

Markovmodulerte Poisson-prosesser som modeller for dykketidsdata



MASTEROPPGÅVE I STATISTIKK – DATAANALYSE

Mona Christin Mossestad | 23. 03. 2010

INNHALD

Forord	3
Innleiing	4
1 Vågequal, kvalfangst og IWC	4
2 Innsamling av data	5
3 Behandling av data	7
4 Modellar for fordeling og tettheit av kval	8
5 Systematisering av dykk	11
6 Matematiske modellar av dykketider	12
7 Grafisk framstilling	13
8 Problemstilling og rammer	17
9 Poisson-prosessen	18
10 Coxprosess, kontinuerligtids Markovkjeder og Poisson-prosess med Markov-intensitet	19
11 Kriterier	21
12 Diskusjon	27
Figurar	28
Tabellar	28
Litteratur	29

FORORD

Eg har alltid vore glad i matematikk. Likevel valde eg *ikkje* allmennutdanning med realfag då eg gjekk i vidaregåande skule, og då eg som ung mor starta mi høgskuleutdanning var det ikkje interesser, men fornuftige resonnement som bestemte vegen. Åra gjekk, og så - i 2004 - oppdaga eg tilfeldigvis at dette var det siste året der søkarar utan bakgrunn frå vidaregåande realfaglinje ville bli tatt opp som realfagstudentar på universitetsnivå. Tanken om å studere matematikk på universitetet hadde leika i hovudet lenge, og brått var det no eller aldri. Eg tok sats og starta med iver på eit årsstudium i matematikk ved Universitetet i Bergen. Studiet gav meirsmak. Eg ville gå vidare, og det var med stor glede eg mottok meldinga om at eg var tatt opp som masterstudent i statistikk – dataanalyse.

Livet går ikkje alltid etter planen. Undervegs i masterstudiet har eg vore mykje plaga med sjukdom. Eg har fått forlenga tida på masterprogrammet med over to år og skjønt eg kunne ha søkt om meir, har eg no funne det på tide å avslutte. Ein konsekvens av dette er at eg ikkje har fått tid til å løyse denne oppgåva slik som eg hadde ønska, og resultatet er difor noko amputert.

Trass sjukdom har masterstudiet vore ei oppleving eg ikkje ville ha vore utan. Eg vil gjerne takke mine studiekameratar for kjekke samtalar og interessante diskusjonar. Ein spesiell takk går til Karl Ove som alltid er så interessert og hjelpsam. Eg vil også takke vegleiaren min for forståing og oppmuntring og familien min for støtte og hjelp til alt praktisk. Ikkje minst takkar eg mine foreldre som også har tatt godt vare på barna når eg har gjort mine mange turar til Bergen.

INNLEIING

Alle sjøpattedyr må til overflata for å puste. Det er stor variasjon mellom artar på kor lenge kvart enkelt dykk varer. Dersom tida tilbrakt i overflaten er liten samanlikna med varigheita av dykket, er det rimelig å beskrive fenomenet matematisk som en stokastisk punktprosess i tid. Kvart " blåst" vil då være eit punkt, og tida mellom blåst vil variere stokastisk.

1 VÅGEKVAL, KVALFANGST OG IWC

Vågekvalen (*Balaenoptera acutorostrata*) kan bli 9 meter lang og 8 tonn tung, og er med dette den minste av kvalane i familien finnkvalar. Den er strømlinjeforma og ein rask symjar, og den vandrar over store havområde. Vågekvalen er ein bardekval, det vil sei at den i staden for tenner har bardar som den nyttar til å sile dyreplankton og stimfisk frå vatnet. I våre farvatn beiter den først og fremst på krill, sild, lodde og sil, men den tek også for seg av fiskeartar som torsk, sei og polartorsk. (Øien 2009)

Det er lange tradisjonar for kvalfangst i Norge. Fangst langs norskekysten er omtalt i skriftlige kjelder allereie på 800-talet, og reglar for kvalfangst finn vi i lover frå 900-talet¹. Det har gjennom åra vorte fangsta ulike artar kval, men etter 1980 er vågekvalen er den einaste kvalen det blir drive fangst av i Norge.



Figur 1.1 Vågekval

Kjelde: Havforskningsinstituttet

¹Det er dessutan funne restar etter både små og store kvalar i møddingane ved norske steinalderbusettingar, men det kan også avstamme frå sjølvdøde dyr. Sjå <http://www.hvalfangerklubben.net/index.php?id=Artikkel+11>

Norge er ikkje åleine om å ha tradisjonar for kvalfangst, og fordi kvalar er trekkande artar vart det etter kvart behov for internasjonal regulering. Avtalen *Den internasjonale konvensjon for regulering av kvalfangst (ICRW)* vart framforhandla i 1946, og *Kvalfangstkommisjonen (IWC)* vart oppretta for å handheve og regulere denne avtalen.

På 60- og 70-talet vaks det fram ei erkjenning av at mange kvalartar var overbeskatta og utryddingstrua. Informasjonsgrunnlaget var likevel mangelfullt, og da IWC klassifiserte den nordatlantiske vågekvalen som ein fredda bestand var Norge sterkt ueinig. I 1986 innførte IWC ein fangststopp (moratorium) for **all** kommersiell kvalfangst. Norge reserverte seg mot dette vedtaket, og kunne difor halde fram med fangstinga. Etter internasjonalt press vart den norske vågekvalfangsten likevel stoppa i 1987 i påvente av betre kunnskap om bestandsstørrelsen. Fem år seinare starta den opp att, og dei årlege fangstkvotane vart fastsett på grunnlag av forvaltningsprosedyrar utvikla av vitskapskomiteen i IWC. IWC sitt moratorium gjeld likevel enno, og av dei 79 medlemslanda i IWC er det berre Norge som i dag driv kommersiell kvalfangst. (Stortingsmelding nr. 27 2003/04)

For vågekval i Nord-atlanteren reknar Den internasjonale kvalfangstkommisjonen (IWC) fire bestandar: den kanadiske austkystbestanden, vestgrønlandbestanden, sentralbestanden og den nordaust-atlantiske bestanden. Det er dei to sistnemnde nordmenn driv fangst på. Den nordaust-atlantiske bestanden lever i Nordsjøen, langs norskekysten, i Barentshavet og ved Svalbard, og basert på telletokt i perioden 2002-2007, er denne estimert å bestå av 78 500 dyr. (Øien 2009)

2 INNSAMLING AV DATA

Når ein skal estimere bestanden av vågekval må ein ta utgangspunkt i observasjonar frå fartøy på havet. For ein båt ute på telletokt er linjetransekt ein vanleg metode, det vil sei at det blir gjort visuelle observasjonar medan båten går med konstant hastighet langs ei definert linje.

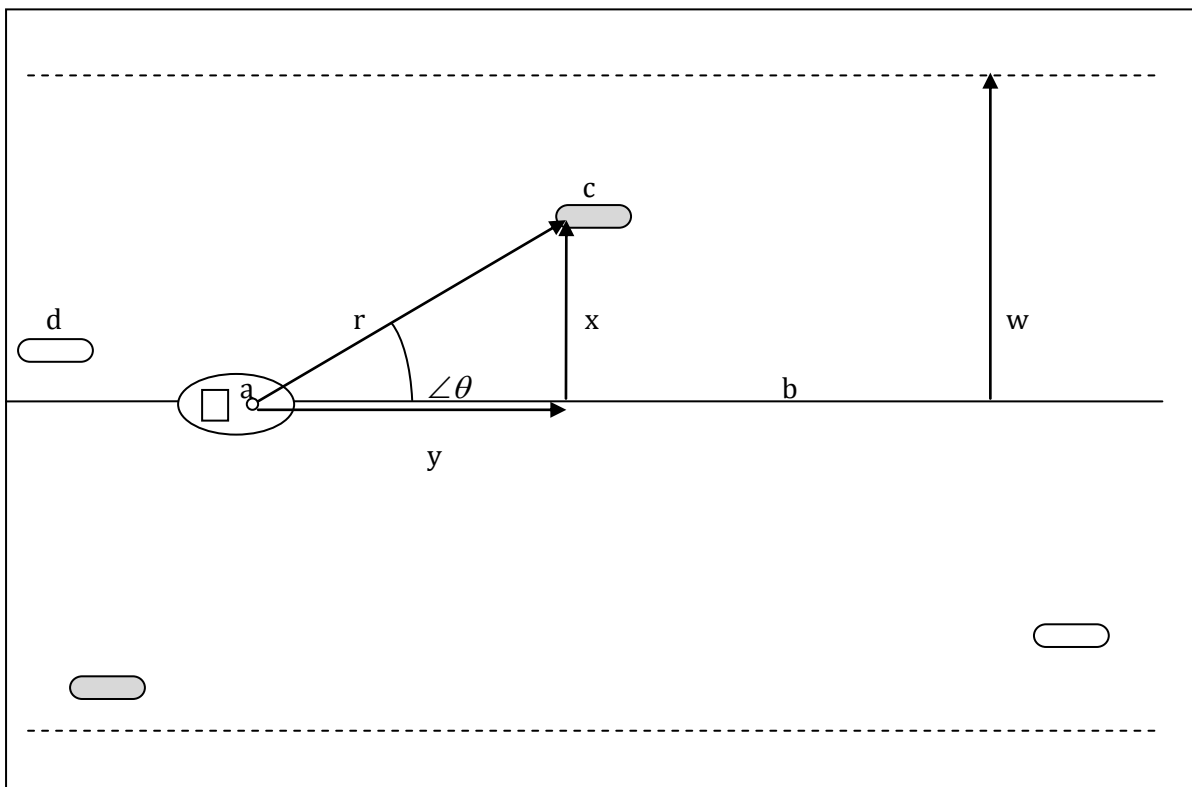
Når ein kval er oppdaga blir den radiale distansen mellom observatør og kval estimert visuelt, og vinkelen mellom transektlinja og linja mellom observatør og kval blir lest av frå eit fastmontert vinkelbord. Desse måla inngår i berekninga av effektiv breidde av søkeområdet. (Raftery og Schweder 1993)

I tradisjonelle linjetransekt er det ein enkelt observatør (eller eit enkelt team av observatørar) som søker etter dyr, og det blir antatt at alle dyr som oppheld seg i søkeområdet blir oppdaga. Men ein slik antakelse vil være feil i søk etter vågekval, for vågekvalen er vanskeleg å observere. (Skaug og Schweder 1999) Han bryt havoverflata berre for å puste, han har ikkje synleg blåst og han er berre synleg i 2-3 sekund. Det at ikkje alle dyr blir oppdaga representerer naturlegvis ei feilkjelde. Vanlege metodar for å estimere størrelsen på denne feilkjelda bygger på observasjonar frå *to* uavhengige observatørar samtidig i same eller parallelle linjetransekt (eng: double-platform line transect)

Fordi vågekvalen er så rask og dessutan sjeldan viser halefinnen er det også vanskeleg å skilje individ frå kvarandre. Dette medfører at same individ kan blir telt fleire gonger. Estimaten blir difor sterkare dersom berekningsmodellen også tek omsyn til dykkeatferden.

Det er to sider ved dykkeatferden som er relevant for dagens berekningsmodellar: *blåstraten* som viser kor ofte ein kval i snitt bryt overflata , og *dykkemønsteret* som viser tidrekka av blåst. (Øien 2009)

Talmaterialet som blir brukt i denne oppgåva er ei tidsrekke av blåst. Observasjonane stammar frå eit forskingstokt utført av havforskningsinstituttet i 1992. Vågekvalen vart påført eit radiomerke den 9 august klokka 09.12.45, og i ca 24 timar og 17 minutt og sendte dette merket signal til VHF-bandet kvar gong kvalen braut vassflata for å puste. Det vart registrert 1078 blåst, noko som gir ei blåstrate på 44 blåst/time. Dykkemønsteret skal vi kikke nærare på i kapittel 7.



Figur 2.1 Skjematisk teikning av eit linjetransekt med ein observatør eller eit team.

