UNIVERSITA' DEGLI STUDI DI SIENA

DIPARTIMENTO DI STUDI AZIENDALI E SOCIALI
SEZIONE: "POPOLAZIONE, AMBIENTE E METODI MATEMATICO-STATISTICI"

Luciano Petrioli

RELATIONAL METHODS IN DEMOGRAPHY

Quaderno n.1 Novembre 1996

RELATIONAL METHODS IN DEMOGRAPHY

Luciano Petrioli

Over the last two centuries, the demographic history of the human population has been characterized by several stages in which natural growth factors like birth and death have clearly shown, even if in different ways, a downward trend.

In the more advanced societies, there has been a generalized decrease of the probability of dying among the young and the elderly with a consequent increase in the average length of life, while, on the other hand, the modification of the age-specific fertility rate values has affected total reproduction at the end of fertility life.

Although, not perfectly appreciable as means of common biometric indicators, the structural variations of survival and fertility may be measured by relational methods which represent a very important instrument for demographic research particularly in developing countries.

With regard to this subject, we will consider the main concepts which make up the basis of such methods, paying particular attention to the problems which arise in the study of fertility.

1. SURVIVAL

1.1 If we have a phenomenon described by the function $f(x)=a+bx$ and we want to compare aspects indicated by $f_1(x)= a_1+b_1x$ and $f_2(x) = a_2+b_2x$, the linking parameters will results from the solution of the system made up by the two equations just seen, that is:

$$\text{where} \quad A = a_1 - a_2 \frac{b_1}{b_2} \; ; \; B = \frac{b_1}{b_2} \qquad [1]$$

or by placing $f_1(x)$ as a function of $f_2(x)$ and looking for the values of A and B from the relation:

$$f_1(x) = A + B\, f_2(x)$$

The function f(x) may be supplied by analysing the phenomenon, that is by directly considering the distinguishing data (as for example in the case of mortality there may be survivors, the probability of dying, the force of mortality, etc.) or their transformations.

1.2 As far as mortality is concerned, W.Brass (1971) has demonstrated a procedure which allows two life-tables to be linked by two parameters. In substance, considering the life-table p(x) and supposing a "logistic" trend for the function of death distribution, that is:

$$q(x) = \frac{1}{1 + a\,e^{-bx}} \qquad [2]$$

we can write, for example:

$$\ln \frac{q(x)}{p(x)} = -\ln a + bx \qquad [3]$$

also called "logit" of the probability of dying. Given two life-tables $p_1(x)$ and $p_2(x)$, we have:

$$\text{logit } q_1(x) = A + B \;\text{logit } q_2(x) \qquad [4]$$

1.3 A relation between the life-tables may be set out by means of the Gompertz function (Petrioli, 1975), according to which the force of mortality may be represented by the relation:

$$\mu(x) = b\,c^{x} \qquad [5]$$

from which $\ln \mu(x) = \ln b + x \ln c$ is obtained [6]

The link between the two life-tables for which:

$$\begin{aligned} \ln \mu_1(x) &= \ln b_1 + x \ln c_1 \\ \ln \mu_2(x) &= \ln b_2 + x \ln c_2 \end{aligned} \qquad [7]$$

will then be given by:

$$\ln \mu_1(x) = A + B \ln \mu_2(x) \qquad [8]$$

having placed: $A = \ln b_1 - (\ln c_1 . \ln b_2)/ \ln c_2$; $B = \ln c_1 / \ln c_2$

Tab.1 Male life-table, approximated instantaneous rate of mortality $\mu(x)$ and logit of q(x). Italy, males, 1950-53.

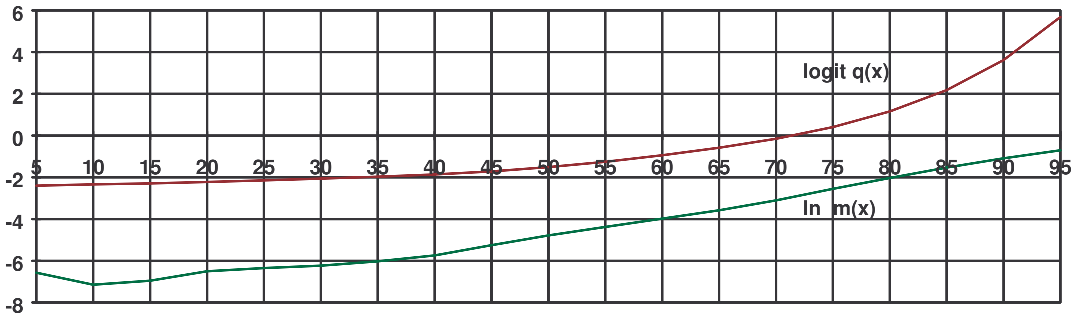
Age x	$l(x)$	μ(X)	ln μ(x)	logit q(x)
1	93247	0.041331	3.186140	-2.62526
5	91630	0.001397	6.573484	-2.39310
10	91184	0.000790	7.143969	-2.33631
15	90811	0.000958	6.950628	-2.29077
20	90237	0.001507	6.497541	-2.22384
25	89486	0.001760	6.342412	-2.14137
30	88651	0.001974	6.227677	-2.05558
35	87701	0.002417	6.025103	-1.96441
40	86505	0.003208	5.742137	-1.85788
45	84737	0.005257	5.248111	-1.71412
50	81944	0.008384	4.781458	-1.51256
55	77797	0.012597	4.374306	-1.25387
60	71945	0.018876	3.969889	-0.94173
65	64133	0.027973	3.576512	-0.58114
70	53669	0.044979	3.101550	-0.14702
75	39782	0.077799	2.553627	0.41455
80	23809	0.131253	2.030630	1.16318
85	10139	0.215159	1.536377	2.18187
90	2602	0.336472	1.089241	3.62252
95	338	0.497041	0.699082	5.68649

We have now observed two “linearizations” of a life-table structure which may be obtained with (4) or with (6). Both allow the relation existing between two life-tables to be established by means of the determination of the linking parameters A e B.

The quality of the relation result depends of the way in which the data in discussion were transformed in view of their linearization. On this subject see table 1 and Fig.1 where the values of the logit of q(x) and the approximated[1] force of mortality are reported. They are calculated for the 1950-53 Italian male life-table.

