

"An almost innumerable number of theatre plays"

The Estimation of Editions on the Basis of Retrieved Copies: Printed Programmes of Jesuit Theatre Plays in the *Provincia Flandro-Belgica* (before 1773)

by Goran PROOT & Leo EGGHE¹

This contribution deals with the question how to estimate the number of ever-existed editions of printed texts on the basis of retrieved copies. This issue will be explained by means of a concrete case: the printed theatre programmes of the Jesuits in the *Provincia Flandro-Belgica* before the suppression of the order in 1773. In general terms, this article offers a mathematical method to estimate the total production of multi-copy artefacts on the basis of retrieved copies.

First, we will discuss the phenomenon of Jesuit drama and its significance within the innovative pedagogical system developed by the fathers shortly after the founding of the order. Then the value of the theatre programmes as a key source for the study of Jesuit drama in Europe and abroad will be explained. This will be followed by a description of the corpus of programmes from the *Provincia Flandro-Belgica* and its characteristics. After a brief evaluation of the classical way to estimate the production of theatre plays at the secondary Jesuit schools, we will introduce a mathematical model to cope with this problem. The conditions for the effectiveness as well as the limits of this method will be discussed from a practical point of view. We also show that this model is capable to (partially) explain the book historical law. Suggestions for related applications of this model within the field of bookhistorical research as well as other fields of interest will conclude this contribution.

I

Shortly after the foundation of the Societas Jesu in 1540, the new order unfolded an unseen activity in the field of the education of the youth.² Ignatius of Loyola (c. 1491-1556) considered not only the thorough education of its own members as a crucial instrument for the creation of a host of independent and reliable propagators of the faith, wherever the needs for it would send them.³ Soon it became clear that there was also a

¹ Leo Egghe is Chief Librarian of Hasselt University and part-time professor in informetrics and information retrieval in the Antwerp University. He has a Ph.D in mathematics and in information science. In 2001 he received the Derek De Solla Price Award in informetrics. He is the founder and Editor-in-Chief of the new *Journal of Informetrics*. Goran Proot is Keeper of the Historical Collections of the Antwerp University. He is bachelor in Philosophy and both master in Germanic Philology and Documentation and Library Sciences. Currently he is preparing a Ph.D on Jesuit theatre plays in the colleges of the *Provincia Flandro-Belgica* before the suppression of the Society of Jesus. In 2005, he got a Special Grant of the National Fund for Scientific Research in Flanders.

² A modern introduction into the Jesuits is Jonathan WRIGHT, *God's Soldiers. Adventure, Politics, Intrigue, and Power. A History of the Jesuits*, New York: Doubleday, 2004.

³ For the objectives of the Jesuits, cf. François COUREL, "La fin unique de la Compagnie de Jésus", in *Archivum Historicum Societatis Iesu*, 35 (1966), pp. 186-211; Saint IGNACE, *Constitutions de la Compagnie de Jésus. Traduction du texte officiel, notes et index par François Courel*, Paris: Desclée de Brouwer, 1976 (*Collection Christus* n. 23. Textes), p. 14; Günter SWITEK, "Die Eigenart der Gesellschaft Jesu im Vergleich zu den anderen Orden in der Sicht des Ignatius und seiner ersten Gefährten", in Michael

great potention in the founding of good secondary schools and education of the youth as a weapon agains moral deterioration and the protruding heretical ideas. The first Jesuit college for both members of the Society and lay students was openend in Gandía in 1546, due to its succes followed by many other colleges all over Europe and abroad.⁴ By the death of Ignatius in 1556, the Society of Jesus counted about 1,000 fathers and controlled fourty colleges, a number that was growing rapidly: in 1608 the Jesuits were over 10,000 members strong and led almost 300 colleges worldwide, numbers that were doubled in 1710.⁵ The enormous respons of the educational project of the fathers, that became within few years the most characteristic works of the order⁶, sprang of the specific method they developed and applied in their schools.⁷ It was amongst other inspired by the pedagogical conceptions of the Brethren of the Common Life and the method at the university of Paris.⁸ Typical for this so called *modus parisiensis* was the great stress on the practising and training of theoretical precepts. In this context the theatrical performance adopted its meaning: it was along with daily home work, the writing poetry, monthly *concertationes* and other exercices an excellent instrument to practice the *eloquentia*, put life into the language of Cicero and establish vivid examples of the christian virtue.⁹

II

From the very beginning of the first Jesuit schools, theatre plays were performed as an integral part of the pedagogical programme.¹⁰ Normally, there were at least two greater performances a year: one around the period of Lent, and one before the beginning of the next school-year that began October first (Saint Bavo). Very often, every of the five or – in the larger colleges– six classes prepared an additional performance during the year. A school-leaver who passed through the complete curriculum of the *humaniora* in a Jesuit

SIEVERNICH & Günter SWITEK, *Ignatianisch. Eigenart und Methode der Gesellschaft Jesu*, Freiburg: Herder, 1991 (second edition), pp. 204-232, here p. 204.

⁴ For the erection of the first colleges, cf. John W. O'MALLEY, *The First Jesuits*, Cambridge, 1994 (third edition), pp. 200-242; for the first college in Gandía, see p. 203. According to BATLLORI (p. 317) the college of Gandía was erected in 1545, cf. Miquel BATLLORI, "Der Beitrag der *Ratio studiorum* für die Bildung des modernen katholischen Bewußtseins", in Michael SIEVERNICH & Günter SWITEK, *Ignatianisch. Eigenart und Methode der Gesellschaft Jesu*, Freiburg: Herder, 1991 (second edition), pp. 314-322. The approbation for the college dates from 4 October 1547, cf. Charles Van de VORST, "Instructions pédagogiques de 1625 et 1647 pour les collèges de la province flandro-belge", in *Archivum Historicum Societatis Iesu* 19 (1950) [publ. in 1951], pp. 181-236, here p. 181.

⁵ Cf. Edmond LAMALLE, "Les catalogues des provinces et des domiciles de la Compagnie de Jésus", in *Archivum Historicum Societatis Iesu* 13 (1944), pp. 77-101;

⁶ Cf. Pietro LETURIA, "Perchè la Compagnia di Gesù divenne un Ordine insegnante", in *Gregorianum* 21 (1940), pp. 350-382, here p. 352.

⁷ The pedagogical plan, system and method of the Jesuits crystallised in the *Ratio studiorum* of 1599, that had a great influence on the development of the secondary schools all over the world up to now, cf. *Ratio Studiorum. Plan raisonné et institution des études dans la Compagnie de Jésus*, Édition bilingue latin-français. Présentée par Adrien Demoustier et Dominique Julia. Traduite par Léone Albrieux et Dolorès Pralon-Julia. Annotée et commentée par Marie-Madeleine Compère, Paris: Belin, 1997, and Vincent J. DUMINUCO (ed.), *The Jesuit Ratio Studiorum: 400th Anniversary Perspectives*, New York: Fordham University Press, 2000.

⁸ Cf. Gabriel CODINA MIR, *Aux sources de la pédagogie des jésuites. Le 'modus parisiensis'*, Roma, 1968 (*Bibliotheca instituti historici S.I.*, vol. 28).