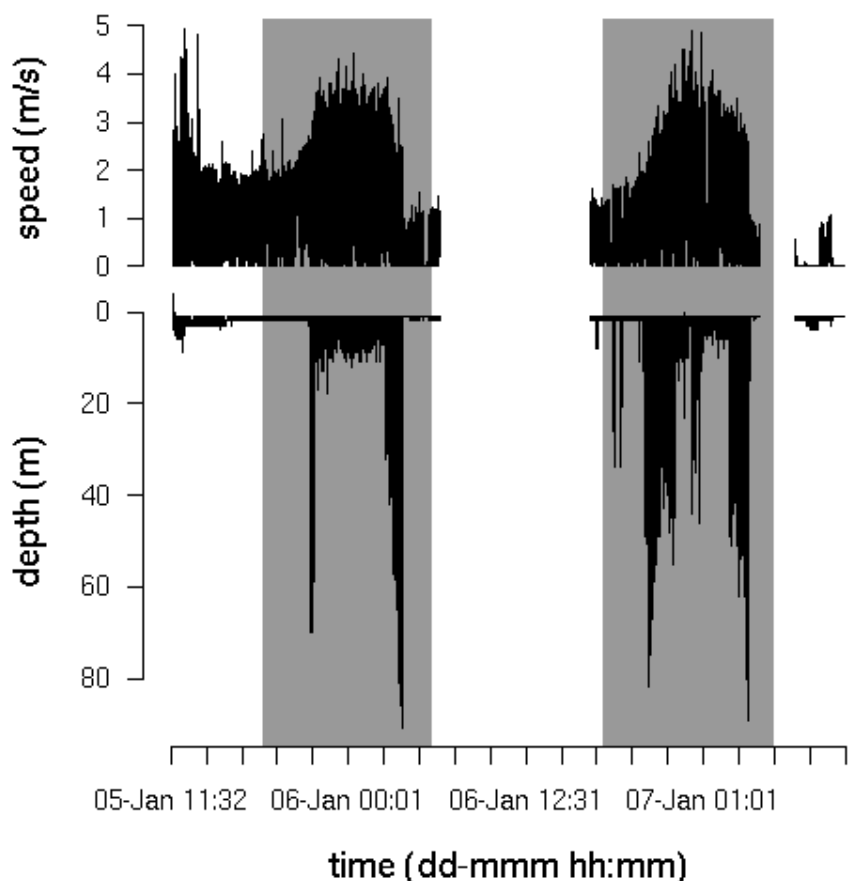
- a) er posisjonen til observatøren, b) er transektlinja, c) er ein observert kval,
- d) er ein uoppdaga kval, r) er den radiale distansen x) er distansen vinkelrett frå transektlinja, y) er distansen langs transektlinja fram til punktet parallelt med kvalen, θ) er vinkelen mellom transektlinja og linja mellom observatør og kval og w) er den effektive halvbreidda av søket.

3 BEHANDLING AV DATA

Det stor interesse for forskning på dykkeatferd, ikkje berre for vågekval, men også på andre dyr og fiskar som lever i og ved vatn. For å lette kartleggingsarbeidet har det blitt tatt i bruk nye elektroniske hjelpemiddel, både for registrering, overføring og lagring av informasjon. Dette har gjort det mogeleg å samle inn store mengder data, og dermed har det også vorte auka behov for programvarer som lettar analysen av all informasjonen. Slik har det blitt utvikla ein del program for avlesing og analyse av data frå spesifikke måleinstrument. Storparten av desse programma krev at du betalar for ein lisens, men det fins også eksempel på fri programvare.

R er eit gratis databehandlingsprogram som er mykje brukt til statistisk analyse. Med dette programmeringsverktøyet kan du lage egne funksjonar, eller du kan installere ferdige pakkar med sett av funksjonar. I 2007 vart det introdusert ein pakke laga for analyse av dykkeatferd. Pakken, som heiter *diveMove*, er spesielt utvikla for å analysere data frå temperatur- og djupneloggar (TDR). Den har funksjonar for til dømes kalibrering av fart - og djupnemålingar, identifisering av våte versus tørre periodar, identifisering av fasane i enkeltdykk (nedstigning, botn og oppstigning), plot for å visualisere dykka, maksimum djupne, tid ved botnen mm.

Alle grafane som blir vist i denne oppgåva er laga i R. For meir informasjon om R sjå <http://www.r-project.org>.



Figur 3.1 Funksjonen `plotTDR()` i R-pakken `diveMove` produserer eit interaktivt plot som tillet zooming.

4 MODELLAR FOR FORDELING OG TETTLEIK AV KVAL

Det har vore mykje aktivitet og stor utvikling innan bestandsestimering etter moratorievedtaket til IWC. Målet er naturlegvis å oppnå eit så eksakt bestandsestimat som mogeleg. I dette kapitlet skal vi sjå på ein mykje brukt metode i dette arbeidet, og eg vil også trekke fram nokre få modellar frå det siste tiåret.

Det er mange variablar å ta omsyn til i estimat av vågekvalbestanden. Linjetransektsøk etter vågekval i nordaustatlanteren er karakterisert ved følgjande (Schweder, et al. 1999):

- Diskret tilgjengelegheit: Dyra er tilgjengelege for observasjon berre når dei bryt overflata for å puste, og blåsta blir behandla som diskrete punkt i tid.
- Redusert sjanse for å oppdage dyr pga vanskane med å observere vågekvalen. Sannsynlegheita for at ein vågekval som oppheld seg i transektområdet blir oppdaga er mindre enn ein.
- Målefeil: Målinga av den radiale distansen og av vinkelen mellom transektlinja og observasjonspunktet kan ha både tilfeldige og systematiske feil.
- Varierende forhold: Observasjonseffektiviteten blir påverka av variablar som sjøgang, vær, observasjonsteam og fartøy.

Arbeidet med bestandsestimat involverer i følgje Borchers og Burnham (Buckland, et al. 2004, side 5-18) tre inferens-steg:

1. Estimering av sannsynligheter for å observere dyr (eng: detection probabilities).
2. Estimering av antall dyr (eller tettheit) i det gjennomsøkte området gitt antall observerte dyr og sannsynlighetene for å observere dyr.
3. Estimering av antall dyr (eller tettheit) i heile målområdet gitt antall dyr (eller tettheit) i det gjennomsøkte området.

Den essensielle parameteren ved bestandsestimat frå linjetransektsøk (Schweder, et al. 1999) er den effektive halv-breidda av transektområdet gitt ved:

$$w = \int_0^{\infty} g(x)dx. \quad (4.1)$$

Funksjonen $g(x)$ (eng: detection function) beskriv sannsynligheten for å oppdage ein vågekval (dvs sjå minst eitt blåst) med avstand x vinkelrett frå transektlinja. La oss kalle $g(x)$ på norsk for deteksjonsfunksjonen. Størrelsen $a = 2wL$, der L er lengda av linjetransektet, gir arealet av det effektive søkeområdet, og den estimerte tettleiken av vågekval i dette området blir då $\hat{D} = n/a$, der n er antall observerte dyr.

For å finne eit estimat av w kan ein sette eit estimat av g inn i formel (4.1) over. Dersom vi antek at g tilhøyrer ein parametrisk familie kan vi estimere parametrane i g ved maximum likelihood.

Sidan vågekvalen er så vanskeleg å observere vil det ofte finnast dyr i transektlinja som aldri blir oppdaga, dvs $g(0) < 1$. Då er det vanleg å ta i betraktning to observatørar som søker i det same området samtidig og uavhengig av kvarandre. Vi kallar den eine observatøren A , den andre B og kombinasjonen av begge for $A \cup B$. Deteksjonsfunksjonane til A , B og $A \cup B$ er henholdsvis g_A , g_B og $g_{A \cup B}$, der $g_{A \cup B}$ er sannsynligheten for at enten A eller B eller begge oppdagar ein kval gitt at dei ikkje er merksame på kvalen i forkant. Observasjonar gjort av både A og B blir kalla duplikate. Fordi det oppstår målefeil frå observatørane er det ikkje trivielt å identifisere duplikate observasjonar, og feil identifisering medfører også feil i estimatet. Merk også at sjølv om A og B observerer uavhengig av kvarandre, så er likevel ikkje observasjonane dei gjer statistisk uavhengige når vi har diskret tilgjengelighet av dyr. (Skaug og Schweder 1999) Dette bør det også takast omsyn til i modellane.

Hazardmodellar er mykje brukt for å estimere g . Slike modellar er verdifulle når linjetransektsøket kan karakteriserast ved følgjande (Schweder, et al. 1999):

- Dyra i populasjonen er tilgjengelege for observasjon berre i korte diskrete tidspunkt.
- Dyra bevegar seg så sakte i forhold til observatøren at bevegelsane til dyra kan ignoreras.
- Det eksisterar to uavhengige observatørar som søker i området samtidig.
- Det same dyret kan være tilgjengeleg for observasjon fleire gonger, og observatøren registrerer estimert posisjon kvar gong han/ho ser dyret.

Hazardmodellar for tilfeller med diskret tilgjengelighet består av to hovudkomponentar. Den første er ei eksplisitt modellering av det fenomenet at observatøren(e) i eit linjetransektsøk gradvis nærmar seg dyra. Vi antar at kvalen blir værande i sin posisjon og at observatøren nærmar seg langs transektlinja med konstant hastighet v . Vi lar x være vinkelrett avstand frå kvalen til punktet p på transektlinja og y distansen fram til p ved eit gitt tidspunkt. Vi definerer hazardfunksjonen $Q(x, y)$ som den betinga sannsynligheten for at ein vågekval blir oppdaga for første gang gitt at han har eit blåst i (x, y) . Hazardfunksjonen reflekterer altså observatørens evne til å oppdage vågekval ved ulike avstandar. Vi antek at $Q(x, y) = 0$ for $y < 0$, dvs at ingen dyr blir oppdaga etter at dei har blitt passerte. Den andre komponenten til den diskrete hazardmodellen er knytt til dykkemønsteret. I diskrete tilfeller, som vi har med vågekval, kallar vi serien av tidspunkt der kvalen er tilgjengeleg for observasjon for $\{T_i, i \geq 0\}$ der T_0 er tidspunktet for det første blåstet etter at kvalen er passert, T_1 er tidspunktet for det siste blåstet før passering osv. Blåstprosessen, som genererer $\{T_i\}$ er altså den andre hovudkomponenten. Deteksjonsfunksjonen for *ein* observatør (eller eit team av observatørar) er då gitt ved:

$$g(x) = 1 - \exp\left\{-\frac{\alpha}{v} \int_0^\infty Q(x, y) dy\right\},$$

der α^{-1} er gjennomsnittleg tid mellom to etterfølgjande blåst. (Skaug og Schweder 1999)

Deteksjonsfunksjonen for $A \cup B$ er gitt ved (Skaug og Schweder 1999):

$$g_{A \cup B}(x) = 1 - \exp \left[-\frac{\alpha}{\nu} \{q_A(x) + q_B(x) - q_{AB}(x)\} \right],$$

Der $q_u(x) = \int_0^\infty Q_u(x, y) dy$ for $u = A, B$ og $q_{AB}(x) = \int_0^\infty Q_A(x, y) Q_B(x, y) dy$.

Modellen for Q må så langt som mogeleg reflektere dei fysiske realitetane, men Q må også være matematisk løyseleg. Dette gjer problemet vanskeleg fordi desse to omsyna ofte står i motsetnad til kvarandre.

Skaug og Schweder (1999) presenterer to ulike modellar for hazardfunksjonen som begge er nyttige i praksis fordi problemet blir matematisk løyseleg. Det blir antatt at blåstprosessen følgjer ein Poissonprosess. Det blir antatt at det ikkje fins målefeil eller varierende forhold under søket.

Schweder et al. (1999) konstruerar ein likelyhoodfunksjon der informasjonen frå $A \cup B$ er i tre grupper slik:

- a) Kvalen sin posisjon i forhold til observatøren ved første observerte blåst, (r, θ) , der r er radial distanse og θ er vinkelen mellom transektlinja og linja mellom kval og observatør.
- b) Vart det første blåstet observert av berre ein av observatørane eller begge? La $u \in \{1, 2\}$ være talet på dette.
- c) Når det første blåstet vart oppdaga av berre den eine observatøren er det ein Bernoulli-trial for kvart etterfølgjande blåst med utfallet "suksess" dersom blåstet også blir observert av den andre observatøren og utfallet "fiasko" elles. La (r_i, θ_i) være kvalen sin posisjon i forhold til observatøren ved det i 'ende etterfølgjande blåstet og definer d_i som $d_i = 1$ dersom den i 'ende Bernoulli-trial'en er ein suksess og $d_i = 0$ elles.

