

# 刺激・反応マトリクスから求まる指標

|                  |          | 出力: 反応<br>観察者が判断したクラス |                 |
|------------------|----------|-----------------------|-----------------|
|                  |          | positive              | negative        |
| 入力: 刺激<br>実際のクラス | positive | TP<br>(       )       | FN<br>(       ) |
|                  | negative | FP<br>(       )       | TN<br>(       ) |

- ✓ Sensitivity (感度)
- ✓ Specificity (特異度)
- ✓ Positive predictive value (陽性的中率: PPV)
- ✓ Negative predictive value (陰性的中率: NPV)

◆ 感度は\_\_\_\_\_とも呼ばれる

◆ 陽性的中率は\_\_\_\_\_とも呼ばれる

- ✓ Accuracy (正診率, 正解率)
- ✓ F score (F値)

Accuracy =

F score =

=

# 各指標を求めてみよう！（１）

|                 |          | 出力：反応<br>観察者が判断したクラス |          |
|-----------------|----------|----------------------|----------|
|                 |          | positive             | negative |
| 入力：刺激<br>実際のクラス | positive | 20                   | 80       |
|                 | negative | 20                   | 10,000   |

*Accuracy*=

*Sensitivity*=

*Specificity*=

*PPV*=

*NPV*=

例えば,

positive: 肺がんを含む画像

negative: 正常画像

と考えると. . . .

*F score*=

# 各指標を求めてみよう！（２）

|                 |          | 出力：反応<br>観察者が判断したクラス |          |
|-----------------|----------|----------------------|----------|
|                 |          | positive             | negative |
| 入力：刺激<br>実際のクラス | positive | 1,000                | 80       |
|                 | negative | 20                   | 10,000   |

*Accuracy*=

*Sensitivity*=

*Specificity*=

*PPV*=

*NPV*=

例えば,

positive: 肺がんを含む画像

negative: 正常画像

と考えると. . . .

*F score*=

## 2018国家試験問題

ある癌診断の画像検査で、真陽性率が98%、偽陽性率が5%であった。この癌の一般的な罹患率は1%である。ある人がこの画像検査を受けて陽性と判断されたとき、実際に癌に罹患している確率に最も近いのはどれか。

1. 10%
2. 17%
3. 25%
4. 67%
5. 95%

|         |          | 出力: 検査結果 |          |
|---------|----------|----------|----------|
|         |          | positive | negative |
| 入力: 被検者 | positive |          |          |
|         | negative |          |          |

# ROC実験の手順

なにを評価するかが大事！

例えば：

- ・ 検出器の違いによる信号検出能
- ・ 撮影条件設定による信号検出能
- ・ 単純X線撮影とCTの病変検出能
- ・ 熟練医と研修医の読影能力
- ・ \_\_\_\_\_ (computer-aided diagnosis: CAD) システムの評価 [without CAD vs. with CAD]

