STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

S 390 (M 90)

Noodzakelijkheid en afdoendheid voor families van discrete en continue kansverdelingen

door

W.N. van Nooten

maart 1968

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 the Advancement of Pure Research (Z.W.O.) and the Central Organization for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

INHOUD:

0. VOORWOORD

Deze scriptie kan worden gezien als het resultaat van een poging om, zonder expliciet gebruik van maattheoretische methoden, onder zo ruim mogelijke regulariteitsvoorwaarden, afdoendheid en noodzakelijkheid te bespreken, en een manier aan te geven om noodzakelijk en afdoende functies te vinden.

De behandeling op deze wijze wijkt in belangrijke mate af van de maattheoretische, in mindere mate van die met regulariteitsvoorwaarden zoals in DYNKIN [7], en FRASER [9] en [10]. Veel uit het derde hoofdstuk is op FRASER [10] gebaseerd.

Het tweede hoofdstuk is in het bijzonder afwijkend van datgene wat mij bekend is uit de literatuur. De aannamen in de drie genoemde artikelen ([7], [9] en [10]) zijn wat strikter. Een gevolg hiervan is dat de geleverde bewijzen soms nieuw, soms aangepaste versies van bestaande bewijzen zijn.

Niet-gedomineerde families van verdelingen zijn niet behandeld. De theorie over dergelijke families is nog niet ver ontwikkeld. Zie echter BURKHOLDER [20].

Verdelingen met een discrete en een continue component (waarbij de discrete component voor elke verdeling in de familie tot dezelfde aftelbare verzameling is beperkt) zijn op eenvoudige wijze geheel analoog te behandelen. Ze worden echter niet genoemd omdat de generalisatie vrijwel triviaal is. Zie FRASER [9].

De regulariteitsvoorwaarde die hier aan families van verdelingen van het continue type wordt opgelegd is topologisch van aard en heel eenvoudig: continue dichtheden met als drager een open deelverzameling van R^n. Samenhang van die open verzameling wordt niet geëist. Aan de dichtheden wordt geen regulariteitseis in de indexverzameling opgelegd, zoals men soms aantreft bij de behandeling van de KOOPMAN-PITMAN klasse van verdelingen. Zie hierover ook 3.1.

Naast het bekende factoriseringscriterium voor afdoendheid (het FISHER-NEYMAN criterium) wordt een tweede criterium ingevoerd en van een naam voorzien: het klassificatiecriterium. Dit criterium blijkt in veel gevallen gemakkelijker te hanteren dan het factoriseringscriterium.

1. VOORBEREIDENDE OPMERKINGEN

1.1. Functies en voortgebrachte klasse-indelingen

Zij X een verzameling met elementen x, t een functie op X. Aan het waardengebied $t(X) = \{t(x) : x \in X\}$ leggen we geen beperkingen op. Zo kan t aan elk punt $x \in X$ een vector of een functie toevoegen.

Punten x en x' in X noemen we equivalent onder t als $t(x) = t(x')$. Dit is natuurlijk een equivalentie-relatie op X en de relatie splitst de ruimte dus op in een aantal disjuncte klassen. De collectie van deze klassen heet de door t voortgebrachte klasse-indeling.
Elke functie op X brengt zo een klasse-indeling voort, maar ook wordt elke klasse-indeling door tenminste één functie voortgebracht. Kies namelijk in elke klasse D van de indeling een punt $x(D)$. De functie $t(x) = x(D)$ voor $x \in D$ voldoet dan.

Laat $\mathcal{D}$ en $\mathcal{E}$ klasse-indelingen van X zijn.

DEF. 1: $\mathcal{E}$ heet een vergroving van $\mathcal{D}$ als geldt:
behoren x en x' tot dezelfde klasse $D \in \mathcal{D}$ dan behoren ze tot eenzelfde klasse $E \in \mathcal{E}$.
$\mathcal{D}$ heet een verfijning van $\mathcal{E}$ als $\mathcal{E}$ een vergroving van $\mathcal{D}$ is.

Zijn t en u functies op X, $\mathcal{D}$ en $\mathcal{E}$ de door t en u voortgebrachte klasse-indelingen, dan is het volgende eenvoudig na te gaan:

STELLING 1:
$\mathcal{E}$ is een vergroving van $\mathcal{D}$ d.e.s.d.a.
(1) voor elk paar x,x' in X geldt: $t(x) = t(x') \implies u(x) = u(x')$,
en ook, d.e.s.d.a.
(2) er een functie ϕ op $t(X)$ bestaat zó dat $\phi(t(x)) = u(x)$ op X.

We kunnen dus zeggen: u is te schrijven als functie van t. Zijn u en t als functie van elkaar te schrijven, dan is $\mathcal{E} = \mathcal{D}$, We noemen u en t dan equivalente functies.

VOORBEELDEN:

1. Zij $X = R^1$. De functies $|x|$ en x^2 zijn equivalent; de bijbehorende klasse-indeling bestaat uit $\{0\}$ en voor alle $a > 0$ de klassen $\{-a,a\}$.
2. Zij $t = (t_1, t_2, \dots, t_r)$ een vectorwaardige functie op een verzameling X, dan is elke component t_i $(1 \leq i \leq r)$ te schrijven als functie van t; t_i brengt dus een niet-fijnere klasse-indeling voort dan t. Hetzelfde geldt voor functies $t = \{t_\lambda : \lambda \in \Lambda\}$ op X, bij willekeurige index-verzameling Λ. Elke t_λ brengt een niet-fijnere klasse-indeling voort dan t.
3. Zij $U = \{u_\lambda : \lambda \in \Lambda\}$ een lineaire ruimte van functies op een verzameling X. Een lineair onafhankelijke basis voor deze ruimte is equivalent met U, als functie op X beschouwd.
4. Zij $x = (x_1, x_2, \dots, x_n) \in R^n$. Laat $t(x) = (t_1(x), t_2(x), \dots, t_n(x))$ een functie op X met zijn continue partiële afgeleiden naar $x_1, x_2, \dots, x_n$ in een omgeving van $x_0 \in X$ en met determinant van Jacobi $J = \left| \left(\frac{\partial t_i}{\partial x_j}\right)_{\substack{i=1,2,\dots,n \\ j=1,2,\dots,n}} \right| \neq 0$ in x_0. In een omgeving van x_0 is dan x equivalent met t, d.w.z. t heeft in een omgeving van x_0 een eenduidig bepaalde inverse. (Zie bv. COURANT [6], p. 142 vv.)

Belangrijk is het nu volgende voorbeeld. We laten het voorafgaan door een lemma, dat voor het bewijs nodig is.

Definiëer de volgende functies:

$t(x) = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$, waarin $x \in R^n$, $x_{(1)} \leqq x_{(2)} \leqq \dots \leqq x_{(n)}$.

$$u(x) = \left(\sum_{i=1}^n x_i,\ \sum_{i=1}^n x_i^2,\ \dots,\ \sum_{i=1}^n x_i^n \right).$$

$$v(x) = \left(\sum_{i=1}^n x_i,\ \sum_{i_1<i_2} x_{i_1} x_{i_2},\ \sum_{i_1<i_2<i_3} x_{i_1} x_{i_2} x_{i_3},\ \dots,\ x_1 x_2 \dots x_n \right).$$

LEMMA:

Tussen de functies u en v bestaan de volgende relaties, de <u>identiteiten van Newton</u>, voor $k \leqq n$, als u_i en v_i de i-de component van u en v aangeven:

$$u_k - u_{k-1}v_1 + u_{k-2}v_2 - \dots + (-)^{k-1}u_1v_{k-1} + (-)^k kv_k = 0 \quad \dots\dots \text{(a)}.$$

BEWIJS:

$$u_l v_{k-l} = a_l + a_{l+1}, \text{ waarin } a_h = \sum_i x_i^h \sum_{\substack{i_1<i_2<\dots<i_{k-h} \\ i_j \neq i}} x_{i_1}x_{i_2}\dots x_{i_{k-h}}$$

en $l < k-1$;

$u_{k-1}v_1 = a_{k-1} + u_k$.

Met behulp hiervan blijkt:

$$kv_k = k \sum_{i_1<\dots<i_k} x_{i_1}x_{i_2} \dots x_{i_k} = a_1 = u_1v_{k-1} - a_2 = \dots =$$

$$u_1v_{k-1} - u_2v_{k-2} + \dots + (-)^{k-2}u_{k-2}v_2 + (-)^{k-1}a_{k-1},$$

waaruit het gestelde volgt.

STELLING 2:

De functies t, u en v zijn equivalent.

BEWIJS:

Merk eerst op:

Voor $z \in R^1$ is

$$(z-x_1)(z-x_2)\dots(z-x_k) = z^k - z^{k-1}v_1 + z^{k-2}v_2 - \dots + (-)^{k-1}zv_{k-1} + \\ + (-)^k v_k \quad \dots\dots \text{(b)}.$$

1. Zijn x en x' equivalent onder t, dan eveneens onder u, want $t(x) = t(x')$ houdt in dat x en x' dezelfde componenten hebben, waarbij slechts de volgorde kan verschillen. Elke functie die symmetrisch is in deze componenten heeft dan dezelfde waarde in x als in x'. In het bijzonder geldt dit voor u.
2. Zijn x en x' equivalent onder u, dan ook onder v. Nu is nl. $u(x) = u(x')$. Uit de identiteiten van Newton blijkt:

$u_1(x) = u_1(x') = v_1(x) = v_1(x')$.
Stel $v_l(x) = v_l(x')$ voor $1 < k \leqq n$, dan volgt uit (a):
$v_k(x) = v_k(x')$. Dus $v(x) = v(x')$ en x is equivalent met x'
onder v.

3. Zijn x en x' equivalent onder v, dan ook onder t. Zij $v(x) = v(x')$. De nulpunten van (b) stemmen dan voor $v(x)$ en $v(x')$ overeen, in waarde en multipliciteit. Het zijn $x_1, x_2, \ldots, x_n$. Dit wil juist zeggen dat x en x' dezelfde componenten hebben, zodat x en x' equivalent zijn onder t.

1.2. Lineaire ruimten

Uitgangspunt is de volgende situatie:
gegeven zijn twee verzamelingen X en Θ en op $X \times \Theta$ is een reëelwaardige functie $u(x,\theta)$ gedefiniëerd. Om de gedachten te bepalen: deze functie zal in eenvoudige relatie staan tot een familie van kansverdelingen $P_\theta(\underline{x} = x)$ op X en voor $\theta \in \Theta$, of van waarschijnlijkheidsdichtheden $f(x|\theta)$ op X en voor $\theta \in \Theta$.

De functie $u(x,\theta)$ willen we nu eens beschouwen als functie op X, dus als een functie die aan elk punt $x \in X$ een functie van θ alleen toevoegt, dan weer als functie op Θ, die dus aan elk punt $\theta \in \Theta$ een functie van x alleen toevoegt. Om het accent aan te geven schrijven we in het eerste geval wel $u(x,.)$ of $\{u(x,\theta) : \theta \in \Theta\}$ op X, in het tweede geval $u(.,\theta)$ of $\{u(x,\theta) : x \in X\}$ op Θ.

