Statistical Inference in Environmental Radioactivity Measurements

Daniel J. Strom
Pacific Northwest National Laboratory
Richland, Washington USA
+1 509 375 2626 strom@pnl.gov
Presented in Madrid, 18 November 2003

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

Statistical Criteria for Decision-Making in Environmental Cleanup

- *a priori* determinations of detection capabilities
- *a posteriori* decisions of whether radioactivity has been detected in a particular sample
- total propagated uncertainty in measurement results
- sampling strategies

Most Difficult: Alpha-Emitters

- problems are most difficult for alpha-emitting radionuclides
  - ^{230}Th (found in uranium mill tailings)
  - Pu (from reprocessing of irradiated nuclear fuel)
- expense of sampling and laboratory analysis
  - α-spectrometry
  - mass spectrometry (ICPMS, TIMS)
- expense and consequences of incorrect decisions

Outline

- The problem of terminology
- The *two* "counting problems"
- A discussion of "total propagated uncertainty"
- Individual sample decision levels and the "Great Leap of Inference"
- Compare measurements with decision threshold, not detection level
- Decision strategies (Bayesian and classical)
- Consequences of wrong decisions: do we need better decision rules?
- Detection capabilities
- Report and record uncensored, un-rounded measurements results with uncertainties
- Utility and limitations of averaging results of many samples
- Probabilistic blank and environmental background subtraction
- Non-Bayesian methods of uncertainty evaluation in modeling
- References

Terminology 1

- **measurand**: the unknown value of a physical quantity representing the "true state of Nature"
- **measured result**: result of a measurement made of a measurand
- **error**: the [unknowable] difference between a measured result the actual value of the measurand
- **uncertainty of measurement**: a "parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measureand."
  - a bound for the likely size of the measurement error
- **spurious errors** include **blunders, mistakes**

Terminology 2

- 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.

- *This presentation focuses on Type A uncertainties.*

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

$u_j \approx$ corresponding standard deviation;

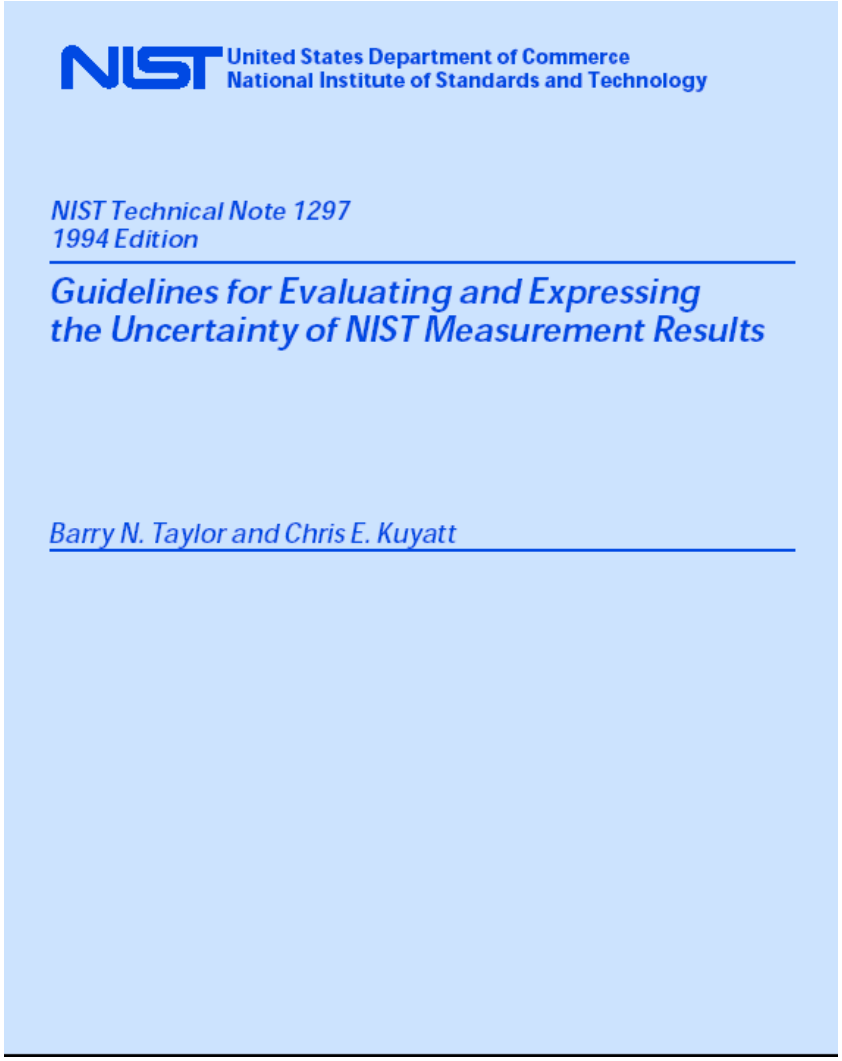
$$u_j = \left|\sqrt{u_j^2}\right|;$$

$u_j^2 \approx$ 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.

NIST Technical Note 1297 (1994)

- same as 1995 ISO Guide to the Expression of Uncertainty in Measurement


United States Department of Commerce
National Institute of Standards and Technology

NIST Technical Note 1297
1994 Edition

Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results

Barry N. Taylor and Chris E. Kuyatt

- http://physics.nist.gov/cuu/pdf/tn1297.pdf

Total Propagated Uncertainty (=Combined Standard Uncertainty)

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

Total Propagated Uncertainty

- Includes
  - counting uncertainty for analyte (A)
  - counting uncertainty for tracer (includes radiochemical recovery and counting yield) (A)
  - uncertainty in tracer calibration (B)
  - uncertainty in tracer volume (A)
  - uncertainty in aliquot volume (B)
  - "system" uncertainty (B) ≈ 3%
- We return to *TPU* later

The Two Aspects of the Counting Problem

The Two Counting Problems

- Radioactive decay is a Bernoulli process described by a binomial or Poisson distribution
- 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!

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 sampled from a Poisson distribution 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

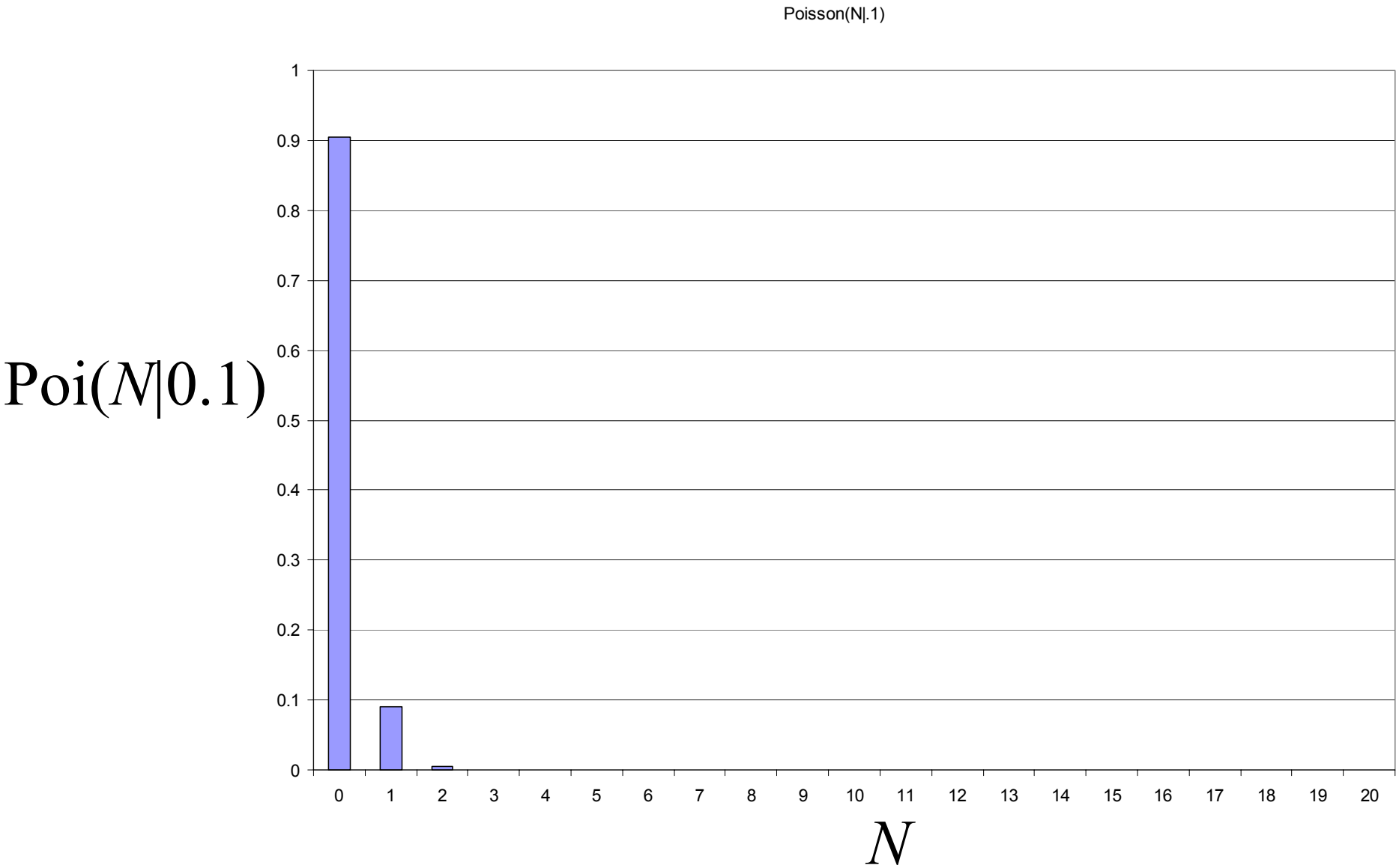
  $p << 1$ (decay probability during counting is low)

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

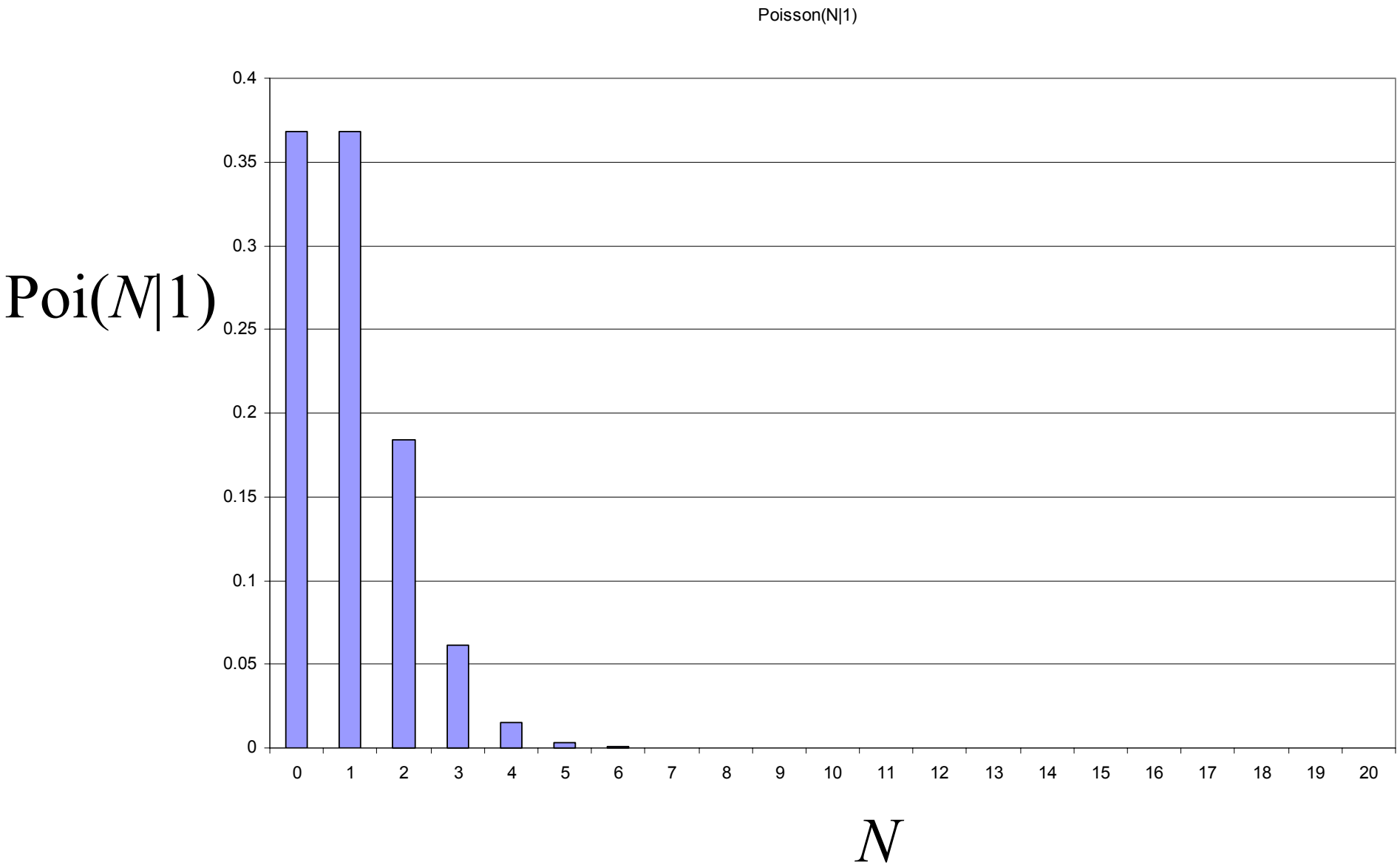
$$Poi(N \mid \mu) = \frac{e^{-\mu}\mu^{N}}{N!}$$