⁹ Cf. Gabriel CODINA MIR, *Aux sources...*, p. 109.

¹⁰ Cf. Jean-Marie VALENTIN, *Theatrum catholicum. Les jésuites et la scène en Allemagne au XVIe et au XVIIe siècles*, Nancy: Presses universitaires de Nancy, 1990, p. 78.

college, might have seen some forty different plays. As most other means of distraction were forbidden for these pupils, the performances played a very important role in the formation of the students, as well in a cultural as in an ideological sense. Most of the shown themes were taken from the Bible, the sacred history, the own missionary activities, history or ancient literature. The type of theatre offered to the public can –with Valentin– best be characterized as *theatrum catholicum*, theatre with a clear religious and didactic goal, and in all it breathes the ideology of the contra-reformation.

The public didn't have to pay to attend the performances, but only had to come in time because the school-theatre was very popular. Alternatives were at least more expensive and sometimes highly questionable. At the end of the sixteenth century, the government in the Southern Netherlands considered the plays of the highly active and popular *retoriciens* as shady and took repeatedly measures to put them down.¹¹ The loyal Jesuits on the other hand obtained in some cities almost the role of master of ceremonies and that helped to put their stamp on the cultural urban life. They were prominent present on special occasions such as entries of the clerical and secular arm, festivities, processions, commemorations and so on. This enormous production of plays and spectacles of all kind is the subject of research of this phenomenon in as many countries the Jesuits were active, which is almost everywhere. The literary part of the Jesuit studies have therefore grown to an independent discipline.

Most of these spectacles are only documented by printed or sometimes handwritten programmes. Integral texts were hardly put in print and full-text manuscripts turn rarely up, too. Already in the nineteenth century bibliographers started to describe these valuable sources, very often from a corporational (De Backer-Sommervogel's *Bibliothèque de la Compagnie de Jésus*), a national or regional point of view.¹² Later on, play lists per college or Jesuit province were compiled.¹³ Titles from other sources such as the annual letters (*Litterae annuae*) of the Jesuits, the histories of the houses (*Historia domus*) and information in town accounts were added to the information offered by the programmes. Thus came about huge repertoria of many Jesuit colleges, on the basis of which emerged a torrent of studies.¹⁴ The printed programmes form the cornerstone of many of these contributions. In contrast to the rest of the material, they inform us about a great number of aspects of the performances in a coherent way. In addition to the title, the performers, the date, the hour, the location, the subject, the sources and the content of the play, the programmes very often mention the occasion, the Maecenas, where and by whom the programme was printed, a list of the roles and actors and the occurrence of

¹¹ Cf. Alfred PONCELET, *Histoire de la Compagnie de Jésus dans les anciens Pays-Bas. Établissement de la Compagnie de Jésus en Belgique et ses développements jusqu'à la fin du règne d'Albert et d'Isabelle*, Bruxelles, 1927 (2 vols.), II, p. 88.

¹² Cf. Augustin de BACKER & Carlos SOMMERVOGEL, *Bibliothèque de la Compagnie de Jésus*, Heverlee-Leuven, 1960 (12 vols., anastatical reprint of the edition 1890-1911. See also László POLGAR, *Bibliographie sur l'histoire de la Compagnie de Jésus 1901-1980*, Roma: Institutum Historicum S.I., 1981-1990 (3 parts in 6 vols.).

¹³ Most famous is Jean-Marie VALENTIN, *Le Théâtre des Jésuites dans les Pays de Langue Allemande. Répertoire chronologique des pièces représentées et des documents conservés (1555-1773)*. Stuttgart: Hiersemann, 1983-1984 (2 vols.).

¹⁴ For an overview of the academic production on Jesuit spectacles, see Nigel GRIFFIN, *Jesuit School Drama. A checklist of critical literature*, Londen, 1976, and Nigel GRIFFIN, *Jesuit School Drama. A checklist of critical literature. Supplement No. 1*, Londen, 1986.

entr'actes (comic plays that were interwoven within the drama), dances, ballets or musical interventions and so on. Other sources may sometimes go into detail about one or two aspects, such as the specific context of the performance or the reception of the play by the public, but remain silent on the rest of these elements.

III

The importance of the theatre programmes can hardly be overestimated. Therefore it is all the more remarkable that a thoroughly bookhistorical and critical description of this source remains in default. Many questions remain unanswered. Only to mention some of them: the greatest lot of the programmes don't mention an author. Many themes travelled around neglecting borders as these theatre was played in Neolatin, the international language of the humanist world. But how did that work, as complete texts are hardly to find and perhaps never conceived as a complete scenario? What were the selection criteria for the actors? In the lower classes, there were many times more students than parts in the play. And what about the comical *entr'actes*, the ballets and the dances? In Flanders, the programmes mention them from the second half of the seventeenth century on. That doesn't mean that they weren't shown before, as the same phenomenon can be observed with the indication of roles and actors. Seventeenth century programmes record them only very rarely, and on the contrary, they are consequently listed in the eighteenth century, which of course does not imply that there weren't any actors before.

From a materialistic point of view there are still many questions to be answered as well, such as how much did the programmes cost? Were they distributed for free and by whom? Who received them? How many programmes of one edition were produced?

The question we will deal in the rest of this article is how many different editions of programmes were printed. Furthermore, we will try to define the known part of this production. The question becomes extremely relevant considering the central place this source occupies in the study of Jesuit theatre. If we don't have a clear idea of the number of editions that have been produced, in fact we don't know on which part of the total production we rely for our research either, a thorny issue when one wants to draw conclusions.

IV

We will deal with this question by means of a concrete case: the printed programmes of plays performed by the students of the colleges in the *Provincia Flandro-Belgica* in pre-suppression time (1574-1773).¹⁵ In the beginning of the nineteenth century DE BACKER-SOMMERVOGEL 1890-1911 listed about 1070 spectacular performances under the direction of the Jesuits in the eighteen colleges of this province, that was slightly greater than present day Flanders, the northern, Dutch-speaking part of Belgium.¹⁶ For 768 of their descriptions, we can still retrieve the primary source, such as a full text or a

