

2021年11月13日（土）  
第68回日本臨床検査医学会 EBLM委員会企画  
臨床検査領域に対する統計・機械学習的アプローチ

# 臨床検査の診断的有用性の評価法 その統計理論と使い方の実際

山口大学大学院医学系研究科  
保健学系学域・生態情報検査学  
市原清志

日本臨床検査医学会  
COI 開示

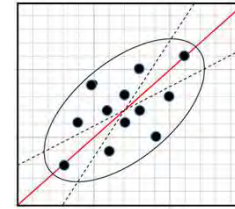
発表者：市原清志

本講演に関し、開示すべきCOI  
関係にある企業などはありません

# 臨床検査領域で求められる統計処理技術

## 1) 測定法の方法間比較 (回帰直線の求め方)

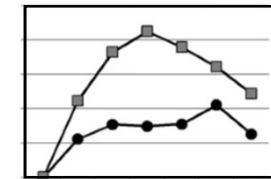
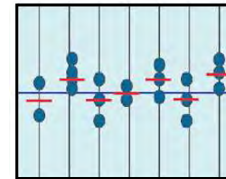
線形関係式と予測回帰式の違い



## 2) 測定精度の評価と変動要因の分析 (分散分析)

検査室内QC : 日間CV、日内CV、機種間CV

検査室間QC : 方法間CV、方法内CV



## 3) 臨床検査の生理的変動分析

個体内変動要因分析 (条件間比較分析 : 日内リズム、体位変化)

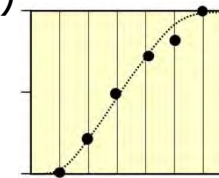
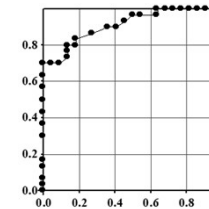
病気以外で検査値が変化する要因の解明 (重回帰分析)

## 4) 臨床検査の病態変動分析 = 診断的有用性評価

単一検査の診断性能評価 (感度・特異度とROC分析)

複数検査の比較評価 (多重ロジスティック分析)

組み合わせ診断方程式の作成



### <利用上のノウハウ>

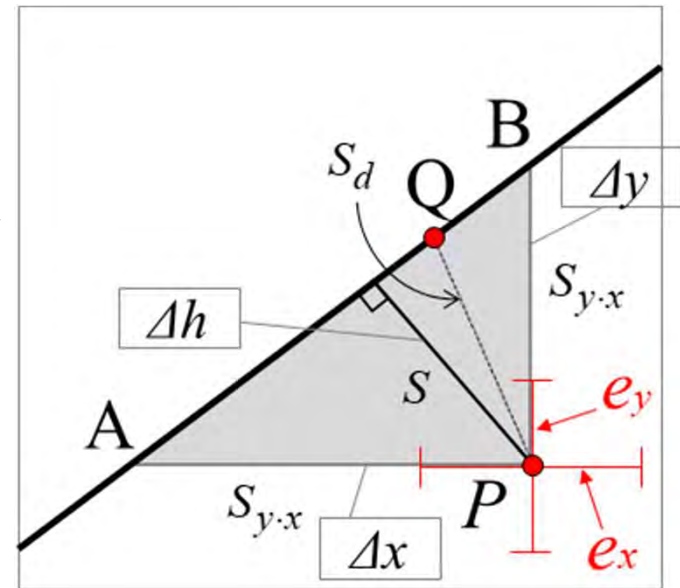
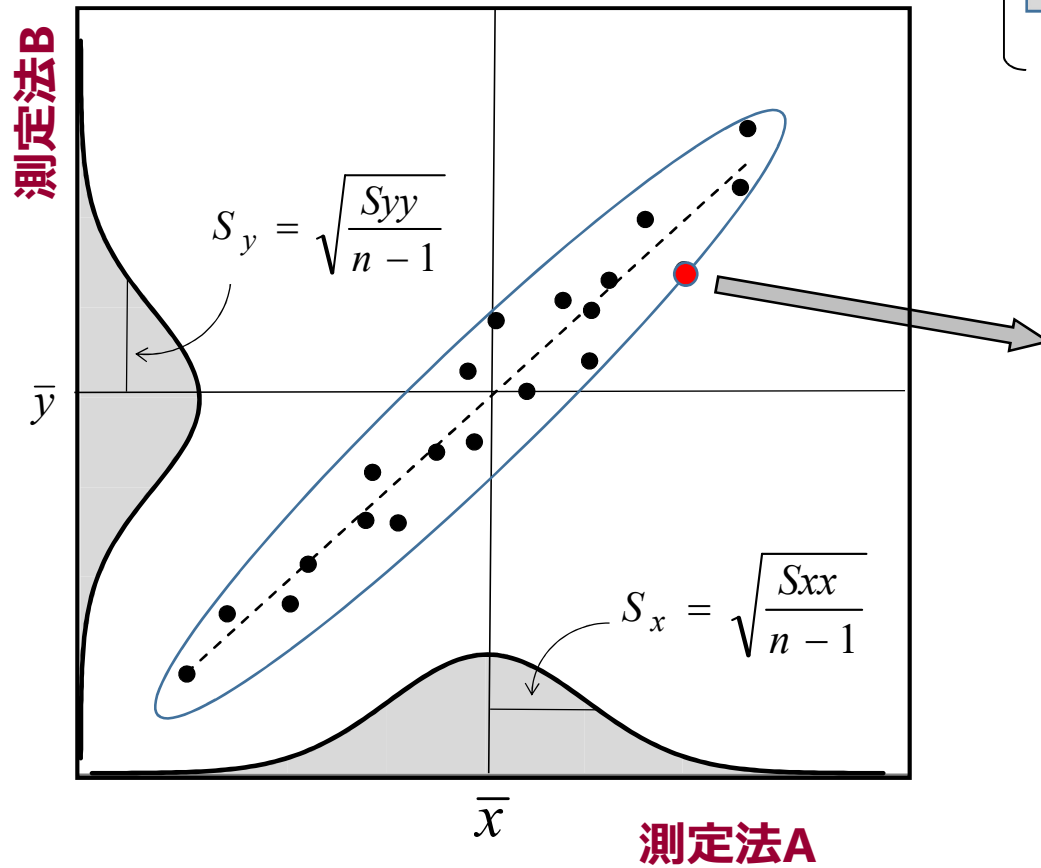
- ・ 特定の要因の影響が強すぎる場合、**傾向スコア**でマッチング
- ・ モデルに組み込める説明変数の上限
- ・ 過剰適合への対応

**【参考1】**

**測定法の方法間比較  
(回帰直線の求め方)**

# 方法間比較のための 回帰直線の求め方

- Model I 回帰**
- ①  $x \rightarrow y$  の回帰  $\Sigma \Delta y^2$  を最少に
  - ②  $y \rightarrow x$  の回帰  $\Sigma \Delta x^2$  を最少に
- Model II 回帰**
- ③ 標準主軸回帰  $\Sigma \Delta x \Delta y$  を最少に
  - ④ 主成分回帰  $\Sigma \Delta h^2$  を最少に
  - ⑤ Deming回帰  $s_d$  を最少に
- ③ Passing-Bablok 回帰・・・ノンパラ法



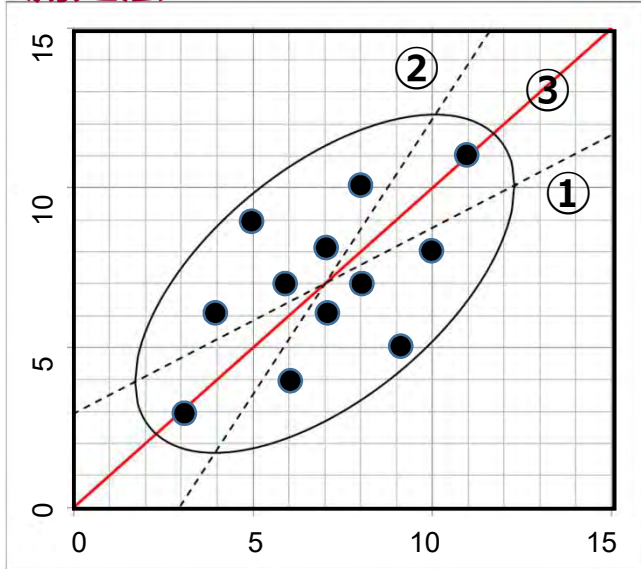
誤差分散比  $\lambda = \frac{e_y^2}{e_x^2}$

# 回帰直線の推定法

方法間比較では測定法AとBを同等に扱うため線形関係式を求める

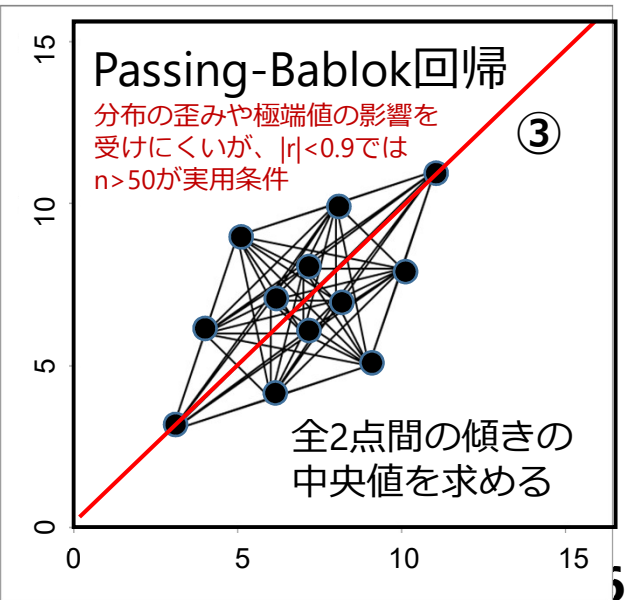
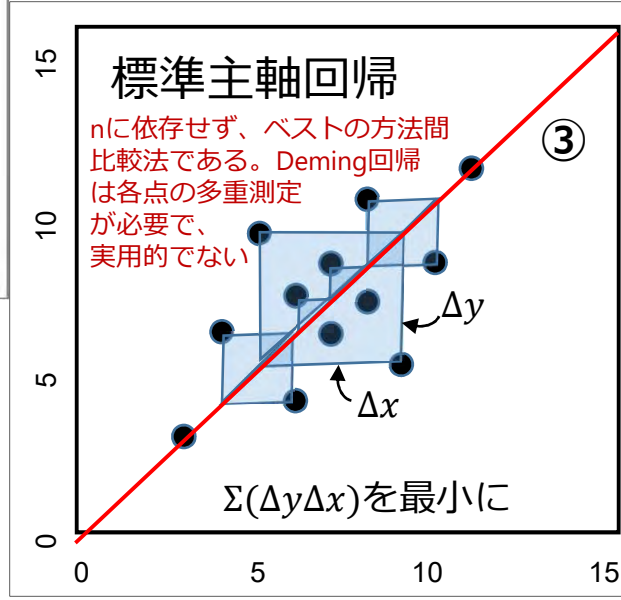
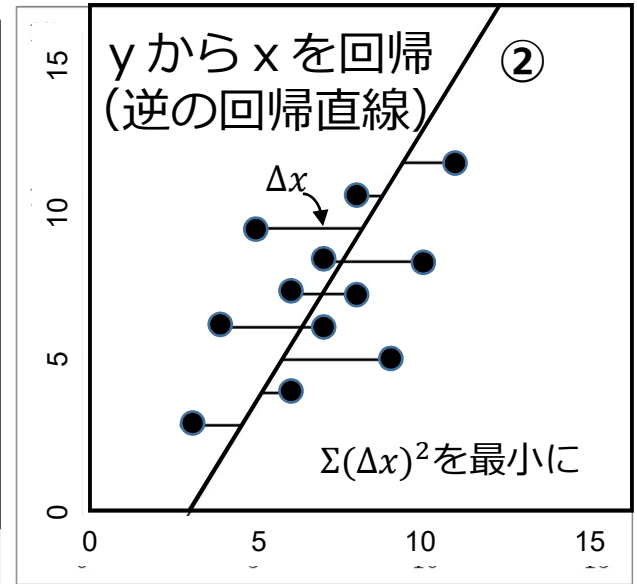
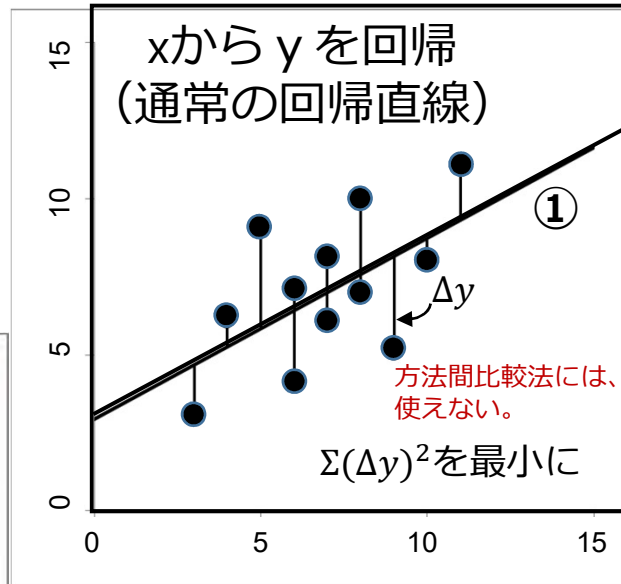
回帰直線 →

測定法B



測定法A

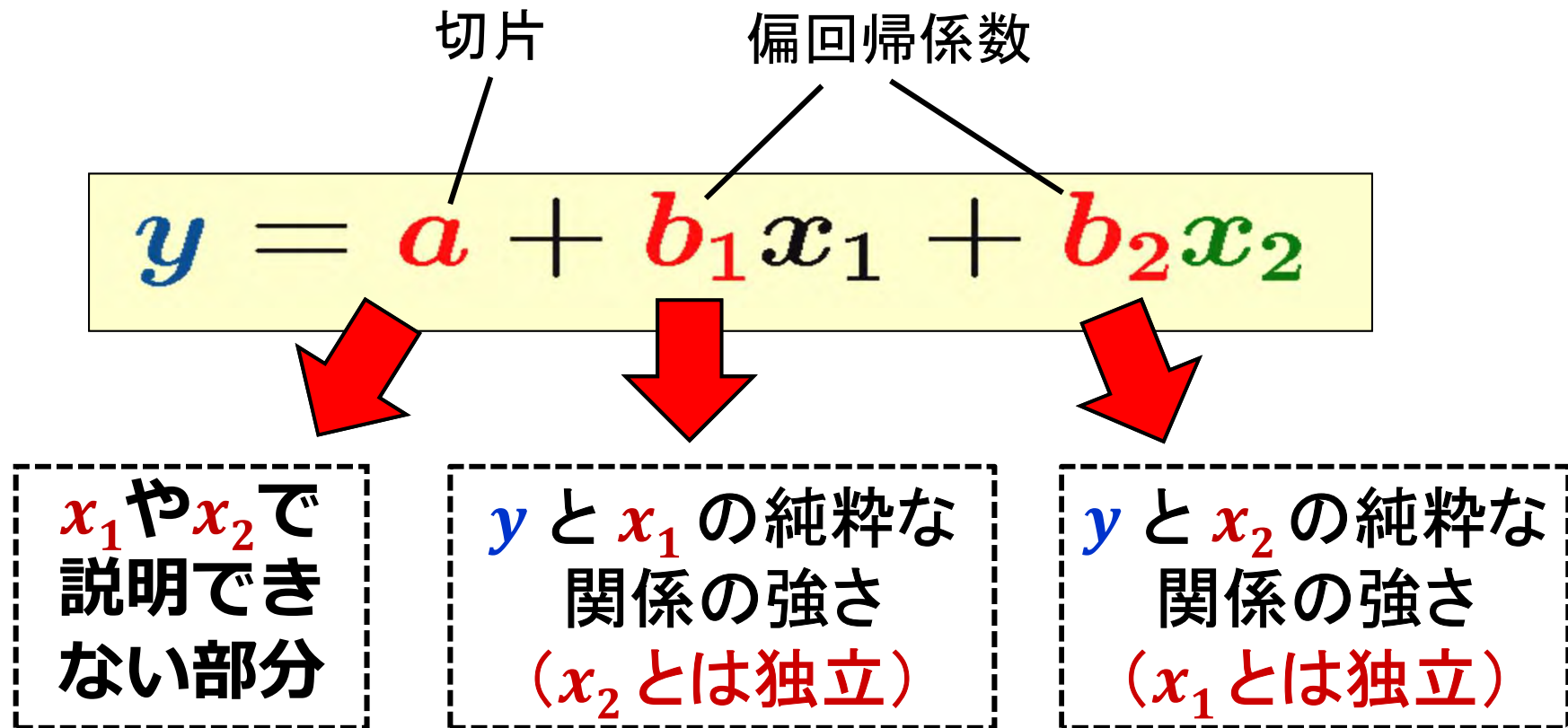
線形関係式 →



**【参考2】**

**重回帰分析の基礎  
(利用のポイント)**

## 重回帰式と回帰係数の意味

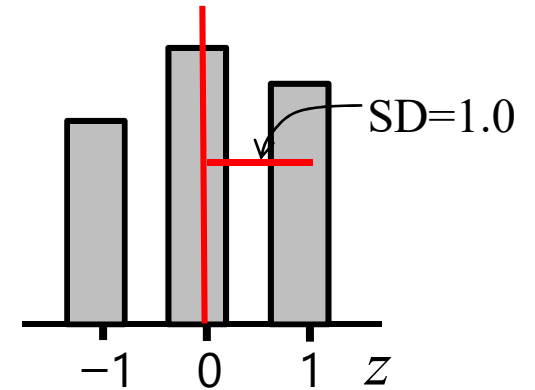
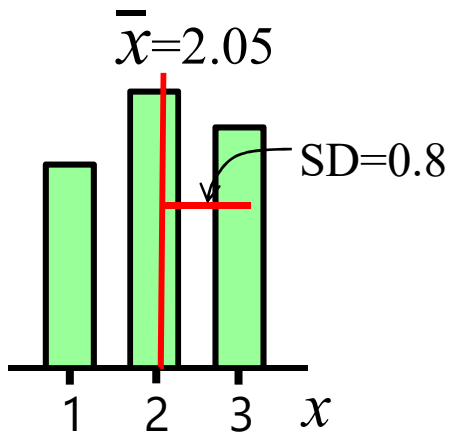


偏回帰係数  $b_1$  と  $b_2$  から各説明変数と  $y$  の関係の強さを表すが、単位に依存する。  
しかし標準偏回帰変数( $\text{std}\beta$ )に変換すれば、単位によらず関係の強さを評価できる。