We breiden $\{u(x,\theta) : \theta \in \Theta\}$ uit tot een lineaire ruimte U_x op X door alle functies $\sum_{i=1}^{s} a_i u(x,\theta_i)$ voor $s = 1, 2, \ldots$, $a_i \in R^1$, $\theta_i \in \Theta$ toe te voegen.

Op dezelfde wijze krijgt men de lineaire ruimte U_θ op Θ door sommen $\sum_{i=1}^{s} b_i u(x_i,\theta)$ met reële coëfficiënten.

Zij de rang van U_x r, die van U_θ r'. Als basis kan men natuurlijk steeds functies $u(x,\theta_\lambda)$ met $\theta_\lambda \in \Theta$ en $u(x_\mu,\theta)$ met $x_\mu \in X$ nemen.

Gemakkelijk is met behulp van 1.1. stelling 1 (1) na te gaan dat een basis voor U_x equivalent is met U_x en met $\{u(x,\theta) : \theta \in \Theta\} = u(x,.)$.
Verder geldt:

STELLING:

<u>r, de rang van U_x en r', de rang van U_θ zijn beide oneindig of beide eindig en gelijk</u>.

BEWIJS:

Als $r < \infty$ en $v_1(x), v_2(x), \ldots, v_r(x)$ een lineair onafhankelijke basis is voor U_x, dan is elke $u(x,\theta)$ te schrijven als:

$u(x,\theta) = \sum_{i=1}^{r} c_i(\theta)v_i(x)$. De functies $c_1(\theta), c_2(\theta), \ldots, c_r(\theta)$ brengen nu U_θ voort en er is een $r' \leqq r$ zo dat $c_1'(\theta), c_2'(\theta), \ldots, c_{r'}'(\theta)$ hieronder lineair onafhankelijk zijn.

Uitgaande van $r' < \infty$ blijkt evenzo dat $r \leqq r'$, zodat $r = r'$.

Tegelijkertijd is duidelijk dat als een van de rangen, bv. r oneindig is, de andere niet eindig kan zijn.

2. AFDOENDHEID EN NOODZAKELIJKHEID VAN KLASSE-INDELINGEN VOOR FAMILIES VAN KANSVERDELINGEN

2.1. De uitkomstenruimte; de drager

2.1.1. Families van discrete kansverdelingen

Onder een familie van discrete kansverdelingen verstaan we een collectie $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ van kansverdelingen, die alle gedefiniëerd zijn op dezelfde verzameling X, die hoogstens aftelbaar veel elementen bevat. Bovendien kiezen we X zó, dat voor elke $x \in X$ geldt: $P_\theta(\underline{x} = x) > 0$ voor tenminste één $\theta \in \Theta$. X heet de uitkomstenruimte; θ indiceert de kansverdelingen, Θ heet de indexverzameling en elke functie $g(\theta)$ op Θ heet een parameter van $\mathcal{P}$.

Zij $\theta \in \Theta$. De verzameling $X_\theta = \{x : P_\theta(\underline{x} = x) > 0\}$ heet de drager van P_θ. Daar voor alle $x \in X$ geldt dat $P_\theta(\underline{x} = x) > 0$ voor tenminste één θ, is $X = \bigcup_{\theta \in \Theta} X_\theta$.

Verder is er klaarblijkelijk een rij indices $\theta_1, \theta_2, \ldots$ in Θ, zó dat $X = \bigcup_{k=1}^{\infty} X_{\theta_k}$.

2.1.2. Families van kansverdelingen met een dichtheid

Voor families van kansverdelingen met een dichtheid kiezen we collecties $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ van waarschijnlijkheidsmaten, elk gedefiniëerd op de (meetbare) deelverzamelingen van een verzameling $X \subset R^n$, en die voldoen aan de volgende regulariteitsvoorwaarde:

REGULARITEITSVOORWAARDE R:

Bij elke P_θ, $\theta \in \Theta$, is er een functie $f(x|\theta)$, eindigwaardig en niet-negatief op X, zó dat voor elke meetbare verzameling $A \subset X$ geldt:

$$\int_A f(x|\theta)dx = P_\theta(A).$$

De dragers $X_\theta = \{x : f(x|\theta) > 0\}$ moeten open deelverzamelingen van X zijn.

Ook nu kiezen we X als $\bigcup_{\theta \in \Theta} X_\theta$, zodat X open is en voor alle $x \in X$ geldt: er is een $\theta \in \Theta$ zó dat $f(x|\theta) > 0$.

Bovendien is X te schrijven als $\bigcup_{k=1}^{\infty} X_{\theta_k}$ voor een geschikte rij indices $\theta_1, \theta_2, \ldots$ in Θ. Dit is de Lindelöf-eigenschap voor tweede axiomaruimten (zie bv. PERVIN [16], p. 81).

2.1.3. Algemene opmerkingen

In de voorbeelden zal men bij beide typen familie de ruimte X vaak kunnen opvatten als de produktruimte $\prod_{i=1}^{n} X_i$ van n kopieën van eenzelfde ruimte. Dat is bijvoorbeeld het geval als X de uitkomstenruimte is bij een steekproef van de grootte n van identiek verdeelde grootheden. Dit wordt echter niet geëist.

Is $X_\theta = X$ voor alle $\theta \in \Theta$, dan spreken we van een familie met constante drager, of met een drager die niet van θ afhangt. Is $X_\theta \neq X$ voor een $\theta \in \Theta$ dan zeggen we dat de drager van θ afhangt.

Aan de regulariteitsvoorwaarde R voldoen de bekende families van continue verdelingen zoals normale verdelingen, Γ- en B-verdelingen, Cauchy-verdelingen; verder de homogene verdelingen op $(0,\theta)$, $\theta > 0$, op (θ_1,θ_2), $-\infty < \theta_1 < \theta_2 < \infty$, exponentiële en lognormale verdelingen met of zonder variabel beginpunt (op (a,∞)) en alle families van n-voudige produkten van deze families.

In het vervolg willen we beschikken over disjuncte splitsingen van de ruimte X van het volgende type:

DEF 1:

Als $X_{\theta_1}, X_{\theta_2}, \ldots$ een rij dragers is die tezamen X overdekken en als voor $k = 1, 2, \ldots$ $X_k^* = X_{\theta_k} \setminus (\bigcup_{i=1}^{k-1} X_{\theta_i})$, dan heet $\{X_1^*, X_2^*, \ldots\}$ een disjuncte splitsing van X door dragers.

Op grond van de opmerkingen aan het eind van 2.1.1 en 2.1.2 is de definitie zinvol. Duidelijk is dat de verzamelingen X_k^* disjunct zijn en X overdekken.
Onder de X_k^* kunnen lege verzamelingen voorkomen; dit is in het bijzonder steeds het geval als de dragers niet van θ afhangen. De enige disjuncte splitsing van X door dragers is dan $\{X, \phi, \phi, \ldots\}$.

Zij $\mathcal{P}$ een collectie van waarschijnlijkheidsmaten van het discrete type of met dichtheden die aan R voldoen: $\{P_\theta : \theta \in \Theta\}$ of $\{f(x|\theta) : \theta \in \Theta\}$.
Zij verder $\{X_1^*, X_2^*, \ldots\}$ een disjuncte splitsing van X door dragers.
Dan geldt:

STELLING 1:

Zijn x en $x' \in X$ zó, dat voor alle θ $P_\theta(\underline{x} = x) > 0$ d.e.s.d.a. $P_\theta(\underline{x} = x') > 0$ (of $f(x|\theta) > 0$ d.e.s.d.a. $f(x'|\theta) > 0$), dan is er een k met x en $x' \in X_k^*$.

BEWIJS:

Uit de definitie van dragers X_θ volgt direct dat voor alle $\theta \in \Theta$, x en x' hetzij beide in X_θ òf beide buiten X_θ liggen.
Zij nu $x \in X_k^*$. Dan is $x \in X_{\theta_k}$ en $x \notin X_{\theta_i}$ voor $i = 1, 2, \ldots, k-1$.
Hetzelfde geldt dan eveneens voor x', zodat $x' \in X_k^*$.

2.2. De functies $L(.|x)$ en $l(.|x)$

Voor elke $\theta \in \Theta$ hebben we een kansverdeling $P_\theta(\underline{x} = x)$, $x \in X$ of een dichtheid $f(x|\theta)$, $x \in X$. Beschouw de hele familie $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ of $\mathcal{F} = \{f(x|\theta) : \theta \in \Theta\}$ als één functie van x. De waarde die bijvoorbeeld $\mathcal{F}$ in $x_0 \in X$ aanneemt is dan de parameter $f(x_0|\theta)$, $\theta \in \Theta$. Om er de nadruk op te leggen dat zo'n functie $\mathcal{F} = \{f(x|\theta) : \theta \in \Theta\}$ als functie van x wordt beschouwd schrijven we: $f(x|.)$.

DEF 1:

De <u>aannemelijkheidsfunctie</u> $L(.|x)$ wordt gegeven door:
$L(\theta|x) = \log P_\theta(\underline{x} = x)$, $\theta \in \Theta$ of door
$L(\theta|x) = \log f(x|\theta)$, $\theta \in \Theta$,
naar gelang van het type van de beschouwde familie.
Voor $L(\theta|x)$ wordt de waarde $-\infty$ toegelaten.

Kies vervolgens een rij dragers $X_{\theta_1}, X_{\theta_2}, \ldots$ die X overdekken en vorm daaruit de disjuncte splitsing van X door dragers $\{X_1^*, X_2^*, \ldots\}$, zoals aangegeven in 2.1.3. def. 1.

DEF 2:

$$P_{\theta^*}(\underline{x} = x) = P_{\theta_h}(\underline{x} = x) \text{ als } x \in X_h^* \text{ of}$$

$$f(x|\theta^*) = f(x|\theta_h) \text{ als } x \in X_h^* \text{ voor } h = 1, 2, \ldots .$$

$$L(\theta^*|x) = \log P_{\theta^*}(\underline{x} = x) \text{ resp.}$$

$$L(\theta^*|x) = \log f(x|\theta^*).$$

GEVOLG 1:

$P_{\theta^*}(\underline{x} = x)$ en $f(x|\theta^*)$ zijn positieve functies op X.

$L(\theta^*|x)$ is een eindigwaardige functie op X.

OPMERKING 1:

De schrijfwijze $P_{\theta^*}(\underline{x} = x)$ en $f(x|\theta^*)$ is zuiver formeel. Er is in het algemeen geen sprake van een kansverdeling of van een dichtheid.

DEF 3:

$l(.|x)$, het <u>aannemelijkheidsverschil</u>, wordt gedefiniëerd als

$l(.|x) = L(.|x) - L(\theta^*|x)$.

OPMERKING 2:

Op X geldt:

$P_.(\underline{x} = x) = P_{\theta^*}(\underline{x} = x) \cdot \exp[l(.|x)]$ in het discrete, en

$f(x|.) = f(x|\theta^*) \cdot \exp[l(.|x)]$ in het continue geval.

