

Monte-Carlo-módszerek a statisztikában*

Kehl Dániel,
a Pécsi Tudományegyetem
Közgazdaságtudományi
Karának tanársegéde
E-mail: kehd@ktk.pte.hu

A tanulmány a Monte-Carlo-módszerek statisztikai alkalmazásáról nyújt áttekintést és néhányat be is mutat, elsősorban a véletlen értékek generálásának technikája és a numerikus integrálás területéről. Az MC-integráláshoz kapcsolódóan ismerteti a főbb varianciacsökkentő módszerek alkalmazását is. Ezek az ismeretek elkerülhetetlenül szükségesek a bayesi statisztika alkalmazása esetén, ahol a poszterior eloszlásban rejlő információk kinyerése gyakran csak szimuláció segítségével lehetséges. A cikkben bemutatott módszerek közös jellemzője, hogy alacsony dimenziószám esetén alkalmazhatók hatékonyan.

TÁRGYSZÓ:
Monte-Carlo-módszer.
Véletlenérték-generálás.
Varianciacsökkentő módszerek.

* A szerző ezúton mond köszönetet a tanulmányhoz fűzött megjegyzésekért és tanácsokért *Hunyadi László* professzornak, *Abaliget Gallusznak* és intézeti kollégáinak, valamint a Rosztoczy Alapítványnak anyagi támogatásáért. Minden esetleges fennmaradó hibáért természetesen a szerzőt terheli a felelősség.

Monte-Carlo-módszerek összefoglaló névvel illetünk számos eljárást, technikát, melyek közös jellemzője, hogy véletlenszám-sorozatok generálásán alapulnak. A módszerek népszerűségének oka rendkívül egyszerű: analitikusan követhetetlen feladatok eredményeit vagyunk képesek tetszőleges közelítéssel meghatározni velük. A robbanásszerű elterjedéshez a matematikai alapok lefektetésén kívül szükség volt egy másik összetevőre, a véletlen értékeket generáló, a számításokat gyorsan elvégző számítógépekre. A bemutatott példákat a népszerű, ingyenes R (2011) programcsomag segítségével oldottam meg, kérésre a kódokat rendelkezésre bocsátom.

Napjainkra a modern statisztika alapvető eszközévé nőtte ki magát ez a széles terület, melyről megkísérlek átfogó, de nem túlságosan mély képet nyújtani. A tanulmány felépítése a témával foglalkozó szakkönyvek (*Albert* [2009]; *Casella–Berger* [2002]; *Rizzo* [2008]; *Robert–Casella* [2004], [2010]) struktúráját követi, azokból néhány példát is átvesz. Az általános bevezető után a különböző eloszlásokból való véletlenérték-generálás egyszerű technikáit szemlélteti, majd az egyik gyakran alkalmazott területet, a Monte-Carlo-integrálást és az ehhez kapcsolódó varianciacsökkentő módszereket tárgyalja. Fő célja a módszerek mögötti intuíció és a lehetséges felhasználási területek bemutatása.

1. A Monte-Carlo-módszerekről általában

A véletlen események felhasználásának ötlete nem új a statisztikában, már a számítógépek megjelenése előtt is voltak alkalmazásai, elég csak a Buffon-féle túproblémára (π közelítése a padlóra dobott tűk segítségével a XVIII. században) vagy a *Gossett* nevéhez fűződő, t -eloszlásról szóló cikkekre (*Student* [1908]) utalni. A véletlen értékek felhasználásának történetéről, a módszerek fejlődéséről az érdeklődő Olvasónak például *Robert–Casella* [2011] nyújt kimerítő irodalomjegyzéket.¹

A valós statisztikai alkalmazások felsorolása, azok sokszínűsége miatt szinte lehetetlen: hagyományos alkalmazási terület a különböző tesztek erőfüggvényeinek kiszámítása, kritikus értékek vagy becslőfüggvények jellemzőinek (paraméterek, MSE, percentilisek stb.), konfidenciaintervallumok takarási valószínűségének meghatározása. Ebbe a körbe tartoznak a Bootstrap- és Jackknife-módszerek is. Ezen kívül fontos szerepet töltenek be a Monte-Carlo- (MC-) és Markov-lánc Monte-Carlo- (MCMC-) módszerek a bayesi statisztikában (*Hunyadi* [2011]), ahol a feladat össze-

¹ Lásd erről *Kehl* [2012] *Statisztikai Szemle* hasábjain megjelent rövid ismertetőjét.

tett, sokdimenziós sűrűségfüggvények (poszteriorok) leírása, amely szinte minden esetben integrálok meghatározását jelenti a gyakorlatban. A harmadik nagy felhasználási terület a sztochasztikus optimalizáció, amely összetett függvények szélsőértékeit, illetve szélsőértékhelyeit keresi. Tipikusan ilyen probléma az összetett likelihood függvények maximumának keresése ML-becslés meghatározásakor.

Tanulmányomban a véletlen számok generálásáról általában, valamint az egyszerűbb MC integrálási technikákról ejtek szót. Az MC optimalizációs, valamint a MCMC-technikák bemutatása jelen írásnak nem célja, utóbbiakat egy későbbi cikkben tervezem tárgyalni. A szimulációk következtetéses statisztikában történő alkalmazása és a sztochasztikus optimalizáció önmagában szintén egy-egy átfogó dolgozat témája lehetne.

2. Véletlen számok egyszerű generálási technikái

A számítógépes véletlenszám-generálás alapja gyakorlatilag az egyenletes eloszlás. Nem kívánok részletesen foglalkozni azzal, hogy a számítógépek csupán ún. pszeudo véletlen szám létrehozására képesek. Az ilyen módszerek tulajdonképp egy hosszú sorozatot állítanak elő, amely matematikai tulajdonságai alapján megfelelő minőségűnek tekinthető.² Valamennyi, számítási célokat is szolgáló programcsomag tartalmaz egyenletes eloszlásból származó véletlenszám-generátort, R-ben ez a függvény a *runif()*, Excelben a *vél()*, Matlabban pedig a *rand()*. A szoftverek alapértelmezésben az ún. Mersenne Twister-eljárást használják, amely egy gyors, jó minőségű, pszeudo véletlen számokat generáló algoritmus (*Matsumoto–Nishimura* [1998]). Véletlen számon tehát ezentúl pszeudo, számítógép által generált véletlen számokat értek. A különböző ismert, gyakran használt eloszlásokból származó véletlen számok hatékony generálásának komoly irodalma van, amivel szintén nem kívánok foglalkozni.

A következőkben röviden bemutatom – egy-egy illusztratív példával kiegészítve – a leggyakrabban alkalmazott véletlenszám-generálási eljárásokat. Hangsúlyozom, hogy a példák döntő többsége pusztán illusztratív, a cél a módszerek mögötti intuíció, az előnyök és hátrányok bemutatása, nem hatékony módszerek fejlesztése.

2.1. Inverz eloszlásfüggvény módszer

A véletlen számok generálásának talán legegyszerűbb módja az ún. inverz eloszlásfüggvény módszer (inverse transform method), hátránya azonban, hogy nem min-

² Létezik olyan R-csomag, mely a www.random.org honlapon keresztül igazi véletlen számokat használ, azonban tudományos célokra a megfelelően jó minőségű pszeudo véletlen számok teljesen elfogadottak.

den esetben alkalmazható (például többváltozós eloszlások). Ha X folytonos véletlen változó $F_X(x)$ eloszlásfüggvénnyel, akkor $U = F_X(X) \sim Unif(0,1)$, azaz egyenletes eloszlású a $[0,1]$ intervallumon (ún. probability integral transform).

