UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
Matematika - uporabna smer

Andreja Kmet

MODELIRANJE KREDITNIH TVEGANJ

Diplomsko delo

Ljubljana, 2005

Na tej strani bi se rada zahvalila vsem, ki so mi na kakršenkoli način pomagali pri diplomi ali pri študiju nasploh.
Hvala mojemu mentorju, profesorju Tomažu Koširju, ki se je odločil za mentorstvo, čeprav nisem izbrala čisto "njegovega"področja. Hvala, ker si je vedno vzel čas in se mi je pomagal pregristi čez težje izpeljave in dokaze. Hvala, ker me je pomiril, ko sem bila preveč zaskrbljena, in ker me je blago priganjal, ko me je minevala vnema. Hvala tudi za vse nasvete v zvezi z nadaljnjim študijem in poklicno kariero.
Hvala Genotu Zavodniku, najbolj zabavnemu ekonomistu, kar jih poznam. Ker mi je predlagal kup zanimivih tem za diplomsko delo in me zalagal z literaturo. In ker mi je z veseljem razložil vse meni nove pojme iz sveta bančništva. Hvala, ker mi je ponudil sodelovanje pri zelo zanimivem projektu, saj sem tako dobila možnost videti, kaj pomeni ravnanje s kreditnimi tveganji tudi v praksi.
Hvala mojemu očetu Andreju, ki me je že zgodaj navdušil za matematiko. In ker je vedno vedel odgovore na vsa vprašanja, tudi ob najbolj nemogočih urah. Hvala moji mami Dragi, ki je mirno prenašala vso to matematiko okrog sebe in me zraven še spodbujala. Staršema hvala tudi za finančno podporo ves čas študija.
Hvala mojemu fantu Primožu za potrpljenje in vzpodbudo, ko sem to najbolj potrebovala. Hvala, ker se je skupaj z mano marsičemu odpovedal. In hvala za nepogrešljivo pomoč pri vseh tehničnih težavah, ki so nastopile pri nastajanju diplome.
Hvala prijateljicama Maji in Ireni za pomoč pri nekaterih prevodih.
Hvala prijateljicam Heleni, Ireni, Lidiji in Urši, ker so razumele, da imam manj časa za druženje z njimi.
Hvala tudi vsem ostalim prijateljem in znancem, ki so me dovolj pogosto spraševali, kdaj bom končala...

Kazalo

Program diplomskega dela

V delu predstavite modeliranje kreditnih tveganj. Osredotočite se na model CreditRisk+.

V Ljubljani, 8.4.2004 Tomaž Košir

Povzetek

Kreditno tveganje predstavlja največji delež med vsemi tveganji, s katerimi se srečuje banka. Zato tehnike za obvladovanje tega tveganja postajajo čedalje pomembnejše. Banko v prvi vrsti zanima potencialna izguba in ekonomski kapital, ki mora zagotavljati ustrezno donosnost glede na prevzeta tveganja.
V diplomskem delu spoznamo osnove modeliranja kreditnih tveganj. Osnovni problem, s katerim se srečamo, je poiskati porazdelitev izgube kreditnega portfelja na podlagi poznavanja verjetnosti neplačil posameznih kreditojemalcev. Na začetku privzamemo, da ima verjetnost neplačila dolga fiksno vrednost, kasneje pa analiziramo primer, ko se verjetnost neplačila dolga s časom spreminja. Na ta način vpeljemo v model korelacije in lahko govorimo o t.i. sistemskem tveganju. V zadnjem poglavju predstavimo model CreditRisk+, ki na podlagi danega kreditnega portfelja preko rodovne funkcije poišče porazdelitev portfeljske izgube.
Diplomskemu delu je priložena zgoščenka s praktičnim prikazom delovanja modela CreditRisk+.

Math. Subj. Class. (2000): 62P20, 91B30, 91B66

Ključne besede: kreditno tveganje, porazdelitev portfeljske izgube, ekonomski kapital, mera tveganja, prispevek k tveganju kreditnega portfelja

Abstract

Credit risk is the largest of all risks met by a bank. For this reason the techniques for controlling credit risk are becoming more and more important. A bank is foremost interested in the potential loss and economic capital which has to assure appropriate profitability according to the risk taken.
In this paper the basis for credit risk modeling is presented. Our primary goal is to find the credit portfolio loss distribution on the basis of obligors' probability of default. First we assume fixed default rates; later on we analyze default losses with variable default rates. In this manner correlations are introduced in the model and we can speak of so called systematic risk. In the last chapter the model CreditRisk+ is presented which finds the loss distribution using probability generating function. This paper is complemented by a CD with a practical illustration of the model CreditRisk+ functioning.

Keywords: credit risk, portfolio loss distribution, economic capital, risk measure, risk contribution

Uvod

Z vidika lastnikov in nadzornikov banke je kapital najpomembnejša ekonomska dobrina, ki mora zagotavljati solventnost, to je sposobnost banke, da ob morebitni likvidaciji lahko izplača upnike in lastnike, in primerno donosnost glede na prevzeta tveganja. Do nedavnega je večina držav sledila baselskemu kapitalskemu sporazumu iz leta 1988, po katerem mora biti lastniški kapital bank enak najmanj 8% uteženih tveganih naložb.
Obstoječi kapitalski predpisi pa ne ustrezajo več praksi bančnega poslovanja, saj so se v zadnjih letih na finančnih trgih pojavila številna nova oz. spremenjena tveganja. Baselski komite [1] je tako junija 1999 izdal predlog nove sheme za merjenje kapitalske ustreznosti, t.i. Basel II, ki bo nadomestil sporazum iz leta 1988. Nova shema bo ohranila idejo starega sporazuma, to je varnost in stabilnost finančnega sistema, hkrati pa bo odpravila njegove pomanjkljivosti. Stara metodologija namreč za izračun uteženih tveganih naložb bank uporablja sistem tehtanja, ki ne omogoča zadostnega razlikovanja med različnimi kreditojemalci banke in je zato pogosto slab približek ekonomskega tveganja, ki so mu izpostavljene banke. Novi kapitalski sporazum se poglobljeno ukvarja s kreditnim tveganjem in s tem v zvezi razvija nove tehnike za ravnanje s kreditnimi tveganji.
Tudi v Sloveniji že nekaj let tečejo priprave na uveljavitev novega kapitalskega sporazuma. Do leta 2006, ko bo stopil v veljavo povsod po svetu, se morajo slovenske banke in Banka Slovenije kot njihov pristojni nadzornik dobro pripraviti na vse spremembe.

Kot rečeno, novi kapitalski sporazum Basel II daje glavni poudarek kreditnim tveganjem. Od vseh tveganj, s katerimi se banke srečujejo, namreč ravno kreditno tveganje predstavlja največji delež - po velikosti in pomembnosti. Glavni vir izgube pri tem tveganju je nesposobnost kreditojemalcev, da bi poravnali svoje obveznosti do banke, ki se zrcali v različnih dogodkih: zamujanjem z odplačili, prisilno poravnavo, stečajem... Plačilna nesposobnost kreditojemalcev lahko zaradi učinkov sistemskega tveganja, ki se pojavi kot

[1] Banka za mednarodne poravnave, Basel, Švica

posledica korelacij med kreditojemalci, privede tudi do nesolventnosti banke. Obvladovati kreditno tveganje za banko med drugim pomeni znati napovedati potencialno izgubo, predvideti potrebno količino ekonomskega kapitala za kritje nepričakovane izgube in izračunati pravo verjetnost neplačila dolga posameznega kreditojemalca. Ravnanje s kreditnimi tveganji je v zadnjem času hitro razvijajoče se področje in je kot tako zanimivo s stališča ekonomije in matematike.

Matematično orodje, ki ga uporabljamo pri modeliranju kreditnih tveganj, najdemo predvsem v verjetnosti in statistiki. Zato v 1. poglavju ponovimo osnovna sredstva, ki jih bomo potrebovali za razumevanje v nadaljevanju.
V 2. poglavju razložimo osnovni problem ravnanja s kreditnimi tveganji. Dan je kreditni portfelj, zanima pa nas potencialna izguba, ki nastane ob kreditiranju. Iščemo torej porazdelitev slučajne spremenljivke, ki pomeni portfeljsko izgubo. Spoznamo naslednje pojme: pričakovana in nepričakovana izguba, verjetnost neplačila dolga, izpostava ob nastanku plačilne nesposobnosti kreditojemalca, izguba ob nastopu izostanka odplačila dolga, ekonomski kapital.
V 3. poglavju se lotimo modeliranja koreliranih izostankov odplačil dolga. Ugotovimo, da se slučajna spremenljivka, ki predstavlja portfeljsko izgubo, največkrat obnaša kot Bernoullijeva ali Poissonova slučajna spremenljivka. Najprej preučimo enostaven primer, ko je verjetnost neplačila dolga fiksna, potem pa še bolj realen primer, ko privzamemo, da je verjetnost neplačila dolga slučajna spremenljivka. Na ta način vpeljemo v model korelacije. Nazadnje še primerjamo Bernoullijev in Poissonov model.
V 4. poglavju definiramo dve glavni meri tveganja, tvegano vrednost VaR in pričakovan izpad, ter pojem koherentnosti, ki velja za lepo lastnost mere tveganja. Dokažemo, da tvegana vrednost ni koherentna, pričakovan izpad pa je. S pomočjo omenjenih mer tveganja izračunamo ekonomski kapital, to je zahtevani kapital za kritje nepričakovanih izgub. Na koncu še definiramo prispevek h kapitalu in pokažemo, kako se izračuna prispevek posameznih kreditojemalcev k tveganju kreditnega portfelja.
5. poglavje, kjer spoznamo model CreditRisk+, predstavlja glavni del tega diplomskega dela, saj s pomočjo znanja, ki smo si ga pridobili do sem, pokažemo, kako se v praksi modelira kreditna tveganja. Začetki modela CreditRisk+ segajo v zgodnja devetdeseta leta, ko je banka Credit Suisse First Boston (CSFB) razvijala nove metode za ravnanje s tveganji. Leta 1993 je skupina Credit Suisse Group vzporedno delala na projektu za razvoj ravnanja s kreditnimi tveganji, kjer je uporabila izkušnje CSFB, v decembru 1996 pa je Credit Suisse Group predstavila model CreditRisk+.
Problem, s katerim se ukvarja model, smo omenili že v drugem poglavju:

dan je kreditni portfelj, v katerem poznamo kredite in verjetnosti neplačil dolga posameznih kreditojemalcev. Definiramo slučajno spremenljivko, ki pomeni portfeljsko izgubo, in iščemo njeno porazdelitev. CreditRisk+ s pomočjo rekurzivne zveze izračuna rodovno funkcijo izgube, verjetnosti izgub posameznih zneskov pa nato prebere iz razvoja funkcije v vrsto. V začetnih primerih privzamemo, da je verjetnost neplačila dolga posameznega kreditojemalca fiksno število. Nato se lotimo sektorske analize: ugotovimo, da na spremenljivost verjetnosti neplačila dolga vplivajo različni dejavniki. Kreditojemalce tako razdelimo v več različnih sektorjev, kjer vsak sektor pomeni množico kreditojemalcev pod skupnim vplivom nekega dejavnika, ki določa verjetnost neplačila dolga. Le-ta tako postane slučajna spremenljivka. Nazadnje predstavimo še posplošeno sektorsko analizo, ki dovoljuje, da na danega kreditojemalca vpliva več dejavnikov. Na ta način vpeljemo v model korelacije.
Žal pa CreditRisk+ ni numerično stabilen, saj v rekurzivni zvezi za izračun rodovne funkcije sešteva približno enako velika, toda nasprotno predznačena števila, kar lahko vodi do sorazmeroma velike relativne napake. Zato predstavimo še alternativo algoritmu, uporabljenemu v modelu CreditRisk+, kjer poskrbimo za pozitiven predznak vseh koeficientov, ki se pojavijo v postopku računanja.
CreditRisk+ poleg porazdelitve izgube izračuna še izgubo na danem percentilu in prispevek posameznih kreditojemalcev k tveganju kreditnega portfelja.
Diplomskemu delu je priložena zgoščenka s praktičnim prikazom delovanja modela CreditRisk+. Predstavljeni so vsi trije primeri: izguba s fiksno verjetnostjo neplačil, sektorska analiza in posplošena sektorska analiza.

Poglavje 1

Sredstva iz verjetnosti in statistike

1.1 Slučajne spremenljivke

1.1.1 Diskretne slučajne spremenljivke

Diskretna slučajna spremenljivka je slučajna spremenljivka, ki zavzame le končno ali števno neskončno vrednosti. Kakšna je njena porazdelitev, nam pove njena verjetnostna funkcija.

Definicija 1.1.1 : *Naj bo X diskretna slučajna spremenljivka, ki zavzame vrednosti $x_1, x_2, \ldots$. Tedaj obstaja funkcija p, tako da velja*

$$p(x_i) = P(X = x_i) \quad \textit{in} \quad \sum_i p(x_i) = 1.$$

Ta funkcija se imenuje frekvenčna funkcija ali verjetnostna funkcija slučajne spremenljivke X.

Poleg verjetnostne funkcije pogosto potrebujemo porazdelitveno funkcijo, ki je definirana na naslednji način:

Definicija 1.1.2 : *Porazdelitvena funkcija slučajne spremenljivke X je enaka*

$$F_X(x) = P(X \leq x), \quad -\infty \leq x \leq \infty.$$

Lastnosti porazdelitvene funkcije:

- je naraščajoča
- $\lim_{x \to -\infty} F_X(x) = 0$

- $\lim_{x\to\infty} F_X(x) = 1$

Oglejmo si nekaj primerov pogostih diskretnih slučajnih spremenljivk:

Zgled 1.1.1 *Bernoullijeva slučajna spremenljivka X, $X \sim Bernoulli(p)$, zavzame samo dve vrednosti: 1 in 0 z verjetnostma p oz. $1-p$. To zapišemo na naslednji način:*

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

Verjetnostno funkcijo p definiramo kot

$$p(x) = \begin{cases} p; & x = 1 \\ 1-p; & x = 0 \\ 0; & sicer \end{cases}$$

ali krajše

$$p(x) = \begin{cases} p^x\,(1-p)^{1-x}; & x = 0 \text{ ali } x = 1 \\ 0; & sicer \end{cases}$$

Zgled 1.1.2 *Naj bo X binomska slučajna spremenljivka, $X \sim Bin(n,p)$. Tedaj velja*

$$p(k) = P(X = k) = \binom{n}{k} p^k\,(1-p)^{n-k}, \quad k = 0, 1, 2, ..., n.$$

Zgled 1.1.3 *Naj bo X Poissonova slučajna spremenljivka, $X \sim Poisson(\lambda), \lambda > 0$. Tedaj velja*

$$p(k) = P(X = k) = \frac{\lambda^k}{k!}\,e^{-\lambda}, \quad k = 0, 1, 2, ...$$

1.1.2 Zvezne slučajne spremenljivke

Pogosto nas bolj kot slučajne spremenljivke, ki zavzamejo le končno ali števno mnogo vrednosti, zanimajo tiste, ki niso porazdeljene diskretno. Kakšna je porazdelitev zvezne slučajne spremenljivke, nam pove njena gostota $f(x)$.

Definicija 1.1.3 : *Zvezna slučajna spremenljivka X ima gostoto, če obstaja taka funkcija f_X, da velja*

$$P(X \in A) = \int_A f_X(x)\,dx.$$

Gostota f ima naslednje lastnosti:

- $f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Zgled 1.1.4 : *Dana je gama slučajna spremenljivka X, $X \sim \Gamma(a, \lambda)$. Njena gostota je definirana za $x \geq 0$ in je enaka*

$$f_X(x) = \frac{1}{\Gamma(a)}\,\lambda^a\,x^{a-1}\,e^{-\lambda x}.$$

Zgled 1.1.5 : *Dana je beta slučajna spremenljivka X, $X \sim Beta(\alpha, \beta)$. Njena gostota je definirana za $x \in [0, 1]$ in je enaka*

$$f_X(x) = \frac{1}{B(\alpha, \beta)}\,x^{\alpha-1}\,(1-x)^{\beta-1}.$$

Zgled 1.1.6 : *Dana je enakomerna slučajna spremenljivka X, $X \sim I(a, b)$. Njena gostota je definirana za $x \in [a, b]$ in je enaka*

$$f_X(x) = \frac{1}{b-a}.$$

Zgled 1.1.7 : *Dana je normalna slučajna spremenljivka X, $X \sim N(\mu, \sigma^2)$. Njena gostota je definirana za vse $x \in \mathbb{R}$ in je enaka*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}}\,e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

V posebnem primeru, ko je $\mu = 0$ in $\sigma = 1$, X imenujemo standardna normalna slučajna spremenljivka.

Ena od posledic definicije je, da zvezna slučajna spremenljivka X zavzame posamezno vrednost z verjetnostjo 0

$$P(X = a) = \int_a^a f_X(x)\,dx = 0$$

Ta posledica se morda zdi nenavadna, vendar je smiselna. Namreč, če bi enakomerna slučajna spremenljivka iz zgleda 1.1.6 s pozitivno verjetnostjo zavzela katerokoli od možnih vrednosti, bi morala s to verjetnostjo zavzeti vse vrednosti iz intervala $[a, b]$. V tem primeru bi bila vsota verjetnosti neskončna.
Enakosti

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$$

veljajo za zvezno slučajno spremenljivko X, v diskretnem primeru pa to ne drži.
Če je f zvezna v x, je verjetnost, da X leži v majhnem intervalu okrog x, sorazmerna z $f(x)$:

$$P(x \leq X \leq x + dx) \approx f(x)\, dx.$$

Porazdelitvena funkcija zvezne slučajne spremenljivke je definirana na enak način kot porazdelitvena funkcija diskretne slučajne spremenljivke:

$$F_X(x) = P(X \leq x), \quad -\infty \leq x \leq \infty.$$

$F(x)$ lahko izrazimo s pomočjo gostote $f(x)$ takole:

$$F(x) = \int_{-\infty}^{x} f(u)\, du. \tag{1.1}$$

Če je f zvezna v x, od tod sledi

$$f(x) = F'(x). \tag{1.2}$$

S pomočjo porazdelitvene funkcije lahko izračunamo verjetnost, da X leži v intervalu (a, b)

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{a}^{b} f(u)\, du \tag{1.3}$$

Ker je porazdelitvena funkcija F naraščajoča, je inverz dobro definiran:

$$x = F^{-1}(y), \quad \text{če} \quad y = F(x).$$

Definicija 1.1.4 : *p-ti kvantil porazdelitvene funkcije F je taka vrednost x_p, da velja*

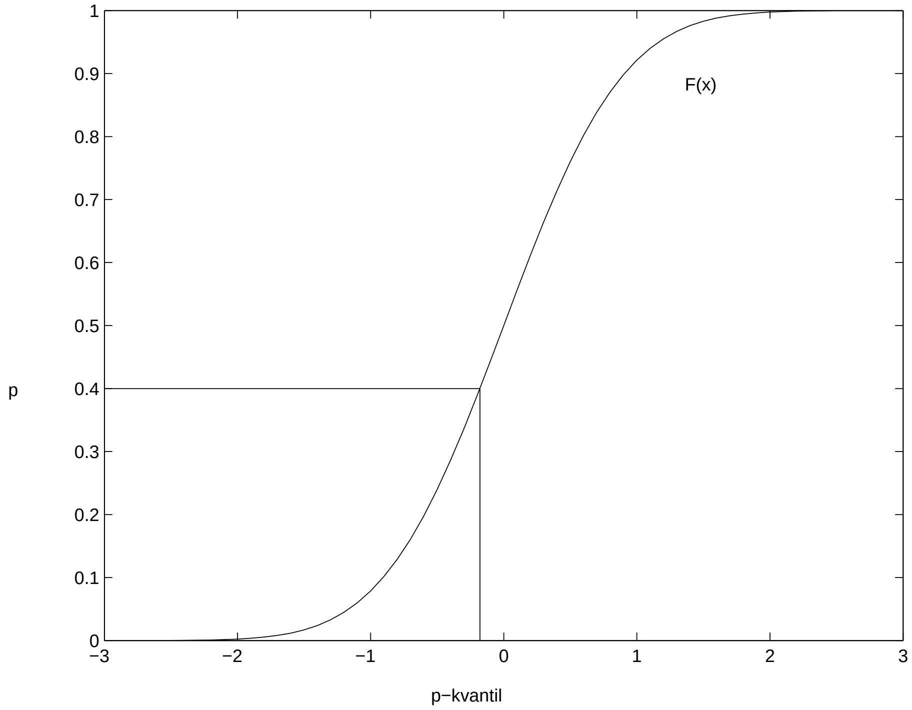
$$F(x_p) = p \quad oz. \quad P(X \leq x_p) = p.$$

Če je F strogo naraščajoča, potem je x_p enolično določen z $x_p = F^{-1}(p)$.

1.2 Večrazsežna porazdelitev

Obnašanje dveh slučajnih spremenljivk, X in Y, je v obeh primerih, zveznem in diskretnem, določeno s porazdelitveno funkcijo

$$F(x, y) = P(X \leq x, Y \leq y).$$



Slika 1.1: p-ti kvantil porazdelitvene funkcije F ($p = 0.4, x_p = -0.1791$)

Podobno lahko zapišemo za n slučajnih spremenljivk

$$F(x_1, x_2, ..., x_n) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n).$$

Za začetek si bolj podrobno oglejmo diskretni primer:

Definicija 1.2.1 : *Naj bosta X in Y diskretni slučajni spremenljivki, definirani na istem prostoru, in naj zavzameta vrednosti $x_1, x_2, ..., x_n$ oz. $y_1, y_2, ..., y_n$. Večrazsežna verjetnostna funkcija slučajnega vektorja (X, Y) je enaka*

$$p(x_i, y_j) = P(X = x_i, Y = y_j).$$

Definicija 1.2.2 : *Verjetnostna funkcija posamezne spremenljivke se imenuje robna verjetnostna funkcija in jo izračunamo takole:*

$$p_X(x) = \sum_i p(x, y_i) \quad oz. \quad p_Y(y) = \sum_i p(x_i, y).$$

Analogno lahko zapišemo tudi večrazsežno verjetnostno funkcijo večih slučajnih spremenljivk:

$$p(x_1, x_2, ..., x_m) = P(X_1 = x_1, X_2 = x_2, ..., X_m = x_m).$$

Robno verjetnostno funkcijo slučajne spremenljivke X_1 dobimo na naslednji način:

$$p_{X_1}(x) = \sum_{x_2,...,x_m} p(x, x_2, ..., x_m).$$

Podobno velja tudi za zvezni primer:

Definicija 1.2.3 : *Naj bodo $X_1, ..., X_n$ zvezne slučajne spremenljivke. Slučajni vektor $(X_1, ..., X_n)$ ima gostoto, če obstaja taka funkcija $f_{(X_1,...,X_n)}(x_1, ..., x_n)$, da velja*

$$P((X_1, ..., X_n) \in A) = \int ... \int_A f(x_1, ..., x_n)\, dx_n...dx_1.$$

V posebnem, če je $A = \{(X_1, ..., X_n); X_1 \leq x_1, ..., X_n \leq x_n\}$, lahko pišemo

$$F(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f(u_1, ..., u_n)\, du_n...du_1.$$

Od tod sledi

$$f(x_1, ..., x_n) = \frac{\partial^n}{\partial x_1 ... \partial x_n} F(x_1, ..., x_n),$$

kjer je odvod definiran.
Če je f zvezna v $(x_1, ..., x_n)$, je verjetnost, da slučajni vektor $(X_1, ..., X_n)$ leži v majhni okolici $(x_1, ..., x_n)$, sorazmerna z $f(x_1, ..., x_n)$. Včasih je uporabna sledeča aproksimacija:

$$P(x_1 \leq X_1 \leq x_1 + dx_1, ..., x_n \leq X_n \leq x_n + dx_n) \approx f(x_1, ..., x_n)\, dx_1...dx_n$$

Definicija 1.2.4 : *Robna porazdelitvena funkcija slučajne spremenljivke X je enaka*

$$F_X(x) = P(X \leq x) = \lim_{y \to \infty} F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(u, y)\, dy du.$$

Od tod dobimo robno gostoto slučajne spremenljivke X:

Definicija 1.2.5 : *Robna gostota slučajne spremenljivke X je enaka*

$$f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f(x, y)\, dy.$$

1.2.1 Neodvisne slučajne spremenljivke

Definicija 1.2.6 : *Slučajne spremenljivke* $X_1, X_2, ..., X_n$ *so neodvisne, če je večrazsežna porazdelitvena funkcija produkt robnih porazdelitvenih funkcij:*

$$F(x_1, x_2, ..., x_n) = F_{X_1}(x_1)\, F_{X_2}(x_2) \dots F_{X_n}(x_n)$$

za vse $x_1, x_2, ..., x_n$.

Definicija velja za diskretne in zvezne slučajne spremenljivke; v zveznem primeru je ekvivalentna

$$f(x_1, x_2, ..., x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \dots f_{X_n}(x_n). \tag{1.4}$$

Iz enakosti (1.4) namreč sledi

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dvdu = \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u) f_Y(v)\,dvdu =$$

$$= \left[\int_{-\infty}^{x} f_X(u)\,du\right]\left[\int_{-\infty}^{y} f_Y(v)\,dv\right] = F_X(x)\,F_Y(y)$$

1.2.2 Pogojna porazdelitev

Definicija 1.2.7 : *Naj bosta* X *in* Y *diskretni slučajni spremenljivki. Pogojna verjetnost* X *glede na* Y *je enaka*

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} =$$

$$= \frac{p_{(X,Y)}(x_i, y_j)}{p_Y(y_j)} \quad za \quad p_Y(y_j) > 0.$$

Označili jo bomo s $p_{X|Y}(x|y)$. Ta funkcija x je prava verjetnostna funkcija, ker je nenegativna in ker se njene vrednosti seštejejo v 1. Velja še: če sta X in Y neodvisna, je $p_{Y|X}(y|x) = p_Y(y)$.
Definicijo pogojne verjetnostne funkcije lahko zapišemo drugače

$$p_{(X,Y)}(x,y) = p_{X|Y}(x|y)\,p_Y(y).$$

Z vsoto po vseh vrednostih y na obeh straneh dobimo formulo za popolno verjetnost:

Definicija 1.2.8 (formula za popolno verjetnost) :

$$p_X(x) = \sum_y p_{X|Y}(x|y)\, p_Y(y)$$

Vse te količine lahko definiramo tudi za zvezne slučajne spremenljivke.

Definicija 1.2.9 : *Če sta X in Y zvezni slučajni spremenljivki, je pogojna gostota X glede na Y dana z*

$$f_{X|Y}(x|y) = \frac{f_{(X,Y)}(x,y)}{f_Y(x)} \quad \textit{za} \quad 0 < f_Y(x) < \infty.$$

Večrazsežno gostoto lahko izrazimo s pomočjo robne in pogojne gostote na naslednji način:

$$f_{(X,Y)}(x,y) = f_{Y|X}(y|x)\, f_X(x).$$

Integriranje obeh strani po x nam da formulo za popolno verjetnost:

Definicija 1.2.10 (formula za popolno verjetnost) :

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)\, f_X(x)\, dx$$

Ker velja $\frac{dF_X}{dx}(x) = f_X(x)$, lahko formulo za popolno verjetnost zapišemo v obliki

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)\, dF_X(x).$$

1.3 Številske karakteristike slučajnih spremenljivk

Večkrat se pojavi potreba, da bi poglavitne značilnosti kakšne porazdelitve karakterizirali z nekaj števili. Tem številom potem pravimo številske karakteristike ali parametri porazdelitve.

1.3.1 Matematično upanje

Definicija 1.3.1 : *Če je X diskretna slučajna spremenljivka z verjetnostno funkcijo $p(x)$, je njeno matematično upanje enako*

$$E(X) = \sum_i x_i\, p(x_i)$$

pri pogoju $\sum_i |x_i|\, p(x_i) < \infty$. Če vsota divergira, matematično upanje ni definirano.

O matematičnem upanju $E(X)$ lahko razmišljamo kot o pričakovani vrednosti oz. povprečju in ga pogosto označimo z μ ali μ_X. Na enak način definiramo matematično upanje zvezne slučajne spremenljivke, le da vsoto nadomestimo z integralom.

Definicija 1.3.2 : *Če je X zvezna slučajna spremenljivka z gostoto $f(x)$, je njeno matematično upanje enako*

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

pri pogoju $\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$. Če integral divergira, matematično upanje ni definirano.

Lema 1.3.1 :

1. $$E(X+Y) = E(X) + E(Y)$$
2. $$E(\alpha X) = \alpha E(X), \quad \alpha = konst.$$
3. $$X \geq 0 \Rightarrow E(X) \geq 0$$
4. *Če sta X in Y neodvisni slučajni spremenljivki in obstajata matematični upanji $E(X)$ in $E(Y)$, obstaja tudi $E(XY)$ in velja*
$$E(XY) = E(X)E(Y).$$

Dokaz : Dokazali bomo samo točko 4., saj lastnosti 1.−3. sledijo iz definicije. Zaradi neodvisnosti X in Y je

$$f_{(X,Y)}(x,y) = f_X(x) f_Y(y).$$

Računamo

$$E(X)E(Y) = \int_{-\infty}^{\infty} x f_X(x)\, dx \int_{-\infty}^{\infty} y f_Y(y)\, dy.$$

Ker obstajata $\int_{-\infty}^{\infty} |x| f_X(x)\, dx$ in $\int_{-\infty}^{\infty} |y| f_Y(y)\, dy$, obstaja dvakratni in zato dvojni integral. Zgornjo enakost lahko tako pišemo kot

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_X(x) f_Y(y)\, dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{(X,Y)}(x,y)\, dxdy = E(XY).$$

□

Pogosto nas zanima $E(g(X))$, kjer je X slučajna spremenljivka in g neka funkcija.

Trditev 1.3.1 : *Naj bo $Y = g(X)$.*

a) *Če je X diskretna slučajna spremenljivka z verjetnostno funkcijo $p(x)$, je*

$$E(Y) = \sum_x g(x)p(x)$$

pri pogoju $\sum |g(x)|p(x) < \infty$.

b) *Če je X zvezna slučajna spremenljivka z gostoto $f(x)$, je*

$$E(Y) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

pri pogoju $\int |g(x)|f(x)\,dx < \infty$.

Podobno velja v večih dimenzijah:

Trditev 1.3.2 : *Naj bo $(X_1, X_2, ..., X_n)$ slučajni vektor in $Y = g(X_1, X_2, ..., X_n)$.*

a) *Če so $X_1, ..., X_n$ diskretne slučajne spremenljivke z verjetnostno funkcijo $p(x_1, ..., x_n)$, je*

$$E(Y) = \sum_{x_1,...,x_n} g(x_1, ..., x_n)p(x_1, ..., x_n)$$

pri pogoju $\sum_{x_1,...,x_n} |g(x_1, ..., x_n)|p(x_1, ..., x_n) < \infty$.

b) *Če so $X_1, ..., X_n$ zvezne slučajne spremenljivke z gostoto $f(x_1, ..., x_n)$, je*

$$E(Y) = \int\int ... \int g(x_1, ..., x_n)f(x_1, ..., x_n)\,dx_1...dx_n$$

pri pogoju $\int\int ... \int |g(x_1, ..., x_n)|f(x_1, ..., x_n)\,dx_1...dx_n < \infty$.

1.3.2 Varianca in standardna deviacija

Matematično upanje slučajne spremenljivke je njena povprečna vrednost in ga lahko razumemo kot centralno vrednost gostote ali verjetnostne funkcije. Sedaj pa bomo spoznali količino, ki pove, kako razpršena je verjetnostna porazdelitev okoli svojega centra - standardno deviacijo slučajne spremenljivke. Najprej bomo definirali varianco slučajne spremenljivke, nato pa z njeno pomočjo še standardno deviacijo.

Definicija 1.3.3 : *Naj bo X slučajna spremenljivka z matematičnim upanjem $E(X)$. Njena varianca je tedaj*

$$var(X) = E\big((X - E(X))^2\big).$$

Standardna deviacija slučajne spremenljivke X je kvadratni koren variance $var(X)$.

