Interne notater

STATISTISK SENTRALBYRÅ

79/6 25. juli 1979

ET OPPLEGG FOR ANALYSE AV KOHORTDØDELIGHETEN I NORGE ETTER 1845

Av

Jens-Kristian Borgan[x)]

INNHOLD

Side

---

x) Diskusjoner med Ørnulf Borgan har vært nyttige ved utarbeidingen av dette notatet. Knut Ø. Sørensen har lest gjennom manuskriptet og gitt nyttige kommentarer.

ET OPPLEGG FOR ANALYSE AV KOHORTDØDELIGHETEN I NORGE ETTER 1845

1. Innledning

De overlevelsetabellene som tidligere er utarbeidd i Statistisk Sentralbyrå er laget på grunnlag av dødeligheten i bestemte tidsperioder. Tabeller over dødeligheten i fødselskohorter har ikke vært utarbeidd. I et fransk arbeide (Delaporte, 1941) er det imidlertid publisert tabeller for dødeligheten for en del utvalgte kohorter i en del europeiske land, og Norge er ett av disse landene.

I Byråets arkiv finnes data over beregnet folkemengde etter kjønn i ettårige aldersklasser for hvert år etter 1845 og tall for døde etter kjønn, fødselsår og dødsår. På grunnlag av dette materialet og tall for levendefødte etter kjønn vil det være mulig å utarbeide en analyse på kohortdødeligheten i Norge fra 1846 og fram til i dag.

I dette arbeidsnotatet vil vi presentere et forslag til metode for beregning av kohortdødeligheten. Videre vil vi gjøre rede for hvilke dødelighetsmål vi vil regne ut, og gi en presentasjon av det datamaterialet som er tilgjengelig. Men vi vil ikke her presentere noen beregninger på datamaterialet.

Arbeidsnotatet er også ment som et diskusjonsgrunnlag for utarbeiding av en analyse av kohortdødeligheten i Norge for tiden etter 1845. Selve analysen vil vi seinere presentere i serien Statistiske analyser (SA) fra Statistisk Sentralbyrå.

De tilgjengelige tabeller over periodedødeligheten i Norge bruker tidsperioder av ulik lengde. De er også utarbeidd etter en metode som avviker noe fra den vi skisserer her. Datamaterialet avviker også noe for enkelte årganger. For å kunne sammenlikne våre tall for kohortdødelighet med tall for dødelighet i tidsperioder vil vi derfor i vårt opplegg også regne ut dødelighetesmålene med periodedata.

2. Dødelighetsmodellen

Vi tar utgangspunkt i framtidig dødsalder, T_o, for en nyfødt person. Denne dødsalder oppfatter vi som en tilfeldig variabel med en gitt dødelighetslov. La denne lov være gitt ved den kumulative sannsynlighetsfunksjon for T_o, dvs.

(2.1) $\quad {}_xq_o = P\left[T_o \leq x\right]$

som altså er sannsynligheten for at en 0-åring skal dø seinest i alder x. Sannsynligheten for at en nyfødt skal oppleve alder x er da

$$ {}_xp_o = 1 - {}_xq_o \quad , $$

Av l_o nyfødte vil forventet antall gjenlevende i alder x være

(2.2) $\quad l_x = l_o \cdot {}_xp_o$

Som funksjon av alder x fra x = 0 til høyeste levealder x = ω (som vi for enkelthets skyld vil anta heltallig) uttrykker l_x altså forventet antall gjenlevende i en bestand av l_o samtidig fødte, den beskriver altså hvorledes man kan "vente" at samtidig nyfødte vil dø ut. Vi kaller l_x dekrementserien.

Vi definerer ${}_tq_x$ som sannsynligheten for at en person som lever i alder x skal dø før alder x + t, dvs.

(2.3) $\quad {}_tq_x = P\left[x < T_o \leq x + t \mid T_o > x\right]$

Dette gir:

(2.4) $\quad {}_tp_x = 1 - {}_tq_x = \dfrac{l_{x+t}}{l_x}$

der ${}_tp_x$ er sannsynligheten for at en x-åring skal leve i minst t år til.

Vi betegner de ettårige sannsynlighetene med q_x og p_x, dvs.

(2.5) $\quad q_x \overset{def}{=} {}_1q_x \quad$ og $\quad p_x \overset{def}{=} {}_1p_x$

Vi definerer d_x som forventet antall døde mellom alder x og x + 1 av l_o nyfødte.

Av (2.4) får vi da

(2.6) $\quad d_x = l_x - l_{x+1} = l_x\,(1 - p_x) = l_x \cdot q_x$

Vi innfører nå dødsintensiteten μ(x) ved:

(2.7) $\quad \mu(x) = \lim\limits_{t \downarrow 0} \dfrac{{}_tq_x}{t}$

Vi har altså at $\mu(x).dx$ er sannsynligheten for at en x-åring skal dø innen alder x+dx.
Av (2.7) finner vi:

$$(2.8) \qquad {}_tp_x = 1 - {}_tq_x = e^{-\int_x^{x+t} \mu(s)ds}$$

Videre er ${}_tp_x \cdot \mu(x+t)\,dt$ sannsynligheten for at en x-åring skal dø i aldersintervallet $\langle x+t, x+t+dt]$. Forventet gjenstående levetid for en x-åring er derfor ved (2.4):

$$(2.9) \qquad \overset{o}{e}_x = \int_o^{\omega-x} t \cdot {}_tp_x \cdot \mu(x+t)\,dt = \int_o^{\omega-x} \frac{l_{x+t}}{l_x}\,dt$$

Spesielt er $\overset{o}{e}_o$ forventet levetid for en nyfødt.