- Expected number of counts N given Poisson mean μ

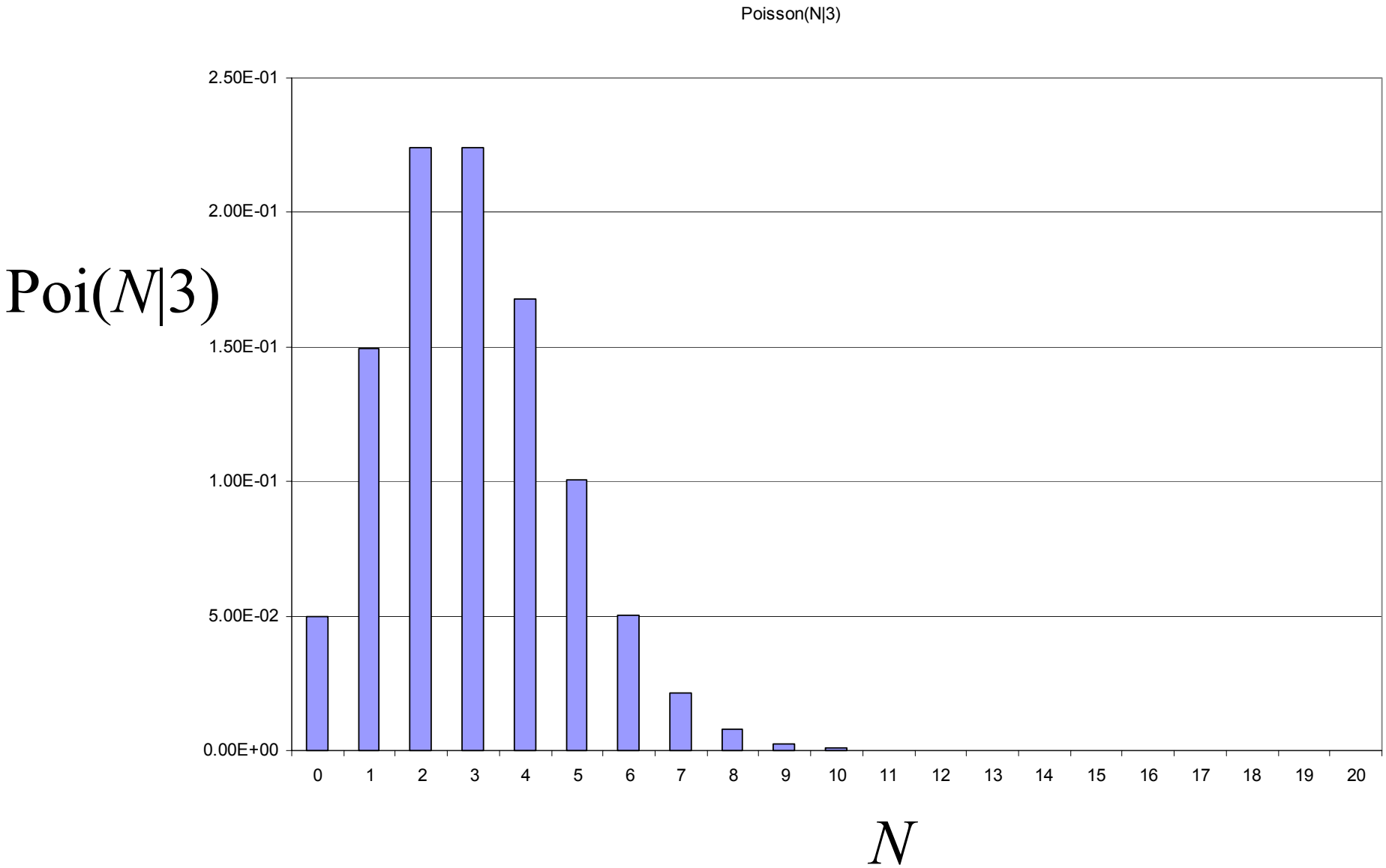
Poisson Distribution, $\mu = \rho t = 0.1$



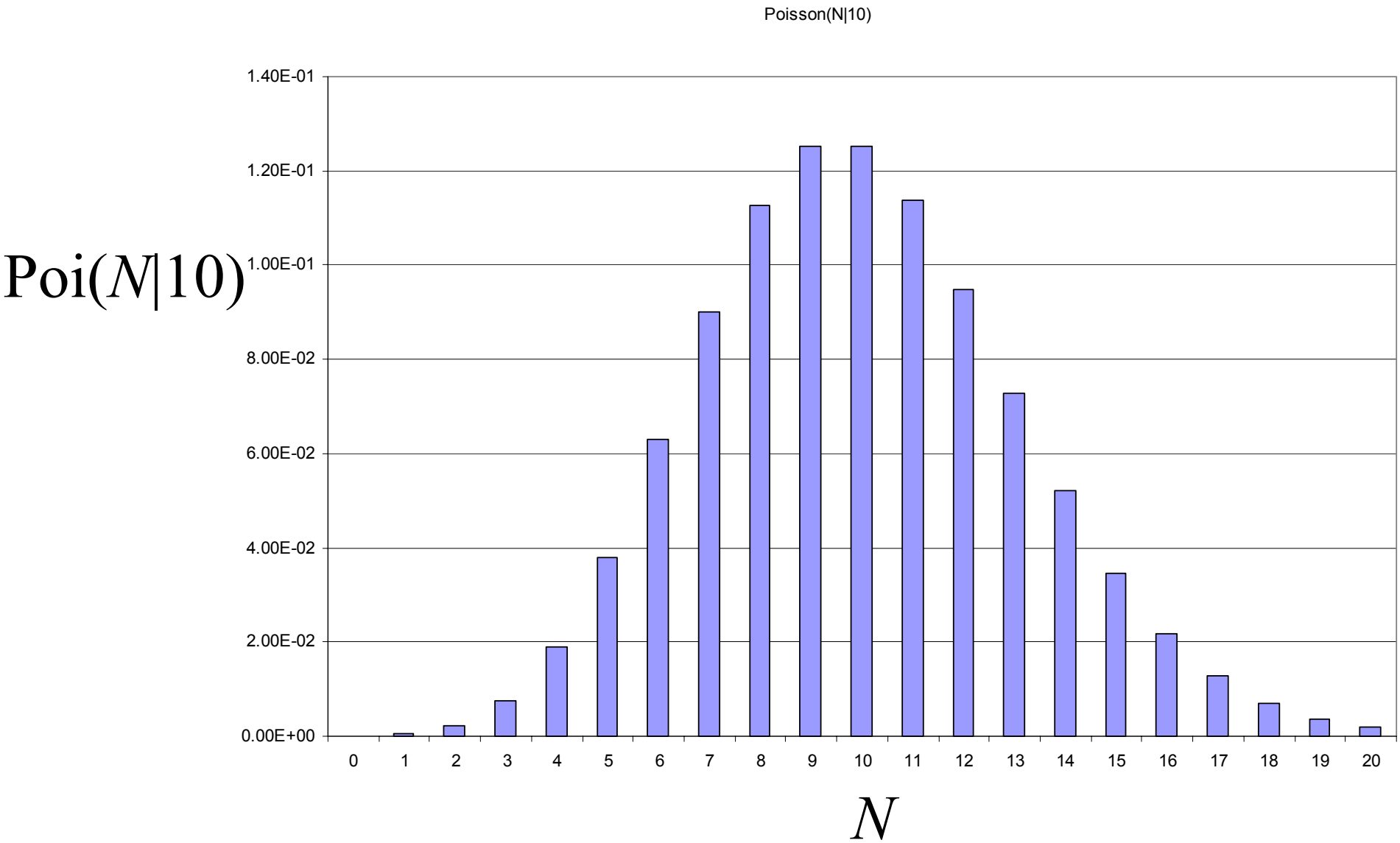
Poisson Distribution, $\mu = \rho t = 1$



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



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



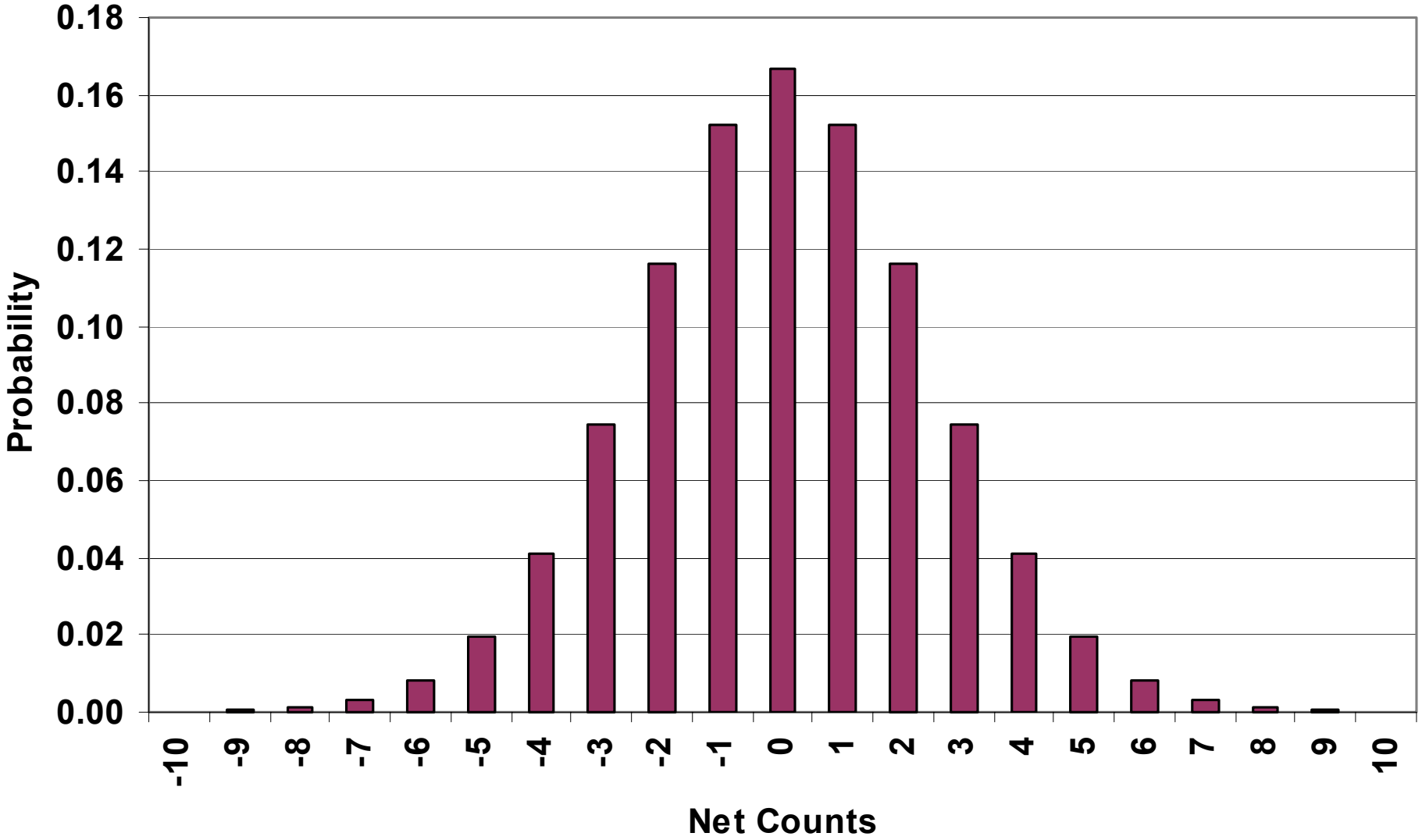
Normal Approximation to the Poisson

- 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

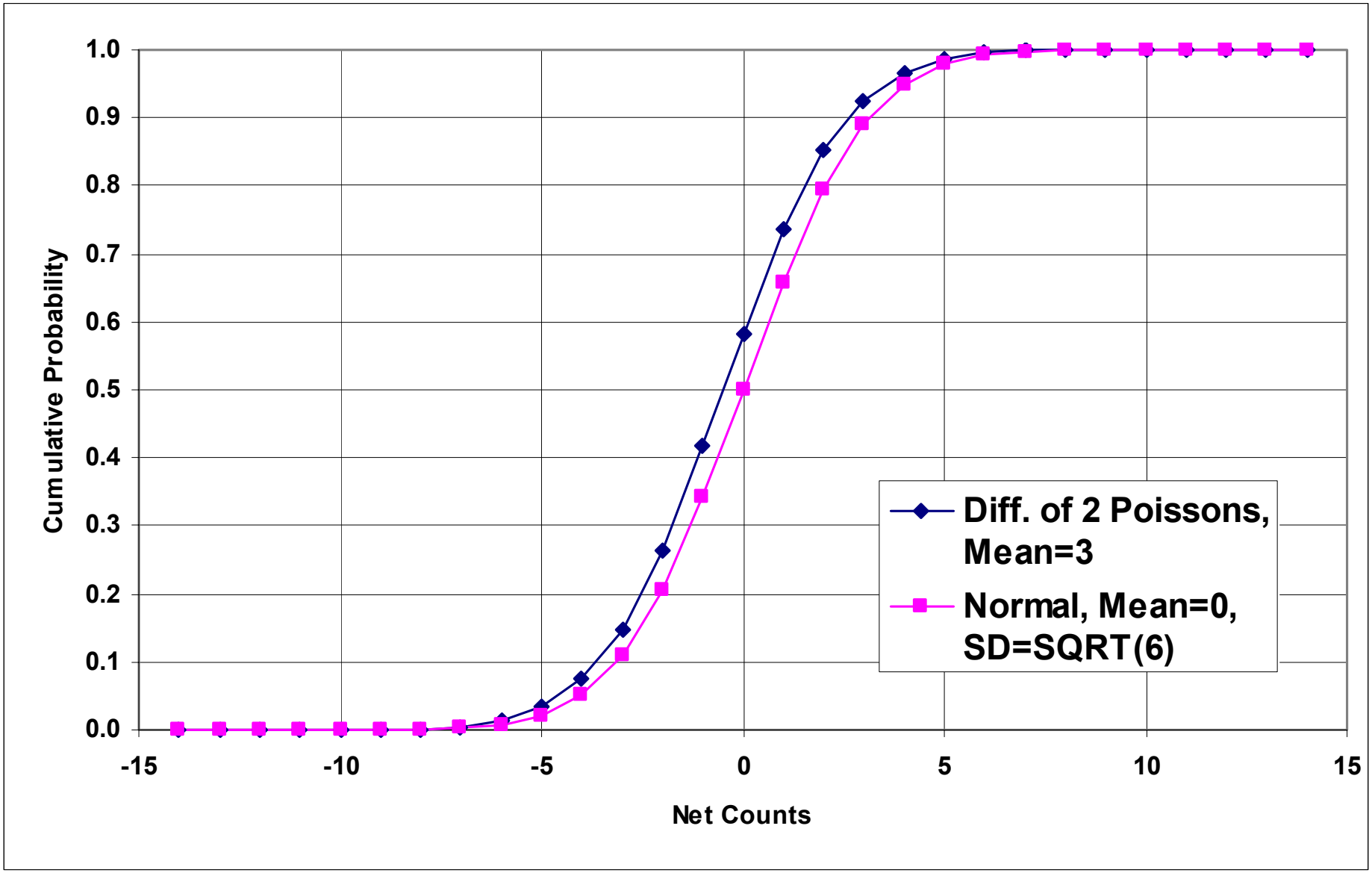
Difference of 2 Poisson Distributions

- 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

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

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

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

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

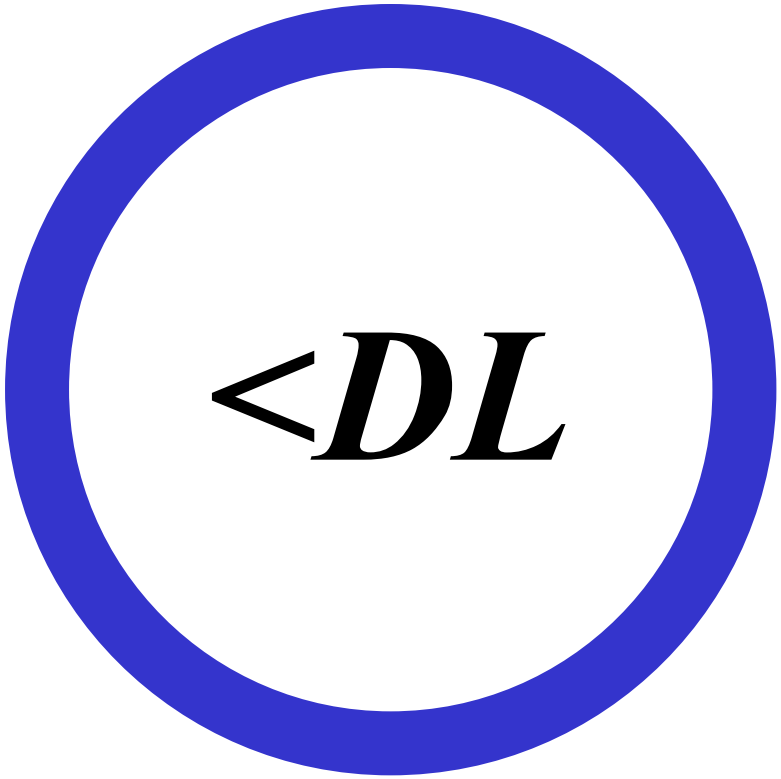
The Reverse Problem: Using Observed Quantities to Estimate Population Parameters (Measurands)

