

# ROC 解析の基礎と 統計的有意差検定

画像セミナー

平成15年2月16日

画像研究委員会 小川 互

## 目次

1. はじめに .....	3
2. ROC 曲線の登場.....	4
3. ROC 曲線を描いてみよう .....	5
4. ROC 解析の実験方法 .....	6
5. ROC 曲線間の差の検定.....	7
1) 平均値の差の t 検定 <sup>7) 8) 10) 11)</sup> .....	10
a) 等しい母分散から得られたサンプルか? .....	10
b) t 値の計算をしよう .....	11
c) 平均値の差の検定.....	11
2) ジャックナイフ法 (Jackknife method) 観察者間変動と観察試料間変動を使った平均値の差の検定.....	12
a) ジャックナイフ法による新しいデータセットの作成方法.....	12
b) 新しいデータセットを使った <b>疑似値</b> の作成 .....	13
6. ROC 曲線に関する疑問.....	14
1) プール法とアベレージ法 <sup>3) 6)</sup> .....	14
2) カーブフィッティング法 <sup>2) 6)</sup> .....	14
参考文献.....	15
信号検出理論、ROC 解析について.....	15
統計的検定について .....	15

## 1. はじめに

ROC 解析はなぜ行われるのであろうか<sup>1) 2) 4) 6)</sup>。診断の精度を評価するため古くは正診率や感度と特異度などが用いられてきた。しかしながら正診率は疾患の分布に大きく依存して変化し、例えば正常50人、異常50人という疾患分布で正診率80%を得ようとする①正常、異常とも50人中40人(正診率=80%)の正診を得るか、または②正常に対して50人中50人(正診率=100%)、異常に対して50人中30人(正診率=60%)の正診を得るか、あるいは、これらの中である必要がある。ところが疾患分布が③正常90人、異常10人の分布になると適当に全員正常であると判断すれば正診率90%を得る。この場合正常に対しては90人中90人(正診率=100%)の正診、異常に対しては10人中0人(正診率=0%)の正診である。

$$\textcircled{1} \text{正診率} = \frac{40 + 40}{50 + 50} \times 100 = 80\%$$

$$\textcircled{2} \text{正診率} = \frac{50 + 30}{50 + 50} \times 100 = 80\%$$

$$\textcircled{3} \text{正診率} = \frac{90 + 0}{90 + 10} \times 100 = 90\%$$

感度とは異常症例に対して正しく異常と判断できる確率であり次の式で表せる

$$\text{感度} = \frac{\text{真陽性数}}{\text{異常症例数}}$$

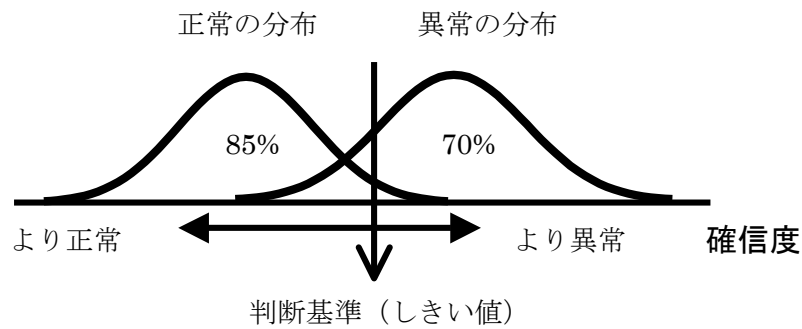
特異度とは正常症例に対して正しく正常と判断できる確率であり次の式で表せる

$$\text{特異度} = \frac{\text{真陰性数}}{\text{正常症例数}}$$

感度と特異度<sup>3) 4) 6)</sup>は判断基準を厳しくするか、あるいは緩くすることで変化する。例えば判断基準を緩くすれば感度は高くなるが、同時に第1種の誤り(誤報確率:異常がないのに異常と判断する誤り)が増えて特異度が低くなるであろう。逆に判断基準を厳しくすれば特異度は高くなるが第2種の誤り(ミス確率:異常があるのに正常と判断してしまう誤り)が増えて感度は低くなるであろう。すなわち  $TP + FN = 1$ 、 $FP + TN = 1$  の関係から、第2種の誤りFNが増えれば感度  $TP = 1 - FN$ 、第1種の誤りFPが増えれば特異度  $TN = 1 - FP$  となり、感度と特異度は低下する。

