Betinget sannsynlighet gitt en nullhendelse?

Gunnar Taraldsen

17. november 1997

Sammendrag

Vi drøfter definisjonen og beregningen av betingede sannsynligheter motivert av et innledende Borel-paradoks. Spesielt utledes en formel (*) som generelt gir den betingede sannsynligheten gitt verdien til en variabel som har kontinuerlig fordeling. Vi mener denne formelen burde være mer kjent enn det den er. Som et eksempel beregnes den betingede forventningsverdien til variabelen $\phi(X_1)$ gitt observatoren X_{maks} tilsvarende et tilfeldig utvalg $X_1, \dots, X_n$ fra en uniform fordeling. Her er ϕ en generell funksjon. Vi gir i tillegg to andre mulige beregningsmetoder. Den ene metoden er gitt ved å skifte variable slik at den elementære metoden gitt ved en simultantetthet kan benyttes. Den andre metoden gir den betingede sannsynligheten som en grenseverdi tilsvarende hendelser som krymper ned til nullhendelsen. Formelen (*) gir et bevis for gyldigheten av denne metoden. Leserne oppfordres til å presentere alternative beregningsmåter. Spesielt er beregningsmetoder for betingede forventningsverdier gitt en tilstrekkelig observator av stor praktisk interesse.

1 Et Borel-paradoks

La (X, Y) være koordinatene til et tilfeldig valgt punkt i enhetskvadratet $(0,1)^2$. Da gjelder

$$P(X < 1/2, Y < 1/2 \,|\, X - Y = 0) = 1/2 \quad \text{og} \quad P(X < 1/2, Y < 1/2 \,|\, X/Y = 1) = 1/4.$$

Dette ser paradoksalt ut fordi hendelsen $(X - Y = 0)$ er lik hendelsen $(X/Y = 1)$. Muligheten for et slikt Borel-paradokset skyldes at hendelsen $(X = Y)$ det betinges med hensyn på er en nullhendelse, dvs $P(X = Y) = 0$. Paradokset krever en nærmere forklaring. I første omgang kan vi beregne de to betingede sannsynlighetene hver for seg ved passende variabelskifte. Denne metoden begrunner vi i siste hovedavsnitt i dette notatet.

La f være tettheten til fordelingen til det tilfeldige punktet. Da er f lik 1 i enhetskvadratet og lik 0 utenfor. Innfør en ny variabel $U = X - Y$. Absoluttverdien til Jacobi-determinanten ved skifte fra koordinatene (x, y) til koordinatene (x, u) er lik 1. Dette gir oss simultantettheten $f_{X,U}$ til X, U og

$$P(X < 1/2, Y < 1/2 \,|\, X - Y = 0) = \int_{x=0}^{1/2} f_{X\,|\,U}(x\,|\,0)dx = \int_{x=0}^{1/2} \frac{f_{X,U}(x,0)}{f_U(0)}dx = \int_{x=0}^{1/2} 1dx = 1/2.$$

Den andre betingede sannsynligheten beregnes enklest ved overgang til polarkoordinater. Arealelementet $dxdy$ erstattes da av $rdrd\theta$. Vi finner

$$P(X < 1/2, Y < 1/2 \,|\, X/Y = 1) = \int_{r=0}^{\sqrt{2}/2} f_{R\,|\,\Theta}(r\,|\,45^\circ)dr = \int_{x=0}^{\sqrt{2}/2} rdr = 1/2(\sqrt{2}/2)^2 = 1/4.$$

Alternativt kan beregningen gjøres ved å innføre variabelen $V = X/Y$, og ved skifte fra (x, y) koordinater til (x, v) koordinater. Beregningen ved polarkoordinater krever strengt tatt litt mer argumentasjon. Grunnen til at beregningen er korrekt er gitt ved at nivåkurvene $x/y =$ konstant er lik nivåkurvene $\theta =$ konstant. Disse nivåkurvene er ikke annet enn rette linjer gjennom origo. Nivåkurvene $x - y =$ konstant er rette linjer med stigningstall 1. Denne forskjellen er årsaken til det tilsynelatende paradokset. For å begrunne dette må vi være mer presise.

