

Theory on Partitioning Reference Values into Subgroups



Ari Lahti
Rikshospitalet, Oslo

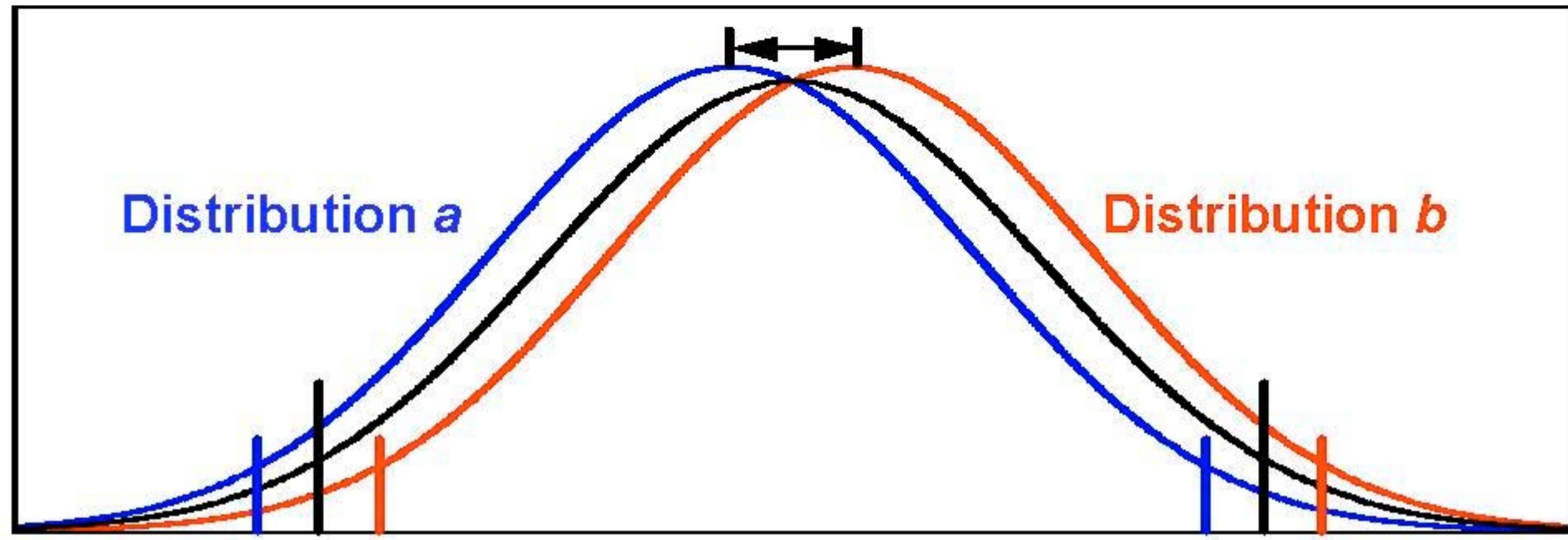
The Present Way of Partitioning or the Harris-Boyd Model

1. Distance criterion

$$z = \frac{|\mu_a - \mu_b|}{\sqrt{\frac{\sigma_a^2}{n_a} + \frac{\sigma_b^2}{n_b}}} > z_{\text{Crit}}$$

Distribution a

Distribution b



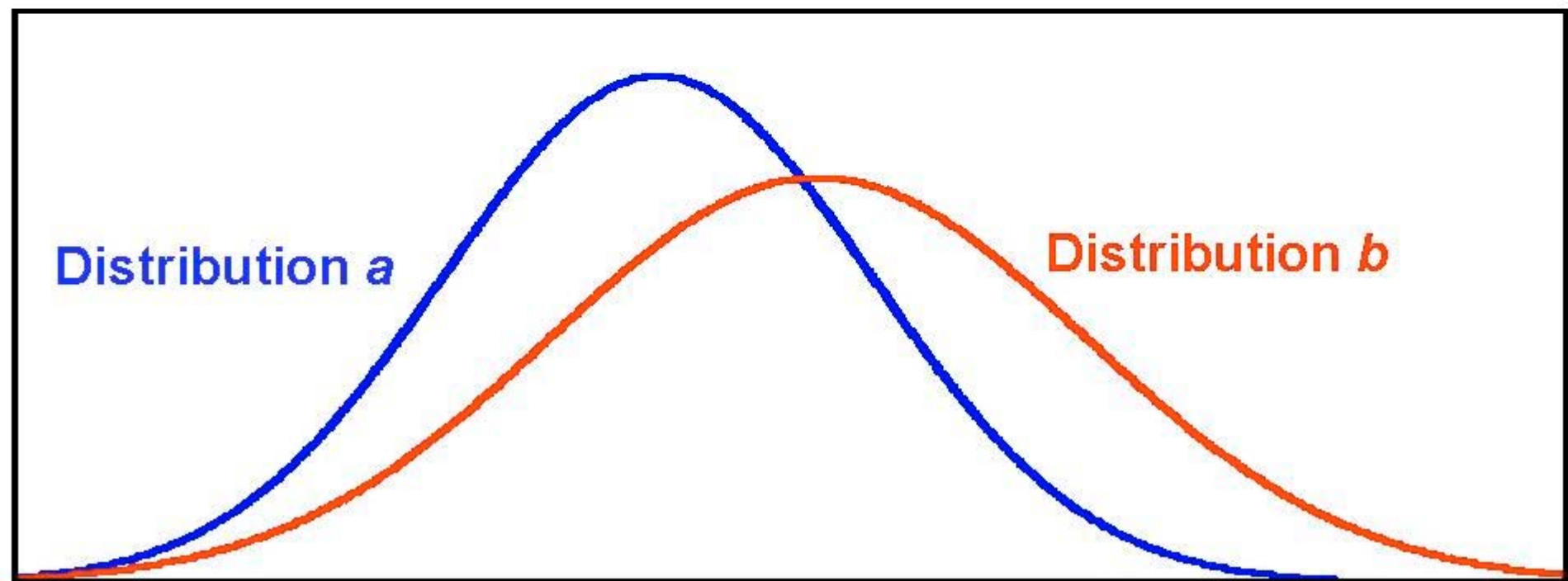
The Present Way of Partitioning or the Harris-Boyd Model

