Rakenteiden Mekaniikka
Vol. 41, Nro. 2, 2008, s. 86 – 89

Ehdollistaminen ja riippuvuus

Ilkka Norros

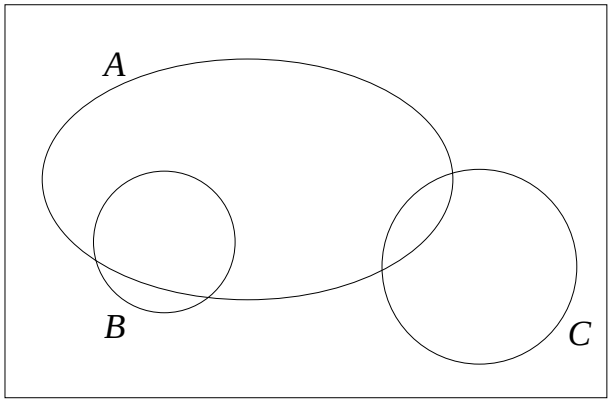
Tiivistelmä. Kirjoitus perustuu Rakenteiden mekaniikan seuran seminaarissa huhtikuussa 2007 pidettyyn esitelmään, jonka tarkoituksena oli virittää kuulijoiden mieliin todennäköisyyslaskennan peruskäsitteitä ja ajattelutapaa.

Avainsanat: todennäköisyys, uhka, informaatio

Ehdollinen todennäköisyys ja informaatio

Todennäköisyyslaskenta kuvaa satunnaista ilmiötä kiinnittämällä ensin jonkin *alkeistapahtumien* joukon Ω, joka sisältää ilmiön kaikki mahdolliset tapaukset. Jos esimerkiksi kysymyksessä on satunnainen jatkuva funktio, Ω:ksi on luontevaa valita jokin funktioavaruus X; jos tuohon funktioon vaikuttaa lisäksi yksi rahanheitto, valitaan $\Omega = X \times \{0, 1\}$, jne. *Todennäköisyys* on Ω:n mitallisille osajoukoille eli *tapahtumille* A määritelty *mitta* $\mathbb{P}(A)$, jolle $\mathbb{P}(\Omega) = 1$.

Todennäköisyyslaskennalle ominaisissa kysymyksenasetteluissa keskeinen rooli on *ehdollisilla* todennäköisyyksillä. Tähän liittyviä käsitteitä ovat myös (tilastollinen) riippuvuus ja informaatio.



$\mathbb{P}(A)$ on keskikokoinen

$\mathbb{P}[A \,|\, B]$ on suuri

$\mathbb{P}[A \,|\, C]$ on pieni

Kuva 1. Ehdollisen todennäköisyyden määritelmä.

Todennäköisyyksien ehdollistaminen on sen huomioon ottamista, mitä tiedetään. Ehdollisen todennäköisyyden määritelmän mukaan tapahtuman A ehdollinen todennäköisyys ehdolla B (tai tapahtuman B suhteen) on (ks. kuva)

$$\mathbb{P}[A \,|\, B] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Ehdollinen todennäköisyys yhtyy ehdottomaan kun tapahtumat A ja B ovat *riippumattomia* eli kun $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$. Informaatioteoria antaa tapahtuman B tapahtumasta A antamalle *informaatiolle* luontevan numeerisen arvon, joka on 0 täsmälleen silloin kun tapahtumat ovat riippumattomia. Sivuutan tähän liittyvät kaavat, mutta intuitiivista informaatio-terminologiaa käytän jatkossakin.

Esimerkki 1. Seuraavan laatikkoleikin analysoinnista kiisteltiin taannoin amerikkalaisilla tiedepalstoilla. Näyttämöllä on kolme laatikkoa, joista yhdessä on palkinto ja muut ovat tyhjiä. Kilpailija valitsee yhden laatikon ja osoittaa sitä. Silloin juontaja avaa toisen ei-valituista laatikoista. Se osoittautuu tyhjäksi. Juontaja kysyy nyt, haluaako kilpailija vaihtaa valintaansa. Kannattaako hänen vaihtaa? Merkitään A:lla tapahtumaa 'valitussa laatikossa on palkinto' ja B:llä tapahtumaa 'juontajan avaama laatikko on tyhjä'. Ongelman ratkaisu edellyttää oikein tehtyä todennäköisyyden ehdollistamista. Itse asiassa se riippuu myös juontajan toiminnan *tulkitsemisesta* — seuraavista vaihtoehdoista ensimmäinen on oikea, mutta muutkin olisivat mahdollisia.

