本書は、Minitab 統計ソフトウェアのアシスタントで使用される方法およびデータチェックを開発するため、Minitab の統計専門家によって行われた調査に関する一連の文書の 1 つです。

単回帰

概要

アシスタントの単回帰手順は、最小二乗法による推定を使用して、1 つの連続予測変数（X）と 1 つの連続応答変数（Y）を含む線形および 2 次モデルを適合します。ユーザはモデルタイプを選択するか、最適なモデルの選択をアシスタントに任せることができます。本書では、アシスタントが回帰モデルの選択に使用する基準について説明します。

さらに、有効な回帰モデルを取得するために重要ないくつかの因子を調査します。まず、サンプルは、検定に十分な検出力を有し、X と Y の関係の強度を推定するために十分な精度を備えた大きさであることが必要です。次に、分析の結果に影響を及ぼす可能性がある異常なデータを特定することが重要です。また、誤差項は正規分布に従っているという仮定を考慮し、全体のモデルと係数の仮説検定に対する非正規性の影響も評価します。最後に、モデルの有用性を確保するために、X と Y の関係を正確に反映するモデルタイプを選択することが重要です。

アシスタントはこれらの因子に基づき、データで次のチェックを行い、レポートカードに結果を表示します。

- データ量
- 異常なデータ
- 正規性
- モデルの適合

本書では、これらの因子が実際にどのように回帰分析に関連するかを調査し、アシスタントでこれらの因子をチェックするためのガイドラインをどのように定めたかについて説明します。

回帰法

モデル選択

アシスタントの回帰分析は、1 つの連続予測変数と 1 つの連続応答変数を含むモデルを適合します。次の 2 つのモデルタイプを適合できます。

- 線形：$F(x) = \beta_0 + \beta_1 X$
- 2 次：$F(x) = \beta_0 + \beta_1 X + \beta_2 X^2$

ユーザーが分析を実行する前にモデルタイプを選択するか、アシスタントにモデルの選択を任せることができます。データに最適なモデルを判断するために使用される、いくつかの方法があります。モデルを確実に役立たせるため、X と Y の関係を正確に反映するモデルタイプを選択することが重要です。

目的

アシスタントでどちらのモデルを使用するかを決定するため、モデル選択に使用できる異なる方法を調査します。

方法

モデル選択に一般的に通常使用される 3 つの方法（Neter et al., 1996）を調査しました。1 つ目の方法は、最高次の項が有意であるモデルを識別します。2 つ目の方法は、R 二乗（調整済み）値が最大のモデルを選択します。3 つ目の方法は、全体の F 検定が有意であるモデルを選択します。詳細は、「付録 A」を参照してください。

アシスタントで使用する方法を決定するため、これらの方法を調査し、互いの計算値を比較しました。また、品質分析の専門家によるフィードバックも収集しました。

結果

この調査の結果、モデル内の最高次の項の統計的有意性に基づくモデルの選択方法を使用することを決定しました。アシスタントは最初に 2 次モデルを調べ、モデル内の 2 次項（β_2）が統計的に有意であるかどうかを検定します。その項が有意でない場合、モデルから 2 次項を削除し、線形項（β_1）を検定します。この方法で選択されたモデルは、モデル選択レポートに表示されます。さらに、ユーザーがアシスタントによって選択されたモデルとは異なるモデルを選択した場合、そのモデルがモデル選択レポートとレポートカードに表示されます。

この方法を選択した理由の一つに、一般に専門家は有意ではない項を除外する単純なモデルを好むというフィードバックへの対応が挙げられます。さらに、各方法の比較結果によると、モデル内の最高次の項の統計的有意性を使用する方法は、最大の R 二乗（調整済み）値に基づいてモデルを選択する方法よりも厳格です。詳細は、「付録 A」を参照してください。

モデルの選択には最高次のモデル項の統計的有意性を使用しますが、モデル選択レポートにはモデルの R 二乗（調整済み）値と全体の F 検定も表示されます。レポートカードに表示されるステータスインジケータについては、次の「モデルの適合」セクションのデータチェックを参照してください。

データチェック

データ量

検出力は、仮説検定が正しくない場合にどの程度の確率で帰無仮説を棄却するかによります。回帰の場合、帰無仮説は「X と Y の間に関係はない」です。データセットが小さすぎる場合、検定の検出力は実際に存在する X と Y の関係を適切に検出できない可能性があります。そのため、実質的に重要な関係を高い確率で検出するために十分な大きさのデータセットが必要です。

目的

X と Y の関係に対する全体の F 検定の検出力、および X と Y の関係の強度の推定値である R 二乗（調整済み）の精度に対して、データ量がどのように影響するかを判断します。この情報は、データで観測される関係の強度が真の関係の基本強度の確実な指標であると信頼するために、十分な大きさのデータセットがあるかどうかを判断するために不可欠です。R 二乗（調整済み）の詳細は、「付録 A」を参照してください。

方法

全体の F 検定の検出力を調べるため、さまざまな R 二乗（調整済み）値とサンプルサイズで検出力を計算しました。また、R 二乗（調整済み）の精度を調べるため、さまざまな母集団調整済みR^2（ρ 二乗（調整済み））の値とさまざまなサンプルサイズで R 二乗（調整済み）の分布をシミュレートしました。R 二乗（調整済み）値の変動性を調べ、R 二乗（調整済み）が ρ 二乗（調整済み）に近くなるために必要なサンプルの大きさを判断しました。計算の詳細は、「付録 B」を参照してください。

結果

やや大きいサンプルの場合、X と Y の関係が実際に重要になるほど強くない場合でも、回帰にはその関係を検出するのに十分な検出力があることがわかりました。具体的には、次のことが判明しました。

