

研究ノート

ビジネスにおける統計分析の分類

東 洵 則 之

目 次

1. はじめに
2. 統計分析の分類と統計分析情報のウェイト

1. は じ め に

統計分析は、データという材料を加工し、意思決定に役立つ情報を生み出すための道具である。情報技術の発達とともに企業活動において情報の果たす役割は大きくなりつつあり、企業における統計分析への関心も高まりつつある。

一般に、統計分析というと、主として、その技術的側面に関心が集中する傾向がある。統計学の教科書においても、度数分布や平均、標準偏差、相関分析、回帰分析、推定、検定など数理技術的な記述が専らである。しかし、ビジネス実践で統計分析を道具として使いこなそうとすると、これでは不十分である。手法の知識だけでなく、データの知識、統計分析のマネジメントの知識が必要である(東洵, 1993)。本稿では、この統計分析のマネジメントに関連して、ビジネスにおける統計分析を分類してみる。

2. 統計分析の分類と統計分析情報のウェイト

統計分析は、関連データを収集し、意味ある事実を明らかにしたり、容易に把握できるようにアレンジしたり、必要なら他の統計的事実や既存の基準と比較できるようにデータを再アレンジすることである。何も整理されていない生データを前に解釈者はデータの意味するところを正確に理解することができず、戸惑うばかりである。データを整理加工することによってデータのもつ情報を見出すことができる。しかし、統計分析と言っても、自然科学分野の場合とビジネス分野の場合では、その位置付けはかなり異なる。また、ビジネス分野に限っても、使われる場面で統計分析の意思決定における位置付けはかなり異なる。

(1) 自然科学における統計分析とビジネスにおける統計分析の違い

統計分析は、そもそも自然科学分野における実験・観測データの分析のために開発・発展させられてきた。しかし、ビジネス分野における統計分析の実施環境は、統計学が前提としてきた自然科学のそれとはかなり異なる。

① データの質

ビジネス分野のデータの質は自然科学の観測データ等と比べ劣っているという点である。経営データのソースには様々なものがある。日常業務の遂行と共に自然発生的に蓄積される業務データが主である。業務データの質が悪くなる原因について、いくつか例を挙げよう。

