

Fejezetek a magas hőmérsékletű kísérleti plazmafizikából

Kocsis Gábor, Bencze Attila, Dunai Dániel, Kálvin Sándor,
Szepesi Tamás & Zoletnik Sándor

2013. szeptember 25.

1. fejezet

Kálvin Sándor: Bayes–adatfeldolgozás alkalmazása a plazmafizikában

1.1. Bevezetés

Az adatfeldolgozás fontos szerepet játszik a kísérleti plazmafizikában. A mérési adatok kiértékelésével vonunk le következtetést a plazma fizikai paramétereiről és az elméleti modellek igaz vagy hamis voltáról. A hagyományos statisztikus módszerek mellett egyre több helyen alkalmazzák a Bayes valószínűségelméleten alapuló adatfeldolgozást. A mérési adatok ilyen módon való kiértékelésének előnye, hogy egyszerű, átgondolt és elméletileg megalapozott választ ad az adatfeldolgozásban felmerülő problémákra. Figyelembe vehetőek a kiértékeléssel kapcsolatos előzetes ismereteink, amelyek nagymértékben javítják a feldolgozás pontosságát. Lehetőséget teremt az úgynevezett érdektelen paraméterek kezelésére, mely a klasszikus statisztikus kiértékelésben nem lehetséges, valamint a modellválasztás (klasszikus elnevezése- hipotézistesztt) problémájának logikailag hibátlan megoldására.

A jegyzetben a Bayes elmélet néhány fontos alkalmazása nem kapott helyet, úgymint a modellválasztás, a numerikus módszerek ismertetése, a kísérletek tervezése [Dreier] és az integrált adatfeldolgozás (IDA–Integrated Data Analysis [IDA]). Ezekről a módszerekről, valamint a jegyzetben található témákról magyarázatok és részletes levezetések a [SIVIA] és a [GREGORY] irodalmakban találhatóak.

1.2. A deduktív és az induktív következtetés – a mérés célja a tudományokban

A fizikában a kísérleti adatokból olyan elméleti modelleket alkotunk, amelyekkel további kísérletek ellenőrizhető eredményeit jósolhatjuk meg. Ezen kísérletek eredményeinek feldolgozásával meghatározható az elmélet igaz vagy hamis volta, illetve az elméletben használt fizikai paraméterek értéke meghatározható.

A deduktív következtetés az, amellyel egy elmélet kísérletileg ellenőrizhető jóslatait alkotjuk meg. Ezek a jóslatok az elmélet és a mérési eljárás ismeretében meghatározhatóak, azaz deduktív következtetésen értjük azt, mikor az okból az általa létrehozott okozatokat határozzuk meg. A deduktív következtetés mindig egyértelmű választ ad a felmerülő kérdésekre.

Mivel tudásunk mindig hiányos és a mérések pontatlanok, az elmélet kísérleti tesztje nem szolgáltat egyértelmű választ annak igaz vagy hamis voltát illetően, így a mért okozatokat más és más hatások is létrehozhatják. A statisztikus vagy valószínűségi következtetés az a folyamat, amivel hiányos információinkból következtetünk egy elmélet igaz vagy hamis voltára, vagy meghatározzuk a fizikai paraméterek lehetséges értékeit. Ezt a folyamatot nevezzük induktív következtetésnek, azaz mikor a mért okozatból következtetünk annak lehetséges okaira. Természetesen a lehetséges elméletekről, vagy egy elmélet fizikai paramétereiről csak valószínűségi kijelentéseket tehetünk.

Az adatfeldolgozása során induktív következtetést kell használnunk, mivel mérési adatokból vonunk le következtetéseket a lehetséges fizikai modellekről illetve a modellek paramétereiről.

1.3. A valószínűség definíciója a Bayes elméletben

A klasszikus valószínűségelméletben egy esemény valószínűségét az esemény bekövetkezésének relatív gyakoriságával definiáljuk. Valószínűséget csak valószínűségi változókra értelmezhetünk, és definíciókon alapuló deduktív következtetéseket használunk. Ez korlátozza alkalmazási lehetőségeit, az induktív következtetés csak közvetett módon lehetséges.

A Bayes elméletben a valószínűség azt fejezi ki, hogy mennyire hiszünk egy esemény igaz voltában az eseménnyel kapcsolatos információk birtokában, így valószínűséget bármely eseményhez vagy elmélethez hozzárendelhetünk. A klasszikus valószínűség a természet objektív tulajdonságát méri, míg a Bayes elméletben logikai következtetés; a természetről rendelkezésünkre álló információinkat/tudásunkat/tudatlanságunkat fejezi ki. Mivel a Bayes elméletben a valószínűség tudásunk mértéke, így mindig feltételes, függ a rendelkezésünkre álló információktól.

1.4. A Konzisztens következtetés algebrája – a Cox axiómák

Azokat a szabályokat, amelyek szükségesek a konzisztens induktív következtetés matematikai leírásához Richard Cox fejlesztette ki.

A matematikai kezelhetőséghez egy valós számot rendelünk minden állításhoz úgy, hogy minél nagyobb ez a szám, annál inkább hiszünk az állítás igaz voltában. A továbbiakban ezt a valós számot $p(A|I)$ -val jelöljük, ami egy A állítás Bayes értelemben vett valószínűsége, valamint $p(A|I) = 0$, ha az A esemény hamis és $p(A|I) = 1$, ha igaz. I a rendelkezésre álló releváns információkat jelöli. Mindent, ami a '|' jel jobb oldalán áll, igaznak tekintjük.

A logikailag következetes elmülethez két követelménynek kell teljesülnie. Az első, ha meghatározzuk, hogy mennyire hiszünk egy esemény igaz voltában, közvetve megadjuk azt is, hogy mennyire hiszünk abban, hogy az esemény hamis. A második követelmény a következő: ha megadjuk, hogy mennyire hiszünk a B esemény igaz voltában, és megadjuk, hogy mennyire hiszünk abban, hogy az A esemény igaz, feltételezve, hogy a B igaz, akkor közvetve megadjuk azt is, hogy mennyire hiszünk abban, hogy az (A, B) esemény igaz. (A vessző a logikai 'és' kapcsolat jele.) Logikailag elvárható, hogy a levonható következtetésnek függetlennek kell lennie attól, hogy milyen úton jutunk el hozzá. Ezen követelmények felhasználásával azt kapta, hogy az esemény hihetőségéhez rendelt szám a valószínűségelmélet szokásos szabályait követi, így a Bayes elméleten a 'hihetőség' joggal nevezhető valószínűségnek.

1.5. A Bayes valószínűségi kalkulus

A műveletek melyeket a Bayes valószínűségi elméletben használunk az összeg és a szorzatszabály, melyeket a Cox axiómákból lehet levezetni. Ha \bar{A} jelöli, hogy az A esemény hamis, az összegszabály a következő:

$$p(A|I) + p(\bar{A}|I) = 1. \quad (1.1)$$

A szorzatszabály a következő alakban írható fel:

$$p(A, B|I) = p(A|B, I)p(B|I) = p(B|A, I)p(A|I). \quad (1.2)$$

Ez a két szabály a Bayes valószínűségelmélet alapegyenlete.

A Bayes-tétel a szorzatszabályból kapható a 1.2. egyenlet átrendezésével:

$$p(A|B, I) = \frac{p(B|A, I)p(A|I)}{p(B|I)}. \quad (1.3)$$

A Bayes-tétel minden adatfeldolgozási folyamat kiinduló egyenlete.

Tekintsünk egy A igaz eseményt, amit fel lehet bontani teljes, egymást kölcsönösen kizáró $\{A_i\}$ eseményrendszerre, azaz az A_i események közül, ha az egyik igaz akkor a többi hamis, de legalább az egyik igaz. Ekkor

$$\sum_i p(A_i|I) = p(A|I) = 1. \quad (1.4)$$

$p(A_i|I)$ az $\{A_i\}$ eseményrendszer valószínűségeloszlása. A 1.4. kifejezés a valószínűségeloszlás normáltságát fejezi ki.

