Further comments on evolution in Hollywood film: the role of models

Mike Baxter[1], 16 Lady Bay Road, West Bridgford, Nottingham, NG2 5BJ, UK
(e-mail: michaelj.baxter@btconnect.com)

July 2014

Abstract

This paper is a further contribution to the debate about the interpretation of the results in an important paper published by Cutting et al. *(2010). Not all the protagonists in the debate believe that claims about evolution in Hollywood film made in the paper can be sustained. The issues are covered in detail in the debate that has taken place on the* Cinemetrics *website and will not be rehearsed in any detail. The paper elaborates on some points raised in Baxter (2013), but the main intention is to stand back from the specifics and muse on some more general issues. These concern aspects of cinemetrics data analysis associated with the 'explosion' of data that the* Cinemetrics *website has made possible, and the increased complexity of statistical methods that have been used to analyze such data. Particular attention is paid to the contrast between deductive and inductive modes of data analysis, and the use of statistical models. Arguably, Cutting* et al. *(2010) engage with both modes of analysis, and their conclusions are heavily dependent on models to which alternatives exist, and which may be wrong. This lies, I think, at the heart of the debate.*

1 Introduction

This paper is prompted by recent debate that has resulted in papers published on the *Cinemetrics* website in 2013–4 (Baxter, Salt and Cutting in order of publication). They stem from a publication by Cutting *et al.* (2010) based on a research project that, among other things, traces the development of editing patterns and structure in Hollywood film, more-or-less since it settled down about the mid-1930s after the introduction of sound. There have been, as Cutting notes in his recent paper, debates between him and Barry Salt on *Cinemetrics* in 2010-11 about how sustainable the claims made in Cutting *et al.* (2010), which have entered the literature, are valid.

These are about the 'evolution' of Hollywood film. I don't think anyone doubts that film has evolved, but Baxter and Salt doubt if has done so in the quite specific ways that Cutting *et al.* (2010) claimed. The work of Redfern (2012a), who has arrived at similar conclusions, should also be mentioned in this respect, but it's not on the *Cinemetrics* site.

The various arguments that have been put forward will not be rehearsed in any detail – they are available in the publications mentioned above. What I want to attempt in this paper is to muse on the role of statistics in cinemetrics, and specifically the use of models. It might be thought of as a form of what is sometimes called *meta*-analysis; you're supposed to say what the *meta* thinking is about. So ...

2 Background

2.1 The context

Cutting *et al.* (2010) published an important paper based on an impressive amount of data collection, eventually interpreted in terms of both filmic development and psychological theories

[1]Emeritus Professor of Statistical Archaeology, Nottingham Trent University, UK

concerned with 'attention'. Others, myself included, don't think their data and analysis sustain their interpretation. My own arguments are in Baxter (2013)[2]. I think this is what the current discussion is about, but want to try and set it in a wider framework.

2.2 The general idea

Ideally the conclusions you draw from numerical data should be driven by ideas. Which are derived from thought. But the conclusions may also be based on models you use to interrogate the numerical data which – and disregarding the technical issues I have already raised in my earlier contribution – is where I think it gets interesting at a higher level. The paper by Cutting *et al.* (2010) will be used as a peg to hang ideas on – to emphasize, the aim here is not to dispute their conclusions (which has been done elsewhere) but to look at the use of statistical methods in cinemetric study beyond the average shot length (ASL), and more particularly the use of models. The aim is to keep the language as non-technical as possible, though technique can't be entirely avoided. A brief and partial summary of the history of the development of statistical thinking (my view of it) follows as a prelude to discussing the use of statistics and models in cinemetric study.

2.3 On statistics

The development of statistical thinking in the first half or so of the 20th century is one of the great intellectual achievements of that period, possibly undersung. Methodological developments were greatly enhanced by practical considerations in the second world war, but it is the earlier thinking of Ronald Fisher (who had predecessors like Karl Pearson) that is often cited as pivotal. Essentially Fisher developed a lot of the mathematics that underpins the statistical methods still taught in university courses, but was driven to do so by practical concerns to do with data analysis.

