

## 原 著

平均および標準偏差補正による連続データに対する  
ROC曲線の適合

小川 互・中谷儀一郎・唐沢 宏

城西放射線技術専門学校

論文受付  
2001年4月16日論文受理  
2003年10月8日

Code No. 520

## 緒 言

ROC解析を行う方法には、評定確信度法と連続確信度法がよく用いられているが、適用範囲の広さと測定の簡便さから連続確信度法の利用されるケースが多くなっている<sup>1,2)</sup>。連続確信度法によるROC曲線の適合<sup>3)</sup>には、連続データをいくつかのカテゴリーに分割して評定確信度法に用いる方法を適用するものと、陰性と陽性のデータの平均と標準偏差を検定し、それを両正規ROCパラメータとして曲線の適合を行う方法が知られている。観察実験から得られるデータは正規分布にはならないことが多いため、平均と標準偏差を両正規ROCパラメータとする曲線適合法は通常行われない。

ROC解析では観察者が信号と雑音を検出する場合の評定基準はともに正規分布であるという仮定がある。しかしながら限られた数の観察資料では評定基準の正

規性を満足し観察実験の結果が正規分布となる保証はない。もし仮に観察実験が無制限行われた場合、その分布は正規分布になるであろうから、この場合には正規分布の裾が0だけ切り取られているとする。通常の観察実験は回数が限られているので、この場合には正規分布の裾が0以上切り取られている。すなわち実験回数が多ければデータ分布は正規分布となるであろう、がしかし実験回数が限られておりデータ数の少ない場合によってはデータ分布が正規分布になることもあるだろうし、また裾の打ち切られた正規分布になることもあるであろう。これを一般化して観察実験で得られたデータの分布は正規分布の裾が0以上切り取られていると仮定する。

われわれはROC解析に用いられる陰性と陽性のデータに対し上述のような仮定のもとに平均と標準偏差の補正を行うことで良好なROC曲線の適合結果を得るこ

## Fitting of ROC Curves for Continuous Data by Using Correction of the Mean and Standard Deviation

WATARU OGAWA, GIICHIROH NAKAYA, and HIROSHI KARASAWA

Josai Radiological Technical College

Received April 16, 2001; Revision accepted Oct. 8, 2003; Code No. 520

## Summary

We examined ROC curve fitting with continuous rating data by using a corrected mean value and corrected standard deviation as binormal ROC parameters. Curve fitting was performed by using corrected categorical data in the same way as the continuous data. A large residual of curve fitting was found in cases of difference between means of positive and negative signal data greater than three times the standard deviation, or with a ratio of the standard deviation of positive and negative signal data above six. We conclude that our proposed method performed well for ROC curve fitting except in some special cases.

**Key words:** Receiver operating characteristic (ROC) curve fitting, Image evaluation, Statistical analysis, Signal detection theory, Bootstrap method

別刷資料請求先：〒171-0044 東京都豊島区千早1-10-26  
城西放射線技術専門学校 小川 互 宛

とができた．さらにブートストラップ法<sup>4-6)</sup>によるシミュレーションを行い，試料間のばらつきを考慮したROC曲線と平均値およびAzなどのROCパラメータの標準誤差を推定することができたので報告する．

1. 理論

1-1 カテゴリカル・データの修正

通常行われる連続確信度法のデータはカテゴリー化できない離散データの形式になっているものが多い．われわれは連続確信度法のデータ形式を拡張することにより，カテゴリカル・データも連続確信度法により得られたデータの特長なケースであると考え，カテゴリカル・データに対し以下のような修正を行った．Fig. 1のようにカテゴリー数を5とした場合，カテゴリーの幅  $h$  は1としてカテゴリー化するのが一般的である．このように連続データをカテゴリー化すると平均値はカテゴリーの幅の大きさによって影響を受けることは少ないが，分散はかなりの影響を受ける．そこで連続分布  $F$  を離散分布  $G$  で近似する場合の修正として知られているシェパードの修正 (Sheppard's correction  $\gamma$ ) を若干変形し，カテゴリー数:  $N_c$ ，連続分布の標準偏差:  $\sigma_F$ ，離散分布の標準偏差:  $\sigma_G$  とした場合，

$$\sigma_F = \sqrt{\sigma_G^2 - \frac{1}{N_c}} \dots\dots\dots (1)$$

(1) 式のような修正を適用することでカテゴリカル・データを連続確信度法のデータと同様に扱った．この修正では  $N_c = \infty$  ならば  $\sigma_F = \sigma_G$  となる．