- サンプルサイズが 15 で X と Y の関係が強い（ρ 二乗（調整済み）= 0.65）場合、統計的に有意な線形関係を検出する確率は 0.9969 です。そのため、15 以上のデータ点の検定で統計的に有意な関係を検出できない場合、真の関係はあまり強くない（ρ 二乗（調整済み）値< 0.65）可能性が高いと言えます。
- サンプルサイズが 40 で X と Y の関係がやや弱い（ρ 二乗（調整済み）= 0.25）場合、統計的に有意な線形関係を検出する確率は 0.9398 です。そのため、40 のデータ点での F 検定は、X と Y の関係がやや弱い場合でもその関係を検出する可能性が高いと言えます。

回帰は X と Y の関係を非常に簡単に検出できます。そのため、統計的に有意な関係を検出する場合、R 二乗（調整済み）を使用して関係の強さも評価する必要があります。サンプルサイズが十分な大きさでない場合、R 二乗（調整済み）の信頼度は低く、サンプルによって大きく異なる可能性があります。ただし、サンプルサイズが 40 以上の場合、R 二乗（調整済

み）値はより安定し、信頼できることがわかりました。40 のサンプルサイズでは、実際の値やモデルタイプ（線形または 2 次）に関係なく、90%の信頼度で R 二乗（調整済み）の観測値が ρ 二乗（調整済み）の 0.20 以内になります。シミュレーション結果の詳細は、「付録 B」を参照してください。

これらの結果に基づき、データ量をチェックするときに、アシスタントレポートカードに次の情報が表示されます。

ステータス	状態
	サンプルサイズ < 40 サンプルサイズが関係の強度を正確に推定するのに十分な大きさではありません。R 二乗や調整済み R 二乗など、関係の強度の測度は大きく変化する場合があります。正確に推定するには、さらに大きなサンプル（通常 40 以上）を使用してください。 **サンプルサイズ ≥ 40** サンプルには関係の強度を正確に推定するのに十分な大きさがあります。

異常なデータ

アシスタントの回帰手順では、異常なデータを標準化残差またはてこ比値の大きい観測値と定義します。通常、これらの測定値は回帰分析の異常なデータを特定するために使用されます（Neter et al., 1996）。異常なデータは結果に大きく影響する可能性があるため、分析を有効にするにはデータの修正が必要な場合があります。ただし、異常なデータは工程の自然変動により発生する可能性もあります。そのため、異常な動作の原因を特定し、そのようなデータ点の処理方法を決定することが重要です。

目的

データ点が異常であるという信号を出すために必要な標準化残差およびてこ比値の大きさを決めようと考えました。

方法

Minitab 標準の回帰手順（[統計] > [回帰] > [回帰]）に基づき、異常な観測値を特定するためのガイドラインを作成しました。

結果

標準化残差

標準化残差は、残差（e_i）をその標準偏差の推定値で割ったものです。一般に、標準化残差の絶対値が 2 より大きい場合に観測値が異常であると見なされます。しかしながら、このガイドラインはやや控え目です。すべての観測値のうち約 5%が偶然の所産としてこの基準を満たすことが予測されます（誤差が正規分布に従う場合）。そのため、異常な動作の原因を調査し、観測値が実際に異常であるかどうかを判断することが重要です。

てこ比値

てこ比値は観測の X 値にのみ関連し、Y 値には依存しません。てこ比値がモデル係数（p）を観測値数（n）で割った数の 3 倍よりも大きい場合、観測値が異常と見なされます。一部

の教科書では$\frac{2 \times p}{n}$を使用していますが、これが一般的に使用される基準値です（Neter et al.， 1996）。

データに高いてこ比値点が含まれる場合、それらの点がデータに適合するために選択したモデルタイプに不適切に影響するかどうかを調査します。たとえば、1 つの極端な X 値によって、線形モデルの代わりに 2 次モデルが選択される可能性があります。2 次モデルで観測された曲線性が、工程に関する理解と一致するかどうかを確認する必要があります。一致しない場合、より単純なモデルをデータに適合するか、追加データを収集してさらに細かく工程を調査します。

異常なデータをチェックするときに、アシスタントレポートカードに次のステータスインジケータが表示されます。

ステータス	状態
	異常なデータ点はありません。異常なデータ点は、結果に強い影響を与える可能性があります。
	1 つ以上の大きな標準化残差または 1 つ以上の高いてこ比値があります。 点の上にマウスを置くか、Minitab のブラッシング機能を使用すると、ワークシートの行を特定できます。異常なデータは、結果に強い影響を与える可能性があるため、異常な性質の原因を特定してください。データ入力または測定のエラーを修正します。特殊原因に関連するデータを削除し、分析を再実行することを検討してください。

正規性

一般に、回帰はランダム誤差（ε）が正規分布に従っているという仮定に基づきます。係数（β）の推定の仮説検定を実施する場合、正規性の仮定が重要になります。ただし、ランダム誤差が正規分布に従っていない場合でも、サンプルが十分に大きければ検定結果は多くの場合信頼できます。

目的

正規分布に基づいて信頼できる結果を得るために必要なサンプルのサイズを判断します。実際の検定結果がその検定の目標有意水準（α、またはタイプ I 過誤率）とどの程度近く一致するか、つまり、異なる非正規分布で期待されるより高い頻度または低い頻度で検定が帰無仮説を誤って棄却するかどうかを判断します。

方法

タイプ I 過誤率を推定するため、正規分布から大きく逸脱する歪んだ分布、裾の重い分布、および裾の軽い分布で複数のシミュレーションを実行しました。これらのシミュレーションは、サンプルサイズを 15 として、線形モデルと 2 次モデルに対して行いました。全体の F 検定とモデル内の最高次の項の検定との両方の調査を行いました。