Ha az A eseményrendszer folytonos (általában egy modell paramétere), akkor annak az állításnak a valószínűsége, hogy A az $[a_1, a_2]$ intervallumba esik $\int_{a_1}^{a_2} p(A|I)dA$, ahol az A eseményrendszer $p(A|I)$ valószínűségi sűrűségfüggvényét a

$$p(A = a|I) = \lim_{\delta a \rightarrow 0} \frac{a \leq A < a + \delta a}{\delta a} \quad (1.5)$$

módon definiáljuk. A továbbiakban függetlenül attól, hogy diszkrét vagy folytonos esetet tekintünk, $p()$ -vel jelöljük mind a valószínűséget, mind a valószínűségi sűrűségfüggvényt. Folytonos esetben az összegzést a megfelelő változóra való integrálással kell helyettesíteni.

Ha a A és a B esemény független azaz, ha

$$p(B|A, I) = p(B|I) \quad \text{és} \quad p(A|B, I) = p(A|I) \quad \text{akkor} \quad p(A, B, |I) = p(A|I)p(B|I). \quad (1.6)$$

Az összeg és a szorzatszabály alkalmazásával belátható, hogy

$$p(B|I) = \sum_i p(B, A_i|I), \quad \text{folytonos esetben} \quad p(B|I) = \int p(B, A|I)dA. \quad (1.7)$$

Ezt az egyenletet marginalizációnak nevezzük. A $p(B, A|I)$ együttes valószínűség eloszlásból meghatározható a B esemény eloszlása, figyelembe véve az A esemény bizonytalanságát. Ez a szabály fontos a Bayes elmélet alkalmazásában, mert lehetőséget teremt olyan paraméterek (A) kezelésére, amelyek szükséges egy probléma megfogalmazásában, de számunkra érdektelenek (nuisance paraméterek). Ez a hibaterjedés Bayes elméletben való általánosítása.

A Bayes-tételt egy teljes, egymást kölcsönösen kizáró $\{A_i\}$ eseményrendszerre alkalmazva annak valószínűségeloszlását határozhatjuk meg:

$$p(A_i|B, I) = \frac{p(B|A_i, I)p(A_i|I)}{\sum_i p(B|A_i, I)p(A_i|I)}. \quad (1.8)$$

A nevezőt a $\sum_i p(A_i|B, I) = 1$ normálási feltételből kapjuk.

1.6. Bayes elmélet: a megismerés leírása

A Bayes-tétel adatkiértékelésben betöltött szerepe világossá válik, ha a mérések feldolgozásában előforduló eseményekre alkalmazzuk. Írja le a D esemény a mérési adatokat, míg a H esemény legyen egy fizikai elmélet, vagy egy igaznak feltételezett modell paramétere. Alkalmazzuk a Bayes-tételt a D és H eseményekre:

$$p(H|D, I) = \frac{p(D|H, I)p(H|I)}{p(D|I)}. \quad (1.9)$$

A Bayes-tételben szereplő kifejezéseknek az elnevezése és értelmezése a következő: a $p(H|I)$ a prior valószínűség, leírja ismereteinket az elmélet igaz voltáról vagy a paraméter értékéről az adatkiértékelés előtt. A $p(D|H, I)$ a likelihood valószínűség, mely megadja a mérési adatok valószínűségét, ha az elmélet igaz, vagy ha egy modell paramétere egy adott érték. Ez módosítja a prior-t és adja meg a $p(H|D, I)$ poszterior valószínűségét, tudásunkat az elmületről vagy a paraméter értékéről a mért adatok figyelembevételével. Ebben az értelemben a Bayes elmélet a megismerés leírása. Meghatározza, hogy tudásunk egy elmületről vagy egy paraméter értékéről hogyan módosul a mérési adatok ismeretében. A nevezőben álló kifejezés $p(D|I)$ a teljes valószínűség (Evidence, Prior Predictive Value), és a modellválasztásban kap lényeges szerepet.

1.7. Paraméterbecslés

Az adatkiértékelésben előforduló egyik feladat mikor a mért adatokból egy adott elmélet egy vagy több paraméterének értékéről akarunk következtetést levonni. A modell paraméterei általában folytonos változók. A Bayes-paraméterbecslés nem egy adott értéket szolgáltat a paraméter értékéről, hanem annak valószínűségeloszlását. Fontos megemlíteni, hogy a modell paramétere konkrét érték, a paraméter valószínűségeloszlása jellemzi ismereteink hiányos voltát. Jelölje \mathbf{D} az adatokat, \mathbf{A} és \mathbf{B} a modell (általában folytonos) paramétervektorait. A modell paramétereit két csoportra osztottuk: \mathbf{A} azok a paraméterek, amelyek számunkra érdekesek, ezen paraméterek értékéről akarunk információt szerezni, míg \mathbf{B} az úgynevezett érdektelen paraméterek (nuisance paraméterek) amelyek szükségesek a fizikai modell leírásához, de különben értéke számunkra érdektelen. Ebben az esetben a Bayes-tétel a következőképpen írható:

$$p(\mathbf{A}, \mathbf{B}|\mathbf{D}, I) = \frac{p(\mathbf{D}|\mathbf{A}, \mathbf{B}, I)p(\mathbf{A}, \mathbf{B}|I)}{\iint p(\mathbf{D}|\mathbf{A}, \mathbf{B}, I)p(\mathbf{A}, \mathbf{B}|I)d\mathbf{A}d\mathbf{B}} \propto p(\mathbf{D}|\mathbf{A}, \mathbf{B}, I)p(\mathbf{A}, \mathbf{B}|I) \quad (1.10)$$

mivel a nevezőben álló kifejezés a paraméterektől független normálási állandó. A $p(\mathbf{D}|\mathbf{A}, \mathbf{B}, I)$ likelihood adja meg a mért adatok valószínűségeloszlását a paraméterek adott értékei esetén. $p(\mathbf{A}, \mathbf{B}|I)$ a paraméterek prior valószínűségeloszlása. A számunkra érdekes \mathbf{A}

paraméterek poszterior eloszlását marginalizációval kaphatjuk meg:

$$p(\mathbf{A}|\mathbf{D}, I) = \int p(\mathbf{A}, \mathbf{B}|\mathbf{D}, I)d\mathbf{B}. \quad (1.11)$$

A 1.10. és a 1.11. egyenletek a Bayes paraméterbecslést teljesen meghatározzák. Azonban ezeknek az egyszerűnek tűnő egyenleteknek a kiszámítása sokszor nem triviális, különösen többdimenziós esetben, mikor az integrálások (a marginalizációk elvégzése) nagymennyiségű numerikus számolást igényelnek.

Jelentős egyszerűsítések lehetségesek, ha a változók függetlenek, a valószínűségek normális eloszlásúak, vagy a fizikai modell lineáris.

A paraméterbecslésénél mind a likelihood, mind a prior szerepet játszik. A likelihood a fizikai modell és a mérési eljárás ismeretében meghatározható. A prior írja le ismereteinket a paraméterekről a mérési adatok feldolgozása előtt. Ennek meghatározásánál felhasználhatjuk előző kísérletek eredményeit: a paraméterek poszterior eloszlását, vagy általános megfontolásokkal élhetünk. Ezeket a módszereket a 1.9. fejezetben tárgyaljuk.

1.7.1. A legjobb becslés és a hibahatárok meghatározása

Egy paraméter poszterior eloszlása következtetésünket a paraméter értékéről teljesen meghatározza. Legyen az α paraméter poszterior eloszlása $p(\alpha|\{d_k\}, I)$, ahol $\{d_k\}$ jelöli a mért adatokat. A poszterior eloszlás maximumát a paraméter legvalószínűbb értékénél veszi föl, a legjobb becslés a paraméter ezen értéke. A poszterior eloszlás ezen érték körüli szélessége megadja a becslés megbízhatóságát. Ezen értékek a poszterior eloszlás lokális tulajdonságát jellemzik. Az eloszlásfüggvény globális tulajdonságait az eloszlás átlagértékével, momentumaival valamint a konfidencia intervallummal lehet megadni.

Közelítsük a poszterior eloszlást a legjobb becslés környezetében Gauss eloszlással:

$$p(\alpha|\{d_k\}, I) \propto \exp \left[-\frac{(\alpha - \alpha_0)^2}{2\sigma^2} \right], \quad (1.12)$$

ahol α_0 az eloszlás maximuma, mely a legjobb becslés, σ paramétere, ami az eloszlás szélességével arányos, a becslés megbízhatóságát jellemzi, így ezen két értékkel jellemezhetjük a poszterior eloszlásfüggvényt.