2. Standard
deviation
criterion

$$\frac{s_b}{s_a} > 1.5$$

Distribution a

Distribution b





Background of the Harris-Boyd Model

- **The model was developed using computer simulated gaussian distributions**
- **The aim was to correlate the test criteria to proportions of the subgroup distributions outside of the common reference limits**

Proportions Outside of the Common Reference Limits

1. Coinciding distributions

$$p_a = 2.5\%$$

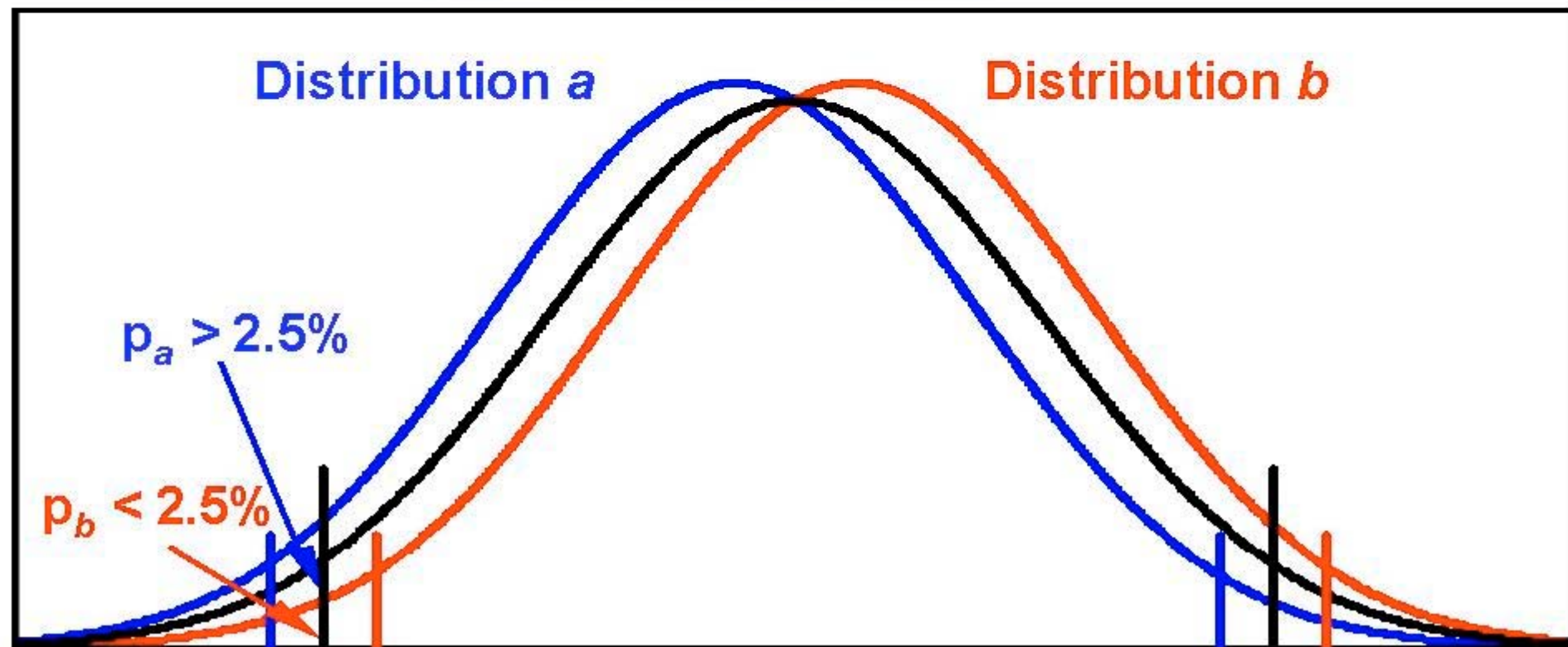
$$p_b = 2.5\%$$

Distribution *a*
Distribution *b*

A normal distribution curve is shown in purple. Two vertical purple lines are drawn on the curve, one on the left tail and one on the right tail. A purple arrow points from the text 'p_a = 2.5%' to the left tail line, and another purple arrow points from 'p_b = 2.5%' to the right tail line. The text 'Distribution a' is in blue and 'Distribution b' is in orange, both centered within the curve.

