Thank you very much for the invitation to come here and tell you something about the early years of Jan Hoem's research. This is a very nice invitation for me to accept, since I had quite a lot of collaboration with Jan during his seven years in Copenhagen.

Jan Hoem's contributions to methodology

Niels Keiding
Section of Biostatistics, University of Copenhagen

Max Planck Institute for Demographic Research
Rostock, 11 April 2017

This is my overview. I was asked to talk about Jan's efforts in statistics. I have made my own interpretation of this topic as you can see, and I think it is important to also mention some persons in the close environment of Jan.

Overview

Pedigree of methodologists

Event history analysis interpretations of demographic concepts

Event history analysis interpretations of demographic techniques

Event history analysis interpretations of actuarial concepts and techniques

Lecture notes

Contributions to statistics

Graduation

Example: Local dependence between life history events

Notice these magic numbers that we enjoyed in Copenhagen. Professor Simonsen, who was the predecessor of Jan, was born exactly 35 years before Jan; so when Simonsen turned 35, Jan was born; and when Jan turned 35, Simonsen turned 70 and retired, and Jan took over. This is the sort of thing that is really appropriate for actuarial mathematicians and demographers.

Jan M. Hoem (1939-2017)

Cand.act., cand.real. (Oslo) 1969 Dr. philos. (Ph.D.) Oslo

1969-1974 Lecturer Univ. Oslo and Researcher, Norwegian Central Bureau of Statistics

1974-1981 Professor of actuarial mathematics, University of Copenhagen 1981-1999 Professor of demometry, University of Stockholm

1999-2007 Director of Max Planck Institute of Demographic Research, Rostock

See comments on previous slide.

Magic numbers

17 April 1904

William Simonsen (professor of actuarial mathematics 1943-1974, University of Copenhagen) was born

17 April 1939

Jan M. Hoem (professor of actuarial mathematics 1974-1981, University of Copenhagen) was born

17 April 1974

Simonsen turned 70 and retired, Hoem turned 35 and took over

This is the pedigree that I will return to. But let me first tell you how I classify Jan's main methodological contributions, by which I mean his work other than the concrete demographic research, of which Gunnar will speak later.

Pedigree of methodologists

Erling Sverdrup

Jan Hoem

Tore Schweder

Odd Aalen

Ørnulf Borgan

Henrik Ramlau-Hansen

Jan did a series of what I would call event history analysis interpretations of demographic concepts. Particularly in the early seventies, I think these were part of his PhD; a couple of them came later on, and I will go into more detail about them.

Jan Hoem's main methodological contributions

Event history analysis interpretations of demographic concepts

1970 Probabilistic fertility models of the life table type. *Theor. Popul. Biol*. **1** (1), 12-38.

A probabilistic approach to nuptiality. *Biométrie-Praximétrie* **11** (1), 3-19.

1971 On the interpretation of the maternity function as a probability density. *Theor. Popul. Biol*. **2**, 319-327. Erratum: **3** (2), 240.

On the interpretation of certain vital rates as averages of underlying forces of transition. *Theor. Popul. Biol.* **2** (4), 454-468.

1972 Inhomogeneous semi-Markov processes, select actuarial tables, and duration-dependence in demography. Pp. 251-296 in T. N. E. Greville (ed.), *Population Dynamics*. New York: Academic Press.

1976 Stochastic stable population theory with continuous time (by Niels Keiding and Jan M. Hoem). *Scand. Act. J*. 1976 (3), 150-175.

1978 Demographic incidence rates. *Theor. Popul. Biol*. **14** (3), 329-337. Bibliographic Note, **18** (2), 195.

1990 Identifiability in hazard models with unobserved heterogeneity: the compatibility of two apparently contradictory results. *Theor. Popul. Biol.* **37** (1), 124-128.

There were also interpretations of demographic techniques; that is, how these calculations were done. You could explain those techniques in Markov chains framework, or what is now called event history analysis. Some of these papers are very, very long, and I do not know how many readers get through them to the end.

Event history analysis interpretations of demographic techniques

1969 Fertility rates and reproduction rates in a probabilistic setting. *Biométrie- Praximétrie* **10** (1), 38-66. Erratum: **11** (1), 20.

1971 Point estimation of forces of transition in demographic models. *J. Roy. Statist. Soc. B* **33** (2). 275-289.

1976 The statistical theory of demographic rates: A review of current developments (with discussion). *Scand. J. Statist*. **3** (4), 169-185.

1982 Multistate life table methodology: A probabilist critique (by Jan M. Hoem and Ulla Funck Jensen). Pp. 155-264 in K. Land and A. Rogers (eds.), *Multidimensional Mathematical Demography*. New York: Academic Press.

1987 Statistical analysis of a multiplicative model and its application to the standardization of vital rates: A review. *International Statistical Review*, 55, 119-152.

Invited contribution (pp. 717-719) to the discussion of David Brillinger's paper "The natural variability of vital rates and associated statistics". *Biometrics* **42** (4), 693-734.

1988 Demographic reproduction rates and the estimation of an expected total count per person in an open population (by Ørnulf Borgan and Jan M. Hoem). *J. American Statist. Assoc*. **83** (403) 886-891.

There was a parallel production of interpretations of actuarial concepts and techniques. These I will have to leave for today, but they are in a style very similar style to the demographic interpretations.

Event history analysis interpretations of actuarial concepts and techniques

1969 Markov chain models in life insurance. *Blätter, Deutsche Ges. für Versich.math*. **9** (2), 91-107.

Some notes on the qualifying period in disability insurance. I. Actuarial values. II. Problems of maximum likelihood estimation. *Mitteil. Verein. schweizer. Versicherungsmath*. **69** (1), 105-116 and **69** (2), 301-317.

1972 A probabilistic theory for collective children's pension insurance. *Skandinavisk Aktuarietidskrift* **54**, 74-79.

Disability income benefits in group life insurance (by Jan M. Hoem, Jan Riis and Rolf Sand). *Skandinavisk Aktuarietidskrift* **54**, 190-203.

1978 Actuarial values of payment streams (by Jan M. Hoem and Odd O. Aalen). *Scand. Act. J*. 1978 (1), 38-48.

1989 The versatility of the Markov chain as a tool in the mathematics of life insurance. *Transactions of the 23rd International Congress of Actuaries*, Helsinki, Finland, 11-16 July 1988. Vol. R, 171-202.

1990 The retrospective premium reserve (by Henk Wolthuis and Jan M. Hoem). *Insurance: Mathematics and Economics* **9**, 229-234.

Lecture notes

There are lecture notes; unfortunately, none of these developed into a real textbook. This is something Jan should have done himself, as he would have done the rest of us an important service, but he never got the time to work through to a real textbook.

1966

Basic concepts of formal demography. Oslo University Press. (In Norwegian.)

1968

Pension funding. Oslo University Press. (In Norwegian.)

1971

An elementary introduction to the mathematics of finance. Oslo University Press. (In Norwegian.)

1973

Non-life insurance mathematics. University of Oslo, Department of Mathematics. (Lecture notes. In Norwegian.)

There is something that could be called contributions to statistics. My favourite is a small paper on purged and partial Markov chains; very elegant, but you do have to be particularly interested. I think it is a very nice piece of work, and I will end this talk with an application of it.

Contributions to statistics

1969

Purged and partial Markov chains. *Skandinavisk Aktuarietidskrift* **52**, 147-155.

The sampling distribution of an estimator arising in connection with the truncated exponential distribution. *Ann. Math. Statist*. **40** (2), 701-703.

1978

Random time changes for multivariate counting processes (by Odd O. Aalen and Jan M. Hoem). *Scand. Act. J*. 1978 (2), 81-101.

And then there are quite a number of papers on graduation; that is, smoothing of curves. It is my impression that this work is not used much anymore; it was sort of derived from the late 19th century, where it was really a computational challenge to graduate curves.

Graduation

1972 On the statistical theory of analytic graduation. *Proc. Sixth Berkeley Symp*. **1**, 569-600.

1975 Some problems in Hadwiger fertility graduation (by Jan M. Hoem and Erling Berge). *Scand. Act. J.* 1975, 129-144.

The demographic interpretation of the basic parameters in Hadwiger fertility graduation (by Jan M. Hoem and Britta Holmbeck). *Statistisk tidskr*. III **13** (5), 369-375.

1976 Theoretical and empirical results on the analytic graduation of fertility rates (by Jan M. Hoem and Erling Berge). Pp. 363-371 in *Proc. 8th Internat. Biometric Conf., Constanta, Romania, 1974*.

On the optimality of modified minimum chi-square analytic graduation. *Scand. J. Statist*. **3** (2), 89-92.

1984 A contribution to the statistical theory of linear graduation. *Insurance: Mathematics and Economics* **3** (1), 1-17

1988 The tails in moving average graduation (by Jan M. Hoem and Per Linnemann). *Scandinavian Actuarial Journal* 4: 193-229.

Let me return to Jan's pedigree of methodologists.

Pedigree of methodologists

Erling Sverdrup

Jan Hoem

Tore Schweder

Odd Aalen

Ørnulf Borgan

Henrik Ramlau-Hansen

Erling Sverdrup was the founder of mathematical statistics in Norway. He was a military cryptographer during the second World War. Then, like many, at least in Scandinavia, he went to the United States in the late forties. He became very fascinated by what was going on over there; he was the main Scandinavian promoter of Neyman-Pearson theory, so much so that in Denmark we thought it was a bit much. He built up the study of mathematical statistics in Oslo, but he was at the same time a professor in actuarial mathematics. So besides his work in mathematical statistics, he wrote survey papers on actuarial mathe–matics as well as demography, and, in particular, important lecture notes published in two editions at University of Oslo. His paper from 1965 was a very important early application of Markov process theory to basic demography models.