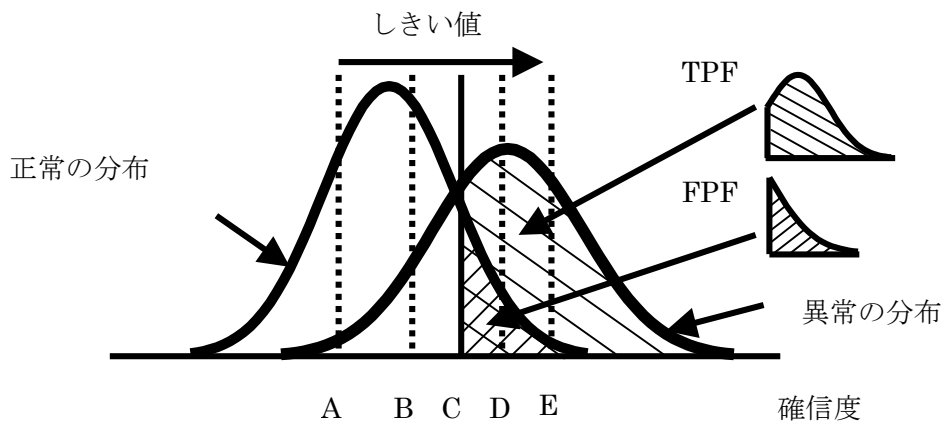
下図のようにある検査において、感度=70%、特異度=85%だったとして、判断基準を“より異常”側に移動させる(判断基準を厳しくする)と感度は低下するが特異度は増加する。

		判 断	
		異常 S	正常 N
症 例	異常 s	$P(S/s)$ TP 的中確率 感度	$P(N/s)$ FN ミス確率 第2種の誤り
	正常 n	$P(S/n)$ FP 誤報確率 第1種の誤り	$P(N/n)$ TN 正しい否定確率 特異度

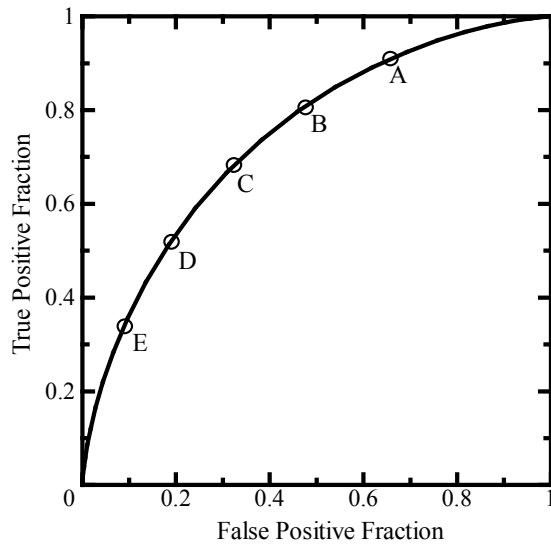


## 2. ROC 曲線の登場

判断基準(しきい値)を変化させながら、偽陽性率(FPF)に対する真陽性率(TPF)をプロットする。TPF=感度、FPF=1-特異度



判断基準 C の時の TPF と FPF は上図の通りである。この判断基準を ABCDE のように変化させると、次に示す ROC 曲線上をこの判断基準が移動するだけであり、ROC 曲線そのものは変化しない。この判断基準は critical value (臨界値、棄却限界値) と呼ばれる。判断基準を変化させても ROC 曲線そのものは変化しないが、ROC 曲線は判断基準をどの程度にすればよいかという基準を示すものではないことに注意が必要である。



### 3. ROC 曲線を描いてみよう

100 枚の異常症例と 100 枚の正常症例の写真を A,B ふたつのシステムにおいて撮影し、ある観察者が観察実験をおこなったところ次のような結果であったと仮定する。A のシステムについてのみ表を完成させておくので、皆さんは B のシステムについて計算してほしい。

