

テーマ	パンの明日の売り上げを予測してみよう				
概要	Excelを用いて回帰分析を行い、パンの売り上げに関する傾向を調べ、明日のパンの売上個数を予測する。				
ねらい	身のまわりの現象や社会現象などを通して、モデル化とシミュレーションの考え方や方法を理解させ、実際の問題解決に活用できるようにする。 その中で、統計的なデータ処理を中心に問題解決を図る。				
関連する主な科目・項目	モデル化とシミュレーション	難易度	易・普・難	想定時間数	約4時間
準備するもの	表計算ソフト（Excel 等） 回帰分析用データ	作成者	和田 修		

1 学習の展開

【実習の前に】

下記の例のように実習の前に統計の基本的な考え方を学習しておくといわれる。

第1時 基本的な統計の考え方

- ・データの要約
- ・代表値（平均・中央値・最頻値）
- ・ばらつき
- ・ヒストグラム
- ・尖度と歪度

第2時 相関と回帰

- ・相関関数
- ・回帰直線
- ・散布図
- ・重回帰分析の概要
- ・数量化理論1類

第3時 Excelにおける回帰分析

回帰分析ツールの基本操作の説明とツールを扱う際の注意点。
簡単な例題演習

【実習】

1. パンの売り上げデータを用意します。
別ファイルの Excel データを参照してください。
2. データを回帰分析にかけるために、月日のデータを無視し、天候と曜日のデータを該当するものを1、該当しないものを0として図のように解析用データを作成する。
3. 作った解析用データの中から、冗長な部分を削除します。(このデータの場合は雨の日のデータと日曜日のデータ) 削除されたデータは、他のデータを見ることにより、復元可能な点に注意してください。

パンの売り上げデータをよく見よう

- unnecessary データを削除。
- 回帰分析用にデータを変換。
- 曜日と天候を0と1で符号化。

年月日	曜日	天候	売上額	
8月2日	水	晴れ	179	
8月3日	金	晴れ	108	
8月4日	土	曇り	123	
8月5日	日	曇り	28	
8月6日	月	曇り	121	
8月7日	火	曇り	144	
8月8日	水	晴れ	97	
8月9日	木	晴れ	108	
8月10日	金	曇り	102	
8月11日	土	晴れ	102	
8月12日	日	晴れ	58	
8月13日	月	曇り	121	
8月14日	火	晴れ	109	
8月15日	水	晴れ	111	
8月16日	木	晴れ	休日	
8月17日	金	曇り	115	
8月18日	土	晴れ	110	
8月19日	日	晴れ	28	
8月20日	月	曇り	80	
8月21日	火	曇り	休日	
8月22日	水	晴れ	休日	
8月23日	木	曇り	休日	
8月24日	金	曇り	107	
8月25日	土	曇り	106	
8月26日	日	曇り	59	
8月27日	月	曇り	114	
8月28日	火	曇り	103	
8月29日	水	曇り	91	
8月30日	木	晴れ	91	
8月31日	金	晴れ	92	

解析用データ

売上	天候				曜日					
	雨	曇	晴	月	火	水	木	金	土	日
179	0	0	1	0	0	0	1	0	0	0
108	0	0	1	0	0	0	0	1	0	0
123	0	1	0	0	0	0	0	0	1	0
28	0	1	0	0	0	0	0	0	0	1
121	0	1	0	1	0	0	0	0	0	0
144	0	1	0	0	1	0	0	0	0	0
97	0	0	1	0	0	0	1	0	0	0
108	0	0	1	0	0	0	1	0	0	0
102	0	0	1	0	0	0	0	0	1	0
102	0	0	1	0	0	0	0	0	1	0
58	0	0	1	0	0	0	0	0	0	1
121	0	0	1	1	0	0	0	0	0	0
109	0	0	1	0	1	0	0	0	0	0
111	0	0	1	0	0	1	0	0	0	0
115	0	1	0	0	0	0	0	1	0	0
110	0	0	1	0	0	0	0	0	1	0
28	0	0	1	0	0	0	0	0	0	1
80	1	0	0	1	0	0	0	0	0	0
107	0	1	0	0	0	0	0	1	0	0
106	0	1	0	0	0	0	0	0	1	0
59	1	0	0	0	0	0	0	0	0	1
114	1	0	0	1	0	0	0	0	0	0
103	1	0	0	0	1	0	0	0	0	0
93	0	1	0	0	0	1	0	0	0	0
91	0	0	1	0	0	0	1	0	0	0
92	0	0	1	0	0	0	0	1	0	0


