

データサイエンスとは

最近、データサイエンスという言葉をよく見聞きしますが、ICT分野の先端的な企業が扱っている印象があり、中小企業さんには縁遠い感じがするのではないのでしょうか。

データサイエンスとは以下の技術とされています。

- ① 数理統計学を基礎とする高度なデータ分析技術
- ② 経営やマーケティングの知見と合わせてビジネスに生かすデータ処理技術

企業の開発、生産、販売現場には多くのデータが蓄積しているはずで、日常業務で収集したデータには大きな付加価値があり、その多くは汎用の表計算ソフト Excel で高度な分析ができ、その結果はビジネスに大いに活用できます。

具体的には、Excel により、重回帰分析、クラスター分析、ロジスティック回帰分析、自己相関関数分析、 χ^2 検定、最適化分析などが実行可能です。

自己相関関数を活用した市販商品の需要予測

以下に自己相関関数を市販商品の需要予測に活用した事例を示します。

市販商品の販売数に何らかの周期性があり、しかも月ごとに2倍以上の差がある場合、あらかじめ精度の高い予測ができれば人員計画や部材発注を効率的にできます。

下図はある企業さんの5年間の出荷台数をグラフ化したものです。

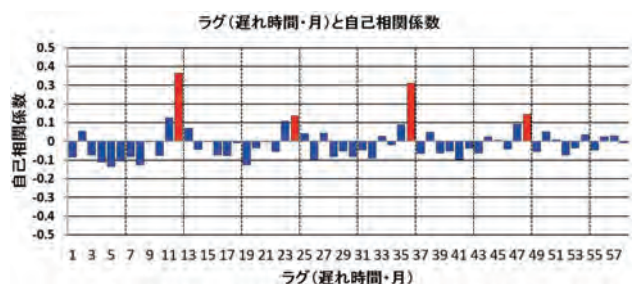


月ごとで見ると、300台/月から850台/月まで、大きく変動していることがわかります。工数を1分単位で管理し、部材の在庫を極小に抑える経営努力をしている企業さんにはこの3倍に近い月ごとの受注変動は経営の大きなリスクになっています。

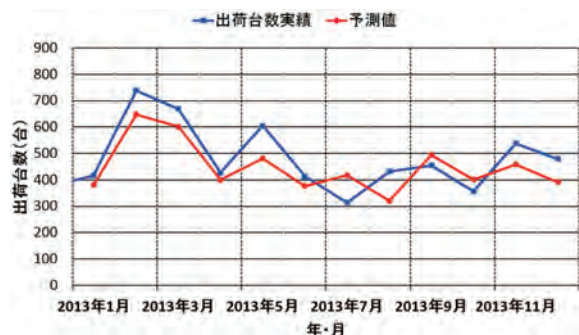
自己相関関数はある時系列データとそのデータ自身にラグ（遅れ時間）を設けたデータとの相関を分析するもので、データに含まれる周期性を検出するのに非常に有効な手法です。

Excelで演算する場合はデータをずらしながら、元データとの相関係数を演算し、横軸をラグ（遅れ時間）、縦軸を相関係数としてグラフ化します。下図に自己相関係数の演算結果を示します。

12月、24月、36月、48月のラグ（遅れ時間）に対して正の相関が顕著に出ており、結局12か月の基本周期であることがわかります。



基本周期が判明すると、それに基づき以後の予測数を計算することができます。本例では2014年の実績がわかっているため、2013年までのデータから算出した予測値との比較ができ、その結果を次図に示します。



このようにして、集積している過去データから今後の事業計画を立てるのに有効な分析データを一般的なパソコンで求めることができます。

データサイエンスは開発から販売まで、業種や専門を問わず、あらゆる場面で有効に使える点で“最強の学問”と言えるのではないのでしょうか。

各企業さんがデータサイエンスを導入され、経営をさらに効率化される活動を ATAC は全面的にサポートをさせていただきますので、ぜひご相談ください。

(坂井記)

編集後記

これまで「企業PRコラム」の欄を設けて ATAC とお付き合いいただいている企業の社長様に会社の経営方針や設備、製造している製品の特徴などを PR していただきましたが、今回、これと平行して、ATAC からみて優れた技術や製品と思えるものを紹介する「新技術・新製品紹介コラム」を始めることにしました。有効にご活用頂ければと願っています。

(白石)