

ROLAND PRESSAT\*

## O PEWNYM PROBLEMIE ANALIZY WZDŁUŻNEJ. OCENA ZDARZEŃ DEMOGRAFICZNYCH METODĄ KOHORT

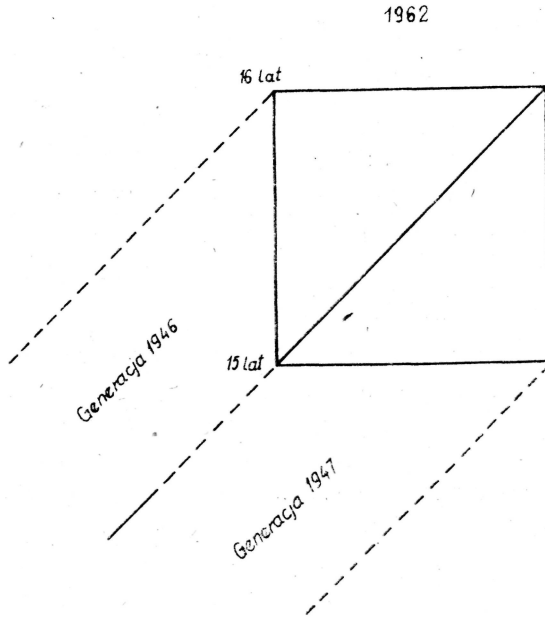
W klasyfikacji zdarzeń demograficznych (urodzenia, śluby, rozwody, zgony) według minionego czasu (urodzenia według wieku itd.) długo przeważało wykorzystanie danych retrospektywnych. Dlatego też dysponuje się urodzeniami, które nastąpiły w danym roku u kobiet obserwowanych według dokładnego wieku, np. między 15 a 16 lat, między 16 a 17, tzn. od 15 ukończonych lat do 16 ukończonych lat.

W odniesieniu do zgonów od dawna podkreślano konieczność dysponowania podwójnym zapisem zdarzeń, by móc ustalić łatwo i dokładnie tablice wymieralności; istnieje podwójna klasyfikacja zgonów, np. 15 ukończonych lat w danym roku (powiedzmy 1962), które miały miejsce według tego, że 15 rocznica została osiągnięta podczas roku rozważanego czy podczas roku poprzedzającego, tzn. w zależności od tego, czy chodzi o zmarłych z generacji 1962—15 = 1947, czy generacji 1961 — 15 = 1946 (ryc. 1). Ta troska o ustalenie tablic wymieralności pociągnęła za sobą dość częste pojawienie się statystyk zgonów o podwójnej tabulacji. O wiele rzadsze są dane tego rodzaju dla urodzeń, ślubów, rozwodów itp. zdarzeń, o ustalenie których mniej się troszczono, ale niezależnie od wszelkich życzeń tego rodzaju statystyki o podwójnej tabulacji są tu ogromnie użyteczne, bo pozwalają na syntezy i analizy wzdlużne, których ważność dobrze jest znana w demografii współczesnej.

Tak więc za każdym razem, kiedy się nie rozporządza podwójną tabulacją, a z reguły jest tak w statystykach nieco dawniejszych, trzeba na ogół szacować metodą grubych przybliżeń. Dlatego chcemy zaproponować metodę dostatecznie

\* Profesor Roland Pressat, wybitny współczesny demograf francuski, dyrektor Departamentu Koniunktury Demograficznej w Institut National d'Etudes Démographiques w Paryżu oraz profesor Uniwersytetu w Montrealu, posiada wielkie osiągnięcia w dziedzinie demografii analitycznej. Przyczynił się do rozwoju i popularyzacji nowoczesnych metod analizy demograficznej. Jest autorem wielu cennych prac z tej dziedziny. Wśród opracowań najbardziej znanych w światowej literaturze demograficznej znajdują się m.in. opracowania: *L'analyse démographique. Concepts — methodes — résultats*, 1969; *Portique de la démographie. Trente Sujets d'analyse*, 1967; *Dictionnaire de démographie*, 1979, które doczekały się licznych wydań i tłumaczeń na języki obce.

wyrafinowaną, by doprowadzić do zadowalających wyników. Przedstawimy tę metodę na przykładzie pierwszych małżeństw i zastosujemy ją do danych francuskich z 1962 r. dotyczących kobiet; ponieważ posiadamy te dane z podwójną tabulacją, będziemy w stanie zmierzyć jakość metody.



