Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Matematički odsjek

Tončica Topić

Bayesova statistika i procjena vrijednosti ulaganja

Diplomski rad

Voditelj rada:
prof.dr.sc. Siniša Slijepčević

Zagreb, studeni 2016.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ______________________________ , predsjednik
2. ______________________________ , član
3. ______________________________ , član

Povjerenstvo je rad ocijenilo ocjenom ______________________ .

Potpisi članova povjerenstva:

1. ______________________________
2. ______________________________
3. ______________________________

Obitelji i prijateljima koji su mi uljepšali studentske dane.
Posebno hvala izvrsnom mentoru prof. dr. sc. Siniši Slijepčeviću.

Sadržaj

1 Uvod

Financijski modeli, koristeći matematički aparat, nastoje što preciznije opisati stanje na svjetskim financijskim tržištima. Pretpostavke su da je model valjan bez obzira koji vremenski period ili vrstu imovine promatrali. Glavni cilj modela je izvući sve relevantne činjenice iz povijesnih podataka, koji, osim važnih informacija, sadrže i slučajnost. Upravo ta slučajnost, koja u konačnici implicira neizvjesnost, stavlja imperativ na korištenje statističkih alata pri stvaranju, procjeni i testiranju modela.

Postoje dva pristupa statističkoj analizi: frekvencionistički i Bayesovski. Osnovna razlika među njima leži u načinu na koji se interpretira pojam vjerojatnosti. Kao što samo ime kaže, frekvencionistička statistika, poznata kao i klasična, vjerojatnost tretira kao limes relativnih frekvencija. Odmah možemo primijetiti da potpuno pridržavanje ovom pristupu nije uvijek moguće u praksi, npr. kada proučavamo događaje koji se rijetko pojavljuju. U tom slučaju, zbog nedovoljne veličine uzorka, prisiljeni smo se oslanjati na teorijske rezultate. S druge strane, zagovornici Bayesovskog pristupa smatraju da je vjerojatnost subjektivna, odnosno da je ona stupanj uvjerenja koji se mijenja dolaskom novih informacija.

Usko vezan uz interpretaciju vjerojatnosti je pojam neizvjesnosti. Pristaše klasičnog pristupa vjeruju kako je izvor neizvjesnosti isključivo realizacija slučajnih varijabli, dok Bayesova statistika ide korak dalje i smatra tu slučajnost nedovoljnom. Stoga kroz parametre modela ugrađuje dodatnu neizvjesnost na način da i njih same promatra kao slučajne (za razliku od ,,klasičara" koji ih uzimaju kao egzaktne).

Bayesova statistika dobila je ime po engleskom matematičaru i teologu Thomasu Bayesu koji je prvi dokazao specijalan slučaj teorema danas poznatog kao Bayesov teorem. Njegov rad "Essay Towards Solving a Problem in the Doctrine of Chances" objavljen je posthumno 1763. godine od strane Royal Societyja. Bez znanja o radu Thomasa Bayesa, francuski matematičar i astronom Pierre-Simon Laplace otkriva Bayesov teorem (1774.), ali u općenitijem i jasnijem obliku te ga sljedećih 40 godina primjenjuje na području statistike, meteorologije, astronomije pa čak i geodezije. Početkom prošlog stoljeća, kada su kvantitativne metode postale najkorišteniji alat pri analiziranju financijskih tržišta i oblikovanju investicijskih strategija, glavnu ulogu igrala je klasična statistika, odnosno frekvencionistički pristup, za što su zaslužni Ronald A. Fisher, Jerzy Neyman i Egon Pearson. Unatoč tome, neki matematičari nastavljaju razvijati Laplaceove ideje zbog čega, nakon nekog vremena, nastaju dvije struje u Bayesovoj statistici: objektivistička i subjektivistička. Zagovornici objektivističke struje smatraju da Bayesovska statistička analiza počiva samo na pretpostavljenom modelu i proučenim podatcima. Subjektivisti s druge strane tvrde da je takva analiza nepotpuna bez subjektivnih uvjerenja. Bayesova statistika dobila je na važnosti tek 80-ih godina dvadesetog stoljea, uglavnom zbog naglog razvoja računskih metoda.