Erling Sverdrup (1917-94)

1939-45 Military cryptographer

1949-50 Rockefeller fellow in the USA, became strong proponent of Neyman-Pearson theory

1953-84 Professor in actuarial mathematics and mathematical statistics at University of Oslo.

1952 Basic concepts in life assurance mathematics. *Skand. Aktuarietidskr.* **35,** 115-131. 1961/1967 *Statistiske metoder ved dødelighetsundersøkelser.* Mimeographed, Institute of mathematics, University of Oslo.

1962 *Forsikring mot invaliditet.* Mimeographed, Inst.of mathematics, University of Oslo.

1962 Actuarial concepts in the theory of financing life insurance activities. *Skand. Aktuarietidskr.* **45,** 190-204.

1965 Estimates and test procedures in connection with stochastic models for deaths, recoveries and transfers between different states of health. *Skand. Aktuarietidskr.,* **48**, 184- 211.

Tore Schweder is an interesting original figure. I put him on the side there because he is not a father of Jan, but he is not a son either; he is a contemporary scientifically speaking, and he has always been a professor at the department of economics. He has always primarily been working with whales – you know that Norway has this unique position of defending the whale-hunting cause; it is good to catch whales according to a Norwegian standpoint, and Tore has had many fights in the scientific com–mittee of the international whaling commission, of which he has been a member since 1989. But his first paper is a deeply original one called 'Composable Markov Processes'; full of causality thoughts but published in the Journal of Applied Probability. Somehow, strangely, he never quoted it himself. When I meet him later this year, I will ask him why he did not go on with the fruitful research in that paper. The reason I mention this is that this paper was definitely one of Jan Hoem's favourite papers.

Tore Schweder (1943-)

Cand. real. (MSc) University of Oslo 1968

PhD. University of California, Berkeley 1974. *Transformation of Point Processes – Applications to Animal Sighting and Catch Processes, with special Emphasis on Whales.*

University of Tromsø 1974-83, since 1981 as professor

Professor at University of Oslo, Department of Economics since 1984.

Schweder, T. 1970. Composable Markov Processes. *Journal of Applied Probability,* **7:** 400-10.

Schweder's first paper. Never quoted by himself.

Rich publication list on foundational problems of statistics and on methods for surveying whales.

Odd Aalen (1947-)

Now I shall mention three of Jan's students. Odd Aalen is the most modest person I know and one of the biggest stars I know. He has revolutionised survival analysis through his PhD from Berkeley in 1975, but one should not forget that he was put on the road to that work by Jan, who supervised his master's thesis on the length of stay of an intrauterine device. The germs of the idea that was later developed in Berkeley were already in Odd's master's thesis; and this is something that we really need to remember about Jan, that he was able to put Odd on this path where he later developed so impressively and strongly. Odd will turn 70 next month, and we are going to have a big celebration of him next fall.

Cand. real. (MSc) in mathematical statistics 1972, University of Oslo. Master's thesis on length of stay of intrauterine device (supervisor J. M. Hoem).

Ph.D. in statistics, University of California, Berkeley 1975: *Statistical inference for a family of counting processes*.

Aalen,O.O (1978). Nonparametric inference for a family of counting processes. *Ann. Statist*. **6**, 701–726.

Aalen,O.O.(1980). A model for non-parametric regression analysis of lifetimes. *Lecture Notes in Statistics*, vol.2, Springer-Verlag, NewYork, pp.1–25.

Aalen, O.O., Borgan, Ø., Keiding, N. & Thormann, J. (1980). Interaction between life history events. Nonparametric analysis for prospective and retrospective data in the presence of censoring. *Scand.J.Statist.* **7,** 161-171.

Aalen, O.O. and Johansen, S. (1978). An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand.J.Statist.* **5,** 141–150.

Aalen, O.O., Borgan, Ø., Gjessing, H.K. (2008). *Survival and Event History Analysis: A Process Point of View*. Springer-Verlag, NewYork.

Aalen, O. O., Andersen, P. K., Borgan, Ø., Gill, R. D.; Keiding, N. (2009). History of applications of martingales in survival analysis. *Journal Électronique d'Histoire des Probabililtés et de la Statistique* **5,** 1-28.

Ørnulf Borgan has also worked in this area. He was also a student of Jan's, and as you can see, he moved to Copenhagen with Jan, and was assistant professor with Jan for two years there. So we got to know each other very well in Copenhagen, and have collaborated quite a bit since then; and as you can see, he has several joint papers with Danes. I shall return to the paper from 1980 based on Schweder's 1970 paper, which I just mentioned. We have also done this book together. Odd did not want to participate in this monograph on statistical models based on counting processes. He thought that it was more important to build up Norwegian biostatistics than to write about mathematics. (Odd has fortunately returned to writing many important papers, as well as his own book.) The whole work behind our book was chronicled in the paper with Odd in the French Electronic Journal of History of Probability and Statistics some years ago.

Ørnulf Borgan (1950-)

Cand.real. (M.Sc.) statistics, University of Oslo, 1976.

Assistant professor in actuarial mathematics, University of Copenhagen 1977-79.

Assistant and Associate professor in statistics, Oslo 1980-92.

Professor in statistics, University of Oslo 1993-.

Borgan, Ø. (1979). On the theory of moving average graduation. *Scand. Actuarial J.* **1979**, 83-105.

Aalen, O.O, Borgan, Ø., Keiding, N., and Thorman, J. (1980). Interaction between life history events. Nonparametric analysis for prospective and retrospective data in the presence of censoring. *Scand. J. Statist.* **7**, 161-171.

Borgan, Ø. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure times data. *Scand. J. Statist*. **11,** 1-16. Correction **11,** 275.

Borgan, Ø. and Ramlau-Hansen, H. (1985). Demographic incidence rates and estimation of intensities with incomplete information. *Ann. Statist.* **13,** 564-582.

Andersen, P.K. and Borgan, Ø. (1985). Counting process models for life history data (with discussion). *Scand. J. Statist.* **12,** 97-158.

Andersen, P.K., Borgan, Ø., Gill, R.D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

Aalen, O.O., Borgan, Ø., and Gjessing, H.K. (2008). *Survival and Event History Analysis: A Process Point of View*. Springer-Verlag, New York.

Aalen, O. O., Andersen, P. K., Borgan, Ø., Gill, R. D.; Keiding, N. (2009). History of applications of martingales in survival analysis. *Journal Électronique d'Histoire des Probabililtés et de la Statistique* **5,** 1-28.

Henrik Ramlau-Hansen was an actuarial student in Copenhagen; he is the only person I know who has published his master's thesis in the leading journal Annals of Statistics. This master's thesis was about using the new counting process approach to smooth counting process intensities, and Jan was his supervisor. Jan put him on this road, but Henrik possibly believed that you could make more money in private insurance than at a university. This belief is unfortunately true, but he has very recently returned to academia at the Copenhagen Business School.

Henrik Ramlau-Hansen (1956-)

1981	Cand. act. Univ. Copenhagen
1982	Cand. scient. MSc Univ. Copenhagen
1985	Ph.D. actuarial mathematics Univ. Copenhagen
1989	Dr. Scient. actuarial mathematics Univ. Copenhagen

1981-88	Assistant and associate professor in actuarial mathematics, Univ. Copenhagen
1988-2016	Managerial posts in private insurance
2016	Associate professor, Copenhagen Business School

Ramlau-Hansen, H. (1983).Smoothing counting process intensities by means of kernel functions. Ann.Statist. 11, 453-466.

Ramlau-Hansen, H. (1983). The choice of a kernel function in the graduation of counting process intensities. Scand.Act.J. 165-182.

Borgan, Ø. & Ramlau-Hansen, H. (1985).Demographic incidence rates and estimation of intensities with incomplete information. Ann. Statist. 13, 564-582.

Back to Jan’s publications. Event history analysis interpretations of demographic concepts: the first paper was called ‘Probabilistic fertility models of the life table type’, and here the abstract indicates what it is; and most of you, I assume, will know what this is. One concept was something called purged measures, and this ‘purge’ really means conditioning, often conditioning on some final states. This is the same as ‘purged’ in the paper on purged and partial Markov chains.

Back to Jan’s publications

Event history analysis interpretations of demographic concepts

1970 Probabilistic fertility models of the life table type.

Theor. Popul. Biol. **1** (1), 12-38.

In this paper, a series of fertility models are presented as progressive extensions of the basic ideas behind the life table. Dimensions like age, parity, marital status, birth interval, and marital status duration are introduced in turn, and interrelations between the various models are indicated. The main arguments are developed in connection with a model which includes age- and parity-specificity.

The various net (influenced) and gross (partial) fertility measures are given some consideration. It turns out that retrospective fertility investigations give rise to a third kind of functions, which are denoted purged measures. These are parallel to but conceptually distinct from the gross measures. The introduction of the purged measures seems to throw some light on aspects of the theory which have previously appeared problematic.

The models are formulated in terms of transition probabilities rather than survivorship functions. In a final section, it is argued that the latter are super-fluous and even potentially harmful.

Note: ‘Purged’ measures are conditional measures, e.g. given survival to some age, cf. Hoem: ‘Purged and Partial Markov chains’, 1969

Event history analysis interpretations of demographic techniques

In a journal in Belgium – I do not know whether it exists any more – Biometrie-Praximétrie, Jan went through these models in great detail. There are applications to estimation in demography, and a Lexis diagram is presented. And this was the way I first learned from a Lexis diagram.

1969 Fertility rates and reproduction rates in a probabilistic setting. *Biométrie-Praximétrie* **10** (1), 38-66. Erratum: **11** (1), 20.