Hasonlóan belátható, hogy (amennyiben az inverz létezik) $F_X^{-1}(U)$ eloszlása megegyezik X eloszlásával, azaz a feladatunk $F_X^{-1}(U)$ meghatározása, majd véletlen $u \sim Unif(0,1)$ generálása és $F_X^{-1}(u)$ kiszámítása. Az elmélet kiterjeszthető (Angus [1994]) – az inverz fogalom általánosításával – többek között a folytonos eloszlásokról diszkrétre.

Példaként tekintsük a Cauchy-eloszlás eloszlásfüggvényét: $F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x-\mu}{\sigma}\right)$, amiből az $F^{-1}(u) = \mu + \sigma \tan\left(\pi\left(u - \frac{1}{2}\right)\right)$ inverz egyszerű átrendezéssel adódik.

Nincs más dolgunk tehát, mint a kívánt számú 0–1 közötti egyenletes eloszlású érték generálása, majd azokon az inverz transzformáció elvégzése. Hasonló módon állítható elő például exponenciális, logisztikus vagy Rayleigh-eloszlású véletlenváltozó-sorozat. A csonkolt normális eloszlás példáját mutatja be *Várpalotai* [2008] dolgozata függelékében.

2.2. Direkt transzformációs módszer

A kívánt véletlen értékek előállítására sok esetben megoldható ismert eloszlások közötti matematikai összefüggések segítségével. Standard normális véletlen változók négyzetre emelésével és összegzésével állíthatunk elő χ_k^2 eloszlást. *Casella* és *Berger* ([2002] 627. old.) átfogó képet adnak a gyakran alkalmazott eloszlások kapcsolati hálójáról, ami alapján a módszer könnyedén implementálható. Példaként említhetném még a lognormális véletlen változó generálását standard normális eloszlásból, vagy F -, illetve Student t -eloszlású értékek létrehozását. Transzformáción alapul a normális eloszlású változókat generáló *Box–Müller* [1958] algoritmus is, amely egy egyenletes változópárból normális eloszlású változópárt állít elő. A direkt transzformációs módszer nyilvánvaló hátránya, hogy nem szokványos eloszlások esetén ilyen lehetőség ritkán áll fent.

2.3. Az elfogadás-elutasítás módszere

Az elfogadás-elutasítás módszer (acceptance-rejection method) alkalmazásához szükségünk van egy olyan eloszlásra (forráseloszlásra – g), melyből könnyedén tu-

dunk véletlen számokat generálni, ráadásul megfelelően „közel” van ahhoz az eloszláshoz, melyből generálni szeretnénk (céleloszlás – f). Legyen X és Y két véletlen változó és jelölje sűrűségfüggvényüket rendre f és g . Tegyük fel továbbá, hogy létezik olyan c konstans, melyre

$$\frac{f(t)}{g(t)} \leq c \quad /1/$$

fennáll minden olyan t -re, ahol $f(t) > 0$. A cél elegendően alacsony, lehetőleg a legalacsonyabb c megtalálása /1/-ben egy olyan g -hez, amely elég hatékony és könnyen generálható. Amennyiben megtaláltuk a megfelelő forráseloszlást és a hozzá tartozó konstans, a következő lépéseket kell elvégeznünk:

1. Generáljunk egy véletlen y számot Y eloszlásból.
2. Generáljunk $u \sim Unif(0, c \times g(y))$ egyenletes eloszlású véletlen értéket.
3. Amennyiben teljesül $u < f(y)$, fogadjuk el y -t X -ből származó véletlen számként, azaz $x := y$, ellenkező esetben utasítsuk el, majd térjünk vissza az 1. pontra.

Adott $Y = y$ feltételhez tartozó elfogadási valószínűség tehát $\frac{f(y)}{cg(y)}$ a 3. lépés és

az egyenletes eloszlás eloszlásfüggvénye alapján. Bármely iteráció összesített (feltétel nélküli) elfogadási valószínűsége $\int_{-\infty}^{\infty} \frac{f(y)}{cg(y)} g(y) dy = \frac{1}{c}$, így egy X -ből származó véletlen szám átlagosan c iterációt, azaz $2c$ (c a forrás, c az egyenletes eloszlásból) véletlen szám generálását igényli. Amennyiben nem találjuk meg a megfelelő (minimális) c -t, a módszer alkalmazható marad, de nem hatékony. A c konstans tartalmilag a javasolt forráseloszlás maximális távolságát méri a céleloszlástól.

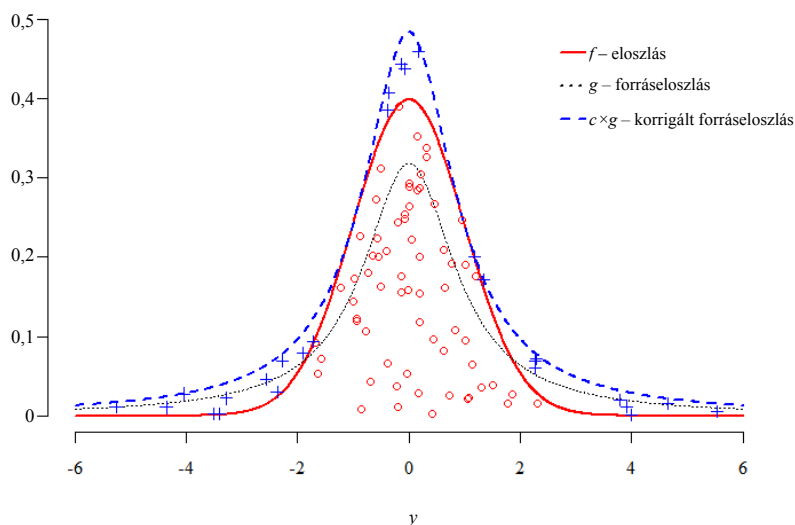
Az elfogadás-elutasítás módszer bemutatására standard normális változókat állítunk elő. Első lépésként egy, a standard normálishoz hasonló, könnyen generálható eloszlást kell keresnünk. Legyen ez a már megismert Cauchy-eloszlás standard változata, hisz abból könnyedén tudunk generálni a megírt inverz eloszlásfüggvény eljárás vagy beépített függvény segítségével. Megfelelő választás lenne természetesen bármely egyéb, a valós tengelyen értelmezett függvény is. Praktikus, ha olyan függvényt választunk, mely vastagabb eloszlásszélességgel rendelkezik, mint a céleloszlás. Második lépésként meg kell határoznunk a lehető legkisebb konstans /1/-ben a kiválasztott g -hez, ehhez írjuk fel a sűrűségfüggvények hányadosát:

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}{\frac{1}{\pi(x^2+1)}} = \frac{\sqrt{\pi}(x^2+1)}{\sqrt{2}} e^{-\frac{1}{2}x^2} \leq c.$$

A hányados felülről korlátos, azaz a Cauchy-eloszlás megfelelő forrás a normális céleloszláshoz. Keressük meg azokat az x_0 értékeket, melyeknél a függvény a maximumát veszi fel. A hányados deriváltja alapján könnyen megállapítható, hogy a függvénynek két maximuma van az $x_0 = \pm 1$ pontokban (valamint lokális minimuma az $x = 0$ pontban). A maximumhelyeken a függvény értéke, azaz a lehetséges minimális konstans $\sqrt{2\pi}e^{-\frac{1}{2}} = c \approx 1,52$.