OPMERKING 3:

Is $x = (x_1, x_2, \ldots, x_n)$ een steekproef uit verdelingen $P_\theta^{(1)}(\underline{x}_1 = x_1), \ldots, P_\theta^{(n)}(\underline{x}_n = x_n)$, dan wel uit verdelingen met dichtheden $f^{(1)}(x_1|\theta), \ldots, f^{(n)}(x_n|\theta)$, waarbij θ een vaste maar onbekende waarde uit een indexverzameling Θ is, dan is

$$l(.|x) = \sum_{i=1}^{n} l^{(i)}(.|x_i), \text{ waarbij } l^{(i)}(.|x) =$$

$$= \log \frac{P_.^{(i)}(\underline{x}_i = x_i)}{P_{\theta^*}^{(i)}(\underline{x}_i = x_i)} \text{ of } l^{(i)}(.|x) = \log \frac{f^{(i)}(x_i|.)}{f^{(i)}(x_i|\theta^*)} .$$

Dit heeft tot gevolg, dat men in het geval van een steekproef slechts de marginale verdelingen hoeft te betrekken in een onderzoek van het aannemelijkheidsverschil.

2.3. Voorwaardelijke waarschijnlijkheid op een eindige verzameling

Zij $P(\underline{x} = x)$ een discrete kansverdeling, $f(x)$ een dichtheid, continu op een open verzameling en daarbuiten 0. Laat E een eindige deelverzameling van X zijn en wel zó dat $P(\underline{x} = x) > 0$ voor een $x \in E$, dan wel $f(x) > 0$ voor een $x \in E$.

DEF 1:

De voorwaardelijke waarschijnlijkheid $P(\underline{x} = x|E)$ van x gegeven E wordt gedefiniëerd als:

$$P(\underline{x} = x|E) = \frac{P(\underline{x} = x)}{P(E)} \quad \text{of als} \quad \frac{f(x)}{\sum_{x_i \in E} f(x_i)} \quad \text{voor } x \in E,$$

$$P(\underline{x} = x|E) = 0 \qquad\qquad \text{voor } x \notin E.$$

TOELICHTING:

Zij in het continue geval met $d(x,y)$ de afstand tussen x en y in X aangegeven. Men kan nu bij de punten x en x_i in E bolvormige omgevingen $B_\varepsilon(x) = \{y : d(x,y) < \varepsilon\}$ en $B_\varepsilon(x_i)$ op dezelfde wijze nemen. $P(\underline{x} = x|E)$ is dan te beschouwen als de limiet voor $\varepsilon \to 0$ van

$$\int_{B_\varepsilon(x)} f(y)dy / \Big(\sum_{x_i \in E} \int_{B_\varepsilon(x_i)} f(y)dy \Big).$$ Dit is voor elke $\varepsilon > 0$ een "gewone" voorwaardelijke waarschijnlijkheid.

Het is duidelijk dat ook in het continue geval $P(\underline{x} = x|E)$ zich op E als een discrete kansverdeling gedraagt. Dit rechtvaardigt tevens de notatie.

Is er geen $x \in E$ met $P(\underline{x} = x) > 0$ of $f(x) > 0$ dan is $P(\underline{x} = x|E)$ niet gedefiniëerd. Hebben we een familie van kansverdelingen $\mathcal{P}$, of van dichtheden $\mathcal{F}$ die aan R voldoen, dan is er voor elke eindige verzameling $E \subset X$ tenminste één $\theta \in \Theta$, waarvoor $P_\theta(\underline{x} = x|E)$ gedefiniëerd is. Hiervan maken we in het volgende gebruik.

2.4. Afdoendheid en criteria voor afdoendheid

Zij $\mathcal{P}$ een familie van discrete verzamelingen, $\mathcal{F}$ een familie van dichtheden die aan R voldoen. X is weer de uitkomstenruimte, Θ de indexverzameling.

DEF 1:

Een klasse-indeling $\mathcal{D}$ van X heet afdoende voor $\mathcal{P}$ of voor $\mathcal{F}$ als voor elke $D \in \mathcal{D}$ en elke eindige verzameling $E \subset D$ geldt:
$P_\theta(\underline{x} = x|E)$ hangt niet van θ af in de zin dat voor alle θ waarvoor $P_\theta(\underline{x} = x|E)$ gedefiniëerd is deze functie dezelfde is.

DEF 1a:

Een functie t op X heet afdoende voor $\mathcal{P}$ (voor $\mathcal{F}$) als de door t voortgebrachte klasse-indeling afdoende is voor $\mathcal{P}$ resp. $\mathcal{F}$.

GEVOLGEN:

1. Is $\mathcal{D}$ afdoende voor $\mathcal{P}$ (voor $\mathcal{F}$) en is $D \in \mathcal{D}$, dan geldt voor alle $\theta \in \Theta$ hetzij $D \subset X_\theta$ of $D \subset X'_\theta = X \setminus X_\theta$.
   Anders gezegd: òf $P_\theta(\underline{x} = x) > 0$ voor alle $x \in D$ òf $P_\theta(\underline{x} = x) = 0$ voor alle $x \in D$, en analoog voor $f(x|\theta)$.

   BEWIJS:

   We geven het bewijs voor $f(x|\theta)$.
   Stel er is een $x \in D$ met $f(x|\theta) > 0$ en zij $x' \in D$ willekeurig.
   Er is een θ' in Θ waarvoor $f(x'|\theta') > 0$. $P_\eta(\underline{x} = x' | \{x,x'\})$ is voor $\eta = \theta$ en voor $\eta = \theta'$ gedefiniëerd en de twee zijn gelijk,

   dus $\dfrac{f(x'|\theta)}{f(x|\theta) + f(x'|\theta)} = \dfrac{f(x'|\theta')}{f(x|\theta') + f(x'|\theta')}$. De tweede is positief en dus ook de eerste zodat $f(x'|\theta) > 0$ voor alle $x' \in D$, en $D \subset X_\theta$. Is er geen $x \in D$ met $f(x|\theta) > 0$ dan is $D \subset X'_\theta$.

2. Zij $(X_1^*, X_2^*, \ldots)$ een disjuncte splitsing van X door dragers.
   Voor de klassen $D \in \mathcal{D}$ geldt dan bovendien: er is een h met $D \subset X_h^*$.

   BEWIJS:

   2.1.3. stelling 1 is voor elk puntenpaar in D van toepassing, gezien gevolg 1. Zijn dus x en $x' \in D$ dan is er een h zodat $x, x' \in X_h^*$. Daar de X_h^* disjunct zijn is dan $D \subset X_h^*$.

3. Op X kan een positieve functie $r(x)$ als volgt worden gedefinieerd: Kies in elke $D \in \mathcal{D}$ een punt x_D. Zoek een $\theta \in \Theta$ met $P_\theta(\underline{x} = x_D) > 0$, of $f(x_D|\theta) > 0$ en schrijf:
$$r(x) = \frac{P_\theta(\underline{x} = x|\{x,x_D\})}{P_\theta(\underline{x} = x_D|\{x,x_D\})} \text{ voor alle } x \in D \text{ en voor alle } D \in \mathcal{D}.$$
Uit gevolg 1 blijkt: $r(x) > 0$ op X en $r(x)$ hangt niet van (de keuze van) θ af.

Voor we de betekenis van afdoendheid nader toelichten geven we twee criteria, equivalent met afdoendheid.

DEF 2:

Voor de familie $\mathcal{P}$ ($\mathcal{F}$) geldt het factoriseringscriterium (FISHER-NEYMAN-criterium) ten opzichte van de klasse-indeling $\mathcal{D}$ van X als op X functies van het volgende type bestaan:

(1) voor alle $\theta \in \Theta$ een functie $k(x,\theta) \geq 0$, constant op elke klasse $D \in \mathcal{D}$;

(2) een functie $h(x) > 0$

terwijl voor alle $\theta \in \Theta$ geldt:

$P_\theta(\underline{x} = x)$ (of $f(x|\theta)$) $= h(x) \times k(x,\theta)$ op X.

DEF 3:

Voor de klasse-indeling $\mathcal{D}$ geldt het klassificatiecriterium ten opzichte van $\mathcal{P}$ (t.o.v. $\mathcal{F}$) als voor elke klasse $D \in \mathcal{D}$ geldt:
zijn x en x' punten van D, dan is er een $c(x,x') > 0$ zó dat voor alle $\theta \in \Theta$:

$P_\theta(\underline{x} = x) = c(x,x')P_\theta(\underline{x} = x')$ dan wel $f(x|\theta) = c(x,x') \cdot f(x'|\theta)$.

STELLING 1:

De klasse-indeling $\mathcal{D}$ is afdoende voor $\mathcal{P}$ (voor $\mathcal{F}$) d.e.s.d.a. voor $\mathcal{P}$ ($\mathcal{F}$) het factoriseringscriterium geldt t.o.v. $\mathcal{D}$ en ook d.e.s.d.a. voor $\mathcal{D}$ het klassificatiecriterium geldt t.o.v. $\mathcal{P}$ ($\mathcal{F}$).

BEWIJS:

We geven het bewijs weer voor dichtheden, het discrete geval wordt bewezen door $f(x|\theta)$ door $P_\theta(\underline{x} = x)$ te vervangen.

a) Zij $\mathfrak{D}$ afdoende voor $\mathcal{P}$, $r(x)$ de in 2.4 gevolg 3 gedefiniëerde fucntie.

Neem $D \in \mathfrak{D}$ en het vaste punt x_D als in de definitie van $r(x)$.

Is $x \in D$, dan is hetzij $f(x|\theta) = 0$ en dan ook $f(x_D|\theta) = 0$ of

$$f(x|\theta) > 0 \text{ en } r(x) = \frac{P_\theta(\underline{x} = x|\{x,x_D\})}{P_\theta(\underline{x} = x_D|\{x,x_D\})} = \frac{f(x|\theta)}{f(x_D|\theta)} .$$

In beide gevallen is $f(x|\theta) = r(x) \cdot f(x_D|\theta)$, zodat het factoriseringscriterium geldt met $h(x) = r(x) > 0$ en $k(x,\theta) = f(x_D|\theta)$, niet negatief en constant op D voor alle $D \in \mathfrak{D}$.

b) Geldt het factoriseringscriterium voor $\mathcal{P}$ t.o.v. $\mathfrak{D}$ en is $D \in \mathfrak{D}$, x en x' in D, dan is:

$$f(x|\theta) = h(x) \cdot k(x,\theta) = h(x) \cdot k(x',\theta) = \frac{h(x)}{h(x')} \cdot f(x'|\theta) =$$

$= c(x,x') \cdot f(x'|\theta)$; hierin is $c(x,x') > 0$ zodat het klassificatiecriterium geldt.

c) Geldt het klassificatiecriterium voor $\mathfrak{D}$ t.o.v. $\mathcal{P}$, is $D \in \mathfrak{D}$ en $E \subset D$ een eindige deelverzameling. We kiezen x_E vast in E en θ zó dat $f(x_E|\theta) > 0$.

$$P_\theta(\underline{x} = x|E) = 0 \text{ als } x \notin E \text{ en } \frac{f(x|\theta)}{\sum\limits_{x_i \in E} f(x_i|\theta)} = \frac{c(x,x_E) \cdot f(x_E|\theta)}{\sum\limits_{x_i \in E} c(x_i,x_E) \cdot f(x_E|\theta)} =$$

$$= \frac{c(x,x_E)}{\sum\limits_{x_i \in E} c(x_i,x_E)}$$ als $x \in E$; hierin komt θ niet voor. Merk op dat uit het klassificatiecriterium direct volgt dat hetzij $f(x|\theta) > 0$ op D óf $f(x|\theta) = 0$ op D, zodat bij een andere keuze van x_E, $P_\theta(\underline{x} = x|E)$ voor dezelfde θ gedefiniëerd is. $\mathfrak{D}$ is dus afdoende voor $\mathcal{P}$.