From Fig.1 in particular, it may be noted that in respect of age x and in the example shown, the trend of logit q(x) from zero to 60-65 years and logit ln μ(x) from 10 to 95 years may be represented quite well by a straight line.

Fig.1 Trend of logarithms of the instantaneous rate of mortality m(i)and of logit q(x) relative to Italian male life-table 1950-53.



Age:x years

1.4 A further transformation may be made (Petrioli, 1983) using the Weibull function in the study of mortality, for which the force of mortality results as being:

$$\mu(x) = a\, b\, (x - \delta)^{b-1} \qquad [9]$$

where δ represents the starting point of the distribution which in this case is placed as equal to zero.

---

(1)Supposing the trend according to a second degree function $f(x)=a^2x+bx+c$ the approximated value of instantaneous rate of mortality may be calculated by the formula:

$$\mu(x) = \left(l_{x-\Delta x} - l_{x+\Delta x}\right) / 2\Delta x l_x$$

where l_x is the number of survivors at age x out of an original cohort of 100,000.

The survival function is given by:

$$p(x) = e^{-\int_0^x \mu(t)dt} \qquad [10]$$

from which, by substituting μ(x) with the value of [10], we obtain:

$$p(x) = e^{-a x^{b}} \qquad [11]$$

By linearizing [11] the result is:

$$\ln [- \ln p(x)] = \ln a + b \ln x \qquad [12]$$

Now,by placing W(x)=ln[-ln p(x)], with reference to the two life-tables, we have:

$$W_1(x) = \ln a_1 + b_1 \ln x \qquad [13]$$
$$W_2(x) = \ln a_2 + b_2 \ln x$$

therefore, as: $A=\ln a_1 - (b_1 .\ln a_2)/b_2;\quad B=b_1 /b_2$ [14]

the following relation is obtained:

$$W_1(x)= A + B.W_2(x) \qquad [15]$$

1.5 If we consider the resistance function, it is possible to set out (Petrioli, 1965, 1975) a type of link such as:

$$\ln [r_1(x).\frac{\omega - x}{x}] = A + B \ln [r_2(x).\frac{\omega - x}{x}] \qquad [16]$$

where, as: $r(x) = \frac{p(x)}{1 - p(x)} \cdot \frac{x}{\omega - x}$ [17]

the terms in the square brackets of [16] would be equal to [4] in the form of logit p(x) if the values of the life-table were used directly. On the other hand, in (16) the function chosen to represent r(x) is used, that is:

$$r(x)= x^{P(1)} (\omega - x)^{P(2)} e^{P(3)x^2 + P(4)x + P(5)} \qquad [18]$$

with the parameters P(1), P(2), P(3), P(4) and P(5). (2)

The parameters linking a model of mortality or a life-table (modelized in accordance with the procedure derived from the transformation with the resistance function) taken as a base, and a model or a reference life-table result as follows:

$P_i(1) = B*[P_b(1)- 1]+1$ $\qquad$ $P_i(4) = B*P_b(4)$ $\qquad$ [19]
$P_i(2) = B*[P_b(2)+ 1]-1$ $\qquad$ $P_i(5) = B*P_b(5)+A$
$P_i(3) = B*P_b(3)$

Evidently [16] is particularly suitable in the research between two or more life-tables when one of those tables is examined as a "base" and the other as reference.

The base life-table is analytically represented by means of [18]. Let us consider two Italian life-tables for females, in the period of time 1960-62 and 1970-72 respectively, which are reported in Table 2.

The survivors of the two tables as well as the logarithm of the ratio between the probability of surviving and the probability of dying from birth to age x can be found in table 2 and in Fig.2.

For example, parameters A and B which refer to the relation between the two life-tables reported in table 2 result as being A=-0.15883544; B=0.91903520, having considered the one of 1970-72 as the base-table.

---

(2) From [16] we can write:

$$r_1(x) = e^A . r_b(x)^B \left[(w-x)/x\right]^{B-1}.$$

where $r_b(x)$ represents a standard function of resistance. This allows us to build models of $r_i(x)$ derived from the standard $r_b(x)$ and parameters A and B, that is:

$$r_i(x) = x^{B[P_b(1)-1]+1} . (w-x)^{B[P_b(2)+1]-1} . e^{BP_b(3)x^2 + BP_b(4)X + BP_b(5)+A}$$

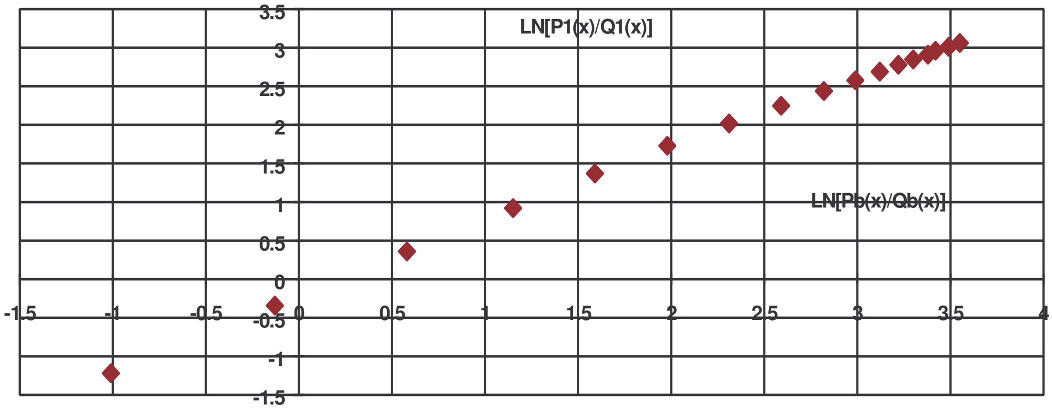
Tab.2 Observed and theoretical survivors of two Italian life-tables for females, 1960-62 and 1970-72.

