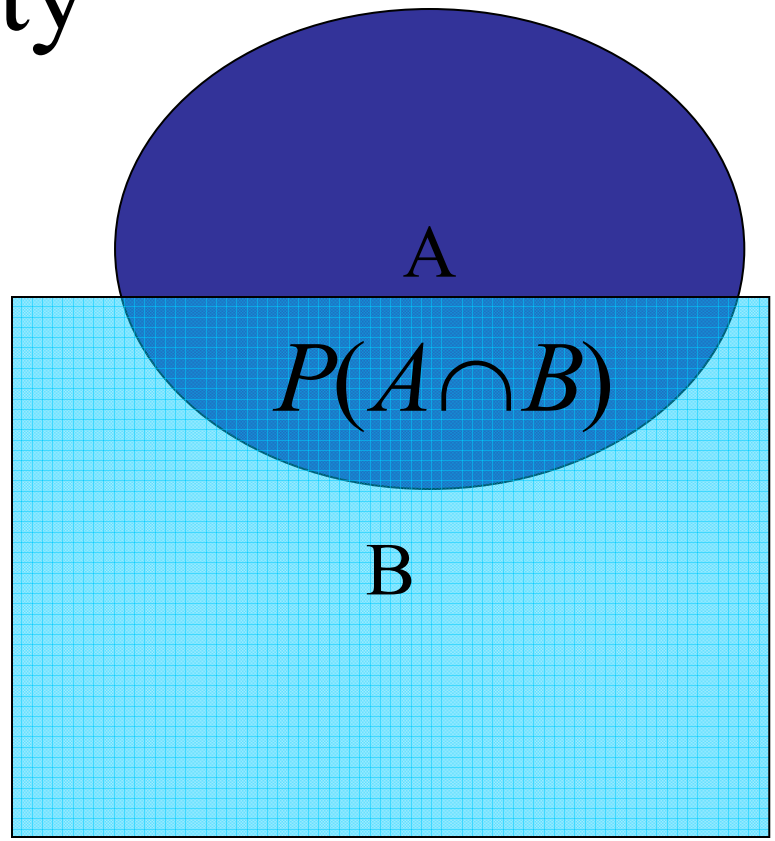
- Even at $\mu_b = 10$, only asymptotically approaches α for larger values
- For very small α, no good even at $\mu_b = 100$!

The Bayesian Approach to the Reverse Problem

The Reverend Thomas Bayes 1702-1761

- *Probability is that degree of confidence dictated by the evidence through Bayes's theorem. -- E.T. Jaynes*

Conditional Probability



Conditional probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

Bayesian Approach: An Identity

- Identity:

$$P(B \mid A)P(A) = P(A \mid B)P(B)$$

- Bayes's Rule (Simple form)

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

- Names:

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalizing Factor}}$$

Bayes's Rule (Simple form)

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

- Names:

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalizing Factor}}$$

Bayesian Approach: Law of Total Probability 1

- Law of total probability:

$$P(A) = P(A \mid B)P(B) + P(A \mid \sim B)P(\sim B)$$

- More useful form of Bayes's rule (denoting likelihood by L):

$$P(B \mid A) = \frac{L(A \mid B)P(B)}{L(A \mid B)P(B) + L(A \mid \sim B)P(\sim B)}$$

Bayesian Approach: Law of Total Probability 2

- Law of total probability:

$$P(A) = \sum_{\text{all } j} P(A \mid B_j) P(B_j)$$

- Even more useful form of Bayes's rule:

$$P(B_i \mid A) = \frac{L(A \mid B_i) P(B_i)}{\sum_{\text{all } j} L(A \mid B_j) P(B_j)}$$

Bayesian Approach: The Prior Probability 1

$$P(B_i \mid A) = \frac{L(A \mid B_i)P(B_i)}{\sum_{\text{all } j} L(A \mid B_j)P(B_j)}$$

- Some form of prior probability is required!
- The prior probability is what you know before you start
- The prior can have more or less effect on the posterior, depending on the precision of the data
- The prior can be subjective
- The prior is the topic of unresolvable arguments

Bayesian Approach: The Prior Probability 2

$$P(B_i \mid A) = \frac{L(A \mid B_i)P(B_i)}{\sum_{\text{all } j} L(A \mid B_j)P(B_j)}$$

- The prior can be "nothing"
  - even "nothing" can take several forms
  - "uniform," "flat," or "uninformative" prior: all values of *B* are "equally probable"
  - "vague" prior: all values of ln(*B*) are equally probable…
- The prior can be hard to nail down
  - "small values of blank are more likely than large ones"

Philosophical Statement of Bayes's Rule

$$P(\text{measurand} \mid \text{evidence}) = \frac{L(\text{evidence} \mid \text{measurand}) P(\text{measurand})}{\text{normalizing factor}}$$

- The measurand or "state of nature" (e.g., count rate from analyte) is what we want to know
- The "evidence" is what we have observed
- The likelihood of the "evidence" given the measurand is what we know about the way nature works
- The probability of the state of nature is what we believed before we obtained the evidence

Probability Density

- probability density is the probability that x lies in an interval between x and $x + \mathrm{d}x$

$$P(x)\,\mathrm{d}\,x = \text{probability that } x$$

$$\text{lies in the interval}$$

$$\text{between } x \text{ and } x + \mathrm{d}x$$

- probability density is a continuous function

Bayes's Rule: Continuous Form

- P's are probability densities

$$P(\mu \mid N) = \frac{L(N \mid \mu) P(\mu)}{\int_0^\infty L(N \mid \nu) P(\nu) \, \mathrm{d}\nu}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalizing Factor}}$$

- We want to determine the posterior probability density

Use of the Posterior Probability Density

$$\text{Expectation value of } \mu = <\mu> = \int_{-\infty}^{\infty} \mu' P(\mu' \,|\, N)\, \mathrm{d}\,\mu'$$

Probability that μ lies between μ_1 and μ_2 =

$$P(\mu_1 \le \mu \le \mu_2) = \int_{\mu_1}^{\mu_2} P(\mu' \,|\, N)\, \mathrm{d}\,\mu'$$

Bayesian Approach for Blank Only

- Assume "uniform," "uninformative," or "flat" prior probability density
- Assume the likelihood probability density is a Poisson

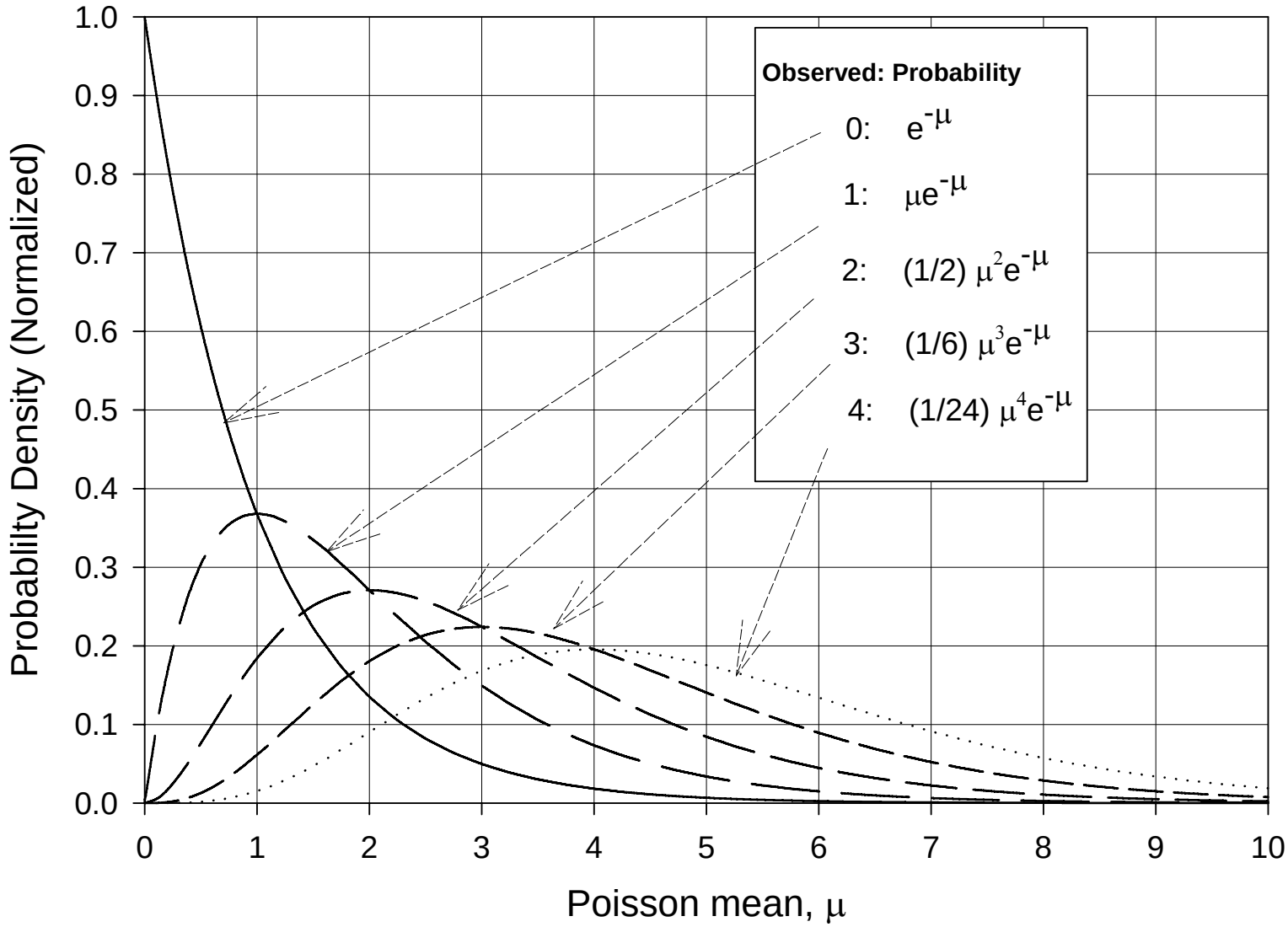
$$L(N \mid \mu) = \frac{e^{-\mu}\mu^{N}}{N!}$$