- 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*
- Bayesian approach shown later

Decision Rules

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

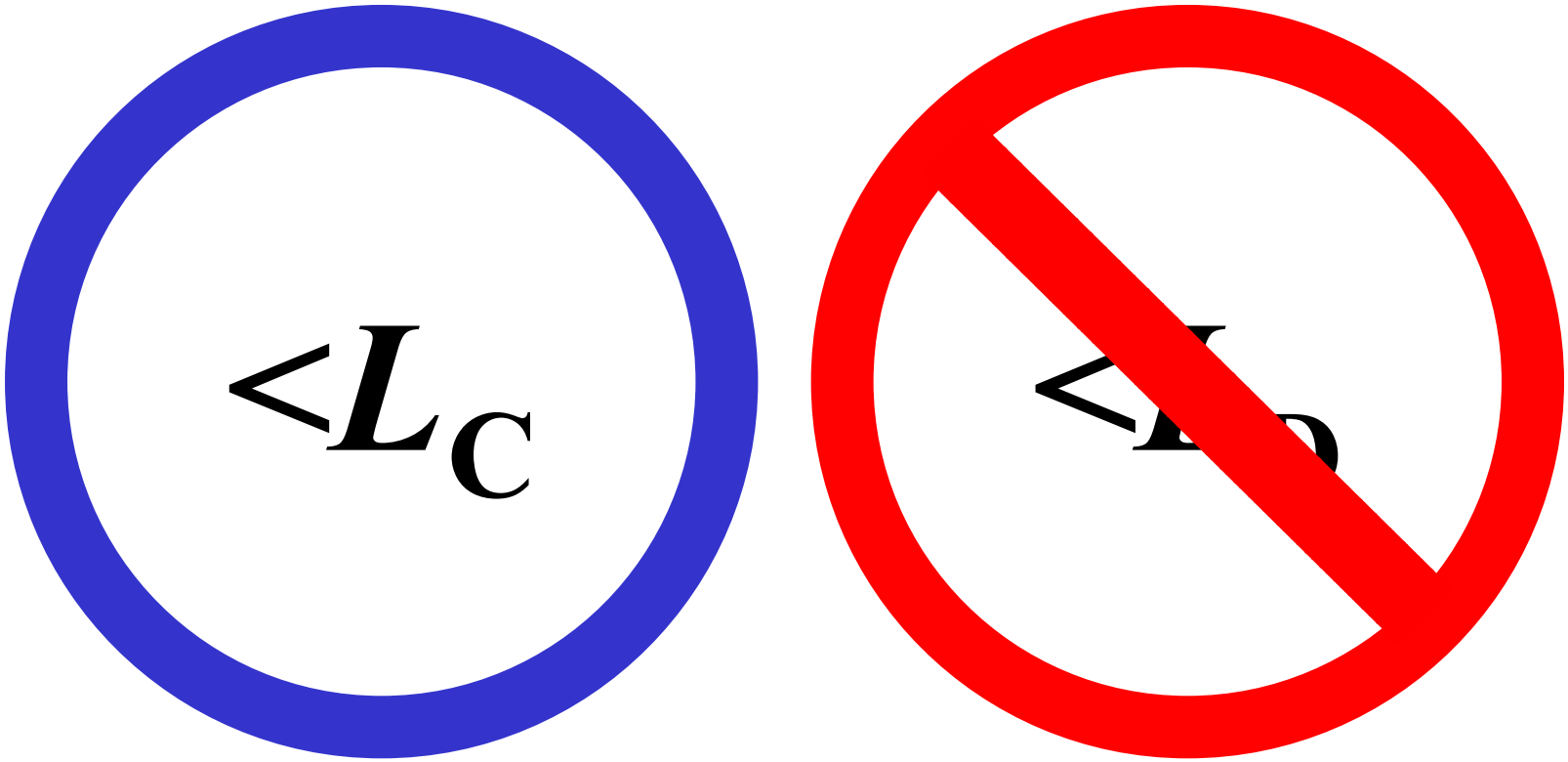
	"DL"	***"MDA"***
Name	decision level	minimum detectable amount
What?	the lowest useable action level	NOT an action level!
Use:	compare measurements to *DL*	Use in planning, advertising or in a statement of work for a contractor: "How much will you charge to provide counting services with this *MDA*?"
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!

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	• made the correct decision (no error)	• “false alarm” • false positive • I’ve committed a Type I error
	No	• the alarm should have sounded, but it 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

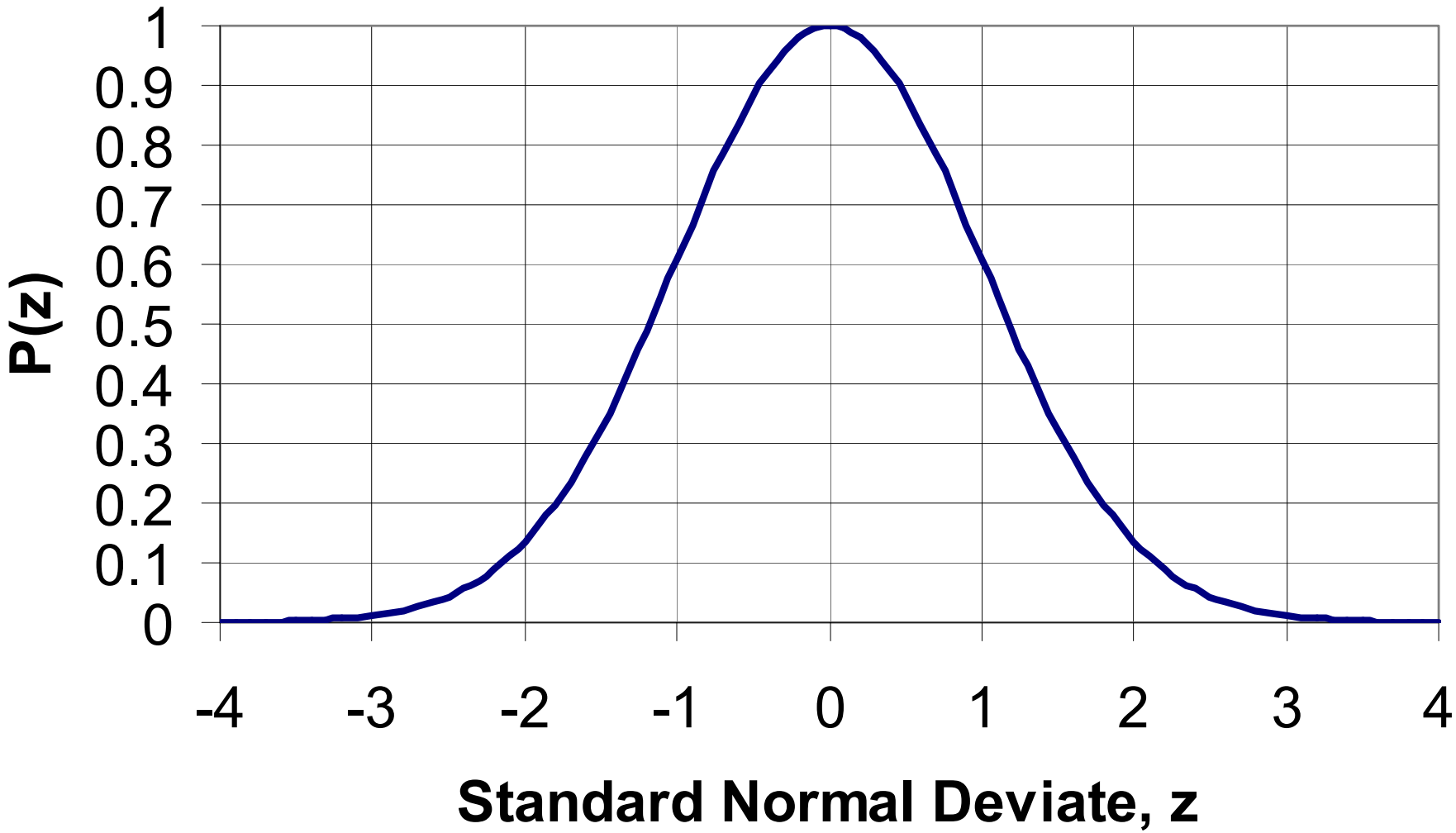
	Assumes Var(μ) = N	Assumes or permits Var(μ) > N	Exact Method or Binomial Distribution
Uses Blank (background) Counts Only	ISO 11929-1 2000 Currie (1968) ANSI N-13.30-1996 Altshuler & Pasternak (1963) Nicholson D_2 (1963)	DL_{N+1} (in Strom & MacLellan 2001)	
Uses Blank and Sample Counts	Nicholson D_1,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

Current Decision Level

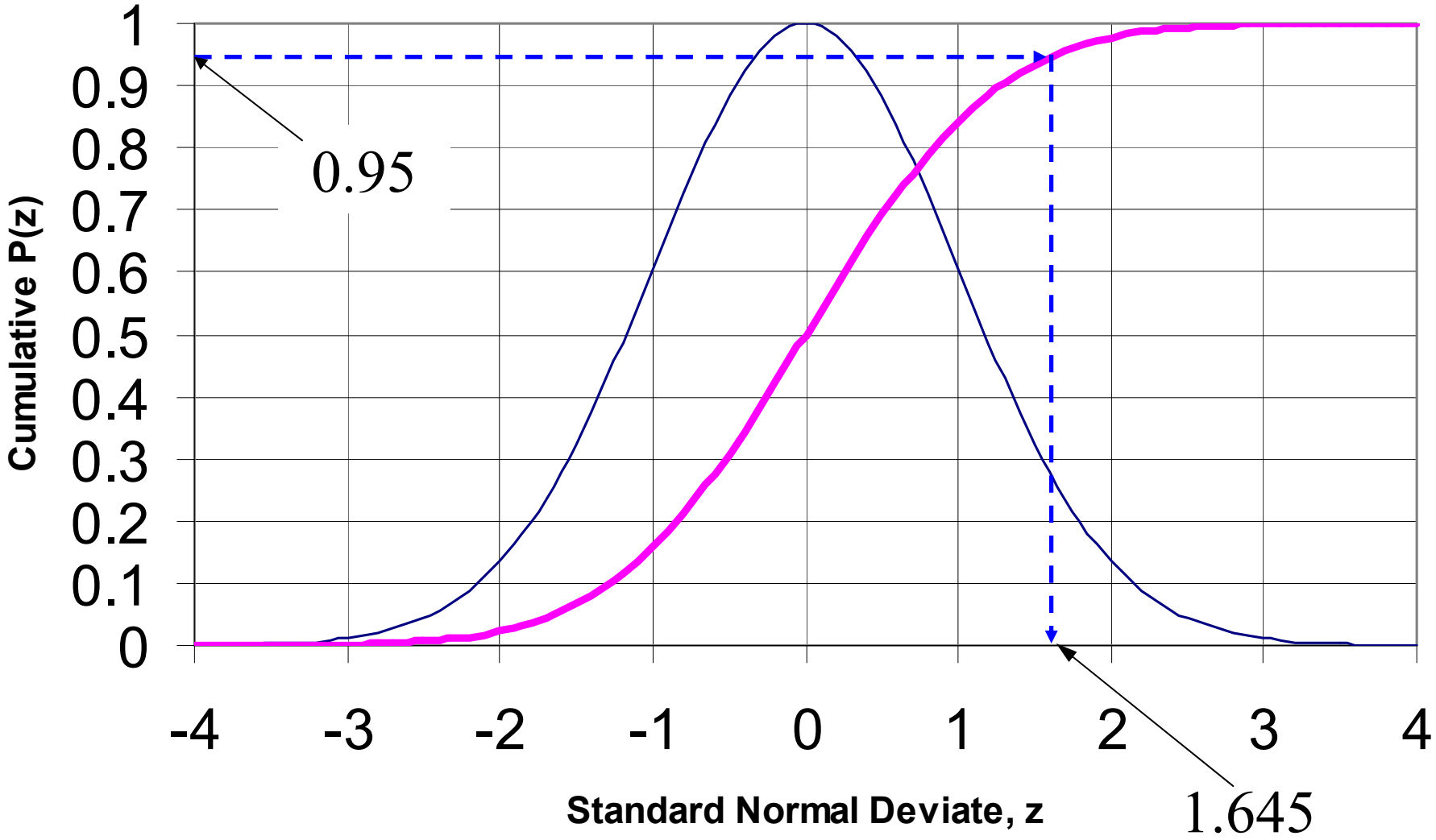
(a.k.a. Critical Level)