各条件に対して 10,000 回の検定を行いました。各検定で帰無仮説が真になるように、ランダムデータを生成しました。次に、目標有意水準に 0.05 を使用した検定を実行し、10,000 回のうち実際に帰無仮説を棄却した回数を数え、この比率を目標有意水準と比較しました。検定が適切に機能する場合、タイプ I 過誤率は目標有意水準に非常に近くなります。このシミュレーションの詳細は、「付録 C」を参照してください。

結果

全体のＦ検定とモデル内の最高次の項の検定の両方で、統計的に有意な結果を得る確率は、どの非正規分布でも大きな差はありません。タイプⅠ過誤率はすべて 0.038～0.0529 で、目標有意水準の 0.05 に非常に近くなっています。

検定は比較的小さいサンプルで適切に機能するため、アシスタントはデータの正規性を検定しません。代わりに、サンプルのサイズをチェックし、サンプルが 15 未満の場合に通知します。アシスタントの回帰レポートカードに次のステータスインジケータが表示されます。

ステータス	状態
	サンプルサイズが 15 以上であるため、正規性は問題になりません。
	サンプルサイズが 15 未満であるため、正規性が問題になる可能性があります。p 値を解釈するときに注意が必要です。小さなサンプルを使用する場合、p 値の正確性は非正規残差誤差の影響を受けやすくなります。

モデルの適合

回帰分析を実行する前に線形モデルまたは 2 次モデルを選択するか、アシスタントにモデルの選択を任せることができます。適切なモデルを選択するために複数の方法を使うことができます。

目的

アシスタントでどちらの方法を使用するかを決定するため、モデルタイプの選択に使用できる異なる方法を調査します。

方法

モデル選択に一般的に使用される 3 つの方法を調査しました。1 つ目の方法は、最高次の項が有意であるモデルを識別します。2 つ目の方法は、R 二乗（調整済み）値が最大のモデルを選択します。3 つ目の方法は、全体のＦ検定が有意であるモデルを選択します。詳細は、「付録 A」を参照してください。

アシスタントで使用する方法を決定するため、これらの方法を調査し、互いの計算値を比較しました。さらに、私たちは品質分析のエキスパートからフィードバックを集めました。

結果

モデル内の最高次の項の統計的有意性に基づいてモデルの選択をする方法を使用することにしました。アシスタントは最初に 2 次モデルを調べ、モデル内の 2 次項（β_3）が統計的に有意であるかどうかを検定します。その項が有意でない場合、線形モデル内の線形項（β_1）を検定します。この方法で選択されたモデルは、モデル選択レポートに表示されます。さらに、ユーザーがアシスタントによって選択されたモデルとは異なるモデルを選択した場合、そのモデルがモデル選択レポートとレポートカードに表示されます。詳細は、上記の「回帰法」セクションを参照してください。

この結果に基づき、アシスタントの回帰レポートカードに次のステータスインジケータが表示されます。

ステータス	状態
	ユーザーのモデルがアシスタントによる最適なモデルと一致する場合 データとモデルが目標値に適合するかどうかを評価する必要があります。適合線プロットを調べて次のことを確認します。 • サンプルに X 値の範囲が適切に含まれている。 • モデルがデータの曲線性に適切に適合する（過剰な適合を回避）。 • 線が特別な関心領域内に適切に適合する。 **ユーザーのモデルがアシスタントによる最適なモデルと一致しない場合** モデル選択レポートに、より適切な可能性がある代替モデルが表示されます。

参考文献

Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996). *Applied linear statistical models*. Chicago: Irwin.

付録 A：モデル選択

応答変数 Y に対する予測変数 X に関する回帰モデルは次の形式です。

$Y = f(X) + \varepsilon$

ここで、関数 f(X)は与えられた X に対する Y の期待値（平均）を表します。

アシスタントでは、関数 f(X)の形式として次の 2 つの選択肢があります。

モデルタイプ	f(X)
線形	$\beta_0 + \beta_1 X$
2 次	$\beta_0 + \beta_1 X + \beta_2 X^2$

係数βの値は未知で、データから推定する必要があります。推定方法は、サンプルの残差の平方和を最小化する最小二乗法です。

$$\min \sum_{i=1}^{n} \left(Y_i - \hat{f}(X_i)\right)^2 .$$

残差は、推定された係数に基づく観測応答Y_iと適合値$\hat{f}(X_i)$間の差です。この平方和の最小化された値は、指定されたモデルの SSE（誤差平方和）です。

アシスタントでモデルタイプの選択に使用する方法を決定するため、次の 3 つのオプションを評価しました。

- モデル内の最高次の項の有意性
- モデルの全体の F 検定
- 調整されたR^2値（R 二乗（調整済み））

モデル内の最高次の項の有意性

この方法では、アシスタントは 2 次モデルから開始し、2 次モデル内の 2 次項に対する次の仮説検定を行います。

$H_0: \beta_2 = 0$

$H_1: \beta_2 \neq 0$

この帰無仮説が棄却された場合、アシスタントは 2 次項の係数がゼロ以外であると結論付け、2 次モデルを選択します。それ以外の場合、アシスタントは線形モデルに対する次の仮説検定を行います。