2 Definisjon og eksistens av betinget forventning

Når $P(B) = 0$, så kan ikke $P(A \mid B)$ defineres med utgangspunkt i $P(A \cap B) = P(A \mid B)P(B)$. Borel-paradokset vi startet med gir at med konvensjonelle definisjoner av betinget sannsynlighet så forblir $P(B \mid A)$ udefinert i tilfellet $P(A) = 0$. Vi vil allikevel gi mening til beregningene vi startet med. Løsningen er gitt ved å innføre den betingede forventningen før den betingede fordelingen. Dette krever litt mer teori som vi skal se på.

Når A er en hendelse, så er indikatorvariabelen $[A]$ definert ved

$$[A](\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \in A^c \end{cases}$$

Når X er en tilfeldig variabel, så er produktet $X\ [A]$ også en tilfeldig variabel og vi kan definere notasjonen $\mathrm{E}(X; A)$ ved $\mathrm{E}(X; A) = \mathrm{E}(X\ \ [A])$. Her må vi anta at forventningsverdien $\mathrm{E}\, X$ er endelig. Leseren kan selv overbevise seg om at funksjonen Q definert ved $Q(A) = \mathrm{E}(X; A)$ er en ikke-normalisert fordeling med fortegn, dvs Q er et fortegnsmål.

La P være en fordeling. Dersom Q er et fortegnsmål slik at

$$P(A) = 0 \text{ gir } Q(A) = 0,$$

så sies Q å være absolutt kontinuerlig med hensyn på P. Radon-Nikodym teoremet (Doob, 1994) sier at i så fall finnes det en tilfeldig variabel X slik at

$$Q(B) = \mathrm{E}(X; B).$$

Her spiller variabelen X rollen som en tetthet for Q i forhold til P. La X være gitt. Det er da lett å vise at forventningsverdien $\mathrm{E}(X; B)$ definerer et fortegnsmål som er absolutt kontinuerlig med hensyn på P. Radon-Nikodym teoremet sier at *alle* absolutt kontinuerlige fortegnsmål fremkommer på denne måten. Vi bemerker at det finnes flere variable enn X som gjør nytten. Enhver Y som oppfyller $P(X = Y) = 1$ er like god.

La X være en tilfeldig variabel og la T være en observator. Et typisk eksempel er gitt ved at T er en vektor av tilfeldige variable, dvs $T = (Y_1, Y_2, \ldots, Y_n)$, men det finnes mange andre muligheter. En annen viktig mulighet er gitt ved $T(\omega) = \omega$ hvor utfallsrommet Ω_T til T er lik utfallsrommet Ω, men hvor Ω_T er utstyrt med en redusert samling av hendelser (en del-σ-algebra). Vi vil definere den betingede forventningen

$$\mathrm{E}(X \mid T = t).$$

Her kan vi ha $P(T = t) = 0$, men vi kan til og med ha tilfeller hvor $(T = t)$ ikke er en hendelse! Først bemerker vi at indikatorvariabelen $[T \in A]$ er gitt ved

$$[T \in A]\ (\omega) = [\{\omega \mid T(\omega) \in A\}](\omega) = \begin{cases} 1 & T(\omega) \in A \\ 0 & T(\omega) \in A^c \end{cases}$$

Vi overlater til leseren å overbevise seg om at vi også har $[T \in A] = [A](T)$ hvor $[A](T)(\omega) = [A](T(\omega))$. Vi kan nå definere hva som menes med en betinget forventning i det generelle tilfellet.