(a) 100 枚の異常症例と 100 枚の正常症例を観察した結果。

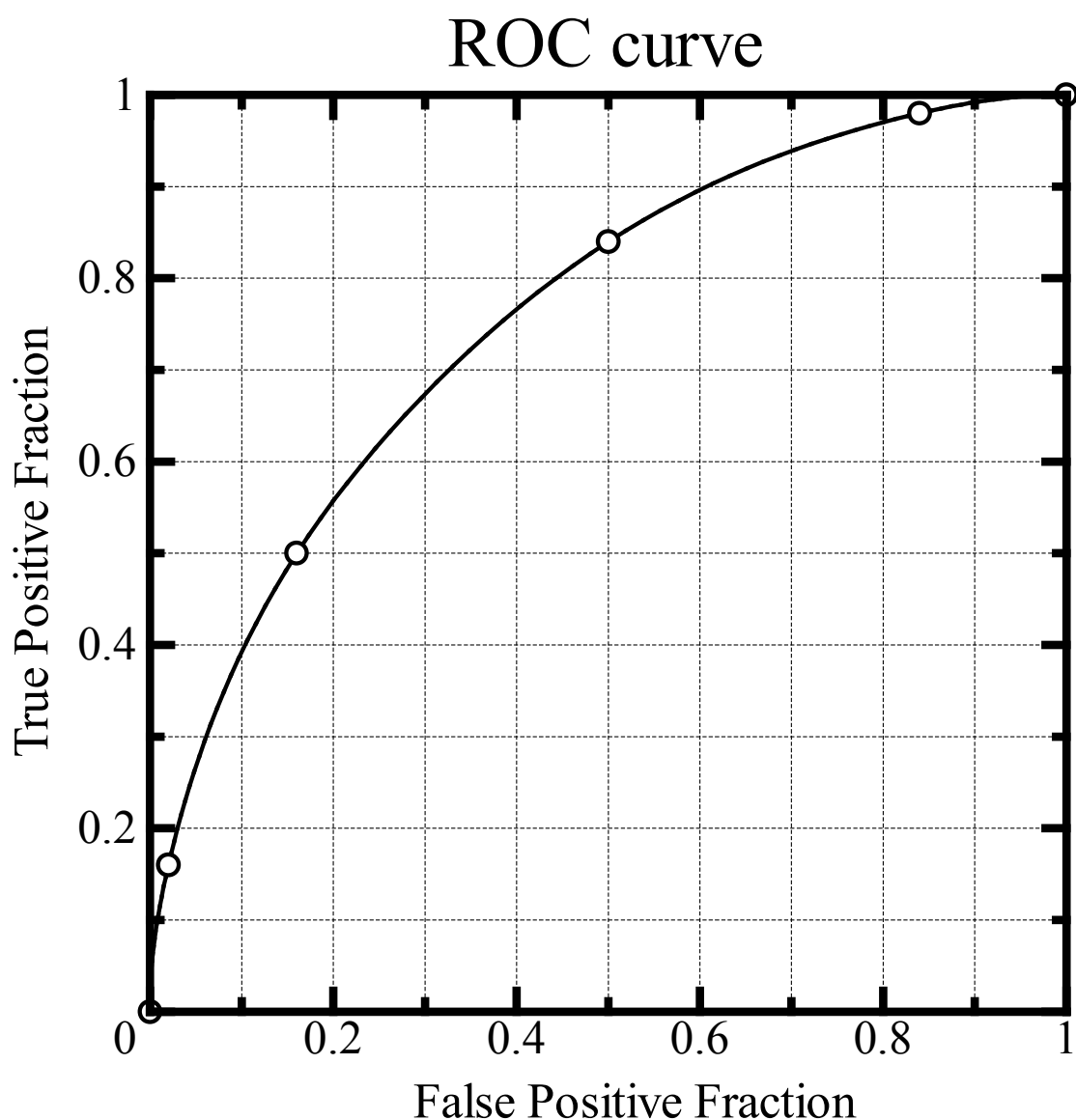
		絶対正常	たぶん正常	判らない	たぶん異常	絶対異常	総数
A	異常症例	2	14	34	34	16	100
	正常症例	16	34	34	14	2	100
B	異常症例	1	5	20	50	24	100
	正常症例	24	50	20	5	1	100