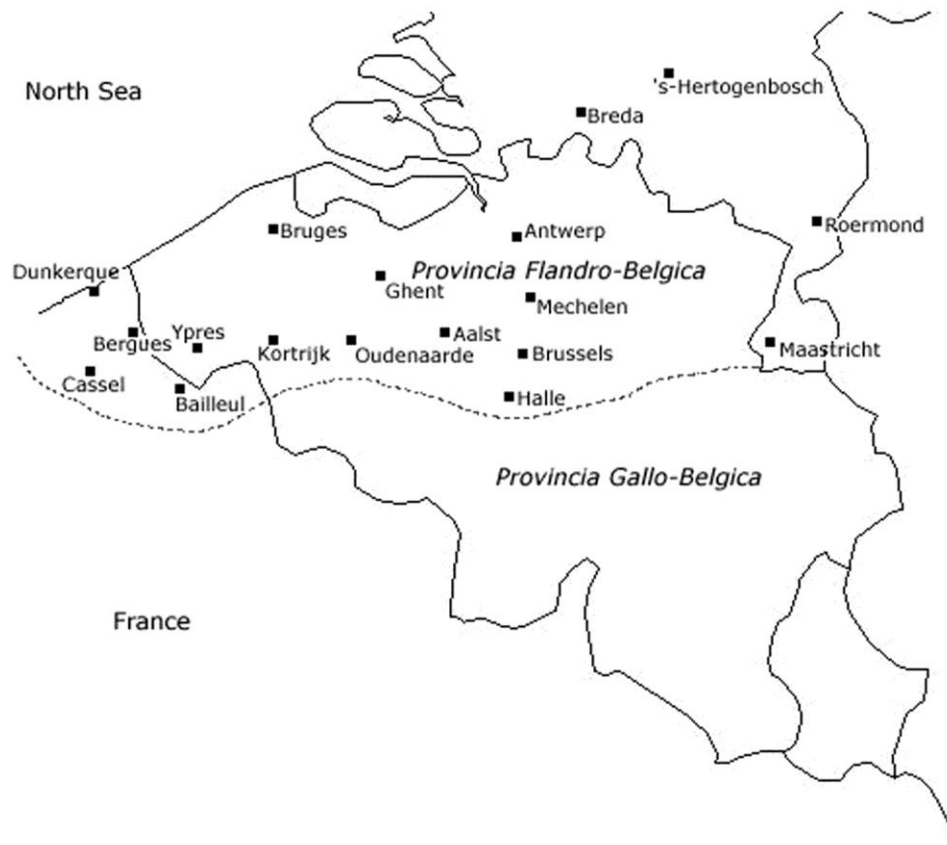
¹⁵ This province arose after the junction of the *Provincia Germania Inferior* in May 1612, cf. Alfred PONCELET, *Nécrologe des jésuites de la province Flandro-Belge*, Wetteren, 1931, p. XXVII-XXVIII. For practical reasons, we use the name of this administrative entity for the preceeding period (1574-1612) as well.

¹⁶ Cf. Augustin de BACKER & Carlos SOMMERVOGEL, *Bibliothèque...*, under the respective colleges.

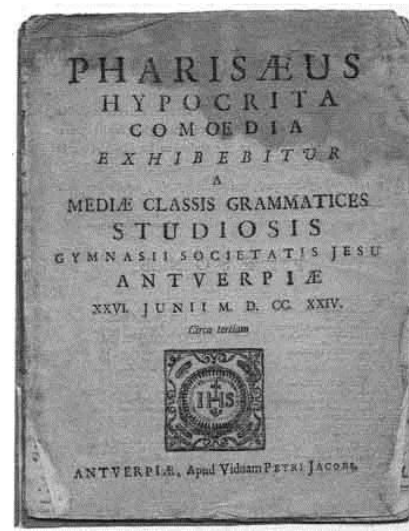
programme. For a great part of the other 302 descriptions in the *Bibliothèque de la Compagnie de Jésus*, the primary sources have meanwhile dissappeared. Two World Wars and other calamities are responsible for the loss of the largest part.¹⁷ Nevertheless, nine years of research in libraries and archives all over Europe have raised the number of performances up to 1907 descriptions (+837), two thirds of them (1276) are documented with a primary source.¹⁸ Amongst the category of primary sources, the programmes build the greatest group (95%); most of them appeared in print (86%). This collection contains as well theatre plays by pupils in the colleges, as by the members of sodalities (congregations of lay men), the catechism groups and members of the *convicts* (boarding schools), beside programme books manufactured for many special occasions. For reasons to be discussed later on, we will limit this group to the printed programme books of theatre plays performed by the pupils in eighteen colleges of the Jesuit *Provincia Flandro-Belgica*: Maastricht (founded in 1574), Antwerp (1575), Bruges (1575), Ieper (Ypres; 1585), Kortrijk (1587), Ghent (1592), Sint-Winoksbergen (*French*: Bergues; 1600), Brussels (1604), 's-Hertogenbosch (1610), Roermond (1611), Mechelen (1615), Oudenaarde (1616), Belle (*Fr.*: Bailleul; 1617), Kassel (*Fr.*: Cassel; 1617), Aalst (Alost; 1620), Duinkerke (*Fr.*: Dunkerque; 1620), Halle (1623) and Breda (1625). The selected editions have to mention the college, the year of production and must fit into the normal school-curriculum. Therefore, theatre programmes that mention special occasions or that were performed on Sundays or holidays are excluded.

¹⁷ In this context, the repetead destruction of the Leuven University Library is notorious, cf. Chris COPPENS, Mark DEREZ & Jan ROEGERS, *Universiteitsbibliotheek Leuven 1425-2000*, Leuven: Universitaire pers, 2005 (p. 135, 160 and 313). Many Jesuit programmes, recorded in the *Bibliothèque de la Compagnie de Jésus* disappeared. A binder of programmes was lost in the fire that devastated the City Library of Dunkerque in 1929 (cf. *Patrimoine des bibliothèques de France. Un guide des régions*, Paris: Payot, 1995 (10 vols. with index, vol. 11), II, pp. 112-113, and Goran PROOT, "Contribution au théâtre des jésuites flamands. Les pièces perdues, les titres retrouvés", in *Archief- en Bibliotheekwezen in België*, 59 (1998), nr. 1-4, pp. 111-171, here p. 111). Doubtless many programmes are lost by the town fires in Roermond in 1665 (cf. J.G.C. VENNER, "De stadsbranden in Roermond (1554 en 1665). Brand, overheidsmaatregelen en herstel", in G.H.A. VENNER, (red.) *Roermond stad met verleden. Negen hoofdstukken over Roermondse geschiedenis*, Roermond, 1985, pp. 116-141) en Bailleul on 8 Mai 1681 (cf. Jérôme FICHEROULLE, *Bailleul. Pages de l'histoire locale*, Bailleul, 1909 (deuxième édition), p. 447). In 1937, R. van Aerde consulted in the Library of the Hogeschool in Ghent a pile of unbound theatre programmes without number, cf. Raymond van AERDE, *Bijdrage tot de geschiedenis van het tooneel te Mechelen. Het schooldrama bij de jezuiteten*, Mechelen, [1937], p. 58. Also Van den Boogerd refers to them, but presumably he hasn't seen the programmes (he refers to an unknown number, cf. Leonardus van den BOOGERD, *Het jezuïetendrama in de Nederlanden*, Groningen, 1961, pp. 248-249). By that time, the programmes must already have been untraceable in 1961.

¹⁸ The inventar of these spectacles is forthcoming: Goran PROOT, *Spectacula Iesuitica Belgica Antiqua. Pars I: Provincia Flandro-Belgica*.



Thus we become a coherent corpus of 804 editions. Coherent not only what concerns the content of the objects, but also in its material appearance. All programmes are printed with black ink as quartos. Usually they count 2 folios (83%) or 4 folios (15%), only exceptionally the number of folios differs. From a typographical point of view, the programmes are printed following a more or less clear concept that varies only a little and rather slowly in function of the changing vogue. Very often these ephemere publications are brought together in binders containing larger numbers of similar documents. But there are also solitary copies kept in collections all over the world. Important numbers of programmes are found in different institutions in Brussels, Ghent, Kortrijk, Antwerp, Louvain, Bruges, Paris, Munich and Boston, but isolated copies are also retrieved in The Hague, London, Münster (Germany), Vienna and in private collections.