A módszer megértését segíti az 1. ábra, melyen a standard normális (f), a Cauchy (g) és a konstanssal szorzott Cauchy ($c \times g$) eloszlásokat, valamint 100 iterációval kapott véletlen értékeket ábrázoltam.

1. ábra. Elfogadás-elutasítás módszer



A folytonos vonallal feltüntetett normális eloszlásból kívánunk generálni, még-hozzá a pontozott vonallal ábrázolt Cauchy-eloszlás segítségével. Ehhez megkeres-tem azt a legkisebb c -t, amellyel a Cauchy-eloszlást szorozva a megnyújtott görbe lefedí a teljes céleloszlást (szaggatott vonal). Ezután a Cauchy-eloszlásból generálunk egy véletlen számot (y), amiről a 0 és $c \times g(y)$ közötti egyenletes eloszlás-

ből származó véletlen érték és az adott pontban érvényes $f(y)$ dönt: a két görbe között helyezkedik-e el (elutasítás – kereszt) vagy a normális eloszlás sűrűségfüggvénye alatt (elfogadás – kör). A körrel jelölt pontok első koordinátái standard normális valószínűségi változóból generált véletlen értékeket képeznek.

Ahogy említettem, c egyben az egy céleloszlásból származó véletlen számhoz szükséges iterációk átlagos számát is jelenti. Amennyiben például 10 000 standard normális valószínűségi változót szeretnénk generálni, úgy átlagosan 15 200 iterációra, azaz 30 400 véletlen szám előállítására van szükség.

Végül megjegyzendő, hogy a módszer abban az esetben is alkalmazható, ha a céleloszlásnak csupán az alakját ismerjük, a normalizáló konstans nem, ahogy ez a bayesi statisztikában gyakran előfordul. Ebben az esetben azonban \hat{c}^{-1} nem az elfogadás valószínűsége, mert az ismeretlen normalizáló konstans „beszivárog” \hat{c} -be.

3. Integrálási módszerek

Első látásra talán szokatlan integrálási módszerekről olvasni statisztikai tanulmányban, mégis gyakran kell élnünk ezzel az eszközzel. Integrálás eredményeképp kaphatjuk meg folytonos valószínűségi változók várható értékét, egyéb momentumaikat, kvantiliseit. A bayesi statisztikában mind a prior, mind a poszterior sűrűségfüggvénnyel írható le, a normalizáló konstans (ami az egységnyi integrálértéket biztosítja) azonban gyakran nem ismert és analitikusan nem is meghatározható. Az ilyen és ehhez hasonló esetek megoldására mutatok be olyan módszereket, melyek analitikusan nem kezelhető határozott integrálok meghatározására szolgálnak. Az ismert determinisztikus módszerek a függvényt egyszerű alakzatokkal közelítik, hátrányuk, hogy magasabb dimenziószám esetén konvergenciájuk lassul. A véletlen értékek generálásán alapuló ún. MC-módszerek implementálásának egyszerűsége magasabb dimenziószám esetén is megmarad, ezért összetettebb, sokváltozós problémák esetén előszeretettel használják őket. A hagyományos MC-becslés varianciáját csökkentő eljárásokat a 4. fejezetben fogom bemutatni.

3.1. Determinisztikus módszerek

Tekintsünk egy egydimenziós integrált, melyet úgy közelítünk, hogy az integrálási intervallumot k részre osztjuk, a részintervallumokra kapott integrálokat pedig összegezzük:

$$\int_a^b f(x) dx = \sum_{i=0}^{k-1} \int_{x_i}^{x_{i+1}} f(x) dx, \quad /2/$$

ahol az $[a, b]$ intervallumot $[x_i, x_{i+1}]$ intervallumokra osztottuk, úgy, hogy $i = 0, 1, \dots, k-1$ és $x_0 = a$, $x_k = b$.

A részintervallumok területének közelítése alapján különböző módszerek léteznek. A legegyszerűbb esetben a bal oldali végpontban vett függvényérték és az $x_{i+1} - x_i$ lépésköz szorzataként, azaz egy téglalappal közelíthetjük a területet. Másként fogalmazva a függvényt minden részintervallumon egy konstans értékkel helyettesítjük, lépcsőssé alakítjuk, ez az ún. Riemann-közelítés. Amennyiben minden részintervallum egyenlő hosszúságú, a számítás tovább egyszerűsödik, /2/ Riemann közelítése ekkor:

$$\hat{R}(k) = \Delta x \sum_{i=0}^{k-1} f(a + i\Delta x). \quad /3/$$

Amennyiben f integrálható, úgy $k \rightarrow \infty$ esetén $\hat{R}(k) \rightarrow \int_a^b f(x) dx$. Természetesen a Riemann-közelítés függ k értékétől. Amennyiben több különböző k -ra kívánjuk megkapni a közelítést, gyakori választás a $k = l$ után az $k' = 2l$ osztópont választása, hisz a második számítás esetén a rácspontok felén már ismerjük a függvényértéket.

A szakaszonkénti konstans érték helyett alkalmazható lineáris vagy magasabb fokú polinom függvénnyel való közelítés is, melyeket trapezoid és Simpson-féle eljárásnak nevezünk. Újfént éljünk azzal az egyszerűsítéssel, hogy egyenlő hosszúságú szakaszokra osztjuk az $[a, b]$ intervallumot, ekkor a trapezoid szabály szerinti becslés:

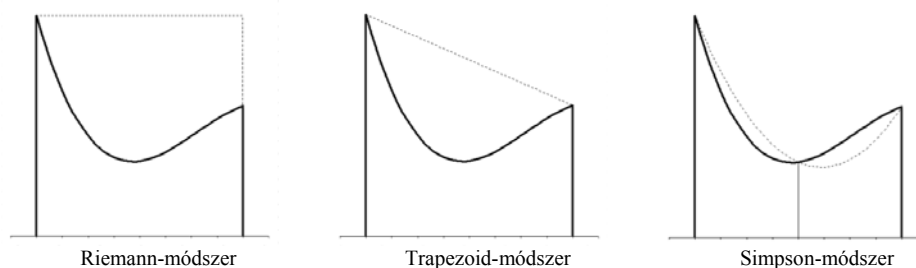
$$\hat{T}(k) = \frac{\Delta x}{2} f(a) + \frac{\Delta x}{2} f(b) + \Delta x \sum_{i=1}^{k-1} f(a + i\Delta x). \quad /4/$$

A Simpson-féle módszer a részintervallumok végpontjain kívül azok középpontjait is felhasználja, a három ponton átmenő másodfokú polinomot alkalmazva a függvény és a terület közelítésére. Az integrál Simpson-féle közelítése felírható:

$$\hat{S}(k) = \frac{\Delta x}{6} \left[f(a) + \sum_{i=1}^{k-1} 4f\left(a + \left(i - \frac{1}{2}\right)\Delta x\right) + 2f(a + i\Delta x) + f(b) \right]. \quad /5/$$

A három módszeren kívül természetesen léteznek egyéb, hasonló megoldások (például Romberg-módszer, Gauss-módszer), melyekre jelen cikkben nem térünk ki. Annak eldöntése, hogy milyen függvény esetén melyik módszer konvergál leggyorsabban, komoly tapasztalatot igényel, az integrálok közelítő meghatározása azonban rendkívül gyors. A 2. ábra a három módszer közelítését mutatja be egy kiválasztott részintervallum ábrázolásával. Az ábra alapján látható, hogy a módszerek pontossága nem azonos a bemutatott esetben, amikor inflexiós pont is van a tartományban, egyértelműen a Simpson-féle módszer tűnik a legpontosabbnak. Azokban az esetekben, amikor a függvény az adott szakaszon egyeneses közeli, az egyszerűbb módszerek konvergálnak gyorsabban.