- 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

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

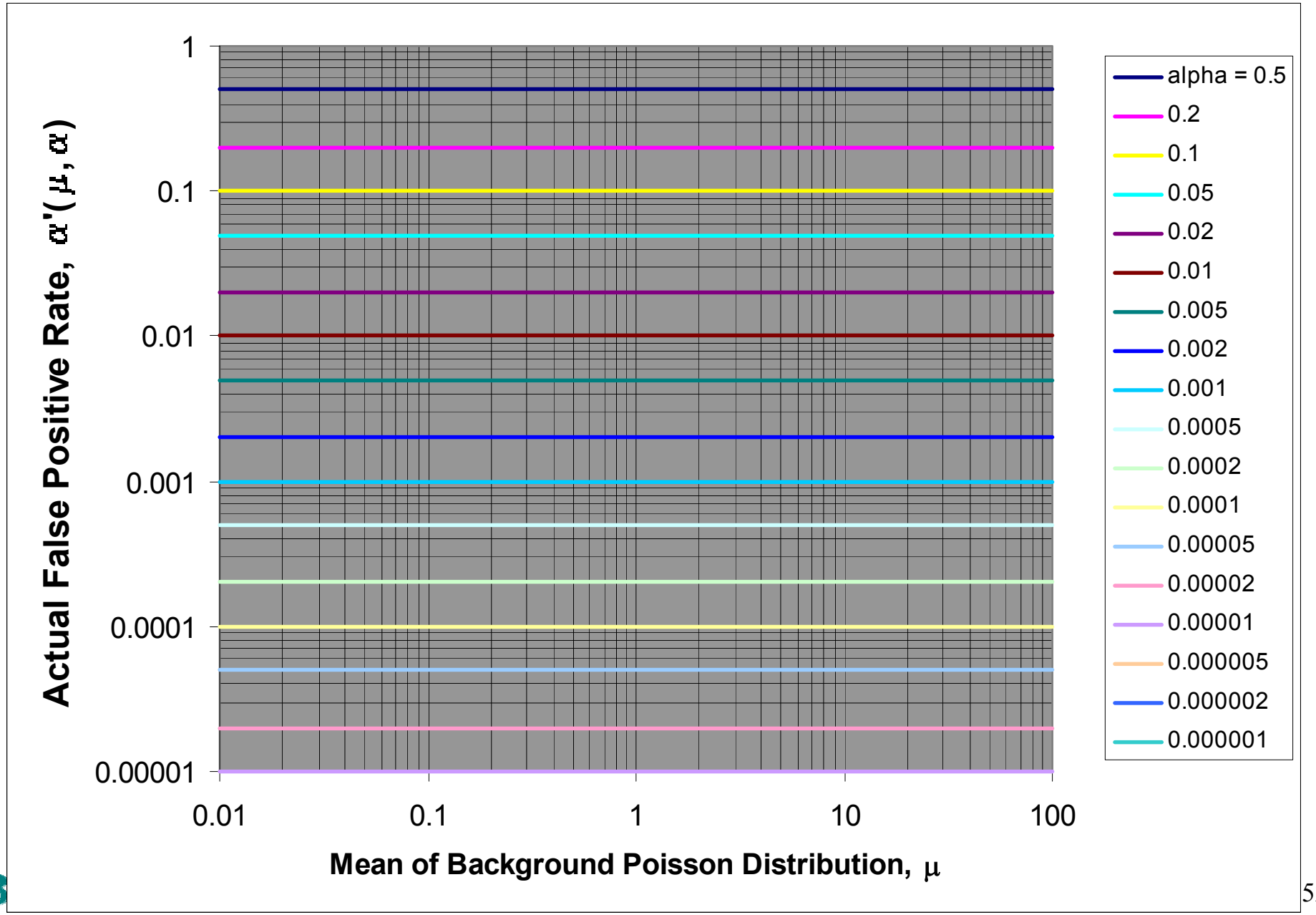
- For $\alpha = 0.05$

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

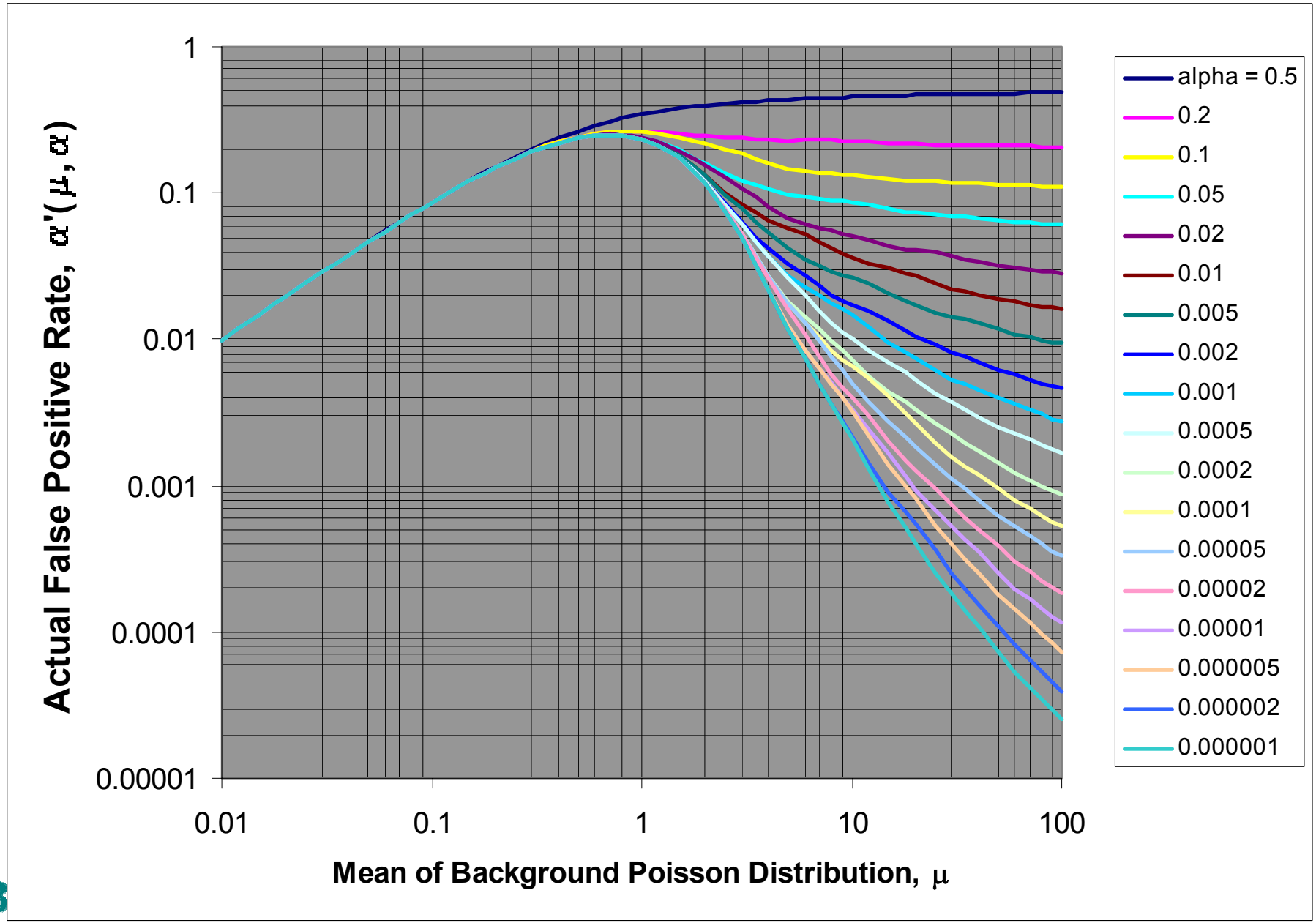
- Expressed as a rate, for non-paired blank:

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

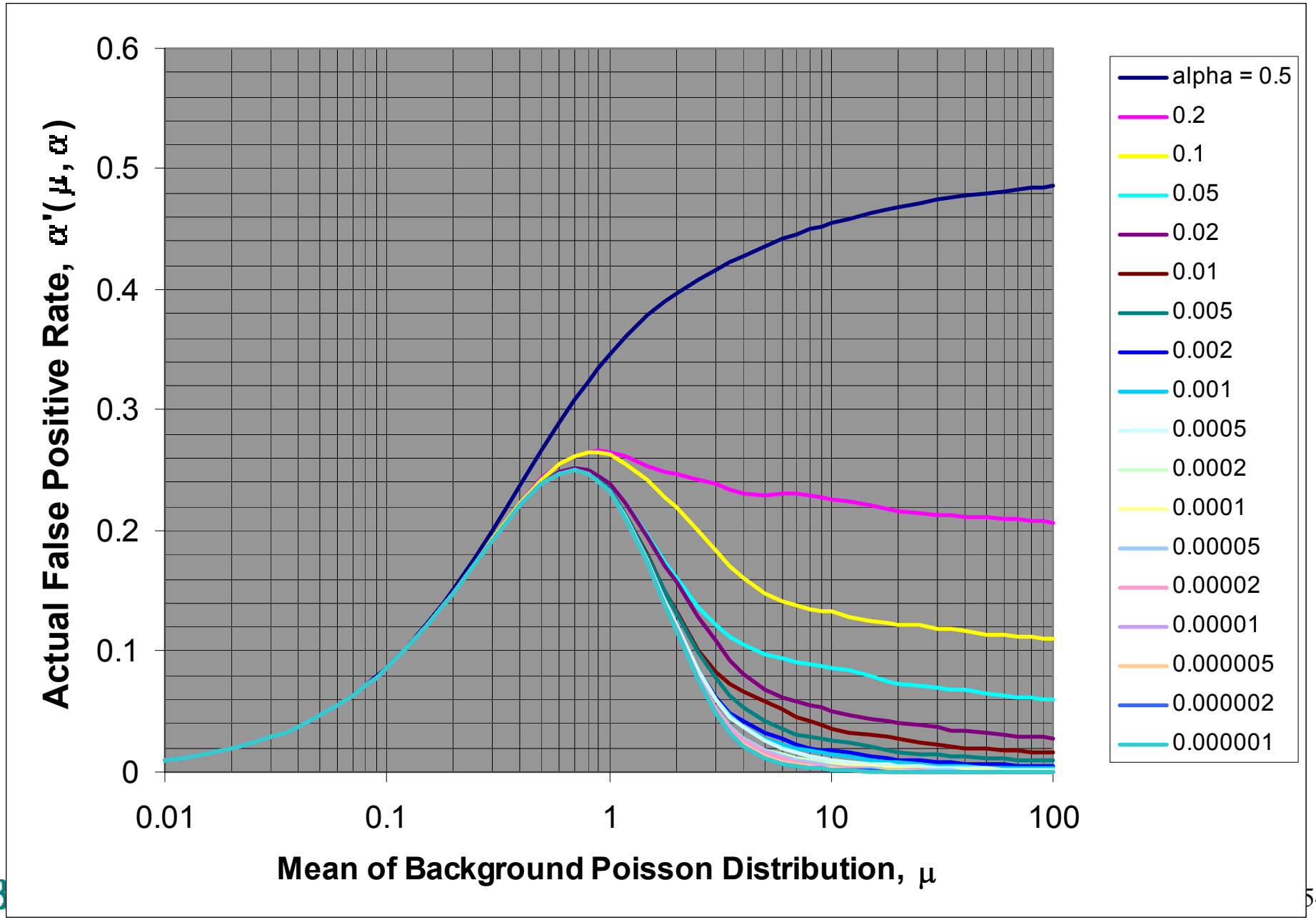
Ideal (Impossible) *DL*



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



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

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

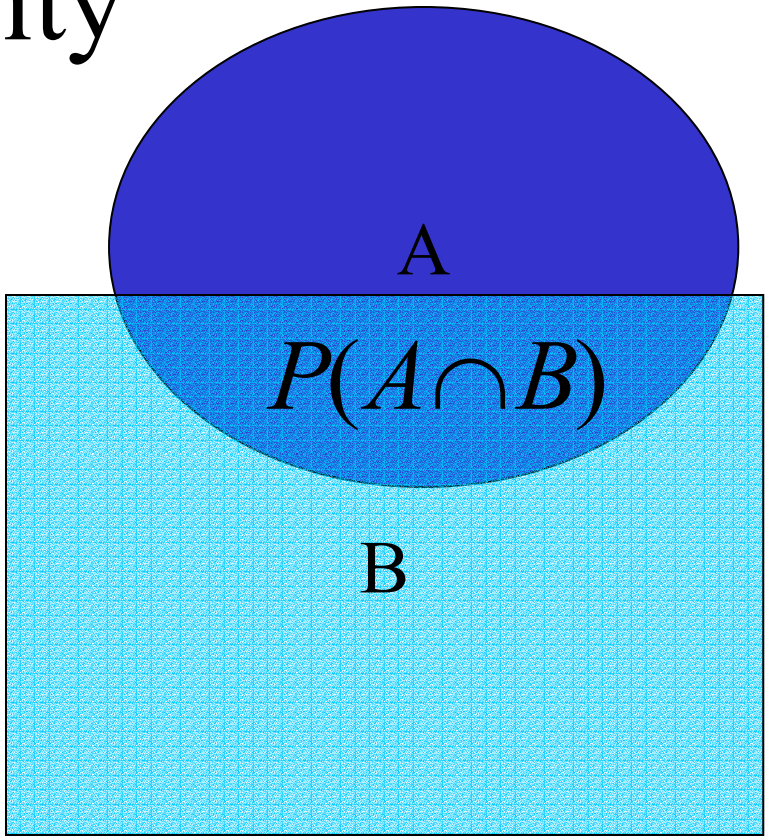
- Even at $\mu_b = 10$, only asymptotically approaches α for larger values
- For very small α, no good even at $\mu_b = 100$!

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:

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

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

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

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

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

Probability that μ lies between μ_1 and μ_2 =

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

Bayesian Approach for Blank Only

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

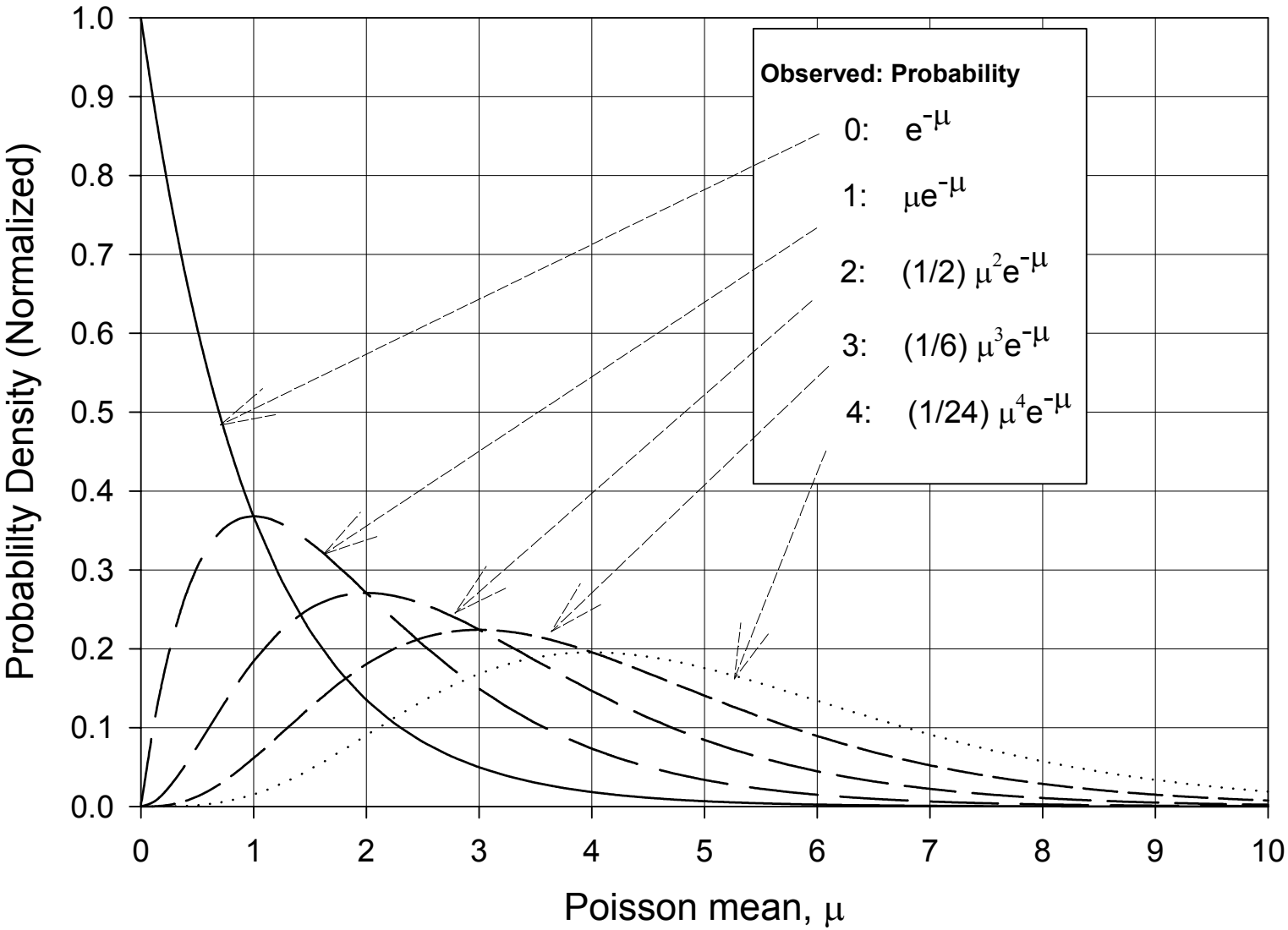
$$L(N \mid \mu) = \frac{e^{-\mu}\mu^{N}}{N!}$$

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 11929-1:2000)
- 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