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

全体の F 検定

この方法は、全体のモデル（線形または 2 次）の検定です。選択された形式の回帰関数 f(X)で次を検定します。

$H_0: f(X)$ は定数です

$H_1: f(X)$ は定数ではありません

調整済みR^2

調整済みR^2（R 二乗（調整済み）、R^2_{adj}）は、モデルによって応答の変動性がどの程度 X に起因するかを測定します。観測された X と Y の関係の強度を測定する一般的な方法は、次の 2 つです。

$$R^2 = 1 - \frac{SSE}{SSTO}$$

および

$$R^2_{adj} = 1 - \frac{SSE/(n-p)}{SSTO/(n-1)}$$

ここで

$$\text{SSTO} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

SSTO は総平方和で、全体の平均$\bar{Y}$での応答の変動を測定します。SSE は回帰関数 f(X)での変動を測定します。R 二乗（調整済み）の調整は完全モデルでの係数の数（p）用で、n - p の自由度によってεの分散を推定します。R^2はモデルに係数が追加されても減少することはありませんが、R 二乗（調整済み）は調整されるため、追加係数によってモデルが改善されない場合は減少する可能性があります。そのため、モデルへの別の項の追加によって応答の追加分散が説明されない場合、R 二乗（調整済み）が減少し、追加項は役に立たないことが示されます。したがって、線形と 2 次の比較には調整された測定を使用する必要があります。

モデル選択方法間の関係

3 つのモデル選択方法の関係、それぞれの計算方法、および互いにどのように影響するかを調べます。

まず、全体の F 検定と R 二乗（調整済み）の計算方法における関係を調べました。全体のモデル検定の F 統計量は、R 二乗（調整済み）の計算でも使用される SSE と SSTO で表すことができます。

$$\text{F} = \frac{(SSTO - SSE)/(p-1)}{SSE/(n-p)}$$

$$= 1 + \left(\frac{n-1}{p-1}\right)\frac{R^2_{adj}}{1 - R^2_{adj}}.$$

上記の計算式は、F 統計量が R 二乗（調整済み）の増加関数であることを示しています。そのため、R 二乗（調整済み）が検定の有意水準（α）で決定された特定の値を超えた場合にのみ、検定で H_0 が棄却されます。これを説明するため、次の表 1 に示されるように、さまざまなサンプルサイズの$\alpha = 0.05$で 2 次モデルの統計的有意性を得るために必要な最小 R 二

乗（調整済み）を計算しました。たとえば、n = 15 の場合に全体の F 検定が統計的に有意になるには、モデルの R 二乗（調整済み）値が 0.291877 以上である必要があります。

表 1　さまざまなサンプルサイズにおいて$\alpha = 0.05$で全体の F 検定が有意になるための最小 R 二乗（調整済み）

サンプルサイズ	最小 R 二乗（調整済み）
4	0.992500
5	0.900000
6	0.773799
7	0.664590
8	0.577608
9	0.508796
10	0.453712
11	0.408911
12	0.371895
13	0.340864
14	0.314512
15	0.291877
16	0.272238
17	0.255044
18	0.239872
19	0.226387
20	0.214326
21	0.203476
22	0.193666
23	0.184752
24	0.176619
25	0.169168
26	0.162318
27	0.155999

サンプルサイズ	最小 R 二乗（調整済み）
28	0.150152
29	0.144726
30	0.139677
31	0.134967
32	0.130564
33	0.126439
34	0.122565
35	0.118922
36	0.115488
37	0.112246
38	0.109182
39	0.106280
40	0.103528
41	0.100914
42	0.098429
43	0.096064
44	0.093809
45	0.091658
46	0.089603
47	0.087637
48	0.085757
49	0.083955
50	0.082227

次に、モデル内の最高次の項の仮説検定と R 二乗（調整済み）間の関係を調べました。2 次モデル内の 2 次項など、最高次の項の検定は、次のように平方和または完全モデル（2 次など）の R 二乗（調整済み）や縮約モデル（線形など）の R 二乗（調整済み）を使用して表すことができます。

$$
\begin{aligned}
F &= \frac{SSE\left(\text{縮約}\right) - SSE\left(\text{完全}\right)}{SSE(\text{完全})/(n-p)} \\
&= 1 + \frac{(n-p+1)\left(R_{adj}^2\left(\text{完全}\right) - R_{adj}^2\left(\text{縮約}\right)\right)}{1 - R_{adj}^2\left(\text{完全}\right)}.
\end{aligned}
$$

この計算式は、F 統計量が$R_{adj}^2\left(\text{縮約}\right)$の固定値に対する$R_{adj}^2\left(\text{完全}\right)$.の増加関数であることを示しています。また、検定統計量は 2 つのR_{adj}^2値間の差に依存することも示しています。特に、統計的に有意になるのに十分な大きさの F 値を得るには、完全モデルの値が縮約モデルの値よりも大きい必要があります。そのため、最適なモデルの選択に最高次の項の有意性を使用するモデルは、R 二乗（調整済み）が最大のモデルを選択する方法よりも厳格です。最高次の項の方法は、より単純なモデルの場合に多くのユーザーに好まれる設定にも対応します。したがって、アシスタントでのモデルの選択には、最高次の項の統計的有意性を使用することに決定しました。

一部のユーザーは、データに最も適合する、R 二乗（調整済み）が最大のモデルを好みます。アシスタントは、モデル選択レポートとレポートカードでこれらの値を提供します。

付録 B：データ量

このセクションでは、全体のモデル検定の検出力、およびモデルの強度の推定値である R 二乗（調整済み）の精度に観測値数 n がどのような影響を及ぼすかを見ていきます。

関係の強度を数量化するため、新しい ρ 二乗（調整済み）を前述のサンプル統計量 R 二乗（調整済み）の母集団に対応する数量として導入します。

$$R_{adj}^2 = 1 - \frac{SSE/(n-p)}{SSTO/(n-1)}$$

これに対して、次のように定義します。

$$\rho_{adj}^2 = 1 - \frac{E(SSE|X)/(n-p)}{E(SSTO|X)/(n-1)}$$

演算子 E(·|X)は、期待値、または X の値が与えられた場合の確率変数の平均を表します。正しいモデルが独立同一分布に従う ε を含む$Y = f(X) + \varepsilon$と仮定すると、次のようになります。

$$\frac{E(SSE|X)}{n-p} = \sigma^2 = Var(\varepsilon)$$

$$\frac{E(SSTO|X)}{n-1} = \sum_{i=1}^{n} \frac{(f(X_i) - \bar{f})^2}{(n-1) + \sigma^2} + \sigma^2 \sum_{i=1}^{n} \frac{(f(X_i) - \bar{f})^2}{(n-1) + \sigma^2}$$

ここで、$\bar{f} = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$.

