

PHUSE EU Connect 2021

*Let's ROC: An empirical
approach to confidence
intervals of Receiver-Operator
Characteristic curves.*

Alexander Ullmann



YEAR

Area

// The Setting

// ROC

// Sensitivity

// Specificity

// Characteristics of ROC

// Algorithms

// Averaging of ROC curves

// d-width confidence bands

// Outlook

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

See

- // Outcome: Binary classification (L) derived from a classification test (a diagnostic test)
- // Classification derived by thresholding a continuous predictor
- // Example: Assume we want to predict heart disease based on a biomarker e.g. NT-proBNP, d-dimer, etc

// Heart disease based a cutoff value of the biomarker.

// Higher values of biomarker indicative of disease

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

See

- // Outcome: Binary classification (L) derived from a classification or a diagnostic test
- // Classification derived by a continuous predictor
- // Example:

// We have a new diagnostic tool that measures this biomarker

// How do we assess the performance of a new diagnostic test somehow in detecting the disease?

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

Se

// Classification is based on threshold or score value taken as cut-off,

no disease

disease

marker scale

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

Assume we know ground truth: sample of 6 (3 disease, 3 no disease)



Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

// Sensitivity and Specificity

	Reference standard	Reference standard
Index test	True positives (TP)	False positives (FP)
Index test	False negatives (FN)	True negatives (TN)
	Sensitivity $TP/(TP+FN)$	Specificity $TN/(FP+TN)$

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

Se

Assume we know ground truth: sample of 6 (3 disease, 3 no disease)



$$\text{Sensitivity} = \frac{1}{3} (33.3\%), \text{Specificity} = \frac{1}{3} (33.3\%)$$

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

Se

Assume we know ground truth: sample of 6 (3 disease, 3 no disease)



$$\text{Sensitivity} = \frac{0}{3} (0.0\%), \text{Specificity} = \frac{3}{3} (100.0\%)$$

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

Se

Assume we know ground truth: sample of 6 (3 disease, 3 no disease)



$$\text{Sensitivity} = \frac{3}{3} (100.0\%), \text{Specificity} = \frac{0}{3} (0.0\%)$$

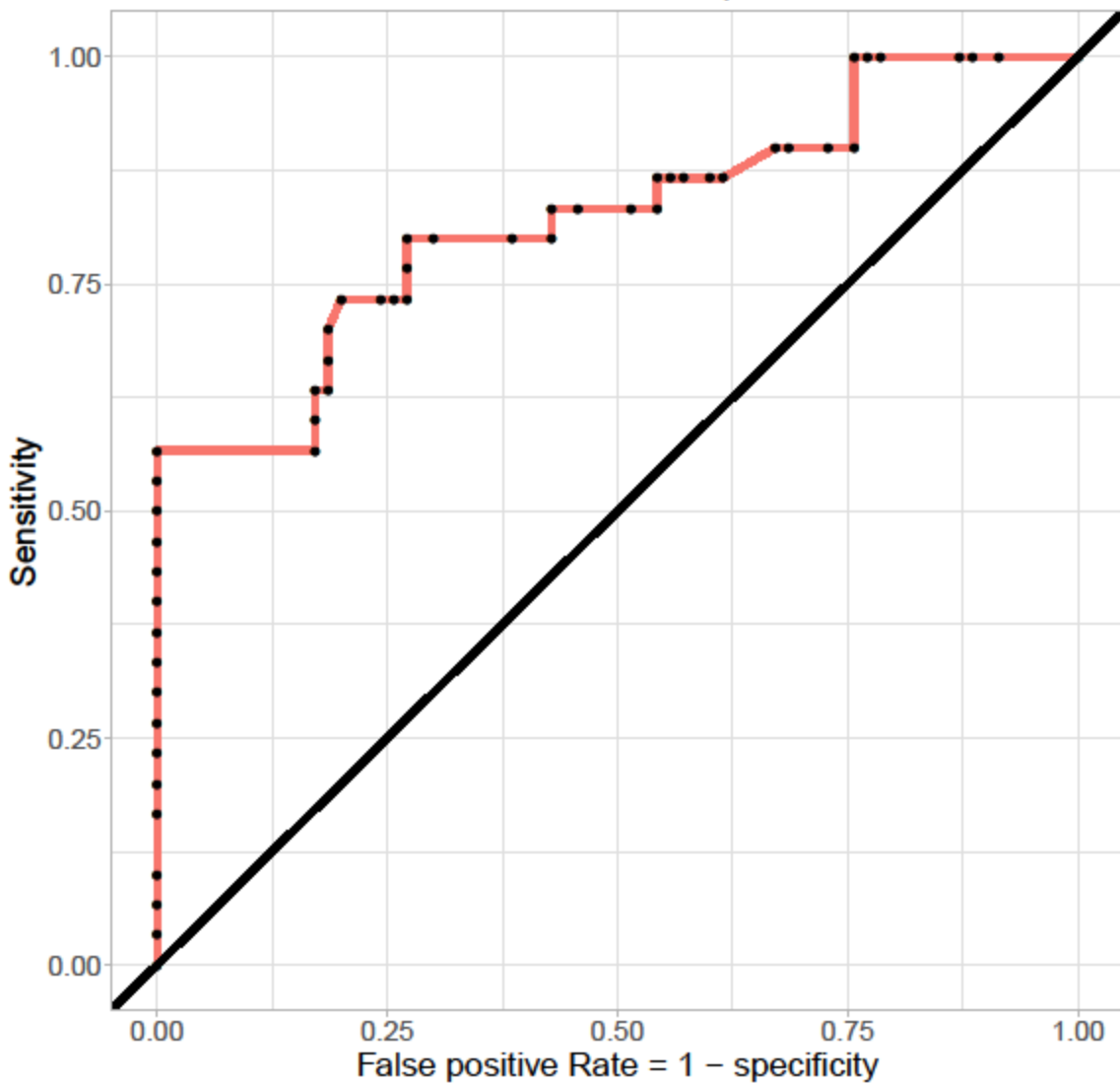
Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

- // The trade-off between sensitivity and specificity depending on the threshold
- // What to do? Well, we can "average" the performance of a diagnostic test using the ROC analysis

Let's ROC: An empirical approach to confidence intervals of Receiver-Operator Characteristic curves.

- // Trade-off between sensitivity and specificity depending on the threshold
- // What to do? Well, we can "average" the performance of a diagnostic test using the ROC analysis
- // Swipe through a range of cuton-points and plot 1-specificity vs. sensitivity
- // False-positive rate vs true positive rate

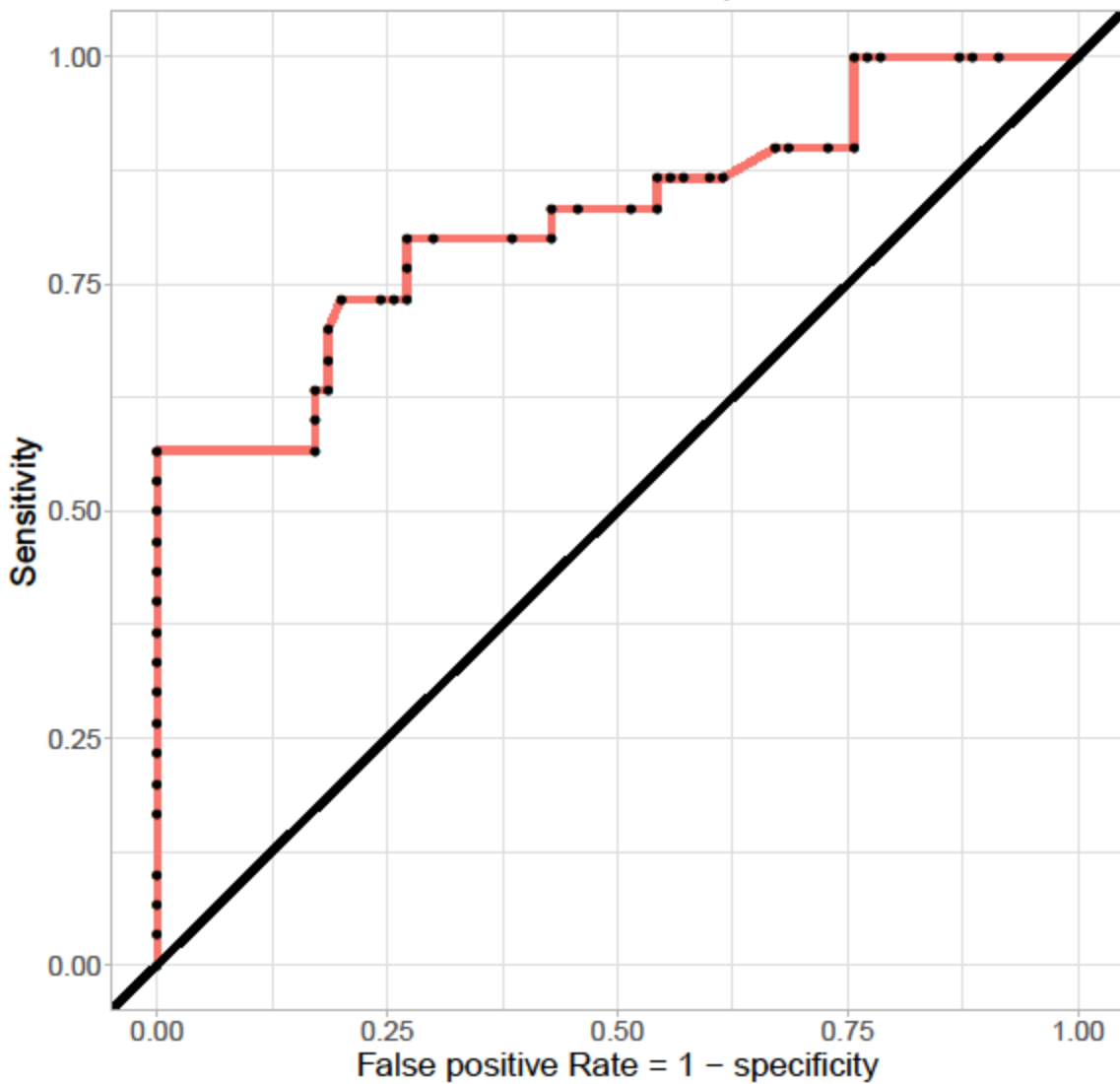
ROC curve example



approach to confidence intervals of
characteristic curves.

