STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1958 - 009
S 234 (V 20)

Het maximaliseren van een functie in een zeker convex gebied

Constance van Eeden

Voordracht in de serie Actualiteiten

29 maart 1958

1958

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

ZW 1958-009
S 234 (V 20)

Het maximaliseren van een functie in een zeker convex gebied

door
Constance van Eeden

Voordracht in de serie Actualiteiten

29 maart 1958

We beschouwen eerst (aan de hand van een voorbeeld) een speciaal geval van het te behandelen probleem.
Laten $\underline{a}_1$ en $\underline{a}_2$ onderling onafhankelijke binomiaal verdeelde stochastische grootheden voorstellen met

(1) $$P[\underline{a}_i = a] = \binom{n_i}{a}\theta_i^a(1-\theta_i)^{n_i-a} \qquad (a=0,\dots,n_i;\ i=1,2).$$

Laat verder bekend zijn dat de kansen θ_1 en θ_2 voldoen aan de ongelijkheid

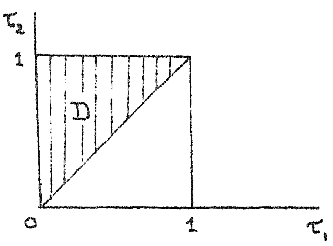
(2) $$\theta_1 \leqq \theta_2 .$$

Gevraagd wordt op grond van deze gegevens de aannemelijkste schattingen t_1 en t_2 van θ_1 en θ_2 te bepalen.
Dit probleem komt neer op het maximaliseren van de aannemelijkheidsfunctie

(3) $$L(\tau_1,\tau_2) \overset{\text{def}}{=} \sum_{i=1}^{2}\{a_i \ln \tau_i + (n_i - a_i)\ln(1-\tau_i)\}$$

in het gebied

(4) $$D \overset{\text{def}}{=} \begin{cases} \tau_1 \leqq \tau_2 \\ 0 \leqq \tau_i \leqq 1 \quad (i=1,2) \end{cases}$$, dus



Dit probleem kan op de volgende manieren gegeneraliseerd worden:

1. door meer dan twee parameters te beschouwen, dus b.v. k $(k>2)$ parameters, die voldoen aan de ongelijkheden

----

1) Stochastische grootheden worden onderscheiden van getallen (b.v. van de waarde die zij bij een experiment aannemen) door hun symbolen te onderstrepen.

(5) $$\theta_1 \leqq \ldots \leqq \theta_k,$$

2. door niet alleen (zoals in (5)) een volledige, maar ook een onvolledige ordering van $\theta_1, \ldots, \theta_k$ te beschouwen. Dus b.v. 3 parameters, die voldoen aan de ongelijkheden

(6) $$\begin{cases} \theta_1 \leqq \theta_2, \\ \theta_1 \leqq \theta_3, \end{cases}$$

3. door bovendien ongelijkheden van de vorm $c_i \leqq \theta_i \leqq d_i$ te beschouwen,

4. door parameters van andere verdelingen te beschouwen, b.v.

a. het gemiddelde van een normale verdeling,

b. de variantie van een normale verdeling,

c. de parameter van een exponentiële verdeling,

d. de parameter van een Poisson-verdeling,

e. de lengte van het interval van een homogene verdeling,

Dit algemene probleem kan als volgt geformuleerd worden. Laten $\underline{x}_1, \ldots, \underline{x}_k$ onderling onafhankelijk stochastische grootheden voorstellen en laten $x_{i,\gamma}$ $(\gamma = 1, \ldots, n_i)$ onderling onafhankelijke waarnemingen van $\underline{x}_i$ zijn $(i = 1, \ldots, k)$. Laat verder, voor iedere $i = 1, \ldots, k$, de verdeling van $\underline{x}_i$ één onbekende parameter θ_i bevatten en laten de parameters $\theta_1, \ldots, \theta_k$ voldoen aan

(7) 1. een aantal niet strijdige ongelijkheden van de vorm $\theta_i \leqq \theta_j$,
2. een aantal, met (7;1) niet strijdige, ongelijkheden van de vorm $c_i \leqq \theta_i \leqq d_i$.

Gevraagd wordt op grond van deze gegevens de aannemelijkste schattingen $t_1, \ldots, t_k$ van $\theta_1, \ldots, \theta_k$ te bepalen.

De ongelijkheden (7;1) kan men schrijven in de vorm

(8) $$\alpha_{i,j}(\theta_i - \theta_j) \leqq 0 \qquad (i,j = 1, \ldots, k),$$

waarbij

(9) 1. $\alpha_{i,j} = -\alpha_{j,i}$,
2. $\alpha_{i,j} = 0$, $+1$ of -1 voor ieder paar (i,j),
3. $\alpha_{i,j} = 1$ dan en slechts dan als θ_i en θ_j voldoen aan de ongelijkheid $\theta_i \leqq \theta_j$.