2. ábra. Numerikus integrálási módszerek



A különböző módszerek konvergenciáját mutatja be a következő, 1. táblázat egy olyan függvény integrálja esetében, ahol ismerjük a tényleges keresett értéket.

1. táblázat

Konvergencia sebessége a három különböző integrálási módszernél

$\int_2^4 e^{-x} dx$	Riemann-	Trapezoid-	Simpson-
	módszer		
$k=10$	0,1291114	0,1174095	0,1143208
$k=100$	0,1181937	0,1170235	0,1167730
$k=1000$	0,1171367	0,1170197	0,1169952
$k=10\ 000$	0,1170313	0,1170196	0,1170172
Tényérték	$e^{-2} - e^{-4} = 0,11701964$		

A numerikus integrálás több nehézsége felmerül már egydimenziós esetben is, például a végtelen intervallumon vett integrál kiszámítása. Általánosságban ez az

eredeti függvény valamilyen transzformálásával kezelhető. Szintén meg kell említeni, hogy az implementálás egyszerűsítése miatt gyakran vesszük a részintervallumokat egyenlő hosszúságúnak, noha jóval pontosabb közelítést kapnánk már alacsony k értékek mellett, ha sűrűn helyeznénk el az osztópontokat azokon a helyeken, ahol a függvény gyorsan változik, és viszonylag ritkán azokon a helyeken, ahol a függvény megközelítőleg állandó.

Többdimenziós függvények esetén a bemutatott determinisztikus módszerek nehezen programozhatók, és a konvergencia egyre lassabb. A következő pontban szereplő, véletlen számok generálásán alapuló módszerek az egyszerűbb esetekben lassabban konvergálnak, azonban az összetettebb integrálok esetén is könnyen implementálhatók és konvergencia tulajdonságaikat megtartják.

3.2. Monte-Carlo-integrálás

A Monte-Carlo-integrálás egy véletlen szám generáláson alapuló statisztikai módszer, mely az 1940-es évek vége óta ismert, elsősorban *Neumann János* és *Stanislaw Ulam* munkásságának köszönhetően. A véletlen kísérletekből való következtetés gondolata – ahogy azt említettem – már sokkal korábban létezett, de az első számítógépek óriási lendületet adtak az alkalmazások elterjedésének.

Tudjuk, hogy ha X véletlen változó $g(x)$ sűrűségfüggvénnyel, akkor $h(X)$ transzformált véletlen változó várható értéke:

$$\mu = E[h(X)] = \int_{-\infty}^{\infty} h(x)g(x)dx. \quad /6/$$

Amennyiben rendelkezünk véletlen mintával X eloszlásából, úgy a függvényértékek átlaga /6/ torzítatlan becslését adja, $n \rightarrow \infty$ esetén

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n h(X_i) \rightarrow \mu, \quad /7/$$

ahol X_i az i -edik mintaelemet reprezentáló véletlen változót jelöli.

Legyen $X \sim Unif(a, b)$, így $E[h(X)] = \frac{1}{b-a} \int_a^b h(x)dx$, azaz az integrálást visszavezetjük egy várható érték meghatározásának problémájára.

A következő lépéseket kell elvégeznünk az $\int_a^b h(x) dx$ integrál Monte-Carlo közelítéséhez:

1. Legyen n független $X_i \sim Unif(a, b)$ véletlen változó.
2. Számítsuk ki az átlagos függvényértéket: $\overline{h(X)} = \frac{1}{n} \sum_{i=1}^n h(X_i)$.
3. A közelítő integrál érték: $\hat{\mu} = (b-a) \overline{h(X)}$.

Egyszerű Monte-Carlo-integrálást alkalmazhatunk az 1. táblázatban példaként felhasznált $\int_2^4 e^{-x} dx$ kiszámítására (Rizzo [2008]). Az integrál közelítését 10 000 véletlen számból számítva, egy lehetséges eredmény 0,1179417, ami ugyan csak a negyedik tizedes jegyben tér el a tényleges értéktől, azonban pontatlanabb a determinisztikus módszereknél. Vegyük észre, hogy a determinisztikus módszerekkel ellentétben most eredményeink a véletlen értékeknek köszönhetően futtatásról futtatásra kis mértékben eltérőek lehetnek!

Az MC-bebecslés varianciája a véletlen számok darabszámának növelésével csökkenthető, ami számításigényes. A konvergencia lassabb, mint a determinisztikus esetben (főként a trapezoid és Simpson-féle módszerhez képest), de magasabb dimenziókban is megmarad a konvergencia sebessége, míg a determinisztikus módszerek egyre lassabbá vagy alkalmazhatatlanná válnak.

Összetettebb problémák esetén jellemzően nem az MC-integrálás implementálása okoz nehézséget, hanem a lassú konvergencia. Fontos olyan módszerek alkalmazása, melyekkel a variancia csökkenthető, viszont a számítási időt egyáltalán nem, vagy alig növelik meg.

4. A variancia csökkentése Monte-Carlo-módszerek esetén

A fejezetben a hagyományos MC-integrálás varianciáját, valamint négy olyan módszert mutatok be, melyek nem a mintaelemszám növelésével csökkentik a MC-bebecslés varianciáját. Belátható, hogy a mintaátlagon alapuló Monte-Carlo-bebecslés torzítatlan és varianciája:

$$Var(\hat{\mu}) = \frac{(b-a)^2}{n^2} Var\left(\sum_i h(X_i)\right) = \frac{(b-a)^2}{n} Var(h(X)), \quad /8/$$

ahol a véletlen értékek függetlenségét használjuk ki. Jelen fejezetben épp ezt a függetlenséget sértjük meg oly módon, hogy a torzítatlanság továbbra is fennálljon, a variancia azonban csökkenjen. Az MC-bebecslés varianciája tehát az integrálási határoktól, a generált véletlen számok számától és a sűrűségfüggvény alakjától függ. A központi határeloszlás-tétel szerint pedig elégségesen nagy mintaelemszám mellett – ami gyakorlatilag mindig igaz – az integrálra (tehát a függvényértékek átlagára) vonatkozó

MC-bebecslések normális eloszlást követnek, azaz $\hat{\mu} \sim N\left(\mu, \frac{(b-a)^2}{n} \text{Var}(h(X))\right)$.

Az $\int_2^4 e^{-x} dx$ integrál érték MC közelítésének eloszlásához meg kell tehát határoznunk $\text{Var}(h(X))$ értékét. A várható értéket, az integrál tényleges értékét, valamint a többi paramétert ismerjük.