En betinget forventning $\mathrm{E}(X \mid T)$ *er en tilfeldig variabel som er en funksjon av* T*, dvs* $\mathrm{E}(X \mid T) = \phi(T)$*, og som oppfyller*

$$\boxed{\mathrm{E}(X; T \in A) = \mathrm{E}(\mathrm{E}(X \mid T); T \in A)}$$

for alle hendelser A *i utfallsrommet til* T*. Videre defineres notasjonen* $\mathrm{E}(X \mid T = t) = \phi(t)$*.*

Denne definisjonen (Rao, 1973, s.96) gir samme resultat som den mer konvensjonelle definisjonen (Lehmann, 1986, s.44) (Doob, 1953, s.18). Konvensjonelt defineres først en betinget forventning med hensyn på en redusert familie av hendelser i utfallsrommet Ω. Vi har valgt definisjonen over fordi den leder raskere til resultater. Det er kanskje også noen som foretrekker vårt valg fordi dette valget skyver σ-algebraene så langt inn i skyggen som mulig.

Radon-Nikodym teoremet gir et kort bevis av at den betingede forventningen finnes: Definer et fortegnsmål Q for utfallsrommet til T ved

$$Q(A) = \mathrm{E}(X; T \in A).$$

Det følger at Q er absolutt kontinuerlig med hensyn på fordelingen P_T til T fordi $P_T(A) = P(T \in A) = 0$ gir $Q(A) = \mathrm{E}(X; T \in A) = 0$. Dermed gir Radon-Nikodym teoremet at det finnes en tetthet ϕ slik at

$$Q(A) = \mathrm{E}_T(\phi; A) = \mathrm{E}(\phi(T); T \in A).$$

Den siste likheten er en konsekvens av det generelle variabelskifteteoremet

$$\mathrm{E}_T\, \psi = \mathrm{E}\, \psi(T)$$

brukt på tilfellet $\psi(t) = \phi(t)\ [A](t)$. Vi har bevist at

$$\mathrm{E}(X; T \in A) = Q(A) = \mathrm{E}(\phi(T); T \in A),$$

som gir

$$E(X \,|\, T) = \phi(T),\ E(X \,|\, T = t) = \phi(t).$$

Dermed har vi bevist at en betinget forventning $\mathrm{E}(X \,|\, T = t)$ finnes. En betinget forventning er ikke entydig fordi enhver $\tilde{\phi}$ som oppfyller $P_T(\tilde{\phi} = \phi) = 1$ også gir en betinget forventning. Fordelingen til $\mathrm{E}(X \,|\, T)$ er imidlertid entydig.

3 Beregning av betingede fordelinger

Et problem så langt er at det ikke uten videre er klart hvordan en betinget forventning kan beregnes. Situasjonen ligner på den en møter i teorien for differensialligninger. Én ting er å vise at en løsning finnes, en helt annen ting er å finne løsningen. Vi begynner med å gi en metode som synes å være lite kjent, dvs vi kjenner ingen referanse til denne metoden.

Utgangspunktet for metoden er den generelle sammenhengen

$$\mathrm{E}(X; T \leq t) = \int_{-\infty}^{t} \mathrm{E}(X \,|\, T = s) P_T(ds),$$

som følger ved inspeksjon av ligningene i foregående hovedavsnitt med $A = (-\infty, t]$. Anta at T er en tilfeldig variabel med en tetthet f. Derivasjon av likheten

$$\mathrm{E}(X; T \leq t) = \int_{-\infty}^{t} \mathrm{E}(X \,|\, T = s) f(s)\ ds$$

gir formelen

$$\mathrm{E}(X \,|\, T = t)\ f(t) = \frac{d}{dt} \mathrm{E}(X; T \leq t),$$

som kan benyttes til beregning av $\mathrm{E}(X \,|\, T = t)$. I det overstående er X en generell tilfeldig variabel så vi kan erstatte X med indikatorvariabelen $[X \in A]$. Da finner vi formelen

$$\boxed{(*) \qquad P(X \in A \,|\, T = t)\ f(t) = \frac{d}{dt} P(X \in A, T \leq t)}$$