1-2 標準偏差の補正

正規分布の裾が打ち切られることにより標準偏差は過小評価される．Fig. 2に示すように平均0，標準偏差1の標準正規分布が  $1\sigma$  で打ち切られると標準偏差は0.857， $2\sigma$  で打ち切られると標準偏差は0.990のように過小評価される．標本  $X$  の標本平均を  $m_s$ ，標本標準偏差を  $\sigma_s$  とする．このとき標本  $X$  の最大値  $X_{max}$  が  $m_s + u\sigma_s$  であった場合，この標本  $X$  は標本平均  $m_s$  標本標準偏差  $\sigma_s$  で与えられる正規分布に対して  $u\sigma_s$  で打ち切られているとする．このとき  $\sigma_s$  は標準正規分布に正規化した場合の  $\sigma$  であるから，その上側確率を  $u\sigma$  から求め1から差し引いたものを上限 (upper) の値とする．標本  $X$  の最小値  $X_{min}$  についても同様に  $-u\sigma$  を求め，これを  $u\sigma$  と読み替えて，その上側確率を  $u\sigma$  から求め1から差し引いたものを下限 (lower) の値とする．これを(2) 式に代入して補正量  $\sigma_{cq}$  を得る．

$$\sigma_{cq} = upper + lower - 1 \dots\dots\dots (2)$$

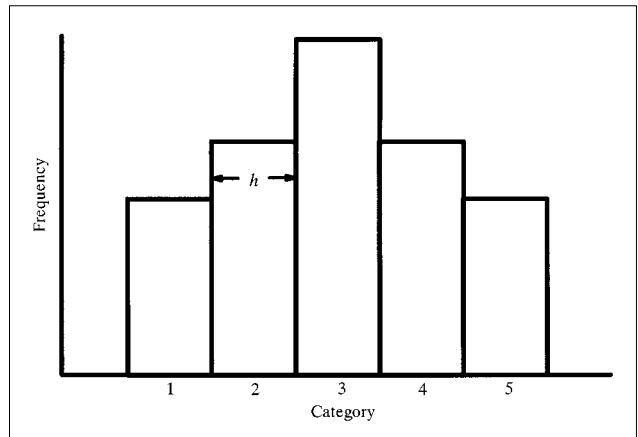


Fig. 1 Histogram for explaining Sheppard's correction. Correction of categorical data.

$$\sigma_F = \sqrt{\sigma_G^2 - \frac{1}{N_c}}$$

Number of categories:  $N_c$ ; Standard deviation of continuous distribution:  $\sigma_F$ ; Standard deviation of discrete distribution:  $\sigma_G$

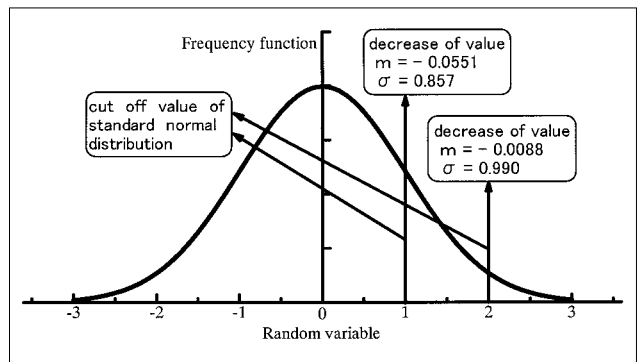


Fig. 2 Mean and standard deviation for the truncated normal distribution.

