CLI in rodovne funkcije momentov

Blaž Erzar

13. julij 2021

1 Uvod

Centralni limitni izrek (CLI) je eden izmed najpomembnejših izrekov iz verjetnosti. Omogoča nam ocenjevanje statističnih parametrov o porazdelitvah. Izrek pravi, da je povprečje $\overline{X}$ enako ali vsaj podobno porazdeljenih slučajnih spremenljivk $X_i (1 \leq i \leq n)$ s končno pričakovano vrednostjo μ in končno varianco σ porazdeljeno približno normalno, tj.

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right), \text{ kjer je } \overline{X} = \frac{X_1 + \ldots + X_n}{n}$$

njihovo povprečje in velja $n \geq 30$[1]. Od tod sledi, da lahko iz samega izreka lahko takoj sklepamo, da je uporaben za aproksimacijo povprečja ali vsote slučajnih spremenljivk poljubnih porazdelitev z normalno. Poleg tega CLI uporabljamo tudi pri računanju *intervalov zaupanja*, s katerimi lahko z uporabo vzorcev ocenjujemo parametre populacije pri dani stopnji zaupanja α. En izmed takih parametrov je že omenjena pričakovana vrednost μ, ki jo lahko precej dobro ocenimo kar z vzorčnim povprečjem:

$$P\left(-z_{\alpha/2} \leq \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1 - \alpha,$$

kjer je kvantil reda $\alpha/2$, tj. $P(X \leq -z_{\alpha/2}) = \alpha/2$ glej sliko. ~~$P(X \leq x) \geq \alpha$ in $P(X \geq x) \geq 1 - \alpha$.~~ To je splošna def., ki pa ji za zv. d. spr. nepotrebno dolga.

To nam močno olajša delo, saj je veliko bolj učinkovito preučevati vzorce, kot pa celotno populacijo, kar je v nekaterih primerih tudi praktično nemogoče. Seveda morajo biti pri tem vzorci primerno izbrani (neodvisni/naključni). Poleg tega pa intervale zaupanja uporabimo tudi pri preverjanju statističnih domnev, s katerimi z uporabo vzorcev testiramo trditve o populaciji.

Vidimo, da ima CLI številne uporabe, a vseeno poglejmo konkreten primer računanja verjetnosti, ki nas bo o tem še bolj prepričal.

[1] Pravilo, da potrebujemo več kot 30 slučajnih spremenljivk, ni strogo določeno. Gre bolj za neko izkustveno pravilo, obstajajo pa tudi nekateri bolj neformalni razlogi za izbor tega števila, glej npr. odgovor na vprašanje A. P. Morffisa [2]. Omenimo še, da v nekateri literaturi lahko namesto števila 30 zasledimo tudi kakšno drugo oceno, npr. 20 (če začnemo z X_i porazdeljeno binomsko) ali 50.

Recimo, da mečemo kovanec in štejemo kolikokrat pade grb (pravzaprav gre za analogen primer naključnemu sprehodu, kjer v vsakem koraku naredimo korak bodisi v levo bodisi v desno). Ker naj bi šlo za pošten kovanec, naj bi bila oba izzida enako verjetna, tj. $p = 1/2$. Vemo, da je slučajna spremenljivka, ki spremlja število grbov, porazdeljena binomsko. Vrednosti njene gostote porazdelitve lahko enostavno izračunamo po binomskem obrazcu, tj.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

kjer je n število metov in k število padlih grbov. Če želimo izračunati verjetnost, da pade vsaj ali največ k grbov? Potrebovali bi formulo za porazdelitveno funkcijo, vendar za binomsko porazdelitev vsota $k+1$ oziroma $n-k$ zgornjih verjetnosti ni učinkovita. Vsota lahko hitro postane dokaj dolga, npr. za $k \geq 30$, računanje pa bi trajalo eksponentno dolgo časa v odvisnosti od dolžine zapisa števila n.

Veliko lažje bi bilo, če bi lahko namesto binomske porazdelitve uporabili normalno, ki v primeru poštenega kovanca (porazdelitev je simetrična) na prvi pogled izgleda podobno. To nam omogoči CLI, saj je binomska porazdelitev pravzaprav samo vsota indikatorskih slučajnih spremenljivk. Te so enako porazdeljene:

$$\begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

in imajo končno varianco $p(1-p)$. To pa pomeni, da lahko vsoto gostote porazdelitve izrazimo s porazdelitveno funkcijo normalne porazdelitve, ki jo lahko pridobimo iz tabele za funkcijo napake, npr. na portalu *Math is Fun* [1].

Ugotovili smo, da je CLI res uporaben in nam pomaga pri računanju. Omenili smo primer meta poštenega kovanca in štetja padlih grbov. Ta porazdelitev je simetrična in za primerno vrednost n že na prvi pogled izgleda podobne oblike kot normalna. Kaj pa če imamo nesimetrično binomsko porazdelitev ($p \neq 1/2$) ali pa neko popolnoma drugačno porazdelitev? Ta na prvi pogled nikakor ne izgleda podobna normalni. Ali torej CLI ne velja? Pokazali bomo, da tudi v tem primeru izrek velja, čeprav dokaz ni tako očiten. Iz tega lahko

sklepamo, da bomo potrebovali še kakšno močno orodje.