(b) (a)のデータを絶対異常の側から絶対正常の方向に累積していく。

		絶対正常	たぶん正常	判らない	たぶん異常	絶対異常
A	異常症例	100	84+14=98	50+34=84	16+34=50	16
	正常症例	100	50+34=84	16+34=50	2+14=16	2
B	異常症例					
	正常症例					

(c) (b)の各セルの値を総数で割って的中確率と誤報確率を求める。

		絶対正常	たぶん正常	判らない	たぶん異常	絶対異常	
A	異常症例	1.00	0.98	0.84	0.50	0.16	TPF
	正常症例	1.00	0.84	0.50	0.16	0.02	FPF
B	異常症例						TPF
	正常症例						FPF



#### 4. ROC 解析の実験方法

たとえどのような実験にしる、実験準備に十分な注意を払わないと、実験データそのものが無駄になってしまう事がある。一般的には実験計画が出来上がり準備が整った時点で実験の 9 割が終了しているといわれるほどである。ROC 解析の場合次のような手順で実験を進めることが良いとされている<sup>5) 9)</sup>。

##### ①実験目的と信号の決定

画像システム間の比較なのか、画像パラメータの比較なのか、画像処理法の比較なのかといったことを決定する。異なったモダリティ間の比較や、比較する因子を単独や、いくつか組み合わせるなども可能だが結果が導けるように注意が必要である。また鮮鋭度を評価するために最適な信号は何かといったことも併せて決定する必要がある。

例) 画像システム間で低コントラスト検出能を比較するために、ビーズを信号とした。

##### ②試料の作成および収集

ROC 解析に用いられる試料は臨床画像とファントム画像が挙げられる。臨床的な評価の場合臨床画像を用いるのが望ましいが、ランダムサンプリングで試料を多く収集することが困難であろう。またファントム画像の場合には信号としてあまりはっきりしたものは望ましくないで、この点は注意を要する。試料の枚数としては多ければ多いほど統計的な有意差検定において検出力が高くなるが、観察実験であるという点を考慮すると、疲労などの観点から(信号+雑音)100 枚(雑音)100 枚程度が望ましいとされる。しかし臨床画像で100 枚収集することは困難であろう。

### ③予備実験

試料の作成や収集の作業が終了する前に、最終的な実験結果を推測するための予備実験をおこなう。実験の手順や、実験計画の誤りなど、事前に確認できることが多く、無駄な本実験が行われるのを未然に防いでくれる。

### ④観察者の選択および学習

観察者も試料と同様に多ければ多いほど望ましいが、観察者間の変動を考慮すると少なくとも5名以上確保することが望ましい。特に臨床画像で観察実験を行う場合、信号の検出に対する十分な知識と能力を持っていなければならないので、診断に関する専門医が観察者になることが望ましい。しかし例えばビーズ玉の検出能を比較するような場合には放射線技師を採用しても特に問題は起こらないだろう。

学習は本実験の目的と実験方法を観察者に理解してもらうための必要不可欠な作業である。

### ⑤観察実験

例えば5段階の評定確信度がよく用いられているが、最近は連続確信度も用いられるようになってきた。観察条件や、観察の順序効果などに配慮して行う。また繰り返し実験を行う場合には最低1ヶ月くらいの間隔をあけて行うようにする。

### ⑥カーブフィッティング

例えば5段階の評定確信度で行った場合にはグラフにプロットした後自在定規などで滑らかに点をつなげる方法がある。またシカゴ大学から公開されているROCKITなどのROC解析ソフトを利用する方法がある。連続確信度の場合には度数分布表を作るか、解析ソフトを利用しないとカーブフィッティングが行えない。

### ⑦統計的有意差検定

カーブフィッティングにより得られたROC曲線を求めてもこれだけではROC解析とは言い難い。①例)のように画像システム間の比較を行うのであれば、ROC曲線下の面積 $A_z$ を用いて、t検定や、ジャックナイフ法を行って有意差を検定する必要がある。

$A_z$ を用いて検定を行う場合には比較するROC曲線が交差しているかどうかで検定の意味が異なってくる。したがってROC曲線は必ず描いて確認する必要がある。

## 5. ROC 曲線間の差の検定

さてここで異常症例10枚と正常症例10枚の写真をSF,CRふたつのシステムで撮影し、A1,A2,A3,A4,A5,A6,の6名の観察者が観察実験をおこなったら次のような結果が得られたと仮