nts

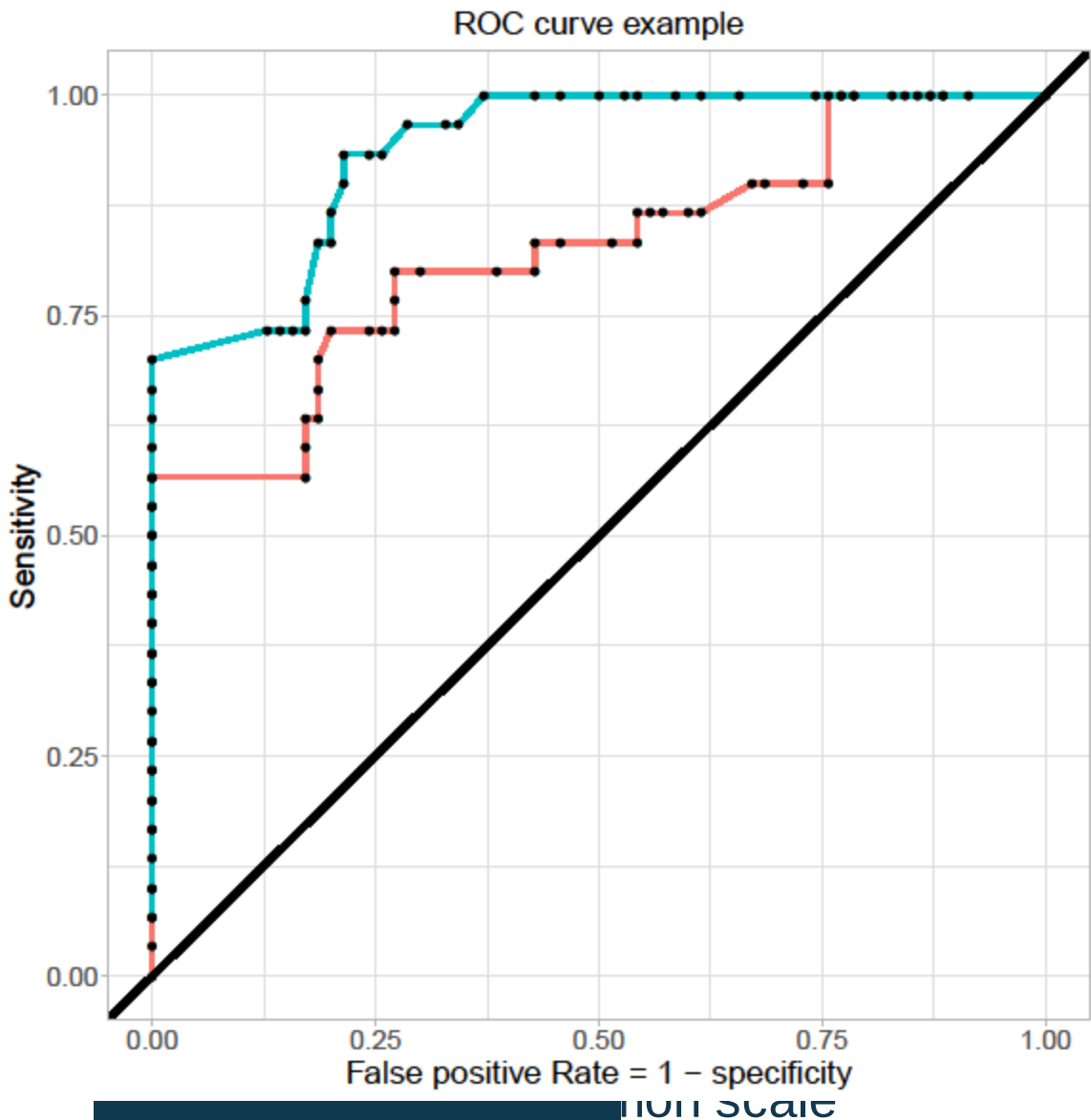
ROC curve example



approach to confidence intervals of
characteristic curves.

ER

a test
y

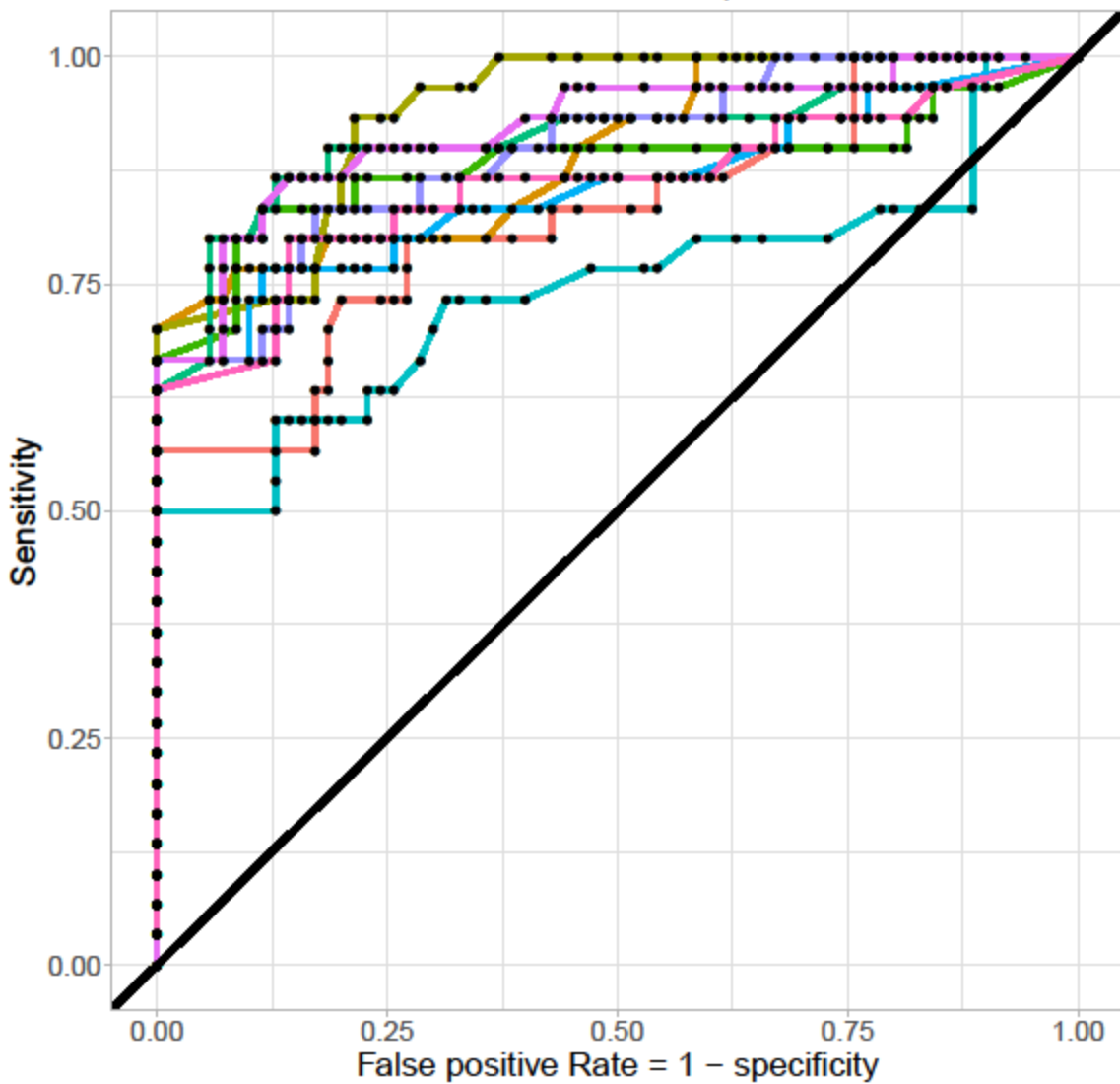


approach to confidence intervals of
characteristic curves.

ER

a test
y

ROC curve example



approach to confidence intervals of
characteristic curves.



ally



like?

;



Let's ROC: Empirical confidence of
Receiver-Operator Characteristic curves.

with

// Q1: how does the average ROC curve look like?

// The averaging along false-positive rate?

// Different for every individual

// Does not account for the 2-dimensionality of the ROC curve

// Threshold averaging as given by Fawcett [1]

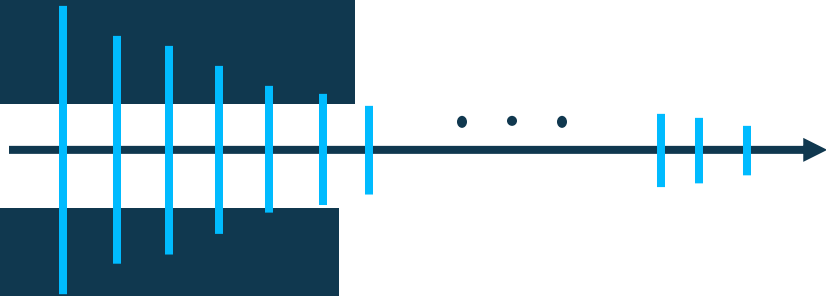
// Basic idea: generate a set of thresholds, then for each threshold find the corresponding ROC curve and averages them

Let's ROC: Empirical confidence of
Receiver-Operator characteristic curves.