	Table 1970-72			Table 1960-62		
Age x	Surviv. observ. $l_b(x)$	Surviv. theoret.	logit $p_b(x)$	Surviv. observ. $l_i(x)$	Surviv. theoret.	logit $p_i(x)$
0	100000	100000	-	100000	100000	-
1	97525	97525	3.673868	96209	96149	3.233893
5	97207	97175	3.549726	95519	95659	3.059479
10	97038	97038	3.489238	95278	95469	3.004566
15	96847	96925	3.424778	95092	95314	2.963978
20	96701	96780	3.378005	94864	95114	2.916170
25	96459	96573	3.304709	94565	94831	2.856428
30	96165	96276	3.221896	94177	94429	2.783360
35	95786	95853	3.123704	93669	93862	2.694309
40	95219	95249	2.991530	92972	93064	2.582396
45	94379	94380	2.820809	91967	91937	2.437872
50	93039	93108	2.592696	90518	90323	2.256154
55	90953	91198	2.307909	88315	87973	2.022604
60	87886	88153	1.981679	84942	84479	1.730059
65	83195	83580	1.599511	79782	79192	1.372725
70	75923	76039	1.148463	71559	71143	0.922691
75	64052	64052	0.577622	58951	59190	0.361940
80	46726	46608	-0.131148	41458	42949	-0.345064
85	26864	26221	-1.001534	22744	24790	-1.222823
P(1)		0.89572			0.90416	
P(2)		5.46027			4.93721	
P(3)		-0.00007			-0.00006	
P(4)		0.06880			0.06323	
P(5)		-26.70223			-24.69913	
Q(0)	0.02475	0.02475		0.03791	0.03851	
E(0)	74.30	74.34		71.85	72.56	
r(xm)	11.638	11.638		8.00081	7.8976	
xm	44.103	44.014		47.820	45.770	
A			-0.15883544			
B			0.91903520			

* P(1),P(2),....,P(5) are the parameters of the resistance function [18];
* Q(0) is the mortality quotient of the first year of life;
* E(0) is the expectation of life at birth;
* r(xm) is the maximum value of resistance function [17];
* xm is the age of maximum r(x);
* A and B are the parameters of [16].

The theoretical survivors of the life-table 1960-62 are therefore obtained from parameters P(1),P(2),P(3),P(4),P(5) reported under the column in Table 2.
These parameters are obtained from relation [19] using parameters $P_b(1)$,$P_b(2)$,$P_b(3)$,$P_b(4)$,$P_b(5)$ of the base-table 1970-72.
The process which proves to be rather satisfactory may be used for various applications as, for example, in population projections.

Fig.2 Trend of ln $[p_1(x)/q_1(x)]$ in function of ln $[p_b(x)/q_b(x)]$



2. REPRODUCTION

2.1 The process of modification concerning female fertility, which for some time has been dependent on and conditioned by the social and economic development of many countries, is reflected in the total number of children a woman has as well as in the structure of the fertility quotient
If we examine the distribution of the age specific fertility rates, it may be noted that, in general, it assumes a particular form which derives from the action of numerous elements. These may be differentiated in endogenous and biological factors and exogenous and evironmental factors. The biological factors depend on the woman herself and appear with varing intensity and length according to how much they are influenced by environmental factors, that is sexual habits, ethnic group, cultural level and economic, social and health reasons, etc.
In modern society,the distribution of age-specific fertility rates in age-groups is affected more by environmental factors than biological factors. This influence

becomes concrete as rational mechanisms are adopted which tend to regulate female fertility.

The evolution of fertility may be examined by analysing indicators which characterize the fenomenon such as total reproduction, mean age at childbearing, the standard deviation, the skewness of age-specific fertility rates, the parity,etc., which may be used both singurarly and in association.

Rather than by single indicators as seen above, the degree of difference between two or more fertility rate distributions may be determined by using all the information made available by rates, with one of the relational methods previously proposed.

2.2 For the fertility distributions it is possible to single out one type of link, after having observed that, as with the probability of dying, the function R(x), which indicates female reproductivity up until the age x, also follows a logistic trend and therefore the following relation is possible (3):

$$R(x) = \frac{R(t)}{1 + e^{a+bx}} \qquad [20]$$

from which

$$\ln \left[\frac{R(t)-R(x)}{R(x)}\right] = a+bx \qquad [21]$$

This could be useful for setting out a relation which is similar to [4].

From other points of view, two theoretical lines have been put forward (Petrioli,1975 e 1983)for studying the relations between fertility distributions: one established by using the Gompertz function, the other by the Weibull function as well as the one based on the log-logistic (Menchiari,1988).

Both of these lines are based on a particular aspect of the distribution of fertility rates.

R(x) indicates female reproduction until age x and we will look for its force, that is, the force of reproduction at age x once the total reproduction R(x) at this age, has already been obtained.

---

(3) R(t) represents the total reproduction observed at the end of the fertile period, but it is different from α (theoretical total reproduction) being $\alpha > R(t)$.

The force of reproduction, already indicated with $\mu(x)$ is given by:

$$\lim_{\Delta x \to o} \frac{R(x+\Delta x) - R(x)}{\Delta x \cdot R(x)} = \mu(x) \qquad [22]$$

where $$\mu(x) = \frac{R'(x)}{R(x)} = \frac{f(x)}{R(x)} \qquad [23]$$

with f(x) which represents the fertility function at age x, therefore we may write:

$$f(x) = \mu(x).\ R(x) \qquad [24]$$

By integrating the first two terms of [23] we have:

$$R(x) = e^{\int \mu(x)dx} \qquad [25]$$

therefore if the force of reproduction is represented by the Gompertz function, it results that:

$$\mu(x) = b\ c^{x} \qquad [26]$$

that is: $$R(x) = e^{\frac{b\ c^{x}}{\ln c} + \delta} \qquad [27]$$

where δ is a constant integration additive.

