

情報エントロピーによる商品購買の偏重の検討

ガイラルディア+
(大東文化大学)

- 本研究ではファッションECサイトの顧客データから様々な分析軸の相違に対して、**情報エントロピー(購入の偏り)**に基づいたデータ分析を行った。
- エントロピーの数値が高ければ購入の偏りが少なく、低ければ購入の偏りが大きい。偏りが小さければ満遍なく商品を購入しているということであり、一般には良い購買状況であるといえる。
- 一般に、小売業においては、季節や曜日、顧客層、商品ジャンル等はできる限り平滑化されている方が、外部要因の変動によらず売り上げが安定し、望ましいと考えられる。
- 今回はデバイス、年代・性別、季節、曜日ごとに対するエントロピーと商品カテゴリに対するエントロピーを求めそれぞれの偏りを求めた。具体的には購入の偏りが大きい項目の平滑化(偏りを無くす)を目指した販売促進方法を考えていく。

- 今回の分析ではアプローチ方法として、商品の購買の乱雑さを測るために情報エントロピーを用いた。
- 「乱雑である」というのは、直観的には「均等にちらばっている」という意味だが、いずれにしても感覚的なものである。乱雑さは数学的には情報エントロピーの数値の大小で判断することができる。算出された数値が高いと乱雑であり、低いと偏りがある。情報エントロピー E は以下のように定義される。

$$E = -\sum p_i \log p_i$$

- ここで p_i はカテゴリ i (例: ジャケット/アウターの購買比率、トップスの購買比率など) の生起確率を表す。

- 以下の4つの観点から分析を行った。
 - 季節(春夏秋冬)による購買の変化
 - 年齢・性別による購買の変化
 - 顧客の購入時の使用デバイス(スマートフォンやPC等)による購買の変化
 - 購買曜日による購買の変化

(1)季節別にみた商品カテゴリに対するエントロピー

各商品カテゴリに対する購入累計金額に対して、季節別のエントロピーをセール品、非セール品（セール品以外）、全体のそれぞれについて算出したもの

商品カテゴリに対するエントロピー			
	セール品のみ	非セール品	全体
春	0.88877	0.95380	0.94246
夏	0.88426	0.97705	0.93586
秋	0.87467	0.96109	0.94110
冬	0.88849	0.99659	0.94297

春:3-5,夏:6-8,秋:9-11,冬:12-2月

全体:夏のみ偏っていて、その他は平滑化されている。

非セール品:春は偏っているが、冬は平滑化されている。

→春はセール(冬物セールなど)によってうまく平滑化できている。

→夏はセールを含めてもやや購買が偏る。

購買の偏りは、一時的なものであるにせよ顧客を離反させ、再び購買を促すには余分なスイッチングコストを必要とすることも多い。このため、購買が偏る傾向のある季節(夏)には、セール等による販売促進策の余地があると考えられる。

(1)商品カテゴリ別にみた季節に対するエントロピー

季節に対するエントロピー			
	セール品のみ	非セール品	全体
トップス	0.55078	0.59033	0.59917
パンツ	0.54764	0.59437	0.59821
シューズ	0.55524	0.59806	0.59911
ジャケット/アウター	0.44512	0.52441	0.51160
ワンピース	0.54645	0.59302	0.59769
バッグ	0.54785	0.59989	0.59790
アクセサリー	0.53436	0.60129	0.59694
アンダーウェア	0.52782	0.59987	0.58935
スカート	0.55642	0.58216	0.59723
ファッション雑貨	0.43415	0.57261	0.54954
レグウェア	0.53848	0.59685	0.59451
帽子	0.53386	0.59268	0.60028
財布/小物	0.53249	0.59093	0.58725
ヘアアクセサリー	0.52509	0.59614	0.59769
雑貨/ホビー/スポーツ	0.55035	0.60068	0.60005
インテリア	0.54690	0.59738	0.59925