Ryc. 1

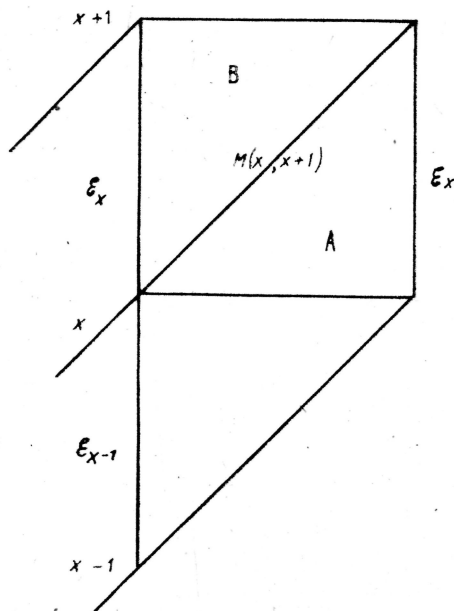
Zanalizujemy czynniki wyjaśniające sposób, w jaki się rozkładają pierwsze małżeństwa w dwu trójkątach diagramu Lexisa, w danym roku, w przedziale wieku  $(x, x+1)$  (ryc. 2).

1) Liczebność obu badanych generacji. Biorąc pod uwagę teoretyczne rozwinienia, na których się oprzemy, będziemy rozważali liczby ogółem (już pozostających w związku małżeńskim i osób stanu wolnego) tych generacji na początku roku, liczby, które oznaczymy przez  $E_{x-1}$  i  $E_x$ .

2) Rozkłady małżeństw według wieku, w generacjach, tak jak podaje tablica małżeństw dla ludzi stanu wolnego. Tak więc pierwsze małżeństwa zachodzące w okresie (trójkącie)  $B$  dotyczą kobiet przeciętnie starszych niż wychodzące za mąż w okresie (trójkącie)  $A$ : wraz z posunięciem się w wieku, caeteris paribus, małżeństwa w  $B$  będą najpierw liczniejsze, a potem mniej liczne niż w  $A$ .

3) Sezonowość pierwszych małżeństw. Różne miesiące roku nie są jednakowo reprezentowane w obu trójkątach, tak więc małżeństwa styczniowe są o wiele częściej reprezentowane w trójkącie  $B$  niż w trójkącie  $A$ . Przeciwnie dzieje się w grudniu

Zobaczymy teraz, jak użyć tych czynników, poczynając od informacji, którą dysponujemy.



Ryc. 2

Zwróćmy najpierw uwagę na drugi czynnik: rozkład pierwszych małżeństw według tablic małżeńskich ludzi stanu wolnego. Znamy ten rozkład na podstawie danych z roku branego pod uwagę, obliczając, dla różnego wieku, współczynnik małżeństw drugiej kategorii<sup>1</sup>, to jest stosunek pierwszych małżeństw  $M(x, x+1)$  do przeciętnej liczebności populacji (wszystkie kategorie stanu cywilnego)  $(E_x + E'_x)$ : 2; współczynniki te są rzeczywiście możliwe do przyjęcia dla pierwszych małżeństw z tablic małżeńskich<sup>2</sup>. W ten sposób otrzymano je z danych francuskich z 1962 r. (tab. 1).

Trzeba teraz odliczyć od tych współczynników, które stosuje się w szeregu kwadratów diagramu Lexisa (tj. w szeregu przedziałów określonych przez kolejne rocznice), współczynniki dotyczące par trójkątów ( $A$  i  $B$ ), z których składają się kwadraty. Wszystkim pierwszym małżeństwom rozrzuconym w jednym trójkącie i przy założeniu równomierności rozkładu odpowiada punkt średni położony w środku ciężkości tego trójkąta. Tak więc trójkątom  $A$  i  $B$  z ryciny 3 odpowiadają średnie punkty wieku  $x + \frac{1}{3}$  i  $x + \frac{2}{3}$ . Pozostaje oszacować, dla tych roczników wieku, promile pierwszych zawartych małżeństw (w terminie częstości dla osób wszystkich stanów matrymonialnych) podczas jednego roku, począwszy od współczynników obliczonych poprzednio. Mamy tu do wykonania interpolacje, których idee przedstawia rycina 4: krzywa współczynnika małżeństw jest zadana przez wartość współczynników

<sup>1</sup> Według naszej terminologii; por. *Vanalyse démographique*, wyd. 3, s. 101.