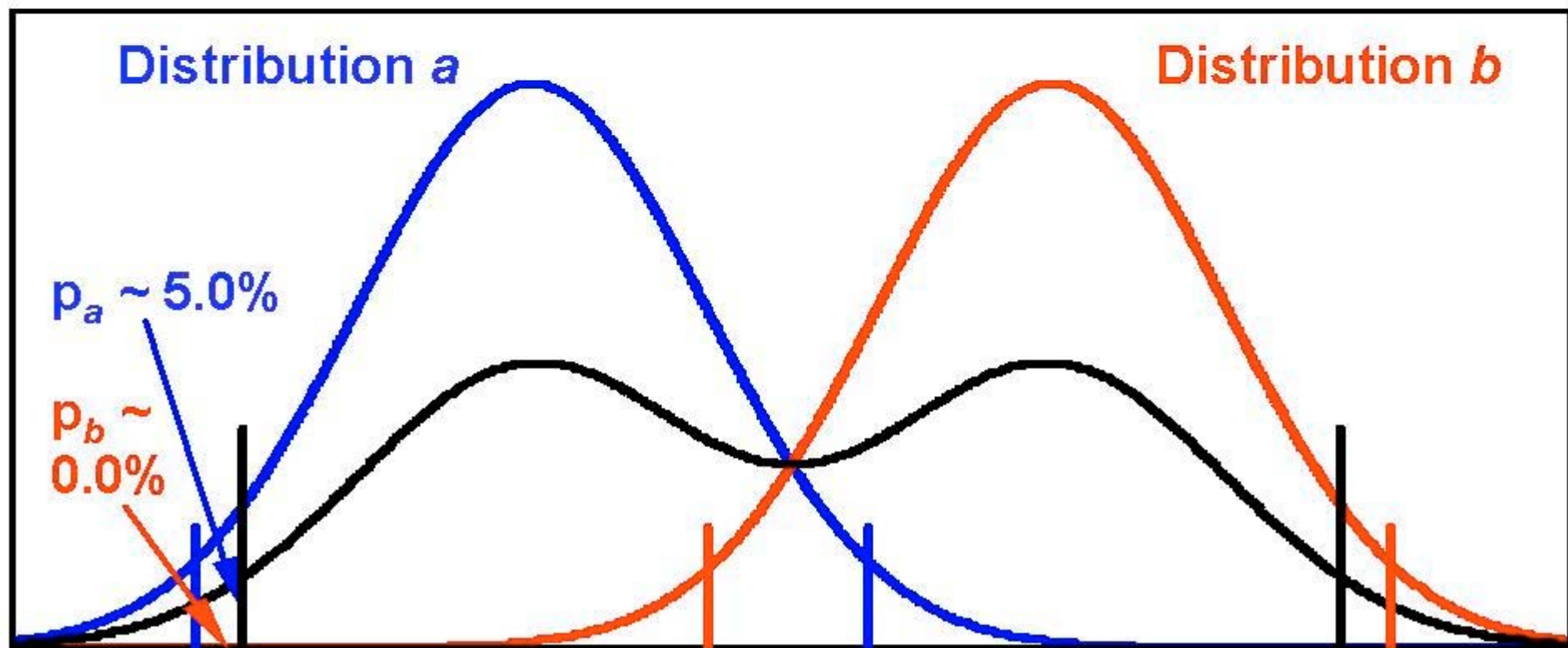
Proportions Outside of the Common Reference Limits

2. Distributions having different means



Proportions Outside of the Common Reference Limits

3. Distributions lying far from each other: extreme proportions



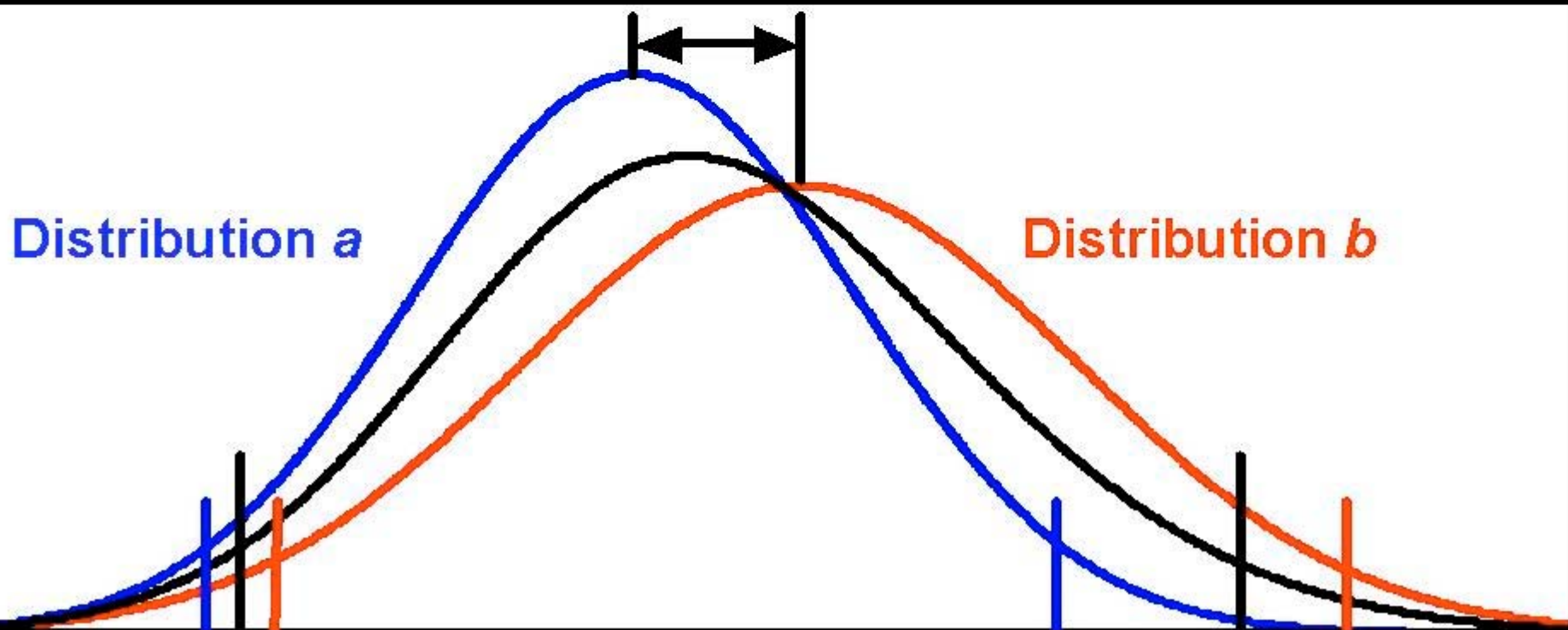


Proportion Criterion of Harris and Boyd

- **The larger proportion at both ends of the distributions lies in the interval [2.5%, 5.0%)**
- **If one of these larger proportions exceeds 4.0%, partitioning is recommended**