定する。

判断結果のカテゴリーは表を簡単にするため次のように取り決める。データは正規乱数を利用して作成した。表中の数値は次のカテゴリーを示している。

- 5: 絶対異常
- 4: たぶん異常
- 3: 判らない
- 2: たぶん正常
- 1: 絶対正常

SF: Negative Case M=2,SD=1, Positive Case M=3,SD=1

CR: Negative Case M=2,SD=1, Positive Case M=3.5,SD=1

正規乱数によって作成されたサンプルデータセット

A1	SF	異常症例										Az
		3	3	2	4	2	2	3	2	2	2	
	CR	正常症例										0.75
		3	1	2	2	1	2	1	1	2	1	
A2	SF	異常症例										Az
		2	5	3	2	4	4	2	3	3	2	
	CR	正常症例										0.81
		2	3	3	1	2	3	2	3	3	3	
A3	SF	異常症例										Az
		2	3	4	2	4	2	4	4	2	5	
	CR	正常症例										0.78
		2	1	1	2	2	3	3	2	2	2	
A3	SF	異常症例										Az
		2	3	4	2	4	2	4	4	2	5	
	CR	正常症例										0.78
		2	1	1	2	2	3	3	2	2	2	
A3	SF	異常症例										Az
		2	3	4	2	4	2	4	4	2	5	
	CR	正常症例										0.78
		2	1	1	2	2	3	3	2	2	2	
A3	SF	異常症例										Az
		2	3	4	2	4	2	4	4	2	5	
	CR	正常症例										0.78
		2	1	1	2	2	3	3	2	2	2	
A3	SF	異常症例										Az
		2	3	4	2	4	2	4	4	2	5	
	CR	正常症例										0.78
		2	1	1	2	2	3	3	2	2	2	

A4	SF	異常症例										Az
		1	4	2	4	1	3	3	4	4	5	
		正常症例										0.71
	1	1	2	2	1	2	4	2	2	4		
	CR	異常症例										Az
		4	2	5	4	2	4	3	3	3	4	
正常症例										0.82		
3	3	1	2	3	2	2	3	1	2			
A5	SF	異常症例										Az
		2	3	3	2	3	2	2	3	3	3	
		正常症例										0.65
	1	2	1	1	2	4	1	3	4	1		
	CR	異常症例										Az
		2	3	5	3	4	1	4	5	5	3	
正常症例										0.77		
3	1	2	2	2	2	3	1	2	4			
A6	SF	異常症例										Az
		4	4	3	2	2	3	2	3	2	2	
		正常症例										0.64
	3	2	3	1	2	3	2	3	1	2		
	CR	異常症例										Az
		2	3	3	5	1	4	2	3	3	3	
正常症例										0.84		
1	2	1	3	2	1	1	2	2	1			

A1～A6の観察者におけるSFシステムとCRシステムのAzの差をまとめると

観察者	Az <sub>CR</sub>	Az <sub>SF</sub>	Az <sub>CR</sub> - Az <sub>SF</sub>
A1	0.81	0.75	0.06
A2	0.75	0.78	-0.03
A3	0.78	0.78	0
A4	0.82	0.71	0.11
A5	0.77	0.65	0.12
A6	0.84	0.64	0.20
標本平均	0.795 ( $\bar{x}$ )	0.718 ( $\bar{y}$ )	0.07667 ( $\bar{x} - \bar{y}$ )
標本標準偏差: $s$	0.0339 ( $s_x$ )	0.0624 ( $s_y$ )	
標本分散: $s^2$	0.00115 ( $s_x^2$ )	0.00389 ( $s_y^2$ )	
プールされた分散: $s_p^2$			0.00252 ( $s_p^2$ )
平均間の差の標準誤差: $s.e._{\bar{x}-\bar{y}}$			0.02900 ( $s.e._{\bar{x}-\bar{y}}$ )

