

# Chebysheven desberdintza esponentziala: probabilista baten orotariko giltza

*Joseba Dalmau*

Université Paris Sud eta ENS

dalmau@dma.ens.fr

Jasota: 15/05/2013

Onartua: 14/07/2013

**Laburpena:**  $X$  zorizko aldagai baterako Chebysheven desberdintza esponentziala aurkeztuko dugu:

$$\forall t \geq 0, \forall \lambda \geq 0 \quad P(X \geq t) \leq e^{-\lambda t} E(e^{\lambda X}).$$

Emaitza honen garrantzia agerian jartzen saiatuko gara, probabilitate eta estatistikaren munduko hainbat arlotan nola agertzen den azalduz. Esate baterako, kontzentrazio desberdintzetan edo desbiderapen handien teorian. Bukatzeko, desbiderapen handien teoria erabiliko dugu mekanika estatistikoko eredu bat aztertzeko.

**Abstract:** In this expository paper we will present Chebyshev's exponential inequality, which for a random variable  $X$  takes the form:

$$\forall t \geq 0, \forall \lambda \geq 0 \quad P(X \geq t) \leq e^{-\lambda t} E(e^{\lambda X}).$$

We will try to convince the reader of the importance of this result by showing how it naturally appears in many areas of probability theory and statistics. For instance, in concentration inequalities and in the theory of large deviations. To finish with, we will use the theory of large deviations to study a model coming from statistical mechanics.

## 1. SARRERA

### Zorizko aldagaiak

Estatistikari eta probabilistak zorizko aldagaiak eta hauen propietateak aztertzen ditugun matematikariak gara. Zorizko aldagai bat

$$X : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$$

aplikazio bat da. Jatorri-espazioa  $\Omega$  da,  $\mathcal{F}$  espazio horren gaineko tribu bat da eta  $P$ ,  $(\Omega, \mathcal{F})$  espazio neurgarrian definitutako probabilitate neurri bat.  $(\Omega, \mathcal{F}, P)$  hirukoteari probabilitate-espazio deritzo.  $X$  aplikazioak, zorizko aldagaia izateko, baldintza bakarra bete behar du:  $\mathbb{R}$ -ko tribu boreldarra-rekiko neurgarria izan behar du. Dei dezagun  $\mathcal{B}$   $\mathbb{R}$ -ko (ohiko topologiako) irekiek sortzen duten tribua;  $X$  neurgarria dela diogu edozein  $B \in \mathcal{B}$  hartuta  $X^{-1}(B) \in \mathcal{F}$  betetzen bada. Probabilitate arloko testuetan ohikoa den bezala,  $\{X \in B\}$  idatziko dugu,  $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$  multzoa adierazteko.

Teorian, zorizko aldagai bat funtzio neurgarri bat besterik ez da. Praktikan ordea, funtzio arrunt bat eta zorizko aldagai bat oso desberdinak dira. Har dezagun  $f: (x, y) \in \mathbb{R}^2 \mapsto x + y \in \mathbb{R}$  funtzioa esate baterako.  $\mathbb{R}^2$  jatorri-espazioko edozein puntu hartuta, zehatz-mehatz jakin dezakegu  $f$ -ren birtartez bere irudia zein den,  $f(1, 2) = 3$  edo  $f(1/2, -1) = -1/2$  adibideetan bezala. Zorizko aldagaiak, auzazko jarrera bat modelizatzeko matematikariok asmatutako tresnak dira. Emandako adibideak ez du horretarako balio; hori dela eta, tipikoki ez dugu zorizko aldagai bat  $X: \omega \in \Omega \mapsto X(\omega) \in \mathbb{R}$  funtzio zehatz eta ezagun bat bezala ulertuko. Aitzitik,  $\omega \in \Omega$  puntu baten irudia esplizituki eman gabe, auza kontutan hartu ahal izateko nahikoa informazio emango digu; orokorrean,  $I \in \mathbb{R}$  tarte bat bada,  $P(X \in I)$  moduko kantitateak kalkulatu ahal izango ditugu.

1. **ADIBIDEA.**  $X$   $p$  parametroko Bernoulli aldagai bat dela esaten dugu, eta  $X \sim \text{Ber}(p)$  adierazten dugu, hurrengo erlazioa betetzen bada:

$$P(X = 1) = 1 - P(X = 0) = p.$$

## Momentuak

Zorizko aldagaiak hobeto ulertzeko, hauekin lotura duten zenbait balio kalkulatu ohi dira. Jo dezagun  $X$  zorizko aldagai bat dela. Balio hauetako bat da bere itzaropena,

$$E[X] = \int_{\Omega} X dP,$$

baita era orokorrago batean  $X$ -ren  $k$ -garren momentu (absolutua) ere,

$$E[X^k] = \int_{\Omega} X^k dP,$$

non integralak Lebesgueren zentzuan ulertzen ditugun.  $X$ -ren  $k$ -garren momentu zentratua,  $E[(X - E[X])^k]$  ere erabilgarria da; bigarren mo-

mentu zentratua da gehien erabiltzen duguna, bariantza izenarekin ezagunagoa:

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2.$$

Nolabait esatearren, itxaropenak adierazten du zorizko aldagaiak zein balioren inguruan fluktuatzen duen. Bariantzak berriz, adierazten digu zorizko aldagaia bere itxaropenetik «hurbil» edo «urrun» pilatzen den.

### **Batezbesteko enpirikoa**

Gehienetan ez dugu zorizko aldagai bakar bat aztertzen; aldagai anitz hartzen ditugu aztergai, eta haien funtzioak aztertzen ditugu<sup>1</sup>. Argitu dezagun puntu hau adibide baten bidez:  $X \sim \text{Ber}(1/2)$  zorizko aldagaia txanpon bat jaurtitzearekin identifika dezakegu, baina noski, txanpona behin eta berriro bota dezakegu, eta luzaroko portaerari errepara diezaiokegu. Gisa honetako adibideak asko dira; hori dela eta, probabilitate-teoria klasikoak, sakon ikertu ditu zorizko aldagaien segidak,  $(X_n)_{n \geq 1}$ . Segida horien artean aztertuenak dira zorizko aldagai aske eta berdinki banatuen segidak, a.b.b. laburtuko ditugunak, txanpona hainbatetan jaurtiz lortzen dugun segidaren modukoak, izan ere. Segidei eurei erreparatu barik, hauen funtzioei edota hauetatik eratorritako segidei egiten diegu kasu, esaterako,

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n} \quad n \geq 1.$$

Honi,  $X_1, \dots, X_n$  aldagaien batezbesteko enpirikoa deritzo, eta probabilitate-teoriako emaitza handienen protagonista da. Emaitza hauen artean bi azpimarratu behar ditugu: zenbaki handien legeak eta limitearen teorema zentrala.

### **Teorema garrantzitsuak**

**2. TEOREMA** (Zenbaki handien legeak). *Izan bedi  $(X_n)_{n \geq 1}$  a.b.b. zorizko aldagaien segida.  $E[|X_1|] < \infty$  bada, orduan,*

**Zenbaki handien lege ahula.**  $\bar{X}_n \rightarrow E[X_1]$  probabilitatean. *Hau da,*

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P(|\bar{X}_n - E[X_1]| > \varepsilon) = 0.$$

---

<sup>1</sup> Orokorrean funtzio hauek «jatorrak» izaten dira, hau da, neurgarritasun propietatea gordetzen dute.