(I Hoem, 1967 , kap. 7, er det gitt en mer generell definisjon av dødsintensiteten hvor det tas hensyn til muligheten for selekt dødelighet, og en utledning av formlene (2.8) og (2.9).)

Ved Euler-Mac Laurins summeformel kan et integral skrives på formen:

$$\int_a^b f(x)dx = \sum_{t=a}^{b} f(t) - \frac{1}{2}[f(a) + f(b)] - \frac{1}{12}[f'(b) - f'(a)]$$

$$+ \frac{1}{720}[f'''(b) - f'''(a)] + \ldots$$

(For bevis av Euler-Mac Laurins summeformel, se Simonsen, 1962 , kap. IX eller Freeman, 1960 , avsnitt IX.8, s. 188-189.)

Ved å bruke 3 ledd i denne formelen får vi følgende tilnærmelse for $\overset{o}{e}_x$:

$$(2.10) \qquad \overset{o}{e}_x = \int_o^{\omega-x} \frac{l_{x+t}}{l_x}\,dt \simeq \sum_{t=o}^{\omega-x} \frac{l_{x+t}}{l_x} - \frac{1}{2} - \frac{1}{12}\,\mu(x)$$

Vi definerer også $\overset{o}{e}_{x/y}$ som forventet levetid mellom to gitte aldre x og y. Av Euler-Mac Laurins summeformel får vi da:

$$(2.11) \qquad \overset{o}{e}_{x/y} = \int_o^{y-x} \frac{l_{x+t}}{l_x} \simeq \sum_{t=o}^{y-x} \frac{l_{x+t}}{l_x} - \frac{1}{2}\left[1 + \frac{l_y}{l_x}\right] - \frac{1}{12}\left[- \frac{l_y}{l_x}\mu(y) + \mu(x)\right]$$

(Ved bruk av Euler-Mac Laurins summeformel har vi antatt at x og y er heltallige.)

3. Estimering

I dette avsnittet skal vi se nærmere på hvordan de grunnleggende parametrene i dødelighetsmodellen kan estimeres. Vi vil her forutsette en forenklet situasjon. I vår anvendelse av modellen i kapitel 4 må vi tillempe resultatene noe.

Vi betrakter en fødselskohort på L nyfødte. Personene i kohorten observeres så lenge de lever, og vi registrerer nøyaktig alder ved dødstidspunktet. Vi skal her gå ut fra at det ikke forekommer inn- og utvandring i bestanden. Estimeringsteorien blir helt tilsvarende om vi regner med inn- og utvandring.

Anta nå at dødsintensiteten $\mu(x)$ kan regnes for konstant over aldersåret, dvs. vi antar at det eksisterer parametre $\mu_o, \ldots, \mu_{\omega-1}$ slik at:

$$\mu(x) = \sum_{j=o}^{\omega-1} \mu_j \, I_j(x)$$

der $$I_j(x) = \begin{cases} 1 & \text{hvis } x \in \langle j, j+1] \\ 0 & \text{ellers} \end{cases}$$

I stedet for å estimere den kontinuerlige funksjonen $\mu(x)$ er problemet med dette redusert til å estimere parametrene $\mu_o, \mu_1, \ldots, \mu_{\omega-1}$

For en enkelt person (person nr. i) definerer vi nå følgende størrelser:

R_{ji} = levetid i aldersintervallet $\langle j, j+1]$

D_{ji} = antall ganger personen dør i aldersintervallet $\langle j, j+1]$

Her er altså R_{ji} lik 1 hvis person nr. i blir minst $j+1$ år, lik 0 hvis personen blir høyest j år, og lik faktisk levetid etter fylte j år hvis personen dør mellom alder j og $j+1$. Videre er D_{ji} lik 1 hvis personen dør mellom alder j og $j+1$ og lik 0 ellers.

Med denne notasjonen kan likelihooden for person nr. i gis ved

$$(3.1) \qquad \Lambda_i = e^{-\sum_{j=0}^{\omega-1} \mu_j R_{ji}} \cdot \prod_{j=0}^{\omega-1} \mu_j^{D_{ji}}$$

mens den totale likelihooden for hele bestanden av L nyfødte blir

$$(3.2) \qquad \Lambda = e^{-\sum_{j=0}^{\omega-1} \mu_j R_{j.}} \cdot \prod_{j=0}^{\omega-1} \mu_j^{D_{j.}}$$

der

$$(3.3) \quad R_j = \sum_{i=1}^{L} R_{ji} \quad \text{og}$$

$$(3.4) \quad D_j = \sum_{i=1}^{L} D_{ji} \quad ,$$

D_j blir altså antall dødsfall observert for hele kohorten mellom alder j og $j+1$, mens R_j er den totale levetid i aldersintervallet $\langle j, j+1]$

Ved å maksimere Λ med hensyn på μ_x, finner vi at sannsynlighets-maksimeringsestimatorene blir

$$(3.5) \quad \hat{\mu}_x = \frac{D_x}{R_x} \qquad x = 0, 1, \ldots, \omega-1$$

Vi ser at $\hat{\mu}_x$ blir forholdet mellom antall døde i aldersintervallet $\langle x, x+1]$ og total levetid i dette intervallet.

(I Hoem, 1971 blir estimatoren (3.5) utledet under mer generelle betingelser. Det vises for eksempel at (3.5) er en god estimator også i tilfelle med flere avgangsårsaker og med ulik eksponeringstid.)

Av estimatorene $\hat{\mu}_x$ for dødsintensiteten kan vi ved (2.8), (2.4), (2.6), (2.10) og (2.11) finne estimatorer for de andre størrelsene i dødelighetsmodellen.