したがって、

$$\rho_{adj}^2 = \frac{\sum_{i=1}^{n}\left(f(X_i) - \bar{f}\right)^2/(n-1)}{\sum_{i=1}^{n}\left(f(X_i) - \bar{f}\right)^2/(n-1) + \sigma^2}$$

全体のモデルの有意性

全体のモデルの統計的有意性を検定する場合、ランダム誤差 ε が独立で、正規分布に従っていると仮定します。Y の平均が定数($f(X) = \beta_0$)であるという帰無仮説では、F 検定統計量に $F(p-1, n-p)$ 分布があります。対立仮説では、F 統計量には非心パラメータを使用した非心$F(p-1, n-p, \theta)$分布があります。

$$\theta = \sum_{i=1}^{n}\left(f(X_i) - \bar{f}\right)^2/\sigma^2$$

$$= \frac{(n-1)\rho_{adj}^2}{1 - \rho_{adj}^2}$$

H_0 を棄却する確率は、n と ρ 二乗（調整済み）の両方で増加している非心パラメータによって高くなります。

上記の計算式を使用して、線形モデルと 2 次モデルに対して n ＝ 15 の場合のさまざまな ρ 二乗（調整済み）値で全体の F 検定の検出力を計算しました。この結果を表 2 に示します。

表 2　n = 15 で ρ 二乗（調整済み）値が異なる線形モデルと 2 次モデルの検出力

ρ 二乗（調整済み）	θ	F の検出力 線形	F の検出力 2 次
0.05	0.737	0.12523	0.09615
0.10	1.556	0.21175	0.15239
0.15	2.471	0.30766	0.21896
0.20	3.500	0.41024	0.29560
0.25	4.667	0.51590	0.38139
0.30	6.000	0.62033	0.47448
0.35	7.538	0.71868	0.57196
0.40	9.333	0.80606	0.66973
0.45	11.455	0.87819	0.76259
0.50	14.000	0.93237	0.84476
0.55	17.111	0.96823	0.91084
0.60	21.000	0.98820	0.95737
0.65	26.000	0.99688	0.98443
0.70	32.667	0.99951	0.99625
0.75	42.000	0.99997	0.99954
0.80	56.000	1.00000	0.99998
0.85	79.333	1.00000	1.00000
0.90	126.000	1.00000	1.00000
0.95	266.000	1.00000	1.00000

全体的に、X と Y の関係が強くてサンプルサイズが 15 以上の場合、検定の検出力は高いことが判りました。たとえば、ρ 二乗（調整済み） = 0.65 の場合、表 2 には$\alpha = 0.05$で統計的に有意な線形関係の検出確率は 0.99688 であることが示されています。F 検定でのこのような強い関係の検出の失敗は、サンプルの 0.5%未満で発生します。2 次モデルの場合でも、F 検定での関係の検出の失敗はサンプルの 2%未満で発生します。そのため、検定が 15 以上の観測値数で統計的に有意な関係を検出できない場合、真の関係があるとしてもその ρ 二乗（調整済み）値は 0.65 よりも低いことを示しています。ただし、ρ 二乗（調整済み）値が実用的な対象となるには 0.65 以上である必要はありません。

次に、大きいサンプルサイズ（n = 40）で、全体のＦ検定の検出力を調べます。サンプルサイズｎ = 40 はＲ二乗（調整済み）の精度における重要なしきい値であると判断し（次の「関係の強度」を参照）、このサンプルサイズの検出力の値を評価することにしました。線形モデルと２次モデルに対して、n = 40 の場合のさまざまな ρ 二乗（調整済み）値で全体のＦ検定の検出力を計算しました。この結果を表３に示します。

表３　n = 40 で ρ 二乗（調整済み）値が異なる線形モデルと２次モデルの検出力

ρ 二乗（調整済み）	θ	Ｆの検出力 線形	Ｆの検出力 ２次
0.05	2.0526	0.28698	0.21541
0.10	4.3333	0.52752	0.41502
0.15	6.8824	0.72464	0.60957
0.20	9.7500	0.86053	0.76981
0.25	13.0000	0.93980	0.88237
0.30	16.7143	0.97846	0.94925
0.35	21.0000	0.99386	0.98217
0.40	26.0000	0.99868	0.99515
0.45	31.9091	0.99980	0.99905
0.50	39.0000	0.99998	0.99988
0.55	47.6667	1.00000	0.99999
0.60	58.5000	1.00000	1.00000
0.65	72.4286	1.00000	1.00000

ＸとＹの関係がやや弱い場合でも、検出力は高いことがわかりました。たとえば、ρ 二乗（調整済み）= 0.25 の場合でも、表３には$\alpha = 0.05$で統計的に有意な線形関係の検出確率は 0.93980 であることが示されています。観測値数が 40 の場合、ＸとＹの関係がやや弱くても、Ｆ検定がその関係を検出できない可能性は低いと言えます。