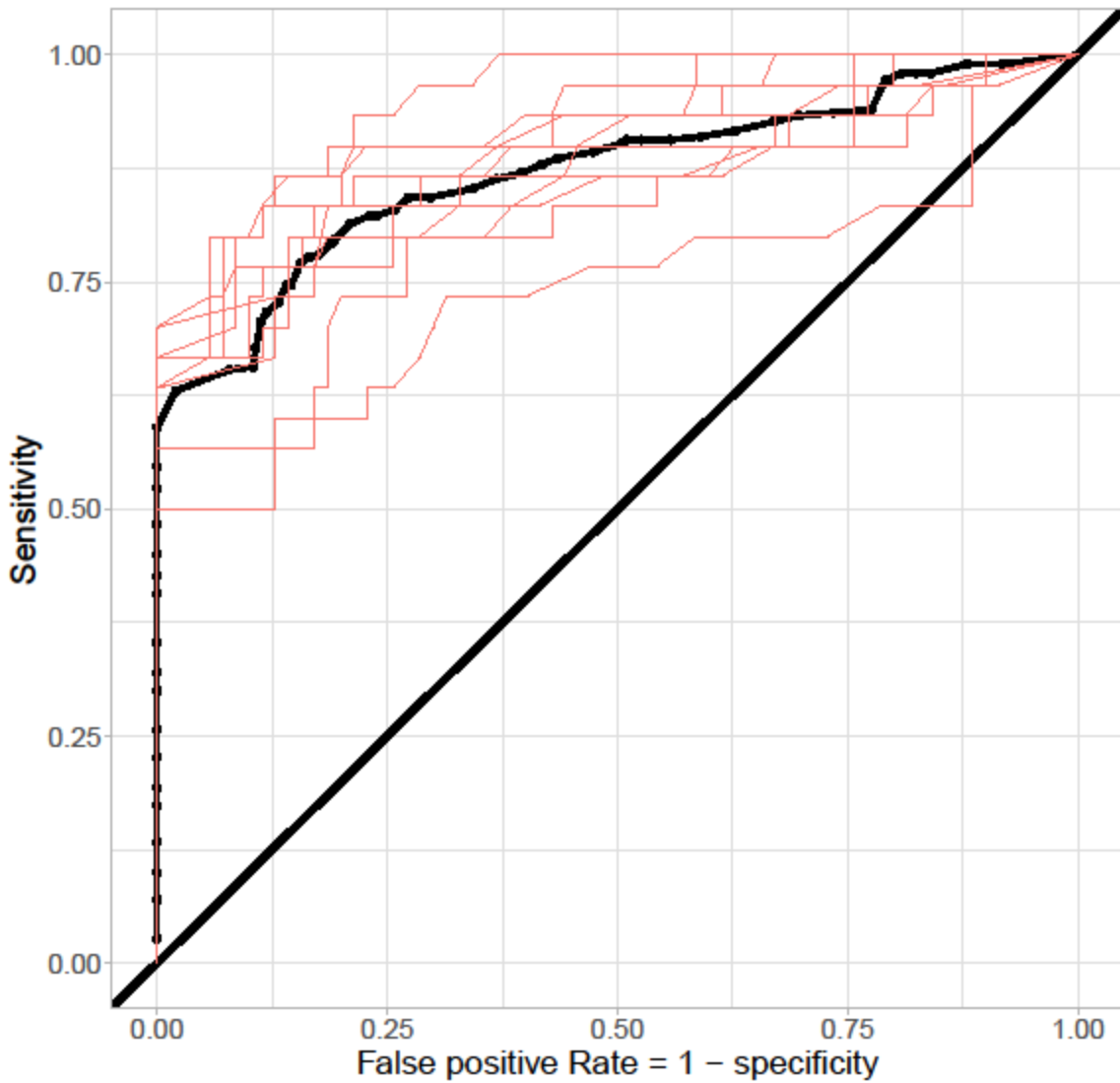
// Q1: how does the average ROC curve look like?

// Threshold averaging as by Fawcett [1]

// Idea: generate a for each threshold find the corresponding
of each ROC curve



Roc average curve together the 10 originals – simulated data



confidence of
these curves.

Is

Is

R

Let's ROC: Empirical confidence of
Receiver-Operator Characteristic curves.
Algorithm: fixed-width bands

// Q2: And how do the confidence bands
look like?

Core idea: Bootstrap

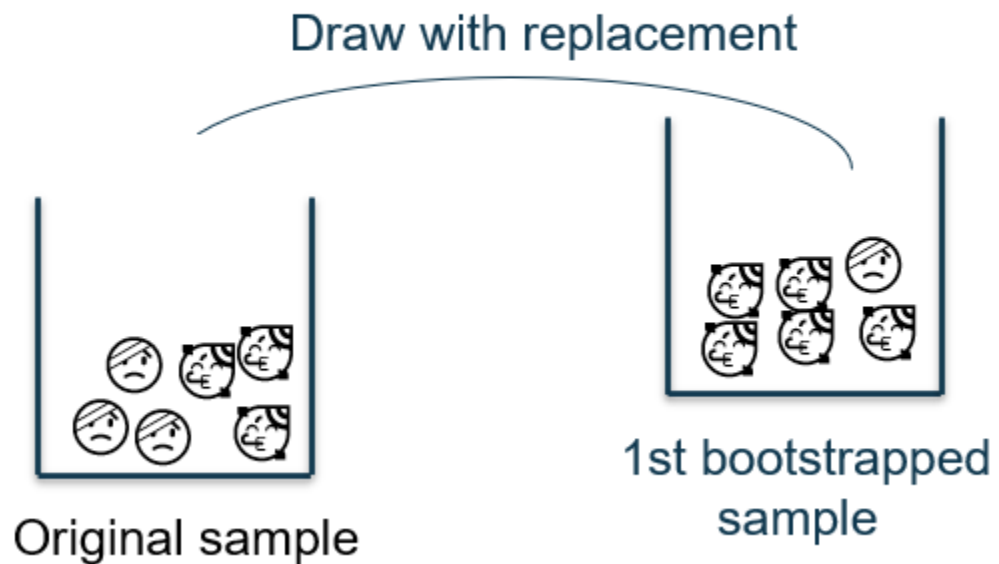


Original sample

Let's ROC: Empirical confidence of
Receiver-Operator Characteristic curves.
Algorithm: fixed-width bands

// Q2: And how do the confidence bands
look like?

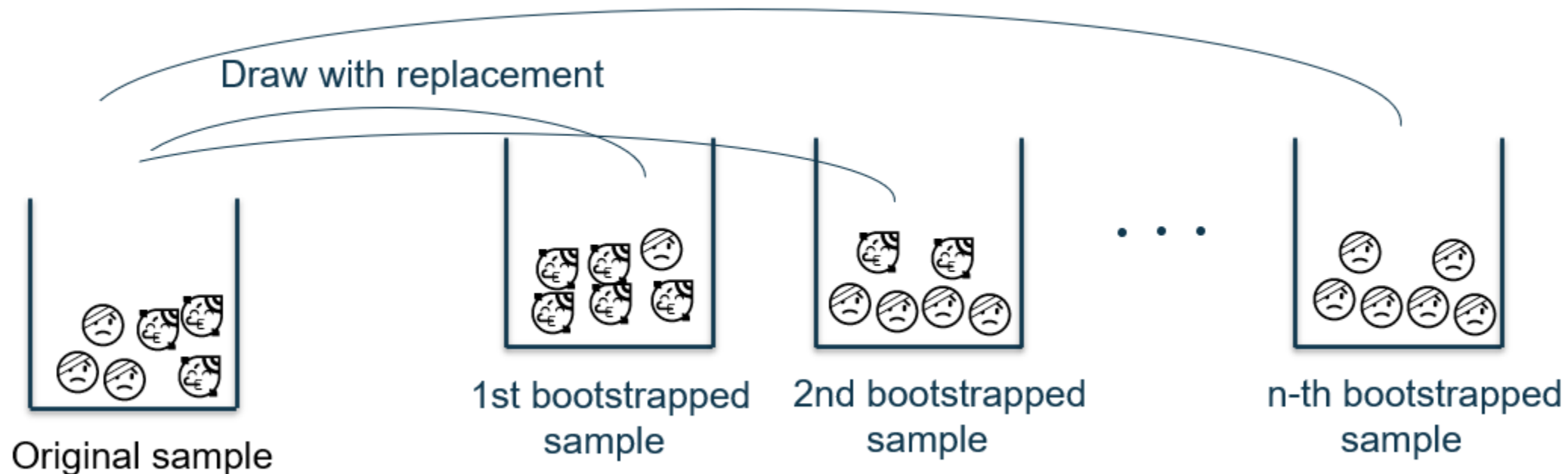
Core idea: Bootstrapping



Let's ROC: Empirical confidence of
Receiver-Operator characteristic curves.
Algorithm: fixed-width bands

// Q2: And how do the confidence bands
look like?

Core idea: Bootstrapping

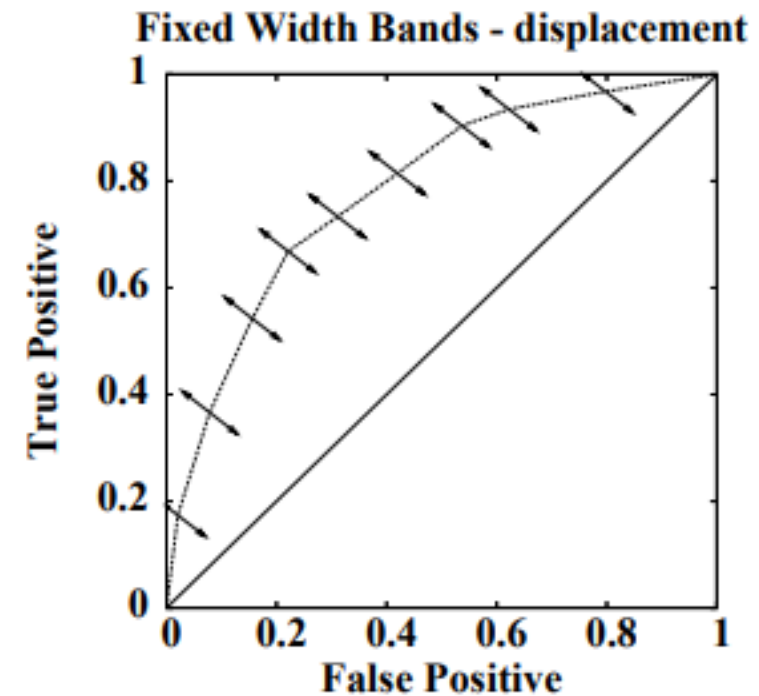


Let's ROC: Campbell confidence of
Receiver-Operator characteristic curves.
Algorithm. Fixed-width confidence bands

// how do the confidence bands

look like?