4. Beregning av antall døde og gjennomlevet tid

I kapitel 3 fant vi at dødsintensiteten i alder x kan estimeres som forholdet mellom antall døde mellom eksakt alder x og x + 1 og total levetid i det samme aldersintervallet.

I vårt materiale har vi observert antall personer i hvert årskull ved utgangen av hvert kalenderår, og dødsfall etter de dødes fødselsår i hvert kalenderår. Vi vil derfor estimere dødsintensiteten som forholdet mellom antall døde av et årskull i løpet av et kalenderår og samlet gjennomlevet tid for dette årskullet i det aktuelle kalenderåret. Vi utfører beregningene for menn og kvinner hver for seg, og alder blir altså definert som differansen mellom kalenderår og fødselsår.

(Se Hoem, 1974 for en grundig diskusjon av ulike alderbegrep og de tilhørende estimatorer for dødsintensiteten.)

I det følgende vil vi benytte notasjonen i Hoem (1974).

Ved beregning av våre folketall ved utgangen av kalenderårene er det tatt hensyn til inn- og utvandring. I oppstillingen nedenfor vil vi ikke skrive ut inn- og utvandringen med symboler. Disse størrelsene går imidlertid inn i de beregnede folketallene.

Fra Hoem (1974) har vi nå for kohorten født i år n:

$F(n)$ = antall levendefødte

$D_o^{\Delta}(n)$ = antall døde i løpet av fødselsåret

$D_x^{\square}(n+x)$ = antall døde i år n+x $x = 1, 2, \ldots, \omega-1$

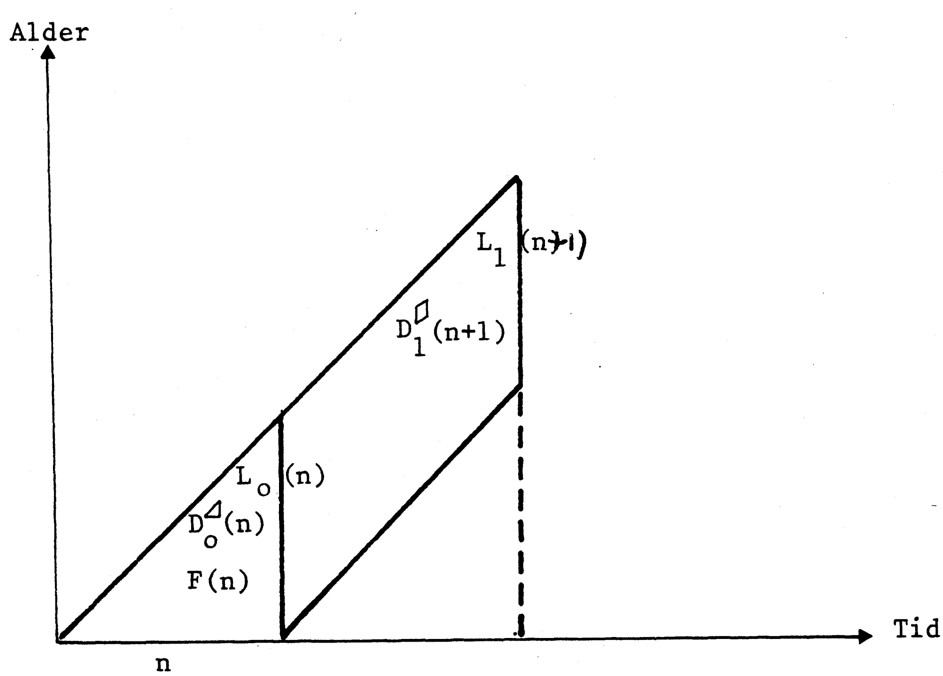
$L_x(n+x)$ = antall levende pr. 31/12 år n+x $x = 0, 1, \ldots, \omega-1$

$R_o^{\Delta}(n)$ = samlet gjennomlevet tid i løpet av fødselsåret

$R_x^{\square}$ = samlet gjennomlevet tid i løpet av år n+x
$x = 1, 2, \ldots, \omega-1$

(Vi bruker ingen indeks for kjønn, da situasjonene for menn og kvinner er helt analoge.)

For denne fødselskohorten kan vi tegne følgende utgave av Lexis skjema (se f.eks. Hoem, 1974) .



FIGUR 1

Vi får følgende formler for antall døde og sum gjennomlevet tid

$$D_o = D_o^{\Delta}(n)$$

(4.1)