Title pages from programmes theatre plays by pupils of the Antwerp Jesuit college, 1615 and 1724 (copies: Brussels Royal Library, II 11797 nr. 21 (left) and Antwerp City Library, B 867119 Hs.)

It is clear that this corpus is only a part of the original corpus of theatre plays performed by the pupils in the colleges. The size of this sample can be calculated in different ways. The classical way to proceed is to multiply the number of colleges (i.e. 18) with the number of years each college has functioned. This number has to be multiplied for the number of classes installed in each school for normally every class produced a play of its own, raised with one for the concluding theatre play at the end of the school-year.¹⁹ For Brussels, this results in the following calculation (**table 1**):

Table 1. Brussels Jesuit college²⁰

<i>School-years</i>	<i>Number of years</i>	<i>Number of classes</i>	<i>Probable number of plays</i>
1604-1620	16	5	80 + 16
1620-1773	153	6	918 + 153
<i>Total</i>	<i>169</i>		<i>1167</i>

This exercise is quite simple for a well documented college as the *Gymnasium Regii* in the court city, but much more difficult for smaller schools such as that in Aalst, for which the data are very difficult to collect due to the changeable number of classes, periods of discontinuity and the lack of sources (**table 2**).²¹

¹⁹ Cf. Jean-Marie VALENTIN, *Theatrum catholicum*, p. 78.

²⁰ Cf. Michel HERMANS, "De la fondation à la suppression", in Bernard STENUIT (ed.), *Les collèges jésuites de Bruxelles. Histoire et pédagogie 1604, 1835, 1905, 2005*, Bruxelles: Éditions Lessius, 2005, blz. 53-83, here p. 75.

²¹ The example of Aalst is not chosen at random. It is one of the best studied colleges from the *Provincia Flandro-Belgica*, from which we also have retrieved a quite large number of programmes.

Table 2. Aalst Jesuit college²²

<i>School-years</i>	<i>Number of years</i>	<i>Number of classes</i>	<i>Probable number of plays</i>
1620-1621	1	1	1 + 1
1621-1622	1	2	2 + 1
1622-1623	1	3	3 + 1
1623-1624	1	4	4 + 1
1624-1625	1	5	5 + 1
1625-1628	3	4 or 5	12 + 3 or 15 + 3
1628-1630 ²³	2	3 (?)	6 (?) + 3
1630-1638	8	4 or 5	32 + 8 or 40 + 8
1638-1650 ²⁴	12	3 (?)	36 (?) + 12
1650-1651	1	5	5 + 1
1651-1657	6	5 (?)	30 (?) + 6
1657-1658	1	5	5 + 1
1658-1666	8	5 (?)	40 (?) + 8
1666-1668 ²⁵	2	0	0
1668-1674	6	5	30 + 6
1674-1675	1	3	3 + 1
1675-1730 ²⁶	55	3 (?)	165 (?) + 55
1730-1773	43	3	129 + 43
<i>Total</i>	<i>153</i>		<i>661 / 672</i>

Our selection contains only 54 printed editions for Brussels and 41 for Aalst. The estimated number of college plays is respectively about 21 and 16 times higher. All eighteen colleges together existed 2,478 years. This number multiplied with an average of six performances each year makes a total of 14,868. For reasons mentioned before, this estimation is not unproblematic.

It would be wrong to state that our selection of 804 programmes could serve as a representative sample of the total college theatre production. In the first place, that sample is not taken at random. Secondly, even when it could be considered as a representative sample, it would *not* be a sample of almost 15,000 plays. As will be stated below, the sources in our selection refer to a much more smaller group of theatre performances. The selected editions of programmes represent an unknown group of editions bearing the same characteristics as those selected: printed programmes of performances by students at Jesuit colleges in the *Provincia Flandro-Belgica*, that mention the college and the year of performance and that were played in the course of the school-year. When research is based on these documents, the results of that research refer to a greater group of similar documents. In second instance and with much more reservation, these documents refer to theatre performances. Not to *all* Jesuit theatre plays by students, but only these theatre plays for which the fathers had had programme books printed. As we already calculated, many more theatrical performances were held, for the greatest part of which sources are lacking. But for how many plays the Jesuits had programmes printed ?

²² Based on Jozef de BROUWER, *De jezuieten te Aalst*, Zingem: Genootschap voor Aalsterse Geschiedenis, 1979-1981 (3 vols), vol. 1.

²³ Much information is lacking, cf. Jozef de BROUWER, *De jezuieten te Aalst*, part 1, p. 49, p. 62.

²⁴ Much information is lacking, cf. Jozef de BROUWER, *De jezuieten te Aalst*, part 1, p. 49, p. 62.

²⁵ College closed, cf. Jozef de BROUWER, *De jezuieten te Aalst*, part 1, p. 49, p. 98.

²⁶ Much information is lacking, cf. Jozef de BROUWER, *De jezuieten te Aalst*, part 1, p. 64.

V

It is the basic challenge of this contribution to estimate the total number N of ever existed editions of theatre plays (hereafter called "pieces") by college students that were printed in the *Provincia Flandro-Belgica* (before 1773). This will be executed based on the number of retrieved printed copies. The model is applicable to different kinds of "multi-copy documents" but here we will keep the terminology "piece(s)" and "copies". Each "piece" refers to one edition of a programme, that is printed on several but identical "copies". Other potential applications will be discussed at the end of this contribution and in EGGHE & PROOT 2007, where the present model will be extended to more general applications.²⁷

The basic unknown variable is the value p ($0 < p < 1$) being the probability that a copy of a piece is lost. The variable p is unknown but fixed (i.e. non-time dependent), being the fraction of the copies (that were ever produced) that are lost, nowadays. Hence p expresses, intuitively, the negative effect of time which is responsible for the fact that, after a long time period (in this case more than 200 years) there are many lost copies of pieces, a fact that is historically very clear²⁸ but will also become clear in the sequel of this mathematical-probabilistic treatment. Note that p is the probability that a copy of a piece to be lost (not that a piece to be lost): this is logical: one can only lose the printed "touchable" copies of a piece; this in turn, might or might not lead to a loss of all copies of a piece in the end and hence of the piece itself (see further).

Another unknown variable is the number a of copies per piece. This variable will be used as an adaptable parameter and it will turn out that the exact knowledge of this parameter is not necessary for the estimation of the number of lost pieces (or the total number of pieces that ever existed). In EGGHE & PROOT 2007 we will present a method to determine a for cases where this is necessary but this is not the case here (see further) and hence this method will not be presented here. The used probabilistic methods are very elementary.²⁹ Given the abstract values of p and a , the probability (or fraction) P_0 to lose a piece with a copies (i.e. to have a piece where all a copies are lost) is equal to

$$P_0 = p^a \quad (1)$$

In other words, P_0 is the fraction (among all pieces) of lost pieces. This means that we work in the universe of "all pieces" (lost and not-lost, i.e. of which some $(1, 2, \dots, a)$ copies remain).

Among all pieces, the probability P_1 to have a piece of which exactly one copy is left (and hence $a-1$ copies are lost) is given by

²⁷ Cf. Leo EGGHE & Goran PROOT, "The estimation of the number of lost multi-copy documents: a new type of informetrics theory", in *Journal of Informetrics* 1 (2007) (forthcoming).