- Jos juontaja avaa aina tyhjän laatikon, valinta kannattaa vaihtaa: tällöin nimittäin B ei anna mitään informaatiota A:sta, joten $\mathbb{P}[\,A \,|\, B] = \mathbb{P}(A) = 1/3$.
- Jos juontaja valitsee avattavan umpimähkään, vaihtamisesta ei ole enempää hyötyä kuin haittaakaan, sillä $\mathbb{P}[\,A \,|\, B] = 1/2$.
- Jos juontaja avaa palkinnon sisältävän laatikon aina kun voi, vaihtaminen ei kannata, sillä $\mathbb{P}[\,A \,|\, B] = 1$.

Taaksepäin päättely: inversio, Bayesin kaava

Syy-seuraus-suhteet ovat reaalimaailmassa yleensä ainakin jossain määrin satunnaisia, ts. 'syy' nostaa 'seurauksen' todennäköisyyttä muttei määrää sitä ehdottoman varmasti. Seurausten ehdolliset todennäköisyydet syyn suhteen on usein helpompi arvioida kuin syyn päätteleminen seurauksista. Periaattellisen ratkaisun tähän 'inversio-ongelmaan' antaa Bayesin kaava, joka yksinkertaisimmassa muodossaan kuuluu:

$$\mathbb{P}[\,A \,|\, B] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}[\,B \,|\, A]\,\mathbb{P}(A)}{\mathbb{P}[\,B \,|\, A]\,\mathbb{P}(A) + \mathbb{P}[\,B \,|\, \Omega \setminus A]\,\mathbb{P}(\Omega \setminus A)}.$$

Esimerkki 2. Valitaan edellisessä kaavassa $A =$ 'kattorakenteet tehty hyvin', $B =$ 'katto romahtaa'.

Bayesläinen tilastotiede (ks. esim. [3]) lähtee ajatuksesta, että kaikella tuntemattomalla on todennäköisyys, joka tulkitaan subjektiiviseksi käsitykseksi asioiden tilasta. Havainnot muuttavat tätä käsitystä, ja uudet, korjatut todennäköisyydet lasketaan Bayesin kaavalla. Yleisessä tapauksessa kaavan nimittäjänä on integraalilauseke.

Esimerkki 3. Tarkastellaan 'vääntyneen lantin' heittoa, jossa kruunan todennäköisyys on tuntematon luku $Q \in [0,1]$. Olkoon subjektiivinen jakaumamme Q:lle aluksi tasainen jakauma $\mathrm{Tas}(0,1)$. Rahaa heittämällä saadaan Q:sta tarkentuvaa tietoa. Jos merkitään $B_k^n =$ 'n heitossa saadaan k kruunaa', Q:lle saadaan ehdollinen jakauma

$$\mathbb{P}[Q \in \mathrm{d}q \,|\, B_k^n] = \frac{\binom{n}{k} q^k (1-q)^{n-k}\,\mathrm{d}q}{\displaystyle\int_0^1 \binom{n}{k} r^k (1-r)^{n-k}\,\mathrm{d}r} = const \cdot q^k (1-q)^{n-k}\,\mathrm{d}q,$$

jonka tiheys on suurimmillaan kohdassa $q = k/n$.

Uhkakäsitteet

Satunnaisen ajanhetken T koittamista on usein osuvaa kuvata *uhkafunktiolla* (hazard rate; ks. esim. [2])

$$r(t) = \frac{\mathbb{P}[T \in [t, t + \mathrm{d}t] \,|\, T \geq t]}{\mathrm{d}t} = \frac{f(t)}{1 - F(t)},$$

missä T:n kertymäfunktio on F ja tiheysfunktio $f = F'$. Uhkafunktion arvo $r(t)$ kertoo, millä *intensiteetillä* kyseinen hetki (uhasta puhuttaessa usein ikävä hetki, kuten vaurioituminen tai kuolema) pyrkii koittamaan seuraavassa silmänräpäyksessä eli aikavälillä $[t, t + \mathrm{d}t]$ sillä ehdolla, että ennen hetkeä t se ei ole vielä koittanut.

