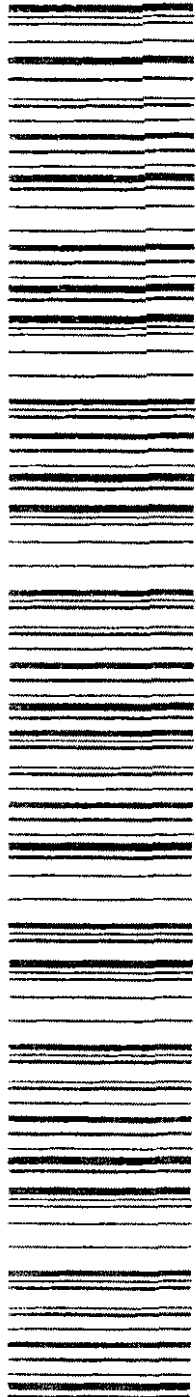


Leonard Shabman

Eric P. Smith

Virginia Tech

October 16, 2000



12703



# The Management Questions



- ◆ Given small sample sizes for monitoring data (8 to - *maybe* – 60 observations)
- ◆ Does a water segment violate water quality standards?
- ◆ Has a TMDL implementation plan achieved water quality standards?

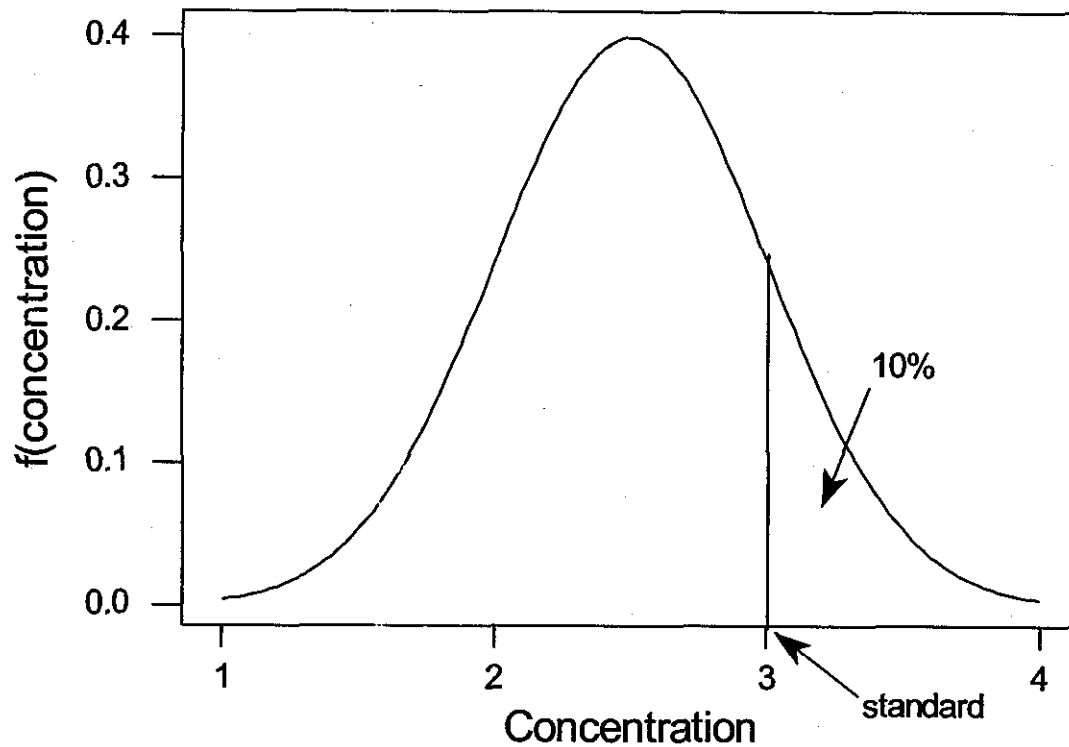


# Sound Statistical Science

---

- ◆ EPA guidelines on interpretation of monitoring data is “*naive*” statistics
- ◆ Sound statistical procedures
  - Recognize the possibility of making data interpretation errors
  - Allows for explicit control over the kind of error that might be made

**The Question:** Does the water meet a water quality standard 90% of time?





---

# Data for the Statistical Analysis

---





- ◆ Data collected a different times

Occasion	Measurement	Gt standard	Score
1	$x_1$	no	0
2	$x_2$	no	0
3	$x_3$	yes	1
4	$x_4$	no	0
5	$x_5$	no	0

# Error rates

- ◆ Type I error = false positive
  - Too many measurements exceed the standard although the site is not impaired
  
- ◆ Type II error = false negative
  - Many measurements less than standard although the site is impaired


The possibility of error is *always* present and *always* addressed – either explicitly or implicitly



