STN
SAS Technical News

For Higher
Customer Satisfaction,
We Bridge
the SAS System
Between
Customer's World.

AUTUMN 2012

SAS 9.3 (TS1M2) のリリースと SAS Analytical Products 12.1について

2012年8月末に米国でSAS 9.3（TS1M2）がリリースされ、日本では、2012年9月より提供を開始しました。このメンテナンスリリースから、SAS/STATを始めとするSAS分析プロダクト群が「SAS Analytical Products 12.1」としてカテゴライズされます。
このリリースからは、主にSASの分析ソフトウェアプロダクトは機能拡張が準備でき次第、お客様に提供されます。更新は12 ～ 18ヵ月ごとを予定しています。なお、この変更に基づき、分析プロダクトのリリース番号は12.1からはじまる、新たな番号になります。
SAS Analytical Products12.1には、以下のプロダクトが含まれます。
SAS/STAT®、SAS/ETS®、SAS® Enterprise Miner™、SAS/QC®、SAS/IML®、SAS/IML® Studio、SAS/OR®、SAS® Simulation Studio、SAS® Forecast Server、SAS® Model Manager、SAS® Sentiment Analysis、SAS® Content Categorization、SAS® Text Miner
各プロダクトの変更・拡張についての詳細は、以下のWebサイトに記載してあります。
http://support.sas.com/rnd/app/analytics/12.1/new.html

SAS9.3 TS1M2の分析機能では、より多くのベイズ分析からより多くのモデル選択手法、多変量モニタリング技法からネットワーク最適化と分析など、多岐にわたり拡張されています。
こちらの詳細は、以下のリンク（英文）よりご確認ください。
http://support.sas.com/rnd/app/analytics/12.1/new.html

本号の特集は、このSAS Analytical Products 12.1に含まれるSAS/STATの拡張機能にフォーカスし掲載しています。
その他、今回のメンテナンスで対応した修正個所については、以下のリンク（英文）よりご確認ください。
http://support.sas.com/techsup/reports/maintSAS93/SAS93_TS1M2_issues_addressed.html

特集
分析における拡張点：SAS/STAT® 12.1

SAS 9.2以降では、プロダクトごとに個別の番号が割り振られるように変更されました。SAS 9.2 Maintenance 3ではSAS/STATのプロダクト番号は9.22となっており、PLMプロシジャの追加、線形モデルに対する機能などが拡張されました。また、2011年9月にリリースのSAS 9.3では、プロダクト番号は9.3となり、有限混合モデルに対するFMMプロシジャが新たに追加されました。
今夏にリリースされましたSAS 9.3の最新メインテナンス版 SAS 9.3 TS1M2では、SAS/STATのプロダクト番号は12.1となり、新たに4つのプロシジャが追加されるなど、機能が追加、拡張されています。今号では、新規のプロシジャと併せ、既存のプロシジャにおける主な拡張点をご紹介します。

特集

分析における拡張点：SAS/STAT® 12.1

1 9.22から9.3、そして12.1へ

SAS/STATのプロダクト番号は9.22から9.3へかわり、プロシジャの追加、機能が拡張されてきました。最新12.1の前に、各バージョンにおける主な拡張点をご紹介します。

1.1 9.22 における拡張点

回帰分析、分散分析などを含む線形モデルに関し、多くの機能が拡張されています。その一つとして、EFFECTステートメント（評価版、9.3では正規版）が追加され、スプライン関数、ラグ関数などをモデルに含めることができます（計11のプロシジャが対応）。 また、モデル推定後に対するステートメントとして、ESTIMATEステートメント、LSMEANSステートメント、TESTステートメントがより多くのプロシジャでサポートされています。この他、EFFECTPLOTステートメント、LSMESTIMATEステートメント、SLICEステートメントも新たに追加されています。なお、CONTRASTステートメントを除き、12.1ではLIFEREGプロシジャ、PROBITプロシジャにてもこれらのステートメントがサポートされます。

推定したモデルの情報をアイテムストアとして保存するためのSTOREステートメントを多くのプロシジャ（計10プロシジャ）にて実行でき、新たに追加されたPLMプロシジャにて呼び出し、LS平均値、スコアの算出などを再度モデル推定することなく実行できるよう、拡張されています。

※より詳細に関しては、Technical News Summer 2010をご参照ください。

http://www.sas.com/jp/periodicals/technews/pdf/10sum.pdf

1.2 9.3における拡張点

9.3では、デフォルトの出力形式がLISTINGからHTML形式に変更されています。また、ODS統計グラフ機能もデフォルトにて有効となっており、分析関連のプロシジャを実行することで、出力結果とともにグラフが表示されます。このODS統計グラフ機能に関し、GLMPOWERプロシジャ、POWERプロシジャなど、新たに9つのプロシジャが対応するよう、拡張されています。

追加されたプロシジャとしてはFMMプロシジャがあり、複数の分布が混合した場合のモデル、有限混合モデルに対応しています。このプロシジャは9.3では評価版となりますが、最新の12.1では正規版となり、切断分布（Truncated Distribution）も指定できるように機能が拡張されています。

カテゴリカルな説明変数を含め、線形モデルの変数選択に対応しているGLMSELECTプロシジャでは、新たにSTOREステートメントが追加されており、変数選択した後のモデル情報をアイテムストアとして保存できます。

※より詳細に関しては、Technical News Autumn 2011をご参照ください。
http://www.sas.com/jp/periodicals/technews/pdf/11aut.pdf

2 12.1での新規プロシジャ

最新のSAS/STAT 12.1では、新たに4つのプロシジャが追加されており、分析機能が拡張されています。これらのプロシジャに関し、以下にて記述します。

2.1 比率、リスク算出における拡張

2.1.1 STDRATEプロシジャ

疫学などにおいて、有病のRATE（比率）、RISK（リスク）を算出します。ただし、集団における、単純なRATE、RISK統計量の算出では、交絡因子などが考慮されておらず、値としてあまり意味をなさないことがあります。例えば、各国における有病率を算出した場合、人口の比率、年齢の構造が異なるため、単純には比較することができません。STDRATEプロシジャでは、標準化手法として層を用いた分析に対応しており、DIRECT（直接法）、INDIRECT（間接法）の2手法にて、比率、リスクを算出できます。算出を行う上では、2つの集団、対象集団（Study Population）、参照集団（Reference Population）が必要となります。直接法では、参照集団における層の割合（重み）を用いて、対象集団における比率、リスクの重み付け平均値を全体に対する統計量として算出します。一方、間接法では、参照集団における層ごとの比率、リスクを元に、対象集団にて予期される度数（期待値）を算出した上で、観測値との比較を行います。間接法を用い、RATE統計量の算出を行っている場合、標準化死亡比（SMR：Standardized Mortality Ratio）を求めることができます。

2.1.2 例題

STDRATEプロシジャを用いた分析の一例として、フロリダ州と米国全体における、皮膚がんによる死亡率の比較を用います。サンプルデータには、年齢ごとにおける死亡数、および人年（Pearson-Year）が含まれます。

サンプルデータ

```
DATA Florida_C43;
  INPUT Age $1-5 Event PYear comma11.;
  DATALINES;
00-04     0     953,785
05-14     0   1,997,935
15-24     4   1,885,014
25-34    14   1,957,573
35-44    43   2,356,649
45-54    72   2,088,000
55-64    70   1,548,371
65-74   126   1,447,432
75-84   136   1,087,524
85+      73     335,944
;
```

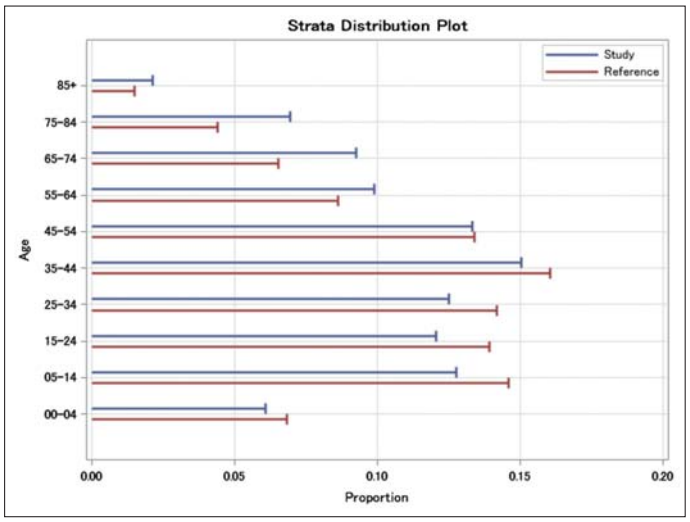
フロリダ州と同じように、米国全体に対しても年齢ごとのデータセット（US_C43）を用い、以下のプログラムにて分析を行います。

例

```
PROC STDRATE DATA=Florida_C43 REFDATA=US_C43
          METHOD=indirect STAT=rate(MULT=100000) PLOTS=all;
  POPULATION EVENT=Event TOTAL=PYear;
  REFERENCE  EVENT=Event TOTAL=PYear;
  STRATA Age / STATS SMR;
RUN;
```

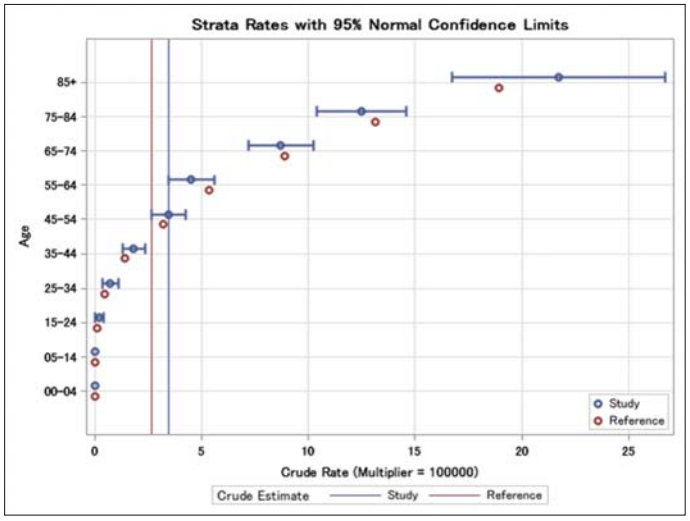
基本的にSTDRATEプロシジャが必要とする2つのデータセットは、DATA=オプション、REFDATA=オプションにて指定します。この例では、対象集団であるフロリダ州のデータセット（DATA=オプション）、参照集団である米国全体のデータセット（REFDATA=オプション）を指定しています。また、METHOD=オプションにて間接法を用いるためにINDIRECTの指定を行い、統計量としてRATE（比率）としています。RATE（比率）の場合には、サブオプションとしてMULT=オプションを指定でき、100,000人年あたりのRATE統計量として値を求めています。POPULATIONステートメント、REFERENCEステートメントでは、それぞれデータセットにおけるイベントを示す変数名、および人年（RATEを算出する上で母数となる変数名）を指定します。STRATAステートメントでは、層を示す変数名AGEを指定した上で、各層ごとのRATEなどの統計量を算出するようにSTATSオプションを追加しています。このプログラムを実行した場合、ODS統計グラフ機能がデフォルトにて有効となっていますので、以下のようなグラフが表示されます。