品切れ …品切れを起こしている売り場でいくらPOSデータを分析しても正しい売れ筋はつかめない。

売上操作 …月末になるとノルマ達成のため商品を小売に無理やり押し込んだり、逆に、今月の売上を来月に繰り延べることもある。これでは真の需要量は得られない。

販売状況 …通常価格での売上と特価での売上を区別して記録できていな

ければ、価格と売上との関係は分析できないが、そうなっていない場合がある。

定義の不一致…「売上」は「注文を受けた時点か」「伝票を切った時点か」「実際納品した時点か」が統一されていない場合がある。

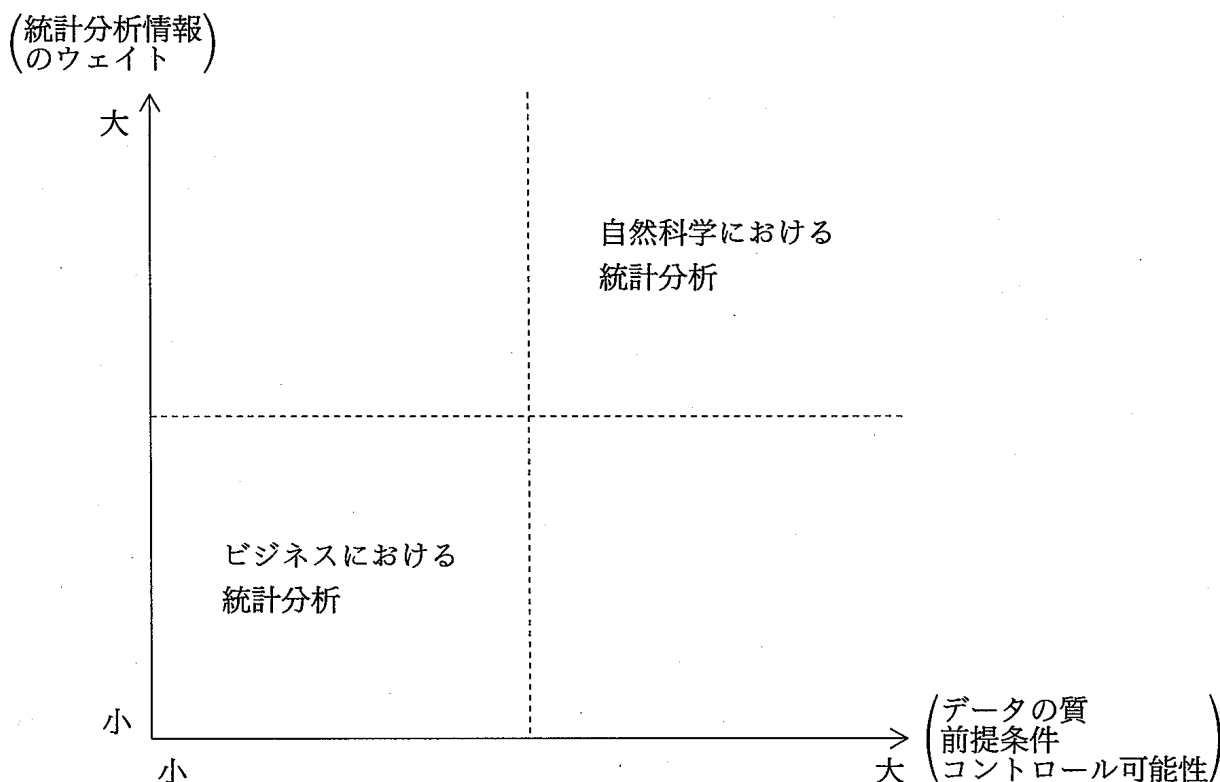
人為的ミス …例えば、注文処理を発注システムで実行せずに電話やファクシミリで済ませてしまいコンピュータ上に未登録となっていることも珍しくない。

また、人間特有のケアレスミスである聞き間違い、記入間違いなども珍しいことではない。以上のような原因のため、一般に業務データ、さらには経営データは自然科学分野のデータより質が悪い傾向にある。