$$D_x = D_x^{\lozenge}(n+x) \qquad x = 1, 2, \ldots, \omega-1$$

Av Hoem (1974) formel (7.5) er

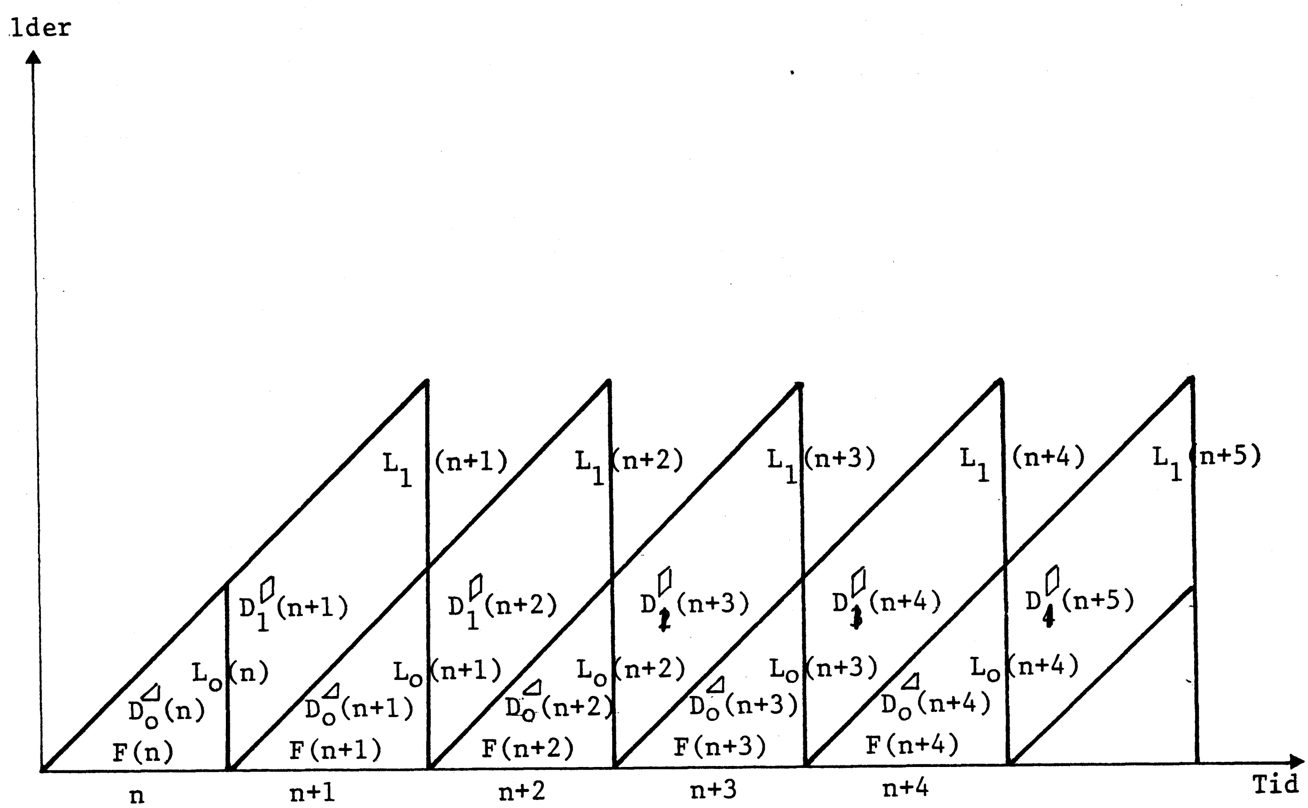
(4.2) $$R_o = R_o^{\Delta}(n) \simeq \frac{1}{4}\{F(n) + L_o(n)\}$$

og av Hoem (1974) formel (6.4)

(4.3) $$R_x = R_x^{\lozenge}(n) \simeq \frac{1}{2}\{L_{x-1}(n+x-1) + L_x(n+x)\} \qquad x = 1, 2, \ldots, \omega-1$$

For å utjevne tilfeldige variasjoner og for å få et mindre materiale som er enklere å analysere, kan det være aktuelt å slå sammen døde og total gjennomlevet tid i 5 og 5 fødselskohorter.

Vi får da følgende utgave av Lexis skjema



FIGUR 2

Formlene for antall døde og sum gjennomlevet tid blir:

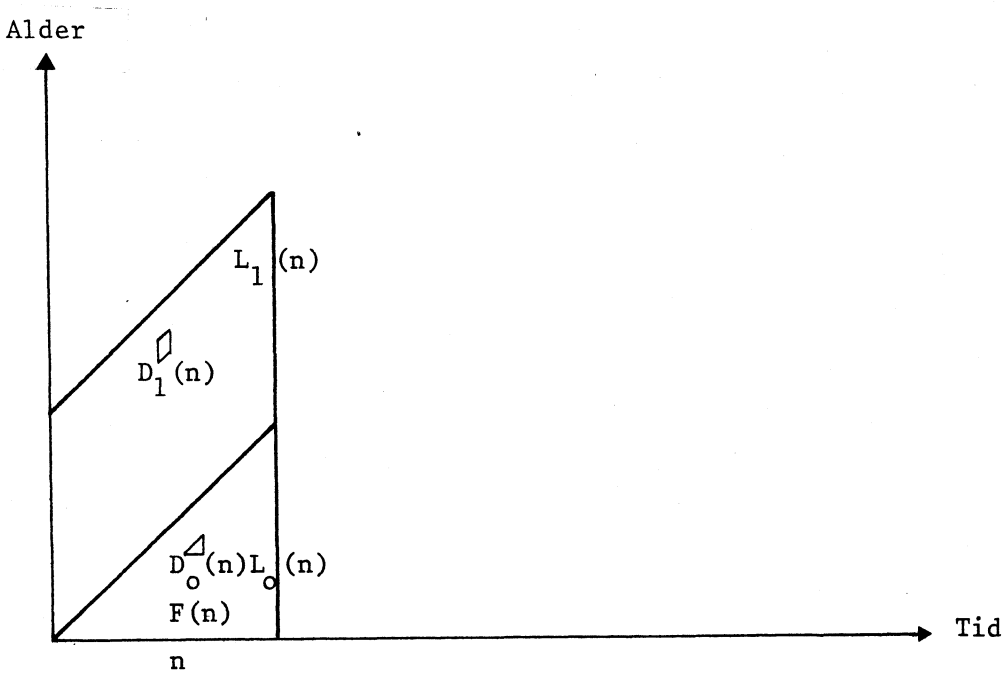
$$(4.1)\quad \begin{aligned} D_o &= \sum_{i=o}^{4} D_o^{\Delta}(n+i) \\ D_x &= \sum_{i=0}^{4} D_x^{\Diamond}(n+x+i) \qquad x = 1, 2, \ldots, \omega-1 \end{aligned}$$

og

$$(4.2)\quad R_o = \sum_{i=o}^{4} R_o^{\Delta}(n+i) \simeq \frac{1}{4}\left[\sum_{i=o}^{4} F(n+i) + \sum_{i=o}^{4} L_o(n+i)\right]$$

$$(4.3)\quad R_x = \sum_{i=o}^{4} R_x^{\Diamond}(n+x+i) \simeq \frac{1}{2}\left[\sum_{i=o}^{4} L_{x-1}(n+x+i-1) + \sum_{i=o}^{4} L_x(n+x+i)\right]$$