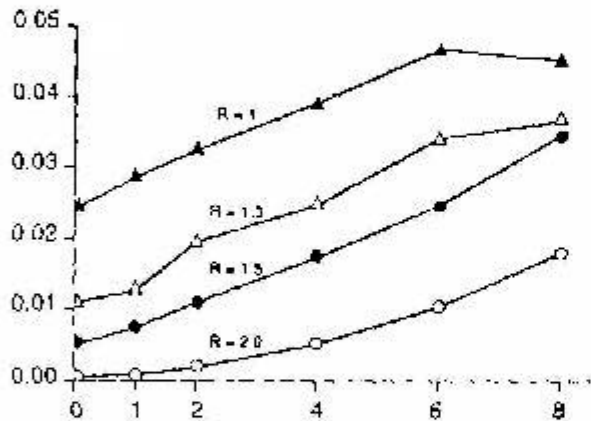
Harris-Boyd Model: Problem

Proportions may be different at
different ends of the distributions

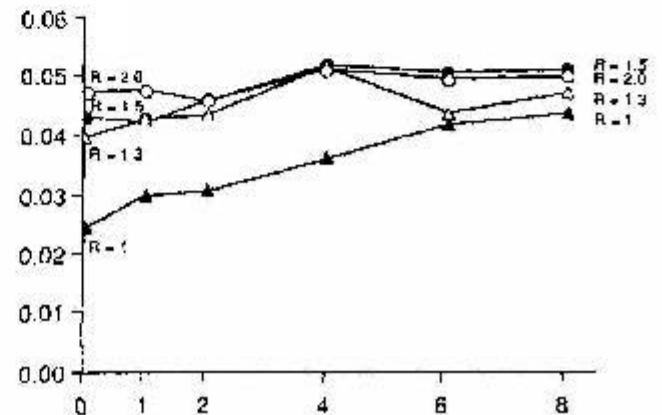
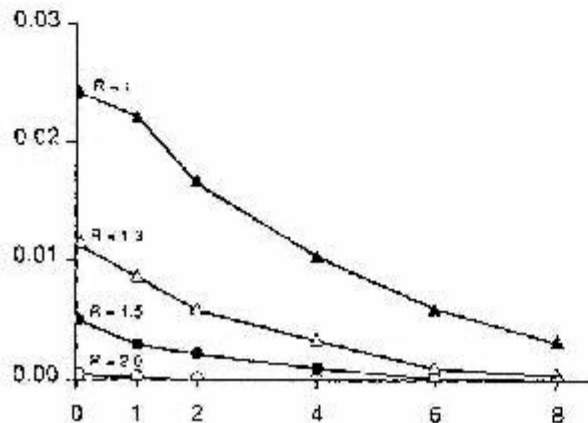
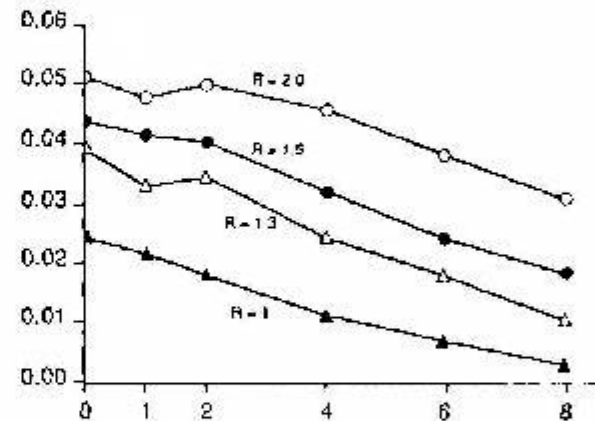


Harris-Boyd Model: Proportion vs. Distance Curves

Distribution a



Distribution b





Harris-Boyd Model: Summary

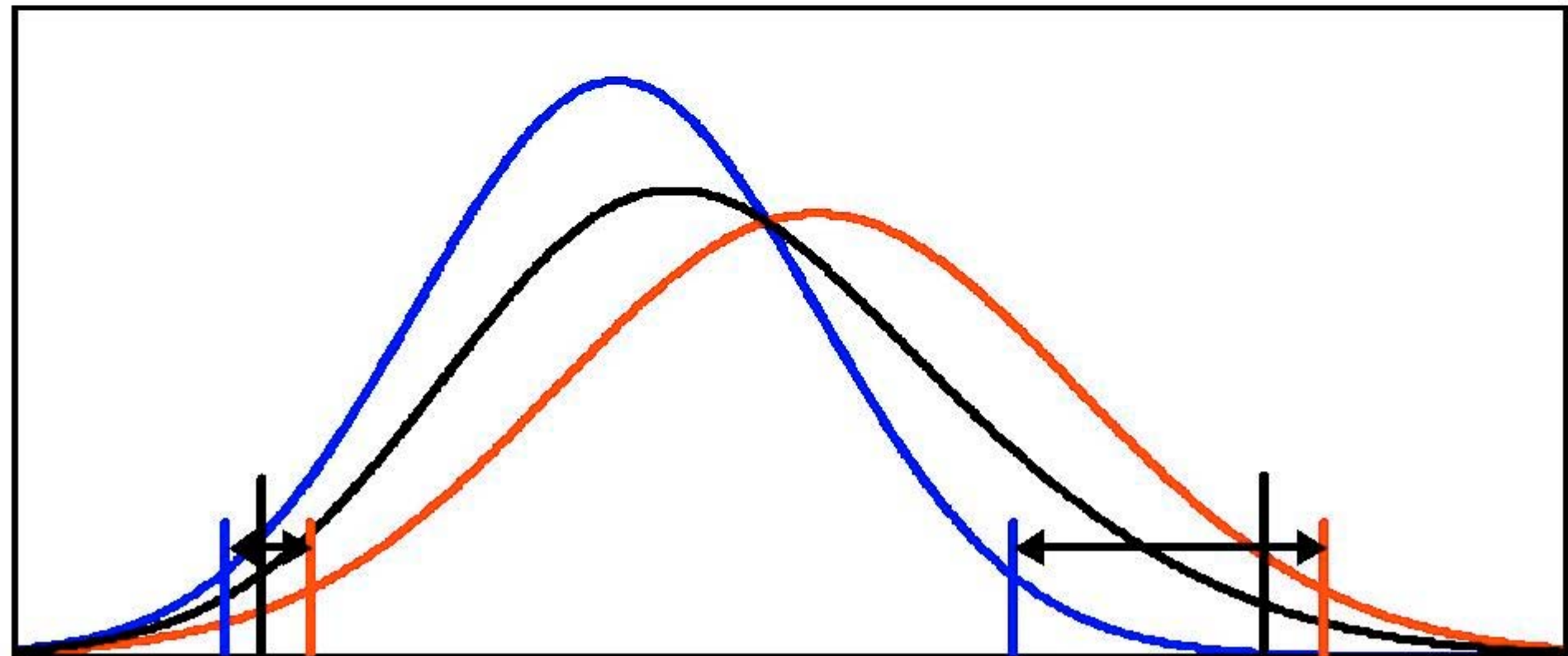
- **Both ends of the distributions cannot be controlled simultaneously**
- **The proportion vs. distance curves were constructed using computer simulations**