この  $\sigma_{cq}$  は分布の裾がどの程度打ち切られているかを示す量であり標準偏差が過小評価される量とはほぼ比例する．したがって標準偏差  $\sigma$  は補正量  $\sigma_{cq}$  を使って(3) 式のように補正され，補正された標準偏差  $\sigma_c$  を得る．

$$\sigma_c = \frac{\sigma}{\sigma_{cq}} \dots\dots\dots (3)$$

1-3 平均値の算出と補正

標本の代表値として相加平均を用いることが一般的であるが，外れ値の影響を受けやすく，また，標本の分布が正規分布でなければ全体の代表値としては適当でないことが知られている．そこで，外れ値の影響を受けにくく標本分布の形にあまり依存しない平均値推定法を用いた．まず外れ値をグラブスの検定<sup>8)</sup>により棄却する．次に標本の各データ ( $x_i$ ) に対し，中央値を用いて(4) 式のようにそれぞれ重み付けを行った値 ( $X_i \gamma$ )<sup>10)</sup> を用いて，その相加平均値を計算し加重平均

を求めた。

$$\begin{cases} X_i > 0 & : z = cw \\ X_i \leq x_i : z = \left(1 - \left(\frac{z}{cw}\right)^2\right)^2 & : z \leq cw \end{cases} \dots\dots\dots (4)$$

$z$  : median- $x_i$   
 $c$  : range  
 $w$  : weight  
 $x_i$  : weighted value

ROC曲線の形状は、両正規ROCパラメータとしての平均値の影響を非常に強く受ける。これは平均値と標準偏差をパラメータとしてROC曲線を作図してみれば容易に確認できる。Fig. 2に示すように平均0、標準偏差1の標準正規分布が1 $\sigma$ で打ち切られると平均値は-0.0551、2 $\sigma$ で打ち切られると-0.0088だけ中心より移動する。このように正規分布の裾が打ち切られた場合の平均値を補正するため、次のような補正を行った。まず(5)式を用いて補正值 $m_{cv}$ を得る。

$$m_{cv} = 1 - |\text{upper} - \text{lower}| \dots\dots\dots (5)$$

この補正值 $m_{cv}$ は分布の偏りを表す量であり標準正規分布が $m_{cv}$ に対してどの程度の補正量となるかを求めた。 $m_{cv}$ に対する補正量の関係から補正関数(6)式を求め、 $m_{cv}$ を(6)式に代入することで平均値の補正量 $m_{cq}$ を得る。

$$m_{cq} = 0.45596 \times m_{cv}^2 - 0.41913 \times m_{cv} - 0.0331 \dots\dots\dots (6)$$

平均値 $m$ は補正された標準偏差 $\sigma_c$ と補正量 $m_{cq}$ を用いて(7)式により、補正された平均値 $m_c$ を得る。

$$m_c = m + \sigma_c \times m_{cq} \dots\dots\dots (7)$$

なおグラフの検定でデータが棄却された場合には、この補正を行っていない。

#### 1-4 ブートストラップ法によるROCパラメータの推定

試料間の変動を含んだ陰性および陽性のデータから得られたROC曲線およびそのパラメータは、誤差範囲をもって推定されるのが望ましい。Azの値の標準誤差は、ブートストラップ法によるシミュレーションを行って推定した。さらに陰性および陽性の平均値と標準偏差および平均値の標準誤差についても同様にブートストラップ法により推定することができる。