**Zenbaki handien lege sendoa.**  $\bar{X}_n \rightarrow E[X_1]$  ia ziur. Hau da,

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_n = E[X_1]\right) = 1.$$

**3. TEOREMA** (Limitearen teorema zentrala). *Izan bedi  $(X_n)_{n \geq 1}$  a.b.b. zorizko aldagaien segida.  $E[X_1^2] < \infty$  bada, orduan  $\sqrt{n}(\bar{X}_n - E[X_1])$  zorizko aldagaiaren banaketak  $\mathcal{N}(0, \sigma^2)$  banaketarantz konbergitzen du,  $\sigma^2 = \text{Var}(X_1)$  izanik. Hau da, edozein  $f: \mathbb{R} \rightarrow \mathbb{R}$  funtzio jarraiki eta bornatu hartuta,*

$$\lim_{n \rightarrow \infty} E\left[f\left(\sqrt{n}(\bar{X}_n - E[X_1])\right)\right] = \int f(x) e^{-\frac{x^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}}.$$

Edo baliokidea dena, edozein tarte  $I \subset \mathbb{R}$  hartuta,

$$\lim_{n \rightarrow \infty} P\left(\bar{X}_n \in E[X_1] + \frac{I}{\sqrt{n}}\right) = \int_I e^{-\frac{x^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}}.$$

## Orokorpenak

Emaitza hauek orokortzeko bide desberdin asko daude. Batetik, emaitza hauek asintotikoak dira; hau da,  $n$  «infinitura doanean» zer gertatzen den esaten dute. Kasu askotan ordea, jakin behar dugu  $n$  finkoa denean  $\bar{X}_n$  zorizko aldagaiari buruz zer esan dezakegun. Honi erantzungo diote hirugarren atalean garatuko ditugun kontzentrazio desberdintzek.

Bestetik, zenbaki handien legeek konbergentzia ematen badigute ere, ez dakigu hau zein abiaduratan gertatzen den. Abiadura hori estimatzeko helburuarekin garatu izan da desbiderapen handien teoria, laugarren atalean jorratuko duguna.

Dauden beste orokorpen posible anitzen artean, bi aipatuko ditugu euren garrantzia dela eta. Batetik, zorizko aldagaien beste funtzioei so egin diezaiekegu, baturei erreparatu beharrean. Bestela esanda,  $f(X_1, \dots, X_n)$  zorizko aldagaiari buruzko estimazioak lortu nahiko genituzke;  $f$  funtzioa  $\sup(X_1, \dots, X_n)$  edota  $X_1 \cdots X_n$  izan daiteke esate baterako. Bestetik, a.b.b. ez diren zorizko aldagaien segidak azter ditzakegu. Hauei buruz ere arituko gara testuan zehar, sakonean izango ez bada ere.

Orokorpen hauetan zein probabilitatearen arloko beste eremu askotan, Chebysheven desberdintza esponenziala erabiltzen da frogapen andana egiteko. Tresna hau hurrengo atalean garatuko dugu, geroago erabili ahal izateko.

Bukatzeko, bosgarren atalean mekanika estatistikoko Ising-Curie-Weiss eredua aurkeztuko dugu, eta desbiderapen handien teoria erabiliz aztertuko dugu.

## 2. CHEBYSHEVEN DESBERDINTZA ESPONENTZIALA

### Markoven eta Chebysheven desberdintzak

Datozen ataletan erabiliko dugun tresnarik oinarrizkoena garatuko dugu hemen, Chebysheven desberdintza esponentziala. Baina aurrena, horren aitzindariak aurkeztuko ditugu.

Izan bitez  $X \geq 0$  zorizko aldagai bat eta  $t > 0$  zenbaki erreal bat.  $1_{X \geq t}$  funtzioa  $\{X \geq t\}$  multzoaren funtzio indikatzailea da,  $\omega \in \Omega$  hartuta

$$1_{X \geq t}(\omega) = \begin{cases} 1, & X(\omega) \geq t \text{ bada,} \\ 0, & \text{bestela.} \end{cases}$$

Hurrengo erlazioa dugu beraz:

$$1_{X \geq t} \leq \frac{X}{t} \cdot 1_{X \geq t} \leq \frac{X}{t}.$$

Desberdintza horretan itxaropenak hartuz *Markoven desberdintza* ospetsua lortzen dugu:

$$P(X \geq t) \leq \frac{E[X]}{t}.$$

Zehazki,  $X$  edozein zorizko aldagai bada,  $|X|$  bere balio absolutua zorizko aldagai positiboa da eta aurreko desberdintza aplikatu diezaiokegu:

$$P(|X| \geq t) \leq \frac{E[|X|]}{t}.$$

Abiapuntutzat hartu dugun desberdintza erraz hobetu dezakegu,

$$1_{|X| \geq t} \leq \frac{|X|^k}{t^k} \cdot 1_{|X| \geq t} \leq \frac{|X|^k}{t^k},$$

eta berriro ere itxaropenak hartuz,

$$P(|X| \geq t) \leq \frac{E[|X|^k]}{t^k}.$$

Desberdintza hau  $|X - E[X]|$  zorizko aldagaiari aplikatuko diogu  $k = 2$  kasuan, *Chebysheven desberdintza* lortuko dugu horrela:

$$P(|X - E[X]| \geq t) \leq \frac{\text{Var}(X)}{t^2}.$$

Goian aipatutako desberdintza horiek guztiak benetan erabilgarriak suerta daitezzen, garrantzitsua da eskuin aldean agertzen diren itxaropenak finituak izatea;  $k$ -garren momentu finitua izan behar du  $X$  zorizko aldagaia.

### Chebysheven desberdintza esponenziala

Markoven eta Chebysheven desberdintzak oso garrantzitsuak dira. Nolanahi ere, kasu askotan ez dira nahikoa izaten; estimazio hobeak nahi izaten ditugu. Amarru txiki bat erabiliz asko irabaz dezakegu. Izan bedi  $\lambda \geq 0$ , Markoven desberdintzaren laguntzaz hau lortuko dugu:

$$P(X \geq t) \leq P(\lambda X \geq \lambda t) = P(e^{\lambda X} \geq e^{\lambda t}) \leq e^{-\lambda t} E[e^{\lambda X}],$$

hau da *Chebysheven desberdintza esponenziala*. Beste behin ere, erabilgarria izan dadin, garrantzitsua da  $E[e^{\lambda X}] < \infty$  betetzea.  $M(\lambda) = E[e^{\lambda X}]$  funtzioari  $X$  zorizko aldagaiaren *momentuen funtzio sortzailea* deritzo, edo baita  $X$ -ren *Laplacen transformatua* ere.  $\Lambda(\lambda) = \ln E[e^{\lambda X}]$  bere logaritmo nepertarrari  $X$ -ren *logLaplace funtzioa* deituko diogu. Azken funtzio hau erabiliz honela berridatz dezakegu Chebysheven desberdintza esponenziala:

$$P(X \geq t) \leq \exp -(\lambda t - \Lambda(\lambda)).$$

Datozen ataletan desberdintza hau ustiatzen saiatuko gara, hainbat norabidetan.

## 3. KONTZENTRAZIO-DESBERDINTZAK

### Helburua

Izan bitez  $X_1, \dots, X_n$  zorizko aldagai independenteak eta  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  funtzio bat. Kontzentrazio desberdintzek neurtzen dute  $Z = f(X_1, \dots, X_n)$  zorizko aldagaia bere itxaropenetik zenbat aldentzen den. Bestela esanda,  $t \geq 0$  hartuta,

$$P(Z - E[Z] \geq t) \quad P(Z - E[Z] \leq -t) \quad P(|Z - E[Z]| \geq t)$$

kantitateak (goitik) bornatu nahi ditugu.