If we place:

$$e^{\frac{b}{\ln c}} = \beta\ ; \qquad e^{\delta} = \alpha\ ; \qquad c = \gamma$$

the fertility distribution function results as being:

$$R(x) = \alpha\beta^{\gamma^{x}} \qquad [28]$$

The density function already seen in [24] will be [4] :

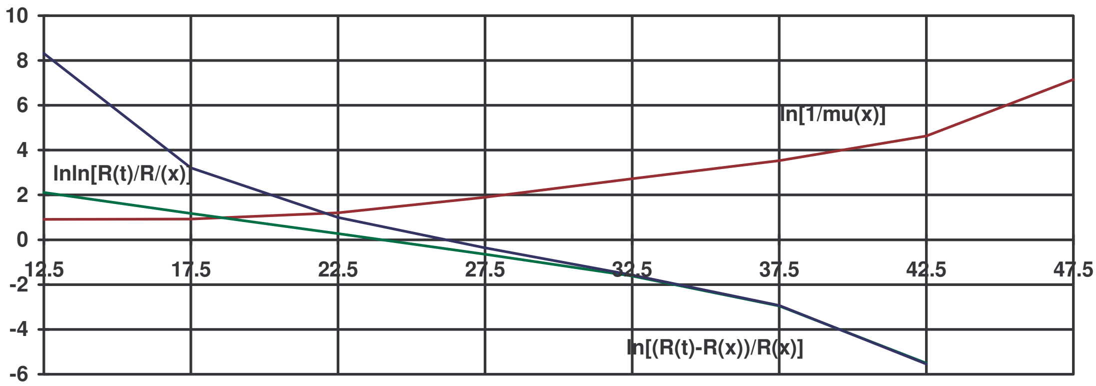
$$f(x)=\alpha.\ln \gamma.\ln \beta.\gamma^{x} . \beta^{\gamma^{x}} \qquad [29]$$

Tab.3 Table of female fertility. Italy, 1960. Transformation for their linearization.

Age groups x,x+Δx	R(x+Δx)	f(x)	Gompertz ln[1/μ(x)]	Weibull lnln[R(t)/R(x)]	Logistic $\ln\left[\frac{R(t)-R(x)}{R(x)}\right]$
10-14	0.6	0.12	0.9163	2.1186	8.3197
15-19	95.2	18.92	0.9289	1.1796	3.2137
20-24	660.0	112.96	1.2068	0.2753	1.0050
25-29	1451.8	158.35	1.8973	-0.6375	-0.3615
30-34	2018.6	113.36	2.7283	-1.6143	-1.5732
35-39	2337.4	63.76	3.5310	-2.9492	-2.9229
40-44	2453.5	23.22	4.6363	-5.5045	-5.5435
45-49	2463.1	1.92	7.1549	-	

(4) To determine the parameters of the Gompertz function in the fertility distributions it is better to use a transformation of the variable making $f(x-x_o)$ or $R(x-x_o)$. In numerous applications x_o has been fixed at 24 years.

Fig.3 Trend of the transformation of fertility rates with the Gompertz, Weibull and logistic functions.



Age: years

As we have seen for mortality, if we have two fertility rate distributions, by using the reciprocal of the force of reproduction we can write the link as follows:

$$\ln [1/\mu_i(x)] = A + B\,[1/\mu_b(x)] \qquad [30]$$

bearing in mind that on the basis of [26] it is $[1/\mu(x)]= K-x.\ln c$, having placed $-\ln b=K$.

As every function chosen to represent fertility requires a transformation of the observed data, an application is shown in table 3 where the Gompertz, Weibull and logistic functions for Italian female rates in 1960 are examined (fig.3).

The transformation graphs show quite a linear trend which is much clearer for the Gompertz and Weibull function than for the logistic one. From [30] we can determine the values of A and B both directly, and after having calculated the linearity parameters of the two observed fertility functions, thereby obtaining.

$$A = K_1 - K_2 \frac{\ln C_1}{\ln C_2}\,; \qquad B = \frac{\ln C_1}{\ln C_2} \qquad [31]$$

If, on the other hand, the two distributions have been analytically represented [28] and [29] we will have:

$$A = -\ln(\ln\gamma_1.\ln\beta_1) - \frac{\ln\gamma_1}{\ln\gamma_2}\,[-\ln(\ln\gamma_2.\ln\beta_2)]\,; \qquad B = \frac{\ln\gamma_1}{\ln\gamma_2} \qquad [32]$$

Likewise, we can again consider the Gompertz function in form [28] and proceed to the linearization of the two age-specific fertility rate distributions, obtaining:

$$\ln \ln \frac{\alpha_i}{R_i(x)} = \ln[-\ln \beta_i] + (x\text{-}TV).\ln \gamma_i$$

$$\ln \ln \frac{\alpha_b}{R_b(x)} = \ln[-\ln \beta_b] + (x\text{-}TV).\ln \gamma_b \qquad [33]$$

from which, having placed the variable transformation as equal to TV we have:

$$\ln \ln \frac{\alpha_i}{R_i(x)} = A_i + B_i \ln \ln \frac{\alpha_b}{R_b(x)} \qquad [34]$$

where $A_i = \ln[-\ln \beta_i]\text{-}B_i.\ln[-\ln\beta_b]$; $\quad B_i = \dfrac{\ln \gamma_i}{\ln \gamma_b}$ [35]

if we only work on the parameters of [28].

However,[34] can be solved by using the observed data directly, placing the total reproduction rate observed at the end of the fertile period t in place of α_i and α_b that is $R_i(t)$ and $R_b(t)$.

To pass from one fertility function taken as standard to other functions linked by means of a set of A and B values, the parameters of the models obtained will be:

$$\gamma_i = \gamma_s^{B_i} \;; \qquad \beta_i = e^{\frac{(\ln \gamma_s .\ln \beta_s)^{B_i}}{B_i \,.\, e^{A_i} \,.\, \ln \gamma_s}} \qquad [36]$$

Obviously the same A and B values may be used directly in [30] so as to have:

$$\ln [1/\mu_i(x)] = A_i + B_i\, [1/\mu_b(x)] \qquad [37]$$

in which the values of the force of fertility can be directly calculated from the observed data through [22].

2.3 The instantaneous reproduction rate [22] or [23] may be represented by the Weibull function, that is:

$$\mu(X) = \lambda\gamma(V-X)^{\gamma-1} \qquad [38]$$

By substituting this expression in [25] and by developing the integral we have:

$$R(x) = \alpha.e^{-\lambda(V-X)^{\gamma}} \qquad [39]$$

that is the reproduction function until age x with parameters λ, γ and $\alpha=e^{C}$ where c is the constant integration additive.

The theoretical fertility rate distribution is then:

$$f(X) = \alpha.\ \lambda.\ \gamma(V-X)^{\gamma-1}.\ e^{-\lambda(V-X)^{\gamma}} \qquad [40]$$

which results as being formed by the product of the instantaneous reproduction rate $\mu(x)$ for the reproduction R(x) had up until age x. In order to determine the link between two or more fertility distributions, which we will indicate as $f_i(x)$ with a as "base" or reference, indicated with $f_b(x)$, we can work along two lines which respectively consider the instantaneous rate and the total reproduction up age x.