Bayesian Approach for Blank Only

- With a uniform prior, Bayes's rule inverts the likelihood to yield the posterior
- μ becomes a function of N, instead of N being a function of μ
- Posterior probability density:

$$P(\mu \mid N) = \frac{e^{-\mu}\mu^{N}}{N!}$$

Posterior Probability Densities for μ
(conditional on observed values)



Bayesian Approach

- Assuming uniform "flat" *prior* probability distribution: any value of N is equally likely
- If N counts observed
  - N is maximum likelihood, but $N + 1$ is expectation value:

$$< N_b > = N_b + 1$$

  - variance and standard deviation are simple:

$$\mathrm{Var}(< N_b >) = N_b + 1$$

$$s(< N_b >) = \sqrt{N_b + 1}$$

Ancient References for Bayesian *N*+1 Result Using a Flat Prior

Rainwater, L.J.; Wu, C.S. Applications of Probability Theory to Nuclear Particle Detection. Nucleonics 1(October):60-69; **1947.**

Friedlander, G.; Kennedy, J.W.; Miller, J.M. Nuclear and Radiochemistry. 2nd edition. New York: John Wiley & Sons, Inc.; **1955** & **1963**. The 1963 reference has a section on "Statistical Inference and Bayes' Theorem" (pp. 178-181).

Stevenson, P.C. Processing of Counting Data. NAS-NS-3109. Livermore, California: National Academy of Sciences -- National Research Council; **1966**.

Little, R.J.A. The Statistical Analysis of Low-Level Radioactivity in the Presence of Background Counts. Health Phys. 43(5):693-703; **1982**.

Flat Prior?

- "True" Bayesians are offended by a flat prior
  - "You always know more than nothing"
- Strom's arguments for the flat prior
  - it is the best of both worlds, classical and Bayesian
    - inverting the prior gives a probability distribution that's not available from classical methods
  - it obeys *Bohr's correspondence principle*: 'Any new theory must correspond to the old theory in the regime in which the old theory is known to be valid.'
  - it's what you use on the first experiment
  - it does not require one to postulate that everything is drawn from the same population

Quasi-Bayesian Statistics:
Relationships Among Observed Quantities

- R_n is the same for the paired blank case, slightly different if $t_b \neq t_g$
- $s(R_n)$ is larger because of N+1
- MARLAP 19D.3?

$$R_b = \frac{N_b + 1}{t_b}; \; R_g = \frac{N_g + 1}{t_g}$$

$$R_n = R_g - R_b = \frac{N_g + 1}{t_g} - \frac{N_b + 1}{t_b}$$

$$s(R_n) = \sqrt{\frac{N_g + 1}{t_g^2} + \frac{N_b + 1}{t_b^2}}$$

Why the N13.30 Decision Rule Fails at Very Low Background Rates

- an observed background count causes a decision of "detected" ($\propto \mu_b$)
- unless a gross count is observed as well ($\propto \mu_b^2$)
- *independent* of α !
- so, for $\mu_b < 0.3$, $\alpha' \propto \mu_b - \mu_b^2$
- reason: false assumption that observed values N_b and $N_b^{1/2}$ are good estimates of the mean and standard deviation of background

What Are Alternative Decision Rules?

- "N_b + 1" Decision Rule
- Altshuler & Pasternak (A&P; 1963) / Turner (1995) Eq. 11.68
- Keith McCroan's generalization of A&P
  - (= ISO 11929-1:2000MARLAP)
- James H. Stapleton's rule
- Nicholson (1963) D_1 rule
- Nicholson (1963) D_3 rule
- Nicholson (1963) D_e "exact" / Sumerling & Darby (1981) rule
- Bayesian approach
- Rigaud's (2003) rule (not evaluated here)

"N_b + 1" Decision Rule

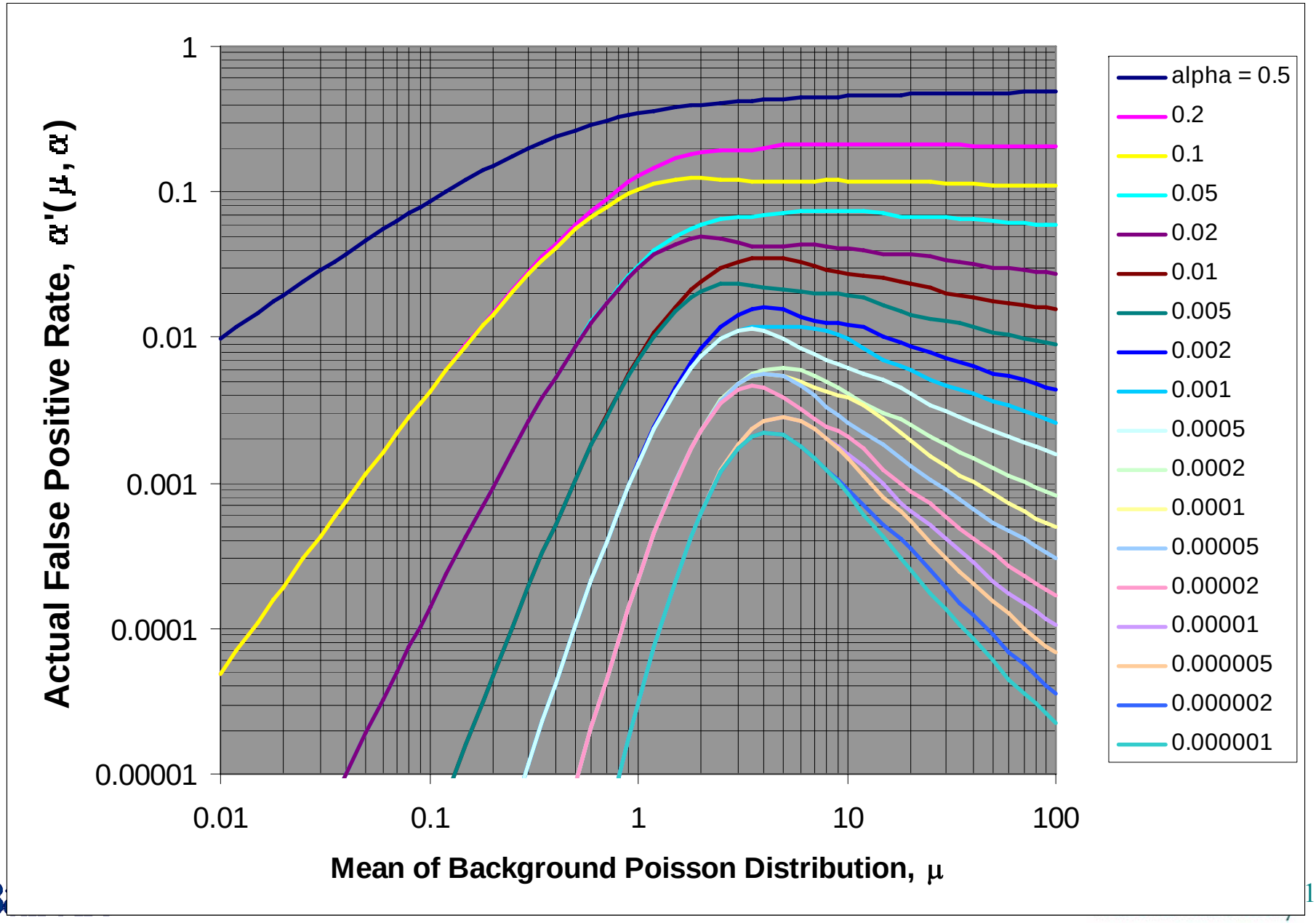
- Bayesian inference of background rate
- Question: If one observes N_b counts, what is the ***expectation value*** of the background distribution that gave rise to this observation (see figure)?
- Bayesian Answer (uniform prior): $\mu_b = N_b + 1$