Naš cilj je bralcem predstaviti rodovne funkcije in približati dokaz CLI. Opazimo lahko, da sta osrednji količini naše študije pričakovana vrednost in varianca. Slednjo lahko vpeljemo s pričakovano vrednostjo ($\sigma^2 = \mathbb{E}[(X - \mathbb{E}(X))^2]$). To sta pravzaprav momenta. *Moment* slučajne spremenljivke X reda $k \in \mathbb{N}$ glede na točko $a \in \mathbb{R}$ je količina definirana z

$$m_k(a) = \mathbb{E}((X - a)^k).$$

Za $a = 0$ dobimo *začetni moment*, za $a = \mathbb{E}(X)$ pa *centralni moment*. Torej je pričakovana vrednost prvi začetni moment, varianca pa drugi centralni. Zaradi tega si bomo pri dokazu pomagali z rodovnimi funkcijami momentov, ki omogočajo računanje omenjenih momentov.

V **2.** in **3. razdelku** bomo podrobneje spoznali rodovne funkcije momentov, njihovo formalno definicijo in osnovne lastnosti. Nekatere bomo tudi dokazali. V **razdelku 4** jih bomo uporabili za izračun konkretnih rodovnih funkcij za nekaj porazdelitev. Omenili bomo tudi Laplaceovo transformacijo (**razdelek 5**) iz katere sledi pomemben izrek, ki ga bomo potrebovali pri glavni nalogi te projektne naloge, dokazu CLI, Tega se bomo lotili v zadnjem **razdelku 6**.

←V dodatku **A** bomo predstavili še karakteristične funkcije, vendar ne tako podrobno kot rodovne funkcije momentov (brez dokazovanja) [2]. Te funkcije bi lahko uporabili tudi za dokaz CLI, vendar njihova obdelava zahteva tudi uporabo kompleksne analize, ki se ji bomo raje izognili. V dodatku **B** pa predstavimo še splošnejše rodovne funkcije. Nekaj jih bomo tudi izračunali ter uporabili za štetje.

2 Definicija rodovne funkcije

Rodovna funkcija momentov slučajne spremenljivke X $M_X : \mathbb{R} \to [0, \infty)$, je definirana z:

$$M_X(t) = \mathbb{E}(e^{tX}),$$

kjer $\mathbb{E}$ predstavlja pričakovano vrednost (v nadaljevanju ji bomo rekli kar rodovna funkcija). Rodovna funkcija $M_X(t)$ obstaja, če je pričakovana vrednost končna v okolici ničle, tj. obstaja pozitivna konstanta a, da je vrednost funkcije $M_X(t)$ končna za vsak $t \in [-a, a]$.

[2] Tako karakteristične funkcije kot tudi rodovne funkcije momentov spadajo v harmonično analizo. Njuno idejo sta nakazala že de Moivre (1667-1754) in Euler (1707-1783). Resnično moč in uporabo karakterističnih funkcij pa je predstavil Ljapunov (1857-1918), ki jih je uporabil tudi za eleganten dokaz centralnega limitnega izreka, glej Mackey [4].

Lahko jo zapišemo tudi v obliki integrala za zvezno slučajno spremenljivko oziroma v obliki vsote za diskretno:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} p_X(x)\, dx \quad \text{in}$$

$$M_X(t) = \sum_{x \in Z_X} e^{tx} P(X = x),$$

kjer je Z_X zaloga vrednosti diskretne slučajne spremenljivke X. Pri osnovni definiciji lahko naredimo tudi razvoj eksponentne funkcije v Taylorjevo vrsto, glej Strang [7, pogl. 10.4], kar je pogosto uporabno pri računanju ali dokazovanju:

$$M_X(t) = \mathbb{E}(e^{tX}) = \mathbb{E}\left[1 + tX + \frac{(tX)^2}{2!} + \cdots\right] =$$
$$= 1 + t\,\mathbb{E}(X) + \frac{t^2}{2!}\,\mathbb{E}(X^2) + \cdots.$$

3 Lastnosti rodovnih funkcij

Sedaj ko smo spoznali definicijo rodovne funkcije, bomo vpeljali še nekaj lastnosti rodovnih funkcij. Te bomo tudi dokazali, uporabili pa jih bomo tako pri računanju primerov v naslednjem razdelku, kot tudi pri dokazovanju CLI.

Izrek 1. *Naj bo S končna linearna kombinacija neodvisnih slučajnih spremenljivk X_i, tj.*

$$S = \sum_{i=0}^{n} a_i X_i\,,\ a_i \in \mathbb{R}.$$

Potem je rodovna funkcija slučajne spremenljvke S enaka $M_S(t) = M_{X_1}(a_1 t) \cdot \ldots \cdot M_{X_n}(a_n t)$.