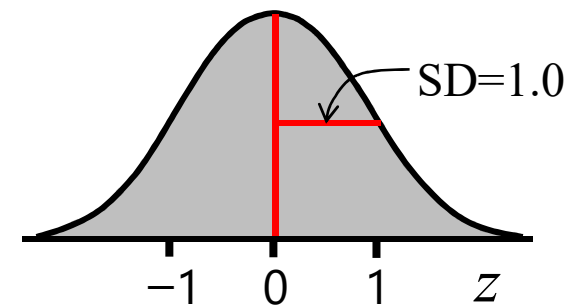
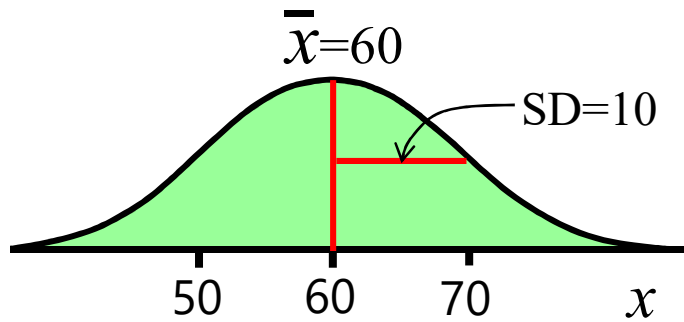


# 標準偏回帰係数(stdβ)の計算

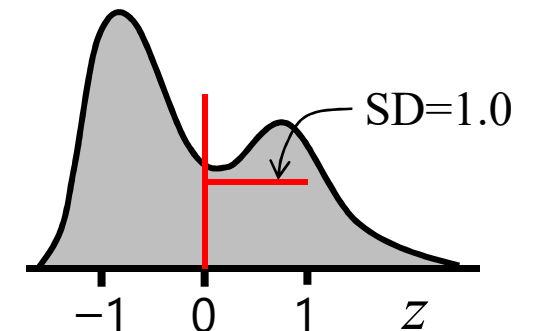
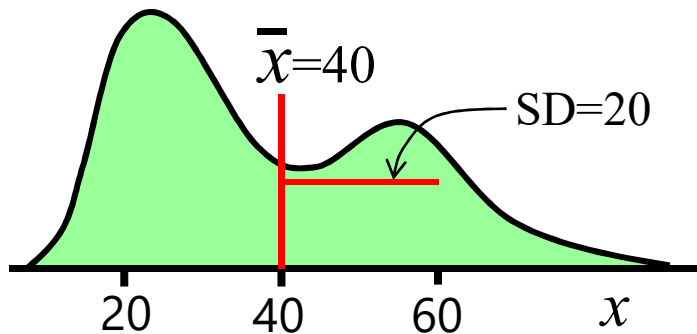
標準偏回帰係数 ≡ 偏相関係数  
 Standard partial regression coefficient ≡ partial correlation coefficient



目的変数も、説明変数も  
 数値を標準化してから  
 重回帰式を作成



$$z = \frac{x - \bar{x}}{SD}$$



多変量回帰では、説明変数が追加される毎に、  
偏回帰係数が再配分される

①  $y = a + b_1 x_1$

②  $y = a' + b'_1 x_1 + b'_2 x_2$

③  $y = a'' + b''_1 x_1 + b''_2 x_2 + b''_3 x_3$

		説明変数		
		性別	年齢	BMI
目的変数		stdβ1	stdβ2	stdβ3
①	HDL-C	<b>0.486</b>		
②	HDL-C	<b>0.485</b>	<b>0.046</b>	
③	HDL-C	<b>0.382</b>	<b>0.097</b>	<b>-0.288</b>

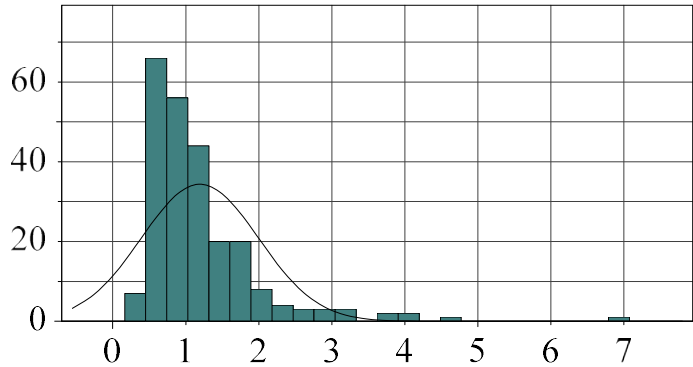
年齢の追加で影響受けず

BMIの追加で大きな変化

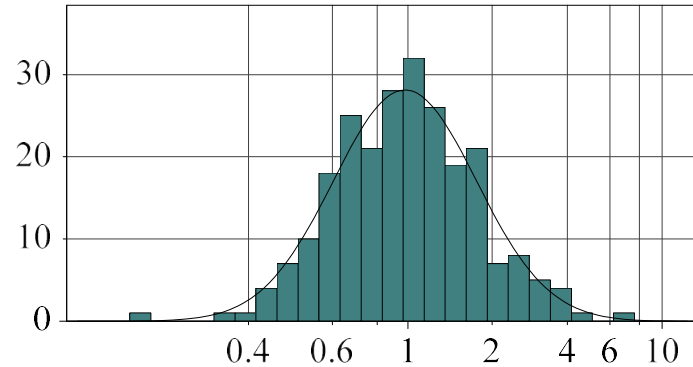
## 例題1：中性脂肪（TG）の生理的変動要因の分析

年齢	BMI	喫煙度	飲酒度	運動度	TG	HDL-C	LDL-C
61	23.3	0	3	2	1.21	1.92	4.21
55	27.0	2	1	0	2.06	1.12	4.17
54	26.8	0	0	3	0.91	1.63	2.96
43	20.3	0	1	0	1.39	1.00	3.41
30	23.0	0	4	1	0.43	1.97	2.95
57	23.2	0	2	1	0.71	1.92	2.68
61	23.1	0	0	0	0.97	1.55	4.31
63	26.0	0	0	3	1.27	1.11	2.84
22	19.5	1	0	0	0.47	1.26	1.89
52	23.8	1	2	1	1.75	1.78	3.15
25	22.8	0	1	0	0.67	1.43	3.63

健常男性：n=240



対数変換



重相関係数



目的変数 TG

説明変数	$\beta$	SE( $\beta$ )	std $\beta$	P値	R
年齢	0.0138	0.0039	0.224	0.00048	0.224

年齢	0.0109	0.0038	0.176	0.00482	0.348
BMI	0.0843	0.0192	0.271	0.00002	

年齢	0.0106	0.0038	0.171	0.00623	0.353
BMI	0.0845	0.0192	0.272	0.00002	
喫煙度	0.0787	0.0816	0.059	0.33570	

年齢	0.0106	0.0039	0.171	0.00645	0.353
BMI	0.0845	0.0193	0.272	0.00002	
喫煙度	0.0794	0.0837	0.059	0.34350	
飲酒度	-0.0008	0.0421	-0.001	0.98435	
運動度	0.0022	0.0426	0.003	0.95921	

目的変数 log(TG)

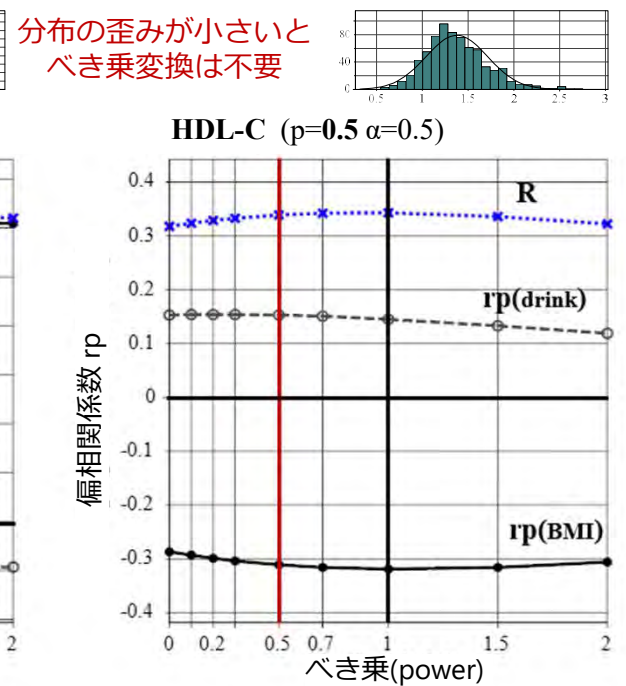
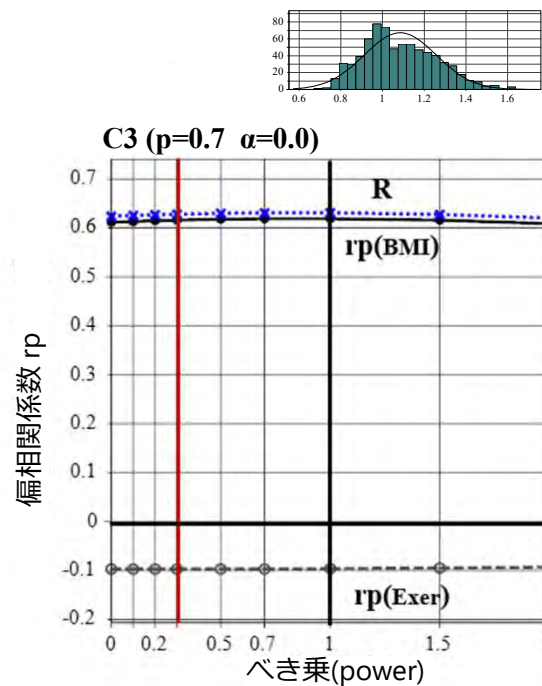
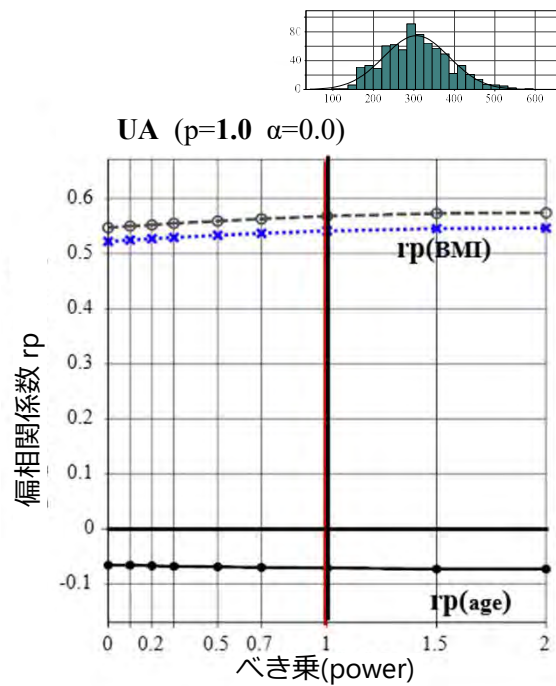
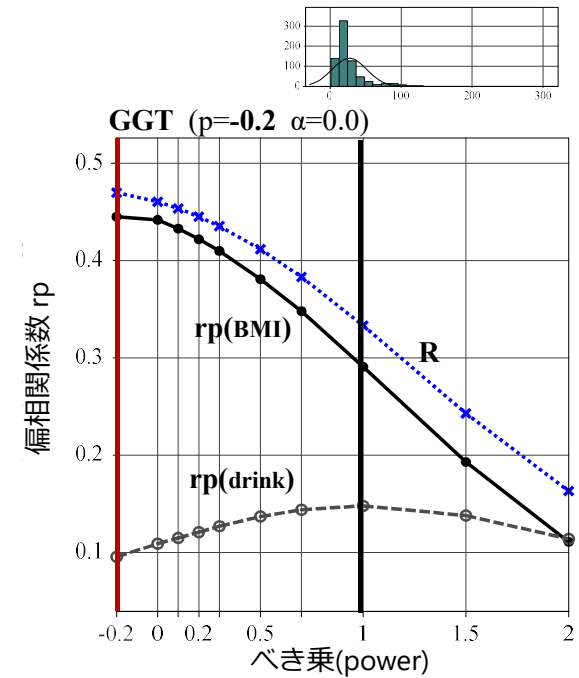
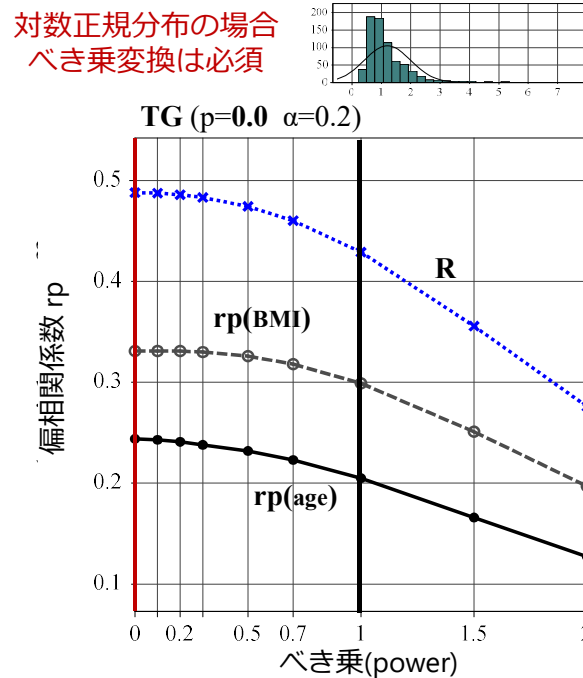
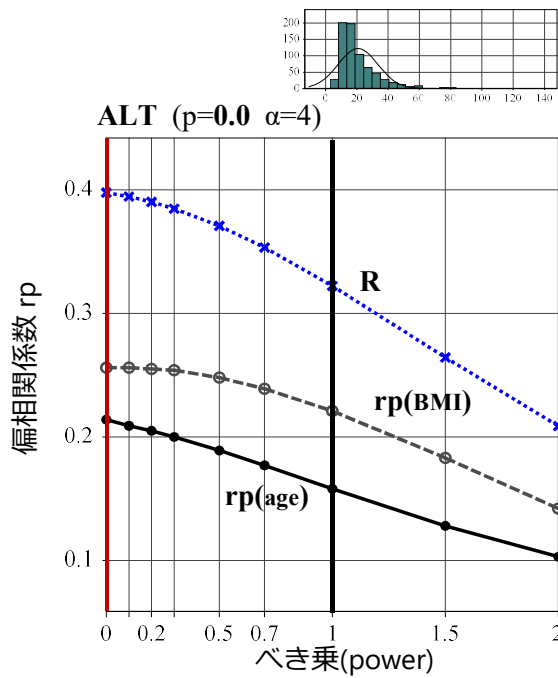
説明変数	$\beta$	SE( $\beta$ )	std $\beta$	P値	R
年齢	0.0167	0.0035	0.300	0.00000	0.300

年齢	0.0136	0.0033	0.244	0.00006	0.433
BMI	0.0885	0.0167	0.316	0.00000	

年齢	0.0130	0.0033	0.234	0.00010	0.453
BMI	0.0890	0.0165	0.318	0.00000	
喫煙度	0.1628	0.0701	0.135	0.02108	

年齢	0.0131	0.0033	0.235	0.00010	0.458
BMI	0.0885	0.0165	0.316	0.00000	
喫煙度	0.1798	0.0716	0.149	0.01275	
飲酒度	-0.0411	0.0360	-0.068	0.25498	
運動度	0.0181	0.0364	0.029	0.62005	

TG分布の歪みを調整すると、std $\beta$ が高くなる 12

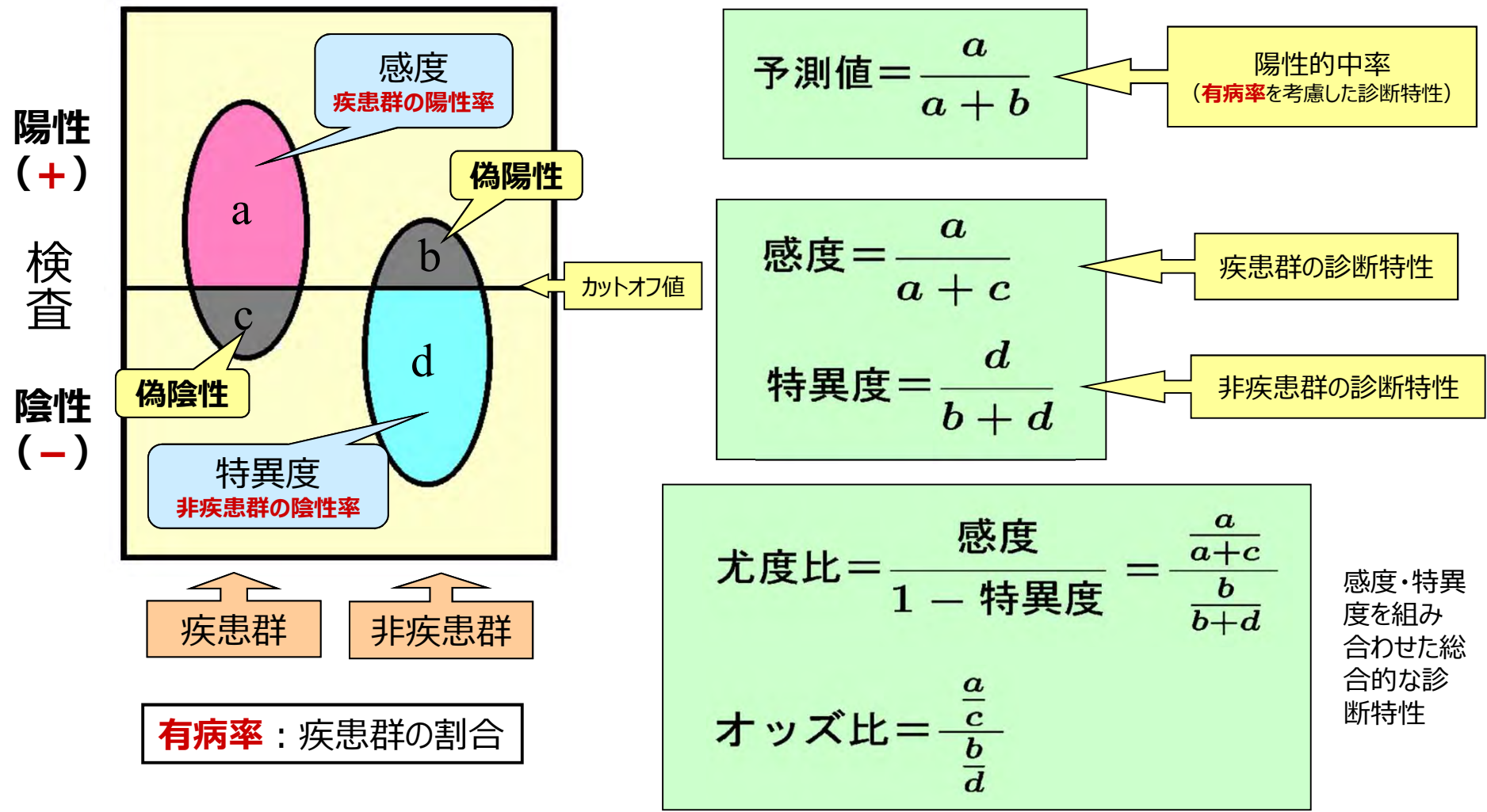


対数正規分布の場合  
べき乗変換は必須

分布の歪みが小さいと  
べき乗変換は不要

# ROC分析による 診断特性の計算

# 検査の診断的有用性に関する基本特性



感度・特異度だけでは、診断能は判断できない。  
両者を組み合わせた、尤度比またはオッズ比を用いる方が良い

# オッズと確率

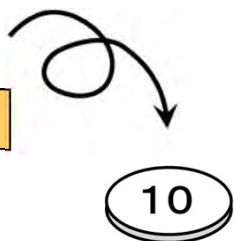
## Odds

オッズ

## Probability 確率

(Likelihood: 尤度)