## 2. 方法

連続データに対する平均および標準偏差補正によるROC曲線の適合度を評価するため、われわれは次のような方法を用いた。1-1~1-4の補正方法を実際の観察

実験に適用するには、観察実験により得られた陰性および陽性のデータをそれぞれ標準化し標準正規分布とする。連続データに対応するROC曲線がどのような形状であるかということは未知であるから、適合度の評価をするためにカテゴリ数20のカテゴリカル・データを利用した。これをカテゴリカル・データとして扱って処理した場合と、連続データとして扱って処理した場合でROC曲線の比較を行った。多くのケースでROC曲線の適合度を評価するために、陰性および陽性のデータそれぞれについて独立に平均と標準偏差を指定して正規乱数による整数のシミュレーションデータを作成した。正規乱数により計算された整数のデータ値がカテゴリ上限の20を超えた場合には、データ値は20となるようにデータを作成する。またカテゴリ下限の1未満となった場合には、データ値が1となるデータを作成する。このようにして平均と標準偏差ならびにデータ数の異なる200程度のデータセットを作成しそれぞれについて評価を行った。

このようにして得られたデータセットはカテゴリカル・データであるから、ROC曲線のグラフ上にカテゴリ数だけデータ点としてプロットされるので、平均および標準偏差補正により得られたROC曲線との適合度を残差平方和などにより評価することが可能である。

## 3. 結果

今回の結果はグラフの検定の有意水準99%、加重平均の重み調節因子 $w=2.0$ で行った。データセットのなかから代表的なデータをピックアップし結果を示す。

Fig. 3の(a)はデータ数が100で標準偏差の大きなケース、(b)はデータ数が100で標準偏差の小さなケース、(c)はデータ数が20と少ないケース、(d)は陰性と陽性のデータ数に偏りがあるケース、(e)は陰性と陽性の標準偏差が異なるケース、(f)はデータ数を500とし陰性、陽性ともほぼ正規分布となる理想的なケースである。

Table 1に今回の補正により推定された陰性および陽性の平均値と標準偏差ならびに残差平方和を示す。ここでいう残差平方和とはグラフ上にプロットされたカテゴリカル・データの点と平均および標準偏差補正により得られたROC曲線との差の平方和である。Fig. 3にはシミュレーションにより得られたデータセットの分布とROC曲線のグラフを示す。Fig. 3のデータセットの分布上に記入してある $m$ は平均値、 $\sigma$ は標準偏差、 $n$ はデータ数を表している。Table 1でみられる特徴として(b)のように平均値の差が2 $\sigma$ より大きいケースや、(c)のようにデータ数がカテゴリ数20に対し

