FISHER–NEYMANOV FAKTORIZACIJSKI IZREK

Martin Raič

Pomlad 2011

1 O zadostnosti

Fisher–Neymanov faktorizacijski izrek je pomemben rezultat pri študiju zadostnosti statistik. Za statistične modele na diskretnih prostorih ga ni težko dokazati. Tu ga bomo dokazali v zelo veliki splošnosti.

V inferenčni statistiki imamo opravka s statističnimi modeli, opažanji in statistikami. Sodoben pogled na statistični model je, da je to družina verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$, definiranih na merljivem prostoru $(\Omega, \mathscr{F})$. Opažanja se tipično predstavijo z določeno merljivo funkcijo X na tem merljivem prostoru; dana statistična spremenljivka je opazljiva, če je merljiva funkcija osnovnega opažanja X. Tak pogled še vedno precej izhaja iz prakse; s teoretičnega stališča pa je ugodneje, če opazljivost definiramo s σ-algebro *opazljivih dogodkov*, ki jo označimo denimo z $\mathscr{G}$: statistična spremenljivka Y je opazljiva, če je $\mathscr{G}$-merljiva.

Ideja zadostne statistike je, da se je pri statističnem sklepanju dovolj omejiti le nanjo. Tipično je to spet neka opazljiva spremenljivka, denimo Y, za katero velja, da je pogojna porazdelitev osnovnega opažanja X glede na Y neodvisna od θ. Toda spet vemo, da lahko namesto pogojne porazdelitve glede na slučajno spremenljivko gledamo pogojno porazdelitev glede na σ-algebro. Zato bomo tu gledali zadostne σ-algebre.

Definicija. Naj bo $(\Omega, \mathscr{G})$ merljiv prostor. Za σ-algebro $\mathscr{H} \subseteq \mathscr{G}$ pravimo, da je pri opažanjih iz $\mathscr{G}$ *zadostna* za družino verjetnostnih mer $\mathbb{P}_\theta, \theta \in \Theta$, definiranih na $\mathscr{G}$, če za vsak $G \in \mathscr{G}$ obstaja taka $\mathscr{H}$-merljiva statistična spremenljivka Y, da za vsak $\theta \in \Theta$ velja $Y = \mathbb{P}_\theta(G \mid \mathscr{H})$.

Opomba. Iz "standardne mašinerije" sledi, da v tem primeru tudi za vsako $\mathscr{G}$-merljivo in omejeno statistično spremenljivko X obstaja taka $\mathscr{H}$-merljiva statistična spremenljivka Y, da za vsak $\theta \in \Theta$ velja $Y = \mathbb{E}_\theta(X \mid \mathscr{H})$.

Pri zgornji definiciji ne zahtevamo *regularnosti*, ki pomeni, da so spremenljivke Y "prav sestavljene". Podrobnosti so podane v spodnji definiciji.

Definicija. Naj bo $(\Omega, \mathscr{F}, \mathbb{P})$ verjetnostni prostor, dani pa naj bosta še σ-algebri $\mathscr{G}, \mathscr{H} \subseteq \mathscr{F}$. Funkcija $Q\colon\ \Omega \times \mathscr{G} \to [0, 1]$ je *regularna pogojna verjetnost* na $\mathscr{G}$ glede na $\mathscr{H}$, če velja:

(1) Za vsak $\omega \in \Omega$ je preslikava $G \mapsto Q(\omega, G)$ verjetnostna mera.

(2) Za vsak $G \in \mathscr{G}$ slučajna spremenljivka, definirana po predpisu $Y(\omega) := Q(\omega, G)$, zadošča $Y = \mathbb{P}(G \mid \mathscr{H})$.

V duhu zgornje definicije lahko definiramo *regularno zadostnost*.

Definicija. Naj bo $(\Omega, \mathscr{G})$ merljiv prostor. Za σ-algebro $\mathscr{H} \subseteq \mathscr{G}$ pravimo, da je pri opažanjih iz $\mathscr{G}$ *regularno zadostna* za družino verjetnostnih mer $\mathbb{P}_\theta, \theta \in \Theta$, definiranih na $\mathscr{G}$, če obstaja funkcija Q, ki je regularna pogojna verjetnost glede na $\mathscr{H}$ pri vseh verjetnostnih merah $\mathbb{P}_\theta$.