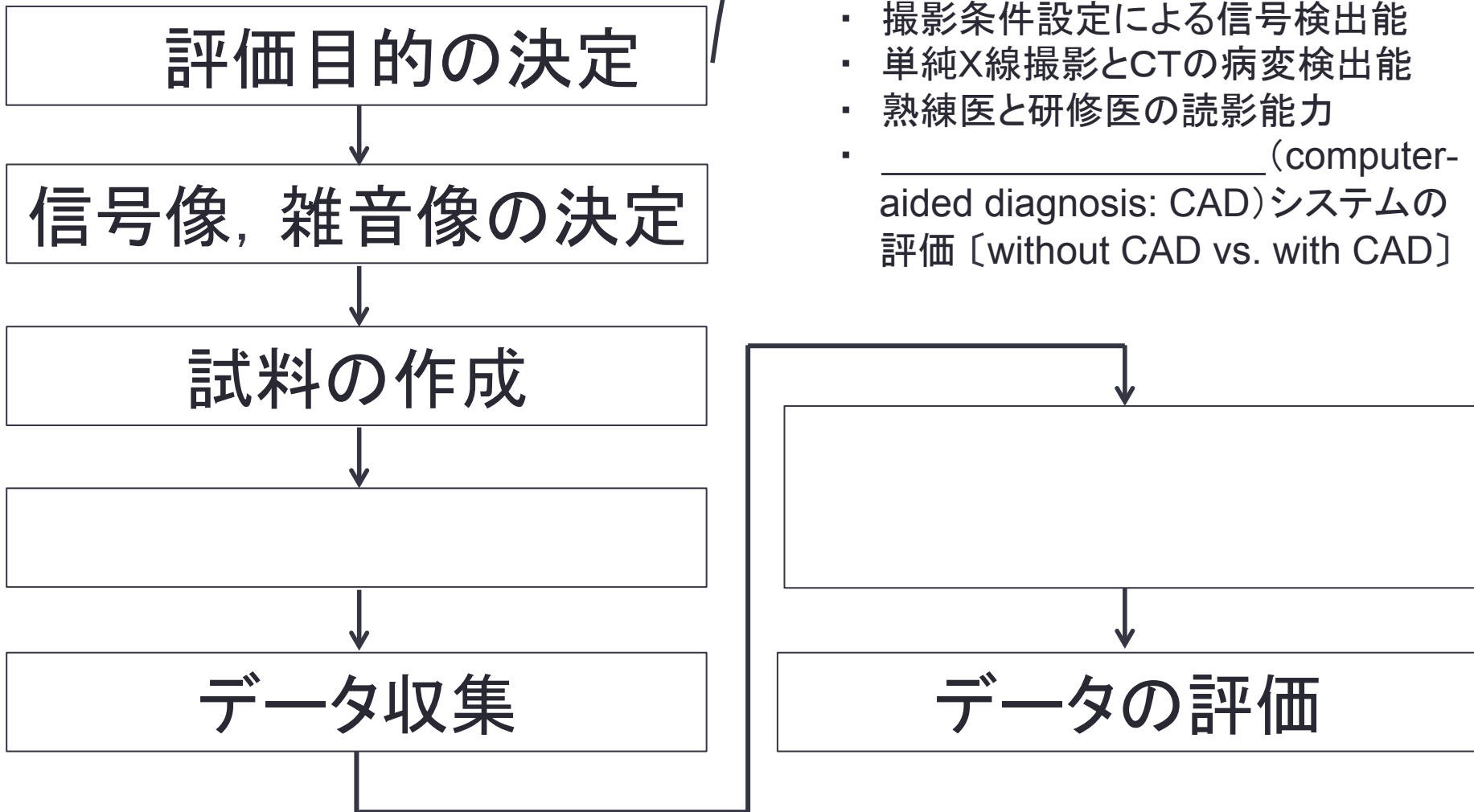
評価目的の決定

信号像，雑音像の決定

試料の作成

データ収集

データの評価



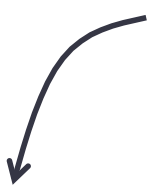
# ROC曲線の作成<sup>1</sup>

## (A) 実験方法

### (1) 画像受光系:

測定の対象となる撮像系を用意する

CR, DR



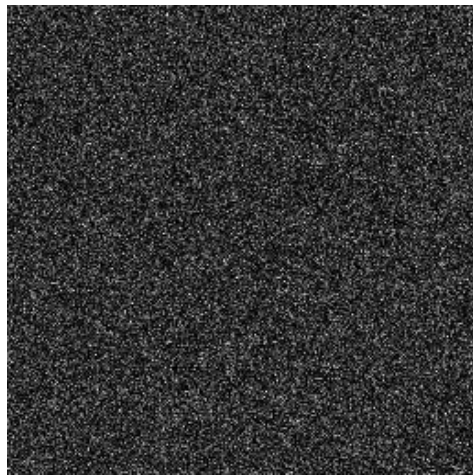
### (2) 微小信号:

微小信号として、アクリル樹脂製の2 mm直径のビーズ玉を数個用意する

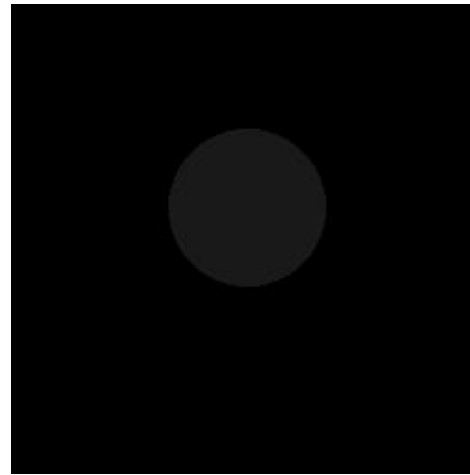
### (3) サンプル作成:

例えば、1つの撮像系に対して、\_\_\_\_\_(微小信号＋雑音の画面)100枚、\_\_\_\_\_(信号を含まない画面)100枚、の合計200枚を撮影する。

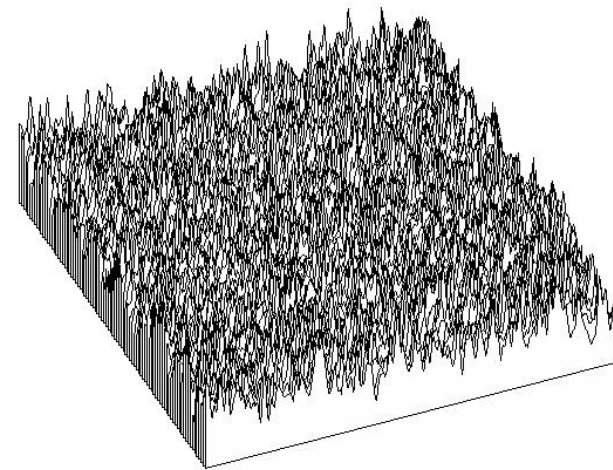
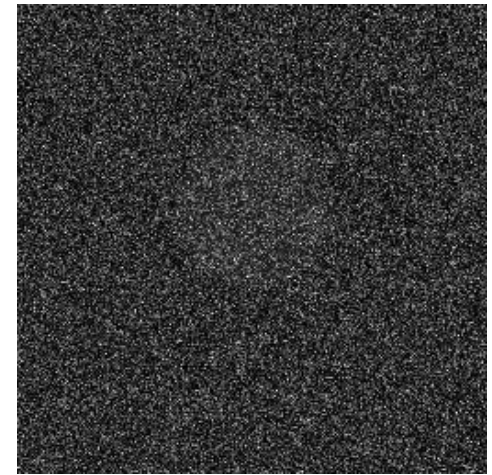
単純なディスク信号を仮定したときの雑音像(negative像)と信号+雑音像(positive像)