図1. 層ごとの人年 (PYear) の割合グラフ



このグラフ（図1）では、各集団における、年齢（AGE）ごとの人年（PYear）の分布をしめしており、対象集団であるフロリダ州の場合、米国全体と比較し、高齢グループにおける割合が高くなっています。

図2. 層ごとのRATE (比率) 統計量、信頼区間のグラフ



STRATAステートメントにてSTATSオプションを追加することで、各層における統計量、比率が算出、表示されます。値は表としても出力されますが、上記のグラフ（図2）も併せて作成され、全体的な傾向を把握しやすくなっています。なお、全体における比率（Crude Rate）は、参照線（対象集団：青、参照集団：赤）としてグラフに表示されています。また、統計量としてRATEを指定している場合には、STRATAステートメントにおけるSMRオプションが有効となり、標準化死亡比（SMR）が算出されます。

図3. 標準化死亡比 (SMR) の出力画面

Standardized Morbidity/Mortality Ratio

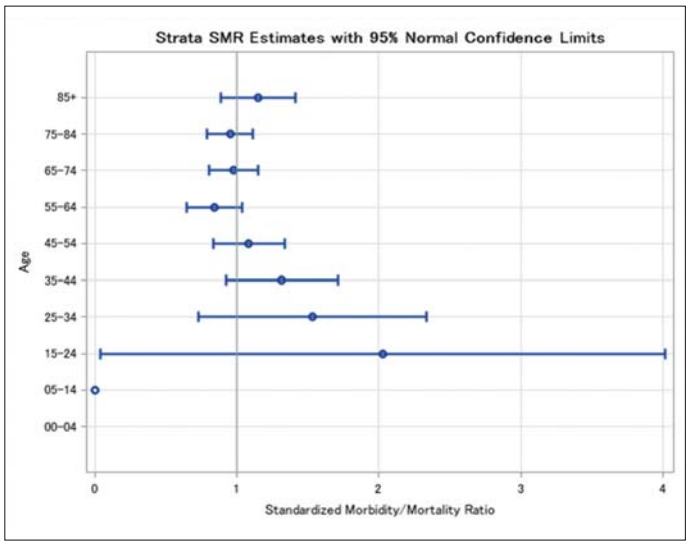
Observed Events	Expected Events	SMR	Standard Error	95% Normal Confidence Limits		Z	Pr > |Z|
538	528.726	1.0175	0.0439	0.9316	1.1035	0.40	0.6893

Indirectly Standardized Rate Estimates
Rate Multiplier = 100000

Study Population			Reference Crude Rate	Expected Events	SMR	Standardized Rate			
Observed Events	Population-Time	Crude Rate				Estimate	Standard Error	95% Normal Confidence Limits	
538	15658227	3.4359	2.6366	528.726	1.0175	2.6829	0.1157	2.4562	2.9096

比率に対してのグラフと同様、標準化死亡比（SMR）に関しても層ごとの統計量を示すグラフ（図4）が作成され、全体としての傾向を捉えられます。

図4. 層ごとの標準化死亡比 (SMR) 統計量、信頼区間のグラフ



2.2 分位点回帰における拡張

REGプロシジャにて対応している回帰分析、GLMプロシジャにて対応している線形モデルなどでは、平均値に対するモデルを仮定し、説明変数に対する係数、パラメータを最小2乗法にて求めます。この平均値に対するモデルを拡張し、分位点に対するモデルを推定する手法として分位点回帰があります。

SAS/STATでは、SAS 9.1.3より評価版、SAS 9.2にて正規版のプロシジャとしてQUANTREGプロシジャがあり、分位点回帰に対応しています。12.1では、分位点回帰における拡張として、2つのプロシジャ（評価版）が追加されています。

2.2.1 QUANTSELECTプロシジャ

線形モデルに対するGLMプロシジャに対して、変数選択を行うGLMSELECTプロシジャがあります。同様に、分位点回帰モデルにおける変数選択に対するプロシジャとして、新たにQUANTSELECTプロシジャ（評価版）が追加されています。

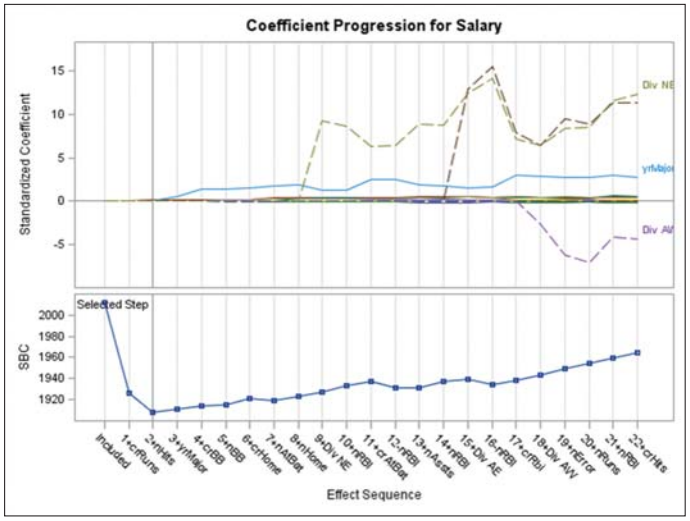
変数選択の手法としては、FORWARD、BACKWARD、STEPWISE、LASSOがサポートされており、MODELステートメントにおけるSELECTION=オプションにて用いる手法を指定します。また、サブオプションとして、変数の追加／削除を行う基準統計量をSTOP=オプションにて指定します。指定できる基準統計量としては、ADJR1、AIC、AICC、SBCがあります。さらに複数のモデル候補より、最適モデルを選択する基準統計量はCHOOSE=オプションにて指定します。プログラムの記述例としては、以下となります。

例

```
PROC QUANTSELECT DATA=baseball PLOTS=all;
  CLASS Div;
  MODEL Salary = nAtBat nHits nHome nRuns nRBI nBB yrMajor crAtBat
                 crHits crHome crRuns crRbi crBB nAssts nError nOuts Div
        / QUANTILES=0.1 SELECTION=lasso(adaptive STOP=aic CHOOSE=sbc SH=7);
RUN;
```

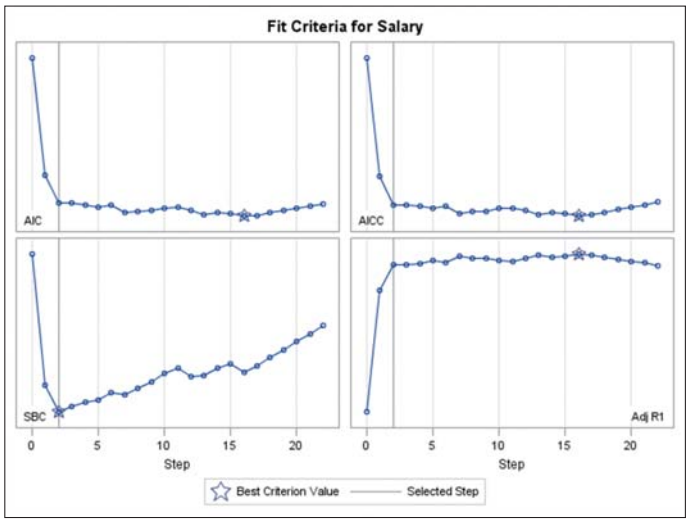
これは、1986年のメジャーリーグ選手の成績（nAtBat、nHits、…）を元に、1987年の年棒（Salary）を推定している例となります。QUANTILES=オプションにて10パーセント点を指定しており、SELECTION=オプションにて変数選択の手法を指定します。ここでは、Adaptive LASSO法を用いて、各ステップにてモデルに追加／削除する変数を選択します。また、AIC統計量（STOP=オプション）を基準にどのステップにて変数選択のプロセスを止めるかを判断し、SBC統計量（CHOOSE=オプション）が最小となるモデルを、最適のモデルとして最後に選択しています。
各ステップにおける変数の追加／削除は表としても出力されます。また、以下のようなグラフ（図5）としても表示されます。

図5. 変数選択プロセスのグラフ



説明変数としては、1986年のヒット数（nHits）、生涯における得点数（crRuns）が選択されていることがわかります。また、各指標における変遷のグラフ（図6）も表示され、変数選択における各ステップにおける変動を把握できます。

図6. 基準統計量のグラフ



2.2.2 QUANTLIFEプロシジャ

打ち切りデータに関しては、生存時間分析に対応しているLIFETESTプロシジャ、PHREGプロシジャ、LIFEREGプロシジャがあります。12.1では新たにQUANTLIFEプロシジャ（評価版）が追加され、打ち切りデータに対し、分位点回帰を行うことができます。
QUANTLIFEプロシジャでは、METHOD=オプションでパラメータ推定手法を選択でき、Kaplan-Meierタイプの手法（KM）（デフォルト）と、Nelson-Aalenタイプの手法（NA）の2つがサポートされています。以下がプログラムの記述例となります。

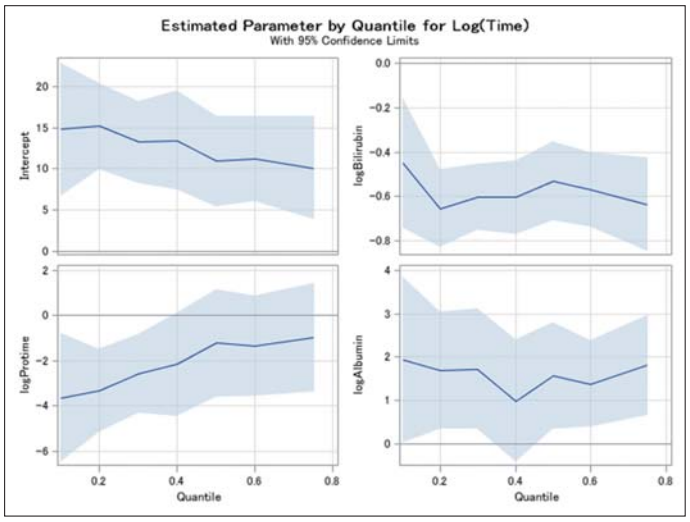
例

```
PROC QUANTLIFE DATA=pbc LOG METHOD=na PLOTS=all SEED=1268;
  MODEL Time*Status(0)=logBilirubin logProtime logAlbumin Age Edema
                      /QUANTILE=(.1 .2 .3 .4 .5 .6 .75);
RUN;
```

これは、FlemmingとHarrington（1991）の肝硬変に関するデータを元にした分析例となります。MODELステートメントでは、応答変数として時間を示す変数TIMEとともに、打ち切りの情報を示す変数STATUSを記述します。また、他の分位点回帰に対する指定と同じように、QUANTILE=オプションにて、任意の分位点を指定します。さらにこのプログラムでは、PROC QUANTLIFEステートメントにていくつかのオプションを指定しています。LOGオプションは、分析を行う前に応答変数の対数変換を行います。推定方法としてNAを用いる場合は、明示的にMETHOD=NAと記述します。QUANTLIFEプロシジャにおける、パラメータ推定値に対する信頼区間は、Resamling法を用いた算出となるため、疑似乱数生成のためのシード値をSEED=オプションで指定しています。

QUANTILE=オプションに指定した各分位点に対し、モデルにおけるパラメータ推定値、信頼区間が表として出力されます。また、PLOTS=ALLオプションを指定していますので、以下のようなパラメータ推定値、および信頼区間の変動を把握しやすいグラフ（一部）（図7）が表示されます。