例 1



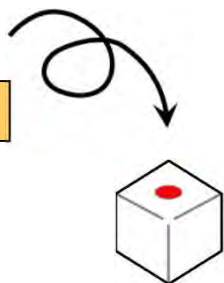
$$\frac{\text{10}}{\text{ }} = \frac{1}{1}$$

オッズは起こる場合と起こらない場合の比

$$\frac{\text{10}}{\text{10} + \text{ }} = \frac{1}{2}$$

確率は起こる場合の全体に対する比

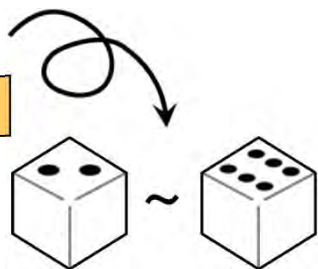
例 2



$$\frac{\text{Red}}{\text{Black} \sim \text{White}} = \frac{1}{5}$$

$$\frac{\text{Red}}{\text{Red} \sim \text{Black} \sim \text{White}} = \frac{1}{6}$$

例 3



$$\frac{\text{Black} \sim \text{White}}{\text{Red}} = \frac{5}{1}$$

$$\frac{\text{Black} \sim \text{White}}{\text{Red} \sim \text{Black} \sim \text{White}} = \frac{5}{6}$$

$$O = \frac{P}{1 - P}$$

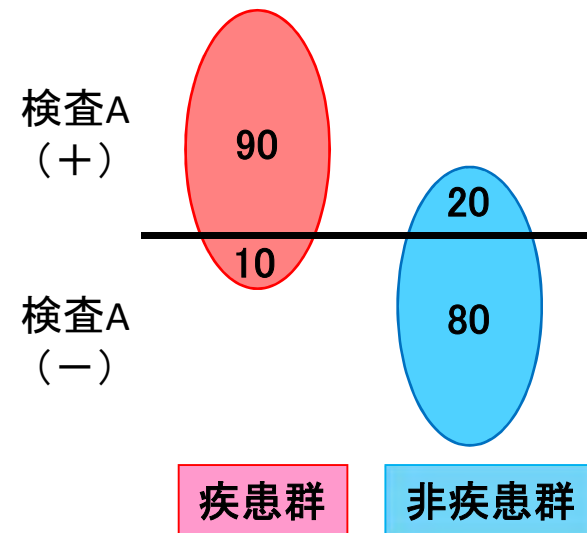
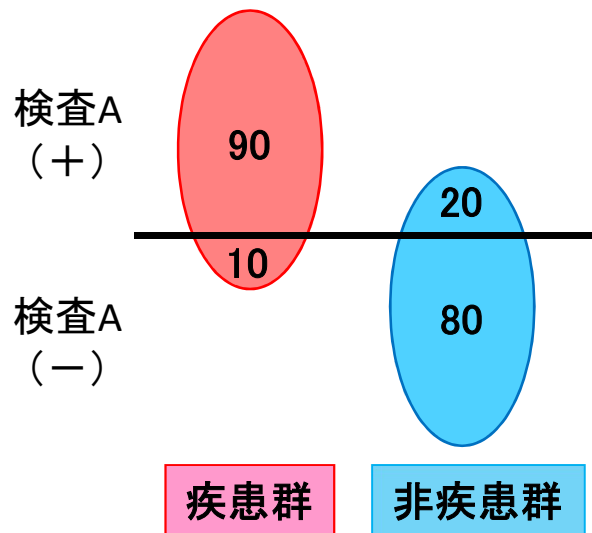
相互変換可能

オッズも確率も起こりやすさの指標



$$\text{尤度比} = \frac{\text{疾患群の陽性率}}{\text{非疾患群の偽陽性率}}$$

$$\text{オッズ比} = \frac{\text{疾患群の陽性オッズ}}{\text{非疾患群の偽陽性オッズ}}$$

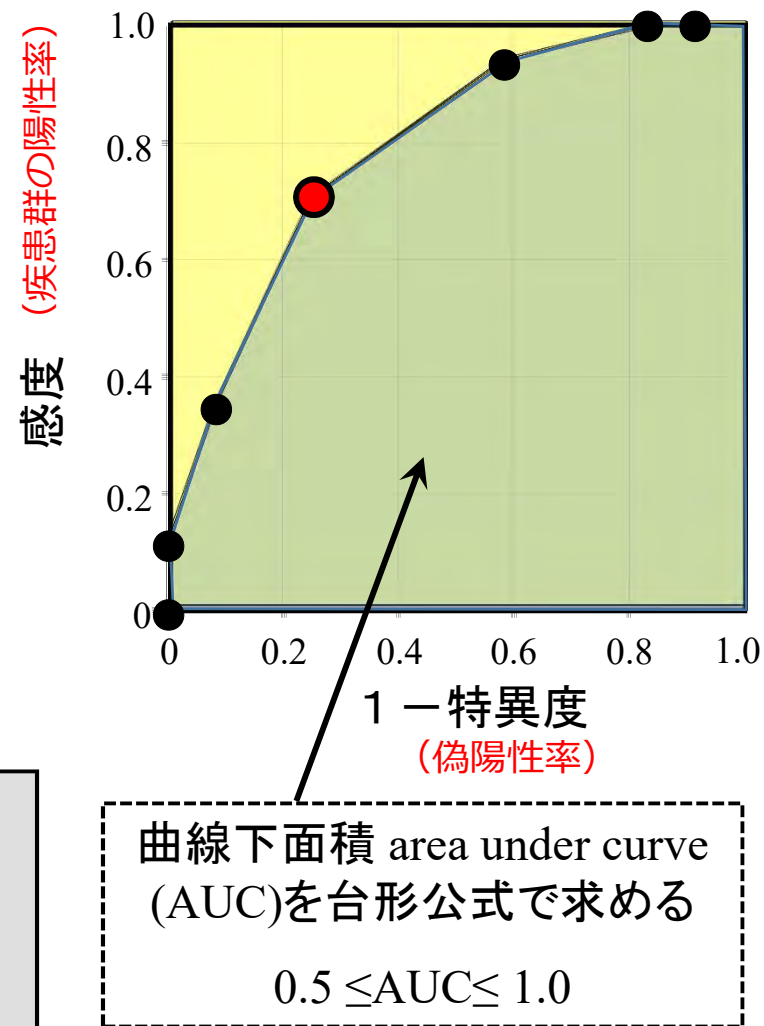
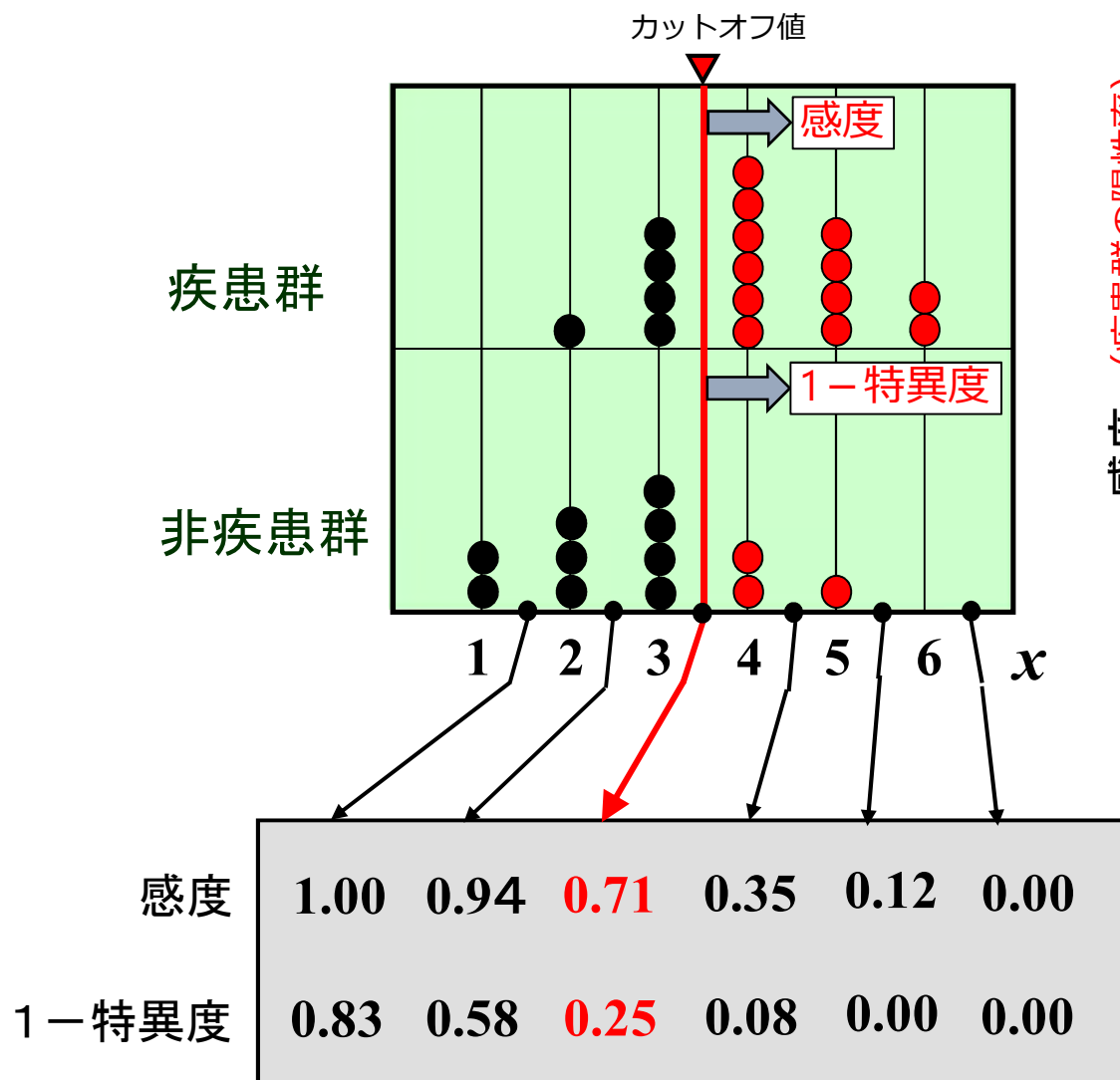


$$\text{尤度比} = \frac{90/100}{20/100} = \frac{0.9}{0.2} = 4.5$$

$$\text{オッズ比} = \frac{90/10}{20/80} = \frac{9.0}{0.25} = 36$$

検査の診断能の評価ではオッズ比が主に用いられる  
 これは、ロジスティック回帰で簡単に求まるため  
 問題は、尤度比もオッズ比も、カットオフ値に依存すること

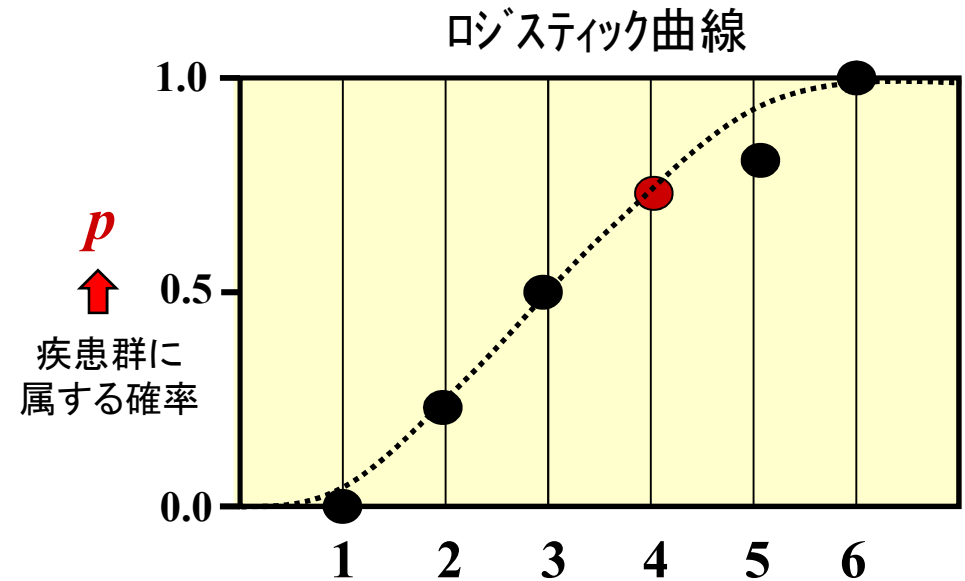
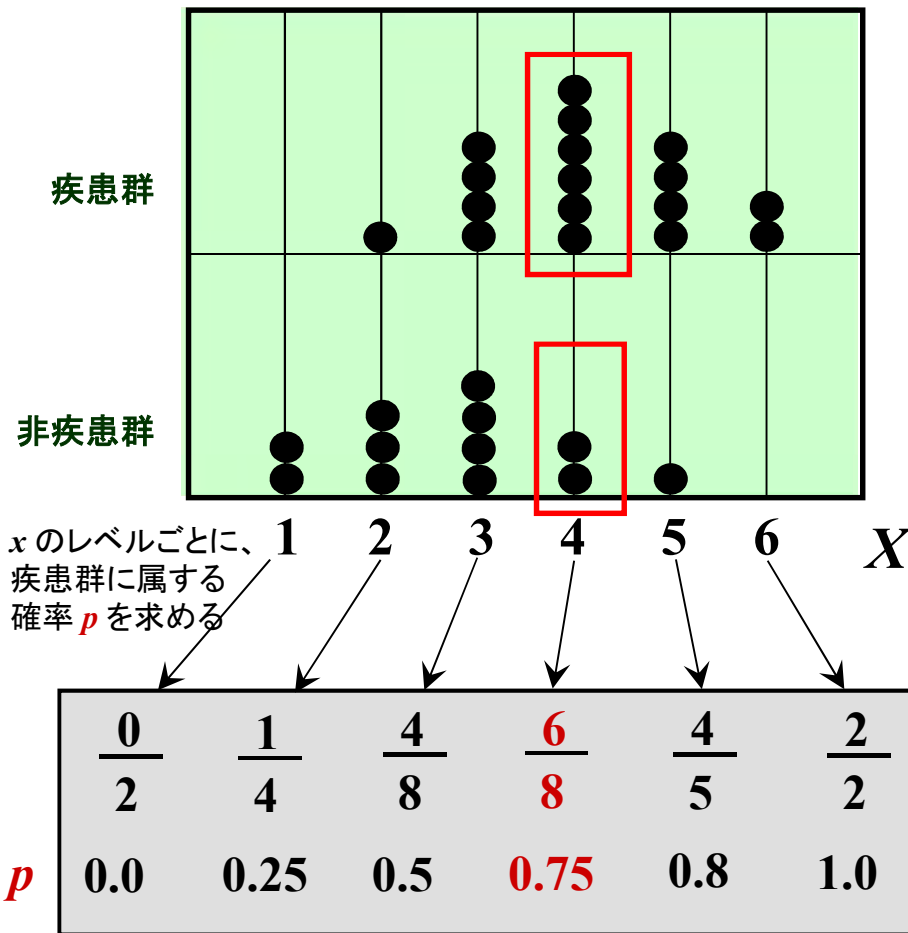
# ROC曲線による判別度の分析



尤度比もオッズ比も、cutoff値に依存したが、  
AUCはcutoff値の段階的な変化で決めているので、診断能をよりの確に表す

# 多重ロジスティック分析による 複数検査の診断特性の比較

# ロジスティック曲線による判別



$$p = \frac{1}{1 + e^{-X}}$$

$$X = a + bx$$

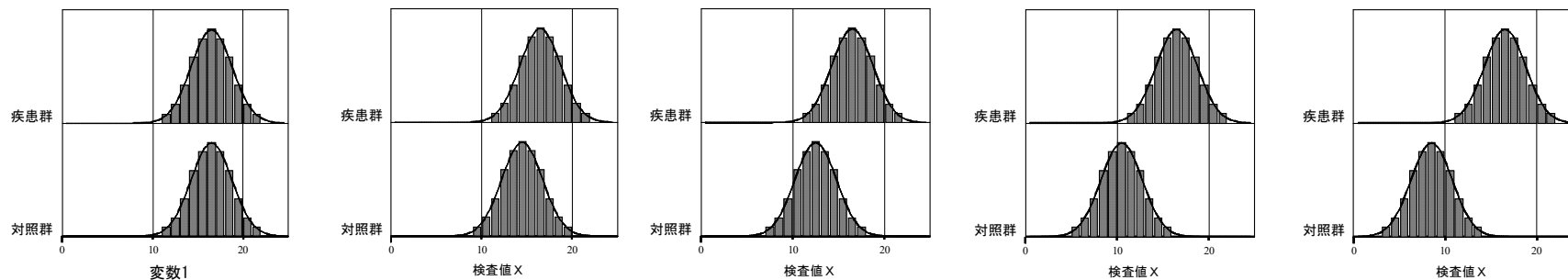
要点

- 1) 回帰係数  $b$  が検査の診断能を表す
- 2)  $b$  が 0 と有意に異なるかを検定
- 3)  $\exp(b)$  は検査のオッズ比を表す