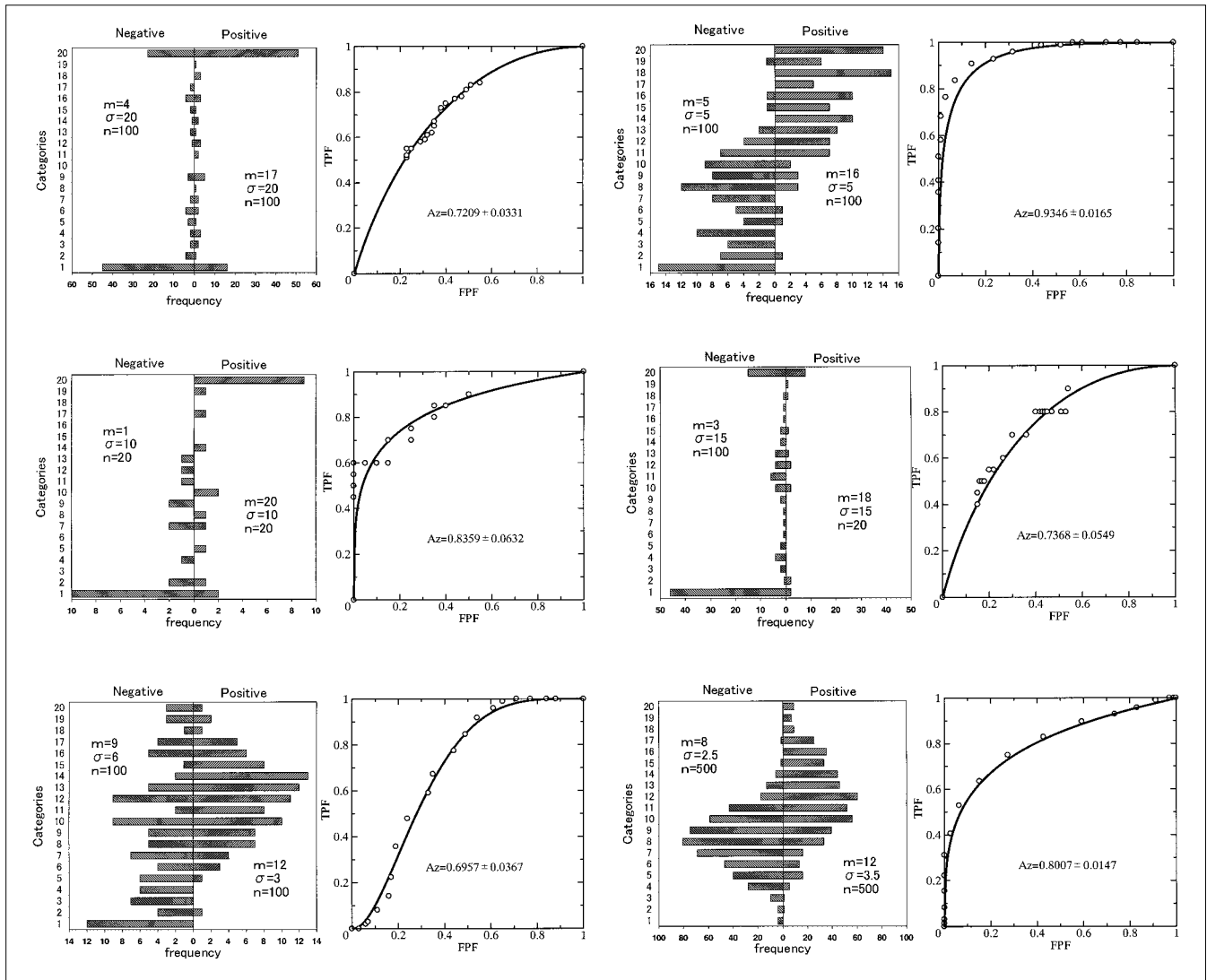


Fig. 3 Relation between the distribution of simulated data sets and the ROC curve  $\alpha(x|f)$

a	b
c	d
e	f

て20と少ないケースで残差平方和が大きい結果となった。Fig. 3でみられる特徴として正規乱数を用いてデータセットを作成したにもかかわらず、(e)のPositiveデータと(f)以外の分布は正規分布とはいえない分布の形状となっている。Azの値はブートストラップ法を行って推定された値と標準誤差を、それぞれROC曲線のグラフ中に示した。

Fig. 4は、陰性と陽性の平均値の差の標準偏差に対する割合と残差平方和の関係である。このケースでは陰性と陽性のデータの標準偏差が等しいものを集めた。

グラフから分かることは、平均値の差が1σ程度の場合、残差平方和の値は非常に小さいということである。しかし2σを超えると残差平方和の値が大きくな

Table 1 Corrected means, corrected standard deviations, and residual sum of squares for simulated data sets shown in Fig. 3. m and σ indicate mean and standard deviation, respectively.

	Negative case		Positive case		residual sum of squares
	m	σ	m	σ	
(a)	5.18	11.4	14.0	9.59	0.0077
(b)	6.15	4.11	14.8	3.85	0.7753
(c)	3.60	5.73	13.8	8.37	1.9946
(d)	5.95	9.63	14.0	8.05	0.0328
(e)	8.27	6.10	11.8	3.32	0.0214
(f)	7.85	2.49	11.5	3.53	0.1839

