ML-METODEN. (Maximum Likelihood-metoden.)

Ex.1 Antag att man har två tärningar där den ena är en "blufftärning" som har sannolikhet 9/10 att ge en sexa och den andra är en vanlig tärning som har sannolikhet 1/6 att ge en sexa. Man väljer en tärning på måfå utan att veta vilket slags tärning det är. När man kastar tärningen får man en sexa. Alternativ 1. är nu att det är "blufftärningen" man kastat och alternativ 2. är att det är den vanliga tärningen. Vilket alternativ bör man välja? Självklart alternativ 1. tycker du säkert.

Det du då har gjort är att du använt en s k Maximum Likelihood skattning. (En "maximal trolikhets" skattning.) För att se hur en ML-skattning fungerar inför vi nu parametern a som är alternativets nummer och tittar på sannolikheten att man får en sexa som funktion av a. Om vi kallar funktionen för L(a) får vi:

L(a) = P(sexa)= 9/10 om a=1

L(a)=P(sexa)= 1/6 om a=2.

Vi ser nu att vi valde det a-värde som gjorde L(a) så stor som möjligt.

Ex.2 Vi gör tre oberoende observationer på en stokastisk variabel som är poissonfördelad med parameter μ och får 2, 3 och 1. Vår uppgift är att skatta μ på bästa sätt. I FS 3 ser vi att parametern μ i en poissonfördelning har en dubbel betydelse, den är både väntevärde och varians på samma gång. Om vi uppfattar μ som väntevärde ska vi skatta μ med medeltalet d v s med (2+3+1)/3=2 och om vi istället ser μ som varians ska vi skatta μ med stickprovsvariansen s^2 som enligt FS 8 blir $= ((2-2)^2+(3-2)^2+(1-2)^2)/(3-1) = 1$. Vilket är bäst? Gör vi på samma sätt som i exempel 1 och ser på sannolikheten för "det som hänt" som funktion av parametern får vi:

$L(\mu) = P((X_1 = 2) \cap (X_2 = 3) \cap (X_3 = 1))$ = (ober. händelser!)= $P(X_1 = 2)\cdot P(X_2 = 3)\cdot P(X_3 = 1)$ = (FS 3 sid 1) = $\frac{\mu^2}{2!}e^{-\mu} \cdot \frac{\mu^3}{3!}e^{-\mu} \cdot \frac{\mu^1}{1!}e^{-\mu}$.

Tänker vi nu på samma sätt som i exempel 1 ska vi välja det μ-värde som gör L(μ) så stor som möjligt. $\frac{dL(\mu)}{d\mu} = 0$ ger efter lite räknande μ=(2+3+1)/3=2 som är det värde som gör L(μ) maximal och (här!) är den bästa skattningen, vilket betyder att vi bör uppfatta μ som väntevärde. Obs. Svaret bör egentligen vara $\mu \approx 2$ eftersom μ är en OKÄND konstant som vi har uppskattat.

Ex.3 En stokastisk variabel X har täthetsfunktion $\frac{1}{m}e^{-x/m}$ om x>0 och =0 annars. Tre oberoende observationer av X ger 1.245, 2.541 och 0.687. Härled ML-skattningen av m.

I exempel 2 var L(m)= produkten av sannolikhetsfunktionerna. Nu när vi har en kontinuerlig fördelning blir L(m) istället = produkten av täthetsfunktionerna d v s vi får:

$L(m) = \frac{1}{m}e^{-1.245/m} \cdot \frac{1}{m}e^{-2.541/m} \cdot \frac{1}{m}e^{-0.687/m}$

Lite räknande ger att L(m) har max för m=(1.245+2.541+0.687)/3=1.491 som är den sökta ML- skattningen. Rimligt eftersom m=E(X). (Jfr exponentialfördelning i FS 4.)

För att slippa derivera produkter kan man utnyttja att L(m) och ln (L(m)) har maximum för samma m-värde.(Detta beror på att ln är en växande funktion.) Eftersom ln(L(m)) är en summa (av logaritmer) förenklar detta deriveringen avsevärt, men man måste se till att man verkligen kan log-lagarna: 1. $\ln(ab) = \ln a + \ln b$, 2. $\ln(a/b) = \ln a - \ln b$, 3. $\ln a^b = b \ln a$. En annan sak att tänka på är att om man känner sig det minsta osäker på hur summa och produkttecken fungerar är det bättre att inte använda dessa tecken utan att istället skriva ut summor och produkter ”med prickar mellan” t ex $x_1+x_2+\cdots+x_n$ istället för $\sum_{j=1}^n x_j$

Det finns gott om tentatal på ML-metoden. Försök först att lösa talen själv utan att titta på lösningen. Att bara läsa igenom lösningen lär man sig mycket lite på. Här får du ett gammalt tentatal att öva på. (Med flit har ingen lösning bifogats.)