Cilj ovog rada je dati pregled teorije koja stoji iza Bayesova statistike i pokazati na koji način se može primijeniti u financijskoj industriji, točnije pri procjeni vrijednosti ulaganja.

2 Osnovni pojmovi teorije vjerojatnosti

Da bismo u potpunosti razumjeli Bayesovu paradigmu potrebno je definirati matematičku strukturu na kojoj ona počiva. Pokazuje se da je za to najprikladniji aparat teorije vjerojatnosti. Stoga poglavlje počinjemo Kolmongorovim aksiomima koji će nas prirodno dovesti do Bayesovog teorema.

2.1 Vjerojatnosni prostor. Uvjetna vjerojatnost

Neka je Ω proizvoljan neprazan skup. Sa $\mathcal{P}(\Omega)$ označimo partitivni skup od Ω.

Definicija 2.1. *Familija $\mathcal{F}$ podskupova od Ω ($\mathcal{F} \subseteq \mathcal{P}(\Omega)$) jest σ-algebra skupova (na Ω) ako je*

F1. $\emptyset \in \mathcal{F}$.

F2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.

F3. $A_i \in \mathcal{F}, i \in \mathbb{N} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definicija 2.2. *Neka je $\mathcal{F}$ σ-algebra na skupu Ω. Uređen par $(\Omega, \mathcal{F})$ zove se izmjeriv prostor.*

Definicija 2.3. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor. Funkcija $P : \mathcal{F} \to \mathbb{R}$ jest vjerojatnost (na $\mathcal{F}$, na Ω) ako vrijedi*

P1. $P(A) \geq 0, A \in \mathcal{F}$.

P2. $P(\Omega) = 1$.

P3. $A_i \in \mathcal{F}, i \in \mathbb{N}$ *i* $A_i \cap A_j = \emptyset$ *za* $i \neq j$

$$(A_i \text{ - međusobno disjunktni}) \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

Definicija 2.4. *Uređena trojka $(\Omega, \mathcal{F}, P)$, gdje je $\mathcal{F}$ σ-algebra na Ω i P vjerojatnost na $\mathcal{F}$, zove se vjerojatnosni prostor.*

Sada ćemo navesti neke osnovne karakteristike funkcije vjerojatnosti koje slijede iz definicije vjerojatnosnog prostora. Elemente σ-algebre $\mathcal{F}$ zovemo događaji, a broj $P(A)$, $A \in \mathcal{F}$, vjerojatnost događaja A.

Teorem 2.5. *Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor. Tada vrijedi* *(a)* $P(\emptyset) = 0$

(b) Ako su $A_i, ..., A_n \in \mathcal{F}$ *međusobno disjunktni, tada je*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

(c) $A,\ B \in \mathcal{F}$ *i* $A \subseteq B \Rightarrow P(A) \leq P(B)$

(d) $A_n \in \mathcal{F},\ n \in \mathbb{N} \Rightarrow P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$

(e) $A,\ B \in \mathcal{F} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(f) $A \in \mathcal{F} \Rightarrow P(A^c) = 1 - P(A)$.