<u>Linearization of the instantaneous rate</u>

Considering again [38], we have:

$$\ln \mu_i(x) = \ln \lambda_i . \gamma_i + (\gamma_i-1).\ln(v-x) \qquad [41]$$
$$\ln \mu_b(x) = \ln \lambda_b . \lambda_b + (\gamma_b-1).\ln(v-x)$$

As we have seen for the Gompertz function, the parameters linking the two straight lines which allow passing from one fertility distribution to another, depend on the condition of parallelism of the straight lines and on the distance of the linearization values:

$$\ln \mu_i(x) = A + B \ln \mu_s(x) \qquad [42]$$

where $$A = \ln \lambda_i\gamma_i - \frac{\gamma_i\text{-}1}{\gamma_b\text{-}1} \ln \lambda_b\gamma_b \ ; \qquad B = \frac{\gamma_i\text{-}1}{\gamma_b\text{-}1} \qquad [43]$$

As mentioned above for the Gompertz function, in order to calculate the instantaneous rate, it is possible to proceed to their determination directly from observed data.

<u>Linearizing total reproduction</u>

Considering the value α in [39] as the total reproduction, at the end of the fertile period the following relation is obtained:

$$\ln \ln \frac{\alpha_i}{R_i(x)} = \ln \lambda_i + \gamma_i.\ln(v\text{-}t) \qquad [44]$$

as well as $$\ln \ln \frac{\alpha_b}{R_b(x)} = \ln \lambda_b + \gamma_b.\ln(v\text{-}t)$$

from which: $$\ln \ln \frac{\alpha_i}{R_i(x)} = A_i + B_i \ln \ln \frac{\alpha_b}{R_b(x)} \qquad [45]$$

where $$A_i = \ln \lambda_i\text{-}B_i.\ln \gamma_b \ ; \qquad B_i = \frac{\gamma_i}{\gamma_b} \qquad [46]$$

The values of A and B may be fixed according to a sequence which translates the transformation process of the standard fertility rates by working directly on [45].

2.4 Let us consider the log-logistic function (Menchiari, 1988) to represent the trend of female reproduction at every age x until total observed reproduction is reached, indicated R(t), for t in any case higher then every x. The log-logistic function has the following functions respectively

repartition: $R(x) = R(t)[1+((x-\theta)/A)^{-B}]^{-1}$ [47]

density: $f(x)=[(R(t).B)/A]*[((x-\theta)/A)^{-(B+1)}]*[1+((x-\theta)/A)^{-B}]^{-2}$ [48]

where θ represents the transformation variable.
If the present relation is fixed as:

$$\ln [(R(t)-R(x))/R(x)] = LL(x) \quad [49]$$

the linearization of [47] may be written as:

$$- \ln [LL(x)]= -B \ln A + B \ln(x-\theta) \quad [50]$$

Menchiari has shown, in this respect, that between two fertility rate distributions, one taken as base of reference (b) and the other as reference (i) the following relation exists.

$$-\ln[LL_i(x)] = A_i - B_i \ln[LL_b(x)] \quad [51]$$

2.5 Determining the fertility rates is easy when the relations between distributions have been made through analytical functions which are representative of the phenomenon, using the relative parameters [28,29,39,40].

If, on the other hand, the relations between fertility rate distributions are established through instantaneous rates [22,30,37,41], as these are based on the knowledge of the values of R(x), the availability of the final value R(t), or the total reproduction in age t, is sufficient for gradually going back to the fertility rate values in age t-1, t-2,...making:

$$R(x) = \frac{R(x+\Delta x).[2-\Delta x.\mu(x+\Delta x/2)]}{[2+\Delta x.\mu(x+\Delta x/2)]} \quad [52]$$

which derives from the following relation with which the approximated value of $\mu(x+\Delta x/2)$ may be calculated:

$$\mu(x+\Delta x/2)=\frac{2.[R(x+\Delta x)-R(x)]}{\Delta x[R(x+\Delta x)+R(x)]} \qquad [53]$$

as we have seen in [22] and [23].

Let us consider, for example, two distributions of quinquennial fertility rates for Italy in 1930 and 1945.

The case shown in table 4 is particular inasmuch as the instantaneous rate is used col.3 which is calculated on the distribution of the 1930 rates. Therefore, the reproduction obtained (col.8) maintains the reference distribution structure with identical mean child-bearing age and empirical variance.

Tab.4 Relations between distributions of fertility rates. Italy 1930 and 1945. Observed and theoretical data.

Age groups	Year 1930			Year 1945				
$x,x+\Delta x$	$\bar{f}(x)$ obse. (1)	$R(x+\Delta x)$ observ. (2)	$\ln[1/\bar{\mu}_1(x)]$ observ. (3)	$\bar{f}(x)$ obse. (4)	$R(x+\Delta x)$ observ. (5)	$\ln[1/\bar{\mu}_1(x)]$ observ. (6)	$f(x)$ theor. (7)	$R(x+\Delta x)$ theor. (8)
10-14	0.02	0.1	0.9162	0.06	0.3	0.9162	0.01	0.07
15-19	15.90	79.6	0.9188	17.26	86.6	0.9232	11.10	55.50
20-24	127.20	715.6	1.1396	88.92	531.2	1.2453	88.70	499.00
25-29	189.80	1664.6	1.8358	129.68	1179.6	1.8865	132.38	1160.90
30-34	172.96	2529.4	2.4952	118.54	1772.3	2.5218	120.62	1764.00
35-39	129.76	3178.2	3.0907	90.44	2224.5	3.0954	90.50	2216.50
40-44	58.32	3469.8	4.0430	40.34	2426.2	4.0543	40.68	2419.90
45-49	7.06	3505.1	6.2025	3.66	2444.5	6.5004	4.92	2444.50

If instead we use the parameters linking two distribu-tions, they will be A=0;B=1 respectively for 1930, which is taken as base and A=-0.03520; B=1.037928 for 1945 (always in respect of 1930), therefore, it results that:

$$\ln[1/\mu_{1940}(x)] = -0.03535 + 1.03798 \ln[1/\mu_{1930}(x)] \qquad [54]$$

from which, starting from the values observed in 1930, the theoretical values for 1940 are obtained (Tab.5).