### Cramér-Chernoff metodoa

Hasteko, kasurik sinpleena aztertuko dugu:

$$Z = \bar{X}_n = \frac{X_1 + \dots + X_n}{n}.$$

$\tilde{Z} = Z - E[Z]$  edo  $\tilde{Z} = E[Z] - Z$  hartuz,  $Z$  zorizko aldagai zentratua dela suposa dezakegu. Har dezagun  $\Lambda_Z(\lambda) = \ln E[e^{\lambda Z}]$ ,  $Z$  zorizko aldagaiaren logLaplace funtzioa. Aurreko atalean ikusi bezala, Chebysheven desberdintza esponentzialari esker hauxe dugu:

$$P(Z \geq t) \leq \exp(-\lambda t - \Lambda_Z(\lambda)).$$

Erabil dezagun  $X_1, \dots, X_n$  zorizko aldagaien arteko independentzia formula hau garatzeko:

$$E[e^{\lambda Z}] = E[e^{\frac{\lambda}{n}(X_1 + \dots + X_n)}] = E[e^{\frac{\lambda}{n}X_1}] \dots E[e^{\frac{\lambda}{n}X_n}],$$

logaritmoak hartuz,

$$\Lambda_Z(\lambda) = \Lambda_{X_1}\left(\frac{\lambda}{n}\right) + \dots + \Lambda_{X_n}\left(\frac{\lambda}{n}\right),$$

zehazki,  $X_1, \dots, X_n$  aldagaiek banaketa bera badute,

$$\Lambda_Z(\lambda) = n\Lambda_{X_1}\left(\frac{\lambda}{n}\right).$$

$\lambda \mapsto n\lambda$  aldagai aldaketa eginez hau lortzen dugu:

$$P(Z \geq t) \leq \exp(-n(\lambda t - \Lambda_{X_1}(\lambda))).$$

Desberdintza hau  $\lambda \geq 0$  guztietarako egia denez, eskuineko aldean infimoa hartu eta *Chernoffen desberdintza* lortuko dugu:

$$P(Z \geq t) \leq \inf_{\lambda \geq 0} \exp(-n(\lambda t - \Lambda_{X_1}(\lambda))) = \exp(-n \sup_{\lambda \geq 0} (\lambda t - \Lambda_{X_1}(\lambda))).$$

Azken supremo horri *Cramérren transformatua* deitu ohi zaio eta  $\Lambda_{X_1}^*$  adierazi ohi da:

$$\Lambda_{X_1}^*(t) = \sup_{\lambda \geq 0} (\lambda t - \Lambda_{X_1}(\lambda)).$$

Laburbilduz, Cramér-Chernoff metodoa  $\{Z \geq t\}$  gertakariaren probabilitatearentzako *borne esponenzialak* bilatzea du helburu. Hurrengo kasuak bereiz ditzakegu:

1.  $X$  zorizko aldagai zentratu bat bada, hauxe dugu:

$$P(X \geq t) \leq \exp -\Lambda_X^*(t).$$

2.  $Z = \frac{1}{n} (X_1 + \dots + X_n)$  zorizko aldagai zentratua bada, eta  $X_1, \dots, X_n$  independenteak badira, hauxe dugu:

$$P(Z \geq t) \leq \exp -(\Lambda_{X_1}^*(t) + \dots + \Lambda_{X_n}^*(t)).$$

3. Gainera,  $X_1, \dots, X_n$  aldagaiek banaketa bera badute, hauxe dugu:

$$P(Z \geq t) \leq \exp -n\Lambda_{X_1}^*(t).$$

### Lehen aplikazioak

**Bernoulli zorizko aldagaiak.** Izan bitez  $(X_k)_{k=1}^n$   $p$  parametroko Bernoulli zorizko aldagai independenteak. Erraz kalkula ditzakegu:

$$E[X_1] = p \quad \text{Var}(X_1) = p(1 - p).$$

Izan bedi  $X = X_1 - p$  aldagai zentratua. Ondokoa dugu:

$$E[e^{\lambda X}] = (pe^\lambda + 1 - p)e^{-\lambda p},$$

eta bere logLaplace funtzioa hauxe da:

$$\Lambda(\lambda) = \ln(pe^\lambda + 1 - p) - \lambda p.$$

Har dezagun  $\varphi(\lambda) = \lambda t - \ln(pe^\lambda + 1 - p) + \lambda p$  eta deriba dezagun,

$$\varphi'(\lambda) = t + p - \frac{pe^\lambda}{pe^\lambda + 1 - p}.$$

Adierazpen hau 0-rekin berdinduz eta  $\lambda$  isolatuz  $\varphi$ -ren maximo bakarra aurkitzen dugu:

$$\lambda = \ln \frac{(1-p)(p+t)}{p(1-p-t)}.$$



Balio hau  $\varphi$ -n ordezkatzuz, hauxe lortzen dugu:

$$\begin{aligned}\Lambda^*(t) &= (t+p) \ln \frac{(1-p)(p+t)}{p(1-p-t)} - \ln \left( \frac{(1-p)(p+t)}{1-p-t} + 1-p \right) = \\ &= (1-t-p) \ln \frac{1-t-p}{1-p} + (t+p) \ln \frac{t+p}{p}.\end{aligned}$$

Noski, honek bakarrik du zentzua  $t \leq 1-p$  denean. Zehazki,  $p = 1/2$  denean hauxe dugu,

$$\Lambda^*(t) = \frac{1+2t}{2} \ln(1+2t) + \frac{1-2t}{2} \ln(1-2t).$$

Bukatzeko, Chernoffen desberdintzari esker,

$$P\left(\frac{X_1 + \dots + X_n}{n} - \frac{1}{2} \geq t\right) \leq \exp -n \left( \frac{1+2t}{2} \ln(1+2t) + \frac{1-2t}{2} \ln(1-2t) \right).$$

4. *OHARRA*. Guk  $X$  zorizko aldagaiak  $\text{Ber}(1/2)$  banaketa duela diogu ondo-koia betetzen denean:

$$P(X=1) = 1 - P(X=0) = \frac{1}{2}.$$

Aitzitik, kasu askotan Bernoulli aldagaiak 0 eta 1 balioak hartu beharrean,  $-1$  eta 1 balioak hartzea komeni zaigu. Hots,  $Y$  zorizko aldagaia, hurrengo erlazioa betetzen duen aldagaia izatea nahi dugu:

$$P(Y=1) = 1 - P(Y=-1) = \frac{1}{2}.$$

Era horretako aldagai bat erraz lortu dezakegu  $X$  zorizko aldagaia transformatuz:

$$Y = 2X - 1.$$

Hau  $X_i$  bakoitzarekin eginez gero,  $X_i$ -rako lortutako emaitzari esker, hauxe dugu:

$$\begin{aligned}P\left(\frac{Y_1 + \dots + Y_n}{n} \geq \frac{t}{2}\right) &= P\left(\frac{X_1 + \dots + X_n}{n} - \frac{1}{2} \geq \frac{t}{2}\right) \leq \\ &\leq \exp -n \left( \frac{1+t}{2} \ln(1+t) + \frac{1-t}{2} \ln(1-t) \right).\end{aligned}$$

Bestela esanda,

$$\Lambda_{Y_1}^*(t) = \frac{1+t}{2} \ln(1+t) + \frac{1-t}{2} \ln(1-t).$$

Adierazpen hau ez dugu ahaztu behar, azken atalean erabiliko baitugu.

