

ウェブスクレイピングによる価格収集（衣料品）の取組状況について（案）

近年の消費者のインターネット通信販売（以下、「ネット販売」という。）の市場規模及び購入割合を踏まえ、順次、POSデータの利活用やウェブスクレイピング技術等を用いた価格の捕捉の検討を進めている。

すでに、テレビ及び旅行サービス（航空運賃、宿泊料、外国パック旅行費）について検討を進めているところであるが、今回、ネット販売の市場規模及び購入割合が比較的大きい衣料品※1について、試験的にウェブスクレイピングによる価格収集※2を行い、将来、ネット販売価格の採用に当たっての課題等について現時点の整理を行った。

※1 「衣類・服装雑貨等」のBtoC-EC市場規模（2017年）は16,454億円
（経済産業省、「平成29年度 電子商取引に関する市場調査」）

※2 衣料品については、市場において、POSデータの存在が確認できなかったことから、スクレイピングによる価格収集により検討を行うこととした。

1. 収集の概要

(1) 対象データ

2018年6月から12月の間の日次収集データ

(2) 対象サイト

ネット販売額の多い7サイト（様々な衣料品を販売するサイト）

(3) 対象品目

衣料品 44 品目（品目の詳細は別紙参照）

(4) 収集条件

小売物価統計調査の基本銘柄※3と可能な限り同品質の商品を収集できるよう条件を設定

※3 小売物価統計調査は、消費者が一般的に購入する代表的な商品を基本銘柄に設定し調査する。

(5) 収集価格数（概数）

品目の価格数は、2018年6月から12月分の7か月間において、一か月当たり数千～数十万件（収集価格数の詳細は別紙参照）

2. データの特徴

収集した衣料品データの特徴は、以下(a)～(c)のとおり。なお、これらの特徴は収集した衣料品全てに該当する。

(a) 類似・関連商品の混入

各サイトの検索条件の設定において、可能な限り基本銘柄に合致するよう調整しているものの、一定程度の商品絞込みしかできず、類似・関連商品のデータが多く混入する。（混入状況は別紙参照。）

例)「背広服」(基本銘柄はスーツ上下)の収集データに「ジャケット」(約46%)、「ツープANTS」(約5%)、「スリーピース」(約5%)の類似・関連商品の混入を確認(括弧内は2018年6月分の割合)。

(b) 商品識別情報の不足

- 商品を識別するための固有の商品名や商品コードが存在しない。

商品名の例) 秋冬物2ボタンスーツ

- 時間の経過とともに、商品名及び商品説明等の情報が変化する場合があります、それらの情報から商品の同一性を容易に判断することはできない。

商品名の例)【クーポン利用で500円オフ】秋冬物2ボタンスーツ

(c) 品質判断の困難さ

- 各商品情報の多くは膨大であり、素材等の品質情報を容易に抽出できない。
- 商品により品質情報の程度に差がある(品質情報がない場合もある)。
- 商品情報の表現が商品によって異なる(同一サイトで異なる場合がある)。

素材情報の例1)「素材:ウール100% 胴裏:ポリエステル/レーヨン 袖裏:キュプラ」

素材情報の例2)「毛100%」

3. 今後の対応について

今回、様々な衣料品を販売するサイトから価格等情報を取集したが、以下①及び②の課題が明らかとなった。

これらの課題を解決し、価格取集後に短期間で指数へ反映できる仕組みを構築しなければならないが、今回の価格取集方法ではその実現性は低いと考えられる。

このため、衣料品については、本検討結果、IT技術の進展及び諸外国の動向等を踏まえつつ、新たな取集方法(ウェブスクレイピング以外の方法を含む)、効率的・効果的なデータ処理の方法等について、中長期的なネット販売価格の採用を視野に入れ、引き続き検討を進めることとしたい。

【価格取集の課題】

- ① 季節やトレンドなどにより商品が入れ替わるなか、多く混在する類似・関連商品を除外し、特定の商品を抽出するデータ処理が必要であること
- ② 各商品情報は膨大で質に差があり、かつ、時間の経過とともに変化するなか、それらの情報から品質情報を抽出し、同一品質の商品と判定するデータ処理が必要であること