$$DL_{N+1}(N_b, \alpha) = k_\alpha \sqrt{2(N_b + 1)}$$

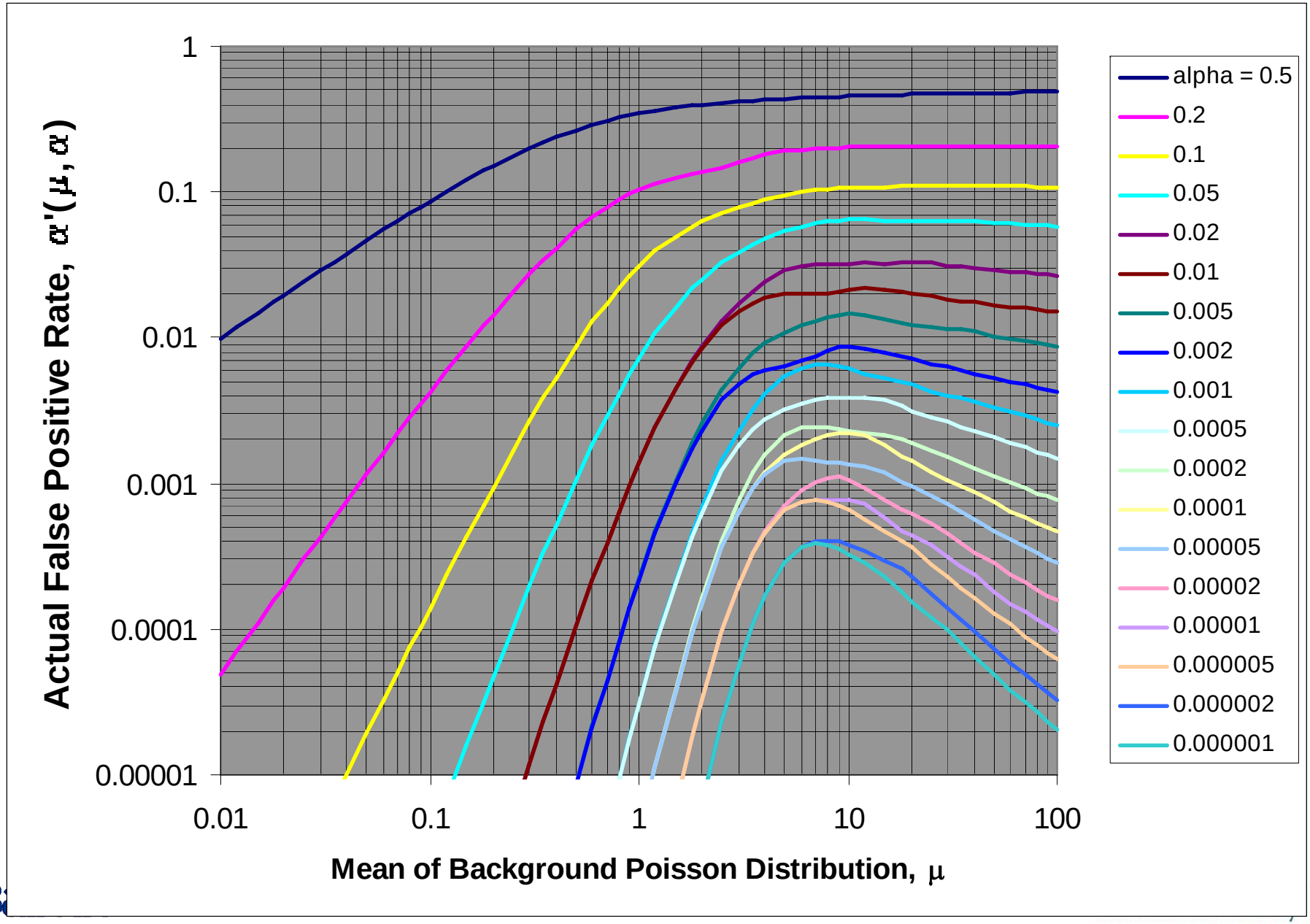
- Idea: Rainwater & Wu 1947; Friedlander & Kennedy 1955; Friedlander et al. 1963; Stevenson 1966; Little 1982

$$DL_{N+1}(R_n, \alpha) = k_\alpha \sqrt{\frac{(N_b + 1)}{t_b}\left(\frac{1}{t_b} + \frac{1}{t_g}\right)}$$

$$DL(\alpha, N_b) = k_\alpha \sqrt{2(N_b + 1)}$$



$$DL\,(\alpha, N_b) = k_\alpha \sqrt{2(N_b + 2)}$$



Altshuler & Pasternak Decision Rule

- 1963; Turner (1995) Eq. 11.68
- the confidence interval of the net activity
- equivalent to Currie's detection level (minimum detectable count), when the decision level is set to zero
- applies to net count rate

$$DL_{\mathrm{A\&P/Turner}}(R_{\mathrm{n}},\alpha)=\frac{k_{\alpha}^{2}}{2t_{\mathrm{g}}}+\frac{k_{\alpha}}{2}\sqrt{\frac{k_{\alpha}^{2}}{t_{\mathrm{g}}^{2}}+4R_{\mathrm{b}}\left(\frac{t_{\mathrm{g}}+t_{\mathrm{b}}}{t_{\mathrm{g}}t_{\mathrm{b}}}\right)}$$

McCroan/MARLAP/ISO Decision Rule

- Generalization of Altshuler & Pasternak

$$DL_{\mathrm{McCroan}}(R_{\mathrm{n}},\alpha)=\frac{k_{\alpha}^{2}}{2t_{\mathrm{b}}}+\frac{k_{\alpha}}{2}\sqrt{\frac{k_{\alpha}^{2}}{t_{\mathrm{b}}^{2}}+4R_{\mathrm{b}}\left(\frac{1}{t_{\mathrm{g}}}+\frac{1}{t_{\mathrm{b}}}\right)}$$

- MARLAP July 2001 *Draft*; same as ISO (*ISO notation*):

$$DL_{\mathrm{ISO}}(R_{\mathrm{n}},\alpha)=R_{\mathrm{n}}^{*}=\frac{k_{1-\alpha}^{2}}{2t_{0}}\left(1+\sqrt{1+\frac{4R_{0}t_{0}}{k_{1-\alpha}^{2}}\left(1+\frac{t_{0}}{t_{\mathrm{S}}}\right)}\right)$$

- Only differs from A&P when count times differ
- notation problem: Strom uses k_{α} where ISO uses $k_{1-\alpha}$

An Obvious Argument?

- using *both* the background and gross sample measurements to estimate the background increases the power of the test

Stapleton's Decision Rule

$$z_{\text{Stapleton}}(N_{\text{g}}, t_{\text{g}}, N_{\text{b}}, t_{\text{b}}, d) = \frac{2\sqrt{\dfrac{N_{\text{g}} + d}{t_{\text{g}}}} - \sqrt{\dfrac{N_{\text{b}} + d}{t_{\text{b}}}}}{\sqrt{\dfrac{1}{t_{\text{g}}} + \dfrac{1}{t_{\text{b}}}}}$$

- d is an arbitrary number, $0 < d < 1$; 0.4 is good
- z is standard normal deviate for this combination of N_{b}, N_{g}, t_{b}, t_{g}, and d
- Compare z to k_{α} to determine whether you've detected activity at your chosen α

Rigaud's Decision Rule

$$L_{\mathrm{C}} = 0.5\left(\frac{1}{t_{\mathrm{g}}} + \frac{1}{t_{\mathrm{b}}}\right) + k_{\alpha}\sqrt{R_{\mathrm{b}}\left(\frac{1}{t_{\mathrm{s}}} + \frac{1}{t_{\mathrm{b}}}\right)}$$

Nicholson (1963) D_1 Decision Rule

$$DL_{\text{Nicholson D}_1}(R_{\text{n}},\alpha) = k_\alpha \sqrt{\frac{N_{\text{b}}}{t_{\text{b}}^2} + \frac{N_{\text{g}}}{t_{\text{g}}^2}}$$

Nicholson (1963) D_3 Decision Rule

$$DL_{\text{Nicholson D}_3}(R_{\text{n}},\alpha) = k_\alpha \sqrt{\frac{N_{\text{b}} + N_{\text{g}}}{t_{\text{b}} t_{\text{g}}}}$$

Nicholson (1963) D_e "Exact" / Sumerling & Darby (1981) Decision Rule

- difference of 2 Poissons is distributed as a binomial
- Number of trials, $N_{\text{total}} = N_{\text{b}} + N_{\text{g}}$
- probability of success = $t_{\text{g}}/(t_{\text{g}}+t_{\text{b}})$
- the null hypothesis that the sample is blank is rejected if a blank sample would have produced a gross count as large or larger than the observed 100α% of the time or less, that is, if

$$\sum_{i=N_{\text{g}}}^{N_{\text{total}}} \binom{N_{\text{total}}}{N_{\text{g}}} Q_0^i (1-Q_0)^{N_{\text{total}}-i} \leq \alpha$$

where

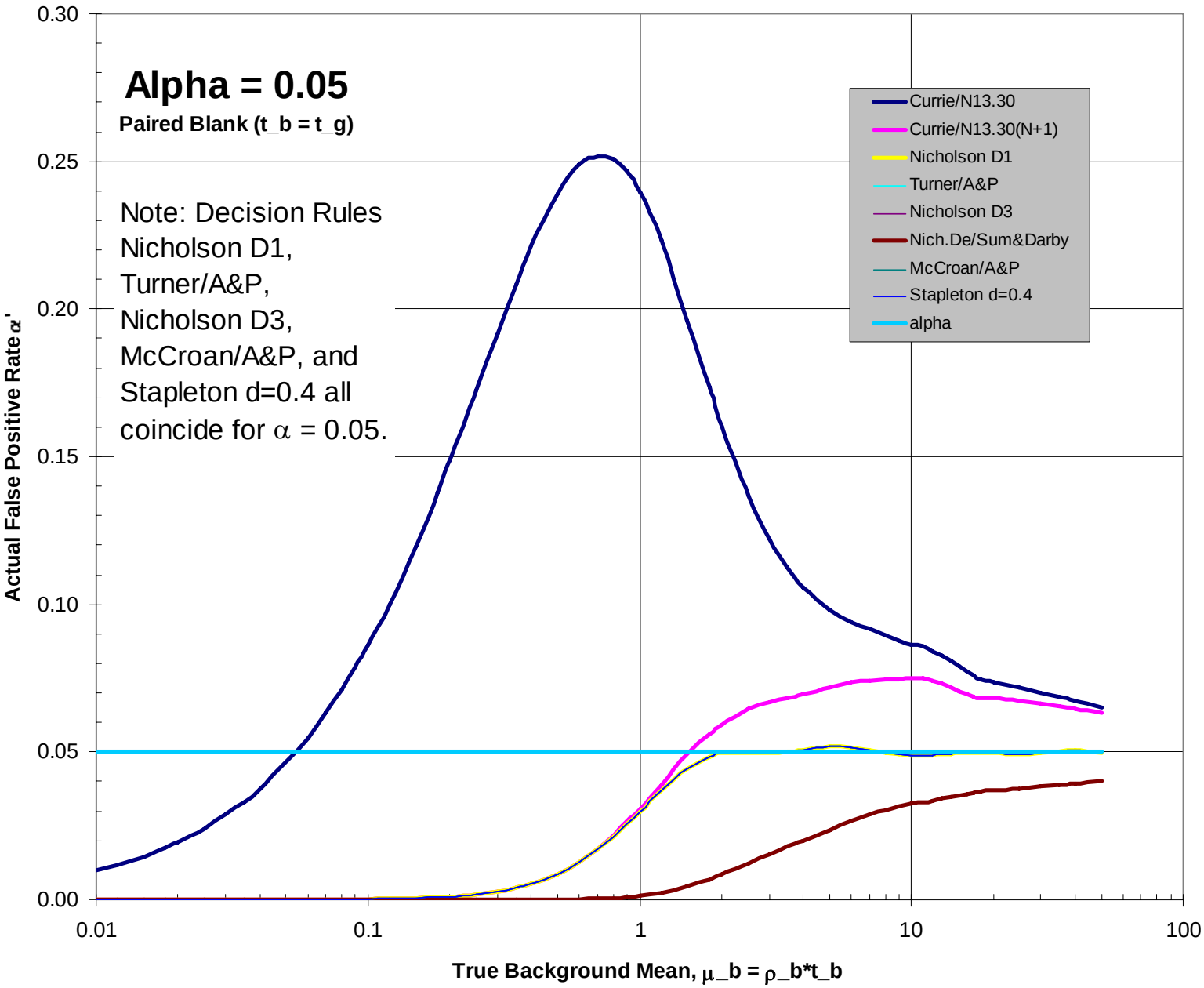
$$Q_o = \frac{\rho_{\mathrm{b}} t_{\mathrm{g}}}{\rho_{\mathrm{b}} t_{\mathrm{g}} + \rho_{\mathrm{b}} t_{\mathrm{b}}} = \frac{t_{\mathrm{g}}}{t_{\mathrm{g}} + t_{\mathrm{b}}}$$