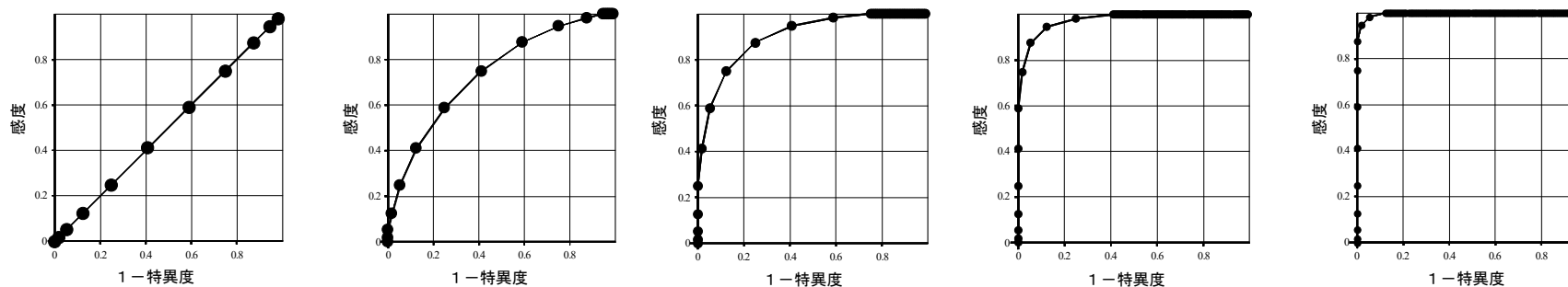
ROC解析と一変量ロジスティック回帰は検査の診断能の評価は、数理的に同等である

# 2群の分離度とROC曲線、ロジスティック曲線

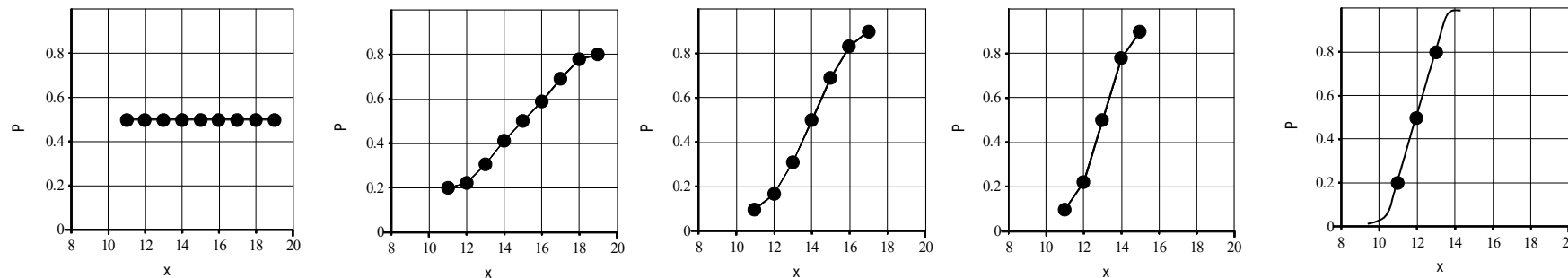
## 2群の検査値



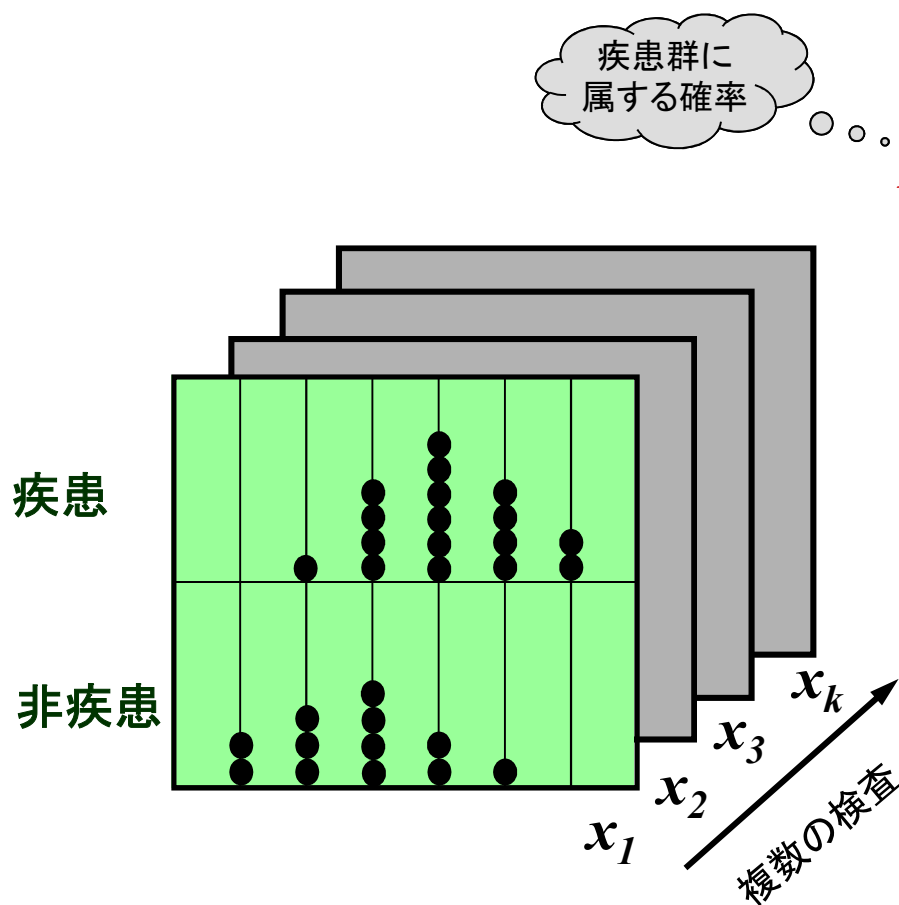
## ROC曲線



## ロジスティック曲線

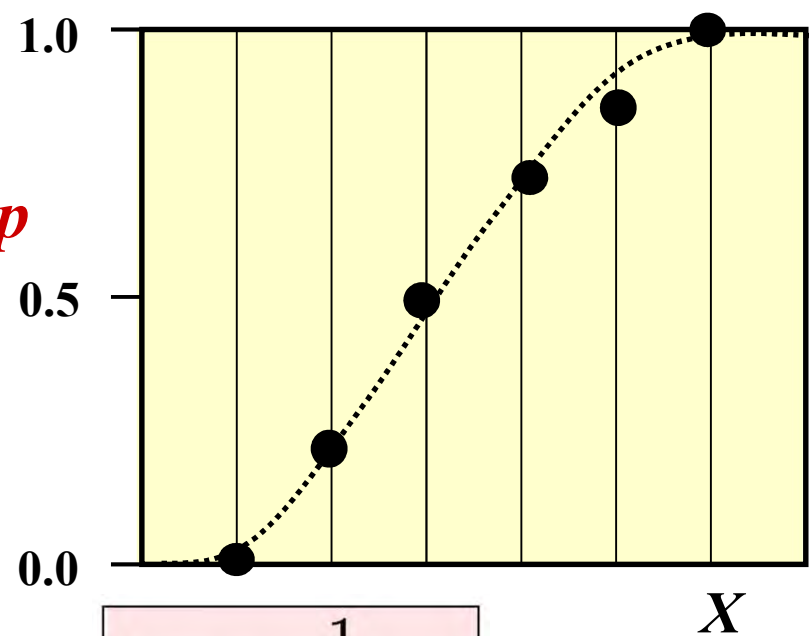


# 多重ロジスティック曲線による判別



多重ロジスティック回帰のメリットは  
複数検査の診断能を同時評価できること

多重ロジスティック曲線



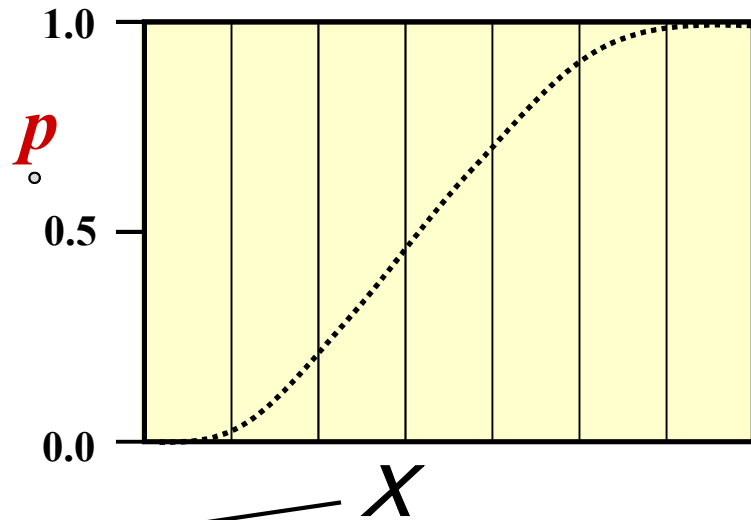
$$p = \frac{1}{1 + e^{-X}}$$

検査値の組み合わせ  
 $X = a_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$

任意の  $x_1 \sim x_k$  を代入 →  
 $p$  が大きいと疾患群と診断

# 多重ロジスティック分析：検査の診断的有用性の総合評価

疾患群に  
属する確率



$$p = \frac{1}{1 + e^{-X}}$$

疾患は高齢者に偏っている。  
検査Aは加齢で値が増加するが、  
検査Bより優っているか？

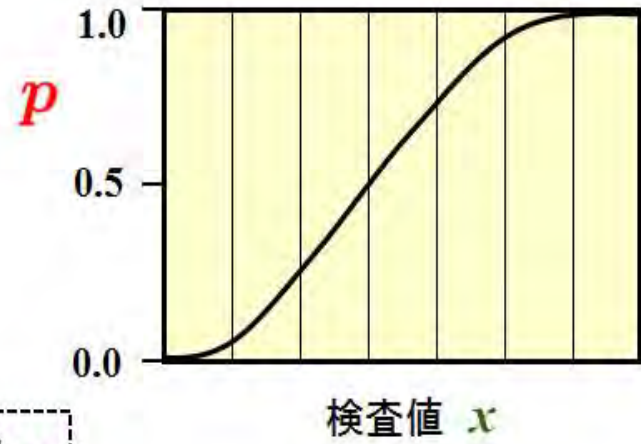
$$X = \beta_0 + \beta_1(\text{年齢}) + \beta_2(\text{検査A}) + \beta_3(\text{検査B})$$

多重ロジスティック回帰のメリットは  
交絡因子を制御しながら、複数検査の診断能を同時評価できること

## ロジスティック曲線の2つの表現型

疾患群に属する**確率**

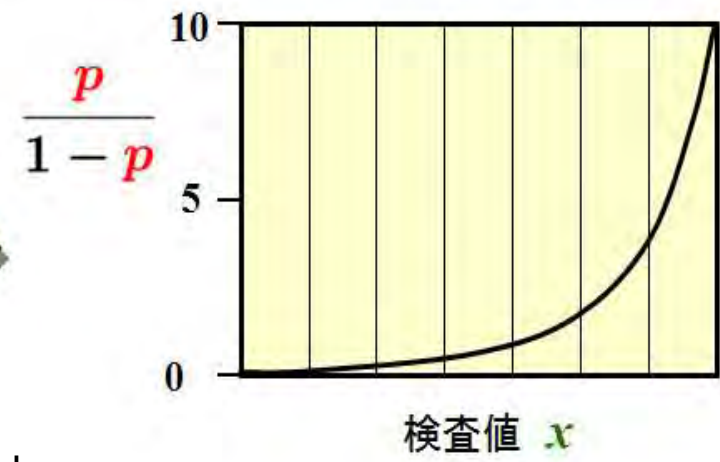
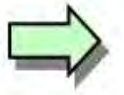
$$p = \frac{1}{1 + e^{-X}}$$



$X = a + bx$

疾患群に属する**オッズ**

$$\frac{p}{1-p} = \frac{1}{1 - \frac{1}{1 + e^{-X}}} = e^X$$



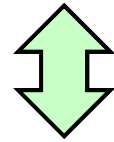
多重ロジスティック回帰の回帰係数の意味は、式を  $p = \frac{1}{1 + e^{-X}}$  でなく、 $Odd = e^X$  の式に置き換えると分かる



## 多重ロジスティック曲線の2つの表現型

疾患群に属する確率

$$p = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}}$$



疾患群に属するオッズ

$$\frac{p}{1 - p} = e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}$$

オッズ (O)

# オッズ比の計算：2値変量（ダミー変数）の場合

$x_1 = 1$  のとき

$$O_{x_1=1} = e^{a+b_1 \cdot 1 + b_2 x_2 + \dots + b_p x_p}$$

$x_1 = 0$  のとき

$$O_{x_1=0} = e^{a+b_1 \cdot 0 + b_2 x_2 + \dots + b_p x_p}$$

演算公式

$$e^0 = 1$$

$$e^{x+y} = e^x \cdot e^y$$

$$e^{x-y} = \frac{e^x}{e^y}$$

$$\text{オッズ比 } OR = \frac{O_{x_1=1}}{O_{x_1=0}} = \frac{e^{a+b_1 \cdot 1 + b_2 x_2 + \dots + b_p x_p}}{e^{a+b_1 \cdot 0 + b_2 x_2 + \dots + b_p x_p}} = e^{b_1}$$

2値変数が1の場合、  
0と比べ疾患群に属す  
オッズが何倍変わるか

指数演算の公式から  
分母、分子を成分分解  
すると、残るのは  $e^{b_1}$

回帰係数の  
指数を取ると  
オッズ比が求まる

## オッズ比の計算：連続変量の場合

$x_1 = x + \Delta$  のとき

$$O_{x_1=x+\Delta} = e^{\alpha + \beta_1(x+\Delta) + \beta_2x_2 + \dots + \beta_kx_k}$$

$x_1 = x$  のとき

$$O_{x_1=x} = e^{\alpha + \beta_1 \cdot x + \beta_2x_2 + \dots + \beta_kx_k}$$

回帰係数に  
変化量  $\Delta$  をかけて  
指数を取ると  
オッズ比が求まる

$$\text{オッズ比 } OR = \frac{O_{x_1=x+\Delta}}{O_{x_1=x}} = \frac{e^{\alpha + \beta_1(x+\Delta) + \beta_2x_2 + \dots + \beta_kx_k}}{e^{\alpha + \beta_1 \cdot x + \beta_2x_2 + \dots + \beta_kx_k}} = e^{\beta_1 \Delta}$$

説明変数が  $\Delta$  変化したとき、  
疾患群に属するオッズが何倍変化するか

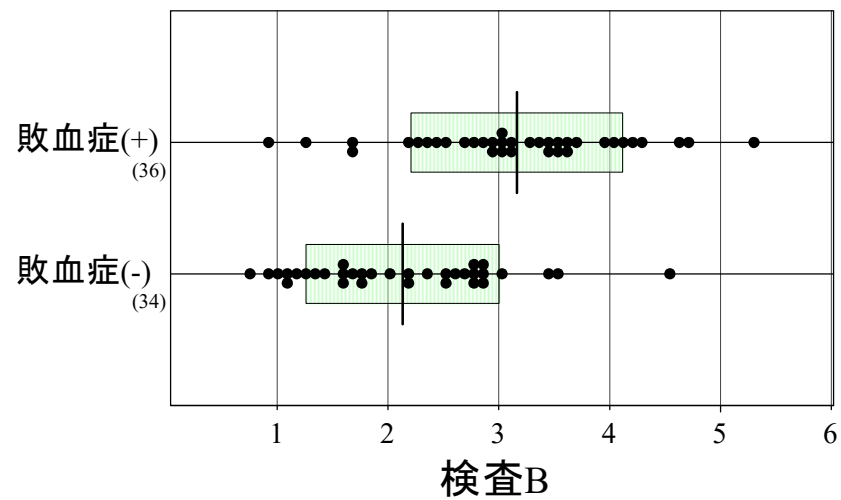
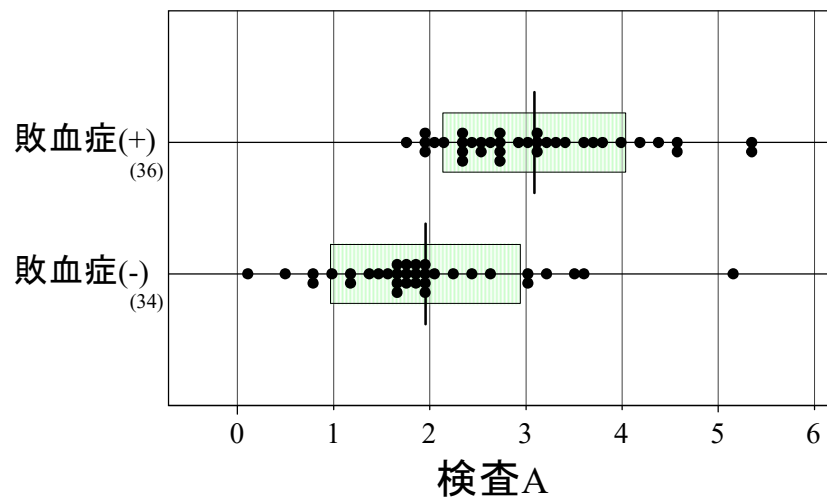
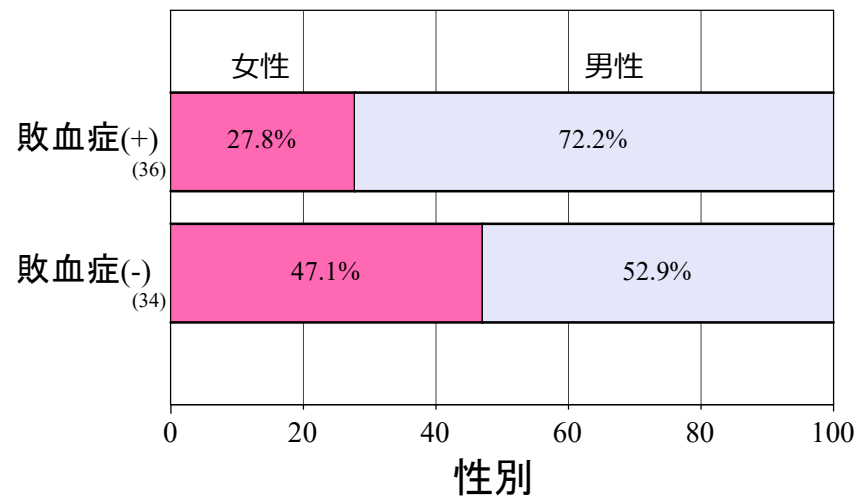
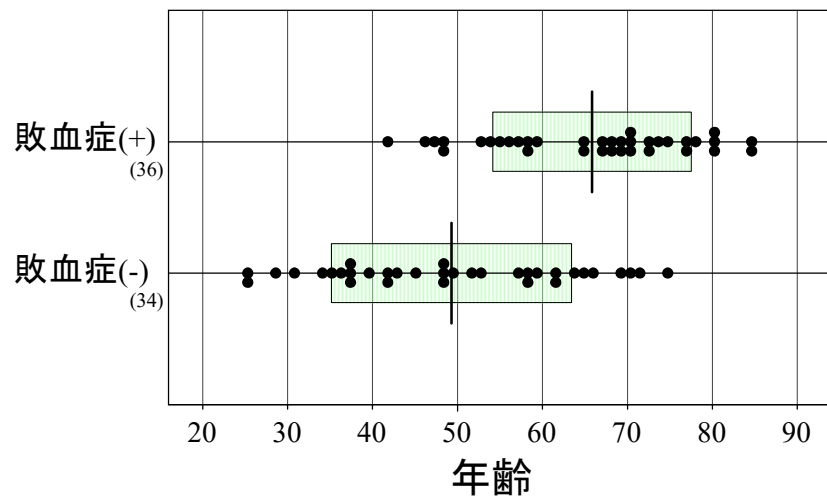
## 例題2：敗血症に対する2つの検査法の診断能の比較

