

Type/Token-Taken informetrics

Non Peer-reviewed author version

EGGHE, Leo (2003) Type/Token-Taken informetrics. In: Journal of the American Society for Information Science and Technology, 54(7). p. 603-610.

DOI: 10.1002/asi.10247

Handle: <http://hdl.handle.net/1942/761>

TYPE/TOKEN-TAKEN INFORMETRICS

by

L. Egghe, LUC, Universitaire Campus, B-3590 Diepenbeek, Belgium¹

and

UIA, Universiteitsplein 1, B-2610 Wilrijk, Belgium

ABSTRACT

Type/Token-Taken informetrics is a new part of informetrics that studies the use of items rather than the items itself. Here items are the objects that are produced by the sources (e.g. journals producing articles, authors producing papers etc.). In linguistics a source is also called a type (e.g. a word) and an item a token (e.g. the use of words in texts). In informetrics, types that occur often e.g. in a database will also be requested often e.g. in information retrieval. The relative use of these occurrences will be higher than their relative occurrences itself, hence the name Type/Token-Taken informetrics.

This paper studies the frequency distribution of Type/Token-Taken informetrics, starting from the one of Type/Token informetrics (i.e. source-item relationships). We are also studying the average number μ^* of item uses in Type/Token-Taken informetrics and compare this with the classical average number μ in Type/Token informetrics. We show that $\mu^* \geq \mu$ always and that μ^* is an increasing function of μ . A method is presented to actually calculate μ^* from μ and a given α which is the exponent in Lotka's frequency distribution of Type/Token informetrics. We leave open the problem of developing non-Lotkaian Type/Token-Taken informetrics.

Key words and phrases : Type/Token, Type/Token-Taken, Lotka, frequency distribution, average.

¹ Permanent address.

I. Introduction

The dual approach of informetrics is very well-known nowadays and forms the basis of the classical informetric laws. Since we will rely heavily on this dual approach, we will briefly present the main aspects of it. We can start with a bibliography consisting of journals that contain articles on a certain topic. We could also say that journals are sources that "produce" articles, generally called items. Using this terminology we go back to the original formulation by Bradford (Bradford (1934) but see also Egghe and Rousseau (1990) for a more complete treatment of dual informetrics). We are hence dealing with what can be called source/item informetrics. Indeed, journals/articles are not the only interpretation of the concept source/item. Lotka (1926) described the author/paper relationships leading to his celebrated frequency law

$$f(n) = \frac{D}{n^\alpha} \quad (1)$$

where $f(n)$ denotes the number of authors with n papers, $n=1,2,3,\dots$. Most classically, $\alpha=2$ but, in general, α can take "any" value superior to 1.

In this article we restrict ourselves to the Lotka-type frequency law (1), although some results (such as Theorem III.1) are more generally valid. The reason is that the new informetrics model (Type/Token-Taken) is best (first) applied to the simplest possible functions f (for reasons of calculation). The power law (1) clearly is a very simple function, certainly when compared to other frequency functions that have been in use in informetrics the last decade (see e.g. Egghe and Rousseau (1990)). It is a widely used law and has historical roots. Furthermore it is the unique function with the scale-free property expressing that, if n is multiplied with a constant (say C , another "scale" for the variable n), we still have, up to a constant, the same function f as in (1). Finally, as shown in many publications, Lotka's law is the basis for derived informetric results giving also an idea what these derived results could be in case we have other frequency functions for which the similar results are too intricate to be calculated. In this way, Lotka's law can be considered as the basis for informetrics with non-Lotkaian frequency functions f (or at least as a first

approximation). Also for comparison reasons of Type/Token informetrics with Type/Token-Taken informetrics it is necessary that a simple concrete function (as the power function) is used.

Other examples of source/item informetrics are

- Key words (amongst which search keys or N-grams) and their occurrence in a database (see e.g. Egghe (1992,2000), Kilgour, Long and Leiderman (1970)),
- Articles and the references they give (synchronous citation analysis) or the citations they receive (diachronous citation analysis) (see e.g. Stinson (1981) and Stinson and Lancaster (1987)),
- Books in a library and their borrowings in a certain time period (see e.g. Burrell and Cane (1982)),
- Services (e.g. at a library circulation or information desk) and their duration (measured e.g. in seconds).
- Words and their occurrences in a text (e.g. a book) (see e.g. Zipf (1949), Herdan (1960,1964)).