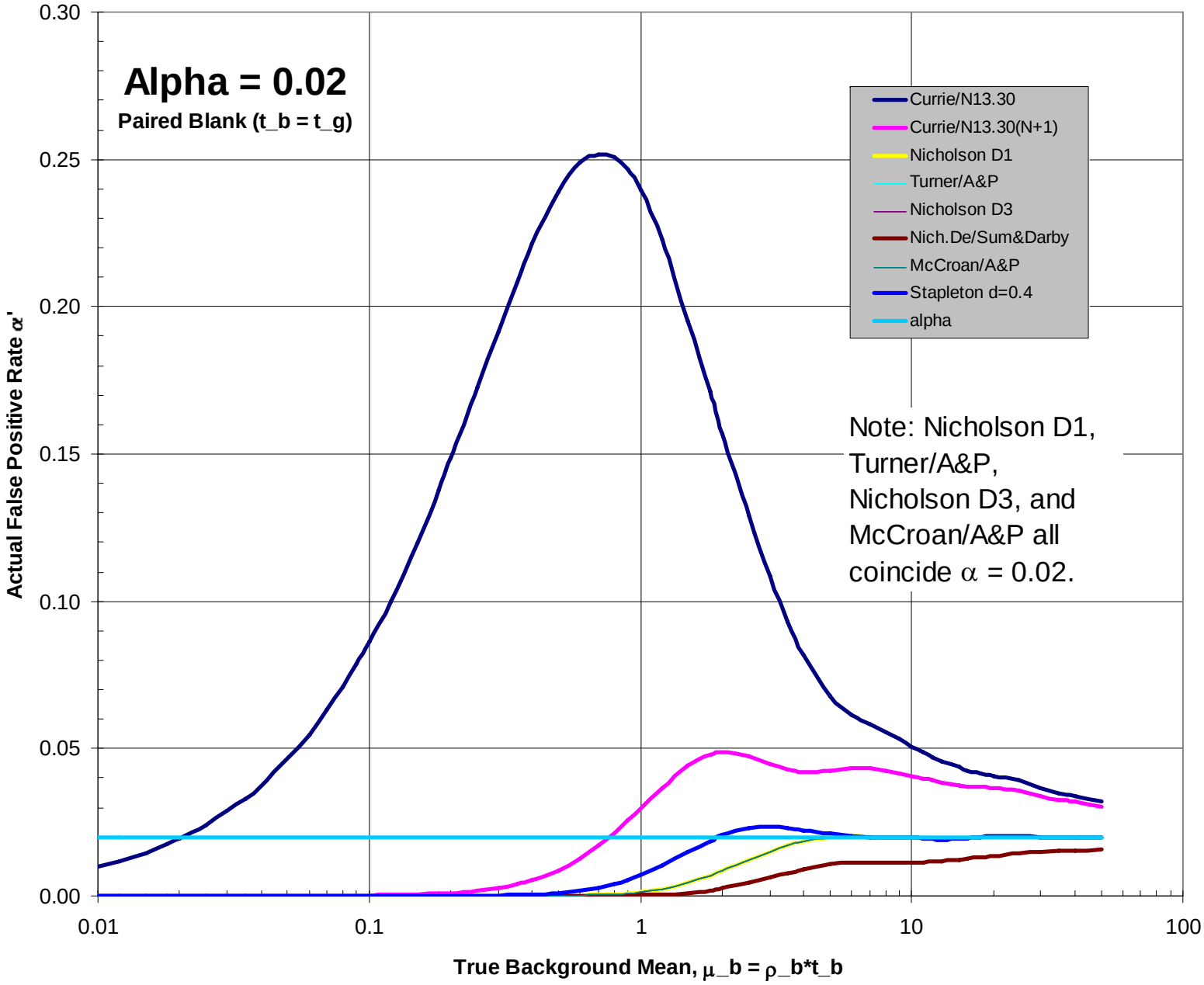
and

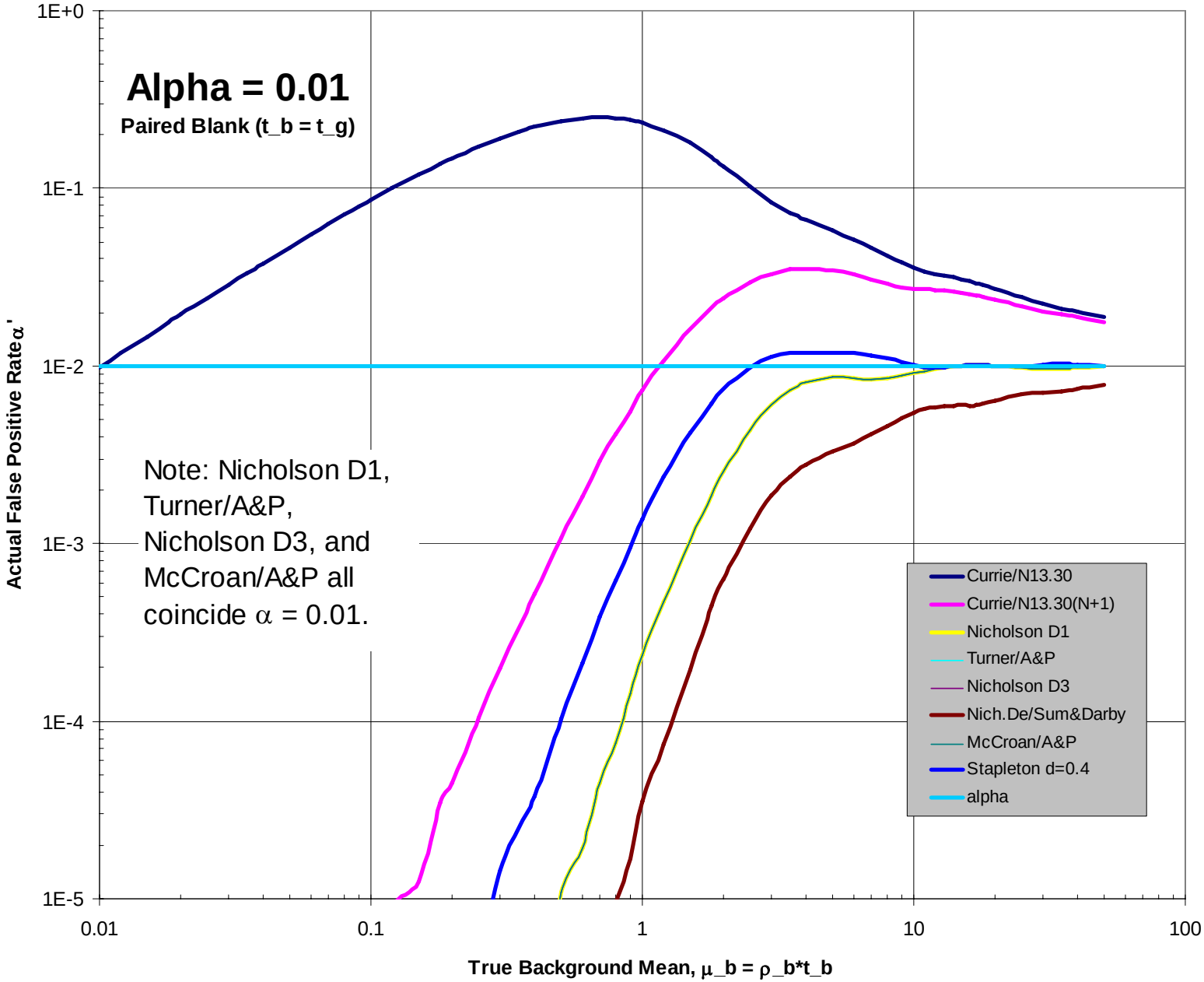
$$\begin{pmatrix} N_{\mathrm{total}} \\ N_{\mathrm{g}} \end{pmatrix} \text{is a binomial coefficient.}$$