Fejtsük Taylor sorba a poszterior eloszlás $L(\alpha) = \ln [p(\alpha|\{d_k\}, I)]$ logaritmusát az α_0 pont környezetében:

$$L(\alpha) = L(\alpha_0) + \frac{dL(\alpha_0)}{d\alpha}(\alpha - \alpha_0) + \frac{1}{2} \frac{d^2L(\alpha_0)}{d\alpha^2}(\alpha - \alpha_0)^2 + \dots \quad (1.13)$$

Mivel L monoton függvénye p -nek, így L maximuma p -nek is maximuma, azaz a legjobb becslést a következő feltétel adja:

$$\frac{dL(\alpha_0)}{d\alpha} = 0. \quad (1.14)$$

A magasabb rendű tagokat elhanyagolva és a 1.14. feltételt kihasználva, majd az egyenlet exponenciálisát véve kapjuk, hogy

$$p(\alpha|\{d_k\}, I) \propto \exp\left[\frac{1}{2}\frac{d^2L(\alpha_0)}{d\alpha^2}(\alpha - \alpha_0)^2\right]. \quad (1.15)$$

Ezt összehasonlítva a 1.12. egyenlettel kapjuk, hogy

$$\sigma = \left[-\frac{d^2L(\alpha_0)}{d\alpha^2}\right]^{-1/2}. \quad (1.16)$$

Következtetésünket az α paraméterről tömören a következőképpen foglalhatjuk össze:

$$\alpha = \alpha_0 \pm \sigma. \quad (1.17)$$

Itt α_0 a paraméter legjobb becslése, σ pedig a becslés megbízhatóságát jellemzi, amit szokás a paraméter hibájának nevezni. A becslés hatékonyságának e módon való jellemzése akkor reális, ha a poszterior Gauss eloszlással közelíthető. Egyszerű problémák, és/vagy több mérési adat esetén ez sokszor jó közelítéssel teljesül.

1.7.2. Klasszikus paraméterbecslés

Az előzőekben megmutattuk, hogy a Bayes-paraméterbecslés hogyan határozza meg egy igaznak feltételezett modell paramétereinek valószínűségeloszlását, a poszterior eloszlást. Ebben a fejezetben a két legismertebb klasszikus paraméterbecslésnek, a maximum likelihood és a legkisebb négyzetek módszerének, a Bayes-paraméterbecsléssel való kapcsolatát tekintjük át. A klasszikus paraméterbecslés esetében nem vehetünk figyelembe előzetes információkat a paraméterek értékeiről, így a 1.10. kifejezésben egyenletes (nem informatív) prior eloszlást kell használni (Az egyenletes prior eloszlás fejezi ki teljes tudatlanságunkat a paraméterről), azaz:

$$p(\mathbf{A}|\mathbf{D}, I) \propto p(\mathbf{D}|\mathbf{A}, I), \quad (1.18)$$

a poszterior a likelihood-dal arányos. Az érdektelen paramétereket nem különböztetjük meg, mert klasszikus esetben nincs lehetőség ezen paraméterek kezelésére (marginalizációra). A

$$\mathcal{L}(\mathbf{A}) = p(\mathbf{D}|\mathbf{A}, I) \quad (1.19)$$

likelihood függvény maximalizálásával kaphatjuk az \mathbf{A} paraméterek legjobb becslését. A likelihood függvény használata helyett maximalizálhatjuk annak logaritmusát, ami sok esetben egyszerűsíti a számolásokat. Ez a maximum likelihood paraméterbecslés.

Tekintsük azt az esetet mikor a zaj nélküli adatokat a $D = f(x, \mathbf{A})$ függvénnyel lehet jellemezni, és az $\{x_i\}$ mintavételi helyeken a mérési adatok $\mathbf{D} = \{d_i\}$. Minden

adat leírható a $p_i(d_i|\mathbf{A}, I)$ eloszlással. Ha a mérések függetlenek, akkor a likelihood a következő:

$$\mathcal{L}(\mathbf{A}) = p(\mathbf{D}|\mathbf{A}, I) = \prod_{i=1} p_i(d_i|\mathbf{A}, I). \quad (1.20)$$

Az adatfeldolgozásban gyakran előforduló eset, mikor a mérések bizonytalansága Gauss eloszlással írható le:

$$p_i(d_i|\mathbf{A}, I) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{(d_i - f(x_i, \mathbf{A}))^2}{2\sigma_i^2} \right], \quad (1.21)$$

ahol σ_i az i . mérés 'hibája'. Ebben az esetben a likelihood függvény logaritmus a következő:

$$\ln [\mathcal{L}(\mathbf{A})] \propto -\frac{1}{2}\chi^2 = -\frac{1}{2} \sum_{i=1} \left[\frac{d_i - f(x_i, \mathbf{A})}{\sigma_i} \right]^2. \quad (1.22)$$

A log-likelihood függvénynek ott van maximuma ahol χ^2 -nek minimuma, így a paraméterek legjobb becslését a χ^2 minimalizációjából határozhatjuk meg. Ebben az esetben χ^2 minimalizálásról vagy legkisebb négyzetek módszeréről beszélünk.

A fenti esetek rámutatnak arra, hogy mind a maximum likelihood mind a legkisebb négyzetek módszere a Bayes paraméterbecslés speciális esete. Mindkét esetben egyenletes prior feltételezéssel élünk. A legkisebb négyzetek módszerénél a mérési adatok függetlenségét és a mérések zajának Gauss eloszlását is fel kellett használnunk. Ha ezen klasszikus paraméterbecslések nem használhatóak, akkor a probléma pontosabb megfogalmazása szükséges, és a Bayes-paraméterbecslés alkalmazása szolgáltatja a megoldást.

1.8. Átlagolás Gauss zaj esetén

A fentebb ismertetett általános elvek szemléltetésére tekintsük azt a problémát, amikor a mérések bizonytalanságát normális eloszlással jellemezhetjük. A normális eloszlást gyakran használjuk a mérési adatok zajának jellemzésére (Látsd a 1.9. fejezetet).

1.8.1. Mérések ismert hibával

Legyen $\mathbf{D} = \{d_i\}$ a független mérési adatok halmaza és tegyük fel, hogy $d_i = \mu + \varepsilon_i$, ahol az ε_i zaj σ paraméterrel jellemezhető Gauss eloszlás, a mérési adatok száma legyen n . Annak a valószínűsége, hogy a i . adat d_i

$$p(d_i|\mu, \sigma, I) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(d_i - \mu)^2}{2\sigma^2} \right]. \quad (1.23)$$

Határozzuk meg a μ paraméter legjobb becslését, és a becslés hibáját abban az esetben amikor σ értéke ismert. A μ paraméter becsléséhez meg kell határozni μ poszterior

eloszlását. Ezt a Bayes-tétel segítségével a következőképpen adhatjuk meg:

$$p(\mu|\mathbf{D}, \sigma, I) \propto p(\mathbf{D}|\mu, \sigma, I)p(\mu|\sigma, I). \quad (1.24)$$

Abban az esetben, ha a mérési adatok függetlenek, a likelihood:

$$p(\mathbf{D}|\mu, \sigma, I) = \prod_{i=1}^n p(d_i|\mu, \sigma, I) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\sum_{i=1}^n \frac{(d_i - \mu)^2}{2\sigma^2}\right]. \quad (1.25)$$

Használjunk egyszerű egyenletes prior eloszlást, azaz $p(\mu|\sigma, I) = p(\mu) = \text{konstans}$. A poszterior logaritmus a következő:

$$L = \ln [p(\mu|\mathbf{D}, \sigma)] = \text{konstans} - \sum_{i=1}^n \frac{(d_i - \mu)^2}{2\sigma^2}, \quad (1.26)$$

ahol a konstans tartalmazza azokat a tagokat, amelyek függetlenek μ -tól. A 1.14. egyenlet szerint μ legjobb becslését a

$$\frac{dL(\mu_0)}{d\mu} = \sum_{i=1}^n \frac{d_i - \mu_0}{\sigma^2} = 0 \quad (1.27)$$

feltételből határozhatjuk meg. Az egyenlet átrendezésével kapjuk, hogy μ legjobb becslése a mérési adatok átlaga:

$$\mu_0 = \frac{1}{n} \sum_{i=1}^n d_i. \quad (1.28)$$

A becslés hibáját a 1.16. egyenletnek megfelelően L második deriváltjából kaphatjuk meg:

$$\frac{d^2L(\mu_0)}{d\mu^2} = -\sum_{i=1}^n \frac{1}{\sigma^2} = -\frac{n}{\sigma^2}. \quad (1.29)$$

A következtetésünket μ -ról tömören összefoglalva:

$$\mu = \mu_0 \pm \frac{\sigma}{\sqrt{n}}. \quad (1.30)$$

A jól ismert eredményt kapjuk, a becslés hibája a mérések számának négyzetgyökével fordítottan arányos.