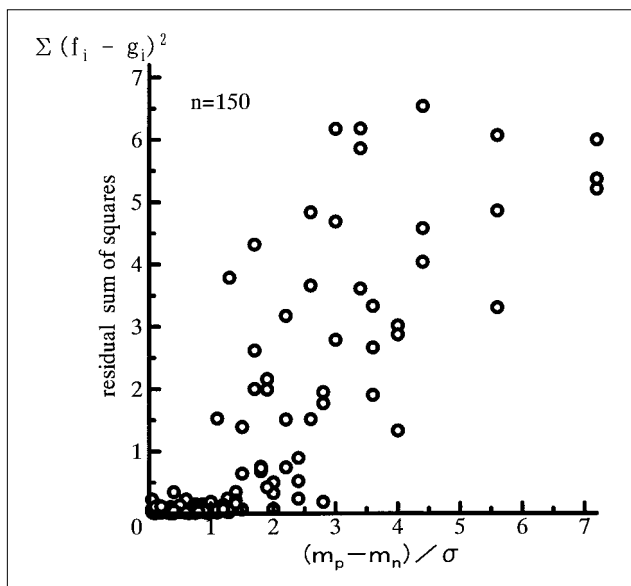


Fig. 4 Relation between the ratio of the differences of mean values to the standard deviation, and the residual sum of squares.

Table 2 Simulated data sets with a large residual sum of squares.

Negative			Positive			ratio	residual sum of squares
m	σ	n	m	σ	n		
4	10	20	17	10	20	1.3	3.7838
2	10	100	19	10	100	1.7	2.6172
2	10	20	19	10	20	1.7	4.3156
2	10	100	19	10	20	1.7	2.0000
1	10	100	20	10	100	1.9	2.1575
1	10	20	20	10	20	1.9	1.9894
5	5	20	16	5	20	2.2	3.1713
4	5	20	17	5	20	2.6	3.6566
4	5	100	17	5	20	2.6	4.8327
7	2.5	100	14	2.5	100	2.8	1.9477
3	5	100	18	5	100	3.0	2.7864
3	5	20	18	5	20	3.0	6.1708
3	5	100	18	5	20	3.0	4.6835

る結果となった。平均値の差が  $3\sigma$  を超える場合、ROC曲線は  $(0,0)-(0,1)-(1,1)$  の軸上に描画され、 $A_z$  の値もほぼ 1 となる。このように平均値の差が  $3\sigma$  を超える場合、すべてのケースで残差平方和の値は 1.32 より大きかった。また平均値の差が  $1\sigma$  以下の場合にはすべてのケースで残差平方和は 0.35 より小さかった。平均値の差が  $1\sigma$  を超え  $3\sigma$  以下の場合で残差平方和の値が大きいケースをピックアップすると Table 2 のようになった。Table 2 から分かることは、データ数が 20 と少ないケースではデータのばらつきが大きいいため残差平方和も大きい結果となっている。

また平均値がカテゴリーの下限付近や上限付近に偏っているケースも残差平方和の値が大きい結果となっている。

Fig. 5 は、陰性と陽性の標準偏差が異なるケースについて、それぞれの標準偏差の比と残差平方和の関係である。標準偏差の比は (8) 式により計算した。

$$\begin{cases} \sigma_p / \sigma_n \geq 1 : \sigma_p / \sigma_n \\ \sigma_p / \sigma_n < 1 : \sigma_n / \sigma_p \end{cases} \dots\dots\dots (8)$$

グラフから分かることは標準偏差の比が 3 未満ならば残差平方和の値は 0.42 以下、6 未満ならば 1.28 以下と比較的小さい値となったことである。標準偏差の比が 6 を超えると残差平方和の値が大きくなるという結果であった。

#### 4. 考 察

連続データに対する ROC 曲線の適合には LABROC や ROCFIT といったソフトウェアがよく用いられてい

る。しかし連続データをカテゴリーに分類するという過程が必要であり、カテゴリー化が客観的かつ標準的なものでなければ再現性を持たないことは明らかである。しかしこの方法による ROC 曲線の適合などの作用についてはコンピュータシミュレーションを用いて非常に広範囲にわたって調べられている<sup>3)</sup>。