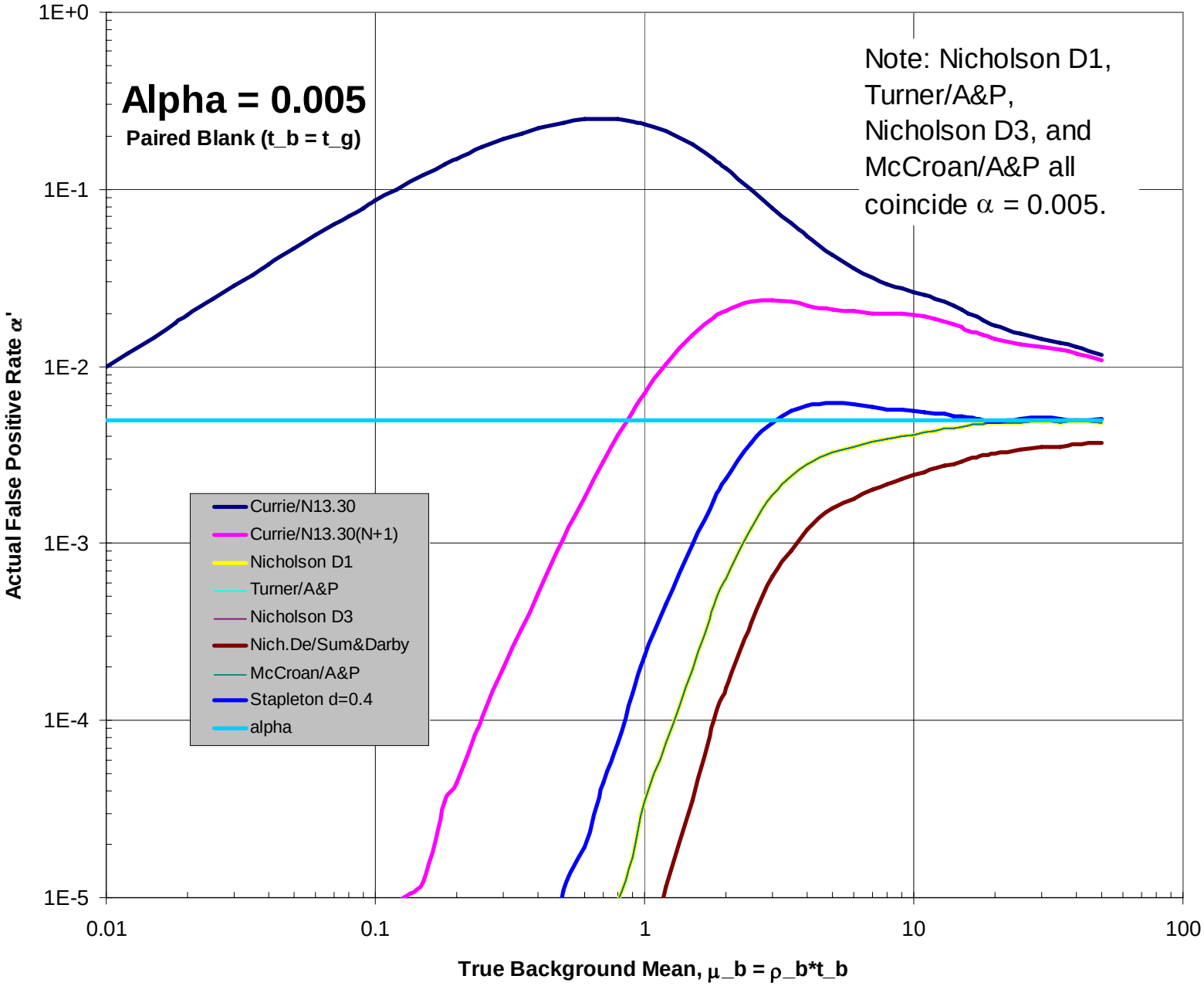
Test of Decision Rules

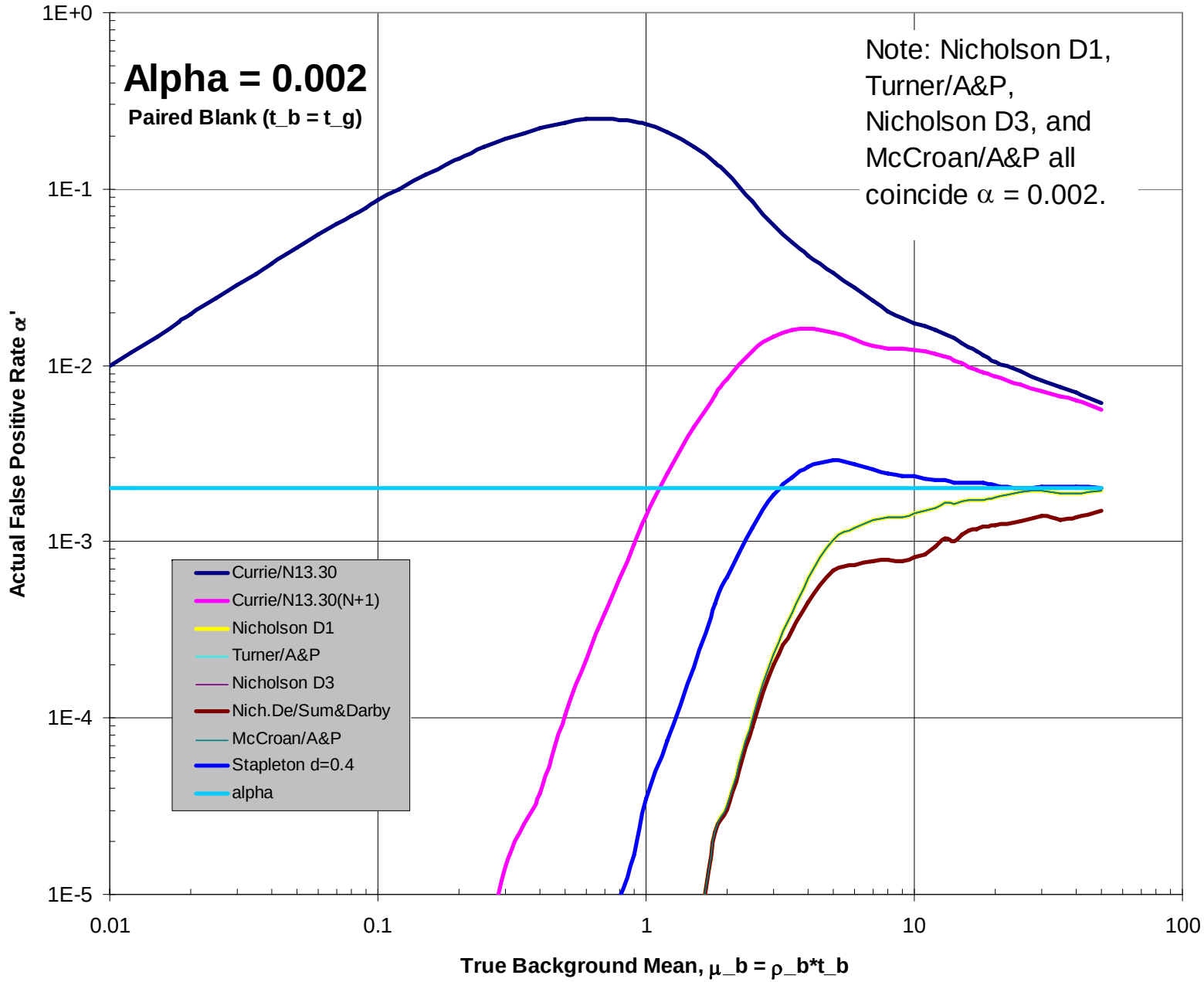
- Monte Carlo simulation (Strom and MacLellan 2001)
- 3,141,593 trials at each of
  - 6 values of α, 0.001 to 0.05
  - 57 values of $\mu_b = \rho_b t_b$ (0.01 to 50)
- MacLellan's exact calculation (MacLellan and Strom 1999) not possible for exact (binomial) or Stapleton's tests or Nicholson's D_1 and D_3 rules, because they use both N_b and N_g.
- Monte Carlo agrees exactly where comparison is possible