	敗血症	年齢	性別	検査A	検査B
1	0	25	1	1.96	1.27
2	0	28	0	1.34	2.69
3	0	26	1	2.44	2.80
4	0	31	0	1.70	1.19
5	0	34	0	1.18	2.90
6	0	36	0	1.78	2.18
7	0	35	1	1.88	4.57
8	0	37	0	1.70	2.00
9	0	38	0	0.48	1.69
10	1	42	0	1.78	2.25
11	0	39	0	1.00	0.73
12	0	38	0	1.70	1.75
13	1	48	1	2.72	3.08
14	0	42	0	1.78	2.52
15	0	42	0	0.08	1.04
16	0	46	1	2.22	2.75
17	0	42	1	0.78	1.61
18	1	49	0	3.42	4.27
19	1	46	0	2.38	2.56
20	1	48	1	3.16	3.54
21	0	49	0	1.18	1.45
22	0	48	1	1.96	2.82
23	1	52	0	1.96	4.05
24	1	57	0	3.74	4.70

	敗血症	年齢	性別	検査A	検査B
25	0	48	1	1.88	1.58
26	1	54	0	3.80	3.00
27	0	52	0	2.68	1.38
28	0	53	1	1.70	1.74
29	1	55	1	2.72	2.95
30	1	56	0	2.76	2.84
31	1	58	0	1.96	3.73
32	0	50	1	1.88	2.82
33	1	59	0	3.00	3.63
34	1	58	0	3.30	1.72
35	0	57	0	0.78	3.04
36	0	58	1	3.06	2.60
37	1	65	0	2.34	2.96
38	0	58	1	3.26	3.43
39	0	60	1	1.58	2.36
40	0	61	1	3.50	1.57
41	1	67	1	4.22	0.95
42	1	69	0	2.44	3.49
43	1	65	0	2.56	4.65
44	1	69	0	2.94	2.37
45	1	67	0	2.56	4.25
46	1	68	1	2.38	2.20
47	1	71	0	3.10	1.29
48	0	62	0	1.96	1.06

	敗血症	年齢	性別	検査A	検査B
49	0	66	1	3.58	1.81
50	0	64	0	3.06	0.95
51	1	70	0	2.68	3.96
52	0	65	0	1.78	2.23
53	1	74	0	3.18	3.44
54	1	73	0	4.58	3.32
55	1	71	0	2.02	3.05
56	1	72	1	3.08	2.99
57	1	70	0	2.60	2.67
58	0	69	0	1.96	2.51
59	0	71	0	2.02	1.12
60	1	77	1	4.34	1.72
61	0	71	1	5.18	2.84
62	1	74	0	2.38	3.34
63	0	75	1	1.48	3.50
64	1	77	0	3.56	5.33
65	1	80	0	4.02	3.10
66	1	79	1	2.10	3.53
67	1	84	0	5.38	4.13
68	1	80	1	5.38	2.74
69	1	81	1	1.96	2.43
70	1	85	0	4.62	3.63

敗血症(+) 36例  
敗血症(-) 34例



目的変数 =  
敗血症の有無

P値から、同等の診断能

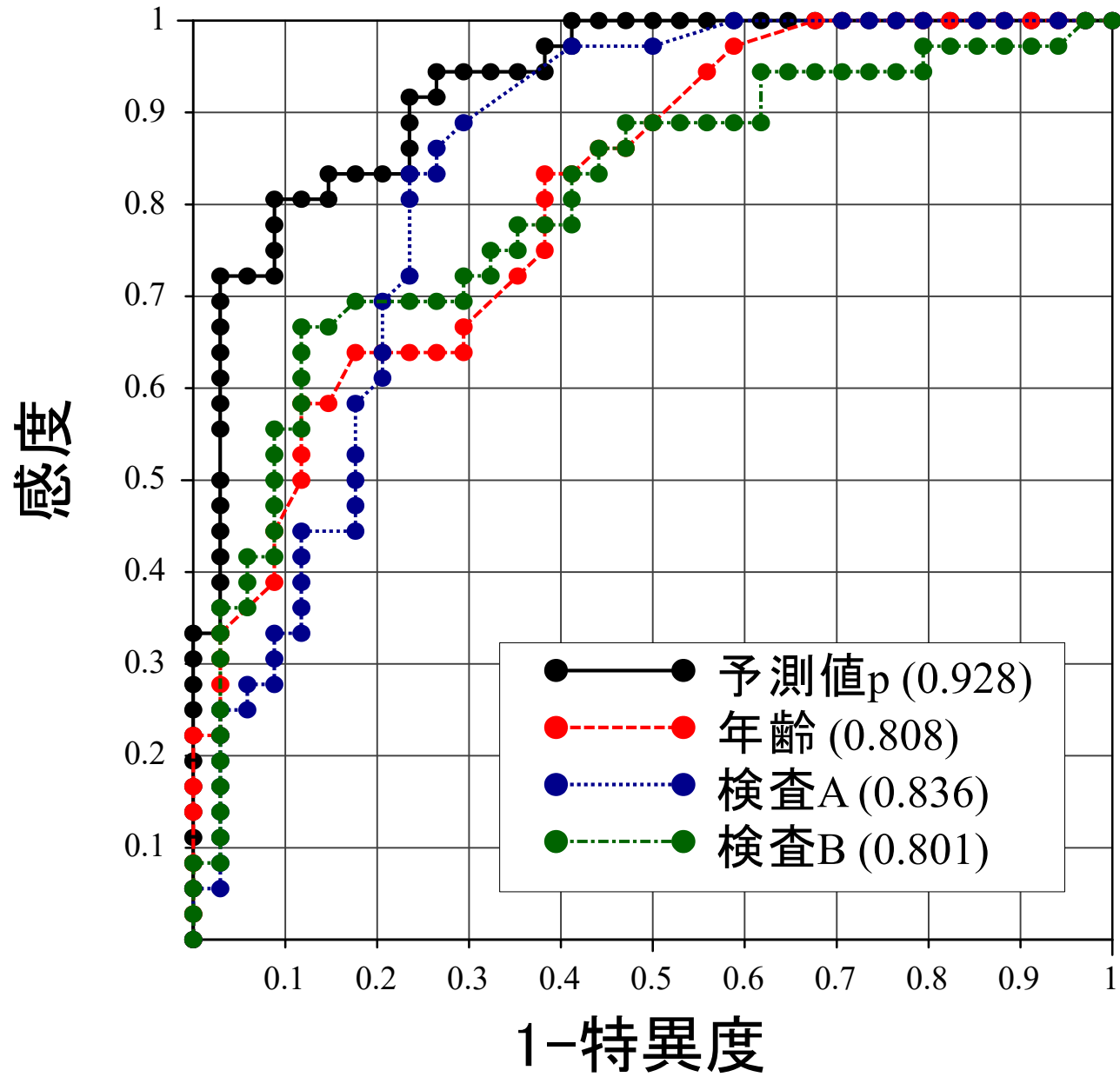
年齢の追加でstdβが低下

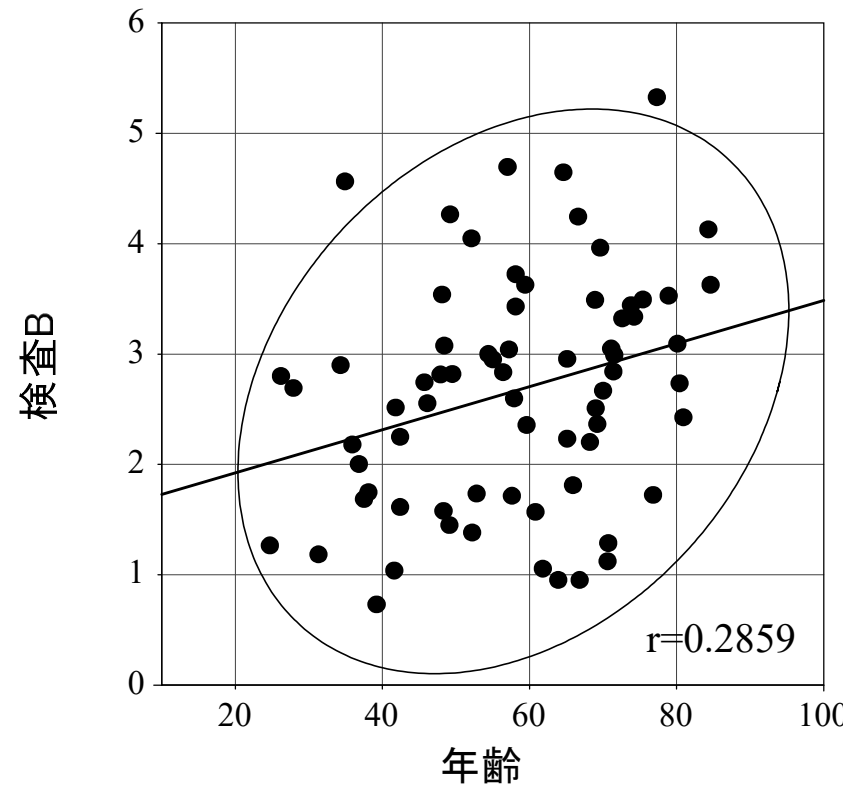
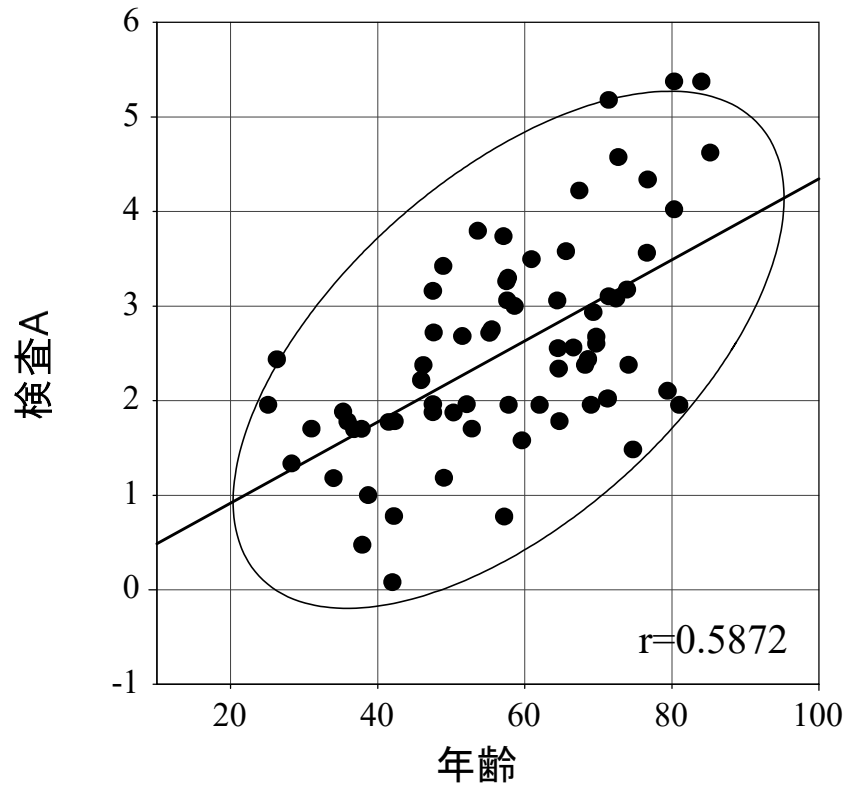
年齢の追加でstdβ変化せず

P値で判断すると、診断能は  
検査B、年齢、検査Aの順となる

P値から性別も診断に貢献しており  
4指標の組合せ診断がベストと言える

	変数名	$\beta$	SE( $\beta$ )	z	P	OR	$\Delta\beta$	95%CI		AUC
1	検査A	1.346	0.369	3.648	0.0003	3.841	1	1.864	7.914	0.836
1	検査B	1.232	0.333	3.693	0.0002	3.426	1	1.782	6.587	0.801
1	年齢	0.067	0.026	2.610	0.0091	1.954	10	1.181	3.231	0.852
2	検査A	0.941	0.406	2.316	0.0206	2.563	1	1.156	5.682	
1	年齢	0.096	0.028	3.457	0.0005	2.611	10	1.515	4.498	0.883
2	検査B	1.168	0.361	3.235	0.0012	3.217	1	1.585	6.529	
1	年齢	0.072	0.030	2.412	0.0159	2.060	10	1.145	3.706	0.903
2	検査A	0.960	0.457	2.101	0.0356	2.612	1	1.067	6.394	
3	検査B	1.211	0.392	3.084	0.0020	3.355	1	1.555	7.241	
1	年齢	0.072	0.030	2.363	0.0182	2.047	10	1.130	3.709	0.928
2	検査A	1.339	0.526	2.545	0.0109	3.814	1	1.361	10.691	
3	検査B	1.442	0.468	3.079	0.0021	4.228	1	1.689	10.584	
4	性別	-1.962	0.832	-2.358	0.0184	0.141	1	0.028	0.718	





検査Aが見かけ上、診断能が高かったのは年齢と相関があったため。  
すなわち、疾患群で年齢が高かったため、交絡現象が生じたといえる。



### 例題3: HCVによる肝硬変(LC)と肝細胞癌(HCC) 鑑別に用いる検査診断マーカの探索

**研究目的** : 線維化の程度を揃えて、臨床検査値だけでLCとHCCを鑑別する

HCC	Age	Sex	FIB-4	PIVKA	AFP	ALT	AST	LDH	GGT	ALP	ChE	TBil	TP	Alb	Glu	Cre	UN	K	Hb	WBC	PLT	PT
1	81	1	4.16		137.2	48	45	205	52	215	107	0.6	6.8	3.65	95	0.87	12	4.05	11.9	4900	12.5	93.6
0	70	1	1.80	11.3	1.0	27	21	229	92			0.7	7.1	4.30		1.13	24		12.0	9760	15.7	90.8
1	73	1	14.75	33.4	95.4	148	145	228	78	325	135	0.9	7.9	3.60	87	0.78	25	4.00	13.5	3900	5.9	78.0
1	74	0	3.93	26.6		27	41	232	18	308	169	0.5	8.2	3.91	93	0.55	12	4.55	12.7	4950	14.9	84.7
0	75	1	6.16	22.8	3.7	16	23	169	34	192	292	1.4	7.7	4.90	103	0.86	21	4.20	14.3	4580	7.0	85.4
0	63	1	1.30		1.7	15	17		31			0.5	6.7	4.00		1.23	27	4.30	14.1	9610	21.3	
0	74	0	3.60		6.1	19	25		57	239		1.1	8.3	4.10	96	0.82	16	4.20	13.9	5280	11.8	
1	69	0	6.62		16.5	39	66	296	42	303	142	1.2	6.0	3.10		1.17	20	3.17	12.3	4890	11.0	
1	72	1	4.90		8.3	24	39	235	23	168	101	0.8	8.5	3.30	87	1.56	66	4.10	9.5	3370	11.7	
0	56	1	0.47		4.0	13	14	131	55	286		0.4	6.8	3.80	95	0.63	13	5.00	13.0	6570	46.0	
0	72	1	1.90	27.7	3.4	27	23	174	37	239	282	0.7	7.2	4.20	174	0.99	15	4.60	13.0	4220	16.8	98.8
0	67	0	1.94	20.8	3.8	8	19	192	17	321		0.8		4.10	135	0.50	15		14.0	3860	23.2	
0	41	0	0.97		3.8	23	24	268	35	409	356	0.6	7.0	3.60	99	0.68	8	3.70	12.9	11450	21.1	
0	74	0	5.36	11.0	6.9	28	51	251	22	262	133	0.8	6.8	3.30	114	0.90	20		6.2	3860	13.3	
0	74	0	6.05		4.9	11	28	505	23	127		0.8	7.3	4.30	92	0.61	12	3.63	11.0	5157	10.2	
0	66	1	1.65	16.3	7.8	29	37	178	24	260	313	0.5	7.5	3.80	114	0.72	12	4.40	13.0	7625	27.1	
0	34	1	0.65		1.0	28	23	194	16	267		0.9	7.0	4.40	87	1.00	12	3.70	15.0	3850	22.7	101.7
0	60	1	2.59	64.5	3.1	32	45	136	86	264	351	0.4	7.3	4.00	108	0.57	16	4.40	12.5	4310	18.4	95.5
0	79	0	1.77	12.1	2.9	12	18		15	261	233	0.5	7.1	3.80	93	0.93	23	4.70	9.3	3960	23.2	98.6

n=452 (HCC 129例、LC 323 例)

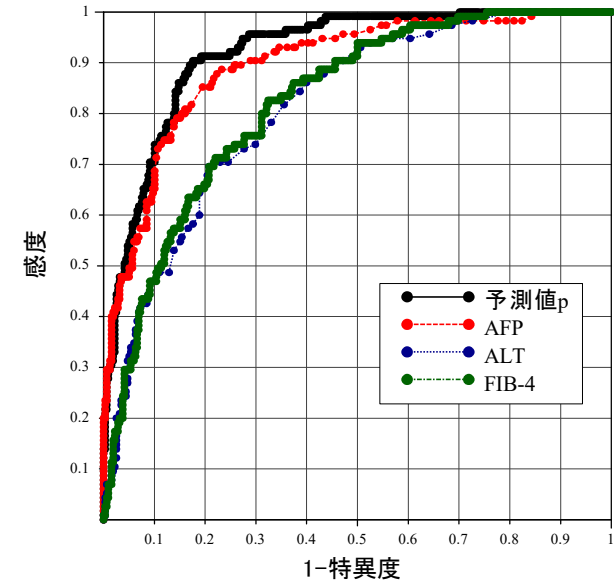
目的変数：HCC 有効データ数=433 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-7.3821	0.84831			
1	AFP	0.91077	0.15146	6.0132	0.000000	2.4862
2	ALT	1.0327	0.25087	4.1165	0.000039	2.8087
3	FIB-4	0.73387	0.24448	3.0018	0.002683	2.0831

< 回帰の適合度指標 >

AIC=286.0616, AUC=0.9197

$$\text{FIB-4} = \frac{\text{年齢} \times \text{AST}}{\text{PLT} \times \sqrt{\text{ALT}}}$$