This procedure may be useful in the projection of fertility rates because it allows us to pass from base $\ln[1/\mu_b(x)]$ to $\ln[1/\mu_i(x)]$ of the distribution we intend to obtain.

This may be done after having introduced a foreseen fertility rate as well as linking parameters A and B estimated so as to obtain the complete distribution of the fertility rates and to vary their structure at the same time.

Tab.5 Results of the relationship between fertility rate distributions through [54]. Italy 1930 and 1945.

Age groups $x,x+\Delta x$	Year 1945 $\ln[1/\bar{\mu}_1(x)]$ theor.	$\bar{f}(x)$ theor.	$R(x+\Delta x)$
10-14	0.9156	0.01	0.06
15-19	0.9183	12.51	62.60
20-24	1.1475	92.68	544.00
25-29	1.8702	136.38	1225.90
30-34	2.5546	118.26	1817.20
35-39	3.1727	85.02	2242.40
40-44	4.1612	36.38	2424.30
45-49	6.4027	0.04	2444.50

In substance, as [52] implicitly indicates, once the total fertility rate value has been fixed, it is possible, by going back in the various ages, to determine the values of the various R(x) in both annual rate distributions and quinquennial distributions.

2.6 The procedure becomes more complicated when looking for a relation between a complete rate distribution and a truncated fertility rate distribution in which the missing rates are to be determined. In such a case the first step is to refer to the estimate of the total fertility rate which may be done with one of the procedures normally indicated for this purpose or by relation [45] in which the linking parameters A_i and B_i result from the linear relation $y(x)=A + Bx$, where the variables are $y(x)=\ln\ln(R(t)_i/R_i(x))$; $x=\ln\ln(R(t)_b/R_b(x))$.

We therefore have:

$$R(t)_i = R_i(x).\mathrm{EXP}\left[\left(\ln\frac{R(t)_b}{R_b(x)}\right)^{B_i} .\, e^{-A_i}\right] \qquad [55]$$

which may be resolved only after having determined the parameters linking the age at the beginning of the rate distribution to the age in which the distribution is "truncated" for lack of subsequent rates.

By developing [55] for some values higher than $R_i(x)$ which are known in the examined rates distribution, we obtain a series of corresponding values of R(t) of which we may observe a tendency to converge towards a limit value assumed as total fertility rate (fig.4).

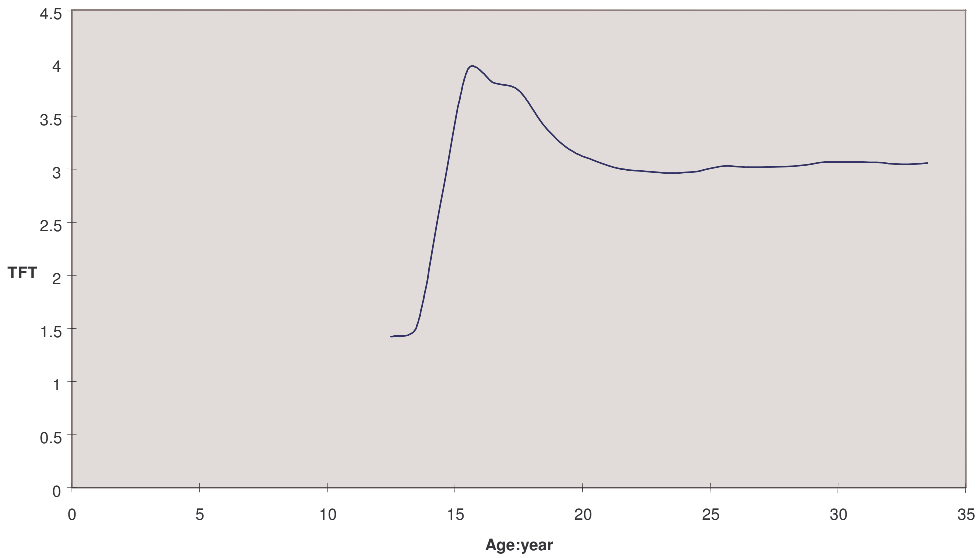
Let us consider, for example, two fertility rate distributions of the Italian population for 1930 e 1935 (tab.6 and 8). As we have already said, by applying [34] to the total observed reproduction rates in the various ages from 12.5 to the age in which the distribution is truncated, it is possible to obtain the values of the parameters which link two considered distributions (of which the 1930 one is taken as base) which are:

$$A = -0.07979; \qquad B = 0.986015$$

$$\ln\left(\ln \frac{R_i(t)}{R_i(x)}\right) = -0.07979 + 0.986015 \ln\left(\ln \frac{R_b(t)}{R_b(x)}\right) \qquad [56]$$

Then, by comparing the degree of matching between the observed values of ln $\mu_i(x)$ and the theoretical ones resulting from [55] we see, for example, that the index of determination r^2 is equal to 0.9996.

Fig.4 Total fertility rate values (TFT) estimated at various ages. Italy , 1930 as base and 1935 as reference.



Now, to evaluate the capacity which [53] offers for estimating the total fertility rate (TFT) relating to a truncated fertility rate distribution, let us consider (tab.5)some pairs of rate distributions observed in Italy from 1930 to 1965 (complete, from 12.5 to 49.5 years of ages), supposing however that for the reference one the age rates from 34.5 years, which we intend to determine, are not known. In table 8 the pairs of distributions studied are indicated.

In each of them the first year is the base distribution while the second is the reference one which is considered "truncated" from 34.5 years inclusive for comparing the degree of correspondence between the observed rates and the estimated ones.

The TFT values, as we have said before, are estimated on the basis of [49] and are necessary for calculating the missing rates using [47]. The total fertility rate however will then result as being different because the rates estimated after 34.5 years are added to the observed ones included between 12.5 and 33.5 years of age. But, as can be seen from the examples in table 6 the values result as being quite close to the age-specific fertility rates as well as both the observed and the theoretical total fertility rates.

For example, supposing the 1935 rate distribution is truncated at 33.5 years and we want to complete it by using the existing link with the 1930 rates up until the above age, we see that we arrive at an estimate of the total fertility rate equal to 3.058, while the true value, drawn from observed data results as being 3.085.