図7. パラメータ推定値、信頼区間のグラフ（一部）



例えば、左下の説明変数LogProtimeに関しては、分位点（Quantile）が大きくなるほど、パラメータ推定値が0に近づき、応答変数に対する効果が小さくなっています。また、以下のように生存関数のグラフ（図8）、および分位点予測値のグラフ（図9）も表示されます。

図8. 生存関数のグラフ

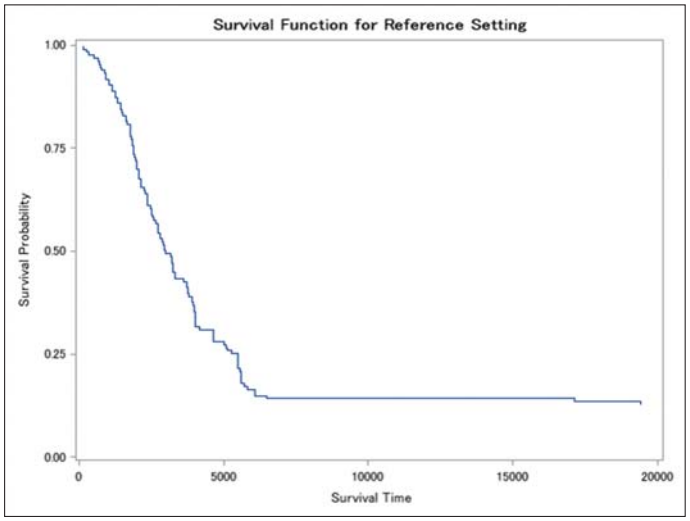
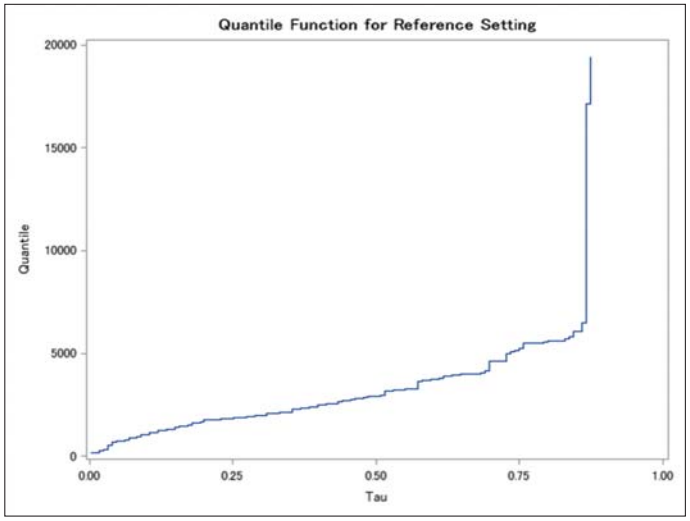


図9. 分位点予測値のグラフ



2.3 ノンパラメトリック回帰における拡張

ノンパラメトリック回帰に関しては、LOESSプロシジャ、TPSLINEプロシジャ、GAMプロシジャが対応しています。LOESSプロシジャは局所的なパラメトリックな関数（3次関数など）にて近似を行い、TPSLINEプロシジャは薄板平滑化スプライン（Thin Plate Spline）に基づくモデリングを行います。これらのアプローチでは多くのパラメータが含まれるため、説明変数の数（次元）が少ないケースに対応しているプロシジャとなります。また、一般化加法モデルに対応しているGAMプロシジャでは、LOESSプロシジャ、TPSPLINEプロシジャより、大きなデータセットにも対応することができますが、加法モデルという制限があります。

2.3.1 ADAPTIVEREGプロシジャ

新たに追加されているADAPTIVEREGプロシジャ（評価版）は、スプライン関数を用いた回帰とモデル選択を組み合わせた手法を用いた、より高次元のデータに対するノンパラメトリックな分析に対応しています。このプロシジャではパラメトリックなモデルを仮定せず、また、スプライン回帰を行う上での結節点（Knot Value）を指定する必要がありません。かわりに、適したスプライン基底関数を生成し、多くの変数に対する結節点（Knot Value）を自動的に選択することになります。その後、モデル選択の手法を活用しながら、モデルの剪定、つまり、モデル削減を行うことによって、より適した、簡略なモデルを選択します。ADAPTIVEREGプロシジャを用いた記述例は以下となります。

例

```
PROC ADAPTIVEREG DATA=autompg PLOTS=all;
  CLASS cylinders year origin;
  MODEL mpg = cylinders displacement horsepower
              weight acceleration year origin / ADDITIVE;
RUN;
```

ここでは、車の燃費（MPG）を研究する上で、車のさまざまな特徴、シリンダー数（CYLINDERS）、排気量（DISPLACEMENT）、馬力（HORSEPOWER）などの変数を用いています。CYLINDERSなどの変数はCLASSステートメントにて指定し、値がカテゴリカルであることを指定しています。この例では、MODELステートメントにADDITIVEオプションを追加し、加法モデルを併せて指定しています。他のプロシジャと同様、以下のような適合度統計量が算出、表示されます（図10）。

図10. 適合度統計量の出力

Fit Statistics	
GCV	11.55804
GCV R-Square	0.81128
Effective Degrees of Freedom	23
R-Square	0.83161
Adjusted R-Square	0.82682
Mean Square Error	10.57977
Average Square Error	10.26079

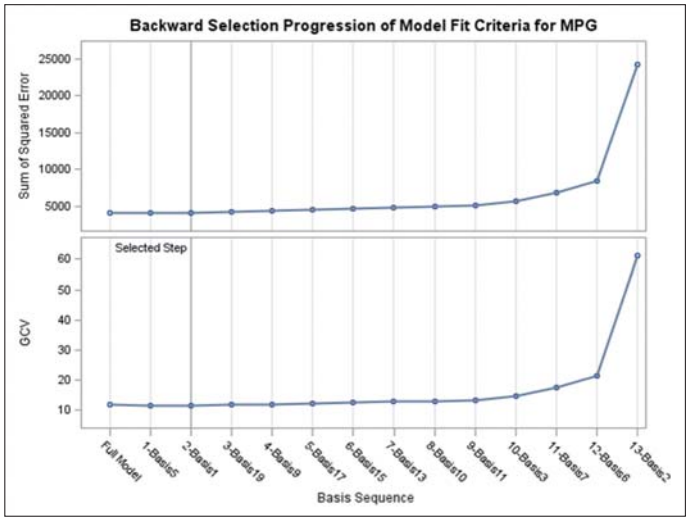
また、モデルの情報を示す、以下の表（図11）が表示されます。

図11. 基底関数（Basis）に関する出力

Regression Spline Model after Backward Selection					
Name	Coefficient	Parent	Variable	Knot	Levels
Basis0	29.4394		Intercept		
Basis2	0.004412	Basis0	Weight	3139.00	
Basis3	-21.2899	Basis0	Horsepower	.	
Basis6	0.1534	Basis3	Horsepower	158.00	
Basis7	2.3920	Basis3	Year		10 12 11 9 8 7 3
Basis9	1.6658	Basis0	Acceleration	21.0000	
Basis10	0.4672	Basis0	Acceleration	21.0000	
Basis11	-8.1766	Basis0	Cylinders		0 3
Basis13	-10.0976	Basis4	Origin		0
Basis15	2.1354	Basis0	Origin		2
Basis17	6.7675	Basis0	Cylinders		3
Basis19	1.4987	Basis0	Year		3 10 12 11 9

ここでは、モデル削減を行った後の基底関数（Basis）に対して、係数の推定値、関連する変数、結節点（Knot）が示されています。基底関数の詳細、式に関しては、PROC ADAPTIVEREGステートメントにDETAILS=BASISオプションを追加することで確認できます。また、併せて以下のグラフが生成されます。

図12. 基底関数（Basis）の選択プロセスグラフ



モデル削減における各ステップを表示したグラフ（図12）となっており、GCV（Generalized Cross Validation Error）統計量が最小となっている2-Basis1までを削除したモデルが選択されています。つまり、この場合、Basis5、Basis1はモデルから削除されています。
また、基底関数ごとではなく、説明変数に対する判断基準として、以下の表（図13）が出力されます。

図13. 説明変数に関する出力

ANOVA Decomposition				
Functional Component	Number of Bases	DF	Change If Omitted	
			Lack of Fit	GCV
Weight	1	2	299.55	0.7165
Horsepower	1	2	1324.81	3.5875
Year	2	4	1183.22	3.0358
Acceleration	2	4	287.76	0.5546
Cylinders	2	4	321.11	0.6470
Origin	2	4	316.04	0.6330

Variable Importance		
Variable	Number of Bases	Importance
Horsepower	1	100.00
Year	2	85.46
Weight	1	21.10
Cylinders	2	19.08
Origin	2	18.67
Acceleration	2	16.38

最初の表では、各変数を削除した場合のモデルに対する影響を示しており、馬力（Horsepower）、車の年代（Year）が大きく影響していることがわかります。また、変数の重要性を示している2番目の表（Variable Importance）においても同様に捉えることができます。

3 その他の主な拡張点

前章では12.1にて追加された4つのプロシジャをサンプルプログラムとともにご紹介しました。この他、さまざまなプロシジャにてステートメント、オプションが追加され、機能が拡張されています。ここでは、主な拡張点についてご紹介します。

● EFFECTPLOTステートメント

箱ひげ図（BOX）、交互作用プロット（INTERACTION）のグラフにおいて、予測値を線でつなぐためのCONNNETオプション、予測値を点で表示するためのCLUSTERオプションが追加されています。

● FREQプロシジャ

リスク差（RISKDIFF）の信頼区間算出の手法として、Agresit-Caffo法（CL=AC）とMiettinen-Nurminen法（CL=MN）が追加されています。また、比率に対するWilson法の信頼区間算出において、連続性に対する修正が追加されています。

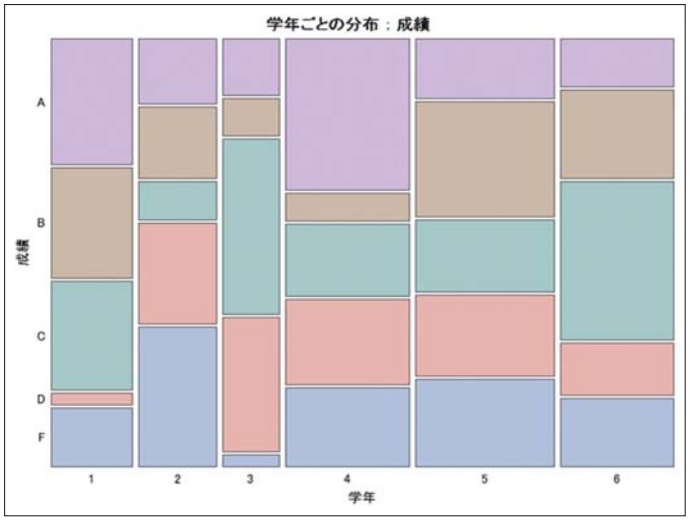
TESTステートメントにPLCORRオプションが追加され、ポリコリック相関係数に対するWald検定、尤度比検定を求められます。

カイ2乗検定のサブオプションとしてDF＝オプションが追加され、検定に用いる自由度を指定できます。また、TESTF=オプション、TESTP=オプションの値としてデータセットを指定でき、一元表の検定における帰無仮説の値をデータセットとして与えることができます。さらに、尤度比検定を求めるためのLRCHISQオプションが追加されています。