Correlation between distance and proportion criteria for partitioning remains poor



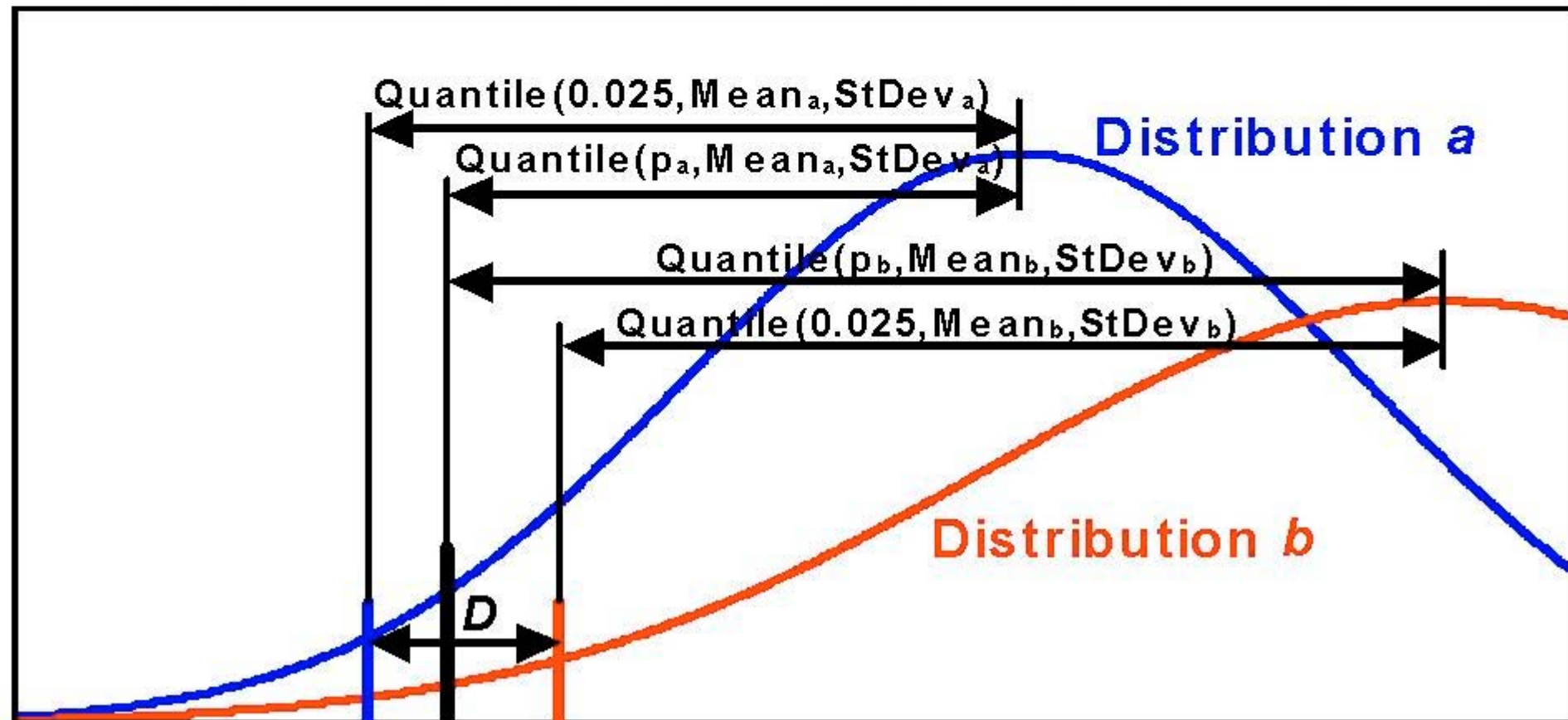
New Model: Principle

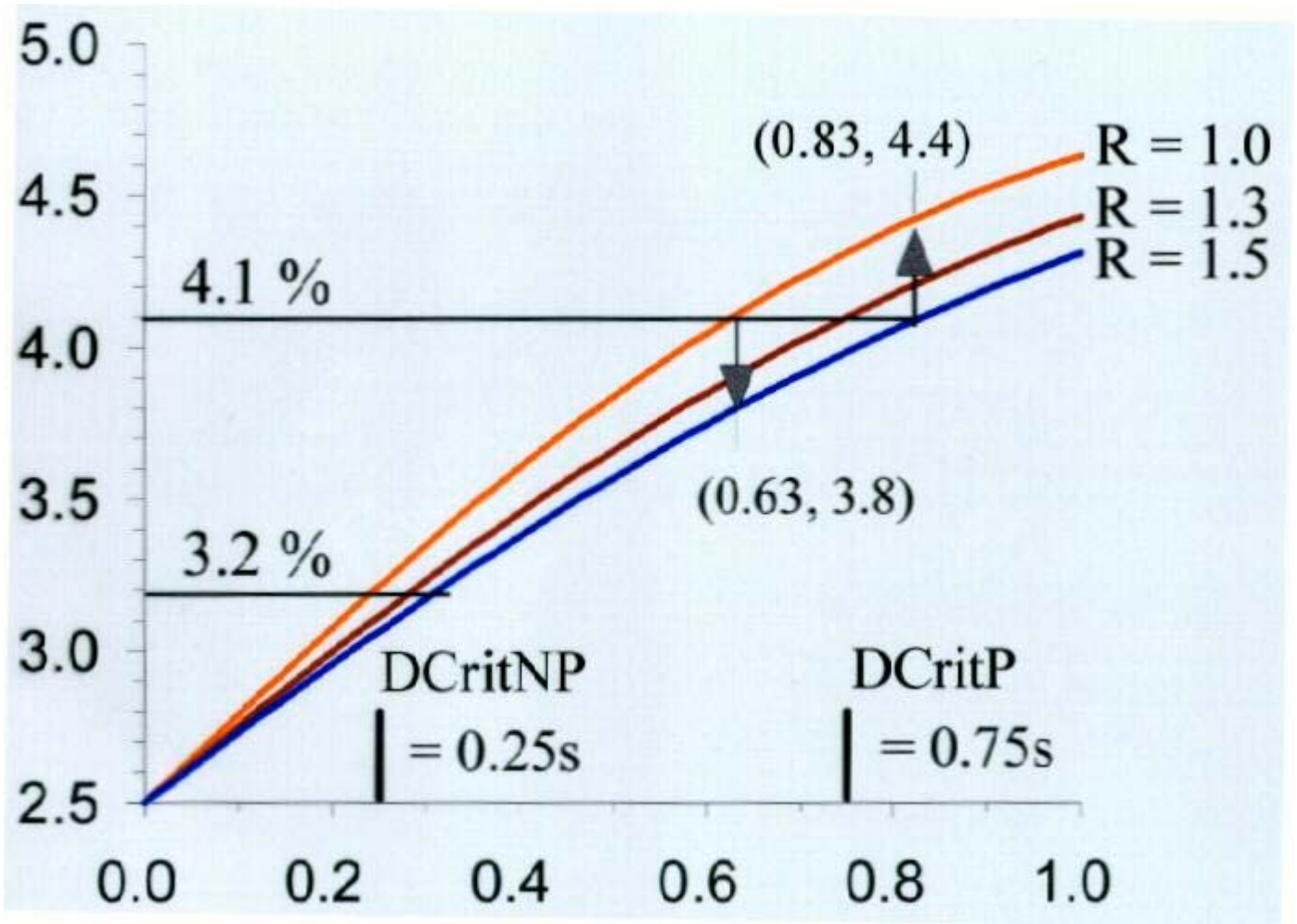
Focus on distances between reference limits



New Model: Correlation Between Proportions and Distances