Bernoulli aldagaietarako erabili dugun eredu berari jarraituz, beste zorizko aldagai askotarako kontzentrazio desberdintzak lor ditzakegu. Hona hemen zenbait:

**Poisson zorizko aldagaiak.** Izan bitez  $(X_k)_{k=1}^n$   $\mu$  parametroko Poisson zorizko aldagaiak. Hau da,

$$P(X_1 = k) = e^{-\mu} \frac{\mu^k}{k!}, \quad k = 0, 1, 2, \dots$$

hauen itxaropena eta bariantza ondokoak dira:

$$E(X_1) = \mu \quad \text{Var}(X_1) = \mu.$$

Eta Cramér-Chernoff metodoarekin lortutako kontzentrazio desberdintza:

$$P\left(\frac{X_1 + \dots + X_n}{n} - \mu \geq t\right) \leq \exp -n \left( (t + \mu) \ln\left(1 + \frac{t}{\mu}\right) - t \right).$$

**Zorizko aldagai normalak.** Izan bitez  $(X_k)_{k=1}^n$  zorizko aldagai normal estandar eta independenteak. Hau da,  $I \subset \mathbb{R}$  edozein tarte izanda,

$$P\left(\frac{X_1 + \dots + X_n}{n} \geq t\right) \leq \exp -n \frac{t^2}{2}.$$

Zorizko aldagai hauen itxaropena 0 da eta bariantza berriz 1. Kontzentrazio desberdintzak itxura dotoreagoa dauka kasu honetan:

$$P\left(\frac{X_1 + \dots + X_n}{n} \geq t\right) \leq \exp -n \frac{t^2}{2}.$$

Suposa dezagun hurrena  $(X_k)_{k=1}^n$  a.b.b. zorizko aldagaiak direla eta euren banaketa  $\mathcal{N}(\mu, \sigma^2)$  dela. Hots,  $I \subset \mathbb{R}$  edozein tarte izanda,

$$P(X_1 \in I) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_I \exp -\frac{(x-\mu)^2}{2\sigma^2} dx.$$

$Y_k = \frac{X_k - \mu}{\sigma}$  jarritz,  $(Y_k)_{k=1}^n$  zorizko aldagai normal estandar eta independenteak dira, aurreko kasuan bezala. Ondorioz,

$$P\left(\frac{X_1 + \dots + X_n}{n} - \mu \geq t\right) = P\left(\frac{Y_1 + \dots + Y_n}{n} - \frac{t}{\sigma}\right) \leq \exp -n \frac{t^2}{2\sigma^2}.$$

### Zenbait teorema garrantzitsu

Ahantz ditzagun une batez zorizko aldagai independenteen baturak eta erabil dezagun Cramér-Chernoff metodoa zorizko aldagai orokorrako aztertzeke. Kasurik bakunenean  $X$  zorizko aldagai bornatua da. Hoeffdingi [1] zor diogu hurrengo emaitza hau.

5. **LEMA.** *Izan bedi  $X$   $[a, b]$  tartean balioak hartzen dituen zorizko aldagai zentratua. Orduan,*

$$P(X \geq t) \leq \exp \frac{2t^2}{(b-a)^2}, \quad \forall t \geq 0.$$

*Frogapena.*  $x \mapsto e^{\lambda x}$  funtzioaren konbexutasunari esker hau dugu:

$$e^{\lambda x} \leq \frac{b-X}{b-a} e^{\lambda a} + \frac{X-a}{b-a} e^{\lambda b},$$

eta itxaropenak hartuz,  $X$  zentratua denez gero,

$$\begin{aligned} E[e^{\lambda X}] &\leq \frac{be^{\lambda a} - ae^{\lambda b}}{b-a} = \frac{b - ae^{\lambda(b-a)}}{b-a} e^{\lambda a} = \\ &= \left(1 + \frac{a}{b-a} - \frac{a}{b-a} e^{\lambda(b-a)}\right) e^{\lambda(b-a)\frac{a}{b-a}}. \end{aligned}$$

$p = -\frac{a}{b-a}$ ,  $h = \lambda(b-a)$  jarritz eta logaritmoak hartuz,

$$\Lambda(\lambda) \leq -ph + \ln(1 - p + pe^h).$$

Azken funtzio honi  $L(h)$  deituz eta birritan deribatuz, hauxe dugu:

$$L''(h) = \frac{pe^h(1-p+pe^h) - (pe^h)^2}{(1-p+pe^h)^2} = \left(\frac{pe^h}{1-p+pe^h}\right) \left(1 - \frac{pe^h}{1-p+pe^h}\right).$$

Eskuineko adierazpen hau  $u(1 - u)$  erakoa da, eta  $0 \leq u \leq 1$  izanez gero,  $u(1 - u) \leq \frac{1}{4}$  ere badugu. Gure kasuan hala suertatzen denez, Taylorren garapena erabil dezakegu  $L(h)$  funtzioarekin hurrengo bornea lortzeko:

$$L(h) = L(0) + L'(0)h + \frac{L''(\xi)}{2}h^2 \leq \frac{h^2}{8}$$

$L(0) = L'(0) = 0$  eta  $0 \leq \xi \leq h \leq 1$  baitira.  $(b - a)\lambda$  adierazpideari deitu diogunez  $h$ , goiko guztia bateratzen badugu, hau lortzen dugu:

$$\Lambda(\lambda) \leq L(h) \leq \frac{(b - a)^2 \lambda^2}{8}.$$

Cramér-Chernoff metodoa aplikatuz, bilatutako emaitza lortuko dugu:

$$P(X \geq t) \leq \exp \frac{2t^2}{(b - a)^2}.$$

□

Hau ikusi eta berehala, zorizko aldagai bornatuen baturarako hurrengo korolariora lor dezakegu.

**6. TEOREMA (Hoeffdingen desberdintza).** *Izan bitez  $X_1, \dots, X_n$  zorizko aldagai independenteak, non  $X_i \sim [a_i, b_i]$  tartean hartzen dituen balioak. Har dezagun*

$$Z = \sum_{i=1}^n X_i - E[X_i].$$

Orduan,  $t \geq 0$  edozein delarik ere,

$$P(Z \geq t) \leq \exp \frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}.$$

Zorizko aldagai independenteen baturarako kontzentrazio desberdintzak hurrengo teorema biek ere ematen dizkigute. Frogarik gabe emango ditugu teorema hauek, nahiko teknikoak baitira, eta atal honen bukaera osatuko dute. Frogapenak [4] liburuan aurki ditzakegu.

7. **TEOREMA** (Bennetten desberdintza). *Izan bitez  $X_1, \dots, X_n$  zorizko aldagai independenteak eta har dezagun  $Z$  aurreko teoreman bezala. Demagun:*

1.  $v = \sum_{i=1}^n E[X_i^2] < \infty$ ,
2.  $\exists b > 0$  non  $\forall i = 1, \dots, n, X_i \leq b$  den.

Orduan,  $\lambda, t \geq 0$  edozein izanda,

$$\Lambda_Z(\lambda) \leq \frac{v}{b^2}(e^{b\lambda} - b\lambda - 1),$$

eta hortaz,

$$P(Z \geq t) \leq \exp -\frac{v}{b^2} \left( \left(1 + \frac{bt}{v}\right) \ln \left(1 + \frac{bt}{v}\right) - \frac{bt}{v} \right).$$

8. **TEOREMA** (Bernsteinen desberdintza). *Izan bitez  $X_1, \dots, X_n$  zorizko aldagai independenteak eta  $Z$  lehen bezala. Jo dezagun ondoko bi baldintzak betzen direla*

1.  $\exists v > 0$  non  $\sum_{i=1}^n E[X_i^2] \leq v$  den,
2.  $\exists c > 0$  non  $\sum_{i=1}^n E[X_i]_+^q \leq \frac{q!}{2} v c^{q-2}$  den,  $q \geq 3$  edozein izanik ere.

Orduan,  $\forall \lambda \in (0, 1/c), \forall t \geq 0$ ,

$$\Lambda_Z(\lambda) \leq \frac{v\lambda^2}{2(1-\lambda c)} \quad \text{eta} \quad P(Z \geq \sqrt{2vt} + ct) \leq e^{-t}.$$

#### 4. DESBIDERAPEN HANDIAK

##### Helburua

Desbiderapen handien teoriak gertaera arraroak aztertzen ditu; teoria bat barik, teknika multzo bat da. Hasteko  $(X)_{n \geq 1}$  a.b.b. zorizko aldagaien segidari erreparatuko diogu. Dei diezaiogun  $\bar{X}_n$ -ren banaketari  $\mu_n$ , zenbaki handien legeengatik,  $(\mu_n)_{n \geq 1}$  segidak  $\delta_{E[X_1]}$  banaketarantz konbergitzen du

( $\delta_x$  puntuko Diracen masa da). Aukera dezagun  $I \subset \mathbb{R}$  tarte bat,  $E[X_1] \notin \bar{I}$  betetzen duena; orduan,

$$P(\bar{X}_n \in I) = \mu_n(I) \xrightarrow{n \rightarrow \infty} 0.$$

Hau da,  $\{X_n \in I\}$  gertaerak gero eta probabilitate txikiagoa du;  $n$  handia de-  
nean gertaera «arraro» bat da. Mota honetako probabilitateak estimatzea da  
desbiderapen handien teoriaren helburua. Lanaren zati bat egina dugu kon-  
tzentrazio ezberdintzekin, nolana ere, bi alderdi nagusik bereiziko dute  
teoria hau aurreko ataleko teorian:

- Emaidza asintotikoak bilatzen ditugu orain. Hots, « $n \rightarrow \infty$ » erregi-  
menean zer gertatzen den jakin nahi dugu.
- Ereduaren abstrakzioaren bidez lortuko dugun orokortasuna.

