Prosjektoppgaver om diffusjonsprosesser og diffusjonstilnærmelse

February 22, 2007

I alle oppgavene skal det skrives litt om hva diffusjonsprosesser er, hvilke spesielle resultater fra diffusjonsteorien man skal se nærmere på i oppgaven, og hva som menes med diffusjonstilnærmelse. De teoretiske resultatene for diffusjoner som benyttes skal demonstreres ved å simulere diffusjonsprosesser. Prosessene som skal tilnærmes med diffusjoner skal simuleres med passende utvalg av parameterverdier, og egenskapene til prosessene skal sammenliknes med de resultatene man finner ved å benytte diffusjonstilnærmelsen. Prøv å finne ut hva som skal til (partameterverdier) for at diffusjonstilnærmelsen blir god eller dårlig. Det blir lagt vekt på god pedagodisk framstilling med utstrakt bruk av grafer.

A Diffusjonstilnærmelse til fødels- og dødsprosesser uten tetthetsregulering

La N betegne antall individer i en populasjon. Hvert individ bidrar med et antall individer til neste generasjon og disse bidragene antas å være uavhengige. Et individ overlever til neste generasjon med sannsynlighet p og gir dermed et overlevelsesbidragbidrag I som blir 0 eller 1. I tillegg bidrar individet med X avkom til neste generasjon. Den variable X har en gitt diskret fordeling og generelt kan I og X være avhengige variable (f. eks. kan X være lik null hvis I er null). En mulig modell er at X er poissonfordelt og uavhengig av I, men det er interessant også å se på andre modeller.

Vi skal studere hvordan antall individer endrer seg over tid og hvor lang tid det tar før populasjonen dør ut. En slik prosess kan tilnærmes med en diffusjon med infinitesimal forventning og varians som begge er proposjonale med populasjonsstørrelsen N.

Veiledning: Denne oppgaven kan gjøres av to studenter. Begge gir diffusjonstilnærmelse på formen Diff(aN, bN). Parameteren a bør i utgangspunktet velges litt mindre enn 0 (f. eks. $a = -0.01$ og $N_0 = 1000$) slik at prosessen helt sikkert dør ut. Men man kan også gjøre kjøringer med $a > 0$.

1. Individenes bidrag er $W = I + X$ der I er indikator for overlevelse og X er antall avkom. La X være uavhengig av I med fordeling som er en Poissonblanding. Vis studenten at Poissonblandinger oppfyller $E(X) = EE(X|\lambda) = E(\lambda)$ og $\text{var}(X) = E(\text{var}(X)|\lambda) + \text{var}(E(X|\lambda)) = E(\lambda) + \text{var}(\lambda)$. Fordelingen til λ kan f. eks. være gamma eller lognormal.

2. Anta at avkommet dør hvis moren dør, dvs. $X = 0$ hvis $I = 0$. Hvis moren overlever har X f. eks. Poissonfordeling (eller Poissonblanding) med forventning μ og varians σ^2. Vi kan da uttrykke W som $J(1+X)$ der J og X nå er uavhengige, og J har samme fordeling som I. Hvis vi innfører $p = P(I = 1)$ finner vi $EW = EJE(1 + X) = p(1 + \mu)$ og $E(W^2) = EJ^2E(1 + X)^2 = p((1 + \mu)^2 + \sigma^2)$ som gir $\text{var}(W) = p(1 - p)(1 + \mu)^2 + p\sigma^2$.

B Multiplikative populasjonsmodeller uten tetthetsregulering

For store populasjoner kan ofte følgende modell være realistisk: Populasjonsstørrelsen neste år er populasjonsstørrelsen forrige år multiplisert med en stokastisk faktor. Verdiene til denne faktoren fra år til år er uavhengige med samme fordeling. Populasjonen dør ut hvis populasjonsstørrelsen et år blir midre eller lik 1.