Het interval $[c_i, d_i]$ geven we aan met $\mathcal{I}_i$ en $L_i(\tau)$ stelt de aannemelijkheidsfunctie voor de i^e steekproef voor $(i=1,\ldots,k)$. Het probleem komt dan neer op het maximaliseren van de functie

(10) $$L = L(\tau_1,\ldots,\tau_k) \overset{def}{=} \sum_{i=1}^{k} L_i(\tau_i)$$

in het gebied

(11) $$D \overset{def}{=} \begin{cases} \alpha_{i,j}(\tau_i - \tau_j) \leq 0, \\ \tau_i \in \mathcal{I}_i \end{cases} \qquad (i,j=1,\ldots,k).$$

We zullen nu voorwaarden geven waaronder dit maximum bestaat en uniek is; verder beschrijven we een methode met behulp waarvan dit maximum gevonden kan worden.
Laat, voor een deelverzameling M van de getallen $\{1,\ldots,k\}$,

(12) $$\mathcal{I}_M \overset{def}{=} \bigcap_{i\in M} \mathcal{I}_i$$

en, voor $\mathcal{I}_M \neq 0$,

(13) $$L_M(\xi) \overset{def}{=} \sum_{i\in M} L_i(\xi)$$

dan onderstellen we dat de volgende voorwaarde vervuld is
Voorwaarde A : De functie $L_M(\xi)$ is, voor iedere M met $\mathcal{I}_M \neq 0$, streng unimodaal in $\mathcal{I}_M$, d.w.z.

1. $L_M(\xi)$ bezit een uniek maximum in $\mathcal{I}_M$, dat bereikt wordt in, stel, het punt v_M,
2. $L_M(\xi)$ is $\begin{cases} \text{a. monotoon stijgend voor } \xi < v_M, \\ \text{b. monotoon dalend voor } \xi > v_M. \end{cases}$

Als voorwaarde A vervuld is dan geldt
Stelling 1 : De functie $L(\tau_1,\ldots,\tau_k)$ bezit een uniek maximum in het gebied D.
Het punt waar dit maximum bereikt wordt geven we aan met $t_1,\ldots,t_k$.
Voor de ongelijkheden $\tau_i \leq \tau_j$ (d.w.z. $\alpha_{i,j}=1$) kan men de volgende twee gevallen onderscheiden:

1. er is een h waarvoor beide ongelijkheden $\tau_i \leq \tau_h$ en $\tau_h \leq \tau_j$ gelden. De ongelijkheid $\tau_i \leq \tau_j$ volgt dan uit deze twee ongelijkheden.

2. er is geen k waarvoor beide ongelijkheden $\tau_i \leq \tau_k$ en $\tau_k \leq \tau_j$ gelden.

De ongelijkheden $\tau_i \leq \tau_j$ waarvoor 1 geldt noemen we de niet-essentiële ongelijkheden; de andere noemen we de essentiële ongelijkheden. Bij een volledige ordening ($\tau_1 \leq \ldots \leq \tau_k$) b.v. zijn de ongelijkheden $\tau_i \leq \tau_j$ voor $j > i+1$ de niet-essentiële ongelijkheden en $\tau_i \leq \tau_{i+1}$ zijn de essentiële.

De essentiële ongelijkheden geven we aan met $R_1, \ldots, R_s$ en de bij R_λ behorende indices i en j geven we aan met i_λ en j_λ.

Dan geldt

<u>Stelling 2 : Als D' het gebied is dat verkregen wordt uit D door één essentiële ongelijkheid, stel R_λ, weg te laten en als $(t'_1, \ldots, t'_k)$ het punt voorstelt waar L zijn maximum bereikt in D' dan is</u>

$$(14) \quad \begin{cases} 1.\ t_i = t'_i \ (i = 1, \ldots, k) & \text{als } (t'_1, \ldots, t'_k) \in D, \\ 2.\ t_{i_\lambda} = t_{j_\lambda} & \text{als } (t'_1, \ldots, t'_k) \notin D. \end{cases}$$

Deze stelling brengt het probleem terug tot de volgende twee problemen:

1. het maximaliseren van L in een gebied met $s-1$ essentiële ongelijkheden,

2. het maximaliseren van L in de doorsnede van D met de diagonaal $\tau_{i_\lambda} = \tau_{j_\lambda}$, dus het maximaliseren van de functie

$$(15) \qquad \sum_{\substack{h \in E \\ h \neq i_\lambda,\, h \neq j_\lambda}} L_h(\tau_h) + L_{\{i_\lambda, j_\lambda\}}(\tau_{i_\lambda})$$

in een gebied met hoogstens $s-1$ essentiële ongelijkheden. Om het eerste probleem op te lossen kan men weer stelling 2 toepassen door een essentiële ongelijkheid van D' weg te laten. Voor het tweede probleem past men stelling 2 toe door een essentiële ongelijkheid van de doorsnede van D met de diagonaal $\tau_{i_\lambda} = \tau_{j_\lambda}$ weg te laten. Door herhaald toepassen van stelling 2 komt men dus tot het probleem van het maximaliseren van de functie