Eri T:n jakaumilla uhkafunktio on erityyppinen:

- tasaisella jakaumalla kasvava
- eksponenttijakaumalla vakio
- Pareton jakaumalla laskeva
- ihmisen eliniän jakaumalla ensin laskeva, sitten kasvava; lisäksi 15-20 vuoden tienoilla uhka käy monissa maissa väliaikaisesti korkeammalla.

Uhka voidaan määritellä myös stokastisena prosessina, ja ehdollistaminen voidaan tällöin tehdä erilaisten *historioiden* suhteen:

$$\mathrm{d}R_t = \mathbb{P}[T \in [t, t + \mathrm{d}t] \,|\, \mathcal{F}_{t-}],$$

missä $\mathcal{F}_t$:llä merkitään tarkasteltavan prosessin koko historiaa ennen hetkeä t. Historiatieto voi olla tarkempaa tai ylimalkaisempaa, ja sen rooli uhkaprosessissa on oleellinen. Perinteinen uhkafunktio $r(t)$ liittyy tällöin siihen tapaukseen, että $\mathcal{F}_t$ sisältää tiedon vain siitä, onko hetki T jo koittanut vai ei. Mitä rikkaamman historian suhteen ehdollistamalla uhka lasketaan, sitä osuvammin se vastaa todellista, 'fysikaalista' uhkaa kyseisenä ajanhetkenä.

Riippuvuudesta

Satunnaisilmiöiden riippumattomuus on vahva ja selkeä käsite: riippumattomat tapahtumat eivät anna toisistaan mitään informaatiota. Tässä yhteydessä on hyvä muistuttaa siitä, että satunnaismuuttujien korreloimattomuus on *paljon heikompi* ehto kuin niiden riippumattomuus.

Esimerkki 4. Olkoon X:llä standardinormaalijakauma $N(0, 1)$, ja olkoon $Y = X$ tai $Y = -X$ todennäköisyyksillä $\frac{1}{2} - \frac{1}{2}$ riippumatta X:n arvosta. Tällöin X:n ja Y:n välinen korrelaatio on nolla (ts. $\mathbb{E}\{XY\} - \mathbb{E}\{X\}\,\mathbb{E}\{Y\} = 0$), vaikka X ja Y sisältävät toisistaan äärettömän paljon informaatiota kertoessaan toistensa itseisarvon tarkalleen.

Uhkaprosessit antavat mahdollisuuden kuvata 'stokastisesti kausaalisia' riippuvuuksia dynaamisesti: yhden tapahtuman sattuminen voi vaikuttaa toisen tapahtuman uhkaan (ks. esim. [1]). Tällainen vaikutus voi olla epäsymmetristä kuten luonnonkin kausaalisuus: auringonpilkut saattavat vaikuttaa joihinkin ilmiöihin maapallon elämässä, mutta maapallon elämä ei vaikuta auringonpilkkuihin mitään. Tilastollinen riippuvuus tarjoaa sen sijaan aina molemminpuolista informaatiota: jos auringonpilkut antavat informaatiota maapallon tapahtumista, myös maapallon tapahtumat antavat informaatiota auringonpilkkujen esiintymisestä.

Viitteet

[1] E. Arjas and I. Norros. Stochastic Order and Martingale Dynamics in Multivariate Life Length Models: A Review. Teoksessa K. Mosler and M. Scarsini (toim.): *Stochastic Orders and Decision Under Risk*. IMS Lecture Notes-Monograph Series, Vol. 19, 1991.

[2] R.E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart and Winston, 1975.

[3] A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin. *Bayesian Data Analysis*. Chapman and Hall, 1995.

Ilkka Norros
VTT, Tietoverkkojen suorituskyky
PL 1000, 02044 VTT
s-posti: `ilkka.norros@vtt.fi`