1.8.2. Mérések ismeretlen hibával

Tekintsük azt az esetet, amikor a mérések hibája (σ) minden adatra azonos, de ismeretlen paraméter. Ebben az esetben a μ poszterior eloszlását marginalizációval kaphatjuk:

$$p(\mu|\mathbf{D}, I) = \int p(\mu, \sigma|\mathbf{D}, I)d\sigma. \quad (1.31)$$

Az integrálandó kifejezés a Bayes-tétel szerint:

$$p(\mu, \sigma | \mathbf{D}, I) \propto p(\mathbf{D} | \mu, \sigma, I) p(\mu, \sigma | I). \quad (1.32)$$

A likelihood eloszlást a 1.25. kifejezés adja. A prior-nak válasszunk egyenletes eloszlást és vegyük figyelembe, hogy σ csak pozitív lehet:

$$p(\mu, \sigma | I) = \begin{cases} \text{konstans} & \text{ha } \sigma > 0 \\ 0 & \text{máshol} \end{cases}. \quad (1.33)$$

A marginalizáció eredményét levezetés nélkül közöljük:

$$p(\mu | \mathbf{D}, I) \propto \left[\sum_{i=1}^n (d_i - \mu)^2 \right]^{-(n-1)/2}. \quad (1.34)$$

A poszterior eloszlás logaritmusából a 1.14. és a 1.16. egyenletek felhasználásával kapjuk, hogy

$$\mu = \mu_0 \pm \frac{S}{\sqrt{n}} \quad \text{ahol} \quad \mu_0 = \frac{1}{n} \sum_{i=1}^n d_i \quad \text{és} \quad S = \sqrt{\frac{V}{n-1}}, \quad (1.35)$$

ahol $V = \sum_{i=1}^n (d_i - \mu_0)^2$. A legjobb becslés ebben az esetben is a mérési adatok átlaga, a becslés hatékonyságát a mérési adatokból meghatározott ($\sigma = S$) szórással jellemezhetjük a 1.30. kifejezéshez hasonlóan.

Marginalizációval meghatározhatjuk a Gauss folyamat σ szórásának poszterior eloszlását is:

$$p(\sigma | \mathbf{D}, I) = \int p(\mu, \sigma | \mathbf{D}, I) d\mu. \quad (1.36)$$

A marginalizáció eredménye:

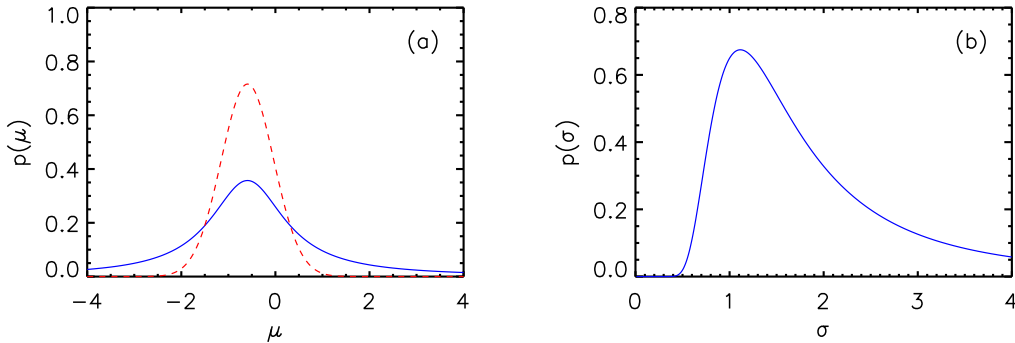
$$p(\sigma | \mathbf{D}, I) \propto \sigma^{1-n} \exp\left(-\frac{V}{2\sigma^2}\right). \quad (1.37)$$

Ezen kifejezés felhasználásával megkaphatjuk σ legjobb becslését és hibáját:

$$\sigma = \sigma_0 \pm \frac{\sigma_0}{\sqrt{2(n-1)}} \quad \text{ahol} \quad \sigma_0 = \sqrt{\frac{V}{(n-1)}}. \quad (1.38)$$

A marginalizációk részletes levezetései megtalálhatóak a [GREGORY] és [SIVIA] irodalmakban.

Az eredményeket a 1.1. ábrán szemléltetjük. Az $n = 4$ mérési adatot véletlenszerűen generáltuk nulla várható értékkel és egységnyi szórással. A 1.1.(a) ábrán a μ paraméter marginális eloszlása látható összehasonlítva egy Gauss eloszlással melynek szórása $\sqrt{V/n(n-1)}$, ami a μ hibáját jellemző Gauss eloszlás paramétere a 1.35. kifejezésnek



1.1. ábra. (a) A $p(\mu|\mathbf{D}, I)$ marginális eloszlás (folytonos vonal) összehasonlítva $\sigma = \sqrt{V/n(n-1)}$ paraméterű Gauss eloszlással (szaggatott vonal). (b) A $p(\sigma|\mathbf{D}, I)$ marginális eloszlás. Az adatok száma $n = 4$.

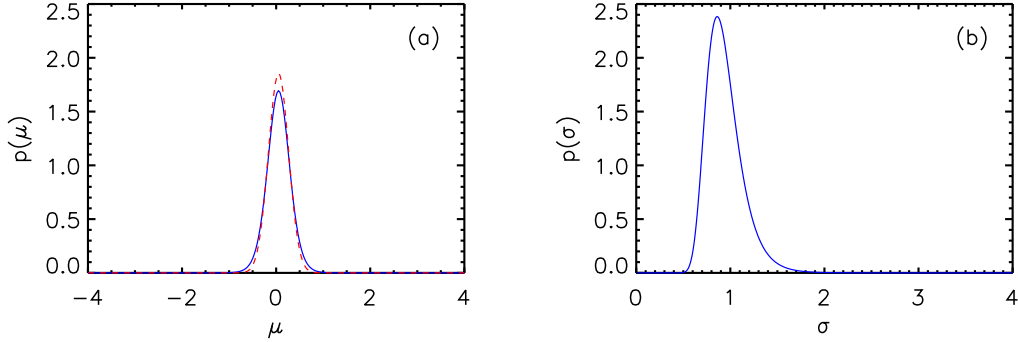
megfelelően. A 1.1.(b) ábrán σ marginális eloszlását ábrázoltuk. Látható, hogy μ marginális eloszlása lassabban cseng le, mint a 1.35. egyenletekben μ hibáját jellemző S paraméterű Gauss eloszlás, azaz a becslés megbízhatóságát csökkenti, ha a Gauss folyamat szórását a mért adatokból kell meghatározni. A fenti kiértékelést elvégeztük $n = 16$ mérési adat esetére is. Az eredmények a 1.2. ábrán láthatóak. Megállapíthatjuk, hogy a paraméterek becslése pontosabb, négyszeres adatmennyiség mintegy felére csökkentette a poszterior eloszlások szélességét (a becslés hatékonysága $\propto 1/\sqrt{n}$), emellett a poszterior eloszlások egyre inkább Gauss eloszláshoz közelítenek.

1.9. A valószínűség meghatározása

Az előző fejezetekben láttuk, hogy a Bayes elmélet hogyan teremt kapcsolatot a valószínűségek között. Míg a likelihood a fizikai modell és a mérési eljárás ismeretében meghatározható, addig a prior eloszlás meghatározásához a rendelkezésünkre álló előzetes információkat kell figyelembe venni. Létezik néhány általános szabály ezen prior valószínűségek meghatározására. A prior meghatározásánál általános elv, hogy olyan eloszlást kell meghatározni, ami leginkább kifejezi tudatlanságunkat az adott paraméterrel kapcsolatban.

