Jan Lorenz, jalorenz@ethz.ch, post@janlo.de
Chair of Systems Design, ETH Zurich, www.sg.ethz.ch

Universality in Continuous Opinion Dynamics

Empirical studies of movie rating distributions (based on preprint arXiv:0806.2305)

Idea

Models of continuous opinion dynamics often lack comparison with empirical data. Nowadays ★-based rating systems are ubiquitous on the internet. These rating scales are a good approximation of continuous opinions because ratings are adjustable fairly fine-grained. Their empirical analysis can uncover funamental universal characteristics of continuous opinion dynamics. This helps to better understand the human usage of rating systems, which is important because rating systems produce the data for recommendation systems.

Universal Empirical Observations

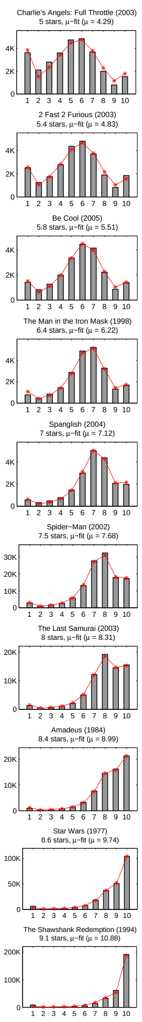
10★-movie-rating histograms on IMDb show two facts:

(i) Histograms have **'Gaussian-like'** shape on the central bins (2★–9★). That means there is only one peak and the decay when deviating from the peak looks like a Gaussian.

(ii) **Extreme** bins (1★ and 10★) are **higher than expected** from extrapolating the Gaussian-like shape on central bins.

These observations contradict claims on the shape from the literature: for example the uniform, U-like, or single-peaked shape. (These claims were either made ad hoc for theoretical reasons, or on the basis of data on the coarser grained 5★-scale.)

IMDb Examles: Histograms and μ-fit



Theoretical Explanation

Both empirical observations can be explained on the basic assumption, that a rating is a random variable:

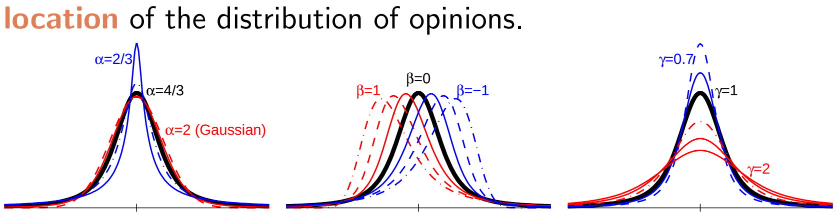
(i) Ratings are **pre-averaged opinions** of opinions of others, the current average rating, or several criteria which are seen as independent random variables. Thus, the distribution of ratings has a Gaussian-like shape due to the **central limit theorem**.

(ii) Ratings are **discretised and confined** continuous opinions. All probability mass which lies beyond the rating scale has to be collected in the extremal bins.

The only distribution which can evolve under averaging is the **Lévy skew α-stable** distribution $S(\alpha,\beta,\gamma,\mu)$. Thus, it is a natural candidate to characterise rating histograms according to (i). Its probability density function is

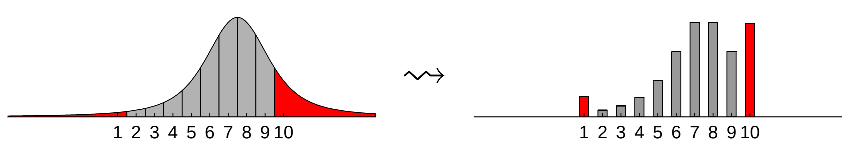
$$P(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi(t)e^{-itx}\,dt, \quad \varphi(t) = e^{\mu - |\gamma t|^{\alpha}\,(1 - i\beta\,\mathrm{sign}(t)\Phi)}.$$

with $\Phi = \tan(\frac{\pi\alpha}{2})$ if $\alpha \neq 1$ and $\Phi = -\frac{2}{\pi}\log|t|$ for $\alpha = 1$. Parameters characterise $\alpha \in (0,2]$ **peakedness**, $\beta \in [-1,1]$ **skewness**, $\gamma \in [0,\infty)$ **scale** (or dispersion), and $\mu \in (-\infty,\infty)$ **location** of the distribution of opinions.



The normal distribution comes out for $\alpha = 2$, in this case β has no effect. For $\alpha < 2$ the distribution has power-law tails which decay on both sides as $P(x) \propto |x|^{-(1+\alpha)}$. The mean of the distribution exists for $\alpha > 1$ and is μ. The generalized central limit theorem states that the average of several random variables which distributionS have power law tails with the same exponent converge to a Lévy skew α-stable distribution.

