

# YouTube Live におけるコメント生成確率分布を用いた 配信者特性の計量化

松元拓也 林高樹

慶應義塾大学大学院経営管理研究科

## 1 導入

いわゆるクリエイターエコノミーの成長が著しい中で、動画配信が職業としても認知され始めている。特に世界的な動画プラットフォームである YouTube で配信活動を行う人々は、ユーチューバー (YouTuber) とも呼ばれている。

このユーチューバーをはじめとする動画配信者を評価する上で用いられやすい指標が、チャンネル登録者数および同時接続者数である。例えば YouTube では、配信者のチャンネルを「登録」することが可能であり、これはブックマーク機能に近い。しかしながら動画を視聴する上でチャンネル登録は必須でないことや、またチャンネル登録は複数のチャンネルに対して行うことも可能であり、カジュアルなファン数を示す指標であると考えられる。一方で、よりコアなファン数を示すと考えられる指標が同時接続者数であり、ライブ配信を視聴した人数を指す。これは放送時間に合わせてリアルタイムに集まった人数であり、より積極的な支持がこの数値に反映されていると考えられる。

しかしながらこれらの指標が示すのは配信者に紐づくファン数のみであり、その人気を簡単に把握することは出来るものの、配信者が本来個別に持っているはず特性を読み取ることは難しい。例えば、同程度のチャンネル登録者数、同時接続者数であったとしても、瞬間的な盛り上げが上手い配信者もいれば、ラジオのように聞きやすいトークを長時間配信しているケースもあるだろう。そこで本研究では、YouTube の生配信におけるコメントを分析対象とし、配信者特性の指標化手法について検討した。

## 2 手法検討

本研究では、ライブ配信におけるコメント生成の背景に、配信者の特性 (キャラクター性や配信内容など) を仮定する。感覚的な理解としては、配信者が面白いことを話せばコメントは盛り上がり、単位時間あたりのコメント数が増加するはずである。言い換えれば、コメントが生成される確率分布のパラメータは、配信者の特性によりもたらされていると想定される。

また厳密には、YouTube Live における視聴者がリアルタイムで行う書き込みは chat (チャット) という名称であるが、本稿では国内にて使われやすいコメントという用語で統一している。

この検証を目的として、本研究ではバウト解析の枠組みを採用する。バウト解析は Shull et al. (2001) [1] で提唱され、動物の反応は不均一に発生するという考えから、ある時間では反応がまとまって出現するとしている。このまとまりをバウト (bout), 反対にバウトでない時間を休止 (pause) と呼び、このバウトの発生間隔や、バウト内における反応の発生率などを分析単位とした。しかしながら実際の動画では、配信時間の中で様々な変化が起きており、この変化もパラメータに折り込むべきであると考えられる。そのため本研究では Matsui et al. (2018) [2] のバウトパターンの時系列モデリングをライブ配信におけるコメント生成への適用を試みた。

## 3 方法および分析データ

Matsui et al., (2018) によるバウトパターンの時系列モデリングの枠組みを応用した。

$$\text{logit}(q_t) = \mu_{q,t}$$

$$\log(\lambda_t) = \mu_{\lambda,t}$$

$$\mu_{q,t} \sim \text{Normal}(\mu_{q,t-1}, \sigma_q^2)$$

$$\mu_{\lambda,t} \sim \text{Normal}(\mu_{\lambda,t-1}, \sigma_\lambda^2)$$

$$z_t \sim \text{Bernoulli}(q)$$

$$r_t = \begin{cases} 0 & (z_t = 0) \\ \text{Poisson}(\lambda) & (z_t = 1) \end{cases}$$

具体的には、単位時間を1秒と設定し、コメント発生パターンにベルヌーイ分布を適用し、さらにコメントが検出された場合に、その内部におけるコメント数をポアソン分布で仮定している。しかしながらこれだけでは、コメント全体に対してこの2つの混合分布(ゼロ過剰ポアソン分布)を仮定しているだけである。換言すれば、各時点 $t$ において、コメント数がゼロ過剰ポアソン分布に従って発生していること仮定している。さらに、システムのパラメータが時間と共に変化する状況を考える。具体的には、状態空間モデルを採用し、潜在変数 $\mu$ が直前の値を平均とした正規分布に従うと仮定する。分析データ\*1は、2022年12月15日時点でYouTube上にて閲覧可能なライブ配信済み動画4本を対象とした。

## 4 分析結果

図の1と2に、それぞれパラメータ $q$ パラメータ $\lambda$ の推定結果を示す。縦軸はそれぞれのパラメータであり、横軸は1秒単位の時間軸である。

## 5 考察と課題

詳細は発表当日にて報告するが、バウトの時系列モデリングにより、ライブ配信内におけるパラメータ推移の差が得られ、この違いは、配信者の違いによりもたらされていると考えられる。

今次提案した手法により得られる配信者ごとの指標を、本研究の目標である各々の特性評価につなげていくことが今後の課題である。また、分析対象を拡大することで、配信者特性をよりの確に表現することが可能な確率モデルのクラスを探索することも課題である。

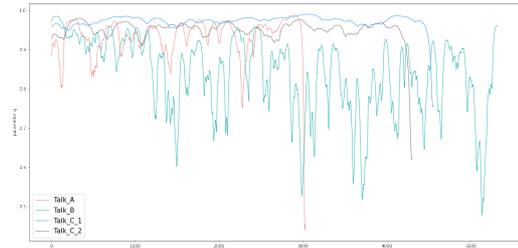


図1 parameter  $q$  の推定結果



図2 parameter  $\lambda$  の推定結果

## 参考文献

- [1] Shull, R. L., and Gaynor, S. T., and Grimes, J. A. (2001). Response rate viewed as engagement bouts: Effects of relative reinforcement and schedule type. *Journal of the Experimental Analysis of Behavior*, 75(3), 247 – 274.
- [2] Matsui, H., Yamada, K., Sakagami, T., and Tanno, T. (2018). Modeling bout – pause response patterns in variable-ratio and variable-interval schedules using hierarchical Bayesian methodology. *Behavioural Processes*, 157, 346 – 353.

\*1 分析データに関しては、YouTubeの利用規約に従い、ソフトウェア等を使用した自動収集は行わず、目視および手作業にて1秒毎のコメント数を独自に集計した。