### Cramérren teorema

Desbiderapen handien teorian ohikoa den formaltasunera eta abstrak-  
ziora pasa aurretik, teoria honen aitzindari den Cramérren teorema aurkez-  
tuko dugu.

**9. TEOREMA (Cramér).** *Izan bedi  $(X_n)_{n \geq 1}$  a.b.b. zorizko aldagaien segida bat, eta  $(X_n)_{n \geq 1}$  euren batezbesteko empirikoen segida.*

$$\forall x \in \mathbb{R} \quad \left( \frac{1}{n} \ln P(\bar{X}_n \geq x) \right)_{n \geq 1}$$

segidak limite bat du  $[-\infty, 0]$  tartean. Are gehiago,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln P(\bar{X}_n \geq x) = - \sup_{\lambda \geq 0} (\lambda x - \ln E[e^{\lambda X_1}]) = -\Lambda^*(x).$$

Ez dugu teorema bere osotasunean frogatuko; limitearen existentzia  
egiaztatuko dugu, hots,  $\forall x \in \mathbb{R}$ ,  $I(x) \in [0, \infty]$  existitzen dela, non

$$\lim_{n \rightarrow \infty} - \frac{1}{n} \ln P(\bar{X}_n \geq x) = I(x)$$

den, eta baita

$$I(x) \geq \sup_{\lambda \geq 0} (\lambda x - \Lambda(\lambda))$$

dela ere. Kontrako desberdintza askoz korapilatsuagoa da; [6] artikuluan aurki daiteke teorema osoaren frogapen ederra.

Aurrena lema azpibatukorra aurkeztu eta frogatuko dugu.

$$f: \mathbb{N} \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$$

*aplikazio azpibatukorra* dela diogu baldin eta ondokoa betetzen bada:

$$\forall m, n \geq 1 \quad f(m+n) \leq f(m) + f(n).$$

10. **LEMA.** *Izan bedi  $f: \mathbb{N} \rightarrow \mathbb{R}^+$  aplikazio azpibatukor bat. Orduan,*

$$\lim_{n \rightarrow \infty} \frac{f(n)}{n} = \inf_{n \geq 1} \frac{f(n)}{n}.$$

*Frogapena.* Izan bitez  $n \geq m \geq 1$ .  $n$  eta  $m$ -ren arteko zatiketa euklidearra eginez

$$n = mq + r, \quad 0 \leq r < m$$

idatz dezakegu. Hortaz,

$$f(n) = f(mq + r) \leq qf(m) + f(r),$$

eta hemendik

$$\frac{f(n)}{n} \leq \frac{q}{n} f(m) + \frac{f(r)}{n} \leq \frac{q}{n} f(m) + \frac{1}{n} \max_{0 \leq k < m} f(k).$$

Azken batura honetan, bigarren gaiak 0-rantz jotzen du  $n \rightarrow \infty$ -ra doanean, eta  $q/n$ -k berriz  $1/m$ -rantz. Hori dela eta,

$$\limsup_{n \rightarrow \infty} \frac{f(n)}{n} \leq \frac{f(m)}{m},$$

baina hau  $m \geq 1$  guztietarako denez egia,

$$\limsup_{n \rightarrow \infty} \frac{f(n)}{n} = \inf_{m \geq 1} \frac{f(m)}{m} \leq \liminf_{n \rightarrow \infty} \frac{f(n)}{n}.$$

□

*Teoremaren frogapena.* Aurrena  $(-\frac{1}{n} \ln P(\bar{X}_n \geq x))_{n \geq 1}$  segidaren konbergentzia frogatuko dugu lema azpibatukorra erabiliz. Izan bitez  $x \in \mathbb{R}$  eta  $n, m \geq 1$ . Ondokoa dugu:

$$\begin{aligned} P(\bar{X}_{n+m} \geq x) &= P\left(\frac{X_1 + \dots + X_n + X_{n+1} + \dots + X_{n+m}}{n+m} \geq x\right) \geq \\ &\geq P\left(\frac{X_1 + \dots + X_n}{n} \geq x, \frac{X_{n+1} + \dots + X_{n+m}}{m} \geq x\right) = \\ &= P(\bar{X}_n \geq x)P(\bar{X}_m \geq x), \end{aligned}$$

non azken berdintza segida a.b.b. bat dugulako idatzi ahal izan dugun. Logaritmoak hartuz,  $(-\ln P(\bar{X}_n \geq x))_{n \geq 1}$  segida azpibatukorra dela ikusten dugu. Gainera,  $x \in \mathbb{R}$  aukeratu badugu  $P(X_1 \geq x) > 0$  izan dadin,  $\forall n \geq 1$ ,  $P(\bar{X}_n \geq x) \geq P(X_1 \geq x)^n > 0$  izango dugu eta gure segidak  $\mathbb{R}^+$ -en hartuko ditu bere balioak. Baditugu lema azpibatukorra aplikatzeko hipotesi guztiak. Erabil dezagun lema bada, eta dei diezaiogun  $I(x)$  honek emango digun limitari, hots,

$$I(x) = \lim_{n \rightarrow \infty} -\frac{1}{n} \ln P(\bar{X}_n \geq x) = \inf_{n \geq 1} -\frac{1}{n} \ln P(\bar{X}_n \geq x).$$

Bukatzeko,

$$I(x) \geq \sup_{\lambda \geq 0} (\lambda x - \Lambda(\lambda))$$

dela ikusiko dugu. Izan ere, Chebysheven desberdintza esponentzialari esker, hau dugu:

$$I(x) = \inf_{n \geq 1} -\frac{1}{n} \ln P(\bar{X}_n \geq x) \geq \inf_{n \geq 1} -\frac{1}{n} \ln \exp -n(\lambda x - \Lambda(\lambda)) = \lambda x - \Lambda(\lambda),$$

eta hau  $\lambda \geq 0$  guztietarako betetzen da, nahi genuen desberdintzari bide emanaz:

$$I(x) \geq \sup_{\lambda \geq 0} (\lambda x - \Lambda(\lambda)).$$

□

## Desbiderapen handien formalismoa

Jarraian desbiderapen handien formalismoa aurkeztuko dugu, ideia eta noziorik garrantzitsuenak aurkeztuz. Orokortasun handiagoa lortzeko ez gara zorizko aldagaietara mugatuko; probabilitate neurriekin egingo dugu lan.



Probabilitate neurri hauek espazio metriko baten gainean hartuko ditugu, nahiz eta kontzeptu asko espazio topologiko orokorragoetan ere baliagarriak izan, baldintza egokien pean betiere. Jo dezagun beraz  $(X, d)$  espazio metriko bat dela, eta dei diezaiogun  $\mathcal{B}$  metrika horretako bola irekiek sortutako tribu boreldarrari.

11. **DEFINIZIOA.**  $I : X \rightarrow [0, \infty]$  *tasa-funtzio* bat dela diogu baldin eta azpitik erdijarraikia bada, hau da,

$$\{x \in X : I(x) \leq \lambda\}$$

multzoa itxia baldin bada  $\lambda \in \mathbb{R}$  guztietarako. Horrez gain multzo horiek guztiak trinkoak badira,  $I$  *tasa-funtzio ona* dela diogu.