Regularne pogojne verjetnosti ne obstajajo vedno – v grobem je mogoče reči, da obstajajo na “dovolj lepih” verjetnostnih prostorih (več o tem si lahko bralec pogleda v [1]). Včasih pa obstaja pogojna porazdelitev.

DEFINICIJA. Naj bo $(\Omega, \mathscr{F}, \mathbb{P})$ verjetnostni prostor, dana naj bo še σ-algebra $\mathscr{H} \subseteq \mathscr{F}$ in naj bo X slučajna spremenljivka z vrednostmi na merljivem prostoru $(S, \mathscr{S})$. Funkcija $Q\colon \Omega \times \mathscr{S} \to [0,1]$ je *pogojna porazdelitev* slučajne spremenljivke X glede na $\mathscr{H}$, če velja:

(1) Za vsak $\omega \in \Omega$ je preslikava $A \mapsto Q(\omega, A)$ verjetnostna mera.

(2) Za vsak $A \in \mathscr{S}$ slučajna spremenljivka, definirana po predpisu $Y(\omega) := Q(\omega, A)$, zadošča $Y = \mathbb{P}(X \in A \mid \mathscr{H})$.

Opomba. Regularna pogojna verjetnost na σ-algebri $\mathscr{G}$ je natanko regularna pogojna porazdelitev identitete, ki slika iz $(\Omega, \mathscr{F})$ v $(\Omega, \mathscr{G})$.

Opomba. Če je $(\Omega, \mathscr{F}, \mathbb{P})$ verjetnostni prostor, X slučajna spremenljivka, $\mathscr{H} \subseteq \mathscr{F}$ še ena σ-algebra in Q regularna pogojna verjetnost na $\mathscr{F}$ glede na $\mathscr{H}$, je s predpisom:

$$(\omega, A) \mapsto Q\big(\omega, \{X \in A\}\big)$$

določena pogojna porazdelitev slučajne spremenljivke X glede na $\mathscr{H}$.

Brž ko torej obstaja regularna pogojna verjetnost, obstaja tudi pogojna porazdelitev. Zanimivo pa je obratno vprašanje:

Če ima slučajna spremenljivka X pogojno porazdelitev glede na $\mathscr{H}$, ali obstaja regularna pogojna verjetnost na $\sigma(X)$ glede na $\mathscr{H}$?

Znano je, da pogojne porazdelitve obstajajo za slučajne spremenljivke, ki slikajo v *Borelove* prostore (glej npr. Shiryaev [3], izrek II.7.5, stran 229). Tako bi lahko zgornje vprašanje modificirali še na naslednji način:

Ali za vsako slučajno spremenljivka X, ki slika iz $(\Omega, \mathscr{F})$ v Borelov prostor, in vsako σ-algebro $\mathscr{H} \subseteq F$ obstaja regularna pogojna verjetnost na $\sigma(X)$ glede na $\mathscr{H}$?

S tovrstnimi vprašanji se tukaj ne bomo ukvarjali. Namesto tega bomo raje definirali še eno različico regularne zadostnosti.

DEFINICIJA. Naj bo $(\Omega, \mathscr{F})$ merljiv prostor in X merljiva preslikava, ki slika iz $(\omega, \mathscr{F})$ v $(S, \mathscr{S})$. Za σ-algebro $\mathscr{H} \subseteq \sigma(X)$ pravimo, da je pri opažanju X *regularno zadostna* za družino verjetnostnih mer $\mathbb{P}_\theta, \theta \in \Theta$, definiranih na $\mathscr{F}$, če obstaja funkcija Q, ki je regularna pogojna porazdelitev statistične spremenljivke X glede na $\mathscr{H}$ pri vseh verjetnostnih merah $\mathbb{P}_\theta$.

Opomba. Regularna zadostnost pri opažanjih iz σ-algebre $\mathscr{G}$ je natanko regularna zadostnost pri opažanju, podanem z identiteto iz $(\Omega, \mathscr{F})$ v $(\Omega, \mathscr{G})$.

Če je torej $\mathscr{H}$ regularno zadostna pri $\sigma(X)$, je regularno zadostna tudi pri X; če velja slednje, pa je $\mathscr{H}$ zadostna pri $\sigma(X)$. Eden od zadostnih pogojev za regularno zadostnost je podan v naslednji trditvi, še enega pa bomo formulirali kasneje.