Dokaz. Pogledati [6]. □

Definicija 2.6. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor i* $A,\ B \in \mathcal{F}$. *Događaji* A *i* B *su nezavisni ako vrijedi*

$$P(A \cap B) = P(A)P(B). \tag{1}$$

Definicija 2.7. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor i* $A \in \mathcal{F}$ *takav da je* $P(A) > 0$. *Funkcija* $P_A : \mathcal{F} \to [0, 1]$ *definirana s*

$$P_A(B) = P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad B \in \mathcal{F} \tag{2}$$

je vjerojatnost na $\mathcal{F}$ *i zovemo je uvjetna vjerojatnost uz uvjet* A. *Broj* $P(B|A)$ *zovemo vjerojatnost od* B *uz uvjet da se* A *dogodio.*

Proširimo definiciju uvjetne vjerojatnosti uz uvjet A ($A \in \mathcal{F}$) na slučaj kada je $P(A) = 0$ tako da stavimo

$$P(B|A) = P(B) \quad \textit{ako je} \quad P(A) = 0, \quad B \in \mathcal{F}. \tag{3}$$

Iz (1) i (3) slijedi

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A), \quad A, B \in \mathcal{F}. \tag{4}$$

Budući da je P_A vjerojatnost na $\mathcal{F}$, iz teorema 2.5 slijedi

Propozicija 2.8. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor i* $A \in \mathcal{F}$ *takav da je* $P(A) > 0$. *Tada vrijedi* *(a)* $P(\emptyset|A) = 0$

(b) $P(B^c|A) = 1 - P(B|A),\ B \in \mathcal{F}$

(c) $A_1, A_2 \in \mathcal{F} \Rightarrow P(A_1 \cup A_2|A) = P(A_1|A) + P(A_2|A) - P(A_1 \cap A_2|A)$

(d) $A_1, A_2 \in \mathcal{F}\ \&\ A_1 \subseteq A_2 \Rightarrow P(A_1|A) \leq P(A_2|A)$

(e) $A_n \in \mathcal{F},\ n \in \mathbb{N} \Rightarrow P(\bigcup_{n=1}^{\infty} A_n|A) \leq \sum_{n=1}^{\infty} P(A_n|A)$.

2.2 Slučajne varijable

Vjerojatnosni prostor $(\Omega, \mathcal{F}, P)$ osnovni je pojam u teoriji vjerojatnosti, a služi nam kao matematički model za proučavanje slučajnih odnosno neizvjesnih pokusa. Često nad rezultatima takvih pokusa provodimo nekakva mjerenja, tj. svakom rezultatu pridružujemo realne brojeve. Prema tome važno je promatrati realne funkcije na Ω, koje ćemo zvati slučajne varijable. Prije nego ih definiramo, uvest ćemo pojam Borelovih funkcija i skupova.

Neka je $\mathbb{R}$ skup svih realnih brojeva. Sa $\mathcal{B}$ označimo σ-algebru generiranu familijom svih otvorenih skupova na $\mathbb{R}$. $\mathcal{B}$ zovemo σ-algebra Borelovih skupova na $\mathbb{R}$, a elemente σ-algebre $\mathcal{B}$ zovemo Borelovi skupovi. Nadalje, funkcija $g : \mathbb{R} \to \mathbb{R}$ jest Borelova funkcija ako je $g^{-1}(B) \in \mathcal{B}$ za svako $B \in \mathcal{B}$, tj. ako je $g^{-1}(\mathcal{B}) \subseteq \mathcal{B}$. U slučaju k - dimenzionalnog realnog prostora Borelovu σ-algebru generiranu svim otvorenim skupovima na $\mathbb{R}^k$ označavamo s $\mathcal{B}(\mathbb{R}^k)$.

Definicija 2.9. *Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor. Slučajna varijabla $(k = 1)$, slučajni vektor $(k \geq 2)$, je izmjerivo preslikavanje $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, tj. preslikavanje $X : \Omega \to \mathbb{R}^k$ za koje vrijedi $X^{-1}(B) \in \mathcal{F}$ za svako $B \in \mathcal{B}(\mathbb{R}^k)$ odnosno $X^{-1}(\mathcal{B}(\mathbb{R}^k)) \subseteq \mathcal{F}$.*

Jedno od osnovnih obilježja slučajnih varijabli jesu pripadne funkcije distribucije s kojima se i najčešće efektivno operira. Kako bi izbjegli smetnje koje izaziva općenitost skupa Ω na kojem je slučajna varijabla definirana, uvest ćemo vjerojatnosni prostor karakterističan za nju.

Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor, $(\mathbb{R}, \mathcal{B})$ izmjeriv prostor i X slučajna varijabla na Ω. Funkciju $P_X : \mathcal{B} \to [0, 1]$ definiranu s

$$P_X(B) = P(X \in B) = P(X^{-1}(B)), \quad B \in \mathcal{B} \tag{5}$$

zovemo vjerojatnosna mjera inducirana sa X (ili zakon razdiobe od X), a vjerojatnosni prostor $(\mathbb{R}, \mathcal{B}, P_X)$ zovemo vjerojatnosni prostor induciran sa X. Lako je provjeriti da je P_X vjerojatnost na $\mathcal{B}$.

Definicija 2.10. *Neka je X slučajna varijabla na Ω. Funkcija distribucije od X jest funkcija $F_X : \mathbb{R} \to [0,1]$ definirana sa*

$$F_X(x) = P_X(\langle -\infty, x]) = P(X \in \langle -\infty, x]) = P(X \leq x), \quad x \in \mathbb{R}. \tag{6}$$

Način na koji računamo vjerojatnosti događaja vezanih za slučajnu varijablu ovisi o karakteru, tj. tipu slučajne varijable. Postoje dva glavna tipa slučajnih varijabli: diskretne i neprekidne.

Definicija 2.11. *Slučajna varijabla X je diskretna ukoliko postoji konačan ili prebrojiv skup $D \subseteq \mathbb{R}$ takav da je $P_X(D) = P(X \in D) = 1$.*

Diskretne slučajne varijable obično zadajemo tako da zadamo skup $D = \{x_1, x_2, ...\}$ i brojeve $p_n = P(X = x_n)$, što zapisujemo u obliku tablice

$$X = \begin{pmatrix} x_1 & x_2 & ... & x_n & ... \\ p_1 & p_2 & ... & p_n & ... \end{pmatrix}. \tag{7}$$

Tablicu (7) zovemo distribucija ili zakon razdiobe slučajne varijable X.

Definicija 2.12. *Neka je X slučajna varijabla na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ i neka je F_X njezina funkcija distribucije. Kažemo da je X apsolutno neprekidna ili, kraće, neprekidna slučajna varijabla ako postoji nenegativna realna Borelova funkcija $f : \mathbb{R} \to \mathbb{R}_+$ takva da je*

$$F_X(x) = \int_{-\infty}^{x} f(t) d\lambda(t), \quad x \in \mathbb{R}. \tag{8}$$

Funkcija f iz (8) zove se funkcija gustoće slučajne varijable X.

Integral u (8) je Lebesgueov integral funkcije f u odnosu na Lebesgueovu mjeru λ. (vidi H. Šikić [7]).

Sada ćemo priču proširiti na slučajne vektore počevši s karakterističnim vjerojatnosnim prostorom. Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i $X : \Omega \to \mathbb{R}^n$ n - dimenzionalni slučajni vektor. Tada je $X = (X_1, X_2, ..., X_n)$, pri čemu su $X_k : \Omega \to \mathbb{R}$, k = 1,2,...,n slučajne varijable.

Za $B \in \mathcal{B}(\mathbb{R}^n)$ stavimo

$$P_X(B) = P(X^{-1}(B)) = P(X \in B). \tag{9}$$

Relacijom (9) definirana je funkcija vjerojatnosti $P_X : \mathcal{B}(\mathbb{R}^n) \to [0,1]$. P_X zovemo vjerojatnosna mjera inducirana slučajnim vektorom X (ili zakon razdiobe od X). Prema tome, svakom n-dimenzionalnom slučajnom vektoru X na prirodan se način preko relacije (9) pridružuje vjerojatnosni prostor $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P_X)$ koji zovemo vjerojatnosni prostor induciran slučajnim vektorom X.