ジャケット/アウター、ファッション雑貨は季節により購買傾向が異なる。

ただし、4分割の場合の最大エントロピー（各季節に全く同額売れた場合）は0.60206であるので全体的にみるとかなり均等に季節ごとに売れているといえる。

(2)性別・年齢層・使用デバイス別にみた商品カテゴリに対するエントロピー

男女・年代別(10代～30代前半・30代後半以上)のほか、使用デバイス(PCのみ使用、スマートフォンのみ使用、PCスマホ併用)でも大きな差が生じているのではないか？

商品カテゴリに対するエントロピー

	PC	スマートフォン	PCスマホ併用
男性(10代～30代前半)	1.30705	1.32085	1.31449
男性(30代後半以上)	1.31234	1.30266	1.31212
女性(10代～30代前半)	1.37174	1.37456	1.37871
女性(30代後半以上)	1.37597	1.38287	1.38369

- 女性は男性よりエントロピーが高い(購買が平滑化されている)。→ワンピースやアクセサリ、スカートなど、女性はファッションの幅が広いいため？
- 特に、30代後半以上のスマートフォンから購入しているユーザーのエントロピーが高い。→子供や家族の物も買っているため？

男性は一般に購買が偏る傾向があり、平滑化の観点からは、男性向けに色々な商品の購買を促すようなことが必要であるといえる。

(3) 購入時の使用デバイスに関するエントロピー

商品カテゴリに対するエントロピー			
	セール品のみ	非セール品	全体
PC	1.20181	1.20146	1.20181
スマートフォン	1.20359	1.20340	1.20349
PCスマートフォン併用	1.20354	1.20315	1.20344

デバイスに対するエントロピー			
	セール品のみ	非セール品	全体
トップス	0.47670	0.47637	0.47656
パンツ	0.47667	0.47666	0.47665
シューズ	0.47611	0.47671	0.47653
ジャケット/アウター	0.47691	0.47672	0.47690
ワンピース	0.47626	0.47624	0.47625
バッグ	0.47646	0.47656	0.47655
アクセサリ	0.47531	0.47290	0.47384
アンダーウェア	0.47460	0.47656	0.47680
スカート	0.47481	0.47447	0.47473
ファッション雑貨	0.47632	0.47712	0.47701
レッグウェア	0.47628	0.47614	0.47630
帽子	0.47656	0.47546	0.47593
財布/小物	0.47678	0.47405	0.47481
ヘアアクセサリ	0.47114	0.47073	0.47134
雑貨/ホビー/スポーツ	0.47555	0.47683	0.47694
インテリア	0.47609	0.47650	0.47651

当初はPCやスマートフォンではインタフェースが異なり、また高額な商品はスマートフォンでは買うのに抵抗があるために、購買パターンには大きな差異があるものと予想していたが、実際にはほとんど変わらない結果であった。

実際、現在のスマートフォンでは、一般に画像の解像度やページのレイアウトを含め、ほぼPCと変わらない、もしくはPC以上に使いやすいインタフェースとなっており、ネットショッピングにおける購買状況の差異にはもはや結びつかなくなっているといえる。

(4) 曜日別にみた商品カテゴリに対するエントロピー

商品カテゴリに対するエントロピー			
	セール品のみ	非セール品	全体
日	0.896175	0.980382	0.952063
月	0.898879	0.981231	0.954160
火	0.901824	0.981240	0.955071
水	0.899428	0.980781	0.948602
木	0.901770	0.983864	0.953911
金	0.897429	0.979441	0.946426
土	0.892728	0.979518	0.948353

- 金曜、土曜に偏りがある
→ 週末であることから、週末に特有の購買パターンが存在するのではないか。
- 対して火曜日は全てのパターンにおいて平滑化がなされており、比較的安定して購買される日であるといえる。