衣料品価格数(概数)の推移及び類似・関連商品の混在状況

NO	品目名	収集時期	2018年 価格数(概数)							計	類似・関連商品の混在等状況(例示)
			6月	7月	8月	9月	10月	11月	12月		
1	婦人用着物	1～12月	16,700	19,000	20,400	20,000	19,700	17,000	18,000	130,500	浴衣、帯・草履・バック等セット・単品、仕立て料
2	婦人用帯	1～12月	3,600	3,300	3,400	3,900	4,100	4,100	4,400	26,400	帯留め等の小物、仕立て料
3	背広服(春夏物)	3～8月	17,300	24,500	27,000					68,700	ジャケット(単品)、ベスト(単品)、ツーパンツ、スリーピース
	背広服(秋冬物)	1、2月、9～12月				29,800	31,400	31,200	33,100	125,300	
4	男子用上着	1、2月、10～12月					27,800	27,500	26,400	81,600	コート、パーカー、ベスト、ピンテージ物
5	男子用ズボン(秋冬物)	3～8月				31,600	33,900	34,600	33,400	133,400	短パン、フォーマルズボン、裾上げテープ
	男子用ズボン(夏物)	1、2月、9～12月	12,500	18,300	27,400					58,200	
	男子用ズボン(ジーンズ)	1～12月	11,300	16,500	19,700	23,200	23,900	22,600	21,500	138,300	
6	男子用コート	1月、11～12月						20,200	22,200	42,300	ジャケット、ベスト、パーカー、カーディガン
7	男子用学校制服	1～3月									コスチューム、中古、上着(単品)、下着(単品)、ネクタイ(上記は2018年5月(試験)の収集数4,094価格の状況)
8	ワンピース(春夏物)	3～8月	53,900	66,300	85,300					205,300	チュニック、カットソー、スカート、上着+スカート
	ワンピース(秋冬物)	1、2月、9～12月				115,000	93,900	111,300	128,700	448,800	
9	婦人用スーツ(春夏物)	3～8月	50,500	61,800	92,800					205,000	ジャケット(単品)、スカート(単品)、ワンピース
	婦人用スーツ(秋冬物)	1、2月、9～12月	0	0	0	115,000	86,000	84,500	108,700	394,200	
10	スカート(春夏物)	3～8月	39,100	111,600	163,600					314,200	パンツ、上着(ブラウス、プルオーバー等)+スカート
	スカート(秋冬物)	1、2月、9～12月				209,200	199,700	218,300	225,900	853,000	
11	婦人用スラックス(ジーンズ)	1、2月、9～12月	25,200	49,600	72,200	94,500	70,100	73,100	75,400	459,900	ガウチョパンツ、テーパードパンツ、スカート付き
	婦人用スラックス(秋冬物)	1～12月				77,000	71,700	82,600	85,400	316,600	
12	婦人用コート	1月、11～12月						75,800	86,300	162,100	カーディガン、パーカー、ジャケット、カーディガン
13	婦人用上着	1～3月、9～12月				84,900	65,800	76,000	94,100	320,600	ジャケット、コート、パーカー、ジャンパー
14	女子用学校制服	1～3月									中古、コスチューム、上着(単品)、下着(単品)、ネクタイ(上記は2018年5月(試験)の収集数29,414価格の状況)
15	男児用ズボン	1～12月	27,200	55,300	60,900	63,200	50,900	45,900	44,900	348,000	レギンス、スウェット、ジャージ、ハーフパンツ
16	女児用スカート	1～12月	35,900	53,100	59,100	56,900	52,100	47,100	52,200	356,100	パンツ、スカツ、スカパン、レギンス、キュロット
17	乳児服	1～12月	17,900	22,600	30,100	33,900	34,400	33,900	37,200	209,800	ロンパース、スポーツカバーオール、パーカー、ギフト用
18	ワイシャツ(長袖)	1～12月	12,200	13,200	24,000	42,000	34,800	31,000	36,800	193,700	アートシャツ、ジーンズ風シャツ、ダウンシャツ、Tシャツ
	ワイシャツ(半袖)	5～8月	7,200	8,400	9,500					25,100	
19	男子用スポーツシャツ(長袖)	1～3月、9～12月				18,800	21,400	20,300	22,600	83,000	Tシャツ、カットソー
	男子用スポーツシャツ(半袖)	4～8月	48,700	63,000	63,100					174,700	
20	男子用セーター	1～3月、9～12月				29,300	57,000	41,500	41,900	169,500	カーディガン、レディース、タートルネック、半袖
21	ブラウス(長袖)	1～3月、9～12月				175,900	124,100	121,700	132,400	553,900	ブルオーバー、キャミソール、Tシャツ
	ブラウス(半袖)	4～8月	46,100	53,800	70,700					170,500	
22	婦人用Tシャツ(長袖)	1～3月、9～12月				193,900	163,600	173,600	176,400	707,400	カットソー、ブルオーバー、チュニック
	婦人用Tシャツ(半袖)	4～8月	38,300	64,400	129,800					232,400	