Definicija. Množica A je *povsem ničelna* v okviru σ-algebre $\mathscr{H}$ za družino mer μ_θ, $\theta \in \Theta$, definiranih na $\mathscr{H}$, če za vsak $\theta \in \Theta$ obstaja tak $H \in \mathscr{H}$, da je $A \subseteq H$ in $\mu_\theta(H) = 0$.

Trditev 1.1. *Naj bo $(\Omega, \mathscr{F})$ merljiv prostor z družino verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$, in naj bo X statistična spremenljivka, ki slika v merljiv prostor s števno generirano σ-algebro. Označimo $\mathscr{G} := \sigma(X)$. Dana naj bo še σ-algebra $\mathscr{H} \subseteq \mathscr{G}$, ki naj vsebuje vse dogodke, ki so v njenem okviru povsem ničelni za verjetnostne mere $\mathbb{P}_\theta$. Nadalje naj za vsak $\theta \in \Theta$ pri verjetnostni meri $\mathbb{P}_\theta$ obstaja pogojna porazdelitev slučajne spremenljivke X glede na $\mathscr{H}$. Če je tedaj $\mathscr{H}$ pri $\mathscr{G}$ zadostna, je pri X tudi regularno zadostna.*

Posledica 1.2. *Naj bo $(\Omega, \mathscr{G})$ merljiv prostor z družino verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$, dana pa naj bo še σ-algebra $\mathscr{H} \subseteq \mathscr{G}$, ki naj vsebuje vse dogodke, ki so v njenem okviru povsem ničelni za verjetnostne mere $\mathbb{P}_\theta$. Nadalje naj bo $\mathscr{G}$ števno generirana in naj pri vseh verjetnostnih merah $\mathbb{P}_\theta$ obstaja regularna pogojna verjetnost na $\mathscr{G}$ glede na $\mathscr{H}$. Če je tedaj $\mathscr{H}$ pri opažanjih iz $\mathscr{G}$ zadostna, je pri teh opažanjih tudi regularno zadostna.*

Posledica 1.3. *Naj bo $(\Omega, \mathscr{F})$ merljiv prostor z družino verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$, in naj bo X statistična spremenljivka, ki slika v Borelov prostor. Označimo $\mathscr{G} := \sigma(X)$. Če je σ-algebra $\mathscr{H} \subseteq \mathscr{G}$ pri opažanjih iz $\mathscr{G}$ zadostna in če vsebuje vse dogodke, ki so v njenem okviru povsem ničelni za verjetnostne mere $\mathbb{P}_\theta$, je $\mathscr{H}$ pri opažanju X tudi regularno zadostna.*

Vprašanje, ki ga puščamo odprto: ali iz zadostnosti in obstoja regularnih pogojnih verjetnosti vedno sledi regularna zadostnost?

Dokaz trditve 1.1. Ker je $\mathscr{S}$ števno generirana, obstaja tudi največ števna družina $\mathscr{T} \subseteq \mathscr{S}$, ki je zaprta za končne preseke in generira $\mathscr{S}$. Zaradi zadostnosti za vsak $A \in \mathscr{T}$ obstaja taka $\mathscr{H}$-merljiva statistična spremenljivka Y_A, da je $Y_A = \mathbb{P}_\theta(X \in A \mid \mathscr{H})$ za vse $\theta \in \Theta$. Nadalje za vsak $\theta \in \Theta$ s Q_θ označimo pogojno porazdelitev slučajne spremenljivke X glede na $\mathscr{H}$ pri verjetnostni meri $\mathbb{P}_\theta$. Zaradi števnosti za vse ω iz nekega dogodka F_θ s $\mathbb{P}_\theta(F_\theta) = 1$ velja $Y_A(\omega) = Q_\theta(\omega, A)$ za vse $A \in \mathscr{T}$. To pomeni, da se da za vse $\omega \in F := \bigcup_{\theta \in \Theta} F_\theta$ preslikava $A \mapsto Y_A(\omega)$ razširiti do verjetnostne mere na $(S, \mathscr{S})$. Ker je množica $\mathscr{T}$ zaprta za končne preseke, so te razširitve enolične. Označimo jih z $A \mapsto Q(\omega, A)$, za $\omega \notin F$ pa definirajmo $Q(\omega, A) := \mathbf{1}(x \in A)$, kjer je x neka točka iz S. Ker $\mathscr{H}$ vsebuje vse množice, ki so popolnoma ničelne v $\mathscr{H}$ glede na verjetnostne mere $\mathbb{P}_\theta$, za vsak $A \in \mathscr{T}$ še vedno velja $(\omega \mapsto Q(\omega, A)) = \mathbb{P}(X \in A \mid \mathscr{H})$. Toda ker je $\mathscr{T}$ zaprta za preseke in generira $\mathscr{S}$, to velja tudi za vse $A \in \mathscr{S}$. To pa pomeni, da je Q iskana pogojna porazdelitev, ki je neodvisna od θ. ■

