"Bootstrapping"—
A Contemporary Statistical Term*

John Aldrich
Economics Division
School of Social Sciences
University of Southampton
Southampton
SO17 1BJ
UK

e-mail: john.aldrich@soton.ac.uk

July 2012

*I am grateful to Bradley Efron and Yves Berger for information.

Introduction

The bootstrap method in statistical analysis uses computer simulation to generate standard errors, confidence intervals and probability values—some of the most frequently calculated entities of statistical inference. The method was described in a 1979 article by the Stanford statistician Bradley Efron and it represents probably the most influential contribution to statistical methodology of recent decades. A search of

JSTOR Statistics journals to 2008 found the bootstrap mentioned in nearly 4,000 articles; a Google search on bootstrapping in statistics found over 800,000 results. The silver anniversary of the bootstrap was marked by a special issue of the journal *Statistical Science* (May 2003). As a linguistic phenomenon the statistical bootstrap is equally impressive. Not only is the word frequently encountered in the writings and utterances of statisticians, it is used both as a noun and a verb—*to bootstrap*—and there are numerous derived constructions such as *bootstrap test*, *wild bootstrap* and *block bootstrap*.

The idiom "to pull oneself up by one's own bootstraps" from which the statistical terminology derives has a history of figurative use—see Quinion (2002) and Zimmer (2005). At first sight it appears that the statisticians simply borrowed an established term from computing. Thus the *OED* has an entry for the verb *to bootstrap* meaning "to make use of existing resources or capabilities to raise (oneself) to a new situation or state; to modify or improve by making use of what is already present" and a quotation from Diaconis and Efron (1983) follows examples dated 1958 and 1960 from computing journals. On examination, however, it appears that the statistical term does not have quite this meaning and that it was not borrowed from computer science.

In the context of word-formation in Statistics—described in Section 1—*bootstrap* is both an unusually inventive coinage and a word of its time. The statistical career of *bootstrapping* is described in Section 2 and added perspectives on the terminology are provided from the way(s) it has gone into other languages, an aspect considered in Section 3. The statistical terms mentioned in this paper are defined in Dodge's *Dictionary of Statistical Terms* (2006) and there is information on the history of

individual terms in David (2001) and Miller (2011). Miller's compendium covers mathematical terms generally as does Aldrich's (2011b) survey of mathematical terminology; the latter is the source of some of the material in the present article derives. Boos and Stefanski (2010) provide an elementary introduction to bootstrap methods in statistics.

1 The language community

Bootstrapping—idea and term—came from the mathematical statistics wing of the American statistical community: bootstrapping was first presented in an article published in the *Annals of Statistics*, a journal of the Institute of Mathematical Statistics. When the American Statistical Association, the umbrella organisation for American statisticians, was set up in 1839 its task was to gather facts about society—'state-facts'—but since then the scope of statistics has broadened enormously and specialisms like mathematical statistics have developed; for a sketch of the history of statistics and of some leading personalities see Aldrich (2011a). The Institute of Mathematical Statistics (IMS) was established in the 1930s and it drew on two traditions, a statistical tradition from Britain and a tradition of probability as pure mathematics from Continental Europe. The chief figures of the British tradition were not social recorders of the early nineteenth-century kind but applied mathematicians interested in biology. Their successors, the mathematical statisticians of the IMS, were not committed to research in any specific subject but aimed to develop mathematical techniques applicable to all: they wanted to produce good mathematics and practically useful techniques.

All this activity required vocabulary and H. A. David's "First (?) Occurrence of Common Terms in Statistics and Probability" (2001) provides a convenient list of the most important terms and a record of when they came into use; the most recent coinage is from 1984. David's list has around 450 terms; the International Statistical Institute's *Multilingual Glossary* (2011) has 5000 entries and there are about 500 statistical entries on the Miller page. Around 90% of David's terms were coined in the twentieth century and coined overwhelmingly in English with only a few borrowed from German, French and Italian; the pattern is rather different from that for mathematics in general with its older word-stock—see Aldrich (2011b). In David's list the ordinary educated reader will find creations like *autoregression* and *unimodal* and eponymous terms, such as *Bayes estimate* and *Studentization*, but mostly familiar words, like *alias* or *association*, used in special senses or fixed in set phrases like *balanced incomplete blocks* or *kernel estimate*.