2x2表の場合、リスク差の検定手法として、Barnard法が追加されています。

PLOTS=MOSAICPLOTの指定が追加され、2次元表に対するモザイクプロット（図14）を作成できます。（モザイクプロットはSURVEYFREQプロシジャにてもサポートされます）。

図14. モザイクプロットの例



● GLIMMIXプロシジャ

KenwardとRoger（2009）に基づく、標準誤差、自由度に対する調整を行うためのオプションDDFM=KENWARDROGER2が追加されています。

● LIFETESTプロシジャ

重みの変数を指定するWEIGHTステートメントがサポートされます。また、層別変数にラベルが指定されている場合、結果、およびグラフの表示において、変数名ではなく、ラベルを用いた表示となります

● LOESSプロシジャ

予測値、残差などをデータセットに出力するためのOUTPUTステートメントが追加されています。

● LOGISTICプロシジャ

部分的な比例オッズモデル、つまり、一部の説明変数に対しては、応答変数の水準ごとに異なるパラメータを許容するモデルを推定するためのオプションUNEQUALSLOPESがMODELステートメントに追加されています。また、STRATAステートメントを用いた層別解析の場合にも、ESTIMATEステートメント、LSMEANSステートメント、LSMESTIMATEステートメント、SLICEステートメント、STOREステートメントがサポートされます。

MODELステートメントにPCORRオプションが追加され、切片以外のモデルパラメータに対し、部分相関係数を求められます。

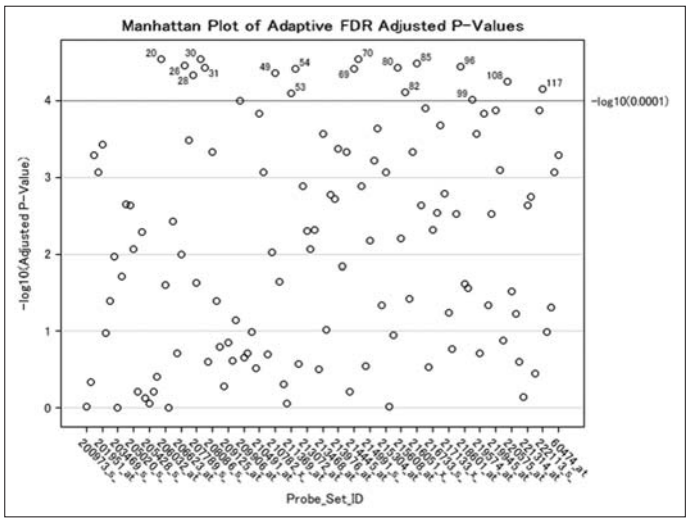
● MCMCプロシジャ

これまでのリリースでは欠損値を含むオブザベーションは分析より削除されていました。12.1ではデフォルトにて応答変数の欠損値を未知のパラメータとして扱い、分析が実行されるよう、拡張されています。また、変量効果に対するRANDOMステートメントでは、複数レベルでの階層的モデルに対応しています。

● MULTTESTプロシジャ

IDステートメントを用いて、結果、グラフ上におけるオブザベーションを判別しやすくなります。また、マンハッタンプロット（図15）の作成にも対応しています。

図15. マンハッタンプロットの例



● NPAR1WAYプロシジャ

多群のデータに対して、多重調整法の一つであるDwass、Steel、Critchlow-Fligner法をおこなうDSCFオプションが追加されています。また、2群の場合は、Flinger-Policello検定に対するFPオプションが追加されています。
中央値の差に対する検定Hodges-Lehmann検定において、参照水準を指定するREFCLASS=サブオプションが追加されています。

● PHREGプロシジャ

RANDOMステートメントにDIST=オプションが追加され、ガンマ分布、対数ガンマ分布を指定できます。また、BAYESステートメントにおけるDISPERSIONPRIOR=オプションで分散(dispersion)に対しても事前分布を指定できます。
BASELINEステートメント、OUTPUTステートメントにて、METHOD=FMオプションの指定を行うことで、Fleming-Harrington法による推定値を求められます。

● POWERプロシジャ

LOGISTICステートメントにおいて、CORR=オプションが追加され、対象としての説明変数と他の共変量との相関係数を考慮した上で、検出力、サンプルサイズが算出できます。

● SURVEYSELECTプロシジャ

サンプリングの手法として、BernoulliとPoisson手法が追加されています。Bernoulli法の場合、SAMPRATE=オプションにて指定した確率に基づき、オブザベーションごとにサンプルとして抽出するかを判断します。この手法を拡張した手法がPoisson手法であり、SIZEステートメントにて指定した変数の値(確率)に基づき、オブザベーションごとに抽出を行います。この2つの手法ではサンプルのサイズは一定とはなりません。また、Mersenne-Twister法を用いた乱数生成のアルゴリズムに変更されています。このため、12.1より前のリリースと同じ結果を必要とする場合には、RANUNIオプションの指定が必要となります。

4 参考文献

12.1に新たに追加された4つのプロシジャに関しては、米国にて開催された SAS Global Forum 2012 にて紹介されており、以下の文献にて詳細を参照できます。

Look Out: After SAS/STAT® 9.3 Comes SAS/STAT 12.1!
http://support.sas.com/resources/papers/proceedings12/313-2012.pdf

また、SAS/STAT 12.1を含む SAS 分析のプロダクトに関しては、以下のページよりドキュメントを参照できます。
http://support.sas.com/rnd/app/analytics/12.1/new.html

5 おわりに

SAS/STAT の最新バージョン 12.1 における拡張点として、新たに追加された 4 つのプロシジャをサンプルプログラムを用いてご紹介しました。また、主な拡張点、変更点に関しても併せてご紹介しました。さまざまな分析手法を考慮し、どの手法を用いるかを検討する上で、今号における内容が参考となれば幸いです。

SAS Academic News
SAS アカデミック・ニュース

朝野先生による「マーケティングとデータ解析」第2回では、仕事の出会いからマーケッターとして、仕事を通じてどのように統計学を学び調査や分析を行い答えを導きだすか書かれています。
新村先生のコラム「SAS／JMPとの歩み」では、SASのミニコン版の代理販売店となりシステムインテグレータとしての事業を展開し、より精度の高い分析によりさまざまな問題の解決に挑んでいきます。

随想
「マーケティングとデータ解析」

コラム
「SAS／JMPとの歩み」

随想 「マーケティングとデータ解析」

朝野 熙彦
多摩大学大学院客員教授

第２回 統計学の独習記

前号では「行き当たりばったり半生記」と題して、私と仕事との幸せな出会いを紹介させていただきました。私は良いデータ・アナリストが育つ条件は「よい本」「よい仲間」「よい仕事」の３つだと考えています。今回は私が若者時代に仕事をしながらどのように統計学の勉強をしたかを披露させていただきます。

1. 実社会への貢献

私が研究しているマーケティングは基礎学問ではなく実学ですので、マーケティングのためのデータ解析も実践活動そのものといえます。

データ解析を学ぶ上でよい本を読むことはもちろん大切ですが、マーケティングはただ本を読んで思索にふけるだけで済む学問ではないことはご承知ください。

書斎で読書しているだけではデータ解析の上手なユーザーにはなれないでしょう。データ解析の本当の教科書は本の中にも教室の黒板の上にもなく、データ解析が応用される実社会にあるのではないでしょうか。もしデータ解析をした結果が社会の現実と乖離がある場合は、データ解析の側が反省するという謙虚な姿勢が大事だと思います。

データ解析に取り掛る前に、データ自体がどのような環境と文脈（コンテクスト）の下で得られたのかを理解しておくことも大切です。そもそもデータ解析で用いる統計モデルが、分析したい現象のモデルとしては不適切であるかもしれません。モデルが不適切ならデータ解析がうまくいくはずがありません。私も自分自身の無知のためにそうした失敗を何度も経験してきました。

理論は理論として整合性がなければなりませんが、その理論が実社会で役立つかどうかは実社会が答えを出すことです。データ解析の価値は現実の世界にどれだけ貢献できたかによって評価すべきでしょう。故鳥井道夫氏（元サントリー名誉会長）が日本マーケティング協会会長の時代に述べられた「マーケティングとは売ってなんぼの実践学だ」という認識は、そのままデータ解析にも当てはまると思います。（鳥井道夫(1997)「大才中才小才」プレジデント社、16頁）

データ解析のおかげでビジネスを成功に導けたのかが問われるべきです。世の中には本当の意味で実践に貢献してきたデータ解析があります。医学や薬学がまさにそうです。また推測統計学の発祥となった農学そして推測統計学を社会に普及させることになった生産管理も、真摯にデータ解析の実践活動を積み重ねて産業界の発展に貢献してきたのです。

それらと比べるとマーケティングの実務の世界では、マーケティング提案を裏付けるための都合のよい道具、あるいは取引先企業を感心させるためのギミックとしてデータ解析を使うことはなかったでしょうか。先輩諸科学の研究姿勢を見習いたいと思います。

2. 素晴らしい本との出会い

前号で自己紹介しましたように私は調査のイロハも知らないままマーケティング・リサーチの会社に入りましたので、入社後にゼロから勉強をしなければなりませんでした。マーケティング・リサーチの仕事で難しい問題が出てくると独りで、あるいは仲間とともにマーケティングや統計学、その他もろもろの勉強をしました。目標志向だといえば恰好はよいのですが、ありていに言えば「泥縄式」でした。会社員時代に数百件の新製品や新規事業の開発に携わりましたが、プロジェクトが始まる前から必要な知識がそろっていた、というような案件はめったになく、プロジェクトが始まってから付け焼刃で勉強を始めた場合がほとんどでした。

大学はその理想的な存在意義としては、

A：大学生の間に社会で役立つ勉強をしっかりと身に着ける
B：社会に出てから学生時代に身に着けた勉強を仕事に生かす
C：そうして大学が社会に貢献できることを実証してみせる

という幸せなストーリーを描いております。本当にA⇒B⇒Cの理想通りに実践できている学生もいないわけではありません。たまたま私個人のケースでは、

A'：大学生の間に社会で役立つ勉強を身に着けなかった
B'：社会に出てから仕事に必要な勉強を始めた
C'：そのため大学が社会に貢献できることは実証できなかった

というパスをたどっただけのことです。

私と一緒に勉強してくれる仲間は職場には少なかったですね。でもマーケティング・リサーチの仕事を通じてだんだんと社外の仲間が増えてきました。そもそも仕事を発注してくれたお客様が真っ先に仲間に加わってくれました。「よい仕事」と「よい仲間」が

混然一体となってグループ学習をしたものです。その後、大学の教員になってゼミを担当するようになってからの話ですが、ゼミ生でグループを作って論文の講究をさせてみました。集団で講究した方が、学生の勉強意欲も持続できるし教育効果が上がるように思われました。もちろん独習が良いのか、それともグループ学習が良いのかは個人の性格にもよるでしょうから、一概にどちらが良いと押し付けるつもりはありません。