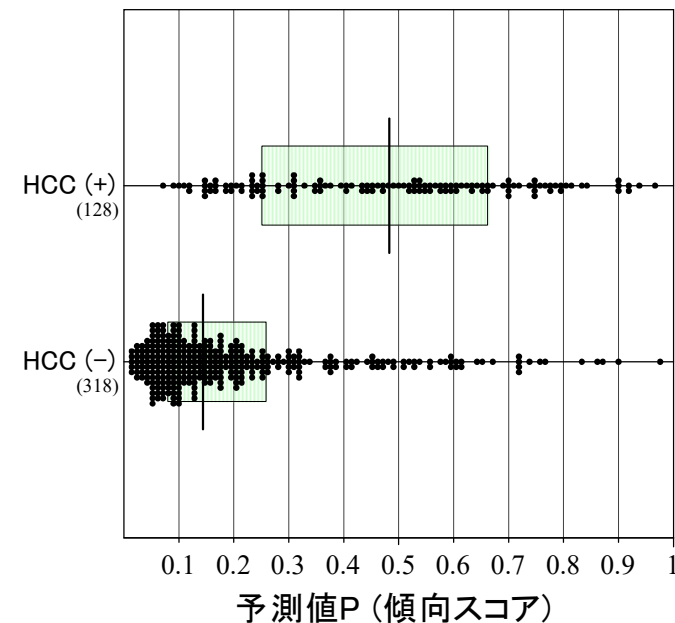
Age, Sex, Fib-4から、各症例のHCCらしさ  
(傾向スコア) を、MLRAで求める

目的変数：HCC 有効データ数=446 [第1頁：群1]

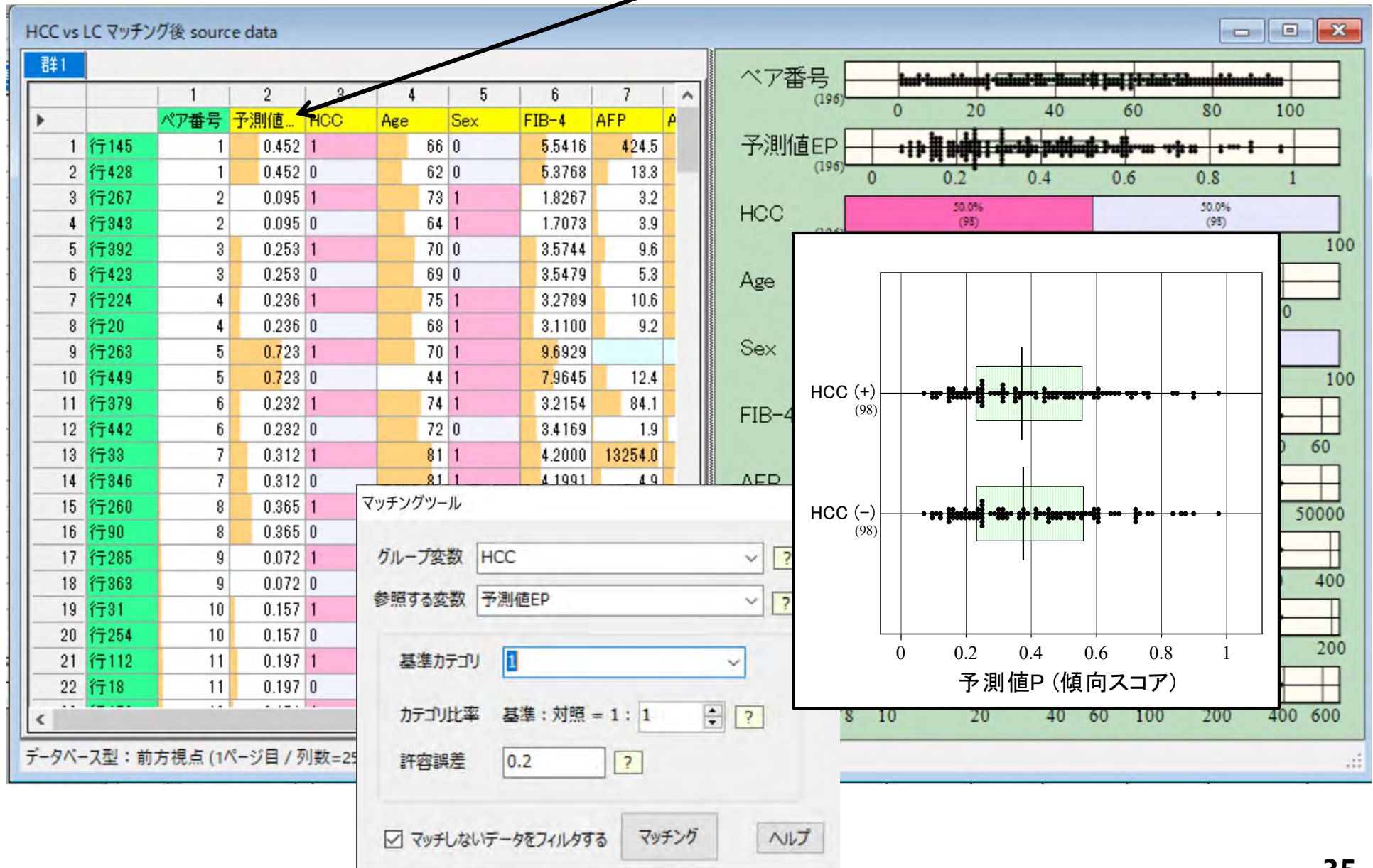
次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-2.4992	0.96608			
1	Age	-0.014346	0.013327	-1.0765	0.281706	0.8664
2	Sex	0.14463	0.24801	0.5832	0.559778	1.1556
3	FIB-4	1.9004	0.20592	9.2288	0.000000	6.6885

< 回帰の適合度指標 >

AIC=418.3315, AUC=0.8288



MLRAにより性別・年齢・FIB-4からHCCの傾向スコア（予測確率）を求め、  
HCC : LC = 1 : 1で症例をマッチング





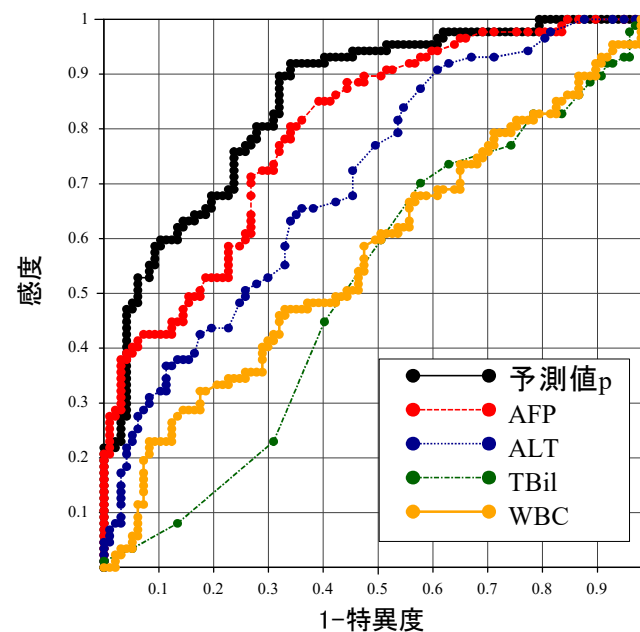
目的変数：HCC 有効データ数=184 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-7.9131	1.4994			
1	AFP	0.98713	0.20234	4.8786	0.000001	2.6835
2	ALT	0.87235	0.32283	2.7022	0.006889	2.3925
3	TBil	-0.92558	0.42226	-2.1920	0.028380	0.3963
4	WBC	0.021824	7.4114E-3	2.9447	0.003233	8.8677

< 回帰の適合度指標 >

AIC=184.7631, AUC=0.8508

肝線維化の程度が同じ条件で、  
HCCとLCの鑑別が、4検査の組合せ  
によりの的確に行える



## 多重ロジスティック分析のポイント

- 回帰係数から個々の要因（検査）の重要性を示す  
調整オッズ比（aOR）と信頼区間が求まる
- aORは、説明変数の単位に依存するので、  
必ず適切な変化幅を指定して計算
- 説明変数の有用性比較にはP値を用いる。aORの比較は困難
- 複数所見の組み合わせで、診断方程式を作成できる  
複合診断の判別度を、ROC分析によりAUCで表せる
- 総症例数が少ないと、過剰適合が生じ再現性に問題あり  
(総症例数  $\geq 100$  が、めやす)
- 疾患群の割合が少ないと、分析精度が落ちる  
疾患症例数  $N$  に対して、説明変数の数の許容限界は  $N/10$

# 省略したスライド

## 例題4：突発性発疹(ES)を対象とした熱性けいれん(FS)と鉄欠乏性貧血の関連性の分析

井上佳也、他：日本小児科学会雑誌 125(6): 883-891, 2021

痙攣	家族歴	月齢	性別	体重	来院時体温	WBC	RBC	Hb	Ht	MCV	MCH	MCHC	PLT	Lym	Mon	Gra	RDW	PCT	MPV	PDW	CRP
0	0	3	1	6.0	38.0	38	364	10.7	31.1	85.4	29.4	34.4	16.7	39.4	9.9	50.7	10.4	0.14	8.2	19.1	0.80
0	0	4	0	6.4	37.7	67	406	10.2	29.2	71.9	25.1	34.9	15.1	67.4	4.2	28.4	12.2	0.15	9.7	17.6	0.40
0	0	6	1	6.5	38.9	109	399	11.0	32.1	80.5	27.6	34.3	13.5	39.4	9.6	51.0	10.6	0.12	8.6	19.4	0.50
0	0	6	1	7.4	39.1	83	454	11.0	32.6	71.8	24.2	33.7	26.3	36.6	2.5	60.9	11.4	0.23	8.6	20.7	0.20
0	0	6	0	8.0	37.8	103	518	14.8	41.7	80.5	28.6	35.5	25.0	54.7	4.8	40.5	11.6	0.20	7.9	19.1	0.20
0	0	7	1	6.8	38.3	102	433	11.3	34.8	80.4	26.1	32.5	19.2	45.2	3.7	51.1	11.2	0.15	7.9	19.1	1.20
0	0	7	1	8.0	39.9	41	357	9.8	28.6	80.1	27.5	34.3	28.8	31.7	11.4	56.9	12.0	0.23	7.9	18.3	1.20
0	0	7	1	9.0	39.2	65	411	11.4	31.8	77.4	27.7	35.8	25.2	39.7	6.3	54.0	11.4	0.19	7.7	18.4	0.90
0	0	7	1	8.3	39.5	78	498	12.2	35.4	71.1	24.5	34.5	16.9	73.7	2.8	23.5	11.4	0.14	8.4	18.7	0.05
0	0	7	1	8.1	40.0	57	347	10.2	29.1	83.9	29.4	35.1	13.7	26.6	4.1	69.3	10.8	0.12	8.6	18.8	0.40
0	0	7	0	8.0	39.8	99	511	13.0	39.2	76.7	25.4	33.2	21.2	30.1	4.5	65.4	11.5	0.17	8.0	19.2	1.90
0	0	7	0	7.0	39.0	58	486	11.9	35.1	72.2	24.5	33.9	19.5	54.4	6.5	39.1	13.2	0.15	7.7	18.4	0.40
0	0	7	0	7.0	39.2	60	384	9.5	27.1	70.6	24.7	35.1	30.3	58.3	12.6	29.1	14.5	0.23	7.7	18.4	0.30
0	0	8	1	10.0	38.5	88	498	13.5	38.6	77.5	27.1	35.0	21.2	58.8	11.0	30.2	12.1	0.16	7.5	17.6	0.05
0	0	8	1	10.0	39.5	55	447	10.9	32.0	71.6	24.4	34.1	24.6	49.6	7.7	42.7	12.2	0.21	8.6	18.2	0.05
0	0	8	1	7.0	39.9	38	417	11.5	34.2	82.0	27.6	33.6	24.4	42.5	14.2	43.3	10.7	0.19	7.7	18.9	0.50
0	0	8	1	9.6	38.6	36	435	12.5	35.2	80.9	28.7	35.5	15.8	45.3	2.9	51.8	11.6	0.12	7.5	18.9	1.10
0	0	8	1	8.3	38.1	78	596	15.0	46.0	77.2	25.2	32.6	7.5	45.4	4.8	49.8	10.8	0.07	9.1	18.6	0.10
0	0	8	1	7.6	37.3	42	475	13.6	38.5	81.1	28.6	35.3	19.2	90.1	6.3	3.6	11.3	0.15	7.7	20.3	0.10
0	0	8	1	7.5	39.0	96	494	10.4	31.8	64.4	21.1	32.7	26.9	41.2	3.2	55.6	13.3	0.23	8.4	16.3	0.05
0	0	8	0	8.0	38.8	89	488	11.4	33.8	69.3	23.4	33.7	35.0	51.6	7.0	41.4	12.5	0.27	7.6	17.1	0.30
0	0	8	0	8.3	38.5	59	540	12.6	36.3	67.2	23.3	34.7	25.5	65.4	5.3	29.3	14.8	0.20	7.7	16.3	0.05
0	0	8	0	9.0	38.7	50	420	11.5	33.9	80.7	27.4	33.9	10.6	41.6	3.4	55.0	10.9	0.09	8.6	18.2	0.50
0	0	8	0	8.2	38.6	62	436	10.9	31.7	72.7	25.0	34.4	24.6	54.0	3.5	42.5	12.6	0.21	8.4	18.1	0.90

n=407 (熱性痙攣 30例、対照374例)

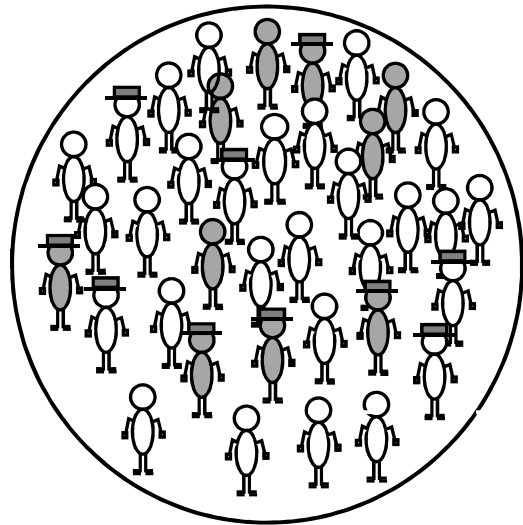
# 横断的研究

(cross sectional study)

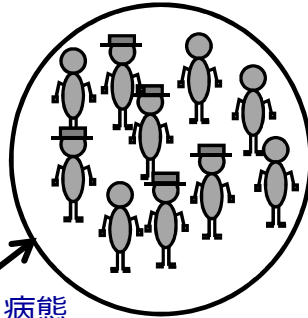
# 集団調査型 (1) 診断マーカの探索

field survey

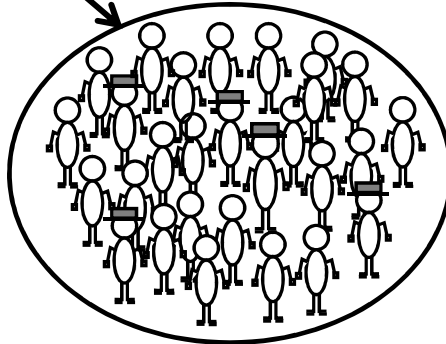
疫学研究の対象となる  
一群の集団 (ES)



疾患 (FS+)

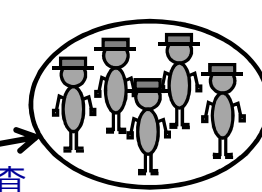


非疾患 (FS-)

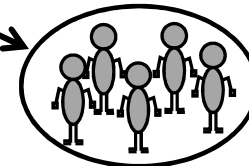


多重ロジスティック分析

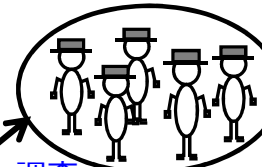
調査



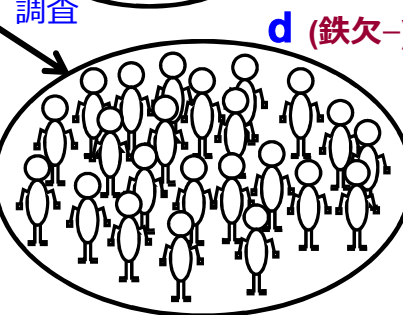
a  
(鉄欠+)



c  
(鉄欠-)



b  
(鉄欠+)

