

Instagram の投稿に関する統計的分析

2019SS080 角田昌駿

指導教員：松田真一

1 はじめに

この研究は、Instagram の「いいね数」に影響を与える要因を分析し、エンゲージメントを高める方法を探求する。「いいね数」は投稿の人気や影響力の指標である。分析には、投稿の内容、時間帯、使用ハッシュタグ、投稿者のフォロワー数などが含まれる。

2 Instagram について

Instagram [2] は、デジタル時代のコミュニケーションとマーケティングに革新をもたらしている。このプラットフォームは、画像と動画の共有に特化している。インフルエンサーや芸能人は、フォロワー数や「いいね」数を通じて影響力を測定している。

3 今回対象とするアカウントについて

本研究では、Instagram におけるユーザー行動とエンゲージメントの分析を目的とし、特に影響力のある日本のインフルエンサー 5 人のアカウントを詳細に調査する。(Web[3] 参照) 対象となるのは、「Naomi Watanabe」(渡辺直美)、「横浜流星」,「ROLA」(ローラ),「i.am.kiko」(水原希子),そして「Takanori Iwata」(岩田剛典)の各アカウントである。

4 データについて

本研究では、Instagram のインフルエンサーの投稿データを分析するために、2021 年 6 月 1 日から同年 12 月末までの期間に投稿されたデータを使用する。収集されるデータは、2024 年 1 月中旬時点の情報を反映している。分析において考慮する説明変数は、1 投稿した時間帯、2 投稿した曜日、3 写真の数や動画の数、4 ハッシュタグ数、5 投稿言語、6 投稿内容、7 人物数、である。統計的手法を用いて、どの要因が「いいね数」に最も強い影響を与えるかを明らかにすることが、本研究の主目的となる。なお、1, 2, 5, 6 はカテゴリ変数であり、1, 2 は「昼」と「火」が基準である。5, 6 はアカウントによって異なるので個別に示す。

5 分析方法

本研究では重回帰分析を用いた。重回帰分析には R の lm 関数を利用した。また、変数選択には step 関数による AIC に基づいた変数減少法を用いた。(浅野 [1], 田村 [4] 参照)

6 分析結果

6.1 各アカウントの結果

渡辺直美の結果: 渡辺直美の結果は表 1 のようになる。決定係数は、0.6352 であった。

表 1 渡辺直美の結果

項目	推定値	標準誤差	p 値
(切片)	335643	46658	0.0000
X1 朝	-82459	55988	0.1538
X1 夜	-107238	69168	0.1341
X7	58621	22407	0.0151

横浜流星の結果: 横浜流星の結果は表 2 のようになる。決定係数は、0.5603 であった。X6 は「プライベート」がカテゴリの基準となる。

表 2 横浜流星の結果

項目	推定値	標準誤差	p 値
(切片)	199448	40335	0.0000
X2 日	-104356	65263	0.1199
X2 月	-24586	31971	0.4476
X2 水	-167919	43000	0.0004
X2 木	-73768	29142	0.0166
X2 金	-15425	29051	0.5992
X2 土	8011	42294	0.8510
X3	19745	8549	0.0277
X4	8292	4045	0.0489
X6 雑誌撮影	-53684	27353	0.0587
X6 宣伝	-97567	25107	0.0005
X7	89072	20927	0.0001

ローラの結果: ローラの結果は表 3 のようになる。決定係数は、0.4596 であった。

表 3 ローラの結果

項目	推定値	標準誤差	p 値
(切片)	67324	5971	0.000
X3	10689	1603	0.000

水原希子の結果: 水原希子の結果は表 4 のようになる。決定係数は、0.2068 であった。X5 は「なし」、X6 は「イベント」がカテゴリの基準となる。

岩田剛典の結果: 岩田剛典の結果は表 5 のようになる。決定係数は、0.2218 であった。X5 は「英語」、X6 は「イベント」がカテゴリの基準となる。

6.2 結果からわかること

1. モデルの適合度: 渡辺直美や横浜流星の決定係数が高いことより、最も良い適合度を示す。また、水原希子と岩田剛典の決定係数が低いことより、適合度は低い。高い適合度のモデルは複数の変数の影響をよく捉えているが、全体的な一貫性は低い。