Reprezentálja λ egy spektrumban a csúcs helyét, és ne legyen előzetes ismeretünk a csúcs pozíciójáról (egy adott intervallumban bárhol elhelyezkedhet). Abban az esetben, ha a koordináta-rendszer kezdőpontját λ_0 -val eltoljuk, akkor ez nem változtathatja meg a csúcs helyének prior valószínűségét, azaz

$$p(\lambda|I)d\lambda = p(\lambda + \lambda_0|I)d(\lambda + \lambda_0). \quad (1.39)$$



1.2. ábra. (a) A $p(\mu|\mathbf{D}, I)$ marginális eloszlás (folytonos vonal) összehasonlítva $\sigma = \sqrt{V/n(n-1)}$ paraméterű Gauss eloszlással (szaggatott vonal). (b) A $p(\sigma|\mathbf{D}, I)$ marginális eloszlás. Az adatok száma $n = 16$.

Mivel λ_0 konstans $d(\lambda + \lambda_0) = d(\lambda)$, amiből következik, hogy

$$p(\lambda|I) = \text{konstans}, \quad (1.40)$$

azaz tudatlanságunk egy helyparaméterről egyenletes prior eloszlással jellemezhető.

Hasonló megfontolással élhetünk abban az esetben is, ha a csúcs A amplitúdójának prior eloszlását akarjuk meghatározni. Ha nem ismerjük az amplitúdó nagyságrendjét, akkor skálájának megváltoztatása (α -val való szorzása), nem változtathatja meg az amplitúdó prior valószínűségét:

$$p(A|I)dA = p(\alpha A|I)d(\alpha A) = \alpha p(\alpha A|I)dA. \quad (1.41)$$

Ezen feltétel abban az esetben teljesül, ha $p(A|I) \propto 1/A$. A pozitív skálaparaméterekre alkalmazandó prior valószínűségi hozzárendelést Jeffreys priornak nevezzük.

Felmerül a kérdés, hogyan határozhatjuk meg a valószínűségi eloszlásokat, ha az eloszlásról csak bizonyos megkötéseink vannak. Ezekben az esetekben a maximum entrópia módszerét használhatjuk. Ezen elv alapján a legkevésbé informatív valószínűségeloszlást az eloszlás úgynevezett entrópiájának maximalizálásával kaphatjuk. Az entrópia pontos értelmezése a [GREGORY] irodalomban megtalálható. Egy $\{p_i = p(X = x_i|I)\}$ eloszlás S entrópiáját a következőképpen definiáljuk:

$$S(\{p_i\}) = - \sum_{i=1} p_i \ln(p_i). \quad (1.42)$$

A maximum entrópia módszerét két példán mutatjuk be.

Az első esetben, csak a valószínűségekre alkalmazandó normálási feltételt használjuk ki, azaz a $\sum_i p_i = 1$ megkötést. A valószínűségeket a Lagrange szorzók módszerével

határozhatjuk meg, azaz a

$$W = - \sum_{i=1} p_i \ln(p_i) + \lambda(1 - \sum_i p_i) \quad (1.43)$$

kifejezést kell maximalizálni és a λ paramétert meghatározni úgy, hogy a normálási feltétel teljesüljön. A szélsőérték feltétele, hogy

$$\frac{\partial W}{\partial p_j} = -1 - \ln(p_j) - \lambda = 0 \quad (1.44)$$

minden j -re. A fenti egyenletekből kapjuk a megoldást:

$$p_j = \exp[-(1 + \lambda)], \quad (1.45)$$

azaz minden esemény valószínűsége ugyanaz. A Lagrange szorzót a normálási feltétel határozza meg. Folytonos eseményrendszer esetén az entrópiát az egyenletes eloszlás maximalizálja, $p(x) = \text{konstans}$.

A második esetben azt vizsgáljuk mikor az eloszlás szórásnégyzetét is ismerjük, azaz

$$\sum_i p_i = 1 \quad \text{és} \quad \sum_i (x_i - \mu)^2 p_i = \sigma^2. \quad (1.46)$$

A maximum entrópia elve alapján a

$$W = - \sum_{i=1} p_i \ln(p_i) + \lambda_0(1 - \sum_i p_i) + \lambda_1(\sigma^2 - \sum_i p_i) \quad (1.47)$$

kifejezést kell maximalizálni ahol λ_0 és λ_1 ismeretlen Lagrange szorzók. A $\partial W / \partial p_j = 0$ feltétel a

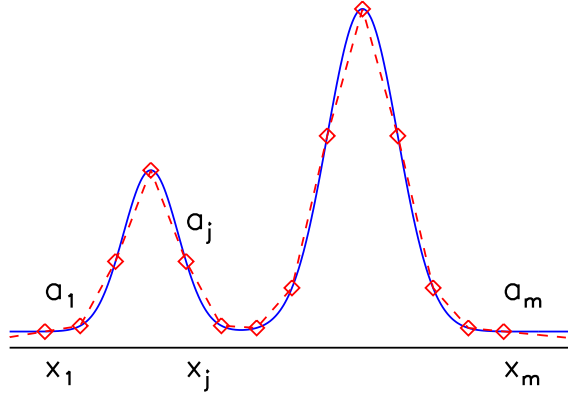
$$p_j = \exp[-(1 + \lambda_0)] \exp(-\lambda_1(x_j - \mu)^2) \quad (1.48)$$

egyenletekre vezet. Az $\exp[-(1 + \lambda_0)]$ normalási állandó és λ_1 a [1.46.](#) feltételekből határozhatóak meg. Folytonos eseményrendszer esetén ez Gauss valószínűségi eloszláshoz vezet:

$$p(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right]. \quad (1.49)$$

Abban az esetben, ha csak az az információ áll rendelkezésünkre, hogy az eloszlás jellemző átlagértékkel és véges szórásnégyzettel akkor a normális eloszlás az amely legjobban kifejezi tudatlanságunkat az adott változóról. Ezen eredmény indokolja a mérési adatok bizonytalanságának Gauss eloszlással való leírását, ha a zaj forrásait nem ismerjük.

A részletes levezetések megtalálhatóak a [\[GREGORY\]](#) irodalomban.



1.3. ábra. Az $f(x)$ (folytonos vonal) függvény parametrizálása a $\{x_j\}$ helyeken felvett $\{a_j\}$ értékeivel. A szaggatott vonal mutatja az $\hat{f}(x)$ parametrizált függvényt .

1.10. Paraméternélküli becslés

A 1.7. fejezetben megmutattuk, hogyan határozhatjuk meg egy modell paramétereit. Ebben a fejezetben azzal a problémával foglalkozunk, mikor a fizikai modellünk nem jellemezhető néhány jól definiált paraméterrel. Ezt a problémát free-form feladatnak nevezzük. Legyen modellünk a folytonos $f(x)$ függvénnyel leírható, ahol x valamilyen helyparaméter, és $f(x)$ nem jellemezhető néhány paraméterrel. Példa lehet erre, ha valamilyen plazmaparaméter folytonos helyfüggését akarjuk meghatározni. A $\{d_k\}$ mérési adatok állnak rendelkezésre ahhoz, hogy következtetést vonjunk le az $f(x)$ -ről, azaz, hogy meghatározzuk a $p(f(x)|\{d_k\}, I)$ poszterior eloszlást. A poszterior meghatározásához meg kell adnunk a likelihood és a prior eloszlásokat, valamint az $f(x)$ függvényt kell jellemezni alkalmas módon. A függvény jellemzésének egyik lehetséges módja, ha az $f(x)$ függvényt bizonyos $\{x_j\}$ helyeken felvett $\{a_j\}$ értékeivel parametrizáljuk, és a függvényt ezen pontok közötti interpolálással adjuk meg. A 1.3. ábra szemlélteti ezt az esetet. Általános esetben a mérési adatokat a következő modellel lehet leírni: $d_k = F_k(f(x), \alpha, \dots) + \varepsilon_k$, ahol F_k valamilyen funkcionál, α, \dots a modell egyéb paramétereit és ε_k a mérés hibája. Ebben az esetben a Bayes-tétel a következő alakban adható meg:

$$p(\{a_j\}, \alpha, \dots | \{d_k\}, I) \propto p(\{d_k\} | \{a_j\}, \alpha, \dots, I) p(\{a_j\}, \alpha, \dots | I). \quad (1.50)$$