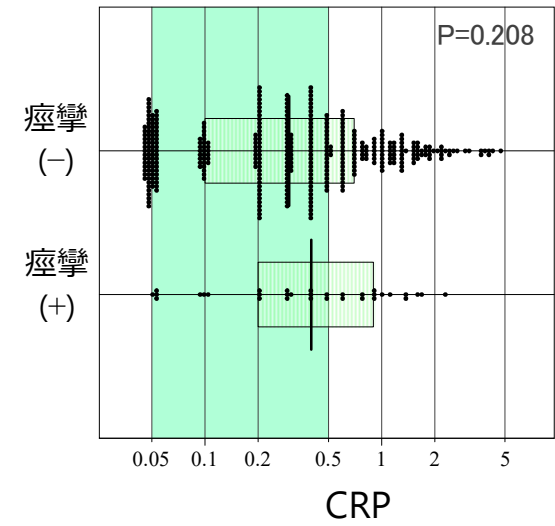
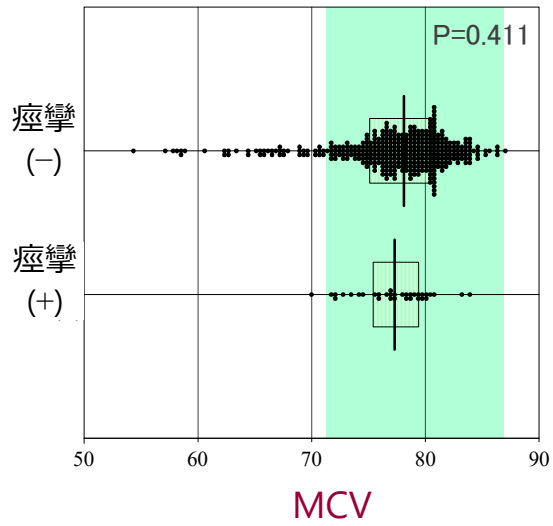
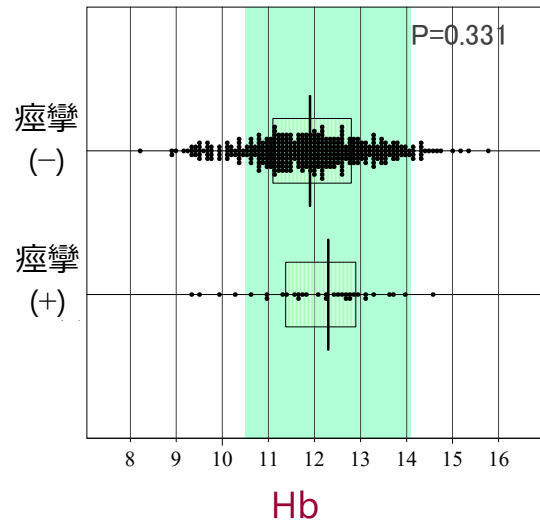
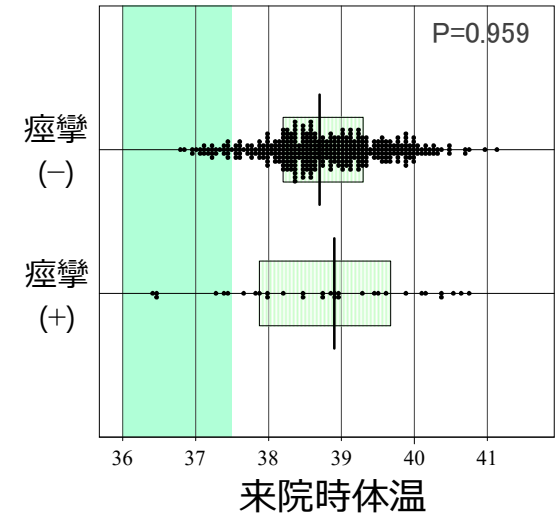
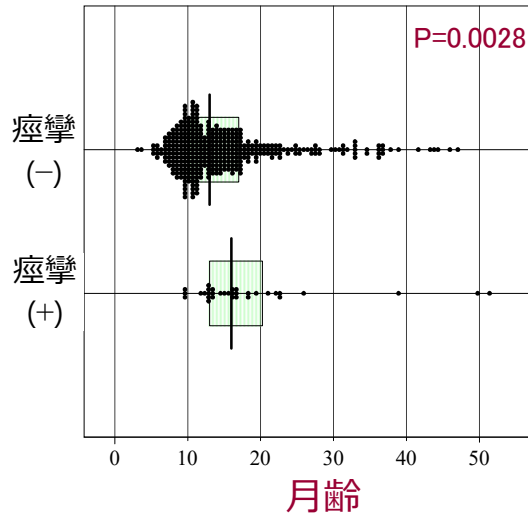
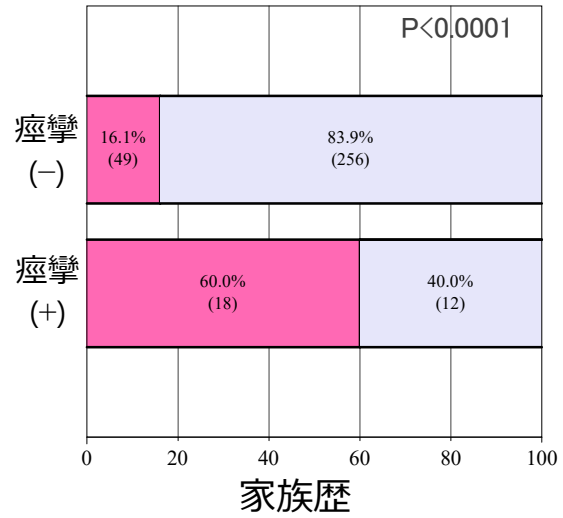


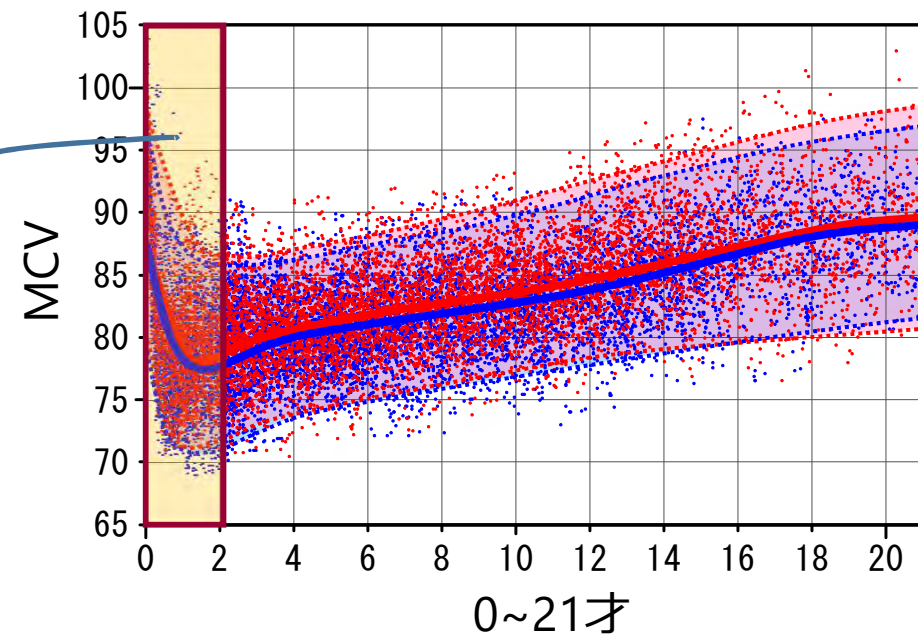
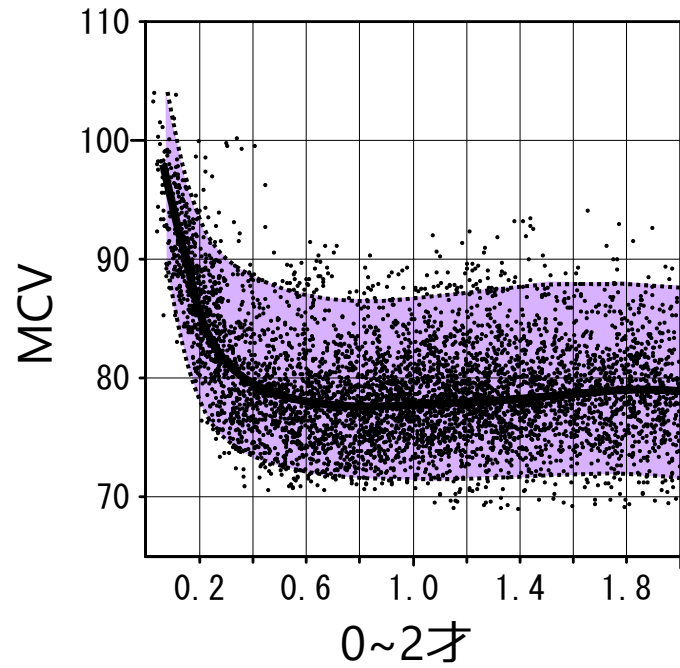
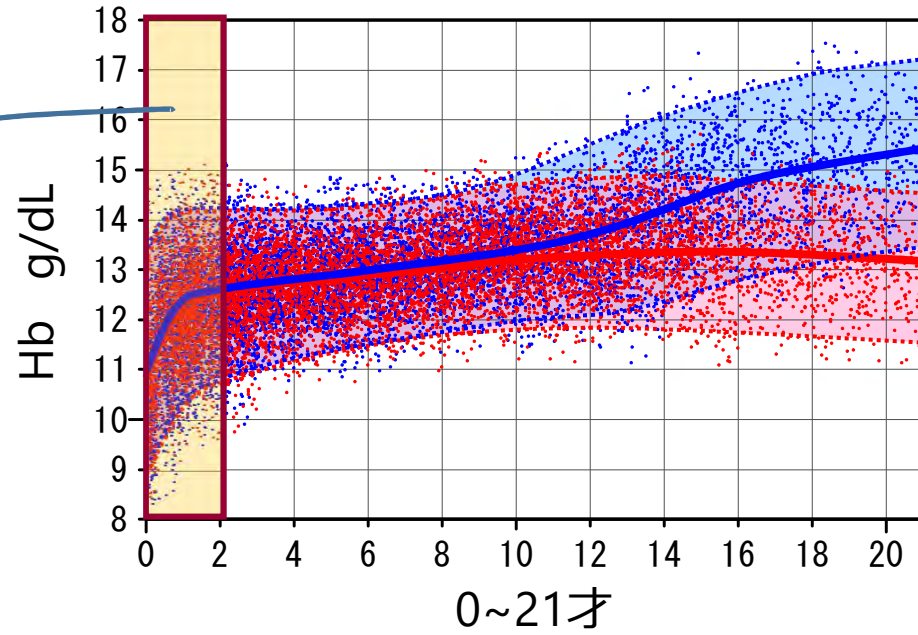
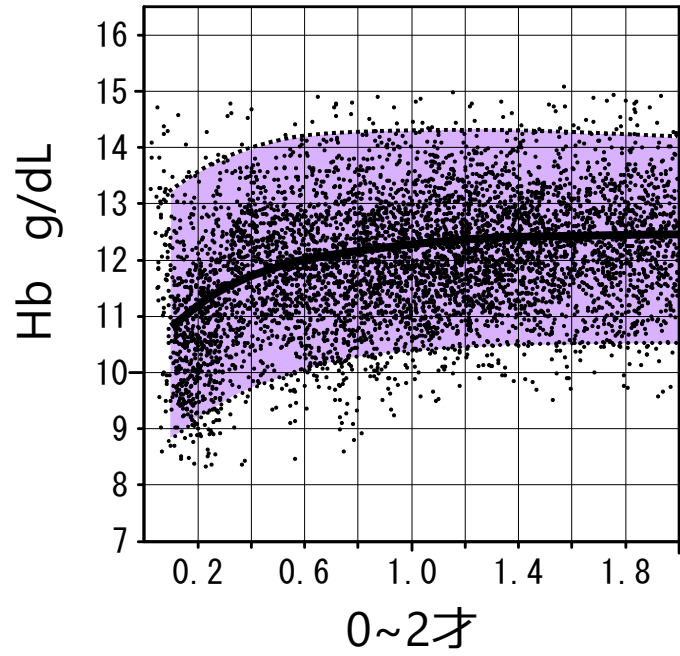
d  
(鉄欠-)

診断に有用な  
臨床・検査所見  
を探索

$$OR = \frac{a/c}{b/d}$$





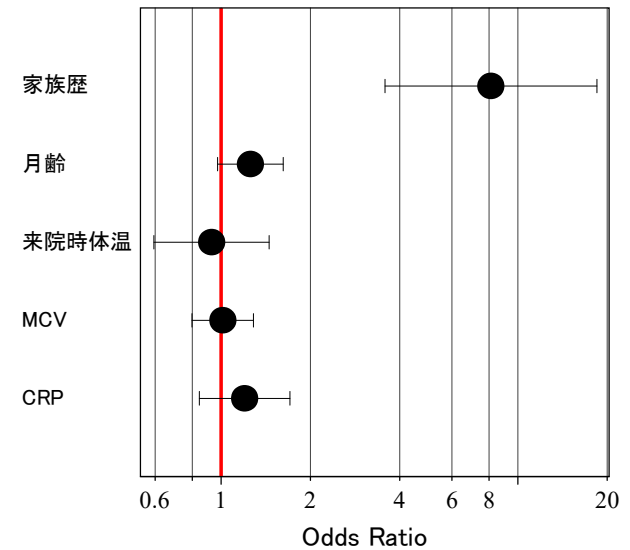


目的変数：痙攣 有効データ数=334 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-1.0304	9.4263			
1	家族歴	2.0930	0.41930	4.9917	0.000001	8.1093
2	月齢	0.038078	0.021624	1.7609	0.078258	1.2567
3	来院時体温	-0.072809	0.22788	-0.3195	0.749348	0.9298
4	MCV	4.5718E-3	0.040654	0.1125	0.910462	1.0138
5	CRP	0.18322	0.17892	1.0240	0.305828	1.2011

< 回帰の適合度指標 >

AIC=182.6863, AUC=0.7846

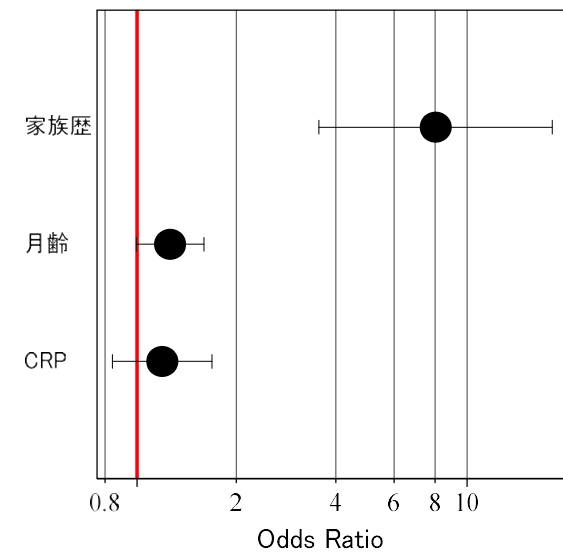


目的変数：痙攣 有効データ数=334 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-3.5110	0.48946			
1	家族歴	2.0821	0.41478	5.0197	0.000001	8.0213
2	月齢	0.038528	0.020010	1.9254	0.054175	1.2601
3	CRP	0.17554	0.17742	0.9894	0.322470	1.1919