Če je X diskretna slučajna spremenljivka z verjetnostno funkcijo $p(x)$ in matematičnim upanjem $\mu = E(X)$, lahko glede na definicijo in trditev 1.3.1 pišemo

$$var(X) = \sum_i (x_i - \mu)^2\, p(x_i).$$

Podobno velja za zvezno slučajno spremenljivko X z gostoto $f(x)$ in matematičnim upanjem $\mu = E(X)$

$$var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx.$$

Varianca slučajne spremenljivke X meri povprečno oddaljenost X od njenega matematičnega upanja $E(X)$ in je zato mera razpršenosti porazdelitve. Navadno jo označimo s σ^2, standardno deviacijo pa s σ.
V nadaljevanju bomo zapisali nekaj preprostih trditev, katerih dokaze lahko najdemo v [8].

Trditev 1.3.3 : *Če $var(X)$ obstaja in je $Y = aX + b$, je*

$$var(Y) = a^2\, var(X).$$

Varianco pa lahko izračunamo tudi malo drugače:

Trditev 1.3.4 : *Če varianca obstaja, je enaka*

$$var(X) = E(X^2) - E(X)^2.$$

1.3.3 Kovarianca in korelacija

Varianca slučajne spremenljivke je mera njena spremenljivosti, kovarianca dveh slučajnih spremenljivk pa mera skupne spremenljivosti oz. stopnje povezanosti med njima.

Definicija 1.3.4 : *Naj bosta X in Y slučajni spremenljivki z matematičnima upanjema $E(X)$ in $E(Y)$. Kovarianca med njima je tedaj enaka*

$$cov(X,Y) = E\big((X - E(X))(Y - E(Y))\big).$$

Kovarianca je torej povprečna vrednost produkta dveh spremenljivk: oddaljenosti spremenljivke X od njenega povprečja in oddaljenosti spremenljivke Y od njenega povprečja. Če sta slučajni spremenljivki pozitivno korelirani (tj. če je spremenljivka X večja od svojega povprečja, je tudi spremenljivka Y večja od svojega povprečja), je kovarianca pozitivna. Jasno je kovarianca negativna, če sta slučajni spremenljivki negativno korelirani (tj. če je spremenljivka X večja od svojega povprečja, je spremenljivka Y manjša od svojega povprečja).
Če razvijemo produkt, lahko kovarianco zapišemo na drug način:

$$cov(X,Y) = E(XY) - E(X)E(Y).$$

Vemo: če sta X in Y neodvisni, je $E(XY) = E(X)E(Y)$. Tedaj je $cov(X,Y) = 0$. Obratno ni nujno res.

Trditev 1.3.5 : *Naj bo $U = a + \sum_{i=1}^{n} b_i X_i$ in $V = c + \sum_{j=1}^{m} d_j Y_j$. Tedaj je*

$$cov(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_i d_j cov(X_i, Y_j).$$

Ta trditev ima mnogo uporabnih posledic. Ker je $var(X) = cov(X,X)$, lahko pišemo

$$var(X+Y) = cov(X+Y, X+Y) = var(X) + var(Y) + 2\,cov(X,Y).$$

V splošnem je formula za varianco linearne kombinacije slučajnih spremenljivk enaka:

Trditev 1.3.6 :

$$var\Big(a + \sum_{i=1}^{n} b_i X_i\Big) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_i b_j cov(X_i, X_j)$$

Če so slučajne spremenljivke X_i med seboj neodvisne, je $cov(X_i, X_j) = 0$ za $i \neq j$ in lahko zapišemo:

Posledica 1.3.1 : *Naj bodo slučajne spremenljivke X_i med seboj neodvisne. Tedaj velja*

$$var\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} var(X_i).$$

S pomočjo kovariance definiramo še eno pomembno količino:

Definicija 1.3.5 : *Naj bosta X in Y slučajni spremenljivki in naj obstajajo $var(X)$,$var(Y)$ in $cov(X,Y)$. Poleg tega naj bosta varianci neničelni. Tedaj definiramo korelacijski koeficient med X in Y takole:*

$$\rho = \frac{cov(X,Y)}{\sqrt{var(X)}\sqrt{var(Y)}}.$$

Korelacijski koeficient je mera za stopnjo korelacije, zato pogosto uporabljamo krajši izraz korelacija namesto korelacijski koeficient.

Korelacijski koeficient ni odvisen od enot in je zato v mnogih primerih bolj uporabna mera povezanosti kot kovarianca.
Ker standardni deviaciji X in Y označimo s σ_X in σ_Y, njuno kovarianco pa s $\sigma_{X,Y}$, lahko zgornjo definicijo zapišemo krajše

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}.$$

Trditev 1.3.7 :

$$-1 \leq \rho \leq 1$$

Dokaz : Ker je varianca slučajne spremenljivke nenegativna, lahko pišemo

$$0 \leq var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) =$$

$$= var\left(\frac{X}{\sigma_X}\right) + var\left(\frac{Y}{\sigma_Y}\right) + 2\,cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right) =$$

$$= \frac{var(X)}{\sigma_X^2} + \frac{var(Y)}{\sigma_Y^2} + \frac{2\,cov(X,Y)}{\sigma_X \sigma_Y} = 2\,(1+\rho)$$

Od tod sledi $\rho \geq -1$. Podobno

$$0 \leq var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 2\,(1-\rho),$$

kar ima za posledico $\rho \leq 1$.

$\square$

1.3.4 Pogojno matematično upanje

Definicija 1.3.6 : *Naj bo A dogodek s $P(A) > 0$. Pogojno matematično upanje glede na dogodek A definiramo kot*

$$E(X|A) = \frac{1}{P(A)}\, E(X \cdot 1_A).$$

Definicija 1.3.7 : *Naj bo X poljubna slučajna spremenljivka z $E(|X|) < \infty$, Y pa diskretna slučajna spremenljivka z vrednostmi $\{y_1, y_2, ...\}$. Definiramo*

$$E(X|Y = y_k) = \frac{1}{P(Y = y_k)}\, E(X \cdot 1_{\{Y=y_k\}}).$$

Pogojno matematično upanje slučajne spremenljivke X glede na slučajno spremenljivko Y s pomočjo zgornje enakosti zapišemo takole:

$$E(X|Y) = \sum_k \frac{1}{P(Y = y_k)}\, E(X \cdot 1_{\{Y=y_k\}})\, 1_{\{Y=y_k\}}.$$

Opombe:

- $E(X|Y)$ je funkcija Y in je zato slučajna spremenljivka.
- Za pogojno matematično upanje slučajne spremenljivke X glede na $\{Y = y\}$ včasih uporabimo tudi naslednjo definicijo:

  Definicija 1.3.8 : *Naj bosta X in Y diskretni slučajni spremenljivki in $p_{X|Y}(x|y)$ pogojna verjetnostna funkcija X glede na Y. Pogojno matematično upanje slučajne spremenljivke X glede na $\{Y = y\}$ je enako*

  $$E(X|Y = y) = \sum_x x\, p_{X|Y}(x|y).$$

  Za zvezni slučajni spremenljivki X in Y ter pogojno gostoto $f_{X|Y}(x|y)$ pogojno matematično upanje definiramo kot

  $$E(X|Y = y) = \int x\, f_{X|Y}(x|y)\, dx.$$

  Bolj splošno, pogojno matematično upanje funkcije $h(X)$ je

  $$E(h(X)|Y = y) = \int h(x)\, f_{X|Y}(x|y)\, dx.$$

Če pogojno matematično upanje slučajne spremenljivke X glede na dogodek $\{Y = y\}$ obstaja za vsako vrednost y, ki jo Y lahko zavzame, je dobro definirana funkcija Y in je zato slučajna spremenljivka. Označimo jo z $E(X|Y)$.

Zgled 1.3.1 : *Imejmo neodvisni slučajni spremenljivki X in Y, obe binomsko porazdeljeni: $X \sim Bin(n,p), Y \sim Bin(m,p)$. Zanima nas pogojno matematično upanje $E(X|X+Y)$.*
Najprej izračunajmo pogojno verjetnostno funkcijo:

$$p_{X|X+Y}(x|x+y) = P(X=k|X+Y=l) = \frac{P(X=k, X+Y=l)}{P(X+Y=l)} =$$

$$= \frac{P(X=k, Y=l-k)}{P(X+Y=l)} = \frac{P(X=k)\,P(Y=l-k)}{P(X+Y=l)} =$$

$$= \frac{\binom{n}{k}\,p^k\,q^{n-k}\,\binom{m}{l-k}\,p^{l-k}\,q^{m-l+k}}{\binom{m+n}{l}\,p^l\,q^{m+n-l}} = \frac{\binom{n}{k}\binom{m}{l-k}}{\binom{m+n}{l}}.$$

S pomočjo le-te dobimo pogojno matematično upanje:

$$E(X|X+Y=l) = \sum_{k=1}^{n} k\,P(X=k|X+Y=l) = \frac{l\cdot n}{n+m}.$$

In nazadnje

$$E(X|X+Y) = \frac{(X+Y)\cdot n}{n+m}.$$

Zanimali nas bosta matematično upanje in varianca slučajne spremenljivke $E(X|Y)$.

Izrek 1.3.1 :

$$E\big(E(Y|X)\big) = E(Y)$$

Dokaz : Ta izrek bomo dokazali za diskreten primer; dokaz za zvezen primer gre podobno. Pokazati moramo

$$E(Y) = \sum_x E(Y|X=x)p_X(x),$$

kjer je

$$E(Y|X=x) = \sum_y y p_{Y|X}(y|x).$$

Če zamenjamo vrstni red obeh vsot, dobimo

$$\sum_x E(Y|X=x)p_X(x) = \sum_y y \sum_x p_{Y|X}(y|x)p_X(x).$$

Spomnimo se formule za popolno verjetnost

$$p_Y(y) = \sum_x p_{Y|X}(y|x)p_X(x).$$

Od tod

$$\sum_y \sum_x p_{Y|X}(y|x)p_X(x) = \sum_y y p_Y(y) = E(Y).$$

□

Izrek 1.3.2 :

$$var(E(Y|X)) = var(Y) - E(var(Y|X))$$

Dokaz : Najprej moramo seveda pojasniti, kako je definirana pogojna varianca:

$$var(Y|X = x) = E(Y^2|X = x) - (E(Y)|X = x)^2 \tag{1.5}$$

za vse vrednosti x.
Pogojno varianco $var(Y|X)$, kot prej $E(Y|X)$, definiramo kot slučajno spremenljivko. Matematično upanje slučajne spremenljivke $var(Y|X)$ je jasno $E(var(Y|X))$. Zapišemo lahko

$$var(Y|X) = E(Y^2|X) - (E(Y|X))^2 \tag{1.6}$$

in

$$E(var(Y|X)) = E(E(Y^2|X)) - E((E(Y|X))^2). \tag{1.7}$$

Velja tudi (po definiciji variance)

$$var(E(Y|X)) = E((E(Y|X))^2) - (E(E(Y|X)))^2. \tag{1.8}$$

Potrebujemo samo še

$$var(Y) = E(Y^2) - (E(Y))^2 = E(E(Y^2|X)) - (E(E(Y|X)))^2. \tag{1.9}$$

Z uporabo enakosti (1.9), (1.7) in (1.8) končno dobimo

$$var(Y) = E(E(Y^2|X)) - (E(E(Y|X)))^2 =$$

$$= E(E(Y^2|X)) - E(E(Y|X)^2) + E(E(Y|X)^2) - (E(E(Y|X)))^2 =$$

$$= E(var(Y|X)) + var(E(Y|X))$$

□

1.3.5 Momenti slučajnih spremenljivk

Definicija 1.3.9 : *Naj bo X diskretna slučajna spremenljivka. k-ti moment ali moment reda k slučajne spremenljivke X glede na vrednost a je enak*

$$m_k(a) = E((X-a)^k) = \sum_i (x_i - a)^k \, p(x_i).$$

Če je X zvezna slučajna spremenljivka, je njen k-ti moment glede na vrednost a enak

$$m_k(a) = E((X-a)^k) = \int_{-\infty}^{\infty} (x-a)^k \, dF_X(x).$$

Nekateri momenti imajo posebno ime: $z_k = m_k(0)$ je začetni moment, $m_k = m_k(E(X))$ pa centralni moment.

1.4 Rodovne funkcije

Pri obravnavanju slučajnih spremenljivk se za zelo koristne izkažejo nekatere transformacije njihovih porazdelitvenih zakonov. V tem poglavju si bomo ogledali eno od njih: rodovno funkcijo. Uporabljamo jo le pri slučajnih spremenljivkah, katerih zaloga vrednosti je množica nenegativnih celih števil ali kakšna njena podmnožica. Takim slučajnim spremenljivkam pravimo celoštevilske.

Definicija 1.4.1 *Naj bo X celoštevilska slučajna spremenljivka. Njeno rodovno funkcijo G_X definiramo kot*

$$G_X(s) = E(s^X) = \sum_{k=0}^{\infty} s^k P(X=k).$$

Opomba:
Za $|s| \le 1$ vrsta konvergira enakomerno in absolutno. Za $|s| < 1$ je $G_X(s)$ neskončnokrat odvedljiva.
Rodovna funkcija $G_X(s)$ vsebuje informacijo o porazdelitvi X.

Zgled 1.4.1 *Naj bo $X \sim Bin(n,p)$. Rodovna funkcija slučajne spremenljivke X je enaka:*

$$G_X(s) = \sum_{k=0}^{\infty} \binom{n}{k} p^k q^{n-k} s^k = (q+ps)^n.$$

Izrek 1.4.1 *Naj bosta X in Y nenegativni, celoštevilski slučajni spremenljivki z rodovnima funkcijama G_X in G_Y. Velja:*

1. *G_X natanko določa porazdelitev $P(X = k)$ za $k = 0, 1, ...$*
2. *Če sta X in Y neodvisni, je $G_{X+Y}(s) = G_X(s)\, G_Y(s)$*

Dokaz :

1. $P(X = k) = \frac{G^{(k)}(0)}{k!}$
2. Če sta X in Y neodvisni, sta tudi s^X in s^Y neodvisni. Sledi

$$G_{X+Y}(s) = E(s^{X+Y}) = E(s^X s^Y) =$$

$$E(s^X)\, E(s^Y) = G_X(s)\, G_Y(s).$$

□

1.5 Vzorčne funkcije

1.5.1 Populacija in vzorec

Naj bo G množica z elementi e. Imenovali jo bomo populacija. Naj bo X neka količina, katere vrednost $X(e)$ je pri vsakem e iz G natanko določena. $X(e)$ razumemo kot enolično realno funkcijo, definirano na G. Velja še, da je pri vsakem realnem x natanko določen delež $F(x)$ tistih elementov e iz množice G, pri katerih je $X(e) < x$. Pri končnih populacijah je to zahtevo vedno mogoče izpolniti: če je n velikost populacije in m število tistih elementov e, za katere je $X(e) < x$, je $F(x) = \frac{m}{n}$. V vsakem primeru je funkcija $F(x)$ nepadajoča, definirana na vsej realni osi in zanjo velja

$$\lim_{x\to\infty} F(x) = 1, \quad \lim_{x\to-\infty} F(x) = 0.$$

Privzemimo, da bomo elemente e izbirali na slepo. To pomeni, da ima pri vsakem izbiranju vsak element enako možnost, da bo izbran. Tedaj je pri vsakem realnem x pojav, da bo pri izbranem elementu e veljala neenakost $X(e) < x$, dogodek z verjetnostjo $F(x)$:

$$F(x) = P(X(e) < x) \tag{1.10}$$

Pri slepem izbiranju elementov iz G je torej količina $X = X(e)$ slučajna spremenljivka, $F(x)$ pa njena porazdelitvena funkcija.
Če je populacija G končna, ima spremenljivka $X(e)$ samo končno mnogo

različnih vrednosti, recimo $x_1, ..., x_r$. Naj bo $f_i, i = 1, ..., r$, delež tistih elementov e, pri katerih je $X(e) = x_i$. Pri slučajnem izbiranju je tako X diskretna slučajna spremenljivka z verjetnostno funkcijo

$$p(x_i) = P(X(e) = x_i) = f_i. \tag{1.11}$$

Namesto merljive lastnosti $X(e)$ imajo elementi e lahko kakšno kvalitativno lastnost $A(e)$. Recimo, da so njena mogoča stanja $A_1, ..., A_r$ in da so njihove relativne frekvence na G števila $f_1, ..., f_r$. Pri slepem izbiranju elementov iz populacije sestavljajo pojavi, da izberemo element, katerega lastnost $A(e)$ je v stanju $A_i, i = 1, ..., r$, popoln sistem dogodkov. Vsak dogodek A_i ima pri tem verjetnost

$$P(A_i) = f_i \tag{1.12}$$

Naj bo X slučajna spremenljivka pri slepem izbiranju elementov iz populacije G. Če bi hoteli poznati porazdelitveni zakon te slučajne spremenljivke, to je porazdelitveno funkcijo iz relacije (1.10) ali verjetnostno funkcijo iz relacije (1.11), bi morali ugotoviti vrednost X pri vseh elementih populacije. Pri neskončni populaciji tega ne moremo narediti, pri končni pa je večkrat nesmotrno, ker je predrago ali pa se populacija pri tem celo uniči. V takih primerih se moramo zadovoljiti s tem, da preučimo spremenljivko X samo na primerno izbranem delu populacije. Temu delu populacije pravimo vzorec. Seveda je pomembno, kako zanesljiv je vzorec: kaj je mogoče reči o slučajni spremenljivki X na populaciji na podlagi informacije, ki jo dobimo iz vzorca. S tega stališča je najboljši vzorec sestavljen tako, da imajo pri vsakem izbiranju za vzorec vsi elementi populacije enako možnost, da bodo izbrani. Tako sestavljen vzorec se imenuje enostaven slučajen vzorec ali kar slučajen vzorec. Če hočemo dobiti slučajen vzorec, je treba že izbrane elemente pred ponovnim izbiranjem vrniti v populacijo, tako da je lahko isti element večkrat izbran.
Vzemimo, da imamo slučajni vzorec

$$\{e_1, ..., e_n\}. \tag{1.13}$$

Ugotavljanje vrednosti $X(e)$ pri slepo izbranem elementu e je s stališča verjetnostnega računa realizacija slučajne spremenljivke X. Vzorec (1.13) nam da n realizacij količine X, recimo $x_1, ..., x_n$. Iz njih sestavimo vektor

$$\mathbf{z} = (x_1, ..., x_n), \tag{1.14}$$

ki ga imamo lahko za realizacijo slučajnega vektorja

$$\mathbf{Z} = (X_1, ..., X_n). \tag{1.15}$$

Če je vzorec (1.13) slučajen, so vse komponente slučajnega vektorja **Z** med seboj neodvisne in enako porazdeljene, namreč tako kot slučajna spremenljivka X na populaciji. V tem je glavna prednost slučajnega vzorca.

1.5.2 Mera pozicije

Mera pozicije določi pozicijo posamezne vrednosti glede na ostale vrednosti v vzorcu ali populaciji. Največkrat uporabimo kvartile ali percentile.

Trije kvartili (Q_1, Q_2 in Q_3) razdelijo urejene podatke na 4 enake dele.

25%	25%	25%	25%
Q_1	Q_2	Q_3	

Zgled 1.5.1 *Dan je urejen vzorec:*

$$7, 16, 18, 18, 19, 20, 20, 22, 31, 34, 38, 58.$$

Razdelimo ga na štiri enake dele: v prvem delu so vrednosti $7, 16, 18$, *v drugem* $18, 19, 20$ *itd. Prvi kvartil* Q_1 *izračunamo kot povprečje med zadnjo vrednostjo prvega dela in prvo vrednostjo drugega dela, torej* $Q_1 = \frac{18+18}{2} = 18$. *Podobno dobimo* $Q_2 = \frac{20+20}{2} = 20$ *in* $Q_3 = \frac{31+34}{2} = 32.5$.

Percentili razdelijo urejene podatke na 100 enakih delov. k-ti percentil označimo s P_k in pomeni vrednost v množici podatkov, tako da je okrog $k\%$ meritev manjših od vrednosti P_k in okrog $(100-k)\%$ večjih od vrednosti P_k. P_k je torej vrednost $\left(\frac{kn}{100}\right)$-tega izraza v urejeni množici podatkov.

Zgled 1.5.2 *Imejmo urejen vzorec kot prej:*

$$7, 16, 18, 18, 19, 20, 20, 22, 31, 34, 38, 58.$$

Pozicijo 62. *percentila izračunamo po formuli* $\frac{kn}{100}$, *torej* $\frac{62 \cdot 12}{100} = 7.44$. 7.44-*ti izraz lahko aproksimiramo s povprečjem med* 7. *in* 8. *izrazom. Tako je* $P_{62} = \frac{20+22}{2} = 21$.

1.5.3 Vzorčna porazdelitvena funkcija

Definicija 1.5.1 : *Vzorec* $\{e_1, ..., e_n\}$ *lahko razumemo kot populacijo, na kateri je definirana slučajna spremenljivka* X. *Za vsak realen* x *označimo s* $k(x)$ *število tistih elementov v vzorcu, pri katerih je* $X(e) < x$. *Tedaj je funkcija*

$$V_n(x) = \frac{k(x)}{n}$$

porazdelitvena funkcija spremenljivke X *na vzorcu. Pravimo ji vzorčna porazdelitvena funkcija ali empirična porazdelitvena funkcija slučajne spremenljivke* X.

Naj bo x poljubno realno število in naj bo pri slučajnem izbiranju iz populacije G

$$F(x) = P(X < x).$$

Sestavljanje slučajnega vzorca velikosti n imamo lahko za n ponovitev istega poskusa. Označimo s $P_n(k)$ verjetnost, da ima pri teh ponovitvah dogodek $\{X < x\}$ frekvenco k. Po Bernoullijevi formuli je

$$P_n(k) = \binom{n}{k}(F(x))^k(1 - F(x))^{n-k}.$$

Dogodek $\{X < x\}$ ima frekvenco k natanko takrat, kadar je $V_n(x) = \frac{k}{n}$. Torej je

$$P\left(V_n(x) = \frac{k}{n}\right) = \binom{n}{k}(F(x))^k(1 - F(x))^{n-k}. \tag{1.16}$$

Empirična porazdelitvena funkcija $V_n(x)$ je torej diskretna slučajna spremenljivka; količina $nV_n(x)$ je porazdeljena po binomskem zakonu $Bin(n, F(x))$. Ker je količina $V_n(x)$ pri danem x slučajna spremenljivka, se v splošnem ne ujema z vrednostjo $F(x)$, velja pa

$$E(V_n(x)) = F(x).$$

Zanima nas, kaj se dogaja z $V_n(x)$, ko n raste čez vse meje. To nam pove Glivenkov izrek, imenovan tudi Osnovni izrek matematične statistike. Za dokaz tega izreka bomo potrebovali še Borelov izrek.

Izrek 1.5.1 (Borelov izrek) : *Če je* $\frac{k}{n}$ *relativna frekvenca dogodka* A *v* n *poskusih in je* p *verjetnost dogodka* A *v posameznem poskusu, velja*

$$p\left(\lim_{n\to\infty} \frac{k}{n} = p\right) = 1.$$

Izrek 1.5.2 (Glivenkov izrek) : *Naj bo* $D_n := \sup_{x\in\mathbb{R}} |V_n(x) - F(x)|$ *za* $n = 1, 2, 3, \ldots.$ *Zaporedje* $D_1, D_2, D_3, \ldots$ *skoraj gotovo konvergira proti* 0*, to pomeni,*

$$P(\lim_{n\to\infty} D_n = 0) = 1.$$

Dokaz : Izberemo poljubno naravno število m in pri vsakem k, $1 \leq k \leq m$, označimo z x_{mk} najmanjše realno število x, pri katerem je

$$F(x) \leq \frac{k}{m} \leq F(x + 0).$$

Naj bo

$$D_n^{(1)} = max_k |V_n(x_{mk}) - F(x_{mk})|,$$

$$D_n^{(2)} = max_k |V_n(x_{mk} + 0) - F(x_{mk} + 0)|.$$

Med količinami $D_n^{(1)}$,$D_n^{(2)}$ in D_n je zveza

$$D_n \leq max\{D_n^{(1)}, D_n^{(2)}\} + \frac{1}{m}. \tag{1.17}$$

Naj bo x fiksen. Število $V_n(x)$ je relativna frekvenca dogodka $\{X < x\}$ v n ponovitvah poskusa, v katerem ima ta dogodek verjetnost $F(x)$. Po Borelovem izreku je

$$P\big(lim_{n\to\infty} V_n(x) = F(x)\big) = 1. \tag{1.18}$$

Podobno je $V_n(x + 0)$ relativna frekvenca dogodka $\{X \leq x\}$ v n ponovitvah poskusa, v katerem je $P(X \leq x) = F(x + 0)$. Zato je

$$P\big(lim_{x\to\infty} V_n(x + 0) = F(x + 0)\big) = 1. \tag{1.19}$$

Iz enakosti (1.17), (1.18) in (1.19) sledi, da je za vsako naravno število m

$$P(\overline{lim_{n\to\infty}} D_n > \frac{1}{m}) = 0,$$

kar smo zahtevali.

$\square$

Vzorčna porazdelitvena funkcija torej pri naraščanju vzorca z verjetnostjo 1 enakomerno konvergira k porazdelitveni funkciji na populaciji.

1.5.4 Statistike

Definicija 1.5.2 : *Naj bo X slučajna spremenljivka in*

$$\mathbf{Z} = (X_1, ..., X_n)$$

njen slučajni vzorec. Naj bo U neka Borelova funkcija (npr. zvezna ali taka s števno zalogo vrednosti)

$$U : R^n \to R^m \quad (m < n), \quad U = U(\mathbf{Z}) = U(X_1, ..., X_n).$$

U imenujemo vzorčna statistika ali na kratko statistika.

Oglejmo si nekaj pomembnih statistik:

1. a) Naj bo r celo nenegativno število in c katerokoli realno število. Količina

   $$M_r(c) = \frac{1}{n} \sum_{k=1}^{n} (X_k - c)^r$$

   se imenuje r-ti vzorčni moment glede na vrednost c. Posebej je

   $$Z_r = \frac{1}{n} \sum_{k=1}^{n} X_k^r$$

   r-ti vzorčni začetni moment.
   Vzorčni začetni moment prvega reda Z_1 imenujemo vzorčno povprečje. Označimo ga z $\overline{X}$ in pišemo

   $$\overline{X} = \frac{1}{n} \sum_{k=1}^{n} X_k.$$

   b) Statistiko

   $$S^2 = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X})^2$$

   zaradi njenih lepih lastnosti, o katerih bomo nekaj več povedali kasneje, nekateri imenujejo vzorčna disperzija, čeprav dejansko to ni.

2. Ogledali smo si statistike, ki so ali vzorčni momenti ali pa njihove funkcije (npr. S^2). Drug pomemben vir statistik je urejanje vzorca $X_1, ..., X_n$ po velikosti.

a) V vzorcu $X_1, ..., X_n$ so vrednosti zapisane po vrstnem redu izbiranja: vrednost X_k dobimo pri k-tem izbiranju v vzorec. Zdaj pa vsako realizacijo vzorca $x_1, ..., x_n$ razvrstimo v naraščajoče zaporedje

$$x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}.$$

Z $X_{(k)}$ označimo slučajno spremenljivko, ki ima pri vsaki realizaciji vzorca za vrednost število $x_{(k)}, k = 1, ..., n$. Slučajne spremenljivke

$$X_{(1)}, X_{(2)}, ..., X_{(n)}$$

se imenujejo vrstilne statistike.
V nadaljevanju bomo pokazali, kako porazdelitveno funkcijo k-te vrstilne statistike izrazimo s pomočjo porazdelitvene funkcije slučajne spremenljivke X. Kot prej z $F(x)$ označimo porazdelitveno funkcijo slučajne spremenljivke X, z $V_n(x)$ pa vzorčno porazdelitveno funkcijo. Naj bo $F_{(k)}(x)$ porazdelitvena funkcija k-te vrstilne statistike

$$F_{(k)}(x) = P(X_{(k)} < x).$$

Oglejmo si jo natančneje. Dogodek $\{X_{(k)} < x\}$ se zgodi natanko takrat, ko je vsaj k vrednosti v vzorcu manjših od x oz. kadar ima vzorčna porazdelitvena funkcija vrednost večjo ali enako $\frac{k}{n}$. Zato je

$$F_{(k)}(x) = P\left(V_n(x) \geq \frac{k}{n}\right) = \sum_{m=k}^{n} P\left(V_n(x) = \frac{m}{n}\right). \tag{1.20}$$

Po formuli (1.16) je

$$P\left(V_n(x) = \frac{m}{n}\right) = \binom{n}{m}(F(x))^m(1 - F(x))^{n-m}.$$

To in enakost (1.20) nam dasta

$$F_{(k)}(x) = \sum_{m=k}^{n} \binom{n}{m}(F(x))^m(1 - F(x))^{n-m}. \tag{1.21}$$

Tako torej izrazimo porazdelitveno funkcijo k-te vrstilne statistike s porazdelitveno funkcijo slučajne spremenljivke X.

b) Naj bo $0 < p < 1$ in

$$r = \begin{cases} np; & np \in \mathbb{Z} \\ [np]+1; & \text{sicer} \end{cases}$$

Statistika

$$Q_p = X_{(r)}$$

se imenuje p-ti vzorčni kvantil.

1.5.5 Ocenjevanje parametrov

Naj ima slučajna spremenljivka X porazdelitveno funkcijo $F(x)$, ki je odvisna od nekega skalarnega ali vektorskega parametra q. Zato namesto $F(x)$ pišemo $F(x,q)$, torej

$$F(x,q) = P(X < x).$$

Če je X zvezna, lahko porazdelitveni zakon opišemo tudi z gostoto verjetnosti

$$p(x,q) = F'(x,q).$$

V glavnem bomo obravnavali situacijo, ko je tip porazdelitvene funkcije F ali njene gostote p znan, parameter q pa neznan.
Naj bo

$$\mathbf{Z} = (X_1, ..., X_n)$$

slučajen vzorec za spremenljivko X. Zalogo vrednosti vzorca $\mathbf{Z}$, to je množico vseh mogočih realizacij $\mathbf{z} = (x_1, ..., x_n)$, imenujemo prostor vzorcev in ga označimo z W. Gostoto slučajnega vzorca zapišemo kot

$$L(\mathbf{z},q) = L(x_1, ..., x_n, q) = p(x_1,q) \; ... \; p(x_n,q)$$

in jo večkrat obravnavamo kot funkcijo parametra q pri fiksirani realizaciji $\mathbf{z} = (x_1, ..., x_n)$ slučajnega vzorca. V tem primeru ji rečemo funkcija zanesljivosti vzorca. To poimenovanje izhaja iz enakosti

$$L(\mathbf{z},q) = P(\mathbf{Z} = \mathbf{z}),$$

ki velja za diskretno slučajno spremenljivko X z verjetnostno funkcijo $p(x,q)$. $L(\mathbf{z},q)$ kot funkcija parametra q tedaj meri, s kolikšno verjetnostjo se lahko zanesemo, da bomo za slučajni vzorec $\mathbf{Z}$ dobili ravno predpisano realizacijo $\mathbf{z}$.
Zanima nas, kako bi informacijo o $F(x,q)$, ki se skriva v slučajnem vzorcu $\mathbf{Z}$, izkoristili za sklepanje o neznanem parametru q. Poznamo dve metodi: točkasto in intervalsko ocenjevanje parametra. Zaenkrat bomo opisali samo prvo.

Definicija 1.5.3 : *Naj bo*

$$U = f(X_1, ...X_n) = f(\mathbf{Z})$$

statistika, ki funkcijsko ni odvisna od q in katere porazdelitveni zakon je odvisen od q. Poleg tega naj prostor vzorcev upodobi v prostor parametrov, to pomeni

$$f(W) \subset Q.$$

Vrednost u take statistike U na dani realizaciji $\mathbf{z}$ *vzorca* $\mathbf{Z}$*,*

$$u = f(\mathbf{z}),$$

vzamemo za oceno parametra q, statistika U pa se zaradi svoje vloge imenuje cenilka parametra q.

Zgled 1.5.3 *Naj bo* $X \sim N(\mu, \sigma^2)$*. Vzorčno povprečje* $\overline{X}$ *je cenilka matematičnega upanja* μ*, statistika* S^2 *pa je cenilka variance* σ^2*.*

Cenilka je dobra, kadar z veliko verjetnostjo daje ocene, ki se malo ločijo od prave vrednosti parametra. V nadaljevanju se bomo ukvarjali z lastnostmi, ki naj bi jih imela dobra cenilka.