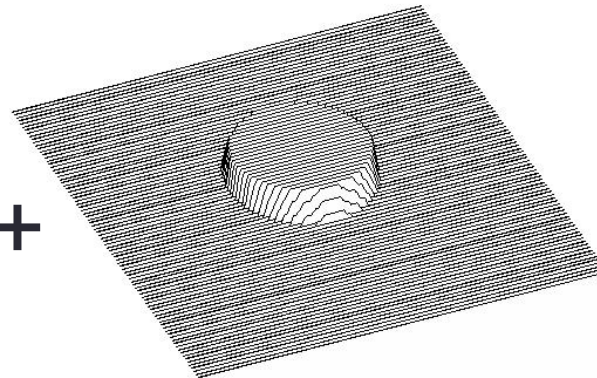
+



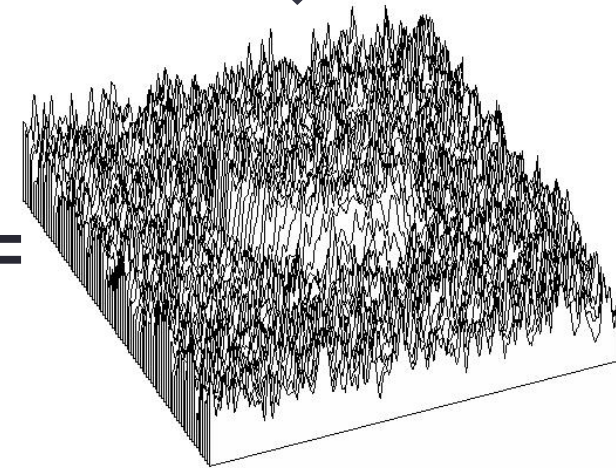
=



+



=



雑音  
(Negative)

信号

雑音+信号  
(positive)

# ROC曲線の作成<sup>2</sup>

\* 画面上，信号影の判読が\_\_\_\_\_を作成する

\* はっきりと\_\_\_\_\_が読み取れる画面は不適當

(4) 撮影 & 表示条件：

例えば，画面の写真濃度は $0.65 \pm 0.03$ になるように設定  
表示条件（ウィンドウ条件）は一定に保つこと

## (B) 実験方法

(1) サンプルの提示：

全サンプル（200枚）の中から，\_\_\_\_\_に1枚ずつを取り出し，モニタに提示する

(2) 観測距離： 40～50 cm

(3) 観測時間： 10～18 sec



# ROC曲線の作成<sup>3</sup>

## (C) 評定実験

- I. 信号(病変)は絶対ある
- II. 信号(病変)は多分ある
- III. わからない
- IV. 信号(病変)は多分ない
- V. 信号(病変)は絶対ない

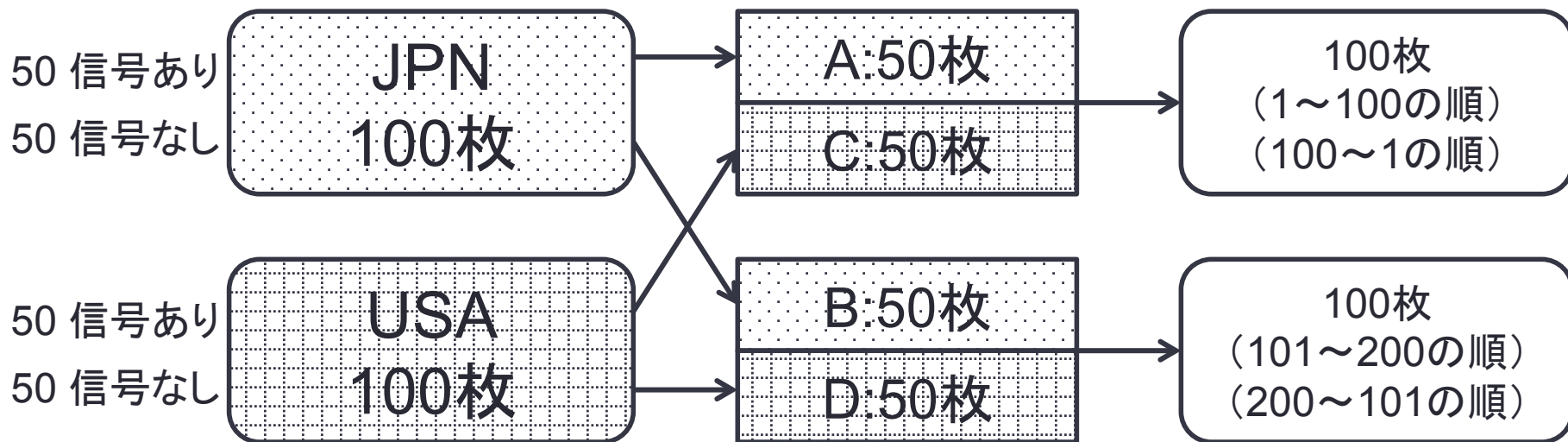
■  
観察者は各観察試料を判断基準ごとにカテゴリ  
( )にそれぞれ分類する

■  
観察者は各観察試料に対して、判断基準に応じたスコアをつける( )



信号(病変)なし

信号(病変)あり



を排除するための観察資料の分割例(8通り)

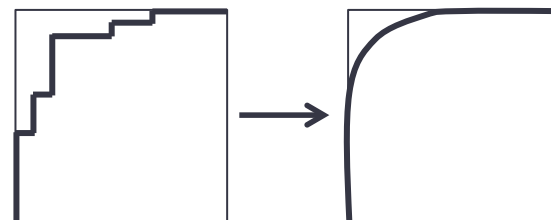
- Independent rating

2つのシステム評価実験を、  
十分な時間的間隔をあけて  
個々に行う

公開されているsoftwareを用いて  
解析・ROC曲線の作成を行う

- Sequential rating

2つのシステム評価実験を、  
時間的間隔をあけずに連続  
に行う



学習 → 実験開始 → データ収集 → ROC曲線  
解析 作成

# ROCカーブフィッティング

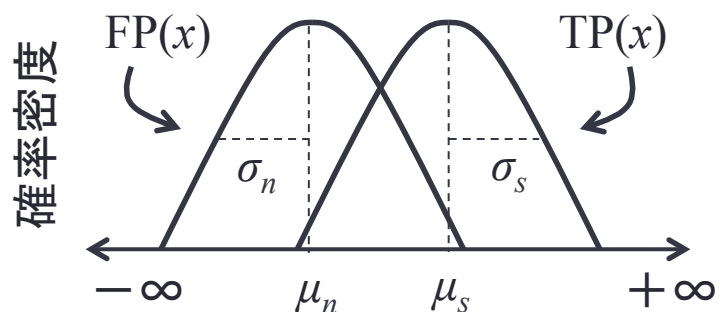
全試料のスコアから\_\_\_\_\_を推定し、  
そこからROC曲線を算出する一連の計算手順