Dokaz.

$$M_S(t) = \mathbb{E}(e^{tS}) = \mathbb{E}\left(e^{t(a_1 X_1 + \ldots + a_n X_n)}\right) =$$
$$= \mathbb{E}\left(e^{ta_1 X_1} \cdot \ldots \cdot e^{ta_n X_n}\right).$$

Ker so slučajne spremenljivke $X_i\,(1 \le i \le n)$ neodvisne, lahko pričakovano vrednost produktov izračunamo kot produkt pričakovanih vrednosti in dobimo:

$$M_S(t) = \mathbb{E}\left(e^{ta_1 X_1}\right) \cdot \ldots \cdot \mathbb{E}\left(e^{ta_n X_n}\right) =$$
$$= M_{X_1}(a_1 t) \cdot \ldots \cdot M_{X_n}(a_n t). \qquad \square$$

Izrek 2. *Naj bo X slučajna spremenljivka z rodovno funkcijo M_X. Potem je rodovna funkcija slučajne spremenljivke $aX + b$, kjer sta a in b realni konstanti, enaka*

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Dokaz.

$$M_{aX+b}(t) = \mathbb{E}\left(e^{t(aX+b)}\right) = e^{bt}\,\mathbb{E}\left(e^{(at)\cdot X}\right) =$$
$$= e^{bt} M_X(at). \qquad \square$$

Izrek 3 (Enoličnost). *Naj bosta X in Y slučajni spremenljivki, F_X in F_Y njuni porazdelitveni funkciji ter M_X in M_Y njuni rodovni funkciji definirani za $t \in [-a, a]$. Potem imata X in Y enako porazdelitev, tj. $F_X(x) = F_Y(x)\ \forall x \in \mathbb{R}$, natanko tedaj, ko imata enako rodovno funkcijo, tj. $M_X(t) = M_Y(t)\ \forall t \in [-a, a]$.*

Skica dokaza. ($\Longrightarrow$) Iz definicije sledi, da imata enako porazdeljeni slučajni spremenljivki enako rodovno funkcijo. ($\Longleftarrow$) Omejimo se na diskretni slučajni spremenljivki s končno zalogo vrednosti, tj. $|Z_X|, |Z_Y| < \infty$. Njuni gostoti porazdelitve označimo s $p_X(x)$ in $p_Y(y)$, z $A = \{a_1, \ldots, a_n\}$ pa unijo njunih zalog vrednosti, tj. $A = Z_X \cup Z_Y$. Sedaj zapišimo dejstvo, da sta rodovni funkciji za vsak $t \in [-a, a]$ enaki:

$$\sum_{i=1}^{n} e^{ta_i} p_X(a_i) = \sum_{i=1}^{n} e^{ta_i} p_Y(a_i)$$

oziroma

$$\sum_{i=1}^{n} e^{ta_i} \left[p_X(a_i) - p_Y(a_i)\right] = 0.$$

Upoštevali smo, da $p_X(a_i) = 0$, če $a_i \notin Z_X$ in podobno za slučajno spremenljivko Y. Za

$$s = e^t \quad \text{in} \quad c_i = p_X(a_i) - p_Y(a_i)$$

zgornja enakost preide v

$$\sum_{i=1}^{n} s^{a_i} c_i = 0$$

za neskončno vrednosti $s > 0$. To je pravzaprav polinom za spremenljivko s s koeficienti c_i. Ta bo enak 0 pri vseh vrednostih spremenljivke s le, če bodo vsi koeficienti enaki 0, tj. $p_X(a_i) = p_Y(a_i)$. Iz tega sledi, da sta spremenljivki X in Y enako porazdeljeni. □

Izrek 4. *Naj bo X slučajna spremenljivka. Rodovna funkcija $M_X(t)$ je končna za $t = 0$ oz. $M_X(0) = 1$.*

Dokaz. $M_X(0) = \mathbb{E}(e^{0 \cdot X}) = \mathbb{E}(e^0) = \mathbb{E}(1) = 1.$ □

Izrek 5. *Naj bo X slučajna spremenljivka in M_X njena rodovna funkcija. Potem za poljubno naravno število k velja*

$$\mathbb{E}(X^k) = M_X^{(k)}(0),$$

kjer $M_X^{(k)}$ predstavlja k-ti odvod.

Dokaz. Po pravilu za odvod posredne funkcije izračunajmo k-ti odvod rodovne funkcije slučajne spremenljivke X:

$$M_X^{(k)}(t) = \frac{d^k}{dt^k} \mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left(X^k e^{tX}\right).$$

Za $t = 0$ dobimo:

$$M_X^{(k)}(0) = \mathbb{E}\left(X^k e^{0 \cdot X}\right) = \mathbb{E}\left(X^k\right),$$

kar smo želeli pokazati. □

4 PRIMERI RODOVNIH FUNKCIJ MOMENTOV

Primer 1. Naj bo X diskretna slučajna spremenljivka podana z verjetnostno tabelo

$$X \sim \begin{pmatrix} x_1 & \ldots & x_n \\ p_1 & \ldots & p_n \end{pmatrix}.$$

Po definiciji lahko rodovno funkcijo M_X izračunamo kot vsoto:

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{x \in Z_X} e^{tx} P(X = x) = \underline{\underline{\sum_{i=1}^{n} p_i e^{tx_i}}}.$$