Definicija 1.5.4 : *Naj ima slučajna spremenljivka X gostoto* $p(x, q)$*, q neznan, in naj bo* $\mathbf{Z} = (X_1, ..., X_n)$ *slučajen vzorec za X. Dalje, naj bo* $U = u(\mathbf{Z})$ *cenilka parametra q. Če je pri vsakem mogočem v pogojna gostota vzorca* $\mathbf{Z}$ *glede na dogodek* $\{U = v\}$ *neodvisna od q, pravimo, da je cenilka U zadostna za q.*

Zgled 1.5.4 *Naj bo X Bernoullijeva slučajna spremenljivka,* $X \sim Bernoulli(p)$*, in parameter p neznan. Oglejmo si pogojno porazdelitev vzorca* $(X_1, ..., X_n)$ *glede na dogodek* $\{\overline{X} = \frac{k}{n}\}$*.*

$$P\left(X_1 = x_1, ..., X_n = x_n | \overline{X} = \frac{k}{n}\right) = \frac{P\big(X_1 = x_1, ..., X_n = x_n, \overline{X} = \frac{k}{n}\big)}{P\big(\overline{X} = \frac{k}{n}\big)} =$$

$$= \frac{p^k\,(1-p)^{n-k}}{\binom{n}{k}\,p^k\,(1-p)^{n-k}} = \frac{1}{\binom{n}{k}}.$$

Pogojna porazdelitev vzorca $\mathbf{Z}$ *glede na dogodek* $\{\overline{X} = \frac{k}{n}\}$ *je torej neodvisna od parametra p, statistika* $\overline{X}$ *je zadostna cenilka za p.*

Definicija 1.5.5 : *Slučajna spremenljivka X naj bo porazdeljena po zakonu $F(x, q)$, $q \in Q$, funkcija slučajnega vzorca $U = U(X_1, ..., X_n)$ pa naj bo cenilka parametra q. Če je*

$$E(U) = q \quad \forall q \in Q,$$

se U imenuje nepristranska cenilka parametra q. Če pa je

$$E(U) \neq q \quad \text{za kak } q \in Q,$$

je cenilka U pristranska.

Če je cenilka U nepristranska in $U_1, U_2, ...$ njene neodvisne realizacije, je po krepkem zakonu velikih števil

$$P\Big(lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{m} U_k = q\Big) = 1.$$

V tem je prednost nepristranske cenilke pred pristransko. Seveda se ta prednost pokaže šele pri velikem številu ocen. Posamezna ocena, dobljena z nepristransko cenilko, pa je lahko manj točna od ocene, ki jo da kakšna pristranska cenilka.

Zgled 1.5.5 *Če ima slučajna spremenljivka X končno matematično upanje μ, je vzorčno povprečje $\overline{X}$ nepristranska cenilka za μ. Statistika S^2 pa je nepristranska cenilka variance σ^2.*

1.6 Vzorčenje

1.6.1 Parametri populacije

V tem podpoglavju definiramo tiste številske parametre populacije, ki jih bomo skušali oceniti s pomočjo vzorca. Naj bo velikost populacije enaka N. Poznamo vrednosti neke lastnosti vseh članov populacije; označimo jih z $x_1, x_2, ..., x_N$. Spremenljivka x_i je tako lahko starost, višina, teža itd. Lahko pa zavzame vrednosti 1 ali 0, kar pomeni prisotnost oz. odsotnost neke lastnosti. Tedaj govorimo o t.i. dihotomnem primeru.

Definicija 1.6.1 : *Povprečje populacije je enako*

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i.$$

V dihotomnem primeru je povprečje μ enako deležu p posameznikov, ki imajo dano lastnost.

Definicija 1.6.2 : *Celota populacije (population total) je enaka*

$$\tau = \sum_{i=1}^{N} x_i = N\mu.$$

V dihotomnem primeru je celota populacije število vseh članov populacije, ki imajo dano lastnost.

Definicija 1.6.3 : *Varianca populacije je enaka*

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2.$$

Zapišemo lahko tudi malo drugače:

$$\sigma^2 = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2 - 2\,\mu\sum_{i=1}^{N} x_i + N\mu^2\right) = \tag{1.22}$$

$$= \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2 - 2\,N\mu^2 + N\mu^2\right) = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2.$$

V dihotomnem primeru se varianca populacije poenostavi v

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 - \mu^2 = p - p^2 = p\,(1-p).$$

Upoštevali smo dejstvo, da je vsak x_i^2 enak 0 ali 1.

Standardna deviacija populacije je kvadratni koren variance populacije in je mera razpršenosti populacije.

1.6.2 Enostavno slučajno vzorčenje

Najbolj elementarna oblika vzorčenja je enostavno slučajno vzorčenje: vsak od $\binom{N}{n}$ vzorcev velikosti n se pojavi z enako verjetnostjo. Privzamemo, da se vzorčenje izvede brez ponavljanja: tako se vsak element populacije v posameznem vzorcu pojavi največ enkrat.
Ker je sestavljanje vzorca naključno, je tako tudi vzorčno povprečje. Analiza natančnosti aproksimacije povprečja populacije z vzorčnim povprečjem mora biti zato verjetnostne narave. V nadaljevanju bomo izpeljali nekaj statističnih lastnosti vzorčnega povprečja.

Matematično upanje in varianca vzorčnega povprečja

Definicija 1.6.4 : *Označimo vzorec velikosti n z $(X_1, X_2, ..., X_n)$. Vzorčno povprečje je enako*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

in pomeni oceno za povprečje populacije. Kot oceno za celoto populacije bomo vzeli

$$T = N\overline{X}.$$

Lastnosti T sledijo iz lastnosti $\overline{X}$. Ker je vsak X_i slučajna spremenljivka, je tudi $\overline{X}$ slučajna spremenljivka; njena verjetnostna porazdelitev se imenuje vzorčna porazdelitev. V splošnem je vsaka številska ali statistična vrednost, dobljena iz slučajnega vzorca, slučajna spremenljivka in ima pripadajočo vzorčno porazdelitev. Vzorčna porazdelitev slučajne spremenljivke $\overline{X}$ določa, kako natančno le-ta oceni μ.
Naslednja lema velja za vzorčenje z ali brez vračanja.

Lema 1.6.1 : *Naj bodo $\zeta_1, \zeta_2, ..., \zeta_m$ različne vrednosti članov populacije in naj bo število članov populacije, ki imajo vrednost ζ_j, enako $n_j, j = 1, 2, ..., m$. Tedaj je X_i diskretna slučajna spremenljivka z verjetnostno funkcijo*

$$P(X_i = \zeta_j) = \frac{n_j}{N}$$

in velja

$$E(X_i) = \mu, \quad var(X_i) = \sigma^2.$$

Dokaz : Edine možne vrednosti, ki jih lahko X_i zavzame, so $\zeta_1, \zeta_2, ..., \zeta_m$. Ker je vsak član populacije z enako verjetnostjo i-ti član vzorca, je $P(X_i = \zeta_j) = \frac{n_j}{N}$. Matematično upanje slučajne spremenljivke X_i je tedaj

$$E(X_i) = \sum_{j=1}^{m} \zeta_j P(X_i = \zeta_j) = \frac{1}{N}\sum_{j=1}^{m} n_j \zeta_j = \mu.$$

Varianca pa

$$var(X_i) = E(X_i^2) - E(X_i)^2 = \frac{1}{N}\sum_{j=1}^{m} n_j \zeta_j^2 - \mu^2 = \sigma^2.$$

Tukaj smo uporabili dejstvo $\sum_{i=1}^{N} x_i^2 = \sum_{j=1}^{m} n_j \zeta_j^2$ in zvezo (1.22).

□

Za center vzorčne porazdelitve bomo vzeli $E(\overline{X})$, za njeno razpršenost pa standardno deviacijo vzorčnega povprečja $\overline{X}$. Pokazali bomo, da je center vzorčne porazdelitve pri μ in da je njena razpršenost obratno sorazmerna s kvadratnim korenom velikosti vzorca n.

Trditev 1.6.1 : *Pri enostavnem slučajnem vzorčenju velja*

a)
$$E(\overline{X}) = \mu$$

b)
$$E(T) = \tau$$

Dokaz :

a) Zaradi aditivnosti matematičnega upanja in enakosti $E(X_i) = \mu$ iz leme 1.6.1, lahko pišemo

$$E(\overline{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu.$$

b)
$$E(T) = E(N\overline{X}) = NE(\overline{X}) = N\mu = \tau$$

□

Poskušajmo poiskati še $var(\overline{X})$. Ker je $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, lahko pišemo

$$var(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_i, X_j). \tag{1.23}$$

Denimo, da gre za vzorčenje z vračanjem. Tedaj so slučajne spremenljivke X_i med seboj neodvisne in za $i \neq j$ velja $cov(X_i, X_j) = 0$. Vemo, da je $cov(X_i, X_i) = var(X_i) = \sigma^2$. Od tod sledi

$$var(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n} var(X_i) = \frac{\sigma^2}{n}. \tag{1.24}$$

Standardna deviacija slučajne spremenljivke $\overline{X}$ oz. standardna napaka je enaka

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}. \tag{1.25}$$

Pri vzorčenju brez vračanja moramo upoštevati odvisnost slučajnih spremenljivk X_i, kar zaplete sicer enostaven rezultat. Videli pa bomo, da je v primeru, ko je velikost vzorca n relativno majhna proti velikosti populacije N, odvisnost šibka in rezultat iz enakosti (1.25) služi kot dober približek.
Da bomo lahko izračunali varianco vzorčnega povprečja pri vzorčenju brez vračanja, moramo najprej poiskati $cov(X_i, X_j)$ za $i \neq j$.

Lema 1.6.2 : *Pri enostavnem slučajnem vzorčenju brez vračanja za $i \neq j$ velja*

$$cov(X_i, X_j) = -\frac{\sigma^2}{N-1}.$$

Dokaz : Zapišimo formulo za kovarianco

$$cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$$

in razvijmo prvo matematično upanje iz te formule:

$$E(X_i X_j) = \sum_{k=1}^{m} \sum_{l=1}^{m} \zeta_k \, \zeta_l \, P(X_i = \zeta_k, X_j = \zeta_l) =$$

$$= \sum_{k=1}^{m} \zeta_k \, P(X_i = \zeta_k) \sum_{l=1}^{m} \zeta_l \, P(X_j = \zeta_l | X_i = \zeta_k).$$

Poiščimo pogojno verjetnost v zgornjem izrazu

$$P(X_j = \zeta_l | X_i = \zeta_k) = \begin{cases} \frac{n_l}{N-1}; & k \neq l \\ \frac{n_l - 1}{N-1}; & k = l \end{cases}$$

Če pišemo

$$\frac{n_k - 1}{N-1} \, \zeta_k = \frac{n_k}{N-1} \, \zeta_k - \frac{\zeta_k}{N-1},$$

izraz za $E(X_i X_j)$ postane

$$\sum_{k=1}^{m} \zeta_k \, \frac{n_k}{N} \left(\sum_{l=1}^{m} \zeta_l \, \frac{n_l}{N-1} - \frac{\zeta_k}{N-1} \right) = \frac{1}{N(N-1)} \left(\tau^2 - \sum_{k=1}^{m} \zeta_k^2 \, n_k \right) =$$

$$= \frac{\tau^2}{N(N-1)} - \frac{1}{N(N-1)} \sum_{k=1}^{m} \zeta_k^2 \, n_k = \frac{N\mu^2}{N-1} - \frac{1}{N-1}(\mu^2 + \sigma^2) =$$

$$= \mu^2 - \frac{\sigma^2}{N-1}.$$

Nazadnje še odštejemo $E(X_i)E(X_j) = \mu^2$ od zadnje enačbe in dobimo

$$cov(X_i, X_j) = -\frac{\sigma^2}{N-1} \quad za \quad i \neq j.$$

$\square$

Vidimo, da je kovarianca $cov(X_i, X_j)$ za velike vrednosti N majhna.

Trditev 1.6.2 : *Pri slučajnem vzorčenju brez vračanja velja*

$$var(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{\sigma^2}{n}\left(1 - \frac{n-1}{N-1}\right).$$

Dokaz :

$$var(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{m} cov(X_i, X_j) = \frac{1}{n^2}\sum_{i=1}^{n} var(X_i) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} cov(X_i, X_j) =$$

$$= \frac{\sigma^2}{n} - \frac{1}{n^2}\, n\,(n-1)\,\frac{\sigma^2}{N-1} = \frac{\sigma^2}{n}\left(1 - \frac{n-1}{N-1}\right).$$

$\square$

Opazimo, da se varianca vzorca pri vzorčenju brez vračanja razlikuje od tiste pri vzorčenju z vračanjem za faktor

$$1 - \frac{n-1}{N-1}\,,$$

imenovanim tudi končna korekcija populacije (finite population correction). Ulomku $\frac{n}{N}$ pravimo delež vzorčenja (sampling fraction). Pogosto je delež vzorčenja zelo majhen in zato za standardno napako (standardno deviacijo) slučajne spremenljivke $\overline{X}$ velja

$$\sigma_{\overline{X}} \approx \frac{\sigma}{\sqrt{n}}.$$

Vidimo, da razpršenost vzorčne porazdelitve in zato točnost slučajne spremenljivke $\overline{X}$ določa velikost vzorca n in ne velikost populacije N. Poleg tega na natančnost vzorčnega povprečja vpliva tudi standardna deviacija σ populacije. Če je le-ta majhna, vrednosti populacije niso močno razpršene in majhen vzorec je skoraj točen. Če pa je razpršenost vrednosti velika, je potreben veliko večji vzorec, da zagotovimo isto natančnost.

Normalna aproksimacija porazdelitve vzorčnega povprečja $\overline{\mathbf{X}}$

Našli smo matematično upanje in varianco vzorčnega povprečja $\overline{X}$. Sedaj pa bi radi poiskali vzorčno porazdelitev, za kar moramo poznati populacijo. Z uporabo centralnega limitnega izreka bomo sklepali o aproksimaciji vzorčne porazdelitve - normalni ali Gaussovi aproksimaciji.
Naj bo $X_1, X_2, ...$ zaporedje neodvisnih, enako porazdeljenih slučajnih spremenljivk. Označimo matematično upanje in varianco

$$E(X_i) = \mu, \quad var(X_i) = \sigma.$$

Vzorčno povprečje spremenljivk $X_1, X_2, ..., X_n$ je enako

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Velja

$$E(\overline{X}_n) = \mu, \quad var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Centralni limitni izrek pravi, da za fiksno število z velja

$$P\left(\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq z\right) \longrightarrow F(z), \quad \text{ko } n \to \infty, \tag{1.26}$$

kjer je F porazdelitvena funkcija standardne normalne porazdelitve. Ideja vzorčenja in centralnega limitnega izreka nista čisto enaki - pri vzorčenju brez vračanja namreč slučajne spremenljivke X_i niso neodvisne. Poleg tega nima smisla, da n pošljemo proti neskončno, saj je velikost populacije N fiksna. Toda za druge limitne izreke je bilo dokazano, da sovpadajo s principom vzorčenja: če je n velik, vendar majhen relativno glede na N, je povprečje enostavnega slučajnega vzorca $\overline{X}_n$ približno normalno porazdeljeno.
Da pokažemo smiselnost uporabe centralnega limitnega izreka, bomo poskušali aproksimirati verjetnost $P(|\overline{X} - \mu| \leq \delta)$, da je napaka pri ocenjevanju μ z $\overline{X}$ manjša od neke konstante δ:

$$P(|\overline{X} - \mu| \leq \delta) = P(-\delta \leq \overline{X} - \mu \leq \delta) =$$

$$= P\left(-\frac{\delta}{\sigma_{\overline{X}}} \leq \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} \leq \frac{\delta}{\sigma_{\overline{X}}}\right) =$$

$$\approx F\left(\frac{\delta}{\sigma_{\overline{X}}}\right) - F\left(-\frac{\delta}{\sigma_{\overline{X}}}\right) = 2\,F\left(\frac{\delta}{\sigma_{\overline{X}}}\right) - 1.$$

Zadnjo enakost smo dobili z upoštevanjem $F(-z) = 1 - F(z)$, ki velja zaradi standardne normalne porazdelitve okrog 0.

Definirajmo še interval zaupanja.

Definicija 1.6.5 : *Interval zaupanja za parameter Θ populacije je iz vzorca izračunani naključni interval, ki vsebuje Θ z dano verjetnostjo.*

Na primer, 95% interval zaupanja za μ je naključni interval, ki vsebuje μ z verjetnostjo 0.95; če bi vzeli veliko število naključnih vzorcev in poiskali interval zaupanja za vsakega od njih, bi 95% teh intervalov vsebovalo μ. Če je verjetnost enaka $1 - \alpha$, se interval imenuje $100(1 - \alpha)$% interval zaupanja, količini $1 - \alpha$ pa pravimo stopnja zaupanja.

Poglavje 2

Osnove ravnanja s kreditnimi tveganji

2.1 Pričakovana izguba

Obresti, ki jih banka zaračuna ob vsakem kreditu, predstavljajo ceno denarja. V del te cene je vračunano tudi tveganje zaradi neplačevanja obveznosti. V obstoječih sistemih se obresti izračunava neodvisno od kvalitete kreditojemalca in so zato za vse enake, v kratkem pa bodo obresti lahko odvisne od tveganosti naložb. Novi kapitalski sporazum Basel II namreč predvideva izračunavanje pričakovane izgube, s pomočjo katere lahko določimo obresti za posameznega kreditojemalca glede na tveganost poslovanja z njim. Banka za zaščito pred izgubo uporabi koncept zavarovanja, kot ga poznamo pri avtomobilskem ali zdravstvenem zavarovanju. Izkušnje kažejo, da lahko celo dobri kreditojemalci propadejo in ne poravnajo svojih finančnih obveznosti do banke. Banka zato za vse kredite zaračuna primerne pribitke za tveganje in s tem ustvari rezervo za pričakovano izgubo. Ta rezerva je varovalo pred možnimi izgubami, nastalimi zaradi slabih kreditov, tj. kreditov, ki bodo vrnjeni z zamudo ali pa nikoli ne bodo odplačani.

V nadaljevanju bomo vpeljali verjetnostni prostor $(\Omega, \mathcal{F}, P)$, kjer je Ω prostor dogodkov, $\mathcal{F}$ je σ-algebra in P verjetnostna mera. Elementi σ-algebre $\mathcal{F}$ so merljivi dogodki, zato je smiselno zahtevati merljivost izostanka odplačila dolga. Navadno $\mathcal{F}$ identificiramo z dostopnimi informacijami, zato naj bi bila informacija o tem, ali kreditojemalec propade ali preživi, vključena v množico merljivih dogodkov.

Banka vsakemu svojemu kreditojemalcu pripiše verjetnost neplačila dolga p. Gledano z vidika matematike to pomeni, da definira Bernoullijevo slučajno spremenljivko D, ki pomeni nesposobnost poplačila dolga posameznega kreditojemalca in je porazdeljena takole:

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

Poleg tega določi še dve količini: izgubo ob nastopu izostanka odplačila dolga η (loss given default), ki predstavlja delež kredita, za katerega se pričakuje, da bo izgubljen v primeru nesposobnosti poplačila dolga, in izpostavo ob nastanku plačilne nesposobnosti kreditojemalca ν (exposure at default), ki predstavlja celoten znesek kredita.
Tako dobimo slučajno spremenljivko L, ki pomeni izgubo terjatve do posameznega kreditojemalca in je definirana takole:

$$L = \nu \cdot \eta \cdot D. \tag{2.1}$$

Sedaj lahko definiramo pričakovano izgubo EL (expected loss) posameznega kreditojemalca kot matematično upanje pripadajoče slučajne spremenljivke L:

$$EL = E(L) = \nu \cdot \eta \cdot E(D) = \nu \cdot \eta \cdot p. \tag{2.2}$$

Za tak zapis matematičnega upanja je potrebna predpostavka, da sta ν in η konstantni vrednosti, npr. matematični upanji nekih slučajnih spremenljivk, ki vplivajo na spremembe (underlying random variables). Če to ni res, moramo privzeti, da so slučajne spremenljivke ν, η in D neodvisne. Vendar je tudi predpostavka o neodvisnosti vprašljiva in v splošnem zelo poenostavljajoča. Lahko rečemo, da je enakost (2.2) najbolj enostavna formula matematičnega upanja. Od zdaj naprej velja dogovor, da je ν vedno deterministična, medtem ko velikost izgube ob nastopu izostanka odplačila dolga S (severity of loss) razumemo kot slučajno spremenljivko z matematičnim upanjem η. Zaradi enostavnosti v tem poglavju privzemimo, da sta S in L neodvisni.

2.1.1 Verjetnost neplačila dolga

Naloga pripisovanja verjetnosti neplačila dolga posameznemu kreditojemalcu kreditnega portfelja banke še zdaleč ni lahka. Oceniti moramo kreditno sposobnost kreditojemalca in ga na podlagi te ocene razvrstiti v ustrezni bonitetni razred. To je t.i. rating, ki lahko temelji na kvalitativni oceni kreditnih analitikov ali pa na kvantitativnem modelu za razvrščanje v bonitetne razrede.

Kalibracija ali umeritev

Kalibracija je proces pripisovanja verjetnosti neplačila dolga posameznemu bonitetnemu razredu. Če množico vseh bonitetnih razredov označimo z $\mathcal{R}$, posamezni bonitetni razred pa z R, lahko definiramo preslikavo:

$$\mathcal{R} \mapsto [0,1], \quad R \mapsto P(R).$$

V nadaljevanju bomo pokazali, kako lahko na enostaven način poiščemo tako preslikavo:
Označimo s $h_i(R)$ povprečno frekvenco izostanka odplačila dolga bonitetnega razreda R v letu i, za $i = 1, ..., n$. Potem lahko izračunamo povprečno vrednost in standardno deviacijo teh frekvenc v vseh letih za posamezni bonitetni razred R:

$$m(R) = \frac{1}{n}\sum_{i=1}^{n} h_i(R)$$

in

$$s(R) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(h_i(R) - m(R))^2}$$

Povprečna vrednost $m(R)$ bonitetnega razreda R je naš prvi približek za verjetnost neplačila dolga, pripisano bonitetnemu razredu R. Standardna deviacija $s(R)$ pomeni odstopanje in nam pove, kolikšno napako naredimo, če imamo $m(R)$ za dobro oceno verjetnosti neplačila dolga kreditojemalca iz bonitetnega razreda R.
Nato nanesemo povprečne vrednosti $m(R)$ v koordinatni sistem, v katerem x-os predstavlja bonitetne razrede. Na podlagi različnih empiričnih študij o nesposobnosti poplačila dolga lahko sklepamo, da frekvence nesposobnosti poplačila dolga naraščajo eksponentno s padajočo kreditno sposobnostjo. Na logaritemski skali, ki je zaradi narave podatkov precej pogosta, to pomeni linearno naraščanje. S pomočjo standardne regresijske teorije poiščemo funkcijo, ki se najbolje prilega danim podatkom. Nazadnje uporabimo vrednost dobljene regresijske funkcije za oceno verjetnosti neplačila dolga p, pripisano posameznemu bonitetnemu razredu.

2.1.2 Izguba ob nastopu izostanka odplačila dolga

Izgubo ob nastopu izostanka odplačila dolga η definiramo kot

$$\eta = 1 - \text{stopnja poplačila}$$

in predstavlja delež izgube, ki jo bo banka v primeru izostanka odplačila dolga v resnici utrpela. Ocenjevanje tega deleža ni lahka naloga, saj je stopnja poplačila odvisna od mnogih faktorjev, npr. od kvalitete sredstev, ki jamčijo za kredit (hipoteke, obveznice, garantna pisma...), in od starosti terjatev banke do kreditojemalca. η razumemo kot matematično upanje velikosti izgube S ob nesposobnosti plačila dolga

$$E(S) = \eta.$$

2.2 Nepričakovana izguba

Za banko je še pomembnejši od pričakovane izgube pojav nepričakovane izgube, ki je posledica dveh lastnosti portfelja: korelacije med komitenti in koncentracije sredstev. Nepričakovana izguba UL (unexpected loss) je znesek, ki prekorači povprečno pričakovano izgubo in je definirana kot standardna deviacija slučajne spremenljivke L:

$$UL = \sqrt{var(L)} = \sqrt{var(\nu\, S\, D)}. \tag{2.3}$$

Trditev 2.2.1 : *Predpostavimo, da sta izostanek odplačila dolga D in velikost izgube ob nastopu izostanka plačila dolga S neodvisna. Tedaj velja:*

$$UL = \nu \sqrt{var(S)\, p + \eta^2\, p\, (1 - p)}.$$

Dokaz :
Računajmo:

$$\begin{aligned} UL &= \nu \sqrt{var(SD)} = \\ &= \nu \sqrt{E((SD)^2) - E(SD)^2} = \\ &= \nu \sqrt{E(S^2)\, E(D^2) - E(S)^2\, E(D)^2} = \\ &= \nu \sqrt{(var(S) + E(S)^2)\, p - \eta^2\, p^2} = \\ &= \nu \sqrt{var(S)\, p + \eta^2\, p - \eta^2\, p^2} = \\ &= \nu \sqrt{var(S)\, p + \eta^2\, p\, (1 - p)} \end{aligned}$$

□

Opomba: Privzetek o neodvisnosti izostanka plačila dolga in velikosti izgube ni vedno realističen. Uporabimo ga samo zato, da dobimo prvo aproksimacijo za pravo nepričakovano izgubo. Pravzaprav je precej verjetno, da v povprečju stopnja poplačil pade, če slabe ekonomske razmere inducirajo porast frekvenc izostanka odplačila dolga.

Do sedaj smo se ukvarjali s kreditnim tveganjem pri enem samem kreditojemalcu. Zdaj pa bomo pogledali modeliranje celotne izgube kreditnega portfelja.
Za ta namen potrebujemo portfelj, sestavljen iz m kreditov, kjer, kot prej, definiramo izgubo terjatve do posameznega kreditojemalca

$$L_i = \nu_i \, S_i \, D_i \tag{2.4}$$

s $P(D_i = 1) = p_i$.
Portfeljska izguba je tedaj definirana kot slučajna spremenljivka

$$L_{PF} = \sum_{i=1}^{m} L_i = \sum_{i=1}^{m} \nu_i \, S_i \, D_i. \tag{2.5}$$

Potrebujemo še količini EL_{PF} in UL_{PF}, ki sta analogni EL in UL in pomenita matematično upanje oz. standardno deviacijo portfeljske izgube. Za EL_{PF} lahko uporabimo aditivnost matematičnega upanja in tako dobimo:

$$EL_{PF} = \sum_{i=1}^{m} EL_i = \sum_{i=1}^{m} \nu_i \, \eta_i \, p_i. \tag{2.6}$$

Za UL_{PF} aditivnost velja, če so slučajne spremenljivke L_i paroma nekorelirane. Žal to ni res, korelacija je namreč glavna lastnost kreditnega tveganja. Nepričakovana izguba portfelja UL_{PF} je prva količina, kjer korelacije oz. kovariance igrajo glavno vlogo:

$$UL_{PF} = \sqrt{var(L_{PF})} \tag{2.7}$$

$$= \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{m} \nu_i \, \nu_j \, cov(S_i D_i \, , S_j D_j)}.$$

V posebnem primeru, ko so velikosti izgube S_i konstantne, lahko nepričakovano izgubo portfelja UL_{PF} izrazimo s pomočjo korelacij med izostanki odplačila:

Trditev 2.2.2 : *Za portfelj s konstantnimi velikostmi izgube S_i velja*

$$UL_{PF}^2 = \sum_{i,j=1}^{m} \nu_i \, \nu_j \, \eta_i \, \eta_j \, \sqrt{p_i \, (1-p_i) \, p_j \, (1-p_j)} \, \rho_{i,j},$$

kjer je $\rho_{i,j} = corr(D_i \, , D_j) = \frac{cov(D_i \, , D_j)}{\sqrt{var(D_i) \, var(D_j)}}$ korelacijski koeficient med izostankoma odplačil i-tega in j-tega kreditojemalca.

Preden nadaljujemo, bi radi razmislili o pomenu in interpretaciji korelacije. Za začetek vzemimo portfelj dveh kreditov z $\eta = 100\%$ in $\nu = 1$. Opravka imamo torej samo z D_i za $i = 1, 2$. Označimo še $\rho = corr(D_i \, , D_j)$ in $p_i = P(D_i = 1)$. Tedaj velja

$$UL_{PF}^2 = p_1 \, (1-p_1) + p_2 \, (1-p_2) + 2 \, \rho \, \sqrt{p_1 \, (1-p_1)} \, \sqrt{p_2 \, (1-p_2)}. \tag{2.8}$$

Ločimo tri primere glede na korelacijski koeficient izostankov odplačil ρ:

$\rho = 0$: V tem primeru tretji izraz v enakosti (2.8) izgine. Lahko rečemo, da $\rho = 0$ pomeni popolno razpršenost. Koncept le-te je lahko razložiti: investiranje v veliko različnih sredstev v splošnem zmanjša portfeljsko tveganje, ker ni zelo verjetno, da bi bilo hkrati veliko število kreditojemalcev nesposobno odplačati dolg. Manj, kot imajo krediti v portfelju skupnega, manjša je verjetnost, da izostanek odplačila dolga enega kreditojemalca bistveno vpliva na ekonomsko prihodnost ostalih kreditov v portfelju. Primer $\rho = 0$ se nanaša na situacijo, ko so krediti v portfelju popolnoma nepovezani. Če interpretiramo UL_{PF} kot mero portfeljskega tveganja, vidimo, da ta primer minimizira tveganje skupne nesposobnosti poplačila.

$\rho > 0$: Kreditojemalca sta povezana: nesposobnost poplačila dolga enega kreditojemalca poveča verjetnost neplačila drugega. Poglejmo si pogojno verjetnost neplačila drugega kreditojemalca pod pogojem, da je prvi kreditojemalec nezmožen poravnati svoje obveznosti:

$$P(D_2 = 1 | D_1 = 1) = \frac{P(D_1 = 1, D_2 = 1)}{P(D_1 = 1)} = \frac{E(D_1 D_2)}{p_1} \tag{2.9}$$

$$= \frac{p_1 \, p_2 + cov(D_1, D_2)}{p_1} = p_2 + \frac{cov(D_1, D_2)}{p_1}.$$

Vidimo, da pozitivna korelacija oz. kovarianca vodi k večji pogojni verjetnosti neplačila, kot je nepogojna verjetnost neplačila p_2 drugega kreditojemalca. Z drugimi besedami, v primeru pozitivne korelacije vsak izostanek odplačila dolga vpliva na ostale kredite v portfelju, namreč poveča se število izgub. Robni primer je primer popolne korelacije ($\rho = 1$). Naj velja $p = p_1 = p_2$. Tedaj enakost (2.8) postane $UL_{PF} = 2\sqrt{p\,(1-p)}$. To pomeni, da je portfeljsko tveganje odvisno samo od enega kreditojemalca, vendar z dvojno jakostjo (govorimo o t.i. koncentraciji tveganja). Izostanek odplačila enega kreditojemalca vodi k izostanku odplačila drugega kreditojemalca skoraj gotovo.

$\rho < 0$: To je zrcalna situacija primera $\rho > 0$. Zato si bomo ogledali samo robni primer negativne korelacije ($\rho = -1$). Če se značilnosti obeh kreditov ujemajo, lahko na investicijo v sredstvo 1 gledamo kot na skoraj popolno zaščito pred investicijo v sredstvo 2. Iz enakosti (2.8) sledi, da v primeru popolne zaščite nepričakovana izguba portfelja UL_{PF} popolnoma izgine ($UL_{PF} = 0$). To pomeni, da popolna zaščita (investiranje v sredstvo 2 pri korelaciji $\rho = -1$) v celoti izniči tveganje zaradi investicije v sredstvo 1.