2 Zadostnost in faktorizacija

Fisher–Neymanov faktorizacijski izrek nam pove, da zadostnost σ-algebre $\mathscr{H}$ pomeni, da imajo verjetnostne mere $\mathscr{H}$-merljive Radon–Nikodymove odvode glede na neko verjetnostno mero, torej so glede na njo tudi absolutno zvezne. Absolutna zveznost glede na verjetnostno mero pa ne sledi avtomatično iz zadostnosti (vsaka σ-algebra je zadostna pri opažanjih iz same sebe), torej jo bomo morali dodatno privzeti. Pokazali bomo, da je dovolj privzeti absolutno zveznost glede na neko σ-končno mero. Za ta namen bomo potrebovali naslednji rezultat, katerega dokaz si bralec lahko pogleda v [2] (izrek A.78, str. 600):

Izrek 2.1 (Halmos, Savage, 1949). *Naj bo μ_θ, $\theta \in \Theta$, družina mer na merljivem prostoru $(\Omega, \mathscr{F})$, ki so vse absolutno zvezne glede na neko σ-končno mero. Tedaj obstajajo taka števila $c_\theta \geq 0$, $\theta \in \Theta$, da je $\sum_{\theta\in\Theta} c_\theta = 1$ in da so mere μ_θ absolutno zvezne tudi glede na mero $\sum_{\theta\in\theta} c_\theta \mu_\theta$.*

Opomba. Če je $\sum_{\theta\in\Theta} c_\theta = 1$, to pomeni, da je kvečjemu števno mnogo števil c_θ različnih od nič.

Opomba. Da je tudi $\nu^* := \sum_{\theta\in\theta} \mu_\theta$ mera, sledi iz Fubinijevega izreka. Če so vse mere μ_θ verjetnostne, je to tudi ν^*.

Izrek 2.2 (Fisher–Neymanov faktorizacijski izrek). *Naj bo $(\Omega, \mathscr{G})$ merljiv prostor z družino verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$. Dana naj bo še σ-algebra $\mathscr{H} \subseteq \mathscr{G}$. Tedaj so naslednje trditve ekvivalentne:*

(1) *Mere $\mathbb{P}_\theta$ so vse absolutno zvezne glede na neko σ-končno mero in $\mathscr{H}$ je pri opažanjih iz $\mathscr{G}$ zadostna.*

(2) *Obstaja σ-končna mera ν^* ter funkcije Z_θ, $\theta \in \Theta$, ki so $\mathscr{H}$-merljive in so Radon–Nikodymovi odvodi mer $\mathbb{P}_\theta$ glede na ν^*, gledano na $\mathscr{G}$.*

(3) *Obstaja verjetnostna mera $\mathbb{P}^*$ ter funkcije Z_θ, $\theta \in \Theta$, ki so $\mathscr{H}$-merljive in so Radon–Nikodymovi odvodi mer $\mathbb{P}_\theta$ glede na $\mathbb{P}^*$, gledano na $\mathscr{G}$.*

DOKAZ.

(1) $\Rightarrow$ (3). Po izreku 2.1 obstajajo taka števila $c_\theta \geq 0$, da je $\sum_{\theta\in\Theta} c_\theta = 1$ in da so mere $\mathbb{P}_\theta$ absolutno zvezne glede na mero $\mathbb{P}^* := \sum_{\theta\in\Theta} \mathbb{P}_\theta$. Za vsak $\theta \in \Theta$ naj bo Z_θ Radon–Nikodymov odvod mere $\mathbb{P}_\theta$ glede na $\mathbb{P}^*$, gledano na $\mathscr{H}$. Pokazali bomo, da je to tudi Radon–Nikodymov odvod na $\mathscr{G}$.