12. **DEFINIZIOA.** Izan bedi  $(\mu_n)_{n \geq 1}$   $(X, \mathcal{B})$  espazioaren gaineko probabilitate neurrien segida bat.  $(\mu_n)_{n \geq 1}$  segidak *desbiderapen handien printzipio (DHP)* bat betetzen duela diogu,  $I$  *tasa-funtzioarekiko*, baldin eta  $B \in \mathcal{B}$  edozein boreldar hartuta, honako hau betetzen bada:

$$\begin{aligned} -\inf \{I(x) : x \in \overset{\circ}{B}\} &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \ln \mu_n(B) \leq \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \ln \mu_n(B) \leq -\inf \{I(x) : x \in \bar{B}\} \end{aligned}$$

Era berean  $(X_n)_{n \geq 1}$  zorizko aldagaien segida batek DHP bat betetzen duela diogu baldin eta zorizko aldagaien banaketan segidak DHP-a betetzen badu.

Oro har, DHP-aren azpi eta goi borneak era oso desberdinean frogatzen dira. Goi bornea izaten da zailena; gehienetan, aurrena multzo trinkoetarako egiten da frogia; ostean segidaren tentsioan oinarritutako argudioak erabiliz, multzo itxietara zabaltzen da bornea. Behe bornea errazagoa izaten da ikusteko, Chebysheven desberdintza esponentzian oinarritutako arrazonamenduak erabiltzen dira kasu honetan.  $(\mu_n)_{n \geq 1}$  segidak behe bornea betetzen badu, eta goi bornea multzo trinkoetarako betetzen bada, *DHP* ahula betetzen duela esaten dugu.

Segituan frogarik gabe emango dugun teoremak dio DHP-ak ez direla funtzio jarraikiak bitartez galtzen. Frogapenak testu honen helburua gaindizten du, eta [3] liburuan aurki dezakegu.

13. **TEOREMA.** *Izan bitez  $(X, d_X)$  eta  $(Y, d_Y)$  espazio metriko bi eta izan bedi  $f : X \rightarrow Y$  funtzio jarraiki bat. Jo dezagun  $(\mu_n)_{n \geq 1}$   $X$ -ren gaineko probabilitate neurrien segidak DHP bat betetzen duela  $I$  *tasa-funtzio onare-**

kiko. Izan bedi  $\nu_n = \mu_n \circ f^{-1}$ ,  $\mu_n$ -ren irudi-neurria  $f$  funtzioaren bitartez. Orduan  $(\nu_n)_{n \geq 1}$   $Y$ -ren gaineko probabilitate neurrien segidak DHP bat betetzen du  $J$  tasa-funtzioarekiko:

$$\forall y \in Y \quad J(y) = \inf\{I(x) : x \in X, f(x) = y\}.$$

### Varadhanen lema

Atal honetako azken bi emaitzak emango ditugu hemen, ondoren landuko dugun mekanika estatistikoko ereduaz aztertzeko oso erabilgarriak izango baitzaizkigu.

14. **LEMA** (Varadhanen lema). *Izan bedi  $(\mu_n)_{n \geq 1}$  probabilitate neurrien segida bat  $X$  espazio metrikoaren gainean. Suposa dezagun  $(\mu_n)_{n \geq 1}$  segidak  $I$  tasa-funtzio onarekiko DHP bat betetzen duela.  $f: X \rightarrow \mathbb{R}$  funtzio jarraiki eta bornatu bat bada, orduan*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_X e^{nf(x)} d\mu(x) \right) = \sup_{x \in X} (f(x) - I(x)).$$

*Frogapena.* Oraingo honetan ere teoremaren zati bat baino ez dugu frogatu: azpi bornea. Hots,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_X e^{nf(x)} d\mu(x) \right) \geq \sup_{x \in X} (f(x) - I(x)).$$

Hori ikusteko, har dezagun  $x_0 \in X$  edozein puntu eta  $V$  puntu horren ingurune ireki bat. Orduan,

$$\int_X e^{nf(x)} d\mu(x) \geq \int_V e^{nf(x)} d\mu(x) \geq \exp\left(n \cdot \inf_{x \in V} f(x)\right) \cdot \mu_n(V)$$

eta DHP-aren azpi-borneari esker, logaritmoak hartuz, ostean  $n$ -rekin zatituz ondokoa dugu,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_X e^{nf(x)} d\mu(x) \right) &\geq \inf_{x \in V} f(x) - \inf\{I(x) : x \in V\} \geq \\ &\geq \inf_{x \in V} f(x) - I(x_0) \geq \sup_V \inf_{x \in V} f(x) - I(x_0) \geq f(x_0) - I(x_0), \end{aligned}$$

non azkenurreko adierazpenean, supremoa  $x_0$ -ren  $V$  ingurune guztien artean hartzen dugun, noski  $\sup_V \inf_{x \in V} f(x) = x_0$  da orduan. Eta emaitza hau  $x_0$  guztietarako egia denez, badugu nahi genuen desberdintza. Goi bornean frogapena [3] liburuan aurki daiteke.  $\square$

Hurrengo emaitza dugu Varadhanen lemarren aplikazio bezala.

**15. TEOREMA.** *Izan bedi  $(\mu_n)_{n \geq 1}$  probabilitate neurrien segida bat  $(X, \mathcal{B})$  espazio metriko neurgarriaren gainean. Suposa dezagun  $(\mu_n)_{n \geq 1}$  segidak  $I$  tasa-funtzio onarekiko DHP bat betetzen duela.  $f: X \rightarrow \mathbb{R}$  funtzio jarraiki eta bornatu bat bada, eta  $(X, \mathcal{B})$  espazioan  $(\nu_n)_{n \geq 1}$  probabilitate neurri berriak honela definitzen baditugu:*

$$\forall B \in \mathcal{B} \quad \nu_n(B) = \frac{\int_B e^{nf(x)} d\nu_n(x)}{\int_X e^{nf(x)} d\nu_n(x)},$$

orduan,  $(\nu_n)_{n \geq 1}$  segidak DHP bat betetzen du  $J$  tasa-funtzioarekiko:

$$\forall x \in X \quad J(x) = I(x) - f(x) - \inf_{y \in X} (I(y) - f(y)).$$

*Frogapena.* Varadhanen leinari esker hau dugu:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_X e^{nf(x)} d\mu(x) \right) = \sup_{x \in X} (f(x) - I(x)).$$

Aurrena azpi bornea frogatuko dugu. Horretarako, nahikoa da ondokoa betetzen dela ikustea:

$$\forall B \subset X \text{ ireki} \quad \forall x \in B \quad \forall \varepsilon > 0$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) \geq f(x) - I(x) - \varepsilon.$$

Har dezagun  $V \subset B$   $x$ -ren ingurune ireki bat.  $f$  jarraikia denez,  $V$  inguruneak honako propietate hau betetzen duela suposa dezakegu:

$$\forall y \in V \quad f(y) \geq f(x) - \varepsilon.$$

Orduan,

$$\int_B e^{nf} d\mu_n \geq \int_V e^{nf} d\mu_n \geq \int_V e^{nf(x-\varepsilon)} d\mu_n \geq \varepsilon^{nf(x-\varepsilon)} \mu_n(V),$$

eta hemendik,  $(\mu_n)_{n \geq 1}$  segidak betetzen duen DHP-ari esker,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) \geq f(x) - \varepsilon - I(x).$$

Froga dezagun hurrena goi bornea, hots,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) \leq -\inf\{I(x) - f(x) : x \in \bar{B}\}.$$

Har ditzagun

$$\alpha < \inf\{I(x) - f(x) : x \in \bar{B}\} \quad \text{eta} \quad A = \{x \in X : I(x) - f(x) \leq \alpha\}.$$