2. The model

§ 2.1. We shall let $\mu(x)$ be the force of mortality and $\varphi(x)$ be the force of fertility for an x year old parent. This will be taken to mean the following :

Let us observe a parent alive at age x during the age interval $(x, x + \Delta x)$ with $\Delta x > 0$. Then

(i) the probability that the parent will die in the age interval without giving birth to any children, equals $\mu(x) \Delta x + o(\Delta x)$, where $o(\Delta x)/\Delta x \to 0$ as $\Delta x \to 0$,

(ii) the probability that the parent will have exactly one birth in the age interval and survive to age $x + \Delta x$, equals $\varphi(x) \Delta x + o(\Delta x)$,

(iii) the probability that the parent will survive to age $x + \Delta x$ without giving birth to any children in the age interval, equals $1 - [\mu(x) + \varphi(x)] \Delta x + o(\Delta x)$, and

(iv) the probability that the parent gives more than one birth, or gives at least one birth and then dies within the age interval, is $o(\Delta x)$.

We shall assume that $\mu(.)$ and $\varphi(.)$ are continuous functions for $x \geqq 0$, with $\mu(x) > 0$ for $0 \leqq x < \omega$, $\varphi(x) > 0$ for $\omega_1 < x < \omega_2$, $\varphi(x) = 0$ otherwise. Thus the fertile period is the period where $\varphi(x) > 0$. Multiple births will be taken care of at a later stage. (See § 2.7.)

See comments on slide 19.

6. Some applications to estimation in demography

§ 6.1. The application of the results of the previous chapters to demographic estimation is fairly straightforward. We shall give three examples to indicate how this may be done. Adaptation to other situations can be made ad hoc.

See comments on slide 19.

To fix our ideas, we shall concentrate on a situation where it is desired to investigate female fertility with respect to (live) children of both sexes in a population. Let x be an integer. We define

$L_x^{(N)}$ as the number of women of age x in the population on January 1 of the year N,

$M_x^{(N)}$ as the number of women who experience their x-th birthday in the population in the year N,

$C_x^{(N)+}$ as the number of live children born in the year N by mothers who experience their x-th birthday in that year and who have the birth no earlier than this birthday, and

$C_x^{(N)-}$ as the number of live children born in the year N by x-year old mothers who have the birth before their own birthday in that year (1).

See comments on slide 19.

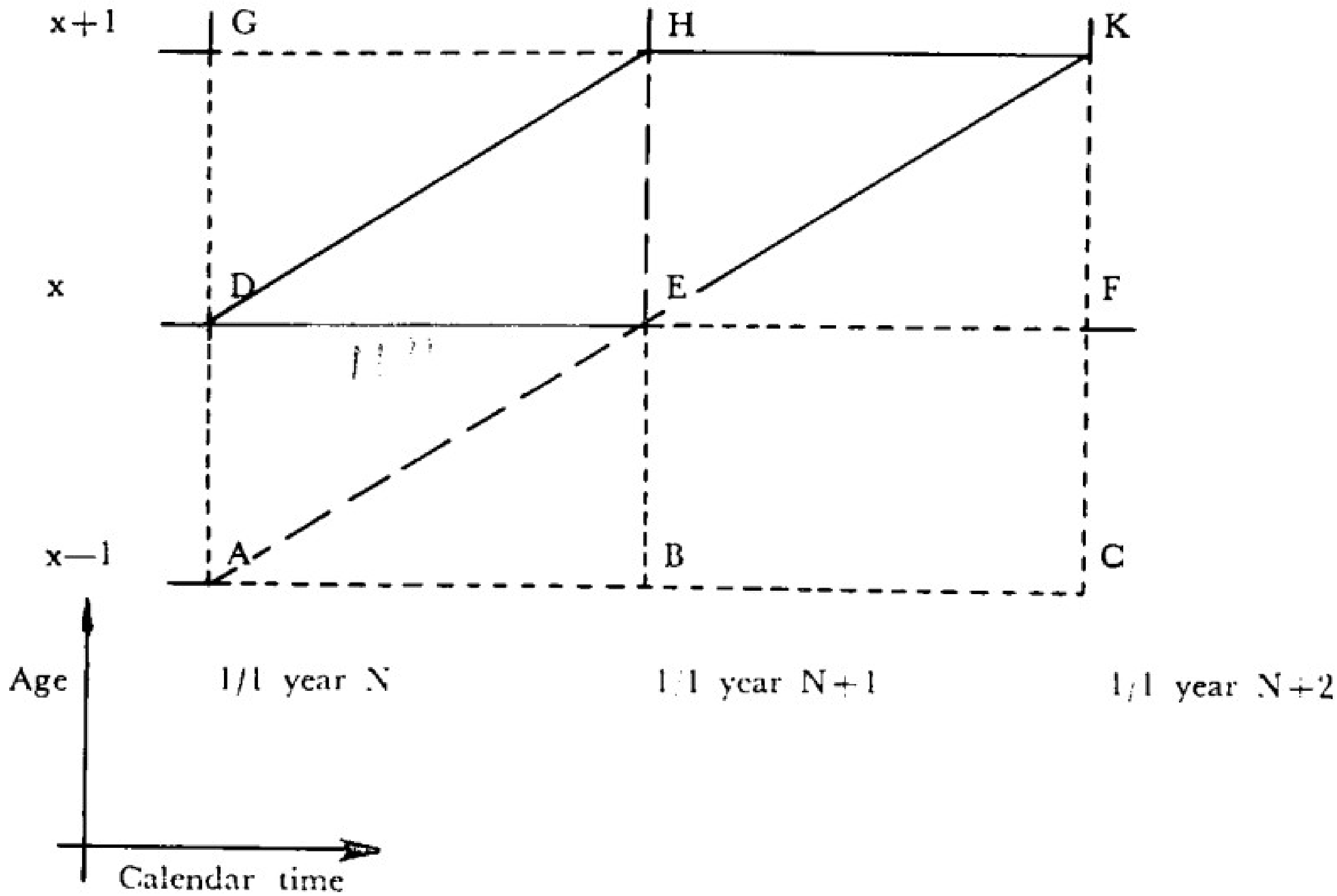


FIG. 6.1. — Section of the Lexis diagram

See comments on slide 19.

§ 6.2. (*Age year method*). It may be desired to analyse fertilty in the parallelogram DEKH. Let $f_x^{(N)}$ be the quantity corresponding to the f_x of (3.5) for a woman whose lifeline has points in the parallelogram, let $f_x^{(N)'}$ be the quantity corresponding to the f'_x of (3.7), and let $\mu_x^{(N)}$ be her force of mortality (regarded as a constant parameter throughout the parallelogram). By (4.3) and § 4.4 one would estimate these three quantities by

$$\hat{\mu}_x^{(N)} = \frac{D_x^{(N)+} + D_x^{(N+1)-}}{U_x^{(N)+} + U_x^{(N+1)-}},$$

$$\hat{f}_x^{(N)'} = \frac{C_x^{(N)+} + C_x^{(N+1)-}}{U_x^{(N)+} + U_x^{(N+1)-}} \tag{6.1}$$

and $\hat{f}_x^{(N)} = \hat{f}_x^{(N)'} \varkappa(\hat{\mu}_x^{(N)})$.

If the data are not available in a form which gives the relatively detailed information required in these estimation formulae, various approximation techniques can be used. For example

$1/2\ [M_x^{(N)} + M_{x+1}^{(N+1)}]$ may be chosen as an approximation to $U_x^{(N)+} + U_x^{(N+1)-}$.

If not even the $M_x^{(N)}$ are available, their values again may be approximated by similar techniques. One may e.g. choose $1/2\ [L_{x-1}^{(N)} + L_x^{(N+1)}]$ as an approximation to $M_x^{(N)}$. In this case, $U_x^{(N)+} + U_x^{(N+1)-}$ will be approximated by

$$\frac{1}{4}\ L_{x-1}^{(N)} + \frac{1}{2}\ L_x^{(N+1)} + \frac{1}{4}\ L_{x+1}^{(N+2)}.$$

The quantity corresponding to the $\tilde{f}_x$ of § 5.1 will be

$$\tilde{f}_x^{(N)} = \frac{C_x^{(N)+} + C_x^{(N+1)-}}{M_x^{(N)}} \tag{6.2}$$

See comments on slide 19.

§ 6.3. (*Calendar year method*). Let $*f_x^{(N)}$, $*f_x^{(N)'}$ and $*\mu_x^{(N)}$ be the quantities corresponding to the f_x of (3.5), the f_x' of (3.7), and the force of mortality, respectively, for a woman whose lifeline has points in AEHD. Then

$$*\hat{\mu}_x^{(N)} = \frac{D_{x-1}^{(N)-} + D_x^{(N)+}}{U_{x-1}^{(N)-} + U_x^{(N)+}}, \tag{6.3}$$

$$*\hat{f}_x^{(N)'} = \frac{C_{x-1}^{(N)-} + C_x^{(N)+}}{U_{x-1}^{(N)-} + U_x^{(N)+}},$$

and $$*\hat{f}_x^{(N)} = *\hat{f}_x^{(N)'} \varkappa(*\mu_x^{(N)}).$$

If necessary, $U_{x-1}^{(N)-} + U_x^{(N)+}$ may be approximated by $M_x^{(N)}$ or by $1/2\,[L_{x-1}^{(N)} + L_x^{(N+1)}]$. The relation corresponding to (6.2) is

$$*\hat{f}_x^{(N)} = \frac{C_{x-1}^{(N)-} + C_x^{(N)}}{L_{x-1}^{(N)}}. \tag{6.4}$$

§ 6.4. Let $\sharp f_x^{(N)}$, $\sharp f_x^{(N)'}$ and $\sharp\mu_x^{(N)}$ be the quantities corresponding to the f_x of (3.5), the f_x' of (3.7), and the force of mortality, respectively, for a woman whose lifeline has points in DEHG. Then

$$\sharp\hat{\mu}_x^{(N)} = \frac{D_x^{(N)+} + D_x^{(N)-}}{U_x^{(N)+} + U_x^{(N)-}},$$

$$\sharp\hat{f}_x^{(N)'} = \frac{C_x^{(N)+} + C_x^{(N)-}}{U_x^{(N)+} + U_x^{(N)-}}, \tag{6.5}$$

and $$\sharp\hat{f}_x^{(N)} = \sharp\hat{f}_x^{(N)'} \varkappa(\sharp\hat{\mu}_x^{(N)}).$$

An approximation to $U_x^{(N)+} + U_x^{(N)-}$ suggested by SVERDRUP (1961) is

$$1/6\, L_{x-1}^{(N)} + 1/3\, L_x^{(N)} + 1/3\, L_x^{(N+1)} + 1/6\, L_{x+1}^{(N+1)}.$$

It is an essential feature of the present situation that not all z_j are

We went through so carefully that one of my colleagues, Søren Tolver Jensen, discovered that one of the approximation formulae was actually not optimal. If you looked up Sverdrup’s lecture notes, you could see that he has a different approximation formula. So Jan felt compelled to publish this correction note with reference to Søren Tolver Jensen.