$$(16) \qquad \sum_{\nu=1}^{N} L_{M_\nu}(\xi_\nu)$$

in een gebied zonder essentiële ongelijkheden, dus in het gebied

$$G_N \overset{\text{def}}{=} \prod_{\nu=1}^{N} \mathcal{Y}_{M_\nu}, \tag{17}$$

waarbij $M_1, \ldots, M_N$ deelverzamelingen van $\{1,\ldots,k\}$ zijn met

$$\left\{\begin{array}{ll} 1.\ \bigcup_{\nu=1}^{N} M_\nu = \{1,\ldots,k\}, & \\ 2.\ M_\nu \cap M_\mu = 0 & \text{voor ieder paar } (\nu,\mu) \text{ met } \nu \neq \mu, \\ 3.\ \mathcal{Y}_{M_\nu} = 0 & \text{voor iedere } \nu. \end{array}\right. \tag{18}$$

De functie (16) bereikt zijn maximum in het gebied G_N in het punt $(v_{M_1}, \ldots, v_{M_N})$ (zie voorwaarde A).

Voor de coördinaten van het punt waar L zijn maximum in D bereikt (dus voor $t_1, \ldots, t_k$) kan ook een expliciete uitdrukking gegeven worden. Laat

$$\left\{\begin{array}{l} S_i \overset{\text{def}}{=} i \cup \mathcal{E}ns\{j \mid \alpha_{j,i} = 1\}, \\ T_i \overset{\text{def}}{=} i \cup \mathcal{E}ns\{j \mid \alpha_{i,j} = 1\} \end{array}\right. \tag{19}$$

en laat, voor een deelverzameling M van $\{1,\ldots,k\}$,

$$\left\{\begin{array}{l} S_M \overset{\text{def}}{=} \bigcup_{i\in M} S_i, \\ T_M \overset{\text{def}}{=} \bigcup_{i\in M} T_i, \end{array}\right. \tag{20}$$

dan is

$$t_i = \max_{M} \min_{M'} \{ v_{T_M \cap S_{M'}} \mid i \in T_M \cap S_{M'} \}. \tag{21}$$

Bij een volledige ordening b.v. bestaat S_i uit de getallen $1,\ldots,i$ en T_i uit de getallen $i, \ldots, k$. Dus als

$$\left\{\begin{array}{l} r_1 \overset{\text{def}}{=} \max_{i\in M'} i, \\ r_2 \overset{\text{def}}{=} \min_{i\in M} i, \end{array}\right. \tag{22}$$

dan bestaat $S_{M'}$ uit de getallen $1, \ldots, r_1$ en T_M uit de getallen $r_2, \ldots, k$. In dit geval is dus

(23) $$t_i = \max_{\tau_2} \min_{\tau_1} \{ v_{\{\tau_2,\ldots,\tau_1\}} \mid \tau_2 \leqq i \leqq \tau_1 \}.$$

Verder gelden nog de volgende stellingen

Stelling 3 : Als (i,j) een paar is met

(24) $$\begin{cases} 1. \ \alpha_{i,j} = 1, \ v_i \geqq v_j, \\ 2. \ \alpha_{i,h} = \alpha_{j,h} \quad \text{voor iedere } h \text{ met } h \neq i, h \neq j \end{cases}$$

dan bereikt L zijn maximum in D voor $\tau_i = \tau_j$.

Bij een volledige ordening geldt (24;2) voor ieder paar (i,j) met $j = i + 1$; dus in dit geval geldt

(25) $$t_i = t_{i+1} \quad \text{voor iedere } i \text{ met } v_i \geqq v_{i+1}.$$

Stelling 4 : Als (i,j) een paar is met

(26) $$\begin{cases} 1. \ \alpha_{i,j} = 0, \ v_i \geqq v_j, \\ 2. \ \alpha_{i,h} \leqq \alpha_{j,h} \quad \text{voor iedere } h \text{ met } h \neq i \text{ en } h \neq j \end{cases}$$

dan bereikt L zijn maximum in D voor $\tau_j \leqq \tau_i$.

Door toepassing van stelling 3 brengt men het probleem terug tot het maximaliseren van een som van k -1 termen met hoogstens s -1 essentiële ongelijkheden. Op grond van stelling 4 kan men de ongelijkheid $\tau_j \leqq \tau_i$ toevoegen, waardoor het probleem overgaat in het maximaliseren van L in een kleiner gebied.

Litteratuur

VAN EEDEN, CONSTANCE, Maximum likelihood estimation of partially or completely ordered parameters, Proc. Kon.Ned.Akad.v.Wet. A 60 (1957), Indagationes Mathematicae 19 (1957) 128-136 and 201-211.

VAN EEDEN, CONSTANCE, Note on two methods for estimating ordered parameters of probability distributions, Proc. Kon.Ned.Akad.v.Wet. A 60 (1957), Indagationes Mathematicae 19 (1957) 506-512.

VAN EEDEN, CONSTANCE, A least squares inequality for maximum likelihood estimates of ordered parameters, Proc. Kon.Ned.Akad.v. Wet. A 60 (1957), Indagationes Mathematicae 19 (1957), 513-521.