$x_1, x_2 \cdots x_n$ är oberoende observationer på en stokastisk variabel som har täthetsfunktionen: $f(x) = \frac{x^2}{\theta} e^{-x^3/3\theta}$ om $x > 0$ och $= 0$ annars. Härled ML-skattningen av parametern θ.
Svar: $\theta \approx (x_1^3 + x_2^3 + \cdots + x_n^3)/3n$

Tillägg: Visa att ML-skattningen ovan är väntevärdesriktig. ML-skattningen är ofta den bästa skattningen i den meningen att den är ”precisast” (dvs har minst medelfel), men den är inte automatiskt väntevärdesriktig.

För att visa att en skattning är väntevärdesriktig måste man tänka sig tillbaka till situationen INNAN observationerna har gjorts. Skattningen av θ är då en stokastisk variabel. Om vi betecknar denna med θ^* och skriver X_j istället för x_j har vi $\theta^* = (X_1^3+X_2^3+\cdots+X_n^3)/3n$ och det vi ska visa är att $\mathrm{E}(\theta^*) = \theta$ gäller. Om vi utnyttjar att väntevärdet för en summa är lika med summan av väntevärdena samt att $\mathrm{E}(X_j^3) = \int_{-\infty}^{\infty} x^3 \cdot f(x)dx = \int_0^\infty x^3 \cdot \frac{x^2}{\theta} e^{-x^3/3\theta} dx$ är detta inte så svårt. (Sätt $u = x^3/3\theta$ och partialintegrera för att lösa integralen.)

Ex. 4. Som vi såg i ex.1 är det inte alltid nödvändigt att derivera för att bestämma en ML-skattning. Här är ett annat lite mer knepigt exempel på detta: x_1, x_2, x_3, x_4 är oberoende observationer på X som har täthetsfunktionen f(x)=1/θ om $0 \leq x \leq \theta$ och $= 0$ annars. Vår uppgift är att bestämma ML-skattningen av θ.

Eftersom fördelningen är kontinuerlig får vi $L(\theta) = f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4)$. För att komma vidare måste vi nu se talen x_i INTE som de observationer de är utan som tal vilka som helst. Vi kan dock anta att alla x_i är ≥ 0 för skulle något x_i vara < 0 så blir L(θ)=0 vilket inte kan vara dess största värde. Om alla x_i är ≥ 0 så är ”alla $0 \leq x_i \leq \theta$” ekvivalent med att ”Max $x_i \leq \theta$”. Max x_i (=det största av x_i:na) är ett tal som vi kallar c.

För $\theta \geq c$ har vi alltså L(θ) $= (1/\theta)^4$ medan $\theta < c$ ger L(θ)=0. Om man ritar upp funktionen L(θ) så ser man att den har sitt maximum för θ=c. (Detta är ett s k spetsmaximum som inte avslöjar sig vid derivering.) ML-skattningen av θ i denna situation är alltså det största av de observerade värdena, vilket du nog inser är ett mycket rimligt resultat. Skattningen är dock inte väntevärdesriktig utan måste ”multipliceras upp lite” för att bli det, närmare bestämt multipliceras med $\frac{n+1}{n}$ där n är antalet observationer. Detta är inte alltför svårt att visa. (Se sid 258 i kursboken.)

Antag exempelvis att observationerna är 3.73, 4.25, 2.67, 4.20. Max x_i=4.25 är ML-skattningen av θ och vill vi att den ska bli väntevärdesriktig så måste vi multiplicera den med 5/4.

SAMMANFATTNING. ML-metoden är en metod för att RÄKNA FRAM en skattning av en okänd parameter i en fördelning. Härvid betraktas parametern (tillfälligtvis) som argument i den s k Likelihood-funktionen som det gäller att maximera. Ofta (men inte alltid) innebär maximeringen en derivering. Denna underlättas vanligen om man istället för att derivera L-funktionen själv deriverar ln för den. I det matematiska genomförandet av maximeringen ligger den största svårigheten och det är viktig att man övar detta. Tänk också alltid efter om ditt svar är rimligt. Skattningen ska t ex alltid vara en klart definierad funktion av observationerna.