Etter at ein observatør har oppdaga ein vågekval er ikkje sannsynligheten for at han observerer etterfølgjande blåst styrt av hazardfunksjonen. Antalet Bernoulli-trials, I , i c) er definert som den første verdien av i som gir $d_i = 1$. Likelyhoodfunksjonen blir da gitt ved:

$$L_{TOT} = f_1(r, \theta) \cdot f_2(u|r, \theta) \cdot \prod_{i=1}^I f_3(d_i|r_i, \theta_i),$$

der det siste leddet er lik ein dersom $I = 0$.

Schweder et al. (1999) seier at sidan det ikkje er tatt omsyn til målefeil vil direkte bruk av denne loglikelyhoodfunksjonen ikkje gje forventningsrette estimat, men at ein korrigerer denne feilen kan bli gjort ved ein transformasjon av argumenta, dvs parametrane, i modellen. Å bestemme denne transformasjonen innebærer simulering av data *med* målefeil. Transformasjonen blir kalla for ei bru, F . Eit estimat av ein glatt funksjon F kan vi finne ved regresjonsanalyse av dei simulerte dataene.

Williams, Leaper, et al. (2007) har gjort undersøkingar av målefeil ved linjetranseksøk. Han viser at målefeil utgjør ei stor feilkilde i berekning av estimatet, og at det er behov for betre målemetodar og måleutstyr under sjølve søket.

Hedley og Buckland (2004) poengterer at dyr i den verkelege verda ikkje er uniformt fordelte, men at tettleiken av dyra varierer med i samsvar med miljø, tilgang på mat o.a. Dei presenterar difor ein modell som let tettleiken til dyra være klustra. Dei antek då at dyra er lokaliserte i henhold til ein ikkje-homogen Poissonprosess med intensitet $D(x, y)$. Waagepetersen og Schweder (2004) foreslår i staden ein shot-noise Cox-prosess medan Skaug (2006) foreslår ein Markovmodulert Poissonprosess som modell for spredningen av vågekvalane.

Williams, Hedley og Hammond (2006) peiker på at vanlege linjetranseksøk medfører store økonomiske kostnader for dei som administrerer det og at dette begrensar kven som mogelegheit til å utføre slike søk. Dei foreslår difor, i staden for å bruke eigne båtar i designa søk, å bruke båtar som er tilgjengelege i området, til dømes fiskebåtar, lastebåtar eller cruise-skip (eng: ships of opportunity). Dei har også gjennomført eit slikt søk.

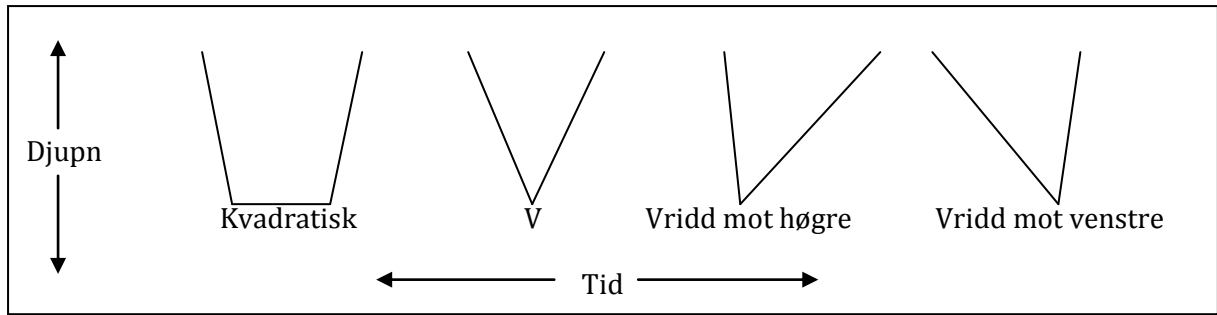
Arbeidet med å utvikle metodar og modellar for bestandsestimat av vågekval er på ingen måte avslutta. Det er stor internasjonal interesse for kval og kvalforskning, og dette saman med IWC sitt krav om vitskapeleg dokumentasjon er faktorar som tilseier at vi kan vente stor utvikling også i åra som kjem.

5 SYSTEMATISERING AV DYKK

Det er mange artar av fuglar og pattedyr som, trass i at dei er avhengige av å puste luft, brukar mykje tid under vatn. Det er også samla inn ein god del informasjon om dykkeatferden til desse artane. Noko av denne informasjonen er henta inn spesielt med tanke på bruk i bestandsestimat. Døme på dette er Øien, Bøthun og Kleivane (2007) som presenterer info om blåstrater og dykketider frå vågekval, og Heide-Jørgensen og Simon (2007) som viser ei oversikt over blåstrater frå vågekval, finnkval og knølkval. Men informasjon om dykkemønster er også interessant for anna forskning innan zoologi og økologi.

Det er gjort mange studier av dykk som med det formål å kartlegge jaktadferd Døme er Watwood, et al. (2006) som har analysert dykkedata frå spermkval og samanlikna jaktatferden til spermkval frå tre ulike havområde: Atlanterhavet, Mexicogulfen og Liguriahavet. Fowler, et al. (2006) brukar dykkedata frå sjøløver for å samanlikne jaktatferden til dyr av ulike aldrar, frå 6-månadars alder til vaksne, og Tinker, et al. (2007) har studert dykkedata for å finne ut om den Californiske havoterer endrar jaktadferd etter kva bytte han fangar.

López, et al. (2008) har analysert dykkeatferden til turistattraksjonen tumlaren for å finne ut om den blir påverka av båttrafikken. Dykkedata ved nærvær av båtar vs. fråvær av båtar har blitt undersøkt, og dei viser at tumlaren responderer på båttrafikk med endra atferd. Liknande resultat har vi også for vågekval.



Figur 5.1 Dykkeprofilar

Brukt av Schreer, Kovacs og Hines til klassifisering av dykk.

Schreer, Kovacs og Hines (2001) har henta inn informasjon om dykkeatferden til tolv ulike artar: Knoppskarv, adeliepingvin, kvitkinnpingvin, bøylepingvin, galápagospelsel, antarktisk pelsel, grønlandssel, klappmyss, weddelsel, elefantsel og hvalross. Dei klassifiserar dykka til alle artane etter dykkeprofil (sjå figur 5.1) og samanliknar deretter ulike dykkevariablar, bl.a. dykketid og tid ved overflata mellom dykk i dei ulike profilane. Samanlikningar blir gjort både mellom artane og mellom individ av same art. Det er også gjort berekningar av forventa aerobisk grense for dykka sin varigheit. Konklusjonen til Schreer, Kovacs og Hines er at det er store likheiter i dykkemønster på tvers av artsopprinnelse.

Her er det definitivt potensiale for å dra ut meir informasjon: Det er samla inn mengder av data, og med modellar kan vi få ei enda betre beskriving av situasjonen.

6 MATEMATISKE MODELLAR AV DYKKETIDER

Vi har sett at modellar for dykketider inngår i bestandsestimat av vågekval. Fordi tida vågekvalen brukar ved overflata er så kort i forhold til dykketida, kan vi sjå blåsta som punkt i tid. Poisson-prosessen har vore mykje brukt for å modellere dette. Det siste tiåret har det imidlertid kome eit par nye forslag, og vi skal berre kort nemne desse.

Schweder (2003) foreslår ein modell med ein blanding av to gammafordelingar. Fordelinga til dykketidene blir med hans modell

$$f(x) = qg(x|s_1, r_1) + (1-q)g(x|s_0, r_0),$$

der g er gammafordelingen.

Schweder og Øien (2007) har valt å først kategorisere dykketidene i korte, mellomlange og lange dykk. Dei antek deretter lognormalfordeling for dei lange (l) og dei korte (s) dykka og uniform fordeling for dei mellomlange dykka. Dykketidene har då følgjande fordeling:

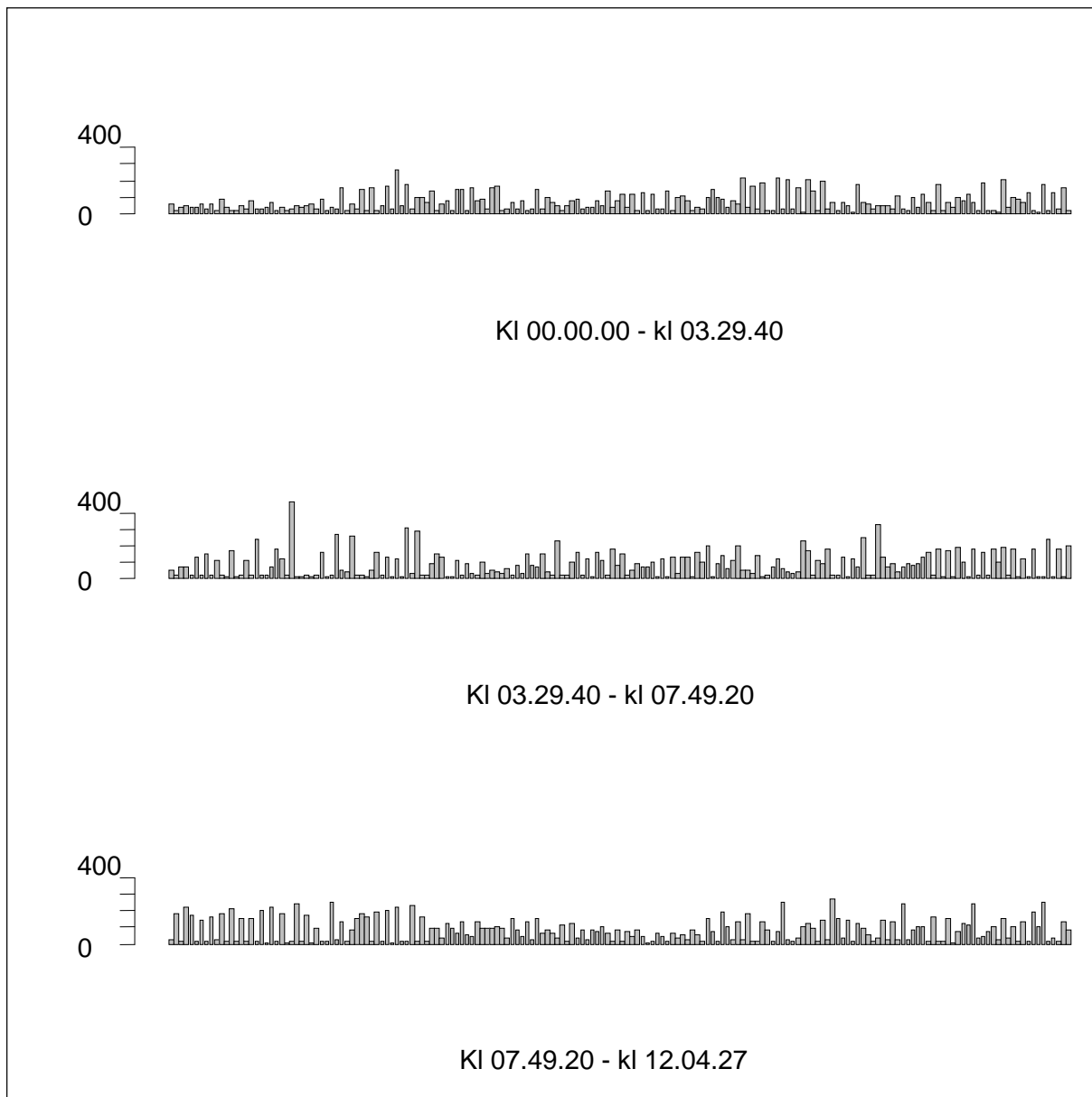
$$f(d) = p \frac{1}{\sigma_s} \Phi\left(\frac{\log(d) - \mu_s}{\sigma_s}\right) / d + q \frac{1}{105} u\left(\frac{d-35}{105}\right) + r \frac{1}{\sigma_l} \Phi\left(\frac{\log(d) - \mu_l}{\sigma_l}\right) / d,$$

der $p + q + r = 1$, Φ er normalfordeling og u er uniform fordeling.