### 1) 平均値の差の t 検定 7) 8) 10) 11)

まず CR と SF の Az の平均値に差があるかどうか、一般的な t 検定をおこなってみよう。

#### a) 等しい母分散から得られたサンプルか？

t 検定では母分散  $\sigma_x^2 = \sigma_y^2$  の場合と  $\sigma_x^2 \neq \sigma_y^2$  の場合とでやり方が異なる。そこで母分散の比の検定をおこなって  $\sigma_x^2 = \sigma_y^2$  かどうかを調べよう。

帰無仮説 (null hypothesis) は

$$H_0: \text{母分散は等しい。} \sigma_x^2 = \sigma_y^2$$

対立仮説 (alternate hypothesis) は

$$H_A: \text{母分散は等しくない。} \sigma_x^2 \neq \sigma_y^2$$

である。ここでフィッシャーの分散比を計算すると (※計算は比が 1 以上になるようにする)

$$F = \frac{s_y^2}{s_x^2} = \frac{0.00390}{0.00115} = 3.39$$

帰無仮説が正しいとすると  $F$  は自由度 ( $m-1, n-1$ ) の  $F$  分布に従うので表より  $F_{0.025}(5,5)=7.15$  (ここで  $m, n$  は観察者 A1-A6 の 6 である)。すなわち母分散が等しい集団からのサンプルであるとするならば自由度 (5,5) において標本分散の比が 7.15 を超える確

率は5%程度しかないことを示している。 $F=3.39 < 7.15$ なので帰無仮説は棄却されず等しい母分散からサンプリングされた標本であるとみなせる。

## b) t 値の計算をしよう

t 値は次の式で計算される

$$t = \frac{\bar{x} - \bar{y}}{s.e._{\bar{x}-\bar{y}}} = \frac{0.07667}{0.02900} = 2.644$$

ここで  $s.e._{\bar{x}-\bar{y}}$  は次のようにして求める

$$s_p = \sqrt{\frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}} = \sqrt{\frac{5 \times 0.00115 + 5 \times 0.00389}{6+6-2}} = \sqrt{\frac{0.02520}{10}} = \sqrt{0.00252}$$

$$= 0.05023$$

このとき  $s_p$  はプールされた分散の平方根である。

$$s.e._{\bar{x}-\bar{y}} = s_p \sqrt{\frac{1}{m} + \frac{1}{n}} = 0.05023 \sqrt{\frac{1}{6} + \frac{1}{6}} = 0.05023 \sqrt{0.33333} = 0.05023 \times 0.57735$$

$$= 0.02900$$

## c) 平均値の差の検定

観察者 A1~A6, 6 人による CR システムの平均値  $\bar{x}$  (0.795) と SF システムの平均値  $\bar{y}$  (0.718) に差があるかどうかを調べよう。

帰無仮説は

$$H_0: \text{平均は等しい。} \bar{x} = \bar{y}$$

対立仮説は両側検定でおこない

$$H_A: \text{平均は等しくない。} \bar{x} \neq \bar{y}$$

帰無仮説が正しいとすると b) で計算した t 値は自由度  $m+n-2(=10)$  の t 分布に従うので、表より  $t_{0.025}(10)=2.228$ 。これは 平均値の差 ( $\bar{x} - \bar{y}$ ) を 平均値の差の誤差 ( $s.e._{\bar{x}-\bar{y}}$ ) で割った値が 2.228 を超える確率は 5%程度しかないことを示している。すなわち t 値が 2.228 を超えた場合、平均値は等しいという帰無仮説を支持できる確率は 5%未満だということである。 $t=2.644 > 2.228$  なので帰無仮説は棄却され、したがって平均値は等しくない。

## 2) ジャックナイフ法 (Jackknife method) 観察者間変動と観察試料間変動を使った平均値の差の検定

t 検定では観察者間の変動から標準誤差を求めて平均値の差の検定に用いた。ジャックナイフ法<sup>8)10)</sup>は、観察試料間の変動を擬似的に創りだすリサンプリング法のひとつである。

### a) ジャックナイフ法による新しいデータセットの作成方法

ジャックナイフ法では、観察者間の変動と観察試料間の変動が平均値の差の検定に用いられる。したがって検定で得られた結果は観察者間と観察試料間の変動が差し引かれた、純粋なシステム間の有意差であるということが出来る。ここでは次のようなデータ形式を扱うことになる。ジャックナイフ法で扱われるデータ形式(表内のデータは例である)。