2.2.1 Ekonomski kapital

Portfeljska izguba ima v levo ukrivljeno porazdelitev. Njena osnovna značilnost so predvsem majhne verjetnosti velikih izgub, ki pa so kritične za varnost oz. solventnost banke. Zato je potrebno definirati ustrezno velikost kapitala, s katerim bi bila sposobna pokriti potencialno izgubo, ki bi lahko povzročila stečaj banke.
Najbolj pogost način za določanje kapitala za tveganje (risk capital) je koncept ekonomskega kapitala EC (economic capital). Za vnaprej predpisano stopnjo zaupanja α je definiran kot

$$EC_\alpha = q_\alpha - EL_{PF}, \tag{2.10}$$

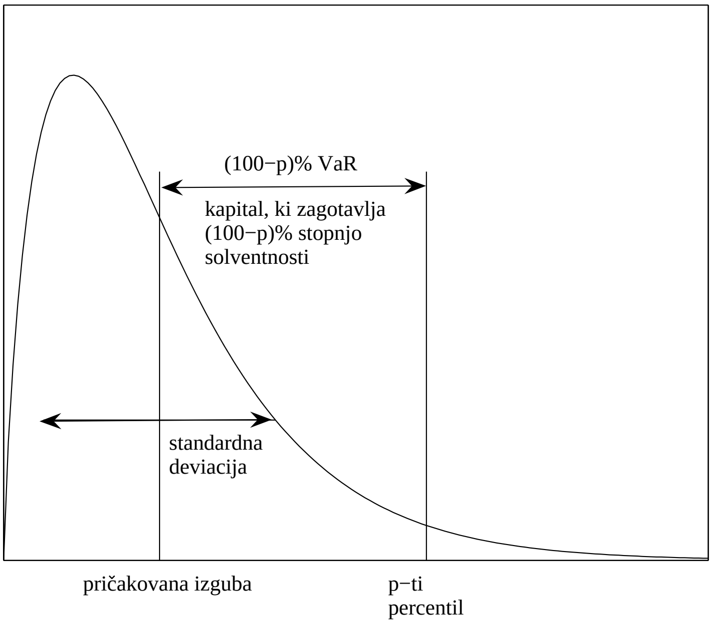
kjer je q_α[1] α-kvantil portfeljske izgube L_{PF}, določen z

$$q_\alpha = \inf\{q > 0; P(L_{PF} \le q) \ge \alpha\}. \tag{2.11}$$

2.2.2 Porazdelitev izgube

Porazdelitev portfeljske izgube igra glavno vlogo pri ravnanju s kreditnimi tveganji. Na sliki 2.1 lahko vidimo glavne količine, povezane s tveganjem kreditnega portfelja. V praksi obstajata dva načina, s katerima generiramo porazdelitev izgube: prva metoda temelji na simulaciji Monte Carlo, druga pa na analitični aproksimaciji.

[1]Za q_α uporabljamo izraz tvegana vrednost (Value-at-Risk ali krajše VaR)



Slika 2.1: Porazdelitev izgube

Simulacija izgub Monte Carlo

Pri simulaciji Monte Carlo izgube simuliramo in prikažemo v obliki histograma, da dobimo empirično porazdelitev izgube danega portfelja. Empirično porazdelitev določimo na naslednji način:
Privzemimo, da smo simulirali n potencialnih portfeljskih izgub $L_{PF}^{(1)}, ..., L_{PF}^{(n)}$, upoštevajoč porazdelitve posameznih slučajnih spremenljivk izgube in korelacije med njimi. Tedaj je empirična porazdelitev izgube enaka

$$F(x) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}_{[0,x]} \, (L_{PF}^{(j)}). \tag{2.12}$$

Iz empirične porazdelitve izgube lahko izpeljemo količine, povezane s portfeljskim tveganjem, o katerih smo govorili v prejšnjih podrazdelkih. Na primer, α-kvantil izgube lahko dobimo neposredno iz naših simulacijskih rezultatov $LP_{PF}^{(1)}, ..., LP_{PF}^{(n)}$ na naslednji način:
Poiščemo vrstilne statistike slučajnih spremenljivk $LP_{PF}^{(1)}, ..., LP_{PF}^{(n)}$. Velja

$$L_{PF}^{(i_1)} \leq L_{PF}^{(i_2)} \leq ... \leq L_{PF}^{(i_n)}.$$

α-kvantil $\widehat{q_\alpha}$ empirične porazdelitve izgub (za poljubno stopnjo zaupanja α) je tedaj

$$\widehat{q_\alpha} = \begin{cases} \alpha L_{PF}^{(i_{[n\alpha]})} + (1-\alpha)\, L_{PF}^{(i_{[n\alpha]+1})}; & n\alpha \notin \mathbb{N} \\ L_{PF}^{(i_{[n\alpha]})}; & n\alpha \in \mathbb{N} \end{cases} \tag{2.13}$$

Ekonomski kapital izračunamo po formuli

$$\widehat{EC_\alpha} = \widehat{q_\alpha} - \frac{1}{n}\sum_{j=1}^{n} LP_{PF}^{(j)} \tag{2.14}$$

Analogno lahko dobimo vse preostale količine, tako da izračunamo ustrezne empirične statistike.

Analitična aproksimacija

Dejanski portfelj z neznano porazdelitvijo izgube preslikamo v ekvivalenten portfelj z znano porazdelitvijo izgube. Porazdelitev izgube ekvivalentnega portfelja potem jemljemo kot nadomestilo za dejansko porazdelitev izgube originalnega portfelja. V praksi se to pogosto izvaja na naslednji način: izberemo družino porazdelitev, določenih s prvim in drugim momentom, ki so tipično ukrivljene in imajo debel rep[2]. Iz znanih značilnosti originalnega portfelja izračunamo prvi (EL) in drugi moment (UL^2). EL originalnega portfelja lahko dobimo iz informacij o ratingu, kreditni izpostavljenosti in porazdelitvi izgube ob nastopu izostanka plačila dolga. Drugega momenta pa ne moremo izračunati brez nekaterih privzetkov v zvezi s korelacijo med izostanki plačila v portfelju. Pogosto uganemo medsebojno odvisnost naložb in iz nje izračunamo pripadajočo korelacijo med izostanki odplačila. Glede na tako določeni prvi in drugi moment lahko iz parametriziane družine porazdelitev izgub izberemo tisto, ki najbolj ustreza originalnemu portfelju. To porazdelitev potem interpretiramo kot porazdelitev izgube ekvivalentnega portfelja, ki je bila izbrana s postopkom ujemanja momentov. Oglejmo si primer:

Denimo, da želimo porazdelitev izgube originalnega portfelja aproksimirati z beta porazdelitvijo, ki se ujema z originalnim portfeljem v prvem in drugem momentu. Z drugimi besedami, iščemo slučajno spremenljivko

$$X \sim Beta(a, b),$$

[2] V naši terminologiji ima porazdelitev debel rep, če so njeni kvantili pri visoki stopnji zaupanja višji kot kvantili normalne porazdelitve z istim prvim in drugim momentom.

ki predstavlja portfeljsko izgubo, tako da parametra a in b rešita naslednji enačbi:

$$E(X) = \frac{a}{a+b} \quad \text{in} \tag{2.15}$$

$$var(X) = \frac{ab}{(a+b)^2\,(a+b+1)}$$

Pri analitični aproksimaciji vzamemo slučajno spremenljivko X kot približek za neznano porazdelitev izgube portfelja, s katerim smo začeli. Če sledimo temu privzetku, lahko vse količine originalnega portfelja, ki so povezane s tveganjem, aproksimiramo z ustreznimi količinami slučajne spremenljivke X. Ker pravo porazdelitev izgube nadomestimo z neko znano porazdelitvijo, so vsi potrebni izračuni narejeni sorazmeroma hitro. Cena, ki jo plačamo za to je, da so vsi izračuni močno odvisni od izbire modela. podvrženost vseh izračunov močnemu modelskemu tveganju. Porazdelitev izgube ima obliko beta porazdelitve. Toda obstaja več različnih dvoparametrskih družin gostot, ki ustrezajo tipični obliki porazdelitve izgube, npr. nekatere gama porazdelitve, F-porazdelitev in ostale. Žal se razlikujejo v repih, tako da, če bi ena od njih res dobro aproksimirala neznano porazdelitev izgube, bi bile vse ostale napačna izbira. Zato je izbira primerne družine porazdelitev za analitično aproksimacijo pomemben vir modelskega tveganja.
Tehnike analitične aproksimacije se da uspešno uporabljati pri t.i. homogenih portfeljih (to so portfelji, v katerih imajo vse transakcije primerljive značilnosti tveganja, npr. ni koncentracije sredstev, verjetnost neplačila dolga v razredu z zmerno širino, ena sama država in vrsta industrije itd.).

Čeprav je simulacija Monte Carlo v primerjavi z analitično aproksimacijo zelo dolgotrajna (pri velikih portfeljih lahko traja več ur), je veliko bolj zanesljiva, saj ne naredi toliko privzetkov in natančno upošteva korelacije znotraj portfelja. Celo več, simulacija Monte Carlo upošteva vse različne karakteristike tveganja kreditov v portfelju.

Poglavje 3

Modeliranje koreliranih izostankov odplačil dolga

V tem poglavju bomo obravnavali portfeljsko izgubo in korelacijo med izostanki plačil v portfelju. Privzeli bomo časovno obdobje enega leta. Denimo, da imamo kreditni portfelj z m kreditojemalci, tako da i-ti kreditojemalec spada v bonitetni razred R_i. S pomočjo kalibracije dobimo, da je njegova verjetnost neplačila dolga enaka p_i. V enem letu se bonitetni razred kreditojemalca lahko spremeni glede na spremembo njegove kreditne kvalitete. Bolj formalno, označimo množico možnih bonitetnih razredov z $\{0, ..., d\}$, kjer $d \in \mathbb{N}$ pomeni stanje nesposobnosti poplačila dolga,

$$R_i \in \{0, ..., d\} \quad \text{in} \quad p_i = P(R_i \to d).$$

Oznaka $R \to R'$ pomeni spremembo iz bonitetnega razreda R v bonitetni razred R' v obdobju enega leta.
V tem poglavju se bomo osredotočili na primer dveh stanj, kar pomeni, da se bomo omejili na

$$d = 1, \quad D_i = R_i \in \{0, 1\}, \quad p_i = P(D_i = 1).$$

Model dveh stanj ne omogoča možnosti spreminjanja bonitetnih razredov; govorimo samo o nezmnožnosti ali zmožnosti odplačila dolga.
V 2. poglavju smo definirali slučajno spremenljivko izgube kot indikator izostanka odplačila dolga. V kontekstu modela dveh stanj je najbolj naraven izbor Bernoullijeva slučajna spremenljivka. Drug uveljavljen pristop je modeliranje izostanka odplačila dolga s Poissonovo slučajno spremenljivko.
Ker do neke mere obstajajo relacije in skupen izvor obeh pristopov, se je večkrat pojavila ideja, da bi združili Bernoullijev in Poissonov model. Toda modela nista v celoti primerljiva, saj ustrezna mešana modela (fiksne parametre porazdelitve nadomestimo s slučajnimi spremenljivkami in na ta način v

modela vpeljemo korelacijo) generirata porazdelitvi izgub z opaznimi razlikami v repih.
V nadaljevanju bo veljalo: Bernoullijeve slučajne spremenljivke bomo označevali z L, Poissonove pa z L'.

3.1 Bernoullijev model

Vektor slučajnih spremenljivk $\mathbf{L} = (L_1, ..., L_m)$ je (Bernoullijeva) statistika izgube, če so vse robne porazdelitve $\mathbf{L}$ Bernoullijeve:

$$L_i \sim Bernoulli(p_i), \quad \text{tj.} \quad L_i \sim \begin{pmatrix} 1 & 0 \\ p_i & 1 - p_i \end{pmatrix}.$$

Absolutna izguba oz. relativna izguba sta definirani kot

$$L = \sum_{i=1}^{m} L_i \quad \text{oz.} \quad \frac{L}{m}.$$

Spremenljivko L imenujemo portfeljska izguba ne glede na to, ali mislimo na absolutno ali relativno vrednost.
Za začetek si oglejmo enostaven primer neodvisnih izostankov odplačil dolga, ki pa ni zelo realističen. Privzemimo enotno verjetnost neplačila dolga p in neodvisnost med kreditojemalci:

$$L_i \sim Bernoulli(p) \quad \text{in} \quad \{L_i\}_{i=1,...,m} \quad \text{neodvisne.}$$

V tem primeru za absolutno portfeljsko izgubo L velja $L \sim Bin(m, p)$.
Če dopustimo različne verjetnosti neplačila dolga p_i, še vedno pa velja, da so kreditojemalci med seboj neodvisni, spet dobimo portfeljsko izgubo L kot konvolucijo posameznih spremenljivk izgube. Matematično upanje in varianca sta enaka:

$$E(L) = \sum_{i=1}^{m} p_i \quad \text{in} \quad var(L) = \sum_{i=1}^{m} p_i\,(1 - p_i). \tag{3.1}$$

Če bi pri ravnanju s kreditnimi tveganji privzeli neodvisnost kreditojemalcev danega portfelja, bi lahko glede na centralni limitni izrek vsaj pri velikih portfeljih sklepali, da je izguba približno Gaussova spremenljivka. Žal neodvisnosti izgub ne smemo pričakovati, saj je pri modeliranju kreditnih tveganj korelacija glavni izziv. Zato nadaljujmo z bolj realistično izvedbo statistike izgube.

3.1.1 Splošen Bernoullijev mešani model

Kot prej dobimo portfeljsko izgubo iz statistike izgube $\mathbf{L} = (L_1, ..., L_m)$ z Bernoullijevimi slučajnimi spremenljivkami $L_i \sim Bernoulli(P_i)$. Toda tokrat so verjetnosti izgube slučajne spremenljivke $\mathbf{P} = (P_1, ..., P_m)$ z neko porazdelitveno funkcijo $\mathbf{F}$ z nosilcem na $[0,1]^m$. Dodatno privzamemo, da so slučajne spremenljivke $L_1, ..., L_m$ pogojno na realizacijo $\mathbf{p} = (p_1, ..., p_m)$ vektorja $\mathbf{P}$ neodvisne. Matematično to pogojno neodvisnost zapišemo takole:

$$L_i|\{P_i = p_i\} \sim Bernoulli(p_i) \quad , \quad \{L_i|\{\mathbf{P} = \mathbf{p}\}\}_{i=1,...,m} \quad \text{neodvisne.}$$

Večrazsežno porazdelitev tedaj dobimo takole:

$$P(L_1 = l_1, ..., L_m = l_m) = \tag{3.2}$$

$$= \int_{[0,1]^m} P(L_1 = l_1, ..., L_m = l_m|\mathbf{P} = \mathbf{p})\, d\mathbf{F}(p_1, ..., p_m) =$$

$$= \int_{[0,1]^m} P(L_1 = l_1|\mathbf{P} = \mathbf{p}) ... P(L_m = l_m|\mathbf{P} = \mathbf{p})\, d\mathbf{F}(p_1, ..., p_m) =$$

$$= \int_{[0,1]^m} \prod_{i=1}^{m} p_i^{l_i} (1 - p_i)^{1-l_i}\, d\mathbf{F}(p_1, ..., p_m),$$

kjer je $l_i \in \{0, 1\}$.

Trditev 3.1.1 : *Matematično upanje in varianca posamezne izgube L_i sta enaka*

$$E(L_i) = E(P_i),$$
$$var(L_i) = E(P_i)(1 - E(P_i)) \qquad (i = 1, ..., m).$$

Dokaz : Prvo enakost pokažemo takole:

$$E(L_i) = E(E(L_i|\mathbf{P}))$$
$$E(L_i|\mathbf{P} = \mathbf{p}) = p_i \quad \text{in zato} \quad E(E(L_i|\mathbf{P})) = E(P_i).$$

Drugo pokažemo podobno:

$$var(L_i) = var(E(L_i|\mathbf{P})) + E(var(L_i|\mathbf{P})) = \tag{3.3}$$

$$= var(P_i) + E(P_i\,(1 - P_i)) = E(P_i)(1 - E(P_i)).$$

□

Kovarianca med posameznimi izgubami je enaka:

$$cov(L_i, L_j) = E(L_i L_j) - E(L_i)E(L_j) = cov(P_i, P_j) \tag{3.4}$$

Da dokažemo to enakost, je dovolj videti

$$E(L_i L_j) = E(P_i P_j),$$

saj po trditvi 3.1.1 velja

$$E(L_i) = E(P_i).$$

Torej:

$$E(L_i L_j) = P(L_i = 1, L_j = 1) = \int_{[0,1]^m} P(L_i = 1, L_j = 1 | \mathbf{P} = \mathbf{p})\, d\mathbf{F}(p_1, ..., p_m) =$$

$$= \int_{[0,1]^m} P(L_i = 1 | \mathbf{P} = \mathbf{p})\, P(L_j = 1 | \mathbf{P} = \mathbf{p})\, d\mathbf{F}(p_1, ..., p_m) =$$

$$= \int_{[0,1]^m} p_i\, p_j\, d\mathbf{F}(p_1, ..., p_m) = E(P_i P_j).$$

Korelacijo med izostanki odplačil v Bernoullijevem mešanem modelu lahko tako zapišemo kot

$$corr(L_i, L_j) = \frac{cov(P_i, P_j)}{\sqrt{E(P_i)(1 - E(P_i))}\sqrt{E(P_j)(1 - E(P_j))}} \tag{3.5}$$

Enakosti (3.4) in (3.5) povesta, da je odvisnost med izgubami v portfelju določena s kovariančno strukturo večrazsežne porazdelitve $\mathbf{F}$ vektorja $\mathbf{P}$.

3.1.2 Enotna verjetnost neplačila dolga in enotna korelacija

Za portfelje, kjer so približno vsi krediti enako veliki in v smislu tveganja istega tipa, je smiselno privzeti enotno verjetnost neplačila dolga in enotno korelacijo med kreditojemalci. Takim portfeljem potem rečemo enotni portfelji in so popolni kandidati za analitično aproksimacijo.
Privzetek o enotnosti vodi v zamenljivost[1] Bernoullijevih spremenljivk $L_i \sim Bernoulli(P)$ s slučajno verjetnostjo neplačila dolga P, kjer je F njena porazdelitvena funkcija z nosilcem na $[0, 1]$. Privzamemo pogojno neodvisnost

[1] Velja $(L_1, ..., L_m) \sim (L_{\Pi(1)}, ..., L_{\Pi(m)})$ za poljubno permutacijo Π.

spremenljivk L_i kot v splošnem primeru.
Večrazsežna porazdelitev je tedaj

$$P(L_1 = l_1, ..., L_m = l_m) = \int_0^1 p^k (1-p)^{m-k} \, dF(p), \tag{3.6}$$

kjer je

$$k = \sum_{i=1}^{m} l_i \quad \text{in} \quad l_i \in \{0, 1\}.$$

Verjetnost, da se zgodi natanko k neplačil dolga, je dana z

$$P(L = k) = \binom{m}{k} \int_0^1 p^k (1-p)^{m-k} \, dF(p). \tag{3.7}$$

Enotna verjetnost neplačila dolga je seveda enaka

$$\bar{p} = P(L_i = 1) = E(L_i) = \int_0^1 p \, dF(p) \tag{3.8}$$

in enotna korelacija dveh različnih kreditojemalcev je enaka

$$corr(L_i, L_j) = \frac{P(L_i = 1, L_j = 1) - \bar{p}^2}{\bar{p}\,(1-\bar{p})} = = \frac{\int_0^1 p^2 \, dF(p) - \bar{p}^2}{\bar{p}\,(1-\bar{p})}, \tag{3.9}$$

kar ima za posledico

$$corr(L_i, L_j) = \frac{var(P)}{\bar{p}\,(1-\bar{p})}. \tag{3.10}$$

To nam pove, da večja spremenljivost verjetnosti P pomeni večjo korelacijo med izostanki plačil. Poleg tega lahko sklepamo tudi, da je odvisnost med spremenljivkami L_i pozitivna ali enaka 0, ker je varianca vedno nenegativna. Primer $corr(L_i, L_j) = 0$ se zgodi natanko tedaj, ko je $var(P) = 0$, kar pomeni, da P sploh ni naključna. V tem primeru je F Diracova mera, skoncentrirana v $\bar{p}$, absolutna portfeljska izguba L pa je porazdeljena binomsko z verjetnostjo neplačila dolga $\bar{p}$.
Drug robni primer, $corr(L_i, L_j) = 1$, implicira "rigidno"obnašanje posameznih izgub v portfelju: bodisi nobeden od kreditojemalcev ne izpolnjuje prevzetih in dogovorjenih pogodbenih obveznosti bodisi te obveznosti izpolnjujejo vsi kreditojemalci. Ustrezna porazdelitev F je tedaj Bernoullijeva porazdelitev, tako da velja: $P(P = 1) = \bar{p}$ in $P(P = 0) = 1 - \bar{p}$.

3.2 Poissonov model

Izostanek odplačila dolga i-tega kreditojemalca, $i = 1, ..., m$, modeliramo s Poissonovimi slučajnimi spremenljivkami

$$L_i^{'} \sim Poisson(\lambda_i), \quad L_i^{'} \in \{0, 1, 2, ...\}, \quad p_i = P(L_i^{'} \geq 1), \tag{3.11}$$

kjer p_i spet pomeni verjetnost neplačila i-tega kreditojemalca. Opazimo, da Poissonova porazdelitev dovoljuje več izostankov plačil posameznega kreditojemalca. Verjetnost dogodka, da i-ti kreditojemalec izostane s plačilom več kot enkrat, je dana z

$$P(L_i^{'} \geq 2) = 1 - e^{-\lambda_i}(1 + \lambda_i)$$

in je ponavadi majhno število.
Parameter λ_i je običajno blizu verjetnosti p_i, saj velja

$$p_i = P(L_i^{'} \geq 1) = 1 - e^{-\lambda_i} \approx \lambda_i \tag{3.12}$$

za majhne vrednosti λ_i.
Vemo: če sta slučajni spremenljivki $L_1^{'} \sim Poisson(\lambda_1)$, $L_2^{'} \sim Poisson(\lambda_2)$ neodvisni, potem za vsoto velja $L_1^{'} + L_2^{'} \sim Poisson(\lambda_1 + \lambda_2)$. Torej je, če privzamemo neodvisnost, skupno število izgub v portfelju dano z

$$L^{'} = \sum_{i=1}^{m} L_i^{'} \sim Poisson\Big(\sum_{i=1}^{m} \lambda_i\Big). \tag{3.13}$$

Kot prej korelacijo vpeljemo v mešanem modelu.

3.2.1 Splošen Poissonov mešani model

Statistika izgube je tokrat slučajni vektor $\mathbf{L}^{'} = (L_1^{'}, ..., .L_m^{'})$ Poissonovih slučajnih spremenljivk $L_i^{'} \sim Poisson(\lambda_i)$, kjer je $\mathbf{\Lambda} = (\lambda_1, ..., \lambda_m)$ slučajni vektor s porazdelitveno funkcijo $\mathbf{F}$, ki ima nosilec na $[0, \infty)^m$. Poleg tega privzamemo, da so pogojno na realizacijo $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)$ slučajnega vektorja $\mathbf{\Lambda}$ spremenljivke $L_1^{'}, ..., L_m^{'}$ neodvisne:

$$L_i^{'}|\{\Lambda_i = \lambda_i\} \sim Poisson(\lambda_i), \quad \{L_i^{'}|\{\mathbf{\Lambda} = \boldsymbol{\lambda}\}\}_{i=1,...,m} \quad \text{neodvisne.}$$

Večrazsežna porazdelitev spremenljivk $L_i^{'}$ je dana z

$$P(L_1^{'} = l_1, ..., L_m^{'} = l_m) = \tag{3.14}$$

$$= \int_{[0,\infty]^m} P(L_1^{'} = l_1, ..., L_m^{'} = l_m | \mathbf{\Lambda} = \boldsymbol{\lambda}) \, d\mathbf{F}(\lambda_1, ..., \lambda_m) =$$

$$= \int_{[0,\infty]^m} P(L_1 = l_1|\mathbf{\Lambda} = \boldsymbol{\lambda})...P(L_m = l_m|\mathbf{\Lambda} = \boldsymbol{\lambda})\, d\mathbf{F}(\lambda_1, ..., \lambda_m) =$$

$$= \int_{[0,\infty)^m} e^{-(\lambda_1+...+\lambda_m)} \prod_{i=1}^{m} \frac{\lambda_i^{l_i}}{l_i!}\, d\mathbf{F}(\lambda_1, ..., \lambda_m),$$

kjer je $l_i \in \{0, 1, 2, ...\}$. Kot v Bernoullijevem primeru dobimo

$$E(L_i^{'}) = E(\Lambda_i) \qquad (i = 1, ..., m) \tag{3.15}$$

$$var(L_i^{'}) = var(E(L_i^{'}|\mathbf{\Lambda})) + E(var(L_i^{'}|\mathbf{\Lambda})) = var(\Lambda_i) + E(\Lambda_i).$$

Spet velja

$$cov(L_i^{'}, L_j^{'}) = cov(\Lambda_i, \Lambda_j) \tag{3.16}$$

in korelacija med izostanki plačila dolga je dana z

$$corr(L_i^{'}, L_j^{'}) = \frac{cov(\Lambda_i, \Lambda_j)}{\sqrt{var(\Lambda_i) + E(\Lambda_i)}\sqrt{var(\Lambda_j) + E(\Lambda_j)}} \tag{3.17}$$

Na enak način kot pri Bernoullijevem modelu nam to pove, da korelacijo določa porazdelitvena funkcija $\mathbf{F}$ slučajnega vektorja $\mathbf{\Lambda}$.

3.2.2 Enoten parameter $\boldsymbol{\lambda}$ in enotna korelacija

Spet se omejimo na enoten parameter λ in enotno korelacijo med kreditojemalci v portfelju. Bolj natančno, enoten portfeljski model v Poissonovem primeru podamo s Poissonovimi spremenljivkami $L_i^{'} \sim Poisson(\Lambda)$ s slučajnim parametrom Λ, ki ima porazdelitveno funkcijo F z nosilcem na $[0, \infty)$. Za spremenljivke $L_i^{'}$ privzemimo, da so pogojno neodvisne. Večrazsežna porazdelitev spremenljivk $L_i^{'}$ je enaka

$$P(L_1^{'} = l_1, ..., L_m^{'} = l_m) = \int_0^{\infty} e^{-m\lambda} \frac{\lambda^{l_1+...+l_m}}{l_1!...l_m!}\, dF(\lambda). \tag{3.18}$$

Verjetnost dogodka, da se zgodi natanko k izostankov plačil, je enaka

$$P(L^{'} = k) = \int_0^{\infty} P(L^{'} = k|\Lambda = \lambda)\, dF(\lambda) \tag{3.19}$$

$$= \int_0^{\infty} e^{-m\lambda} \frac{m^k \lambda^k}{k!}\, dF(\lambda).$$

Zaradi neomejenosti nosilca Poissonove porazdelitve lahko absolutna izguba $L^{'}$ preseže število možnih izostankov plačil dolga. Že na začetku tega razdelka smo povedali, da je za tipične parametrizacije verjetnost dogodka večih izostankov plačil majhna. V Poissonovem modelu je enotna verjetnost neplačila vseh kreditojemalcev v portfelju definirana z

$$\bar{p} = P(L_i^{'} \geq 1) = \int_0^\infty P(L_i^{'} \geq 1 | \Lambda = \lambda)\, dF(\lambda) \tag{3.20}$$

$$= \int_0^\infty (1 - e^{-\lambda})\, dF(\lambda).$$

Enakost (3.17) lahko pišemo drugače

$$corr(L_i^{'}, L_j^{'}) = \frac{var(\Lambda)}{var(\Lambda) + E(\Lambda)} \quad (i \neq j). \tag{3.21}$$

To formulo lahko izrazimo tudi s pomočjo disperzije, ki je definirana kot

$$D_X = \frac{var(X)}{E(X)}. \tag{3.22}$$

Disperzija Poissonove slučajne spremenljivke je enaka 1. Zato je Poissonova porazdelitev neke vrste merilo, ko se odločamo o prerazpršenosti (overdispersion) ($D_X > 1$) oz. podrazprešenosti (underdispersion) ($D_X < 1$). V splošnem so glede na enakost (3.15) nedegenerirane[2] spremenljivke iz Poissonovega modela prerazpršene.
Formula (3.21) nam pove, da korelacija med številom izostankov plačil različnih kreditojemalcev narašča z disperzijo slučajne spremenljivke Λ. Za dokaz te trditve zapišimo

$$corr(L_i^{'}, L_j^{'}) = \frac{D_\Lambda}{D_\Lambda + 1} \quad (i \neq j). \tag{3.23}$$

Iz zgornjega sledi, da porast disperzije poveča efekte mešanega modela, ki okrepijo odvisnost med izostanki plačil različnih kreditojemalcev.

3.3 Primerjava Bernoullijevega in Poissonovega modela

Zaradi zakona majhnih števil[3] za velike m in majhne p velja

$$Bin(m, p) \approx Poisson(p\,m).$$

[2] Slučajni parameter Λ ni skoncentriran v eni točki, $P_\Lambda \neq \varepsilon_\lambda$

[3] aproksimacija binomske porazdelitve s pomočjo Poissonove porazdelitve

Označimo $\lambda = p\,m$ in privzemimo neodvisnost izostankov plačil dolga. Tedaj nam zgornja enakost pove, da lahko absolutno izgubo portfelja L, $L = \sum L_i$, Bernoulijevih statistik izgube $(L_1, ..., L_m)$ z enotno verjetnostjo neplačila p aproksimiramo s Poissonovo spremenljivko $L' \sim Poisson(\lambda)$. Toda zakon majhnih števil v nobenem primeru ni dovolj močan argument za dokaz večinskega mnenja, da sta Bernoullijev in Poissonov pristop bolj ali manj kompatibilna. Da pokažemo, da imata oba pristopa opazne sistematične razlike, se vrnimo k korelacijam med izostanki plačil; poglejmo zveze (3.5), (3.3) in (3.17). V Bernoullijevem primeru imamo

$$corr(L_i, L_j) = \tag{3.24}$$

$$= \frac{cov(P_i, P_j)}{\sqrt{var(P_i) + E(P_i\,(1 - P_i))}\,\sqrt{var(P_j) + E(P_j\,(1 - P_j)}},$$

medtem ko v Poissonovem primeru dobimo

$$corr(L_i, L_j) = \frac{cov(\Lambda_i, \Lambda_j)}{\sqrt{var(\Lambda_i) + E(\Lambda_j)}\,\sqrt{var(\Lambda_i) + E(\Lambda_j)}}. \tag{3.25}$$

Če se osredotočimo samo na glavne slučajne spremenljivke P_i, P_j oz. Λ_i, Λ_j, vidimo, da v imenovalcih enakosti (3.24) in (3.25) primerjamo

$$var(P_i) + E(P_i(1 - P_i)) = var(P_i) + E(P_i) - E(P_i^2) \tag{3.26}$$

$$\text{z} \quad var(\Lambda_i) + E(\Lambda_i).$$

Analogno determinističnemu primeru (3.12) lahko pričakujemo, da bosta P_i in Λ_i iste velikosti. Zaradi enostavnosti privzemimo, da imata P_i in Λ_i isto matematično upanje in varianco. V tem primeru enakosti (3.26), (3.24) in (3.25) povejo, da Bernoullijev model vedno povzroči večjo korelacijo med izostanki plačil dolga kot Poissonov. Toda večja korelacija med izostanki plačil se odraža v debelejših repih ustreznih porazdelitev izgub. Z drugimi besedami, če imamo dana prvi in drugi moment spremenljivk P_i in Λ_i, se bosta matematični upanji spremenljivk L_i in L_i' ujemali, varianca L_i' pa bo vedno presegla varianco L_i in s tem povzročila manjšo korelacijo med izostanki plačil.
Torej obstaja sistematična razlika med Bernoullijevim in Poissonovim mešanim modelom.