For more source/item examples - even beyond the field of informetrics - we refer the reader to Egghe and Rousseau (1990).

The last example is found in quantitative linguistics. Sources (words) are there called "types" and the occurrences of these words are called "tokens" (see e.g. Herdan (1960)). They hence talk about Type/Token linguistics. Generalizing, all the above examples of dual informetrics (i.e. source/item relations) could be called Type/Token informetrics. We will use this terminology henceforth since it will be more handy to describe the purpose of this paper : the introduction of what we can call Type/Token-Taken informetrics.

The problem with Type/Token (T/T) informetrics is that we only describe sources and the items in these sources. In other words : sources and their production. T/T informetrics (important in itself of course) does not study the use of these items, hence these items as they are encountered by information professionals or other information users. Let us give some examples.

- Search keys (e.g. N-grams) are used by cataloguers in a library network to quickly find out if a newly arrived book (hence that must be catalogued) is already in the database. For a description of search keys (or N-grams as a special case) see e.g. Egghe (1992, 2000), Kilgour, Long and Leiderman (1970). The distribution of the occurrences (tokens) of the N-grams (types) in the catalogue will be different from the distribution of the occurrences of these N-grams as experienced by the cataloguer. Indeed the more an N-gram occurs in the catalogue, the more this N-gram will also be used by the cataloguer. This will lead to larger (this will be shown in the sequel) average screen lengths than what would be expected from the occurrences of the N-grams in the database itself. Cataloguers are using ("taking") the occurrences (token) of the N-grams (type), hence this is an example of a Type/Token-Taken (T/TT) informetrics.
- Generalizing the above example : every user seeking information from a database, using a key-word, will use more a "popular" (i.e. heavily used) key-word than a less popular one and these popular key-words give rise to more hits in the database. Averaged over the users, the number of hits per search will be longer (again this will be proved in the sequel) than the average number of documents containing this key-word (and, of course, the same for their distributions from which these averages are derived). This is, hence, another example of T/TT informetrics.
- Some books in a library are borrowed heavily but many books are borrowed seldomly. But a library user will seek these popular books more frequently than the other ones. So, on the average, the user will encounter more borrowed cases than the average number of borrowings per book. Also this is an example of T/TT informetrics.
- Service times at a desk. Since long service times have a higher probability to be encountered by a person seeking service, an average such person will experience longer service times (i.e. waiting times) than what could be expected from the average service time itself. Again a T/TT informetrics example.

It is clear that the reader can add several other examples of the difference between T/T informetrics and T/TT informetrics. Surprisingly, T/TT informetrics has never been

studied. The only attempts this author could find in the literature is the description of Rousseau (1990,1992) of positive reinforcement (indeed also the effect of use of items (there called : availability)) but only in the context of its effect on Lorenz curves (i.e. concentration theory) and not in the general informetrics laws context (as we will do in this paper).

T/TT informetrics studies the laws of informetrics as experienced by a user and is in this sense more important than T/T informetrics.

In the next section the T/TT frequency distribution will be derived from the one in T/T informetrics. The third section proves results on the average of these distributions (μ^* for T/TT and μ for T/T), namely that always $\mu^* \geq \mu$ and that μ^* is an increasing function of μ . In fact, using Lotka's law (1) in T/T informetrics we will show that μ^* is a function of μ and α , increasing in μ (α fixed) and we conjecture that μ^* is an increasing function of α (μ fixed).

It is not easy to determine the function $\mu^* = \mu^*(\alpha, \mu)$. In the fourth section a (numeric) method to do so is explained and an explicite table of $\mu^*(\alpha, \mu)$ is presented. We leave open for further study non-Lotkaian T/TT informetric developments (hence replacing α by another (other) parameter (s)).