Definicija 2.13. *Neka je* $X : \Omega \to \mathbb{R}^n$ *slučajni vektor,* $X = (X_1, X_2, ..., X_n)$. *Funkcija distribucije od* X *jest funkcija* $F_X : \mathbb{R}^n \to [0,1]$ *definirana s*

$$F_X(x) = P_X(\langle -\infty, x]) = P(X \in \langle -\infty, x]) = P(X \leq x), \quad x \in \mathbb{R}^n. \tag{10}$$

Definicija 2.14. *Neka je* $X : \Omega \to \mathbb{R}^n$ n *- dimenzionalni slučajni vektor,* $X = (X_1, X_2, ..., X_n)$ *i neka je* F_X *njegova funkcija distribucije.*

X je diskretan slučajni vektor ako postoji konačan ili prebrojiv podskup E *od* $\mathbb{R}^n$ *takav da je* $P(X \in E) = 1$.

X je apsolutno neprekidan ili, kraće, neprekidan slučajni vektor ako postoji nenegativna Borelova funkcija $f : \mathbb{R}^n \to \mathbb{R}_+$ *takva da je*

$$F_X(x) = \int_{-\infty}^{x} f(t) d\lambda(t), \quad x \in \mathbb{R}^n. \tag{11}$$

Funkciju f u (11) zovemo funkcija gustoće slučajnog vektora X.

Neka je $X = (X_1, X_2, ..., X_n)$ slučajni vektor s funkcijom gustoće f i neka je $x_1 \in \mathbb{R}$. Tada vrijedi:

$$\begin{aligned} F_{X_1}(x_1) &= P(X_1 \leq x_1) = P(X_1 \leq x_1, X_2 \in \mathbb{R}, ..., X_n \in \mathbb{R}) \\ &= \int_{-\infty}^{x1} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(t_1, ..., t_n) dt_1 ... dt_n. \end{aligned}$$

Odavde prema definiciji neprekidne slučajne varijable slijedi da X_1 ima gustoću danu s

$$f_1(x_1) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_2, ..., x_n) dx_2 ... dx_n.$$

Analogno, svaka slučajna varijabla X_i ima gustoću f_i koja se dobije integriranjem funkcije f po svim varijablama osim x_i. Kažemo da su F_i i f_i marginalna funkcija distribucije odnosno gustoće slučajne varijable X_i, i = 1,...,n.

Nezavisnost događaja u vjerojatnosnom prostoru definirali smo u 2.1. Sada ćemo promotriti nezavisnost slučajnih varijabli.

Definicija 2.15. *Neka su* $X_1, ..., X_n$ *slučajne varijable na vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$. *Kažemo da su* $X_1, ..., X_n$ *nezavisne ako za proizvoljne* $B_i \in \mathcal{B}$ $(i = 1, ..., n)$ *vrijedi*

$$P(X_1 \in B_1, ..., X_n \in B_n) = P(\bigcap_{i=1}^{n} (X_i \in B_i)) = \prod_{i=1}^{n} P(X_i \in B_i). \tag{12}$$

Za kraj ovog dijela teorije vjerojatnosti, koji će nam služiti kao podloga Bayesovom teoremu pa kasnije i Bayesovoj paradigmi, definirat ćemo uvjetnu funkciju gustoće i zakon razdiobe slučajnih varijabli.