// from Campbell,

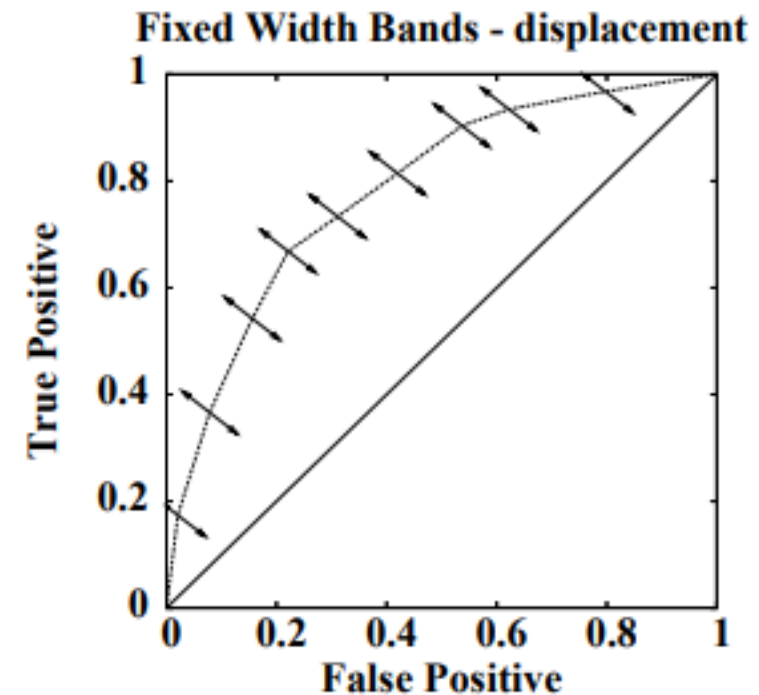


Let's ROC: Empirical confidence of
Receiver-Operator Characteristic curves.
Algorithm. Fixed-width confidence bands

// How do the confidence bands
look like?

// 3 simple steps:

1. Identify the slope by which to displace
the original ROC curve



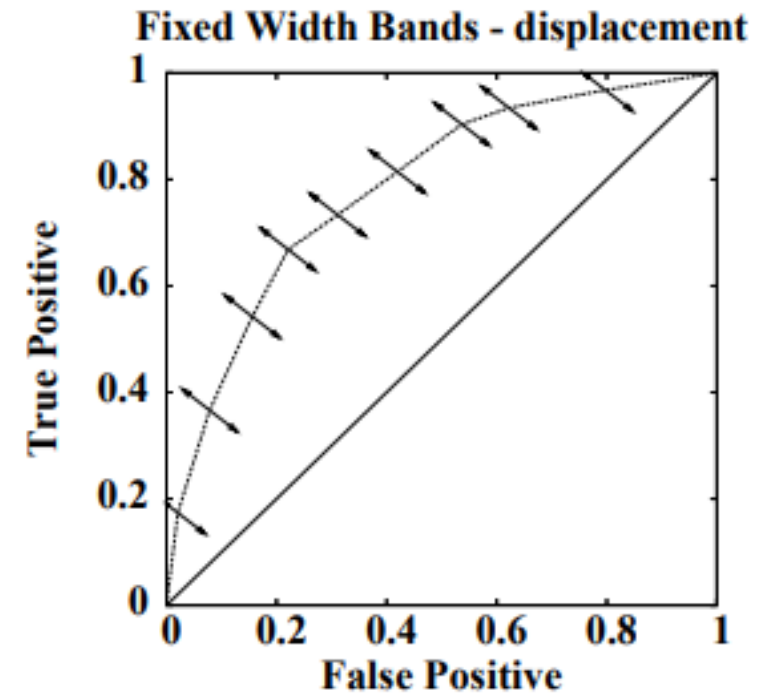
Let's ROC: Empirical confidence of Receiver-Operator Characteristic curves. Algorithm. Fixed-width confidence bands

// How do the confidence bands look like?

// 3 simple steps:

1. Identify the slope by which to displace

2. From original sample D , we generate D^* by B bootstraps. Calculate aveROC for each D^*



Let's ROC: Empirical confidence of Receiver-Operator Characteristic curves. Algorithm. Fixed-width confidence bands

// How do the confidence bands look like?

// 3 simple steps:

1. Identify the slope by which to displace

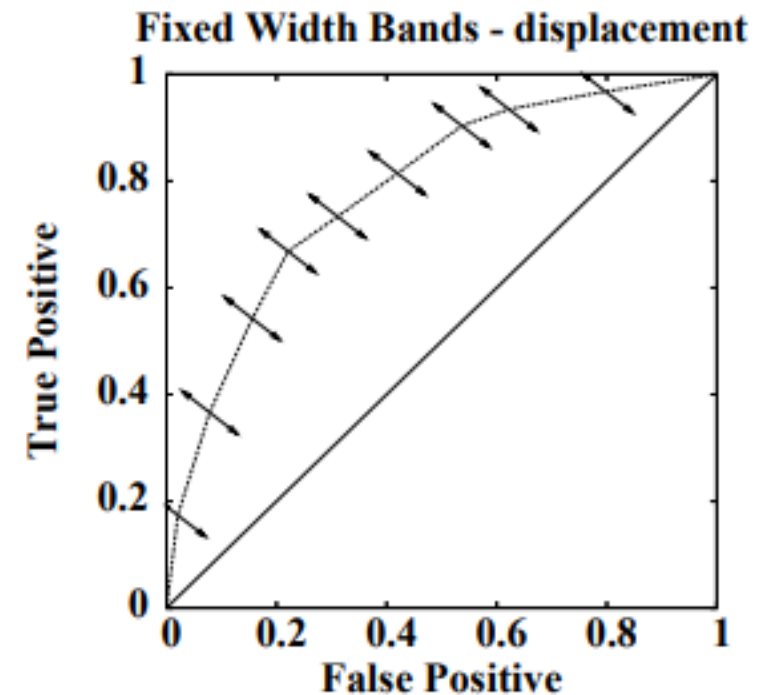
ROC curve

2. From original sample D , we generate D^*

by B bootstraps. Calculate aveROC for each D^*

3. Find maximum distance from original aveROC

curve to each of the aveROC curves of the B bootstrap samples



Let's ROC: Empirical confidence of Receiver-Operator Characteristic curves. Algorithm. Fixed-width confidence bands

// How do the confidence bands look like?

// 3 simple steps:

1. Identify the slope by which to displace

ROC curve

2. From original sample D , we generate D^*

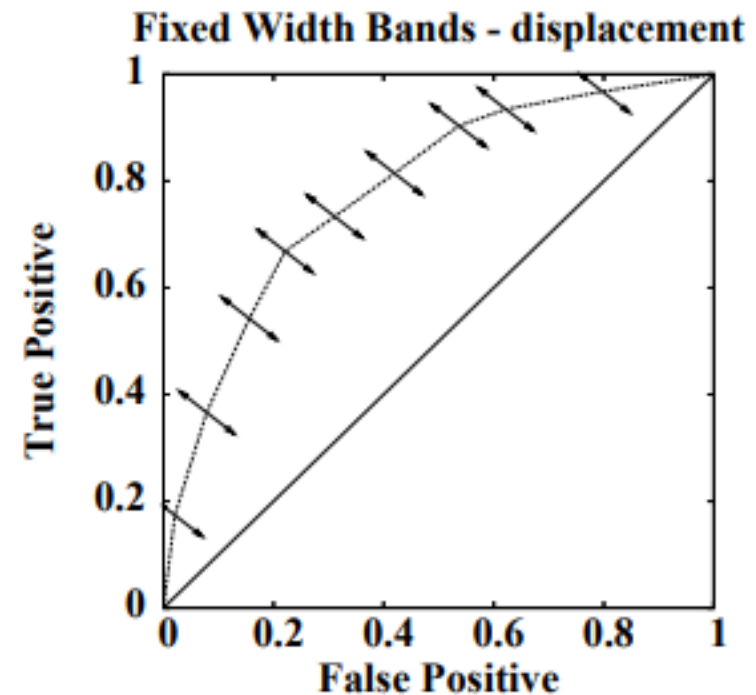
by B bootstraps. Calculate aveROC for each D^*

3. Find maximum distance from original aveROC

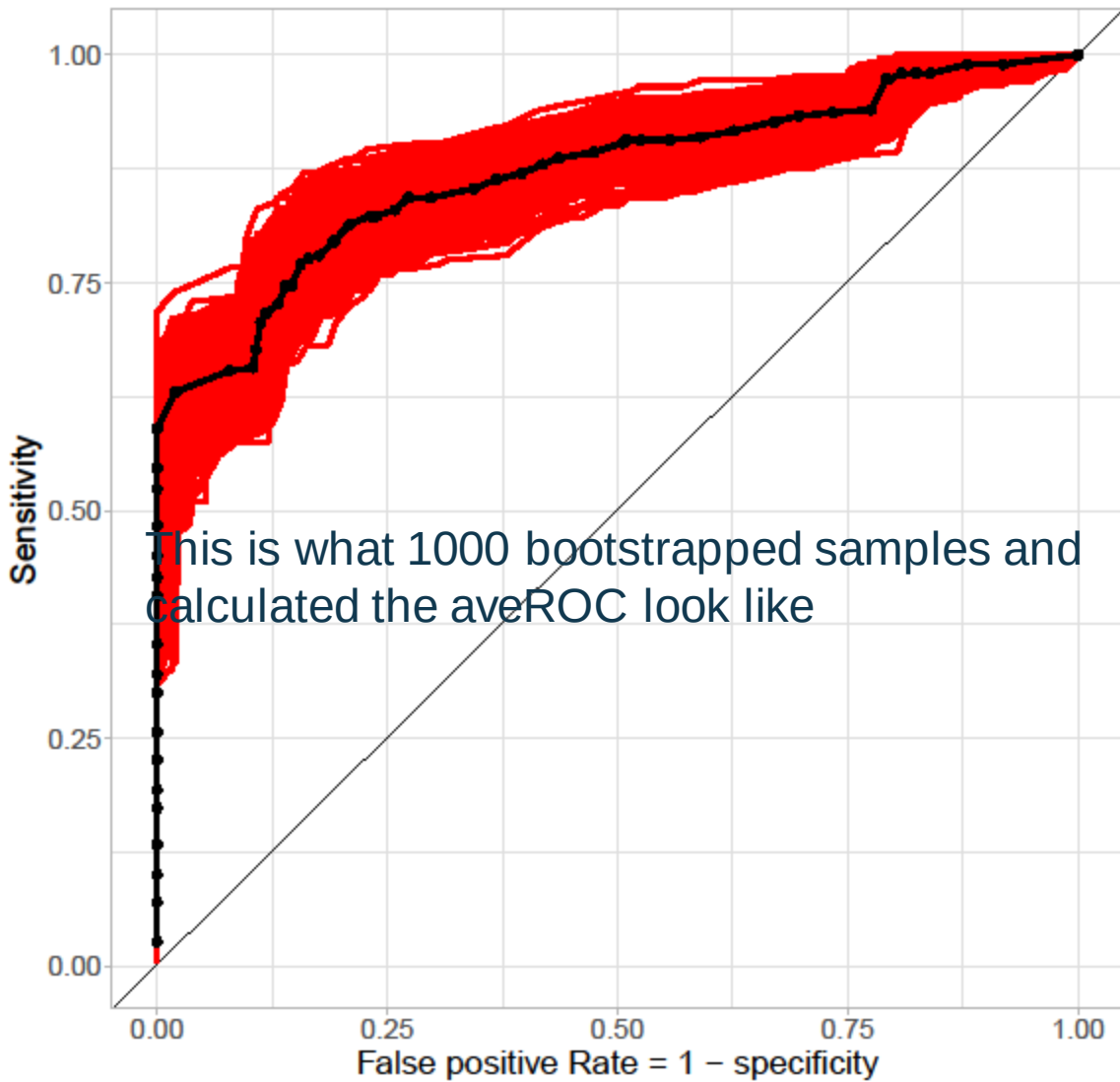
curve to each of the aveROC curves of the B bootstrap samples