II. Type/Token-Taken informetrics

We start from Lotkaian T/T informetrics where we adopt the continuous setting as was also done in Egghe (1989,1990). We denote by A the total number of items and by T the total number of sources. Classical Lotkaian T/T informetrics is given by the law of Lotka (continuous setting) : for $\alpha > 1$

$$f(j) = \frac{D}{j^\alpha} \quad (2)$$

for $j \in [1, \rho(A)]$. Here $f(j)$ denotes the density of the number of sources (types) with a density of j items (tokens) (compare with the discrete setting (1)). Note that, by definition,

$$\int_1^{\rho(A)} f(j) dj = T \quad (3)$$

and

$$\int_1^{\rho(A)} j f(j) dj = A. \quad (4)$$

This is the only formalism of the dual theory of informetrics as developed in Egghe (1989,1990) that we need. The function f describes the occurrence of items in sources, otherwise stated, the Type-Token relationship, hence T/T informetrics. Note that the average number of items per source (tokens per type) is given by

$$\mu = \frac{A}{T} = \frac{\int_1^{\rho(A)} j f(j) dj}{\int_1^{\rho(A)} f(j) dj}. \quad (5)$$

Type/Token-Taken informetrics is obtained by replacing the function f by the function g :

$$g(j) = j f(j) \quad (6)$$

$j \in [1, \rho(A)]$ describing token densities instead of type densities : $g(j)$ denotes the density of the tokens in the types with token-density j . In the discrete setting this would mean (replacing j by $n \in \mathbb{N}$) : the fraction of the items that belong to sources with n items.

Rationale of formula (6)

Formula (6) is basic for Type/Token-Taken informetrics: it is the defining function and requires explanation. We go back, to fix the ideas, to the example (given in the previous section) on search keys as used by cataloguers. Of course this explanation can be given for any application. So suppose that f is the frequency law of search keys in a catalogue, i.e.

$f(n)$ is the number of search keys that occur $n \in \mathbb{N}$ times in the catalogue. Now, the cataloguer used these search keys and here we come on the third level (the "Taken" level). Each time the cataloguer wants to make a catalogue description of a new book the search key of this book is checked against the already existing catalogue (expressed in terms of search keys) i.e. the items are used here as sources (hence the name "taken"): in the catalogue we have that $nf(n)$ is the number of used keys that appear n times. In other words our third level informetrics has the function g (formula (6)) as frequency function and not f .

The average use μ^* is obtained as follows : denote by

$$P^*(j) = \frac{g(j)}{\int_1^{\rho(A)} g(k) dk} = \frac{jf(j)}{\int_1^{\rho(A)} kf(k) dk} \quad (7)$$

the probability of these tokens (in the types with token-density j). So, such a type has a token-density of j . Hence the average number μ^* of used tokens per type equals

$$\mu^* = \int_1^{\rho(A)} j P^*(j) dj \quad (8)$$

which hence leads to the formula

$$\mu^* = \frac{\int_1^{\rho(A)} j g(j) dj}{\int_1^{\rho(A)} g(j) dj} = \frac{\int_1^{\rho(A)} j^2 f(j) dj}{\int_1^{\rho(A)} j f(j) dj}. \quad (9)$$

Formulae (6), (7) and (9) form the basis for Type/Token-Taken informetrics (T/TT informetrics). It measures the observed average number of items per source by users, rather than the average number μ of items per source in the sense of occurrence.

We will close this section by giving concrete formulae for the underlying distributions of f and g . Since (3) implies (since $\alpha > 1$)

$$\int_1^{\rho(A)} f(j) dj = \frac{D}{\alpha-1} \left(1 - \frac{1}{\rho(A)^{\alpha-1}} \right) = T \quad (10)$$

we have that the probability distribution of f is given by

$$P(j) = \frac{f(j)}{\int_1^{\rho(A)} f(k) dk} = \frac{1-\alpha}{\rho(A)^{1-\alpha}-1} \frac{1}{j^\alpha} . \quad (11)$$