Definicija 2.16. *Neka je (X,Y) dvodimenzionalni slučajni vektor s funkcijom gustoće* $f = f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$. *Ako je* $y \in \mathbb{R}$ *takav da je* $f_Y > 0$, *onda definiramo*
a) Uvjetnu funkciju gustoće $f(\cdot|y) = f_{X|Y}(\cdot|y) : \mathbb{R} \to \mathbb{R}$ *slučajne varijable* X *uz uvjet* $Y = y$ *formulom*

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \tag{13}$$

b) Uvjenu razdiobu $P_{X|Y} : \mathcal{B} \to [0,1]$ *od* X *uz uvjet* $Y = y$ *s*

$$P_{X|Y}(X \in B|Y = y) = \int\limits_B f_{X|Y}(x|y)dx. \tag{14}$$

Ako u (14) stavimo $B = \langle \infty, x]$, $x \in \mathbb{R}$ dobijemo

$$P_{X|Y}(X \leq x|Y = y) = \int\limits_{-\infty}^{x} f_{X|Y}(t|y)dt, \quad f_Y(y) > 0. \tag{15}$$

Prema tome, za svako $x \in \mathbb{R}$ vrijedi

$$F_{X|Y}(x|y) = \int\limits_{-\infty}^{x} f_{X|Y}(t|y)dt, \quad f_Y(y) > 0 \tag{16}$$

čime smo dobili uvjetnu distribuciju od X za dano Y = y.

Napomena 2.17. *a) Ako je (X,Y) diskretan slučajni vektor tada vrijedi*

$$f_{X|Y} = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{P(X = x, Y = y)}{P(Y = y)} = P(X = x|Y = y), \quad f_Y(y) > 0.$$

b) Neka je (X,Y) neprekidan slučajni vektor. Primjetimo da je funkcija $P_{X|Y}(X \in B|Y = y)$ *definirana samo za* $f_Y > 0$. *Dokažimo da je, u slučaju neprekidnog vektora, vjerojatnost da* (X,Y) *padne u skup* $A = \{(x,y); f_Y(y) = 0\}$ *jednaka nula. Zaista, vrijedi*

$$P((X,Y) \in A) = \iint\limits_A f_{X,Y}(x,y)dxdy = \int\limits_{\{y; f_Y(y)=0\}} [\int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)dx]dy =$$

$$= \int\limits_{\{y; f_Y(y)=0\}} f_Y(y)dy = 0.$$

Prema tome možemo zanemariti one (x,y) *za koje (14) nije definirana.*

2.3 Bayesov teorem

Dosadašnja teorija oslanjala se na općeniti vjerojatnosni prostor. U ovom poglavlju ograničit ćemo se na slučaj konačnog ili prebrojivog skupa Ω i u tom svjetlu iskazati i dokazati Bayesov teorem.

Definicija 2.18. *Konačna ili prebrojiva familija* $(H_i, i = 1, 2, ...)$ *događaja u vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$ *jest potpun sistem događaja ako je* $H_i \neq \emptyset$ *za svako* i*,* $H_i \cap H_j = \emptyset$ *za* $i \neq j$ *(tj. događaji* H_i *se međusobno isključuju) i* $\bigcup\limits_i H_i = \Omega$*.*

Teorem 2.19 (Formula potpune vjerojatnosti)**.** *Neka je* $(H_i, i = 1, 2, ...)$ *potpun sistem događaja u vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$*. Tada za proizvoljno* $A \in \mathcal{F}$ *vrijedi*

$$P(A) = \sum_i P(H_i)P(A|H_i). \tag{17}$$

Dokaz. Za $A \in \mathcal{F}$ vrijedi

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap (\bigcup_i H_i)) = P(\bigcup_i (A \cap H_i)) \\ &= \sum_i P(A \cap H_i) = \sum_i P(H_i)P(A|H_i). \end{aligned}$$

□

Teorem 2.20 (Bayesova formula)**.** *Neka je* $(H_i, i = 1, 2, ...)$ *potpun sistem događaja u vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$ *i* $A \in \mathcal{F}$ *takav da je* $P(A) > 0$*. Tada za svako* i *vrijedi*

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum\limits_j P(H_j)P(A|H_j)}. \tag{18}$$

Dokaz. Iz (2), (4) i teorema 2.19 slijedi

$$P(H_i|A) = \frac{P(H_i \cap A)}{P(A)} = \frac{P(H_i)P(A|H_i)}{\sum\limits_j P(H_j)P(A|H_j)}$$