3*. Take 1-delta (delta i.e. 0.05) quantile of these maxima,

(1-delta) % of the curves lie completely within the generated bands.



original Roc curve with all 1000 bootstrap samples – simulated data



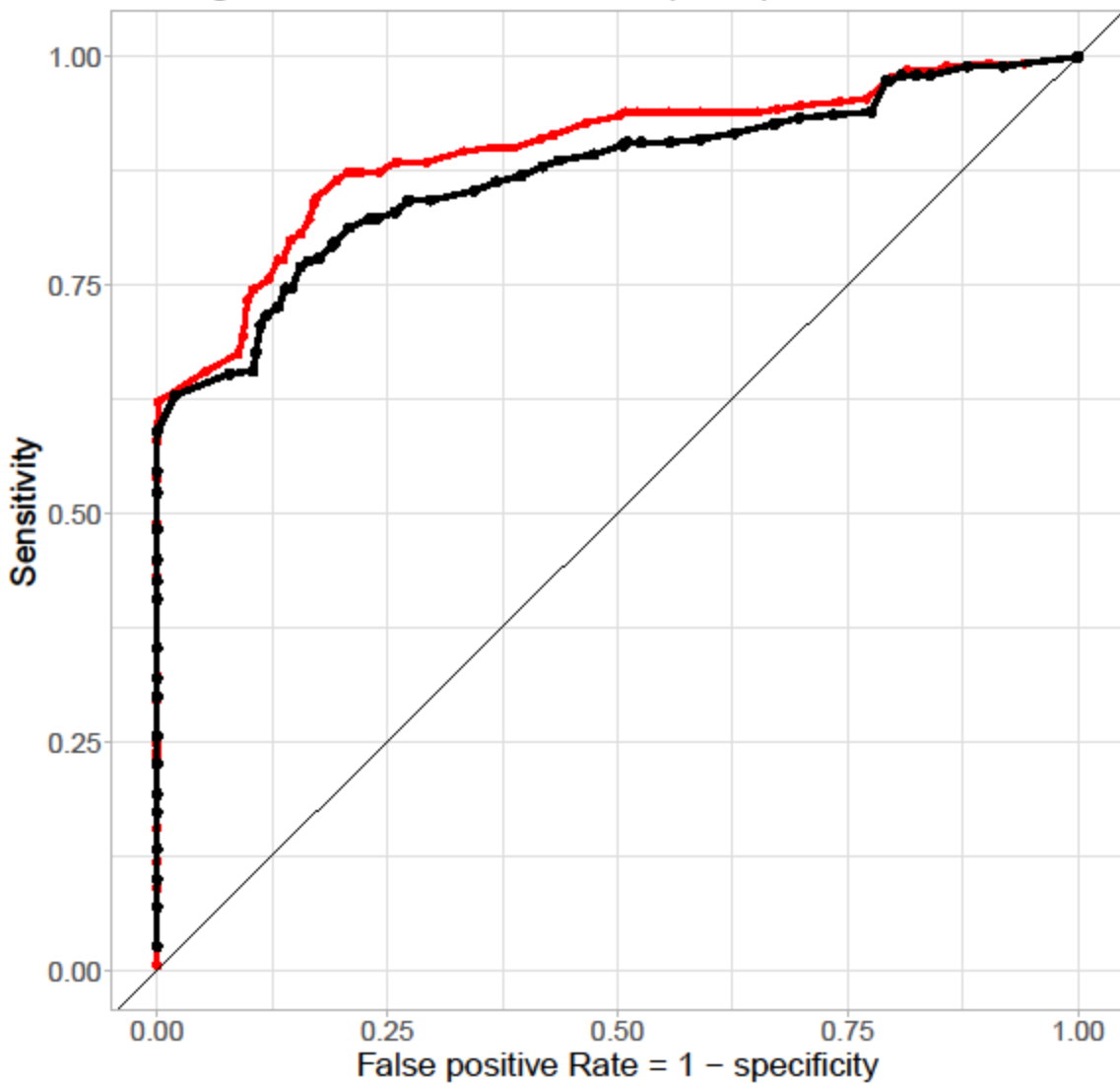
confidence bands of
these curves.

confidence bands

distance between 2 ROC

R

original Roc curve with 1 bootstrap sample – simulated data



confidence bands of
these curves.

confidence bands

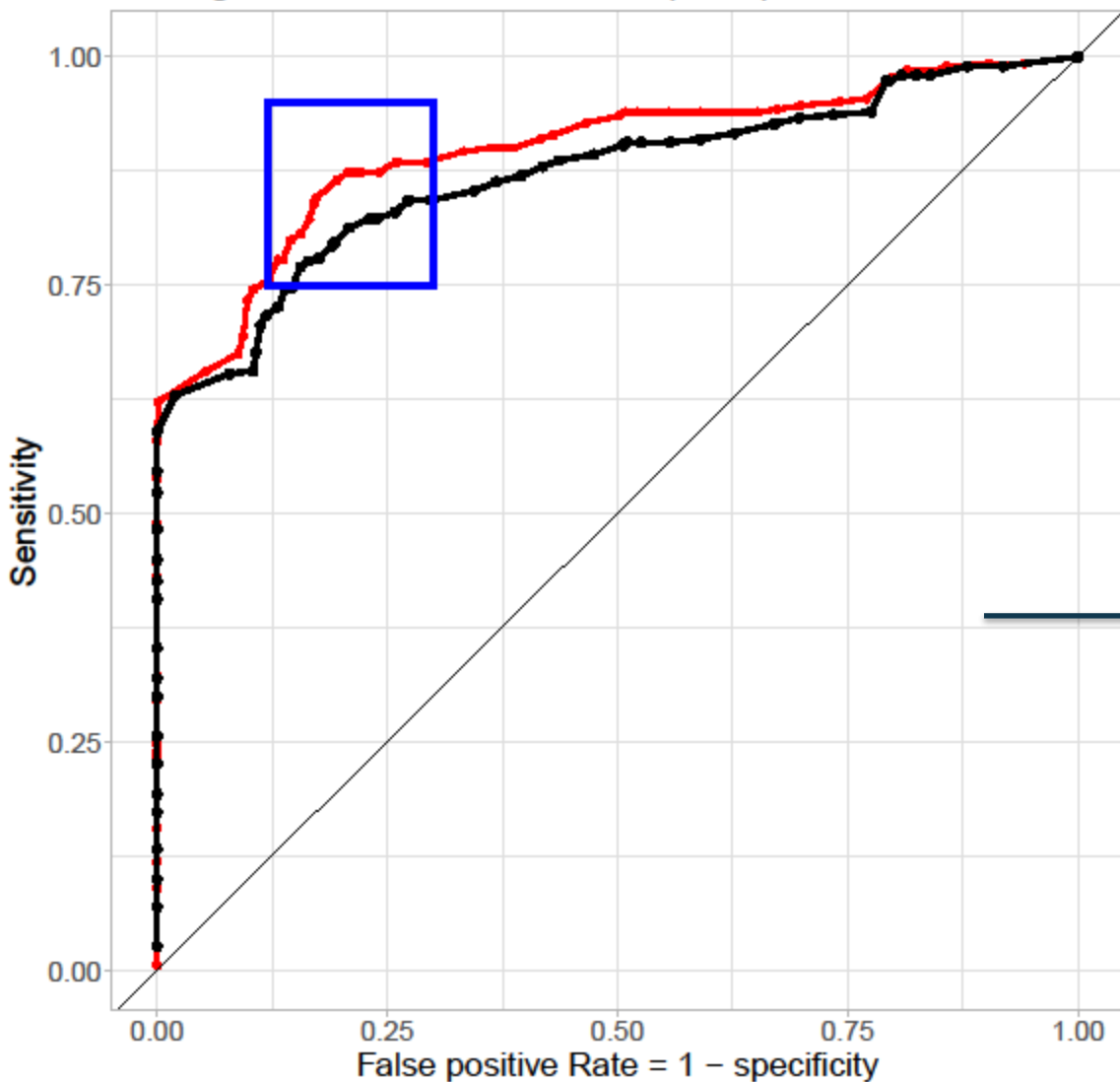
distance between 2 ROC



R

Where would you expect the
maximum distance in direction of
the slopes to be?

original Roc curve with 1 bootstrap sample – simulated data



confidence bands of
these curves.