7 GRAFISK FRAMSTILLING

La oss anta at tid brukt ved overflata lik null. Da blir varigheita av dykka lik tida mellom to målepunkt, og dykketida varierer frå 7 sekund til 7 minutt og 43 sekund.

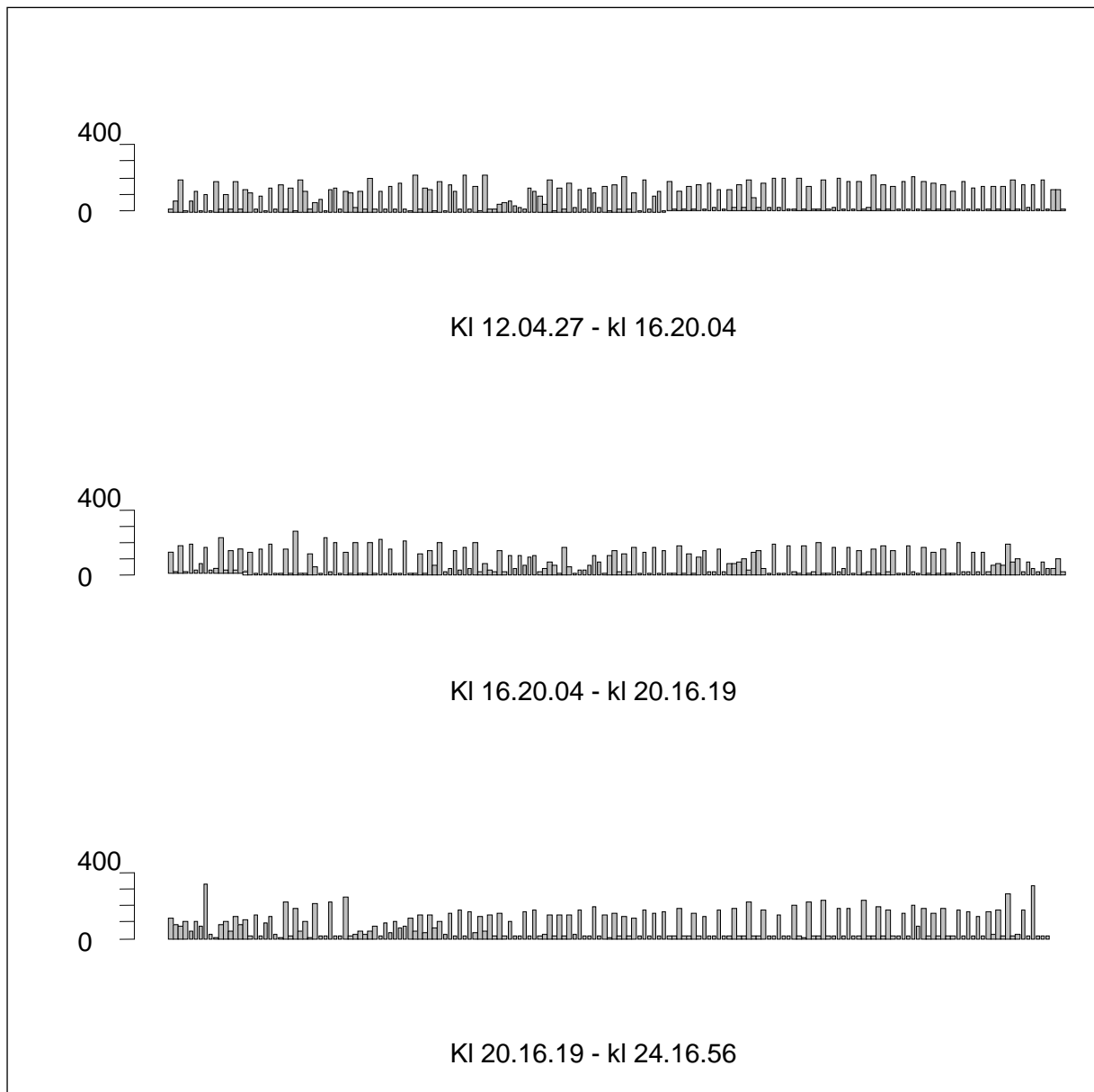
Figurane 7.1 og 7.2 viser vågekvalen sine dykk det døgnet han vart observert. Dykka, representerte ved søyler med høgde svarande til varigheit, er vist i kronologisk rekkefølge. Klokka er satt til 00.00.00 ved start. Dei første 20 minutta etter målestart utfører vågekvalen ei serie dykk på mellom 20 og 90 sek. Denne atferden ser vi ikkje att seinare i observasjonen, og det er rimeleg å anta at kvalen her har ein stressreaksjon etter påføringa av radiomerket.



Figur 7.1 Suksessivt påfølgande dykk

det første halvdøgnet av observasjonen. Kvart dykk er representert ved ei søyle, og høgda på søyla viser varigheita til dykket i sekund.

Dei neste tolv timane trer det ikkje fram noko klart mønster i dykkeatferden, bortsett frå i enkelte små intervall der kvalen etter kvart lange dykk har 1-2 påfølgjande korte dykk.

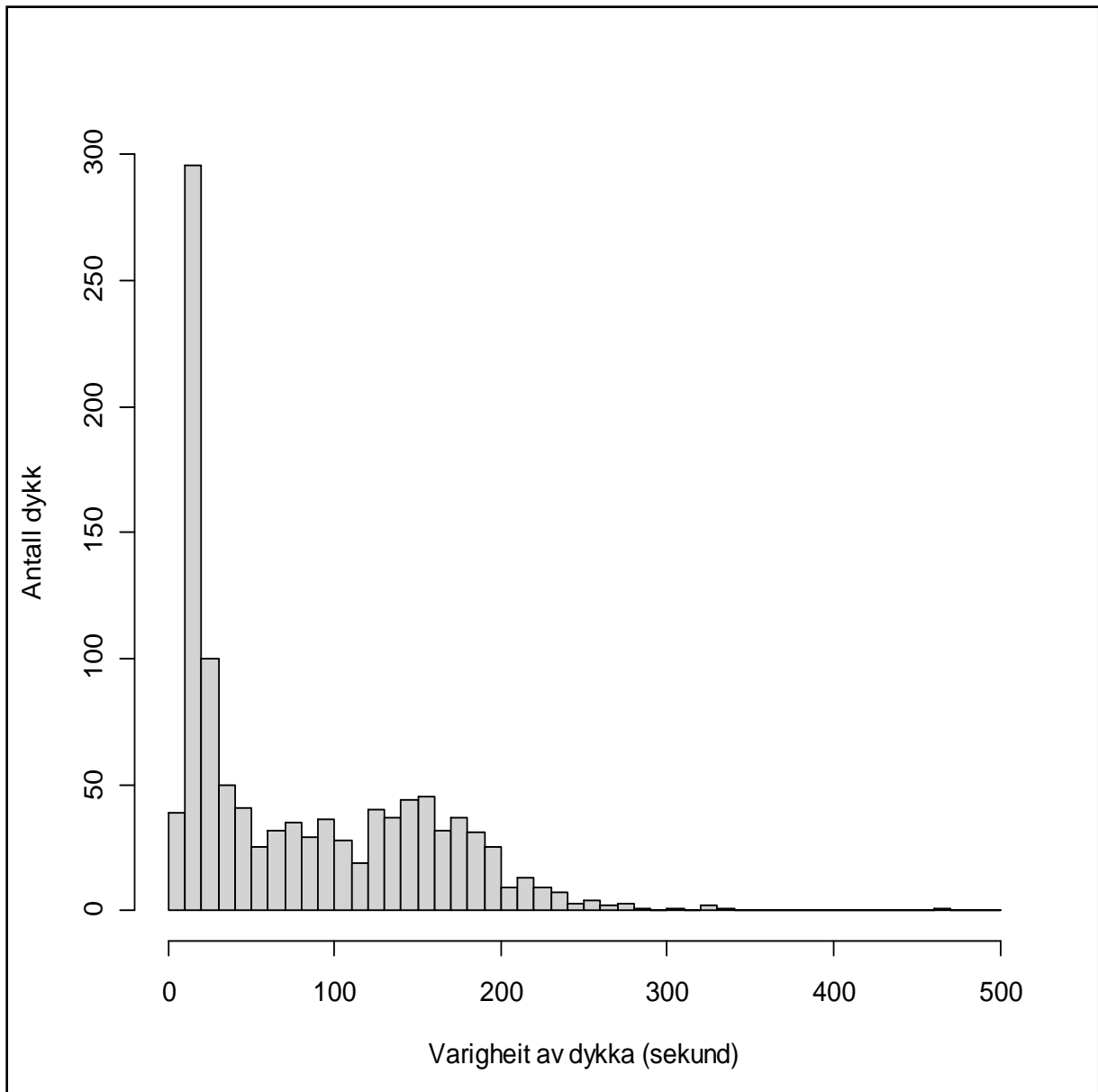


Figur 7.2 Suksessivt påfølgjande dykk

det andre halvdøgnet av observasjonen. Kvart dykk er representert ved ei søyle, og høgda på søyla viser varigheita til dykket i sekund.

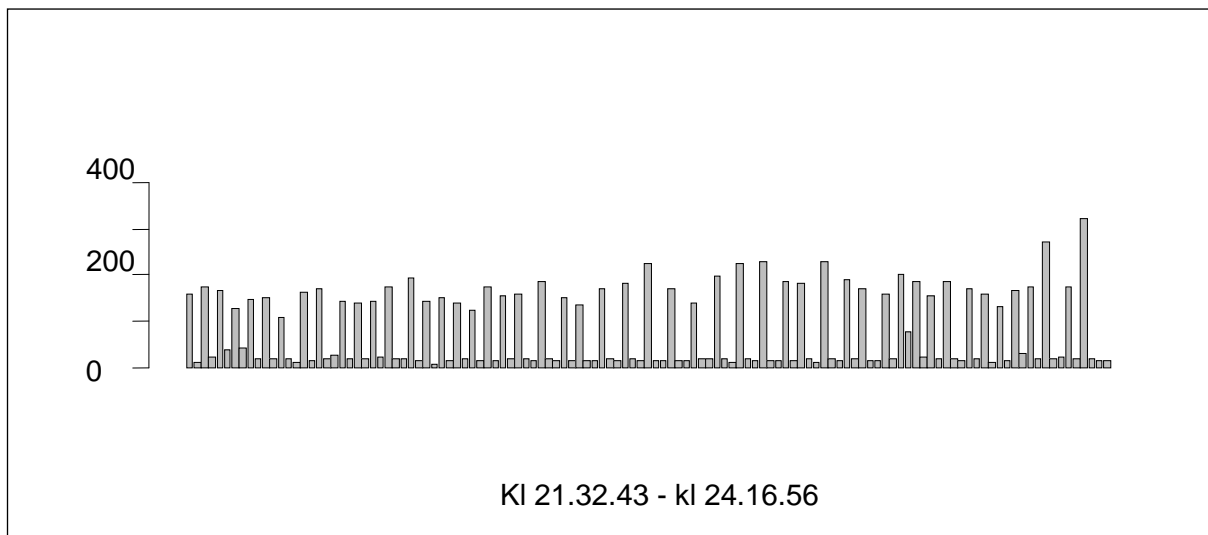
I dei siste tolv timar og tolv minutt av observasjonen har atferden endra seg. Det som førekom i korte intervall den første halvdel av observasjonen er no det dominerande mønsteret: kvart lange dykk blir følgt av 1-2 korte. No ser vi berre kortare intervall utan klart dykkemønster.

Lat oss sjå på frekvensen av dei ulike dykketidene. Histogrammet i figur 7.3 viser at dykk av varigheit 10-20 sekund opptre svært ofte, nesten 300 gonger på 24,28 timar. Vågekvalen har også hyppige (ca 100) dykk på mellom 20 og 30 sekund. Dykk på over 200 sekund er lite vanlege, og dykk på over 300 sekund er sjeldne.