<sup>2</sup> *Ibidem*, s. 110-112.

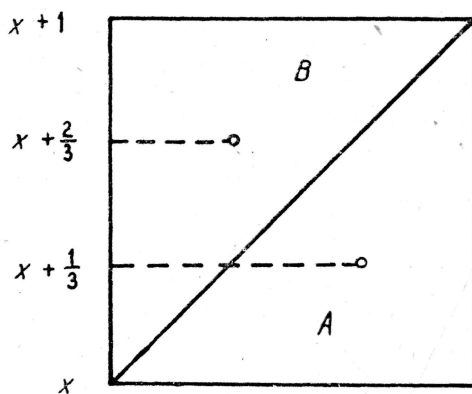
Tabela 1

Szacowanie współczynników małżeństw drugiej kategorii

Wiek (w latach ukończonych)	Pierwsze małżeństwa	Przeciętna liczba lud- ności w połowie roku	Współ- czynnik (w ‰)	Wiek (w latach ukończonych)	Pierwsze małżeństwa	Przeciętna liczba lud- ności w połowie roku	Współ- czynnik (w ‰)
1	2	3	2 : 3	1	2	3	2 : 3
14	34	415 500	8	25	14 372	289 800	4959
15	1 190	405 700	293	26	10 644	293 800	3622
16	4 220	348 700	1 210	27	7 899	302 500	2611
17	10 397	298 800	3 489	28	6 004	309 100	1942
18	22 176	294 600	7 527	29	4 803	316 500	1518
19	31 359	282 200	11 112	30	3 870	325 400	1189
20	34 488	257 900	13 373	31	3 147	330 400	952
21	34 166	250 700	13 628	32	2 612	328 500	795
22	33 558	270 700	12 397	33	2 037	323 800	629
23	27 157	283 400	9 583	34	1 716	324 800	529
24	20 180	284 800	7 086				

poprzednich związanych przeciętnie ze środkami przedziałów wieku ( $x - \frac{1}{2}$ ,  $x + \frac{1}{2}$ ,  $x + \frac{2}{3}$ , ...), skąd wyprowadza się wartość  $t_A$  przy  $x + \frac{1}{3}$  i  $t_B$  przy  $x + \frac{2}{3}$ .

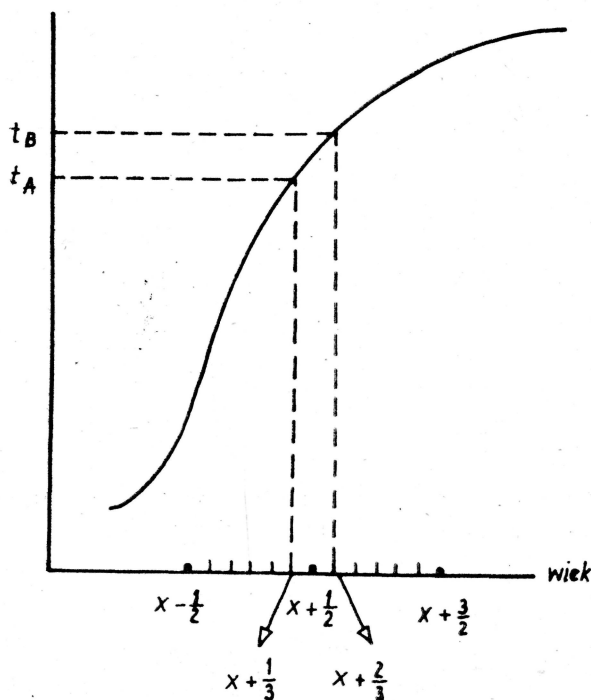
Z uwagi na to, że tak określone współczynniki mają wymiar roczny (odsetki małżeństw w przeliczeniu na jeden rok), wystarczyłoby podzielić je przez 2, żeby