Other very bright people came along; disagreed with Fisher on the theoretical front and developed the mathematics[3]. My personal view is that statistics should be focused on the need to analyze data, but mathematical statistics became a discipline in it's own right, still practised by scholars some of who probably wouldn't recognize real data if it jumped up and said 'boo' to them. This isn't entirely prejudice on my part; see George Box's autobiography for an account of the split between mathematical and applied statisticians in US university departments in the later 20th century (Box, 2013). You can see this in some of the journal literature, devoted to 'mathematizing', in a way that most people can't read, things which everyone knows to be true. Mathematical rigor has its place but can be overdone.

This is by way of pointing out that there are theoretical divisions in the statistical world and a lot of what gets published is pretty useless from a practical view. It doesn't differ from other disciplines in this respect – academics are obliged to publish whether they have something useful to say or not (and I do not exclude myself from this judgment).

So you have what might be perceived as a dichotomy between mathematical and applied statistics. I don't necessarily believe in dichotomies but will adopt the idea for a moment. On the applied statistics front you have what might be viewed as a dichotomy between hypothesis-testing methodologies (of the sort originally developed by Fisher and others) and exploratory data analysis (EDA), often associated with the seminal work of John Tukey (1977). As far as cinemetric study goes EDA methodology has been championed by Nick Redfern in numerous papers, though he is a bit inconsistent in this since he has also done this in conjunction with quite explicit hypothesis-testing ideas (Redfern, 2012b).

I recall reading Tukey (1977) a bit after its publication. It was undoubtedly important, has been influential, and unfortunately has been afforded cult status at times in some disciplines (I'm

[2] I offered this to the journal in which Cutting *et al.* (2010) was published; was rejected promptly without review because it was too long; offered it elsewhere and was rejected (without full review) because it was judged to be too complex. I actually understand the reasons why this happens and have no complaints. It is a pity though that in the world of film scholarship, and in the judgment of journal editors, anything quantitative is considered to be too challenging for their audience.

[3] Statistics has a lot of 'theory' about which there has been vigorous and frequently acrimonious dispute. I'm not going into this here, but see McGrayne (2011) for an entertaining account.

thinking statistics, archaeology, geography etc. – people have built careers on exploiting the ideas). EDA methods are just one way of looking at data that can avoid hypothesis-testing ideas but can feed into them. In terms of 'building careers' a lot of my own research eventually concentrated on applications of multivariate analysis to archaeological data. Much of this was exploring data for structure – so it's 'exploratory' with a small 'e'. But you can think of it as a form of EDA if you wish. What sometimes emerges is that what you can think of as inductive pattern seeking methodologies can suggest structure in data that can then be examined more closely using what have been called 'hypothetico-deductive' methods.

Superficially, you have another dichotomy here between 'inductive' and 'hypothetico-deductive' methodology. I think all these dichotomies, which some perceive, are false. As far as data analysis goes different ideas feed off each other and the 'dichotomy' just represents two ends of a continuum. There are plenty of scholars (and I'm thinking of fields other than cinemetrics) who take an eclectic or pragmatic view to analyzing their data and use whatever methods help to answer their questions – think of them as theoretically and methodologically promiscuous. But then you also have the theoretical monagamists who are absolutely convinced that there is one way to look at things – theirs – that is correct. You also have serial monagamists who change their ideas and then assert that their latest idea (which may be in complete opposition to their previous one) is 'correct'.

The above assertions could be justified with reference to what I know about the development of quantitative analysis in geography and archaeology. I won't bore the reader with the detail, but Baxter (2003) touches on the subject as far as archaeology is concerned. And what is the relevance of all the above to cinemetric study?