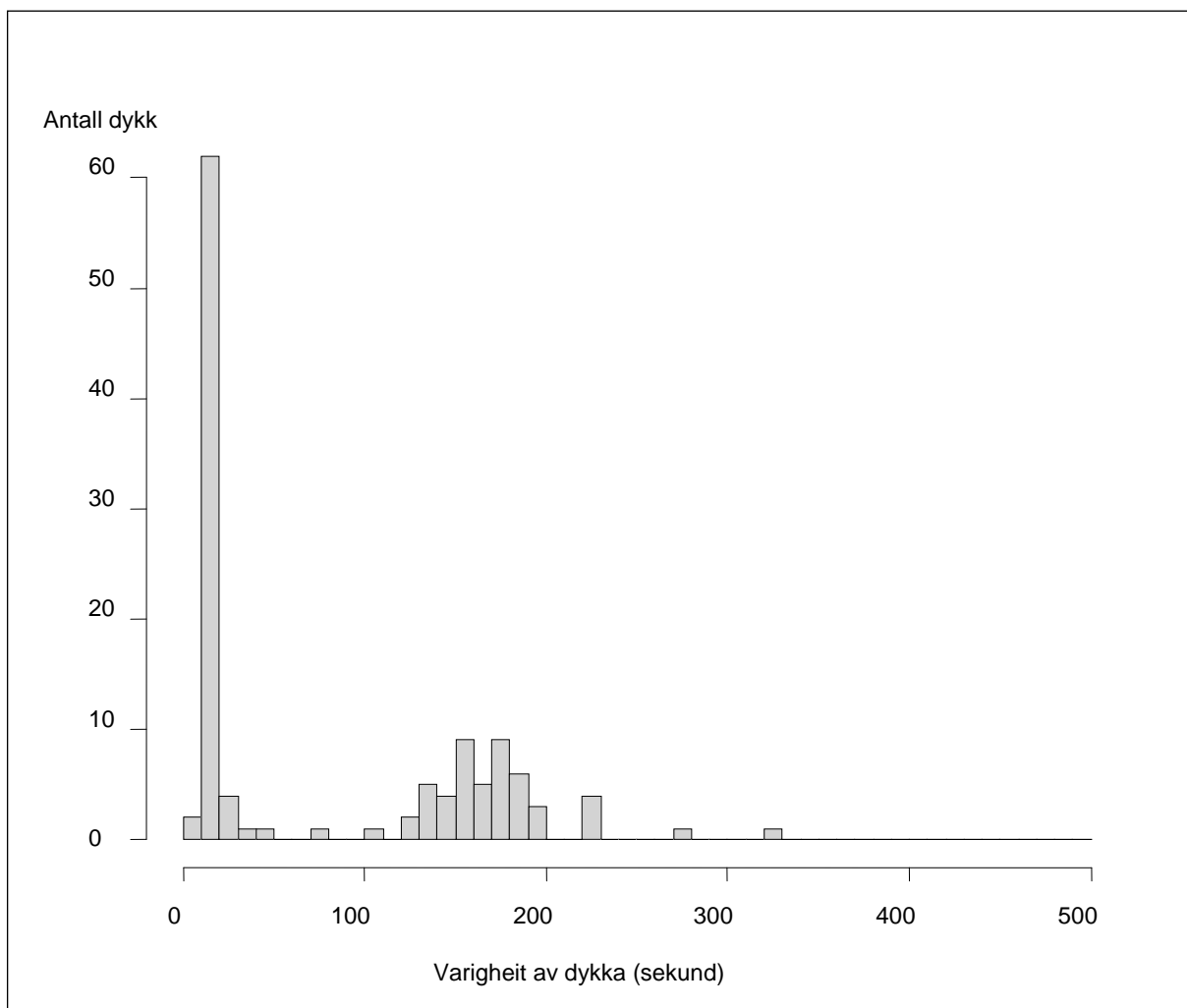
Figur 7.3 Dykketid, frekvens.

I figur 7.4 kan vi sjå nærare på mønsteret vi fann i vågekvalen sine dykk. Figuren viser eit utsnitt av dykka dei siste 2 timar og 3 kvarter av observasjonstida. Vi ser nesten ingen dykk som avviker frå mønsteret på eit langt dykk og 1-2 korte, og skilnaden i varigheit er distinkt. Dette blir bekrefta i figur 7.5 der vi ser at varigheita på dykka klumpar seg i to grupper. Dei korte dykka er i hovudsak på under $\frac{1}{2}$ minutt og dei lange ligg på mellom 2 og $3\frac{5}{6}$ minutt. Dei korte dykka varierer mindre i varigheit enn dei lange.



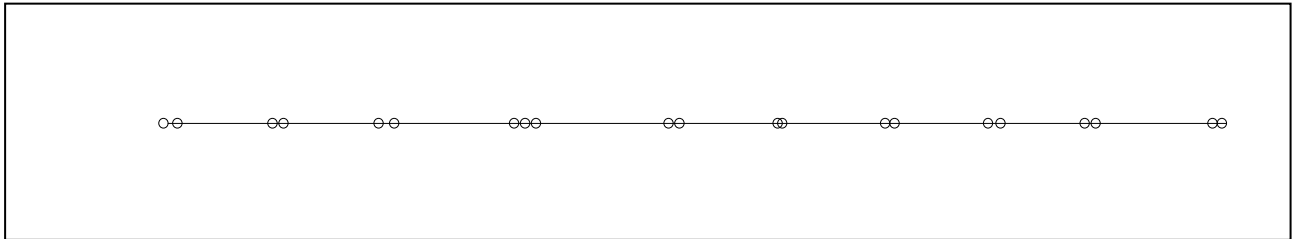
Figur 7.4 Dykkemønster vist som ei serie dykk.

Dykkene er frå dei siste 2 timar 44 minutt og 13 sekund av observasjonstida. Høgda på søylene viser varigheita til dykkene i sekund



Figur 7.5 Dykketid dei siste 2 timar 44 minutt og 13 sekund av observasjonstida.

I staden for å sjå vågekvalens dykkemønster som ei serie dykk, skal vi no sjå mønsteret som ei serie blåst. Figur 7.6 viser alle blåst i tidsrommet 22.00 - 22.30 etter observasjonsstart. Blåsta blir vist som punkt på ei tidslinje, og sett på denne måten vil dykketida representere ventetida mellom blåst. Vi ser mønsteret: 2-3 blåst på kort tid og deretter lang ventetid fram til neste gruppe med 2-3 blåst.



Figur 7.6 Dykkemønster vist som ei serie blåst.

Blåsta er plasserte som punkt på ei tidslinje. Vi ser blåsta i tidsrommet frå kl 22.00.00 til kl 22.30.00 etter start.

8 PROBLEMSTILLING OG RAMMER

Serien av blåst kan beskrivast matematisk som ein stokastisk punktprosess i tid. Vi ser kvart blåst som eit punkt, medan tida mellom blåst varierer stokastisk.

Den vanlege modellen for punktprosessdata er Poisson-prosessen. Den forutset at dykketidene er eksponentialfordelte og uavhengige. Som vi har sett i figurane 7.4 og 7.6 fins det eit mønster i dykkeatferden til vågekvalen. Dette indikerer at varigheita av eit dykk *ikkje* er uavhengig av varigheita av førre dykk, og vi må vurdere om ein homogen Poisson-prosess vil passe som modell for våre dykketidsdata.

Ein Markovmodulert Poisson-prosess tillet data å være "klustra". Denne kan difor fungere som ein betre modell for våre data. Da kan vi la intensiteten λ i Poisson-prosessen variere mellom to nivå: λ_1 (djupdykksfase) og λ_2 (overflatefase) slik at $\lambda_2 > \lambda_1$. Ein kontinuerligtids

Markovprosess med intensitetar μ_1 og μ_2 vil bestemme kva for ein tilstand kvalen er i til ei kvar tid.

Rammene for oppgåva er enkle: Vi har dykketidsdata frå *ein* vågekval, og vi antar at tid brukt ved overflata er lik null.

9 POISSON-PROSESSEN

Definisjon 11.1

Ein Poisson-prosess med intensitet $\lambda > 0$ er ein stokastisk prosess $\{X(t); t \geq 0\}$ av heiltalsverdi der følgjande påstandar er sanne:

- i. For alle tidspunkt $t_0 = 0 < t_1 < t_2 < \dots < t_n$ er

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

uavhengige stokastiske variable.

- ii. For $s \geq 0$ og $t > 0$ har den stokastiske variabelen $X(t+s) - X(s)$ poissonfordeling

$$P[X(s+t) - X(s) = k] = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad \text{for } k = 0, 1, \dots$$

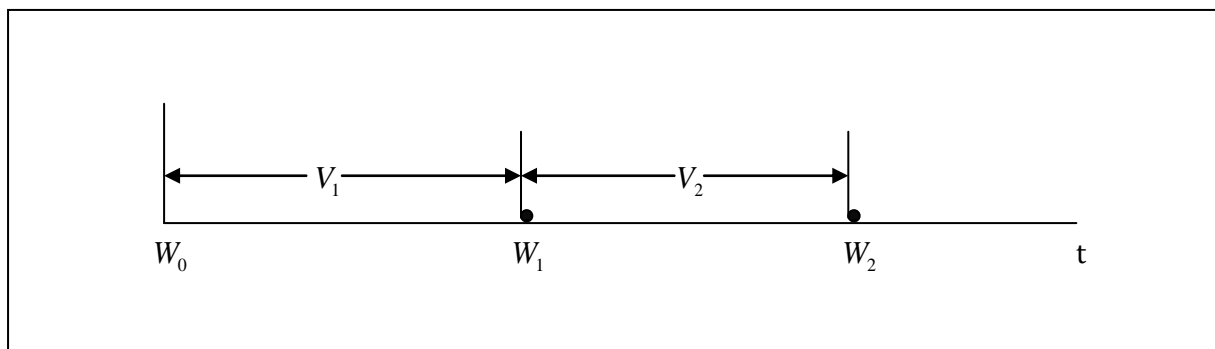
- iii. $X(0) = 0$

Intensiteten λ i ein Poisson-prosess er proporsjonalitetskonstanten i sannsynligheten for at ei hending skal oppstå i eit bitte lite intervall. Det vil sei

$$\begin{aligned} P[X(t+h) - X(t) = 1] &= \frac{e^{-\lambda h} (\lambda h)}{1!} \\ &= (\lambda h)(1 - \lambda h + \frac{1}{2} \lambda^2 h^2 - \dots) \\ &= \lambda h + o(h), \end{aligned}$$

der $o(h)$ er eit uspesifisert restledd av orden mindre enn h .

Figur 9.1 viser gangen i ein typisk poisson punktprosess. W_n er tidspunktet for den n 'te hendinga og $V_n = W_n - W_{n-1}$ er ventetida mellom hendingane.



Figur 9.1 Døme på Poisson punktprosess.

Ventetidene V_n i Poisson-prosessen har følgjande sannsynlighetsfordeling:

Teorem 11.1

Ventetidene V_0, V_1, \dots, V_{n-1} er uavhengige og eksponentialfordelte stokastiske variable

$$f_{V_k}(v) = \lambda e^{-\lambda v}, \quad v \geq 0$$

For bevis sjå side 291-293 i Taylor og Karlin (1998).