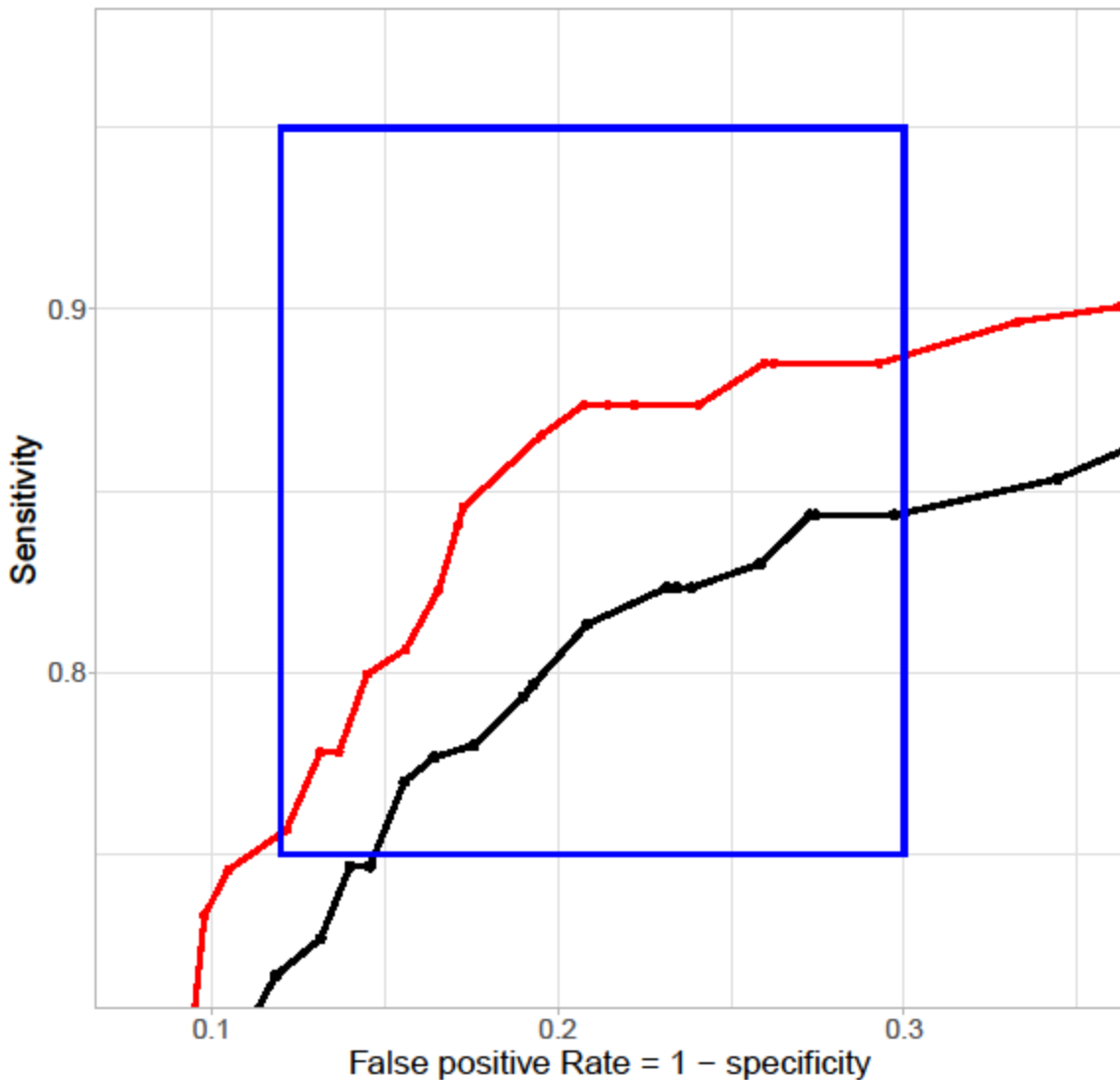
confidence bands

distance between 2 ROC

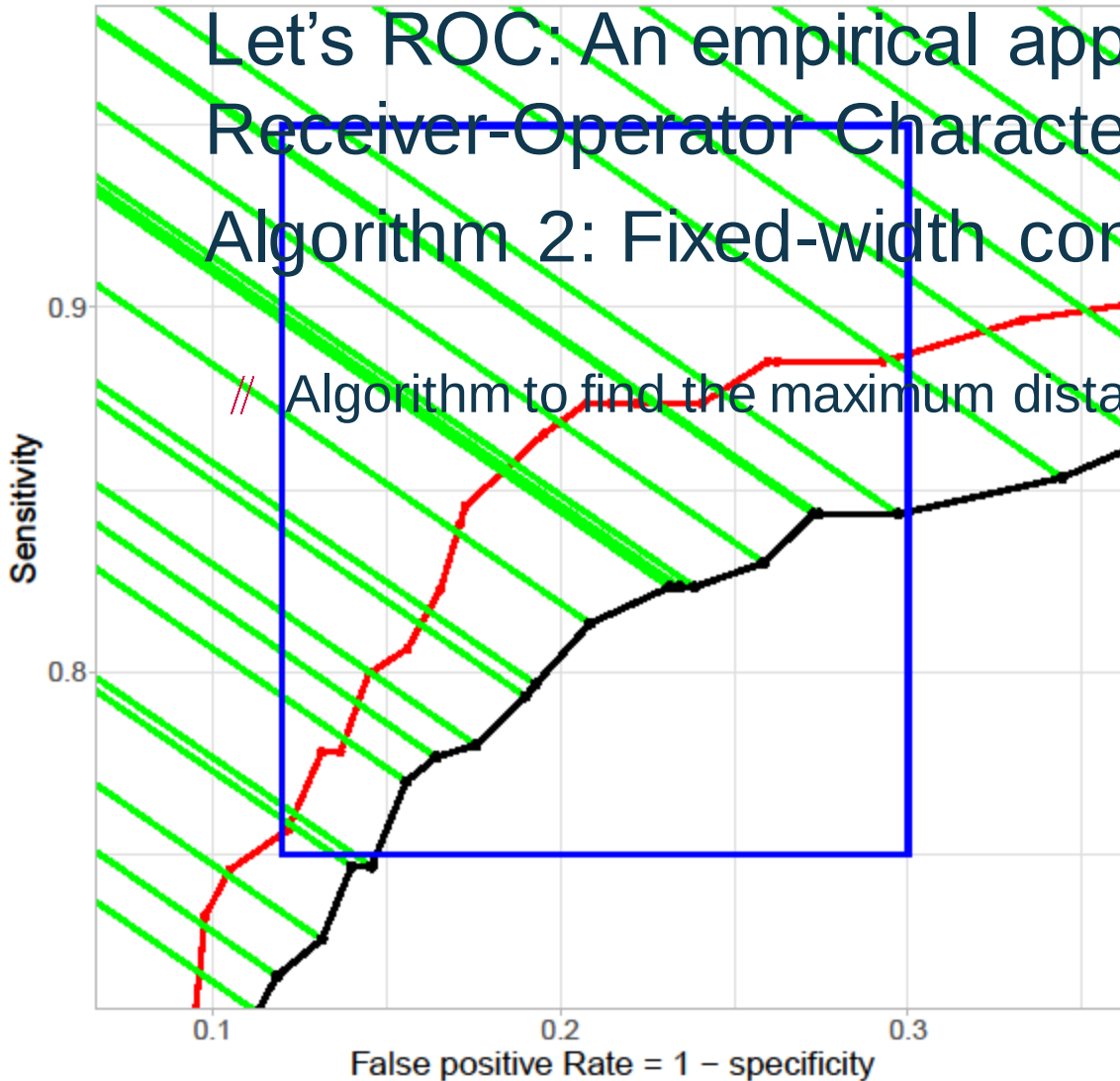
R

Where would you expect the
maximum distance in direction of
the slopes to be?

original Roc curve with 1 bootstrap sample / Zoomed – simulated data



confidence bands of
these curves.
confidence bands
difference between 2 ROC
curves.
Zoom on the



Let's ROC: An empirical approach of confidence of Receiver-Operator Characteristic curves.

Algorithm 2: Fixed-width confidence bands

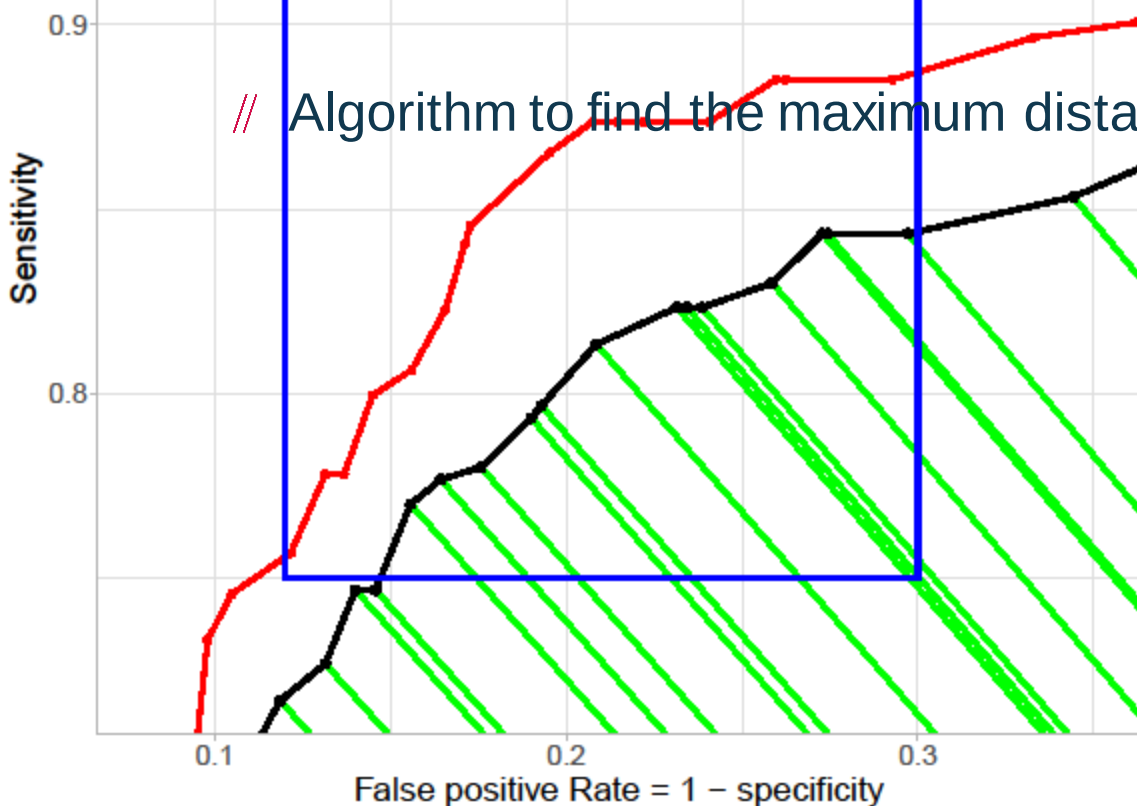
// Algorithm to find the maximum distance between 2 ROC

R

- on the area
- add lines above the original aveROC in the direction of the slope
- Where would the green lines cross the red ones?