関係の強度

すでに示したように、データの統計的に有意な関係はＸとＹの強い基本関係を示すとは限りません。そのため、多くのユーザーは関係の実際の強度をＲ二乗（調整済み）などの指標で確認します。Ｒ二乗（調整済み）を ρ 二乗（調整済み）の推定値とする場合、その推定値が真の ρ 二乗（調整済み）値に適度に近いことを信頼できる必要があります。

Ｒ二乗（調整済み）と ρ 二乗（調整済み）の関係を説明するため、異なる ρ 二乗（調整済み）値に対するＲ二乗（調整済み）の分布をシミュレートし、異なるｎの値で変数Ｒ二乗（調整済み）がどのようになるかを確認しました。次の図 1～4 のグラフは、シミュレート

した 10,000 の R 二乗（調整済み）値のヒストグラムを示しています。サンプルサイズ 15 と 40 での R 二乗（調整済み）の変動性を比較できるように、各ヒストグラムのペアで ρ 二乗（調整済み）の値は同じになっています。0.0、0.30、0.60、および 0.90 の ρ 二乗（調整済み）値を検定しました。すべてのシミュレーションは線形モデルで実行されています。

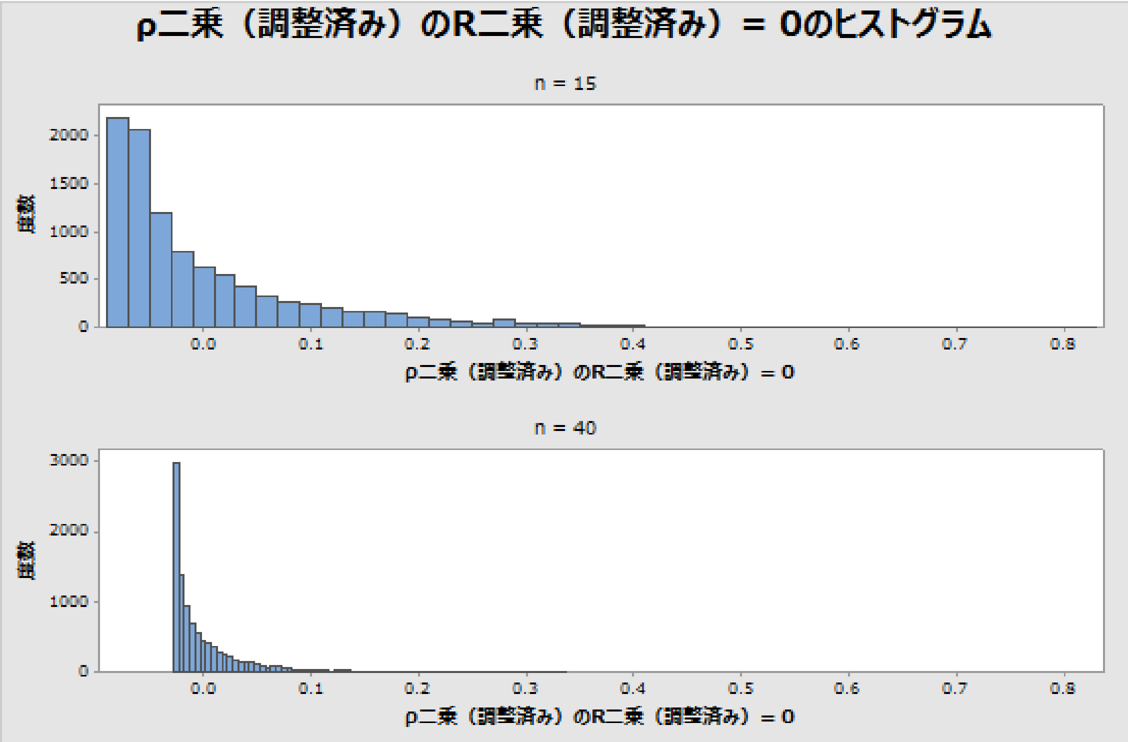

図 1　n = 15 と n = 40 で ρ 二乗（調整済み）= 0.0 の場合にシミュレートした R 二乗（調整済み）値

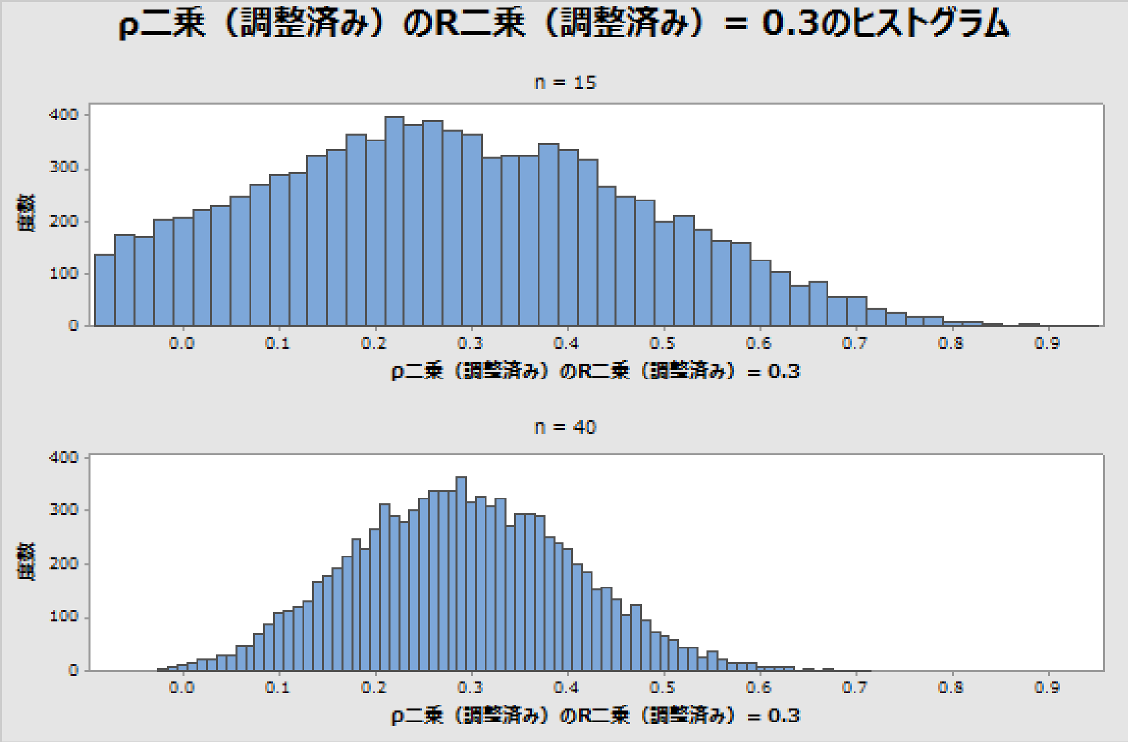

図 2　n = 15 と n = 40 で ρ 二乗（調整済み）= 0.30 の場合にシミュレートした R 二乗（調整済み）値

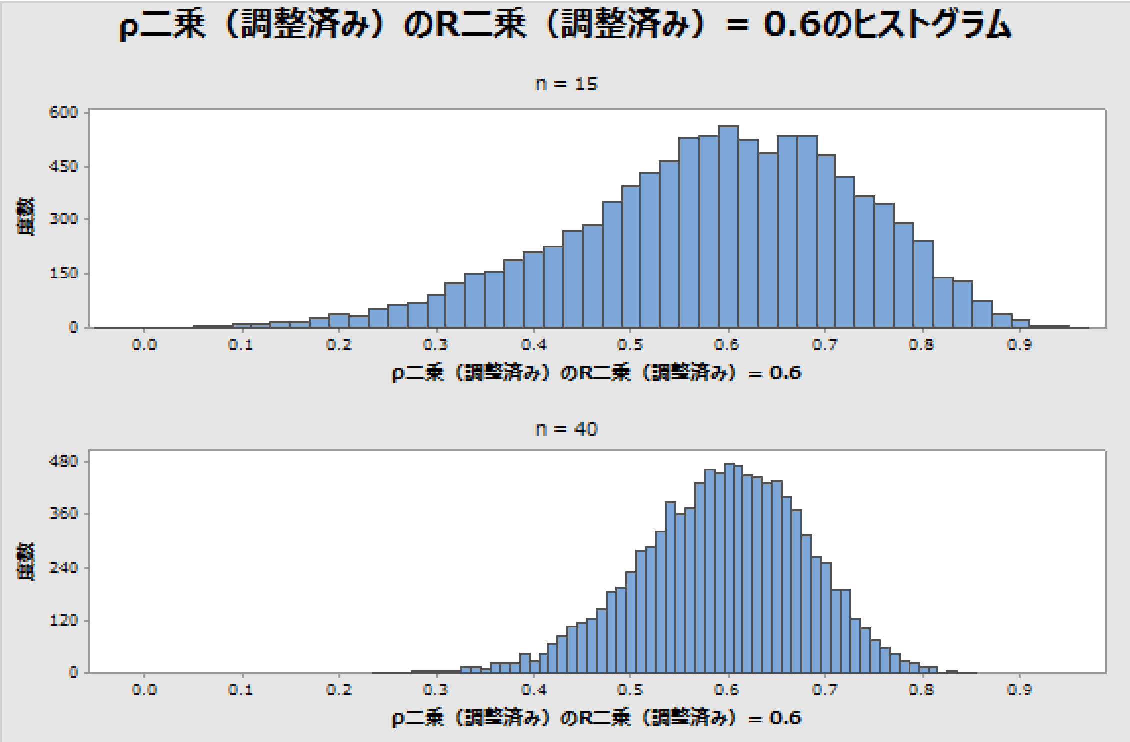

図 3　n = 15 と n = 40 で ρ 二乗（調整済み）= 0.60 の場合にシミュレートした R 二乗（調整済み）値

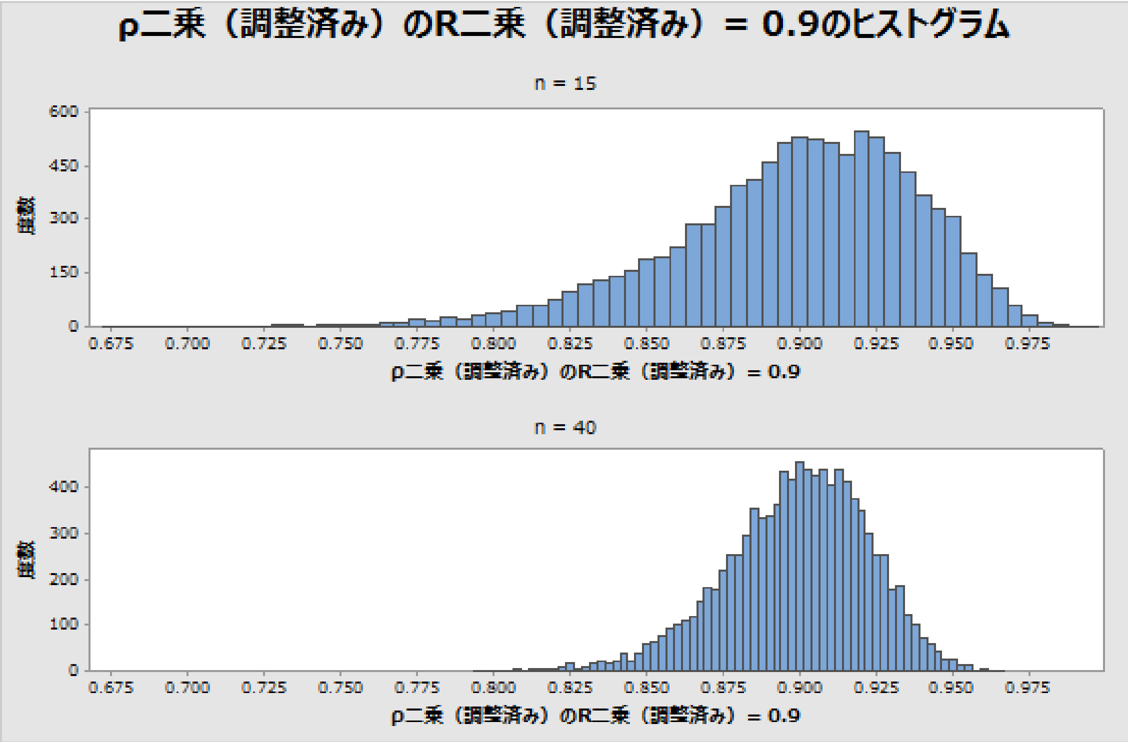

図 4　n = 15 と n = 40 で ρ 二乗（調整済み）= 0.90 の場合にシミュレートした R 二乗（調整済み）値