統計分析の際満たされるべき前提条件であるデータ数や変数の独立性などが、手法が要求するレベルまで満たされる程度もビジネス分野では相対的に小さい。自然科学分野のデータは、例えば実験室など外部環境から遮蔽された中で計画的に収集される場合が多いのに対して、各種業務データは、日常業務の遂行と共に自然発生的に蓄積される履歴データであることに起因する。

② 非統計分析情報のウェイト

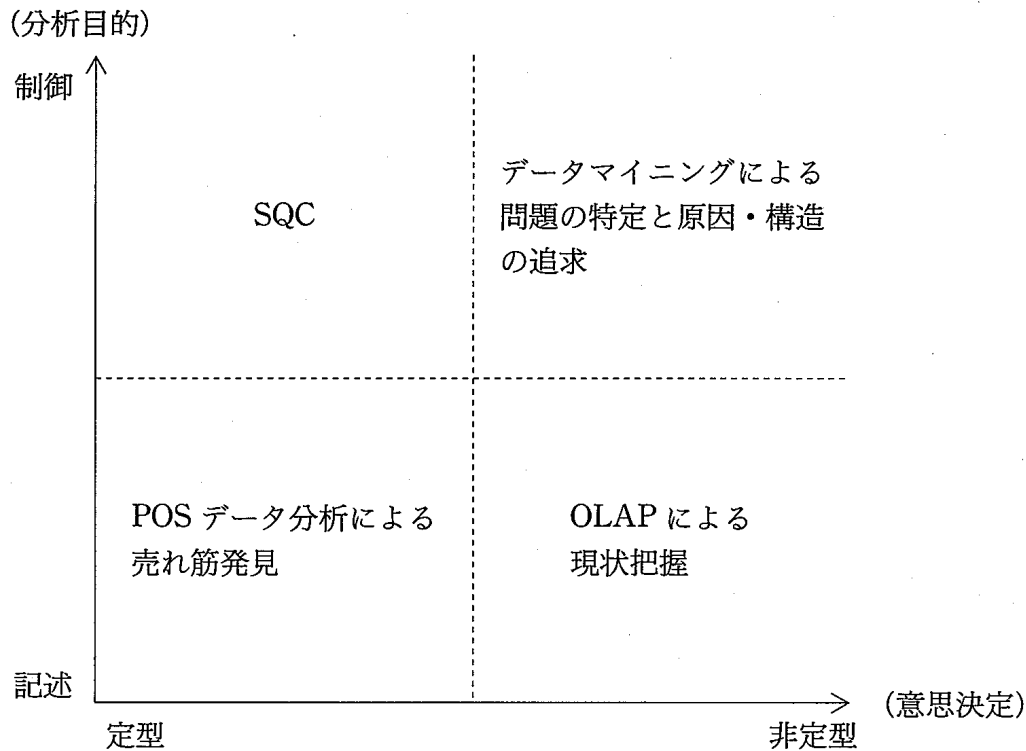
統計分析情報は、ある事象の一側面を測定したものであり、定性情報・事実情報、経験、直観（これらを非統計分析情報と呼ぶ）と脳裏で統合され意思決定につなげられるべきものである。しかし、このウェイトが自然科学分野とビジネス分野では異なる。一般に、自然科学分野では、統計分析結果が非統計分析情報より重視されるのに対して、ビジネス分野では必ずしもそうはならない。意思決定において非統計分析情報が併用される余地は、ビジネス分野では相対的に大きい。以上の①、②をまとめ、自然科学とビジネスにおける統計分析の違いを図示すると次のようになる。