Correction Note

Correction to “Fertility rates and reproduction rates in a probabilistic stetting” by Jan M. Hoem (*Biométrie-Praximétrie*, 1969, **10** (1), pp. 38-66).

In § 6.2 of the paper, the approximations

$$M_x^{(N)} \approx \tfrac{1}{2}\left(L_{x-1}^{(N)} + L_x^{(N+1)}\right) \text{ and } U_x^{(N)+} + U_x^{(N+1)-} \approx \tfrac{1}{2}\left(M_x^{(N)} + M_{x+1}^{(N+1)}\right)$$

were combined to give the formula

$$(1) \qquad U_x^{(N)-} + U_x^{(N+1)} \approx \tfrac{1}{4} L_{x-1}^{(N)} + \tfrac{1}{2} L_x^{(N+1)} + \tfrac{1}{4} L_{x+1}^{(N+2)}.$$

It turns out that one can do better than this. According to Sverdrup (1961, p. 54, (15) and (16); a reference given in the paper),

$$(2) \qquad U_x^{(N)-} \approx \tfrac{1}{3} L_x^{(N)} + \tfrac{1}{6} L_{x+1}^{(N+1)},$$

and

$$(3) \qquad U_x^{(N)+} \approx \tfrac{1}{6} L_{x-1}^{(N)} + \tfrac{1}{3} L_x^{(N+1)}.$$

In place of (1), this gives

$$(4) \qquad U_x^{(N)-} + U_x^{(N+1)} \approx \tfrac{1}{6} L_{x-1}^{(N)} + \tfrac{2}{3} L_x^{(N+1)} + \tfrac{1}{6} L_{x+1}^{(N+2)}.$$

The author is grateful to Mr. Søren Tolver Jensen, who suggested (4).

In the journal of The Royal Statistical Society, we see for the first time in this literature a diagram with the boxes and the arrows.

1971 Point estimation of forces of transition in demographic models. *J. Roy. Statist. Soc. B* **33** (2). 275-289.

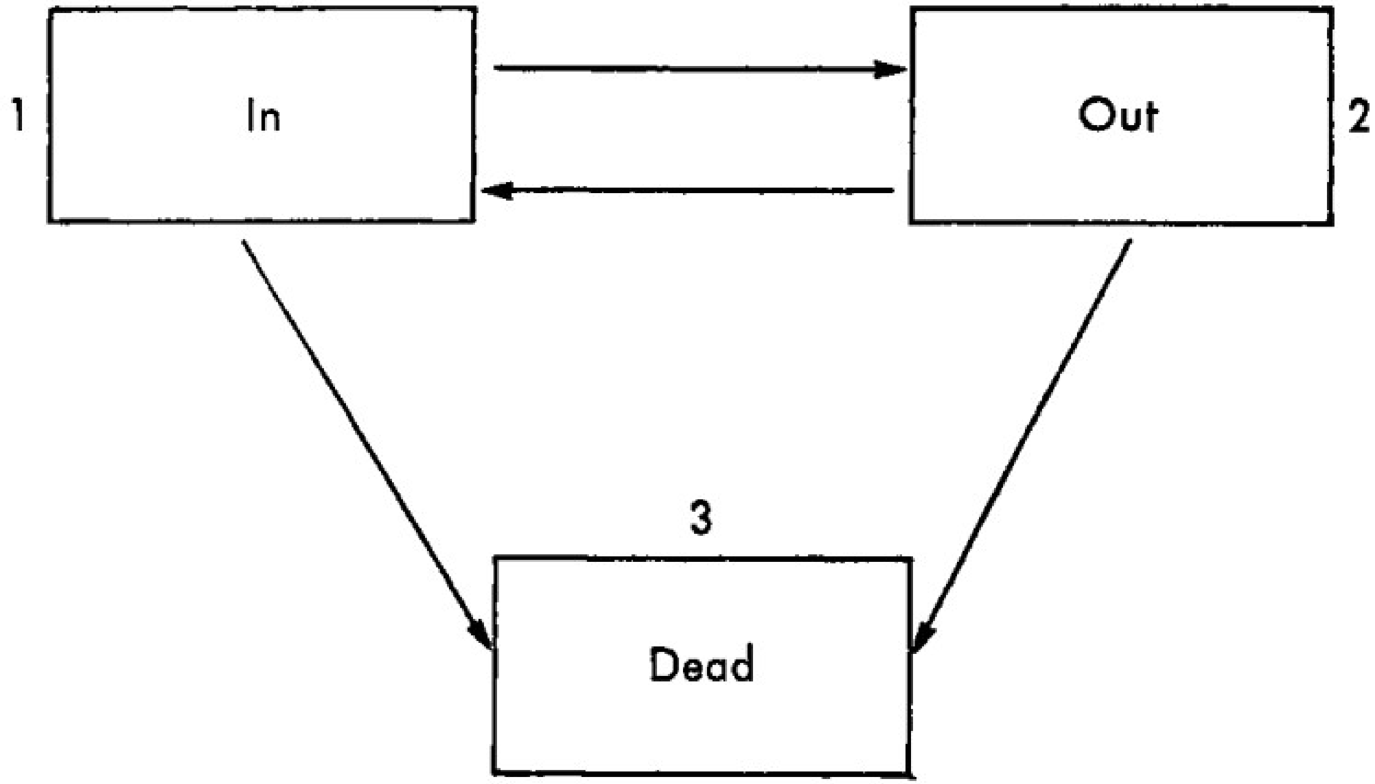


FIG. 1. The states of the model of work-force participation.

Five years later, Jan was persuaded to give a survey of his work to the Scandinavian statisticians, and there were by then a lot of these figures. I do not think that this is new to you at all, but from a historical point of view, it is of some interest to see when these diagrams were first published, first used. Nowadays, we have them everywhere, but it was not like that back in the seventies.

Event history analysis interpretations of demographic techniques

1976 The statistical theory of demographic rates: A review of current developments (with discussion). *Scand. J. Statist.* **3** (4), 169-185.

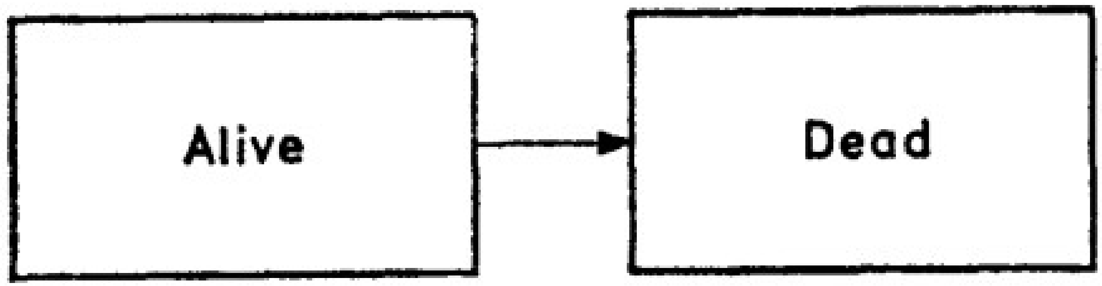


Fig. 1. Pure mortality.

See comments on slide 27.

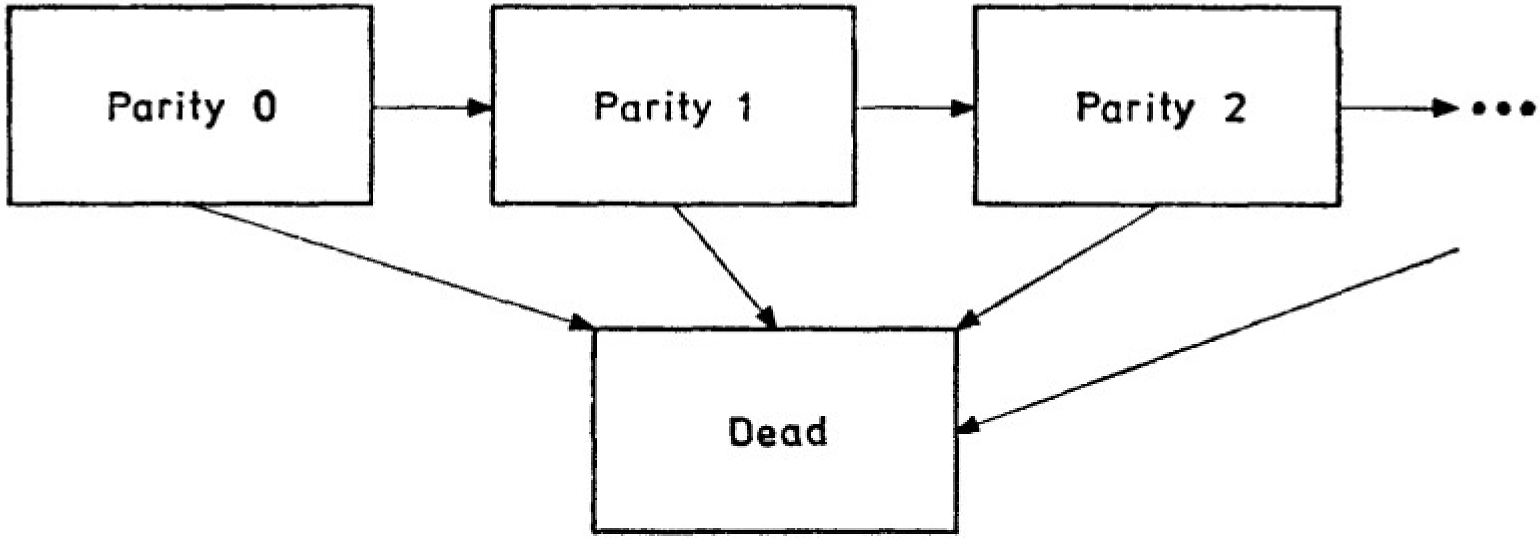


Fig. 2. Fertility statuses.