The leading contributors to David's list are the British statisticians Karl Pearson (1857-1936) and Ronald Fisher (1890-1962). Pearson, like other writers of his time, was apt to produce classical neologisms—terms like *kurtosis* and *heteroscedasticity*—while Fisher's practice was to invest existing English words of a certain weight with new meaning—thus *efficiency*, *fiducial* and *variance* became technical terms and Fisher's new words, such as *covariance* and *statistic*, were usually variations on existing ones. The most influential figure in the IMS of the 1940s was Abraham Wald (1902-1950). Wald personified the confluence of British statistics and Continental mathematics: a Viennese-trained geometer, he moved to the United States in 1938 where he encountered the British tradition in mathematical statistics. His coinings, such as *admissible* and *risk function*, did not have the ingenuity of Fisher's

adaptations from ordinary language but they were effective and remain typical of linguistic creation in mathematical statistics. Conflicting views on what makes mathematical terminology effective are reviewed in Aldrich (2011b).

Some neologisms from the bootstrap era illustrate the regular processes of vocabulary enlargement: *kriging* is an eponymous term, *p* formula* derives from the notation used to express the formula, and *Markov chain Monte Carlo* is a mechanical compound of the eponymous term *Markov chain* and the jokey *Monte Carlo* for computer simulation techniques based on random number generation. The most prominent statistical word-smith of the bootstrap era was John Tukey (1915-2000). Tukey's way with words was completely different from Pearson's or even from Fisher's: he favoured the informal, the colloquial. While *histogram* reflected Pearson's classical taste, *box and whisker plot* reflected Tukey's and Tukey even tried to refresh the existing language by replacing the august *median* by the quotidian *hinge*. One of Tukey's terms—*jackknife*—is part of the bootstrap story, as we see below. Tukey's explanation of the term—"a boy-scout's jackknife is symbolic of a rough-and-ready instrument capable of being utilized in all contingencies and emergencies"—speaks volumes for the instrument's strengths and limitations but conveys no idea of what it does.

Tukey's taste for the striking and the accessible was in tune with the times for English language mathematicians were becoming less stuffy in matters of terminology: topologists talked about *the hairy ball theorem* while group theorists investigated an object they called the *monster* and contemplated conjectures known collectively as *monstrous moonshine*; see Conway and Norton (1979) and Quinion (1998). Physicists

were playful too in investing the words *boojum* and *quark* from *Finnegan's Wake* and *The Hunting of the Snark* with technical meaning: see David Mermin (1981) and Murray Gell-Mann (2001). Classical neologisms were becoming rare: in 1969 Alan H. Schoen was pleased to draw on his classical schooling to create the term *gyroid* but in the nineteenth century that private pleasure would have been a duty.

2 The bootstrap

For the origin of the term *bootstrap* in statistics there are two main sources of information: the publications that launched the term/method and the fuller and more informal accounts Efron has given since. Other statisticians have not written about the word although it has evidently piqued their curiosity. The chorus of foreign language translations given in Section 3 below provides a commentary on the word.

Bradley Efron (born 1938) introduced the bootstrap in a paper published in the *Annals of Statistics* in 1979 and based on his 1977 Rietz lecture to the Institute of Mathematical Statistics. Efron (1979a: 3) has formal definitions of *bootstrap sample* and *bootstrap distribution* while *bootstrap methods* and *bootstrap* (noun and verb) acquire their meanings through association with these defined concepts. The Rietz lecture was an important lecture—Henry Rietz was a founding father of the Institute—and by 1977 Efron was established enough to be invited to give it. Efron's offering has proved to be the most influential of the series so far but it confounded expectations: Efron recalled to Holmes et al. (2003: 271) that his paper was initially turned down by the *Annals* because it "didn't have any theorems in it."