Since

$$A = \int_1^{\rho(A)} g(j) dj = \int_1^{\rho(A)} j f(j) dj = \frac{D}{2-\alpha} (\rho(A)^{2-\alpha} - 1) \quad (12a)$$

if $\alpha \neq 2$ and

$$A = \int_1^{\rho(A)} g(j) dj = \int_1^{\rho(A)} j f(j) dj = D \ln \rho(A) \quad (12b)$$

if $\alpha=2$ we have for the probability distribution P^* of g the following formula (by (7)) :

if $\alpha \neq 2$:

$$P^*(j) = \frac{\frac{D}{j^{\alpha-1}}}{\frac{D}{2-\alpha} (\rho(A)^{2-\alpha} - 1)}$$

and if $\alpha=2$

$$P^*(j) = \frac{\frac{D}{j^{\alpha-1}}}{D \ln \rho(A)} .$$

Hence,

$$P^*(j) = \frac{2-\alpha}{\rho(A)^{2-\alpha}-1} \frac{1}{j^{\alpha-1}} \quad (13a)$$

if $\alpha \neq 2$ and

$$P^*(j) = \frac{\ln \rho(A)}{j} \quad (13b)$$

if $\alpha=2$, for all $\alpha > 1$.

T/T informetrics describes the source-item relationship as such while T/TT informetrics describes the source-item relationship as it is experienced by users (information professionals as well as information seekers). It is clear that a further elaboration of the above model is important. In the next section we will investigate relations between μ^* and μ . The fourth section then shows how μ^* can be calculated from μ (numerically).

III. Relations between μ and μ^*

In this section, T/T and T/TT informetrics will be compared using the important T/T and T/TT averages μ and μ^* respectively. We have the following results.

Theorem III.1 : $\mu^* \geq \mu$ in all cases.

Proof : One proof can be given by expressing that the variance of the distribution P (equation (11)) is positive :

$$\int_1^{\rho(A)} (j - \mu)^2 P(j) dj \geq 0 .$$

Another simple proof, suggested by one of the referees, goes as follows: express that the next quadratic form (in x) is always positive:

$$\begin{aligned} 0 &\leq \int_1^{\rho(A)} (xj + 1)^2 f(j) dj \\ &= x^2 \int_1^{\rho(A)} j^2 f(j) dj + 2x \int_1^{\rho(A)} j f(j) dj + T, \end{aligned} \quad (14)$$

hence the discriminant of (14) must be negative, yielding that $\mu^* \geq \mu$. \square

The result of Theorem III.1 can be evaluated in informetrics terms as follows (cf. the examples in the previous section) : the average number of items per source, as observed by the user is larger than the real existing average number of items per source. Following the

given examples, the cataloguer will find larger screen lengths (on average) than the average number in the database (a negative constatation). The library user, seeking information from a database, using a key word will find more hits than could be expected from the database itself (according to the IR evaluation criteria "recall" and "precision" - see e.g. Salton and Mc Gill (1987) - this could be interpreted as a positive resp. negative statement). The other two examples given in section II obviously give a negative interpretation : the user will encounter more borrowed books (hence unavailable) than could be expected and the same goes for the waiting times at a service desk !

We will now continue with the evaluation of the exact values of μ and μ^* . From (10) and (12) we have

$$\mu = \frac{A}{T} = \frac{1-\alpha}{2-\alpha} \frac{\rho(A)^{2-\alpha}-1}{\rho(A)^{1-\alpha}-1} \quad (15)$$

if $\alpha \neq 2$ (as always we take $\alpha > 1$). If $\alpha = 2$ we have

$$\mu = \frac{A}{T} = \frac{\ln \rho(A)}{1 - \frac{1}{\rho(A)}}. \quad (16)$$

For μ^* we evaluate :

$$\int_1^{\rho(A)} j^2 f(j) dj = \int_1^{\rho(A)} \frac{D}{j^{\alpha-2}} dj. \quad (17)$$

Suppose $\alpha \neq 2$, $\alpha \neq 3$. Then (17) equals

$$\int_1^{\rho(A)} j^2 f(j) dj = \frac{D}{3-\alpha} (\rho(A)^{3-\alpha} - 1). \quad (18)$$

If $\alpha = 2$ we have

$$\int_1^{\rho(A)} j^2 f(j) dj = D(\rho(A) - 1) \quad (19)$$

and if $\alpha=3$ we have

$$\int_1^{\rho(A)} j^2 f(j) dj = D \ln \rho(A). \quad (20)$$

This gives for μ^* : if $\alpha \neq 2, \alpha \neq 3$

$$\mu^* = \frac{2-\alpha}{3-\alpha} \cdot \frac{\rho(A)^{3-\alpha}-1}{\rho(A)^{2-\alpha}-1} \quad (21)$$

(by (9), (12) and (18)). If $\alpha=2$ we find (using now also (19)) :

$$\mu^* = \frac{\rho(A)-1}{\ln \rho(A)} \quad (22)$$

and, finally, if $\alpha=3$ (using (20)) we have

$$\mu^* = \frac{\ln \rho(A)}{1 - \frac{1}{\rho(A)}}. \quad (23)$$

From the formulae for μ and μ^* we see that a direct relation between μ and μ^* is not derivable (since (15) cannot be solved exactly for $\rho(A)$ - it can be done numerically and this will be the topic of the next section, giving concrete values of μ^* in function of μ and α).