See comments on slide 27.

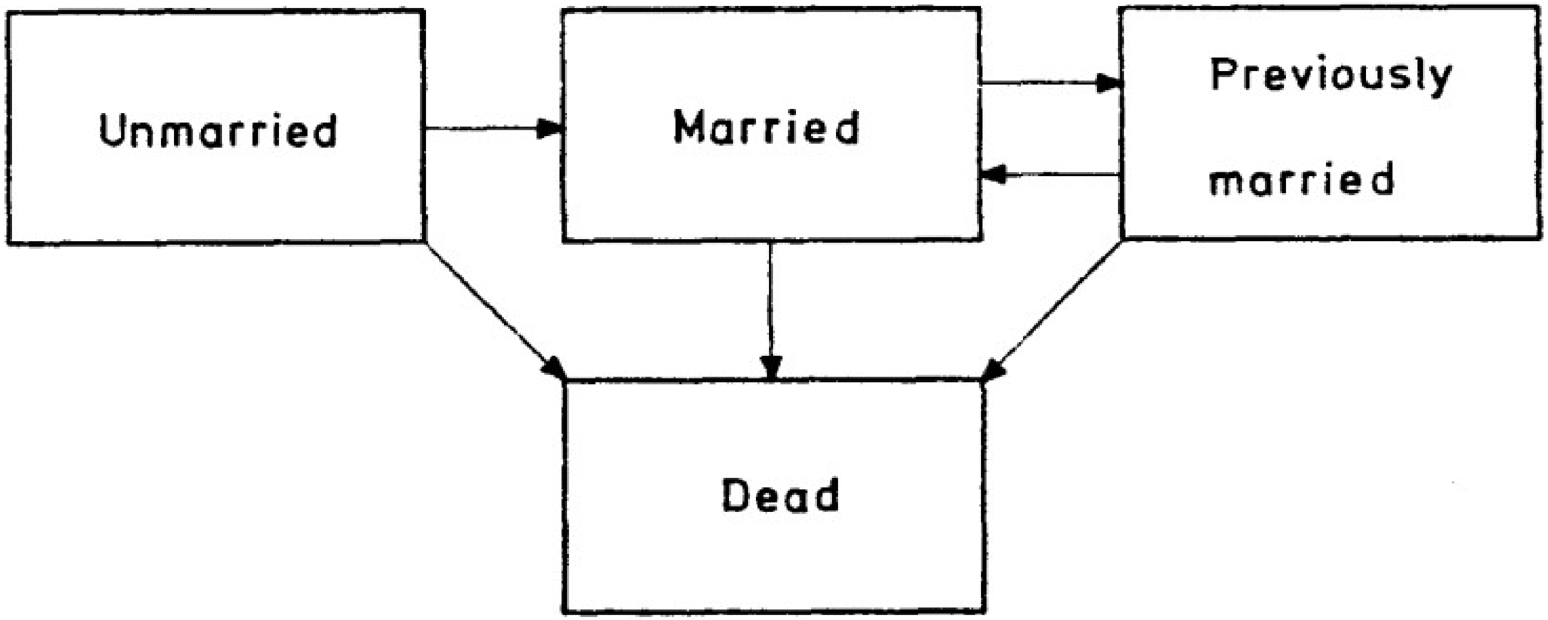


Fig. 3. Marital statuses.

See comments on slide 27.

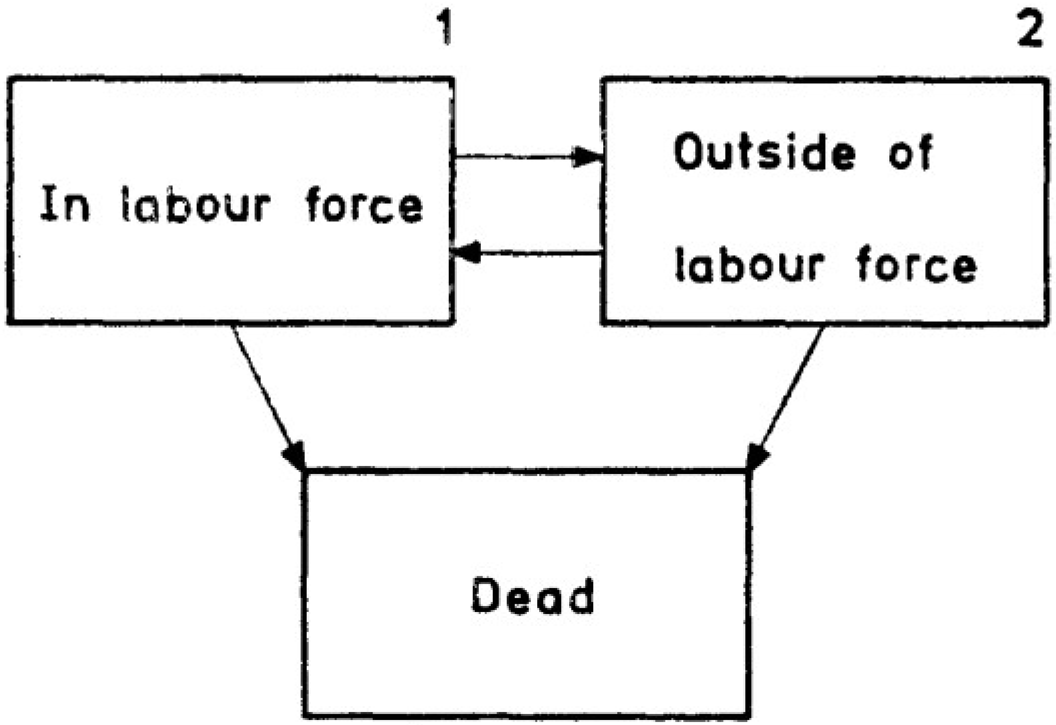


Fig. 4. Labour force statuses.

See comments on slide 27.

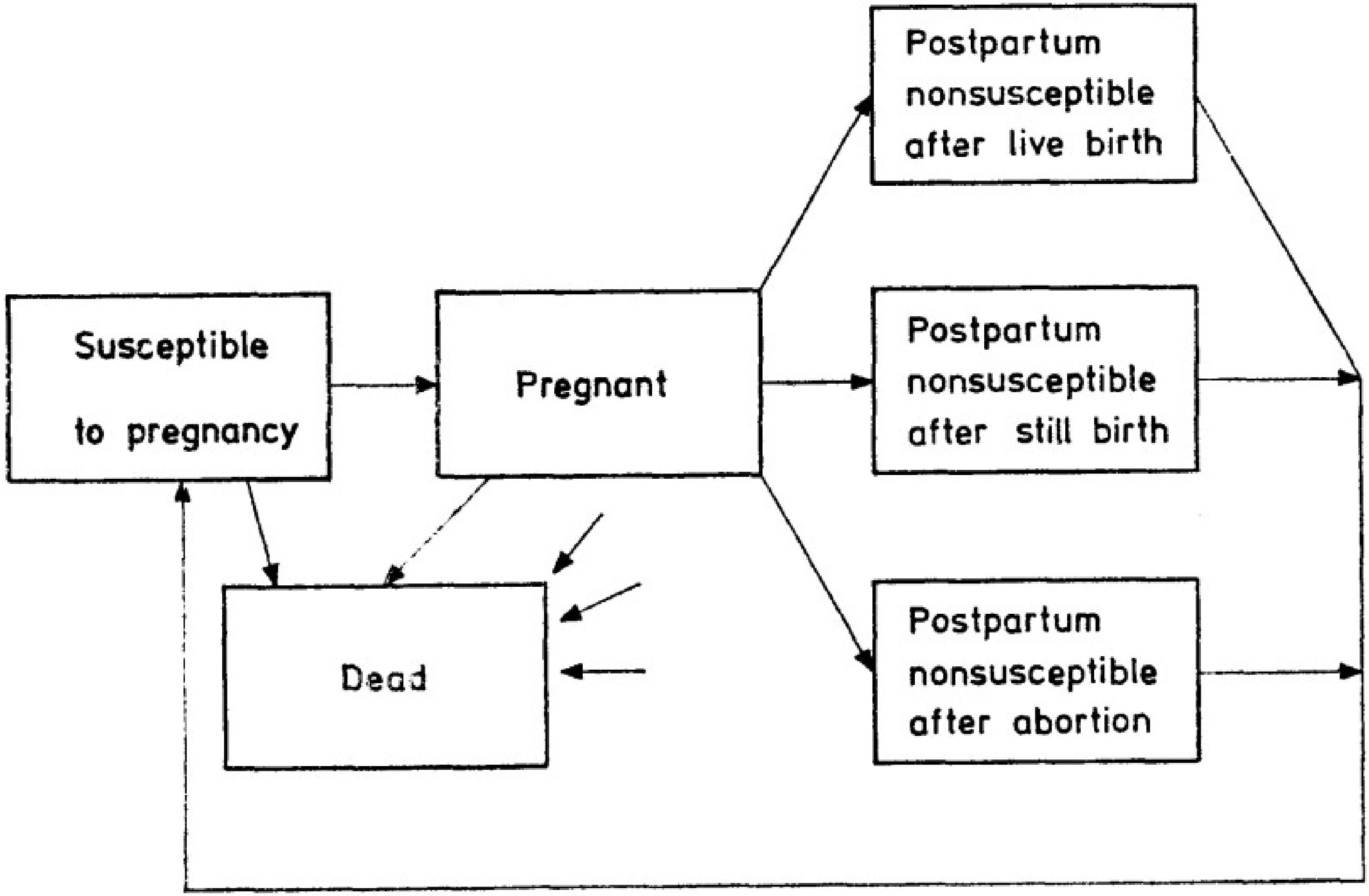


Fig. 5. **Human reproduction.**

See also comments on slide 27.
Note this diagram on intrauterine devices inspired by Odd Aalen's master's thesis. Now these diagrams are all mainstream.