Hvis vi skal regne med periodedata, observerer vi alle årskullene over et kalenderår. Estimeringen blir selvsagt helt analog som for kohorter. Hvis vi tegner denne situasjonen inn i et Lexis skjema får vi:



FIGUR 3

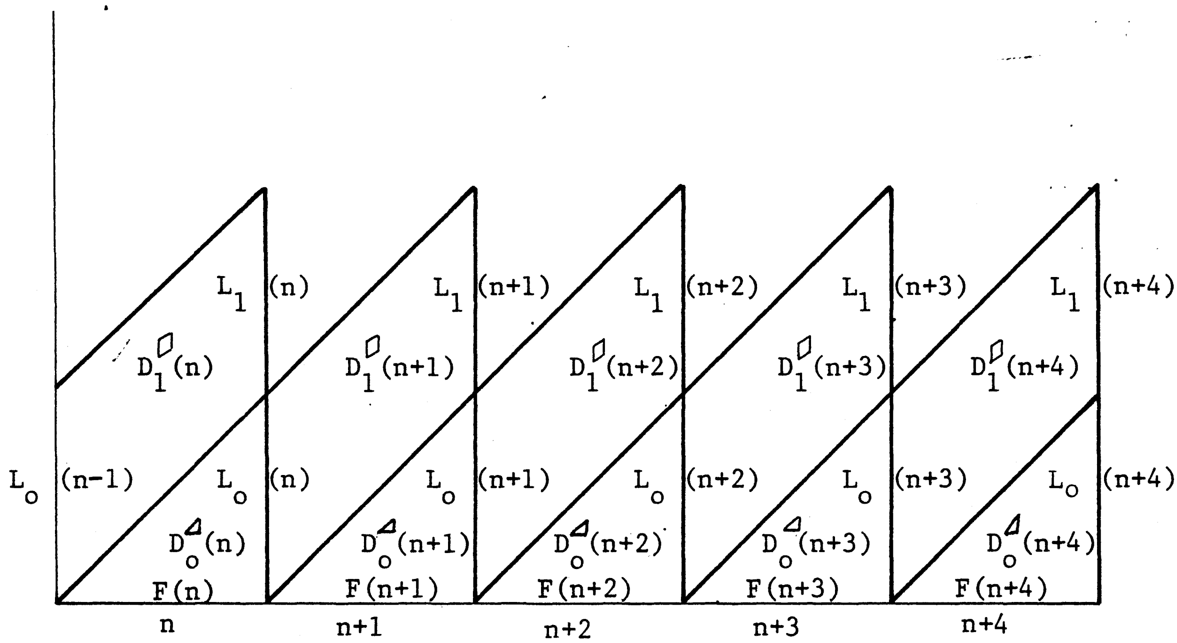
Vi får da følgende formler for antall døde og sum gjennomlevet tid:

$$(4.4)\quad \begin{aligned} D_o &= D_o^{\Delta}(n) \\ D_x &= D_x^{\Diamond}(n) \end{aligned}$$

(4.5) $\quad R_o^{\Delta}(n) \simeq \frac{1}{4}\{F(n) + L_o(n)\}$

(4.6) $\quad R_x = R_x^{\Diamond}(n) \approx \frac{1}{2}\{L_{x-1}(n-1) + L_x(n)\} \qquad x = 1, 2, \ldots, \omega-1$

Hvis vi regner med perioder, hver på 5 kalenderår kan vi tegne følgende Lexis skjema



FIGUR 4

og vi får følgende formler for antall døde og sum gjennomlevet tid

(4.4)'
$$D_o = \sum_{i=o}^{4} D_o^{\Delta}(n+i)$$
$$D_x = \sum_{i=o}^{4} D_x^{\Diamond}(n+i)$$