$$D = |Q(0.025, a) - Q(p_a, a) + Q(0.05 - p_a, b) - Q(0.025, b)|$$







New Model: Summary

- **Each end of the distributions is treated separately**
- **The proportion vs. distance curves are determined by using mathematical calculations**

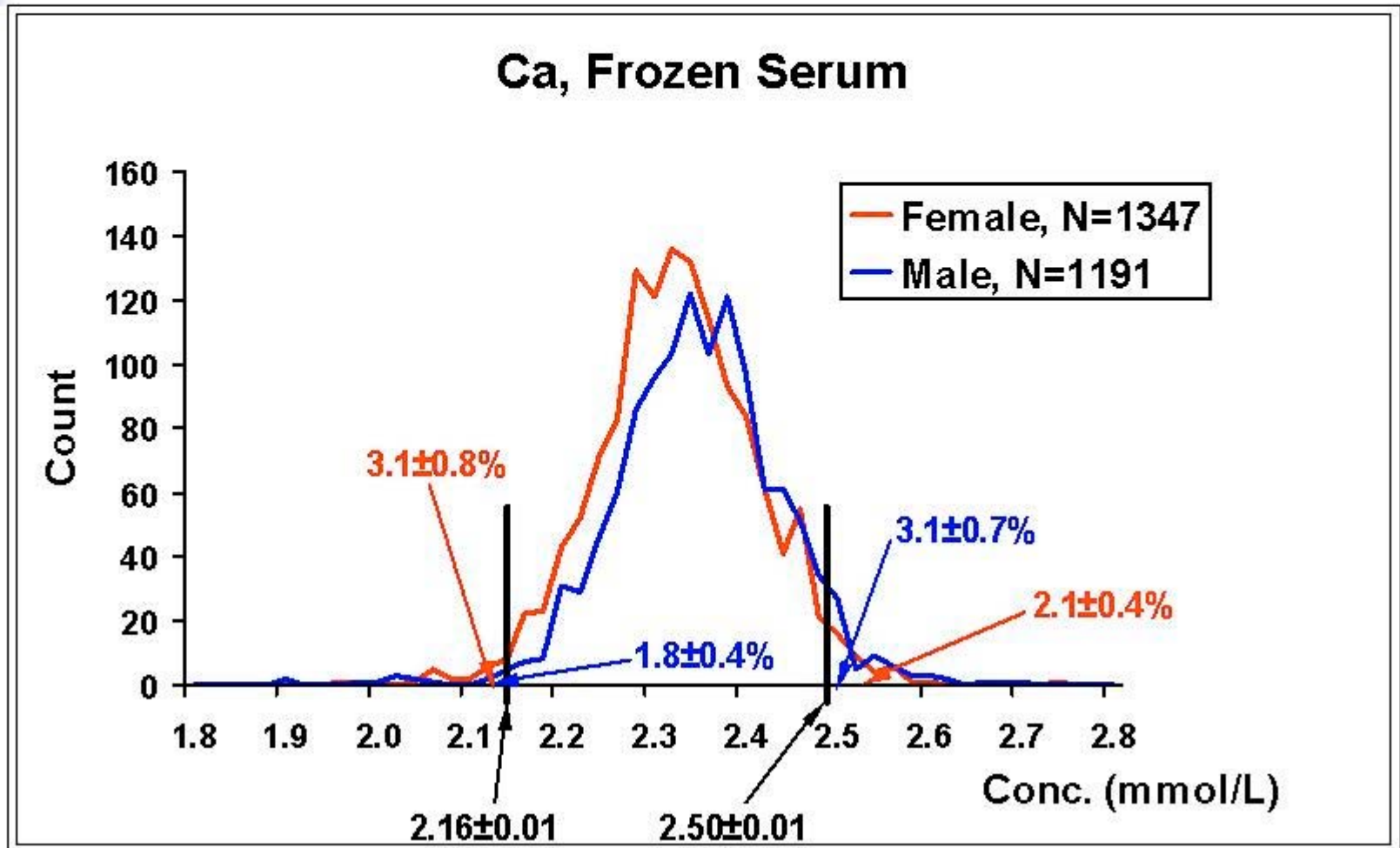
Correlation between distance and proportion criteria for partitioning is good