(4) 購買されたアイテムの曜日（日曜～土曜）に対するエントロピー

購買されたアイテムの曜日（日曜～土曜）に対する

曜日に対するエントロピー			
	セール品のみ	非セール品	全体
トップス	0.844784	0.842447	0.844491
パンツ	0.844729	0.841994	0.844436
シューズ	0.844916	0.842242	0.844777
ジャケット/アウター	0.844929	0.839937	0.844016
ワンピース	0.844712	0.841366	0.844228
バッグ	0.844826	0.842885	0.844777
アクセサリ	0.844653	0.841754	0.844731
アンダーウェア	0.843929	0.841719	0.844192
スカート	0.844235	0.842149	0.844368
ファッション雑貨	0.844986	0.842016	0.844787
レッグウェア	0.844499	0.842462	0.844658
帽子	0.844783	0.842804	0.84479
財布/小物	0.843743	0.843105	0.844281
ヘアアクセサリ	0.844425	0.839920	0.844208
雑貨/ホビー/スポーツ	0.842814	0.841334	0.843393
インテリア	0.844507	0.839898	0.844657

当初の予想:曜日別の購入パターンには違いがあるのではないか?

セール品と全体のエントロピーは大きな差異がなく、曜日別の偏りは認められない。

非セール品については、ジャケット/アウターとヘアアクセサリ、インテリアで曜日別の偏りが存在する。

ただし、セールを含めれば偏りが無くなることから、セール以外では余り購買されることが少ないアイテムであると推察される。

まとめ(1)全体の総括

- 春は、非セール品の偏りが大きい。夏は全体において偏りが大きい。
- ジャケット/アウター、ファッション雑貨は季節毎の偏りが大きい。
- 男性の偏りは大きく、女性は平滑化されている(特に30代女性)
- 購入デバイス別では大きな偏りが無い。
- 曜日別では、金曜日、土曜日の偏りが大きく、火曜日は偏りが無い。
- 偏りが無い方が小売業の販売形態としては望ましいが、一部の切り口では偏りが見られた。
- 今後、セール等によりさらなる平滑化が求められる。

まとめ(2) 結果から戦略を練る

- 例えば、**季節**という要素から見るならば、セールをすることで平滑化されているのが春、最終的に偏るのが夏ということになる。
- 夏に向けたセールやキャンペーンなどが必要と考えられるが、具体的なセールやキャンペーンの内容・方法については、**季節性**を考慮した検討が必要。
- 要素の性質上**偏りもある程度許容し**、戦略を練らなくてはならない。

- ◆ これに対して、**年齢・性別、曜日**といった要素に関しては、比較的戦略は練りやすい。
- ◆ 男性に対する平滑化戦略は、継続的・包括的な購買層の獲得に結び付き、大きな効果を生むのではないかと予測できる。
- ◆ 偏りのある**金土**の二日間に購買パターンをフォローするキャンペーンを実施する。
- ◆ **給料日直後や年金支給日**など狙って購買の平滑化(いつもと異なるジャンル商品を推薦)を促す。
- ◆ 全体的には、偏りを無くしつつ、**各商品カテゴリの特色**を生かして、カテゴリ毎の戦略を立てる方が有効な戦略や結果に結びつくことが出来るのではないかと考えられる。

まとめ(3)終わりに

- 我々の分析手法は購買の偏りを無くし、**安定した購買パターンを構築**するということに前提を置いている。
- 多くの小売業においては、単に数が売れるというだけではなく、安定した顧客層に継続的に購買してもらえることが将来的なビジネスの安定性にもつながる。
- 一方、一般的な小売業ではセールや商品等の企画は、多くの場合**属人的**である。
- 今回、こうした平滑化の視点から購買状況の偏りを**定量的に**分析できた点は大きな成果である。
- 今回のデータでは多くの商品カテゴリにセール等の戦略を絡めながら購買の平滑化がある程度実現されている。
- しかし、男性や曜日別に対してのアプローチが足りない点も多い。
- 衣類に関する男性の購買が平滑化していないことは、アパレル産業においてはある種常識的なことであって、さして大きな影響は出ないであろうと想定される。しかし、時代の潮流が変わっていることも事実であり、多くの人々の趣味嗜好が多様に変化する中で、これからの時代に合わせた多様かつ柔軟な戦略が求められるだろう。