Def. 2 en 3 zijn natuurlijk ook te formuleren voor een functie t door voor $\mathfrak{D}$ de door t voortgebrachte klasse-indeling te nemen.

Wil men in het geval van een familie $\mathcal{P}$ van discrete verdelingen met behulp van de definitie van afdoendheid nagaan of een klasse-indeling $\mathfrak{D}$ afdoende is voor $\mathcal{P}$, dan is het niet nodig voor $D \in \mathfrak{D}$ en elke eindige deelverzameling $E \subset D$ na te gaan of $P_\theta(\underline{x} = x|E)$ niet van θ afhangt. In dit geval is, zodra $P_\theta(D) > 0$ ook

$$P_\theta(\underline{x} = x|D) = \frac{P_\theta(\underline{x} = x)}{P_\theta(D)}$$ als $x \in D$ en 0 buiten D gedefiniëerd.

Nu geldt:

STELLING 2:

De klasse-indeling $\mathcal{D}$ is afdoende voor $\mathcal{P}$ d.e.s.d.a. voor elke $D \in \mathcal{D}$ $P_\theta(\underline{x} = x|D)$ niet van θ afhangt voor die θ waarvoor de functie bestaat.

BEWIJS:

De voorwaarde is noodzakelijk. Voor eindige D is dat triviaal; is D aftelbaar oneindig dan neemt men een stijgende rij eindige deelverzamelingen $A_1, A_2, \ldots$ van D met $\bigcup_{i=1}^{\infty} A_i = D$ en $P(A_1) > 0$

$$P_\theta(\underline{x} = x|D) = \frac{P_\theta(\underline{x} = x)}{P_\theta(A_1)} \times \frac{P_\theta(A_1)}{P_\theta(A_2)} \times \ldots$$ Geen van de factoren hangt van θ af.

Is nu aan de voorwaarde voldaan en is $E \subset D \in \mathcal{D}$, terwijl E eindig en $P_\theta(E) > 0$. Dan is $P_\theta(\underline{x} = x|E) = P_\theta(\underline{x} = x|D) \times \frac{P_\theta(D)}{P_\theta(E)}$.

$P_\theta(\underline{x} = x|D)$ en $\frac{P_\theta(E)}{P_\theta(D)}$ hangen niet van θ af.

2.5. Stellingen over afdoendheid

STELLING 1:

Is de klasse-indeling $\mathcal{D}$ afdoende voor $\mathcal{P}$ (voor $\mathcal{F}$) en is $\mathcal{E}$ een verfijning van $\mathcal{D}$, dan is $\mathcal{E}$ afdoende voor $\mathcal{P}$ (voor $\mathcal{F}$).
Zijn u en t functies op X en is u afdoende voor $\mathcal{P}$ ($\mathcal{F}$), terwijl u te schrijven is als functie van t, dan is t afdoende voor $\mathcal{P}$ ($\mathcal{F}$).

BEWIJS:

Een factorisering bij $\mathcal{D}$ is zonder meer te gebruiken als factorisering bij $\mathcal{E}$.

OPMERKING:

Voor elke familie $\mathcal{P}$ of $\mathcal{F}$ is er tenminste één afdoende klasse-indeling, nl. die waarbij de klassen de punten van X zijn.
Anders gezegd: de functie $u(x) = x$ (de uitkomst van een experiment) is altijd afdoende voor $\mathcal{P}$ of $\mathcal{F}$.

STELLING 2:

De functies $l(.|x)$, $L(.|x)$ en $P.(\underline{x} = x)$ of $f(x|.)$ zijn afdoende voor $\mathcal{P}$ of $\mathcal{F}$, ongeacht de keuze van θ^* bij de definitie van $l(.|x)$.

BEWIJS:

We geven het bewijs voor $l(.|x)$. Daar $l(.|x)$ te schrijven is als functie van elk van de andere completeert stelling 1 het bewijs.

Nu geldt:

$$P.(\underline{x} = x) = P_{\theta^*}(\underline{x} = x) \cdot \exp[l(.|x)], \text{ of}$$

$$f(x|.) = f(x|\theta^*) \cdot \exp[l(.|x)], \text{ volgens 2.2 opm. 2.}$$

Daarmee voldoet $\mathcal{P}$ resp. $\mathcal{F}$ aan het factoriseringscriterium ten opzichte van de door $l(.|x)$ voortgebrachte klasse-indeling. Beide gelden ongeacht de keuze van θ^*, dat wil zeggen ongeacht de keuze van een disjuncte splitsing van X door dragers, ook al leidt zo'n andere keuze tot een andere factorisering.

STELLING 3:

Is $(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ een steekproef uit een verdeling behorende tot een familie $\mathcal{P}$ of $\mathcal{F}$, zodat ook de simultane verdeling tot een familie van dit type behoort, dan is de "order statistic" $(\underline{x}_{(1)}, \underline{x}_{(2)}, \ldots, \underline{x}_{(n)})$ afdoende voor de familie.

BEWIJS:

Voor uitkomsten $x = (x_1, x_2, \ldots, x_n)$ en $x' = (x_{i_1}, x_{i_2}, \ldots, x_{i_n})$ met dezelfde geordende uitkomst is voor alle $\theta \in \Theta$ $f(x|\theta) =$
$= f(x_1|\theta) \times f(x_2|\theta) \times \ldots \times f(x_n|\theta) = 1 \times f(x_{i_1}|\theta) \times f(x_{i_2}|\theta) \times \ldots$
$\ldots \times f(x_{i_n}|\theta)$, zodat het klassificatiecriterium geldt met $c(x,x') = 1$.

2.6. Grofste afdoende klasse-indeling

DEF 1:

Twee punten x en x' heten gelijkwaardig onder $\mathcal{P}$ of $\mathcal{F}$ (notatie: $x \equiv x'$), als er een constante $c(x,x') > 0$ bestaat zó dat

$$P.(\underline{x} = x) = c(x,x')P.(\underline{x} = x') \quad \text{ofwel}$$
$$f(x|.) = c(x,x')f(x'|.).$$

De collectie $\mathfrak{D}^*$ van equivalentie-klassen onder deze relatie voldoet dan aan:

STELLING 1:

Voor $\mathfrak{D}^*$ gelden de volgende beweringen:

(1) $\mathfrak{D}^*$ is afdoende voor $\mathcal{P}$ of $\mathcal{F}$;

(2) elke afdoende klasse-indeling $\mathfrak{D}$ voor $\mathcal{P}$ of $\mathcal{F}$ is een verfijning van $\mathfrak{D}^*$.

Beide beweringen zijn op grond van het klassificatiecriterium triviaal.

$\mathfrak{D}^*$ noemen we de grofste afdoende klasse-indeling.
Klasse-indelingen die een vergroving zijn van $\mathfrak{D}^*$ heten noodzakelijke klasse-indelingen. $\mathfrak{D}^*$ kan dus ook genoemd worden: de noodzakelijke en afdoende klasse-indeling. We spreken evenzo van noodzakelijke en van noodzakelijk en afdoende functies.

STELLING 2 (HOOFDSTELLING):

De functie $l(.|x)$ brengt, ongeacht de keuze van θ^*, $\mathfrak{D}^*$ voort, en is dus noodzakelijk en afdoende.

BEWIJS:

Zij $\mathfrak{D}$ de door $l(.|x)$ voortgebrachte klasse-indeling. Volgens 2.5 stelling 2 is $\mathfrak{D}$ afdoende voor $\mathcal{P}$ of $\mathcal{F}$ en dus op grond van stelling 1 een verfijning van $\mathfrak{D}^*$.
Is omgekeerd $x \equiv x'$, dan is voor een familie $\mathcal{F}$ (en analoog voor $\mathcal{P}$)

$f(x|.) = c(x,x')f(x'|.)$, en voor elke keuze van een disjuncte splitsing van X door dragers, voor $x \in X$ en een k

$f(x|\theta^*) = f(x|\theta_k) = c(x,x')f(x'|\theta_k) = c(x,x')f(x'|\theta^*)$.

Hieruit volgt dat $l(.|x) = l(.|x')$. Daarmee is $\mathfrak{D}$ ook een vergroving van $\mathfrak{D}^*$ gebleken, en $\mathfrak{D} = \mathfrak{D}^*$.

2.7. Noodzakelijke functies

Iedere functie die constant is op elke klasse van $\mathfrak{D}^*$ heet noodzakelijk. Verder wordt $\mathfrak{D}^*$ voortgebracht door $l(.|x)$. Hieruit leiden we een aantal beweringen af:

BEWERING 1:

Voor elke $\theta \in \Theta$ is $l(\theta|x)$ noodzakelijk; voor elke deelverzameling $\Theta^* \in \Theta$ is $\{l(\theta|x) : \theta \in \Theta^*\}$ noodzakelijk. De lineaire ruimte U van sommen $\sum_{i=1}^{n} a_i l(\theta_i|x)$ met coëfficiënten $a_i \in R^1$ en met $\theta_i \in \Theta$ is equivalent met $l(.|x)$ evenals elke basis voor U. De basis en U zijn dus beide noodzakelijk en afdoende.

BEWERING 2:

$a(x) = \sup_{\theta \in \Theta} l(\theta|x)$ is noodzakelijk voor de familie $\mathcal{P}$ of $\mathcal{F}$; de waarde $+\infty$ wordt toegestaan.

BEWERING 3:

Heeft voor alle x $l(\theta|x)$ een maximum $a(x)$, dan is er een noodzakelijke functie $\theta(x)$ waarvoor geldt: $l(\theta(x)|x) = a(x)$.

BEWIJS:

Zij $D \in \mathfrak{D}^*$, dan is $l(.|x)$ constant op D. Voor elk punt $x \in D$ wordt het maximum voor dezelfde verzameling van θ-waarden aangenomen. Kies hieruit een $\theta(x) = \theta_D$ en gebruik deze waarde op de hele klasse D.

BEWERING 4:

<u>Bestaat $a(x) = \max_{\theta \in \Theta} l(\theta|x)$ op X, dan is er een noodzakelijke meest aannemelijke schatter voor θ.</u>

BEWIJS:

Bestaat $a(x) = \max_{\theta \in \Theta} l(\theta|x)$, dan bestaat ook $\max_{\theta \in \Theta} L(\theta|x) = L(\theta^*|x) + a(x)$. De maxima van $l(\theta|x)$ en van $L(\theta|x)$ worden voor dezelfde functie $\theta(x) = \theta_D$ aangenomen.