Utledningen kan enkelt generaliseres til observatorer T som er flerdimensjonale og resultatet er likt, men derivasjonen erstattes av derivasjon mhp hver av komponentene til t. Et viktig poeng er at metoden ikke forutsetter at det finnes en simultantetthet for X, T. Et annet poeng er at utledningen viser at funksjonen vi skal derivere automatisk er deriverbar. Dette betyr desverre ikke at den deriverte eksisterer for alle t, men snarere for alle t i en mengde D som oppfyller $P(T \in D) = 1$. Spesielt eksisterer den deriverte for alle t dersom $\mathrm{E}(X \,|\, T = t) f(t)$ er kontinuerlig som funksjon av t. Formelen (*) kan omskrives til den intuitivt tiltalende formelen:

$$P(X \in A \,|\, T = t)\; f(t) = f(t \,|\, X \in A)\; P(X \in A).$$

Vi kan gå tilbake til det helt generelle tilfellet, dvs utfallsrommet Ω_T til T er ikke nødvendigvis lik mengden av reelle tall. Utgangspunktet for metoden gir også

$$E(X; T = t) = E(X \,|\, T = t) P(T = t).$$

Her må vi anta at $(T = t)$ er en hendelse. Med $X = [A]$ og $B = (T = t)$ finnes den velkjente

$$P(A \cap B) = P(A \,|\, B) P(B).$$

Vi ser på noen eksempler på bruk av beregningsformelen (*) for tilfellet hvor $P(T = t) = 0$. La $X_1, \ldots, X_n$ være uavhengige og uniformt fordelt på intervallet $(0, \theta)$. Vi vil beregne

$$\mathrm{E}(2X_1 \,|\, X_{\mathrm{maks}})$$

som en illustrasjon av metoden over. Metoden en møter i de fleste lærebøker er ikke umiddelbart anvendelig fordi det ikke finnes noen simultantetthet for X_1, X_{maks}, eller X, X_{maks}. La $0 < t < \theta$. Derivasjon av

$$P(X_{\mathrm{maks}} \leq t) = (t/\theta)^n$$

gir tettheten

$$f(t) = nt^{n-1}/\theta^n \qquad 0 < t < \theta$$

til X_{maks}. Integrasjon mhp simultantettheten til $X_1, \ldots, X_n$ gir

$$\mathrm{E}(2X_1; X_{\mathrm{maks}} \leq t) = (t/\theta)^{n-1} \int_0^t 2s/\theta \; ds = t^{n+1}/\theta^n.$$

Derivasjon og divisjon med f gir $\mathrm{E}(2X_1 \,|\, X_{\mathrm{maks}} = t) = \frac{n+1}{n} t$ og

$$\mathrm{E}(2X_1 \,|\, X_{\mathrm{maks}}) = \frac{n+1}{n} X_{\mathrm{maks}}.$$

La en statistisk modell være gitt ved en familie $\{P^\theta\}$, $\theta \in \Theta$, av fordelinger for et utfallsrom Ω. En observator T (for Ω) er en funksjon $T : \Omega \to \Omega_T$ slik at $(T \in A) = \{\omega \,|\, T(\omega) \in A\}$ er en hendelse (i Ω) for enhver hendelse A i utfallsrommet Ω_T til T. Dersom $\Omega_T = \mathbb{R}$ med den konvensjonelle hendelsesfamilien, så er T en tilfeldig variabel. Her er det viktig å merke seg at parameteren θ ikke inngår på noen måte i definisjonen av en observator. Fordelingen til T vil imidlertid avhenge av parameteren θ ved at

$$P_T^\theta(A) = P^\theta(T \in A).$$

Et eksempel på en statistisk modell er gitt ved å la utfallsrommet være $\Omega = \mathbb{R}^n$, ved å la parameterrommet være $\Theta = \mathbb{R}$, og ved å la fordelingen P^θ være gitt ved en tetthet som er lik $1/\theta^n$ i området $(0, \theta)^n$ og lik 0 ellers. Ved å definere $X_j(\omega_1, \ldots, \omega_n) = \omega_j$ følger det at vi er tilbake i samme situasjon som i vårt regneeksempel.