However, from the above, we are able to prove the following result :

Theorem III.2 : For each α fixed we have that μ^* is an increasing function of μ .

The proof is given in the Appendix.

As will become clear from the next section we can conjecture the following, but we are not able to find a proof for it.

Conjecture III.3 : For each fixed μ we have that μ^* is an increasing function of α .

We also state the following

Problem III.4 : In the above T/TT informetrics theory we used Lotka's law as the basic frequency law. Develop a T/TT informetrics theory for other acceptable frequency distributions (continuous or not) such as the Poisson distribution, the Negative Binomial distribution, the Lognormal distribution,...

IV. Concrete calculations of the T/TT average μ^* in function of the T/T average μ and Lotka's α

The calculation of μ^* in function of μ and α requires the solution of the following equation in $x = \rho(A)$ - see (15) : if $\alpha \neq 2$

$$x^{2-\alpha} \left(\frac{1}{\mu} \frac{\alpha-1}{\alpha-2} \right) - x^{1-\alpha} + 1 - \frac{1}{\mu} \frac{\alpha-1}{\alpha-2} = 0 \quad (24)$$

and if $\alpha = 2$ (see (16))

$$\ln x + \frac{\mu}{x} - \mu = 0 . \quad (25)$$

Once x is found we have then μ^* using the formulae (cf. (21), (22), (23)) : if $\alpha \neq 2$, $\alpha \neq 3$

$$\mu^* = \frac{2-\alpha}{3-\alpha} \frac{x^{3-\alpha}-1}{x^{2-\alpha}-1} \quad (26)$$

and if $\alpha = 2$

$$\mu^* = \frac{x-1}{\ln x} \quad (27)$$

and if $\alpha = 3$

$$\mu^* = \frac{\ln x}{1 - \frac{1}{x}} . \quad (28)$$

From (24), (25) it is clear that determining a direct functional relationship $\mu^* = \mu^*(\mu, \alpha)$ is not possible : given α and μ , we can only solve (24), (25) in a numerical way. This we have done, using the MATHCAD 4.0 software package. One important remark must be made : α and μ cannot be taken fully independent from each other : there is a lot of freedom but the following restrictions must be respected :

$$\begin{aligned}
 A &= \int_1^{\rho(A)} j f(j) dj \\
 &= \int_1^{\rho(A)} \frac{D}{j^{\alpha-1}} dj \\
 &< \int_1^{\infty} \frac{D}{j^{\alpha-1}} dj .
 \end{aligned} \tag{29}$$

The last integral equals $D \frac{j^{2-\alpha}}{2-\alpha} \Big|_1^{\infty}$ if $\alpha \neq 2$ and $D \ln j \Big|_1^{\infty}$ if $\alpha = 2$. Hence we have ∞ if $\alpha \leq 2$ and $\frac{D}{\alpha-2}$ if $\alpha > 2$.

If we approximate T by (since $\alpha > 1$ and $\rho(A)$ is large)

$$\begin{aligned}
 T &= \int_1^{\rho(A)} f(j) dj \\
 &\approx \int_1^{\infty} f(j) dj \\
 &= \frac{D}{\alpha-1}
 \end{aligned} \tag{30}$$

we find, combining (29) and (30) that there are no restrictions for μ , given $\alpha \geq 2$ but that, if $\alpha < 2$ we must have that

$$A < T \frac{\alpha-1}{\alpha-2} ,$$

hence

$$\mu < \frac{\alpha-1}{\alpha-2}. \quad (31)$$

For the table of μ^* in function of μ and α (which will be given in the sequel) we hence have the following restrictions

$$\alpha = 2.5 \Rightarrow \mu < 3$$

$$\alpha = 3 \Rightarrow \mu < 2$$

$$\alpha = 3.5 \Rightarrow \mu < 1.6667 .$$

This explains the empty parts in the following table.