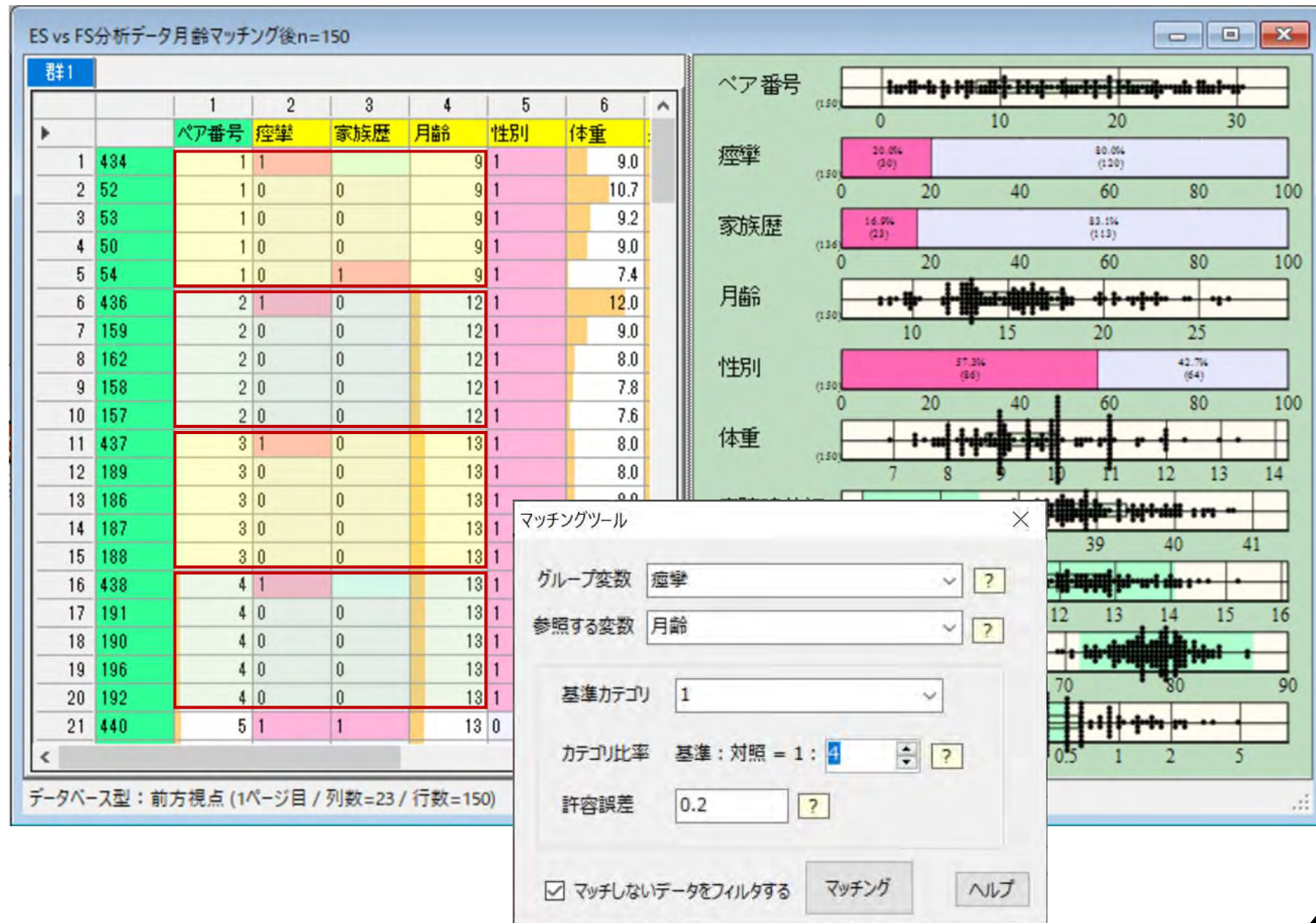
< 回帰の適合度指標 >

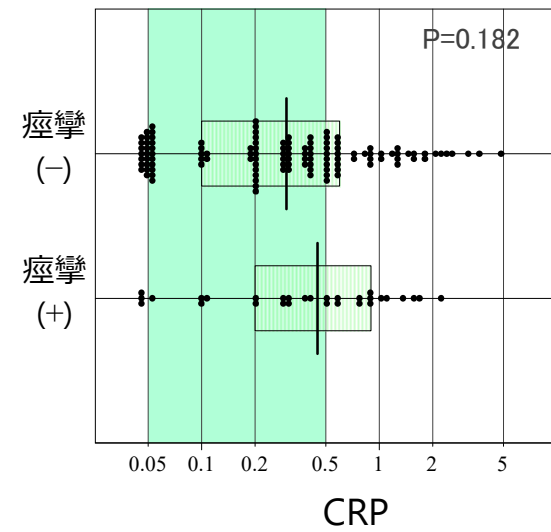
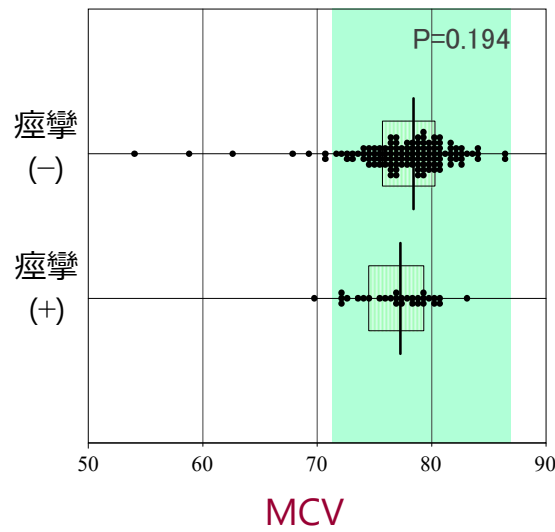
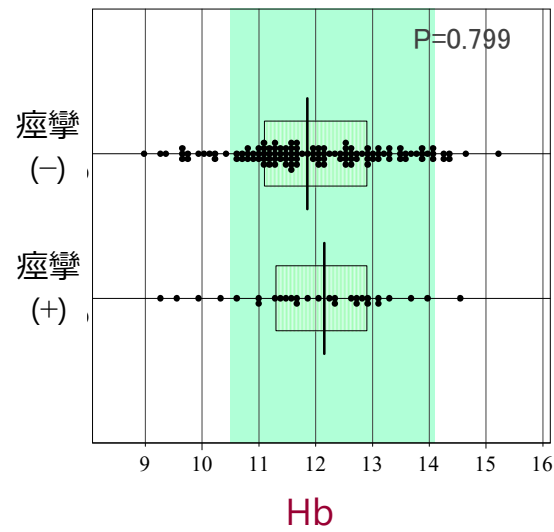
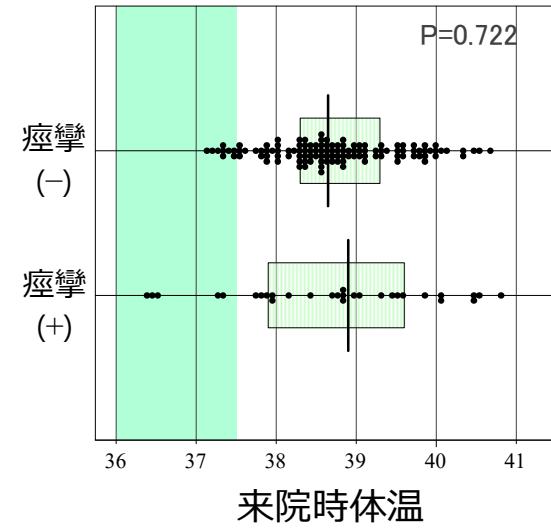
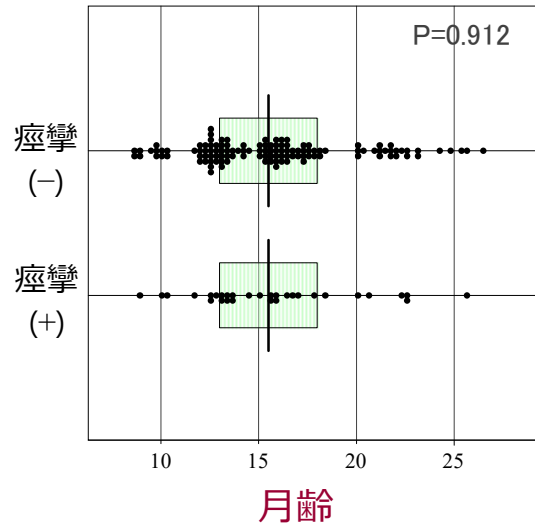
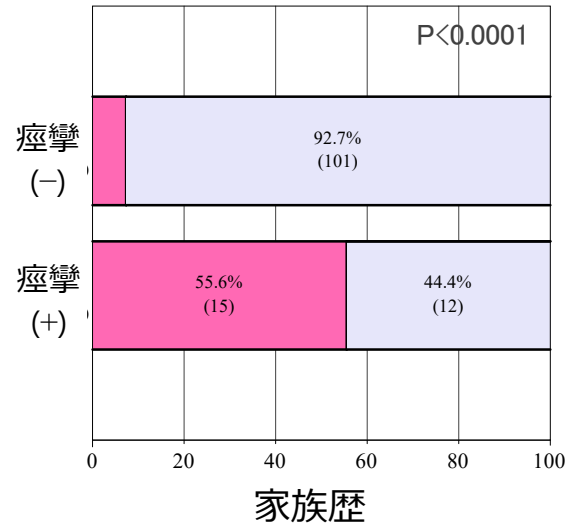
AIC=178.8033, AUC=0.7847





## 月齢をマッチングさせたデータによる再分析



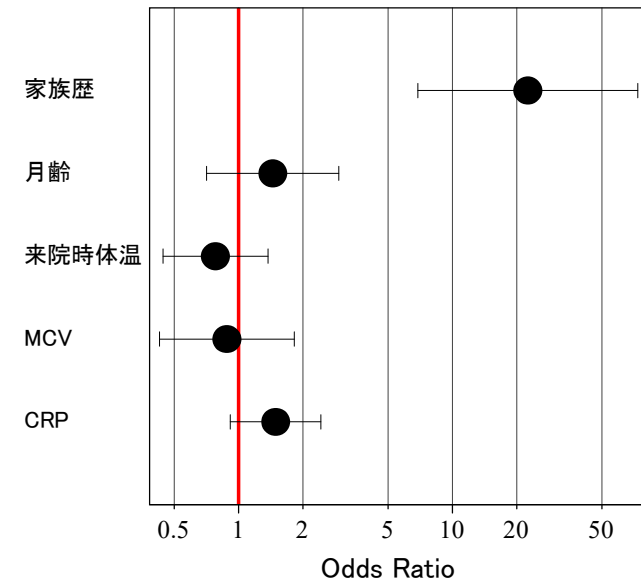


目的変数：痙攣 有効データ数=136 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		8.4905	12.538			
1	家族歴	3.1148	0.60391	5.1577	0.000000	22.5284
2	月齢	0.061439	0.060523	1.0151	0.310040	1.4458
3	来院時体温	-0.24856	0.28836	-0.8620	0.388698	0.7799
4	MCV	-0.020898	0.061772	-0.3383	0.735135	0.8822
5	CRP	0.39969	0.24840	1.6090	0.107606	1.4914

< 回帰の適合度指標 >

AIC=113.9606, AUC=0.8050

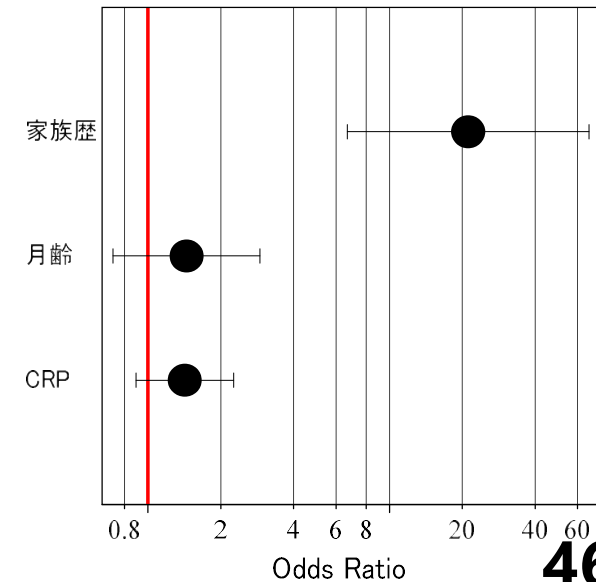


目的変数：痙攣 有効データ数=136 [第1頁：群1]

次数	変数名	$\beta$	SE( $\beta$ )	z	P	オッズ比
0		-2.7900	1.0783			
1	家族歴	3.0511	0.58811	5.1879	0.000000	21.1377
2	月齢	0.061184	0.059539	1.0276	0.304123	1.4435
3	CRP	0.35006	0.23741	1.4745	0.140349	1.4191

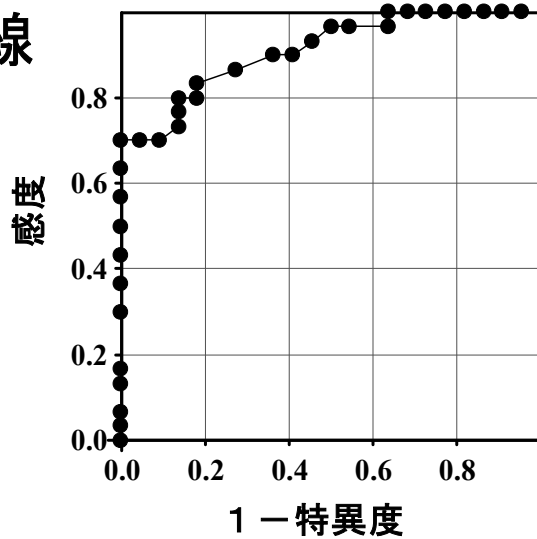
< 回帰の適合度指標 >

AIC=110.7915, AUC=0.7992



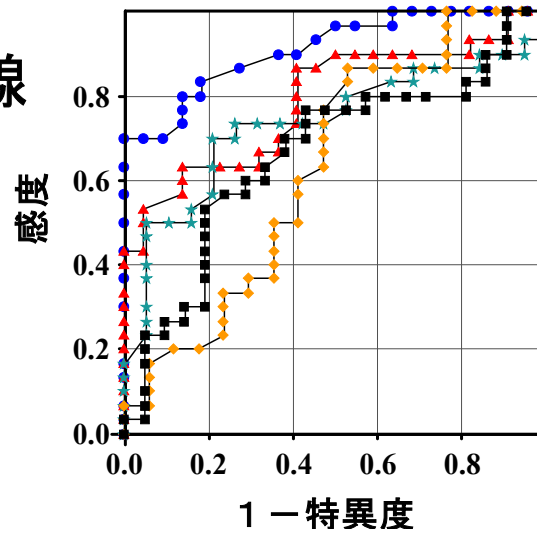
# 補足スライド

# ROC曲線

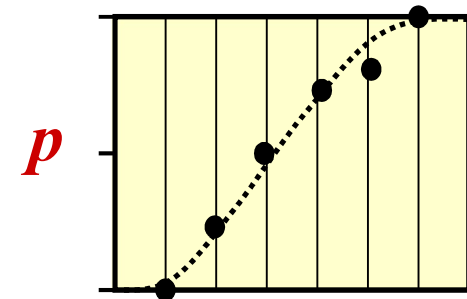


= 1変量ロジスティック曲線

# 複合 ROC曲線

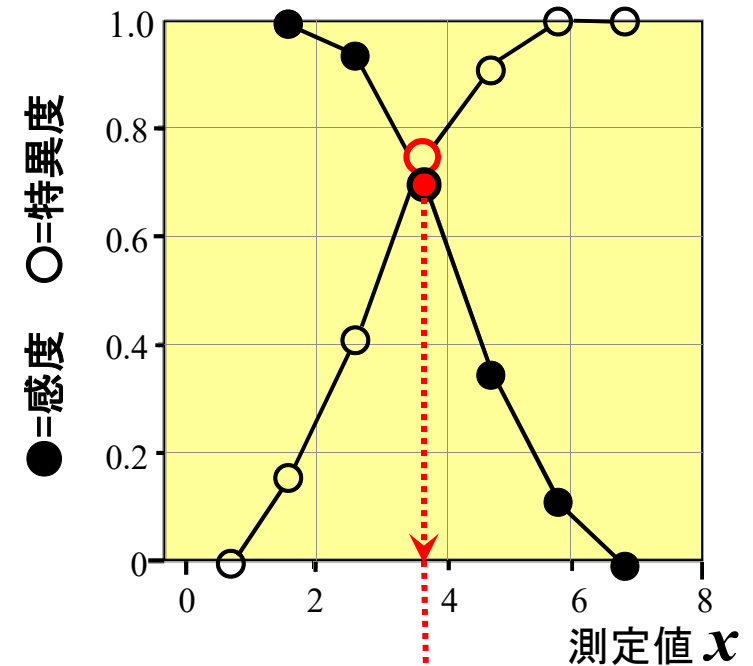
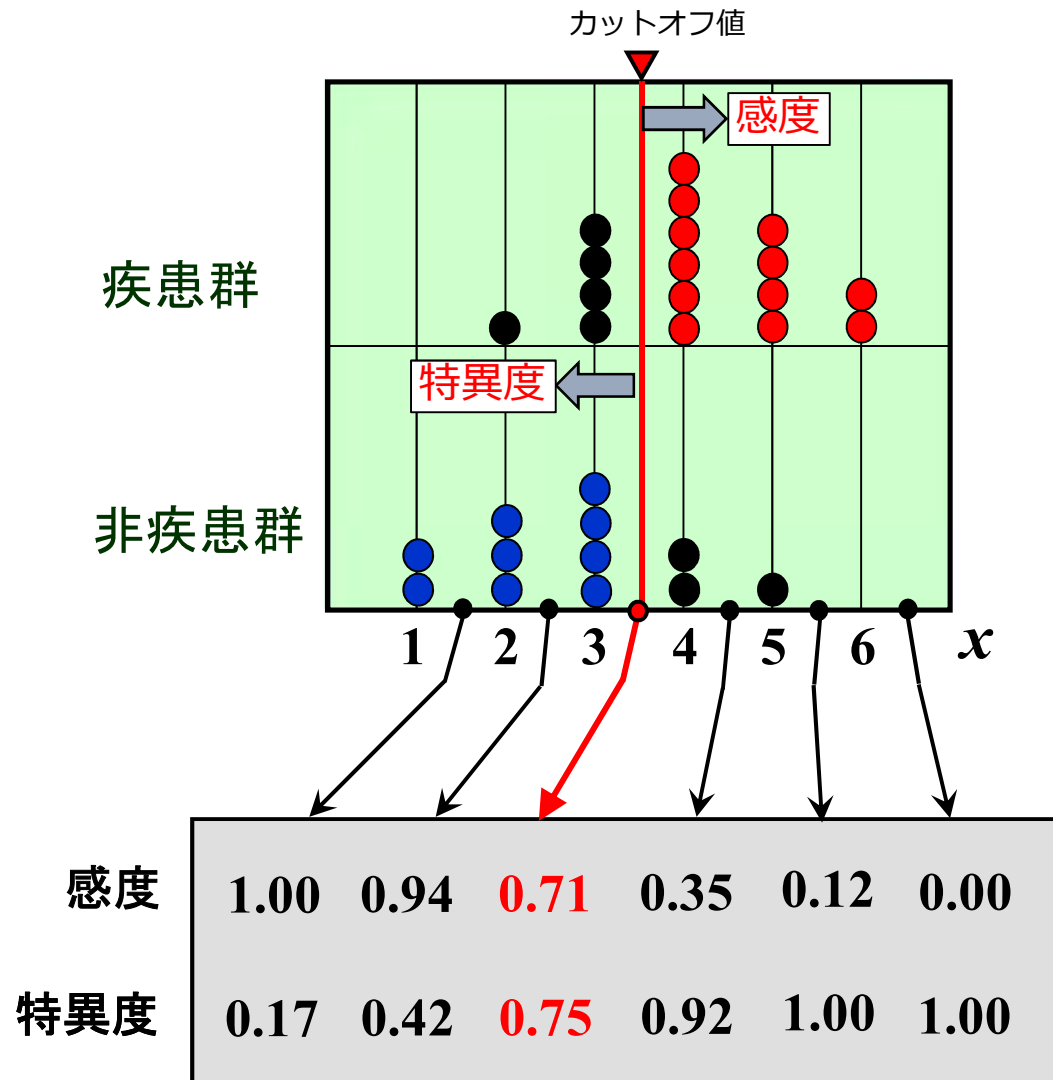


≠ 多重ロジスティック曲線





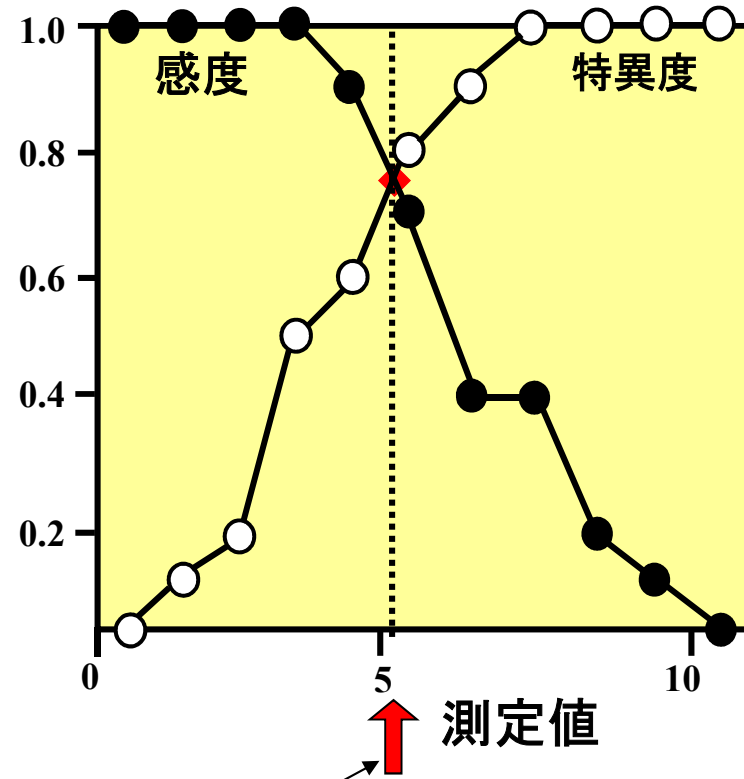
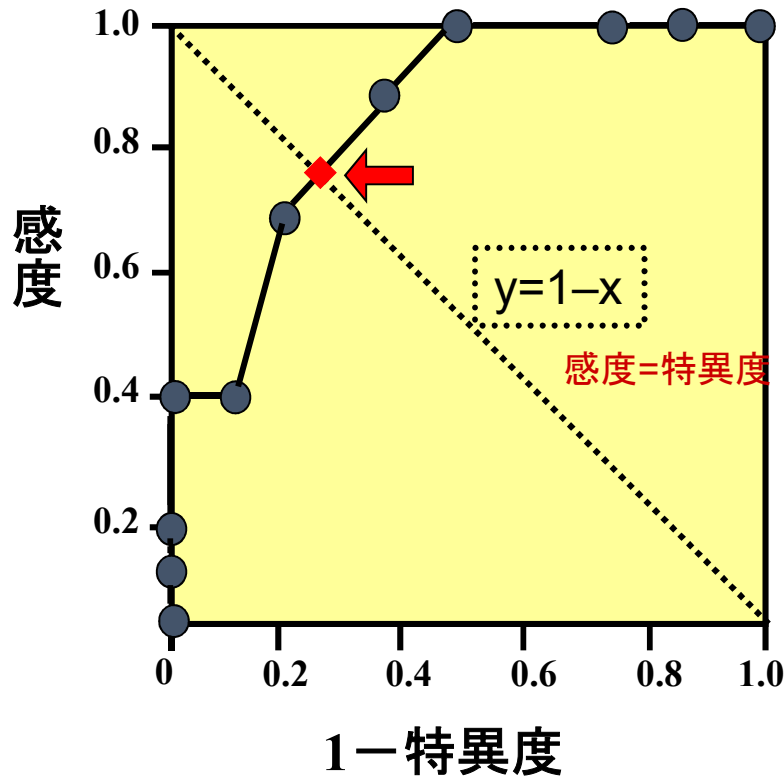
# 感度・特異度曲線によるカットオフ値の設定



通常、感度曲線と特異度曲線の交点をカットオフ値とする

# 感度・特異度曲線とカットオフ値の決め方

ROC曲線と対角線との交点のカットオフ値 = 感度特異度曲線の交点のカットオフ値



感度 = 特異度となる値