| 資料ID | 観察者score | class    |
|------|----------|----------|
| 1    | 23       | negative |
| 2    | 9        | negative |
| 3    | 11       | positive |
| 4    | 75       | positive |
| ⋮    | ⋮        | ⋮        |
| ⋮    | ⋮        | ⋮        |
| ⋮    | ⋮        | ⋮        |
| 198  | 54       | positive |
| 199  | 39       | positive |
| 200  | 93       | negative |

⇒  
⇒  
 $\mu$ と $\sigma$ を  
推定

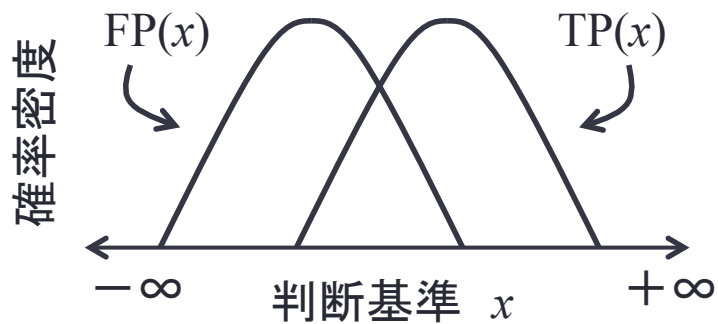
(最尤度比推定法)

両正規分布

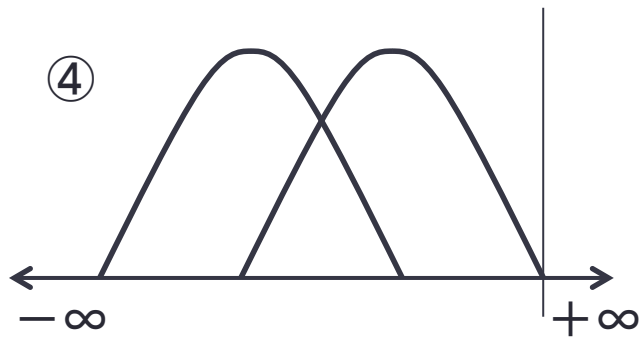
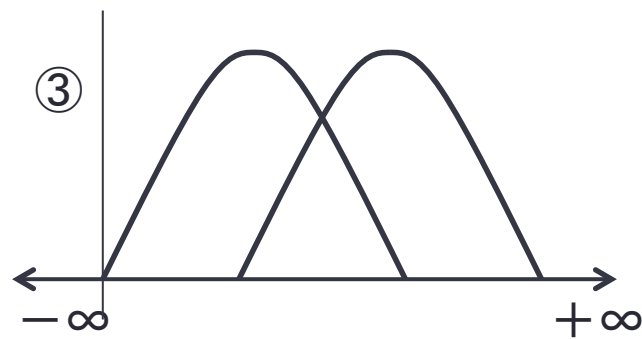
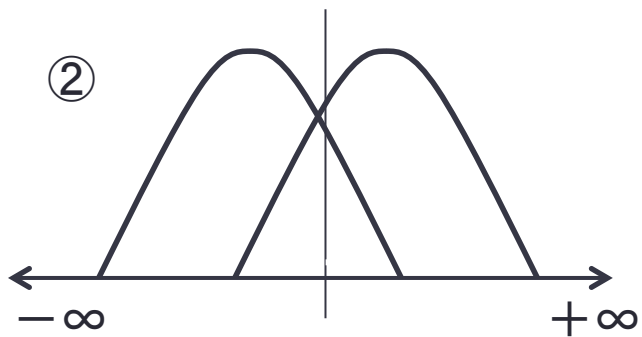
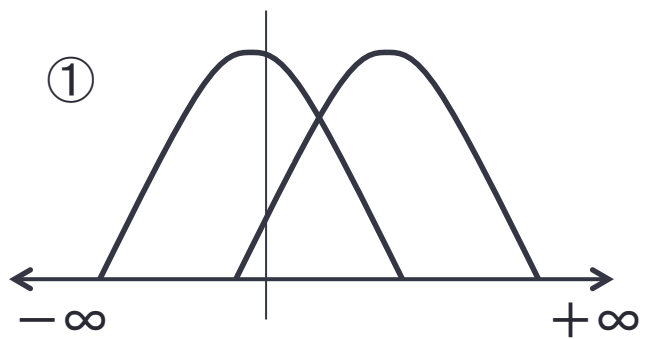
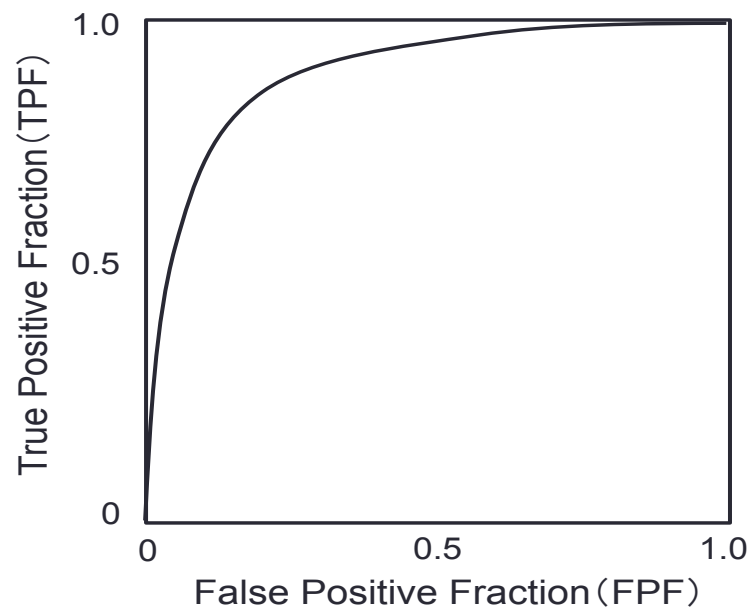