3 Statistics and Cinemetrics

Cinemetrics has a prehistory that is documented in discussions on the *Cinemetrics* website, and there is a video available of Yuri Tsivian discussing this[4]. I think though that most scholars interested in this sort of thing would regard Barry Salt's (1974) paper *Statistical Style Analysis of Motion Pictures* as the seminal paper in the development of what is now viewed as cinemetric data analysis. Salt's book (in several editions – the latest is Salt, 2009) on *Film Style & Technology* is also important.

I've given the title of Salt's paper in full to emphasize the term *Statistical Style Analysis*. Salt's work, as I understand it, inspired what might be considered to be the next major development which was the creation by Yuri Tsivian of the *Cinemetrics* website in the mid-2000s. You will notice that there is about a 30 year gap between Salt's original paper and this. My thinking is that Salt was seriously ahead of things and that you really needed the computational power that developed somewhat later to explore his ideas more thoroughly. You need data, and preferably quick ways of looking at it. *Cinemetrics* has made available an enormous amount of data of varying quality (which is a recognized problem being dealt with). What you can do with such data you choose to use, with statistical analysis, has increased tremendously since Salt first published on this.

But you have to start, possibly, with ideas. If you read Salt (2006) you will realize that, before his serious involvement with film, he was a trained physicist. This is a 'hard' science and one aspect of it is that you use experimental data to test or explore specific ideas. This can be called 'hypothetico-deductive' and characterizes 'hard' scientific methodology. That is, you start with a theory, hypothesis, or simply an idea and collect data to examine its validity. My understanding is that Salt (1974) was conceived in this sort of spirit. In passing I should note that the term 'hypothesis' used here is perhaps a bit more general than the way it is used in the statistical literature in the context of 'hypothesis-testing' methodology.

Then along came powerful computing facilities, tools like *Cinemetrics* and – compared to what was available before – enormous amounts of data and very powerful statistical software for exploring the data. It is, in fact, only fairly recently, that the usefulness of this 'power' has begun

[4] http://neubauercollegium.uchicago.edu/events/uc/Cinemetrics-Conference/

to be investigated. A lot of cinemetric analysis has been based on the use of descriptive statistics such as the ASL or median shot length (MSL). I'll return to this shortly.

The main thought here is that there has been an 'explosion' in the amount of data available since the establishment of the *Cinemetrics* website without this necessarily being accompanied by a similar explosion in ideas about what to do with it. It is surprisingly hard to find published examples of what people do with the data that has been collected.

But one way you can go about things, if you have a large body of data that is 'coherent' in some sense, is to simply look at it and see what happens. I think this is what Salt would regard as inductive analysis. That is, you start off with the data to see if it throws up any ideas, rather than using the data to interrogate an idea. If I can indulge myself and use some of my own work to illustrate, I am guilty of using inductive methodology (multivariate statistical pattern seeking methods) to explore structure in the Biograph films of D.W. Griffith, which is published on the *Cinemetrics* website. This was just a pattern-seeking exercise, and I eventually decided I wasn't entirely satisfied with the outcome.

After I'd done this work Yuri Tsivian directed me to an article on *Pace in the Movies* by D.W. Griffith[5]. This, in effect, states how films (of a certain kind) should be structured in terms of a series of 'climaxes' of increasing intensity. I wondered if his own films were so structured; investigated his major feature films between 1914 and 1921; and concluded they were. Armed with this discovery I returned to the Biographs with the specific intention of seeing if they had a similar structure, and they mostly don't. Technical details are available in Baxter (2014) with a more informal videoed conference presentation[6].

One point, possibly just a thought, to make here is that you can interpret *Pace in the Movies* as generating an 'hypothesis', in the general sense of the word, about film structure that can be investigated using empirical data (shot-lengths). So perhaps you can think of this as hypothetico-deductive. As far as the Biographs go my first attempts at looking at them used undirected pattern-seeking methods (sometimes called unsupervised learning methods). My second attempt was much more directed in the sense that I had a specific pattern in mind and was investigating if there was any evidence in the data for it, so it's more at the hypothetico-deductive end of the scale. What you have here is a distinction between directed and undirected methods of pattern seeking (and other terminology exists). This probably can't be equated exactly with the distinction between hypothetico-deductive and inductive methodology but analogies undoubtedly exist.