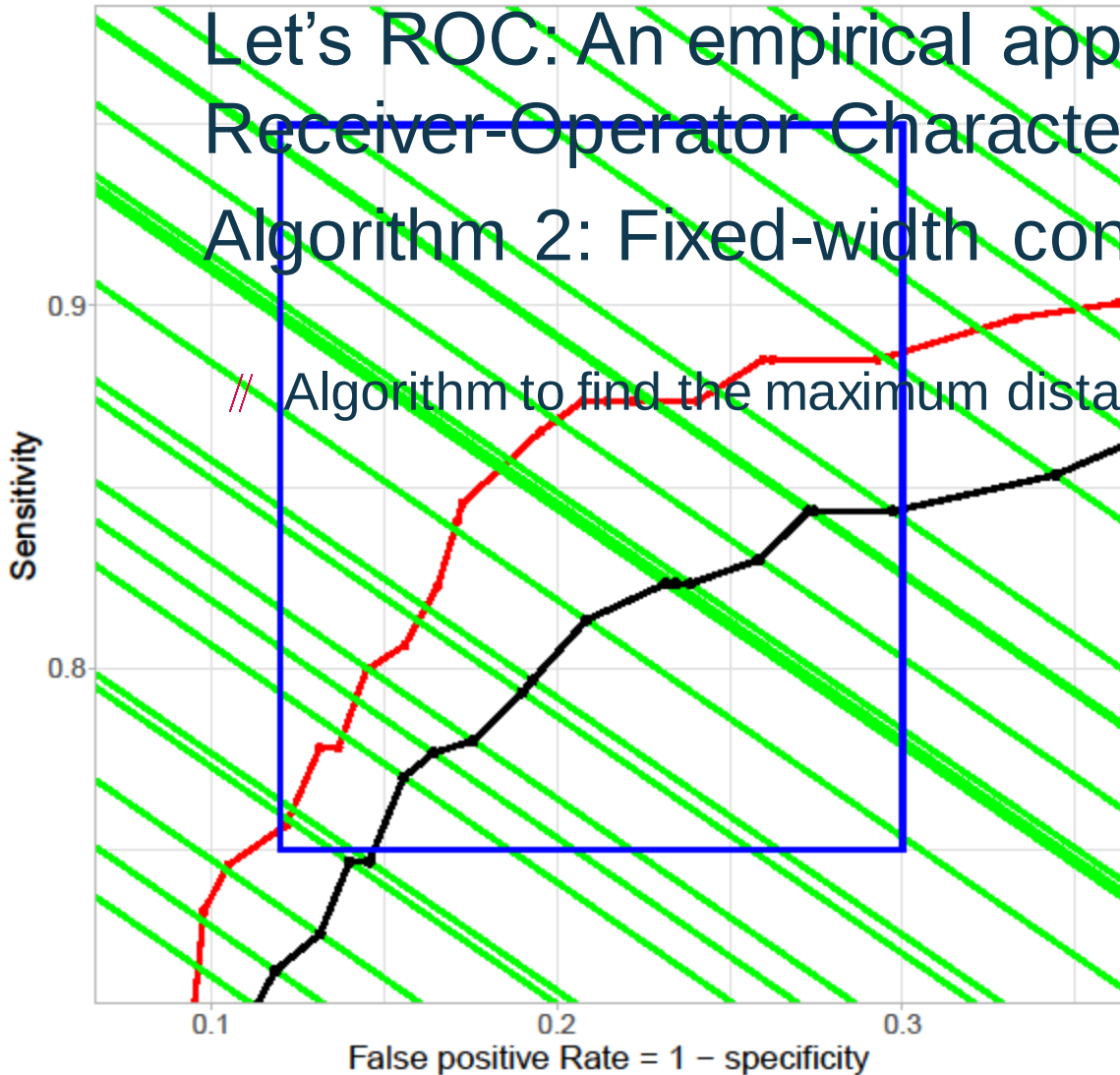
Let's ROC: An empirical approach to confidence of Receiver-Operator Characteristic curves.

Algorithm 2: Fixed-width confidence bands



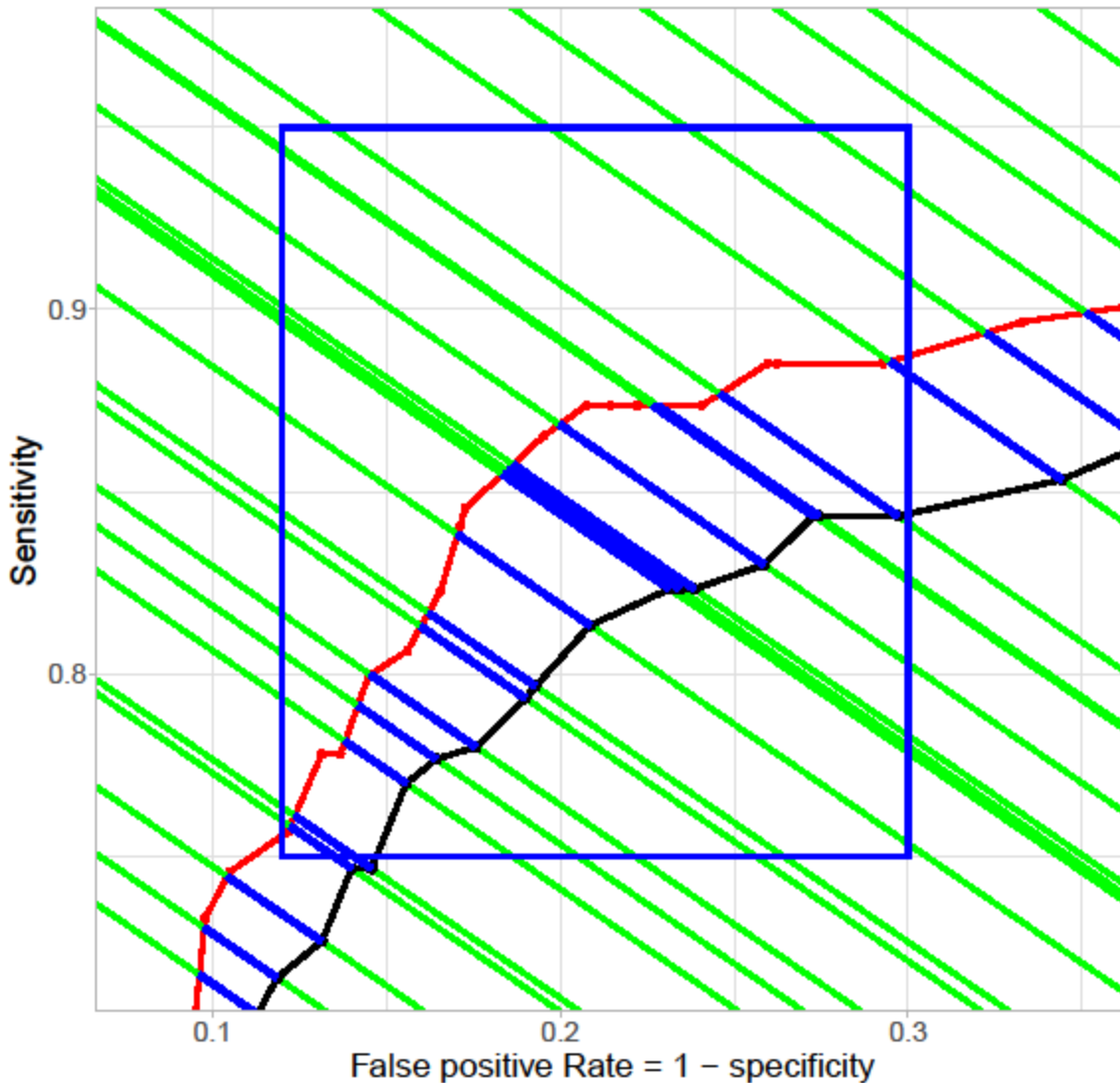
ROC

- add lines on the area
- add lines below the original aveROC in the direction of the slope, in cases where the red ROC would lie below the black ROC



- Make it easier
- Just add lines that run through the original ROC points in direction of the slope

original Roc curve with 1 bootstrap sample – simulated data



confidence of
these curves.

confidence bands

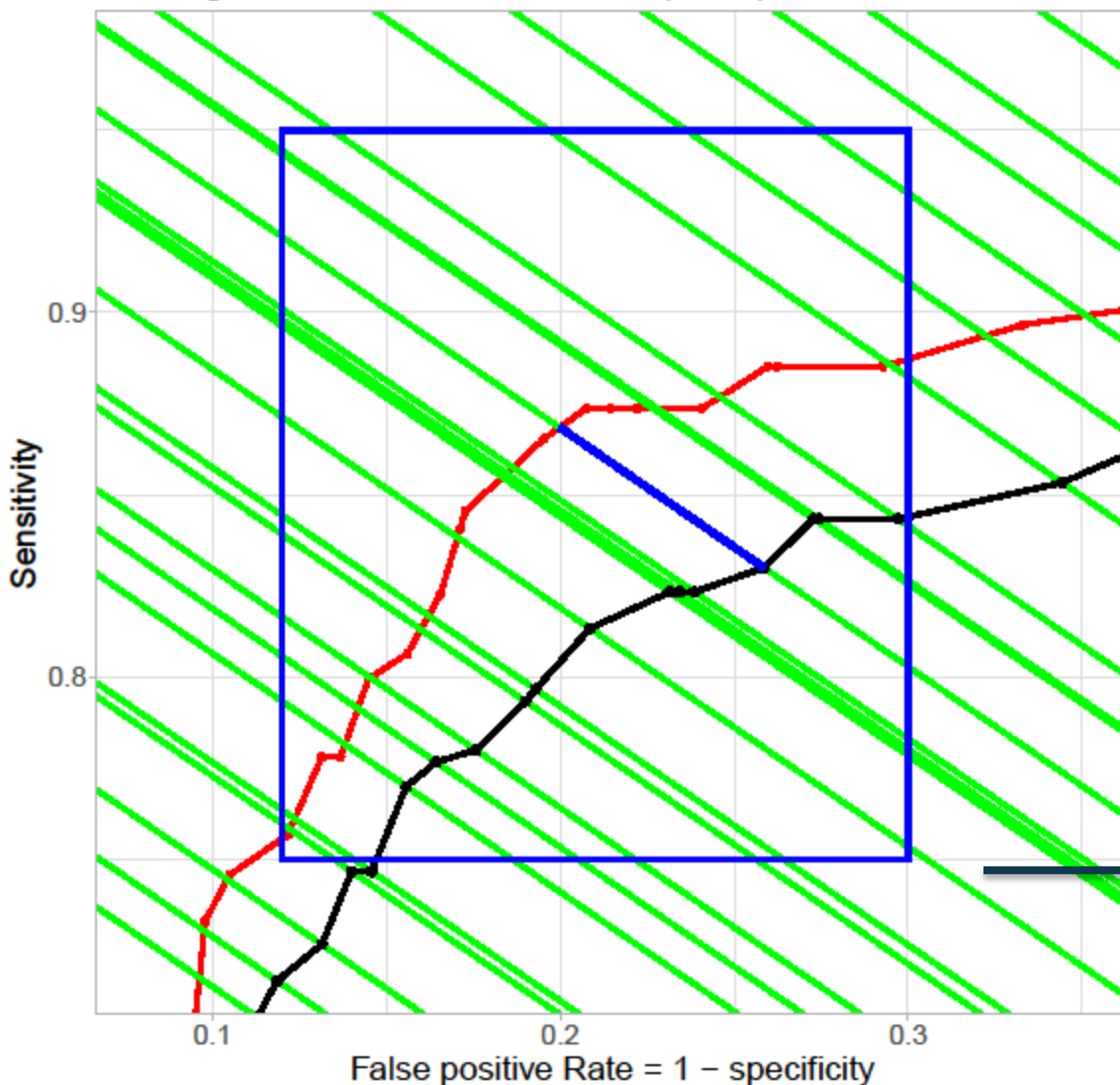
distance between 2 ROC

- Make it easier
- Just add lines that run through the original ROC points in direction of the slope