Primer 2. Naj bo X_i Bernoullijeva slučajna spremenljivka, tj. $X_i \sim B(1, p)$. Kot v prejšnjem primeru, rodovno funkcijo izračunamo po definiciji kot vsoto:

$$\begin{aligned} M_{X_i}(t) = \mathbb{E}(e^{tX_i}) &= \sum_{x \in Z_{X_i}} e^{tx} P(X_i = x) = \\ &= e^{t \cdot 0}(1 - p) + e^{t \cdot 1} p = \underline{\underline{1 - p + pe^t}}. \end{aligned}$$

Primer 3. Naj bo X diskretna slučajna spremenljivka, ki je pozradeljena binomsko, tj. $X \sim B(n, p)$. Lahko jo zapišemo kot vsoto n neodvisnih Bernoullijevih slučajnih spremenljivk X_i, tj. $X = X_1 + \ldots + X_n$. Po izreku 1 in primeru 2 izračunamo rodovno funkcijo momentov:

$$\begin{aligned} M_X(t) &= M_{X_1}(t) \cdot \ldots \cdot M_{X_n}(t) = (M_{X_i}(t))^n = \\ &= \underline{\underline{(1 - p + pe^t)^n}}. \end{aligned}$$

Primer 4. Naj bo X zvezna slučajna spremenljivka, ki je porazdeljena eksponentno, tj. $X \sim \text{Exp}(\lambda)$, kjer je njena gostota verjetnosti enaka $p_X(x) = \lambda e^{-\lambda x}$ za $x \geq 0$. Njeno rodovno funkcijo lahko izračunamo kot integral:

$$\begin{aligned} M_X(t) = \mathbb{E}(e^{tX}) &= \int_0^\infty e^{tx} p_X(x)\ dx = \\ &= \int_0^\infty e^{tx} \lambda e^{-\lambda x}\ dx = \lambda \int_0^\infty e^{x(t-\lambda)}\ dx = \\ &= \frac{\lambda}{t - \lambda} \int_0^{-\infty} e^z\ dz = \frac{\lambda}{t - \lambda} (e^z) \Big|_0^{-\infty} = \underline{\underline{\frac{\lambda}{\lambda - t}}}. \end{aligned}$$

Če želimo, da zgornji integral obstaja, mora biti stopnja potence $e^{x(t-\lambda)}$ v neskončnosti enaka 0, tj. $t - \lambda < 0$, iz česar sledi $t < \lambda$, kar je pogoj za obstoj rodovne funkcije. Pri integriranju smo uvedli tudi novo spremenljivko $z = x(t - \lambda)$, hkrati pa smo spremenili predznak zgornje meje, saj velja $t - \lambda < 0$.

Primer 5. Izračunajmo še rodovno funkcijo standardizirane normalne spremenljivke, saj bomo to funkcijo uporabili tudi pri dokazu CLI. Naj bo torej X zvezna slučajna spremenljivka, ki je porazdeljena standardno normalno, tj. $X \sim N(0,1)$, kjer je njena gostota verjetnosti $p_X(x) = (1/\sqrt{2\pi}) \cdot e^{-x^2/2}$. Tako kot v prejšnjem primeru se lotimo izračuna z integralom:

$$\begin{aligned} M_X(t) = \mathbb{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = \\ &= \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}}\, dx = \\ &= e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz = \underline{\underline{e^{\frac{t^2}{2}}}}. \end{aligned}$$

Pri integriranju smo uvedli novo spremenljivko $z = x - t$, a meje ostanejo nespremenjene. Na koncu smo v integralu dobili gostoto verjetnosti normalne porazdelitve, katere integral je 1, s čimer smo dobili končni rezultat.

5 Laplaceova transformacija

Pri rodovni funkcije gre pravzaprav za *dvostrano Laplaceovo transformacijo*, ki je za funkcijo f definirana kot

$$\mathcal{B}\{f\}(s) = \int_{-\infty}^{\infty} e^{-st} f(t)\, dt.$$

Ta ima lastnost enoličnosti. To pomeni, da za zvezni funkcije f in g velja $f(t) = g(t)\ \forall t \geq 0$, če obstaja konstanta C, da za njuni transformaciji velja $\mathcal{B}\{f\}(s) = \mathcal{B}\{g\}(s)\ \forall s > C$, glej Widder [9]. Na tej lastnosti temelji naslednji izrek, katerega dokaz zaradi tehničnih podrobnosti raje opustimo.

Izrek 6 (Konvergenca rodovnih funkcij). *Naj bodo $X_i,\ i = 1, 2, \ldots$ slučajne spremenljivke in M_{X_i} njihove rodovne funkcije, ki so končne za $t \in [-a, a]$. Naj velja*

$$\lim_{i\to\infty} M_{X_i}(t) = M_X(t), \quad t \in [-a, a],$$

kjer je M_X rodovna funkcija neke slučajne spremenljivke X. Potem obstaja enolična porazdelitvena funkcija F_X, ki ima momente določene z rodovno funkcijo M_X in za vsak x, kjer je F_X zvezna velja

$$\lim_{i\to\infty} F_{X_i}(x) = F_X(x).$$

To pomeni, da konvergenca rodovnih funkcij za $|t| < h$ implicira konvergenco porazdelitvenih funkcij.

Pri tem je h neko pozitivno število, da integral

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} p_X(x)\, dx$$

obstaja za vse $|t| < h$ in iz tega po lastnostih uporabljene transformacije sledi, da dani rodovni funkciji M_X ustreza natanko določena gostota porazdelitve p_X.