The starting point for the lecture both conceptually and linguistically was the jackknife, an old and somewhat arcane technique—originating in Quenouille (1949) and Tukey (1958)—but then enjoying a vogue. The title of Efron's paper was "Bootstrap methods: another look at the jackknife" and he has since described how for several years before 1977 he worked at understanding the jackknife. Efron (2000: 1295-6; 2010: 179-81) reports how his first name for the published *bootstrap distribution* was the *combination distribution*; however, no documents using pre-bootstrap terms appear to have survived. In the published lecture Efron (1979a: 1) gave his aim as "to explain the jackknife in terms of a more primitive method, named the "bootstrap" for reasons which will become obvious." More recently Efron (2010: 180) has reported a competitive relationship with the jackknife and with Tukey. "Tukey was full of terrible names for mathematical things, and I wanted to kid him a bit. Also I didn't want students to be lectured on 'the combination distribution *versus* the jackknife' because they'd obviously prefer the jackknife…"

Not wishing to labour the "obvious" Efron (1979a) does not discuss the choice of terminology except in the concluding acknowledgement where he (25) writes: "I also wish to thank the many friends who suggested names more colorful than Bootstrap, including Swiss Army Knife, Meat Axe, Swan-Dive, Jack-Rabbit, and my personal favorite, the Shotgun, which, to para-phrase Tukey, "can blow the head off any problem if the statistician can stand the resulting mess." All the alternatives have a Tukeyesque flavour and would not have been considered appropriate in the time of Pearson, Fisher or Wald.

In a second paper from the same year Efron (1979b: 465n) wrote that "The name "bootstrap" is meant to be euphonic with "jackknife," the two methods being closely related as we shall see, and also to convey the self-help nature of the bootstrap algorithm." The reference to "self-help" links the word to the expression *to bootstrap* found in the *OED* and mentioned in the Introduction above. Both that term and one that the statisticians certainly *did* use—that of *booting* a computer—are rooted in the expression "to pull oneself up by one's own bootstraps." Quinion (2002) has given some history for the computer booting usage, speculating that Robert Heinlein's science fiction story *By His Bootstraps* of 1941 played a part in establishing it in the 1950s. Quinion and also Zimmer (2005) discuss the remoter history of the idiom and its association with Raspe's Baron Munchausen. Zimmer (2005) also traces a shift away from the initial meaning of performing an impossible or preposterous task to performing a possible one: in becoming a cliché the expression lost its edge.

Efron first gave an extended account of the term in *An Introduction to the Bootstrap* (with R. J. Tibshirani) where they (1993: 5) explain that "the use of the term bootstrap derives from the phrase to pull oneself up by one's own bootstrap, widely thought to be based on one of the eighteenth century *Adventures of Baron Munchausen*, by Rudolph Erich Raspe. (The Baron had fallen to the bottom of a deep lake. Just when it looked like all was lost, he thought to pick himself up by his own bootstraps.)" Most recently Efron (2010: 180) has reinforced the notion that his bootstrap was directly inspired by the Raspe idiom: "I always liked the Baron Munchhausen stories."

One element of the bootstrap metaphor that Efron has *not* written on is the manifest *impossibility* of such self-propulsion. If impossibility is an active ingredient in Efron's

coinage then it is a tease coming from the promoter of a technique which will *not* work! It is hard to find any such note in the *OED* quotation from 1958, viz., "Some interesting techniques have been developed whereby it would be possible for a new computer group to 'boot-strap' itself into a position of automatic programming capability upon receiving a new machine." Defiance and/or self-mockery are rare in the formation of technical terms—there are no others in David's statistical list—but the contemptuous phrase *monstrous moonshine* (see Section 1 above) was coined by those *making* the conjectures to which the term refers; see Quinion (1998) and Miller (2011).

3 Bootstrapping into other languages

The world statistical community took to bootstrapping and the ISI's (2011) *Multilingual Glossary* and the bootstrap entries in the *Wikipedias* of different languages show how the terminology has travelled. The foreign language terms illustrate different borrowing strategies and, by re-opening the issue of how to name the construction, provide a commentary on Efron's original decision. David's list contains many eponymous terms, such as *Fisher information* and *Pearson curve*, but no translator of *bootstrapping* to have appears to have taken this option and called the method after its inventor. This would not have been a bad choice for Efron is so closely associated with the method that the term would work as an identifier; from Hall's pre-history (2003) of the method it is clear that no earlier statistician could contest ownership of the method.