Tekintsük általánosan az $X \sim \text{Unif}(a, b)$ valószínűségi változót és határozzuk meg $h(X) = e^{-X}$ valószínűségi változó első és másodrendű momentumait! Alkalmazhatjuk /6/-ot, ahol tudjuk, hogy $g(x) = \frac{1}{b-a}$.

$$E(e^{-X}) = \int_a^b e^{-x} \frac{1}{b-a} dx = \frac{1}{b-a} [-e^{-x}]_a^b = \frac{e^{-a} - e^{-b}}{b-a},$$

valamint

$$E\left(\left(e^{-X}\right)^2\right) = E(e^{-2X}) = \int_a^b e^{-2x} \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{e^{-2x}}{2}\right]_a^b = \frac{e^{-2b} - e^{-2a}}{2(b-a)},$$

azaz a variancia:

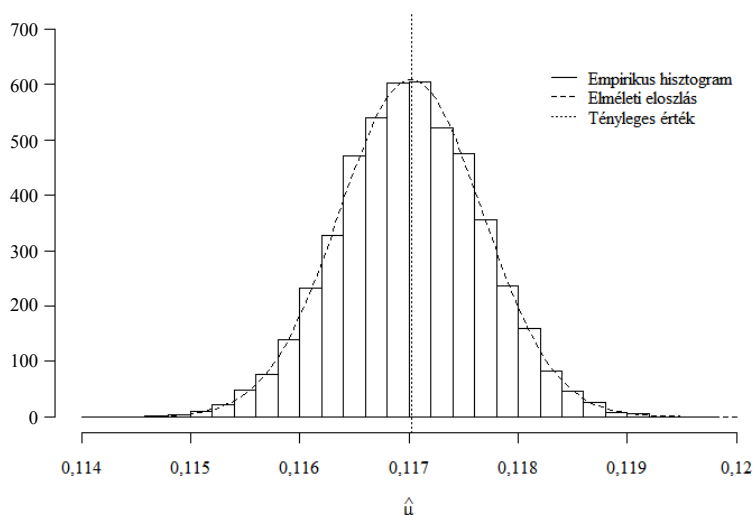
$$\text{Var}(e^{-X}) = E(e^{-2X}) - [E(e^{-X})]^2 = \frac{e^{-2b} - e^{-2a}}{2(b-a)} - \frac{(e^{-a} - e^{-b})^2}{(b-a)^2}. \quad /9/$$

A /9/ képletet alkalmazva $a=2$, $b=4$ esetre azt kapjuk, hogy $\text{Var}(e^{-X}) = \frac{e^2 - 1}{2e^8}$, azaz /8/ alapján a 10 000 elemű mintából álló becslés elméleti eloszlása:

$$\hat{\mu} \sim N\left(\mu, \frac{(b-a)^2}{n} \times \text{Var}(h(X))\right) = N\left(e^{-2} - e^{-4}, \frac{2e^2 - 2}{10\,000e^8}\right).$$

A 3. ábrán az elméleti és a 20 000-szer megismételt, egyenként 10 000 véletlen számot felhasználó MC-beclés eredményei láthatók. A 3. ábrára és a beclés varianciájára később, a varianciacsökkentő eljárások bemutatása során visszautalok, ugyanez az elméleti sűrűségfüggvény szerepel a 4. ábrán is.

3. ábra. Az integrál érték átlagoláson alapuló MC-beclésének elméleti és empirikus eloszlása, valamint a tényleges érték



Az eloszlás ismerete lehetőséget teremt arra, hogy beclésünk köré konfidenciaintervallumot építsünk a szokásos módon, ám most csupán benchmarkként fogjuk azt felhasználni.

Amint az ismert, θ_1 beclőfüggvény hatásosabb θ_2 -nél, ha $\frac{\text{Var}(\theta_1)}{\text{Var}(\theta_2)} < 1$ és mindkét beclőfüggvény torzítatlan módon becsüli θ -t. Ebben az esetben θ_1 használata θ_2 helyett a varianciában

$$\frac{\text{Var}(\theta_1) - \text{Var}(\theta_2)}{\text{Var}(\theta_1)} \times 100 \quad /10/$$

százalékos csökkenést eredményez.

A variancia csökkentésére a következőkben négy fontos eljárást mutatok be, az ellentétes (antitetikus) változók (antithetic variables), az ellenőrző változók (control variables), a fontossági mintavételezés (importance sampling) és a rétegző mintavételezés (stratified sampling) módszereit. Ezek alapötlete nem ismeretlen, az ellenőrző változók módszere regressziós becslés, a fontossági mintavételezés a hányadosbecslés, a rétegző mintavételezés a rétegzett mintavétel (*Galambosné* [2011]) logikáját alkalmazza a variancia csökkentésére.

4.1. Antitetikus változók módszere

Tekintsük két azonos eloszlású, X_1 és X_2 valószínűségi változó átlagát. Az átlag varianciája:

$$\text{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}(\text{Var}(X_1) + \text{Var}(X_2) + 2\text{Cov}(X_1, X_2)). \quad /11/$$

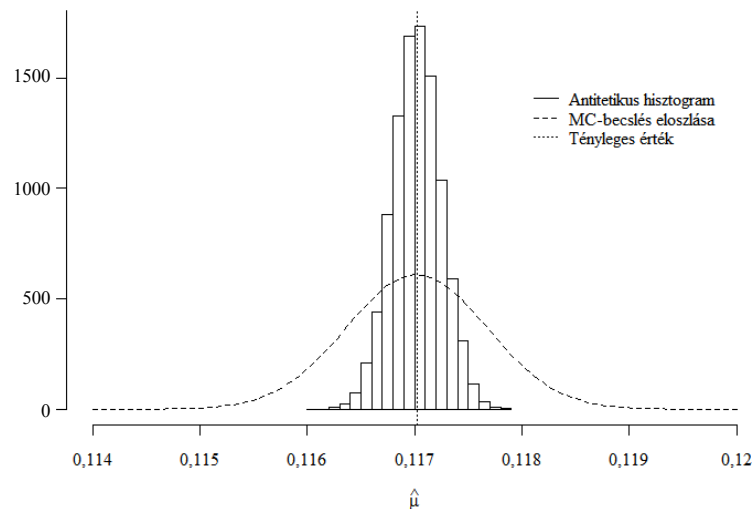
Amennyiben X_1 és X_2 függetlenek, úgy a kovarianciatag /11/-ben 0. Ha tehát olyan változókat használunk, ahol a kovariancia negatív, az átlag varianciája csökkenthető a független esethez viszonyítva. Ez az alapvető ötlet húzódik az antitetikus változók módszere mögött. A Monte-Carlo-szimulációk esetén a keresett integrál becslése a $[0,1]$ egyenletes véletlen változók valamilyen függvénye: $X_1 = h(U_1, U_2, \dots, U_n)$. Tekintsük $X_2 = h(1 - U_1, 1 - U_2, \dots, 1 - U_n)$ antitetikus becslést, ekkor a két valószínűségi változó eloszlása megegyezik. Páronként a véletlen változók közötti kovariancia negatív, értéke $-\frac{1}{12}$. Ekkor bizonyítható (*Rizzo* [2008] 129. old.), hogy bármely monoton h függvényre

$$\text{Cov}(h(U_1, U_2, \dots, U_n), h(1 - U_1, 1 - U_2, \dots, 1 - U_n)) \leq 0.$$