$$FP(x) = \frac{1}{\sqrt{2\pi}\sigma_n} \cdot e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}}$$
$$TP(x) = \frac{1}{\sqrt{2\pi}\sigma_s} \cdot e^{-\frac{(x-\mu_s)^2}{2\sigma_s^2}}$$

# 兩正規分布



$$FP(x) = \frac{1}{\sqrt{2\pi}\sigma_n} \cdot e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}}, \quad TP(x) = \frac{1}{\sqrt{2\pi}\sigma_s} \cdot e^{-\frac{(x-\mu_s)^2}{2\sigma_s^2}}$$

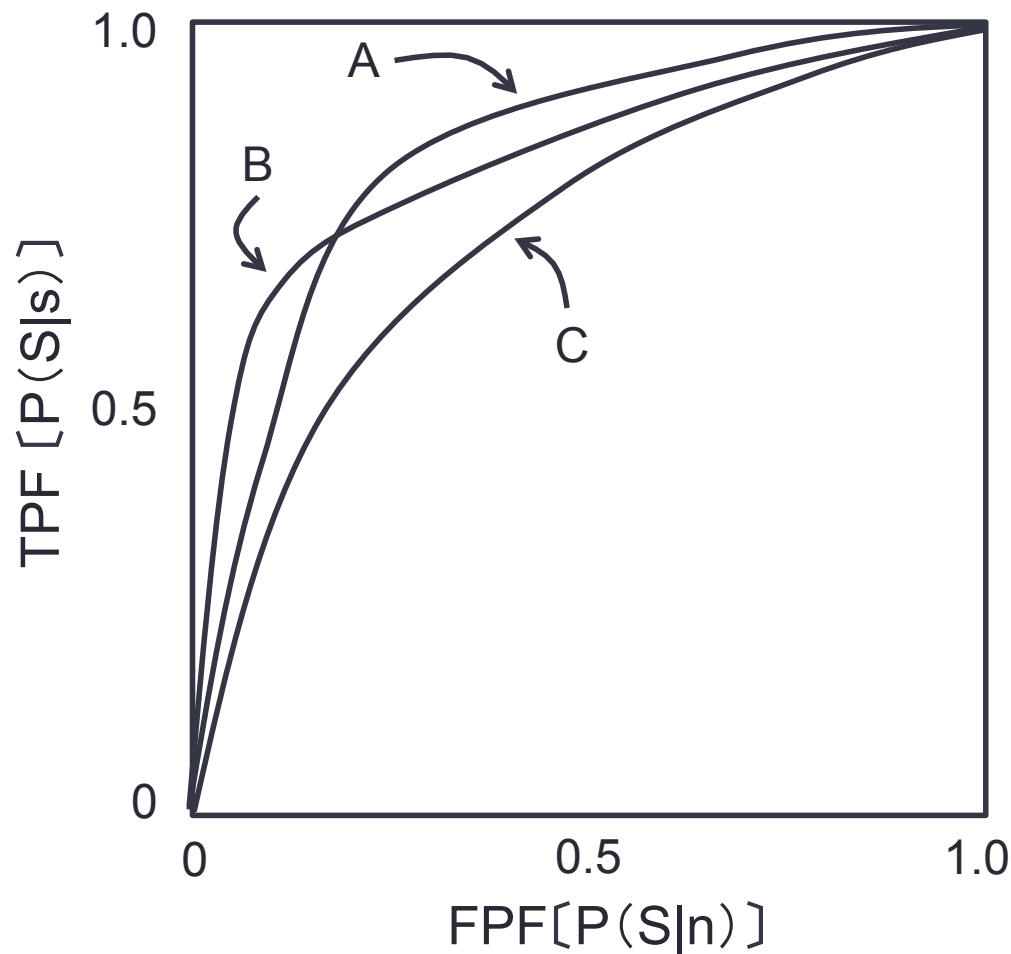


感度: \_\_\_\_\_

→  $P(S|s) + P(N|s) = 1$

特異度: \_\_\_\_\_

→  $P(N|n) + P(S|n) = 1$



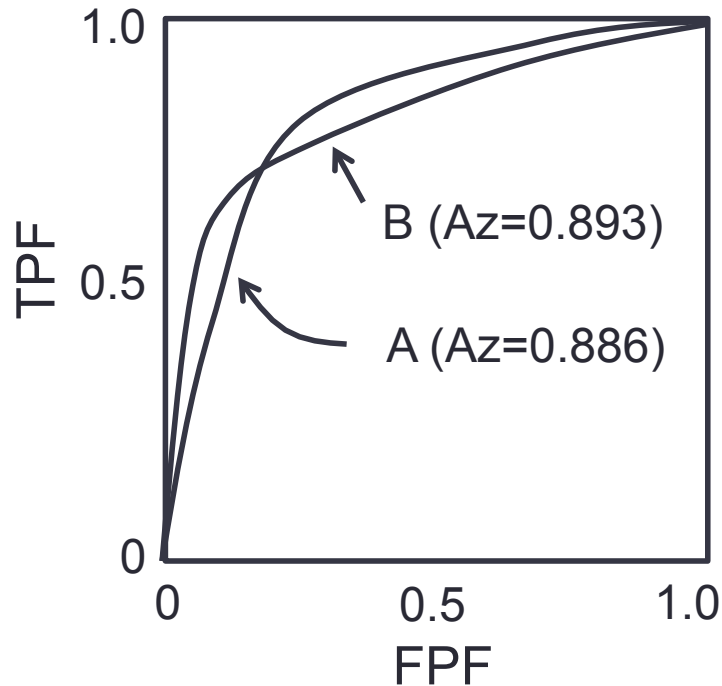
ROC曲線下の面積

Area Under ROC Curve

( )

→ 比較評価に使用  