²⁸ Important reasons are the low value of the programmes, their material vulnerability –counting only one or two uncovered folded sheets–, their ephemere function, i.e. to document a very temporal performance. As some surviving copies prove, paper was very scarce and therefore often reused to make notes on it or to pack things in it. Copies that were meant to be kept had to survive more than ten generations, wars, water, fire and all kinds of vermin.

²⁹ They can be found in G.R. GRIMMETT & D.R. STIRZAKER, *Probability and random processes*, Oxford: Clarendon Press, 1985, and G.C. CANAVOS, *Applied probability and statistical methods*, Boston: Little, Brown and Company, 1984.

$$P_1 = ap^{a-1}(1-p) \quad (2)$$

Indeed, supposing independency in (not) losing copies, the probability to lose $a-1$ copies is p^{a-1} and $1-p$ is the probability *not* to lose one copy. This has to be multiplied by a since the not-lost copy can be situated at places 1, 2, ..., a (i.e. at any of the a copies of a piece). Among all pieces, the probability P_2 to have a piece of which exactly two copies are left (and hence $a-2$ copies are lost), is given by

$$\begin{aligned} P_2 &= \frac{a!}{2!(a-2)!} p^{a-2} (1-p)^2 \\ &= \frac{a(a-1)}{2} p^{a-2} (1-p)^2 \end{aligned} \quad (3)$$

Indeed, again supposing independency in (not) losing copies, the probability to lose $a-2$ copies is p^{a-2} and $(1-p)^2$ is the probability *not* to lose 2 copies. This has to be multiplied by the combination of a over 2, i.e. the number of ways a places can be combined in sets of 2:

$$\begin{aligned} \frac{a!}{2!(a-2)!} &= \frac{a!}{2!(a-2)!} \\ &= \frac{a(a-1)(a-2)...3.2.1}{2.1(a-2)(a-3)...3.2.1} \\ &= \frac{a(a-1)}{2} \end{aligned} \quad (4)$$

In general, for each $j = 0, 1, 2, \dots, a-1, a$, the probability P_j to have a piece of which j copies are left (and hence of which $a-j$ copies are lost) is equal to

$$P_j = \frac{a!}{j!(a-j)!} p^{a-j} (1-p)^j \quad (5)$$

where

$$\frac{a!}{j!(a-j)!} = \frac{a!}{j!(a-j)!} \quad (6)$$

Indeed, again supposing independency in (not) losing copies, the probability to lose $a-j$ copies is p^{a-j} and $(1-p)^j$ is the probability *not* to lose j copies. This has to be multiplied by

the combination of a over j , i.e. the number of ways a places can be combined in sets of j , being (6).

This case covers all cases $P_0, P_1, \dots, P_a$, the latter one being

$$\begin{aligned} P_a &= \sum_{j=0}^a \frac{a!}{j!(a-j)!} (1-p)^j p^{a-j} \\ &= (1-p)^a \end{aligned} \quad (7)$$

So we have described all fractions in the universe of all pieces of which now $0, 1, 2, \dots, a-1$, a copies are left. Note the necessary requirement

$$\begin{aligned} \sum_{j=0}^a P_j &= \sum_{j=0}^a \frac{a!}{j!(a-j)!} p^j (1-p)^{a-j} \\ &= (p + (1-p))^a = 1 \end{aligned} \quad (8)$$

hence, hereby covering all possible cases.

It follows from (2) and (3) that

$$\frac{P_2}{P_1} = \frac{a-1}{2} \cdot \frac{1-p}{p} \quad (9)$$

yielding for the unknown p :

$$\begin{aligned} \frac{1-p}{p} &= \frac{2P_2}{(a-1)P_1} \\ 1-p &= \frac{2P_2}{(a-1)P_1} p \\ p \left(1 + \frac{2P_2}{(a-1)P_1} \right) &= 1 \\ p &= \frac{1}{1 + \frac{2P_2}{(a-1)P_1}} \end{aligned} \quad (10)$$

If a is fixed and known (which is not the case, but see further), then (10) gives us an estimation of p since the value of $\frac{P_2}{P_1}$ is known. Indeed, P_1 and P_2 are not known, but $\frac{P_2}{P_1}$ represents the fraction of pieces of which 2 copies are left divided by the fraction of pieces of which 1 copy is left. This is equal to

$$\frac{P_2}{P_1} = \frac{NP_2}{NP_1} \quad (11)$$

being the total number of pieces of which 2 copies are left, divided by the number of pieces of which 1 copy is left and these numbers are known; the concrete example of Jesuit theatre programmes will be given in Section VI.

Formula (10) is the key to find the number of lost pieces. First, the fraction of lost pieces, P_0 is, using (1) and (10)

$$P_0 = \frac{1}{1 + \frac{2P_2}{(a-1)P_1}} \quad (12)$$

To determine the actual (estimated) number of lost pieces we base ourselves on P_0 , $1-P_0$ and the actual total number or not-lost pieces (i.e. found pieces), denoted N_f . Using the "rule of three", we estimate the total number N_l of lost pieces by

$$N_l = N_f \frac{P_0}{1 - P_0} \quad (13)$$

This follows clearly from the obvious relations:

$$NP_0 = N_l \quad (14)$$

$$N(1 - P_0) = N_f \quad (15)$$

(the total number of lost pieces (N_l) equals its fraction multiplied by the total number of pieces (N) and the total number of found pieces (N_f) equals its fraction (being $1 - P_0$) multiplied by the total number of pieces (N)).

Note that, for (13) (as well as for (11)) we do not need the value of N of the total number of pieces that ever existed (since N cancels in the division of (14) and (15) and in (11)). In fact, from (13) we even have an estimate for this total number N :

$$N = N_f + N_l \quad (16)$$

which is obvious.

VI

In this section we will apply this theoretical model to the case of Jesuit theatre programmes. Our sample contains the following numbers concerning pieces (editions of programmes documenting Jesuit theatre plays) and their copies (**table 3**):

Table 3. Distribution of editions and copies (Jesuit theatre programmes)

<i>Number of editions</i>		<i>Number of copies</i>
714	pieces are found with	1 copy
82	pieces are found with	2 copies
4	pieces are found with	3 copies
3	pieces are found with	4 copies
1	piece is found with	5 copies
no	pieces are found with	6 or more copies

This gives a total of $N_f = 804$ found pieces or editions.

Note that in the above theory we only need to know $\frac{P_2}{P_1}$ which is, by (11), equal to

$$\frac{P_2}{P_1} = \frac{82}{714} \quad (17)$$

It is clear that the other (small) values are much more unstable.