さて、若いころ私が読んだ懐かしい本に、竹内啓・柳井晴夫「多変量解析の基礎」東洋経済新報社（1972年）、という本がありました。

この本は線形空間への射影という大変すっきりした概念で多変量解析を解き明かした本です。各種の多変量解析は、それぞれの方法ごとに目的関数が設定できて、その最大化をはかると多くの場合、固有値・固有ベクトルを計算する問題に帰着することが知られていました。しかし、あれはあれ、これはこれの計算問題という感じで統一的な見通しに欠けていたのです。それに対して線形空間に別の線形空間を射影するという唯一のアイデアだけですべての多変量解析が一気に説明できるという透徹した原理は爽快でした。もちろん非線形の多変量解析については同じアイデアでは対応できませんが、1970年代のマーケティング界では非線形の分析はめったに使われていませんでした。この本がきっかけになって、当時若者だった私は統計解析に興味を持つようになったのです。全くの初心者ではありましたが、データ解析の面白さに目覚めてしまった瞬間です。

ですからこの本がその後の自分の人生を導いてくれた本であることは間違いありません。心から感謝できる本にめぐり合えたことは私の幸せでした。

ところでこの本の初版には数式展開や記号に関する誤植がたくさんありました。プライム（'）が抜けているだの＋と－が逆だのといった些細な印刷ミスが100個所くらいはあったでしょう。ミスを訂正しながら精読することはとても楽しいものです。論理にあいまいさが無く明瞭に書かれた本だからこそ間違いにも気づくのです。著者は読者がミスを訂正しながら勉強できるように、親切心で校正モレを残しておいてくれたのかもしれません。（違うか）

なお、誤植について一般論を言わせていただくと、最先端の研究をされている学者は、常に新しい研究に没頭しているために、脱稿後の本の校正をしている余裕がないという事情があります。もう一つ「多変量解析の基礎」の場合は、この本の大部分を執筆された柳井先生の字が達筆すぎて、印刷所の人に判読しがたかったという事情があります。私の知る限りですが、優れた学者はとかく手書きでは読みづらい字を書かれる傾向にありました。そう、当時は原稿用紙のマス目に鉛筆や万年筆で原稿を書いていたものでした。

後日談ですが、柳井晴夫・竹内啓「射影行列・一般逆行列・特異値分解」東京大学出版会（1983年）、というこれも素晴らしい本がその後出版されました。相変わらず100個所どころではなく校正モレがありまして、そのことを柳井先生に伝えました。すると印刷ミスを書きこんだ本を貸してくれ、こんど増刷する時に参考にするから、という読者にとっては誠に光栄なご依頼を受けました。さすがに一流の研究者だけあって、汚名にかかわるなどと立腹しないものだと感心した次第です。

3. 疑いつつ読めば勉強になる本

ただの印刷ミスの問題ではなくて論旨そのものが間違っている本も存在します。特に調査のための実務マニュアル本の中には統計学的な観点からみて間違いといってよい記述が少なくありません。ひとつ例題を出してみましょう。

【例題】平均値の95%信頼区間

母集団での平均値をμ、分散をσ^2とする。母集団の規模をN、サンプルの大きさをnとする。サンプルの平均値$\bar{x}$の分布はN−nとnが大きければ正規分布に近似する。母集団から無作為抽出して標本平均$\bar{x}$を求めると、平均値の95%は

$$\mu-2\sqrt{\frac{N-n}{N-1}\frac{\sigma^2}{n}}$$ より大きく

$$\mu+2\sqrt{\frac{N-n}{N-1}\frac{\sigma^2}{n}}$$ より小さい。

つまり信頼度95%で

$$\left|\bar{x}-\mu\right|<2\sqrt{\frac{N-n}{N-1}\frac{\sigma^2}{n}}$$ ……①である。

このことから

$$\bar{x}-2\sqrt{\frac{N-n}{N-1}\frac{\sigma^2}{n}}<\mu<\bar{x}+2\sqrt{\frac{N-n}{N-1}\frac{\sigma^2}{n}}$$ ……②

が導ける。通常Nは極めて大きいので、近似的に真の平均値μは信頼度95%で

$$\bar{x}\pm2\sqrt{\frac{\sigma^2}{n}}$$ ……③の範囲に入る。

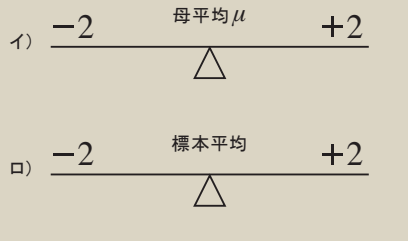
これを平均値のサンプリング誤差という。母集団の分散σ^2は未知なので、標本分散s^2に置き換えて

$$\bar{x}-2\sqrt{\frac{s^2}{n}}<\mu<\bar{x}+2\sqrt{\frac{s^2}{n}}$$ ……④

を平均値の95%信頼区間と呼ぶ。

信頼度ではなく信頼係数という言い方もあります。正規分布なら上の①～④の2は1.96が正しい、などという細かいことを問題にしているわけではありません。

例題に書かれた平均値の信頼区間に関する考え方は、図に描くとイ）からロ）が導けるという論理です。



【大いなる誤解】

イ）の方は、母集団のパラメータが未知なのですから、実務的には何も教えてくれない絵空事です。問題はロ）です。ロ）の意味はある調査から求めた標本平均があって、その上下の一定幅の区間に母平均が分布するということなのでしょうか。数値例でいうとNはとても大きいとして

n=400、$\bar{x}$=1000、s=600

で④を適用すると、

$$2\sqrt{\frac{s^2}{n}}=2\sqrt{\frac{600^2}{400}}=60$$

したがって母平均が次の区間に入る確率が0.95だといえるのかどうかという問題です。

$$P\left[\ 940 < \mu < 1060\ \right] = 0.95 \quad \cdots\cdots ⑤$$

調査で問題にしている確率変数をXとして、n個のXが独立同一分布に従うと仮定しますと、n個のXの標本平均の統計量は

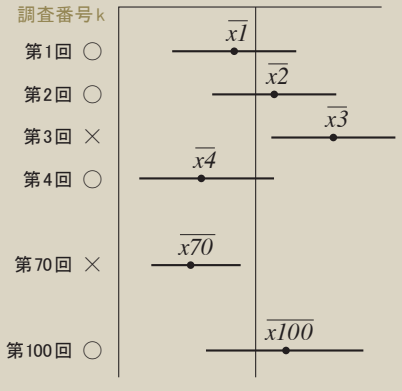
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

であって、それ自体が確率変数になります。ですから1回の調査から得られた平均値は母平均μなのではなくて確率変数$\bar{X}$の1つの実現値

$$\bar{x}_k = \frac{1}{n}\left(x_1 + x_2 + \cdots + x_n\right)$$

に過ぎないのです。$\bar{x}$につけた下付き添字のkは第k回目の調査結果であることを表します。ですからもし第2回目の調査を行えば、n個の違ったデータが観測されるわけで、その平均値も$\bar{x}_2$に変わりますし、当然ながら標本分散s_2^2もs_1^2とは違った値になります。ですからロ）の区間は中心の位置だけでなく区間の幅までも調査のたびに変動するのです。次の図をご覧ください。調査回数がk=1,2，…100の場合の模式図です。

調査を繰り返せば信頼区間が変動する（×の回数は母平均が区間外）



さてμは定数ですし、特定の調査から得られた⑤の区間は固定されますから、μは区間[940,1060]に入るか否か（1か0）のどちらかであって、95%の確率でこの区間に入るというような確率的な意味合いは持ちません。つまり⑤式の確率的解釈は明確な誤りです。確率変数と実現値の混同が間違ったコメントを生んだのでしょう。全く同一の調査をk=1,2,…; Kと繰り返すことは実際にはほとんどなく、大抵調査は第1回で終わってしまいます。そのため、調査をやりなおすたびに信頼区間が変動する、という事実に出会う機会がなかったのでしょう。でも想像してみれば信頼区間が変動するのは当然だと思いませんか？

1970年～1980年代にリサーチャーが読んでいた実務書には例題と同じ誤りがよくありました。比較的よく知られた調査の実務書のどれをみてもみな同じパターンで書いてあるので、逆にさかのぼってみると1950年代に出版された統計調査の本にまでさかのぼれます。どの本が間違いの元祖かという責任追及が本コラムの目的ではありませんが、誤った理解が普及してしまったのは困ったものです。

きっと推測統計学に詳しくない読者のために、物事を分かりやすく説明しようという親切心からの記述だったのだろうと思います。「テキストを疑いながら読む」ことは学びの第1歩ですから好ましいことですが、誤解したままの人がいてもいいのか？という疑問は残ります。

4. お勧めの統計学の本

リサーチの実務マニュアルには、質問文の作り方だとかインタビューの仕方など、まさにリサーチの専門的な業務の進め方が書かれています。ですから、そうした本は必要だし価値はあるのです。しかしながら手軽なビジネス書や実務マニュアルに統計理論や数式まで任せるのが不安であれば、いっそのこと統計学の専門家が書いた専門書を読んだらどうでしょうか。

前節で指摘した問題に関しては、たとえば竹村彰通（2007）「統計第2版」協立出版の109頁では「μはパラメータの真値であり、これは固定されている。確率的に変動するのはIという区間である。つまり、区間を何度も作ると、その形の区間がμを含む割合が95%になるという意味合いである」、ときちんと書いてあります。

このネイマン流の確率言明について、蓑谷千凰彦（2009）「これからはじめる統計学」東京図書（253頁）においても同様に、100回調査をすれば、それぞれの調査ごとに1本ずつ違った信頼区間が得られ、計100本の信頼区間のうち平均して95本がμの真の値を含むであろう、という意味の説明をしています。第1回目の調査でいきなり最終的な信頼区間が確定するわけではないし、$\bar{x}_1$がμに一致するという根拠もないのです。

上記した2冊の本は、どちらもすっきりと分かりやすく統計学を解説されているので、初心者の方にもお勧めしたいと思います。

略歴：

千葉大文理学部卒業後市場調査会社に就職、埼玉大大学院修了、千葉大・筑波大講師、専修大・都立大・首都大教授を経て多摩大学大学院客員教授。学習院マネジメントスクール講師、日本マーケティング・サイエンス学会論文誌編集委員、日本行動計量学会理事。

主な著書に『アンケート調査入門』東京図書（編著）、『最新マーケティング・サイエンスの基礎』講談社、『Rによるマーケティング・シミュレーション』同友館、『入門共分散構造分析の実際』講談社、『魅力工学の実践』海文堂出版、『入門多変量解析の実際第2版』講談社、『新製品開発』朝倉書店（朝野熙彦・山中正彦著）などがある。

コラム「SAS／JMPとの歩み」

新村 秀一
成蹊大学 経済学部教授 理学博士

第3回 SASのミニコン版代理店とシステムインテグレータとしての事業展開

前号のつづきです。

9. SASのミニコン版代理店（1985年～1989年）

9-1 ソフトウエアの価格

1984年頃に当時のSASジャパン社長の辻本氏の訪問を受けました。要件は、SASのミニコン版が汎用機版に比べ営業マンが販売に苦戦しており、SCS社に代理店になってほしいという依頼でした。私はかねてより予感していたので引き受けることにしました。ソフトウエアの価格は、稼働機種のハードウエア価格の制限を受けます。汎用機で成功したソフトウエアは、ミニコン版では価格をミニコンのハードウエア価格に合わせて汎用機と同じ機能でも価格を低く抑えなければなりません。それがPC版になると、汎用機版の価格体系を破壊的に創造する必要があります。