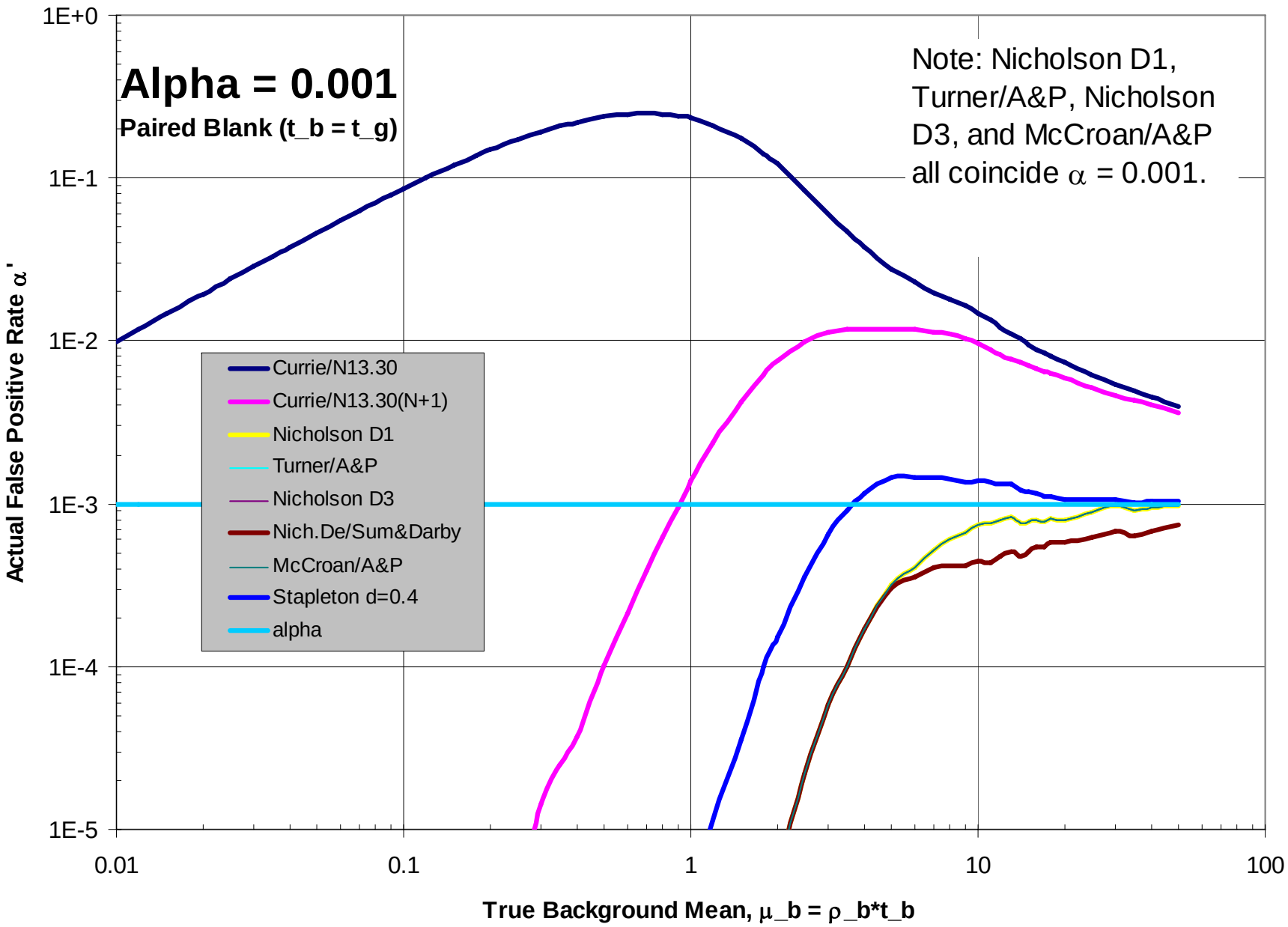

Alpha = 0.05
Paired Blank (t_b = t_g)
Note: Decision Rules Nicholson D1, Turner/A&P, Nicholson D3, McCroan/A&P, and Stapleton d=0.4 all coincide for α = 0.05.
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha
Actual False Positive Rate α'
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.01
0.1
1
10
100
True Background Mean, μ_b = ρ_b*t_b


Alpha = 0.02
Paired Blank (t_b = t_g)
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha
Note: Nicholson D1, Turner/A&P, Nicholson D3, and McCroan/A&P all coincide α = 0.02.
Actual False Positive Rate α'
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.01
0.1
1
10
100
True Background Mean, μ_b = ρ_b*t_b


Alpha = 0.01
Paired Blank (t_b = t_g)
Note: Nicholson D1, Turner/A&P, Nicholson D3, and McCroan/A&P all coincide α = 0.01.
Actual False Positive Rate α'
True Background Mean, μ_b = ρ_b*t_b
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
0.01
0.1
1
10
100
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha


Alpha = 0.005
Paired Blank (t_b = t_g)
Note: Nicholson D1, Turner/A&P, Nicholson D3, and McCroan/A&P all coincide α = 0.005.
Actual False Positive Rate α'
True Background Mean, μ_b = ρ_b*t_b
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
0.01
0.1
1
10
100
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha


Alpha = 0.002
Paired Blank (t_b = t_g)
Note: Nicholson D1, Turner/A&P, Nicholson D3, and McCroan/A&P all coincide α = 0.002.
Actual False Positive Rate α'
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
0.01
0.1
1
10
100
True Background Mean, μ_b = ρ_b*t_b
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha


Alpha = 0.001
Paired Blank (t_b = t_g)
Note: Nicholson D1, Turner/A&P, Nicholson D3, and McCroan/A&P all coincide α = 0.001.
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha
1E+0
1E-1
1E-2
1E-3
1E-4
1E-5
0.01
0.1
1
10
100
Actual False Positive Rate α'
True Background Mean, μ_b = ρ_b*t_b

Results when $t_b = t_g$, $N_b < 10$

- Nicholson D_1, Turner/A&P, Nicholson D_3, and McCroan/ISO all coincide when $t_b = t_g$
- Nicholson D_2/Currie/N13.30/MARSSIM is poorest
- "$N + 1$" rule is much better, but not adequate
- Stapleton's rule is best, followed by the quartet, followed by D_e/S&D
- No rule is good below $N_b = 3$; smaller α is worse
- Need further work for different count times, $t_b \neq t_g$
- ANSI/HPS N13.30 under revision; so is MARLAP

Comparison of Decision Rules for α = 0.05

$R_n t_g$	Currie (1968)	Altshuler & Pasternak Eq15	Sumerling and Darby (1981)	Currie using (N+1)	Nicholson (1963)
0	0 (1)	3	5	3	5
1	3	5	6	4	6
2	4	5	7	5	7
3	5	6	7	5	7
4	5	7	8	6	8
5	6	7	8	6	8
6	6	8	9	7	9
7	7	8	9	7	9
8	7	9	10	7	10
9	7	9	10	8	10
10	8	9	10	8	10

Software Utility under Development

- Freeware Windows 9x/2000/XP 32-bit GUI application
- Shows decisions for all 8 rules for any N_b, N_g, t_b, t_g.
- Not for public release yet – beta available
- Handles unequal background and gross count times
- Shows the amazing diversity of the decision rules
- Shows weakness of current "N13.30" rule

Software Utility under Development

Decision Level Calculator

ISO 11929 Series

ISO 11929	Determination of the detection limit and decision threshold for ionizing radiation measurements --
-1:2000	Part 1: Fundamentals and application to counting measurements without the influence of sample treatment
-2:2000	Part 2: Fundamentals and application to counting measurements with the influence of sample treatment
-3:2000	Part 3: Fundamentals and application to counting measurements by high resolution gamma spectrometry, without the influence of sample treatment
-4:2001	Part 4: Fundamentals and application to measurements by use of linear-scale analogue ratemeters, without the influence of sample treatment
-5:2005	Part 5: Fundamentals and applications to counting measurements on filters during accumulation of radioactive material
-6:2005	Part 6: Fundamentals and applications to measurements by use of transient mode
-7:2005	Part 7: Fundamentals and general applications
-8:2005	Part 8: Fundamentals and application to unfolding of spectrometric measurements without the influence of sample treatment

Software Utility under Development

Counting Statistics for Radioactivity Measurements - Strom@PNL.GOV

DISCLAIMER: Daniel J. Strom, Battelle, the Pacific Northwest National Laboratory, and the U.S. Department of Energy hereby disclaim any and all liability for any use or misuse of this software, and disclaim any and all liability for any errors in this software. USE AT YOUR OWN RISK.

Enter data in green boxes or use default values.

Counting Yield	1	+/-	0.1
Radiochemical Yield	1	+/-	0.1

DL-5_Project Version 2007-07-14
Daniel J. Strom PNNL
strom@pnl.gov (509)375-2626

Future Features: units for time & activity, MDA's, necessary gross counts to have S&D prob < alpha, S&D DL, save your defaults, Stirling N! for S&D, select any alpha, Bayesian with uniform prior.

	Background	Gross	"Net*"
Counts	0	1	1
	N_b-1 N_b+1	N_g-1 N_g+1	
Time	1	1	
	t_b-1 t_b+1	t_g-1 t_g+1	

Calculate

*Note: N_n = N_g - (t_g/t_b)*N_b may not be an integer if t_b not= t_g.

			Net
Rate	0.00E+00	1.00E+00	1.00E+00
Std. Dev.	0.00E+00	1.00E+00	1.00E+00
CV			1.00E+00

Activity	Total Propagated S.D. of Activity	CV of Activity
1.00E+00	1.01E+00	1.01E+00

Acceptable Probability of a False Positive (alpha)

Scroll and click on desired value.

0.5
0.2
0.1
0.05

alpha being used: 0.05

Standard Normal Deviate for (1 - alpha) 1.644853

	Decision Level for Net Count Rate	Detected?
Nicholson D2, ANSI N13.30	0.00E+00	Yes
ANSI N13.30 with (N_b + 1)	2.33E+00	No
A&P, Turner 2nd ed. Eq. 11.68	2.71E+00	No
ISO 11929-1; MARLAP	2.71E+00	No
James H. Stapleton's test	3.00E+00	No
Nicholson D1	1.64E+00	No
Nicholson D3	1.64E+00	No
Nicholson De, Sumerling & Darby	not defined	No

Stapleton's p-value	Stapleton's d, 0 <= d <= 0.5	
0.218021	0.4	Z = 0.778893

Below is the probability that the gross counts could occur from the background distribution using Nicholson's De or Sumerling & Darby

0.5

View Formulas

View McCroan's E-mails

View References

Monte Carlo Proofs

- Crystal Ball is an add-in to Microsoft Excel
  - www.decisioneering.com
- Poisson simulation