Ha a meghatározandó $\{a_j\}$ paraméterek száma nagyobb mint a mérési adatok száma, vagy ha a mérési adatok egy halmaza alig függ a meghatározandó $f(x)$ függvénytől, az

$\{a_j\}$ paraméterek hibája nagy lehet, ha az $\{a_j\}$ paraméterekre egyenletes prior eloszlást használunk. Nemegeyenletes prior használatával elérhető, hogy a poszterior eloszlás egyértelmű maximummal legyen jellemezhető. Ennek egyik lehetséges módja az úgynevezett regularizáció. Regularizáció esetén fizikailag indokolható megkötést adunk meg az $f(x)$ függvényre. Az $\{a_k\}$ paraméterek becslését a

$$K \cdot S(\hat{f}(x)) + \ln [p(\{d_k\}|\{a_j\}, \alpha, \dots, I)p(\alpha, \dots | I)] \quad (1.51)$$

kifejezést maximalizálásával kaphatjuk meg, ahol $\hat{f}(x)$ az $\{a_j\}$ paraméterek által meghatározott függvény. Itt K a regularizációs paraméter, $S()$ a regularizációs függvény, ami az $f(x)$ fizikailag elvárt tulajdonságát jellemzi. A regularizációs függvényt úgy kell megválasztani, hogy értéke annál kisebb legyen minél inkább eltér az $\hat{f}(x)$ függvény a fizikailag elvárt tulajdonságtól. A 1.51. kifejezés exponenciálisát véve, és összehasonlítva a 1.50 Bayes-tétellel megállapítható, hogy a paraméterek prior eloszlását a

$$p(\{a_j\}, \alpha, \dots | I) = p(\alpha, \dots | I) \exp \left[K \cdot S(\hat{f}(x)) \right] \quad (1.52)$$

alakban adtuk meg, azaz a regularizáció az $\{a_j\}$ paraméterekre egy nemegeyenletes prior eloszlás alkalmazása ($\hat{f}(x)$ csak az $\{a_j\}$ paraméterektől függ).

Független, Gauss zajjal jellemezhető mérések esetén a regularizációs paraméter értékét úgy kell meghatározni, hogy $\chi^2 = \sum_k^n (d_k - F_k)^2 / \sigma_k^2 \approx n$ legyen, ahol n a mérések száma. Ezen feltétel biztosítja, hogy a megoldás összeegyeztethető legyen a mért adatokkal.

1.11. Alkalmazások a plazmafizikában

1.11.1. Csúcs helyének és amplitúdójának meghatározása

A plazmafizikában gyakran előforduló feladat, hogy meg kell határozni egy adott alakú jel helyét és amplitúdóját a rendelkezésre álló zajjal terhelt mérési adatokból. Példaként tekintünk azt az esetet amikor a jel (D) adott w szélességű Gauss csúcs, azaz

$$D(x, A, x_0) = A \exp \left[-\frac{(x - x_0)^2}{2w^2} \right], \quad (1.53)$$

ahol x a mérési változó, A és x_0 a meghatározandó paraméterek, a csúcs amplitúdója és helye. Méréseinket végezzük az $\{x_k\}$ helyeken, a mérési adatok legyenek $\{d_k\}$. Az ideális (zaj nélküli) mérési értékek:

$$D_k(x_k, A, x_0) = A \exp \left[-\frac{(x_k - x_0)^2}{2w^2} \right]. \quad (1.54)$$

Zajjal terhelt mérés esetén $d_k = D_k + \epsilon$. Tételezzük fel, hogy az ϵ mérési zaj Gauss eloszlású σ szórással, így a k . mérés valószínűsége:

$$p(d_k|A, x_0, I) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(d_k - D_k)^2}{2\sigma^2} \right]. \quad (1.55)$$

A Bayes-tétel szerint következtetésünket a meghatározandó paraméterekről a

$$p(A, x_0|\{d_k\}, I) \propto p(\{d_k\}|A, x_0, I)p(A, x_0|I). \quad (1.56)$$

poszterior eloszlás írja le. Ha a mérések függetlenek akkor a likelihood függvényt az egyes mérések valószínűségeinek szorzata adja:

$$\begin{aligned} p(\{d_k\}|A, x_0, I) &= \prod_{k=1}^n p(d_k|A, x_0, I) = \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left[-\sum_{k=1}^n \frac{(d_k - D_k)^2}{2\sigma^2} \right]. \end{aligned} \quad (1.57)$$

A kiértékelést az alábbi példán szemléltetjük. Véletlenszerűen generáltunk mérési adatokat amelyek a 1.4. ábrán láthatóak. Feltüntettük a 1.53. egyenlet szerinti jelet is. A csúcs amplitúdóját $A = 0,1$ -nek, helyét $x_0 = 0,3$ -nak és a zaj szórást $\sigma = 0,04$ -nek választottuk. Mivel a csúcs amplitúdója csak pozitív lehet, a prior eloszlást válasszuk a következőnek:

$$p(A, x_0|I) = \begin{cases} \text{konstans} & \text{ha } 0 < A < 0,2 \text{ és } 0,2 < x_0 < 0,4 \\ 0 & \text{máshol,} \end{cases} \quad (1.58)$$

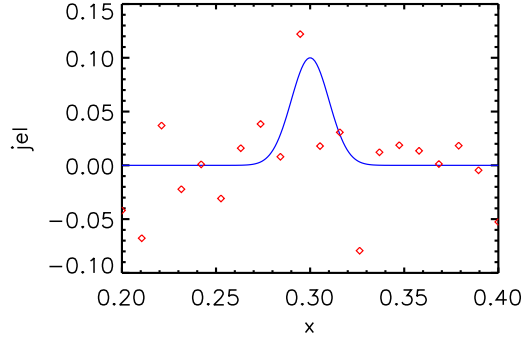
azaz mind az amplitúdóra, mind a csúcs pozíciójára egyenletes priort használunk. A kiértékelést elvégeztük az amplitúdóra alkalmazott Jeffreys prior alkalmazásával is. A prior eloszlás ebben az esetben:

$$p(A, x_0|I) = \begin{cases} \propto \frac{1}{A} & \text{ha } 0,01 < A < 0,2 \text{ és } 0,2 < x_0 < 0,4 \\ 0 & \text{máshol} \end{cases}. \quad (1.59)$$

A 1.5. ábrán a $p(A, x_0|\{d_k\}, I)$ poszterior eloszlásokat ábrázoltuk.

Az $p(A, x_0|\{d_k\}, I)$ poszterior eloszlás teljesen leírja következtetésünket a csúcs helyéről és amplitúdójáról. Gyakran azonban csak a csúcs helyére van szükségünk függetlenül az amplitúdójától. Az amplitúdót fel kellett használni a kiértékelés során, mert a likelihood meghatározásához szükséges volt. Szeretnénk tehát meghatározni a $p(x_0|\{d_k\}, I)$ eloszlást. Ezt a marginalizációval kaphatjuk meg, azaz

$$p(x_0|\{d_k\}, I) = \int p(A, x_0|\{d_k\}, I)dA. \quad (1.60)$$



1.4. ábra. A véletlenszerűen generált mérési adatok (\diamond). A folytonos vonal a 1.53. egyenlet szerinti jel.

Hasonlóan kaphatjuk meg az amplitúdó $p(A|\{d_k\}, I)$ marginális eloszlását.

Az integrálásokat numerikusan elvégezve a marginális eloszlások a 1.6.(a) és (b) ábrákon láthatóak. Annak ellenére, hogy a jel/zaj arány kicsi, megállapítható, hogy a csúcs helyének meghatározása viszonylag pontos, a csúcs alakjának előzetes ismerete miatt. Az amplitúdó bizonytalansága viszont jelentős. Valószínűsége nulla negatív értékek esetén, az alkalmazott priornak megfelelően. A 1.6.(b) ábrán az amplitúdó marginális eloszlását Jeffreys prior alkalmazása esetén is ábrázoltuk. Mivel az amplitúdó pozitív skálaparaméter, a Jeffreys prior alkalmazása realisabb eredményt szolgáltat. Az amplitúdó ebben az esetben kis értékekre sokkal valószínűbb, mint egyenletes prior alkalmazása esetén.