6 Dokaz CLI

Z uporabo rodovnih funkcij se sedaj lotimo dokazovanja CLI. Najprej pa ponovno zapišimo že izračunano rodovno funkcijo za standardno normalno porazdelitev iz primera 5, ki jo bomo potrebovali na koncu dokaza:

$$M_{N(0,1)}(t) = e^{\frac{t^2}{2}}.$$

Naj bodo $X_1, \cdots, X_n$ neodvisne in enako porazdeljene slučajne spremenljivke, kjer sta njihova pričakovana vrednost μ in standardni odklon σ končna, torej velja $\mu, \sigma < \infty$. Naj bo S njihova vsota:

$$S = X_1 + \cdots + X_n.$$

Izračunajmo rodovno funkcijo standardizirane spremenljivke X_i, ki je $Z_i = \frac{X_i - \mu}{\sigma}$, hkrati pa lahko predpostavimo $\sigma \neq 0$ (saj bi sicer slučajna spremenljivka bila konstantna, za katero pa CLI ne potrebujemo):

$$M_{Z_i}(t) = 1 + \frac{t^2}{2!} + \frac{t^3}{3!}\,\mathbb{E}(Z_i^3) + \cdots. \qquad (*)$$

Na koncu smo upoštevali $\mathbb{E}(Z_i) = 0$ in $\mathbb{E}(Z_i^2) = 1$. Definirajmo še povprečje $\overline{X} = S/n$, ki ima pričakovano vrednost μ in standardni odklon $\sigma/\sqrt{n}$, ter izračunajmo standardizirano povprečje Y:

$$\begin{aligned} Y = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} = \frac{1}{\sqrt{n}} \cdot \frac{n\overline{X} - n\mu}{\sigma} = \\ = \frac{1}{\sqrt{n}} \cdot \frac{\sum_{i=1}^n X_i - n\mu}{\sigma} = \frac{1}{\sqrt{n}} \cdot \sum_{i=1}^n Z_i = \sum_{i=0}^n \frac{Z_i}{\sqrt{n}}. \end{aligned}$$

Na koncu smo dobili vsoto slučajnih spremenljivk $Z_i/\sqrt{n}$, ki pa so neodvisne in enako porazdeljene, kar pomeni, da lahko izračunamo rodovno funkcijo standardnega povprečja:

$$\begin{aligned} M_Y(t) = \prod_{i=0}^n M_{\frac{Z_i}{\sqrt{n}}}(t) = \prod_{i=0}^n M_{Z_i}\left(\frac{t}{\sqrt{n}}\right) = \\ = \left[1 + \frac{t^2}{2n} + \frac{t^3}{3!\ n^{3/2}} E(Z_i^3) + \ldots\right]^n. \end{aligned}$$

V zadnji enakosti smo uporabili ($*$), kjer moramo upoštevati, da je tokrat parameter po izreku 1 enak $t/\sqrt{n}$. Ker pa so slučajne spremenljivke Z_i enako porazdeljene, imajo po izreku 3 tudi enake rodovne funkcije, zato lahko produkt združimo v en člen s stopnjo n v potenci.

Obe strani enačbe logaritmiramo ter uvedemo novo spremenljivko:

$$x = \frac{t^2}{2n} + \frac{t^3}{3!\ n^{3/2}}\mathbb{E}(Z_i^3) + \cdots,$$
$$\ln M_Y(t) = \ln(1+x)^n = n\ln(1+x).$$

Tudi za $\ln(1+x)$ naredimo razvoj v Taylorjevo vrsto in nazaj vstavimo x:

$$\begin{aligned}
&\ln M_Y(t) = n\ln(1+x) = \\
&= n\left[x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots\right] = \\
&= n\left[\left(\frac{t^2}{2n} + \frac{t^3}{3!\ n^{3/2}} + \cdots\right)\right. \\
&\quad \left. - \frac{1}{2}\left(\frac{t^2}{2n} + \frac{t^3}{3!\ n^{3/2}} + \cdots\right)^2 + \cdots\right].
\end{aligned}$$

Sedaj izračunamo limito funkcije $\ln M_Y(t)$, ko gre $n \to \infty$:

$$\lim_{n\to\infty} \ln M_Y(t) = \frac{t^2}{2}$$

in na obeh straneh uporabimo eksponentno funkcijo, ki je inverzna funkcija logaritma $\ln$, hkrati pa tudi zvezna, kar pomeni, da lahko zamenjamo vrstni red eksponentne funkije in limitiranja:

$$\lim_{n\to\infty} e^{\ln M_Y(t)} = e^{\frac{t^2}{2}}.$$

Kompozitum funkcije in njenega inverza je identiteta in zato velja

$$\lim_{n\to\infty} M_Y(t) = e^{\frac{t^2}{2}}.$$

To je pravzaprav rodovna funkcija standarne normalne porazdelitve, kar pomeni, da $M_Y(t)$ limitira proti $M_{N(0,1)}(t)$. Od tod po izreku 6 sledi, da porazdelitvena funkcija Y konvergira proti porazdelitveni funkciji standardne normalne, kar smo želeli pokazati.