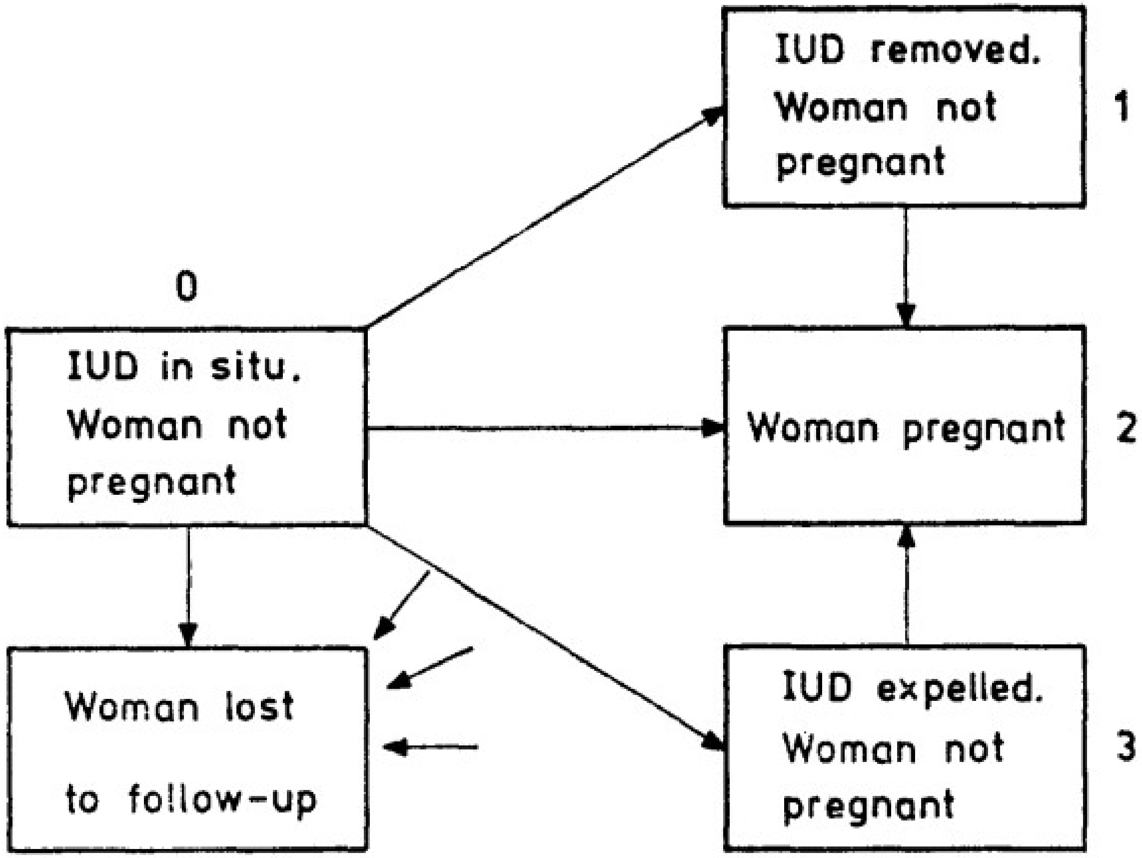


Fig. 6. The use-effectiveness of an intra-uterine contraceptive device.

Jan also worried about the meaning of randomness in demographic models. An issue that has long been settled among statisticians, but that still seems to arise occasionally among demographers, is the question of whether statistical significance is relevant to data that cover a complete population, or whether it only pertains to the sampling error that arises in sample surveys. Jan once reflected on this issue by stating ‘firmly that individual life histories are seen most fruitfully as realizations of stochastic processes each of which is subject to random variation, and that this should be taken into account even when the set of observations contains all members of a population or population segment’. I do agree so far, but some problems follow that Jan – as far as I can tell – never really took up; in particular, overdispersion.

Stochastic processes and the full population: Goodness of fit?

2008: The reporting of statistical significance in scientific journals. Reflexion. *Demographic Research* **18** (15), 437-442.

‘…..an issue that has long been settled among statisticians but that still seems to arise occasionally among demographers, namely the question whether statistical significance is relevant to data that cover a complete population or whether it only pertains to the sampling error that arises in sample surveys. I want to state firmly that individual life histories are seen most fruitfully as realizations of stochastic processes each of which is subject to random variation, and that this should be taken into account even when the set of observations contains all members of a population or population segment.’

<u>*Problem*</u>*: most of these stochastic processes are piecewise constant Poisson processes. Their variance is by assumption equal to the mean, and is estimated from the empirical mean, not from the empirical variability.*

Possible tool to capture overdispersion: the ‘sandwich estimator’ based on the empirical variability.

Daniel Courgeau has written a fascinating historical paper, again in the French online journal History of Probability and Statistics. The English version is called ‘Dispersion of measurements in demography: a historical view’. It is a very interesting discussion of where those randomnesses come in at various times through history. Speaking of demography and other works, Courgeau, for example in this book on multilevel synthesis, has started using multilevel modelling for demography; and this is something I do not think Jan was doing.

Stochastic processes and the full population:

Courgeau on random heterogeneity and multilevel models

Courgeau D. 2010. Dispersion of measurements in demography: a historical view. *Electr. J. Hist. Prob. Stat.* **6:**1–19

Courgeau D. 2010. *Multilevel Synthesis. From the Group to the Individual.* Dordrecht, Neth.: Springer

My final example is to illustrate Schweder's brilliant idea of local dependence in the application that we did, with good advice from Jan. There are textbook presentations of Schweder's idea by Courgeau and Lelièvre, by Blossfeld and Rohwer, and in our monograph.
The model says that we have two life events, A and B. We assume that the intensity that B happens is the same whether or not A has happened, so the intensity that B happens before A is the same as the intensity that B happens after A, but the intensity that A happens before B is smaller than the intensity that A happens after B. This way, one can combine two events, and one gets an asymmetric dependence context that is made possible by including time. Schweder did have some hypothetical examples in his paper in the Journal of Applied Probability, but we had a concrete problem that we started with and that we wanted to solve.

Example: Local dependence between life history events

Idea: Schweder (1970)

Textbook presentations:

Courgeau D, Lelièvre E. 1992. *Event History Analysis in Demography*. Oxford, UK: Clarendon

Andersen PK, Borgan Ø, Gill RD, Keiding N. 1993. *Statistical Models Based on Counting Processes.* New York: Springer

Blossfeld H, Rohwer G. 1995. *Techniques of Event History Modeling.* Mahwah, NJ: Lawrence Erlbaum

See comments on slide 35.

Composable Markov processes, local dependence

T. Schweder (1970), *J.Appl.Prob.*, **7**, 400-410.

$$\begin{array}{ccccc} & & \alpha_{OB}(t) & & \\ & \text{O} & \longrightarrow & \text{B} & \\ \alpha_{OA}(t) & \downarrow & & \downarrow & \alpha_{B,AB}(t) \\ & \text{A} & \longrightarrow & \text{AB} & \\ & & \alpha_{A,AB}(t) & & \end{array}$$

Assume $\alpha_{OB} = \alpha_{A,AB}$ but $\alpha_{OA} < \alpha_{B,AB}$

A locally dependent on B

B not locally dependent on A

Asymmetric dependence concept made possible by including time

See comments on slide 35.

Example: Interaction between life history events

Pustulosis palmo-plantaris and menopause

Aalen, Borgan, Keiding, Thormann (1980), *Scand.J.Statist.* **7**, 161-171.