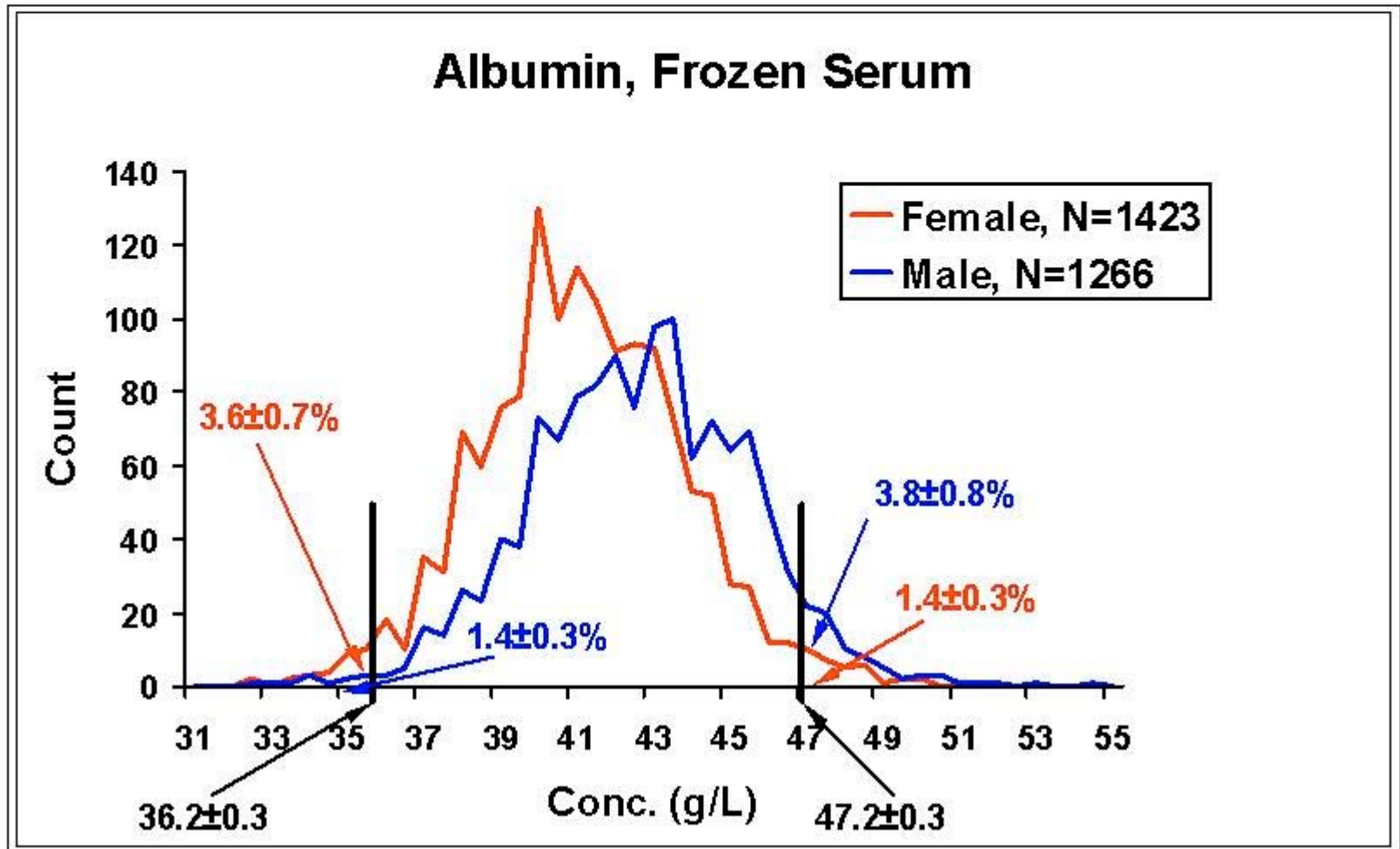
Suggestion for Proportion Criteria

- **If both of the larger proportions lie below 3.2%, combining is recommended**
- **If at least one of the larger proportions lies in the interval $[3.2\%, 4.1\%)$ and neither exceeds 4.1%, the case is a marginal one**
- **If at least one of the larger proportions exceeds 4.1%, partitioning is recommended**