---

# Naïve statistical analysis: The EPA raw score approach

---



*Statistic* = number of measurements > standard /  
number of measurements

*Test:* List as impaired if *statistic*  $\geq$  10%

10% of samples does **NOT** represent 10% of the time



---

# Sound Statistics: The Binomial

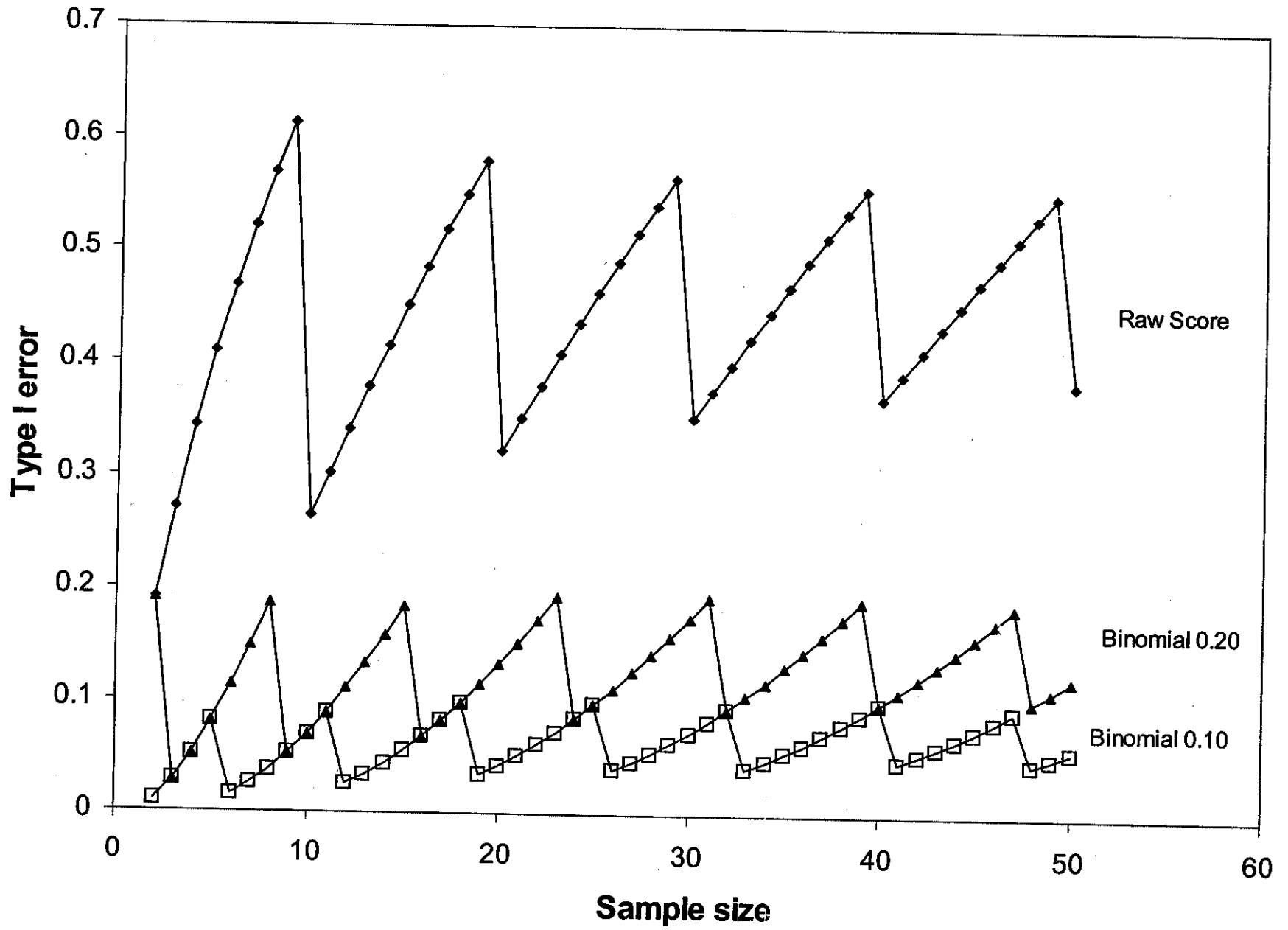
---



- ◆ Statistic: number of measurements above standard or proportion above standard
- ◆ Test: reject the hypothesis that the water is not impaired if a “large” number exceed the standard
  - “Large” is determined by the Binomial distribution and error rates
  - Typical: select acceptable Type I error and increase sample size to limit Type II error