It is now clear that formula (12) for P_0 can be applied (and consequent formulae (13) and (16)) if we know a . In the eighteenth century, rhetoriciens and other companies in Flanders ordered runs between 100 and 3,200 copies from one edition.³⁰ Between 1761 and 1769, the Ghent printer Petrus Franciscus de Goesin delivered 79 differend orders with an average run of 979 copies. From bills of two other Ghent printers, Bauduin Manilius and Hendrik Saetreuver, who continued the printing house, we got information on the size of the runs ordered by the Ghent Jesuits and the fathers from Aalst for their theatre programmes at the end of the seventeenth and the beginning of the eighteenth century.³¹ For the small town of Aalst, the lowest order counted 150 copies of one edition. In Ghent, a run could count up to 950 copies, sometimes broken down into two

³⁰ Cf. Ferd(inand) VANDERHAEGHEN, *Bibliographie Gantoise. Recherches sur la vie et les travaux des imprimeurs de Gand*, Gent: Impr. Eug. Vanderhaeghen, 1858-1869 (7 vols.), vol. 3 (1864), blz. 406, nrs. 4148-4152 (1761), nrs. 4176-4182 (1762), nrs. 4208-4219 (1763), nrs. 4248-4255 (1764), nrs. 4282-4286 (1765), nrs. 4315-4327 (1766), nrs. 4360-4368 (1767), nrs. 4399-4402 (1768), nrs. 4425-4432 (1769).

³¹ Cf. *Comptes des impressions et fournitures faites à Bauduin Manilius et Henri Saetreuver* (Ghent University Library, G 13367, nrs. 102, 111, 154, 174 and 194), two bills in a binder of Jesuit theatre plays (Ghent University Library, G 6145/1-43 (nrs. [0] and nr. 1)), and *Handboek van Hendrik Saetreuver* (Ghent University Library, Hs G 13160, unnumbered, two bills). On Hendrik Saetreuver, see Ferd(inand) VANDERHAEGHEN, *Bibliographie Gantoise...*, vol. 2, p. 320; Sylvie de SMET, *Sociaal-economische analyse van een beroepsgroep: de Gentse boekdrukkers (1450-1700)*, Gent, 1999 (2 vols., unpublished master's thesis), vol. 1, pp. 165-172 and vol. 2, pp. 196-207; Koen de VLIETGER-DE WILDE (ed.), in cooperation with Joost DEPUYDT, Goran PROOT & Stijn van ROSSEM, *Directory of seventeenth-century Printers, Publishers and Booksellers in Flanders = Adresboek van zeventiende-eeuwse drukkers, uitgevers en boekverkopers in Vlaanderen*, Antwerp, 2004 (Uitgaven van de Vereniging der Antwerpse Bibliotheek Nieuwe reeks, 1), p. 121; and the on line database of the STCV (<http://www.stcv.be> under 'database').

or more variant editions of one programme. The bill of 6 September 1686 mentions the expenditure of 16 schillings ("schellingen") for the printing of 800 engravings. This was enough to illustrate the greatest part of the three different editions that were prepared: one edition of 350 Latin programmes of 2 folios, 271 Latin programmes of 4 folios and one Dutch edition of the same text in 2 folios, counting 300 copies.³² 121 copies remained that year thus unillustrated.

The complexity of the historical practice, that is in most cases subject of speculation, leaves us in doubt what to do with a , the parameter that represents the number of copies of a piece. For a small Jesuit college, runs between 150 and 200 copies were common, for larger colleges as that one in Ghent, numbers between 600 and 950 could be of application. However, we are very lucky here: it turns out that formula (12) is extremally stable (constant) in a .

Indeed, using (17), we have for $a = 150$:

$$P_0 = \frac{1}{1 + \frac{2}{149} \cdot \frac{82}{714}} = 0.7936955$$

i.e. 79.4% of all pieces is lost.

For $a = 200$ we have

$$P_0 = \frac{1}{1 + \frac{2}{199} \cdot \frac{82}{714}} = 0.7939673$$

still yielding 79.4% of all pieces being lost.

Even for $a = 750$ we have

$$P_0 = \frac{1}{1 + \frac{2}{749} \cdot \frac{82}{714}} = 0.7945627$$

or 79.5% of all pieces being lost.

³² Cf. Ghent University Library, G 13367 nr. 102.

In short, we can say that function P_0 (formula (12)), for values of a in the above ranges, is practically a constant function of a , a very remarkable finding! The value of a can be determined (using also the more unstable quotient $\frac{P_3}{P_2}$) but, as indicated here, there is no need for it in this case since the printed editions of Jesuit programmes are always printed in relatively high runs ($a \geq 150$ in any case).³³

Finally we adopt formula (13) to determine the (estimated) actual number of lost editions of Jesuit theatre programmes.

$$N_1 = N_f \cdot \frac{P_0}{1 - P_0}$$

As we have $N_f = 804$ (see table 3), and $P_0 \approx 0.794$, hence we estimate the total number of lost pieces as

$$N_1 \gg 804 \cdot \frac{0.794}{0.206} \text{ pieces}$$

$$N_1 \gg 3,099 \text{ pieces} \quad (18)$$

The total number of pieces that was produced is hence estimated (by (16)) as

$$N = N_f + N_1$$

$$N \gg 804 + 3,099 \text{ pieces}$$

$$N \gg 3,903 \text{ pieces} \quad (19)$$

VII

How sure can we be of this value for N ? We leave open the mathematical solution to this problem, but we have dealt with this question from a deductive and practical point of view.

We explored the reliability of the mathematical model as follows. In a database, we created 10 different fictitious corpora, each corpus containing a different number of editions (from 1,000 until 10,000). Every edition is present in 150 copies ($a = 150$). Every copy of each edition is represented by one record in the database. The largest corpus of 10,000 edition counts therefore 1,500,000 unique records, the smallest one 150,000 records or "copies of editions". Firstly, we had the computer pick out $n = 1,000$ records (or copies) at random from every corpus. Then the sample was analysed: how many

³³ Cases with an unknown value for a are dealt with in Leo EGGHE & Goran PROOT, "The estimation of the number of lost multi-copy documents: a new type of informetrics theory", in *Journal of Informetrics* 1 (2007) (forthcoming).

editions were represented by one copy, how many by two copies and so on (see **table 4**). This exercise provided us with the values for P_1 and P_2 , needed in formula (12).

Table 4. Distribution of editions and copies (corpus = 10,000 editions, $a = 150$, $n = 1,000$)

	<i>Number of editions</i>		<i>Number of copies</i>
P_1	906	pieces are found with	1 copy
P_2	44	pieces are found with	2 copies
P_3	2	pieces are found with	3 copies
$P_4, \dots, P_a$	no	pieces are found with	4 or more copies

The total number of found editions is 952. These numbers result for P_0 in

$$P_0 = \frac{1}{1 + \frac{2}{149} \cdot \frac{44}{906}} = 0.90867535$$

Using this result in formula (13) and in formula (16) results in an estimated total number (N) of 10,223 editions. As the corpus from which the sample is taken counts 10,000 editions, this sample gives a very precise idea of its size.

For each corpus, we repeated this exercise 30 times and took 30 samples in order to get an estimation of the precision of the estimations in practice. The results of the tests are presented in **table 5**.