SASの創業者の一人のJohn Sall副社長がJMPを開発したことは理にかなっています。もしJMPが開発されなければ、私見ながら汎用機やミニコン市場を捨ててPCに特化したSPSS社の戦略に負けてPC市場を失っていたことでしょう。SASは企業向けのシステム開発として汎用機からWSまで、JMPは研究者や個人を主体としたPC市場と住み分けた方が全体としての販売戦略として良いと思います。その場合、大学教育にも工夫が要ります。理工学部や医学部を擁する大学へは、将来統計機能を含むシステム開発を見据え、PC版SASを提案し、文科系学部の優位な大学ではJMPを提案して住み分けるべきでしょう。その上で、SASとJMPの機能の互換性をさらにはかり、企業のSASユーザーが個人的にJMPのユーザーになり、大学でJMPの教育を受けた学生が社会人になりSASへ比較的移行しやすくすべきでしょう。

数値計算とグラフ機能を含む数学ソフトのSpeakeasy[24]は、早い時点で配列、行列、時系列、集合の4つのオブジェクトに対し各種演算機能をもっていて、数値計算ソフトのMatLabや数式処理ソフトのMathematicaの源流です。川崎製鉄所の水島製鉄所で開かれたOR学会の中国支部の研究会で、数理計画法ソフトLINDO[25]に関する発表の講師に呼ばれました。その後の見学会で、富士通汎用機で稼働する製鉄システムのアプリケーションの一つとして、Speakeasyが数値計算のシステムに使われている説明を受けた際、SAS/IMLを使用すれば、そのシステムに統計処理も組み込めるとコメントしたことがありました。私は、当時一部の大学の研究者からSAS坊と呼ばれていた所以でもあります。

Speakeasyは汎用機で成功していたためPCへの対応に乗り遅れ、MatLabやMathematicaに完全に負けていたソフトウエアでした。それにもかかわらず、1989年にSCSがSpeakeasyの代理店になったのはアルゴンヌ研究所の原子力研究者で、有馬元文部大臣の研究仲間であり、能などの日本文化にも造詣の深い創業者のStan博士の人柄にほれたことと、Mathematicaなどに比べ使いやすく価格が安いためでした。結局、PC市場戦略の遅れにより、Speakeasyの販売は不成功といわざるをえません。

9-2 販売体制

SASのミニコン版の代理店の件を上司の役員に相談すると、「君の趣味に使っていい社員は2人、いや4人まで、ただし東京ガスなどの重要顧客の課員を課長に無理を言って使わないこと」と釘を刺されました。内心私は「4人も使っていいのか」と安堵しました。また、ソフト販売に必要な資質はSEやプログラマーのそれとは違うと考えていました。そこで、ブラブラ部長の私と市川君の2人で立ち上げることとし、SASなどの受託営業をやっていた塗課長と事務職で事務処理を行う体制にしました。ただし、SASの膨大なテクニカル資料の翻訳を私がやっているわけにはいかなくなり、私の席の前に英語のできる女性2名を配置し、翻訳させることを2年ほど行いました。

SASのミニコン版は、DEC、DGとPrimeという米国のミニコン御三家の版がありました。これらの日本法人と打ち合わせをもち、デモ用のミニコンを各社のセンターで無償利用させてもらうことと、各社の営業部隊と共同で販促の営業を行いました。私が発表用のスライドの原稿を作成し、市川さんが完成スライドにリメイクしました。そして、彼が3社のミニコンを器用に操作し、私が説明を行うという全国行脚を行いました。また品質管理の学会誌[26]や数誌の商業誌への寄稿を行いました[27-33]。

しばらくして市川さんがDECのパフォーマンスが群を抜いているので、DECに絞りませんかと提案してきました。しかし、SAS社からは3社の販売を期待されているので、半年ほど悩んだ末、DECに営業を注力することにしました。そして、DECのミニコンとSASの販売を行うSIサービスを行うことにしました。DECと交渉すると、すでに住商エレクトロニクス（株）が代理店なのでここから仕入れてほしいということで、ミニコンのサポート要員を増やす必要がないので受け入れました。最初は業種を区別なく行っていましたが、製薬企業の臨床試験部門からの問い合わせが増えてきました。製薬企業向けの営業に注力しました。比較的早い時期のユーザーに科研製薬（株）があります。納品後、渡辺さんが面白いものを見せますということで、後でSASレーザーフォームという付加価値製品の切り札になったデモを見せてもらいました。SASの出力結果をお化粧直しし、厚生省へ申請するために、日本語と罫線をレーザープリンターに重ね書きする帳票システムです。科研製薬と契約し販売権を取りました。そして、価格を50万円の売りきりとし、SASとVAXの購入企業のみに販売することにしました。これらの仕組みづくりとSAS社の製薬担当との共同作業がうまくいき、記憶違いがなければ32社の製薬企業にSAS/VAXを販売し、今日日本の統計ユーザーの中で一番強力な統計解析の専門家集団の誕生に貢献できたと考えています。

9-3 筆の力

SASを統計の個人家庭教師のごとく使用した成果として、単に統計書で勉強していた苦難の20代前半に比べ、学習成果を書籍や論文として発表できるようになりました。

[34]はSASのテクニカルレポートの翻訳を

行っていて見つけたJ.Sall博士の回帰分析に関するレポートを本文とし、Goodnight社長の「掃き出し演算子と変数選択法」のテクニカルレポートを付録に付けた翻訳書です。回帰分析から非線形回帰分析、さらに線形計画法で実現できるL1ノルム回帰を重み付き回帰で解説するなど、既存の回帰分析の良書でこれまでに触れられていない広範な内容です。さらに掃き出し演算子の考え方は、それまで100冊以上の内外の統計書に目を通していた私に全く知らない新しい世界を教えてくれました。これはぜひ日本に紹介すべきと意気込み強い意志で朝倉書店と交渉し出版にこぎつけました。[35-36]は、これらに刺激を受けて回帰分析を行列表現で紹介したものです。[37-38]は、掃き出し演算子と変数選択法の解説です。[39]は「オペレーションズ・リサーチ」の編集委員をしていた時、編集長の柳井浩慶応大学元教授から「だれか、筑波で開催された科学万博のデータを入手したので、寄稿しませんか?」といわれ、私が引き受けました。恥ずかしい話、もらった原票の注釈を読まず、来場者数をシャトルバス、団体バス、自家用車で回帰すると、自家用車の回帰係数が8になり、不思議だと指摘しました。後になって、ミニバンを含むことが分かりました。[9]はこのデータを用いて、SASの主要な統計手法を春号の3章で紹介したデータを調べる手法と予測手法にまとめて紹介しています。

私はグラフに関しては詳細に勉強するのが時間の無駄と考えてデフォルト主義を貫いています。ですから[40]は、SASのグラフの基本的な機能でも有用であることを説明しています。

10. 太閤殿下の愚行を繰り返す

いま振り返ればSASミニコン版の販売の成功で、少し調子に乗りすぎたようです。曲がりなりにも営業成果が出てきていたので、新規事業部の部長になり、新規事業を行う1年生の新人の部下も増え、彼らに仕事を確保する必要がありました。豊臣秀吉が、日本制定後、朝鮮出兵という愚行の歴史的事実を理解していたのに、それを繰り返してしまいました。

100人くらいのシステム開発の課員(派遣を含む)がいたシステム担当の海野課長から、部長の新規事業を行うにはあと100人の部員がいるといわれていました。彼の判断は正しかったようです。とりあえずは技術者の増員はかなわないので、販売主体で先行し、仮に大きく育てば技術者を後で割り振る予定でした。

しかし私の一生を顧みれば、自分の研究や教育に役立ち現在も手元に残ったソフトウエアは、統計のSAS、数理計画法のLINGO、数学のSpeakeasyに関するものだけです。

11. 三宅先生間違いで誤分類数最小化(MNM)基準による最適線形判別関数の研究を行う

東大医学部の開原先生が主催する統計研究会で、他の研究者から三宅さんと呼ばれている先生と知り合いました。研究会終了後、「SPSS普及の旗振り役の三宅先生ですか?」ということで名刺交換すると日本医科大学の数学科の三宅章彦教授でした。その後、彼の研究を手伝うことになりました。

最初の研究は、判別分析の標本誤分類確率と母誤分類確率の関係に関する研究です。データ数が少ないほど、説明変数が多いほど、標本誤分類確率は母誤分類確率に比べ過小評価されることを、母誤分類確率を0.5から0の間で変えて標本誤分類確率の5%点から95%点をグラフで分かり易く説明した論文を三宅先生がDijonの学会で発表されました[41]。

ここで母誤分類確率を0.5から0の間でとっていることは、判別分析を正しく知る上で重要です。線形判別分析は、2群が正規分布し分散共分散が等しいというFisherの仮説から出発します。この場合、2群の平均がm_1とm_2で分散共分散がσであるとすれば、2群は正規分布$f1(m_1,\sigma)$と$f_2(m_2,\sigma)$で表わされます。そして, 判別境界は$f_1(m_1, \sigma)= f_2(m_2,\sigma)$であり、群1を群2に間違う誤分類確率$e_{12}$と、群2を群1に間違う誤分類確率$e_{21}$とすれば、$e_{12}= e_{21}$になり、誤分類確率は$e= e_{12}+ e_{21}$になります。そしてFisherの線形判別関数の誤分類確率eは、判別境界を動かして得られる誤分類確率の中で最小になります(e=MNM)。しかし、現実のデータはFisherの仮説をほどんと満たさないので、この前提が崩れてしまいます。SASに限らず統計ソフトはこの前提から出発しているので、分析に用いた2群のケース数が等しくなくても等確率と考えて計算することをデフォルトとしています。ケース数に比例して考える場合は、事前確率をp_1とp_2として、事前確率のオプションを指定することで使い分けています。この場合、判別境界は$p_1*f_1(m_1, \sigma)= p_2 *f_2(m_2,\sigma)$になります。さらに医学診断からの影響と考えていますが、正常群を群1とし、異常群を群2とすれば圧倒的に正常群が多いことになります($p_1>p_2$)。しかし異常群を間違って誤分類するリスクが高くなるので、その程度をリスク($r_1<r_2$)として、判別境界を$r_1*p_1*f_1(m_1, \sigma)= r_2* p_2 *f_2(m_2,\sigma)$で考えます。最初の段階だけが確率分布の議論で、事前確率やリスクを導入したものは、恣意的に正規確率分布を何倍かして変形していることを忘れている人が多いようです。私の誤分類数最小化基準は事前確率を考えた分析で、試験の合否判定を実証研究しました。合否判定は自明な誤分類数が0の判別問題です。得点分布の10%未満を不合格、10%点以上を合格とした合否判定で、LDFで誤分類確率が0.3、2次判別関数が0.9という驚くべき結果が出ました。その時、旧知の統計の教官が誤分類確率は0.5を超えないのではといったのに驚きました。この結果から、少なくとも3つの異なった誤分類確率が得られることと、データはFisherの仮説を満たさないので判別境界を動かすと統計ソフトの出力結果より良い判別結果が得られることを理解すべきです。また、判別結果をROCで描いて評価すれば、判別境界の変化に対応した判別結果の評価と、異なった判別手法の評価にも利用できます。