Dersom intensiteten $\lambda = \lambda(t)$ endrar seg over tid, seier vi at vi har ein ikkje-homogen Poisson-prosess. Den stokastiske variabelen $X(t) - X(s)$ har då poissonfordeling med parameter $\int_s^t \lambda(t) dt$. (Taylor og Karlin 1998)

10 COXPROSESS, KONTINUERLIGTIDS MARKOVKJEDER OG POISSON-PROSESS MED MARKOV-INTENSITET

Anta at $X(t)$ er ein ikkje-homogen Poisson-prosess der intensitetsfunksjonen $\{\lambda(t), t \geq 0\}$ sjølv er ein stokastisk prosess. Da har vi ein dobbelt stokastisk prosess, også kalla Cox prosess. I motsetning til det som er tilfelle i ein poissonprosess gjeld det ikkje generelt for alle $t_0 = 0 < t_1 < t_2 < \dots < t_n$ at $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ er uavhengige stokastiske variablar (Taylor og Karlin 1998). Det gjeld imidlertid generelt for Coxprosessar i intervallet $[0, V]$ at (Skaug 2006):

$$E[n] = \int_0^V E[\lambda(w)] dw, \text{ og}$$

$$Var[n] = E[n] + Var\left[\int_0^V \lambda(w) dw\right]$$

Ein Markov-prosess $\{X_t\}$ er ein stokastisk prosess med den eigenskapen at verdien av X_s for $s > t$ gitt verdien av X_t ikkje blir påverka av verdien til X_u for $u < t$. Det vil sei at sannsynlegheita for neste hending berre er avhengig av den noverande tilstanden og ikkje av tidlegare hendingar (Taylor og Karlin 1998).

Ei kontinuerlegtids Markov-kjede $X(t)$ ($t > 0$) er ein Markov-prosess med tilstandene $0, 1, 2, \dots$ Vi antek at overgangs-sannsynlegheitene er stasjonære slik at

$$P_{ij}(t) = \Pr\{X(t+s) = j | X(s) = i\}.$$

Dersom tilstandsrommet er endeleg og tilstandane merka $\{0, 1, 2, \dots, N\}$ vil $P_{ij}(t)$ tilfredsstillje følgjande påstander (Taylor og Karlin 1998):

a) $P_{ij}(t) \geq 0$

b) $\sum_{j=0}^N P_{ij}(t) = 1, \quad i, j = 0, 1, 2, \dots, N$

c) $P_{ik}(s+t) = \sum_{j=0}^N P_{ij}(s)P_{jk}(t)$ for $t, s \geq 0$ (Chapman-Kolmogorov-likheten)

d) $\lim_{t \rightarrow 0^+} P_{ij}(t) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Ein Markovmodulert Poisson-prosess er ein Cox-prosess der intensitetsfunksjonen er ein Markov-prosess med endeleg tilstandsrom (Skaug 2006). I denne oppgåva skal vi anta at det berre finst to tilstandar: λ_1 og λ_2 ($\lambda_1 < \lambda_2$). Lat

$$\mathbf{Q} = \begin{matrix} & \lambda_1 & \lambda_2 \\ \begin{matrix} \lambda_1 \\ \lambda_2 \end{matrix} & \begin{bmatrix} -\mu_1 & \mu_1 \\ \mu_2 & -\mu_2 \end{bmatrix} \end{matrix}$$

være generatormatrisa for Markov-prosessen $\lambda(t)$. Ventetidene i tilstand λ_1 er uavhengige og eksponentialfordelte med parameter μ_1 , og ventetidene i tilstand λ_2 er uavhengige og eksponentialfordelte med parameter μ_2 (Taylor og Karlin 1998). Lat $\boldsymbol{\pi} = (\pi_1, \pi_2)$ være initialfordelingane slik at $P(\lambda(0) = \lambda_i) = \pi_i, \quad i = 1, 2$. Desse initialfordelingane er gitt ved:

$$\pi_1 = \frac{\mu_2}{\mu_2 + \mu_1} \quad \text{og} \quad \pi_2 = \frac{\mu_1}{\mu_2 + \mu_1}$$

Vi har tidlegare definert ventetidene mellom blåst som $v_i = w_i - w_{i-1}, \quad i = 1, 2, \dots, n$ med $w_0 = 0$. MMPP gjer det mogeleg å finne uttrykket for fordelinga til v_1, v_2, \dots, v_k slik (Skaug 2006): Definer

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad \text{og} \quad \mathbf{C} = \mathbf{Q} - \mathbf{\Lambda}.$$

Definer også $\exp[\mathbf{C}v] = \sum_{i=0}^{\infty} \mathbf{C}^i v^i / i!$ der \mathbf{C}^0 er identitetsmatrisa \mathbf{I} . Sannsynlighetsfordelinga til dykketidene v_1, v_2, \dots, v_k er da gitt ved:

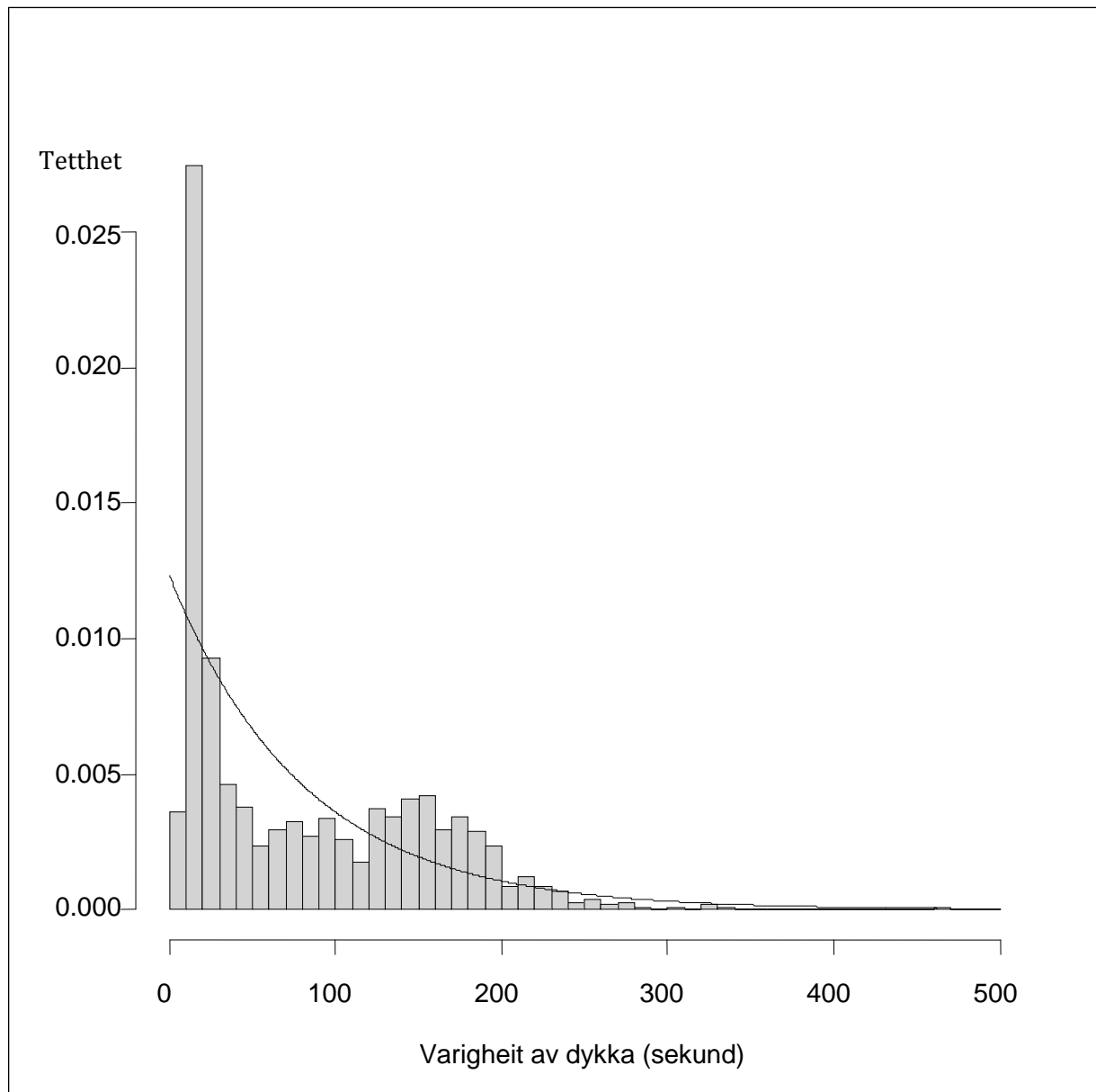
$$f(v_1, v_2, \dots, v_k) = \boldsymbol{\pi} \exp[\mathbf{C}v_1] \mathbf{\Lambda} \exp[\mathbf{C}v_2] \mathbf{\Lambda} \cdots \exp[\mathbf{C}v_k] \mathbf{\Lambda} \mathbf{e}, \quad v_i > 0, \quad (10.1)$$

Der $\mathbf{e} = (1, 1)$. Denne fordelinga er ubetinga, og vi treng ikkje å ha gitt at $\sum_{i=1}^k v_i \leq V$.

11 KRITERIER

I ein Poisson-prosess er ventetidene mellom hendingar uavhengige og eksponentialfordelte slik at $f_{V_k}(v) = \lambda e^{-\lambda v}$, $v \geq 0$. Lat oss undersøke om våre data oppfyller desse kriteria.

Vi hugsar at ventetidene mellom blåst er lik dykketidene. Histogrammet i figur 7.3 side 15 viser frekvensen av dei ulike dykketidene. Figur 11.1 viser dei same dataene, men her med forholdsvis frekvens slik at arealet av dei grå søylene i alt er lik 1.



Figur 11.1 Dykketid, fordeling.

Histogrammet viser fordeling av dykka sin varighet.

Den svarte kurven viser eksponentialfordeling med parameter λ estimert ved MLE.

Histogrammet i figur 11.1 er altså fordelinga av dykketidene i vårt datasett. Denne fordelinga vil vi no samanlikne med eksponentialfordeling. Vi brukar maximum likelyhood estimation og finn

$$\text{parameteren } \hat{\lambda} = \frac{n}{\sum_{i=1}^n v_i} = 0.0123204 .$$

Eksponentialfordelinga med parameter $\hat{\lambda}$ er plotta inn i figur 11.1 som ei svart kurve. Vi ser tydelig avvik mellom denne og histogrammet . I intervallet (0,10] når ikkje søyla i histogrammet lenger opp enn til ca $\frac{1}{3}$ av eksponentialkurva, medan det i neste intervall (10,20] er søyla i histogrammet som er langt over kurva. Dykk med varighet frå 30 til 120 sekund har *sjeldnare* forekomst i våre data enn det som er forventa ved eksponentialfordeling, og dykk på mellom 120 og 200 sekund skjer *oftare* enn forventa. Eksponentialfordelinga ser altså ikkje ut til å beskrive våre data så godt. Vil ein Goodness-of-fit-test bekrefte inntrykket?

Goodness-of-fit-testen er basert på verdien av

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i},$$

der χ^2 er tilnærma kji-kvadratfordelt med $k - 1$ frihetsgrader. Symbola o_i og e_i representerer henholdsvis observerte og forventa frekvenser i det i'ende intervallet.