全体的に、シミュレーションでは実際の関係の強度（ρ 二乗（調整済み））とデータで観測された関係の強度（R 二乗（調整済み））の差はかなり大きくなる可能性があることが示されています。サンプルサイズを 15 から 40 に増やすことで、差が大きくなる可能性は大幅に低くなります。差分の絶対値|R 二乗（調整済み）- ρ 二乗（調整済み）|が 0.20 より大きい場合の確率が 10%以下となる観測値数 n の最小値を同定することにより、40 が適切なしきい値であると決定しました。これは、対象となるすべてのモデルにおける真の ρ 二乗（調整済み）値には関係ありません。線形モデルの場合、最も難しいケースは ρ 二乗（調整済み）= 0.31 で、n = 36 が必要になります。2 次モデルの場合、最も難しいケースは ρ 二乗（調整済み）= 0.30 で、n = 38 が必要になります。40 の観測値数では、その値や、線形モデルと 2 次モデルのどちらを使用するかに関わらず、90%の信頼度で R 二乗（調整済み）の観測値が ρ 二乗（調整済み）の 0.20 以内になります。

付録 C： 正規性

アシスタントで使用される回帰モデルは、すべて次の形式です。

$$Y = f(X) + \varepsilon$$

ランダム項ε平均ゼロで共通の分散σ^2による独立した同じような分布に従う正規確率変数であると一般的に仮定されます。βパラメータの最小二乗推定は、εが正規分布に従うと仮定しない場合でも、最良の線形不偏推定です。正規性の仮定は、f(X)の仮説検定と同様にこれらの推定に確率を追加する場合にのみ重要になります。

正規性の仮定に基づく回帰分析の結果を信頼できるのに十分な n の大きさを判断するため、さまざまな非正規誤差分布で仮説検定のタイプ I 過誤率を調べるシミュレーションを実行しました。

次の表 4 は、線形モデルと 2 次モデルのさまざまな ε の分布において、全体の F 検定が$\alpha = 0.05$で有意であった 10,000 回のシミュレーションの比率を示しています。これらのシミュレーションでは、X と Y の間に関係は無いという帰無仮説は真でした。X 値は区間に等間隔に配置されています。すべての検定で使用したサンプルサイズは n = 15 です。