最近、若手研究者で判別境界をどう選択すれば、誤分類確率を最小化できるかの研究を行っている例も見られますが無駄な試みであることを理解すべきです。

12. CPDの3群判別と多重共線性[42]

鈴村産婦人科教授の自然分娩、かんし分娩、帝王切開という分娩法を予測する研究では、主として帝王切開するか自然分娩にするかの簡便法を鈴村教授が考案しました。これを出産の前に得られる計測値から判別し予測に役立てようという研究です。このデータは多重共線性があり、3個の計測値を省けば多重共線性が解消されることが分かりました。これはSASの誇る全ての説明変数の組み合わせで回帰モデルを検証できるRSQUAREプロシジャで分析しました。19個の説明変数があるので、自然分娩群と帝王切開群を1/-1のダミ変数として分析すると、$(2^{19}-1)\fallingdotseq 2^{20}/2=1024^2/2=524288$個のモデルが検証できます。IBMの汎用機で日中処理依頼をしました。しばらくして、計算センターから親会社の経理処理の業務に多大な影響

を及ぼすのでキャンセルし夜間処理に回してほしいということです。私の趣味で会社の高価なIBM機を使用したので、本来であれば始末書を出すべき内容です。しかしこのデータはその後、私の判別分析の研究に大いに貢献しています。

13. 丸山ワクチンと大阪府立成人病センターとの別れ [45-48]

丸山ワクチンの分析を、三宅先生から依頼されました。三宅先生は高校の同窓である東大医学部教授の開原先生に相談したが色々議論した結果、私を推薦したとのことでした。私はすでに大阪府立成人病センターの疫学部の鈴木隆一郎先生、中西克己先生らと、がんの疫学調査の分析を継続し厚生省の梅垣班で数年報告していました。当時大阪府立成人病センターから依頼された研究も行っていました。両先生に状況を説明し、丸山ワクチンの統計分析に特化することにしました。

仕事が終わってから日本医科大学で、夜間に丸山先生の門下生の数名のボランティアの医師と三宅先生らと検討会をもちました。ある薬の薬効を検証するためには、その薬を投与した群と投与しない群を、医者も患者もどちらに割りつけられたかを知らさずに、投与しない群には偽薬を投与して行う2重盲検法が行われています。しかし、丸山ワクチン研究施設にあるデータは、32万件の患者さん全てに丸山ワクチンが投与されています。ある晩、丸山先生が慰労のため参加され、「丸山ワクチンは副作用がないが、水のように効果がない」と批判された悔しさを淡々と説明されました。しばらくして、閃きました。「丸山ワクチンは副作用のない水のようなものであると認めましょう。そして、手術後1年以内に投与開始した患者さんを、3カ月単位で4群に分け、それらの生存時間の平均値に差があるか否かを調べましょう。もし水であれば差がなく、早く投与した患者群の生存期間が長ければ、水であるという帰無仮説を棄却できます」ということで研究を始めました。そして術後3カ月以内に投与した群の生存時間が9カ月以降1年以内に投与した患者群の平均余命より平均が長いことが分かりました。医療情報学会での発表数日前に、人生で、最初で最後の新聞記者のインタビューを受けました。発表当日は多くのマスコミがくるので事前のインタビューとのことです。ところが多分前日に、認可見送りで有償治験薬の継続が決まってマスコミの騒ぎが収まりました。丸山ワクチンに関しては、その後長期生存例の患者像の特定を試みましたが、どうしても多くの人が納得する成果は得られませんでした。その後、東京大学の大橋先生らがゼリア新薬との共同研究で、PhaseⅡの研究成果を報告されていました[49]。

14. 決定木分析と介護保険

決定木分析は、私の人生に深くかかわっています。私の一生の研究テーマである判別分析は、野村医師の作った枝分かれ論理にかなわなかった点です。野村医師は、医学診断の知識に基づいて枝分かれ論理を作成しましたが、統計手法として「決定木分析（パーティション）」で実現できます。そして次が、統計手法が日本の社会的インフラに使われた介護保険です。ただし、私自身がもっと積極的に関与していればよかったのですが、当初混乱を生みました。一つは、分析に用いていない在宅の介護対象にも適用したことです。言ってみれば、対象外の母集団に適用したことです。次に、順序尺度を名義尺度として扱ったことです。最後は、まだ正式には検証していませんが、多分岐のCHAIDが必ずしも2分岐に比べて良いわけではない点です。これは、野村医師から聞かされた「分岐の上位水準で分岐されたものであっても、下位で他の可能性が出てくれば、上位の他の別の分岐にフィードバックする」という経験知が思い出されます。多分岐が必ずしも良くないことと、小標本の場合に回帰木は一元配置の分散分析、分類木の場合は分割表の独立性の検定で行うことは、息子の卒業研究のテーマとして与えました。私の人生で、最初で最後の数時間勉強の面倒を見たのですが、もう少し教育に関与しておけばと反省しています。

文献：

[9] 新村秀一（1989）. 易しく実践 データ解析の進め方. 共立出版.

[24] 新村秀一（1999）. パソコン楽々数学. 講談社ブルーバックス.

[25] 新村秀一（1992）. 実践数理計画法. 朝倉書店.

[26] 新村秀一（1987）. 体験に基づく汎用統計パッケージの紹介. 品質、17-3、261-268.

[27] 新村秀一（1985）. アメリカから吹き寄せる新しい高級言語の風SAS、ソフトウエア流通、28、54-58.

[28] 新村秀一（1987）. 新しい高級言語の風SAS・1－SAS開発の背景－. ビジネスコミュニケーション、24-1、122-125.

[29] 新村秀一（1987）. 新しい高級言語の風SAS・2－－. ビジネスコミュニケーション、24-2.

[30] 新村秀一（1987）. 新しい高級言語の風SAS・3－簡単なSASジョブ－. ビジネスコミュニケーション、24-3、133-136.

[31] 新村秀一（1987）. 新しい高級言語の風SAS・4－データ解析について－. ビジネスコミュニケーション、24-4、140-143.

[32] 新村秀一（1987）. 新しい高級言語の風SAS・5－SAS言語の重層構造－. ビジネスコミュニケーション、24-5、135-139.

[33] 新村秀一（1987）. 新しい高級言語の風SAS・6－SAS/本体のプロセジャーについて－. ビジネスコミュニケーション、24-6、108-112.

[34] J. Sall（新村訳）（1986）. SASによる回帰分析の実践、朝倉書店、東京.

[35] 新村秀一（1983）. 行列表現による重回帰分析（1）、オペレーションズ・リサーチ、28-9、439-445.

[36] 新村秀一（1983）. 行列表現による重回帰分析（2）、オペレーションズ・リサーチ、28-10、506-512.

[37] 新村秀一（1983）. 重回帰分析における掃出し演算子、オペレーションズ・リサーチ、28-11、565-569.

[38] 新村秀一（1983）. 重回帰分析におけるモデル決定、オペレーションズ・リサーチ、28-12、620-626.

[39] 新村秀一（1986）. 科学万博データの解析、オペレーションズ・リサーチ、30-12、754-766.

[40] 新村秀一（1988）. データ解析に見るグラフ、オペレーションズ・リサーチ、38-4、172-177.

[41] A.Miyake & S.Shinmura (1976). Error rate of linear discriminant function, F.T. de Dombal & F.Gremy , editors 435-445, North- Holland Publishing Cmpany

[42] 新村秀一、三宅章彦（1983）. 重回帰分析と判別解析のモデル決定（1）— 19変数をもつC.P.Dデータの多重共線性の解消 —、医療情報学、3-3、107-124.

[45] 新村秀一、飯田和美、丸山千里（1987h）. SSM（人型結核菌鯛抽出物質、丸山ワクチン）の癌治療における帰無仮説モデルによる評価. 医療情報学、7-3、263-276.

[46] 新村秀一、飯田和美、岩城弘子、丸山千里、三宅章彦（1984）. SSM（丸山ワクチン）の癌治療における統計的評価、第4回医療情報連合大会論文集、614-619.

[47] 飯田和美、丸山千里、新村秀一（1985）. SSM（丸山ワクチン）の癌治療における統計的評価の追跡調査、第5回医療情報連合大会論文集、619-622.

[48] 新村秀一、飯田和美、三宅章彦、岩城弘子、丸山千里（1985）. SSM（丸山ワクチン）の癌治療における統計的評価（2）、第5回医療情報連合大会論文集、623-626.

[49] K.Noda,Y.Ohashi,etal. (2006). Randomized PhaseⅡ Sturdy of Immunomodulator Z-100 in Patients with Stage ⅢB Cervical Cancer with Radiation Therapy. Jpn J Clin Oncol 2006,36 (9),570-578.

Q&A

- SMTP SSL/TLS について
- 不正な文字コードを取り除く方法
- MEMSIZEシステムオプションのデフォルト設定
- IML実行時に他のプロシジャを呼び出す方法
- IML実行時にR言語を呼び出す方法
- SASADMユーザーのパスワードについて
- ログ行番号を1からカウントしなおしたい
- 前の処理で扱ったデータセットの オブザベーション数を取得する方法について

Q SAS の FILENAME EMAIL エンジンは SMTP SSL/TLS をサポートしていますか。

A FILENAME EMAIL エンジンでは SSL/TLS による SMTP 通信の暗号化をサポートしておりません。従って SSL/TLS で暗号化された通信のみを許可する SMTP サーバに対して、SAS プログラムから電子メールを送信することはできません。

Q 文字変数の中に SAS のセッションエンコーディングでは表示できない文字コード、いわゆる不正な文字が含まれている可能性があります。このような不正な文字を、簡単に変数から取り除く方法はありますか。

A SAS® 9.2 より前のリリースでは、不正な文字を簡単に取り除く方法はありませんでした。SAS 9.2 では、新しく追加された KPROPDATA 関数を使うことで不正な文字を別の文字に置き換えることができます。次のサンプルプログラムでは、KPROPDATA 関数を使って、変数 a に含まれる不正な文字を半角スペースに置き換えています。

例

```
DATA dirty;
  /* 不正な文字を含む文字列 */
  a=" データ " || "93"X;
RUN;
DATA clean;
  SET dirty;
  /* 不正な文字を半角スペースに置き換える */
  a=KPROPDATA(a,'BLANK');
RUN;
```

なお、KPROPDATA関数の構文については次のページをご覧ください。

KPROPDATA Function

http://support.sas.com/documentation/cdl/en/nlsref/61893/HTML/default/viewer.htm#a003285543.htm

Q SAS® 9.3 にて前のリリースでは実行できていたプログラムがメモリ不足のため、エラーとなります。MEMSIZE システムオプションの値を確認すると、2G となっており、以前の設定よりは小さくなっていました。MEMSIZE システムオプションのデフォルトは変更されたのでしょうか。

A Windows 環境では、MEMSIZE システムオプションのデフォルトの値は 0(MAX と同等) となっており、実行環境が許容する限り、最大のメモリを用います。SAS 9.3 では、CFG ファイルに MEMSIZE システムオプションが追記され、デフォルトでは 2G の設定に変更されています。MEMSIZE システムオプションの値を変更する場合、CFG ファイルの最後の部分にあります MEMSIZE システムオプションの記述を変更します。

```
/* 変更前 */
-MEMSIZE 2G

/* 変更後 */
-MEMSIZE 0
```