われわれはカテゴリー分類を行わず、連続データからパラメータの点推定を行う方法を試行錯誤しながら模索し今回の補正法を考案した。連続データに対する ROC 曲線の適合に対し、例えば両正規確率紙の標準偏差軸上にデータ点が得られるようにカテゴリー化した場合は最尤法や最小ロジット  $\chi^2$  法および最小二乗法といったアルゴリズムの適用<sup>11)</sup> が可能である。しかし連続データのまま扱おうとすると標準偏差軸上にデータ点は得られず今回のような算術法を用いる必要がある。この 2 種類の方法の間にはデータの扱い方に根本的な違いがあるのでその優劣はつけがたいが、算術法はデータ数が多く得られなければ推定の精度に問題を残す<sup>12)</sup>。そこでわれわれはブートストラップ法を用い推定値の誤差範囲を求めることで推定精度に対する問題の対策とした。今回のデータセットの中でも Fig. 3 の (f) のようにほぼ理想的な正規分布をするデータの場合には、グラフの検定も加重平均も行わない方が適合結果は良い。しかし実際の観察実験で得られたデータが、理想的な正規分布である場合はほとんどないので、われわれの考案した平均および標準偏差補正による ROC 曲線の適合法も理想的な正規分布はほとんどないという前提を基にしている。ただしデータの分布が正規分布に従うということが明らかであれば、今回

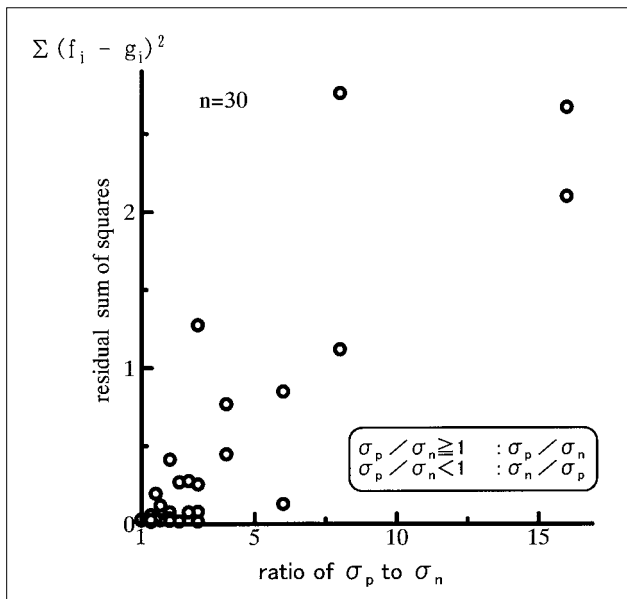


Fig. 5 Relation between the ratio of standard deviations and the residual sum of squares.

のような補正は必要なく、単に標本平均と標本標準偏差を両正規ROCパラメータとしてROC曲線を描くことが可能である<sup>13)</sup>。われわれの行った補正法では、データが正規分布であればグラフの有意水準を99.99%等のように100%近くに設定し加重平均の重みをおおむね3.0以上の値とすれば得られる推定量は相加平均および不偏分散の平方根である標準偏差に等しくなる。

今回のシミュレーションでROC曲線の適合度が悪かったのは、平均の差が $3\sigma$ を超える場合や、標準偏差の比が6を超える場合、またデータ数が少ない場合であった。データ数が少ない場合は本質的にデータのばらつきが多くなるので、これを避けるためにはデータ数を多くすればよい。しかし平均の差が $3\sigma$ を超えたり標準偏差の比が6を超えるケースというのは、ROCのデータとして本来ふさわしくないデータであると考えられる。今回のような平均および標準偏差補正による方法の改良が進めば前述のようなケースでも良好な適合結果が得られるかもしれない。しかし平均の差が $3\sigma$ もあるデータのROC解析に意味があるかどうかと