The following table is obtained by solving (24), (25) for x and then putting x in (26), (27), (28).

Table : Values of μ^* in function of μ and α

	$\alpha=1.5$	$\alpha=2$	$\alpha=2.5$	$\alpha=3$	$\alpha=3.5$
$\mu=1.2$	1.2148	1.2151	1.2157	1.2164	1.2178
$\mu=1.5$	1.5838	1.5984	1.6177	1.6479	1.7010
$\mu=2$	2.3317	2.4612	2.7320	-	-
$\mu=2.5$	3.2492	3.7279	5.8541	-	-
$\mu=3$	4.3333	5.5996	-	-	-
$\mu=3.5$	5.5894	8.3955	-	-	-

Note that $\mu^* > \mu$ and that μ^* increases in μ and α . Note also that the difference $\mu^* - \mu$ increases with μ . Other values of μ^* in function of given μ s and α s can be calculated similarly.

References

- S.C. Bradford (1934). Sources of information on specific subjects. *Engineering*, 137, 85-86. Reprinted in : *Collection Management*, 1, 95-103 (1976-1977). Also reprinted in : *Journal of Information Science*, 10, 148 (facsimile of the first page) and 176-180 (1985).
- Q.L. Burrell and V.R. Cane (1982). The analysis of library data (with discussion). *Journal of the Royal Statistical Society A*, 145, 439-471, 1982.
- L. Egghe (1989). The Duality of informetric Systems with Applications to the empirical Laws. Ph. D. Thesis, The City University, London (UK).
- L. Egghe (1990). The duality of informetric systems with applications to the empirical laws. *Journal of Information Science*, 16(1), 17-27, 1990.
- L. Egghe (1992). Theory of search keys and applications to retrieval techniques used by catalogers. *Mathematical and Computer Modelling*, 16(4), 69-90, 1992.
- L. Egghe (2000). The distribution of N-grams. *Scientometrics*, 47(2), 237-252, 2000.
- L. Egghe and R. Rousseau (1990). Introduction to Informetrics. *Quantitative Methods in Library, Documentation and Information Science*. Elsevier, Amsterdam, 1990.
- G. Herdan (1960). Type-Token Mathematics. A Textbook of mathematical Linguistics, Mouton, 's Gravenhage, 1960.
- G. Herdan (1964). *Quantitative Linguistics*, Butterworths, London, 1964.
- F.G. Kilgour, P.L. Long and E.B. Leiderman (1970). Retrieval of bibliographic entries from a name-title catalog by use of truncated search keys. *Proceedings of the American Society for Information Science*, 7, 79-82, 1970.
- A.J. Lotka (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16, 317-323, 1926.
- R. Rousseau (1990). A bibliometric study of Nieuwenhuysen's bibliography of microcomputer software for online information and documentation work. *Journal of Information Science*, 16, 45-50, 1990.
- R. Rousseau (1992). Concentration and diversity of availability and use in information systems : a positive reinforcement model. *Journal of the American Society for Information Science*, 43(5), 391-395, 1992.

- G. Salton and M.J. Mc Gill (1987). Introduction to modern Information Retrieval. McGraw-Hill, New York, 1987.
- E.R. Stinson (1981). Diachronous versus synchronous Study of Obsolescence. Ph. D. Thesis. University of Illinois, 1981.
- E.R. Stinson and F.W. Lancaster (1987). Synchronous versus diachronous methods in the measurement of obsolescence by citation studies. Journal of Information Science, 13, 65-74, 1987.
- G.K. Zipf (1949). Human Behavior and the Principle of Least Effort. Addison-Wesley, Cambridge, 1949. Reprinted: Hafner, New York, 1965.