さらにデータを最適化する

- 回帰分析用に冗長なデータを削除します。
- 雨の日のデータと日曜日のデータを削除します。
- 削除しても情報は保存されています。

売上	雨	曇	晴	月	火	水	木	金	土	日
179	0	1	0	0	0	0	1	0	0	0
108	0	1	0	0	0	0	0	1	0	0
123	1	0	0	0	0	0	0	0	1	0
28	1	0	0	0	0	0	0	0	0	0
121	1	0	1	0	0	0	0	0	0	0
144	1	0	0	1	0	0	0	0	0	0
97	0	1	0	0	1	0	0	0	0	0
108	0	1	0	0	0	0	0	0	0	0
102	0	1	0	0	0	0	0	0	1	0
102	0	1	0	0	0	0	0	0	1	0
58	0	1	0	0	0	0	0	0	0	0
121	0	1	1	0	0	0	0	0	0	0
109	0	1	0	1	0	0	0	0	0	0
111	0	1	0	0	1	0	0	0	0	0
115	1	0	0	0	0	0	0	0	1	0
110	0	1	0	0	0	0	0	0	0	1
28	0	1	0	0	0	0	0	0	0	0
80	0	0	1	0	0	0	0	0	0	0
107	1	0	0	0	0	0	0	0	0	1
106	1	0	0	0	0	0	0	0	0	1
59	0	0	0	0	0	0	0	0	0	0
114	0	0	0	1	0	0	0	0	0	0
103	1	0	0	0	1	0	0	0	0	0
93	1	0	0	0	1	0	0	0	0	0
91	0	1	0	0	0	1	0	0	0	0
92	0	1	0	0	0	0	0	1	0	0

4. Excel のメニューより、「分析ツール」を選択し起動させます。もし、「ツール」メニューに「分析ツール」が存在しない場合は、Excel のインストールCD-ROMを使って追加のインストールをしてください。
- Excel 以外の表計算ソフトにもこの「分析ツール」に相当したものがありますので、それを起動させてください。

Excelの回帰分析ツールを使う



- 「ツール」メニューより、「分析ツール」を選択します。
- Yの範囲にパンの売り上げ数のデータを1列いれます。売り上げのラベルも一緒に選択します。
- Xの範囲には天候と曜日のデータをすべて指定します。ラベルも範囲に入れます。

- 5 実行結果の画像を右に表示しますので、確認してください。

回帰分析ツールの実行結果

概要			
回帰統計			
重相関係数	0.851173		
重決定 R2	0.724495		
補正 R2	0.594846		
標準誤差	20.71642		
観測数	26		
分散分析表			
	自由度	変動	分散
回帰	8	19185.99	2398.249
残差	17	7295.891	429.17
合計	25	26481.88	
係数			
	係数	標準誤差	t
切片	30.8	14.418	2.136
曇	15.3	14.384	1.065
晴	9.3	14.020	0.660
月	72.1	15.062	4.784
火	79.7	15.905	5.010
水	58.3	16.172	3.603
木	85.9	16.548	5.194
金	61.9	14.296	4.332
土	67.2	15.082	4.454

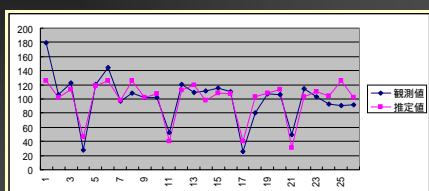
6. 「分析ツール」の実行結果をもとにして、パンの売り上げの予測式を作成します。

7. この式をもとにして、グラフを作り、売り上げに大きく関わっている要因をチェックします。また、予測式から求められるデータと実際のデータを比較して予測式が適切かどうかを検討します。

予測式を完成させよう

売上個数 =	切片	+	雨のとき	+	火曜日	+	水曜日	+	木曜日	+	金曜日	+	土曜日	+	日曜日
	30.8		0.0		79.7		58.3		85.9		61.9		67.2		0.0

モデルを評価しよう



- 青線が実際の売り上げデータです。
- 紫線が計算による推定値です。
- 大体あっていますが、気になる点もあります。

8. 最後に予測式から、明日のパンの売り上げ個数を計算して求めます。

明日のパンの売り上げは何個？

明日の天気	明日の曜日	売り上げ予測
雨	土曜日	個

売り上げ個数 = $30.8 + 0 + 67.2$
計算結果は、98個でした。

2 発展的学習

1. 年鑑・インターネット等を使って回帰分析をしたいデータを集め、そのデータを各自で分析・予測する。

3 その他

1. 重回帰分析など高度な内容には、概念だけにおさえ、数式は扱わない。回帰分析ツール (Excel) についても、その裏づけになる理論には深入りせず、ブラックボックス的に扱う。
2. 回帰分析は計算に時間がかかるために Excel を使うことに注意をうながす。
3. この事例は、下記の書籍を参照させていただいた。

「データマイニング事例集」 上田太一郎 著 共立出版 2800 円 (税別)