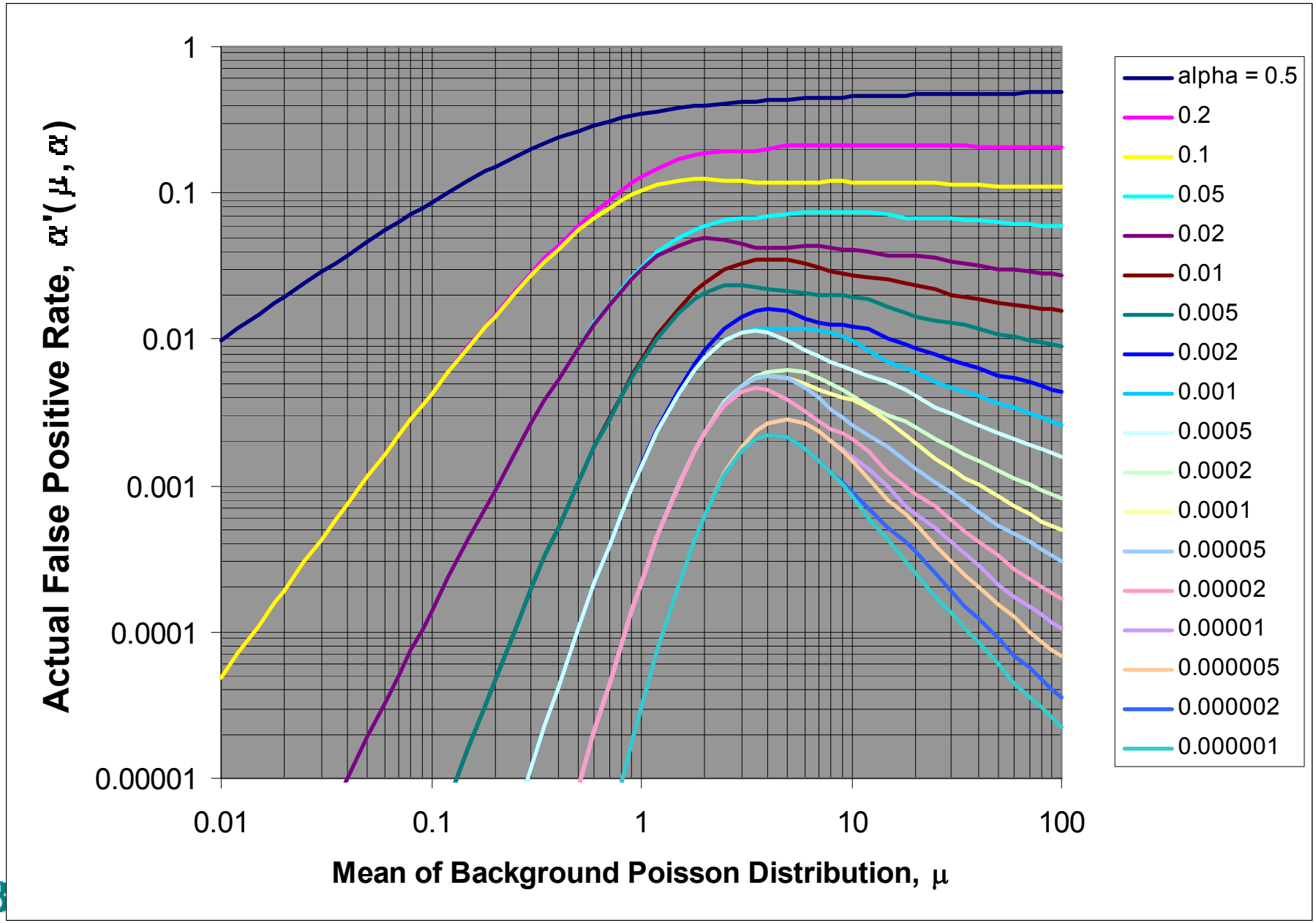
- Bayesian inference of background rate
- Question: If one observes N_b counts, what is the ***expectation value*** of the background distribution that gave rise to this observation (see figure)?
- Bayesian Answer (uniform prior): $\mu_b = N_b + 1$

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

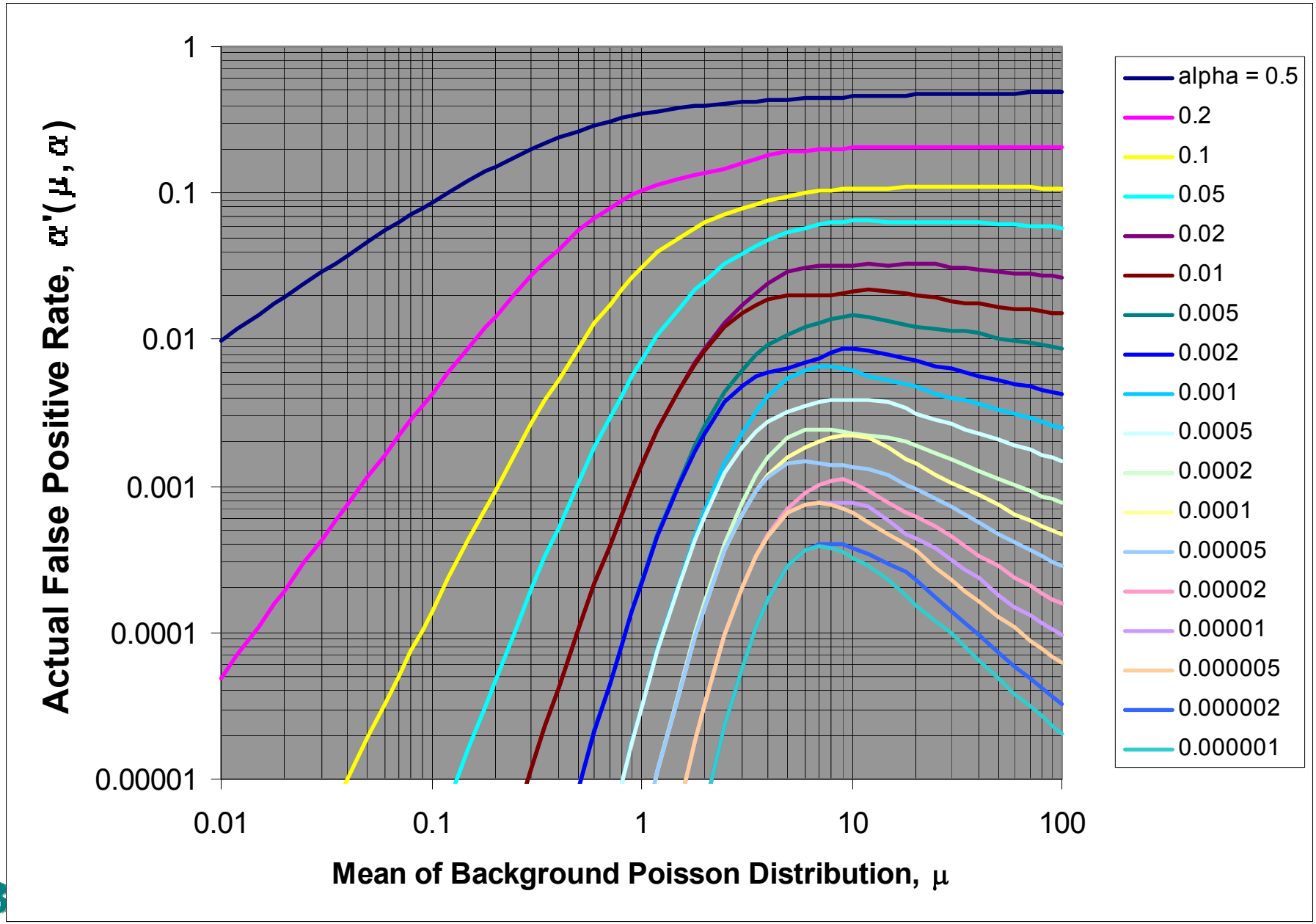
- Idea: Rainwater & Wu 1947; Friedlander & Kennedy 1955; Friedlander et al. 1963; Stevenson 1966; Little 1982

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

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



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



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

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, $N_{total} = N_b + N_g$
- probability of success = $t_g/(t_g+t_b)$
- 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

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

where

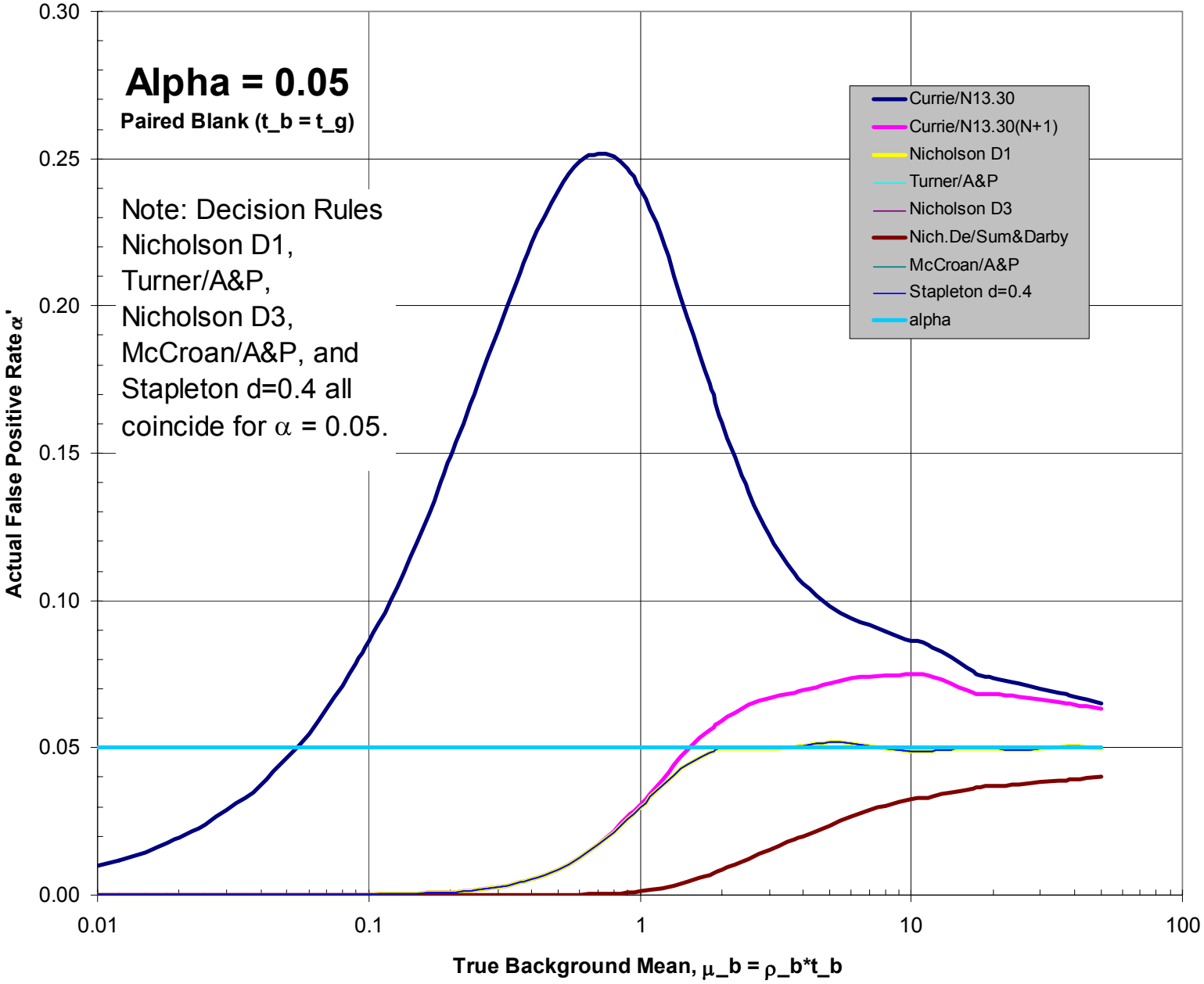
$$Q_o = \frac{\rho_b t_g}{\rho_b t_g + \rho_b t_b} = \frac{t_g}{t_g + t_b}$$