7 Zaključek

Naš cilj je bil predstavitev dokaza centralnega limitnega izreka. Pri tem so nam močno pomagale rodovne funkcije, iz česar vidimo, da so močno orodje pri dokazovanju. Poleg samega dokaza smo nakazali tudi pomembne lastnosti rodovnih funkcij, ki pa smo jih nato uporabili tudi pri računanju teh funkcij za konkretne porazdelitve. Te funkcije lahko nato uporabimo za računanje momentov porazdelitev, ki je po tej poti lahko bistveno lažje kot po običajni poti z vsoto oziroma integralom.

Z zgoraj napisanim smo se le dotaknili resnične moči rodovnih funkcij, v dodatku pa bomo predstavili še karakteristične funkcije. Te so pravzaprav še močnejše, vendar tudi nekoliko zahtevnejše za obravnavo. Njihova resnična moč se skriva v možnosti pretvarjanja med omenjenima oblikama funkcij ter samo porazdelitvijo slučajne spremenljivke. To nam omogočata Fourierjeva in Laplaceova transformacija. V primeru karakterističnih funkcij gre za Fourierjevo transformacijo porazdelitvene funkcije ali gostote porazdelitve slučajne spremenljivke, pri rodovnih funkcijah pa za dvostrano Laplaceovo transformacijo gostote porazdelitve.

Omeni nekje še Romanovo pripombo, da so F. in L. transformiranke uporabne pri reševanju D.E.

A Karakteristične funkcije

Karakteristična funkcija slučajne spremenljivke X je funkcija $\varphi_X : \mathbb{R} \to \mathbb{C}$, ki je odvisna od parametra $t \in \mathbb{R}$ in je definirana z:

$$\varphi_X(t) = \mathbb{E}(e^{itX}),$$

kjer je i imaginarna enota z lastnostjo $i^2 = -1$. Po definiciji pričakovane vrednosti lahko karakteristično funkcijo zapišemo tudi kot integral pri pri zvezni slučajni spremenljivki oziroma vsoto pri disktreni:

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} p_X(x)\,dx = \quad \text{in}$$

$$\varphi_X(t) = \sum_{x\in Z_X} e^{itx} P(X = x),$$

kjer je Z_X zaloga vrednosti slučajne spremenljivke X. Eksponentno funkcijo še razvijemo v Taylorjevo vrsto:

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}\left[1 + itX + \frac{(itX)^2}{2!} + \cdots\right],$$

ter uporabimo linearnost pričakovane vrednosti in poračunamo vrednosti i^k:

$$\varphi_X(t) = 1 + it\,\mathbb{E}(X) - \frac{t^2}{2!}\,\mathbb{E}(X^2) - \cdots.$$

Z uporabo Eulerjeve formule lahko karakteristično funkcijo zapišemo tudi kot

$$\begin{aligned}
\varphi_X(t) &= \mathbb{E}([\cos(tX) + i\sin(tX)] = \\
&= \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)].
\end{aligned}$$

Izrek 1. *Naj bo X slučajna spremenljivka s porazdelitveno funkcijo F. Za njeno karakteristično funkcijo $\varphi(t) = \mathbb{E}(e^{itX})$ velja:*

- $|\varphi(t)| \le \varphi(0) = 1$,
- $\varphi(-t) = \varphi(t)$,
- *funkcija $\varphi(t)$ je enakomerno zvezna,*
- $\varphi_{aX+b}(t) = \varphi_X(at) \cdot e^{itb}$.

Izrek 2. *Naj bo X slučajna spremenljivka s karakteristično funkcijo $\varphi(t)$. Če za neko naravno število n velja $\mathbb{E}(|X|^n) < \infty$, potem za vsako $k \leq n$ obstaja odvod $\varphi^{(k)}(t)$ in veljajo naslednje enakosti:*

- $\varphi^{(k)}(t) = \int_{-\infty}^{\infty} (ix)^k e^{itx}\, dF(x), \quad k = 0, 1, \ldots, n,$
- $\mathbb{E}(X^k) = \dfrac{\varphi^{(k)}(0)}{i^k}, \quad k = 0, 1, \ldots, n,$
- $\varphi(t) = \sum_{k=0}^{n} \dfrac{(it)^k}{k!}\mathbb{E}(X^k) + \dfrac{(it)^n}{n!} r_n(t),$

  kjer je $|r_n(t)| \leq 3\mathbb{E}(|X|^n)$ *in* $\lim_{t\to 0} r_n(t) = 0.$

Izrek 3 (**Produkt karakterističnih funkcij**). *Naj bodo $X_1, X_2, \ldots, X_n$ neodvisne slučajne spremenljivke. Potem velja*

$$\varphi_{X_1+X_2+\ldots+X_n}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \ldots \cdot \varphi_{X_n}(t).$$

Iz definicije karakteristične funkcije sledi, da porazdelitvena funkcija enolično določa karakteristično funkcijo. Velja tudi obratno. To pomeni, da karakteristična funkcija enolično določa porazdelitveno funkcijo, kar nam pove naslednji izrek.

Izrek 4 (**Enoličnost**). *Če imata neki porazdelitveni funkciji isto karakteristično funkcijo, potem sta ti porazdelitveni funkciji enaki.*

Ker karakteristična funkcija enolično določa porazdelitveno funkcijo, lahko iz prve izračunamo drugo. To nam omogoča formula inverzije.