Når vi har gitt en statistisk modell så vil en betinget forventning generelt være på formen

$$\mathrm{E}^\theta(X \mid T)(\omega) = \psi(\theta, T(\omega)),$$

dvs en betinget forventning er en funksjon $\phi : \Theta \times \Omega \to \mathbb{R}$. Dermed er en betinget forventning strengt tatt ikke en tilfeldig variabel pga avhengigheten av parameteren. I vårt regneeksempel fant vi imidlertid at den betingede forventningen

$$\mathrm{E}^\theta(2X_1 \mid X_{\mathrm{maks}}) = \frac{n+1}{n} X_{\mathrm{maks}},$$

dvs i dette tilfellet er den betingede forventningen en tilfeldig variabel. Dette skyldes at observatoren X_{maks} er tilstrekkelig.

Vanligvis defineres observatoren T til å være tilstrekkelig (for parameteren θ i forhold til observatoren X) dersom den betingede fordelingen til X gitt T ikke avhenger av parameteren. Her kan det bemerkes at pga komplikasjoner i forbindelse med eksistensen av betingede fordelinger så synes denne definisjonen å være lite hensiktsmessig. *Vi definerer tilstrekkelighet ved å kreve at $\mathrm{E}^\theta(\phi(X) \mid T)$ er en tilfeldig variabel for enhver ϕ som er en tilfeldig variabel for utfallsrommet til X.* Vi krever altså at det finnes en funksjon $\psi : \Omega_T \to \mathbb{R}$ slik at

$$\mathrm{E}^\theta(\phi(X) \mid T) = \psi(T).$$

En grunn til dette valget er at det finnes betingede forventninger som ikke er gitt ved integrasjon med hensyn på en betinget fordeling (Doob, 1953, s.27) (Lehmann, 1986, s.48), dvs det er ikke alltid gitt at en betinget fordeling finnes. En annen grunn til dette valget er gitt ved at en estimator betinget med hensyn på en tilstrekkelig observator gir en estimator.

I vårt eksempel finnes imidlertid en betinget fordeling. Vi kan bevise at X_{maks} er tilstrekkelig for θ i forhold til X_1 ved å beregne fordelingsfunksjonen til X_1 gitt X_{maks}. I beregningen over erstattes $2X_1$ av indikatorvariabelen $[X_1 \leq x]$. For $0 < x < t$ finner vi

$$P(X_1 \leq x, T \leq t) = (t/\theta)^{n-1} \int_0^x 1/\theta \, ds = \frac{t^{n-1} x}{\theta^n}$$

$$P(X_1 \leq x \mid T = t) = \frac{(n-1)t^{n-2} x}{\theta^n} \cdot \frac{\theta^n}{n t^{n-1}} = \frac{n-1}{n} \frac{x}{t},$$

I tillegg har vi

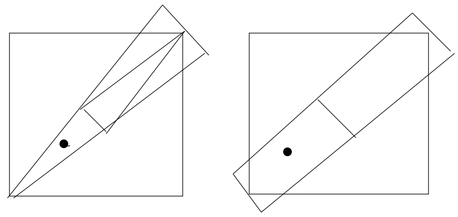
$$P(X_1 \leq t \mid T = t) = 1,$$

som er intuitivt meget rimelig, men dette kan også beregnes slik som over. Konklusjonen er at fordelingen til X_1 gitt $X_{\mathrm{maks}} = t$ er en blanding av diskret og uniform fordi fordelingsfunksjonen har et sprang for $x = t$. Spesielt er sannsynligheten for $X_1 = t$ lik $1/n$. Dette resultatet kan også finnes ved andre argumenter (Lehmann, 1983, s.81). Fordelingsfunksjonen gir

$$\mathrm{E}^\theta(\phi(X_1) \mid X_{\mathrm{maks}}) = \frac{1}{n}\phi(X_{\mathrm{maks}}) + \frac{n-1}{n} \frac{1}{X_{\mathrm{maks}}} \int_0^{X_{\mathrm{maks}}} \phi(x) dx,$$

som beviser tilstrekkeligheten. Med $\phi(x_1) = 2x_1$ gjenvinnes det tidligere resultatet. Andre valg av ϕ gir forventningsrette estimatorer basert på X_{maks} for f eks variansen eller standardavviket til den uniforme fordelingen.