From a Lévy skew α-stable distribution a $(\alpha,\beta,\gamma,\mu)$-histogram is produced by discretisation according to (ii) as:

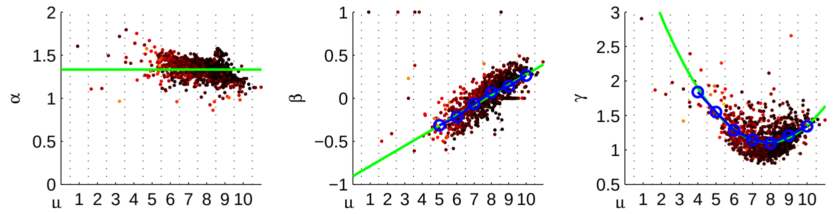


Movie Rating Systems

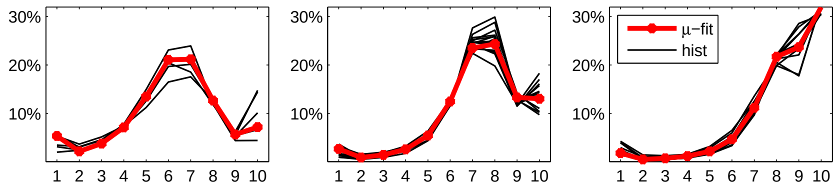
Almost everything can be rated with some ★'s on the Web. The costs of rating is just one click. Movies gain several 10,000 ratings on popular sites: The "Internet Movie Database" IMDb uses a 10★-scale. The DVD-rental service "Netflix" NETFLIX uses 5★-scale. Of most interest is the average number of ★s. Usually it visible close to the rating module. The user-movie-matrix of ratings is the basis for recommendation systems. Netflix awards a prize for a considerable improvement of its recommendation system. Only the histograms of all ratings for a particular movie are of interest here.

Data Fitting and Empirical Validation

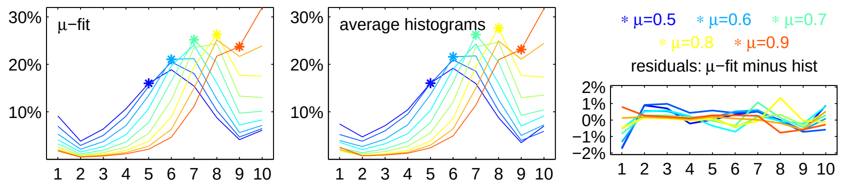
Fitting $(\alpha,\beta,\gamma,\mu)$-histograms against empirical IMDb-histograms reveals values with strong correlation with respect to μ.



Blue lines show averages between gridlines. Minimal γ and $\beta = 0$ appears both for $\mu \approx 7.5$ which is the overall average μ. Green lines show constant, linear and quadratic fits. They are used to construct a one-parameter fit called **μ-fit** (examples at left margin). Its quality is shown by comparison with IMDb-histograms,



and by comparing against average empirical histograms.



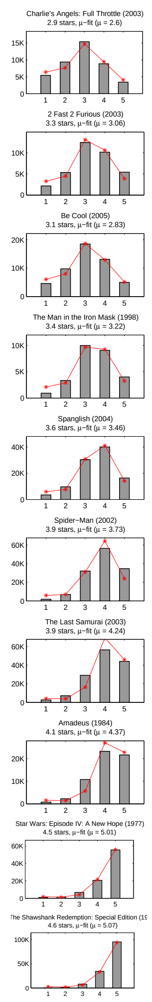
Data Description

IMDb Histograms of all movies with more than 20,000 ratings (1,086). Used for the analysis of this study.

Netflix-Prize data Histograms with more than 20,000 ratings are used to validate μ-fit constructed on the IMDb dataset.

Using μ-fit on Netflix data (see margin) is still good. A four-parameter fit (not shown) indicates that a one-parameter fit is possible with the same characteristics regarding scale and skewness, but less peaked ($\alpha \approx \frac{5}{3}$). These thinner tails may reflect that on a DVD rental website rating is not as hot as on IMDb.

Netflix Examples: Histograms and μ-fit*

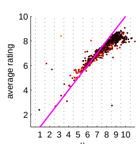


* with natural transformation $\mu_5 = \frac{1}{2}\mu_{10}$, $\gamma_5 = \frac{1}{2}\gamma_{10}$

Conclusion

Assuming that ratings are discretised pre-averaged continuous opinions leads to good approximations of real-world rating histograms. The underlying **opinion distribution is not normal but fat-tailed**. The one-parameter μ-fit is still good and shows surprising regularities: an average movie ($\mu \approx 7.5$) has most narrow distribution and no skewness. **Deviation from an average movie makes the opinion distribution broader and skew** pronouncing the deviation. (The better the fatter the right tail and vice versa). The **universality** $\alpha \approx \frac{4}{3}$ is reasonable.

The results can be used to **characterise movies** in a new way by quantifying their underlying opinion distribution. The quality measure μ for example spreads better than the classical mean (see margin). Further on, differences from average histograms with a given quality can be detected and characterised.



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Departement Management, Technology, and Economics
MTEC