OPMERKING:

De functies $\theta(x)$ in bewering 3 en 4 bestaan, ongeacht de aard van de verzameling Θ. Men krijgt steeds een $\theta(x)$ met waarden in Θ.

BEWERING 5:

Zij $\theta = (\theta_1, \theta_2, \ldots, \theta_r) \in R^r$. Bestaat $\frac{\partial L(\theta|x)}{\partial \theta_i}$ voor een $i \leq r$ op een verzameling $\Theta^* \subset \Theta$, dan is $\{\frac{\partial L(\theta|x)}{\partial \theta_i} : \theta \in \Theta^*\}$ noodzakelijk voor $\mathcal{F}$ of $\mathcal{P}$.

BEWIJS:

Is x equivalent met x' onder $l(.|x)$, of: $l(\theta|x) = l(\theta|x')$ voor alle $\theta \in \Theta$, dan geldt op Θ^*:

$$\frac{\partial L(\theta|x)}{\partial \theta_i} = \frac{\partial l(\theta|x)}{\partial \theta_i} = \frac{\partial l(\theta|x')}{\partial \theta_i} = \frac{\partial L(\theta|x')}{\partial \theta_i}.$$

BEWERING 6:

Is $\chi(x,\theta)$ de karakteristieke functie van de drager X_θ, dan is de functie $\chi(x,.)$ noodzakelijk voor $\mathcal{P}$ of $\mathcal{F}$.

BEWIJS:

Zij x equivalent met x' onder $l(.|x)$. Dan geldt:
x en $x' \in X_\theta$ òf x en $x' \notin X_\theta$, voor elke $\theta \in \Theta$, want $l(\theta|x) = l(\theta|x')$ en dus is $f(x|\theta) = 0$ d.e.s.d.a. $f(x'|\theta) = 0$ (voor discrete verdelingen evenzo). Dus is $\chi(x,\theta) = 1$ d.e.s.d.a. $\chi(x',\theta) = 1$ zodat $\chi(x,.) = \chi(x',.)$.

2.8. Voorbeelden

De hier gegeven voorbeelden hebben alle betrekking op families met constante drager. Voorbeelden waarin families met niet constante drager optreden worden gegeven in 3.4.

Onder wat sterkere regulariteitsvoorwaarden geeft DYNKIN [7] een systematische benadering voor het probleem, de noodzakelijke en afdoende functies voor een familie $\mathcal{F}$ te vinden.

Hier zijn de voorbeelden in hoofdzaak illustratief bedoeld.

In het geval van families met constante drager geeft het klassificatie-criterium doorgaans de eenvoudigste methode voor het vinden van de noodzakelijk en afdoende functies. Daarnaast laten we in enkele gevallen de factorisering zien en, in het geval van discrete kansverdelingen, de voorwaardelijke waarschijnlijkheid op de klassen van $\mathcal{D}^*$, die dus de index niet meer bevat.

Uit $f(x|\theta) = c(x,x')\cdot f(x'|\theta)$ voor alle θ bij equivalente punten x en x' zien we dat gelden moet:

$$\frac{f(x|\theta)}{f(x'|\theta)} \text{ is onafhankelijk van } \theta.$$

1. Zij $(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ een steekproef uit een binomiale verdeling $B(\theta,k)$; $\theta \in (0,1)$; $x = (x_1, x_2, \ldots, x_n)$ waarin $x_i = 0, 1, \ldots, k-1$ of k. Dan is

$$P_\theta(\underline{x} = x) = \prod_{i=1}^{n} \binom{k}{x_i} \cdot \theta^{x_i} \cdot (1-\theta)^{k-x_i} \quad \text{voor alle } \theta \text{ en op } X$$

$$\frac{P_\theta(\underline{x} = x)}{P_\theta(\underline{x} = x')} = \left[\prod_{i=1}^{n} \binom{k}{x_i} / \binom{k}{x_i'}\right] \cdot \left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i'} .$$

Dit quotiënt is onafhankelijk van θ d.e.s.d.a. $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i'$, zodat de functie $\sum_{i=1}^{n} x_i$ noodzakelijk en afdoende is voor $\mathcal{P}$.

$\mathcal{D}^*$ bestaat uit de verzamelingen waarvoor $\sum_{i=1}^{n} x_i$ constant is.

Inderdaad is $P_\theta(\underline{x} = x)$ als volgt te factorizeren:

$$P_\theta(\underline{x} = x) = \left[\prod_{i=1}^{n} \binom{k}{x_i}\right] \cdot \left\{(1-\theta)^{kn} \cdot \left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n} x_i}\right\}.$$

De voorwaardelijke waarschijnlijkheid vindt men bv. aldus:

$\sum_{i=1}^{n} x_i = c$ kan gezien worden als het resultaat van nk trekkingen uit een alternatieve verdeling met kans θ op een uitkomst 1, en $1-\theta$ op een uitkomst 0, dus

$$P_\theta\left(\sum_{i=1}^{n} \underline{x}_i = c\right) = \binom{nk}{c} \theta^c (1-\theta)^{nk-c}$$

$$P_\theta\left(\underline{x} = x \,\middle|\, \sum_{i=1}^{n} \underline{x}_i = c\right) = \frac{P_\theta(\underline{x} = x)}{P_\theta\left(\sum_{i=1}^{n} x_i = c\right)} \quad \text{(als } x \text{ voldoet aan}$$

$$\sum_{i=1}^{n} x_i = c) = \frac{\prod\limits_{i=1}^{n} \binom{k}{x_i}}{\binom{nk}{c}} \text{, inderdaad onafhankelijk van } \theta.$$

2. Zij $(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ een steekproef uit een Poissonverdeling met index θ, $\theta \in (0,\infty)$; $x = (x_1, x_2, \ldots, x_n)$ met $x_i = 0, 1, 2, \ldots$.
Dan is

$$P_\theta(\underline{x} = x) = e^{-n\theta} \cdot \frac{\theta^{\sum\limits_{i=1}^{n} x_i}}{\prod\limits_{i=1}^{n} (x_i!)}$$

$$\frac{P_\theta(\underline{x} = x)}{P_\theta(\underline{x} = x')} = \frac{\prod\limits_{i=1}^{n} (x_i'!)}{\prod\limits_{i=1}^{n} (x_i!)} \cdot \theta^{\sum\limits_{i=1}^{n} x_i - \sum\limits_{i=1}^{n} x_i'} .$$

Dit quotiënt is onafhankelijk van θ d.e.s.d.a. $\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i'$ en $\mathcal{D}^*$ bestaat opnieuw uit de verzamelingen waarvoor $\sum\limits_{i=1}^{n} x_i$ constant is.

Met $h(x) = \dfrac{1}{\prod\limits_{i=1}^{n} (x_i!)}$ verkrijgen we de gewenste factorisering.

De som van n grootheden met Poissonverdeling met index θ heeft zelf weer een Poissonverdeling met index $n\theta$.
Dus is

$$P_\theta(\sum_{i=1}^{n} \underline{x}_i = c) = e^{-n\theta} \frac{(n\theta)^c}{c!}, \text{ zodat}$$

$$P_\theta(\underline{x} = x) | \sum_{i=1}^{n} \underline{x}_i = c) = \frac{c!}{n^c \prod\limits_{i=1}^{n} (x_i!)} = \frac{1}{n^c} \binom{c}{x_1\; x_2 \ldots x_n},$$

waarin θ niet meer voorkomt.

OPMERKING: $\binom{c}{x_1\, x_2 \ldots x_n}$ geeft het aantal wijzen waarop c objecten in groepjes ter grootte $x_1, x_2, \ldots, x_n$ verdeeld kunnen worden, $\frac{1}{n^c}$ het totaal aantal mogelijke manieren waarop c objecten in (hoogstens) n groepen kunnen worden verdeeld.

3. Zij $(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ een steekproef uit een verdeling met dichtheid van de vorm $c(\theta)e^{-y^\theta}$, $\theta \in (0,\infty)$, $y \in (0,\infty)$ en $c(\theta) = 1/(\int_0^\infty e^{-y^\theta}\, dy)$, $x = (x_1, x_2, \ldots, x_n)$, $x_i \in (0,\infty)$.

Dan is

$$f(x|\theta) = c^n(\theta)e^{-\sum_{i=1}^n x_i^\theta}.$$

$$\frac{f(x|\theta)}{f(x'|\theta)} = e^{-\{\sum_{i=1}^n x_i^\theta - \sum_{i=1}^n x_i'^\theta\}}.$$

Dit quotiënt hangt niet van θ af d.e.s.d.a. $\sum_{i=1}^n x_i^\theta = \sum_{i=1}^n x_i'^\theta$, voor alle θ, in het bijzonder voor $\theta = 1, 2, \ldots, n$.
Volgens 1.2 voorbeeld 5 zijn dan $(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ en $(x'_{(1)}, x'_{(2)}, \ldots, x'_{(n)})$ gelijk, zodat de geordende steekproef noodzakelijk is voor de familie. De geordende steekproef is echter ook afdoende (2.5 stelling 3).

Een factorisering staat er als we stellen: $h(x) = 1$ en in plaats van $\sum_{i=1}^n x_i^\theta$ schrijven $\sum_{i=1}^n x_{(i)}^\theta$.

4. Zij $(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ een steekproef uit een normale verdeling $N(\mu,\sigma^2)$, $\mu \in (-\infty,\infty)$, $\sigma^2 \in (0,\infty)$; $x = (x_1, x_2, \ldots, x_n)$, $\theta = (\mu,\sigma^2)$.

$$f(x|\theta) = \left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n \cdot \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2\right],$$

$$\frac{f(x|\theta)}{f(x'|\theta)} = \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i'^2\right) + \frac{\mu}{\sigma^2}\left(\sum_{i=1}^n x_i - \sum_{i=1}^n x_i'\right)\right].$$

Het quotiënt hangt niet van $\theta = (\mu,\sigma^2)$ af d.e.s.d.a. $\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i'^2$, èn $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i'$.

De klassen van $\mathcal{D}^*$ zijn dus de doorsnijding van vlakken $\sum_{i=1}^{n} x_i = c$ met boloppervlakken $\sum_{i=1}^{n} x_i^2 = r$.

Een factorisering wordt eenvoudig gevonden door uitschrijven van de som in de exponent van $f(x|\theta)$.

2.9. Interpretatie van afdoendheid

2.9.1. Families van discrete verdelingen

Zij $\mathcal{D}$ een afdoende klasse-indeling voor $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$.

1. Het factoriseringscriterium

Uit 2.4 stelling 2 blijkt dat, als $P_\theta(D) > 0$, $P_\theta(\underline{x} = x|D) = \dfrac{P_\theta(\underline{x} = x)}{P_\theta(D)}$ voor $x \in D \in \mathcal{D}$ niet van θ afhangt.

We hebben nu de volgende factorisering:

$P_\theta(\underline{x} = x) = P_\theta(D) \times P(\underline{x} = x|D)$ voor $x \in D$. Deze is natuurlijk eveneens juist als $P_\theta(D) = 0$, daar dan ook $P_\theta(\underline{x} = x) = 0$. De eerste factor geeft een kansverdeling over de ruimte $\mathcal{D}$ van klassen, en deze kansverdeling is verschillend voor verschillende θ. De tweede factor geeft binnen elke klasse $D \in \mathcal{D}$ een kansverdeling die dezelfde is voor alle θ.