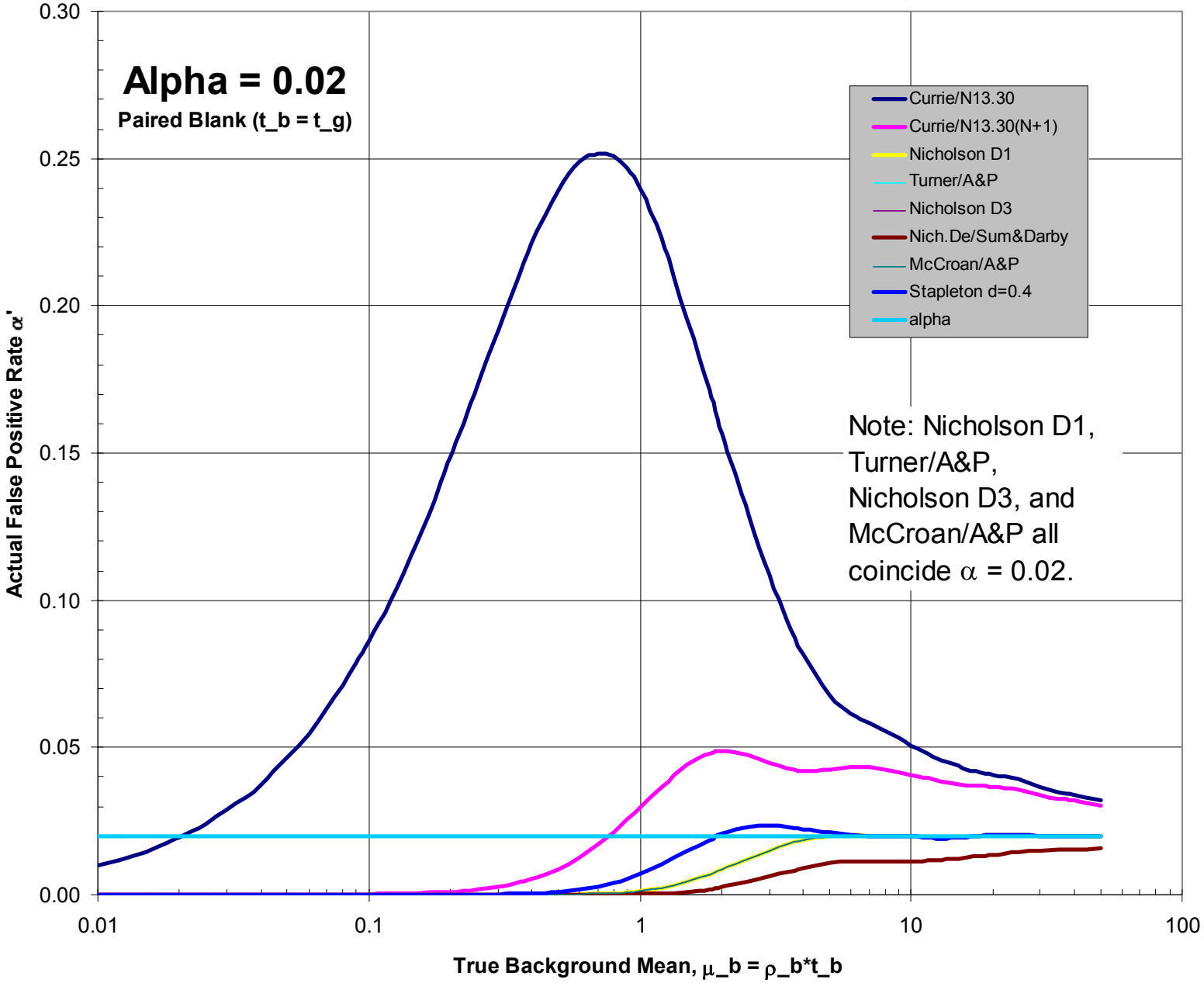
and

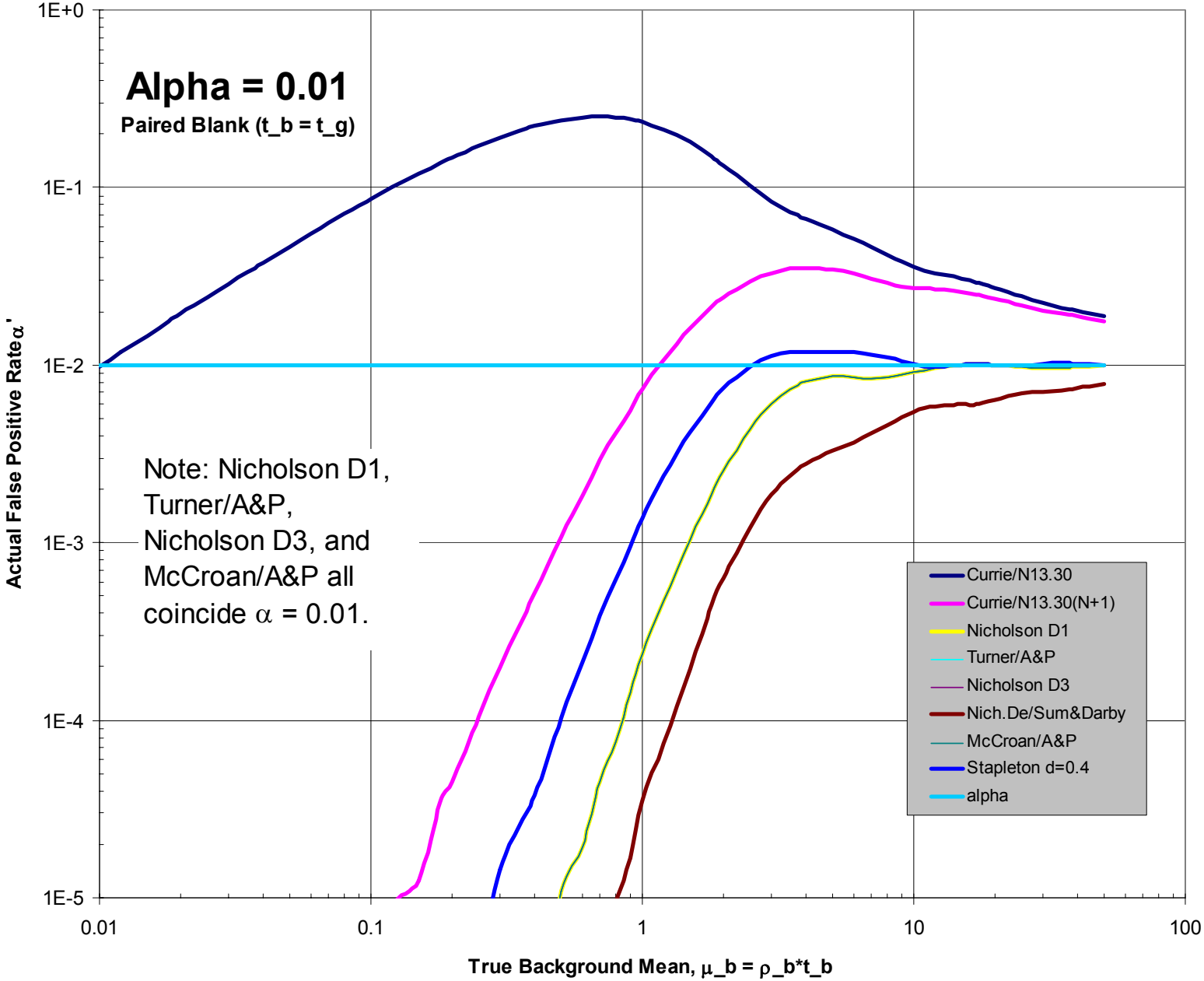
$$\begin{pmatrix} N_{\text{total}} \\ N_g \end{pmatrix} \text{ is a binomial coefficient.}$$

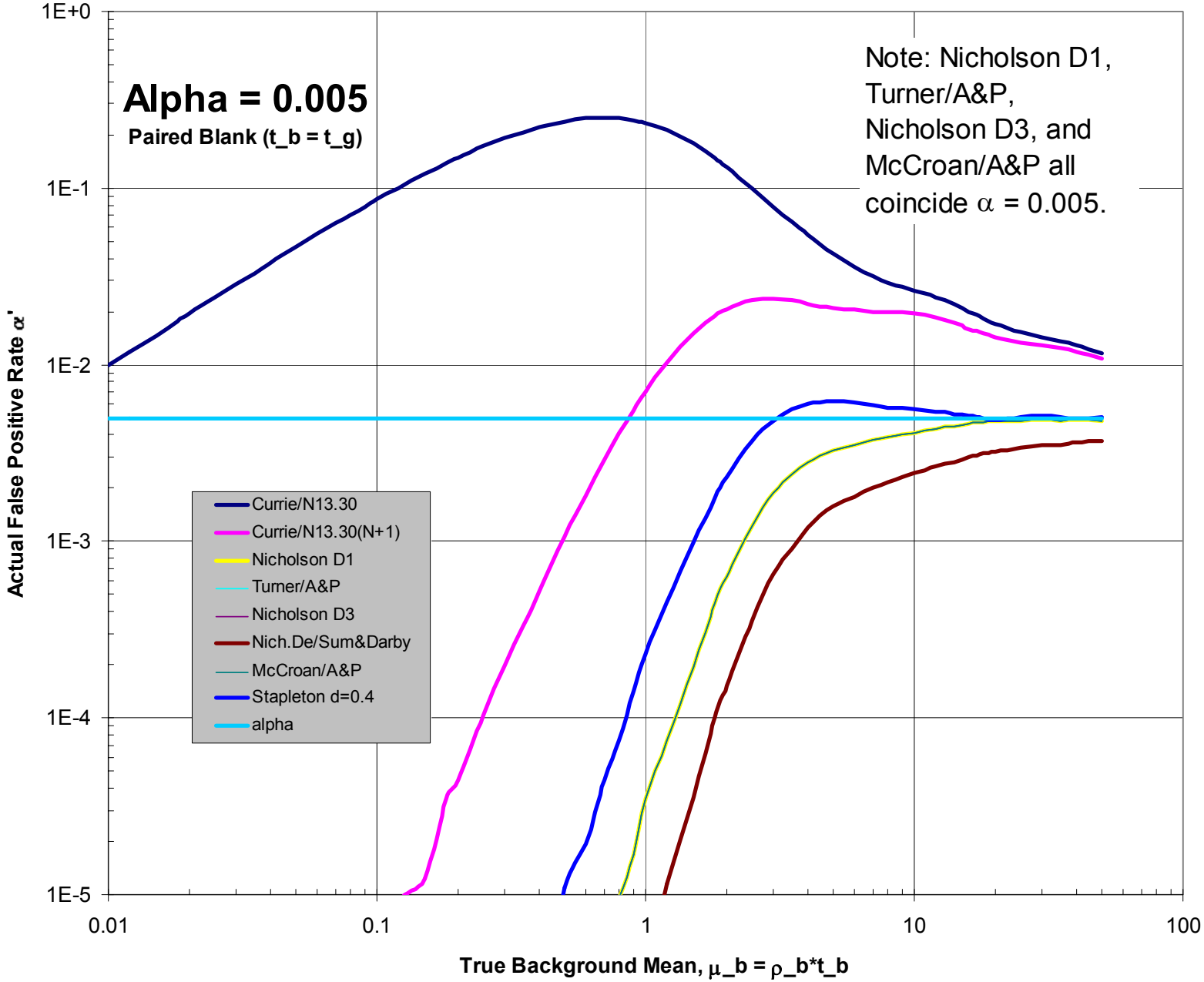
Test of Decision Rules

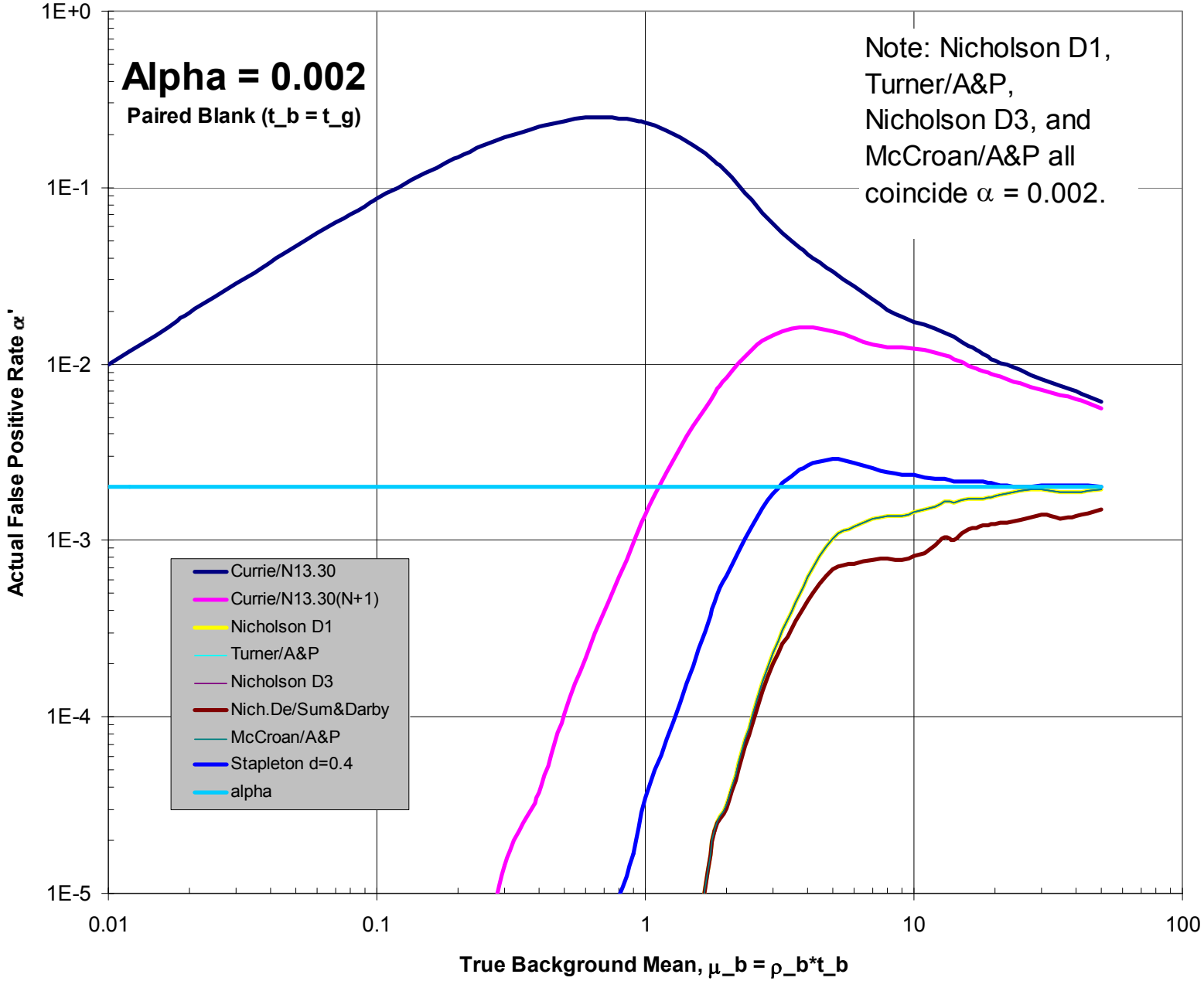
- Monte Carlo simulation (Strom and MacLellan 2001)
- 3,141,593 trials at each of
  - 6 values of α, 0.001 to 0.05
  - 57 values of $\mu_b = \rho_b t_b$ (0.01 to 50)
- MacLellan's exact calculation (MacLellan and Strom 1999) not possible for exact (binomial) or Stapleton's tests or Nicholson's D_1 and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

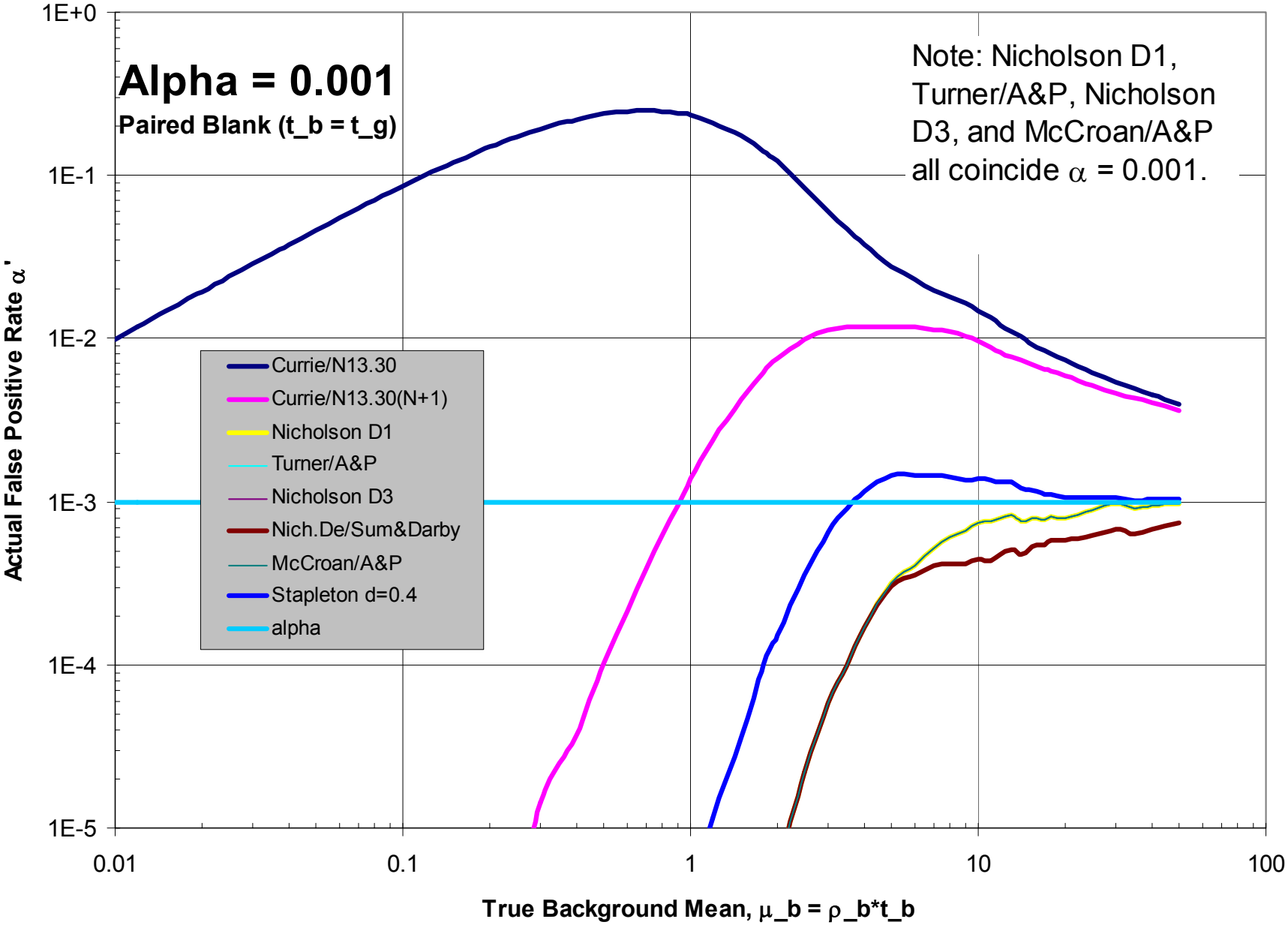

Alpha = 0.05
Paired Blank (t_b = t_g)
Note: Decision Rules Nicholson D1, Turner/A&P, Nicholson D3, McCroan/A&P, and Stapleton d=0.4 all coincide for α = 0.05.
Currie/N13.30
Currie/N13.30(N+1)
Nicholson D1
Turner/A&P
Nicholson D3
Nich.De/Sum&Darby
McCroan/A&P
Stapleton d=0.4
alpha
Actual False Positive 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