Izrek 5 (**Formula inverzije**). *Naj bosta a in b, kjer je $a < b$, točki zveznosti porazdelitvene funkcije F in φ ustrezna karakteristična funkcija. Potem velja*

$$F(b) - F(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)\, dt.$$

Pri zveznih slučajnih spremenljivkah velja

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t)\, dt.$$

B SPLOŠNE RODOVNE FUNKCIJE

Rodovne funkcije se v splošnem uporabljajo kot način zapisa zaporedja števil (tu ne mislimo več na rodovne funkcije momentov). Namesto zaporedja raje zapišemo neko funkcijo, kjer je n-ti člen zaporedja koeficient pri spremenljivki x^n. Te funkcije so torej zapisane kot potenčne vrste, glej Strang [7, pogl. 10.5]. Konvergenca pri teh vrstah ni nujno potrebna, ker teh funkcij ne obravnavamo nujno kot običajne funkcije. Tudi spremenljivka x ostane nedoločena, namesto njene vrednosti ali vrednosti funkcije, pa nas zanimajo koeficienti.

Naj bo $(a_n)_{n=0}^{\infty}$ zaporedje števil. Potem je rodovna funkcija tega zaporedja:

$$f(x) = \sum_{n\geq 0} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots.$$

Vpeljimo še notacijo $[x^k]\, f(x)$, ki predstavlja koeficient funkcije f pri k-ti potenci spremenljivke, tj. x^k.

Primer 1. Najprej izračunajmo rodovno funkcijo za preprosto zaporedje $1, 1, \ldots$ oziroma $a_n = 1$. Po definiciji je to funkcija:

$$f(x) = 1 + x + x^2 + x^3 + \cdots.$$

Vidimo, da gre pravzaprav za geometrično zaporedje z začetnim členom $a_0 = 1$ in koeficientom $q = x$. Po formuli za vsoto neskončnega geometričnega zaporedja $\sum_{i\geq 0} a_i = \frac{1}{1-q}$ lahko izračunamo vrednost te funkcije:

$$f(x) = 1 + x + x^2 + \cdots = \underline{\underline{\frac{1}{1-x}}}.$$

Kot smo omenili na začetku dodatka, nas konvergenca pri teh funkcijah ne zanima. To je tudi razlog, da smo lahko uporabili formulo za vsoto geometričnega zaporedja, čeprav zahteva $|x| < 1$.

Primer 2. Podobno lahko naredimo tudi za končno zaporedje $1, 1, \ldots, 1$ oziroma $a_i = 1,\ i \leq n$ in $a_i = 0,\ i > n$. Ponovno dobimo geometrično zaporedje z začetnim členom $a_0 = 1$ in koeficientom $q = x$, tokrat pa uporabimo formulo za vsoto končnega geometričnega zaporedja $\sum_{i=0}^{n} a_i = \frac{1-q^{n+1}}{1-q}$ in dobimo:

$$f(x) = 1 + x + x^2 + \ldots + x^n = \underline{\underline{\frac{1 - x^{n+1}}{1-x}}}.$$

Primer 3. Sedaj vzemimo zaporedje, kjer je začetnih k členov enakih 0, tj. zaporedje $0, 0, \ldots, 0, 1, \ldots$ oziroma $a_i = 0,\ i < k$ in $a_i = 1,\ i \geq k$. Rodovna funkcija tega zaporedja je:

$$f(x) = x^k + x^{k+1} + \cdots.$$

Iz vseh členov lahko izpostavimo x^k, s čimer v oklepaju dobimo rodovno funkcijo iz primera 1 in uporabimo njen rezultat:

$$f(x) = x^k(1 + x + x^2 + \ldots) = \underline{\underline{\frac{x^k}{1-x}}}.$$

Iz tega tudi sledi, da x^k lahko *izpostavimo* in iščemo manjši koeficient:

$$[x^n]\, x^k \frac{1}{1-x} = [x^{n-k}] \frac{1}{1-x}.$$

Pri računanju koeficientov neke rodovne funkcije si lahko pomagamo z Binomskim izrekom, ki velja tudi za negativna števila [8].

Binomski izrek za negativna števila.

$$(x+y)^{-n} = \sum_{k=0}^{\infty} \binom{-n}{k} x^k y^{-n-k}$$
$$= \sum_{k=0}^{\infty} (-1)^k \binom{n+k-1}{k} x^k y^{-n-k}.$$

Druga enakost sledi iz definicije binomskega simbola:

$$\binom{-n}{r} = \frac{(-n)(-n-1)\cdot\ldots\cdot(n-r+1)}{r(r-1)\cdot\ldots\cdot 2\cdot 1} =$$
$$= \frac{(-1)^r n(n+1)\cdot\ldots\cdot(n+r-1)}{r(r-1)\cdot\ldots\cdot 2\cdot 1} =$$
$$= (-1)^r \frac{(n+r-1)!}{r!(n-1)!} = (-1)^r \binom{n+r-1}{r}.$$