Appendix

Proof of Theorem III.2 : Denoting $x=\rho(A)>1$ in (15) and (21) we have the following equations :

$$\frac{x^{2-\alpha}-1}{x^{1-\alpha}-1} \frac{1}{\mu} \frac{\alpha-1}{\alpha-2} = 1 \quad (\alpha \neq 2) \quad (A.1)$$

$$\mu^* = \frac{2-\alpha}{3-\alpha} \frac{x^{3-\alpha}-1}{x^{2-\alpha}-1} \quad (\alpha \neq 2, \alpha \neq 3) \quad (A.2)$$

Let $1 < \alpha < 2$

Then μ increasing implies $\frac{1}{\mu} \frac{\alpha-1}{\alpha-2}$ increasing since $\frac{\alpha-1}{\alpha-2} < 0$. Hence, by (A.1),

$$\frac{x^{2-\alpha}-1}{x^{1-\alpha}-1}$$

decreases. By the Lemma (below) (since $1-\alpha$ and $2-\alpha$ have opposite sign) we have that x increases. Again by the same Lemma we have that

$$\frac{x^{3-\alpha}-1}{x^{2-\alpha}-1}$$

increases (since $3-\alpha$ and $2-\alpha$ have the same sign), hence, by (A.2), μ^* increases.

$2 < \alpha < 3$

On the same lines as the above proof, using the same Lemma and (A.1) and (A.2) we now have

$$\mu \nearrow \Rightarrow \frac{1}{\mu} \frac{\alpha-1}{\alpha-2} \searrow \Rightarrow \frac{x^{2-\alpha}-1}{x^{1-\alpha}-1} \nearrow \Rightarrow x \nearrow \Rightarrow \frac{x^{3-\alpha}-1}{x^{2-\alpha}-1} \searrow \Rightarrow \mu^* \nearrow.$$

$\alpha > 3$

Now

$$\mu \nearrow \Rightarrow \frac{1}{\mu} \frac{\alpha-1}{\alpha-2} \searrow \Rightarrow \frac{x^{2-\alpha}-1}{x^{1-\alpha}-1} \nearrow \Rightarrow x \nearrow \Rightarrow \frac{x^{3-\alpha}-1}{x^{2-\alpha}-1} \nearrow \Rightarrow \mu^* \nearrow.$$

$\alpha = 2$

From (16) and (22) we have

$$\mu = \frac{\ln x}{1 - \frac{1}{x}} \tag{A.3}$$

$$\mu^* = \frac{x-1}{\ln x} \tag{A.4}$$

$$\frac{d\mu}{dx} = \frac{\frac{1}{x} - \frac{1}{x^2}(1+\ln x)}{\left(1 - \frac{1}{x}\right)^2} > 0$$

since $\ln x < x-1$ always (since $e^x > x+1$, for all $x > 1$). Also

$$\frac{d\mu^*}{dx} = \frac{\ln x - 1 + \frac{1}{x}}{(\ln x)^2} > 0$$

since $\ln x > 1 - \frac{1}{x}$. This follows from the fact that the function $\varphi(x) = \ln x + \frac{1}{x} - 1$ is minimal in $x=1$ and $\varphi(1)=0$. Hence $\mu \nearrow \Rightarrow \mu^* \nearrow$.

$\alpha=3$

From (15) and (23) we have

$$\mu = 2 \frac{x}{x+1} \quad (A.5)$$

$$\mu^* = \frac{\ln x}{1 - \frac{1}{x}}. \quad (A.6)$$

Obviously μ increases in x and by the above ((A.3)) μ^* increases in x . So $\mu \nearrow \Rightarrow \mu^* \nearrow$.

□

Lemma : The function

$$f(x) = \frac{x^{a+1} - 1}{x^a - 1} \quad (A.7)$$

increases if a and $a+1$ have the same sign and decreases if a and $a+1$ have opposite sign.

Proof :

$$f'(x) = \frac{x^{a-1}(x^{a+1} - ax - x + a)}{(x^a - 1)^2}$$

So $f'(x)$ has the same sign as

$$\varphi(x) = x^{a+1} - ax - x + a.$$

Now

$$\varphi'(x) = (a+1)x^a - (a+1)$$

which is zero in $x=1$. Furthermore

$$\varphi''(x) = a(a+1)x^{a-1},$$

so

$$\varphi''(1) = a(a+1).$$

So if $a(a+1) < 0$ then φ has a maximum in 1 and $\varphi(1)=0$. So $\varphi(x) < 0$ for all x and hence f decreases. If $a(a+1) > 0$ then φ has a minimum in 1. So $\varphi(x) > 0$ for all x and hence f increases. \square