Wil men nu iets over θ te weten komen dan kan worden volstaan met een beschouwing van $P_\theta(D)$. De andere factor bevat θ niet en levert daarmee geen bijdrage tot kennis omtrent θ.

2. Het klassificatiecriterium

Laat x en x' twee punten zijn in dezelfde klasse D van $\mathcal{D}$. Heeft men bij een experiment x als uitkomst gekregen, dan heeft men daarbij $P.(\underline{x} = x)$ "waargenomen" veeleer dan alleen x zelf. Op deze waarneming zal men iedere gevolgtrekking omtrent θ moeten baseren. Vindt men x', d.w.z. $P.(\underline{x} = x')$, dan heeft men essentiëel dezelfde kennis over θ, want $P.(\underline{x} = x') = c(x,x')P.(\underline{x} = x)$ en $c(x,x')$ is een constante in het model $\mathscr{P}$. Bij de uitkomsten x en x' zal men dus dezelfde conclusies over θ trekken.

3. De definitie van afdoendheid

We stellen ons twee experimenten voor. Het eerste bestaat uit het doen van een waarneming $\underline{x} = x$. Het tweede uit het doen van een waarneming $\underline{D} = D$ gevolgd door het genereren van een waarde x' volgens de verdeling $P(\underline{x} = x|D)$.
$\underline{x}$ en $\underline{x}'$ hebben nu dezelfde kansverdeling, te weten $P_\theta(\underline{x} = x)$ en $P_\theta(D) \times P(\underline{x}' = x|D)$ voor $x \in D \in \mathscr{D}$.

Men is dus indien men D kent niet slechter af (en ook niet beter) dan indien men de precieze positie van x in D weet, afgezien van de mogelijkheid dat men in de eerste situatie nog aselect een x' in D moet kiezen (volgens de verdeling $P(\underline{x} = x|D)$). In deze zin bevat D evenveel informatie omtrent θ als de uitkomst x (zie HALMOS en SAVAGE [11]).

Zonder het begrip "hoeveelheid informatie omtrent θ" te hebben gedefiniëerd kan men dus wel zeggen: iedere zinvolle definitie voor dit begrip zal aan punten, equivalent voor $\mathscr{P}$, dezelfde waarde toekennen. Het zal dus een noodzakelijke functie zijn.

2.9.2. Families van verdelingen met een dichtheid

Zij $\mathscr{D}$ een afdoende klasse-indeling voor $\mathscr{F} = \{f(x|\theta) \ : \theta \in \Theta\}$.

Het klassificatiecriterium levert direct een interpretatie, volkomen analoog aan die in het discrete geval.
Een interpretatie van de definitie vereist een andere benadering van de theorie en wordt hier niet gegeven. Voor een interpretatie van het factoriseringscriterium geldt hetzelfde; we zullen nochtans laten zien, dat onder strikte regulariteitsvoorwaarden op eenvoudige wijze een analogon van 2.9.1.1 te verkrijgen valt.

Laat $t(x) = (t_1(x), t_2(x), \ldots, t_n(x))$ een vector van reëelwaardige functies zijn die continue partiële afgeleiden hebben naar $x_1, x_2, \ldots, x_n$. Laat verder bij $t(x) = t$ inversen bestaan zodat $x = u(t)$.

Voor de eenvoud nemen we aan dat de toevoeging eeneenduidig is op een open verzameling $\mathcal{O}$ met $P_\theta(\mathcal{O}) = 1$ voor alle θ, en dat de determinant van Jacobi $\mathcal{J} = \left| \left(\frac{\partial x_i}{\partial t_j}\right)_{i,j=1,2,\ldots,n} \right| \neq 0$ op $\mathcal{O}$. Is $t^* = (t_1(x), t_2(x), \ldots$ $\ldots, t_r(x))$, $r \leq n$, afdoende voor de familie $\mathcal{F}$ van dichtheden $\{f(x|\theta) : \theta \in \Theta\}$, dan is $f(x|\theta) = h(x) \cdot k(t^*(x),\theta)$ voor geschikte functies $h(x)$ en $k(t^*(x),\theta)$.
Bij transformatie van X met $t(x)$ ontstaat een dichtheid op $t(X) = T$, $g(t|\theta) = h(u(t)) \cdot k(t^*,\theta) \cdot |\mathcal{J}|_t$; $|\mathcal{J}|_t$ staat voor $|\mathcal{J}|$, uitgedrukt in t.
De marginale verdeling van $t_1, t_2, \ldots, t_r$ is dan:

$$\int_{t_{r+1}} \cdots \int_{t_n} h(u(t))k(t^*,\theta) \cdot |\mathcal{J}|_t \, dt_{r+1} \ldots dt_n = g^*(t^*|\theta).$$

$$g(t|\theta)/g^*(t^*|\theta) = \frac{|\mathcal{J}|_t}{\int_{y_{r+1}} \cdots \int_{y_n} \frac{h(u(y))}{h(u(t))} |\mathcal{J}|_y \, dy_{r+1} \ldots dy_n} = H(t|t^*).$$

$H(t|t^*)$ hangt niet meer van θ af en is een voorwaardelijke dichtheid op de verzamelingen met vaste waarde van t^*, terwijl $g^*(t^*|\theta)$ als wh. dichtheid over de ruimten van klassen $\{t : t^* \text{ vast}\}$ kan worden beschouwd.

Zie verder bv. HOGG en CRAIG [13].

2.9.3. Het model

Tot nu toe hebben we steeds verondersteld dat de familie $\mathcal{P}$ of $\mathcal{F}$, dat wil zeggen de functie $P_\theta(\underline{x} = x)$ of $f(x|\theta)$, als functie op $X \times \Theta$ beschouwd, bekend was. Gegeven dit "model" besloten we dat $l(.|x)$ noodzakelijk en afdoende was.
Stel daarentegen dat we zouden willen onderzoeken of de bewering $P \in \mathcal{P}$ onjuist is. We kunnen dan deze bewering beschouwen als nulhypothese H_0 en we kunnen daarmee voor elke $x \in X$ de verdeling $P(\underline{x} = x|D)$ onder H_0, die niet meer afhangt van de index θ. Dit levert mogelijkheden voor toetsen van het model $\mathcal{P}$ tegen verschillende alternatieve modellen.

3. EXPONENTIËLE FAMILIES VAN KANSVERDELINGEN

3.1. Inleiding

Dit hoofdstuk zal worden geformuleerd voor dichtheden $f(x|\theta)$, hoewel alles onverminderd geldt voor discrete verdelingen.

Tot nu toe hebben we één regulariteitsvoorwaarde (R) ingevoerd, die betrekking had op de dichtheden $f(x|\theta)$. Men kan zich afvragen of het niet voor de hand zou liggen ook aan Θ eisen op te leggen. In 2.7 bewering 5 is bijvoorbeeld een eis geformuleerd.

Het is echter duidelijk dat, gegeven een familie $\mathcal{F}$ van dichtheden, een klasse-indeling $\mathcal{D}$ afdoende is voor $\mathcal{F}$ ongeacht de gekozen indicering. Het is om deze reden dat we er de voorkeur aan hebben gegeven niets over Θ aan te nemen. Een andere reden is, dat bij de nu volgende behandeling van exponentiële families van verdelingen blijken zal dat bij dit heel belangrijke type verdelingen steeds een herindicering mogelijk is die, eventueel na uitbreiding van $\mathcal{F}$ met dichtheden die van hetzelfde type zijn, vanzelf leidt tot een indexverzameling waarop $f(x|\theta)$ voor elke x aan sterke regulariteitseisen voldoet. Dit laten we in 3.2.2 zien.

Daarna zal het geval van niet-constante dragers voor het geval van een aselecte steekproef nader worden beschouwd.

3.2. Constante drager

3.2.1. KOOPMAN-PITMAN klassen

DEF 1:

Een familie van dichtheden $\mathcal{F}$ met constante drager $X \subset R^n$ heet een exponentiële familie of een KOOPMAN-PITMAN klasse als voor alle $\theta \in \Theta$, $f(x|\theta)$ te schrijven is in de vorm

$$(1) \qquad f(x|\theta) = \exp\left[c(x) + \phi(\theta) + \sum_{i=1}^{s} c_i(x)\phi_i(\theta)\right] \quad \text{op } X.$$

DEF 2:

Een exponentiële familie $\mathcal{F}$ is in <u>kanonieke vorm</u> met betrekking tot $\theta_0 \in \Theta$, $x_0 \in X$ als in (1):

1. $\phi(\theta_0) = 0$
2. $\phi_1(\theta), \phi_2(\theta), \ldots, \phi_s(\theta)$ zijn lineair onafhankelijk
3. $(\phi_1(\theta_0), \phi_2(\theta_0), \ldots, \phi_s(\theta_0)) = (0, 0, \ldots, 0)$
4. $c_1(x), c_2(x), \ldots, c_s(x)$ zijn lineair onafhankelijk
5. $(c_1(x_0), c_2(x_0), \ldots, c_s(x_0)) = (0, 0, \ldots, 0)$.

Elke exponentiële familie kan in kanonieke vorm worden gebracht. Zij namelijk:

$$f(x|\theta) = \exp\left[b(x) + \psi(\theta) + \sum_{i=1}^{S} b_i(x)\psi_i(\theta)\right], \text{ en } \theta_0 \in \Theta,\ x_0 \in X;$$

dan is

$$l(\theta|x) - l(\theta|x_0) = \sum_{i=1}^{S} (b_i(x) - b_i(x_0))(\psi_i(\theta) - \psi_i(\theta_0)) \overset{\text{def}}{=} u(x,\theta).$$

Overeenkomstig 1.2 vormen we uit de $u(x,\theta)$ de lineaire ruimte U_x. Deze heeft klaarblijkelijk een eindige rang $s \leq S$. Een basis zij $c_1(x), c_2(x), \ldots, c_s(x)$, waarvoor dus 4. geldt.
Daar $u(x_0,\theta) = 0$ voor alle θ is ook $c_i(x_0) = 0$ voor $i = 1, 2, \ldots, s$, zodat aan 5. voldaan is.
Schrijf nu: $u(x,\theta) = \sum_{i=1}^{s} c_i(x)\phi_i(\theta)$.
Daar $u(x,\theta_0) = 0$ en de $c_i(x)$ lineair onafhankelijk zijn volgt hieruit dat $\phi_i(\theta_0) = 0$, zodat aan 3. is voldaan. Verder blijkt uit 1.2 stelling 1 dat de $\phi_i(\theta)$ eveneens lineair onafhankelijk zijn, zodat 2. geldt.

Definiëer tenslotte $\phi(\theta)$ als $l(\theta|x_0)$ en $c(x)$ als $L(\theta_0|x)$, dan is $\phi(\theta_0) = 0$ en $f(x|\theta) = \exp\left[c(x) + \phi(\theta) + \sum_{i=1}^{s} c_i(x)\phi_i(\theta)\right]$ is in kanonieke vorm.