If we then determine the estimated values of the fertility rates from the age of 34.5 and we add them to those of the previous age, which are known, we obtain a total fertility rate at the 50th year which is equal to 3.097, quite close to the real one which, as we said above, is 3.085. Not very large differences are also noted in the other distribution pairs considered(Tab.8).

Finally, it may be noted that the analysis now developed for distributions of available rates for contemporaries may be useful when applied to the study of truncated distributions for real generations.

Tab.6 Fertility rates per age. Comparison between the rate distribution for 1930 and for 1935 supposed as truncated from the age of 33.5 onwards.

Age X	F(X) 1930 (a)	TF(X) 1930 (b)	F(X) 1935 (c)	F(X) 1935 (d)	TFT 1935 (e)	TFS (f)	LLBASE (g)	LLREF (h)
12.5	0.00002	0.00002	0.00003	0.00003	0.00003	1.42200	2.46009	2.41494
13.5	0.00007	0.00009	0.00009	0.00009	0.00012	1.51396	2.32256	2.28267
14.5	0.00010	0.00019	0.00030	0.00030	0.00042	2.74275	2.24650	2.14593
15.5	0.00060	0.00079	0.00170	0.00170	0.00212	3.93538	2.08313	1.93602
16.5	0.00320	0.00399	0.00650	0.00650	0.00862	3.81691	1.85787	1.70990
17.5	0.00950	0.01349	0.01630	0.01630	0.02492	3.74289	1.64710	1.49668
18.5	0.02340	0.03689	0.03060	0.03060	0.05552	3.40616	1.43173	1.29904
19.5	0.04280	0.07969	0.04770	0.04770	0.10322	3.18471	1.22839	1.11372
20.5	0.07060	0.15029	0.07280	0.07280	0.17602	3.07856	1.02291	0.92105
21.5	0.09690	0.24719	0.09240	0.09240	0.26842	3.00447	0.82580	0.73715
22.5	0.13340	0.38059	0.12410	0.12410	0.39252	2.98051	0.61634	0.53645
23.5	0.15760	0.53819	0.14060	0.14060	0.53312	2.96226	0.40922	0.33916
24.5	0.17750	0.71569	0.16070	0.16070	0.69382	2.97994	0.19935	0.13129
25.5	0.17920	0.89489	0.16960	0.16960	0.86342	3.02909	-0.00285	-0.08163
26.5	0.20230	1.09719	0.17150	0.17150	1.03492	3.01771	-0.23149	-0.30052
27.5	0.19950	1.29669	0.17190	0.17190	1.20682	3.02293	-0.46795	-0.53312
28.5	0.19500	1.49169	0.16940	0.16940	1.37622	3.03404	-0.72115	-0.78654
29.5	0.17300	1.66469	0.16210	0.16210	1.53832	3.06803	-0.97694	-1.06691
30.5	0.19700	1.86169	0.16530	0.16530	1.70362	3.06713	-1.32948	-1.41880
31.5	0.17580	2.03749	0.14540	0.14540	1.84902	3.06448	-1.74651	-1.83193
32.5	0.19060	2.22809	0.14700	0.14700	1.99602	3.04702	-2.46563	-2.48166
33.5	0.15000	2.37809	0.13150	0.13150	2.12752	3.05851	-3.92195	-3.92195
34.5	0.15140	2.52949	0.13680	0.13033	2.25785	-	-	-
35.5	0.15000	2.67949	0.13370	0.12912	2.38697	-	-	-
36.5	0.14760	2.82709	0.11650	0.12706	2.51403	-	-	-
37.5	0.12840	2.95549	0.10950	0.11053	2.62456	-	-	-
38.5	0.12040	3.07589	0.09960	0.10364	2.72820	-	-	-
39.5	0.10240	3.17829	0.08630	0.08815	2.81635	-	-	-
40.5	0.09840	3.27669	0.08130	0.08470	2.90105	-	-	-
41.5	0.06570	3.34239	0.05920	0.05656	2.95761	-	-	-
42.5	0.05990	3.40229	0.05030	0.05156	3.00917	-	-	-
43.5	0.04020	3.44249	0.03480	0.03460	3.04377	-	-	-
44.5	0.02740	3.46989	0.02340	0.02359	3.06736	-	-	-
45.5	0.01760	3.48749	0.01350	0.01515	3.08251	-	-	-
46.5	0.00960	3.49709	0.00700	0.00826	3.09078	-	-	-
47.5	0.00460	3.50169	0.00330	0.00396	3.09474	-	-	-
48.5	0.00230	3.50399	0.00150	0.00198	3.09672	-	-	-
49.5	0.00120	3.50519	0.00050	0.00103	3.09775	-	-	-

a) Non cumulated rates all observed in the base-year.
b) Cumulated rates all observed in the base year.
c) Non cumulated rates all observed in year of reference.
d) Theoretical non cumulated rates in year of reference.
e) Theoretical cumulated rates in year of reference.
f) Total fertility rates estimated in age x through [54].
g) LLBASE and LLRIF are respectively LOG(LOG(RBN(N)/RBN(X))) on base data and LOG(LOG(RIN(N)/RIN(X))) on reference data.
h) The rates indicated in columns b) and e) refer to the age in the first column, plus half a year as they are cumulated rates.

Tab.7 Fertility rates for the Italian population from 1935 to 1965 (Rates multiplied by 100000).