□

3 Bayesova paradigma

Postavili smo matematičke temelje na kojima ćemo graditi daljnju teoriju. Na kraju prošlog poglavlja iskazali smo i dokazali Bayesov teorem u slučaju diskretnog vjerojatnosnog prostora kojeg ćemo kasnije proširiti na slučaj općenitog skupa Ω. Nadalje, primjere slučajnih varijabli, odnosno distribucija uvodit ćemo po potrebi kroz razne primjene teorije. U nastavku se fokusiramo na ključne principe Bayesovske analize.

3.1 Funkcija vjerodostojnosti

Definicija 3.1. *Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i neka je X slučajna varijabla na Ω s funkcijom distribucije F. Kažemo da $X_1, X_2..., X_n$ čine slučajni uzorak duljine n iz populacije s funkcijom distribucije F ako su $X_1, X_2..., X_n$ nezavisne i jednako distribuirane slučajne varijable sa zajedničkom funkcijom distribucije F.*

Ako funkcija F ima gustoću f kažemo da je uzorak uzet iz gustoće f. Intuitivno slučajni uzorak duljine n odgovara nizu od n nezavisnih mjerenja slučajnog svojstva neke populacije koji se opisuje slučajnom varijablom X. Neka je $X = (X_1, X_2..., X_n)$ slučajni uzorak iz gustoće f. S $f(x|\theta) = f(x_1, ..., x_n|\theta)$ označavamo zajedničku funkciju gustoće slučajnog uzorka uz parametar $\theta \in \Theta$, gdje je Θ parametarski prostor.

Definicija 3.2. *Neka je $x = (x_1, x_2..., x_n)$ jedna realizacija slučajnog uzorka $X = (X_1, X_2..., X_n)$ sa zajedničkom funkcijom gustoće $f(x|\theta)$, $x \in \mathbb{R}^n$ te neka je $\theta \in \Theta$ nepoznat parametar. Tada je vjerodostojnost funkcija $L : \Theta \to \mathbb{R}$ definirana s*

$$L(\theta) = f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta). \tag{19}$$

Primjer 3.3 (Funkcija vjerodostojnosti Poissonove distribucije)**.** *Kažemo da slučajna varijabla ima Poissonovu distribuciju ili da je X Poissonova slučajna varijabla, što označavamo sa $X \sim P(\theta)$, $\theta \in R$, $\theta > 0$, ako joj je funkcija gustoće dana s*

$$P(X = k) = \frac{\theta^k}{k!} e^{-\theta}, \quad k = 0, 1, 2, ... \quad . \tag{20}$$

Poissonova distribucija opisuje slučajan broj događaja u nekom vremenskom periodu, pri čemu je parametar θ stopa učestalosti odnosno prosjek događaja po jedinici vremena.

Zanima nas godišnji broj defaulta sjevernoameričkih korporacija koje su izdale obveznice te smo prikupili informacije za posljednjih 20 godina. Nadalje, pretpostavimo da spomenuti defaulti prate Poissonovu razdiobu, $X \sim P(\theta)$. Sa $x_i, i = 1, 2, ..., 20$, označimo prikupljene podatke. Tada je funkcija vjerodostojnosti za nepoznati parametar θ oblika:

$$\begin{aligned} L(\theta|x_1, ..., x_{20}) = \prod_{i=1}^{20} P(X = x_i|\theta) &= \prod_{i=1}^{20} \frac{\theta^{x_i}}{x_i!} e^{-\theta} \\ &= \frac{\theta^{\sum_{i=1}^{20} x_i}}{\prod_{i=1}^{20} x_i!} e^{-20\theta}. \end{aligned} \tag{21}$$