観察者	CR					SF				
	B1	B2	...	B9	B10	C1	C2	...	C9	C10
A1	0.9	0.8		0.9	0.8	0.7	0.7		0.8	0.9
A2	0.8	0.7		0.7	0.7	0.5	0.7		0.6	0.6
A3	0.9	0.8		0.8	0.9	0.6	0.8		0.5	0.5
A4	0.8	0.9		0.7	0.6	0.8	0.6		0.7	0.7
A5	0.9	0.7		0.6	0.7	0.9	0.9		0.8	0.8
A6	0.8	0.8		0.7	0.8	0.6	0.5		0.6	0.5

今回のデータは試料数が 10 枚であり、ジャックナイフ法では、この 10 枚の試料から新たに B1~B10 と C1~C10 までのデータをリサンプリングにより創り出す。CR と SF の試料に対してそれぞれ 1~10 の番号をふって、それぞれ CR1~CR10 (CR による試料)、SF1~SF10 (SF による試料) によるリサンプリング法を次に示す。アンダーバーはデータが無い。

例) 観察者 A1 のケース

B1: \_\_\_\_, CR2, CR3, CR4, CR5, CR6, CR7, CR8, CR9, CR10

B2: CR1, \_\_\_\_, CR3, CR4, CR5, CR6, CR7, CR8, CR9, CR10

• •

• •

• •

B9: CR1, CR2, CR3, CR4, CR5, CR6, CR7, CR8, \_\_\_\_, CR10

B10: CR1, CR2, CR3, CR4, CR5, CR6, CR7, CR8, CR9, \_\_\_\_

C1: \_\_\_\_, SF2, SF3, SF4, SF5, SF6, SF7, SF8, SF9, SF10

C2: SF1, \_\_\_\_, SF3, SF4, SF5, SF6, SF7, SF8, SF9, SF10

• •

• •

• •

C9: SF1, SF2, SF3, SF4, SF5, SF6, SF7, SF8, \_\_\_\_, SF10

C10: SF1, SF2, SF3, SF4, SF5, SF6, SF7, SF8, SF9, \_\_\_\_

これを観察者 A1~A6 についておこなって新たにデータセットを作成する。ジャックナイフ法はリサンプリング法の中でも比較的単純で、観察試料が n 枚ならば n-1 枚の観察試料で順次 n 組の観察試料を創り出す。このように観察者 A1~A6 と B1~B10 ならびに C1~C10 の多重比較により CR システムと SF システムの差の検定をおこなう。このようなケースには三元配置分散分析<sup>10)</sup>が利用される。三元配置の場合には A,B の交互作用、B,C の交互作用、C,A の交互作用の 2 因子交互作用のほか、3 因子交互作用である A,B,C の交互作用の考慮される場合があるが、ROC 曲線の検定では 3 因子交互作用は考慮されていない。分散分析は各因子間の分散比 (F 値) を検定するのであるが、平均の差の検定であることに注意が必要である。

b) 新しいデータセットを使った疑似値の作成

それでは実際に検定に用いる疑似値の作成方法をみてみよう。まず観察者 A1 について、データセット B1 を用いた疑似値 YB<sub>11</sub> と、データセット C1 を用いた疑似値 YC<sub>11</sub> を求めてみよう。

元のデータセット

A1	SF	異常症例										Az
		3	3	2	4	2	2	3	2	2	2	
		正常症例										0.75
		3	1	2	2	1	2	1	1	2	1	
	CR	異常症例										Az
		4	5	5	4	4	3	2	3	3	4	
正常症例										0.81		
2		3	3	1	2	3	2	3	3		3	

ジャックナイフ法により作成されたデータセット

A1	B1	異常症例										Az
		3	2	4	2	2	3	2	2	2		
		正常症例										0.78
		1	2	2	1	2	1	1	2	1		
	C1	異常症例										Az
		5	5	4	4	3	2	3	3	4		
正常症例										0.77		
3		3	1	2	3	2	3	3	3			

$$\begin{aligned} YB_{11} &= 10A_{ZCR} - (10-1)A_{ZB1} \\ &= 10 \times 0.75 - 9 \times 0.78 \\ &= 0.48 \end{aligned}$$

$$\begin{aligned} YC_{11} &= 10A_{ZSF} - (10-1)A_{ZC1} \\ &= 10 \times 0.81 - 9 \times 0.77 \\ &= 1.17 \end{aligned}$$