Poglavje 4

Alternativne mere tveganja in alokacija kapitala

Na prvi pogled se zdi definicija ekonomskega kapitala, kot smo jo vzeli v 2. poglavju, popolnoma zadovoljiva. Žal pa ima koncept tvegane vrednosti VaR nekaj pomanjkljivosti, saj ne zadosti vsem lastnostim, ki naj bi jih imela mera tveganja. O t.i. koherentnih merah tveganja bomo nekaj več povedali v naslednjem razdelku, sedaj pa na hitro ponovimo glavne oznake.
Portfeljska izguba je dana z

$$L_{PF} = \sum_{i=1}^{m} \omega_i \, \eta_i \, D_i,$$

kjer je

$$\omega_i = \frac{\nu_i}{\sum_{j=1}^{m} \nu_j}$$

utež i-tega kredita. Velikost izgube ob nastopu izostanka odplačila dolga označujemo z η_i, D_i pa je Bernoullijeva slučajna spremenljivka, ki pomeni nesposobnost plačila dolga i-tega kreditojemalca. Verjetnost neplačila dolga i-tega kreditojemalca (npr. v enem letu) označimo s

$$P(D_i = 1) = p_i.$$

Korelacija med izostanki plačil je dana z

$$\rho_{ij} = corr(D_i, D_j).$$

Privzemimo, da so velikosti izgub η_i deterministične spremenljivke. Slučajno spremenljivko L_i, ki pomeni izgubo terjatve do i-tega kreditojemalca, definirajmo kot

$$L_i = \eta_i \, D_i.$$

4.1 Koherentne mere tveganja

Označimo z $L^\infty = L^\infty(\Omega, \mathcal{F}, P)$ prostor omejenih realnih slučajnih spremenljivk, definiranih na verjetnostnem prostoru $(\Omega, \mathcal{F}, P)$. Tedaj lahko definiramo koherentno mero tveganja na naslednji način:

Definicija 4.1.1 : *Preslikava $\gamma : L^\infty(\Omega, \mathcal{F}, P) \mapsto \mathbb{R}$ se imenuje koherentna mera tveganja, če zadošča naslednjim lastnostim:*

1. *subaditivnost:* $\forall X, Y \in L^\infty : \gamma(X + Y) \leq \gamma(X) + \gamma(Y)$
2. *monotonost:* $\forall X, Y \in L^\infty : X \leq Y \Rightarrow \gamma(X) \leq \gamma(Y)$
3. *pozitivna homogenost:* $\forall \lambda > 0, \forall X \in L^\infty : \gamma(\lambda X) = \lambda \gamma(X)$
4. *translacijska invarianca:* $\forall x \in \mathbb{R}, \forall X \in L^\infty : \gamma(X + x) = \gamma(X) + x$

X razumemo kot portfeljsko izgubo in $\gamma(X)$ kot zahtevani kapital, ki služi kot zaščita proti izgubi X v skladu s politiko ravnanja s kreditnimi tveganji v banki.

Poskušajmo sedaj intuitivno opisati lastnosti iz definicije:
subaditivnost: Zaradi razpršenosti je tveganje pri uniji dveh portfeljev manjše od vsote tveganj pri dveh portfeljih. Kasneje bomo videli, da kvantili v splošnem niso subaditivni, zato tvegana vrednost VaR, ki smo jo omenili v 2. poglavju, ni koherentna mera tveganja.
monotonost: Denimo, da opazujemo portfelja A in B z izgubama X_A in X_B. Če je izguba portfelja A manjša od izgube portfelja B s.g., je zahtevani kapital $\gamma(X_A)$ za portfelj A manjši od zahtevanega kapitala $\gamma(X_B)$ za portfelj B.
homogenost: Vzemimo kreditni portfelj z izgubo X in pomnožimo vse kredite s faktorjem λ. Tedaj se, seveda, izguba X spremeni v izgubo λX in zahtevani kapital zaradi tveganja $\gamma(X)$ se spremeni v $\lambda \gamma(X)$.
translacijska invarianca: Če je x kapital, ki bo gotovo zgubljen/dobljen pri nekem portfelju v določenem časovnem obdobju, se v skladu s tem poveča/ zmanjša zahtevani kapital za kritje izgub v tem portfelju. Translacijska invarianca ima za posledico lastnost $\gamma(X - \gamma(X)) = 0$ za vsako izgubo $X \in L^\infty$.

Tipični meri tveganja sta tvegana vrednost in pričakovani izpad, o katerih bomo nekaj več povedali v nadaljevanju.

Definicija 4.1.2 : *Dana je verjetnostna mera P in stopnja zaupanja α. Tvegana vrednost (Value-at-Risk ali krajše VaR) definiramo kot α-kvantil izgube X:*

$$VaR_\alpha(X) = \inf\{x \geq 0; P(X \leq x) \geq \alpha\}.$$

Kadar je X zvezna slučajna spremenljivka, se definicija poenostavi v:

$$P(X \leq VaR_\alpha(X)) = \alpha.$$

Trditev 4.1.1 : *Tvegana vrednost VaR, definirana na L^∞, ni koherentna mera tveganja.*

Dokaz : VaR je monotona, pozitivno homogena in translacijsko invariantna, ni pa subaditivna. Dokažimo, da zadošča prvim trem lastnostim, in poiščimo protiprimer za četrto.

1. monotonost:
   Vzemimo slučajni spremenljivki X in $Y, X \leq Y$. Za vsak $x \in \mathbb{R}$ velja

   $$P(X \leq x) \geq P(Y \leq x).$$

   Zanimajo nas samo primeri, ko je

   $$P(X \leq x) \geq P(Y \leq x) \geq \alpha. \tag{4.1}$$

   Poiščemo najmanjši tak x, da velja drugi del neenakosti:

   $$\inf\{x \geq 0; P(Y \leq x) \geq \alpha\}.$$

   Ta izraz je ravno enak $VaR_\alpha(Y)$.
   Ker neenakost (4.1) velja za vsak x, velja tudi za $VaR_\alpha(Y)$. Torej:

   $$P(X \leq VaR_\alpha(Y)) \geq \alpha.$$

   Sedaj lahko zapišemo:

   $$VaR_\alpha(X) = \inf\{x \geq 0; P(X \leq x) \geq \alpha\} \leq VaR_\alpha(Y).$$

2. homogenost:
   Vzemimo slučajno spremenljivko X in $\lambda \in \mathbb{R}, \lambda > 0$. Tedaj velja:

   $$\begin{aligned} VaR_\alpha(\lambda X) &= \inf\{x \geq 0; P(\lambda X \leq x) \geq \alpha\} = \\ &= \inf\Big\{x \geq 0; P\Big(X \leq \frac{x}{\lambda}\Big) \geq \alpha\Big\} = \\ &\underset{x'=\frac{x}{\lambda}}{=} \inf\{\lambda x'; P(X \leq x') \geq \alpha\} = \\ &= \lambda \inf\{x' \geq 0; P(X \leq x') \geq \alpha\} = \lambda VaR_\alpha(X). \end{aligned}$$

3. translacijska invarianca:
   Vzemimo slučajno spremenljivko $X, a \in \mathbb{R}$ in si oglejmo $VaR_\alpha(X+a)$:

$$VaR_\alpha(X+a) = \inf\{x \geq 0; P(X+a \leq x) \geq \alpha\} =$$

$$= \inf\{x \geq 0; P(X \leq x-a) \geq \alpha\} =$$

$$\underset{\substack{x'=x-a \\ x' \geq 0}}{=} \inf\{x'+a \geq 0; P(X \leq x') \geq \alpha\} =$$

$$= \inf\{x' \geq 0; P(X \leq x') \geq \alpha\} + a = VaR_\alpha(X) + a.$$

4. subaditivnost:
   Oglejmo si preprost protiprimer. Vzemimo dva neodvisna kredita z verjetnostjo nevračila p, $0.006 \leq p \leq 0.01$. Privzemimo, da je izguba η ob nastopu izostanka odplačila dolga 100% in znesek kredita enak 1. Definirajmo portfelja A in B, tako da je vsak od njiju sestavljen iz enega od obeh kreditov. Za portfeljski izgubi X_A in X_B, ki sta porazdeljeni $X_A \sim X_B \sim Bernoulli(p)$, velja

$$VaR_{99\%}(X_A) = VaR_{99\%}(X_B) = 0.$$

Vzemimo sedaj portfelj C, ki je definiran kot unija portfeljev A in B, in označimo z $X_C = X_A + X_B$ ustrezno portfeljsko izgubo. Tedaj velja

$$P(X_C = 0) = (1-p)^2 < 99\%.$$

Zato je

$$VaR_{99\%}(X_C) > 0.$$

Od tod sledi

$$VaR_{99\%}(X_A + X_B) > VaR_{99\%}(X_A) + VaR_{99\%}(X_B).$$

$\square$

Definicija 4.1.3 : *Dana je stopnja zaupanja α. Pogojno matematično upanje v repu porazdelitve (the tail conditional expectation) ali pričakovan izpad (expected shortfall) je definiran kot*

$$TCE_\alpha(X) = E(X|X \geq VaR_\alpha(X)).$$

Definirajmo s $c = VaR_\alpha(X)$ kritični prag izgube glede na neko stopnjo zaupanja α. Pričakovan izpad zagotavlja zaščito pred povprečno vrednostjo izgube, ki prekorači kritični prag c. Z drugimi besedami, TCE se osredotoči na pričakovano izgubo v repu porazdelitve izgube portfelja. Kritični prag c, ki je odvisen od stopnje zaupanja α, mora fiksirati uprava banke in je del politike ravnanja s kreditnimi tveganji v banki.

Za dokaz trditve v nadaljevanju bomo potrebovali naslednjo lemo:

Lema 4.1.1 : *Naj bo Y slučajna spremenljivka, definirana na prostoru $L^1(\Omega, \mathcal{F}, P)$, in $y \in \mathbb{R}$ tak, da je verjetnost $P(Y \geq y)$ pozitivna. Tedaj za vse $F \in \mathcal{F}$ s $P(Y \geq y) \leq P(F)$ velja*

$$E(Y|Y \geq y) \geq E(Y|F).$$

Dokaz : V primeru, ko je $P(F \cap \{Y \leq y\}) = 0$, je neenakost trivialna. Privzemimo torej, da je verjetnost $P(F \cap \{Y \geq y\})$ pozitivna. Tedaj velja:

$$E(Y|Y \geq y) = E(y + (Y-y)|Y \geq y) = y + \frac{E((Y-y)\,1_{\{Y\geq y\}})}{P(Y \geq y)} =$$

$$= y + \frac{E((Y-y)\,1_{\{Y\geq y\}\cap F}) + E((Y-y)\,1_{\{Y\geq y\}\cap\Omega\setminus F})}{P(Y \geq y)} \geq$$

$$\geq y + \frac{E((Y-y)\,1_{\{Y\geq y\}\cap F})}{P(Y \geq y)} =$$

$$= y + \frac{E((Y-y)|\{Y \geq y\}\cap F)\,P(\{Y \geq y\}\cap F)}{P(Y \geq y)} =$$

$$= y + E((Y-y)|\{Y \geq y\}\cap F)\,P(F|\{Y \geq y\}).$$

Zaradi privzetka $P(F) \geq P(Y \geq y)$ je $P(F|\{Y \geq y\}) \geq P(\{Y \geq y\}|F)$. Vsoto lahko zato ocenimo navzdol takole:

$$y + E((Y-y)|\{Y \geq y\}\cap F)\,P(F|\{Y \geq y\}) \geq$$

$$\geq y + E((Y-y)|\{Y \geq y\}\cap F)\,P(\{Y \geq y\}|F),$$

kar je enako

$$y + \frac{E((Y-y)|\{Y \geq y\}\cap F)\,P(\{Y \geq y\}\cap F)}{P(F)} = y + \frac{E((Y-y)\,1_{\{Y\geq y\}\cap F})}{P(F)} \geq$$

$$\geq y + \frac{E((Y-y)\,1_{\{Y\geq y\}\cap F}) + E((Y-y)\,1_{\{Y<y\}\cap F})}{P(F)}.$$

Zadnja neenakost velja, ker je pogojno matematično upanje $E((Y-y)\,1_{\{Y<y\}\cap F})$ negativno. Ta izraz lahko zapišemo tudi kot

$$y + \frac{E((Y-y)\,1_F)}{P(F)} = y + E((Y-y)|F) = E(Y|F).$$

$\square$

Trditev 4.1.2 : *Pričakovan izpad je koherentna mera tveganja.*

Dokaz :

1. subaditivnost:
   Vzemimo slučajni spremenljivki X in Y in si oglejmo pričakovani izpad vsote:

   $$TCE_\alpha(X+Y) = E(X+Y|X+Y \geq VaR_\alpha(X+Y)) =$$

   $$= E(X|X+Y \geq VaR_\alpha(X+Y)) + E(Y|X+Y \geq VaR_\alpha(X+Y)).$$

   Če za definicijo mere VaR vzamemo $P(X \leq VaR_\alpha(X)) = \alpha$, velja

   $$P(X \geq VaR_\alpha(X)) = P(Y \geq VaR_\alpha(Y)) = P(X+Y \geq VaR_\alpha(X+Y)).$$

   Po lemi 4.1.1 tedaj sledi

   $$E(X|X+Y \geq VaR_\alpha(X+Y)) \leq E(X|X \geq VaR_\alpha(X))$$

   in

   $$E(Y|X+Y \geq VaR_\alpha(X+Y)) \leq E(Y|Y \geq VaR_\alpha(Y)).$$

   Zato lahko ocenimo

   $$E(X|X+Y \geq VaR_\alpha(X+Y)) + E(Y|X+Y \geq VaR_\alpha(X+Y)) \leq$$

   $$\leq E(X|X \geq VaR_\alpha(X)) + E(Y|VaR_\alpha(Y)) = TCE_\alpha(X) + TCE_\alpha(Y).$$

2. monotonost:
   Naj bosta X in Y slučajni spremenljivki, $X \leq Y$. Zato velja

   $$E(X|X \geq VaR_\alpha(X)) \leq E(Y|X \geq VaR_\alpha(X)).$$

   Kot prej definicijo mere VaR pišemo kot $P(X \leq VaR_\alpha(X)) = \alpha$ in ugotovimo

   $$P(X \geq VaR_\alpha(X)) = P(Y \geq VaR_\alpha(Y)).$$

S pomočjo leme 4.1.1 pogojno matematično upanje $E(Y|X \geq VaR_\alpha(X))$ ocenimo navzgor:

$$E(Y|X \geq VaR_\alpha(X)) \leq E(Y|Y \geq VaR_\alpha(Y)).$$

Zapišimo vse skupaj:

$$TCE_\alpha(X) = E(X|X \geq VaR_\alpha(X)) \leq E(Y|X \geq VaR_\alpha(X)) \leq$$

$$\leq E(Y|Y \geq VaR_\alpha(Y)) = TCE_\alpha(Y),$$

kar je neenakost, ki smo jo hoteli dokazati.

3. pozitivna homogenost:
Vzemimo slučajno spremenljivko X in $\lambda \in \mathbb{R}, \lambda > 0$. Upoštevajmo pozitivno homogenost mere VaR in pišimo:

$$TCE_\alpha(\lambda X) = E(\lambda|\lambda X \geq VaR_\alpha(\lambda X)) = E(\lambda X|\lambda X \geq \lambda VaR_\alpha(X)) =$$

$$= E(\lambda X|X \geq VaR_\alpha(X)) = \lambda E(X|X \geq VaR_\alpha(X)) = \lambda TCE_\alpha(X)$$

4. translacijska invarianca:
Vzemimo slučajno spremenljivko $X, a \in \mathbb{R}$ in upoštevajmo translacijsko invarianco mere VaR. Tedaj velja:

$$TCE_\alpha(X+a) = E(X+a|X+a \geq VaR_\alpha(X+a)) =$$

$$= E(X+a|X+a \geq VaR_\alpha(X)+a) = E(X+a|X \geq VaR_\alpha(X)) =$$
$$= E(X|X \geq VaR_\alpha(X)) + a = TCE_\alpha(X) + a$$

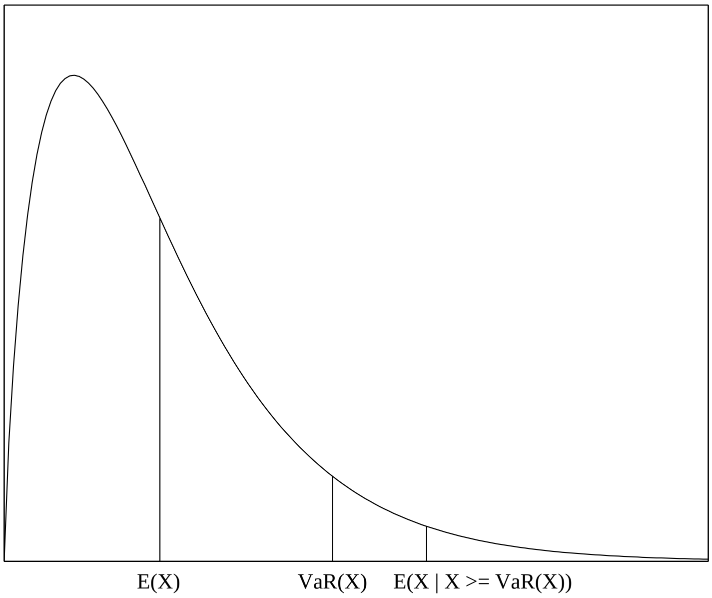
□

Pričakovan izpad zato predstavlja alternativo klasični tvegani vrednosti VaR. Ekonomski kapital, ki temelji na pričakovanem izpadu, lahko definiramo kot

$$EC_{TCE}(VaR_\alpha(X)) = E(X|X \geq VaR_\alpha(X)) - E(X). \tag{4.2}$$

Ta metoda računanja ekonomskega kapitala vključuje tudi rezultate nad kritičnim pragom izgube. Zaradi enostavnejših oznak v nadaljevanju pišimo

$$EC_{TCE_\alpha} = EC_{TCE}(VaR_\alpha(X)).$$



Slika 4.1: Pričakovan izpad

4.2 Prispevek h kapitalu

Prispevek h kapitalu oz. prispevek k tveganju kreditnega portfelja (risk contribution) je del ekonomskega kapitala, ki ga prispeva posamezni kreditojemalec. Definiramo ga kot odvod dane mere tveganja v smeri uteži naložbe.
Ko torej govorimo o alokaciji kapitala, se srečamo s problemom odvedljivosti mere tveganja: le-ta mora biti "dovolj gladka", da iskani odvodi obstajajo. Primer take mere tveganja je standardna deviacija. Na tem dejstvu sloni znana variančno-kovariančna metoda. Z vidika optimizacije portfelja sta kvantil in standardna deviacija usklajena samo v primeru normalne porazdelitve. Iz prejšnjih poglavij pa vemo, da je porazdelitev izgube kreditnega portfelja tipično ukrivljena z debelim repom. Zato moramo pri klasičnemu pristopu zaradi enostavnosti uporabe vzeti v zakup nedoslednost.
K sreči pa variančno-kovariančna metoda za alokacijo kapitala, prilagojena za uporabo pri kreditnih portfeljih, v večini primerov vodi k sprejemljivim rezultatom in jo zaradi njene enostavnosti uporabljajo v večini standardnih paketov. V naslednjem podrazdelku bomo razložili podrobnosti.

4.2.1 Variančno-kovariančna metoda

Jedro variančno-kovariančne metode je vprašanje, kolikšen je prispevek posameznega kredita k standardni deviaciji portfelja UL_{PF}. Da bi odgovorili na to vprašanje, po variančno-kovariančnem pristopu razdelimo portfeljsko tveganje UL_{PF} na prispevke k tveganju kreditnega portfelja RC_i, tako da velja

$$\sum_{i=1}^{m} \omega_i \, RC_i = UL_{PF}. \tag{4.3}$$

Na ta način se tehtani prispevki k tveganju kreditnega portfelja seštejejo v skupno tveganje portfelja, ki ga identificiramo s standardno deviacijo. Če namesto zneska kredita ν_i vzamemo delež kredita $\omega_i = \frac{\nu_i}{\sum_j \nu_j}$, lahko analogno trditvi 2.2.2 zapišemo

$$\begin{aligned} UL_{PF} &= \frac{1}{UL_{PF}} \sum_{i=1}^{m} \omega_i \sum_{j=1}^{m} \omega_j \, \eta_i \, \eta_j \, \sigma_{D_i} \, \sigma_{D_j} \, \rho_{ij} = \\ &= \frac{1}{UL_{PF}} \sum_{i=1}^{m} \omega_i \sum_{j=1}^{m} \omega_j \, UL_i \, UL_j \, \rho_{ij}, \end{aligned} \tag{4.4}$$

kjer s $\sigma_{D_i} = \sqrt{p_i \, (1 - p_i)}$ označimo standardno deviacijo izostanka plačila in upoštevamo enakost $UL_i = \eta_i \, \sigma_{D_i}$. Od tod

$$RC_i = \frac{UL_i}{UL_{PF}} \sum_{j=1}^{m} \omega_j \, UL_j \, \rho_{ij}. \tag{4.5}$$

Veljata tudi naslednji dve zvezi

$$RC_i = \frac{1}{UL_{PF}} \, cov(L_i, L) \tag{4.6}$$

in

$$RC_i = \frac{\partial UL_{PF}}{\partial \omega_i}. \tag{4.7}$$

Zveza (4.7) nam pove: če se utež i-tega kredita v portfelju poveča za h, se UL_{PF} poveča za hRC_i. Sledi, da se uteženi parcialni odvodi seštejejo v UL_{PF}. Z opazovanjem razmerja med prispevkom k tveganju kreditnega portfelja in standardno deviacijo ugotovimo, da v večini primerov velja

$$\frac{RC_i}{UL_i} \leq 1. \tag{4.8}$$

Ker skupen kapital zaradi tveganja določimo s pomočjo kvantilov

$$EC_{VaR_\alpha} = VaR_\alpha(L) - E(L), \tag{4.9}$$

moramo posamezen prispevek k tveganju kreditnega portfelja pomnožiti s t.i. kapitalskim množiteljem

$$CM_\alpha = \frac{EC_{VaR_\alpha}}{UL_{PF}}. \tag{4.10}$$

Prispevek h kapitalu za i-ti kredit je tedaj enak

$$\delta_i = CM_\alpha RC_i, \quad \text{kjer velja} \quad \sum_{i=1}^{m} \omega_i\, \delta_i = EC_{VaR_\alpha}. \tag{4.11}$$

Količina δ_i se imenuje analitični prispevek kapitala i-te transakcije h kapitalu portfelja.
Kapitalski množitelj je pomožna količina, odvisna od posameznega portfelja. Upoštevati moramo namreč dejstvo, da kvantili porazdelitve izgube kreditnega portfelja, v nasprotju z normalno porazdelitvijo, niso odvisni samo od standardne deviacije, ampak tudi od drugih vplivov kot npr. korelacije, verjetnosti neplačil in uteži naložb.
V nadaljevanju si bomo ogledali alokacijo kapitala glede na tvegano vrednost oz. pričakovan izpad.

4.2.2 Alokacija kapitala v odvisnosti od tvegane vrednosti

Računanje prispevka k tveganju s pomočjo tvegane vrednosti VaR je težko, saj v splošnem kvantili niso odvedljivi po utežeh naložb. Pri predpostavki, da je večrazsežna gostota slučajne spremenljivke X_i zvezna, odvod $VaR_\alpha(X)$, kjer je $X = \sum_i \omega_i\, X_i$, obstaja. Velja

$$\frac{\partial VaR_\alpha}{\partial \omega_i}(X) = E(X_i | X = VaR_\alpha(X)) \tag{4.12}$$

(za dokaz glej [10]).
Žal pa je porazdelitev izgube portfelja $L = \sum_i \omega_i\, L_i$, kot smo jo definirali na začetku tega poglavja, nezvezna. Zato odvodi mere VaR v tem smislu ne obstajajo. Lahko pa definiramo prispevek k tveganju kreditnega portfelja preko desne strani enakosti (4.12) na naslednji način

$$\gamma_i = E(L_i | L = VaR_\alpha(L)) - E(L_i). \tag{4.13}$$

Za lažje razumevanje se spomnimo

$$\frac{\partial E(L)}{\partial \omega_i} = E(L_i) \quad \text{in} \quad \sum_{i=1}^{m} \omega_i \, \gamma_i = EC_{VaR_\alpha}.$$

Pri velikih portfeljih in na primerni skali je porazdelitev izgube L blizu zvezni funkciji, vendar tudi v takih primerih žal ni analitična.
Alokacija kapitala, dobljena s tvegano vrednostjo VaR, ni zadovoljiva. Čeprav je morda $\{RC_i\}_{i=1,\ldots,m}$ smiselna particija standardne deviacije portfelja, ne pove kaj dosti o tveganju v repu porazdelitve izgube.

4.2.3 Alokacija kapitala v odvisnosti od pričakovanega izpada

Pri prispevkih k tveganju kreditnega portfelja, dobljenih s pomočjo pričakovanega izpada, se srečamo z istimi tehničnimi težavami kot prej, namreč mera TCE v splošnem ni odvedljiva. Vendar pa velja: če je porazdelitev izgube dovolj gladka, je TCE_α parcialno odvedljiv po utežeh naložb

$$\frac{\partial TCE_\alpha}{\partial \omega_i}(X) = E(X_i | X \geq VaR_\alpha(X)). \tag{4.14}$$

V primeru, da parcialni odvodi ne obstajajo, lahko prispevke, izračunane iz pričakovanega izpada, za nezvezno spremenljivko portfeljske izgube definiramo z

$$\zeta_i = E(L_i | L \geq VaR_\alpha(L)) - E(L_i). \tag{4.15}$$

Analogno primeru $VaR - EC$ lahko zapišemo

$$\sum_{i=1}^{m} \omega_i \, \zeta_i = EC_{TCE_\alpha},$$

kar nam pove, da lahko EC v odvisnosti od pričakovanega izpada dobimo kot uteženo vsoto ustreznih prispevkov.

Poglavje 5

Model CREDITRISK+

5.1 Uvod

Model CreditRisk+ obravnava tveganja zaradi neplačevanja dolgov na osnovi informacij o velikosti in zapadlosti kredita, kreditni kvaliteti danega kreditojemalca in sistemskem tveganju, ki smo mu izpostavljeni ob kreditiranju tega kreditojemalca. CreditRisk+ je statističen model tveganja zaradi neplačevanja, ki verjetnost neplačila obravnava kot slučajno spremenljivko. Vključuje standardno deviacijo verjetnosti neplačila, da tako poudari njeno spremenljivost. Dejavniki iz okolja, na primer ekonomsko stanje v državi, pogosto pogojujejo korelacijo med izostanki plačil, čeprav med njimi morda ni nobene vzročne povezave. Učinke teh dejavnikov v model CreditRisk+ vključujemo preko spremenljivosti verjetnosti neplačila in preko sektorske analize. Zato eksplicitnega podatka o korelaciji med izostanki plačil ne potrebujemo.
Za modeliranje nenadnega izostanka odplačila dolga danega kreditojemalca uporabljamo matematične tehnike, ki so zelo razvite v zavarovalništvu.

5.2 Izostanki plačil pri fiksnih verjetnostih neplačil

V začetnih razdelkih predpostavimo, da je verjetnost neplačila dolga vsakega kreditojemalca fiksna. S to predpostavko in dejstvom, da ni vzročne povezave med posameznimi izostanki plačil dolga, le-te interpretiramo kot neodvisne. Izostanki poplačila dolga se pojavijo kot zaporedje dogodkov na tak način, da je nemogoče napovedati točen čas pojavitve kateregakoli izmed izostankov plačil ali točnega števila vseh izostankov plačil.

5.2.1 Izostanki odplačil dolga

Dan je portfelj N kreditov. Predpostavimo, da poznamo verjetnost neplačila v razdobju enega leta za vsak kredit. Označimo z A slučajno spremenljivko, ki pomeni neplačilo dolga posameznega kreditojemalca. To je Bernoullijeva slučajna spremenljivka

$$A \sim \begin{pmatrix} 1 & 0 \\ p_A & 1 - p_A \end{pmatrix}.$$

Njena rodovna funkcija je enaka

$$F_A(z) = \sum_{k=0}^{1} P(A = k) z^k = 1 + p_A\,(z - 1). \tag{5.1}$$

Sedaj lahko definiramo slučajno spremenljivko P, ki pomeni portfeljsko izgubo, na naslednji način:

$$P = A_1 + A_2 + ... + A_N.$$

Rodovno funkcijo slučajne spremenljivke P definiramo kot

$$F(z) = \sum_{n=0}^{\infty} P(\text{zgodi se } n \text{ izostankov plačil})\, z^n. \tag{5.2}$$

Ker je kreditojemalcev končno, točneje N, rodovna funkcija pa je neskončna vsota, vzamemo $P(\text{zgodi se } n \text{ izostankov plačil}) = 0$ za $n > N$.
Zaradi neodvisnosti med izostanki plačila posameznih kreditojemalcev lahko pišemo

$$F(z) = \prod_A F_A(z) = \prod_A (1 + p_A\,(z - 1)). \tag{5.3}$$

Zapišimo to še drugače:

$$\log F(z) = \sum_A \log(1 + p_A\,(z - 1)). \tag{5.4}$$

Predpostavimo, da so verjetnosti neplačil p_A majhne. To je značilnost kreditnih portfeljev. S to predpostavko lahko potence verjetnosti p_A zanemarimo in tako logaritem nadomestimo z

$$\log(1 + p_A\,(z - 1)) = p_A\,(z - 1). \tag{5.5}$$

V limiti enakost (5.4) postane

$$F(z) = e^{\sum_A p_A\,(z-1)} = e^{\mu\,(z-1)}, \tag{5.6}$$

kjer je

$$\mu = \sum_A p_A \tag{5.7}$$

in pomeni pričakovano število izostankov poplačil dolga celotnega portfelja v enem letu. Velja namreč

$$E(P) = \sum_A E(A) = \sum_A p_A.$$

Rodovno funkcijo $F(z)$ razvijemo v Taylorjevo vrsto

$$F(z) = e^{\mu\,(z-1)} = e^{-\mu}\,e^{\mu\,z} = \sum_{n=0}^{\infty} \frac{e^{-\mu}\,\mu^n}{n!}\,z^n \tag{5.8}$$

in tako dobimo verjetnost

$$P(\text{zgodi se } n \text{ izostankov plačil}) = \frac{e^{-\mu}\,\mu^n}{n!}. \tag{5.9}$$

V enakosti (5.9) smo za število izostankov plačil dobili znano Poissonovo porazdelitev. Opazimo naslednje:

- Porazdelitev ima samo en parameter in sicer pričakovano število izostankov plačil μ. Porazdelitev ni odvisna od števila kreditov v portfelju ali od posameznih verjetnosti neplačila dolga.
- Ni potrebe, da bi imeli krediti enake verjetnosti neplačil dolga.

5.3 Izgube pri fiksnih verjetnostih neplačil

Pri začetnih predpostavkah smo dobili porazdelitev števila izostankov plačil dolga kreditnega portfelja v enem letu. V bistvu pa nas zanima porazdelitev portfeljske izgube. Porazdelitvi sta različni, ker lahko v enem letu pride do iste izgube zaradi enega velikega izostanka plačila ali zaradi več majhnih izostankov plačil. Medtem ko različne verjetnosti neplačil dolga posameznih kreditov ne vplivajo na porazdelitev skupnega števila izostankov plačil, je porazdelitev izgube, ki v splošnem ni Poissonova, odraz različnih zneskov kreditov. Še več, za skupno porazdelitev je informacija o porazdelitvi posameznih kreditov ključnega pomena.