STELLING 1:

Een familie van dichtheden met constante drager is een exponentiële familie d.e.s.d.a. de rang r van U_x eindig is.

STELLING 2:

Is een familie van dichtheden met constante drager van de vorm (1), dan is $(c_1(x), c_2(x), \ldots, c_s(x))$ noodzakelijk en afdoende voor $\mathcal{F}$.

Beide stellingen zijn eenvoudig te bewijzen. Bij stelling 2 bedenke men dat $l(.|x) = u(x,.) + l(.|x_0)$ en dat $l(.|x_0)$ zich bij de klassevorming als een constante gedraagt.

VOORBEELDEN:

1. De families van normale verdelingen $N(\mu,1)$, $N(0,\sigma^2)$ en $N(\mu,\sigma^2)$; de familie van de χ^2-verdelingen, de Γ- en B-verdelingen en de exponentiële verdelingen met vast beginpunt zijn exponentiële families met constante drager. Vergelijk 2.8 voorbeeld 4.

2. De families van de binomiale verdelingen bij vaste steekproefomvang, de negatief binomiale verdelingen bij vast voorgeschreven laatste succes, en de Poissonverdelingen zijn exponentiële families (2.8 voorbeeld 1 en 2).

3. Niet-exponentiële families zijn:

Cauchyverdelingen: $f(x|\theta) = \dfrac{\theta_1}{\pi(\theta_1^2 + (x - \theta_2)^2)}$, $\theta = (\theta_1,\theta_2)$, $x \in R^1$.

Verdelingen van het type $c(\theta)\exp[-x^\theta]$, $x > 0$, $\theta > 0$.
(Zie 2.8 voorbeeld 3).

3.2.2. Natuurlijke indexverzameling voor exponentiële families

STELLING 1:

Zij $f_0(x)$, $f_1(x)$ een paar wh-dichtheden dat aan R voldoet, terwijl de dragers X_0 en X_1 een niet-lege doorsnede hebben. Er is dan een dichtheid $f_\lambda(x) = \dfrac{1}{I_\lambda} f_1^\lambda(x) f_0^{1-\lambda}(x)$ voor elke $\lambda \in (0,1)$ en deze dichtheden voldoen aan R.

BEWIJS:

Zij $I_\lambda = \int_{X_0 \cap X_1} f_1^\lambda(x) f_0^{1-\lambda}(x)dx$ volgens een integratie-methode.
De integrand is positief op de open verzameling $X_0 \cap X_1$ en I_λ is dus positief.

Verder is er voor positieve a en b en $\lambda \in (0,1)$ de volgende ongelijkheid welbekend: $a^{\lambda} b^{1-\lambda} \leq \lambda a + (1-\lambda) b$.

Dus $f_1^{\lambda}(x) f_0^{1-\lambda}(x) \leq \lambda f_1(x) + (1-\lambda) f_0(x)$ op $X_0 \cap X_1$. Het rechterlid heeft een eindige integraal, dus ook het linkerlid.

Daarmee is gebleken dat $f_\lambda(x)$ een dichtheid is; de continuïteit en de dragervoorwaarde in R zijn natuurlijk vervuld.

STELLING 2:

$\{f_\lambda(x) : \lambda \in (0,1)\}$ uit stelling 1 is een familie van dichtheden die differentiëerbaar is naar λ, en wel willekeurig vaak.

Alleen I_λ levert problemen op. Voor een bewijs dat I_λ willekeurig vaak differentiëerbaar is naar λ en dat de differentiatie onder het integraalteken kan worden uitgevoerd zie LEHMANN [15], p. 52-54.

Zij nu $\mathcal{F}$ een exponentiële familie met constante drager $X \subset R^n$, waarvoor R geldt, en met kanonieke vorm ten opzichte van x_0 en τ_0:

$f(x|\tau) = \exp\left[c(x) + \psi(\tau) + \sum_{i=1}^{r} c_i(x)\psi_i(\tau)\right]$ voor $\tau \in T$, zodat

$u(x,\tau) = l(\tau|x) - l(\tau|x_0) = \sum_{i=1}^{r} c_i(x)\psi_i(\tau)$.

$(\psi_1(\tau), \psi_2(\tau), \ldots, \psi_r(\tau))$ karakteriseert de verdeling, daar de waarde van $\psi(\tau)$ vastligt door de conditie $\int_X f(x|\tau)dx = 1$.

Toepassen van stelling 1 op $f(x|\tau_0)$ en $f(x|\tau_1)$ laat zien dat er dichtheden bestaan van de vorm

$$f_\lambda(x) = \frac{1}{I_\lambda} \cdot \exp\left[c(x) + \lambda\psi(\tau_1) + (1-\lambda)\psi(\tau_0) + \sum_{i=1}^{r} c_i(x)(\lambda\psi_i(\tau_1) + (1-\lambda)\psi_i(\tau_0))\right] \text{ met } \lambda \in (0,1).$$

Het punt $(\lambda\psi_1(\tau_1) + (1-\lambda)\psi_1(\tau_0), \ldots, \lambda\psi_r(\tau_1) + (1-\lambda)\psi_r(\tau_0))$ doorloopt een lijnstuk in R^r ($\lambda \in [0,1]$) met de eindpunten $(\psi_1(\tau_1), \ldots, \psi_r(\tau_r))$ en $(\psi_1(\tau_0), \ldots, \psi_r(\tau_0))$.

Verder zijn de toegevoegde dichtheden eveneens van exponentiële vorm. Door dit herhaald toe te passen blijkt dat men bij elke θ in een convexe verzameling $\Theta \subset R^r$ een dichtheid $f(x|\theta)$ kan construeren. De verkregen (uitgebreide) familie is een exponentiële familie en Θ met de punten $(\theta_1, \theta_2, \ldots, \theta_r)$ kan als nieuwe indexverzameling worden gekozen. Deze wordt de "natuurlijke" indexverzameling genoemd.

OPMERKING 1:

Men kan eenvoudig nagaan dat Θ een open deelverzameling van R^r bevat, want $\psi_1(\tau), \psi_2(\tau), \ldots, \psi_r(\tau)$ zijn lineair onafhankelijk en Θ is convex.

OPMERKING 2:

$(c_1(x), c_2(x), \ldots, c_r(x))$ is ook na de uitbreiding noodzakelijk en afdoende voor $\mathcal{F}$.

Uit stelling 2 volgt verder: $f(x|\theta)$ is in het inwendige van Θ willekeurig vaak differentiëerbaar naar θ_i, $i = 1, 2, \ldots, r$.
Men kan aantonen dat ook alle gemengde hogere afgeleiden bestaan, zoals
$\dfrac{\partial^2 f(x|\theta)}{\partial\theta_i \partial\theta_j}$.
Zonder enige eis op te leggen aan de indexverzameling verkrijgt men dus in het geval van een exponentiële familie met constante drager door uitbreiding van de oorspronkelijke familie en her-indicering analyticiteit in de index.

VOORBEELD:

Uitgaande van de χ^2-verdelingen $f_k(x) = \{\Gamma(k)\cdot 2^k\}^{-1} x^{k-1} e^{-\frac{x}{2}}$, $x \in (0,\infty)$, te schrijven in de vorm:
$f_k(x) = \exp\left[-\log \Gamma(k) - k \log 2 - \frac{x}{2} + (k-1)\log x\right]$ krijgen we
$f(x|\theta) = \exp\left[\phi(\theta) - \frac{x}{2} + (\theta - 1)\log x\right]$ voor $\theta \in [1,\infty)$, bij geschikte keuze van $\phi(\theta)$ dichtheden op $(0,\infty)$. Dit zijn de Γ-verdelingen; uitbreiding tot $\theta \in (0,1)$ wordt zo niet gekregen.

3.2.3. Afdoendheid en noodzakelijkheid voor deelverzamelingen van Θ

Bij exponentiële families heeft, zoals te verwachten was, elk van de functies $c_i(x)$ uit een kanonieke vorm een speciale relatie tot de bijbehorende θ_i. Neem bijvoorbeeld $c_1(x)$, kies $\theta_2^*, \ldots, \theta_r^*$ zó dat $\{\theta : \theta_2 = \theta_2^*, \ldots, \theta_r = \theta_r^*\} = \Theta_1 \neq \emptyset$.
De familie $\mathcal{F}_1 = \{f(x|\theta) : \theta \in \Theta_1\}$ is dan een exponentiële familie, te schrijven als:

$$f(x|\theta) = \exp[\phi^{(1)}(\theta) + c^{(1)}(x) + \theta_1 c_1(x)] \text{ voor } x \in X \text{ en } \theta \in \Theta_1.$$

Het volgende woordgebruik is nu duidelijk:
$c_1(x)$ is afdoende en noodzakelijk voor θ_1 (voor $\mathcal{F}_1$).

Verder krijgt het factoriseringscriterium de volgende vorm:
voor θ_i geldt het factoriseringscriterium ten opzichte van $c_i(x)$ als er een functie $h(x; \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_r) > 0$ bestaat en voor elke $\theta_i \in \Theta_i$ functies $k(c_i(x), \theta_i)$ op X zó dat
$f(x|\theta) = h(x; \theta_1, \theta_2, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_r) k(c_i(x), \theta_i)$.

Het klassificatiecriterium wordt:
voor c_i geldt het klassificatiecriterium ten opzichte van θ_i als voor x en x' in X met $c_i(x) = c_i(x')$ een functie $q(x, x'; \theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_r)$ bestaat, zó dat $f(x|\theta) = q f(x'|\theta)$.

3.3. Niet-constante drager

3.3.1. Functies equivalent met de dragerfunctie

We beperken ons tot het geval van een aselecte steekproef $\underline{x} = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n)$ uit een verdeling met dichtheid $f(y|\theta) \in \mathcal{F}$, waarbij bovendien de drager X_θ, beschreven met behulp van de karakteristieke functies $\chi(y,\theta)$ van één van de volgende gedaanten is:

1. $\chi(y,\theta) = 1$ op $(-\infty, \theta)$ en 0 elders, $\theta > -\infty$;
2. $\chi(y,\theta) = 1$ op (θ, ∞) en 0 elders, $\theta < \infty$;
3. $\chi(y,\theta) = 1$ op (θ_1, θ_2) en 0 elders, waarbij $\theta = (\theta_1, \theta_2)$ en $-\infty < \theta_1 < \theta_2 < \infty$.

Volgens 2.7 bewering 6 zijn de functies $\chi(x,.) = \prod_{i=1}^{n} \chi(y_i,.)$ noodzakelijk en we zoeken een functie, equivalent met $\chi(x,.)$. $x_{(1)}$ geeft de kleinste, $x_{(n)}$ de grootste component van $x = (x_1, x_2, \ldots, x_n)$ aan.

STELLING 1:

<u>Voor families $\mathcal{F}$ van simultane dichtheden bij steekproeven groot n uit verdelingen met dragers van de typen 1, 2 en 3 geldt achtereenvolgens:</u>

<u>1. $x_{(n)}$ is equivalent met $\chi(x,.)$;</u>

<u>2. $x_{(1)}$ is equivalent met $\chi(x,.)$;</u>

<u>3. $(x_{(1)}, x_{(n)})$ is equivalent met $\chi(x,.)$.</u>

BEWIJS:

Het geval 3 zal alleen bewezen worden. De andere zijn dan direct duidelijk.