Intervall	o_i	e_i	Intervall	o_i	e_i	Intervall	o_i	e_i
(0,10]	39	124,8	(90,100]	36	41,2	(180,190]	31	13,6
(10,20]	296	110,4	(100,110]	28	34,4	(190,200]	25	12,0
(20,30]	100	97,6	(110,120]	19	32,2	(200,210]	9	10,6
(30,40]	50	86,3	(120,130]	40	28,5	(210,220]	13	9,4
(40,50]	41	76,3	(130,140]	37	25,2	(220,230]	9	8,3
(50,60]	25	67,4	(140,150]	44	22,2	(230,250]	10	13,8
(60,70]	32	59,6	(150,160]	45	19,7	(250,280]	9	15,3
(70,80]	35	52,7	(160,170]	32	17,4	(280,500]	6	31,9
(80,90]	29	46,6	(170,180]	37	15,4			

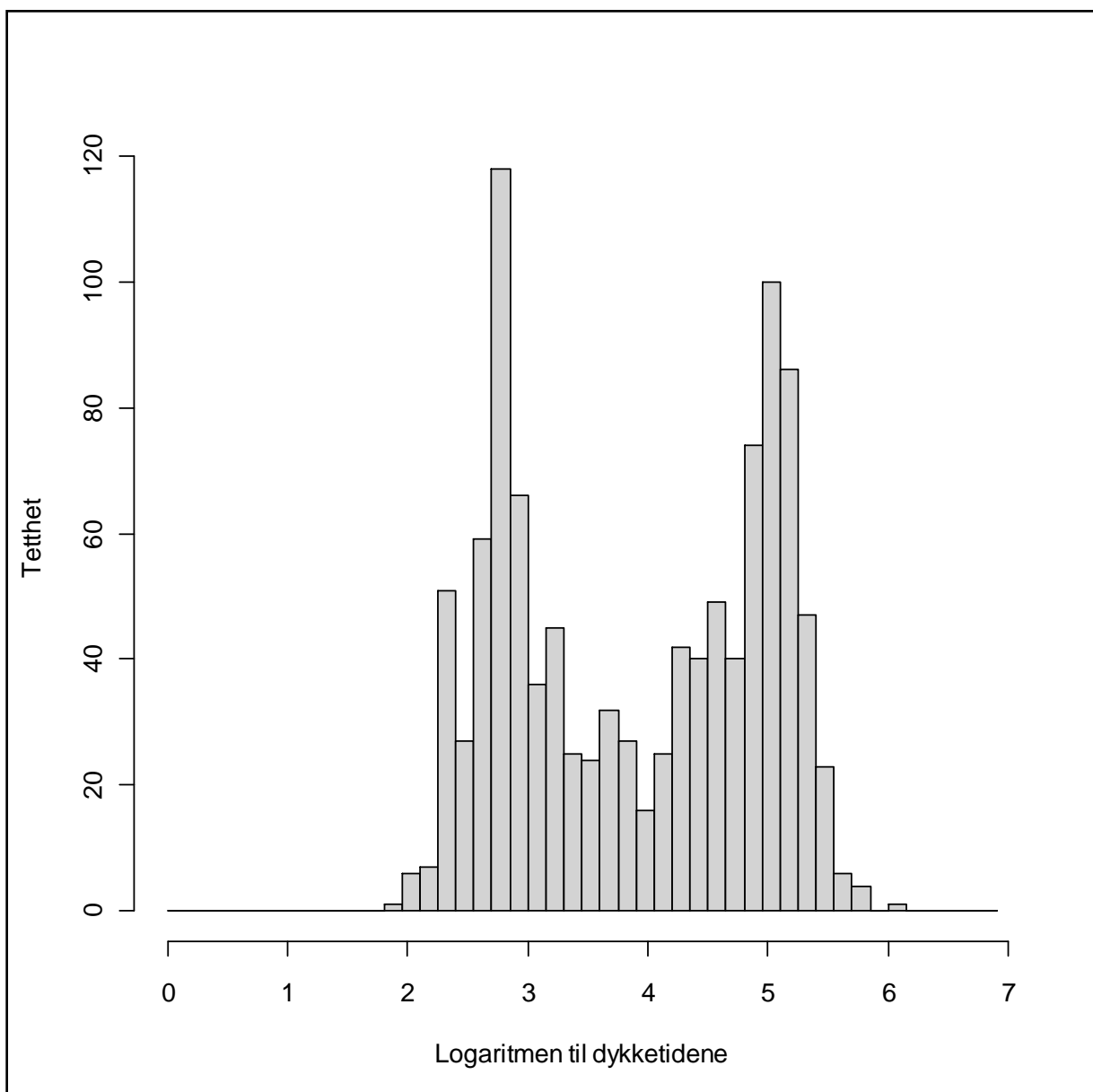
Tabell 11.1 Observerte og forventa frekvensar av dykketider.

Intervalla er for dykketid i sekund, o_i er observert frekvens og e_i forventa frekvens ved eksponentialfordeling.

Tabell 11.1 gir ei oversikt over alle frekvensane, både observerte og forventa. Dei forventa frekvensane er lik arealet under eksponentialkurva gitt antall observasjonar (1077), og er avrunda til ein desimal.

Frå tabell finn vi $\chi^2_{0,001} = 52,620$ for 25 frihetsgrader, og dersom eksponentialfordelinga beskriv dykketidene på ein god måte vil χ^2 ha lågare verdi enn dette. Vi kalkulerar og finn $\chi^2 = 632.0269$. χ^2 er altså mykje større enn $\chi^2_{0,001}$, og vi kan konkludere med at dykketidene ikkje har eksponentiell fordeling.

Dersom vi tek logaritmen til dykketidene får vi data med fordeling som vist i Figur 11.2:



Figur 11.2 Fordeling av logaritmen til dykketidene.

Vi ser at logaritmen til dykktidene ligg nærare normalfordeling enn dykktidene.

Lat oss finne ut om dykktidene er uavhengige av kvarandre. Ved å kalkulere autokorrelasjonen får vi eit mål på den lineære samanhengen mellom dykka. Vi ynskjer også å finne autokorrelasjonen mellom logaritmen til dykktidene. Autokorrelasjonen har verdier i intervallet [-1,1], og er gitt ved:

$$\rho_k = \frac{\text{Cov}(V_t, V_{t+k})}{\text{Var}(V_t)}$$

Vi estimerer autokorrelasjonane ved:

$$r_k = \frac{\sum_{i=1}^{n-k} (v_i - \bar{v})(v_{i+k} - \bar{v})}{\sum_{i=1}^n (v_i - \bar{v})^2},$$

der \bar{v} er middelerdien. Vi får følgjande autokorrelasjonskoeffisientar :

Lag (k)	1	2	3	4	5	6	7	8	9	10
r_{D_k}	-0.511	0.248	-0.066	0.054	-0.005	0.036	0.036	-0.038	0.060	-0.039
$r_{\log D_k}$	-0.543	0.309	-0.106	0.078	-0.016	0.050	0.022	-0.022	0.052	-0.043

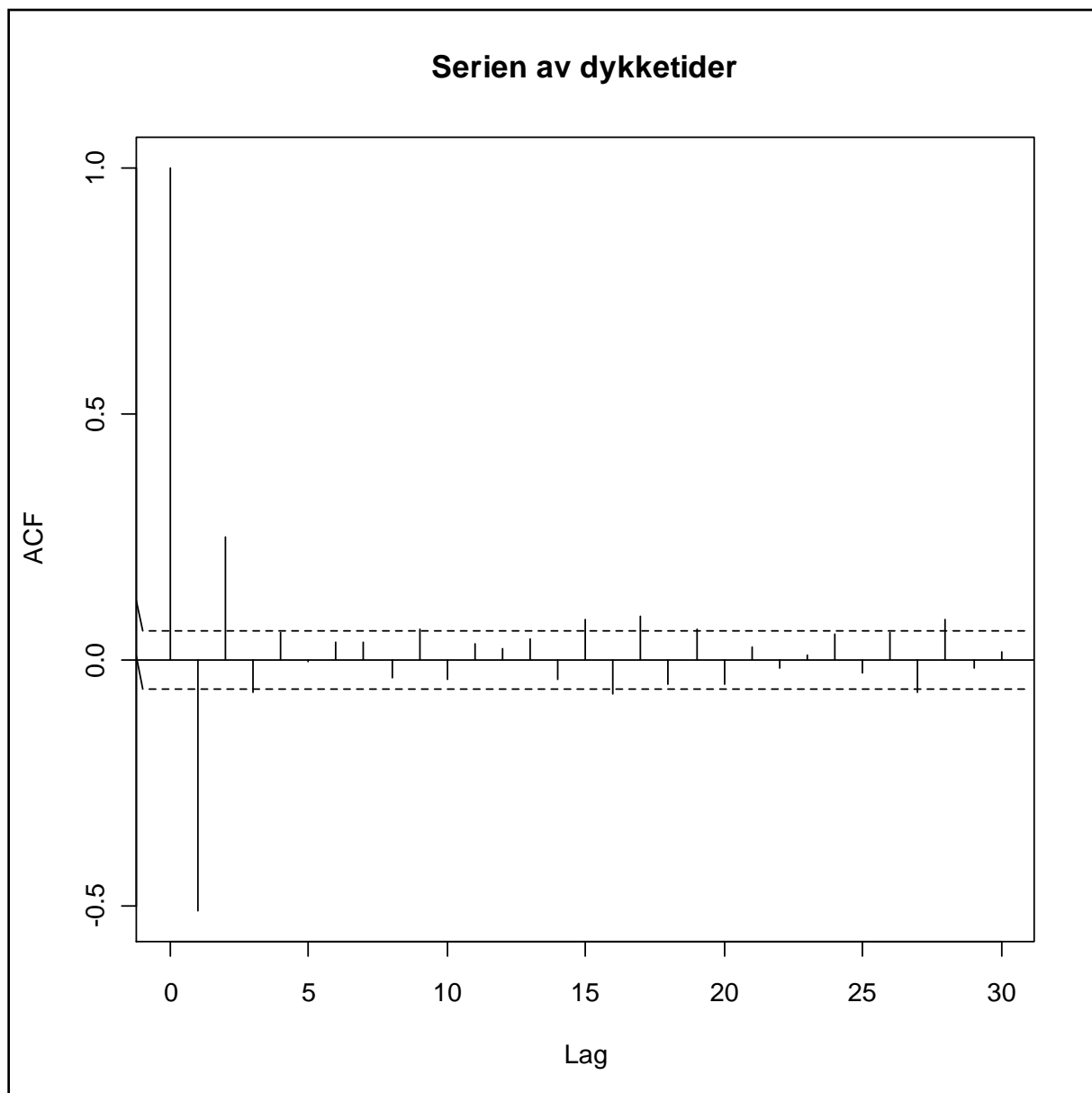
Lag (k)	11	12	13	14	15	16	17	18	19	20
r_{D_k}	0.031	0.022	0.043	-0.040	0.082	-0.071	0.088	-0.052	0.063	-0.052
$r_{\log D_k}$	0.047	0.014	0.030	-0.028	0.077	-0.061	0.092	-0.038	0.068	-0.049

Lag (k)	21	22	23	24	25	26	27	28	29	30
r_{D_k}	0.025	-0.018	0.008	0.051	-0.026	0.054	-0.066	0.080	-0.019	0.016
$r_{\log D_k}$	0.047	-0.017	-0.005	0.048	-0.041	0.054	-0.057	0.065	-0.017	0.003

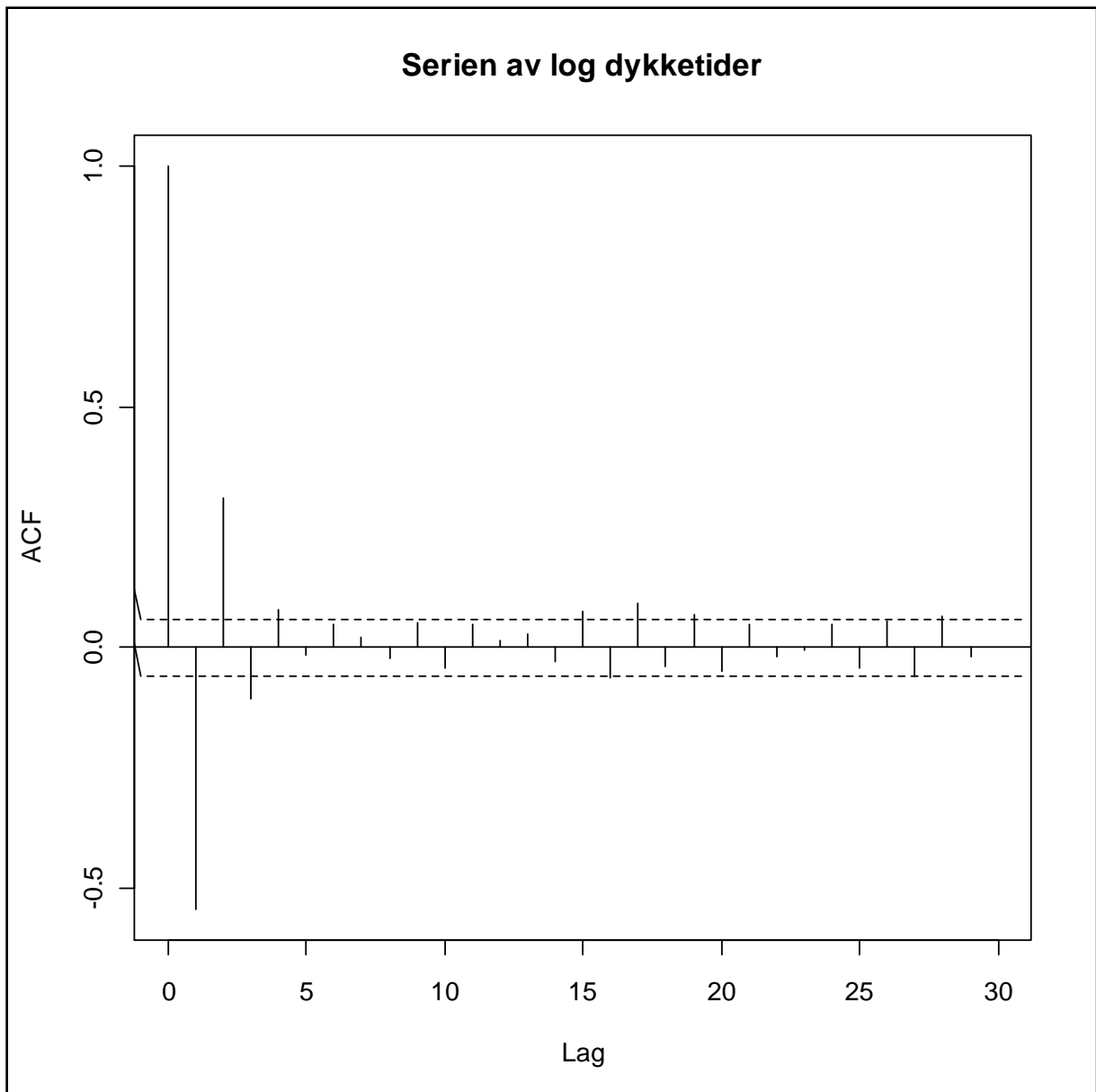
Tabell 11.2 Autokorrelasjonskoeffisientar