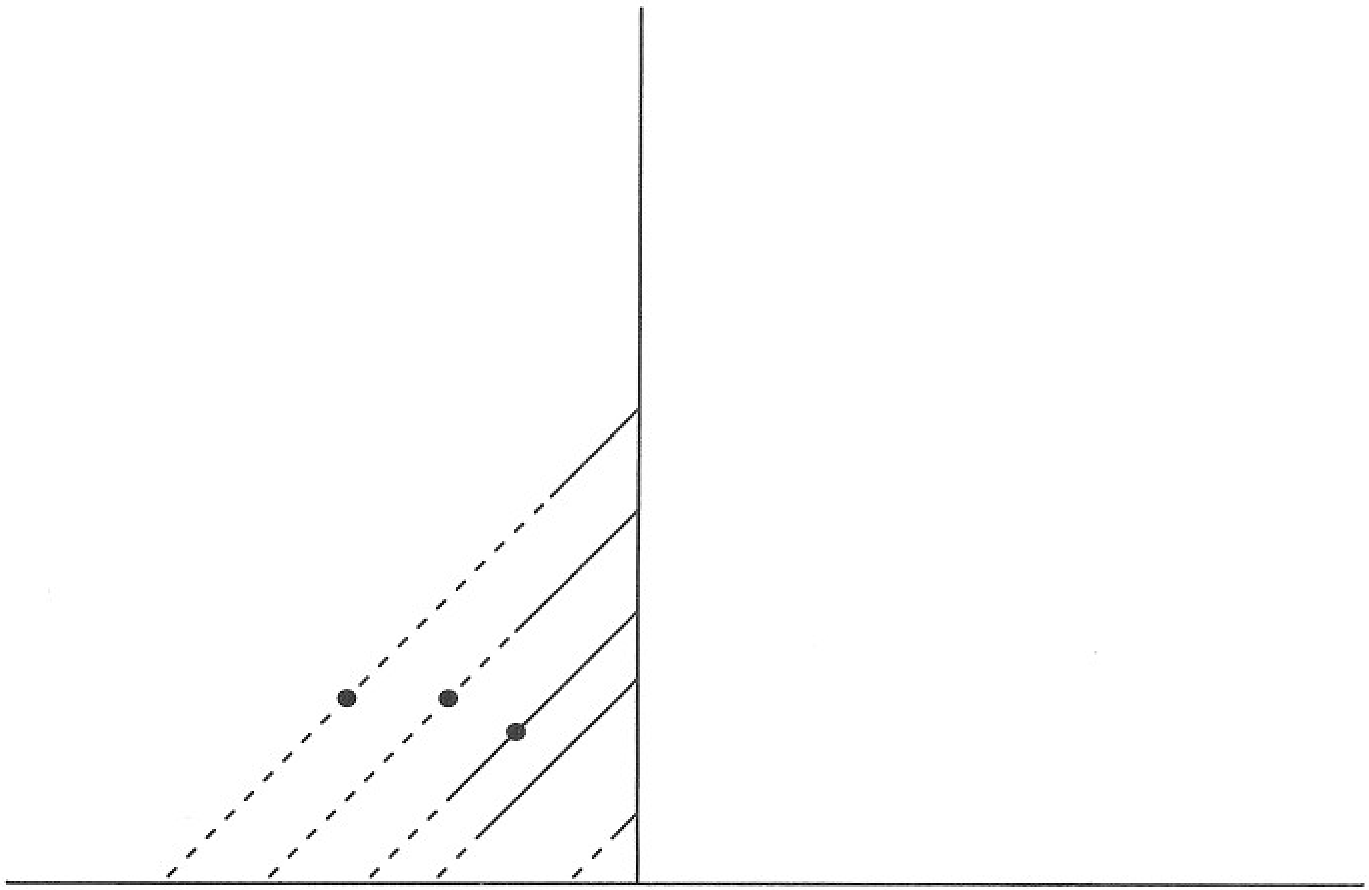
Pustulosis palmo-plantaris: chronically recurrent skin disease localized to palms of the hands and soles of the feet.

Prevalence .05 percent.

Most common among women, first appearance usually between 35 and 55 years of age. Unknown etiology. *Menopause*?

Data 100 consecutive prevalent patients from out-patient clinic.
85 women considered.

From an out-patient dermatological clinic, we got data concerning female patients with a chronic recurring skin disease. This is a Lexis diagram trying to illustrate that for each woman we had the first occurrence of the disease; this is a dot which could be before menopause or after menopause. (The full line is after menopause.) So some of the occurrences were before menopause, and some were after menopause.



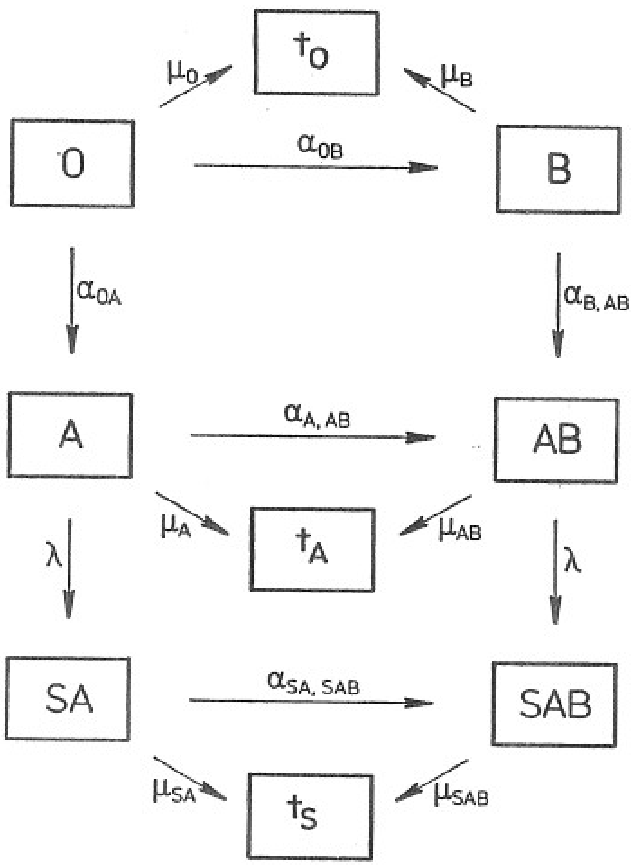
But the data were collected retrospectively, so the question is how one could analyse whether there was any influence of meno–pause on the incidence of this disease. Data came from 85 female patients with first appea–rance, menopause, last seen. And it is one of these satisfactory examples where the whole dataset is just in this table, but the work it takes to really get through these data runs quite a bit more than a page.

Table I. *Age of first appearance, age of natural or induced (*) menopause and age last seen for 85 female patients with pustulosis palmo-plantaris*

First appear.	Meno-pause	Last seen	First appear.	Meno-pause	Last seen	First appear.	Meno-pause	Last seen	First appear.	Meno-pause	Last seen
17	—	18	40	50	74	49	46	49	54	43*	56
19	—	22	40	47	48	49	49	55	56	50	63
23	—	33	40	50	52	49	45	74	57	47	64
24	—	26	40	46	59	50	52	54	57	37	60
27	—	38	40	51*	55	50	44*	53	58	48	60
27	—	40	41	—	41	50	50*	50	59	50	61
29	—	30	41	40	54	50	45*	50	59	47	64
29	—	32	41	38*	43	51	50	53	60	50	62
30	—	31	43	43	44	51	48*	52	60	51	62
30	—	32	43	47	50	51	49	51	61	55	66
31	—	33	44	—	46	51	48	66	62	54*	67
31	—	35	44	—	44	51	46	62	64	48	66
33	—	37	45	46*	47	52	47	64	64	42	68
34	31*	69	45	45*	47	53	48	56	64	40	66
35	—	36	45	43	47	53	48	54	64	49*	64
35	32*	67	45	—	46	53	52	56	64	47*	65
36	—	37	45	45	55	53	46*	56	66	47	68
36	—	40	47	43	51	53	50*	54	67	52*	71
37	28*	38	48	49	70	53	43	54	67	49	79
37	—	37	48	40	50	54	48	54	70	45	71
38	39	39	48	47	49	54	46	54	73	45	74
39	37*	42									

So, we generalised the basic model of Schweder by taking account of the retrospective sampling. We conditioned on terminating in the sampled state using Jan's paper on purged and partial Markov chains, which shows that we still get a Markov chain with new intensities. Therefore, we were able to get qualitative information on the original intensity rates: the intensity before menopause and the intensity after menopause. We test the equality of the corresponding intensities in the conditional Markov Chain. If we reject this hypothesis, this means menopause increases the risk of the disease. We performed this test, and we had to develop a three-sample version of the test because we not only had menopause, we also had induced menopause; that is, if you operate with the medical treatment you could create an artificial menopause. The tests showed clearly significant effects.

Cross-sectional sampling of prevalent cases



Condition on terminating in SA or SAB ($\Leftrightarrow X(\infty) = \dagger_S$) leads to new Markov process with intensities

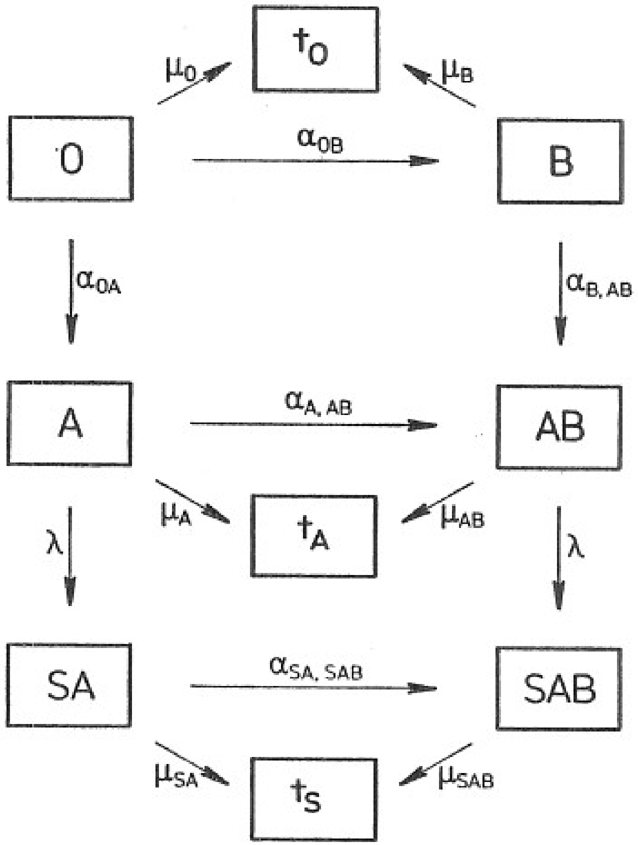
$$\gamma_{ij}(t) = \alpha_{ij}(t)\frac{\Pi_j(t)}{\Pi_i(t)}$$

where $\Pi_k(t) = P_{k\dagger_S}(t,\infty)$
$= P$(patient in state k at age t will eventually get sampled).

Notice Horvitz-Thompson weights

See comments on slide 40.