Zij $\chi(x,.) = \chi(x',.)$ en stel $x_{(n)} < x'_{(n)}$. Er is dan een $\theta = (\theta_1,\theta_2)$, met $x_{(n)} < \theta_2 \leqq x'_{(n)}$, en $\theta_1 < \min(x_{(1)}, x'_{(1)})$.
Dan is $\chi(x,\theta) = 1$ en $\chi(x',\theta) = 0$, wat niet mogelijk is. Dus is $x_{(n)} \geqq x'_{(n)}$ en evenzo $x_{(n)} \leqq x'_{(n)}$, dus $x_{(n)} = x'_{(n)}$. Voor $x_{(1)}$ geldt een analoge redenering.

Laat nu $x_{(1)} = x'_{(1)}$, $x_{(n)} = x'_{(n)}$. Voor alle $\theta = (\theta_1,\theta_2)$ geldt dan hetzij $\theta_1 < x_{(1)} \leqq x_{(n)} < \theta_2$ en $\theta_1 < x'_{(1)} \leqq x'_{(n)} < \theta_2$ zodat $\chi(x,\theta) = \chi(x',\theta) = 1$ òf er is een component $x_{(1)} = x'_{(1)}$ of $x_{(n)} = x'_{(n)}$ buiten het interval (θ_1,θ_2) en $\chi(x,\theta) = \chi(x',\theta) = 0$.
Dit houdt juist in dat x en x' equivalent zijn ten opzichte van $\chi(x,.)$.

VOORBEELDEN:

1. $x_{(1)}$ is noodzakelijk voor steekproeven van de grootte n uit:
   exponentiële verdelingen: $\lambda e^{-\lambda(x-\theta)}$ op (θ,∞) en 0 elders,
   lognormale verdelingen: $\{(x-\theta)\sigma\sqrt{2\pi}\}^{-\frac{1}{2}} \exp[-\frac{1}{2\sigma^2}\{\log(x-\theta)-\mu\}^2]$
   voor $x \in (\theta,\infty)$ en 0 elders.
2. $x_{(n)}$ is noodzakelijk en afdoende voor steekproeven van de grootte n uit homogene verdelingen: $H(0,\theta) = \frac{1}{\theta}$ op $(0,\theta)$ en 0 daarbuiten.
   Noodzakelijkheid van $x_{(n)}$ is zojuist gevonden. Afdoendheid blijkt als volgt: de (afdoende) aannemelijkheidsfunctie $L(.|x)$ is equivalent met $\chi(x,.)$.
3. $(x_{(1)}, x_{(n)})$ is noodzakelijk en afdoende voor steekproeven groot n uit de homogene verdeling:
   $H(\theta_1,\theta_2) = \frac{1}{\theta_2 - \theta_1}$ op (θ_1,θ_2) en 0 daarbuiten. $L(.|x)$ is ook hier equivalent met $\chi(x,.)$.

3.3.2. Exponentiële families

Families van dichtheden met niet-constante drager noemen we exponentiële families, als er lineair onafhankelijke functies $c_1(x)$, $c_2(x)$, ..., $c_r(x)$ op X bestaan, zó dat

$$f(x|\theta,\eta) = \chi(x,\eta)\exp\left[\phi(\theta,\eta) + c(x) + \sum_{i=1}^{r} c_i(x)\phi_i(\theta,\eta)\right].$$

We nemen hierin dragers van de typen 1, 2 of 3 uit de vorige paragraaf. Bovendien veronderstellen we dat $c_1(x)$, $c_2(x)$, ..., $c_r(x)$ een lineair onafhankelijk stelsel is op elke X_η. Dan is ook $\phi_1(\theta,\eta)$, $\phi_2(\theta,\eta)$, ..., $\phi_r(\theta,\eta)$ voor elke η lineair onafhankelijk.

BEWERING:

$(c_1(x), c_2(x), \ldots, c_r(x), x_{(n)})$ is noodzakelijk en afdoende voor een exponentiële familie $\mathcal{F}$ van dichtheden met dragers van het type 1 uit 3.3.1.

$(c_1(x), c_2(x), \ldots, c_r(x), x_{(1)})$ is noodzakelijk en afdoende voor een exponentiële familie met dragers van het type 2.

Zijn de dragers van type 3, dan is $(c_1(x), c_2(x), \ldots, c_r(x), x_{(1)}, x_{(2)})$ noodzakelijk en afdoende voor $\mathcal{F}$.

BEWIJS:

De tweede bewering wordt als voorbeeld gekozen.

1. $x_{(1)}$ is noodzakelijk. Ook elke $c_i(x)$ is noodzakelijk:

Neem een disjuncte splitsing van X door dragers $(X_1^*, X_2^*, \ldots)$. Is nu $l(.|x) = l(.|x')$ dan is er precies één k zodat x en $x' \in X_k^*$ en $X_k^* \subset X_{\eta_k}$ voor de index η_k die bij de constructie van $(X_1^*, X_2^*, \ldots)$ gebruikt is. Is nu $\eta < \eta_k$ dan is ook $X_k^* \subset X_\eta$.

$$l(\theta,\eta|x) = \phi(\theta,\eta) - \phi(\theta_k,\eta_k) + \sum_{i=1}^{r} c_i(x)\{\phi_i(\theta,\eta) - \phi_i(\theta_k,\eta_k)\} \text{ op } X_k^* \quad \ldots(a)$$

Bovendien is $l(\theta,\eta|x') = l(\theta,\eta|x)$ en dus

$$\sum_{i=1}^{r} (c_i(x) - c_i(x'))(\phi_i(\theta,\eta) - \phi_i(\theta_k,\eta_k)) = 0 \text{ op } X_k^*.$$

Uit de lineaire onafhankelijkheid van de $\phi_i(\theta,\eta)$ volgt nu:

$c_i(x) = c_i(x')$ op X_k^* voor i = 1, 2, ..., r en voor k = 1, 2, ...

2. Zij $(c_1(x), c_2(x), \ldots, c_r(x), x_{(1)}) = (c_1(x'), c_2(x'), \ldots$ $\ldots, c_r(x'), x'_{(1)})$. Daar $x_{(1)} = x'_{(1)}$ liggen x en x' in dezelfde verzameling X_k^*, en bovendien voor alle η of beide in X_η of beide buiten X_η. Zijn x en $x' \notin X_\eta$, dan is $l(\theta,\eta|x) = l(\theta,\eta|x') = -\infty$ voor alle θ. Stel x en $x' \in X_\eta$. Op $X_\eta \cap X_k^*$ is dan $l(\theta,\eta|x)$ zowel als $l(\theta,\eta|x')$ van de vorm (a) en dus gelijk voor x en x'. Hiermee is de afdoendheid van $(c_1(x), c_2(x), \ldots, c_r(x), x_{(1)})$ bewezen.

VOORBEELDEN:

1. $(\sum_{i=1}^{n} x_i, x_{(1)})$ is noodzakelijk en afdoende voor steekproeven van de grootte n uit een exponentiële verdeling $\lambda e^{-\lambda(x-a)}$ met $a \in (0,\infty)$, $\lambda \in (0,\infty)$, $x > a$, en 0 voor $x \leq a$.
2. Voor steekproeven groot n uit een verdeling $C(\theta) \cdot e^{-\theta_1(x-a)^{\theta_2}}$ als $x > a$ en 0 als $x \leq a$ ($\theta = (\theta_1, \theta_2)$, $\theta_1 > 0$, $\theta_2 > 0$, $a \in [0,\infty)$) is $(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ noodzakelijk en afdoende. Dit volgt direct uit een beschouwing van de deel-familie met $a = 0$ en $\theta_1 = 1$, waardoor 2.8 voorbeeld 3 wordt verkregen. Voor deze deel-familie is $(x_{(1)}, \ldots, x_{(n)})$ noodzakelijk, dus zeker voor de gehele familie. Tenslotte is $(x_{(1)}, \ldots, x_{(n)})$ steeds afdoende.
3. Is in voorbeeld 4. θ_2 een natuurlijk, vast getal, dan is $(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2, \ldots, \sum_{i=1}^{n} x_i^{\theta_2}, x_{(1)})$ noodzakelijk en afdoende. Is tenslotte ook θ_1 constant, dan is $(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2, \ldots, \sum_{i=1}^{n} x_i^{\theta_2 - 1}, x_{(1)})$ noodzakelijk en afdoende.

Voor een andere benadering zie men DYNKIN [7].

4. Geraadpleegde literatuur

[1] Bahadur, R.R. (1954): Sufficiency and statistical decision functions.
AMS, 25, 423-462.

[2] Bahadur, R.R. (1955): Statistics and subfields.
AMS, 26, 139-141.

[3] Barankin, E.W. (1949): Locally best unbiased estimates.
AMS, 20, 477-501.

[4] Barankin, E.W. en Maitra, A.P. (1963): Generalisations of the Fisher-Darmois-Koopman-Pitman theorem on sufficient statistics.
Sankhyā (A), 25, 217-244.

[5] Brown, L. (1964): Sufficient statistics in the case of independent random variables.
AMS, 35, 1456-1474.

[6] Courant, R. (1956): Differential and integral calculus, vol. II.
Blackie & son, London-Glasgow.

[7] Dynkin, E.B. (1951, vertaling 1961): Necessary and sufficient statistics for a family of probability distributions.
Select. Transl. Math. Statist. and Probab., 1, 23-41.

[8] Fraser, D.A.S. (1957): Nonparametric methods in statistics.
Wiley & sons, New York.

[9] Fraser, D.A.S. (1963): On sufficiency and the exponential family.
J. of Royal Stat. Soc., 25 (1), 115-123.

[10] Fraser, D.A.S. (1966): Sufficiency for regular models.
Sankhyā, 28, 137-144.

[11] Halmos, P.R. en Savage, L.J. (1949): Application of the Radon-Nikodym theorem to the theory of sufficient statistics.
AMS, 20, 225-241.

[12] Hogg, R.V. en Craig, A.T. (1956): Sufficient statistics in elementary distribution theory.
Sankhyā, 17, 209-220.

[13] Hogg, R.V. en Craig, A.T. (1959): Introduction to mathematical statistics.
Macmillan, New York.

[14] Kendall, M.G. en Stuart, A. (1961): The advanced theory of statistics, vol. 2.
Griffin, London.

[15] Lehmann, E.L. (1964): Testing statistical hypotheses.
Wiley & sons, New York.

[16] Pervin, W.J. (1964): Foundations of general topology.
Academic press, New York-London.

[17] Rasch, G. (1960): Probabilistic models for some intelligence and attainment tests.
Danmarks paedagogiske institut, studies in math. psych. I.
Nielsen & Lydiche, Kopenhagen.

[18] Richter, H. (1956): Wahrscheinlichkeitstheorie.
Springer, Berlin-Göttingen-Heidelberg.

[19] Wilks, S.S. (1962): Mathematical statistics.
Wiley & sons, New York.

[20] Burkholder, D.L. (1961): Sufficiency in the undominated case.
AMS, 32, 1191-2000.