A módszer gyakorlati alkalmazása egyszerű. Generáljunk $n/2$ mintaelemet a szükséges egyenletes eloszlásból, majd ezekből további $n/2$ ellentétes változót. A negatív kovariancia miatt az így előállított becslés varianciája alacsonyabb lesz, mint a hagyományos, n darab véletlen számból álló MC-becslésé. Korábban az $\int_2^4 e^{-x} dx$ integrál becsléséhez 10 000 véletlen számot használtunk fel amelyek,

$U_1 \sim Unif(2,4)$ eloszlásúak voltak. Amennyiben 5 000 darab véletlen számot generálunk U_1 -ből, majd az $U_2 = 6 - U_1 \sim Unif(2,4)$ véletlen értékeket párosítjuk hozzájuk, a becslés variációjára alacsonyabb lesz. A 20 000 alkalommal elvégzett becslés eredményeinek empirikus eloszlását az 4. ábra mutatja be hisztogramon, feltüntetve a 3. ábrán bemutatott (eredeti MC-) becslés elméleti eloszlását is. Az összehasonlíthatóság érdekében a korábbi ábra vízszintes tengelyét (0,114–0,120) megtartottam.

4. ábra. Az antitetikus változók használatának hatása az MC-becslés eloszlására



A /10/ arány alapján könnyedén meghatározhatjuk a módszer hatékonyságjavulását, ami 88 százalék körüli értéket mutat, azaz antitetikus változók használatával a becslésünk variációját nagymértékben sikerült csökkenteni, úgy, hogy a számítási költség, a generált véletlen értékek száma nem lett nagyobb. A szóródás csökkenése a 4. ábra alapján is érzékelhető. Hasonló ábrát rajzolhatnánk a további bemutatandó módszerek esetén is, ettől azonban eltekintünk, csupán a variancia csökkenésének arányát számítjuk ki.

4.2. Kontrollváltozók módszere

Tegyük fel, hogy a célunk $\int_a^b h(x) dx$ becslése, és van egy olyan l függvénnyel leírható (kontroll-) változó, mely $\eta = E[l(X)]$ várható értékét ismerjük, és a két vál-

tozó korrelál. A két függvényből konstruálható olyan becslőfüggvény, mely torzítatlan³ bármely c konstans esetén:

$$\hat{\mu}_{cont} = h(X) + c(l(X) - \eta). \quad /12/$$

Ekkor /12/ varianciája felírható, és célunk c függvényében ennek legkisebb értékét megtalálni:

$$Var(\hat{\mu}_{cont}) = c^2 Var[l(X)] + 2c Cov(h(X), l(X)) + Var[h(X)]. \quad /13/$$

A /13/ összefüggés c -ben másodfokú és konkáv, így minimális értékét a

$$c^* = -\frac{Cov(h(X), l(X))}{Var[l(X)]} \quad /14/$$

helyen veszi fel, ahol a variancia értéke

$$Var(\hat{\mu}_{cont}(c^*)) = Var[h(X)] - \frac{[Cov(h(X), l(X))]^2}{Var[l(X)]}. \quad /15/$$

Láthatjuk, hogy a kivonandó taggal csökken a variancia. Ezt tudva kiszámíthatjuk /10/ alapján a variancia százalékos csökkenését:

$$\frac{Var(\theta_1) - Var(\theta_2)}{Var(\theta_1)} = \frac{[Cov(h(X), l(X))]^2}{Var[h(X)] Var[l(X)]} = [Corr(h(X), l(X))]^2. \quad /16/$$

A /16/ képletből egyértelműen látszik, hogy olyan $l(\cdot)$ függvényre van szükségünk, hogy $l(X)$ erősen korrelál $h(X)$ -szel. Amennyiben a valószínűségi változók között nincs korreláció, a módszer nem használható, más kontrollváltozót kell keresnünk. A feladat tehát a helyes változó megtalálása, majd az optimális c^* kiszámítása, melyhez /14/ szerint a varianciára és a kovarianciára van szükségünk. Amennyiben ezek az értékek analitikusan nem meghatározhatók, szimuláció segítségével találhatjuk meg a megfelelő értékeket.

³ $E(\hat{\mu}_{cont}) = E[h(X) + c(l(X) - \eta)] = \mu + c \times 0 = \mu.$

Tekintsük újra az $\int_2^4 e^{-x} dx$ integrált. Keressük azt az $l(X)$ függvényt, amely momentumait könnyen meg tudjuk határozni és erős korrelációt mutat $h(X)$ -szel. A legegyszerűbb választás egy lineáris függvény: $l(X) = \frac{X-2}{2}$. Ekkor $Cov(h(X), l(X)) = -\frac{e^{-4}}{2}$, $Var[l(X)] = \frac{1}{12}$, azaz /14/ alapján $c^* = 6e^{-4}$. A variancia értéke itt /15/ és a már korábban levezetett $h(X)$ -re vonatkozó momentumok segítségével:

$$Var(h(X) + c(l(X) - \eta)) = \frac{e^2 - 1}{2e^8} - 3e^{-8} \approx 0,0000653,$$

ami azt jelenti, hogy a variancia csökkenése /16/ alapján közel 94 százalékos:

$$\frac{Var(\theta_1) - Var(\theta_2)}{Var(\theta_1)} = \frac{6}{e^2 - 1} \approx 0,939.$$

Célunk közvetlenül nem csupán $\mu = E[h(X)]$, hanem az integrál közelítése volt, de a variancia elért csökkenése természetesen az integrál becslésében is hasonlóan jelentkezik. A kontrollváltozó használatával, azaz $h(X) = e^{-X}$ helyett, a szintén torzítatlan $h(X) + c(l(X) - \eta) = e^{-X} + 6e^{-4} \left(\frac{X-2}{2} - \frac{1}{2} \right)$ becslőfüggvényt alkalmazva az integrál becslésének varianciáját jelentősen sikerült csökkenteni.

Az antitetikus változók módszere a kontrollváltozó módszer speciális esete, ahol mindkét becslőfüggvény független azonos eloszlású, és a változópárok közötti korreláció -1 , ekkor $c^* = \frac{1}{2}$ optimális érték adódik. Annak ellenére, hogy az antitetikus változók módszere speciális esetként is felfogható, az irodalomban a két módszert külön tárgyalják.

Gyakran alkalmazott technika több kontrollváltozó felhasználása, hiszen

$$\hat{\mu} = h(X) + \sum_j c_j^* (l_j(X) - \mu_j)$$

szintén torzítatlan becslést ad. Az optimális $\mathbf{c}^* = (c_j^*)$ vektort a l_j és az h függvények közötti maximális korrelációval érhetjük el. Gyakran alkalmazott módszer,

hogy az optimális \mathbf{c}^* meghatározásához egyszerű lineáris regressziót illesztünk, amiből a kontrollváltozós módszer legfontosabb jellemzőit azonnal megkapjuk. A varianciában bekövetkező csökkenés pontosan a lineáris regresszió R^2 értékével egyezik meg, a regressziós paraméterekből pedig \mathbf{c}^* adódik.