5.3.1 Uporaba kreditnih razredov

Na prvem koraku razdelimo kredite portfelja v razrede. S tem opazno zmanjšamo količino podatkov, s katerimi bomo računali.
Zaradi delitve v razrede seveda pride do samo približnega računanja, vendar pa je število kreditov veliko, širina razredov pa je majhna v primerjavi s povprečnim kreditom v portfelju, zato je ta aproksimacija dobra. To ustreza dejstvu, da točno število kreditov v portfelju ni ključnega pomena pri določanju skupnega tveganja.
V nadaljevanju si oglejmo oznake, ki jih uporabljamo pri delitvi v razrede:

spremenljivka	**simbol**
kreditojemalec	A
kredit	L_A
verjetnost neplačila dolga	p_A
pričakovana izguba	λ_A

Izberimo še enoto za kredit L in z njeno pomočjo za vsakega kreditojemalca A definirajmo števili ε_A in ν_A

$$L_A = L\,\nu_A \tag{5.10}$$

in

$$\lambda_A = L\,\varepsilon_A.$$

Glavna ideja je, da vsak ν_A zaokrožimo na najbližje celo število. S tem nadomestimo vsak kredit L_A z najbližjim večkratnikom enote L. Če smo za enoto L izbrali primerno število, ostane, potem ko izvedemo zaokrožanje na velikem portfelju, relativno majhno število možnih vrednosti za ν_A, ki sedaj pomenijo kredit za več kreditojemalcev.
Portfelj tako razdelimo na m kreditnih razredov, indeksiranih po j, kjer $1 \leq j \leq m$. Vpeljemo nove oznake:

spremenljivka	**simbol**
skupen kredit v j-tem kreditnem razredu $[L]$	ν_j
pričakovana izguba v j-tem kreditnem razredu $[L]$	ε_j
pričakovano število izostankov plačil v j-tem kreditnem razredu	μ_j

Izrazimo pričakovano izgubo s pomočjo pričakovanega števila izostankov plačil

$$\varepsilon_j = \nu_j \mu_j, \tag{5.11}$$

od tod

$$\mu_j = \frac{\varepsilon_j}{\nu_j} = \sum_{A:\nu_A=\nu_j} \frac{\varepsilon_A}{\nu_A}.$$

Kot v zvezi (5.7) naj μ pomeni pričakovano število izostankov plačil portfelja v enem letu. Ker je μ vsota pričakovanih števil izostankov plačil v posameznih razredih, velja

$$\mu = \sum_{j=1}^{m} \mu_j = \sum_{j=1}^{m} \frac{\varepsilon_j}{\nu_j} \tag{5.12}$$

5.3.2 Porazdelitev izgube

Zaradi razlike v znesku kredita nekateri izostanki plačil vodijo k večjim izgubam kot drugi. Tudi pri analizi izgube si pomagamo z rodovno funkcijo. Naj bo torej $G(z)$ rodovna funkcija izgube, izražene kot večkratnik enote L

$$G(z) = \sum_{n=0}^{\infty} P(\text{izguba} = nL)\, z^n. \tag{5.13}$$

Predpostavimo, da so krediti v portfelju neodvisni. Torej so kreditni razredi neodvisni in rodovno funkcijo lahko zapišemo kot produkt po kreditnih razredih

$$G(z) = \prod_{i=1}^{m} G_i(z). \tag{5.14}$$

Posamezen kreditni razred lahko obravnavamo kot samostojen portfelj. Upoštevamo še enakost (5.8) in dobimo

$$G_j(z) = \sum_{n=0}^{\infty} P(\text{zgodi se } n \text{ izostankov plačil})\, z^{n\,\nu_j} = \tag{5.15}$$

$$= \sum_{n=0}^{\infty} \frac{e^{-\mu_j}\, \mu_j^n}{n!}\, z^{n\,\nu_j} = e^{-\mu_j + \mu_j\, z^{\nu_j}}.$$

Zato velja

$$G(z) = \prod_{j=1}^{m} e^{-\mu_j + \mu_j\, z^{\nu_j}} = e^{-\sum_{j=1}^{m} \mu_j + \sum_{j=1}^{m} \mu_j\, z^{\nu_j}}. \tag{5.16}$$

To je željena formula za rodovno funkcijo izgube celotnega portfelja. V nadaljevanju bomo pokazali, kako iz rodovne funkcije izpeljemo dejansko porazdelitev izgube.
Relacijo (5.16) bomo poskušali zapisati v malo drugačni obliki. Najprej definirajmo polinom $P(z)$ na naslednji način

$$P(z) = \frac{\sum_{j=1}^{m} \mu_j \, z^{\nu_j}}{\mu} = \frac{\sum_{j=1}^{m} \left(\frac{\varepsilon_j}{\nu_j}\right) z^{\nu_j}}{\sum_{j=1}^{m} \left(\frac{\varepsilon_j}{\nu_j}\right)}, \tag{5.17}$$

kjer smo uporabili zvezi (5.11) in (5.12) za skupno število izostankov plačil μ v portfelju. Rodovno funkcijo iz enakosti (5.16) lahko sedaj zapišemo kot

$$G(z) = e^{\mu\,(P(z)-1)} = F(P(z)). \tag{5.18}$$

$G(z)$ je odvisna samo od podatkov $\boldsymbol{\nu} = (\nu_1, ..., \nu_m)$ in $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_m)$. Da torej dobimo porazdelitev izgube za velik portfelj kreditov, moramo poznati samo različne višine kreditov $\boldsymbol{\nu}$ znotraj portfelja in ustrezne pričakovane izgube $\boldsymbol{\varepsilon}$. To je navadno zelo majhna količina podatkov, celo za velike portfelje.

5.4 Porazdelitev izgube pri fiksnih verjetnostih neplačil

5.4.1 Postopek računanja

V tem podrazdelku bomo razvili učinkovito računsko sredstvo kot pomoč za izpeljavo dejanske porazdelitve izgube iz rodovne funkcije, dane v zvezi (5.16).
Označimo z B_n verjetnosti izgube nL za $n \geq 0$ in jih poskušajmo čim bolj učinkovito izračunati. Vemo

$$P(\text{izguba} = nL) = \frac{1}{n!} \frac{d^n G(z)}{dz^n}\bigg|_{z=0} = B_n. \tag{5.19}$$

Vzamemo $G(z)$ iz enakosti (5.16) in uporabimo Leibnitzovo formulo

$$\frac{1}{n!} \frac{d^n G(z)}{dz^n}\bigg|_{z=0} = \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(G(z) \frac{d}{dz} \sum_{j=1}^{m} \mu_j \, z^{\nu_j} \right)\bigg|_{z=0} = \tag{5.20}$$

$$= \frac{1}{n!} \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{d^{n-k-1}}{dz^{n-k-1}} G(z) \frac{d^{k+1}}{dz^{k+1}} \left(\sum_{j=1}^{m} \mu_j \, z^{\nu_j} \right)\bigg|_{z=0}.$$

Velja

$$\frac{d^{k+1}}{dz^{k+1}}\left(\sum_{j=1}^{m}\mu_j\, z^{\nu_j}\right)\Bigg|_{z=0} = \begin{cases} \mu_j\,(k+1)!; & k=\nu_j-1 \quad \text{za nek } j \\ 0; & \text{sicer} \end{cases} \tag{5.21}$$

Po definiciji pa še

$$\frac{d^{n-k-1}}{dz^{n-k-1}}\,G(z)\Bigg|_{z=0} = (n-k-1)!\,B_{n-k-1} \tag{5.22}$$

Zato sledi

$$B_n = \sum_{\substack{k\le n-1\\ k=\nu_j-1}} \frac{1}{n!}\binom{n-1}{k}(k+1)!\,(n-k-1)!\,\mu_j\,B_{n-k-1} = \tag{5.23}$$

$$= \sum_{j:\nu_j\le n}\frac{\mu_j\,\nu_j}{n}\,B_{n-\nu_j}$$

Uporabimo relacijo $\varepsilon_j = \nu_j\mu_j$ iz enakosti (5.11) in dobimo rekurzivno zvezo za B_n, s katero lahko hitro poiščemo porazdelitev:

$$B_n = \sum_{j:\nu_j\le n}\frac{\varepsilon_j}{n}\,B_{n-\nu_j}. \tag{5.24}$$

Naslednjo formulo uporabimo za izračun prvega člena, ki pomeni verjetnost, da v portfelju ne bo prišlo do izgube:

$$B_0 = G(0) = F(P(0)) = e^{-\mu} = e^{-\sum_{j=1}^{m}\frac{\varepsilon_j}{\nu_j}} \tag{5.25}$$

Spet opazimo, da potrebujemo samo podatka $\boldsymbol{\varepsilon}$ in $\boldsymbol{\nu}$. V praksi to predstavlja majhno količino podatkov, celo če gre za velike portfelje.

5.4.2 Natančnost pri uporabi kreditnih razredov

S procesom razvrščanja v razrede vpeljemo v podatke neko aproksimacijo. Če se osredotočimo na portfeljsko povprečje in standardno deviacijo, lahko pokažemo, da napaka zaradi aproksimacije ni bistvena.
Pričakovana izguba portfelja ε in kvadrat standardne deviacije portfelja σ, izraženi v izbrani enoti L, sta enaka

$$\varepsilon = \sum_{j=1}^{m}\varepsilon_j, \tag{5.26}$$

$$\sigma^2 = \sum_{j=1}^{m} \nu_j \varepsilon_j.$$

Prva enakost je očitna, drugo pa dobimo takole:

$$\sigma^2 = G''(1) + \varepsilon - \varepsilon^2 =$$

$$= \sum_{j=1}^{m} \nu_j \mu_j \sum_{j=1}^{m} \nu_j \mu_j + \sum_{j=1}^{m} \nu_j \mu_j (\nu_j - 1) + \varepsilon - \varepsilon^2 =$$

$$= \varepsilon^2 + \sum_{j=1}^{m} \nu_j \varepsilon_j - \varepsilon + \varepsilon - \varepsilon^2 = \sum_{j=1}^{m} \nu_j \varepsilon_j.$$

Predpostavimo, da sta to točni vrednosti povprečja in standardne deviacije, komponente $\boldsymbol{\nu}$ pa zaokrožimo na večkratnik enote. Zaradi tega procesa se pojavi napaka

$$\hat{\nu}_j = \nu_j + \tau_j, \quad 0 \leq \tau_j \leq 1. \tag{5.27}$$

Vsak τ_j je po absolutni vrednosti največ 1. Predpostavimo, da so krediti zaokroženi navzgor, zato so vsi τ_j pozitivni.
Izbrana metoda zaokroževanja ne vpliva na pričakovano izgubo, saj je le-ta neodvisna od zneskov kreditov posameznih razredov. Za standardno deviacijo pa velja

$$\sigma^2 \leq \hat{\sigma}^2 = \sum_{j=1}^{m} \hat{\nu}_j \varepsilon_j = \sigma^2 + \sum_{j=1}^{m} \tau_j \varepsilon_j \leq \sigma^2 + \sum_{j=1}^{m} \varepsilon_j = \sigma^2 + \varepsilon. \tag{5.28}$$

Korenimo in zanemarimo člene višjega reda v Taylorjevem razvoju izraza $\sqrt{\sigma^2 + \varepsilon}$:

$$\sigma \leq \hat{\sigma} \leq \sigma\left(1 + \frac{\varepsilon}{2\sigma^2}\right) = \sigma + \frac{\varepsilon}{2\sigma}. \tag{5.29}$$

Pričakovana izguba ε in količina σ sta istega reda. Zaključimo lahko:

- Na pričakovano izgubo, ki jo izračunamo z modelom, razvrščanje v razrede ne vpliva.
- Standardna deviacija se poveča za znesek, ki je primerljiv z izbrano enoto.

5.5 Uporaba za izračun izgub v več letih

Rekurzivno relacijo smo izpeljali na osnovi porazdelitve izgube v enem letu. V tem razdelku bomo pokazali, kako lahko začetni model razširimo na časovno obdobje večih let.
Kot prej imamo portfelj kreditojemalcev z majhnimi verjetnostmi neplačila dolga. Zaradi enostavnosti predpostavimo, da je prihodnost portfelja razdeljena na leta. Krediti se lahko spreminjajo iz leta v leto. V posebnem ima vsak kredit individualno zapadlost, ki ustreza običajnim zapadlostim obveznic, kreditov ali ostalih inštrumentov.
Za večletno obdobje bomo za vsako leto potrebovali robne verjetnosti neplačil dolga za vsakega kreditojemalca v portfelju.
Potrebovali bomo še naslednje oznake:

spremenljivka	**simbol**
verjetnost nevračila j-tega kredita v letu t	$p_j^{(t)}$
znesek j-tega kredita v letu t	$L_j^{(t)} = L\,\nu_j^{(t)}$
pričakovana izguba j-tega kredita v letu t	$\lambda_j^{(t)} = L\,\varepsilon_j^{(t)}$

Upoštevamo, da se izostanki plačil istega kredita v različnih letih med seboj izključujejo, in zapišemo rodovno funkcijo izgube posameznega kredita v več letih

$$G_j(z) = 1 - \sum_{t=0}^{T} p_j^{(t)} + \sum_{t=0}^{T} p_j^{(t)} z^{\nu_j^{(t)}} = 1 + \sum_{t=0}^{T} p_j^{(t)} (z^{\nu_j^{(t)}} - 1). \tag{5.30}$$

Rodovna funkcija skupne izgube je enaka

$$G(z) = \prod_j G_j(z) = \prod_j \left(1 + \sum_{t=0}^{T} p_j^{(t)} (z^{\nu_j^{(t)}} - 1)\right). \tag{5.31}$$

Logaritmiramo

$$\log G(z) = \sum_j \log\left(1 + \sum_{t=0}^{T} p_j^{(t)} (z^{\nu_j^{(t)}} - 1)\right). \tag{5.32}$$

Za majhne verjetnosti neplačila dolga kot v enakosti (5.5) velja

$$\log\left(1+\sum_{t=1}^{T} p_j^{(t)}(z^{\nu_j^{(t)}}-1)\right)=\sum_{t=1}^{T} p_j^{(t)}(z^{\nu_j^{(t)}}-1). \tag{5.33}$$

Tako dobimo

$$\log G(z)=\sum_j\sum_{t=1}^{T} p_j^{(t)}(z^{\nu_j^{(t)}}-1). \tag{5.34}$$

Vemo

$$\sum_{j,t} p_j^{(t)}=\sum_{j,t}\frac{\varepsilon_j^{(t)}}{\nu_j^{(t)}}. \tag{5.35}$$

Sledi

$$G(z)=e^{\sum_{j,t}\frac{\varepsilon_j^{(t)}}{\nu_j^{(t)}}(z^{\nu_j^{(t)}}-1)}. \tag{5.36}$$

Opazimo, da je oblika te rodovne funkcije enaka kot tista v enakosti (5.16). Zato lahko rekurzivno relacijo za porazdelitev izgube v enem letu iz enakosti (5.24) razširimo na porazdelitev izgube v več letih

$$B_n=\sum_{\substack{j,t\\ \nu_j^{(t)}\le n}}\frac{\varepsilon_j^{(t)}}{n}B_{n-\nu_j^{(t)}}. \tag{5.37}$$

5.6 Sektorska analiza

V naslednjih podpoglavjih bomo govorili o slučajni verjetnosti neplačila in sektorski analizi.
Spremenljivost verjetnosti neplačila je odvisna od relativno majhnega števila dejavnikov na kreditojemalce znotraj portfelja. Zanima nas stopnja vpliva, ki ga ima vsak od dejavnikov na dani portfelj kreditojemalcev. Dejavnik, kot je npr. ekonomsko stanje posamezne države, ima enoten vpliv na kreditojemalce, ki živijo v tej državi, in sorazmeroma majhen vpliv na ostale kreditojemalce portfelja. V tem razdelku se bomo merjenja vplivov dejavnikov lotili z delitvijo kreditojemalcev med različne sektorje. Vsak sektor je množica kreditojemalcev pod skupnim vplivom nekega dejavnika, ki določa verjetnost neplačila dolga. Kasneje bomo predstavili bolj splošno sektorsko analizo, ki dovoljuje, da na posameznega kreditojemalca vpliva več dejavnikov.
Potrebovali bomo nove oznake, da bomo lahko spremljali delitev portfelja na sektorje in spremenljivost števila izostankov plačil za vsak sektor. Označimo

s $S_k, 1 \leq k \leq n$, sektorje, ki za zdaj pomenijo podmnožico množice kreditojemalcev.
Model CreditRisk+ predpostavlja, da vsak sektor temelji na enem dejavniku. Le-ta vpliva na sektor preko skupnega števila izostankov plačila v sektorju, ki ga modeliramo kot slučajno spremenljivko. Definirajmo torej slučajno spremenljivko X_k, ki pomeni število izostankov plačil v sektorju. Njeno matematično upanje označimo z μ_k, standardno deviacijo pa s σ_k.
Potrebovali bomo še naslednje oznake, ki so precej podobne tistim iz podrazdelka 5.3.1.

podatki o kreditu znotraj sektorja	**prejšnja oznaka**	**nova oznaka**
enota za kredit	L	L
velikost kredita $[L]$	$L_j = L_j\,\nu_j$ $1 \leq j \leq m$	$L_j^{(k)} = L\,\nu_j^{(k)}$ $1 \leq k \leq n$ $1 \leq j \leq m(k)$
pričakovana izguba v posameznem kreditnem razredu $[L]$	$\lambda_j = L\,\varepsilon_j$ $1 \leq j \leq m$	$\lambda_j^{(k)} = L\,\varepsilon_j^{(k)}$ $1 \leq k \leq n$ $1 \leq j \leq m(k)$

Matematično upanje μ_k je povezano s podatki o pričakovani izgubi z relacijo, ki je analogna enakosti (5.12)

$$\mu_k = \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}. \tag{5.38}$$

5.6.1 Ocenjevanje spremenljivosti verjetnosti neplačila dolga

Zanima nas tudi standardna deviacija σ_k skupnega števila izostankov plačil v posameznem sektorju. Poiskati moramo primeren način za oceno standardne deviacije s pomočjo matematičnega upanja. Čeprav v zvezi (5.38) nastopajo količine, določene s kreditnimi razredi, lahko matematično upanje izrazimo kot vsoto po vseh kreditojemalcih v sektorju

$$\mu_k = \sum_{A \in S_k} \frac{\varepsilon_A}{\nu_A}. \tag{5.39}$$

Z enakostjo

$$\frac{\varepsilon_A}{\nu_A} = p_A \tag{5.40}$$

izrazimo povprečno verjetnost neplačila dolga kreditojemalca A v določenem časovnem obdobju. Da dobimo oceno standardne deviacije za vsak sektor, privzemimo, da poznamo verjetnost neplačila p_A in standardno deviacijo σ_A za vse kreditojemalce znotraj posameznega sektorja. Lahko rečemo, da je standardna deviacija odvisna od kreditne kvalitete kreditojemalca. Na ta način privzamemo, da ima kreditna kvaliteta kreditojemalcev znotraj posameznega sektorja večji vpliv na spremenljivost frekvence pričakovanih neplačil kot narava samega sektorja.
Oceno za σ_k dobimo iz standardnih deviacij kreditojemalcev σ_A s procesom povprečenja. Dejansko verjetnost neplačila dolga posameznega kreditojemalca v sektorju bomo modelirali kot slučajno spremenljivko, ki je proporcionalna X_k in ima matematično upanje enako povprečni verjetnosti neplačila tega kreditojemalca. Z X_A označimo slučajno verjetnost neplačila dolga kreditojemalca A. Sedaj lahko zapišemo

$$X_A = \frac{\varepsilon_A}{\nu_A}\frac{X_k}{\mu_k}. \tag{5.41}$$

Po tej enačbi je matematično upanje slučajne spremenljivke X_A enako p_A. Velja tudi

$$\sum_{A\in S_k}\sigma_A = \sum_{A\in S_k}\frac{\varepsilon_A}{\nu_A}\frac{\sigma_k}{\mu_k} = \frac{\sigma_k}{\mu_k}\sum_{A\in S_k}\frac{\varepsilon_A}{\nu_A} = \sigma_k. \tag{5.42}$$

Standardno deviacijo povprečnega števila izostankov plačil v sektorju dobimo kot vsoto standardnih deviacij za vsakega kreditojemalca v sektorju. Lahko pa jo opišemo tudi z zvezo

$$\frac{\sigma_k}{\mu_k} = \frac{\sum_{A\in S_k}\sigma_A}{\sum_{A\in S_k}p_A} = \frac{\sum_{A\in S_k}p_A\left(\frac{\sigma_A}{p_A}\right)}{\sum_{A\in S_k}p_A}. \tag{5.43}$$

Glede na prejšnje izkušnje je ulomek $\frac{\sigma_A}{p_A}$ navadno reda 1, zato je standardna deviacija števila izostankov plačil, opazovanih iz leta v leto med kreditojemalci podobne kreditne kvalitete, istega reda kot povprečno letno število izostankov plačil dolga. Relacija (5.43) pove, da isto velja za posamezne sektorje.

5.7 Izostanki plačil pri spremenljivih verjetnostih neplačil

5.7.1 Pogojna verjetnost neplačila

V tem podrazdelku privzamemo, da je verjetnost neplačila slučajna spremenljivka in s pomočjo rodovne funkcije izpeljemo porazdelitev izostankov plačil celotnega portfelja. Pomagamo si z rodovno funkcijo iz enakosti (5.6)

$$F(z) = e^{\mu\,(z-1)},$$

kjer je število izostankov plačil μ fiksno število. Rodovno funkcijo izostankov plačil celotnega portfelja v splošnem zapišemo kot

$$F(z) = \sum_{n=0}^{\infty} P(\text{zgodi se } n \text{ izostankov plačil})\, z^n.$$

Zaradi neodvisnosti sektorjev lahko $F(z)$ pišemo kot produkt po sektorjih

$$F(z) = \prod_{k=1}^{n} F_k(z).$$

Zato moramo predvsem poiskati $F_k(z)$ za posamezen sektor. Ker X_k pomeni število izostankov plačil v k-tem sektorju, lahko zapišemo rodovno funkcijo izostankov plačil pogojno na vrednost slučajne spremenljivke X_k

$$F_k(z)|\{X_k = x\} = e^{x\,(z-1)}. \tag{5.44}$$

Denimo, da ima X_k gostoto $f_k(x)$, tako da velja

$$P(x \leq X_k \leq x + dx) \approx f_k(x)\, dx. \tag{5.45}$$

Tedaj je rodovna funkcija izostankov plačil v enem sektorju povprečje pogojnih rodovnih funkcij, danih z zvezo (5.44), po vseh možnih vrednostih števila izostankov plačil. To zapišemo v naslednji zvezi

$$F_k(z) = \sum_{n=0}^{\infty} P(\text{zgodi se } n \text{ izostankov plačil})\, z^n = \tag{5.46}$$

$$\sum_{n=0}^{\infty} z^n \int_0^{\infty} P(\text{zgodi se } n \text{ izostankov plačil}|x_k = x)\, f_k(x)\, dx =$$

$$= \int_0^{\infty} e^{x\,(z-1)} f_k(x)\, dx.$$

Da bi dobili eksplicitno formulo za rodovno funkcijo, moramo izbrati primerno porazdelitev za X_k. Na tem mestu naredimo ključni privzetek, namreč, da je X_k gama slučajna spremenljivka z matematičnim upanjem μ_k in standardno deviacijo σ_k. Oglejmo si nekaj njenih osnovnih lastnosti.

Gama porazdelitev $\Gamma(\alpha, \beta)$ je nesimetrična porazdelitev, ki aproksimira normalno, če je njeno matematično upanje veliko. Gostota slučajne spremenljivke X, $X \sim \Gamma(\alpha, \beta)$ je za $x \geq 0$ dana z

$$P(x \leq X \leq x + dx) \approx f(x)\, dx = \frac{1}{\beta^{\alpha}\,\Gamma(\alpha)}\, e^{-\frac{x}{\beta}}\, x^{\alpha-1}\, dx,^{1} \tag{5.47}$$

[1]Za slučajno spremenljivko s tako gostoto ponavadi pišemo $X \sim \Gamma\left(\alpha, \frac{1}{\beta}\right)$

kjer je $\Gamma(\alpha) = \int_0^\infty e^{-x}\, x^{\alpha-1}\, dx$ Gama funkcija.
Gama porazdelitev $\Gamma(\alpha, \beta)$ je enolično določena z matematičnim upanjem in standardno deviacijo, ki sta s parametroma α in β v naslednji relaciji

$$\mu = \alpha\,\beta, \quad \sigma^2 = \alpha\,\beta^2. \tag{5.48}$$

Od tod lahko izračunamo parametre gama porazdelitve za k-ti sektor

$$\alpha_k = \frac{\mu_k^2}{\sigma_k^2}, \quad \beta_k = \frac{\sigma_k^2}{\mu_k}. \tag{5.49}$$

5.7.2 Porazdelitev izostankov plačil v posameznem sektorju

Z izbiro gama porazdelitve za funkcijo $f_k(x)$ lahko izraz za rodovno funkcijo

$$F_k(z) = \int_0^\infty e^{x\,(z-1)} f_k(x)\, dx \tag{5.50}$$

direktno izračunamo (indeks pri α in β izpustimo zaradi enostavnosti izrazov)

$$F_k(z) = \int_0^\infty e^{x\,(z-1)}\, \frac{1}{\beta^\alpha\, \Gamma(\alpha)}\, e^{-\frac{x}{\beta}}\, x^{\alpha-1}\, dx = \tag{5.51}$$

$$= \frac{1}{\beta^\alpha\, \Gamma(\alpha)} \int_0^\infty e^{-y} \left(\frac{y}{\frac{1}{\beta} + 1 - z} \right)^{\alpha-1} \frac{dy}{(\frac{1}{\beta} + 1 - z)} =$$

$$= \frac{1}{\beta^\alpha\, \Gamma(\alpha)}\, \frac{1}{(\frac{1}{\beta} + 1 - z)^\alpha} \int_0^\infty e^{-y}\, y^{\alpha-1}\, dy = \frac{1}{\beta^\alpha\, (\frac{1}{\beta} + 1 - z)^\alpha}.$$

Preuredimo in pišemo

$$F_k(z) = \left(\frac{1 - p_k}{1 - p_k z} \right)^{\alpha_k}, \tag{5.52}$$

kjer je

$$p_k = \frac{\beta_k}{1 + \beta_k}.$$

To je rodovna funkcija porazdelitve izostankov plačil, ki se pojavijo v k-tem sektorju.
Radi bi določili porazdelitev, ki pripada tej rodovni funkciji. Z razvojem $F_k(z)$ v Taylorjevo vrsto

$$F_k(z) = (1 - p_k)^{\alpha_k} \sum_{n=1}^{\infty} \binom{n + \alpha_k - 1}{n} p_k^n\, z^n \tag{5.53}$$

dobimo verjetnost

$$P(\text{zgodi se } n \text{ izostankov plačil}) = (1-p_k)^{\alpha_k} \binom{n+\alpha_k-1}{n} p_k^n. \tag{5.54}$$

Vidimo, da gre za negativno binomsko porazdelitev $NegBin(\alpha_k, 1-p_k)$. Porazdelitev izostanka plačil celotnega portfelja v splošnem ni negativna binomska, je pa neodvisna vsota negativnih binomskih porazdelitev posameznih sektorjev. Rodovna funkcija portfelja je zato enaka

$$F(z) = \prod_{k=1}^{n} F_k(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-p_k z} \right)^{\alpha_k}. \tag{5.55}$$

5.8 Izgube pri spremenljivih verjetnostih neplačil

Rodovna funkcija iz enakosti (5.55) nam da popolno informacijo o pojavu izostankov plačil dolga v portfelju. Radi pa bi prešli iz izostankov plačil dolga na izgube zaradi teh izostankov. Zato potrebujemo informacijo o porazdelitvi kreditov. V podrazdelku 5.3.2 smo ta prehod izvedli s pomočjo fiksnega števila izostankov plačil, sedaj pa upoštevamo slučajnost števila izostankov plačil.

5.8.1 Porazdelitev izgube

Po analogiji z enakostjo (5.13) zapišemo drugo rodovno funkcijo $G(z)$, rodovno funkcijo za izgube v portfelju

$$G(z) = \sum_{n=0}^{\infty} P(\text{izguba} = nL)\, z^n.$$

Iščemo izraz za $G(z)$ in sredstva za učinkovito računanje $G(z)$.
Zaradi neodvisnosti sektorjev kot pri porazdelitvi izostankov plačil obstaja dekompozicija rodovne funkcije

$$G(z) = \prod_{k=1}^{n} G_k(z),$$

kjer z $G_k(z)$ označimo rodovno funkcijo izgube k-tega sektorja, $1 \leq k \leq n$. Kot v relaciji (5.17) definiramo polinome $P_k(z)$, $1 \leq k \leq n$, z

$$P_k(z) = \frac{\sum_{j=1}^{m(k)} \left(\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} \right) z^{\nu_j^{(k)}}}{\sum_{j=1}^{m(k)} \left(\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} \right)} = \frac{1}{\mu_k} \sum_{j=1}^{m(k)} \left(\frac{\varepsilon_j^{(k)}}{\nu_k^{(k)}} \right) z^{\nu_j^{(k)}}, \tag{5.56}$$

kjer smo uporabili izraz za μ_k iz enakosti (5.38). S $P_k(z)$ dobimo zvezo med izostanki plačil dolga in izgubami, saj velja naslednja zveza

$$G_k(z) = F_k(P_k(z)).$$

Le-ta je spet analogna formuli $G(z) = F(P(z))$ iz enakosti (5.18), le da imamo sedaj po eno tako zvezo za vsak sektor. Da dokažemo veljavnost zveze tudi za ta primer, zapišemo enakost (5.56) kot vsoto po vseh kreditojemalcih iz k-tega sektorja

$$P_k(z) = \frac{\sum_{A\in S_k}\left(\frac{\varepsilon_A}{\nu_A}\right)z^{\nu_A}}{\sum_{A\in S_k}\frac{\varepsilon_A}{\nu_A}} = \frac{1}{\mu_k}\sum_{A\in S_k}\left(\frac{\varepsilon_A}{\nu_A}\right)z^{\nu_A}. \tag{5.57}$$

Uporabimo relacijo (5.41) in dobimo

$$e^{-\sum_{A\in S_k} X_A + \sum_{A\in S_k} X_A z^{\nu_A}} = e^{\sum_{A\in S_k} X_A (z^{\nu_A}-1)} = \tag{5.58}$$

$$e^{\frac{X_k}{\mu_k}\sum_{A\in S_k}\frac{\varepsilon_A}{\nu_A}(z^{\nu_A}-1)} = e^{X_k\,(P_k(z)-1)}.$$

Leva stran zgornje enakosti je rodovna funkcija porazdelitve izgub, kjer je X_A verjetnost izostanka plačila kreditojemalca A.
Kot v enakosti (5.50), ki izrazi $F_k(z)$ kot integral Poissonove rodovne funkcije po prostoru vseh možnih vrednosti slučajne spremenljivke X_k, je $G_k(z)$ integral leve strani enakosti (5.58) po istem prostoru. Torej

$$G_k(z) = \sum_{n=0}^{\infty} z^n \int_0^{\infty} P(\text{izguba} = nL|X_k = x_k) f_k(x_k) dx_k = \tag{5.59}$$

$$= \int_0^{\infty} e^{\sum_{A\in S_k} x_A\,(z^{\nu_A}-1)} f_k(x_k)\,dx_k = \int_0^{\infty} e^{x_k\,\left(P_k(z)-1\right)} f_x(x_k)\,dx_k.$$

S substitucijo v enakosti (5.52) in produktom po vseh sektorjih dobimo

$$G(z) = \prod_{k=1}^{n} G_k(z) = \prod_{k=1}^{n}\left(\frac{1-p_k}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}}\right)^{\alpha_k}. \tag{5.60}$$

To je izraz za rodovno funkcijo. V naslednjem razdelku bomo izpeljali zvezo za računanje porazdelitve izgube iz tega izraza.

5.9 Porazdelitev izgube pri spremenljivih verjetnostih neplačil

5.9.1 Splošna rekurzivna zveza

Denimo, da vrsta

$$G(z) = \sum_{n=0}^{\infty} B_n \, z^n \tag{5.61}$$

definira funkcijo $G(z)$, ki zadošča naslednji diferencialni enačbi

$$\frac{d}{dz}\Big(\log G(z)\Big) = \frac{1}{G(z)} \frac{dG(z)}{dz} = \frac{P(z)}{Q(z)}, \tag{5.62}$$

kjer sta P in Q polinoma, dana z

$$P(z) = p_0 + ... + p_r \, z^r \tag{5.63}$$

$$Q(z) = q_0 + ... + q_s \, z^s.$$

Zahtevamo torej, da je odvod logaritma funkcije $G(z)$ racionalna funkcija. Tedaj členi vrste v zvezi (5.61) zadoščajo naslednji rekurzivni zvezi

$$B_{n+1} = \frac{1}{q_0(n+1)} \left(\sum_{i=0}^{min(r,n)} p_i \, B_{n-i} - \sum_{j=0}^{min(s-1,n-1)} q_{j+1} \, (n-j) \, B_{n-j} \right). \tag{5.64}$$

Za dokaz te enakosti preoblikujmo diferencialno enačbo (5.62) v

$$Q(z) \, \frac{dG(z)}{dz} = P(z) \, G(z) \tag{5.65}$$

Razpišemo

$$\left(\sum_{j=0}^{s} q_j \, z^j \right) \left(\sum_{n=0}^{\infty} (n+1) \, B_{n+1} \, z^n \right) = \left(\sum_{i=0}^{r} p_i \, z^i \right) \left(\sum_{n=0}^{\infty} B_n \, z^n \right). \tag{5.66}$$

Za $n \geq 0$ je člen pred z^n na levi oz. desni strani enak

$$\sum_{j=0}^{min(s,n)} q_j \, (n+1-j) \, B_{n+1-j} \quad \text{oz.} \quad \sum_{i=0}^{min(r,n)} p_i \, B_{n-i}. \tag{5.67}$$

Izenačimo in preoblikujemo tako dobljeni izraz

$$q_0 \, (n+1) \, B_{n+1} = \sum_{i=0}^{min(r,n)} p_i \, B_{n-i} - \sum_{j=1}^{min(s,n)} q_j \, (n+1-j) \, B_{n+1-j}, \tag{5.68}$$

kar je ekvivalentno

$$q_0\,(n+1)\,B_{n+1} = \sum_{i=0}^{min(r,n)} p_i\,B_{n-i} - \sum_{j=0}^{min(s-1,n-1)} q_{j+1}\,(n-j)\,B_{n-j}. \qquad (5.69)$$

5.9.2 Uporaba

V enakosti (5.60) smo rodovno funkcijo izgub zapisali kot

$$G(z) = \prod_{k=1}^{n} G_k(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \right)^{\alpha_k}.$$

Zapišimo odvod logaritma funkcije $G(z)$

$$\frac{G'(z)}{G(z)} = \sum_{k=1}^{n} \frac{G_k'(z)}{G_k(z)} = \sum_{k=1}^{n} \frac{\frac{p_k\alpha_k}{\mu_k}\sum_{j=1}^{m(k)} \varepsilon_j^{(k)} z^{\nu_j^{(k)}-1}}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}}. \qquad (5.70)$$

S tem $\frac{G'(z)}{G(z)}$ izrazimo kot racionalno funkcijo. Potem ko iz enakosti

$$\frac{P(z)}{Q(z)} = \sum_{k=1}^{n} \frac{\frac{p_k\alpha_k}{\mu_k}\sum_{j=1}^{m(k)} \varepsilon_j^{(k)} z^{\nu_j^{(k)}-1}}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \qquad (5.71)$$

izračunamo polinoma $P(z)$ in $Q(z)$, lahko uporabimo rezultate prejšnjega podrazdelka in tako dobimo porazdelitev izgube.