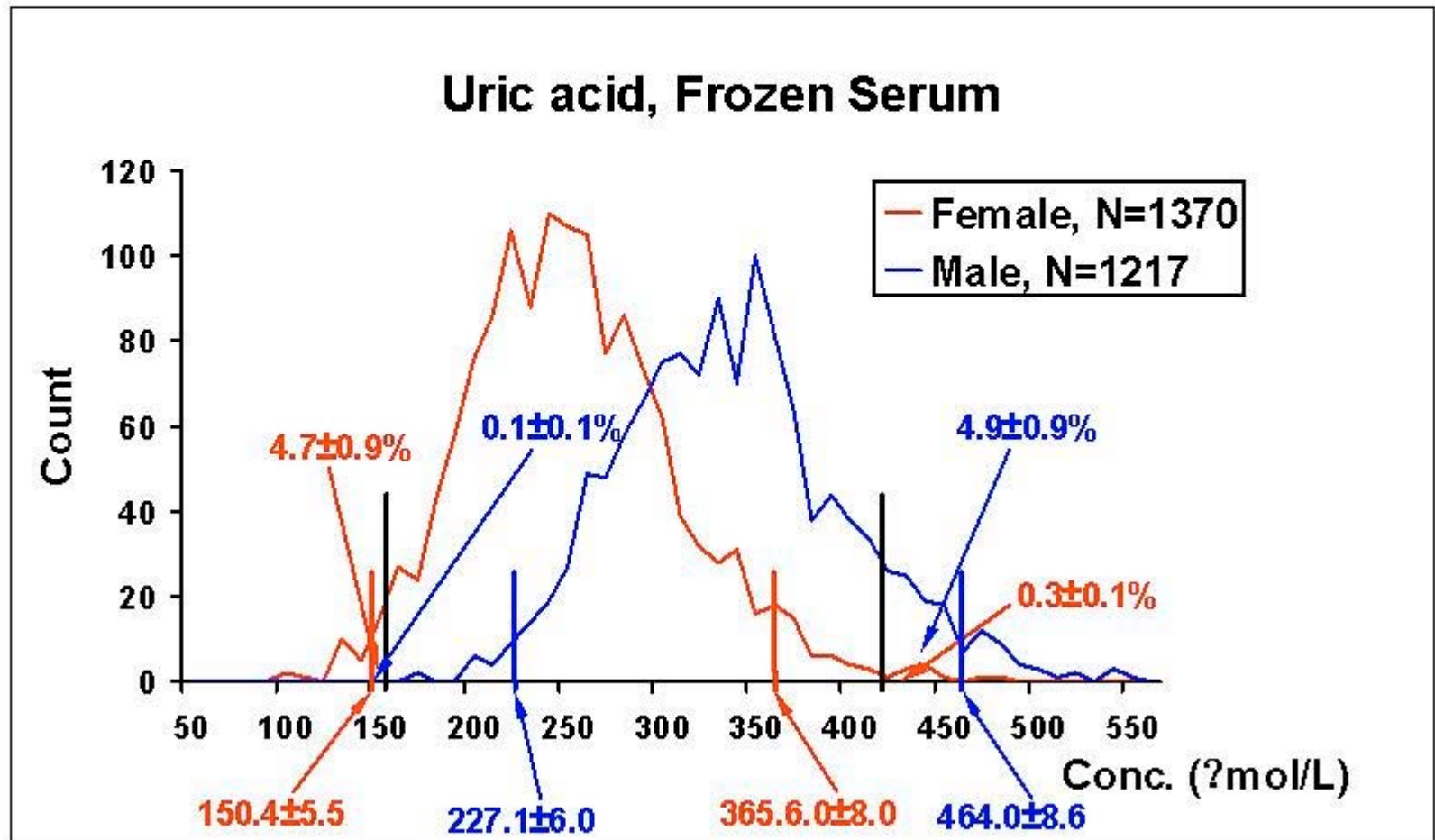
Example on Use of the Proportion Criteria: Combining Case



Example on Use of the Proportion Criteria: Marginal Case



Example on Use of the Proportion Criteria: Partitioning Case





Summary

- **The new model for distance criteria is**
 - **more accurate than the Harris-Boyd model**
 - **easily adjustable to any new values for the critical proportions**
- **Use of distance criteria should be restricted to gaussian distributions**
- **Proportion criteria are applicable to both gaussian and nongaussian distributions**



Clin Chem (2002) 48:2, 338-352

Clinical Chemistry 48:2
338-352 (2002)

Laboratory
Management

Objective Criteria for Partitioning Gaussian-distributed Reference Values into Subgroups

ARI LAHTI,^{1*} PER HYLTOFT PETERSEN,^{2,3} JAMES C. BOYD,⁴ CALLUM G. FRASER,⁵ and NILS JØRGENSEN⁶

Background: The aim of this study was to develop new and useful criteria for partitioning reference values into subgroups applicable to gaussian distributions and to distributions that can be transformed to gaussian distributions.

Methods: The proposed criteria relate to percentages of the subgroups outside each of the reference limits of the combined distribution. Critical values suggested as partitioning criteria for these percentages were derived from analytical bias quality specifications for using common reference intervals throughout a geographic area. As alternative partitioning criteria to the actual percentages, these were transformed mathematically to critical distances between the reference limits of the

between these two values should be classified as marginal, implying that nonstatistical considerations are required to make the final decision on partitioning. The correlation between the critical percentages and the critical distances was mathematically precise in the new model, whereas this correlation is rather approximate in the Harris-Boyd model because focus on the difference between means in this model makes high precision hard to achieve. The application examples suggested that the new model is more radical than the Harris-Boyd model. **Conclusions:** New percentage and distance criteria, to be used for partitioning gaussian-distributed data, have been developed. The distance criteria, applied separately to both reference limit pairs of the subgroup