Bayesian Approach for Blank and Gross Counts: Joint Likelihood 1

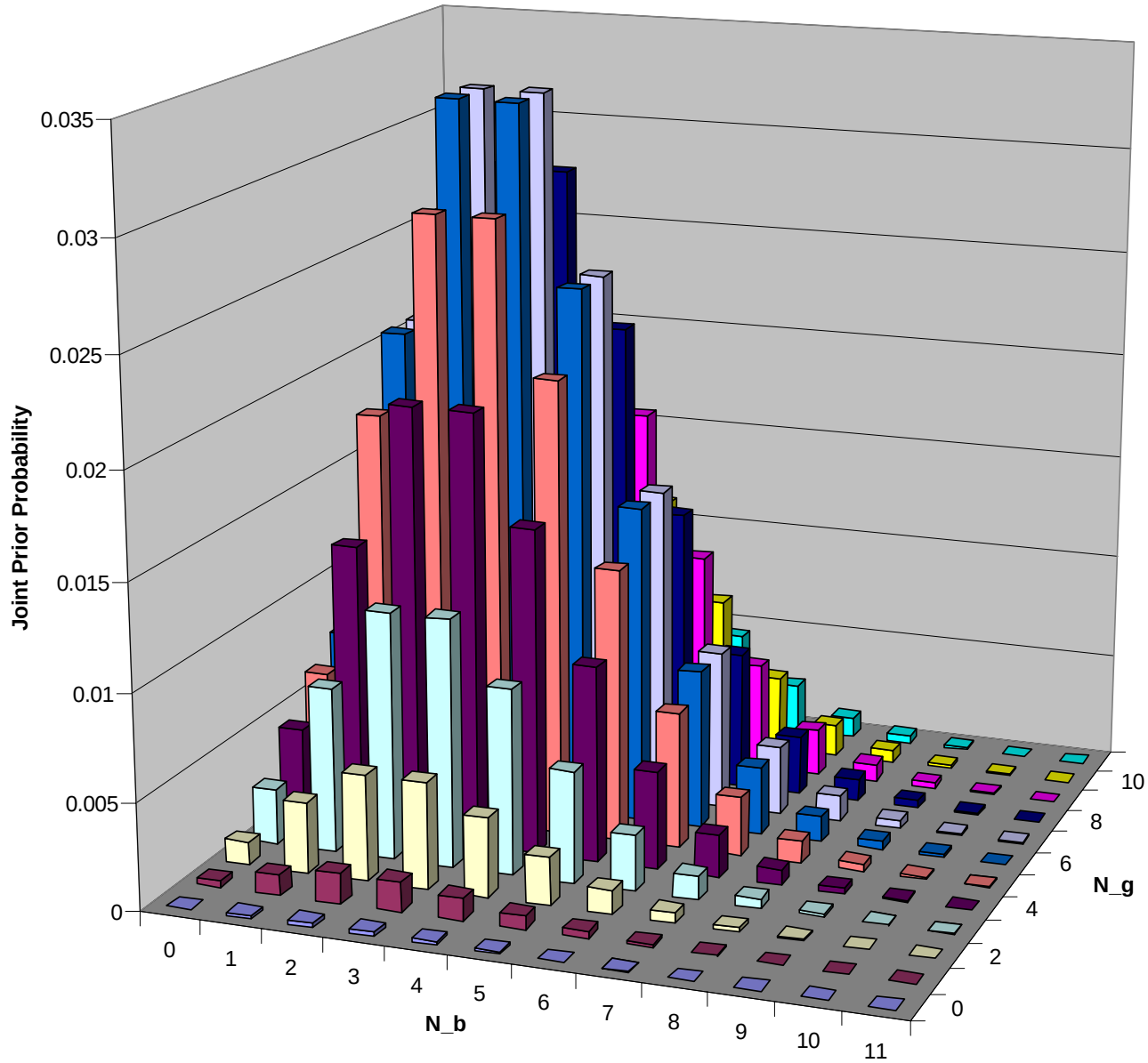
- Assume the likelihood probability densities are Poisson
- Likelihood for N_b and N_g is a function of μ_b and μ_g:

$$L(N_g, N_b \mid \mu_g, \mu_b) = \frac{e^{-\mu_g} \mu^{N_g}}{N_g!} \frac{e^{-\mu_b} \mu^{N_b}}{N_b!}$$

- Not useful because we want N_n, not N_g

Example of Joint Likelihood:

$\mu_b = 3$,
$\mu_n = 3$, so
$\mu_g = 6$



Bayesian Approach for Blank and Gross Counts: Joint Likelihood 2

- Likelihood for N_n is a function of N_g, μ_b, and μ_n:

$$L(N_n \mid \mu_n, \mu_b, N_g) = e^{-(\mu_n + \mu_b)} \sum_{x=0}^{\infty} \frac{\mu_n^{(N_g - x)} \mu_b^x}{(N_g - x)!\,x!}$$

- Don Berry, 11/14/1997
- See Rolf Michel's wipe test example:

  http://www.kernchemie.uni-mainz.de/downloads/saagas21/michel_2.pdf

Bayesian Approach: Limitations

- No matter what prior you pick, people will attack you
  - "Principle of Flat Priors: Use flat priors as an approximation when the prior precision is small in comparison with the sample precision" (Berry 1996, p. 354)
  - "Using flat priors is conservative when the actual prior evidence is consistent with the sample data" (Berry 1996, p. 354)
- Not simple when doing joint probability
- Bayesian results cannot be averaged, added, etc.
  - an average or a sum must be recomputed from first principles

Back to the Real World…
Decision Strategies

Coefficients of Variation Using N+1

- a better "guess"
- t denotes tracer

$$CV^2(R_n) = \frac{\dfrac{N_g+1}{t_g^2} + \dfrac{N_b+1}{t_b^2}}{\left(\dfrac{N_g+1}{t_g} - \dfrac{N_b+1}{t_b}\right)^2}$$

$$CV^2(R_t) = \frac{\dfrac{N_t+1}{t_g^2} + \dfrac{N_{bt}+1}{t_b^2}}{\left(\dfrac{N_t+1}{t_g} - \dfrac{N_{bt}+1}{t_b}\right)^2}$$

Additional *CV*s

- $CV(\text{tracer calibration}) = s_{\text{tracer}}/A_{\text{tracer}}$
- $CV(\text{tracer volume}) = s_{\text{tv}}/V_{\text{tracer}}$
- $CV(\text{aliquot volume}) = s_{\text{aliqout vol.}}/V_{\text{aliquot}}$
- $CV(\text{Type B system performance}) = 0.03$

Uncertainty Propagation Formula

- Simple form of uncertainty propagation ignoring covariances:

$$CV_{\text{total}} = \sqrt{\sum_{i=1}^{N} (CV_i)^2}\text{, where } CV_i = \frac{u(x_i)}{x_i}$$

- Combined standard uncertainty:

$$u_{\text{C}} = A \times CV_{\text{total}}$$

- Hanford uses $2 \times u_{\text{C}}$ as a starting point for decisions

Practical Decision Making: A Multi-step Process Using Classical Statistics

- If a bioassay result is **unexpectedly** > $2 \times u_C$, several possibilities
  - if result not excessively elevated, **recount** sample for 4× as long (i.e., 10,000 min for Pu α-spec at Hanford)
  - if recount result > $2 \times u_C$, obtain new sample (if possible and meaningful) and count new sample
  - if new sample result > $2 \times u_C$, decide "detected"
- Result: negligible false positive rate
- Result: minimization of false negative rate
- Moral of the story: you don't have to decide on only one measurement of one sample!

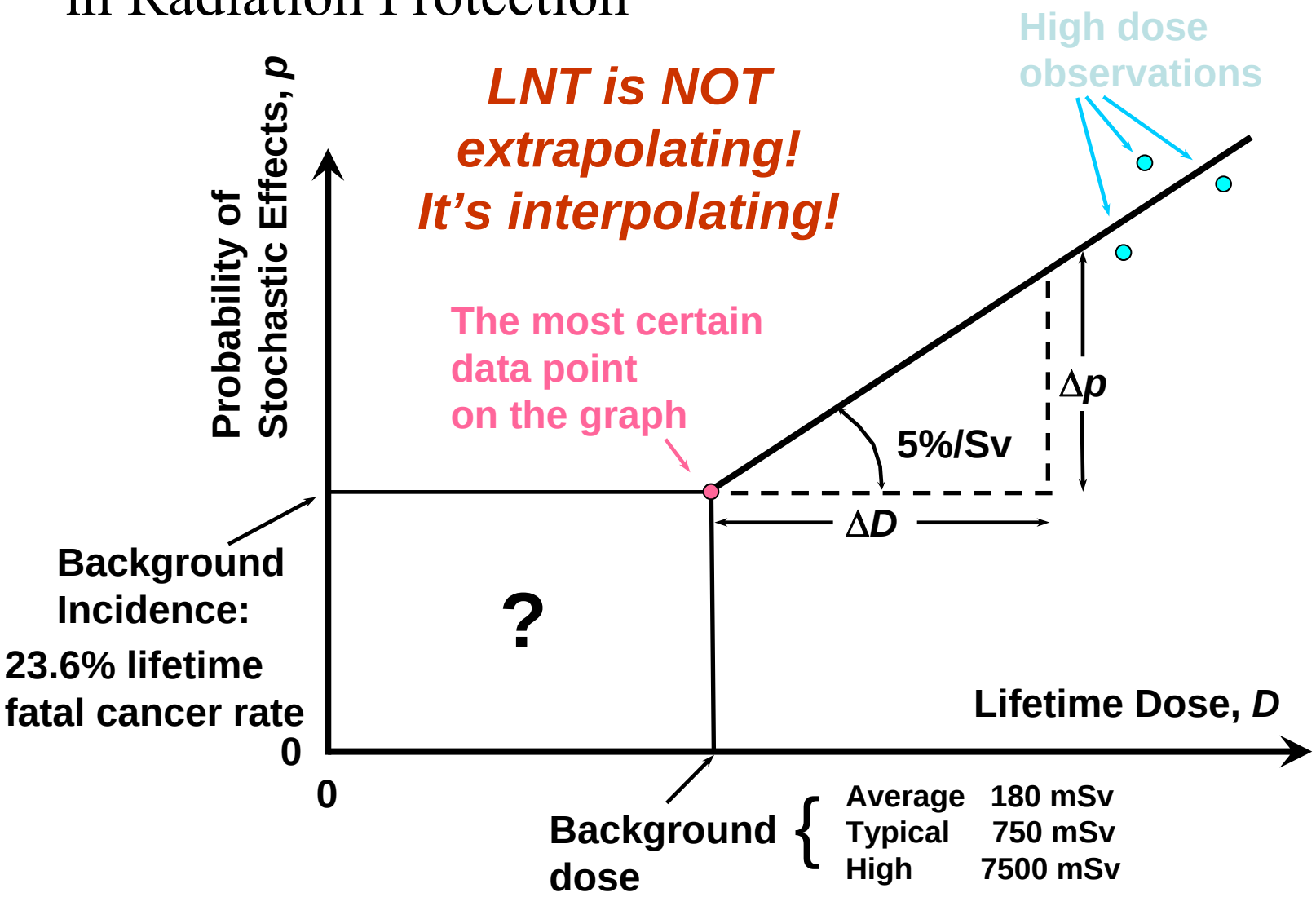
Bayesian Decision Process

- Depends on whose prior is used
  - "Management Prior" or "Optimist's Prior:" There's no way that there's activity in this sample (zero and low values are heavily weighted)
  - "Pessimist's Prior:" Of course there's activity in this sample, and it's bigger than you think (high values weighted)
  - "Ignorant (or Uniform) Prior:" Anything is possible
- Bayesians can recount and resample, too!
- More information is better, but takes time and €€

LNT: Extrapolation or Interpolation?