最大値は \_\_\_\_\_

## statistical significance



ROC曲線間の比較では\_\_\_\_\_

- ・Azを用いる →
- ・各観察者の各試料のscoreを用いて三元配置分散分析を行う

帰無仮説 $H_0$ :「2つのシステム間に統計的な有意差はない」

対立仮説 $H_1$ :「2つのシステム間に統計的な有意差はある」

仮説 $H_0$ が棄却されれば, 対立仮説 $H_1$ が成り立つ



# その他の注意事項

- 観察者には\_\_\_\_\_が必ず必要
  - 実験の目的と実験方法を観察者に十分に理解してもらう
  - 回答の方法, 実験中に注意事項(読影時間, 再読影の禁止, など), 試料の構成(総枚数, 「信号+雑音」と「雑音」の画像のだいたいの割合, など)を説明したうえで, 学習用に用意された試料数枚を提示する
- 観察者の数は多いほどよい
  - 実験の目的に応じて, 十分な知識と経験をもった観察者を選ぶ → 観察者間の変動が少なくなるようにする
- 実験データの意図的な排除はダメ
  - いったん選択した観察者のデータを, その観察者の実験データが悪かったという理由で削除してはいけない

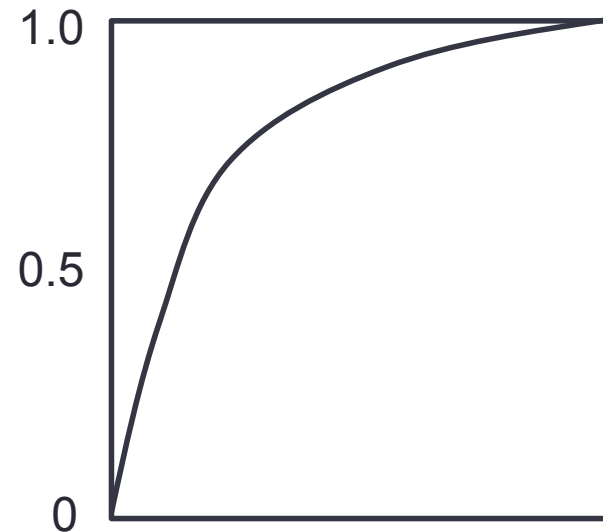
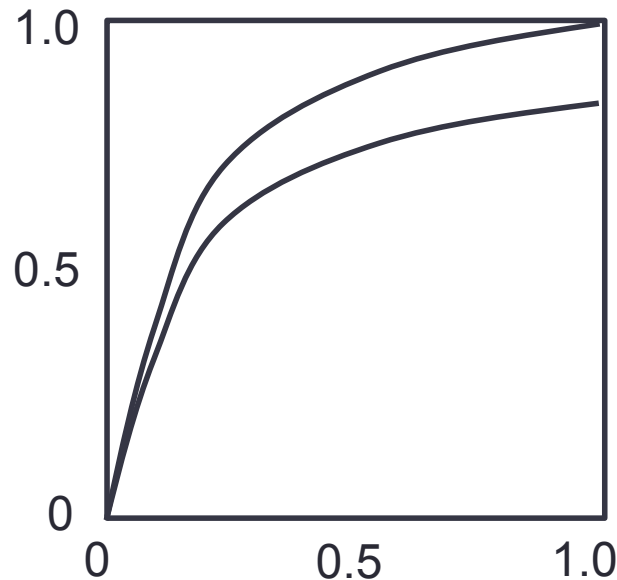