12711





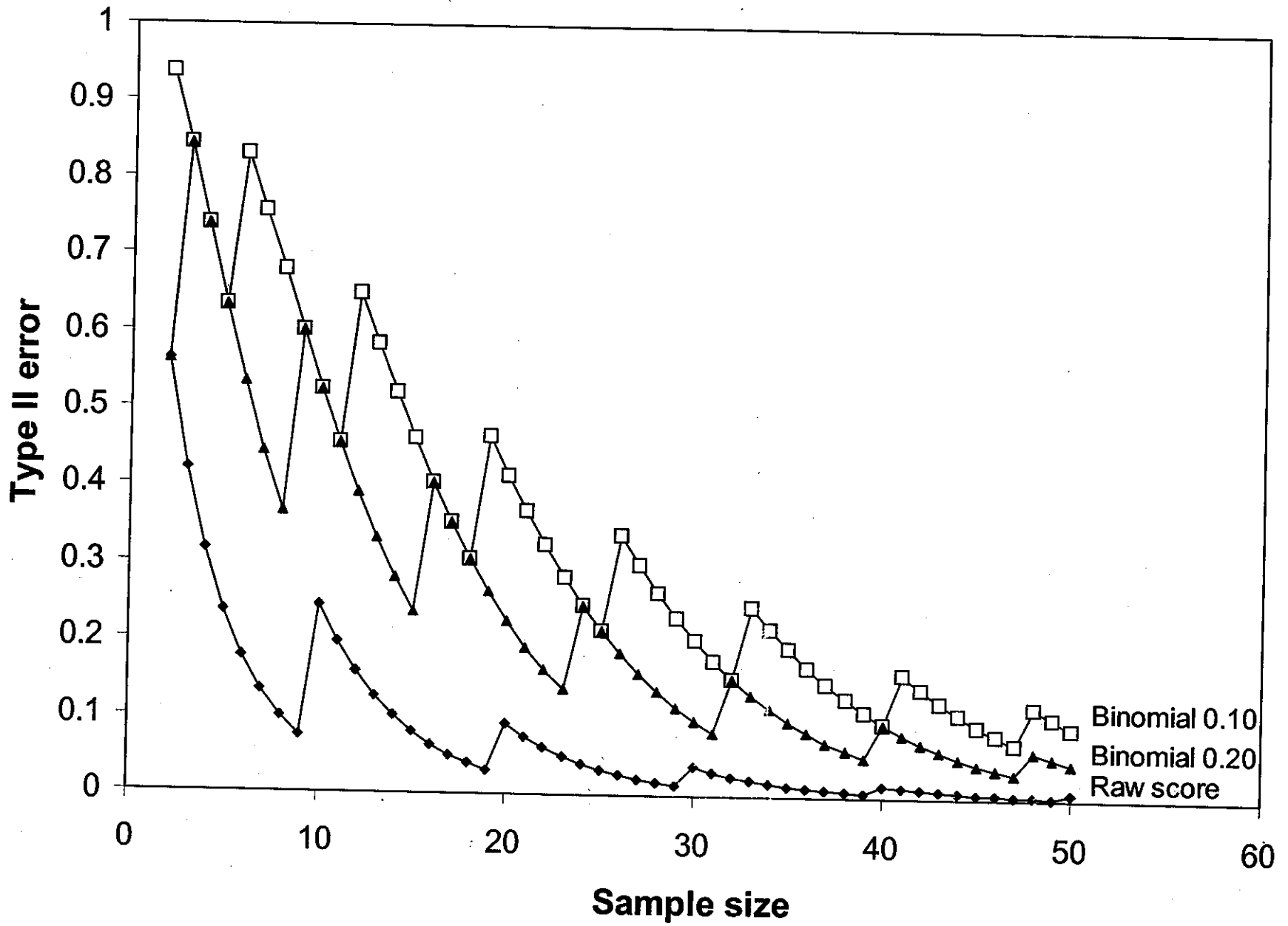
---

# Implication

---



- ◆ Raw score method is prone to high Type I error
- ◆ Binomial method controls Type I error rate by bounding it below a specified value
  - Raw score is similar to a Binomial with Type I error rate (alpha level) of 0.55





# Implication



- ◆ Type II error decreases with increased sample size for *all* methods
  - As Type II error decreases with raw score Type I does *not* decrease



# Managing Errors: Be Transparent

- choose cutoffs in consideration of sample sizes in specific watersheds
- choose cutoffs in consideration of use of the waters in specific watersheds



# Enhanced statistical approaches



- ◆ Bayesian Binomial
  - Makes use of prior information about site
- ◆ Acceptance sampling by variables
  - Based on means and uses information about the degree that standards are exceeded
  - Other methods are available



# Recommendations

- ◆ Adopt binomial as sound science for listing and de-listing
- ◆ Expand sample sizes by using data from all time periods (adjust for trend)
- ◆ Treat binominal as a modest step beyond raw score method; adopt more sophisticated methods to learn more from available data



# Sound Statistical Science



- ◆ EPA guidelines on interpretation of monitoring data is “*naive*” statistical science
- ◆ Sound statistical procedures
  - Recognize the possibility of making data interpretation errors
  - Allows for explicit control over the kind of error that might be made