A multzo trinkoa da eta  $B$ -rekin disjuntua, badago  $\varphi : X \rightarrow \mathbb{R}$  funtzio jarraiki bat non

$$\begin{cases} \varphi(x) = 0, & \forall x \in \bar{B}, \\ \varphi(x) = 1 & \forall x \in A, \\ 0 \leq \varphi \leq 1. \end{cases}$$

Zehazki,  $1_B \leq e^{-nt\varphi}$  dugu  $t \geq 0$  edozein izanik ere, eta hortaz,

$$\int_B e^{nf} d\mu_n \leq \int_X e^{n(f-t\varphi)} d\mu_n.$$

Orain Varadhanen lema erabiliz  $f - t\varphi$  funtzioarekin,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) &\leq \sup_{x \in X} (f(x) - t\varphi(x) - I(x)) \leq \\ &\leq \max \left( \sup_{x \in A} (f(x) - t\varphi(x) - I(x)), \sup_{x \notin X} (f(x) - t\varphi(x) - I(x)) \right) \leq \\ &\leq \max \left( \|f\|_{\infty} - t, -\alpha \right). \end{aligned}$$

$t$  infinitura bidaliz, hau lortzen dugu:

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) \leq -\alpha,$$

eta hau egia da  $\alpha < \inf \{I(x) - f(x) : x \in \bar{B}\}$  edozein izanik; beraz,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \left( \int_B e^{nf} d\mu_n \right) \leq -\inf \{I(x) - f(x) : x \in \bar{B}\}.$$

□

## 5. ISING-CURIE-WEISS EREDUA

### Isingen eredua

Mekanika estatistikoak haxe du helburu, sistema fisiko baten baitan gertatzen diren elkarrekintza mikroskopikoak aztertu, gerora maila makroskopikoan ondorioak atera ahal izateko.

*Isingen eredua* da mekanika estatistikoko eredu sonatuenetako bat. Wilhelm Lorenz fisikariak 1920. urtean asmatu zuen ferromagnetismoa modelizatzeko, eta bere ikasle izan zen Ernst Isingek honen inguruan egin zuen tesia, hortik ereduaren izena. Eredu honek,  $\mathbb{Z}^d$  sarea hartzen du espaziotzat, non  $d \geq 1$  sarearen dimentsioa den.  $\mathbb{Z}^d$ -ko erpin bakoitza atomo bat bezala irudika dezakegu, eta atomo bakoitzak «spin» bat izango du, +1 edo -1. Sistema honek har ditzakeen egoerei *konfigurazio* deritze. Konfigurazio hauek

$$\omega : \mathbb{Z}^d \rightarrow \{-1, +1\}$$

aplikazioak dira, eta konfigurazio bakoitzari *Hamiltoniar* bat dagokio, edo konfigurazio jakin horretan sistemak duen energia:

$$H(\omega) = - \sum_{i,j \in \mathbb{Z}^d : |i-j|=1} \omega(i)\omega(j) - h \sum_{i \in \mathbb{Z}^d} \omega(i),$$

$h$  kanpotik eragindako eremu magnetiko bat da. Onar dezagun temperatura  $T > 0$  dela eta dei diezaigun  $\beta = 1/T$  honen alderantzizkoari.  $\beta$  konstanteari *Boltzmannen faktorea* deritzo maiz.  $\{-1, +1\}^{\mathbb{Z}^d}$  espazioan  $H$  hamiltoniarrari dagokion *Gibbsen neurria* honako haxe da:

$$\mu(\omega) = \frac{1}{Z} e^{-\beta H(\omega)}$$

non

$$Z = \sum_{\omega \in \{-1,+1\}^{\mathbb{Z}^d}} \exp(-\beta H(\omega))$$

*partiketa-funtzioa* den, hau da,  $\mu$  neurria  $\{-1, +1\}$  espazioaren gainean probabilitate neurri bat izateko beharrezkoa zaigun normalizazio-konstantea.

Eredu honetan, partikula edo atomo bakoitzak, bere auzokiderik hurbilenekin baino ez du harremanik, hots, eskatzen dugu atomo biren arteko elkarrekintza kontutan izateko, euren arteko distantzia 1 izatea. Eredu hau, nahiz eta fisikoki sinplea izan, oso aberatsa da, eta matematikaren ikuspegitik ez da aztertzeke erraza.  $d$  sarearen dimentsioa handia denean, batez besteko eremu hurbilpen bat egiten ohi da, hots, beste eredu batekin hurbiltzen dugu Isingen eredu, non partikula guztiek, beste guztiekin elkarrekintzak pairatzen dituzten. Azken honi Ising-Curie-Weiss eredu deritzo eta erazagoa da aztertzen.

### Ising-Curie-Weiss eredu

Izan bitez  $n \geq 1$  eta  $\Omega_n = \{-1, +1\}^n$ . Adieraz dezagun  $\omega = (\omega_1, \dots, \omega_n)$ -ren bitartez  $\Omega_n$ -ko konfigurazio edo puntu orokor bat. Sistema honetan *Hamiltoniarra* honela definitzen da:

$$H_n(\omega) = -\frac{1}{2n} \sum_{1 \leq i, j \leq n} \omega_i \omega_j - h \sum_{i=1}^n \omega_i,$$

non  $h \in \mathbb{R}$  konstanteak kanpoko eremu magnetikoa adierazten duen. Gibbsen neurria Isingen eredu bezala definitzen dugu:

$$\mu_n(\omega) = \frac{1}{Z_n} e^{-\beta H_n(\omega)},$$

non  $Z_n = \sum_{\omega \in \Omega_n} \exp(-\beta H_n(\omega))$  *partiketa-funtzioa* den, hau da,  $\mu_n$  neurria  $\Omega_n$  espazioaren gainean probabilitate neurri bat izan dadin beharrezkoa zaigun normalizazio-konstantea.

Izan bedi  $S_n(\omega) = \omega_1 + \dots + \omega_n$ .  $S_n$  zorizko aldagaia aztertzea dugu helburua,  $\mu_n$  banaketaren pean eta  $n \rightarrow \infty$  erregimen asintotikoan.

### Desbiderapen handien printzipioa

Ohar gaitezen Hamiltoniarra  $S_n/n$ -ren funtzioan adieraz dezakegula:

$$H_n(\omega) = -\frac{n}{2} \left( \frac{S_n(\omega)}{n} \right)^2 - nh \frac{S_n(\omega)}{n} = -nf \left( \frac{S_n(\omega)}{n} \right),$$

$f: x \mapsto \frac{x^2}{2} + hx$  funtzioa izanik. Izan bedi  $B, [-1, +1] \subset \mathbb{R}$  tarteko boreldar bat, orduan, honakoa dugu:

$$\begin{aligned} \mu_n \left( \frac{S_n}{n} \in B \right) &= \frac{1}{Z_n} \sum_{\omega \in \Omega_n} 1_{\left\{ \frac{S_n(\omega)}{n} \in B \right\}} \cdot e^{-\beta H_n(\omega)} = \\ &= \frac{1}{Z_n} \sum_{\omega \in \Omega_n} 1_{\left\{ \frac{S_n(\omega)}{n} \in B \right\}} \cdot \exp \left( \beta n f \left( \frac{S_n(\omega)}{n} \right) \right). \end{aligned}$$

$\Omega_n = \{-1, +1\}^n$  espazioko banaketa uniformea ( $\omega \in \Omega_n$  bakoitzari  $2^{-n}$ -ko probabilitatea ezartzen dion banaketa)  $U_n$  bidez adierazten badugu, aurreko adierazpena berridatz dezakegu,

$$\begin{aligned} \mu_n \left( \frac{S_n}{n} \in B \right) &= \frac{2^n}{Z_n} \int_{\Omega_n} 1_{\left\{ \frac{S_n(\omega)}{n} \in B \right\}} \cdot \exp \left( \beta n f \left( \frac{S_n(\omega)}{n} \right) \right) dU_n(\omega) = \\ &= \frac{2^n}{Z_n} \int_{-1}^{+1} 1_{\{x \in B\}} \cdot e^{n\beta f(x)} d\nu_n(x). \end{aligned}$$

non  $\nu_n$  banaketa,  $U_n$  legearen irudia den  $S_n/n$  aplikazioaren bitartez, hau da,

$$\forall A \in \mathcal{B}([0, 1]) \quad \nu_n(A) = U_n \left( \frac{S_n}{n} \in A \right).$$