2. 変数の統計的有意性: 時間帯と曜日は一部のモデルで有意だが、全体的には不一致。ローラのモデルでは写真ま

表 4 水原希子の結果

項目	推定値	標準誤差	p 値
(切片)	43001	58973	0.4692
X1 朝	15431	14408	0.2892
X1 夜	34110	12070	0.0067
X3	2512	1789	0.1663
X5 英語	-54994	39935	0.1745
X5 英語日本語	25664	44652	0.5679
X5 絵文字	-27290	41210	0.5108
X5 日本語	-55139	41756	0.1925
X5 日本語英語	-76958	40810	0.0650
X6 その他	23422	45976	0.6126
X6 私的	56667	40946	0.1724
X6 景色	-17439	43717	0.6916
X6 雑誌撮影	22568	41967	0.5930
X6 宣伝	19591	40115	0.6273
X7	24368	13518	0.0773

表 5 岩田剛典の結果

項目	推定値	標準誤差	p 値
(切片)	153247	28624	0.0000
X5 絵文字	43974	16703	0.0149
X5 日本語	18785	16735	0.2732
X5 文字なし	43621	38456	0.2683
X6 その他	-9810	33918	0.7750
X6 私的	52244	26066	0.0570
X6 雑誌撮影	9060	29276	0.7597
X6 宣伝	6040	27668	0.8291

たは動画の数が重要。ハッシュタグの数は一般的には重要ではない。投稿文字の影響はモデルによって異なる。投稿の内容はモデルにより結果が異なり、テーマの影響は多様。人物数の影響は不一致だが、時に重要な要因になる。

7 いいね数の比較

各アカウントから、「いいね数」が最も多い5つと最も少ない5つの投稿を選出し、選ばれた投稿の時間帯、曜日、画像や動画の数、ハッシュタグの数、投稿言語、内容を分析する。

7.1 データ比較の分析結果

データ比較の結果は表 6 のようになる。決定係数は、0.769 であった。1, 2 は「昼」と「火」が、X5 は「なし」、X6 は「グルメ系」がカテゴリーの基準となる。

7.2 結果からわかること

時間帯 (X1) はいいね数にはっきりとした影響を与えず、昼と比べて朝は負の影響、夜は正の影響の可能性がある。曜日 (X2) では、火曜日に比べて、月曜日に負の影響が見られるが、他の曜日では統計的に有意な影響はない。写真または動画の数 (X3)、ハッシュタグの数 (X4)、投稿の文字タイプ (X5)、内容 (X6)、人物数 (X7) は、いいね数に有意な影響を与えない。また、モデルの適合度は 0.77 で高いと言える。

表 6 データ比較の変数選択後の表

項目	推定値	標準誤差	p 値
(切片)	-58934	201907	0.7727
X1 朝	-128137	65823	0.0625
X1 夜	43472	47519	0.3687
X2 日	-56318	69885	0.4276
X2 月	-107239	57253	0.0723
X2 水	-8828	56247	0.8765
X2 木	-33085	50672	0.5195
X2 金	-61112	55390	0.2800
X2 土	80378	78473	0.3151
X3	13036	8144	0.1215
X5 ハッシュタグのみ	-130364	144539	0.3754
X5 英語	-142764	124317	0.2613
X5 英語と日本語	39869	132814	0.7664
X5 絵文字のみ	-111961	139122	0.4283
X5 日本語	-64544	126299	0.6136
X5 日本語と英語	110648	155804	0.4839
X6 コスメ系	55086	175587	0.7562
X6 その他	114633	177204	0.5234
X6 プライベート	251333	155583	0.1183
X6 景色	16052	183144	0.9308
X6 雑誌 or 撮影	181154	144990	0.2226
X6 思想や名言	-112256	139765	0.4292
X6 宣伝	28787	156109	0.8551

8 考察

曜日の影響: 月曜日と比較すると土曜日の投稿のいいね数は多い傾向がある。

投稿の種類とコンテンツ: 思想や名言の投稿と比較してプライベートな内容の投稿はいいね数が多い。

時間帯の影響: 朝と比較すると夜の投稿はいいね数が多い。戦略的な投稿タイミングは、特に土曜日や夜間の投稿が効果的であることがわかった。

9 おわりに

本研究から、インスタグラム投稿の「いいね数」は、投稿内容はアカウントごとに個性があるが、固い話は敬遠される。また、曜日や時間帯の影響が大きいことがわかった。

参考文献

- [1] 浅野真哉:『R で学ぶ統計解析入門 重回帰分析から実験計画法まで』, 森北出版, 2019.
- [2] Instagram: <https://www.instagram.com/?hl=ja> (2024 年 1 月 15 日最終閲覧)
- [3] 芸能人・有名人インスタグラムランキング: <https://insta.refetter.com/> (2024 年 1 月 10 日最終閲覧)
- [4] 田村周一:『R によるデータ解析の実践入門』, オライリージャパン, 2019.