Vi skal her studere tiden det tar til populasjonen dør ut. Dette skal gjøres både ved å se på diffusjonstilnærmelsen for populasjonsstørrelsen og logaritmen til denne. Den stokastiske faktoren kan velges log-normal fordelt slik at logarimen til faktoren blir normalfordelt men andre fordelinger kan også undersøkes. .

Veiledning: Den stokastisk faktoren Λ_t må være en positiv stokastisk variabel, f.eks. gamma, lognormal eller rektanguler. Den stokastiske vekstraten $s = E\Lambda_t - 1 - \frac{1}{2}\text{var}(\Lambda_t)$ bør i utgangspunktet velges litt mindre enn null slik at utdøing er sikret, men man kan også gjøre kjøringer med positiv stokastisk vekstrate.

C Diffusjon av partikler, diskret tid

To kammer er adskilt med en vegg men en liten åpning. Tilsammen befinner det seg N partikler i de to kamrene, X i det ene og $N - X$ i det andre. Hvert sekund vil en vilkårlig partikkel (en trukket tilfeldig blant alle N partikler) flytte seg over i motsatt kammer. Vi skal studere hvordan X varierer over tid.

Generelt vil $\text{var}(\Delta X|X = x)$ være avhengig av x, men for store verdier av N vil X ligge nær $N/2$ og vi kan da tilnærme $\text{var}(\Delta X|X = x)$ med $\text{var}(\Delta X|X = N/2)$. Prosessen vil da kunne tilnærmes med en Ornstein-Uhlenbeck process. Egenskaper ved Ornstein-Uhlenbeck prosessen skal først demonstreres ved stokastiske simuleringer. Videre skal man studere diffusjonstilnærmelsen til prosessen X basert på Ornstein-Uhlenbeck prosessen for forskjellige verdier av N.

D Diffusjon av partikler, kontinuerlig tid

Vi betrakter samme modell som skissert i oppgave C og skal studere de samme problemene, men antar nå at hver partikkel har sannsynlighet λdt for å forflytte seg over i motsatt kammer i tidsintervallet $(t, t + dt)$. Vi antar videre at individene beveger seg uavhengig av hverandre. Vi kan også her for store N tilnærme variansen med verdien vi finner ved å sette $X = N/2$. Prosessen kan da tilnærmes med en Ornstein-Uhlenbeck prosess.

Veiledning: Forklar studenten at $P($ overgang i $(t, t + dt)) = N\lambda dt$. Gitt overgang i $(t, t + dt)$ har alle partikler samme sannsynlighet $1/N$ for overgang. Forklar hvordan fordelingen til dX da kan skrives opp og at dette gir oss diffusjonstilnærmelsen.

E Gompertz type tetthetsregulering

Vi betrakter den multiplikative populasjonsmodellen $N_{t+1} = \Lambda_t N_t$ og antar at den stokastisk faktoren kan skrives på formen $\Lambda_t = a_t - b \ln N_t$ der a_t er en sekvens av uavhengige identisk fordelte variable. Diffusjonstilnærmelsen blir da en Gomperz type modelle. Velg passende parametre og fordeling til a_t (f. eks. normalfordeling). Gi en presentasjon av egenskapene til denne diffusjonen og studer hvor god diffusjonstilnærmelsen er for forskjellige valg av parametre.

Veiledning: Man kan f. eks. velge $a_t \sim N(\mu, \sigma^2)$ der $\sigma^2 = 0.01$ og $K = \exp((\mu - 1)/b) = 1000$.

F Logistisk modell

Vi betrakter samme modell som i oppgave E, men setter $\Lambda_t = a_t - bN_t$. Diffusjonstilnærmelsen er da en logistisk type modell og stasjonærfordelingen er gammafordelingen. Demonstrer at dette er riktig ved hjelp av stokastiske simuleringer og studer om stasjonærfordelingen vi finner fra diffusjonstilnærmelsen er en god tiln ærmelse for den diskrete modellen for forskjellige valg av parametre og fordeling for a_t.