for serien av dykktider (r_{D_k}) og for serien av log dykktider ($r_{\log D_k}$).

I figurane 11.3 og 11.4 ser vi autokorrelasjonen til henholdsvis dykketidene og logaritmen til dykketidene. (Autokorrelasjonen for lag 0 er satt til 1.) Konfidensintervallet på $[-0.05, 0.05]$ er markert med stipla linjer. For variablar som ikkje har nokon lineær samanheng ventar vi korrelasjonskoeffisientar nær null. Vi ser imidlertid at det er høg korrelasjonen i dykketidene i lag 1 og den er også relativt høgt i lag 2. Det same gjeld også for korrelasjonen til serien av log dykketider. Dette indikerer at dykketidene slett ikkje er uavhengige av kvarandre.



Figur 11.3 Autokorrelasjon for serien av dykketider



Figur 11.4 Autokorrelasjon for serien av log dykketider.

For at ein Poisson-prosess skal være ei god beskriving for våre dykketidsdata må dykketidene oppfylle kriteria om uavhenghet og eksponentialfordeling. Vi har sett at dette ikkje er tilfellet. Konklusjonen må bli at Poisson-prosessen ikkje passar som modell for våre dykketidsdata.

12 DISKUSJON

Som tidlegare nemnt gir ein Markovmodulert Poisson-prosess (MMPP) rom for klustra data. Med dette utgangspunktet og resultatet frå førre tør eg påstå at ein MMPP vil passe betre som modell for våre dykketidsdata enn ein homogen Poisson-prosess. Vi lar intensiteten λ i Poisson-prosessen variere mellom to nivå: λ_1 (djupdykksfase) og λ_2 (overflatefase) slik at $\lambda_2 > \lambda_1$, og lar ein kontinuerligtids Markovprosess med intensitetar μ_1 og μ_2 bestemme kva for ein tilstand kvalen er i til ei kvar tid.

Skaug (2006) brukar ein MMPP for å modellere vågekvalen si fordeling i havet (sjå kapittel 4), og i den forbindelsen gir han ei beskriving av likelihoodbaserte metodar for å finne estimata av parametraner $\theta = (\lambda_1, \lambda_2, \mu_1, \mu_2)$. Likelihood'en er basert på 1) at det er nøyaktig n blåst i intervallet $[0, L]$ og 2) at blåsta kom i tidspunkta w_1, \dots, w_n . Likelihoodfunksjonen er :

$$Lik(\theta) = \pi \exp[Cv_1] \Lambda \cdots \exp[Cv_n] \Lambda \exp[C(L - w_n)] e,$$

og Skaug anbefalar programvaren AD Modelbuilder som verktøy for å optimalisere denne funksjonen. Dette er framgangsmåten som eg ville nytte. Men den krev mykje tidsbruk, blant anna for å sette seg inn i ny programvare, og eg har dessverre ikkje fått tid til å gjennomføre dette arbeidet.

Datamateriellet som eg har brukt i oppgåva er marginalt. Vi har sett på dykketidene for berre ein vågekval over eit drygt døgn. Det er rimeleg å tenke at dykketidene kan variere både over tid og mellom individ. Schweder og Øien (2007) presenterar data frå i alt ti vågekvalar, og vi ser som venta at det *finst* variasjonar mellom dei. Øien (Vågehval 2009) fortel at radiomerking har gjort det mogeleg å følgje enkeltkval i opptil fem døgn i strekk, og at blåstraten målt over kortare intervall i løpet av ein slik periode varierer like mykje hjå *ein* kval som *mellom* kvalar over tid. Ut frå histogramma i arbeidet til Schweder og Øien (2007) er desse variasjonane likevel ikkje større enn at våre data kan kallast representative, og slike variasjonar kan det korrigerast for i modellen.

Er ein Markovmodulert Poisson-prosess ein god, evt den beste modellen, for dykketidsdata? For å finne svara treng vi eit godt estimat av parametraner våre, og vi treng ei grundig samanlikning av denne modellen i forhold til modellane til Schweder (2003) og Schweder og Øien (2007) (sjå kap 6). Av årsak som tidlegare nemnt har vi ikkje desse resultatane. Det må difor meir arbeid til før vi kan gje ein konklusjon på om Markovmodulerte Poisson-prosessar er gode modellar for dykketidsdata.

FIGURAR

Figur 1.1 Vågekval	4
Figur 2.1 Skjematisk teikning av eit linjetransekt med ein observatør eller eit team.	6
Figur 3.1 Funksjonen plotTDR() i R-pakken diveMove	7
Figur 5.1 Dykkeprofilar	12
Figur 7.1 Suksessivt påfølgande dykk	13
Figur 7.2 Suksessivt påfølgjande dykk.....	14
Figur 7.3 Dykketid, frekvens.	15
Figur 7.4 Dykkemønster vist som ei serie dykk.	16
Figur 7.5 Dykketid dei siste 2 timar 44 minutt og 13 sekund av observasjonstida.	16
Figur 7.6 Dykkemønster vist som ei serie blåst.	17
Figur 9.1 Døme på Poisson punktprosess.....	18
Figur 11.1 Dykketid, fordeling.....	21
Figur 11.2 Fordeling av logaritmen til dykketidene.....	23
Figur 11.3 Autokorrelasjon for serien av dykketider	25
Figur 11.4 Autokorrelasjon for serien av log dykketider.	26

TABELLAR

Tabell 11.1 Observerte og forventa frekvensar av dykketider.	22
Tabell 11.2 Autokorrelasjonskoeffisientar.....	24

LITTERATUR

Asmussen, Søren. «Matrix-analytic Models and their analysis.» *Scandinavian Journal of Statistics* 27, nr. 2 (2000): 193-226.

Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, og L. Thomas. *Advanced Distance Sampling - Estimating abundance of biological population*. New York: Oxford University Press Inc., 2004.

Elliott, Kyle Hamish, Gail K. Davoren, og Anthony J. Gaston. «Increasing energy expenditure for a deep-diving bird alters time allocation during the dive cycle.» *Animal Behaviour* 75 (2008): 1311-1317.

Fowler, Shannon L., Daniel P. Costa, John P. Y. Arnould, Nicholas J. Gales, og Carey E. Kuhn. «Ontogeny of diving behavior in the Australian sea lion: trials of adolescence in a late bloomer.» *Journal of Animal Ecology* 75 (2006): 358-367.

Hedley, Sharon L., og Stephen T. Buckland. «Spatial Models for Line Transect Sampling.» *Journal of Agricultural, Biological and Environmental Statistics* 9, nr. 2 (2004): 181-199.

Heide-Jørgensen, Mads Peter, og Malene Simon. «Cue rates for common minke, fin and humpback whales in west Greenland.» 2007.

López, Bruno Días, Julia Andrea Bernal Shirai, Alberto Bilbao Prieto, og Paula Méndez Fernández. «Diving activity of a solitary wild free ranging bottlenose dolphin (*Tursiops truncatus*).» *Journal of the Marine Biological Association of the United Kingdom* 88, nr. 6 (2008): 1153-1157.

Mori, Yoshihisa, Akinori Takahashi, Fridtjof Mehlum, og Yutaka Watanuki. «An application of optimal diving models to diving behavior of Brünnich's guillemots.» *Animal Behaviour* 64 (2002): 739-745.

Raftery, Adrian E., og Tore Schweder. «Inference About the Ratio of Two Parameters, With Application to Whale Censusing.» *The American Statistician* 47, nr. 4 (1993): 259-264.

Schreer, Jason F., Kit M. Kovacs, og R. J. O'Hara Hines. «Comparative Diving Patterns of Pinnipeds and Seabirds.» *Ecological Monographs* 71, nr. 1 (Februar 2001): 137-162.

Schweder, Tore. «Point process models in line transect studies.» 2003.

Schweder, Tore, Hans J. Skaug, Mette Langaas, og Xenia K. Dimakos. «Simulated Likelihood Methods for Complex Double-Platform Line Transect Surveys.» *Biometrics* 55 (1999): 678-686.

Schweder, Tore, og Nils Øien. «On the surfacing rate in minke whales in the Northeastern Atlantic.» Vitenskapelig rapport IWC/SC/59/RMP3, 2007.

Skaug, Hans J. «Markov modulated Poisson processes for clustered line transect data.» *Environmental and Ecological Statistics* 13, nr. 2 (2006): 199-211.

Skaug, Hans J., og Tore Schweder. «Hazard Models for Line Transect Surveys with Independent Observers.» *Biometrics* 55 (1999): 29-36.

«Stortingsmelding nr. 12.» *Om rent og rikt hav*. 2001/02.

«Stortingsmelding nr. 27.» *Norsk sjøpattedyrpolitikk*. 2003/04.

Taylor, Howard M., og Samuel Karlin. *An Introduction To Stochastic Modeling*. 3. 1998.

Tinker, M. T., D. P. Costa, J. A. Estes, og N. Wieringa. «Individual dietary specialization and dive behavior in the California sea otter: Using archival time-depth data to detect alternative foraging strategies.» *Deep-Sea Research II* 54 (2007): 330-342.

Watwood, Stephanie L., Patrick J. O. Miller, Mark Johnson, Peter T. Madsen, og Peter L. Tyack. «Deep-diving foraging behavior of sperm whales (*Physeter macrocephalus*).» *Journal of Animal Ecology* 75 (2006): 814-825.

Williams, Rob, Russel Leaper, Alexandre N. Zerbini, og Phillip S. Hammond. «Methods for investigating measurement error in cetacean line-transect surveys.» *Journal of the Marine Biological Association of the United Kingdom* 87 (2007): 313-320.

Williams, Rob, Sharon L. Hedley, og Phillip S. Hammond. «Modeling Distribution and Abundance of Antarctic Baleen Whales Using Ships of Opportunity.» *Ecology and Society* 11, nr. 1 (2006).

Waagepetersen, Rasmus, og Tore Schweder. «Likelihood-based inference for clustered line transect data.» Vitenskapelig rapport R-2004-29, Aalborg universitet, 2004.

Øien, Nils. «Vågehval.» *Havets ressurser og miljø*, 2009: 40-42.

Øien, Nils, Gjermund Bøthun, og Lars Kleivane. «Update on available data on surfacing rates of Northeastern Atlantic minke whales.» Vitenskapelig rapport IWC/SC/59/RMP6, 2007.