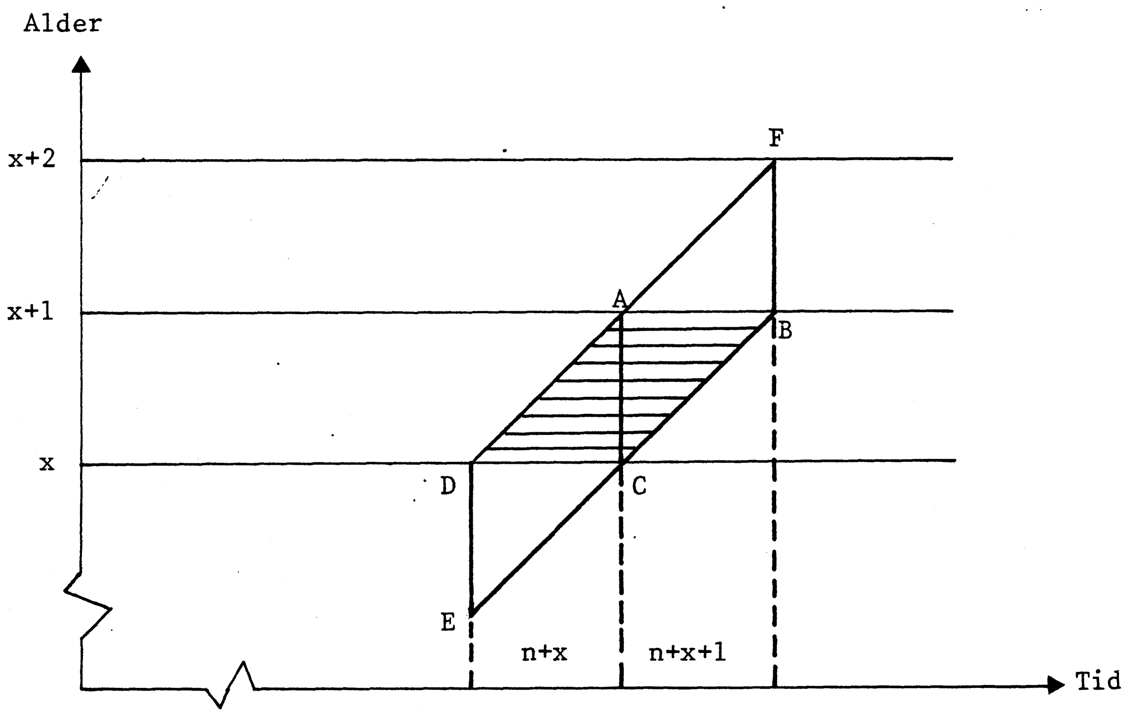
(4.5)' $\quad R_o = \sum_{i=o}^{4} R_o^{\Delta}(n+i) \simeq \frac{1}{4}\left[\sum_{i=o}^{4} F(n+i) + \sum_{i=o}^{4} L_o(n+i)\right]$

(4.6)' $\quad R_x = \sum_{i=o}^{4} R_o^{\Diamond}(n+i) \simeq \frac{1}{2}\{\sum_{i=o}^{4} L_{x-1}(n+i-1) + \sum_{i=o}^{4} L_x(n+i)\}$

$x = 1, 2, \ldots, \omega-1$

På grunnlag av beregningene foran blir μ_x estimert med alder definert i fylte år ved utgangen av året. Hvis vi bruker disse verdiene for $\hat{\mu}_x$ i formlene (2.8), (2.4), (2.6), (2.10) og (2.11), vil vi også få estimatorer for de andre dødelighetsmålene med alderen definert i fylte år ved utgangen av året.

I våre beregninger ønsker vi imidlertid å estimere disse dødelighetsmålene med alder i fylte år. Vi kan belyse sammenhengen mellom de to tilfellene i Lexis skjemaet i figur 5. I figuren har vi fulgt fødselskohorten i år n gjennom de to kalenderårene n+x og n+x+1. Vi har over antatt at dødsintensiteten er konstant over kalenderåret,



FIGUR 5

altså lik μ_x i parallellogrammet ACED og lik μ_{x+1} i parallellogrammet FBCA. I kapitel 3 har vi regnet med konstant dødelighet over aldersåret, altså at dødsintensiteten skal være konstant i det skraverte parallellogrammet ABCD. Det er rimelig å anta at denne vil være gjennomsnittet av dødsintensiteten i ACED og FBCA, altså får vi av (2.8)

$$(4.7) \qquad \hat{p}_x = e^{-\frac{1}{2}[\hat{\mu}_x + \hat{\mu}_{x+1}]} \qquad x = 0, 1, \ldots, \omega-1$$

For estimering av p_x med periodedata blir situasjonen helt analog bortsett fra at vi benytter dødsintensitetene for to fødselskohorter i samme år. Denne formelen er neppe særlig god for $\hat{p}_o$. For de øvrige verdier av $\hat{p}_x$ skulle den være god nok. Av (2.4) finner vi så estimatorer for l_x rekursivt ved

$$(4.8) \quad \hat{l}_{x+1} = \hat{l}_x \cdot \hat{p}_x \qquad x = 0, 1, \ldots, \omega-1$$

$$\hat{l}_o = l_o$$

Videre får vi av (2.6)

$$(4.9) \quad \hat{d}_x = \hat{l}_x - \hat{l}_{x+1} = \hat{l}_x (1 - \hat{p}_x) = \hat{l}_x \cdot \hat{q}_x$$

Bortsett fra i de aller høyeste aldrene blir tredje ledd i formel (2.10) forsvinnende lite. Vi kan derfor se bort fra det ved beregningene. Tilsvarende ledd i formel (2.11) blir enda mindre.

Hvis vi bare tar med to ledd i formlene (2.10) og (2.11) får vi følgende estimatorer for $\mathring{e}_x$ og $\mathring{e}_{x/y}$

$$(4.10) \quad \hat{\mathring{e}}_x = \sum_{t=o}^{\omega-x} \frac{\hat{l}_{x+t}}{\hat{l}_x} - \frac{1}{2}$$

og

$$(4.11) \quad \hat{\mathring{e}}_{x/y} = \sum_{t=o}^{y-x} \frac{\hat{l}_{x+t}}{\hat{l}_x} - \frac{1}{2} \left[1 + \frac{\hat{l}_y}{\hat{l}_x}\right]$$

5. Glatting av dødelighetsrater

I overlevelsestabeller med periodedata er det vanlig å glatte dødelighetsratene for å utjamne tilfeldige variasjoner. Dette blir gjort fordi de beregnede ratene skal brukes som mål på dødelighetsutviklingen i framtida. Ratene blir også glattet fordi de glattede ratene normalt gir et klarere bilde, og letter sammenlikningen fra år til år.