Methodological details can be found in the references cited, but I should emphasize that I did not use any formal hypothesis testing methodology as a statistician would normally understand it. I developed a graphical approach to looking at structure in films using shot-length data and interpreted the results subjectively in terms of the 'hypothesis' I had read into Griffith's article.

This brings us, in a very roundabout way, to the idea of statistical models and cinemetric data analysis. This will be discussed in the specific context of Cutting *et al.* (2010), but the ideas are intended to be more general. The paper, and its conclusions, are heavily dependent on the use of statistical models, which might be regarded as at the hypothetico-deductive end of the spectrum, but are also used for what could be seen as exploratory (inductive) purposes.

The use of models operates at several levels. One is that the time series structure of shot-lengths in films can be modeled using an autoregressive (AR) model. This is a quite specific model and, for different reasons, both Salt (2014) and Baxter (2013) have questioned its validity – there are other models available. I did not go into this in great detail in my paper, but the evidence from Cutting *et al.* and their data is that the AR model is simply not appropriate for many films – I'll come back to this.

A second level at which modeling is used in Cutting *et al.* (2010) is that, having used models to generate statistics descriptive of film structure they then use other forms of statistical modeling to investigate temporal patterning in the statistics. This could be regarded as at the inductive end of the scale. Specifically what they do is fit linear or quadratic regression models to the derived data. These are parametric models that impose pattern on the data that are interpreted in terms

[5]There is some dispute about whether Griffith wrote the article or simply lent his name to it, but the assumption is that he was complicit with the sentiments expressed whoever wrote it.

[6]http://neubauercollegium.uchicago.edu/events/uc/Cinemetrics-Conference/

of evolution in their paper. What Baxter (2013) and Redfern (2012a) have suggested is that if you don't impose this sort of pattern the evidence for evolution of the kind that has been claimed disappears if you use more flexible modeling methods.

Most of this has been said, in possibly different form, elsewhere. The general point to make is that the Cutting *et al.* (2010) paper is highly dependent on the use of specific models, and models can be wrong (which I think they are in this specific case). To repeat, this is a general issue if you engage in the cinemetric analysis of quantified filmic data at a level beyond calculating average (or median) shot-lengths. The problem about models is that they can be wrong, and if you don't allow for this in data analysis you can end up with the wrong conclusions.

Specifically I think that some of the analysis and interpretation in Cutting *et al.* (2010) depends on models (the AR assumption) that are demonstrably unsustainable, or which can be replaced by other models that lead to different conclusions. There are undoubtedly patterns in the data, but the variation is so great that they can be interpreted in different ways, and the use of specific models obscures this fact. Salt (2014) discusses the relationship between the undoubted decrease in ASL in recent years and the first order auto-correlation (AC) coefficient in shot-lengths. This, effectively, is what some of the analysis in Cutting *et al.* demonstrates though it isn't interpreted in this way. The correlation between the decrease in ASL and the AC coefficient is high, but not perfect, particularly in recent years, and I find this intriguing.

What is interesting though (bearing in mind I've only been involved in this kind of thing for about two years and a half) is my perception of how things are evolving from a statistical point of view. A lot of this is down to initiatives on the *Cinemetrics* website, but it seems to me that other individual scholars and groups are getting involved, and you hope (and expect) this will increase. The use of statistical analysis in cinemteric study is becoming more 'sophisticated'; see Redfern (2013) for a recent example.

'Models' can be thought of in different ways, and it can depend on your academic background, or way of thinking. They can be quite specific mathematical models about the way things happens or they can be 'looser' ideas, capable of experimental investigation, that don't necessarily need to be formally expressed in mathematical terms, in the first instance at least. And you can start with the data and see what it tells you – inductive, and don't be ashamed about this – which may throw up ideas that can be cast in the form of a 'model' that can be investigated more rigorously.