To je pravzaprav formula za število kombinacij s ponavljanjem, le da imamo tukaj še negativni predzank pri lihem številu izbir. Opazimo lahko še, da se ta minus pri funkciji $(1-x)^{-n}$ pokrajša:

$$(1-x)^{-n} = (-x+1)^{-n} =$$
$$= \sum_{k=0}^{\infty} (-1)^k \binom{n+k-1}{k} (-x)^k 1^{-n-i} =$$
$$= \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k.$$

Za konec poskusimo rodovne funkcije uporabiti za izračun nekega kombinatoričnega problema. Recimo, da želimo izračunati število izbir 500 žog izmed rdečih, modrih in belih, kjer želimo, da izberemo 50 do 200 modrih in kvečjemu 400 belih. Za vsako barvo bomo naredili rodovno funkcijo, ki bo spremljala število izbir. Koeficient pri k-ti potenci spremenljivke bo predstavljal število izbir k žog. Iz rodovnih funkcij za vsako barvo bomo nato naredili rodovno funkcijo za naš primer. Ker iščemo število izbir 500 žog, pravzaprav iščemo koeficient pri spremenljivki x s stopnjo 500.

Naredimo najprej rodovno funkcijo za rdeče žoge. Žog ne razlikujemo, torej imamo 1 možnost izbora 0 žog, 1 možnost izbora 1 žoge itd. To pomeni, da bodo vsi koeficienti pri tej rodovni funkciji enaki 1:

$$R(x) = 1 + x + \ldots = \frac{1}{1-x}.$$

Ker želimo 50 do 200 modrih žog, imamo 1 možnost izbora 50 žog, 1 možnost izbora 51 žog itd. do 200. Za preostale izbire imamo 0 možnosti, torej bo rodovna funkcija enaka:

$$M(x) = x^{50} + \ldots + x^{200} = x^{50}(1 + \ldots + x^{150}) =$$
$$= x^{50} \frac{1-x^{151}}{1-x}.$$

Podobno naredimo še za bele žoge. Želimo kvečjemu 400 belih žog. Torej imamo 1 možnost izbora 0 žog, 1 možnost izbora 1 žoge itd. do 400. Za preostale izbire imamo 0 možnosti. Rodovna funkcija je enaka:

$$B(x) = 1 + x + \ldots + x^{400} = \frac{1-x^{401}}{1-x}.$$

To so rodovne funkcije za izbor žog ene barve. Nas pa zanima izbor izmed vsemi tremi barvami. Gre za produkt posameznih dogodkov, zato je končna rodovna funkcija enaka:

$$f(x) = R(x)\,M(x)\,B(x) =$$
$$= \frac{1}{1-x} \cdot x^{50} \frac{1-x^{151}}{1-x} \cdot \frac{1-x^{401}}{1-x} =$$
$$= x^{50} \cdot \frac{(1-x^{151})(1-x^{401})}{(1-x)^3} =$$
$$= x^{50}(1-x^{151})(1-x^{401})(1-x)^{-3} =$$
$$= x^{50}(1-x^{151}-x^{401}+x^{552})(1-x)^{-3}.$$

Že na začetku smo povedali, da pravzaprav iščemo koeficient pri spremenljivki s stopnjo 500. Ta koeficient bo predstavljal naš končni rezultat.

$$[x^{500}]\, x^{50}(1-x^{151}-x^{401}+x^{552})(1-x)^{-3} =$$
$$= [x^{450}]\,(1-x)^{-3} - [x^{299}]\,(1-x)^{-3} -$$
$$- [x^{40}]\,(1-x)^{-3} =$$
$$= \binom{3+450-1}{450} - \binom{3+299-1}{299} -$$
$$- \binom{3+49-1}{49} =$$
$$= 101\,926 - 1\,275 - 45\,150 = 55\,501$$

To pomeni, da imamo 55 501 možnosti za izbiro žog.

LITERATURA

[1] *Tabela standardne normalne porazdelitve*. https://www.mathsisfun.com/data/standard-normal-distribution-table.html.

[2] *What is the rationale behind the magic number 30 in statistics*. https://www.researchgate.net/post/What_is_the_rationale_behind_the_magic_number_30_in_statistics.

[3] Casella, G.: *Statistical Inference*. Cengage Learning, 2001.

[4] Mackey, G. W.: *Harmonic analysis as the exploitation of symmetry–a historical survey*. Bulletin (New Series) of the American Mathematical Society, 3(1.P1):543 – 698, 1980. https://doi.org/10.1090/S0273-0979-1980-14783-7.

[5] Mladenović, P.: *Verovatnoća i statistika*. Matematički fakultet Beograd, 2008.

[6] Pishro-Nik, H.: *Introduction to probability, statistics, and random processes*. Kappa Research LLC, 2014. https://www.probabilitycourse.com.

[7] Strang, G.: *Calculus*. Wellesley-Cambridge Press, 1991.

[8] Weisstein, E. W.: *"negative binomial series." from mathworld–a wolfram web resource*. https://mathworld.wolfram.com/NegativeBinomialSeries.html.

[9] Widder, D. V.: *The Laplace Transform*. Princeton University Press, 1946.

[10] Wilf, H. S.: *Generatingfunctionology*. A. K. Peters, Ltd., USA, 2006, ISBN 1568812795.