Age	1930	1935	1940	1945	1950	1955	1960	1965
12.5	2	3	3	3	3	3	4	5
13.5	7	9	9	8	8	9	10	10
14.5	10	30	30	20	20	30	50	70
15.5	60	170	170	170	110	150	230	300
16.5	320	650	610	580	470	500	680	900
17.5	950	1630	1540	1350	1200	1210	1430	2010
18.5	2340	3060	3380	2530	2570	2370	2570	3720
19.5	4280	4770	5840	4000	4470	4080	4550	5740
20.5	7060	7280	9520	5690	6700	6320	6690	8590
21.5	9690	9240	10550	7380	9090	8730	9050	11390
22.5	13340	12410	13860	9040	11510	10950	11750	13970
23.5	15760	14060	15260	10830	13250	12550	13530	15620
24.5	17750	16070	16930	11520	14590	14080	15460	17720
25.5	17920	16960	17740	13450	15070	14750	16160	18020
26.5	20230	17150	17860	12140	15810	14950	16810	18200
27.5	19950	17190	17960	13430	15550	14940	16050	17750
28.5	19500	16940	17200	12910	15530	14640	15320	16550
29.5	17300	16210	17010	12910	14100	13700	14840	15030
30.5	19700	16530	16330	12640	15120	13020	13610	14560
31.5	17580	14540	15890	12270	11980	12190	12260	13430
32.5	19060	14700	14660	12080	12260	11230	11340	12050
33.5	15000	13150	13880	11180	11220	10450	10280	10690
34.5	15140	13680	13090	11110	10720	9210	9190	9670
35.5	15000	13370	12550	10270	9900	9310	8260	8520
36.5	14760	11650	11540	10040	8880	7270	7320	7370
37.5	12840	10950	10680	9090	8340	7120	6260	6260
38.5	12040	9960	9560	8330	7120	5990	5550	5290
39.5	10240	8630	8520	7490	6370	5310	4490	4420
40.5	9840	8130	7440	6410	5380	4340	4180	3470
41.5	6570	5920	5830	5090	4420	3400	2800	2620
42.5	5990	5030	4480	4000	3270	2570	2270	1890
43.5	4020	3480	3060	2860	2320	1680	1450	1310
44.5	2740	2340	2000	1810	1320	1080	910	730
45.5	1760	1350	1180	1020	850	620	530	500
46.5	960	700	550	490	420	280	250	240
47.5	460	330	260	210	210	100	110	130
48.5	230	150	90	80	110	20	50	50
49.5	120	50	80	30	50	2	20	30

Tab.8 Functional link between pairs of fertility distribution rates, estimate of fertility rates of the reference distribution from 33.5 years of age onwards. The rates are multiplied by 100000. Observed (Obs) and theoretical (Theo) data.

Pairs of distributions	1930-1935		1935-1940		1950-1955		1955-1960		1960-1965	
Parameters of link.										
Ai	-0.079790		-0.020572		-0.012378		-0.031799		-0.055299	
Bi	0.986015		1.004801		0.997484		1.002262		0.999019	
RR	0.9996		0.9998		0.9999		0.9998		0.9998	
	Obs.	Theo.	Obs.	Theo.	Obs.	Theo.	Obs.	Theo.	Obs.	Theo.
Age	year 1935		year 1940		year 1955		year 1960		year 1965	
12.5	3	3	3	3	3	3	4	4	10	10
13.5	9	9	9	9	9	9	10	10	10	10
14.5	30	30	30	30	30	30	50	50	70	70
15.5	170	170	170	170	150	150	230	230	300	300
16.5	650	650	610	610	500	500	680	680	900	900
...	...	...	...	...	...	...	...	...	...	...
33.5	13150	13150	13880	13880	10450	10450	10280	10280	10690	10690
34.5	13680	13033	13090	14343	9210	10000	9190	9179	9670	9595
35.5	13370	12912	12550	14018	9310	9223	8260	9279	8520	8624
36.5	11650	12706	11540	12215	7270	8284	7320	7246	7370	7643
37.5	10950	11053	10680	11481	7120	7780	6260	7096	6260	6536
38.5	9960	10364	9560	10443	5990	6642	5550	5970	5290	5795
39.5	8630	8815	8520	9048	5310	5942	4490	5292	4420	4688
40.5	8130	8470	7440	8524	4340	5019	4180	4326	3470	4364
41.5	5920	5656	5830	6207	3400	4123	2800	3389	2620	2923
42.5	5030	5156	4480	5274	2570	3050	2270	2561	1890	2370
43.5	3480	3460	3060	3649	1680	2164	1450	1674	1310	1514
44.5	2340	2359	2000	2453	1080	1231	910	1076	730	950
45.5	1350	1515	1180	1415	620	793	530	618	500	553
46.5	700	826	550	734	280	392	250	279	240	261
47.5	330	396	260	346	100	196	110	100	130	115
48.5	150	198	90	157	20	103	50	20	50	52
49.5	50	103	80	52	2	47	20	2	30	21
TFT	3.085	3.098	3.172	3.266	2.391	2.458	2.463	2.508	2.688	2.723

REFERENCES

BRASS W.(1971)," On the Scale of Mortality", in Biological Aspects of Demography, W.Brass-Taylor & Francis LTD(eds), London.

BRASS W.(1973)," The Demography of Tropical Africa" , Princeton University Press, New Jersey.

BRASS W. (1977)," The assessment of the validity of fertility trend estimates from maternity histories", International Conference: Mexico.1,89-100 IUSSP, Liege.

BRASS W.(1981)," The use of the Gompertz relational model to estimate fertility",General Confe= rence IUSSP:9-16 Decembre, Manila.

BERKSON J.(1944)," Application of the Logistic Function to Bioassay", Journal of the American Statistical Association.

FINNEY D.J.(1952)," Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve", Cambridge University Press.

KEYFITZ N.(1975)," Application of Demographic Models to Methodology of Projections: new dimensions, gaps and pitfalls", UN Population Division Ad Hoc meeting of Experts on Demographic Models.

MENCHIARI A. (1988)," Development of relational techniques by the log-logistic function", Working Paper n.69, Istituto di Statistica, Università degli Studi di Siena (Italy).

PETRIOLI L.(1965),"Particular characteristics of mortality in Italy from 1881 to 1962 and determi= nation of a new function to represent the survival curve.", Rivista italiana di economia, demografia e statistica, Vol.XIX-Fasc.3-4, Roma.

PETRIOLI L.(1975),"Connection among fertility distribution by the Gompertz function", Working Paper n.19. Istituto di Statistica Università degli Studi di Siena, Siena (Italy).

PETRIOLI L.(1977),"A model for the analysis and the projection of the age specific fertility rates", Rivista italiana di economia demografia e statistica, Vol.XXXI, Roma.

PETRIOLI L.(1978)," Fertility models for the italian population", Workin Paper n.31, Istituto di Statistica, Università degli Studi di Siena (Italy).

PETRIOLI L. (1983)," A new relational method among fertility distributions", Working Paper n.53, Istituto di Statistica, Università degli Studi di Siena (Italy).

PETRIOLI L.(1984)," An analysis of finnish fertility in the last two centuries", IUFO Scientific Conference on Family and Population, Espoo, Hanasaari (Finland).