So I think that the idea of 'models' in cinemetric data analysis is interesting, but would not wish to prescribe what the term means. I'd almost like to use something like ideas capable of empirical investigation. But the point about models, when expressed in mathematical or statistical form, is that they can be wrong or (at the statistical level) there may be more than one model that can 'explain' the data which may be contradictory. Models can depend on assumptions that limit the range of conclusions you can draw from them; if the assumptions are wrong then the conclusions may also be. It is, I think, a problem with analyses in Cutting *et al.* (2010) that comes down to the use of inappropriate models. If you are going to use models it is necessary to be aware of their limitations; that they depend on assumptions; and, ideally, you should put mechanisms into place that involve checking the validity of the model[7].

Acknowledgements

I hope the text makes clear the intellectual inspiration I have derived from the works of Barry Salt and Yuri Tsivian. Some of the ideas that I've mused on above were specifically prompted by e-mail discussions I've had with Barry. The paper, and the thought behind it, has also benefited from a visit to the University of Chicago in early 2014, which enabled me to meet Yuri for the first time and Daria Khitrova and discuss ideas with them. Along with others they maintain the *Cinemetrics* website which is an invaluable resource – should you acknowledge websites? My visit

[7] I won't go into this in detail, but I'd distinguish between 'models' that are driven by mathematical or theoretical thinking as in Physics, for example, and those from other areas of study that embody ideas that you can think of as a model that lack a theoretical background in mathematical terms.

to Chicago was supported and funded by the Neubauer Collegium, which facilitates inter/multi-disciplinary research in the humanities (in my particular case statistics and film studies) that is, I think, unusual and great! Of the staff of the Collegium I met, Jamie, Josh and Sarah have my thanks for their support; Jamie arranged the conference that resulted from the visit. Finally, and as ever, thanks to Hilary Cool for reading earlier drafts and telling me where it didn't make sense.

References

Baxter M.J. (2003) *Statistics in Archaeology*. London: Arnold.

Baxter M. (2013) Evolution in Hollywood editing patterns?
http://www.cinemetrics.lv/dev/Evolution_paper_for_Cinemetrics.pdf

Baxter M. (2014) Cutting patterns in D.W. Griffiths silent feature films.
www.academia.edu/5684527/Cutting_patterns_in_D.W._Griffiths_silent_feature_films

Box G.E.P (2013) *An Accidental Statistician*. Hoboken, New Jersey: Wiley.

Cutting J.E. (2014) More on the evolution of popular Hollywood film editing
http://www.cinemetrics.lv/dev/cuttingcinemetricx3.pdf

Cutting J.E., DeLong J.E. and Nothelfer C.E. (2010) Attention and the evolution of Hollywood film. *Psychological Science* **21**, 440-447.

McGrayne S.B. (2011) *The Theory That Would Not Die*. New Haven: Yale University Oress

Redfern N. (2012a) Robust estimation of the modified autoregressive index for high grossing films at the US box office, 1935-2005.
http://nickredfern.files.wordpress.com/2012/11/nick-redfern-the-mar-index-for-hollywood-films1.pdf

Redfern N. (2012b) The lognormal distribution is not an appropriate parametric model for shot length distributions of Hollywood films. *Literary and Linguistic Computing*, doi:10.1093/llc/fqs066

Redfern N. (2013) Film style and narration in Rashomon. *Journal of Japanese and Korean Cinema* **5**, 21-36.

Salt B, (1974) Statistical style analysis of motion pictures. *Film Quarterly* **28**, 13-22.

Salt B. (2006) *Moving Into Pictures*. Starword, London.

Salt B. (2009) *Film Style & Technology: History & Analysis, 3rd edition*. Starword, London, 2009.

Salt B. (2014) Salt on Baxter on Cutting.
http://www.cinemetrics.lv/dev/CutThoughtC.pdf

Tukey J. (1977) *Exploratory Data Analysis*. Addison-Wesley