図表1. 自然科学の統計分析とビジネスの統計分析

(2) ビジネス分野における統計分析の分類と統計分析情報のウェイト

ビジネス分野における統計分析に限ってみても、意思決定に占める統計分析のウェイト、逆に言うと「定性情報・事実情報、経験、直観」のウェイトは様々である。この点を詳しく見るため、ビジネスにおける統計分析の利用場面を、「意思決定が定型的か非定型的か」、「分析目的が記述か制御か」の2つの軸で4つのセルに分類してみよう。何度も繰り返される意思決定を定型的意思決定と言い、同じような手続きで処理ができる。過去に同じようなケースがなくその都度手続きを考えなければならない意思決定を非定型的意思決定と言う。もう一方の軸は、統計分析の目的である。統計分析の目的は、記述 (reporting) → 予測 (prediction) → 説明 (explanation) → 制御 (control) の4つの段階があり、先の段階ほどより詳細な情報が得られる。記述段階は「事象を記述する」レベル、予測段階は、「予測ができるように変数間の統計的関係を求める」レベル、説明段階は「変数間の因果関係を明確にする」レベル、制御段階では「操



図表 2. ビジネスにおける統計分析の分類

作変数を操作することによって目的変数を意図した値に変化させる」レベルである。

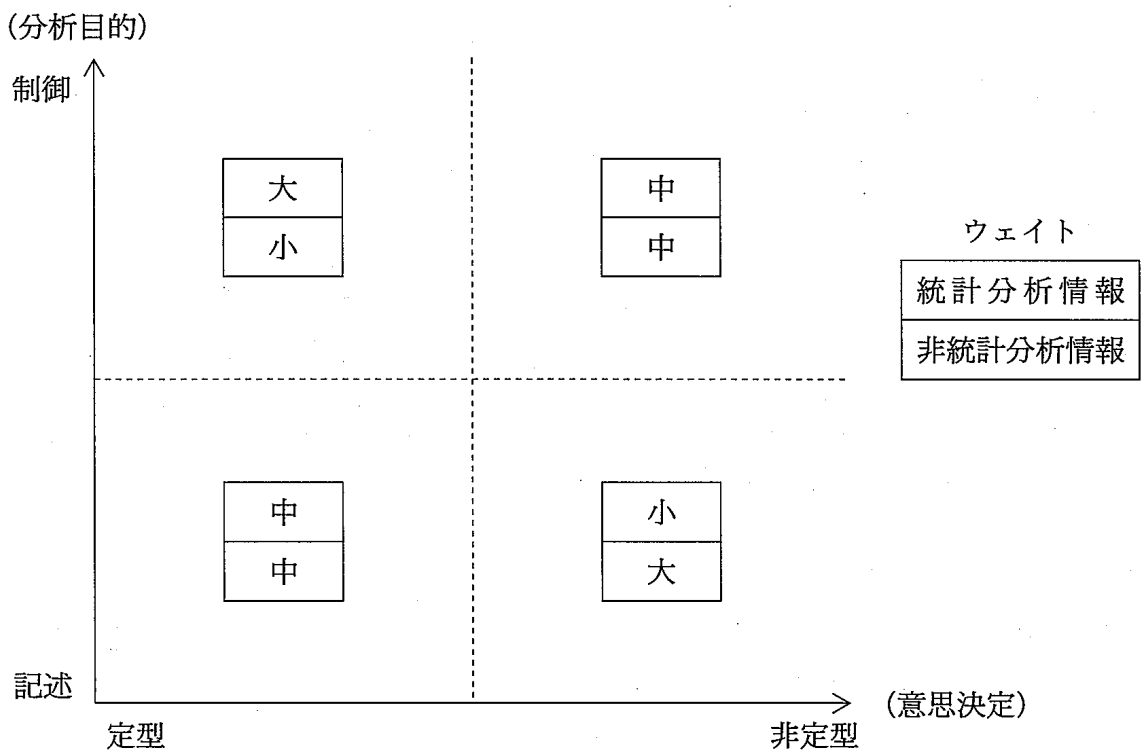
定型的意思決定・記述的手法としては、「POS データ分析による売れ筋発見」が好例としてあげられる。POS データ分析により、商品の売れ行き状況が記述される。分析者はそれを見て品揃えを検討する。売れ筋や品揃えの検討の際は、POS による統計分析情報以外に、定性情報・事実情報、経験、直観なども併用されるため、意思決定に占める統計分析の相対的なウェイトを大・中・小の3段階で表現すると「中」である。

定型的意思決定・制御的手法としては、「SQC(統計的品質管理)」が好例となる。品質や工程の進捗状況を示す数値が示され、規定値を外れていれば警告を発する。分析者はこれにより異常を察知し、操作変数を操作し制御を行う。この場合、統計分析情報が極めて重要な情報となるので、意思決定に占める統計分析の相対的ウェイトは「大」である。逆に、非統計分析情報のウェイトは相

対的に小さい。

非定型的意味決定・記述的手法としては、「OLAP (online analytical processing) による現状把握」が好例である。OLAP はデータウェアハウスの活用形態のひとつである。例えば、必要に応じて分析者が関連のありそうな変数のクロス分析結果やグラフを見ながら、変数間の関連を推測し、試行錯誤を繰り返し、アイデアや企画を練るというものである。OLAP による統計分析は仮説を確かめたり、探ったりする手段となる。OLAP では分析者による試行錯誤が不可欠となるが、試行錯誤の指針として、非統計分析情報が非常に重要となる。このため意味決定に占める統計分析の相対的ウェイトは「小」であり、非統計分析情報の相対的ウェイトは「大」である。

非定型的意味決定・制御的手法では「データマイニングによる問題の特定と原因・構造の追及」が好例である。データマイニングはOLAPと同じくデータウェアハウスの活用形態のひとつである。そこでは、例えば、非常に多数の変



図表3. 意思決定における統計分析情報と非統計分析情報のウェイト

数の因果関係をコンピュータが自動的に調べ、分析者にその結果を提示する。分析者は提示された結果を、非統計分析情報と擦り合せて判断を形成し意思決定を行っていく。意思決定に占める統計分析の相対的ウェイトは「中」である。以上をまとめると、4つのセルごとのウェイトは図表3のようになっている。

ビジネスにおいて統計分析を利用する際、ウェイトの大小を把握しておくことは重要である。なぜなら、目指すべき統計分析の規模、精度、費用、時間等の決定に大きな影響を及ぼす、言いかえると統計分析のスタンスを決定するからである。

参 考 文 献

東測則之、「経営統計学の方向性に関する序論」、松山大学論集，第5巻第3号，1993.7，pp. 449-469.

(本稿は、平成9年度松山大学特別研究助成の成果の一部である。)