この計算を観察者数 6, SF の試料枚数 10, CR の試料枚数 10, について全て計算しジャックナイフ法で扱われるデータ形式の表にまとめ検定をおこなう。今回の場合  $6 \times (10+10) = 120$  個の疑似値を計算する。この疑似値は各試料間の変動から ROC 曲線の変動に直接関係ある値であり、かつ統計的にも独立した値であるとされる。したがって、この疑似値を用いてシステム間と観察者間ならびに観察試料間の変動を分析すればよい。また今回は分散分析の説明は省略する。(D.D.Dorfman らが LABMRMC を無償配布している)

## 6. ROC 曲線に関する疑問

### 1) プール法とアベレージ法<sup>3) 6)</sup>

例えば画像システムを比較するために 6 名の ROC 評定が得られたとしよう。このときシステムを比較するために代表となる ROC 曲線はどのようにして得ればよいのであろうか。これには 2 種類の方法が知られている。

- ①プール法 : 6 名の評定値(1~5)を全て平均し、プールされた評定値で ROC 曲線を作成する方法。
- ②アベレージ法 : 6 名の評定値それぞれで 6 本の ROC 曲線を作成し、横軸(FPF)に対して例えば 0.1 間隔で縦軸(TPF)の値を平均して ROC 曲線を作成する方法。

一般にはほぼ同様の結果が得られるが、プール法で得られた ROC 曲線は真の ROC 曲線よりも悪くなる傾向を示すので、アベレージ法を用いることが推奨されている。

### 2) カーブフィッティング法<sup>2) 6)</sup>

評定確信度については手計算も可能ですが、シカゴ大学で配布されている ROCKIT は最尤推定アルゴリズムを用いて両正規確率紙の標準偏差軸(ND 軸)上で a と b の値を推定しています。連続確信度のデータについてはカテゴリー分けして(度数分布表にしてヒストグラムを描くのと同じ操作)評定確信度で得られた形式と同じにし、同様の方法で a と b の値を推定しています。

異常と正常症例の評定結果の平均と標準偏差から両正規の ROC 曲線を描く方法は、一般には用いられていません。ただしサンプル数が多い場合には有効な方法です。また最尤推定アルゴリズムの他にも最小二乗法やその他の適合法を利用することは可能です。

## 参考文献

### 信号検出理論、ROC 解析について

- 1) 山下一也, 若松孝司: 信号検出理論のX線撮影系への適用. 日放技学誌, 33(2), 111-117, (1977).
- 2) Charles E. Metz, 訳: 畑川政勝, 他: ROC 解析の基礎. 日放技学誌, 46(6), 831-840, (1990).
- 3) 山下一也: 信号検出理論による画質評価. 日放技学誌, 48(12), 2081-2092, (1992).
- 4) 藤田広志, 他: ROC 解析の基礎と最近の進歩. 日放技学誌, 49(9), 1685-1703, (1993).
- 5) 白石順二: 診断能の評価—ROC 解析の実験方法—. 日放技学誌, 55(4), 362-368, (1999).
- 6) 藤田広志, 他: ROC 解析の基礎と応用. 日放技学会放射線医療技術学叢書(8), (1994).

### 統計的検定について

- 7) Sylvia Wassertheil-Smoller: Biostatistics and Epidemiology(2nd edition). Springer-Verlag, 34-75, (1995).
- 8) 白石順二, 宇都宮あかね: ROC 解析における画像システム間の統計的有意差の検定方法—Jackknife 法とその適用—. 日放技学誌, 53(6), 691-698, (1997).
- 9) 佐伯 胖, 松原 望: 実践としての統計学. 東京大学出版会, 147-156, (2001).
- 10) 松原 望: 統計の考え方. 放送大学教育振興会, 157-182, 215-250, (2000).
- 11) 新濃清志: 一般数学. 森北出版株式会社, 161-178, (1984).