Veiledning: Man kan f. eks. velge $a_t \sim N(\mu, \sigma^2)$ der $\sigma^2 = 0.01$, $1.01 < \mu < 1.1$ og $K = (\mu - 1)/b = 1000$.

G Beverton-Holt type modell

Anta at individene i en populasjon (vi betrakter bare hunnene) av størrelse N i gjennomsnitt produserer $Nf_0/(1 + \beta N)$ avkom som overlever til neste generasjon og at de voksne individene overlever med sannsynlighet s. Finn $E(\Delta N|N = n)$ for denne modellen. Anta videre at variansen er $\mathrm{var}(\Delta N|N = n) = \sigma^2 n^2$. Vi at difusjonstilnærmelsen bli samme modell som Beverton-Holt modellen beskrevet i teksten. Undersøk ved simuleringer om det teoretiske uttrykket for stasjonærfordelingen er en brukbar tilnærmelse og studer hvordan forskjellige parametrene påvirker fordelingen.

Veiledning: Overlevelsen s bør velges rundt 0.7-0.95, miljøvariansen σ^2 rundt 0.01, og $K = (f_o/s - 1)/\beta$ av størrelsesorden 1000. Forskjellige valg av f_o (og samsvarende verdi for β) vil bestemme hvor fort prosessen vender tilbake til likevekten. Man kan her se på den deriverte av $f_o/(1 + \beta N) - s$ i $N = K$ som vil være den inverse av returtiden.

H Ornstein-Uhlenbeck prosess som lineær tilnærmelse

Modellen er av samme type som E og F, med $\Lambda_t = a_t - g(N_t)$ der $g(N_t)$ er voksende i N_t. Velg denne funksjonen og studer hvordan det fungerer å tilnærme denne med en lineær funksjon i $\ln N_t$ funnet ved å å bruke første leddet i rekkeutviklingen rundt bæerekapasiteten (den verdien som gir $E\Lambda = 1$). Den lineariserte modellen kan nå tilnærmes med en Ornstein-Uhlenbeck prosess for $\ln N_t$.

Veiledning: Forklar studenten hvordan Λ lineariseres nær $N_0 = K$.

I Gompertz type prosess som lineær til nærmelse

Samme oppgave som H, men lineariseringen skal gjøres i N_t istedet for $\ln N_t$. Modellen kan da tilnærmes med en logistisk type modell.

Veiledning: Skriv $g(N)$ som $h(X)$ der $X = \ln N$ og forklar hvordan Λ lineariseres nær $X_0 = \ln K$.

J Mutasjonsload

Fra populasjonsgenetikken vet vi at reduksjonen i gjennomsnittlig relativ fitness som følge av enveis mutasjon (mutasjonsload) og for recessive mutasjoner er uavhengig av seleksjonskoeffisienten s og lik mutasjonsrate u. Dette forutsetter at populasjonstørrelsen er uendelig. Hvis populasjonsstørrelsen er endelig vil ikke genfrekvensen gå mot noen deterministisk likevekt—genfrekvensen vil i stedet gå mot en stasjonærfordeling dersom vi også forutsetter en viss tilbakemutasjonsrate v. Velg passende parametere for u, v, s, N og undersøk hvor stor forventet mutasjonsload blir i forhold til mutasjonsload i den enklere deterministiske modellen ved hjelp av simuleringer.

K Mutasjonsload

Samme som oppgave J men bruk i stedet modell med delvis dominans.

L Stasjonærfordeling i øymodell med genetisk drift

I den såkalte øymodellen kan det vises at stasjonærfordelingen på hver enkelt øy vil være beta-fordelt om vi tilnærmer prosessen med en diffusjon. Dersom immigrasjonsratene er m, genfrekvensen blant immigranter er p_c, og den lokale populasjonsstørrelsen er N_e er parameterne i denne beta-fordeling $\alpha = 4N_e m p_c$ og $\beta = 4N_e m(1 - p_c)$. Undersøk hvor god denne tilnærmingen er for en mer realistisk modell med diskrete generasjoner og genfrekvenser.