I det foregående er det verdt å merke seg at X_{maks} ikke er en funksjon av X_1 slik at faktoriseringsteoremet (Lehmann, 1986, s.55) ikke er direkte tilgjengelig. Løsningen ved faktoriseringsteoremet er gitt ved at dette kan brukes til å vise at X_{maks} er tilstrekkelig for $X_1, \ldots, X_n$ og derved for X_1.



Figur 1: Relativt areal er hhv 1/4 og 1/2.

4 To vanlige metoder

Intuitivt er det rimelig å gjette (Larsen and Marx, 1986, s.149)(Feller, 1971, s.157) på at

$$\boxed{P(X \in A \,|\, T = t) = \lim_{h \to 0+} P(X \in A \,|\, t \leq T \leq t + h).}$$

Et bevis av at denne formelen gjelder helt generelt finnes indirekte i (Doob, 1994, s.158-160), og en lignende formel bevises ved martingalteori i (Doob, 1953, s.343-348). Vi nøyer oss med å bevise at denne formelen gjelder under to forutsetninger. For det første antar vi at T har en tetthet f som er kontinuerlig og ulik 0 i punktet t. For det andre antar vi at $P(X \in A \,|\, T = t)$ er kontinuerlig i punktet t. Følgende beregning kan begrunnes ved L'Hopitals regel, kjerneregelen for derivasjon og ved beregningsformelen (*) fra forrige hovedavsnitt:

$$\begin{aligned}
\lim_{h \to 0+} P(X \in A \,|\, t \leq T \leq t + h) &= \lim_{h \to 0+} \frac{\frac{d}{dh} P(X \in A \cap (t \leq T \leq t + h))}{\frac{d}{dh} P(t \leq T \leq t + h)} \\
= \lim_{h \to 0+} \frac{\frac{d}{dh} P(X \in A \cap (T \leq t + h))}{\frac{d}{dh} P(T \leq t + h)} &= \lim_{h \to 0+} \frac{\frac{d}{dt} P(X \in A \cap (T \leq t + h))}{f(t + h)} \\
&= \frac{\frac{d}{dt} P(X \in A \cap (T \leq t))}{f(t)} = P(X \in A \,|\, T = t).
\end{aligned}$$

Forutsetningene i beregningen over er spesielt oppfylt for de betingede sannsynlighetene i Borel-paradokset vi innledet med. Leseren oppfordres herved til å overbevise seg om at formelen vi utledet nettopp leder til en geometrisk beregning av de nevnte betingede sannsynlighetene ved hjelp av areal mellom nivåkurver. Svarene kan finnes uten integrasjon fra figur 1.

En tredje metode er gitt ved variabelskifte (Lehmann, 1986, s.20) (Rao, 1973, s.99). Ideen er ganske enkelt å skifte koordinater slik at en finner en simultantetthet hvor variabelen det betinges med hensyn på er en av koordinatene. Deretter kan den elementære metoden benyttes. Denne metoden ble benyttet i det første hovedavsnittet i dette notatet. La $X = (X_1, \ldots, X_n)$ være en tilfeldig vektor hvor fordelingen er gitt av en tetthet f. Anta at vi ønsker å beregne $\mathrm{E}(\phi(X) \,|\, T = t)$ hvor $T = \psi(X) = (\psi_1(X), \ldots, \psi_m(X))$ med $m \leq n$. Anta videre at det finnes en stykkevis glatt en-en transformasjon $\chi : \mathbb{R}^n \to \mathbb{R}^n$ slik at $t = (\psi_1(\chi(t, s)), \ldots, \psi_m(\chi(t, s)))$. Da kan vi skifte fra $x_1, \ldots, x_n$ koordinater til $t_1, \ldots, t_m, s_{m+1}, \ldots, s_n$ koordinater ved at $x = \chi(t, s)$. La

$$\left| \frac{\partial(x_1, \ldots, x_n)}{\partial(t_1, \ldots, s_n)} \right|$$

være absoluttverdien til Jacobi-determinanten til variabelskiftet. Det følger at

$$\mathrm{E}(\phi(X) \,|\, T = t) = \int \phi(x) \frac{f(x)}{g(t)} \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(t_1, \ldots, s_n)} \right| ds$$

hvor g er tettheten til T. Den er gitt ved

$$g(t) = \int f(x) \left| \frac{\partial(x_1, \dots, x_n)}{\partial(t_1, \dots, s_n)} \right| ds$$