表 4　非正規分布での n = 15 を使用した線形モデルおよび 2 次モデルに対する全体の F 検定のタイプ I 過誤率

分布	線形の有意比率	2 次の有意比率
正規	0.04770	0.05060
t(3)	0.04670	0.05150
t(5)	0.04980	0.04540
ラプラス	0.04800	0.04720
一様	0.05140	0.04450
Beta(3, 3)	0.05100	0.05090
指数	0.04380	0.04880
Chi(3)	0.04860	0.05210
Chi(5)	0.04900	0.05260
Chi(10)	0.04970	0.05000
Beta(8, 1)	0.04780	0.04710

次に、最適なモデルの選択に使用される最高次の項の検定を調べました。各シミュレーションで、2 次項が有意であるかどうかを確認します。2 次項が有意ではなかった場合、線形が有意であるかどうかを確認します。これらのシミュレーションでは、帰無仮説は正しく、目標$\alpha = 0.05$、n=15 でした。

表 5　非正規分布での n ＝ 15 を使用した線形モデルまたは 2 次モデル内の最高次の項に対する検定のタイプ I 過誤率

分布	2 次	線形
正規	0.05050	0.04630
t(3)	0.05120	0.04300
t(5)	0.04710	0.04820
ラプラス	0.04770	0.04660
一様	0.04670	0.04900
Beta(3, 3)	0.05000	0.04860
指数	0.04600	0.03800
Chi(3)	0.05110	0.04290
Chi(5)	0.05290	0.04490
Chi(10)	0.04970	0.04610
Beta(8, 1)	0.04770	0.04380

シミュレーション結果は、両方の全体の F 検定とモデル内の最高次の項の検定で、統計的に有意な結果を得る確率はどの誤差分布でも大きな差はないことを示しています。タイプ I 過誤率はすべて 0.038～0.0529 の範囲に入っています。