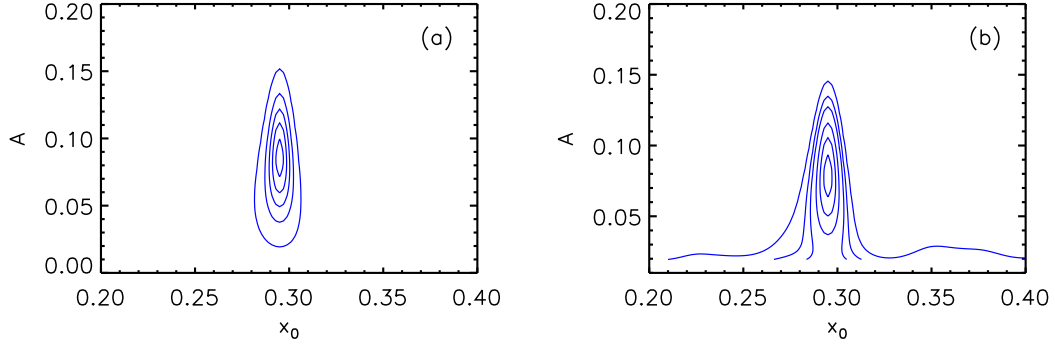
1.11.2. Paraméternélküli becslés

A plazma elektronsűrűség meghatározásának egyik lehetséges módja a plazmába injektált semleges lítiumnyaláb 2s-2p atomi átmenetéhez tartozó nyalábmenti fényeloszlás mérése, mivel a fényeloszlás függ a nyalábmenti elektronsűrűségtől. A mérési módszer részletes leírása a [FISCHER] található. A mérés fizika modellje egyszerűsítve a következőképpen írható le: a $S(z)$ nyalábmenti fényintenzitást a

$$S(z, n_e(z), \alpha|I) = \alpha N_{2p}(z, n_e(z)|I) \quad (1.61)$$

egyenlet határozza meg, ahol N_{2p} a 2p atomi energiaszint betöltöttsége, z a nyalábmenti koordináta. Az α paraméter kapcsolja össze a 2p szint betöltöttségét a mért fényintenzitással. Az N_i atomi szintek betöltöttségét a

$$\frac{dN_i(z)}{dz} = \sum_{j=1}^{M_{Li}} [n_e(z) a_{ij}(T_e(z), Z_{eff}(z), q(z), v_{Li}) + b_{ij}] N_j(z) \quad (1.62)$$



1.5. ábra. (a) A $p(A, x_0 | \{d_k\}, I)$ poszterior eloszlás egyenletes prior esetén. (b) A $p(A, x_0 | \{d_k\}, I)$ poszterior eloszlás Jeffreys prior esetén. A kontúrvonalak a maximális valószínűség 10%, 30%, 50%, 70%, 90%-ának felelnek meg.

differenciálegyenlet rendszer határozza meg [PUSZTAI]. Itt a_{ij}, b_{ij} atomfizikai állandók, $T_e(z)$, $Z_{eff}(z)$, $q(z)$ a plazma hőmérséklete, effektív rendszáma és a szennyezők átlagos töltése és v_{Li} az atomok nyalábmenti sebessége. Ezen paramétereket ismertnek tételezzük fel. A modell meghatározandó paraméterei a nyalábmenti elektronsűrűség és az α paraméter. A $\{d_k\}$ mérési adatok a fényintenzitás értékei adott z_k helyeken, azaz $d_k = S(z_k, n_e(z), \alpha | I) + \varepsilon_k$. ε_k a mérések zaja, amiről feltételezzük, hogy Gauss eloszlású σ_k szórással.

A mérések függetlenségét felhasználva a likelihood a következő:

$$p(\{d_k\} | n_e(z), \alpha, I) = \frac{1}{\prod_k \sqrt{2\pi\sigma_k^2}} \exp \left[-\frac{1}{2} \sum_k \left(\frac{d_k - S(z_k, n_e(z), \alpha | I)}{\sigma_k} \right)^2 \right] \quad (1.63)$$

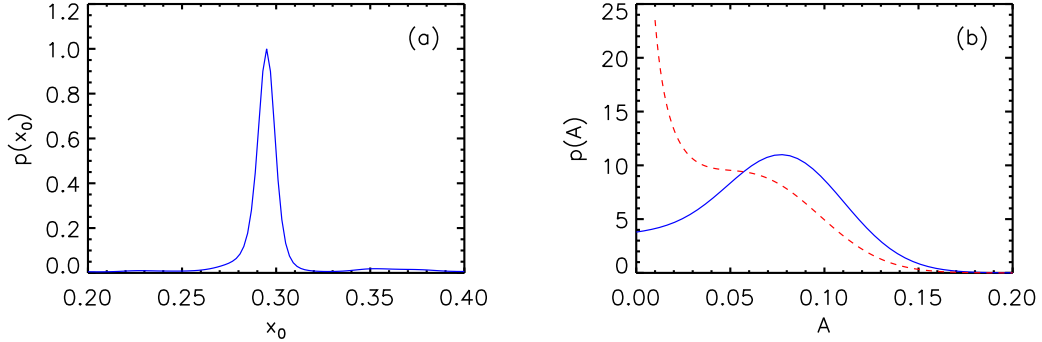
Az $n_e(z)$ függvényt parametrizáljuk a 1.10. fejezetben leírt módon az $\{a_j\} = n_e(\{z_j\})$ értékekkel. A prior meghatározásánál azt a feltételt kell figyelembe venni, hogy mind a sűrűség mind az α paraméter csak pozitív értékeket vehet fel

$$p^a(\{a_j\}, \alpha | I) = \begin{cases} \text{konstans} & \text{ha } a_j > 0 \text{ és } \alpha > 0 \\ 0 & \text{máshol.} \end{cases} \quad (1.64)$$

A meghatározandó paraméterek poszterior eloszlása a Bayes-tételnek megfelelően:

$$p(\{a_j\}, \alpha | \{d_k\}, I) \propto p(\{d_k\} | n_e(z), \alpha, I) p^a(\{a_j\}, \alpha | I). \quad (1.65)$$

Az $\{a_j\}$ és az α paraméterek legjobb becslését numerikusan határoztuk meg a poszterior eloszlás logaritmusának maximalizálásával.



1.6. ábra. (a) A csúcs helyének $p(x_0|\{d_k\}, I)$ marginális eloszlása. (b) A csúcs amplitúdójának $p(A|\{d_k\}, I)$ marginális eloszlása, folytonos vonal: egyenletes prior – szaggatott vonal: Jeffreys prior alkalmazása esetén.

A feldolgozást egy numerikus szimuláción keresztül mutatjuk meg, azaz egy feltételezett elektronsűrűség profilt használva kiszámoltuk a mérési adatokat (a mérés hibáját Gauss zaj hozzáadásával modelleztük), majd ezen adatokat felhasználva végeztük el a kiértékelést. Az elektronsűrűség profil és az $S(z)$ fényprofil a 1.7. ábrán láthatóak.

A 1.8.(a) ábrán a meghatározott az $\{a_j\}$ paraméterek legjobb becsléséből meghatározott elektronsűrűség profil látható. A profil parametrizálásához használt pontokat a \diamond jel mutatja. A 1.8.(b) ábra a mérési adatokat és a visszaállított sűrűségprofilból számolt zaj nélküli mérést mutatja.

Látható, hogy profil fizikailag nem várt oszcilláló jellegű, mivel egyenletes priort használva nem vettük figyelembe a megoldás fizikailag elvárható tulajdonságait. Megmutatható, hogy ezen probléma esetén a mérési adatok egy halmaza alig függ a sűrűségprofiltól. Esetünkben két feltétel adható meg a profil fizikailag elvárt tulajdonságára: az első, hogy a profil általában monoton növekszik a nyaláb mentén. A második: hogy a profil lokálisan sima. Regularizációval ezeket a tulajdonságokat figyelembe tudjuk venni a kiértékelés során.