Vzemimo poljuben dogodek $G \in \mathscr{G}$. Zaradi zadostnosti obstaja statistična spremenljivka Y, za katero je $Y = \mathbb{P}_\theta(G \mid \mathscr{H})$ za vsak $\theta \in \Theta$. Če z $\mathbb{E}_\theta$ in $\mathbb{E}^*$ označimo matematično upanje glede na $\mathbb{P}_\theta$ oziroma $\mathbb{P}^*$, sledi:

$$\mathbb{P}_\theta(G) = \mathbb{E}_\theta\, Y = \mathbb{E}^*\, Y Z_\theta = \sum_{t\in\Theta} c_t\, \mathbb{E}_t\, Y Z_\theta = \sum_{t\in\Theta} c_t\, \mathbb{E}_t\, Z_\theta\, \mathbf{1}_G = \mathbb{E}^*\, Z_\theta\, \mathbf{1}_G\,,$$

torej je Z_θ res Radon–Nikodymov odvod tudi na $\mathscr{G}$.

(3) $\Rightarrow$ (2). Očitno.

(2) $\Rightarrow$ (1). Naj bo $G \in \mathscr{G}$. Konstruirati moramo tako statistično spremenljivko Y, da bo $Y = \mathbb{P}_\theta(G \mid \mathscr{H})$ za vse θ. Ideja je, da, če je $\nu^* = \mathbb{P}^*$ kar verjetnostna mera, vzamemo $Y = \mathbb{P}^*(G \mid \mathscr{H})$. V splošnem pa za Y vzamemo Radon–Nikodymov odvod mere $H \mapsto \nu^*(G \cap H)$ glede na ν^*, gledano na $\mathscr{H}$. Tedaj za vsak $\mathscr{H}$-merljiv dogodek H velja:

$$\mathbb{P}_\theta(G \cap H) = \int_{G \cap H} Z_\theta \, \mathrm{d}\nu^* = \int_H Y Z_\theta \, \mathrm{d}\nu^* = \mathbb{E}_\theta \, Y \, \mathbf{1}_H \,,$$

torej je res $Y = \mathbb{P}_\theta(G \mid \mathscr{H})$. ■

Posledica 2.3. *Naj bo $(\Omega, \mathscr{F})$ merljiv prostor z družino verjetnostnih mer $\mathbb{P}_\theta$, $\theta \in \Theta$, ki so absolutno zvezne glede na neko σ-končno mero. Dana naj bo še σ-algebra $\mathscr{H} \subseteq \sigma(X)$. Nadalje naj X statistična spremenljivka, za katero pri vsaki verjetnostni meri $\mathbb{P}$ na $(\Omega, \mathscr{F})$ obstaja pogojna porazdelitev glede na $\mathscr{H}$. Če je tedaj $\mathscr{H}$ pri $\sigma(X)$ zadostna, je pri X tudi regularno zadostna.*

DOKAZ. Po Fisher–Neymanovem faktorizacijskem izreku obstaja verjetnostna mera $\mathbb{P}^*$ ter funkcije Z_θ, $\theta \in \Theta$, ki so $\mathscr{H}$-merljive in so Radon–Nikodymovi odvodi mer $\mathbb{P}_\theta$ glede na $\mathbb{P}^*$, gledano na $\mathscr{G}$. Naj bo Q pogojna porazdelitev slučajne spremenljivke X glede na $\mathscr{H}$ pri meri $\mathbb{P}^*$. Pokazali bomo, da to velja tudi pri merah $\mathbb{P}_\theta$. Dovolj je pokazati, da za vsak relevanten A velja $(\omega \mapsto Q(\omega, A) = \mathbb{P}_\theta(X \in A \mid \mathscr{H})$. To pa je res, ker za vsak $\mathscr{H}$-merljiv dogodek H velja:

$$\mathbb{P}_\theta\big(\{X \in A\} \cap H\big) = \mathbb{E}^* \, Z_\theta \, \mathbf{1}_{\{X \in A\} \cap H} = \int_H Q(\omega, A) \, Z_\theta(\omega) \; \mathbb{P}^*(\mathrm{d}\omega) = \int_H Q(\omega, A) \; \mathbb{P}_\theta(\mathrm{d}\omega) \,.$$

■

Literatura

[1] A. M. Faden: *The existence of regular conditional probabilities: necessary and sufficient conditions*, Ann. Probab. **13** (1985), 288–298.

[2] M. J. Schervish: *Theory of statistics*, Springer Series in Statistics, Springer-Verlag, New York, 1995.

[3] A. N. Shiryaev: *Probability*, Graduate Texts in Mathematics, vol. 95, Springer-Verlag, New York, 1996.