The terms in David's list are so basic that all have foreign language forms. Pearson's classical neologisms, such as *kurtosis* and *heteroscedasticity*, were international

words and passed into other European languages without essential change. By contrast *bootstrap* is so embedded in an English idiom that it is almost untranslatable. Efron (2010: 180) reports, "I get complaints all the time about the name because it is not a story or a reference that is well known in the US, or outside the German and English worlds."

Statistical words from English have been borrowed into foreign languages in a number of ways and the classification in Durkin's *Guide to Etymology* (2009: 134ff) can be used to organise them; I have used the scheme in a study of English borrowing from German mathematics, Aldrich (2011c). A "loanword" for Durkin shows "borrowing of a word form and its associated meaning, or a component of its meaning." The loanword option is always available and, for the *bootstrap*, borrowing has generally taken this form: *méthode bootstrap*, *Bootstrap-Methoden*, *metodo bootstrap*, etc. The Afrikaans *skoenlusmetode* is the only example in the ISI list of a "loan translation," i.e. a construction that "shows replication of the structure of a foreign-language word or expression by the use of synonymous word forms in the borrowing language." The Spanish *método autosuficiente* may be considered a loan translation but one where the bootstrap element has been de-mythologised; Efron might have chosen such a colourless term in the first place. Two translations, the French *la méthode de Cyrano* (see Genest and Rivest (2008: 5)) and the Chinese 自助法, use local versions of the notion of impossible self-propulsion though neither equivalent quite enjoyed the currency of *bootstrapping* in American English. These constructions are perhaps dangerously over-stretched examples of a "semantic loan" which "shows extension of the meaning of a word as a result of association with the meaning of a partially synonymous word in another language." Because the English

statistical lexicon is rooted in ordinary English the main forms of borrowing are the loan translation and the semantic loan: more typical examples are *zulässige Entscheidungsfunktion*, *fonction de décision admissible*, etc. from (Wald's) *admissible decision function* and *Effizienz*, *efficacité*, etc. from (Fisher's) *efficiency*.

It is also interesting to see how the terms, *jackknife* and *to boot* (a computer), have fared in other languages. The latter has usually been translated: the official French word for *to boot* is *amorcer* while in German there is the formal *hochfahren* and the informal *booten*. The *jackknife* appears to have gone into most statistical languages as a loanword—even into British English for British boy scouts do not use jackknives. A few francophone statisticians use the exotic *la méthode d'Eustache* after Eustache Dubois who invented such a knife at the end of the seventeenth century but I do not know whether this and the companion *méthode de Cyrano* illustrate the French way with Anglo-Saxonisms or parody it.

4 Summary

The present account is an addition to the small literature on the history of specific statistical terms or of families of terms—see e.g. Aronson (2001) and Aldrich (2003). Looking beyond the *bootstrap*, it includes general information on word-formation in Statistics both in the bootstrap era and before and tries to show how the *bootstrap* was at once an exceptional term and a word of its time.

References

Aldrich, J. (2003) The Language of the English Biometric School, *International Statistical Review*, **71**, (1), 109-131.

_____ (2011a) *Figures from the History of Probability and Statistics*

http://www.economics.soton.ac.uk/staff/aldrich/Figures.htm

Accessed June 20, 2012.

_____ (2011b) *Mathematical Words: Origins and Sources*

http://www.economics.soton.ac.uk/staff/aldrich/Mathematical%20Words.htm

Accessed June 20, 2012.

_____ (2012) German Words in English Mathematics, unpublished manuscript.

Aronson, J. K. (2001) Francis Galton and the Invention of Terms for Quantiles, *Journal of Clinical Epidemiology*, **54**, 1191-1194.

Boos, D and L. Stefanski (2010) Efron's Bootstrap, *Significance*, **7**, (4), 186-188.