I begge formlene har vi forkortet notasjonen noe ved at koordinatene x skal oppfattes som en funksjon av koordinatene (t, s) ved $x = \chi(t, s)$. Betingede sannsynligheter finnes ved

$$\boxed{P(X \in A \,|\, T = t) = \int [x \in A] \frac{f(x)}{g(t)} \left| \frac{\partial(x_1, \dots, x_n)}{\partial(t_1, \dots, s_n)} \right| ds}$$

Metoden gir gode eksempler på bruk av metodene som benyttes i kurs i flerdimensjonal analyse, men blir fort litt arbeidskrevende.

Den kritiske leser vil kunne innvende at $P(X \in A \,|\, T = t)$ kun er definert som en ekvivalensklasse av funksjoner ϕ av t hvor ψ er ekvivalent med ϕ dersom $P(\phi(T) = \psi(T)) = 1$. Slik sett gir det desverre ikke mening å snakke om funksjonsverdien $P(X \in A \,|\, T = t)$ for en t som oppfyller $P(T = t) = 0$, dvs konklusjonen er at $P(X \in A \,|\, T = t)$ forblir udefinert. Et svar til denne innvendingen er gitt ved at i tilfellene hvor ekvivalensklassen inneholder en kontinuerlig funksjon, så er denne entydig og formelen (*) gir funksjonsverdien til denne kontinuerlige funksjonen. Et bedre svar er gitt ved en drøfting av hvorvidt $Q(A) = P(X \in A \,|\, T = t)$ definerer en fordeling Q for gitt t, og om denne i så fall er unik. Spørsmålet om eksistens av en slik betinget fordeling er drøftet i (Doob, 1953), (Feller, 1971), (Lehmann, 1986), (Rao, 1973).

På tampen benytter vi anledningen til å utfordre leserne til å formidle alternative metoder for beregning av betingede forventningsverdier. Metodene bør kunne brukes på eksemplene vi har gitt. Dersom noen av leserne vet om en algoritme for simulering fra en betinget fordeling, så vil vi gjerne høre om dette. Vi tenker her selvsagt på tilfellet hvor en ikke kan ta utgangspunkt i en gitt betinget tetthet $f(x \,|\, t)$ fordi formen til $f(x \,|\, t)$ er ukjent. Problemet er å finne en algoritme som er effektiv i tilfeller hvor $P(T = t)$ er liten. En algoritme som virker i spesielle tilfeller er gitt i (Engen and Lillegård, 1997). Har noen av leserne andre forslag?

References

Doob, J. L. (1990 (1953)). *Stochastic Processes.* Wiley Classics Library. Wiley.

Doob, J. L. (1994). *Measure Theory.* Springer.

Engen, S. and Lillegård, M. (1997). Stochastic simulations conditioned on sufficient statistics. *Biometrika*, 84(1):235–240.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, vol. II.* Wiley series in probability and mathematical statistics. Wiley, second edition.

Larsen, R. and Marx, M. (1986). *An introduction to mathematical statistics and its applications.* Prentice-Hall.

Lehmann, E. L. (1983). *Theory of point estimation.* Wiley.

Lehmann, E. L. (1986). *Testing statistical hypotheses.* John Wiley & Sons Inc., second edition.

Rao, C. (1973). *Linear statistical inference and its applications.* Wiley.

Gunnar Taraldsen
Institutt for matematiske fag, seksjon Lade
Norges teknisk-naturvitenskapelige universitet
7034 Trondheim
gunnar@matstat.unit.no
http://www.matstat.unit.no/~gunnar