Ryc. 3

otrzymać odsetki małżeństw na trójkącie, gdyż ta powierzchnia odpowiada faktom zaobserwowanym w przeciętnym czasie trwania 6 miesięcy, gdyby nie było sezonowości małżeństw. Ale ta sezonowość powoduje nierównomierny rozkład zdarzeń i trzeba by uwzględnić jej rozmiary. Dochodzi się do tego rozkładając kwadrat na  $12 \times 12$  małych kwadracików służących do zapisania wagi różnych miesięcy w  $A$  i  $B$ ; tak więc w  $A$  styczeń liczy się za 0,5 miesiąca, luty za 1,5, marzec za 2,5, ...

Stopy płodności



Ryc. 4

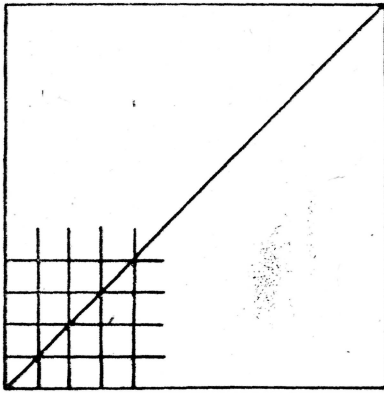
do tych czasów trwania trzeba przeznaczyć współczynniki uwzględniające sezonowość dla roku, o którym mowa<sup>3</sup>. Dołącza się również do trójkąta *A* ciąg wielkości:

$56 \times 0,5 = 28$	$132 \times 6,5 = 858$
$70 \times 1,5 = 105$	$118 \times 7,5 = 885$
$79 \times 2,5 = 197,5$	$126 \times 8,5 = 1071$
$138 \times 3,5 = 483$	$106 \times 9,5 = 1007$
$57 \times 4,5 = 256,5$	$73 \times 10,5 = 766,5$
$128 \times 5,5 = 704$	$117 \times 11,5 = 1345,5$
	<hr style="width: 100%; border: 0; border-top: 1px solid black; margin-bottom: 5px;"/> 7707

Jego suma 7707 oznacza wagę danego trójkąta, ponieważ środek ciężkości trójkąta *B* stanowi różnicę od  $1200 \times 12 = 14\,400$ , czyli 6693. Sprawdzone do jednolitej sumy te wartości wynoszą

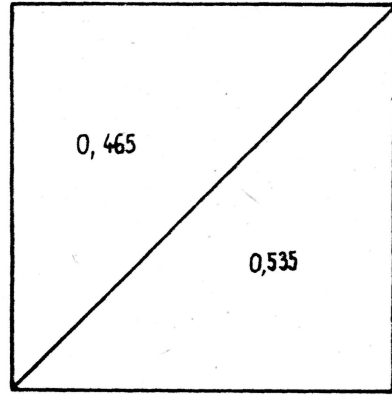
$$\frac{7707}{14400} = 0,535 \quad \text{i} \quad \frac{6693}{14400} = 0,465$$

<sup>3</sup> Współczynniki sezonowości są średnimi miesięcznymi na podstawie 1200 małżeństw rocznie. Pewna niedokładność wynika z tego, że dysponujemy sezonowością wszystkich małżeństw (pierwszych małżeństw i małżeństw ponownych), z braku możliwości odróżnienia, a również i z tego, że nie możemy odróżnić ewentualnych zmienności według wieku zawierania małżeństwa.



$j, f, m, a$

Ryc. 5



Ryc. 6

i oznaczają, ceteris paribus, że 53,5 % pierwszych małżeństw mieści się w trójkącie  $A$ , podczas gdy 46,5% mieści się w trójkącie  $B$ .

