Are Social Media Opinions Normally Distributed?

A white paper by Conversition

Introduction

It’s a common perception that people chatting about brands in social media either love or hate the brands. Why else would they voice their opinions online? Of course, we know that people who have neutral or slightly positive or slightly negative opinions also share their opinions online, but to what extent?

This paper will illustrate the distribution of social media sentiment in three areas:

1. On an individual level for 12 unique datasets
2. A selection of data representing thousands of brands
3. A selection of random data that doesn’t necessarily represent any brands

How is Sentiment Defined?

Conversition’s sentiment system incorporates both a 5 point Likert scale (box scores) and a continuous scale reflecting any sentiment score between 1 (negative) and 5 (positive). This research was conducted using the continuous scores such that a single verbatim might be assigned a score of 1.34 (very negative), 2.97 (neutral), or 4.57 (very positive). Examples of various types of sentiment are shown here.

Negative Sentiment
- So I updated to ios6 & now I have no contacts? Fuck apple.
- Apple maps need to pay for my gas since it keeps getting me lost
- eff you blackberry my phone is so effing slow

Neutral Sentiment
- I drink water or diet coke with dinner.
- Classes, library, fundraiser, flat tire, Walmart, and a little homework
- does US Airways fly to Cayman

Positive Sentiment
- Adidas! The #best make ever!!!
- Wouldn’t be great if Apple just did everything in the world? #apple
- THANK YOU APPLE!!!!! you r my life, you have amazing products and I just got my new ipad!
Sentiment for 12 Unique Datasets

Twelve unique sets of data were selected in order to determine the distribution of sentiment for brands that have a generally negative, generally positive, or generally neutral connotation.

Many of these datasets suffers from floor or ceiling effects. No matter how positive or how negative an opinion is, its score cannot be greater than 5 or smaller than 1. As such, a brand cannot register perfectly normal distributions when most of its opinions are extremely positive or extremely negative.

Though each of the 12 datasets isn’t a perfectly normal distribution, each of them do approximate a normal distribution in their own way. Is social media sentiment normally distributed? Yes!
Sentiment Across Thousands of Brands

It’s clear that opinions about brands or categories at an individual level are normally distributed. Now let’s look at opinions across thousands of brands and categories. In this case, we selected a random sample of over 120 million opinions from the Conversition database and examined the distribution of sentiment scores.

In this case, the distribution is somewhat positively skewed, though it is generally normal. What this distribution indicates is that the companies that come to Conversition for their social media insights are doing well in social media. They’re ahead of the curve and keeping their brand chatter in the right place – the positive place. Is social media data normally distributed? Mostly.

![Branded Social Media Chatter](image)

Sentiment of Random Data

Finally, let’s take a step back and consider the entire world of social media opinions. Forget whether the data reflects a specific brand or a category or a celebrity. Forget the focus on forward thinking brands and smart brand managers. Let’s simply collect a random sample of any messages expressed in the social media space.

The following dataset reflects about 1 million messages and is a lovely, perfectly centered, perfectly normal, unskewed distribution of social media sentiment. Is social media data normally distributed? Absolutely.
Conclusions

We know that people take to social media when a brand has angered or delighted them. These are the types of messages that people love to read and share with their friends because they are so memorable. But, this research demonstrates that social media opinions reflect the entire range of emotions leading to a fully normal distribution of sentiment. Essentially, most people express opinions that are generally neutral and unemotional, and, as the opinions become more and more heated or excited, their frequency becomes less and less. The goal, now, is to ensure your brand focuses on the right side of the normal curve, the positive side.