Overlevelsestabeller på grunnlag av kohortdata belyser det som virkelig er skjedd, og skal ikke brukes som anslag for framtida. Det er derfor kanskje unødvendig å glatte dødelighetsratene. Presentasjonen av materialet i form av fem-årsgrupper gir jo også en viss glattingseffekt.

Den endelige avgjørelse med hensyn til om vi skal glatte dødelighetsratene bør utstå til vi har sett forløpet til råratene. Hvis disse får et meget takkete forløp, kan det bli aktuelt å glatte dem for lettere å kunne sammenlikne de ulike kohortene.

Ved glatting av periodedata er det vanlig å glatte med ratene for de omkringliggende aldersklasser i samme periode. Hvis vi skal glatte ratene for kohorter kan vi glatte med ratene for den samme kohorten i de omkringliggende kalenderår.

Hva slags glatting en vil velge kan avhenge av hvilke forhold en ønsker å belyse. I en situasjon med "krise" i en bestemt periode (f.eks. spanskesyken i 1918) kan det være interessant både å se hvordan dette påvirker personer i perioden i ulike aldergrupper, og å undersøke om kohorten, f.eks. nyfødte i "krise"-tida vil dø annerledes i hele livsløpet enn folk født i andre perioder. (Det siste er imidlertid neppe tilfelle for fødselskohorten i 1918). Glatting for en kohort kan tilsløre den første typen effekter, mens glatting for en periode kan tilsløre den andre typen.

Glatting kan imidlertid også være av interesse for en situasjon som er beskrevet over. Ved å sammenlikne de observerte ratene med de glattede kan vi gi en antydning om hvordan det hadde gått hvis det ikke hadde vært "krise".

Når vi ser bort fra situasjonen med "krise", kan det muligens være riktig å glatte ratene for å lette sammenlikningen fra kohort til kohort og fra periode til periode. Hvis vi glatter kan det likevel være aktuelt å bruke råratene ved analyse av enkelte forløp.

6. Aktuelle dødelighetsmål

Det kan være aktuelt å tabulere dødsintensiteten $\hat{\mu}_x$, dekrementserien, $\hat{l}_x$, antall døde i en overlevelsestabell, $\hat{d}_x$, og forventet gjenstående levetid $\hat{\mathring{e}}_x$. Størrelsene regnes ut både for sammendrag av 5 fødselskohorter og for 5-års perioder. Det kan også være aktuelt å regne ut disse størrelsene for enkeltår, men det er neppe aktuelt å publisere disse tallene.

Med det datamaterialet vi har kan $\hat{\mathring{e}}_o$ bare regnes ut for fødselskohorter som har gjennomlevet hele livet etter 1846. Dette gjelder bare fødselskullene fra 1846 til ca. 1875. For å kunne gi mål for forventet levetid for flere fødselskohorter, kan det være aktuelt å regne ut disse levetidene mellom gitte aldre, f.eks.

$$\hat{\mathring{e}}_{0/1},\ \hat{\mathring{e}}_{1/50},\ \hat{\mathring{e}}_{15/65},$$

dvs. forventet levetid for en nyfødt mellom 0 og 1 år, for en ettåring mellom 1 og 50 år og for en 15 åring mellom 15 og 65 år.

Med vårt datamateriale er det ikke mulig å beregne $\hat{e}^{o}_{0/1}$ særlig nøyaktig. Det kan likevel være interessant å belyse endringer i denne størrelsen over tid.

7. Datagrunnlaget

For å beregne de ulike dødelighetsmål brukes tall for folkemengden etter kjønn og fødselsår ved utgangen av hvert kalenderår og tall for levendefødte etter kjønn og for døde etter kjønn og fødselsår i løpet av de enkelte kalenderårene.

A. Folkemengde

Før 1970 er folkemengden etter kjønn og fødselsår beregnet på grunnlag av de 10-årlige folketellingene og tall for fødte, døde og tilgjengelige tall for inn- og utvandring. Fra og med 1970 er folketallene framkommet ved en opptelling i det sentrale personregisteret.

For årene 1846-1900 er slike beregnede folketall publisert i Statistisk Sentralbyrå (1910). For årene 1901-1909 finnes tilsvarende tall i Statistisk Sentralbyrå (1915). Fra og med 1910 vil vi benytte tall for folkemengde fra en datafile som er etablert i Byrået. Metoden for innsamling og bearbeiding av disse dataene er beskrevet i Statistisk Sentralbyrå (1978).

B. Fødsler

Tall for levendefødte etter kjønn er publisert i Folkemengdens bevegelse. Før 1916 gjelder tallene fødte blant den tilstedeværende befolkning, og en del fødsler er registrert i året etter at fødselen fant sted.

I Statistisk Sentralbyrå (1926) er det gitt tall for levendefødte totalt for begge kjønn. Før 1916 avviker disse tallene noe fra dem som ble publisert i de tidligere utgaver av Folkemengdens bevegelse. Det er rimelig å anta at de sist publiserte tallene er mer korrekte enn dem som tidligere ble publisert i Folkemengdens bevegelse. Vi vil derfor bruke tallene for fødte i 1846-1915 fra Statistisk Sentralbyrå (1926) i våre beregninger.

For årene før 1916 vil vi anta at det hvert år ble født 106 gutter for hver 100 fødte jenter. Bortsett fra dødssannsynligheten i første leveår har tallet på levendefødte liten betydning i våre beregninger, så denne tilnærmingen vil ikke ha noen praktisk betydning.