Pozostaje jeszcze wprowadzenie łącznie drugiego i trzeciego czynnika, co się rozwiązuje przez mnożenie: oznaczamy przez  $t_A$  i  $t_B$  współczynniki w odnośnych trójkątach; stosując do populacji  $E_{x-1}$  współczynnik  $0,535 t_A$ , otrzyma się pierwsze małżeństwa w  $A$ , a do populacji  $E_x$  — współczynnik  $0,465 t_B$  — pierwsze małżeństwa w  $B$ . Wyliczenia zostały w ten sposób doprowadzone od 15 do 31 lat ukończonych. W drugiej kolumnie tabeli figurują współczynniki interpolowane według poprzednich uwag (i usystematyzowane na ryc. 4). Pierwszy z tego rocznika wieku odpowiada trójkątowi  $A$ . W ten sposób dochodzi się do pierwszego szacunku (kolumna: pierwsze małżeństwa wyliczone). Ale ponieważ suma małżeństw znalezionych w ten sposób, odnoszących się do tego samego wieku, nie odpowiada dokładnie temu, czym były małżeństwa zarejestrowane (i tak w wieku 15 lat  $440+700=1140$  zamiast 1190) trzeba dokonać poprawki, co pozwoli dojść do szacunków końcowych.

\*

Wyniki okazują się dość zadowalające i w każdym przypadku o wiele lepsze od pochodzących z niekiedy stosowanych, pobieżnych sposobów postępowania. Najgorsze są wyniki w wieku maksymalnym zawierania małżeństw, co jest naturalne, gdyż interpretacje w sąsiedztwie maksimum jakiejś krzywej, jeśli to maksimum jest źle ustalone, są bardzo przypadkowe.

Zauważmy ponadto, że rozkład liczby małżeństw odpowiadających różnemu wiekowi według dwóch generacji, które są za każdym razem implikowane, nie jest na ogół ostatecznym celem operacji. W istocie zmierza się — i tak właśnie udowodnialiśmy nasz pogląd — do ułożenia danych statystycznych wedle różnych generacji, do których się odnoszą, w taki sposób, by umożliwiły ewentualne analizy wzdłuż-

Tabela 2

Oszacowanie pierwszych małżeństw kobiet według wieku i generacji. Dane wyjściowe według wieku w latach ukończonych — Francja 1962 r.

Wiek w latach ukończonych	Współczynnik małżeństw (w 10 tys.)	Ludność na 1 I 1962 r. (w tys.)	Współczynniki	Pierwsze małżeństwa wyliczone	Pierwsze małżeństwa zarejestrowane	Szacunek ostateczny	Wartości dokładne
15	20	411 100	0,535	440		464	495
	38	393 600	0,465	700	1 190	726	695
16	98		0,535	2 060		2 120	2 091
	148	297 100	0,465	2 040	4 220	2 100	2 129
17	296		0,535	4 700		4 786	4 918
	404	293 200	0,465	5 510	10 397	5 611	5 479
18	674		0,535	10 570		10 872	10 983
	818	289 000	0,465	10 990	22 176	11 304	11 193
19	1 064		0,535	16 450		16 667	16 825
	1 164	267 900	0,465	14 500	31 359	14 692	14 534
20	1 310		0,535	18 780		19 128	19 176
	1 348	240 500	0,465	15 080	34 488	15 360	15 312
21	1 372		0,535	17 650		17 937	19 076
	13 352	254 000	0,465	15 970	34 166	16 229	15 090
22	1 270		0,535	17 260		17 637	18 742
	1 196	280 200	0,465	15 580	33 558	15 921	14 816
23	1 014		0,535	15 200		15 193	15 750
	920	279 900	0,465	11 970	27 157	11 964	11 407
24	948		0,535	11 200		11 290	11 864
	670	283 100	0,465	8 820	20 180	8 890	8 316
25	530		0,535	8 030		8 020	8 383
	472	289 900	0,465	6 360	14 372	6 352	5 989
26	384		0,535	5 960		5 962	6 184
	346	291 000	0,465	4 680	10 644	4 682	4 460
27	276		0,535	4 300		4 316	4 482
	250	307 000	0,465	3 570	7 899	3 583	3 417
28	206		0,535	3 380		3 360	3 523
	188	304 500	0,465	2 660	6 004	2 644	4 461
29	160		0,535	2 610		2 595	2 747
	146	321 700	0,465	2 200	4 803	2 208	2 056
30	121		0,535	2 080		2 124	2 138
	114	322 400	0,465	1 710	3 870	1 746	1 732
31	99		0,535	1 710		1 703	1 770
	94	331 900	0,465	1 450	3 147	1 444	1 377