5.10 Konvergenca primera s spremenljivo verjetnostjo neplačila k primeru s fiksno verjetnostjo neplačila

Čeprav naj bi model CreditRisk+ upošteval spremenljivost verjetnosti neplačila, se v nekaterih okoliščinah obnaša, kot bi bila fiksna. Gre za naslednja primera:

- Standardna deviacija števila izostankov plačil v vsakem sektorju gre proti nič.

5.10 Konvergenca primera s spremenljivo verjetnostjo neplačila k primeru s fiksno verjetnostjo neplačila

- Število sektorjev gre proti neskončno.

V tem razdelku bomo poskušali dokazati, da rodovna funkcija izgube pri spremenljivi verjetnosti neplačila

$$G(z) = \prod_{k=1}^{n} G_k(z) = \prod_{k=1}^{n} \left(\frac{1 - p_k}{1 - \frac{p_k}{\mu_k} \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \right)^{\alpha_k}$$

v prvem primeru, torej, ko je standardna deviacija števila izostankov plačil poljubno majhna, konvergira k rodovni funkciji izgube pri fiksni verjetnosti neplačila

$$G(z) = e^{\sum_{j=1}^{m} \left(\frac{\varepsilon_j}{\nu_j}\right)(z^{\nu_j} - 1)}.$$

Od prej vemo

$$\alpha_k = \frac{\mu_k^2}{\sigma_k^2}, \quad \beta_k = \frac{\sigma_k^2}{\mu_k}, \quad p_k = \frac{\beta_k}{1 + \beta_k}, \quad \mu_k = \sum_{k=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}.$$

p_k lahko pišemo tudi drugače:

$$p_k = \frac{\sigma_k^2}{\mu_k (1 + \frac{\sigma_k^2}{\mu_k})} = \frac{\sigma_k^2}{\mu_k + \sigma_k^2}.$$

Upoštevamo $\sigma_k \to 0$ in si ogledamo limite:

$$\lim_{\sigma_k \to 0} \beta_k = 0, \quad \lim_{\sigma_k \to 0} p_k = 0, \quad \lim_{\sigma_k \to 0} (\alpha_k \, p_k) = \lim_{\sigma_k \to 0} \frac{\mu_k^2 \, \sigma_k^2}{\sigma_k^2 \, (\mu_k + \sigma_k^2)} = \mu_k.$$

Zadnja limita nam da približek za α_k:

$$\alpha_k \approx \frac{\mu_k}{p_k}.$$

Sledi

$$\prod_{k=1}^{n} \left(\frac{1 - p_k}{1 - \frac{p_k}{\mu_k} \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \right)^{\alpha_k} \approx \prod_{k=1}^{n} \left(\frac{1 - p_k}{1 - \frac{p_k}{\mu_k} \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \right)^{\frac{\mu_k}{p_k}}. \tag{5.72}$$

Zaradi preglednosti pišimo

$$c_k = \left(\frac{\frac{p_k}{\mu_k} \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}} - p_k}{1 - \frac{p_k}{\mu_k} \sum_{j=1}^{m(k)} \frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}} z^{\nu_j^{(k)}}} \right).$$

Zgornji produkt tako postane

$$\prod_{k=1}^{n}(1+c_k)^{\frac{1}{c_k}c_k\frac{\mu_k}{p_k}}.$$

Ko gre $\sigma_k \to 0$, gre $c_k \to 0$, zato velja

$$\prod_{k=1}^{n}(1+c_k)^{\frac{1}{c_k}c_k\frac{\mu_k}{p_k}} \approx \prod_{k=1}^{n} e^{c_k\frac{\mu_k}{p_k}}.$$

Vstavimo izraz za c_k in produkt zapišemo kot

$$\prod_{k=1}^{n} e^{\left(\frac{\frac{1}{\mu_k}\sum_{j=1}^{m(k)}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}z^{\nu_j^{(k)}}-1}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}z^{\nu_j^{(k)}}}\right)\cdot\mu_k}.$$

V limiti tako dobimo

$$G(z)=\prod_{k=1}^{n}e^{-\mu_k}e^{\sum_{j=1}^{m(k)}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}z^{\nu_j^{(k)}}} = e^{-\sum_{j,k}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}+\sum_{j,k}\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}z^{\nu_j^{(k)}}}. \tag{5.73}$$

Če v eksponentu zberemo skupaj izraze iz istega kreditnega razreda in različnih sektorjev (pri danem j vsi možni k), vsota po k izgine in dobimo

$$G(z)=e^{\sum_j\frac{\varepsilon_j}{\nu_j}(z^{\nu_j}-1)}.$$

5.11 Splošna sektorska analiza

Pri izpeljavi rodovne funkcije za porazdelitev izgube v razdelku 5.8 smo privzeli, da je portfelj razdeljen na sektorje, ki predstavljajo podmnožice množice kreditojemalcev. To ustreza situaciji, v kateri kreditojemalce razdelimo v skupine, na katere vpliva en dejavnik in so medsebojno neodvisne.
Sedaj si bomo ogledali bolj posplošeno situacijo: na sistematično spremenljivost števila izostankov plačil sicer vpliva relativno majhno število dejavnikov, ni pa nujno, da je verjetnost neplačila dolga posameznega kreditojemalca odvisna samo od enega. V takih okoliščinah portfelja ne moremo opisati s sektorji, sestavljenimi iz množice kreditojemalcev. Model CreditRisk+ vključi to situacijo na enak način kot prej, s tem da koncept sektorjev nadomesti s sistematičnimi dejavniki.

Da bomo lažje razumeli posplošeno sektorsko analizo, si še enkrat oglejmo izpeljavo rodovne funkcije. V enakosti (5.60) smo rodovno funkcijo izrazili kot produkt po sektorjih, rodovne funkcije posameznih sektorjev pa kot integral po prostoru možnih vrednosti števila izostankov plačil

$$G(z) = \prod_{k=1}^{n} G_k(z) = \prod_{k=1}^{n} \int_0^\infty e^{x_k(P_k(z)-1)} f_k(x_k)\, dx_k. \tag{5.74}$$

Ta izraz lahko zapišemo tudi kot večkratni integral

$$G(z) = \int_0^\infty \dots \int_0^\infty e^{\sum_{k=1}^{n} x_k(P_k(z)-1)} \prod_{k=1}^{n} f_k(x_k)\, dx_k. \tag{5.75}$$

Integrand imamo za uteženo gostoto sestavljene Poissonove porazdelitve za dano množico vrednosti povprečij $x_k, 1 \leq k \leq n$.
Uporabimo zvezo (5.57) in eksponent zapišemo na drug način

$$\sum_{k=1}^{n} x_k \left(P_k(z) - 1\right) = \sum_{k=1}^{n} x_k \left(\frac{1}{\mu_k} \sum_{A \in S_k} \left(\frac{\varepsilon_A}{\nu_A} \right) z^{\nu_A} - 1 \right) = \tag{5.76}$$

$$= \sum_{k=1}^{n} x_k \left(\frac{1}{\mu_k} \sum_{A \in S_k} \left(\frac{\varepsilon_A}{\nu_A} \right) z^{\nu_A} - \frac{1}{\mu_k} \sum_{A \in S_k} \frac{\varepsilon_A}{\nu_A} \right) = \sum_{A,k} \delta_{Ak} \frac{x_k}{\mu_k} \frac{\varepsilon_A}{\nu_A} \left(z^{\nu_A} - 1\right),$$

kjer je

$$\delta_{Ak} = \begin{cases} 0; & A \notin S_k \\ 1; & A \in S_k \,. \end{cases} \tag{5.77}$$

Posplošitev koncepta sektorjev pomeni, da na vsakega kreditojemalca vpliva več kot en dejavnik X_k. Zato delta funkcijo nadomestimo z alokacijo kreditojemalcev med sektorje, tako da za vsakega kreditojemalca A izberemo število Θ_{Ak}, da velja

$$\sum_{k=1}^{n} \Theta_{Ak} = 1. \tag{5.78}$$

Število Θ_{Ak} predstavlja intenziteto, s katero k-ti dejavnik vpliva na verjetnost neplačila kreditojemalca A. V razdelku 5.6 smo razpravljali o posebnem primeru, ko je veljalo kar

$$\Theta_{Ak} = \delta_{Ak}. \tag{5.79}$$

V splošnem primeru enakost (5.76) nadomestimo z

$$\sum_{k=1}^{n} x_k \left(P_k(z) - 1\right) = \sum_{A,k} \Theta_{Ak} \frac{x_k}{\mu_k} \frac{\varepsilon_A}{\nu_A} \left(z^{\nu_A} - 1\right). \tag{5.80}$$

Vsak kreditojemalec tako prispeva izraz

$$x_A(z^{\nu_A} - 1), \quad \text{kjer je} \quad x_A = \frac{\varepsilon_A}{\nu_A} \sum_{k=1}^{n} \Theta_{Ak} \frac{x_k}{\mu_k}. \tag{5.81}$$

Enakost (5.57) postane

$$P_k(z) = \frac{1}{\mu_k} \sum_A \Theta_{Ak} \frac{\varepsilon_A}{\nu_A} z^{\nu_A}, \quad \text{kjer je} \quad \mu_k = \sum_A \Theta_{Ak} \frac{\varepsilon_A}{\nu_A}. \tag{5.82}$$

5.11.1 Izvajanje sektorske dekompozicije

Pokazali bomo, kako prilagodimo podatke za posplošeno sektorsko analizo modela CreditRisk+. Privzemimo, da imamo za vsakega kreditojemalca v portfelju oceno, s kolikšno mero k-ti dejavnik vpliva na spremenljivost njegove verjetnosti neplačila. Vemo, da to izrazimo s števili

$$\Theta_{Ak}, \quad \text{tako da velja} \quad \sum_{k=1}^{n} \Theta_{Ak} = 1$$

za vsak sektor k in vsakega kreditojemalca A v portfelju.
Poiskati moramo oceno matematičnega upanja in standardne deviacije za vsak sektor. Ogledali si bomo metodo za ocenjevanje teh parametrov ob privzetku, da oceno za obe količini za posameznega kreditojemalca dobimo iz podatka o njegovi kreditni kvaliteti.
Matematično upanje je za vsak sektor vsota prispevkov vseh kreditojemalcev tehtanih s strukturnimi deleži Θ_{Ak}

$$\mu_k = \sum_A \Theta_{Ak}\, \mu_A. \tag{5.83}$$

Potem lahko po analogiji z zvezo (5.43) izrazimo ulomek $\frac{\sigma_k}{\mu_k}$ kot tehtano povprečje prispevkov vseh kreditojemalcev

$$\frac{\sigma_k}{\mu_k} = \frac{\sum_A \Theta_{Ak}\, \mu_A \frac{\sigma_A}{\mu_A}}{\sum_A \Theta_{Ak}\, \mu_A}. \tag{5.84}$$

Od tod

$$\sigma_k = \sum_A \Theta_{Ak}\, \sigma_A.$$

To je ocena standardne deviacije za vsak dejavnik.

5.12 Prispevki k tveganju kreditnega portfelja in medsebojna korelacija

O prispevkih k tveganju kreditnega portfelja smo govorili že v 4. poglavju. Vemo, da jih računamo kot odvod izbrane mere tveganja. Model CreditRisk+ za mero tveganja vzame kar standardno deviacijo. V prvem razdelku bomo tako izpeljali formulo za prispevek posameznega kreditojemalca k standardni deviaciji porazdelitve izgube.
V drugem razdelku bomo izpeljali formulo za medsebojno korelacijo med izostanki plačil. Le-ta daje mero obsega, do katere je v portfelju prisotna koncentracija tveganja.

5.12.1 Prispevki k tveganju kreditnega portfelja

CreditRisk+ definira prispevek h kapitalu kreditojemalca A s kreditom E_A kot robni učinek prisotnosti kredita E_A bodisi na izgubo na 99. percentilu bodisi na standardno deviacijo porazdelitve izgube.
V prvem primeru analitična formula ni mogoča. Namesto tega uporabimo aproksimacijo, opisano v nadaljevanju.
Naj bodo ε,σ in X pričakovana izguba, standardna deviacija izgube in izguba na danem percentilu porazdelitve. Le-to lahko zapišemo kot

$$X = \varepsilon + \xi\,\sigma. \tag{5.85}$$

Tedaj lahko definiramo prispevek k tveganju kreditnega portfelja na percentilu v smislu prispevka k standardni deviaciji

$$\widehat{RC_A} = \varepsilon_A + \xi\,RC_A. \tag{5.86}$$

Sledi

$$\sum_A \widehat{RC_A} = \sum_A (\varepsilon_A + \xi\,RC_A) = \varepsilon + \xi\,\sigma = X. \tag{5.87}$$

V drugem primeru, torej ko prispevek k tveganju kreditnega portfelja računamo kot robni učinek na standardno deviacijo, analitična formula obstaja.

$$RC_A = E_A \frac{\partial\sigma}{\partial E_A} \tag{5.88}$$

ali ekvivalentno

$$RC_A = \frac{E_A}{2\sigma}\frac{\partial\sigma^2}{\partial E_A}.$$

5.12 Prispevki k tveganju kreditnega portfelja in medsebojna korelacija

V večini modelov, vključno s CreditRisk+, se prispevki k tveganju kreditnega portfelja, definirani z enakostjo (5.88), seštejejo v standardno deviacijo. To je res zaradi variančno-kovariančne formule

$$\sigma^2 = \sum_{A,B} \rho_{AB}\, E_A\, E_B\, \sigma_A\, \sigma_B, \tag{5.89}$$

kjer sta σ_A in σ_B standardni deviaciji izostankov plačil kreditojemalcev A in B. Če so korelacijski koeficienti neodvisni od kreditov, z relacijo (5.89) izrazimo varianco kot homogen kvadratičen polinom kreditov. Zaradi splošnih lastnosti homogenih polinomov od tod sledi

$$\sum_A RC_A = \frac{1}{2\sigma} \sum_A E_A \frac{\partial \sigma^2}{\partial E_A} = \frac{1}{2\sigma}\, 2\sigma^2 = \sigma. \tag{5.90}$$

V nadaljevanju se bomo osredotočili na prispevke k standardni deviaciji. Da bi izračunali desno stran enakosti (5.88), bomo razvili analitično formulo za matematično upanje in varianco porazdelitve izostankov plačil in izgube. Uporabili bomo oznake, ki bodo konsistentne s tistimi od prej. Ker sta matematično upanje in varianca aditivna po sektorjih, lahko naredimo analizo po enem samem sektorju. Zaradi lažje pisave bomo tako spustili sklicevanje na k-ti sektor, kjer to ni res potrebno.

slučajna spremenljivka	**rodovna funkcija**	**gostota**	**matematično upanje**	**varianca**
izostanki plačil v sektorju pogojno na vrednost števila izostankov plačil v sektorju	$E(z,x)$		$\mu_E(x)$	$\sigma_E^2(x)$
število izostankov plačil v k-tem sektorju	$F(z)$		μ_F	σ_F^2
izguba v k-tem sektorju	$G(z)$		μ_G	σ_G^2

Opombe:

- $E(z,x) = e^{x\,(z-1)}$
- $F(z) = \left(\frac{1-p_k}{1-p_k z}\right)^{\alpha_k}$
- $P(z) = \frac{1}{\mu_k} \sum_A \Theta_{Ak} \frac{\varepsilon_A}{\nu_A} z^{\nu_A}$

- $G(z) = \left(\frac{1-p_k}{1-\frac{p_k}{\mu_k} \sum_A \Theta_{Ak} \frac{\varepsilon_A}{\nu_A} z^{\nu_A}} \right)^{\alpha_k}$
- $f(x)$ je gostota izostankov plačil, μ_k in σ_k^2 pa pričakovano število in standardna deviacija izostankov plačil v k-tem sektorju pri fiksni verjetnosti neplačila

Velja

$$G(z) = F(P(z)) \tag{5.91}$$

in

$$F(z) = \int E(z,x) f(x)\, dx. \tag{5.92}$$

Zaradi splošnih lastnosti rodovnih funkcij za E, F in G velja

$$\mu_E(x) = \frac{dE}{dz}(1,x), \quad \mu_F = \frac{dF}{dz}(1), \quad \mu_G = \frac{dG}{dz}(1) \tag{5.93}$$

in

$$\sigma_E^2 + \mu_E^2 = \frac{d^2E}{dz^2}(1) + \frac{dE}{dz}(1), \quad \sigma_F^2 + \mu_F^2 = \frac{d^2F}{dz^2}(1) + \frac{dF}{dz}(1), \tag{5.94}$$

$$\sigma_G^2 + \mu_G^2 = \frac{d^2G}{dz^2}(1) + \frac{dG}{dz}(1).$$

Ker je $E(z,x)$ rodovna funkcija Poissonove porazdelitve, velja

$$\mu_E(x) = \sigma_E^2(x) = x. \tag{5.95}$$

S pomočjo enakosti (5.91), (5.92) in (5.93) dobimo

$$\mu_F = \frac{dF}{dz}(1) = \int \frac{dE}{dz}(1,x) f(x)\, dx = \tag{5.96}$$

$$\int \mu_E(x) f(x)\, dx = \int x f(x)\, dx = \mu_k.$$

Podobno iz zvez (5.92), (5.94) in (5.95) sledi

$$\sigma_F^2 + \mu_F^2 = \frac{d^2F}{dz^2}(1) + \frac{dF}{dz}(1) = \tag{5.97}$$

$$\int \left(\frac{d^2E}{dz^2}(1,x) + \frac{dE}{dz}(1,x) \right) f(x)\, dx =$$

$$= \int (\sigma_E^2(x) + \mu_E^2(x)) f(x)\, dx = \int (x + x^2) f(x)\, dx = \mu_k + \sigma_k^2 + \mu_k^2.$$

Od tod
$$\sigma_F^2 = \mu_k + \sigma_k^2. \tag{5.98}$$
Enakosti (5.96) in (5.98) predstavljata matematično upanje in varianco porazdelitve izostankov plačil. Da dobimo zvezo med momenti porazdelitve izgube, uporabimo enakost (5.91), ki nam s pomočjo verižnega pravila da
$$\frac{dG}{dz}(z) = \frac{dF}{dz}(P(z))\,\frac{dP}{dz}(z) \tag{5.99}$$
$$\frac{d^2G}{dz^2}(z) = \frac{d^2F}{dz^2}(P(z))\left(\frac{dP}{dz}(z)\right)^2 + \frac{dF}{dz}(P(z))\,\frac{d^2P}{dz^2}(z).$$
Od tod
$$\sigma_G^2 = \frac{d^2G}{dz^2}(1) + \frac{dG}{dz}(1) - \mu_G^2 = \frac{d^2G}{dz^2}(1) + \frac{dG}{dz}(1) - \left(\frac{dG}{dz}(1)\right)^2 = \tag{5.100}$$
$$= \frac{d^2F}{dz^2}(P(1))\left(\frac{dP}{dz}(1)\right)^2 + \frac{dF}{dz}(P(1))\frac{d^2P}{dz^2}(1) + \frac{dF}{dz}(P(1))\frac{dP}{dz}(1) - \left(\frac{dF}{dz}(P(1))\frac{dP}{dz}(1)\right)^2$$
Spomnimo se enakosti (5.82) in izračunajmo vrednosti, ki jih potrebujemo:
$$P(1) = 1, \quad \frac{dP}{dz}(1) = \frac{1}{\mu_k}\sum_A \Theta_{Ak}\,\varepsilon_A = \frac{\varepsilon_k}{\mu_k}, \tag{5.101}$$
$$\frac{d^2P}{dz^2}(1) = \frac{1}{\mu_k}\sum_A \Theta_{Ak}\,\varepsilon_A\,(\nu_A - 1).$$
Vstavimo zveze iz (5.101) v enakost (5.100):
$$\sigma_G^2 = (\sigma_F^2 + \mu_F^2 - \mu_F)\left(\frac{\varepsilon_k}{\mu_k}\right)^2 + \mu_F\,\frac{1}{\mu_k}\sum_A \Theta_{Ak}\,\varepsilon_A\,(\nu_A - 1) + \mu_F\,\frac{\varepsilon_k}{\mu_k} - \left(\mu_F\,\frac{\varepsilon_k}{\mu_k}\right)^2 =$$
$$= \sigma_F^2\left(\frac{\varepsilon_k}{\mu_k}\right)^2 - \mu_F\left(\frac{\varepsilon_k}{\mu_k}\right)^2 + \sum_A \Theta_{Ak}\,\varepsilon_A\,\nu_A = (\sigma_F^2 - \mu_k)\left(\frac{\varepsilon_k}{\mu_k}\right)^2 + \sum_A \Theta_{Ak}\,\varepsilon_A\,\nu_A =$$
$$= \sigma_k^2\left(\frac{\varepsilon_k}{\mu_k}\right)^2 + \sum_A \Theta_{Ak}\,\varepsilon_A\,\nu_A.$$
Vsota po sektorjih nam da standardno deviacijo portfeljske izgube
$$\sigma^2 = \sum_{k=1}^{n} \varepsilon_k^2\left(\frac{\sigma_k}{\mu_k}\right)^2 + \sum_A \varepsilon_A\,\nu_A. \tag{5.102}$$

Prispevek k tveganju kreditnega portfelja lahko sedaj dobimo direktno z odvajanjem enakosti (5.102)

$$RC_A = E_A \frac{\partial \sigma}{\partial E_A} = \frac{E_A}{2\sigma} \frac{\partial \sigma^2}{\partial E_A}.$$

Upoštevamo, da za kredit kreditojemalca A poleg oznake E_A uporabljamo tudi ν_A, in zvezi $\varepsilon_A = \nu_A \, \mu_A$ ter $\varepsilon_k = \sum_A \Theta_{Ak} \, \varepsilon_A$. Od tod

$$RC_A = \frac{E_A}{2\sigma} \frac{\partial}{\partial E_A} \left(\sum_B \varepsilon_B \, \nu_B + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \varepsilon_k^2 \right) =$$

$$= \frac{E_A}{2\sigma} \left(2 \, E_A \, \mu_A + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 2 \, \varepsilon_k \, \Theta_{Ak} \, \mu_A \right).$$

In končno

$$RC_A = \frac{E_A \, \mu_A}{\sigma} \left(E_A + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \varepsilon_k \, \Theta_{Ak} \right). \tag{5.103}$$

To je zahtevana formula za prispevek k standardni deviaciji. Da se pokazati, da se ti prispevki seštejejo v standardno deviacijo porazdelitve izgube portfelja:

$$\sum_A RC_A = \frac{E_A \, \mu_A}{\sigma} \left(E_A + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \varepsilon_k \, \Theta_{Ak} \right) =$$

$$= \sum_A \frac{E_A^2 \, \mu_A}{\sigma} + \sum_A \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \frac{E_A \, \mu_A}{\sigma} \Theta_{Ak} \, \varepsilon_k.$$

Uporabimo še enakost (5.102) in dobimo

$$\sum_A RC_A = \sum_A \frac{\varepsilon_A \, \nu_A}{\sigma} + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \varepsilon_k \sum_A \Theta_{Ak} \frac{E_A \, \mu_A}{\sigma} = \tag{5.104}$$

$$= \frac{1}{\sigma} \left(\sum_A \varepsilon_A \, \nu_A + \sum_k \left(\frac{\sigma_k}{\mu_k} \right)^2 \varepsilon_k^2 \right) = \frac{\sigma^2}{\sigma} = \sigma$$

Torej je res vsota prispevkov k tveganju kreditnega portfelja enaka standardni deviaciji porazdelitve portfeljske izgube.

5.12.2 Medsebojna korelacija

V tem podrazdelku bomo izpeljali formulo za medsebojno korelacijo oz. korelacijski koeficient med izostanki plačil. V ta namen vsakemu kreditojemalcu A najprej priredimo indikatorsko funkcijo I_A izostanka plačila, ki je slučajna spremenljivka in zavzame naslednje vrednosti:

$$I_A = \begin{cases} 1; & \text{če } A \text{ v čas. obd. } \triangle t \text{ ni sposoben vrniti dolga} \\ 0; & \text{sicer} \end{cases} \tag{5.105}$$

Slučajno spremenljivko I_{AB}, ki pomeni izostanek plačila dveh kreditojemalcev, A in B, pa definiramo takole:

$$I_{AB} = \begin{cases} 1; & \text{če } A \text{ in } B \text{ v čas. obd. } \triangle t \text{ nista sposobna vrniti dolga} \\ 0; & \text{sicer} \end{cases}$$

Korelacija ρ med izostankoma plačil kreditojemalcev A in B v časovnem obdobju $\triangle t$ je enaka

$$\rho_{AB} = \rho(I_A, I_B). \tag{5.106}$$

To je statistična korelacija med indikatorskima funkcijama kreditojemalcev A in B v časovnem obdobju $\triangle t$. Če je matematično upanje slučajnih spremenljivk I_A, I_B in produkta I_{AB} enako μ_A, μ_B in μ_{AB}, so μ_A, μ_B in μ_{AB} pričakovana števila izostankov plačil kreditojemalcev A, B oz. obeh v časovnem obdobju $\triangle t$. Ker indikatorske funkcije zavzamejo samo vrednosti 0 ali 1, lahko standardni izraz za korelacijo poenostavimo v

$$\rho_{AB} = \frac{\mu_{AB} - \mu_A\,\mu_B}{\sqrt{\mu_A - \mu_A^2}\,\sqrt{\mu_B - \mu_B^2}}. \tag{5.107}$$

Poiskati moramo izraz za μ_{AB}. Vzemimo dva različna kreditojemalca A in B in upoštevajmo splošno sektorsko dekompozicijo na n sektorjev. Tako lahko definiramo:

spremenljivka	**kreditojemalec A**	**kreditojemalec B**
časovno obdobje	$\triangle t$	$\triangle t$
trenutna verjetnost izostanka plačila	p_A	p_B
pričakovano št. izostankov plačil	$\mu_A = p_A\,\triangle t$	$\mu_B = p_B\,\triangle t$
sektorska dekompozicija	$\Theta_{Ak}; 1 \leq k \leq n$	$\Theta_{Bk}; 1 \leq k \leq n$

Ker sta kreditojemalca A in B različna za katerokoli vrednost sektorskega povprečja $X_k, 1 \leq k \leq n$, so izostanki plačil med seboj neodvisni. Tako imamo

$$\mu_{AB} = \int_{x_1} ... \int_{x_n} x_A\, x_B \prod_{k=1}^{n} f_k(x_k)\, dx_k, \tag{5.108}$$

kjer sta x_A in x_B kot v zvezi (5.81):

$$x_A = \sum_{k=1}^{n} \frac{x_k}{\mu_k}\, \Theta_{Ak}\, \mu_A \quad \text{in} \quad x_B = \sum_{k=1}^{n} \frac{x_k}{\mu_k}\, \Theta_{Bk}\, \mu_B.$$

Definirajmo koeficiente $\omega_{kk'}$ na naslednji način

$$\omega_{kk'} = \frac{\Theta_{Ak}\, \Theta_{Bk'}}{\mu_k\, \mu_{k'}}\, \mu_A\, \mu_B. \tag{5.109}$$

Sedaj lahko pišemo

$$\mu_{AB} = \int_{x_1} ... \int_{x_n} \sum_{k,k'} \omega_{kk'}\, x_k\, x_{k'} \prod_{k=1}^{n} f_k(x_k)\, dx_k \tag{5.110}$$

oziroma

$$\begin{aligned} \mu_{AB} = \sum_{k \neq k'} \omega_{kk'} \int_{x_k} \int_{x_{k'}} x_k\, x_{k'}\, f_k(x_k)\, f_{k'}(x_{k'})\, dx_k\, dx_{k'} \int_{\substack{x_j \\ j \neq k,k'}} \prod_{\substack{j=1 \\ j \neq k,k'}}^{n} f_j(x_j)\, dx_j + \\ \sum_{k=1}^{n} \omega_{kk} \int_{x_k} x_k^2\, f_k(x_k)\, dx_k \int_{\substack{x_j \\ j \neq k}} \prod_{\substack{j=1 \\ j \neq k}}^{n} f_j(x_j)\, dx_j. \end{aligned} \tag{5.111}$$

Od tod

$$\begin{aligned} \mu_{AB} &= \sum_{\substack{k=1 \\ k \neq k'}}^{n} \omega_{kk'}\, \mu_k\, \mu_{k'} + \sum_{k=1}^{n} \omega_{kk}\, (\mu_k^2 + \sigma_k^2) = \\ &= \sum_{k,k'=1}^{n} \omega_{kk'}\, \mu_k\, \mu_{k'} + \sum_{k=1}^{n} \omega_{kk}\, \sigma_k^2. \end{aligned} \tag{5.112}$$

Prvi del zgornje vsote lahko preoblikujemo in sicer

$$\begin{aligned} \sum_{k,k'=1}^{n} \omega_{kk'}\, \mu_k\, \mu_{k'} &= \sum_{k,k'=1}^{n} \frac{\Theta_{Ak}\, \Theta_{Bk'}}{\mu_k\, \mu_{k'}}\, \mu_A\, \mu_B\, \mu_k\, \mu_{k'} = \\ &= \left(\sum_{k=1}^{n} \frac{\Theta_{Ak}\, \mu_A}{\mu_k}\, \mu_k \right) \left(\sum_{k=1}^{n} \frac{\Theta_{Bk}\, \mu_B}{\mu_k}\, \mu_k \right) = \mu_A\, \mu_B. \end{aligned} \tag{5.113}$$

Torej velja

$$\mu_{AB} = \mu_A\, \mu_B + \sum_{k=1}^{n} \omega_{kk}\, \sigma_k^2. \tag{5.114}$$

Vstavimo ω_{kk} in zapišemo korelacijo ρ_{AB} kot

$$\rho_{AB} = \frac{\mu_{AB} - \mu_A\,\mu_B}{\sqrt{\mu_A - \mu_A^2}\,\sqrt{\mu_B - \mu_B^2}} \approx \frac{1}{\sqrt{\mu_A\,\mu_B}} \sum_{k=1}^{n} \Theta_{Ak}\,\Theta_{Bk}\,\mu_A\,\mu_B \Big(\frac{\sigma_k}{\mu_k}\Big)^2, \tag{5.115}$$

kjer smo zanemarili člena μ_A^2 in μ_B^2 v imenovalcu. Poenostavimo in dobimo

$$\rho_{AB} = \sqrt{\mu_A\,\mu_B} \sum_{k=1}^{n} \Theta_{Ak}\,\Theta_{Bk} \Big(\frac{\sigma_k}{\mu_k}\Big)^2. \tag{5.116}$$

Relacija (5.116) je formula za korelacijo med izostanki plačil dveh različnih kreditojemalcev A in B. Ta enačba je veljavna, kadar so verjetnosti neplačila v časovnem obdobju majhne. Vrednost korelacije ρ_{AB} je lahko večja od 1, če izberemo prevelike vrednosti matematičnih upanj in standardnih deviacij.
Iz enakosti (5.116) lahko razberemo dve vidni značilnosti:

- Če kreditojemalca A in B nimata nobenega skupnega sektorja, je korelacija med njima enaka 0. To je res, ker nobeden od sistematičnih faktorjev ne vpliva na oba.
- Če privzamemo, da je ulomek $\frac{\sigma_k}{\mu_k}$ reda 1 (to lahko potrdijo zbrani podatki), je v odvisnosti od sektorske dekompozicije korelacija istega reda kot izraz $\sqrt{\mu_A\,\mu_B}$. To je geometrijska sredina dveh verjetnosti neplačil. Zato v splošnem lahko pričakujemo, da bo korelacija med izostanki plačil istega reda kot verjetnost neplačil.