NO	品目名	収集時期	2018年							計	類似・関連商品の混在等状況(例示)
			6月	7月	8月	9月	10月	11月	12月		
23	婦人用セーター(カーディガン)	1～3月、9～12月				150,700	106,700	104,900	116,600	478,800	スカート・パンツセット、カーディガン・ブルオーバーセット、スーツ
	婦人用セーター(ブルオーバー)	4～8月	41,700	47,900	52,900					142,400	
24	子供用Tシャツ(長袖)	1～3月、9～12月				92,400	85,600	72,700	79,900	330,500	カットソー、ベビー服、複数枚セット
	子供用Tシャツ(半袖)	4～8月	34,300	88,200	97,700					220,000	
25	男子用シャツ	1～12月	8,600	15,500	18,600	18,000	16,600	17,900	22,500	117,400	Tシャツ、タンクトップ、機能性(ヒート、クール等)シャツ、複数枚セット
26	男子用パンツ	1～12月	19,600	24,100	28,000	32,900	29,600	29,100	32,100	195,200	ボクサー、トランクス、ショーツ、アンダーウェア、複数セット
27	男子用パジャマ	1～5月、9～12月				21,600	30,200	28,500	29,400	109,600	ルームウェア、パンツ(単品)、ナイトガウン
28	ブラジャー	1～12月	25,900	189,800	225,500	263,200	271,200	251,700	268,500	1,495,500	ブラキャミソール、ロングブラジャー、シームレスブラジャー
29	婦人用ショーツ	1～12月	25,000	117,400	142,200	155,800	135,400	130,400	137,000	843,000	機能性(ナプキン、骨盤補正、吸水、腹巻き付き等)ショーツ、ビキニタイプ、複数セット
30	ランジェリー	1～12月	27,300	27,000	28,600	32,200	31,600	30,600	33,500	210,700	機能性(ブラ、骨盤ベルト付き補正等)キャミソール、ペチコート、ワンピース
31	子供用シャツ	1～12月	1,900	2,500	2,700	3,300	4,000	4,100	4,600	22,700	タンクトップ、長袖シャツ、女児用、複数セット
32	帽子	1～12月	19,600	22,600	21,900	21,200	20,200	15,900	16,600	137,800	ニット帽、ハット、キャスケット
33	ネクタイ	1～12月	18,800	22,600	26,900	32,200	45,000	38,100	33,300	216,600	ネクタイピン、タオルセット、蝶ネクタイ
34	男子用靴下	1～12月	23,000	26,300	33,200	39,900	36,800	35,700	39,200	233,800	フットカバー、スポーツ用ソックス、複数セット
35	婦人用ストッキング	1～12月	36,900	46,700	58,900	76,700	67,200	53,800	53,900	393,700	靴下、タイツ、複数セット
36	婦人用ソックス	1～12月	18,100	28,500	34,700	41,300	35,600	31,800	29,700	219,400	レッグウォーマー、フットカバー、複数セット
37	ベルト	1～12月	27,000	31,800	35,300	35,300	35,600	20,600	14,200	199,500	財布などの革製製品、カジュアルベルト
38	マフラー	1～12月	14,100	17,100	18,000	21,100	26,300	26,500	22,600	145,400	スヌード、スヌール、大型ストール
39	男子靴	1～12月	15,800	18,300	21,600	26,400	29,900	22,400	21,400	155,500	カジュアルシューズ、特殊形状シューズ
40	婦人靴	1～12月	31,100	35,700	60,900	102,600	117,300	93,900	93,100	534,300	ローファー、パブーシュ、パレエ用シューズ
41	運動靴	1～12月	28,200	91,700	100,600	95,600	75,000	65,400	70,500	526,800	スポーツ専用(ランニング、マウンテン等)
42	子供靴	1～12月	9,000	11,700	12,900	12,500	11,100	9,600	10,000	76,500	機能性運動靴、フォーマルシューズ
43	スリッパ	1～12月	41,500	49,800	53,100	61,000	49,800	43,600	44,200	342,800	サンダル、レッグウォーマー、パブーシュ、他の商品付属(トートバック等)
44	サンダル	1～12月	38,600	80,700	89,000	90,000	39,000	27,300	26,400	390,800	スリッパ、ナースシューズ、厚底
総計			969,500	1,700,600	2,122,100	2,843,700	2,565,800	2,548,100	2,707,000	15,447,000	