C. Dødsfall

Statistikken over døde etter fødselsår, som ble publisert i Folkemengdens bevegelse før århundreskiftet, har visse svakheter som det er justert for ved utarbeiding av de tidligere publiserte dødelighetstabeller med periodedata. Tallene som er brukt er imidlertid ikke publisert i en slik form at vi kan bruke dem i våre beregninger.

For enkelte kalenderår i forrige århundre vil vi derfor i våre beregninger ikke ta hensyn til følgende feil:

- en mindre del av de døde er i virkeligheten døde i året før registreringsåret
- det er reigstrert en overvekt av døde på runde aldersår og tilsvarende lavere tall på de nærmeste aldersår
- opp til 3-4 promille av de døde har uoppgitt fødselsår. Disse dødsfallene er holdt utenom våre beregninger.

For årene 1846-1865 er døde etter kjønn og fødselsår publisert i Statistisk Sentralbyrå (1910). Disse tallene er beregnet ut fra døde i grovere aldersgrupper. Metoden for oppsplitting er beskrevet i publikasjonen. Tall for døde i 1866-1870 og 1871-1875, er publisert i henholdsvis Statistisk Sentralbyrå (1873) og Statistisk Sentralbyrå (1881).

For årene 1876-1910 er det publisert tall for døde i Norge etter kjønn og fødselsår i Statistisk Sentralbyrå (1876-1910). For disse årene er tall for døde blant den hjemmehørende befolkningen bare gitt for grupper på 5 fødselsår. Forskjellen mellom døde i Norge og dødsfall blant den hjemmehørende befolkningen i Norge er i hovedsak dødsfall blant norske sjømenn mellom 15 og 60 år.

For årene 1886-1890 og 1896-1910 finnes det håndskrevne tabeller i Byråets arkiv som gjør det mulig å beregne dødsfall blant den hjemmehørende befolkningen på de enkelte fødselsår. Med små tillempninger er dette også mulig for årene 1876-1881.

For årene 1882-1885 er avviket mellom dødsfall i den hjemmehørende og den tilstedeværende befolkningen regnet ut for 5-års gruppene. Disse tallene er så splittet på de enkelte aldersår i samme forhold som gjennomsnittet av tallene i perdiodene 1876-1881 og 1886-1890. På samme måte er tallene for årene 1891-1895 splittet etter gjennomsnittet av årene 1886-1890 og 1896-1900. Disse tallene er så lagt til tallene over dødsfall blant den tilstedeværende befolkningen.

For menn i aldersklassene under 15 år og 60 år og over og for kvinner i alle aldre vil vi bruke tall for dødsfall inntruffet i Norge. Disse tallene avviker ubetydelig fra tallene for døde i den hjemmehørende befolkningen.

I Folkemengdens bevegelse for 1901 og 1902 er tall for døde i Norge i aldrene 30-94 år bare gitt i sammendrag på 2 og 3 fødselsår. Disse tallene har vi fordelt på de enkelte fødselsår i samme forhold som tallet på levendefødte i de samme årene. Tallene for døde menn 15-59 år er siden justert for dødsfall blant nordmenn midlertidig i utlandet og utlendinger midlertidig i Norge.

Tall for døde i 1911 og seinere år vil vi ta fra datafilen som er beskrevet i Statistisk Sentralbyrå (1978).

REFERANSER

[1] Delaporte, P. (1941): Evolution de la mortalité en Europe depuis l'origine des statistiques de l'etat civil (Tables de mortalité de générations.) Statistique générale de la France. Études demographiques No 2, 1941.

[2] Freeman, H. (1960): Finite Differences for Actuarial students. Cambridge University . Press for the Institute of Actuaries.

[3] Hoem, Jan M. (1967): Grunnbegreper i formell befolkningslære. Memorandum fra Sosialøkonomisk Institutt, Universitetet i Oslo.

[4] Hoem, Jan M. (1971): "Point Estimation of Forces of Transition in Demographic Models". Journal of the Royal Statistical Society 1971, series B, s. 275-289.

[5] Hoem, Jan M. (1974): Beregning av befolkningsrater. Statistisk Sentralbyrå, Arbeidsnotat 74/22.

[6] Simonsen, W. (1962): Forelæsningsnoter til numerisk analyse (interpolasjonsregning). Københavns Universitets Fond til Tilvejebringelse af Læremidler.

[7] Statistisk Sentralbyrå (1873): Tabeller vedkommende Folkemængdens Bevægelse 1851-1870. Det statistiske Centralbureau, NOS I C. No. 1.

[8] Statistisk Sentralbyrå (1881): Tabeller vedkommende Folkemængdens Bevægelse 1871-1875. Det statistiske Centralbureau, NOS I C. NO. 1.

[9] Statistisk Sentralbyrå (1876-1910): Statistik over Folkemængdens Bevægelse: 1876-1880, Hefte 1, 1881-1885, Hefte 1. Tabeller vedkommende Folkemængdens Bevægelse: 1881,..., 1885, Folkemængdens bevægelse: 1886,..., 1910. NOS Det statistiske Centralbureau.

[10] Statistisk Sentralbyrå (1910): Norges Folkemængde 1846-1901, NOS Det Statistiske Centralbyraa.

[11] Statistisk Sentralbyrå (1915): Dødelighetstabeller for det Norske Folk 1901/02-1910/11, NOS Det Statistiske Centralbyraa.

[12] Statistisk Sentralbyrå (1926): Statistiske oversikter 1926, Det Statistiske Centralbyrå.

[13] Statistisk Sentralbyrå (1978): Utviklingen i giftermål og dødsfall 1911-1976. Statistisk Sentralbyrå, Statistiske analyser nr. 35.