ne. Analiza tabeli 2 wykazuje nam, że przegrupowanie na generacje jest jak najbardziej zadowalające, ponieważ zestaw kompensacji, prawie stały, redukuje najczęściej odchylenia od nieistotnych. Należy zresztą zaznaczyć, że 60 % odchyłeń pochodzi z dwu generacji (1939 i 1941) na 16 rozpatrywanych. Należy także podkreślić, że 1962 r., oznaczający koniec wojny w Algierii, odznaczył się skutkiem tego perturbacjami na rynku matrymonialnym (demobilizacja wojskowych, repatriacje z Algierii), co utrudnia podjęte szacunki.

Tabela 3

Pierwsze małżeństwa kobiet według generacji, szacowane na podstawie danych według wieku w latach ukończonych — Francja 1962 r.

Generacja	Wyniki		Odchylenia	Generacja	Wyniki		Odchylenia
	szacunkowe	dokładne			szacunkowe	dokładne	
1946	2 846	2 786	+60	1938	23 254	23 271	-17
1945	6 886	7 047	+61	1937	16 910	16 699	+211
1944	16 483	16 462	+21	1936	12 314	12 173	+141
1943	27 971	28 018	-47	1935	8 998	8 942	+56
1942	33 820	33 710	+110	1934	6 943	6 960	-17
1941	33 297	34 388	-1091	1933	5 239	5 208	+31
1940	33 866	33 832	+34	1932	4 332	4 194	+138
1939	31 114	30 566	+548	1931	3 449	3 502	-53

Znaczenie liczebności generacji (a ogólniej kohort), ewolucja z wiekiem (a ogólniej z rozkładem wieku w kohorcie) rozpatrywanego zjawiska i wreszcie cechy charakterystyczne sezonowości — oto trzy czynniki determinujące rozkład zdarzeń między dwiema generacjami lub kohortami, które przejawiają się w wieku (lub rozkładzie wieku w kohorcie) danym w ciągu określonego roku kalendarzowego. Na przykładzie liczby małżeństw osób stanu wolnego widzimy, że uwzględnienie tych czynników pozwala oszacować bardzo dokładnie to, co należy się każdej generacji lub kohorcie, wychodząc od danych ukończonego wieku (lub trwania). Trzeba rozszerzyć tę metodę na inne zjawiska, by osądzić eksperymentalnie jej ogólną wartość. Można, a priori, wątpić w to, że jest ona odpowiednia w pewnych sytuacjach tak więc, jeśli chodzi o urodzenia, sezonowość zachodząca w podobnym wypadku jest na pewno bardzo zmienna, zależnie od wieku kobiety, podczas gdy te zmiany są z pewnością o wiele mniej odczuwalne jeśli chodzi o liczbę małżeństw. Ale być może, w sumie sezonowość w sprawach dotyczących urodzeń jest tylko czynnikiem najzupełniej drugorzędym w rozdziale zdarzeń według generacji, ponieważ ewolucja płodności według wieku ma zupełnie inne znaczenie. W tej dziedzinie, jak też w bardziej ogólnej dziedzinie aproksymacji, z którą ciągle styka się demograf podczas swych badań, ostatnie słowo należy do weryfikacji na drodze doświadczalnej.

ON CERTAIN PROBLEM OF LONGITUDINAL ANALYSIS; APPRAISAL  
OF DEMOGRAPHIC EVENTS BY THE COHORT METHOD

## Summary

A longitudinal analysis requires utilization of statistics of demographic events located in cohorts history. And so, too often, and in practice always, these statistics are prepared by age (in completed years) of units or by antiquity (in completed duration periods) of cohort's elements. That should be then separated what refers to age given in numbers of events regarding each of both cohorts contributing to event's appearance at that age.

The article is devoted just to a method allowing to do such a separation, explaining it on the ground of statistics referring to first marriages (of single people) of French women in 1962. Such a re-grouping of data by generations delivers to satisfactory results.