いう問題も生じてくると思われる。

今回われわれは観察実験により得られるデータが正規分布の裾を打ち切られたデータであると仮定した。しかしこの仮定にはいくつかの問題がある。第一点は観察試料が正規母集団からの無作為抽出であるのかどうかという標本抽出に関する問題である。第二点は観察者の反応が正規分布を仮定できるのかという問題である。この2点に関してROC解析の文献の多くに対策が論じられている。第三点は上記2点を守った場合、観察実験のデータが正規分布になるのかどうかという問題である。われわれの仮定は実際の観察実験により得られたデータが正規分布ではあるけれども裾が打ち切られた形状をしており、データの標本平均および標本標準偏差が本来の平均および標準偏差に対してズレを生じる、というものである。われわれの仮定は、今回の正規乱数を利用したデータセットにおいて良好な結果を得ることができた。しかし実際の観察実験により得られたデータの分布がどのような形状であるかということは、慎重に多くの観察実験を行って検証する必要がある。

## 5. 結 語

平均および標準偏差の補正値を両正規ROCパラメータとして連続データに対するROC曲線の適合を試みた。修正されたカテゴリカル・データを連続データと同様に扱ってROC曲線の適合度を評価した。ROC曲線の適合度が悪かったケースは、平均の差が $3\sigma$ 上のデータか、または、標準偏差の比が6以上のデータであった。われわれの行った方法は特異なケースを除き、良好な結果を得ることができた。さらに、ブートストラップ法を用いて試料間のばらつきを考慮した平均とROCエリアの指標であるAzの標準誤差が求められた。

## 謝 辞

本研究に対してご助言およびデータを提供して下さった大阪市立大学医学部附属病院中央放射線部の白石順二氏に感謝いたします。

## 参考文献

- 1) 川村慎二, 下新原茂, 稲津 博: SPECT画像のROC解析. 日放技学誌, 52(12), 1633-1637, (1996).
- 2) 藤田広志: ROC解析の基礎と応用; 6. 2 コンピュータ支援診断システムにおけるROC解析. 日放技学叢書(8), pp.65-69, 日放技学会, 京都, (1994).
- 3) 白石順二: ROC解析の基礎と応用; 5. 連続確信度法. 日放技学叢書(8), pp.36-54, 日放技学会, 京都, (1994).
- 4) 松原 望: 統計の考え方. 放送大学教育振興会, pp.215-223, (2000).
- 5) 村上征勝, 田村義保: パソコンによるデータ解析. 朝倉書店, pp.123-142, (1999).
- 6) 丹後俊郎: 統計モデル入門. 朝倉書店, pp.18-34, (2000).
- 7) 田中良久: 心理学的測定法. 東京大学出版会, pp.249-250, (1999).
- 8) 石村貞夫: すぐわかる統計解析. 東京図書, pp.76-79, (1998).
- 9) 中川 徹, 小柳義夫: 最小二乗法による実験データ解析. 東京大学出版会, pp.163-168, (1989).
- 10) 新村秀一: パソコンによるデータ解析. 講談社, pp.57-59, (1996).
- 11) Hanley JA: Receiver Operating Characteristic (ROC) Methodology. Critical Reviews Diagnostic Imaging, Volume 29, 315-332(1989)
- 12) 田中良久: 心理学的測定法. 東京大学出版会, pp.68-70, (1999).
- 13) 白石順二, 上田容子, 宇都宮あかね, 他: ROC解析を応用したCT装置の密度分解能の測定. 日放技学誌, 53(4), 431-436, (1997).

## 図表の説明

Fig. 1 変形したシェパードの修正の説明  
カテゴリーカル・データの修正

$$\sigma_F = \sqrt{\sigma_G^2 - \frac{1}{N_C}}$$

カテゴリー数:  $N_C$ , 連続分布の標準偏差:  $\sigma_F$ , 離散分布の標準偏差:  $\sigma_G$

Fig. 2 平均と標準偏差の補正量の説明

Fig. 3 データセットの分布とROC曲線の関係(a)(f)

Fig. 4 平均の差と標準偏差の比に対する残差平方和の関係

Fig. 5 標準偏差の比と残差平方和の関係

Table 1 Fig. 3で示したカテゴリーカル・データの補正後の平均値, 標準偏差および残差平方和

Table 2 残差平方和の値が大きなデータセット