Zehazki,  $B = [-1, +1]$  tarte osoa aukeratzten badugu,  $\mu_n \left( \frac{S_n}{n} \in [-1, 1] \right) = 1$  dugu, eta hemendik,

$$\int_{-1}^{+1} e^{n\beta f(x)} d\nu_n(x) = \frac{Z_n}{2^n},$$

bestela esanda,  $\forall B \in \mathcal{B}([-1, +1])$ ,

$$\mu_n \left( \frac{S_n}{n} \in B \right) = \frac{\int_B e^{n\beta f} d\nu_n}{\int_{-1}^{+1} e^{n\beta f} d\nu_n}.$$

Nolanahi ere,  $\nu_n$  banaketa  $1/2$  parametroko  $n$  Bernoulli aldagaien batez-besteko enpirikoaren banaketa besterik ez da. Cramérren teoremagari

esker badakigu  $\nu_n$ -k DHP bat betetzen duela honako tasa-funtzio honen pean:

$$I(x) = \frac{1+x}{2} \ln(1+x) + \frac{1-x}{2} \ln(1-x).$$

Honek esan nahi du  $\mu_n(S_n/n \in B)$  balioaren estimazio asintotikoa lortzeko, aurreko atalaren amaieran emandako teorema erabil dezakegula. Hori eginda hurrengoa lortzen dugu:

**16. PROPOSIZIOA.**  $\mu_n$  banaketaren pean,  $(S_n/n)_{n \geq 1}$  zorizko aldagaien segidak DHP bat betetzen du

$$\begin{aligned} J(x) &= I(x) - \beta f(x) - \inf_{-1 \leq y \leq 1} (I(y) - \beta f(y)) = \\ &= \frac{1+x}{2} \ln(1+x) + \frac{1-x}{2} \ln(1-x) - \beta \left( \frac{x^2}{2} + hx \right) - C \end{aligned}$$

tasa-funtzioarekiko, non  $C = \inf(I - \beta f)$  konstantea den.

### Ondorioak

Kontsidera dezagun  $K, J$  funtzioaren minimoen multzoa, har dezagun  $\delta > 0$  eta izan bedi  $A, [-1, 1]$  tarteko boreldar hau:

$$A = \{x \in [-1, 1] : d(x, K) \geq \delta\}.$$

Har dezagun  $\bar{x} = \inf \{J(x) : x \in A\}$ .  $J$  funtzio jarraikia denez eta bere minimoetan 0 balioa hartzen duenez, hauxe dugu DHP-ari esker:

$$-\bar{x} = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mu_n \left( d \left( \frac{S_n}{n}, K \geq \delta \right), \right),$$

non  $\bar{x} \in (0, +\infty)$  den. Hemendik,

$$\lim_{n \rightarrow \infty} \ln \mu_n \left( d \left( \frac{S_n}{n}, K \geq \delta \right) \right) = -\infty,$$

eta halaberrez,

$$\lim_{n \rightarrow \infty} \mu_n \left( d \left( \frac{S_n}{n}, K \geq \delta \right) \right) = 0.$$



Nolabait esatearren, zenbaki handien lege ahula berreskuratu dugu: ( $\mu_n$ -ren menpeko)  $S_n/n$ -ren banaketa  $K$  multzoaren inguruan kontzentratzen da.  $K$  zein den jakiteko, gogora dezagun  $J$  tasa-funtzioaren adierazpena,

$$J(x) = \frac{1+x}{2} \ln(1+x) + \frac{1-x}{2} \ln(1-x) - \beta \left( \frac{x^2}{2} + hx \right) - C.$$

Deribatuz,

$$J'(x) = \arg \tanh x - \beta(x+h).$$

Ekuazio honen erroak bilatu behar ditugu  $h$  eta  $\beta$ -ren arabera. Hurrengo hiru kasuak bereiz ditzakegu:

$h = 0, \beta \leq 1$ . Kasu honetan  $K = \{0\}$  dugu.

$h = 0, \beta > 1$ . Oraingoan  $K = \{-x^*, x^*\}$  dugu,  $x^*$  eta  $x^*$  ekuazioaren erro simetriko biak dira.

$h \neq 0, \beta > 0$ . Gure ekuazioak erro bakarra du,  $x^*$  deituko duguna.

Hau ikusita, ondokoa baieztatu dezakegu  $S_n/n$ -ren konbergentziari buruz:

**17. TEOREMA.** *1.  $0 \leq \beta \leq 1$  eta  $h = 0$  edo  $\beta > 0$  eta  $h \neq 0$  bada,*

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \delta_{x^*},$$

*non  $x^*, \arg \tanh x = \beta(x+h)$  ekuazioaren erro bakarra den.*

*2.  $\beta > 1$  eta  $h = 0$  bada,*

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \frac{1}{2} (\delta_{x^*} + \delta_{-x^*}),$$

*non  $-x^*, x^*, \arg \tanh x = \beta(x+h)$  ekuazioaren erro simetriko biak diren.*

*Frogapena.* Lehenengo puntua berehala ondorioztatzen da DHP-tik. Bigarrena frogatzeko ideia besterik ez dugu emango. Lehenik, neurri finituen segiden trinkotasun erlatiboa erabiltzen da limitearen existentzia frogatzeko. Suposa dezagun limite hau  $\nu$  dela. DHP-ari esker, badakigu probabilitate-neurri honen euskarria  $\{-x^*, x^*\}$  multzoa dela, hau da,  $\nu = a\delta_{-x^*} + (1-a)\delta_{x^*}$ ,  $a \in [0, 1]$  zenbakiren batentzat. Aitzitik,  $S_n/n$ -ren banaketa simetrikoa da  $\mu_n$ -ren pean:

$$\mu_n \left( \frac{S_n}{n} \right) = 0.$$

Hori dela eta,  $n$  infiniturantz joaraziz ondokoa lortzen dugu,

$$\int_{-1}^{+1} x dv = 0,$$

eta  $a = 1/2$  izan behar dela ondorioztatzen dugu.  $\square$

Eskerrak eman nahi dizkiet Raphaël Cerf eta Pascal Massart irakasleei. Paris-Sud unibertsitateko «M2 Probabilités et Statistiques» masterrean irakatsi dituzten «desbiderapen handiak eta fase-trantsizioak» eta «ereduen kontzentrazio eta aukeraketa» ikasgaietan baitago inspiratuta lan hau hein handi batean. Era berean, begirale anonimo bat eta euskara-begiralea eskertu nahi ditut, haien iruzkinak eta ekarpenak testua hobetzen lagundu baitidate.

## ERREFERENTZIAK

- [1] Hoeffding W. 1963. «Probability Inequalities for Sums of Bounded Random Variables». *Journal of the American Statistical Association*, **58**(301), 13-30.
- [2] Billingsley P. 1995. *Probability and Measure*. Wiley Interscience.
- [3] Dembo A. eta Zeitouni O. 1998. *Large Deviations Techniques and Applications*. Springer.
- [4] Boucheron S., Lugosi G. eta Massart P. 2013. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press.
- [5] Bennett G. 1962 «Probability Inequalities for the Sum of Independent Random Variables». *Journal of the American Statistical Association*, **57**(297), 33-45.
- [6] Cerf R. eta Petit P. 2011. «A short proof of Cramér's theorem». *American Mathematical Monthly*, **118**(10), 925-931.