Mini steps:

- where do the green lines cross the red ones?
- and what is the distance between these points?

original Roc curve with 1 bootstrap sample – simulated data



confidence bands of
ROC curves.

confidence bands

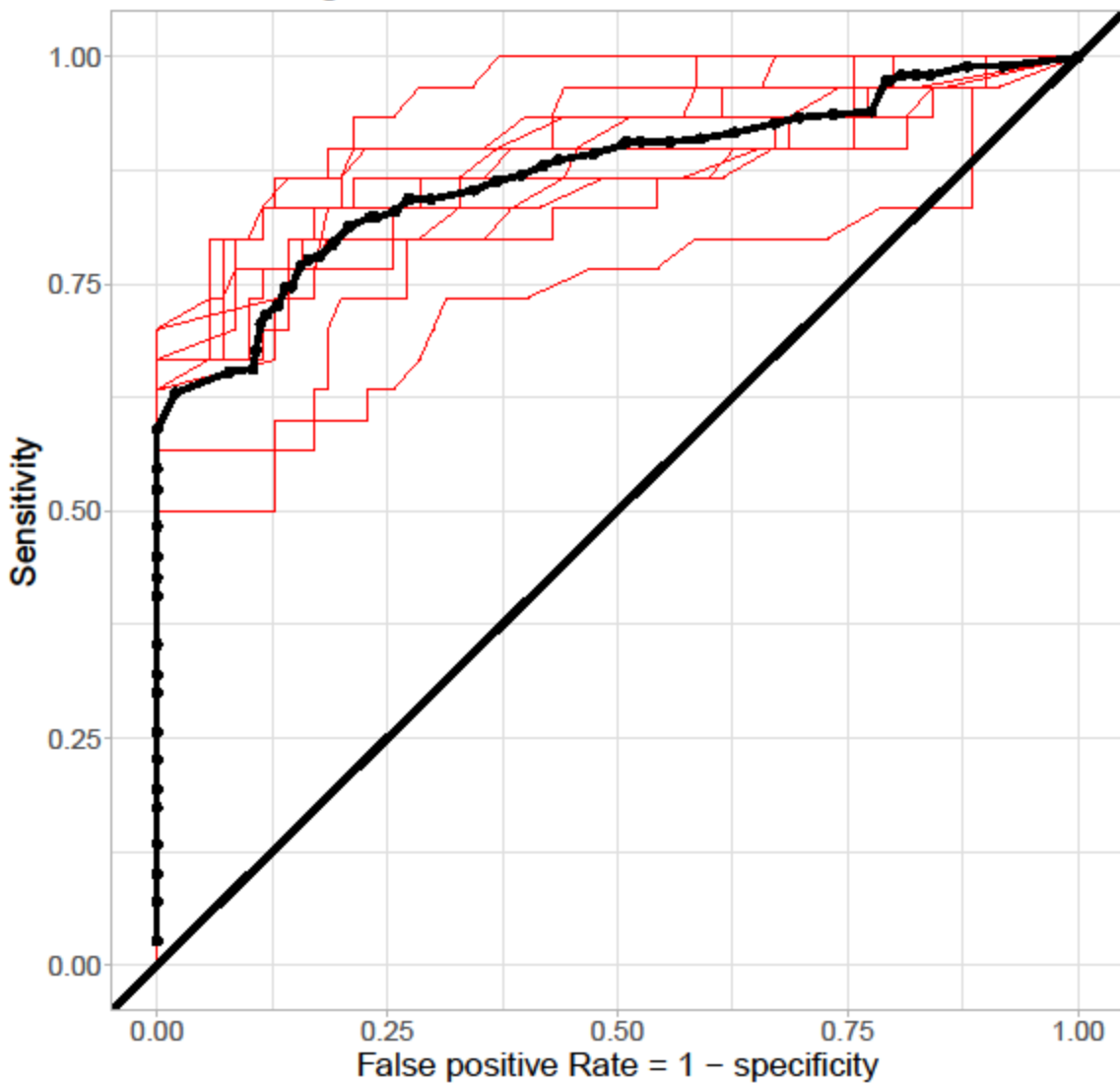
distance between 2 ROC

Mini steps

- where green lines cross red ones?
- and what is the distance between these points?

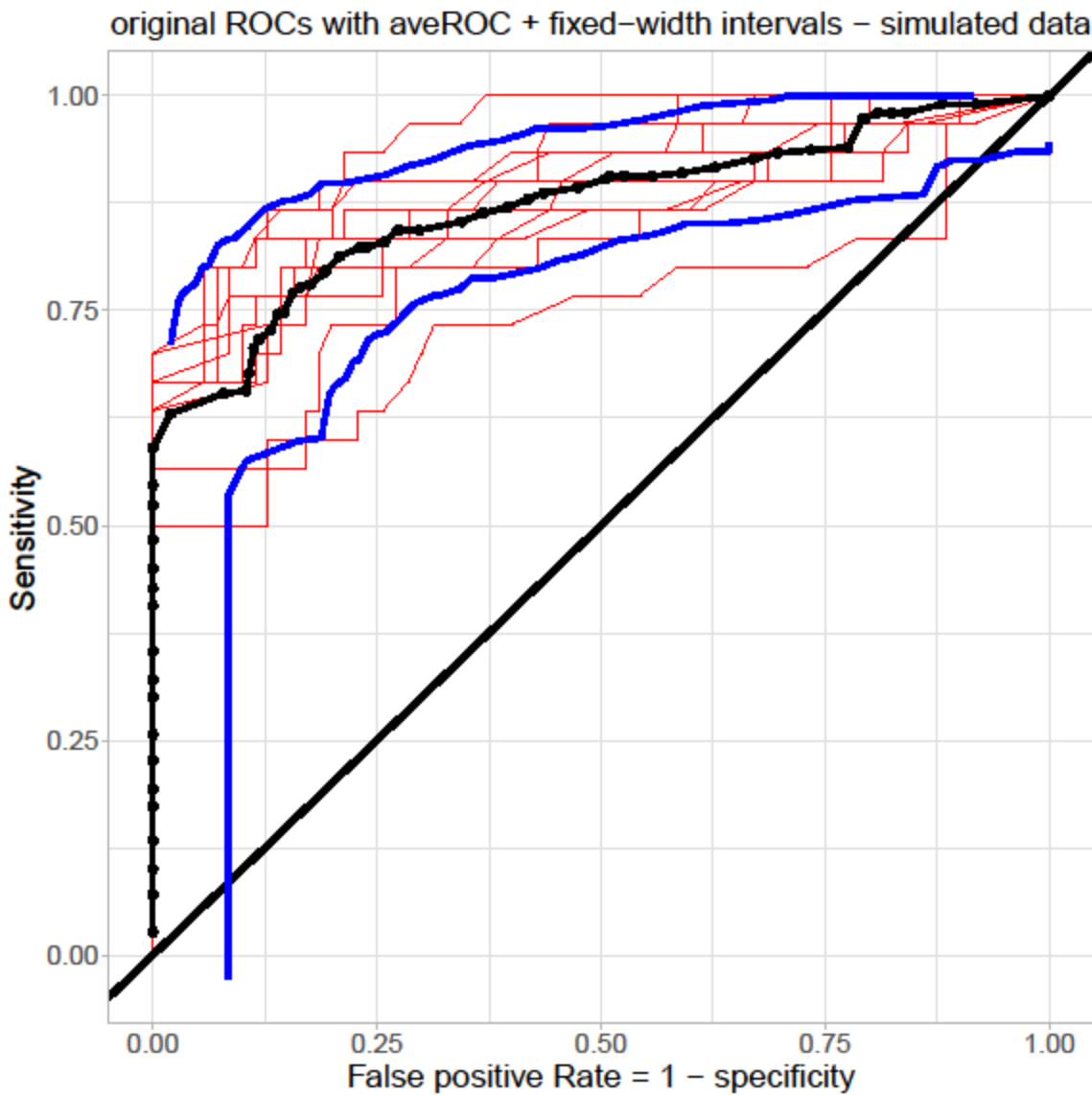
Once the maximum between original and
bootstrapped ROC curve is found
→ store it!
→ and repeat for all other **B** bootstraps,
i.e. find the maxima to all the other
bootstrapped aveROC curves.

original ROCs with aveROC – simulated data



confidence bands of
ROC curves.
confidence bands
bands
bands

the
selection of
fixed-



confidence bands of
individual curves.

confidence bands

bands

$\epsilon_1 = 0$

the
selection of
fixed-

R

Let's ROC: Empirical confidence of Receiver-Operator Characteristic curves.

Command and Output: Fixed-width confidence bands

- Introduced the concept of a stochastic test with
- Introduced ROC curve as a means of accessing performance of a stochastic test
- Introduced how to compare ROC curves for multiple assessments
- We introduced bootstrapping and how to calculate the fixed width bands around ROC curve

- Details will be given in a paper or subsequent presentation(s)
 - How to get the math equations of all the necessary pieces and put them together
 - Pay attention to and the process to get to
 - Pseudo-code for the algorithm

Let's ROC: Empirical confidence of
Receiver-Operator Characteristic curves.
Ratio

- [1] Fawcett, J. (2006) An Introduction to ROC Analysis. Pattern Recognition Letters, 27, 861-874.
<https://doi.org/10.1016/j.patrec.2005.10.010>
- [2] Colton, T., G. (1994). Advanced statistics for the evaluation of diagnostic and laboratory tests. Statistics in medicine, 13, 459-500.

YES

Thank you!

FE