Eddigiekben egyenletes eloszlású valószínűségi változó felhasználásával közelítettünk függvények várható értékén keresztül integrálokat. A következő alfejezetben az egyenletes véletlen számoknál hatékonyabb módszert ismerünk meg.

4.3. Fontossági mintavétel

A bemutatott klasszikus MC-módszer és a variancia csökkentésére irányuló eljárások hátránya, hogy nem használhatók közvetlenül olyan esetekben, amikor valamely integrálási határ nem véges, ráadásul egyenletes véletlen számok alkalmazása nem hatékony, ha $h(\cdot)$ nagyon távol esik az egyenletestől. Mivel az integrálást visszavezettük egy átlagolási problémára, általánosíthatjuk megközelítésünket a súlyozatlan átlag helyett súlyozott átlag (azaz az egyenletestől eltérő sűrűségek) alkalmazásával. Az általános módszer neve fontossági mintavétel (importance sampling).

Tekintsük az X véletlen változót g sűrűségfüggvénnyel, ahol bármely x esetén, melyre $h(x) > 0$, szükségképpen $g(x) > 0$. Legyen továbbá $\tilde{h}(x) = h(x)$, ha $a \leq x \leq b$, ezen kívül $\tilde{h}(x) = 0$, valamint Y véletlen változó $\frac{h(X)}{g(X)}$. Ekkor

$$\int_a^b h(x) dx = \int_a^b \frac{h(x)}{g(x)} g(x) dx = \int_{-\infty}^{\infty} \frac{\tilde{h}(x)}{g(x)} g(x) dx = E\left(\frac{\tilde{h}(X)}{g(X)}\right) = E(Y). \quad 17/$$

Közelítsük $E(Y)$ értékét hagyományos Monte-Carlo-integrálással, azaz számítsuk ki az

$$\frac{1}{n} \sum_{i=1}^n Y_i = \frac{1}{n} \sum_{i=1}^n \frac{\tilde{h}(X_i)}{g(X_i)} \quad 18/$$

átlagot, ahol az X_i -k a $g(x)$ sűrűségfüggvényből származó véletlen értékek. A $g(x)$ függvény neve fontossági függvény (importance function).

A közelítés varianciáját n és $Var(Y)$ határozza meg, ezért a gyakorlatban célszerű, hogy a fontossági függvény $h(x)$ -hez hasonló, a hányados közelítőleg kons-

tans legyen. Hasonlóan fontos szempont, hogy $g(x)$ alapján X könnyen szimulálható legyen.

Bizonyítható (Rizzo [2008] 143. old.), hogy a variancia minimalizálása a

$$f^*(x) = \frac{|h(x)|}{\int_A |h(x)| dx}$$

fontossági függvény alkalmazásával érhető el, ahol $A \in \mathbb{R}$ az a halmaz, ahol integrálni kívánunk. Mivel valószínűtlen, hogy ez a kifejezés rendelkezésre áll, gyakorlati probléma esetén leggyakrabban olyan függvényt választunk, amely elégségesen közel van $|h(x)|$ -hez A -n.

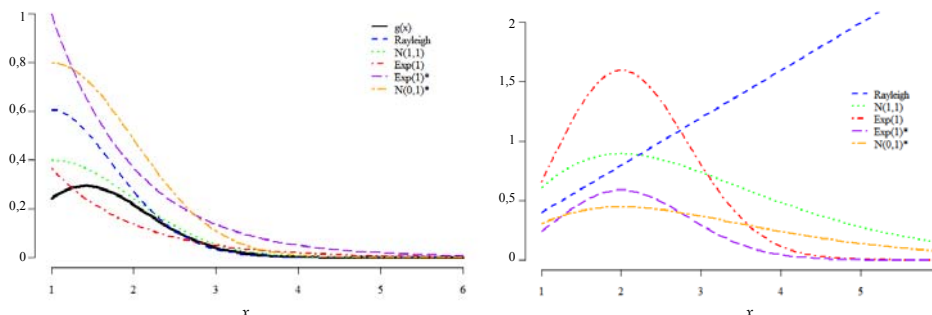
A fontossági mintavételt a korábbiaktól eltérő $h(x) = \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}}$ függvény $(1, \infty)$ intervallumon vett integrálja segítségével mutatom be, még hozzá öt fontossági függvény felhasználásával, azok hatékonyságát összehasonlítva.

1. Standard Rayleigh-eloszlás: $g_1(x) = x e^{-\frac{x^2}{2}} \quad x \geq 0$.
2. Normális eloszlás $N(1,1)$: $g_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}$.
3. Exponenciális eloszlás $Exp(1)$: $g_3(x) = e^{-x}, x \geq 0$.
4. Módosított exponenciális eloszlás $Exp(1)^*$: $g_4(x) = e^{-x+1}, x \geq 1$.
5. Módosított standard normális eloszlás: $N(0,1)^*$: $g_5(x) = 2 \times \phi(x-1), x \geq 1$.

A választott eloszlások egy része nem kizárólag az $x \geq 1$ tartományon értelmezett. A módosított sűrűségfüggvényeket olyan módon alakítottam ki, hogy belőlük könnyű legyen véletlen értékeket nyerni és hasonlítsanak a céleloszlásra (mind akara, mind értelmezési tartományra). A Rayleigh-eloszlásból az inverz eloszlásfüggvény módszerrel, a többi eloszlásból pedig az R beépített függvényei segítségével vettem mintát. Az 5. ábra $h(x)$ integrálandó szakaszán mutatja be a függvényeket

(bal oldal), valamint az $\frac{h(x)}{g_j(x)}$ hányadosokat (jobb oldal).

5. ábra. A fontossági mintavétel sűrűségfüggvényei és a függvények hányadosai



Az 5. ábra alapján a módosított standard normális eloszlás tűnik a legjobb választásnak. Ez a fontossági függvény a módosításnak köszönhetően csak az $x \geq 1$ helyeken értelmezett, míg például g_2 a teljes x tengelyen, g_1 és g_3 pedig a pozitív féltengelyen. Mindez azt jelenti, hogy az g_1, g_2, g_3 függvények esetén az $\frac{\tilde{h}}{g_j}$ hányados sok esetben zérus, az eljárás nem hatékony.

A mintavételek 2000 alkalommal történt elvégzése után az eredményeket a 2. táblázatban foglaltam össze ($n = 10\,000$).

2. táblázat

Az integrál közelítésének eredményei öt fontossági függvénnyel

Becslés jellemzői	1.	2.	3.	4.	5.
	fontossági függvény				
Az integrál becslt értékeinek átlaga	0,40076	0,40059	0,40026	0,40065	0,40063
Az integrál becslt értékeinek szórása	0,00357	0,00412	0,00584	0,00155	0,00044
Nullák átlagos aránya (százalék)	39,3	50,0	63,25	0,00	0,00

Az öt fontossági függvény közül a módosított normális eloszlással készült közelítés rendelkezik a legkisebb varianciával. Az első három jelölt esetén az integrálás határaitól eltérő értelmezési tartomány miatt a generált véletlen értékek döntő többsége nem hasznosul (a táblázatban nullák aránya), mivel az $\frac{\tilde{h}(x)}{g(x)}$ hányados zérus értékű.