Q IML プロシジャを実行したまま、他のプロシジャを実行し、その結果を IML プロシジャの処理に使えますか。

A SAS® 9.2 TS2M3 (SAS/IML 9.22) では、IML プロシジャに新たに SUBMIT ステートメント、ENDSUBMIT ステートメントが追加されています。これらのステートメントを用いることで、IML プロシジャの実行中のまま、他のプロシジャを実行できます。また、プロシジャの結果を ODS OUTPUT ステートメントにてデータセットに出力し、行列として読み込むことで、その後の IML プロシジャでの処理に使えます。以下のサンプルでは、REG プロシジャを実行し、パラメータ推定値、標準誤差の値を行列として読み込んでいます。

例

```
PROC IML;
  SUBMIT;
                          /* プロシジャを用いた処理を記述 */
    PROC REG DATA=sashelp.class;
      MODEL weight=height;
      ODS OUTPUT ParameterEstimates=out;
    RUN;
    QUIT;
  ENDSUBMIT;
                  /* 出力したデータセット ( 一部 ) を読み込む */
  USE out;
  READ all VAR {"Estimate" "StdErr"} INTO result;
  CLOSE out;
                                  /* 行列の内容を確認 */
  PRINT result;
QUIT;
```

Q IMLプロシジャにて、R言語を呼び出し、実行することはできますか。

A SAS® 9.2 TS2M3（SAS/IML 9.22）にてSUBMITステートメントが追加されています。このステートメントにRオプションを追記することでR言語を呼び出し、実行できます。

例

```
PROC IML;
  SUBMIT / R;
                              # R言語のプログラムを記述 #
  ENDSUBMIT;
QUIT;
```

なお、この実行には以下が必要となります。

1. 実行環境でのR言語のインストール

R言語に関しては、以下のページをご参照ください。

http://cran.r-project.org/

2. RLANGシステムオプションの指定

デフォルトではNORLANGとなっていますので、起動時に-RLANGシステムオプションを追加します。または、CFGファイルにこのシステムオプションを追加します。

※SAS 9.2 TS2M3 Windows x64環境の場合、以下のProblem NoteからHotFixをダウンロード、適用してください。

http://support.sas.com/kb/40/252.html

Q 現在SAS® 9.2を利用しておりsasadm以外の無制限ユーザーを用意したいと考えています。どのように設定すればよいでしょうか。

A SAS 9.2以降では従来のSAS 9.1.3とは、必須ユーザーに対する管理方法が異なります。設けたい無制限ユーザーが外部ユーザー（OS上に実際に存在するID）とするのか、SASの内部ユーザーとするのかでも管理や設定の方法は異なります。単に管理者ユーザーとしてMetadataを管理可能なユーザーを用意するだけであれば、内部ユーザーを新規に作成されてはいかがでしょうか。

1. 管理コンソールに管理者ユーザーでログイン
2. 新規にユーザーを作成
3. 「アカウント」タブにて「内部アカウントの作成」を選択
4. パスワードなど必要な設定を行う
   ※ユーザーは自動的に作成したID@saspw のログイン情報を持ちます。
5. 「グループと役割」のタブにて、「SAS Administrators group」に所属させるか、明示的に定義するのであれば、「グループと役割」のタブにて、「Metadata Server: 無制限」と「Management Console: 詳細」の役割を与えます。
6. 管理コンソールに2.で作成したID@saspwとパスワードでログインします。
7. 上記で管理コンソール上からのMetadataの情報管理が可能となります。

詳細は次のマニュアルにて記載しております。

SAS 9.2 インテリジェンスプラットフォーム

http://www.sas.com/japan/service/documentation/onlinedoc/intellplatform/index.html#intell92

・SAS 9.2 Intelligence Platform: Security Administration Guide

・SAS(R) 9.2 Management Console: Guide to Users and Permissions => 「User Administration Tasks」

=> 「Add Administrators」

Q SASログに記載されている行番号を1から再度採番したいのですが、可能でしょうか。

A SAS® 9.3より、RESETLINEステートメントが追加されました。採番しなおしたい際に、下記のステートメントを記述することで、再度1からログ行が表示されます。

```
RESETLINE;
```

Q 前の処理で扱ったデータセットのオブザベーション数を取得する際、例えば、SQLプロシジャで変数に対してCOUNT関数を使用してマクロ変数へ格納する方法、DATAステップにて自動変数 _N_ をCALL SYMPUTにてマクロ変数へ格納する方法があるかと思いますが、もっと簡単に取得する方法はないのでしょうか。

A SAS® 9.3では、前の処理で扱ったデータセットのオブザベーション数をSYSNOBS自動マクロ変数へ格納します。このため、事前にマクロ変数として定義しておく必要がありません。

従来の方法

```
DATA sample;
  DO i = 1 TO 15;
    OUTPUT;
  END;
RUN:

DATA _NULL_;
  SET sample;
  CALL SYMPUTX('obs_count',_N_);
RUN;

DATA _NULL_;
  PUT " オブザベーション数は、 &obs_count. です。 ";
RUN;
```

SYSNOBSを用いた方法

```
DATA sample;
  DO i = 1 TO 15;
    OUTPUT;
  END;
RUN:

DATA _NULL_;
  PUT " オブザベーション数は、 &SYSNOBS. です。 ";
RUN;
```

SAS Marketing News
マーケティングニュース

SAS Executive Briefing

SASでは、現在、全社を挙げて取り組んでおりますHigh-Performance Analytics（以下HPA）について、弊社経営幹部、自らご紹介するためにワールドツアーを実施いたしました。5月のEMEA（ヨーロッパ・中東・アフリカ地域）ツアーに続き、アジア地域では 7月、ムンバイ、シンガポール、香港、ソウルで実施し、日本ではそのツアーの最終日として、7月13日（金）グランドハイアット東京にて実施しました。このセミナーにはDr. グッドナイトCEO、ミカエル・ハグストローム上級副社長、ジム・デイビス上席副社長兼CMOをはじめとするSASのエグゼクティブが来日し、日本の顧客企業に向けてHPAのテクノロジーの先進性とビジネス価値を訴求しました。当日は、SASユーザー企業やパートナーのエグゼクティブ約150名の参加者を迎え、HPA、VA（SAS Visual Analytics）の最新デモンストレーションを中心としたセミナーと、レセプションの2部構成で実施されました。レセプションにおいて、来日したSASのエグゼクティブは参加者と積極的に交流し、情報交換をしました。

<SAS High-Performance Analytics（HPA）について>
http://www.sas.com/offices/asiapacific/japan/platform/hpa/index.html

<SAS Visual Analytics（VA）について>
http://www.sas.com/offices/asiapacific/japan/platform/bi/va.html

Dr. グッドナイトCEOによるHPAのプレゼンテーション

ユーザー総会

8月2日（木）～3日（金）にタワーホール船堀にてSASユーザー総会2012が開催されました。期間中の延べ人数が500名を超えるSASユーザーの方に参加いただき、さまざまな業種や分野にわたった利用方法などについての発表が行われました。また、教育機関やシンクタンク等で活躍する各業界のスペシャリストによる基調講座やSASの経験豊かな講師によるハンズオントレーニング、お客様同士の情報交換を積極的に行っていただく懇談会やワークショップ等、SASの活用について熱い議論が交わされました。

また、一般参加のお客様にはアンケートにお答えいただき、昨年と比較すると年代別使用者が20代で見ると20%増、使用年数別では、1年未満が32%増となり、次世代のアナリストの存在を窺わせる結果がみられました。
今年の傾向として、業種に関わらず多くの方がデータ分析・予測の高度化や、SASの得意とするデータ加工やデータ統合に興味を寄せられました。特に人気を集めたセッションは、東京理科大学大学院浜田知久馬先生のSASによる2値データ解析「ここまでできるFREQプロシジャv9.3」で、271名の来場者を集め盛況でした。数々のセッションでも200名近い来場者で人気だったのが最新手法を紹介したセッションやビックデータ関連でした。論文発表は45本を超える発表があり、優秀者にはSASユーザー会より、表彰ならびに副賞が授与されました。2013年も今年以上の知識の共有の場であるようユーザー総会を益々活性化していきます。

<SASユーザー総会Webサイト>
http://www.sas.com/offices/asiapacific/japan/usergroups/index.html

<SASユーザー総会2012 セッション資料>
http://www.sascom.jp/campaign/usergroups2012/?NM201208
（※参照には、SASプレミアムラウンジへの登録が必要です。）


アナリティクスセッションの様子


懇親会で情報交換する皆様

SAS Publications

新刊書籍のご案内

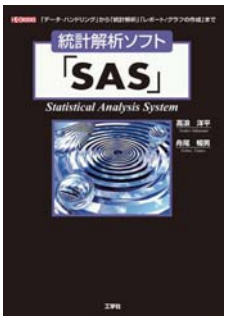
「データ・ハンドリング」から、「文法解説」「統計解析」「レポート/グラフの作成」まで

新刊書籍をご紹介します。
以下、出版元の工学社Webサイトより抜粋です。

「SAS」は、米国SAS Institute社の「統計解析ソフト」です。
「R」や「SPSS」と並んで、世界的に人気の高い統計ソフトで、1960年代から主に大学や研究所で科学・工学分野の研究に用いられてきました。近年ではビジネス向けの機能や関連製品も充実し、日本でも「医薬関連」をはじめ、「自動車」「通信」「金融」「保険」など、さまざまな業種で使われています。
また「SAS」には、有用な機能が豊富に用意されており、統計解析に留まらず、高品質な「グラフ」や「レポート」も作れます。
本書は、これから「SAS」を始める方のための入門書としてはもちろん、SAS中級者や上級者にとっても、時間をかけずに必要な情報を得ることができる内容になっています。

本書の詳細は、以下のURLにてご確認ください。
http://www.kohgakusha.co.jp/books/detail/978-4-7775-1710-7

統計解析ソフト「SAS」
●ISBN：
978-4-7775-1710-7 C3041 ¥3500E
●著者：
高浪 洋平・舟尾 暢男（共著）
●発売日：2012年9月27日
●サイズ：B5版
●ページ数：352ページ
●価格：3,675円（本体:¥3,500）



Latest Releases

最新リリース情報

PCプラットフォーム

Windows版(32-bit/64-bit)	SAS 9.1.3 / 9.2 / 9.3
64-bit Windows(Itanium)版	SAS 9.1.3 / 9.2

メインフレームプラットフォーム

IBM版(OS/390,z/OS)	SAS 9.1.3 / 9.2 / 9.3

UNIXプラットフォーム

SunOS/Solaris版	SAS 9.1.3 / 9.2 / 9.3
HP-UX版	SAS 9.1.3 / 9.2 / 9.3
HP-UX(Itanium)版	SAS 9.1.3 / 9.2 / 9.3
AIX版	SAS 9.1.3 / 9.2 / 9.3
Linux(Intel) 版	SAS 9.1.3 / 9.2 / 9.3

SAS Technical News入手
SAS Technical Newsは、右記のURLから入手できます。 http://www.sas.com/jp/periodicals/technews/index.html

発行:SAS Institute Japan株式会社

STN
SAS Technical News
AUTUMN 2012

■テクニカルニュースに関するお問い合わせ先

テクニカルサポートグループ
TEL:03-6434-3680　FAX:03-6434-3681

SAS Institute Japan株式会社

本社
〒106-6111
東京都港区六本木6-10-1
六本木ヒルズ森タワー 11F
Tel 03(6434)3000
Fax 03(6434)3001

大阪支店
〒530-0004
大阪市北区堂島浜1-4-16
アクア堂島西館 12F
Tel 06(6345)5700
Fax 06(6345)5655

www.sas.com/jp