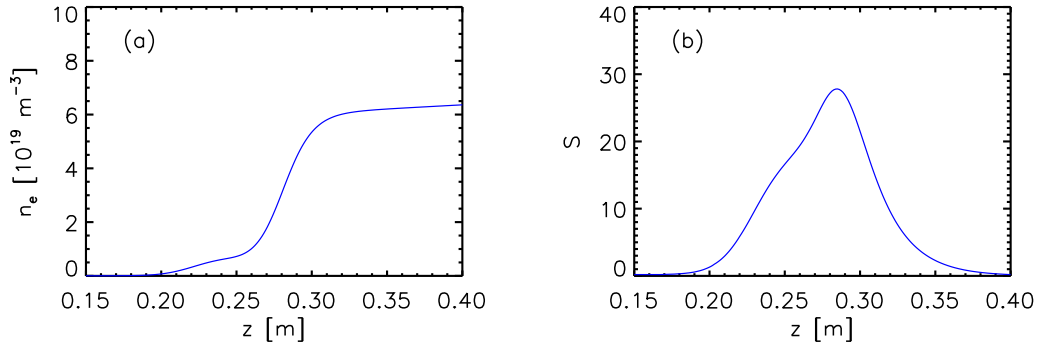
Az első feltétel a

$$p^m(\{a_j\}|I) = \exp \left\{ -R^m \left[\sum_i \frac{(a_{i+1} - a_i)^2}{(a_{i+1} + a_i)^2} (\text{sgn}(a_i - a_{i+1}) + 1) \right] \right\}, \quad (1.66)$$

a második feltétel a

$$p^s(\{a_j\}|I) = \exp \left\{ -R^s \int \left[\frac{d^2 \hat{n}_e(z)}{dz^2} \right]^2 dz \right\} \quad (1.67)$$

prior alkalmazásával érhető el. Itt $\hat{n}_e(z)$ a $\{a_j\}$ paraméterek által meghatározott profil.



1.7. ábra. (a) Az $n_e(z)$ elektronsűrűség profil és (b) az $S(z)$ fényprofil.

A Bayes-tételnek megfelelően a poszterior eloszlás

$$p(\{a_j\}, \alpha | \{d_k\}, I) \propto p(\{d_k\} | \hat{n}_e(z), \alpha, I) p^a(\{a_j\}, \alpha | I) p^m(\{a_j\} | I) p^s(\{a_j\} | I). \quad (1.68)$$

A 1.68. egyenlet szerinti poszterior eloszlás logaritmusát véve belátható, hogy R^m és R^s a 1.51. egyenletnek megfelelően regularizációs paraméterek. Ezen paraméterek alkalmas megválasztásával elérhető, hogy a meghatározott eloszlás kompatibilis legyen a mért adatokkal és fizikailag elvárható tulajdonságú legyen. A 1.9. ábrán az így meghatározott profil látható. Megállapíthatjuk, hogy a regularizáció lehetővé tette fizikailag értelmes profil meghatározását, annak ellenére, hogy mind regularizáció nélkül, mind regularizációval a mért adatok és a visszaállított zaj nélküli mérés a hibahatárokon belül megegyezik, azaz mindkét visszaállított sűrűségprofil összeegyeztethető a mérési adatokkal.

1.12. Feladatok

1.12.1. 1. feladat

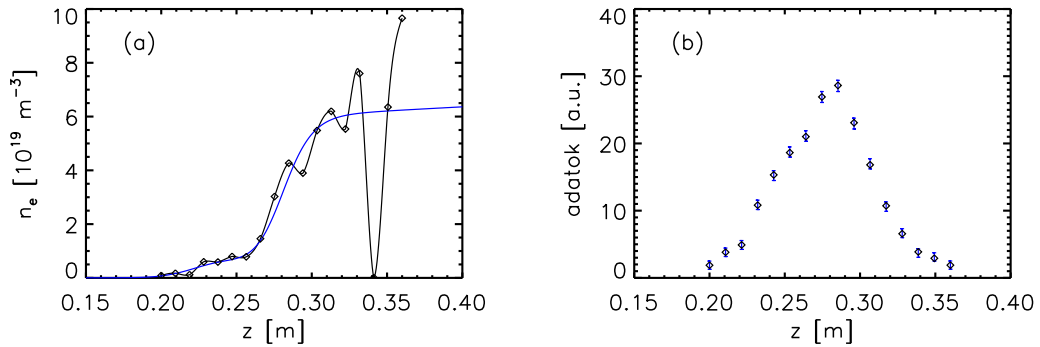
Az adatfeldolgozás során felmerülhet az a feladat, hogy ugyanazon α paraméter becsléséhez több mérési adathalmaz is rendelkezésünkre áll. Az egyszerűség miatt tekintsünk csak két adathalmazt, jelöljük ezeket \mathbf{D}_1 és \mathbf{D}_2 -vel. A Bayes-tétel szerint az α paraméter poszterior eloszlása a következő:

$$p(\alpha | \mathbf{D}_2, \mathbf{D}_1, I) \propto p(\mathbf{D}_2, \mathbf{D}_1 | \alpha, I) \cdot p(\alpha | I). \quad (1.69)$$

Ebben az esetben az adatokat úgy tekintjük mint egységes adathalmazt. Mutassa meg, hogy

$$p(\alpha | \mathbf{D}_2, \mathbf{D}_1, I) \propto p(\mathbf{D}_2 | \alpha, I) \cdot p(\alpha | \mathbf{D}_1, I) ! \quad (1.70)$$

Mi ezen egyenlet értelme, és milyen feltételek esetén alkalmazható?



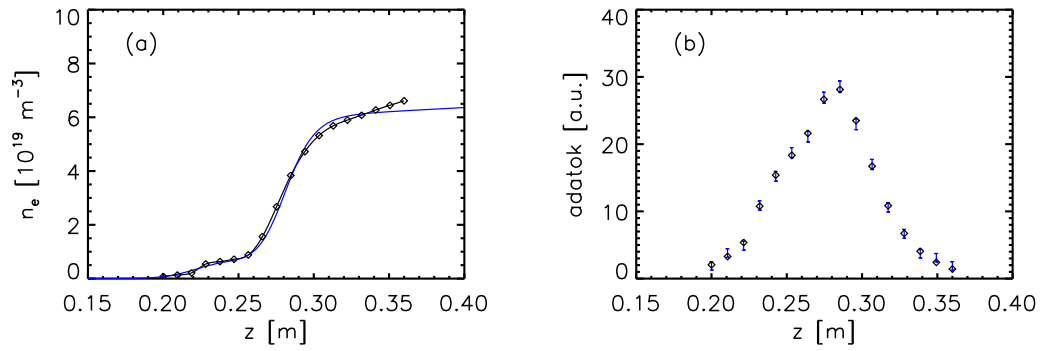
1.8. ábra. (a) A meghatározott sűrűségprofil regularizáció nélkül. A kék vonal a numerikus szimulációnál használt eredeti profil. A \diamond a profil parametrizálásához használt interpolációs pontok. (b) A mért adatok és a visszaállított sűrűségprofilból számolt zaj nélküli mérés (\diamond).

1.12.2. 2. feladat

A 1.8.1. fejezetben tárgyaltaknak megfelelően határozza meg a μ legjobb becslését abban az esetben, ha a mérések hibája nem azonos, azaz a d_k adat hibája σ_k paraméterű normális eloszlással jellemezhető!

1.12.3. 3. feladat

A 1.34. egyenletből kiindulva vezesse le az 1.35. kifejezéseket!



1.9. ábra. (a) A meghatározott sűrűségprofil regularizáció alkalmazásával. A kék vonal a numerikus szimulációnál használt eredeti profil. A \diamond a profil parametrizálásához használt interpolációs pontok. (b) A mért adatok és a visszaállított sűrűségprofilból számolt zaj nélküli mérés (\diamond).

Irodalomjegyzék

- [Dreier] H. Dreier et al.: Bayesian design of plasma diagnostics Rev. Sci. Instrum. 77, 10F323 (2006)
- [IDA] A. Dinklage et al.: Integrated Approaches in Fusion Data Analysis, AIP Conference Proceedings / Volume 735
- [SIVIA] D. S. Sivia and J. Skilling: Data Analysis – a Bayesian Tutorial, Plenum Press, New York, 2006
- [GREGORY] Phil Gregory: Bayesian Data Analysis for the Physical Sciences, Cambridge University Press, 2005
- [PUSZTAI] I. Pusztai et al.: Deconvolution-based correction of alkali beam emission spectroscopy density profile measurements, Rev. Sci. Instrum. 80, 083502 (2009)
- [FISCHER] R. Fischer et al.: Probabilistic lithium beam data analysis, Plasma Phys. Control. Fusion 50 (2008) 085009.