PNNL Counting Statistics Utility

Software Utility under Development

Counting Statistics for Radioactivity Measurements - daniel.j.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-4_Project Version 1999-09-13.1
Daniel J. Strom PNNL
danie.j.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	1	2	1.666666666666
	N_b-1 / N_b+1	N_g-1 / N_g+1	
Time	3	1	
	t_b-1 / t_b+1	t_g-1 / t_g+1	
Rate	3.33E-01	2.00E+00	1.67E+00
Std. Dev.	3.33E-01	1.41E+00	1.45E+00
CV			8.72E-01

Calculate

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

Activity	Total Propagated S.D. of Activity	CV of Activity
1.67E+00	1.47E+00	8.83E-01

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?
Currie, Nicholson D2, N13.30	1.10E+00	Yes
Currie using (N_b + 1)	1.55E+00	Yes
A&P, Turner 2nd ed. Eq. 11.68	3.09E+00	No
McCroan's generlaization of A&P	1.64E+00	Yes
James H. Stapleton's test	2.67E+00	No
Nicholson D1	2.39E+00	No
Nicholson D3	1.64E+00	Yes
Nicholson De, Sumerling & Darby	not defined	No

Stapleton's p-value	Stapleton's d, 0 <= d <= 0.5	
0.066799	0.4	Z = 1.500066

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

0.15625

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

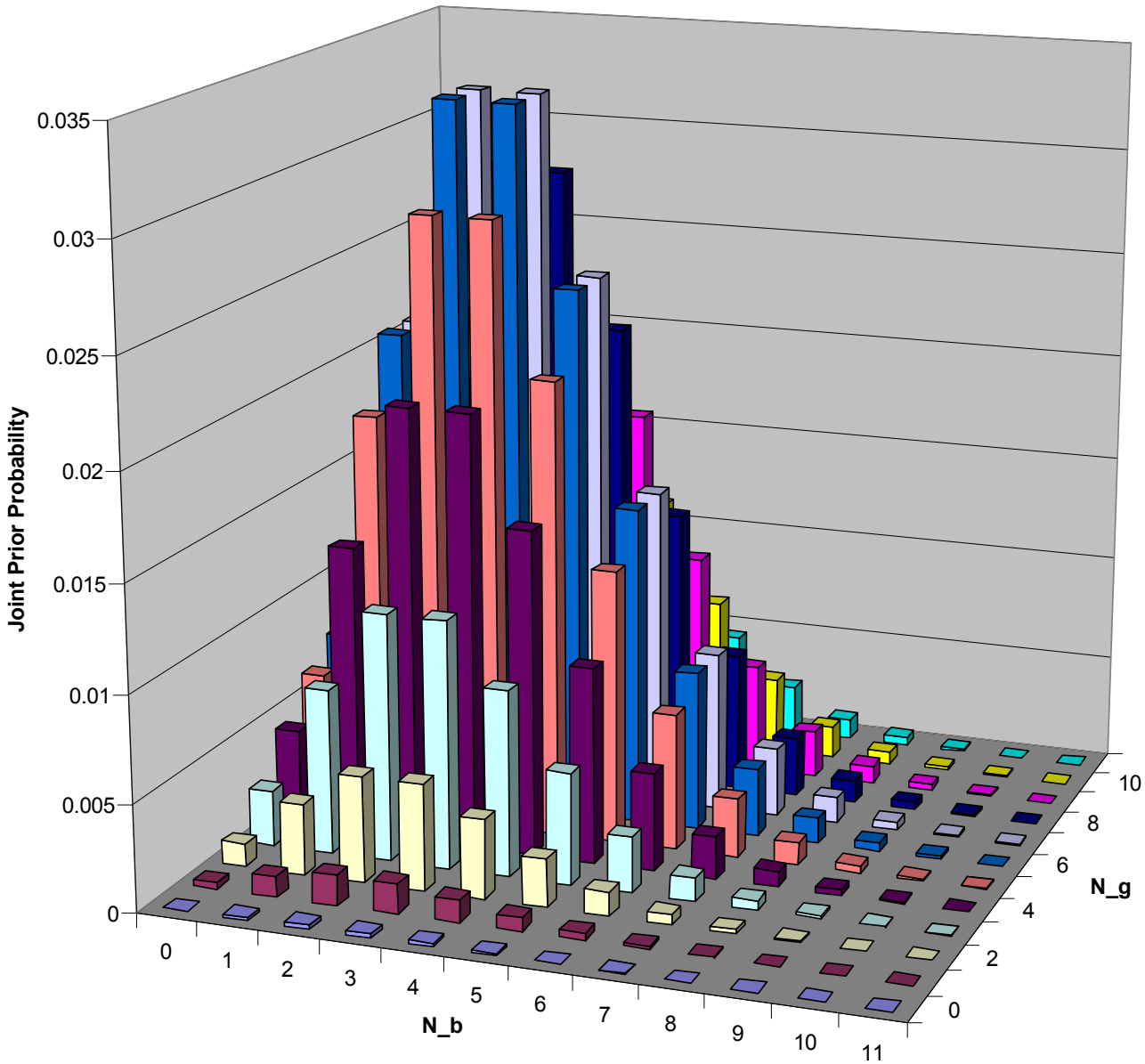
- Assume the likelihood probability densities are Poisson
- Likelihood for N_b and N_g is a function of μ_b and μ_g:

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

- 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

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

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

Back to the Real World…
Decision Strategies

Coefficients of Variation Using N+1

- a better "guess"
- t denotes tracer

$$CV^2(R_\mathrm{n}) = \frac{\dfrac{N_\mathrm{g}+1}{t_\mathrm{g}^2} + \dfrac{N_\mathrm{b}+1}{t_\mathrm{b}^2}}{\left(\dfrac{N_\mathrm{g}+1}{t_\mathrm{g}} - \dfrac{N_\mathrm{b}+1}{t_\mathrm{b}}\right)^2}$$

$$CV^2(R_\mathrm{t}) = \frac{\dfrac{N_\mathrm{t}+1}{t_\mathrm{g}^2} + \dfrac{N_\mathrm{bt}+1}{t_\mathrm{b}^2}}{\left(\dfrac{N_\mathrm{t}+1}{t_\mathrm{g}} - \dfrac{N_\mathrm{bt}+1}{t_\mathrm{b}}\right)^2}$$

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:

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

- "Total Propagated Uncertainty," *TPU*:

$$TPU = A \times CV_{\text{total}}$$

- Hanford uses $2 \times TPU$ as a starting point for decisions

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

- If a bioassay result is **unexpectedly** > 2×*TPU*, 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×*TPU*, obtain new sample (if possible and meaningful) and count new sample
  - if new sample result > 2×*TPU*, 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 €€

Consequences of Wrong Decisions

Do We Need Better Decision Rules?

- False positive bioassays
  - needlessly place worker on work restriction
  - needlessly alarm worker
  - needlessly spend money on unneeded re-sampling and analyses
  - needless investigations
- False negative bioassays
  - not protective of worker
  - if later discovered, can destroy trust and communication

Do We Need Better Decision Rules - 2?

- False positive environmental samples
  - unnecessary, costly cleanup
  - needlessly alarm public
  - needlessly spend money on unneeded resampling and analyses
  - needless investigations
  - political consequences
- False negative environmental samples
  - not protective of public and environment
  - if later discovered, can destroy trust and communication
  - political consequences

Averaging to Improve Detection Capabilities

Applicable Under Some Circumstances

- Examples
  - Monthly, weekly, or daily samples of discharges from a pipe or stack that must meet an annual limit
  - Multiple samples from a "risk analysis unit"
  - Any time a standard applies to a volume that is large compared to sample volume
  - Evaluation of "missed dose" from pooled bioassay samples
- Why it works: mass or number of samples scales with n, counting uncertainty scales with $\sqrt{n}$

Air Sample Example (NUREG-1400)

- n is number of samples
- T_b is blank count time
- T_g is gross (sample) count time
- R_b is background count rate
- E is filter efficiency
- T_S is sampling time
- F is air flow rate; FT_S is volume sampled
- K is counting efficiency (counts s^{-1} Bq^{-1})
- $MD\overline{C}$ is "minimum detectable *average* concentration"

Air Sample Example (NUREG-1400)

$$MD\overline{C} = \left(\frac{\sum_{i=1}^{n} T_{\mathrm{S},i}}{\sum_{i=1}^{n} E_i F_i K_i T_{\mathrm{s},i}^2} \right) \left(\frac{3}{\sum_{i=1}^{n} T_{\mathrm{g},i}} \right) + 3.29 \sqrt{\frac{\sum_{i=1}^{n} \frac{R_{\mathrm{b},i} \left(\frac{1}{T_{\mathrm{g},i}} + \frac{1}{T_{\mathrm{b},i}} \right)}{E_i^2 F_i^2 K_i^2}}{\left(\sum_{i=1}^{n} T_{\mathrm{S},i} \right)^2}}$$

- and if all efficiencies, flow rates, sample times, etc. are equal,

$$MD\overline{C} = \frac{3}{nEFKT_{\mathrm{S}}T_{\mathrm{g}}} + 3.29 \frac{\sqrt{R_{\mathrm{b}} \left(\frac{1}{T_{\mathrm{g}}} + \frac{1}{T_{\mathrm{b}}} \right)}}{\sqrt{n} EFKT_{\mathrm{S}}}$$

Method Works to Extract Signal from Noise

- Summing γ-Spectra
  - 46-keV (4%) ^{210}Pb γ-photon not visible or detected by IGe in vivo counting of 20 uranium miners
  - summing 20 spectra revealed a 46-keV peak (Palmer 1984)
  - can't tell which miner it's in, but can tell what average and total is!
- $1/\sqrt{n}$ principle is widely used
  - kinetic phosphorimetry for chemical determination of U
  - ion cyclotron resonance mass spectroscopy for proteomics
  - any multiple-interrogation technique
- Can add counts of same sample made at different times
  - e.g., recount

Probabilistic Background Subtraction

The Problem: 2 Kinds of "Background"

- Question 1: Is there activity in the sample above blank? If "no," stop.
- Question 2: If the answer to Question 1 is "yes," how much of the activity is due to ambient levels and how much is due to a particular source (e.g., the workplace, release from a nuclear power plant)?

Problem Radionuclides: ^{137}Cs and ^{nat}U

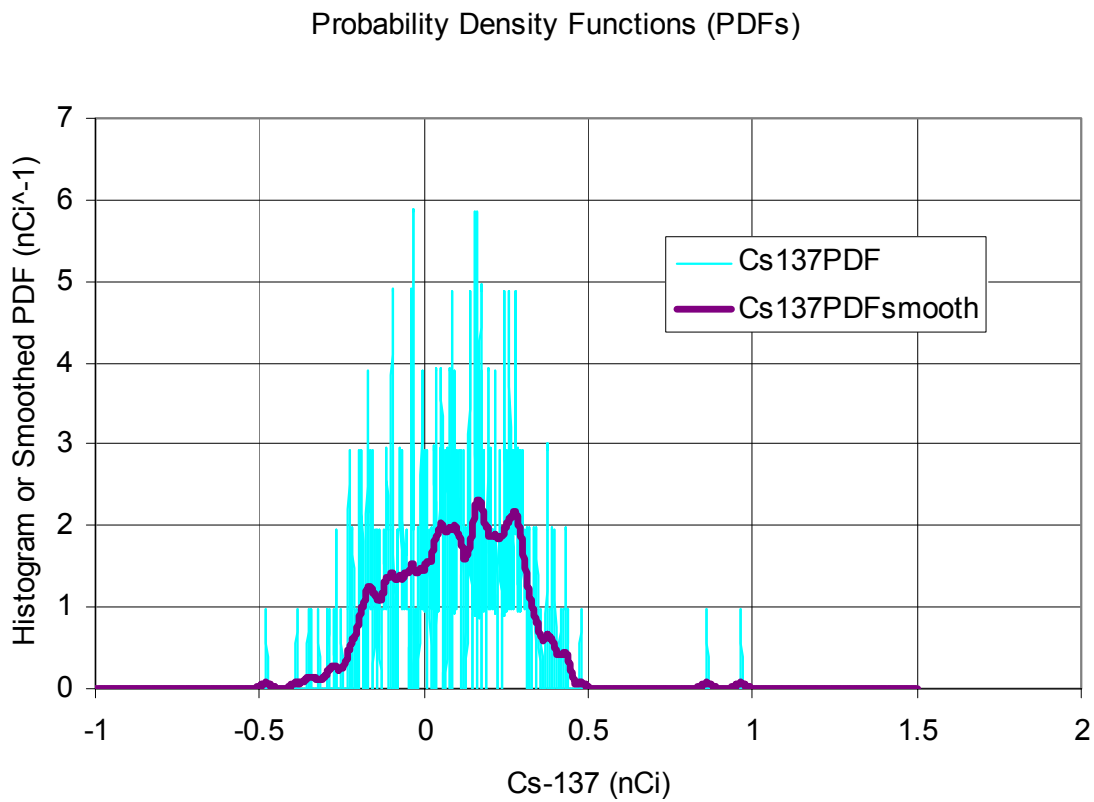
- Your job: analyze environmental or bioassay samples and determine what is "natural" and what is occupational or anthropogenic (human-made)
- ^{137}Cs is ubiquitous (fallout and orphan source meltings at steel mills) as is ^{nat}U
- Solution(s)
  - tricks like ^{236}U as a tracer for "recycled" U
  - baseline measurements (not possible after the fact)
  - analysis of nearby environment or similar workers
  - probabilistic background subtraction (makes the most of what you've got)

Example: ^{137}Cs in a Worker

- Intrinsic germanium (IGe) in vivo chest count
- ^{137}Cs is a tracer for Hanford tank waste, a mixture of fission products, traces of U and transuranics (TRU)
- ^{137}Cs produces much less dose than TRU
- *Question: Is the ^{137}Cs from fallout or the workplace?*
- Answer: make a probabilistic statement about the measurement