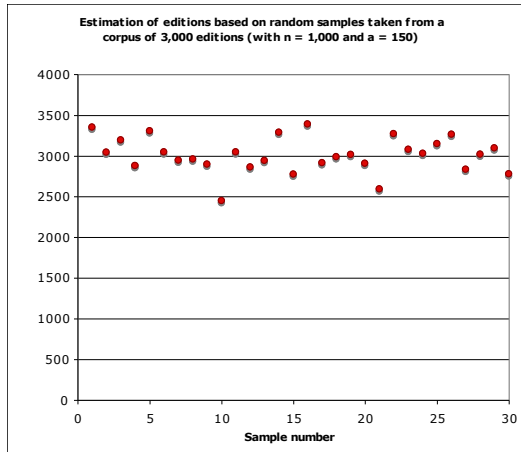
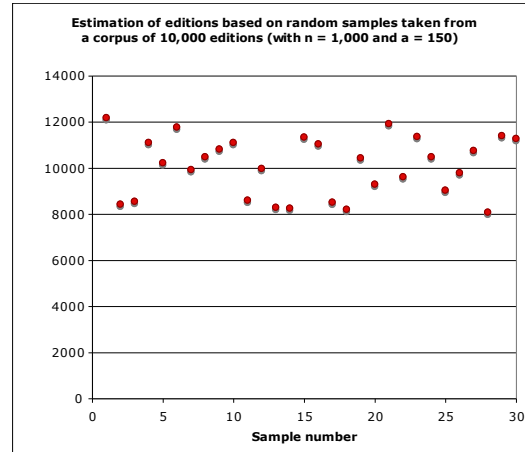
Table 5. Estimation of N based on 30 random samples with $n = 1,000$ and $a = 150$

<i>Size corpus</i>	<i>Average result for N</i>	<i>Min. value for N</i>	<i>Max. value for N</i>	<i>Standard deviation (%)</i>
10,000	10,073	8,074	12,178	12.7%
9,000	9,007 (8,801)	6,437 (6,437)	14,978 (11,096)	17.4% (12,5%)*
8,000	7,890	6,520	11,278	12.3%
7,000	6,947	5,070	9,099	10.8%
6,000	6,126	5,238	8,117	10.5%
5,000	5,102	4,165	6,485	10.4%
4,000	4,059	3,183	5,456	11.6%
3,000	3,010	2,449	3,389	7.1%
2,000	2,011	1,796	2,425	7.1%
1,000	998	873	1,106	5.6%

* These values are due to one sample. The numbers between brackets give the results when this sample is omitted.

The size of the samples being constant ($n = 1,000$), it is obvious that they are more exact, the more the size of the corpus decreases. For corpora between 1,000 and 3,000 editions, the standard deviation is not higher than 7.1%, for larger corpora the average estimated size of the corpus deviates not more than 12.7%.³⁴ Of course, these results are average results of each time 30 samples. **Graphs 1 and 2** show the individual results of each sample for a corpus of 3,000 editions (left) and 10,000 editions.

³⁴ Only one sample out of 300 gave a totally wrong image of the estimated size, see **table 5**. When we abstract of that sample, we get a normal series for the corpus containing 9,000 editions.

**Graph 1.** Corpus of 3,000 editions**Graph 2.** Corpus of 10,000 editions

Overall, the results of the individual estimations are very satisfactory.

In practice, it is most of the time impossible to control the size of the sample. In general, scholars try to take the sample as large as possible. It is useful to know when the size of a sample is sufficient. We tested the variation of the size of the sample in relation to a constant corpus of 4,000 editions each consisting of 150 copies.

Table 6. Estimation of N based on samples with a decreasing size (corpus of 4,000 editions, a = 150)

Size Corpus	Size sample	Average result for N	Standard deviation (%)
4,000	1,000	4,059	11.6%
4,000	900	4,022	9.7%
4,000	800	3,915	14.9%
4,000	700	3,902	11.0%
4,000	600	4,059	16.1%
4,000	500	4,254	19.9%
4,000	400	3,829	20.1%
4,000	300	4,470	30.0%
4,000	200	4,842	70.4%
4,000	100*	-	-

*Samples with $n = 100$ result sometimes in $P_2 = 0$, where the model requires a value $P_2 > 0$.

The greater the sample sizes, the more exact the estimations. Table 6 shows that the exactness of random samples between 700 and 1,000 is quite high: the standard deviation is between 9.7 and 14.9%. Decreasing the sample to 600, 500 or 400 has already tangible consequences: the estimated size of the corpus shows an average mistake of ca. 20%, which is –under circumstances– acceptable. Smaller samples result in unreliable ($200 < n < 300$) or even unworkable ($n = 100$) estimations.

These simulations provide us with a guideline for the interpretation of the result based on our sample of Jesuit theatre programmes. The estimation of the total production of theatre programmes bearing the same characteristics as the programmes in the sample, led to the number of $\approx 3,903$ editions (cf. (19)). Given the fact that the sample consisted of 907 single copies ($714 + 2 \cdot 82 + 3 \cdot 4 + 4 \cdot 3 + 5$ copies, cf. **table 3**), this estimation is probably

$\approx 90\%$ exact. Therefore it is safe to state that the total production of similar printed programmes of Jesuit theatre in the *Provincia Flandro-Belgica* amounted between ≈ 3513 and ≈ 4293 editions. From a methodological point of view, this information is of high importance. It means that, if we analyse the theatrical production of the Jesuits in Flanders before 1773 on the basis of retrieved programmes, we may extrapolate the results to a group of pieces five times larger than the retrieved number. On the other hand, that also means that there is a very large part of the presumed theatrical production of which we do not have a source under the form of a theatre programme. For the largest part of the Jesuit theatre plays, ca. 15,000 productions, we must rely on scarcely available secondary sources, such as town bills, printer's bills, indication in *Historia domus* or *Litterae annuae* and the like. The fundamental question is, whether the ca. 9,000 theatre productions for which presumably no programmes were printed, can be compared to the ones that are documented with a printed programme.

VIII

The above model, elaborated in Section VI, also presents a partial rationale for the “book historical law”. It is not easy to determine the original formulation of this law and its author. We believe (advised by Jan Bos to whom our sincerest thanks) that the oldest formulation of this “law” appeared in the article of Willard (1943)³⁵ and reads as follows: “Since copies from the largest editions are now generally much less common than those from smaller editions, one may conclude, somewhat paradoxically, that the less there are, the more there were: that books were destroyed primarily by the diligence of their original readers” (p. 172-173). Also Salman (1997)³⁶ refers to this “law” as follows (translated from the Dutch language): “There exists a book historical law saying that the probability to save (keep) a copy of a printed piece is reversely proportional to the size of the edition”. In this connection, Salman (e-mail communication) also refers to van Selm, but without providing a concrete reference. Finally, we also found a reference to Salman (1997) in van Rossem (2005)³⁷ in the connection of this book historical law.

It is clear, as also stated in the above references, that not only the size of the edition is determining the probability for a copy to be saved. Also its price and even more its temporary value (e.g. almanacs, newspapers,...) determines this probability.

Yet, in this article, we can give a partial rationale for this book historical law, when we only consider the size of the edition (as it is formulated originally).

Since in this article the parameter p stands for the probability to lose a copy and since the number a stands for the size of the edition we will reformulate the book historical law as

³⁵ Oliver M. Willard. The survival of English books printed before 1640: a theory and some illustrations. The Library, Fourth Series, vol. 23, 171-190, 1943.