**ROC** → Receiver Operating Characteristic

**LROC** → Localization ROC

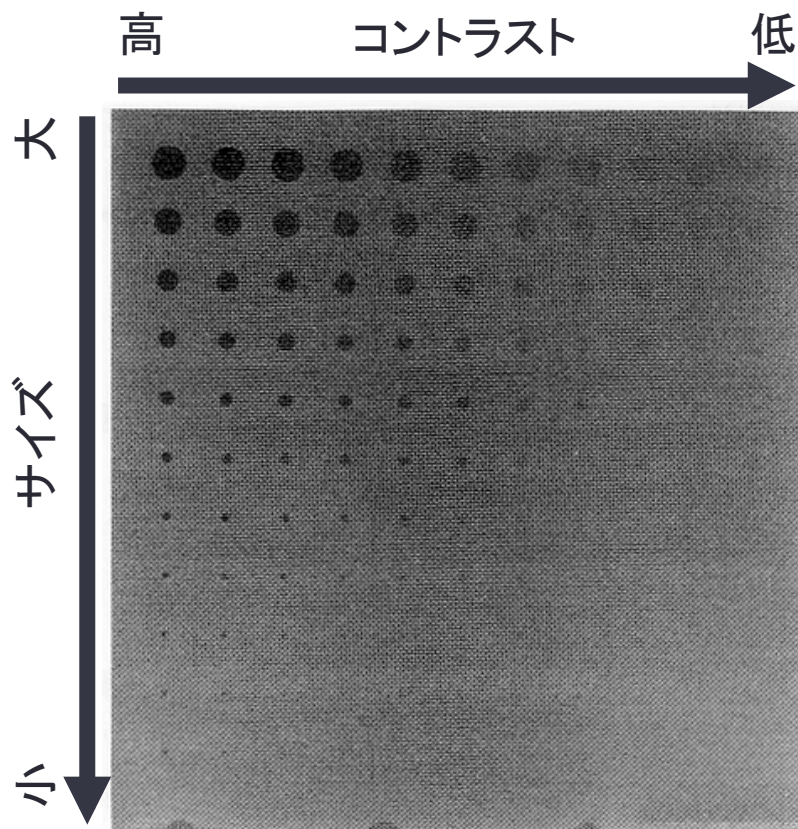
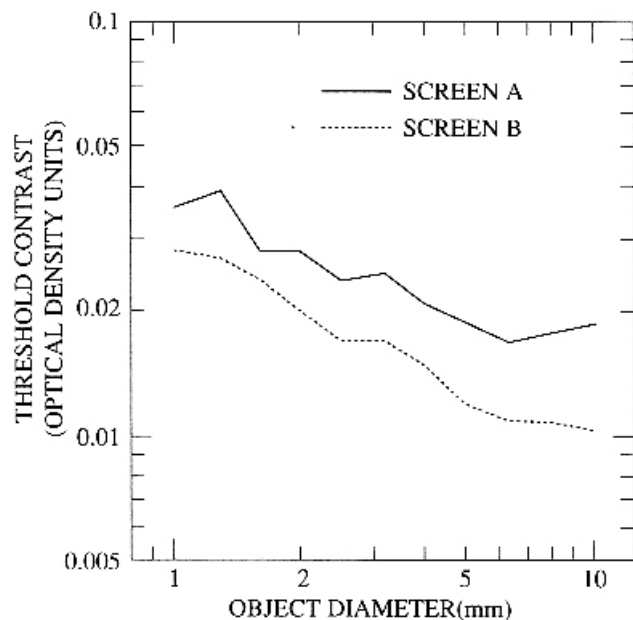
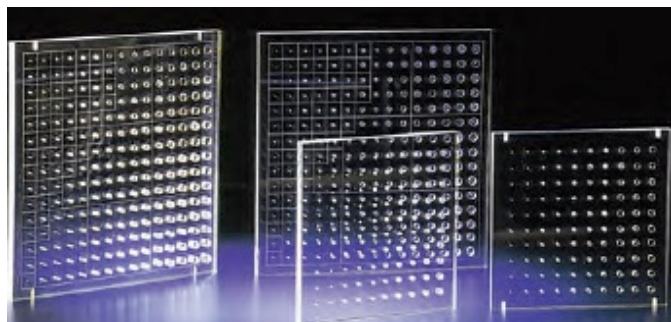
(ROC-type curve for task of detection and localization)

**FROC** → Free-response ROC



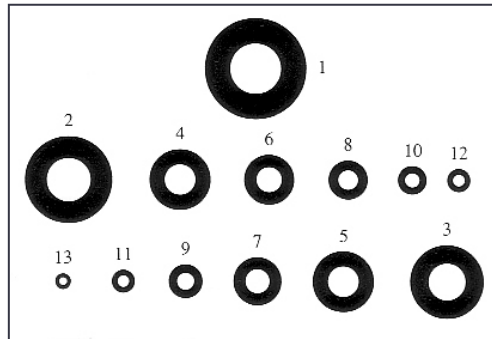
# 医用画像系に導入されている視覚評価

- （バーガーファントム法）

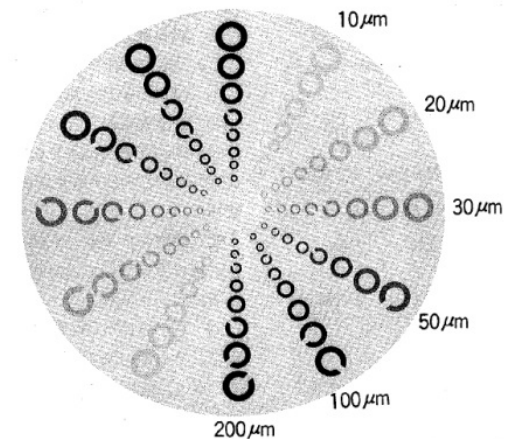


バーガーファントムのX線写真

## ・ハウレットチャート法



## ・ランドル環法



## の方法

### ・二点比較法

- AとBの2種類の試料を比較して、質問事項に当てはまるものを選択させる方法
  - 二点識別法, 二点嗜好法

### ・順位法

- 試料A, B, C, ... を同時に表示して、ある特性または嗜好について順位を付けさせる方法
  - ウィルコクソンの順位和検定, スピアマンの順位相関係数

- 試料A, B, C, ... を2個ずつ組み合わせて比較する方法

- シェッフェの一対比較の原法, シェッフェの一対比較の変法(芳賀, 浦, 中屋)

## 2005国家試験問題

ROC解析について誤っているはどれか.