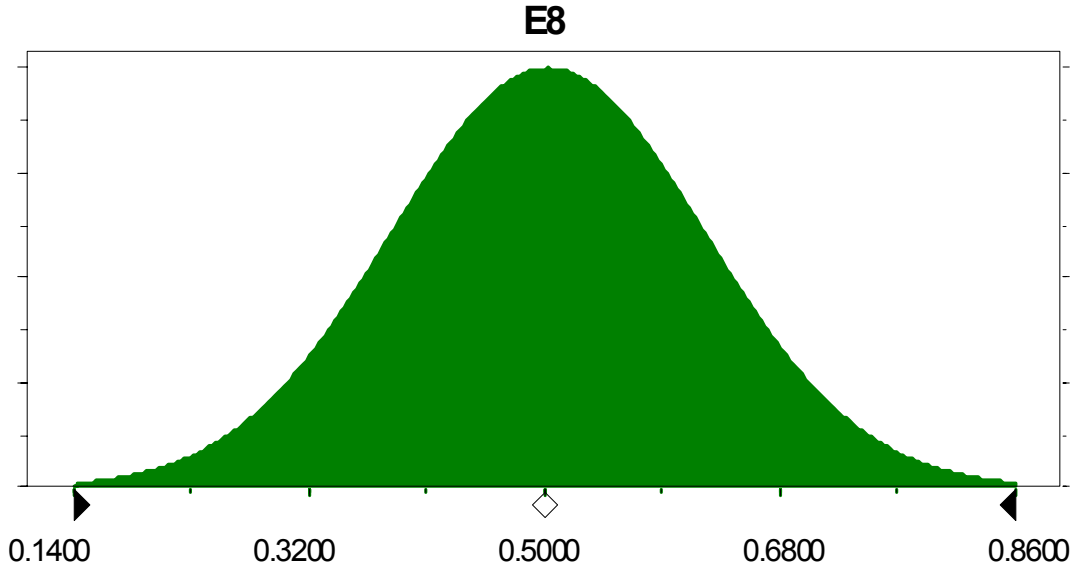
Whole-Body Counts of 409 Unexposed Workers

- Average 3.46 Bq (0.094 nCi)
- Chomentowski & Kellerer (2000) smoothing shown
- For a given worker result with its stated uncertainty, one can subtract "background" 409 different times (not shown)

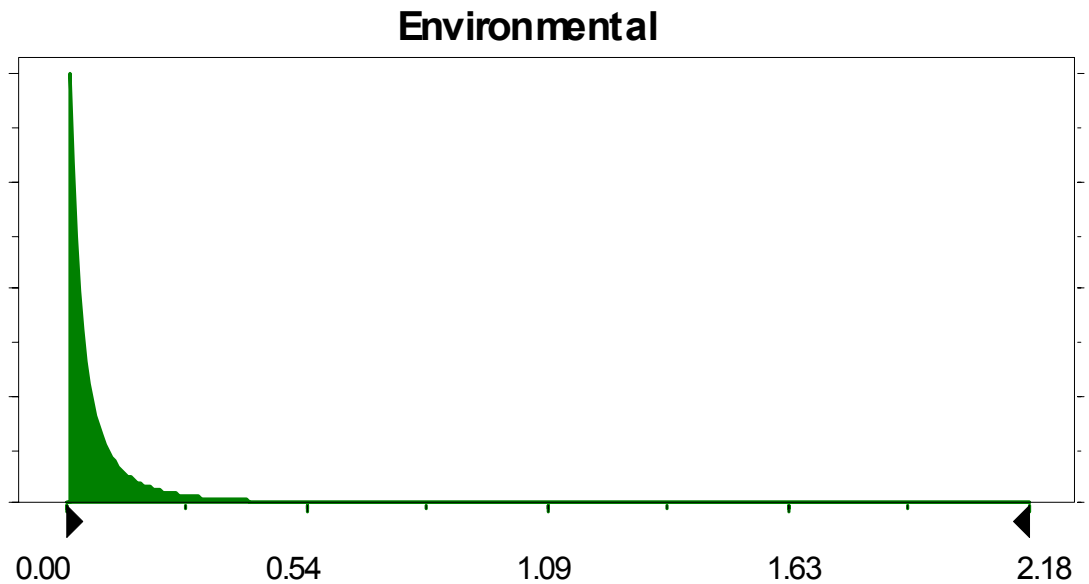


2 Distributions: Worker and Environmental

- Worker: Normal distribution, 0.50 ± 0.12 dpm/d

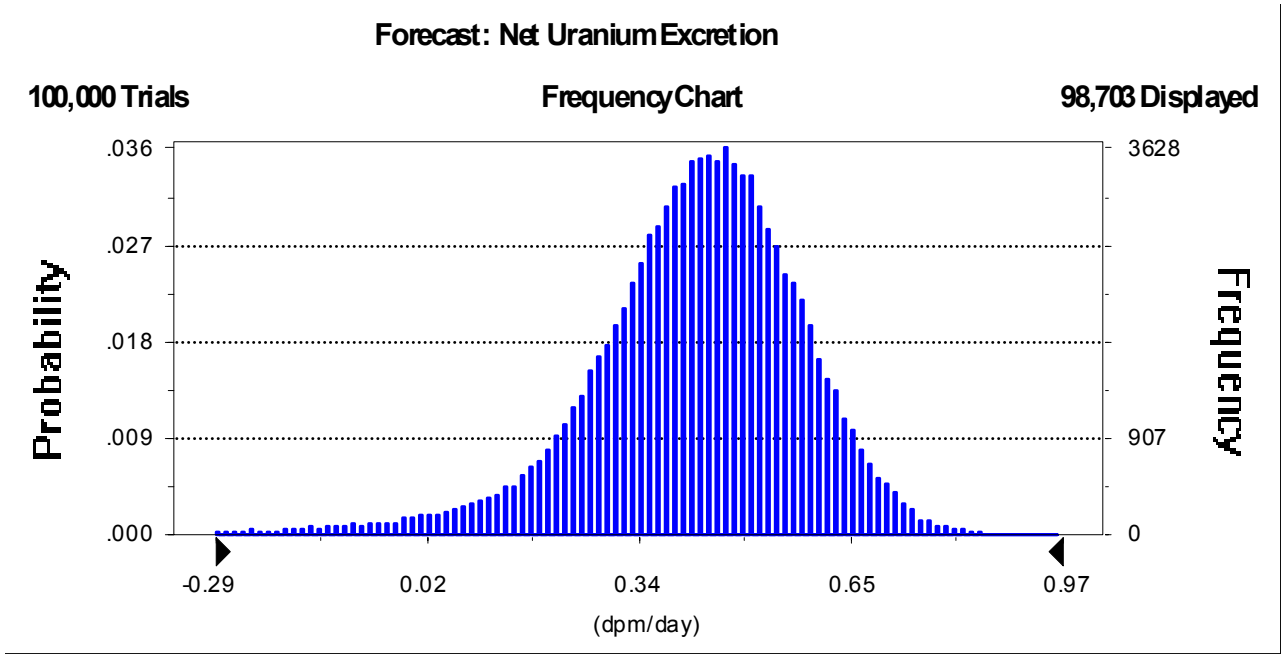


- Environmental: lognormal dist., geo. mean = 0.037 dpm/d, GSD = 3.9



Probabilistic Subtraction of Environmental U from Single Worker Bioassay

- mean 0.407 ± 0.241 dpm/d
- median = 0.436 dpm/d
- 95th %ile = 0.655 dpm/d
- 99th %ile = 0.695 dpm/d



The Future: A Personal Forecast

- Distributions are the way of the future
- In USA, probability of causation is being tied to 95th %ile; this will spread to other fields
- Regulatory bodies, competent authorities, will begin to require compliance to upper percentiles, not means or medians, as "protective" of public, workers, and the environment
- Probabilistic methods will become the norm

Non-Bayesian Methods for Evaluating Uncertainty in Complex Models

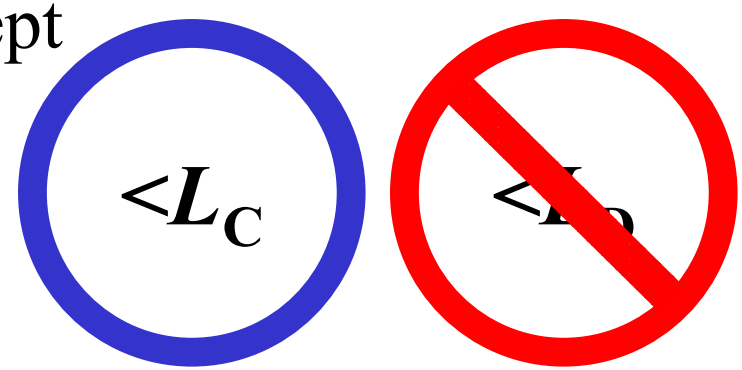
Non-Bayesian Methods for Evaluating Uncertainty in Complex Models

- A variety of methods are available for evaluating Type B uncertainty (e.g., model suitability)
- Biokinetic models have been prototyped, but environmental models are similar
- Methods include
  1. A "resampling" method
     - doesn't mean "obtaining another sample," but rather, it means "sampling subsets of all data"
  2. Updating: Recalculating every time you get another data point
  3. Use of multiple radionuclides one at a time

Conclusions and Recommendations

Conclusions 1

- There are severe problem of terminology
  - many names for the same concept
  - concepts mis-named
  - concepts mis-used
  - compare measurements with decision threshold L_C, not detection level L_D



- There are 2 "counting problems"
  - Forward problem done well by classical statistics
  - Reverse problem only done by Bayesian methods
  - Bayesian result: expectation value is N+1, so uncertainty is greater

Conclusions 2

- There are at least 8 decision rules
  - Most use a "Great Leap of Inference:" N is a good estimate of ρt & Var(ρt)
  - False positive rates are not as claimed for small numbers of counts
- ANSI N13.30/MARSSIM/Currie/etc. DL formula
  - yields too many false positives
  - fairly good only if counts > 100

Conclusions 3

- There are much better approximate & exact solutions
  - $N + 1$ rule is better, but not correct
  - Altshuler & Pasternak/Turner better (??? stay tuned for MARLAP)
  - Detailed Bayesian method works if a flat prior is acceptable
  - Nicholson's (1963) and Sumerling's & Darby's (1981)exact (classical) solution can be implemented - probably the soundest theoretical foundation, but very few false positives
  - McCroan/ISO 11929 ok for equal count times
  - Stapleton's rule performs the best
  - Rigaud (2003) not yet evaluated
  - **No rule works well at very low counts!**

Conclusions 4

- MARLAP is eagerly awaited
- More work to be done on software – this shouldn't be so hard!
- Consensus needed
- Bayes's methods need to be mainstreamed

Conclusions 5

- Decision strategies (Bayesian and classical)
  - You can re-count a sample and decide again
  - You can count the other half of a split or re-sample
  - Combinations of these lead to essentially no false positives
- An enumeration of the consequences of wrong decisions shows that we need the best decision rules we can get
- Detection capabilities: a decision on decision rules needed first
- Report and record uncensored, un-rounded

  positive, zero, and negative measurements

  results with uncertainties

  to avoid falsely exaggerating levels

Conclusions 6

- There is great utility in **averaging results** of many samples when appropriate
  - If a regulatory guideline applies to a year, average samples over a year
- **Probabilistic blank and environmental background subtraction** maximizes information content while distinguishing human-made signal from environmental background
- There exist several **non-Bayesian methods of uncertainty evaluation** in modeling

References

- Visit http://bidug.pnl.gov/ (DOE Bioassay and Internal Dosimetry Users Group)
- Visit http://www.pnl.gov/bayesian for many Bayesian links, including Los Alamos National Laboratory (LANL) site and its links

Acknowledgement

- Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle under Contract DE-AC06-76RLO 1830

- Contact: strom@pnl.gov

The opinions expressed in this presentation are those of Daniel J. Strom may or may not represent those of Battelle, the Pacific Northwest National Laboratory, or the U.S. Department of Energy

References

- Altshuler B and BS Pasternack. 1963. "Statistical Measures of the Lower Limit of Detection of a Radioactivity Counter." Health Physics 9:293-298.
- Berry DA. 1996. *Statistics: A Bayesian Perspective*. Wadsworth Publishing Company, Belmont, California.
- Brodsky A. 1992. "Exact Calculation of Probabilities of False Positives and False Negatives for Low Background Counting." Health Physics 63(2):198-204.
- Currie LA. 1968. "Limits for Qualitative Detection and Quantitative Determination. Application to Radiochemistry." Analytical Chemistry 40(3):586-593.
- d'Agostini G. 1999. *Bayesian reasoning in high energy physics. Principles and applications*. CERN Yellow Report 99-03. http://www-zeus.roma1.infn.it/~agostini/index.html, Dip. di Fisica Università "La Sapienza" and INFN, Rome.
- Gilbert RO. 1987. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.
- Hickey EE, GA Stoetzel, DJ Strom, GR Cicotte, CM Wiblin, and SA McGuire. 1993. *Air Sampling in the Workplace*. Final Report. NUREG-1400; PNL-7814, U.S. Nuclear Regulatory Commission, Washington, DC.

References

- Howson C and P Urbach. 1991. "Bayesian Reasoning in Science." Nature 350:371-374.
- International Organization for Standardization (ISO). 1995. *Guide to the Expression of Uncertainty in Measurement. Corrected and Reprinted, 1995*. International Organization for Standardization, Geneva.
- Jaynes ET. 2003. *Probability Theory : The Logic of Science*. Cambridge University Press, New York.
- Kathren RL. 2001. "Everything But the Counting Statistics: Practical Considerations in Instrumentation and Its Selection and Use." Chapter 17 in *Radiation Instruments*, ed. H Cember, pp. 361-384. Medical Physics Publishing, Madison, Wisconsin.
- Little RJA. 1982a. "The Statistical Analysis of Low-Level Radioactivity in the Presence of Background Counts." Health Physics 43(5):693-703.
- Little RJA. 1982b. "Hypothesis Tests Based on Counts from Blank and Radioactive Samples." Health Physics 42(2):230-231.
- Nicholson WL. 1963. *Fixed Time Estimation of Counting Rates with Background Corrections*. AEC Research and Development Report HW-76279, Hanford Works, Richland, Washington.

References

- Ott WR. 1995. *Environmental Statistics and Data Analysis*. CRC Press LLC, Boca Raton, Florida.
- Palmer HE. 1984. "In-Vivo Counting of Uranium." in *Biokinetics and Analysis of Uranium in Man*. Proceedings of a Colloquium held at Richland, Washington, August 8-9, 1994. USUR-05 HEHF-47, ed. RH Moore, p. I-1-I-29. Hanford Environmental Health Foundation, Richland, Washington.
- Strom DJ and JA MacLellan. 2001. "Evaluation of Eight Decision Rules for Low-Level Radioactivity Counting. PNNL-SA-32868." Health Physics 81(1):27-34.
- Taylor BN and CE Kuyatt. 1994. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST Technical Note 1297 http://physics.nist.gov/cuu/pdf/tn1297.pdf, National Institute of Standards and Technology, Physics Laboratory, Gaithersburg, MD 20899.
- Thomas J. 1963. *Interpretation of Low-Activity Counting*. Risø Report No. 70, Danish Atomic Energy Commission, Risø, Denmark.
- Turner JE. 1995. *Atoms, Radiation, and Radiation Protection*. John Wiley & Sons, New York.

References

- U.S. Environmental Protection Agency (EPA), U.S. Nuclear Regulatory Commission (NRC), U.S. Department of Energy (DOE), and U.S. Department of Defense. 2003. *Multi-Agency Radiological Laboratory Analytical Protocols Manual. Draft*. http://www.epa.gov/radiation/marlap/about_marlap.htm, U.S. Environmental Protection Agency, Washington, DC.
- U.S. Nuclear Regulatory Commission (NRC), U.S. Environmental Protection Agency (EPA), U.S. Department of Energy (DOE), and U.S. Department of Defense. 2000. *MultiAgency Radiation Survey and Site Investigation Manual (MARSSIM; http://www.epa.gov/radiation/marssim/)*. NUREG-1575, U.S. Government Printing Office, Washington, DC.