³⁶ Jeroen Salman. Een handdruk van de tijd. De almanak en het dagelijks leven in de Nederlanden 1500-1700. Stedelijk Museum Het Prinsenhof, Delft, Historisch Museum De Waag, Deventer, Provinciale Bibliotheek fan Fryslân, Leeuwarden. Waanders Uitgevers, Zwolle, 1997.

³⁷ Stijn van Rossem. Drukkersbelangen en politiek verlangen. Hieronymus II Verdussen en de controle op de productie van almanakken in Antwerpen, 1626-1642. De Gulden Passer, 83, 175-200, 2005.

“The probability p to lose a copy of a printed edition in an increasing function of the size a of the edition”.

We think that this is the clearest formulation that is possible: it is equivalent with saying that the probability to keep a copy (i.e. $1-p$ in our notation) is a decreasing function of the size a of the edition and this comprises cases as formulated above: “inversely proportional to the size (a) of an edition” which is, mathematically, not completely clear.

Based on our model, developed in Section V, we can give a mathematical rationale for this law as follows. The relation between p and a is not a mathematical function but a mathematical relation. This means that one a -value can be linked with several p -values.

This is seen in equation (10): denoting $x = \frac{P_2}{P_1}$ we have

$$p = \frac{1}{1 + \frac{2}{a-1}x} \quad (20)$$

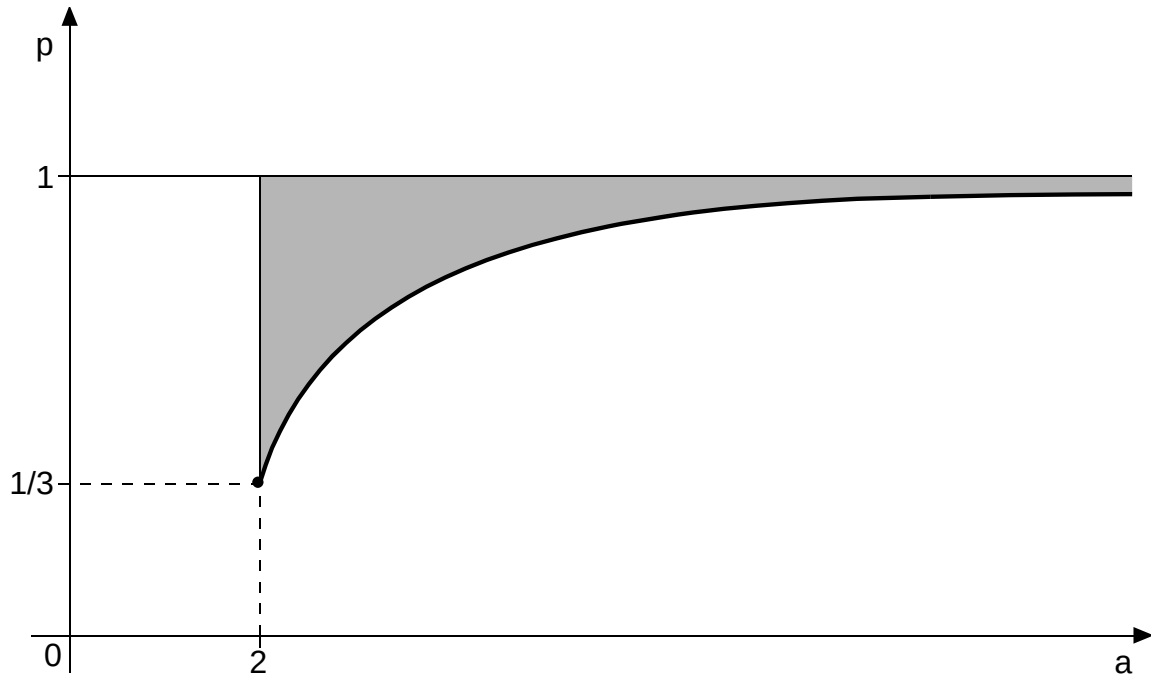
It is clear that $x < 1$ (see the above example: equation (17) shows that $x < 1$) and hence, since $x \in]0, 1[$ we have that

$$\frac{1}{1 + \frac{2}{a-1}} \leq p \leq 1 \quad (21)$$

Evidently, p has an upper bound 1 and a lower bound

$$f(a) = \frac{1}{1 + \frac{2}{a-1}} \quad (22)$$

which is a concavely increasing function since $f'(a) > 0$, $f''(a) < 0$ for all $a \geq 2$, the absolute lower bound for a (the size of the edition) in order to have an applicable model (we need P_2 in Section V implying $a \geq 2$). Formula (21) and (22) imply that the relation between p and a is as in graph 3.



Graph 3. The relation between p (the probability to lose a copy) and a (the size of the edition) is given by the shaded area.

The interpretation of Fig. 1 is as follows. For low values of a (which is not the case in the example of Jesuit theatre plays studied here) we have that p can range in (maximally) the interval $\frac{1}{3}, 1$, but for larger values of a we see that p can only be large (close to 1). So, the higher a , the more limited is the range in which p can vary and the higher (i.e. close to 1) this range is situated.

We stress that this does not completely explain the book historical law. It certainly does not explain the human (evident) attitude to lose copies that have only a temporary value (such as almanacs or newspapers) or (correlated) that have only a low money value. We only noticed that, when data show that $\frac{P_2}{P_1} < 1$ (or in general that $\frac{P_2}{P_1}$ has an upper bound), then a high value of a (the size of the edition) forces p (the probability to lose a copy) to be high. The intuition for this is as follows: a high value of a implies that it will be difficult to have pieces with a low number of copies unless p is very high (close to 1). Now $\frac{P_2}{P_1} < 1$ expresses that we have (relatively) more pieces with 1 copy than with 2 copies, which can only be understood (when a is large) if p is large.

We further underline that any explanation of the book historical law, of a different nature than above, is strengthened by the regularity shown in Fig. 1.

IX

The mathematical method presented in this article has proven to be applicable in practice. Under conditions, it is fairly possible to estimate the total number of ever produced pieces on the basis of a sample of retrieved copies of these pieces. The reliability of the results depends both on the presumed size of the total corpus and the size of the sample. The exact number of produced copies for each piece is less important: the model is very stable for printed media, were the number of produced copies of an edition is normally high enough ($a > 150$).

In general, the way in which the data are collected is highly important for the results of the research. Obviously, the coherence of the sample has its repercussions on the size of the estimated number of pieces that were produced. The more data that are filtered out – because they are possibly corrupt or doubtful – the smaller the estimation of the total corpus will be. On the other hand, the obtained results are much more reliable to work with.

The model presented in this contribution lends itself to many other applications than the single one explored here. It can be used to estimate the total number of different multi-copy artefacts over time, from which a critical number of copies has been retrieved. In a bookhistorical context, the number of editions within genres can be measured. Reliable bibliographies of all kind can serve as a main point to research the production of popular literature, theatre plays, bishop books, etc. These surveys may lead to objective and reliable figures on the total book production in regions and countries. Samples may be taken from electronic databases as the STCN and the STCV and alike to estimate resp. the Dutch and the Flemish book production. Also other objects of art that were produced in limited series are qualified, such as engravings, etchings, maps, stamps, furniture, design, coins, and the more. The field of interest seem to be innumerable...