Retrospective inference from prevalent case-only data



$\gamma_{ij}(t)$ complex function of $\alpha_{ij}(t)$

$\Rightarrow$ estimation hard

Under non-differential mortality

$\mu_0 \equiv \mu_B \equiv \mu_A \equiv \mu_{AB}$ it holds that

$\alpha_{0A}(t) = \alpha_{B,AB}(t) \Rightarrow \gamma_{0A}(t) = \gamma_{B,AB}(t)$

$\alpha_{0B}(t) = \alpha_{A,AB}(t) \Rightarrow \gamma_{0B}(t) = \gamma_{A,AB}(t)$

$\Rightarrow$ (conservative) hypothesis tests direct

Results:

$\alpha_{0B}(t) = \alpha_{A,AB}(t)$ accepted

(Pustulosis does not affect menopause)

$\alpha_{0A}(t) = \alpha_{B,AB}(t)$ rejected

(Menopause increases risk of pustulosis)

I include here the beginning of a careful pencil-written memo by Jan on the use of Schweder's model for this study.

Hoem 2/2-78.

1

Kjære Niels,

Pustulose-prosjektet.

Jeg har tenkt litt på hva det egentlig er dere estimerer på grunnlag av dataene i tabell 1 på s. 9 i KUIMS Preprint 1977/8, og her er resultatet av mine funderinger.

La oss ta utgangspunkt i modellen på s2 i artikkelen, skrevet på flg. måte:

$$\begin{tikzcd} \boxed{O}^{\,0} \arrow[r, "\kappa(x)"] \arrow[d, "\delta(x)"] & \boxed{M}^{\,1} \arrow[d, "\Delta(x)"] \\ \boxed{D}^{\,2} \arrow[r, "\kappa(x)"] & \boxed{DM}^{\,3} \end{tikzcd}$$

See comments on slide 42.

Test of $\gamma_{0D} \equiv \gamma_{I,ID} \equiv \gamma_{M,MD}$.

Three-sample version of generalized log rank test. $\chi^2 = 22.4 \quad f = 2$.

$$\text{If} \quad \alpha_{I,ID} \equiv \alpha_{M,MD} \geq \alpha_{0D} \quad \text{then}$$
$$\gamma_{I,ID} \equiv \gamma_{M,MD} \geq \gamma_{0D}$$

and if the α's are equal, then so are the γ's.

After the rejection of $\alpha_{0D} \equiv \alpha_{I,ID} \equiv \alpha_{M,MD}$

we cannot compare α_{0M} with $\alpha_{D,MD}$

or α_{0I} with $\alpha_{D,ID}$

This was a surprising medical result obtained using a non-standard statistical technique. Fortunately, we were able to repeat the analysis on an independently collected sample where the conservative three-sample test gave similar results, which we finally got published in the book, The Demography of Europe, which was published in commemoration of the departure of Jan from MPIDR here in Rostock.

Surprising medical result, obtained by non-standard statistical technique.

Repeat the analysis on independently collected sample:

Marie Cramers, Marselisborg Hospital, Aarhus

70 women   11 men

The conservative three-sample test this time yielded P = 0.032 . For details see

Keiding, N. (2013). Event history analysis: Local dependence and cross-sectional sampling. In: *The Demography of Europe* (ed. G. Neyer, G. Andersson, H. Kulu, L. Bernardi & C. Bühler). Dordrecht: Springer, 207-220.

See comments on slide 44.

Table 9.1 Age of first appearance, age of menopause or induced (*) menopause and age last seen for 70 women with pustulosis palmo-plantaris seen at Marselisborg Hospital

First appear	Menopause	Last seen
14	–	16
20	–	30
21	–	22
21	–	27
21	–	50
22	–	27
24	–	28
25	–	43
32	–	36
33	52	62
35	–	41
35	–	45
35	36*	69
37	–	37
40	–	47
40	45	60
42	37	45
42	40	58
43	33	58
44	–	50
45	–	47
45	–	48
45	49	53
46	45	48
46	52	70
47	46	48
47	44	52
47	49	54
48	–	52
48	40	63
49	52	52
49	–	51
50	49	51
50	49	52
50	49	60
51	–	51
51	30*	53
51	45*	54
51	48	56
51	50	64
52	47	54
52	43	57
52	51	64
53	43	53
53	46*	54
54	53	56
54	50	65
55	36	62
55	50*	59
56	37	57
56	53	57
56	49	58
56	50	69
57	48	58
57	51	67
58	48*	59
58	46	61
59	56	60
59	49	79
60	44	62
61	51	63
63	41	66
63	49	71
63	50	63
65	59	70
70	45	71
70	51*	72
70	35*	79
73	50	74
74	55	78

(continued)

Table 9.1 (continued)

of marriage and migration, using the modelling presented here. To conclude, there are two issues to consider: (1) the conceptual modelling of the interaction pattern, going back to Schweder (1970) and being very similar in spirit to that of Granger causality, and (2) the actual statistical analysis of these models, taking into account the precise observational pattern. Once the modelling is in place, most of the analysis may be performed by means of standard event history analysis allowing for censoring (generalized survival analysis), now a standard tool in demography. However, possible non-standard sampling patterns may require modifications, such as reviewed above and illustrated by Aalen et al. (1980).

9.4 Case Series Analysis Approach

In this section, I outline an alternative approach to the analysis of *'case-only data'* such as those of the skin disease patients analysed by Aalen et al. (1980).

See comments on slide 44.

Case series analysis

CP Farrington & HJ Whitaker (2006). Semiparametric analysis of case series data. *Appl. Statist.* **55**, 553-594 (with discussion)

Assume that natural menopause modifies the incidence of pustulosis palmo-plantaris by a multiplicative factor μ and induced menopause modifies it by a factor ν, then these methods yield

		original data (Thormann)	confirmatory study (Cramers)
natural menopause	$\hat{\mu}$	4.2 (1.2, 14.9)	3.4 (1.0, 11.2)
induced menopause	$\hat{\nu}$	14.0 (1.9, 104.4)	1.6 (0.2, 17.3)

T^2-test for identity of original and confirmatory study: $P = 0.4$.

The local independence concept has been generalised in very fruitful work by Vanessa Didelez, now a professor in Bremen.

Local independence: general formulation in the bivariate case

Mykland (1986). Res. Rep. Bergen

Aalen (1987). Scand.Act.J.

Two processes $Y_i(t)$ $i = 1, 2$ with histories $\left(\mathcal{F}_t^i\right)$

where $\mathcal{F}_t^i = \sigma\left(Y_i(s)|0 \leq s \leq t\right), \mathcal{F}_t = \mathcal{F}_t^1 \vee \mathcal{F}_t^2$,

compensators $\Lambda_i(t) : M_i(t) = Y_i(t) - \Lambda_i(t)$ is martingale wrt $\left(\mathcal{F}_t^i\right)$.

Assume $M_1(t)$ and $M_2(t)$ are *orthogonal*

(actually a no unmeasured confounder assumption).

Definition. Y_1 locally independent of Y_2 $(Y_2 \not\rightarrow Y_1)$ if $\Lambda_1(t)$ is measurable wrt $\mathcal{F}_t^1$ for all t.

See comments on slide 47.

Local independence: multivariate case

Didelez 2001

k-variate process $(Y(t)) = (Y_1(t), \cdots, Y_k(t))$.

Define subprocesses $Y_A(t) = (Y_i(t), i \in A)$ for $A \subset \{1, \cdots, k\}$. Assume histories given by $\mathcal{F}_t^i = \sigma\left(Y_i(s) | 0 \leq s \leq t\right)$ and define $\mathcal{F}_t^A = \vee_{i \in A} \mathcal{F}_t^i$. For all subsets $A, B \subset \{1, \cdots, k\}$ define (vector) compensators Λ_A, Λ_B and assume that the martingales $Y_A - \Lambda_A$ and $Y_B - \Lambda_B$ are orthogonal.

Definition. Y_B is *locally independent of* Y_A *given* Y_C if all $\mathcal{F}_t^{A \cup B \cup C}$-compensators $\Lambda_i, i \in B$, are measurable wrt $\mathcal{F}_t^{B \cup C}$. Write $Y_A \not\to Y_B | Y_C$ or $A \not\to B | C$. Otherwise Y_B is *locally dependent of* Y_A *given* Y_C.

See comments on slide 47.

Didelez's graphical model theory for event history processes

Let $V = \{1, \cdots, K\}$ be a local independence graph: directed (not necessarily acyclic) graph E defined by *the pairwise dynamic Markov property:*

$$\text{no edge from } j \text{ to } k \text{ ©} \quad Y_j \; 6 \not\rightarrow Y_k | Y_{V \setminus \{j,k\}}.$$

Local dynamic Markov property:

$\sum i \; a \;\; V : V \setminus \text{closure}(i) \; 6 \not\rightarrow \{i\} \mid \text{parents}(i)$

Global dynamic Markov property: For subsets *A, B, C* ¢ *V*

C δ-separates *A* from *B* in the directed graph $A \not\rightarrow B|C$.

δ-separation is a generalization of d-separation to directed graphs.

Theorem. Under regularity conditions the three Markov properties are equivalent.

Let me conclude:
Jan has shown us how to use event history analysis in analysing demographic data. He was generous in guiding and promoting promising young statisticians with an interest in demography and actuarial mathematics. He displayed enormous energy in carrying out a large body of empirical work using these methods.

THANK YOU, JAN!

Jan Hoem

Showed many of us how to use event history analysis in analysing demographic data

Was generous in guiding and promoting promising young statisticians with

interests in demography and actuarial mathematics

Displayed enormous energy in carrying through large bodies of empirical studies using these methods

Thank you, Jan!

THANK YOU, JAN!

Keiding, N. (2014). Event history analysis. *Annu. Rev. Stat. Appl.* **1,** 333-360.