Az $Exp(1)$ eloszlás sűrűségének 63,2 százaléka esik a $[0,1]$ intervallumba, az $N(1,1)$ eloszlás esetén 50 százalék, a függvényforma tekintetében hasonló Rayleigh-eloszlásnál pedig a $[0,1]$ közötti integrál alapján $1 - e^{-\frac{1}{2}} \approx 0,3935$ a felesleges húzások aránya. Ez utóbbi választás további hátránya, hogy az $\frac{h(x)}{g(x)}$ egy pozitív meredekségű lineáris egyenes, azaz nem stabil. A transzformált eloszlások jobban teljesítettek ebben az esetben, általánosságban hátrányuk, hogy a véletlenszám-generálás nem minden esetben triviális.

4.4. Rétegző mintavétel

A rétegző mintavétel⁴ a variancia csökkenését úgy éri el, hogy az integrálandó területet rétegekre bontja, és ezeken a rétegeken belül kis varianciával próbál becsülni. A k darab rétegben rögzített számú véletlen értéket húzunk, úgy, hogy $n = n_1 + n_2 + \dots + n_k$, azzal a céllal, hogy

$$Var(\hat{\mu}_1(n_1) + \hat{\mu}_2(n_2) + \dots + \hat{\mu}_k(n_k)) \ll Var(\hat{\mu}(n)),$$

ahol a bal oldalon a rétegző mintavételt alkalmazó, a jobb oldalon pedig a standard MC-becselőfüggvény látható.

A variancia csökkentése akkor hatékony, ha a rétegekben az integrálandó függvény átlaga jelentősen eltérő, azaz sikerül heterogén rétegeket kialakítani. Amennyiben az integrálandó függvény monoton, ezt könnyű teljesíteni. Jól érzékelhető a hagyományos rétegzett mintavétellel való analógia abból a tényből adódóan is, hogy a rétegző mintavétel mindig kisebb varianciát szolgáltat, kivéve abban az esetben, ha a rétegek átlagai megegyeznek.

A már ismerős $\int_2^4 e^{-x} dx$ példán mindössze két egyenlő hosszúságú réteg alkalmazá-

sával a becslés varianciája kevesebb mint harmadára, négy réteg esetén kevesebb mint 10 százaléka esik azonos mintaelemszám mellett. Ennek eléréséhez csupán annyi a feladatunk, hogy $Unif(2,4)$ véletlen számok helyett például $Unif(2,3)$ és $Unif(3,4)$ véletlen értékek segítségével becsüljük a megfelelő területeket, majd becsléseinket összegezzük.

⁴ A rétegző mintavétel angol terminológiával stratified sampling, azaz az elnevezés megegyezik a rétegzett mintavétel elnevezéssel. A könnyebb megkülönböztetőség érdekében nevezem rétegző mintavételnek az eljárást.

A röviden bemutatott rétegző mintavétel előnye, hogy tetszőlegesen kombinálható a többi varianciacsökkentő eljárással, így a szakirodalom ismeri és használja a rétegző fontossági mintavétel (stratified importance sampling) fogalmát is, ahol a különböző rétegekben akár különböző fontossági függvények is alkalmazhatók.

5. Összefoglalás

A tanulmány a különböző eloszlásokból való véletlenérték-generálás, valamint a Monte-Carlo-integrálás legfontosabb módszereit tekinti át. A számítógép által szolgáltatott pszeudo véletlen számok és az inverz eloszlásfüggvény, direkt transzformációs, valamint elfogadás-elutasítás módszerek segítségével alacsony dimenziószám esetén tudunk sűrűségfüggvényekből szimulálni. Egy másik fontos terület az MC-integrálás, ami az integrálást egy várható érték becslés problémájára vezeti vissza. A becslési variancia csökkentésére szolgáló eljárások bemutatása zárja a cikket. A magasabb dimenziószám esetén alkalmazott, Markov-lánc felállításán alapuló ún. Markov-lánc Monte-Carlo- (MCMC-) módszereket egy későbbi tanulmányomban mutatom be.

Irodalom

- ALBERT, J. H. [2009]: *Bayesian Computation with R*. Springer. Heidelberg, London, New York.
- ANGUS, J. E. [1994]: The Probability Integral Transform and Related Results. *SIAM Review*. Vol. 36. No. 4. pp. 652–654.
- BOX, G. E. P. – MÜLLER, M. E. [1958]: A Note on the Generation of Random Normal Deviates. *The Annals of Mathematical Statistics*. Vol. 29. No. 2. pp. 610–611.
- CASELLA, G. – BERGER, R. L. [2002]: *Statistical Inference*. Duxbury Press. Belmont.
- GALAMBOSNÉ TISZBERGER M. [2011]: A rétegzett mintavételről. *Statisztikai Szemle*. 89. évf. 9. sz. 909–929. old.
- HUNYADI L. [2011]: Bayesi gondolkodás a statisztikában. *Statisztikai Szemle*. 89. évf. 10–11. sz. 1150–1171. old.
- HUNYADI L. – VITA L. [2008]: *Statisztika I-II*. Aula Kiadó. Budapest.
- KEHL D. [2012]: Robert, C. – Casella, G.: Szemelvények a Markov-lánc Monte-Carlo-módszerek történetéből. *Statisztikai Szemle*. 90. évf. 4. sz. 352–354. old.
- KOTZ, S. – READ, C. B. – BALAKRISHNAN, N. – VIDA KOVIC B. (eds.) [2006]: *Encyclopedia of Statistical Sciences*. John Wiley & Sons Wiley. New York.
- MATSUMOTO, M. – NISHIMURA, T. [1998]: Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. *ACM Transactions on Modeling and Computer Simulation*. Vol. 8. No. 1. pp. 3–30.

- PINTÉR J. – RAPPAI G. (szerk.) [2007]: *Statisztika*. Pécsi Tudományegyetem. Pécs.
- R DEVELOPMENT CORE TEAM [2011]: *R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing*. Vienna. <http://www.R-project.org/>
- RIZZO, M. L. [2008]: *Statistical Computing with R*. Chapman & Hall/CRC. Boca Raton.
- ROBERT, C. P. – CASELLA, G. [2004]: *Monte Carlo Statistical Methods (2nd edition)*. Springer. New York.
- ROBERT, C. P. – CASELLA, G. [2010]: *Introducing Monte Carlo Methods with R*. Springer. New York.
- ROBERT, C. P. – CASELLA, G. [2011]: A Short History of Markov Chain Monte Carlo: Subjective Recollections from Incomplete Data. *Statistical Science*. Vol. 26. No. 1. pp. 102–115.
- STUDENT [1908]: The Probable Error of a Mean. *Biometrika*. Vol. 6. No. 1. pp. 1–25.
- VÁRPALOTAI, V. [2008]: *Modern Bayes-i ökonometriai elemzések – Simasági priorok alkalmazása az üzleti ciklusok szinkronizációjának mérésére és az infláció előrejelzése*. PhD-értekezés. Budapesti Corvinus Egyetem. Budapest.

Summary

This paper gives an overview on the statistical applications of generating random variables, as well as on numerical and Monte Carlo integrations. It also describes how methods of variance reduction can be applied. These ideas are inevitably necessary when dealing with Bayesian statistics, especially in cases where the information contained in posterior distributions can only be obtained by simulations. The common feature of the methods herein is that each of them can be effectively applied in low dimensional cases. Another means to determine the features of multivariate and complex density functions is to apply Markov chains and Markov chain Monte Carlo (MCMC) methods, which will be described in a later study.