Conway, J. H. and S. P. Norton (1979) Monstrous Moonshine, *Bulletin of the London Mathematical Society*, **11**, (), 308-339.

David, H. A. (2001) First (?) Occurrence of Common Terms in Statistics and Probability in H. A. David & A. W. F. Edwards (eds.) (2001) *Annotated Readings in the History of Statistics*, New York: Springer.

Diaconis, P. and B. Efron (1983) Computer Intensive Methods in Statistics, *Scientific American,* **248**, (May), 96-108.

Dodge, Y. (ed.) (2006) *Oxford Dictionary of Statistical Terms*, Oxford: Oxford University Press.

Durkin, P. (2009) *The Oxford Guide to Etymology*, Oxford: Oxford University Press.

Efron, B. (1979a) Bootstrap Methods: Another Look at the Jackknife, *Annals of Statistic*s, **7**, (1), 1-26.

_____ (1979b) Computers and the Theory of Statistics: Thinking the Unthinkable, *SIAM Review*, **21**, (4), 460-480.

_____ (1982) *The Jackknife, the Bootstrap, and Other Resampling Plans*, Philadelphia: SIAM.

_____ (1987) Better Bootstrap Confidence Intervals (with discussion), *Journal of the American Statistical Association*, **82**, 171-200.

_____ (2000) The Bootstrap and Modern Statistics, *Journal of the American Statistical Association*, **95**, (Dec.), 1293-1296.

_____ (2003) Second Thoughts on the Bootstrap, *Statistical Science*, **18**, (2), Silver Anniversary of the Bootstrap, 135-140.

_____ (2010) A Life in Statistics: Conversation with Julian Champkin, *Significance*, **7**, (4), 186-188.

Efron, B., and G. Gong (1983) A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation, *American Statistician*, **37**, (1), 36-48.

Efron, B. and R. J. Tibshirani (1993) *An Introduction to the Bootstrap*, New York: Chapman & Hall.

Gell-Mann, M. (1995) *The Quark and the Jaguar: Adventures in the Simple and the Complex*, London: Abacus.

Genest, C. and L.-P. Rivest (2008) Comment compter les cerfs? *Accromath*, **3**, (Été-Automne), 2-7. http://accromath.ca/

Accessed June 20, 2012.

Hall, P. (2003) A Short Prehistory of the Bootstrap, *Statistical Science*, **18**, (2), 158-167.

Holmes, S., C. Morris, R. Tibshirani (2003) Bradley Efron: A Conversation with Good Friends, *Statistical Science*, **18**, (2), 268-281.

International Statistical Institute (2011) *ISI Multilingual Glossary of Statistical Terms* http://isi.cbs.nl/glossary.htm

Accessed June 20, 2012.

Mermin, N. D. (1981) E Pluribus Boojum: the Physicist as Neologist, *Physics Today*, **34**, (4), 46-53.

Miller, J. (ed.) (2011) *Earliest Known Uses of Some of the Words of Mathematics*, http://jeff560.tripod.com/mathword.html

Accessed June 20, 2012.

Oxford English Dictionary (2011) on-line up-dated version of the third, 1989, edition.

Accessed June 20, 2012.

Quenouille, M. (1949) Approximate Tests of Correlation in Time-series, *Journal of the Royal Statistical Society Series B*, **11**, (1), 68-84.

Quinion, M. (1998) *World Wide Words: Moonshine* http://www.worldwidewords.org/topicalwords/tw-moo1.htm

Accessed June 20, 2012.

_____ (2002) *World Wide Words: Boot* http://www.worldwidewords.org/qa/qa-boo2.htm

Accessed June 20, 2012.

Tukey, J. (1958) Bias and Confidence in Not-quite Large Samples (abstract), *Annals of Mathematical Statistics*, **29**, (2), 614.

Zimmer, Benjamin (2005) Figurative "bootstraps" (1834) ADS-L, American Dialect Society email discussion list http://listserv.linguistlist.org/cgi-bin/wa?A2=ads-l;xkNPcQ;200508111624160400b

Accessed June 20, 2012.