5.13 Numerična stabilnost

Algoritem za izračun verjetnosti $P(\text{izguba} = nL)$, ki ga uporablja model CreditRisk+, t.i. Panjerjeva rekurzijska shema, je numerično nestabilen. Panjerjevo rekurzijo izpeljemo ob predpostavki, da je odvod logaritma rodovne funkcije izgub $G(z)$ racionalna funkcija $\frac{P(z)}{Q(z)}$, kjer sta $P(z)$ in $Q(z)$ polinoma. Vzrok numerične nestabilnosti je v kopičenju napak zaradi zaokroževanja pri seštevanju nasprotno predznačenih števil. Oba polinoma, $P(z)$ in $Q(z)$, pa imata pozitivne in negativne koeficiente.

5.13.1 Zaokrožitvene napake

Naj bosta ε_x in ε_y relativni napaki x in y, ε_{x+y} pa relativna napaka seštevanja. Velja

$$\varepsilon_{x+y} = \frac{x}{x+y}\,\varepsilon_x + \frac{y}{x+y}\,\varepsilon_y, \quad \text{če } x + y \neq 0.$$

Če sta sumanda x in y istega predznaka, velja ocena:

$$|\varepsilon_{x+y}| \leq 2\, \max\{|\varepsilon_x|, |\varepsilon_y|\}.$$

Če pa sta različnega predznaka, je vsaj eden od izrazov $\left|\frac{x}{x+y}\right|$ in $\left|\frac{y}{x+y}\right|$ večji od 1 in zato se relativna napaka vsote ε_{x+y} poveča. Če je $x \approx -y$, izraz $x+y$ v imenovalcu izgine in tedaj je relativna napaka ε_{x+y} največja.
Vidimo, da seštevanje dveh števil istega predznaka ne povzroči opaznega večanja napake. Tudi če večkrat ponovimo operacijo, se prvotna napaka ne poveča. Po drugi strani, če se pri ponavljajočem seštevanju (npr. pri rekurziji) samo enkrat zgodi, da sta sumanda podobne velikosti, vendar nasprotnega predznaka, je lahko relativna napaka nekontrolirano velika.
Dalje, relativna napaka zaradi množenja je približno enaka

$$\varepsilon_{x\,y} \approx \varepsilon_x + \varepsilon_y.$$

Relativni napaki argumentov x in y se torej seštejeta.
Zaključimo lahko, da je rekurziven algoritem, ki temelji izključno na seštevanju in množenju nenegativnih števil, numerično stabilen.

5.13.2 Numerično stabilen izračun rodovne funkcije izgub

Za začetek nekoliko prilagodimo oznake. Kot prvo, polinom iz enakosti (5.56) pišimo kot

$$P_k(z) = \sum_{j=1}^{m(k)} \left(\frac{\varepsilon_j^{(k)}}{\nu_j^{(k)}}\right) z^{\nu_j^{(k)}} \tag{5.117}$$

oz., če nas zanima vsota po kreditojemalcih in ne po razredih

$$P_k(z) = \sum_{A\in S_k} \Theta_{Ak}\, \frac{\varepsilon_A}{\nu_A}\, z^{\nu_A} = \sum_{A\in S_k} \Theta_{Ak}\, p_A\, z^{\nu_A}. \tag{5.118}$$

Za nadaljnjo analizo je pomembno, da opazimo, da so koeficienti polinomov P_k nenegativni.
Rodovno funkcijo izgub $G(z)$ pišemo kot

$$G(z) = e^{-\sum_{k=1}^{n} \frac{\mu_k^2}{\sigma_k^2} \ln\left(1+\frac{\sigma_k^2}{\mu_k^2} P_k(1) - \frac{\sigma_k^2}{\mu_k^2} P_k(z)\right)}, \tag{5.119}$$

kar je enako

$$G(z) = \prod_{k=1}^{n} \left(\frac{1}{1+\frac{\sigma_k^2}{\mu_k^2} P_k(1) - \frac{\sigma_k^2}{\mu_k^2} P_k(z)}\right)^{\frac{\mu_k^2}{\sigma_k^2}}. \tag{5.120}$$

Dokažimo, da je rodovna funkcija izgub $G(z)$ iz zveze (5.60)

$$G(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-\frac{p_k}{\mu_k} P_k(z)} \right)^{\alpha_k}$$

res enaka tej iz zveze (5.120).
Spomnimo se naslednjih enakosti:

$$p_k = \frac{\beta_k}{1+\beta_k}, \quad \alpha_k = \frac{\mu_k^2}{\sigma_k^2}, \quad \beta_k = \frac{\sigma_k^2}{\mu_k}.$$

Od tod sledi

$$\beta_k = \frac{p_k}{1-p_k} \quad \text{in} \quad \frac{p_k}{(1-p_k)\,\mu_k} = \frac{\beta_k}{\mu_k} = \frac{\sigma_k^2}{\mu_k^2}.$$

Sedaj lahko pišemo

$$G(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-\frac{p_k}{\mu_k} P_k(z)} \right)^{\alpha_k} = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-p_k+p_k-\frac{p_k}{\mu_k} P_k(z)} \right)^{\frac{\mu_k^2}{\sigma_k^2}} =$$

$$= \prod_{k=1}^{n} \left(\frac{1-p_k}{1-p_k+\frac{p_k}{\mu_k} P_k(1)-\frac{p_k}{\mu_k} P_k(z)} \right)^{\frac{\mu_k^2}{\sigma_k^2}} =$$

$$= \prod_{k=1}^{n} \left(\frac{1}{1+\frac{p_k}{\mu_k(1-p_k)} P_k(1)-\frac{p_k}{\mu_k(1-p_k)} P_k(z)} \right)^{\frac{\mu_k^2}{\sigma_k^2}} =$$

$$= \prod_{k=1}^{n} \left(\frac{1}{1+\frac{\sigma_k^2}{\mu_k^2} P_k(1)-\frac{\sigma_k^2}{\mu_k^2} P_k(z)} \right)^{\frac{\mu_k^2}{\sigma_k^2}},$$

kar je natanko rodovna funkcija izgub iz enakosti (5.120).
Enakost (5.119) lahko interpretiramo kot razvoj analitične funkcije $G(z)$ v vrsto okrog $z = 0$ s konvergenčnim polmerom strogo večjim od 1. Zato je zelo naravno, da koeficiente, tj. verjetnosti $P(\text{izguba} = nL)$, izračunamo neposredno z uporabo standardnih algoritmov za logaritem in eksponentno funkcijo vrste. Razvili bomo metodo za izračun koeficientov razvoja v vrsto enakosti (5.120) in predstavili rekurzijo v dveh korakih. Predznaki koeficientov so taki, da lahko s pomočjo dveh lem zagotovimo numerično stabilnost rekurzivne sheme.

Lema 5.13.1 : *Dano je zaporedje* $\{a_k\}_k$ *z* $a_0 > 0$ *in* $a_k \geq 0 \quad \forall k \geq 1$. *Naj bo* $g(z) := -\ln(a_0 - f(z))$, *kjer je* $f(z) := \sum_{k=1}^{\infty} a_k z^k$. *Privzemimo, da ima* f *pozitivni konvergenčni polmer. Funkcija* g *je zato analitična na disku* $\{z; |z| < R\}$ *za nek* $R > 0$ *in jo lahko razvijemo v vrsto* $g(z) = \sum_{k=0}^{\infty} b_k z^k$ *na tem disku. Tedaj za koeficiente* b_k *velja*

$$b_k \geq 0 \quad \text{za} \quad k \geq 1$$

in rekurzivna shema

$$b_0 = -\ln(a_0),$$

$$b_k = \frac{1}{a_0}\left(a_k + \frac{1}{k}\sum_{l=1}^{k-1} l\, b_l\, a_{k-l}\right)$$

je numerično stabilna.

Dokaz : Z odvajanjem funkcije g dobimo

$$g'(z) = \frac{f'(z)}{a_0 - f(z)}.$$

Od tod

$$\left(a_0 - \sum_{k=1}^{\infty} a_k\, z^k\right) \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, z^k = \sum_{k=0}^{\infty} (k+1)\, a_{k+1}\, z^k.$$

Računajmo

$$\sum_{k=1}^{\infty} a_0\, (k+1)\, b_{k+1}\, z^k - \sum_{k=1}^{\infty} a_k\, z^k \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, z^k = \sum_{k=0}^{\infty} (k+1)\, a_{k+1}\, z^k.$$

Izenačimo koeficienta pred z^k leve in desne strani enačbe:

$$a_0\, (k+1)\, b_{k+1} - \sum_{j=1}^{k} (k+1-j)\, a_j\, b_{k+1-j} = (k+1)\, a_{k+1}.$$

Dalje

$$a_0\, k\, b_k = k\, a_k + \sum_{j=1}^{k-1} (k-j)\, a_j\, b_{k-j} \quad \text{oz.}$$

$$a_0\, k\, b_k = k\, a_k + \sum_{l=1}^{k-1} l\, a_{k-l}\, b_l.$$

In končno

$$b_k = \frac{1}{a_0}\left(a_k + \frac{1}{k}\sum_{l=1}^{k-1} l\, b_l\, a_{k-l}\right).$$

Prvi člen dobimo tako, da vstavimo $z = 0$:

$$b_0 = g(0) = -\ln(a_0).$$

Zaradi privzetka o zaporedju $\{a_k\}_k$ sledi, da je $b_k \geq 0$ za $k \geq 1$. Torej je res rekurzivni izračun koeficientov $\{b_k\}_k$ numerično stabilen.

□

Lema 5.13.2 : *Naj bo* $f(z) := \sum_{k=0}^{\infty} a_k z^k$ *in* $g(z) := e^{f(z)} = \sum_{n=0}^{\infty} b_n z^n$ *na disku* $\{z; |z| < R\}$ *za nek* $R > 0$. *Tedaj velja*

$$b_0 = e^{a_0},$$

$$b_n = \sum_{k=1}^{n} \frac{k}{n}\, b_{n-k}\, a_k \quad \text{za} \quad n \geq 1.$$

Rekurzija je numerično stabilna, če so koeficienti a_k *nenegativni za* $k \geq 1$.

Dokaz : Relacijo $b_0 = e^{a_0}$ dobimo, ko v g vstavimo $z = 0$.
Za j-ti odvod v točki $z = 0$ velja

$$f^{(j)}(0) = j!\, a_j \quad \text{in} \quad g^{(j)} = j!\, b_j. \tag{5.121}$$

Zapišimo še n-ti odvod funkcije g:

$$g^{(n)}(z) = \frac{d^n}{dz^n}\, e^{f(z)} = \frac{d^{n-1}}{dz^{n-1}}\Big(g(z)\, f'(z)\Big).$$

Od tod po Leibnitzovem pravilu za višje odvode produkta dobimo

$$g^{(n)}(z) = \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k+1)}(z)\, g^{(n-k-1)}(z). \tag{5.122}$$

V enakost (5.122) vstavimo $z = 0$ in upoštevamo zvezo (5.121):

$$n!\, b_n = \sum_{k=0}^{n-1} \binom{n-1}{k} (k+1)!\, a_{k+1}\, (n-k-1)!\, b_{n-k-1}.$$

Sledi

$$b_n = \sum_{k=1}^{n} \frac{k}{n}\, a_k\, b_{n-k}.$$

Rekurzija je numerično stabilna, saj iz $a_k \geq 0$ za $k \geq 1$ in $b_0 > 0$ sledi $b_n \geq 0$.

□

Porazdelitev portfeljske izgube lahko s pomočjo rekurzivnih zvez iz zgornjih lem sedaj izračunamo takole:

Definiramo

$$a_0^{(k)} := 1 + \frac{\sigma_k^2}{\mu_k^2} P_k(1),$$

$$a_j^{(k)} := \sum_A \Theta_{Ak}\, p_A\, 1_{\{\nu_j = j\}}, \quad j = 1, ..., M, \quad k = 1, ..., n$$

za vnaprej določen red M. $a_j^{(k)}$ so koeficienti polinoma $\frac{\sigma_k^2}{\mu_k^2} P_k(z)$, s katerim izrazimo funkcijo $g_k(z)$:

$$g_k(z) := -\ln(1 + \frac{\sigma_k^2}{\mu_k^2} P_k(1) - \frac{\sigma_k^2}{\mu_k^2} P_k(z)).$$

Po lemi 5.13.1 lahko funkcijo g_k razvijemo v vrsto

$$g_k(z) = \sum_{j=1}^{M} \beta_j z^j + O(z^{M+1}).$$

Spomnimo se, da lema 5.13.1 zagotavlja, da je $\beta_j \geq 0$ za $j \geq 1$. V naslednjem koraku definiramo funkcijo $G_k(z)$

$$G_k(z) := e^{g_k(z)},$$

ki je analitična in zato oblike

$$G_k(z) = \sum_{n=0}^{M} \gamma_n\, z^n + O(z^{M+1}).$$

Po lemi 5.13.2 koeficiente $\gamma_n, n = 0, ..., M$, dobimo s pomočjo rekurzivne zveze iz koeficientov $\beta_j, j = 0, ..., M$.
$G_k(z)$ je rodovna funkcija izgub k-tega sektorja. Da dobimo koeficiente rodovne funkcije $G(z)$ celotnega portfelja, moramo med seboj zmnožiti vse $G_k(z)$.
Numerična stabilnost algoritma sledi iz lem zaradi predznaka koeficientov $a_j^{(k)}$ in β_j. Koeficienti γ_n ustrezajo verjetnostim $P(\text{izguba} = nL)$ in so točni do $n = M$.

5.13.3 Algoritem

V nadaljevanju si bomo ogledali postopek izračuna porazdelitve portfeljske izgube po zgornjem algoritmu. Vhodni parametri so podatki o portfelju, na izhodu pa dobimo verjetnosti izgub $P(\text{izguba} = nL)$.

1. **Podatki:**
   število kreditojemalcev: `nK`
   število sektorjev: `nS`
   kredit: `kredit(nK)`[2]
   verjetnost neplačila dolga: `verjNepl(nK)`
   standardna deviacija: `stdDev(nK)`
   utež sektorja: `utezSekt(nK,nS)`
   enota zneska: `L`
   stopnja vmesnega polinoma: `stP`

2. **Spremenljivke:**
   stopnja zadnjega polinoma: `stZ`
   kreditni razred: `razred(nK)`
   vrednost polinoma P_k v z=1: `polinom1(nS)`
   matematično upanje sektorja: `matUpanjeSekt(nS)`
   standardna deviacija sektorja: `stdDevSekt(nS)`
   verjetnost izgube: `verjetnost(nS,stZ)`
   začetni koeficienti: `alfa(nS,stP)`
   koeficienti, dobljeni iz prve rekurzivne zveze: `beta(nS,stP)`[3]
   koeficienti, dobljeni iz druge rekurzivne zveze: `gama(nS,stZ)`[4]
   kvocient med standardno deviacijo in matematičnim upanjem sektorja:
   `kvocient(nS)`

3. **Postopek:**

```
stZ := nS·stP;
za k = 1 po 1 do nS
  matUpanjeSekt(k) := 0;
  stdDevSekt(k) := 0;
  za j = 1 po 1 do stP
    alfa(k,j) := 0;
```

[2]Oznaka `ime(n)` pomeni seznam dolžine n, `ime(n,m)` pa dvodimenzionalen seznam.
[3]Lema 5.13.1
[4]Lema 5.13.2

```
    konec za
konec za

za i = 1 po 1 do nK
    razred(i) := kredit(i)/L;
    če razred(i) == 0 potem
        razred(i) := 1;
    konec če
    za k = 1 po 1 do nS
        matUpanjeSekt(k) += utezSekt(i,k)·verjNepl(i);
        stdDevSekt(k) += utezSekt(i,k)·stdDev(i);
    konec za
konec za

za k = 1 po 1 do nS
    če matUpanjeSekt(k) == 0 potem
        zaključi[5];
    sicer
        (kvocient(k) := stdDevSekt(k)/matUpanjeSekt(k))^2;
        polinom1(k) := matUpanjeSekt(k);
    konec če
konec za

za k = 1 po 1 do nS
    alfa(k,1) := 1 + kvocient(k)·polinom1(k);
konec za

za i = 1 po 1 do nK
    za j = 2 po 1 do stP
        če razred(i) == j-1 potem
            alfa(k,j) += kvocient(k)·utezSekt(i,k)·verjNepl(i,k);
        konec če
    konec za
konec za

za k = 1 po 1 do nS
    beta(k,1) := - log(alfa(k,1))/kvocient(k);
```

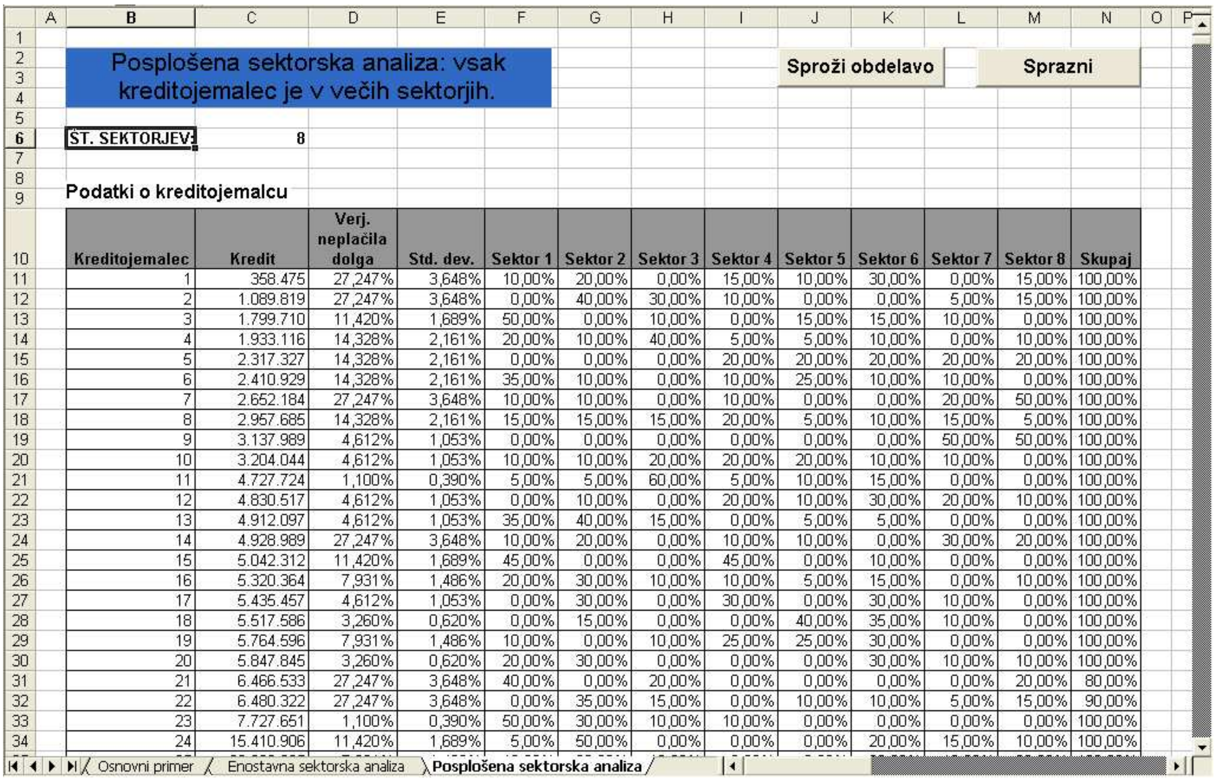
[5] Program se ustavi, če je matematično upanje katerega od sektorjev enako nič. V tem primeru so namreč uteži za vse kreditojemalce enake nič, kar ni v skladu z navodili.

```
    za j = 2 po 1 do stP
      beta(k,j) := alfa(k,j)/alfa(k,1);
      za l = 2 po 1 do j-1
        beta(k,j) +=
          (1/(j-1)·alfa(k,1))·(l-1)·beta(k,l)·alfa(k,j-l+1);
      konec za
    konec za
  konec za

  za k = 1 po 1 do nS
    za j = 2 po 1 do stP
      (beta(k,j) /= kvocient(k));
    konec za
  konec za

  za k = 1 po 1 do nS
    gama(k,1) := e^beta(k,1);
    za j = 2 po 1 do stP
      za l = 2 po 1 do j
        gama(k,j) += (l-1)/(j-1)·gama(k,j-l+1)·beta(k,l);
      konec za
    konec za
    za j = stP+1 po 1 do stZ
      gama(k,j) := 0;
    konec za
  konec za

  za j = 1 po 1 do stZ
    verjetnost(1,j) := gama(1,j);
  konec za

  za k = 1 po 1 do nS
    za j = 1 po 1 do stZ
      za l = 1 po 1 do j
        verjetnost(k,j) += verjetnost(k-1,l)·gama(k,j-l+1);
      konec za
    konec za
  konec za
```

4. **Rezultati:**
   Rezultati so shranjeni v spremenljivki verjetnost, kjer verjetnost(nS,m) pomeni P(izguba = mL).

Dodatek A

Ekranske slike demonstracijskega programa



Posplošena sektorska analiza: vsak kreditojemalec je v večih sektorjih.

Sproži obdelavo | Sprazni

ŠT. SEKTORJEV: 8

Podatki o kreditojemalcu

Kreditojemalec	Kredit	Verj. neplačila dolga	Std. dev.	Sektor 1	Sektor 2	Sektor 3	Sektor 4	Sektor 5	Sektor 6	Sektor 7	Sektor 8	Skupaj
1	358.475	27,247%	3,648%	10,00%	20,00%	0,00%	15,00%	10,00%	30,00%	0,00%	15,00%	100,00%
2	1.089.819	27,247%	3,648%	0,00%	40,00%	30,00%	10,00%	0,00%	0,00%	5,00%	15,00%	100,00%
3	1.799.710	11,420%	1,689%	50,00%	0,00%	10,00%	0,00%	15,00%	15,00%	10,00%	0,00%	100,00%
4	1.933.116	14,328%	2,161%	20,00%	10,00%	40,00%	5,00%	5,00%	10,00%	0,00%	10,00%	100,00%
5	2.317.327	14,328%	2,161%	0,00%	0,00%	0,00%	20,00%	20,00%	20,00%	20,00%	20,00%	100,00%
6	2.410.929	14,328%	2,161%	35,00%	10,00%	0,00%	10,00%	25,00%	10,00%	10,00%	0,00%	100,00%
7	2.652.184	27,247%	3,648%	10,00%	10,00%	0,00%	10,00%	0,00%	0,00%	20,00%	50,00%	100,00%
8	2.957.685	14,328%	2,161%	15,00%	15,00%	15,00%	20,00%	5,00%	10,00%	15,00%	5,00%	100,00%
9	3.137.989	4,612%	1,053%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	50,00%	50,00%	100,00%
10	3.204.044	4,612%	1,053%	10,00%	10,00%	20,00%	20,00%	20,00%	10,00%	10,00%	0,00%	100,00%
11	4.727.724	1,100%	0,390%	5,00%	5,00%	60,00%	5,00%	10,00%	15,00%	0,00%	0,00%	100,00%
12	4.830.517	4,612%	1,053%	0,00%	10,00%	0,00%	20,00%	10,00%	30,00%	20,00%	10,00%	100,00%
13	4.912.097	4,612%	1,053%	35,00%	40,00%	15,00%	0,00%	5,00%	5,00%	0,00%	0,00%	100,00%
14	4.928.989	27,247%	3,648%	10,00%	20,00%	0,00%	10,00%	10,00%	0,00%	30,00%	20,00%	100,00%
15	5.042.312	11,420%	1,689%	45,00%	0,00%	0,00%	45,00%	0,00%	10,00%	0,00%	0,00%	100,00%
16	5.320.364	7,931%	1,486%	20,00%	30,00%	10,00%	10,00%	5,00%	15,00%	0,00%	10,00%	100,00%
17	5.435.457	4,612%	1,053%	0,00%	30,00%	0,00%	30,00%	0,00%	30,00%	10,00%	0,00%	100,00%
18	5.517.586	3,260%	0,620%	0,00%	15,00%	0,00%	0,00%	40,00%	35,00%	10,00%	0,00%	100,00%
19	5.764.596	7,931%	1,486%	10,00%	0,00%	10,00%	25,00%	25,00%	30,00%	0,00%	0,00%	100,00%
20	5.847.845	3,260%	0,620%	20,00%	30,00%	0,00%	0,00%	0,00%	30,00%	10,00%	10,00%	100,00%
21	6.466.533	27,247%	3,648%	40,00%	0,00%	20,00%	0,00%	0,00%	0,00%	0,00%	20,00%	80,00%
22	6.480.322	27,247%	3,648%	0,00%	35,00%	15,00%	0,00%	10,00%	10,00%	5,00%	15,00%	90,00%
23	7.727.651	1,100%	0,390%	50,00%	30,00%	10,00%	10,00%	0,00%	0,00%	0,00%	0,00%	100,00%
24	15.410.906	11,420%	1,689%	5,00%	50,00%	0,00%	0,00%	0,00%	20,00%	15,00%	10,00%	100,00%

Osnovni primer | Enostavna sektorska analiza | Posplošena sektorska analiza

Slika A.1: Podatki o kreditojemalcu

Porazdelitev izgube

Izguba	Verjetnost
0	0,04927228
200.000	0,00000000
400.000	0,01329838
600.000	0,00000000
800.000	0,00180314
1.000.000	0,01329375
1.200.000	0,00016377
1.400.000	0,00359744
1.600.000	0,00001121
1.800.000	0,00606519
2.000.000	0,00880254
2.200.000	0,00155358
2.400.000	0,01638367
2.600.000	0,01350365
2.800.000	0,00562072
3.000.000	0,01267381
3.200.000	0,00546069
3.400.000	0,00672558
3.600.000	0,00523933
3.800.000	0,00240465
4.000.000	0,00389970
4.200.000	0,00325439
4.400.000	0,00489672
4.600.000	0,00328184
4.800.000	0,00701897
5.000.000	0,02761880
5.200.000	0,00477995
5.400.000	0,01649160
5.600.000	0,00644169
5.800.000	0,01175883
6.000.000	0,00940325

Izguba na danem percentilu

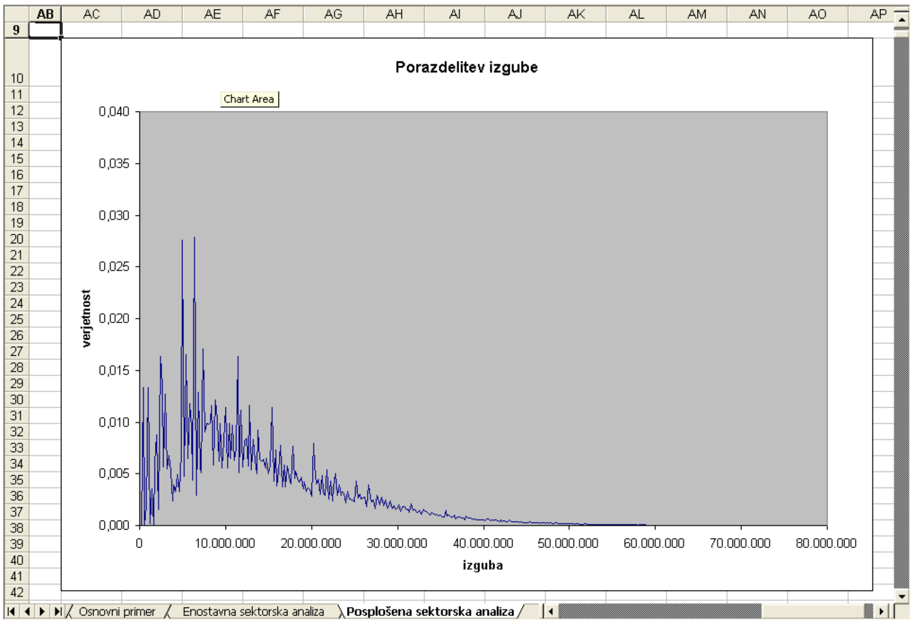
Percentil	Izguba
Povprecje	13.891.762
50	11.200.000
75	18.800.000
95	33.400.000
97,5	38.800.000
99	45.800.000
99,5	50.800.000
99,75	55.400.000
99,9	61.600.000

Prispevek k tveganju kreditnega portfelja

Kreditojemalec	Pričakovana izguba	Prisp. k tveganju kred. portfelja
1	97.673,19	3.727,52
2	296.941,47	32.022,87
3	205.529,82	35.892,85
4	276.973,36	51.856,76
5	332.022,41	74.107,00
6	345.433,54	80.326,89
7	722.636,90	184.976,02
8	423.771,75	120.561,93
9	144.734,40	43.674,34
10	147.781,08	45.425,53
11	52.007,11	23.503,67
12	222.799,37	102.970,48
13	226.562,11	106.734,98
14	1.342.994,80	633.772,26
15	575.840,27	277.671,83
16	421.962,70	214.993,38
17	250.701,20	130.391,79
18	179.850,98	94.822,62
19	457.195,13	251.576,47
20	190.616,09	106.727,32
21	1.761.927,28	1.087.377,26
22	1.765.684,35	1.093.309,55
23	85.007,67	62.757,36
24	1.759.950,64	2.583.687,52
25	1.605.164,39	3.091.138,53

Osnovni primer / Enostavna sektorska analiza / Posplošena sektorska ana

Slika A.2: Porazdelitev izgube, izguba na danem percentilu, prispevek k tveganju kreditnega portfelja



Slika A.3: Graf porazdelitve izgube

Dodatek B

Slovar

default - nesposobnost plačila dolga
economic capital - ekonomski kapital
expected loss - pričakovana izguba
expected shortfall - pričakovan izpad
exposure at default - izpostava ob nastanku plačilne nesposobnosti kreditojemalca
loss given default - izguba ob nastopu izostanka plačila dolga
probability of default - verjetnost neplačila dolga
risk capital - kapital za tveganje
risk contribution - prispevek k tveganju kreditnega portfelja
severity of loss - velikost izgube ob nastopu izostanka odplačila dolga
tail conditional expectation - pogojno matematično upanje v repu porazdelitve
underlying random variables - slučajne spremenljivke, ki vplivajo na spremembe
unexpected loss - nepričakovana izguba
Value-at-Risk (VaR) - tvegana vrednost

Literatura

[1] Bluhm C., Overbeck L., Wagner C., An Introduction to Credit Risk Modeling, Chapman & Hall/CRC Financial Mathematics Series, 2003

[2] CREDIT SUISSE FIRST BOSTON, CreditRisk+ : A Credit Risk Management Framework, http://www.csfb.com/creditrisk, 1997

[3] Haaf H., Reiss O., Schoenmakers J., Numerically Stable Computation of CreditRisk+, http://www.wias-berlin.de/publications/preprints/846/wias_preprints_846.pdf, 2003

[4] http://www.bsi.si/html/basel2/default.htm

[5] Jamnik R., Matematična statistika, DZS, Ljubljana, 1984

[6] Mann P.S., Statistics for Business and Economics, John Wiley & Sons, Inc., 1995

[7] Ribnikar I., Monetarna ekonomija I (Denar, finančne institucije in denarna politika), Ljubljana: Ekonomska fakulteta, 2003

[8] Rice J.A., Mathematical Statistics and Data Analysis, Duxbury Press, 1995

[9] Sušnik S., Predlagane spremembe mednarodnih standardov kapitala in kapitalske ustreznosti bank, http://www.bsi.si/html/basel2/03_aktivnosti/dokumenti/Posvet-Portoroz1_bs-nbp.pdf, 2000

[10] Tasche D., Conditional Expectation as Quantile Derivative, http://www-m4.mathematik.tu-muenchen.de/m4/pers/tasche/quant.pdf, 2000