- The use of the linear non-threshold dose response model as a basis for radiation protection
  - extrapolation from high doses? or
  - interpolation between high dose data to zero added dose data?

The Linear Nonthreshold (LNT) Dose-Response Model in Radiation Protection



After Abel González, IAEA, presented at 1999 Bridging Radiation Policy and Science

Addendum: Software & Reference Links

Daniel J. Strom

Pacific Northwest National Laboratory

Richland, Washington USA

strom@pnl.gov +1 (509) 375-2626

Work partially supported by the U.S. Department of Energy under Contract No. DE-AC06-76RLO 1830

Units and Constants

- There is a “right” way to do things
- National Institute of Standards and Technology (NIST) is the authority in the USA
- http://physics.nist.gov/cuu
- NBS Special Publication 330 (1991) http://physics.nist.gov/Document/sp330.pdf
- NIST Special Publication 811 (1995) http://physics.nist.gov/Document/sp811.pdf

Web Resources: Documents & Data

- National Nuclear Data Center http://www.nndc.bnl.gov/
- *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/, 2007 July 15

Web Resources: Software

- NRC's Radiological Toolbox http://ordose.ornl.gov/downloads.html
- Crystal Ball is an add-in to Microsoft Excel www.decisioneering.com
- Strom's Lognormal Codes: Both LOGNORM4 and the Lognormal Fitting Utility can be downloaded at http://qecc.pnl.gov/LOGNORM4.htm
- RESRAD http://web.ead.anl.gov/resrad/home2/

Recent Papers that Were Not Mentioned (But May of Great Interest)

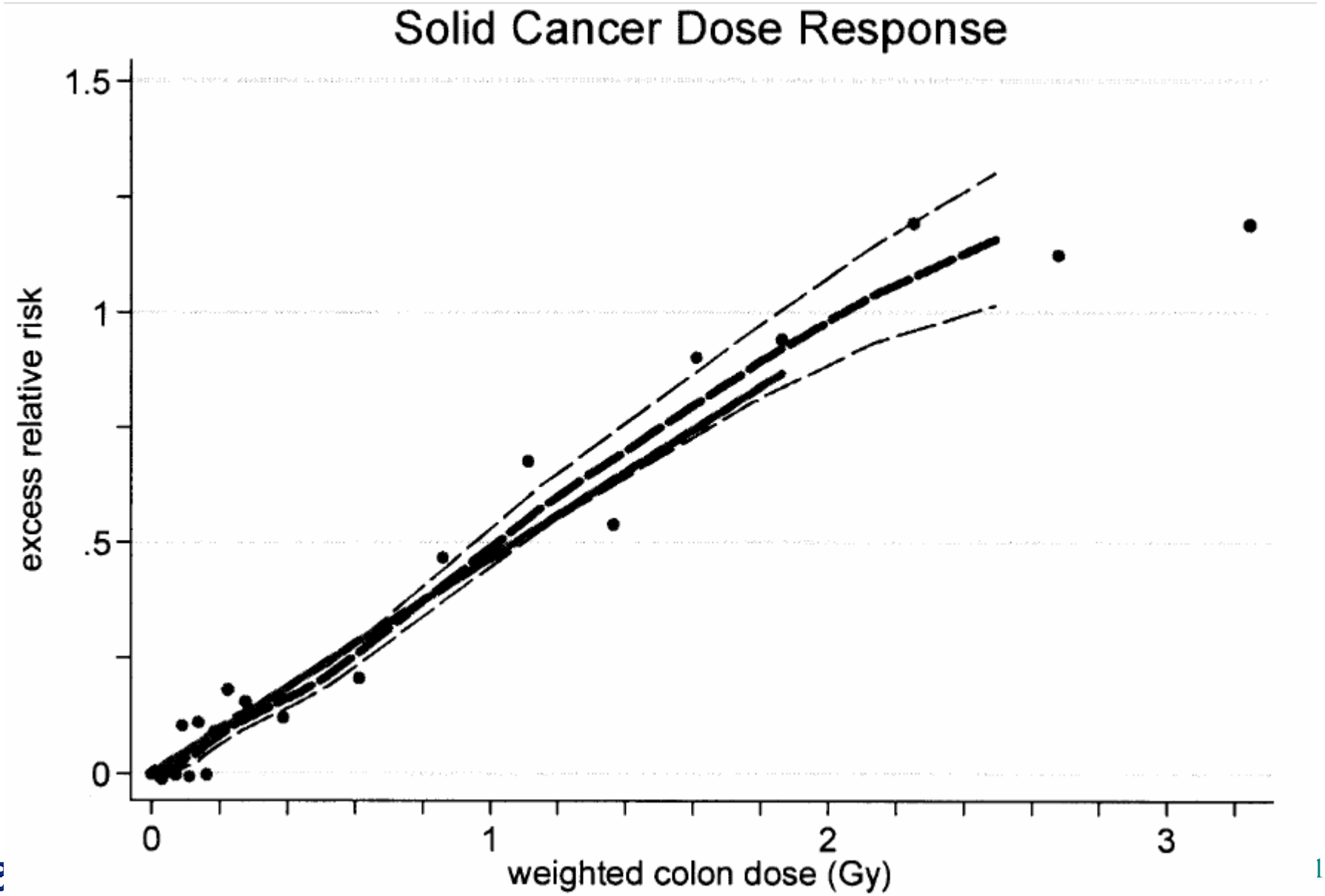
Medical Doses: May 2007

- Amis ES, Jr., PF Butler, KE Applegate, SB Birnbaum, LF Brateman, JM Hevezi, FA Mettler, RL Morin, MJ Pentecost, GG Smith, KJ Strauss, and RK Zeman. 2007. "American College of Radiology white paper on radiation dose in medicine." J.Am.Coll.Radiol. 4(5):272-284
- Addresses huge increase in diagnostic imaging
- Proposed action plans for
    - Measurements
    - Referring Physicians
    - Radiologists
    - Technologists
    - Patients
    - Medical Physicists
    - Vendors
    - Regulatory Agencies, Accrediting Bodies, and Third-party Payers

May, 2007: Cancer *Incidence* in LSS

- Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, Mabuchi K, Kodama K. 2007. Solid Cancer Incidence in Atomic Bomb Survivors: 1958–1998. Radiat. Res. **168**, 1–64
- "The data were consistent with a linear dose response over the 0- to 2-Gy range, while there was some flattening of the dose response at higher doses.
- "Furthermore, there is a statistically significant dose response when analyses were limited to cohort members with doses of 0.15 Gy or less.
- "Significant radiation-associated increases in risk were seen for most sites, including oral cavity, esophagus, stomach, colon, liver, lung, non-melanoma skin, breast, ovary, bladder, nervous system and thyroid."

Preston et al. 2007: LSS CA Incidence



Genetic Susceptibility to Radiation in a Human Population (April 2007)

- Flint-Richter P and S Sadetzki. 2007. "Genetic predisposition for the development of radiation-associated meningioma: an epidemiological study." Lancet Oncol. 8(5):403-410
- Hall EJ. 2007. "Cancer caused by x-rays--a random event?" Lancet Oncol. 8(5):369-370.

Flint-Richter and Sadetzki (2007)

- Ionising radiation is an established risk factor for meningioma, yet less than 1% of irradiated individuals develop this tumour. Familial aggregation of meningioma is rare. We aimed to assess whether genetic factors can modify the risk for meningioma formation after the initiating eff ect of radiation, by comparison of the frequency of meningiomas in families that included irradiated and unirradiated siblings.
- This study was based on a larger epidemiological, genetic case-control study, and included 525 families that were divided according to irradiation and disease status of each of the family's index participant: 160 had radiation-associated meningioma (RAM); 145 were irradiated and did not develop meningioma; 85 had meningioma with no previous history of irradiation; and 135 were unirradiated and did not develop meningioma. Data were collected by questionnaires.

Flint-Richter and Sadetzki (2007)

- We found additional first-degree relatives with meningioma in 17 families (11%) in the RAM group, whereas only between one and two such families (1%) were found in the other groups ($p<0{\cdot}0001$). All meningiomas seen in the families of the RAM group were in irradiated participants. Also, 22 families (10%) in the RAM group had members with cancers in irradiated sites (including head, neck, and chest) compared with 9 (5%) of irradiated controls ($p=0{\cdot}04$).
- This dataset of families, which included irradiated and unirradiated, and also affected and unaffected family members, created a natural experiment. Our results support the idea that genetic susceptibility increases the risk of developing meningioma after exposure to radiation. Further studies are needed to identify the specific genes involved in this familial sensitivity to ionising radiation. DNA repair and cell-cycle control genes, such as the ataxia-telangiectasia gene, could be plausible candidates for investigation.

References

Bureau International des Poids et Mesures (BIPM). Le Système International d'Unités (SI), 6th Ed. French and English Texts. Sèvres, France: Bureau International des Poids et Mesures; 1991.

International Organization for Standardization (ISO). Quantities and Units - Part 0: General Principles. ISO 31-0:1992. Geneva, Switzerland: International Organization for Standardization; 1992.

McGlashan, M.L. Amount of Substance and the Mole. Metrologia 31(6):447-455 1995.

Strom DJ. 2007. Default Assumptions and Methods for Atomic Weapons Employer Dose Reconstruction. Battelle TIB-5000 rev. 00-F; PNWD-3741 Rev. 0; http://www.cdc.gov/niosh/ocas/awedocs.html, NIOSH-Cincinnati.

Taylor, B.N. Guide for the Use of the International System of Units. The Modernized Metric System. NIST Special Publication 811. Springfield, Virginia: National Technical Information Service; http://physics.nist.gov/SI ; 1995.