1. 資料の難易度に応じてROC曲線は変動する
2. 臨床画像試料では病変が信号に対応する
3. ROC曲線下の面積の最大値は1である
4. 偽陽性率が1のとき真陽性率は0になる
5. 画像処理が診断能に与える効果を評価できる

## 2004国家試験問題

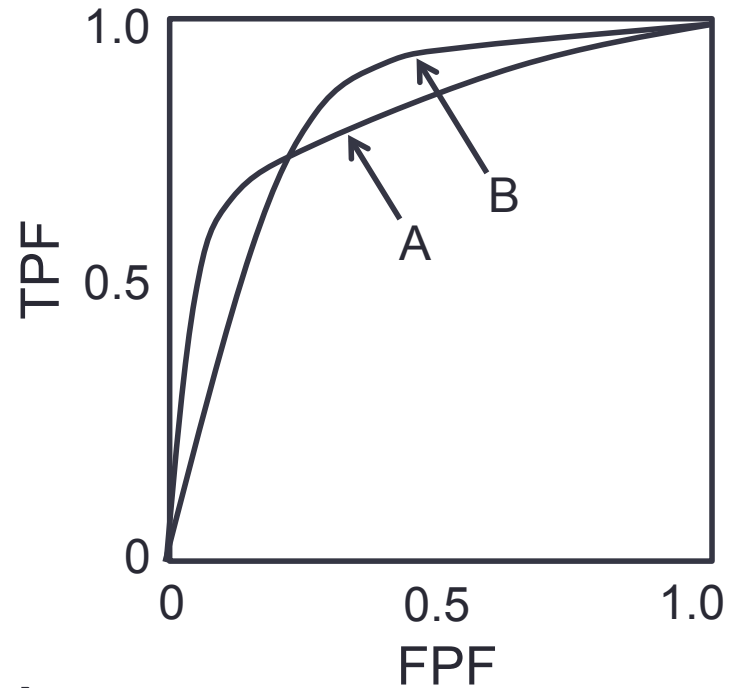
ROC解析について正しいのはどれか.

1. 解析結果は物理的評価に一致する
2. 真陰性の確率をグラフの横軸にとる
3. 試料枚数が少ないほど統計的な変動は小さい
4. MR画像とCT画像との病変検出能を比較できる
5. 雑音画像試料のみの観察でもROC曲線を描ける

ROC解析で誤っているはどれか. 2つ選べ.

1. ROC曲線の縦軸は真陰性である
2. ROC曲線の横軸は偽陽性である
3. 特異度は有病正診率と呼ばれている
4. 真陽性はtrue positiveである
5. 偽陰性はfalse negativeである

右に示す2本のROC曲線  
A, Bについて正しいはど  
れか.



1. 診断の正確さはBよりAの方が高い
2. 2つの検診システムの診断能を評価したROC曲線である場合, BよりもAの方が評価が高い
3. 感度が50%の時はAよりもBの方が特異度が高い
4. 信号像に対する反応の条件付確率密度関数の広がり  
はBよりもAの方が小さい
5. 特異度が90%の時はBよりもAの方が感度が高い

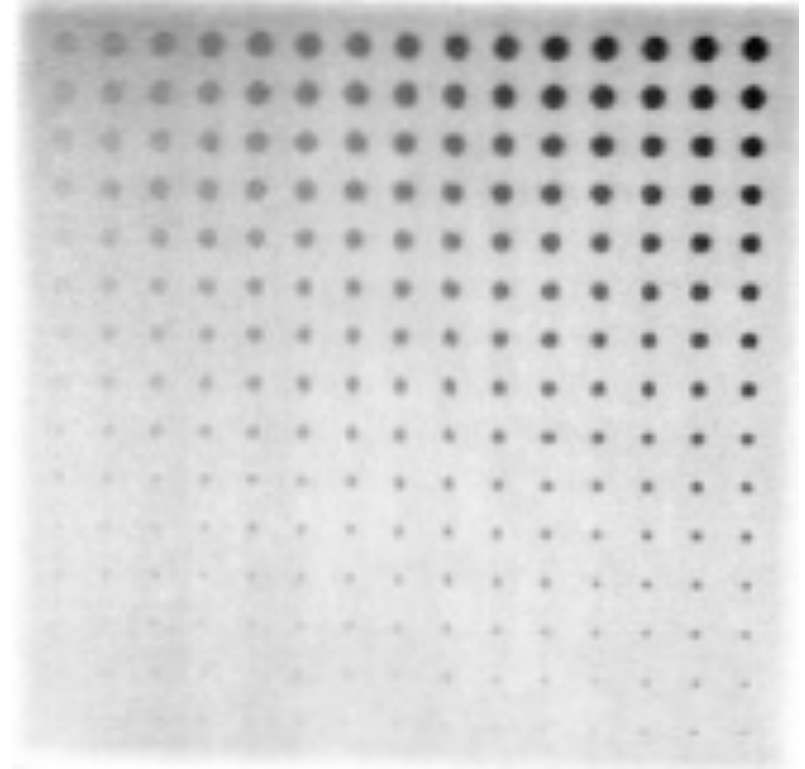
# ROC解析の実験目的として不適当なのはどれか.

1. ビーズ玉(2 mm  $\phi$ )を用いた増感紙の鮮鋭度の比較
2. 特定の疾患についてCRとMRIによる診断能の比較
3. ボケマスク処理の診断能に対する効果
4. マンモグラフィにおけるコンピュータ支援診断(CAD)システムの利用効果
5. 造影剤検査などの付加的因子の効果



## 2013国家試験問題

X線画像特性を調べる  
ファントム画像を示す。  
評価結果として得られるのはどれか



1. MTF
2. 特性曲線
3. ROC曲線
4. C-Dダイアグラム
5. ウィーナースペクトル