S obzirom da dvije proporcionalne funkcije vjerodostojnosti nose istu informaciju o θ, možemo zanemariti one izraze koji ne sadrže nepoznati parametar. U našem slučaju to bi značilo da je (21) oblika

$$L(\theta|x_1, ..., x_{20}) \propto \theta^{\sum_{i=1}^{20} x_i} e^{-20\theta}. \tag{22}$$

pri čemu $\propto$ označava proporcionalnost. Maksimizirajući (21) ili (22) (u odnosu na θ) dobivamo procjenitelj metodom maksimalne vjerodostojnosti parametra θ dan s

$$\hat{\theta} = \frac{\sum_{i=1}^{20} x_i}{20}.$$

3.2 Diskretna verzija Bayesovog teorema

Jedan od osnovnih mehanizama učenja jest prihvaćenje novih informacija s kojima nadograđujemo postojeću bazu znanja. Upravo ovaj proces leži u srži Bayesove paradigme odnosno teorema koji nam, kao formalni alat, omogućuje da ga provedemo. Sama riječ paradigma nas upućuje da se radi o posebnom referentnom okviru, okviru koji se može nositi sa većinom problema za koje znanstvenici smatraju da su bitni [4].

Zagovornici Bayesovskog pristupa vjerojatnost tretiraju kao stupanj uvjerenja, te je mijanjaju („nadograđuju") dolaskom novih informacija. Sada ćemo, kroz diskretnu verziju teorema, objasniti kako se takav pristup općenito realizira. Kasnije ćemo priču prirodno proširiti na neprekidan oblik Bayesovog teorema.

Sa E označimo informacije koje imamo prije nego smo proučili podatke i pretpostavimo da subjektivnu vjerojatnost od E možemo izraziti kao $P(E)$. Bayesov teorem nam govori da, nakon što proučimo podatke D, vjerovanje u E se mijenja po formuli

$$P(E|D) = \frac{P(D|E)P(E)}{P(D)} = \frac{P(D|E)P(E)}{P(D|E)P(E) + P(D|E^c)P(E^c)}, \tag{23}$$

pri čemu je:

1. $P(D|E)$ uvjetna vjerojatnost danih podataka D uz uvjet da je vjerovanje u E istinito.
2. P(D) marginalna gustoća podataka D, $P(D) > 0$.

Definicija 3.4. *Vjerojatnost događaja E prije analize podataka, P(E), zovemo apriori vjerojatnost, a prilagođenu vjerojatnost $P(E|D)$ aposteriori vjerojatnost.*

Primjetimo da se ovim pristupom slučajnost opisuje kombinirajući informacije dobivene iz dva izvora: obrađenih podataka i apriori uvjerenja. Primjerom ćemo to slikovitije opisati.

Primjer 3.5. *Promatramo menadžera fonda koji testira strategiju pronalaženja akvizicijskih meta te ispitivanja njihove efikasnosti, preciznije pripdane cijene oslobađanja novčanog toka (eng. PFCF). Definiramo dva događaja:*

$$\begin{aligned} D &= \textit{Tvrtki X, u protekle 3 godine, PFCF je više od tri puta manji} \\ &\quad\ \textit{od prosjeka u tom sektoru.} \\ E &= \textit{Tvrtka X postala je akvizicijska meta u toku promatrane godine.} \end{aligned}$$

Na početku, nezavisno od pokazatelja, menadžer pretpostavlja da je vjerojatnost da tvrtka postane meta 40%, što znači

$$P(E) = 0.4 \quad odnosno \quad P(E^c) = 0.6.$$

Nadalje pretpostavimo da nakon analize podataka menadžer ima vjerojatnost od 75% da tvrtka koja je postala meta, ima više od tri puta manji PFCF nego prosjek u tom sektoru. Dok je vjerojatnost da tvrtka koja nije meta, ima više od tri puta manji PFCF nego prosjek u tom sektoru jednaka 35%:

$$P(D|E) = 0.75 \quad i \quad P(D|E^c) = 0.35.$$

Sada je menadžer u stanju nadograditi apriori uvjerenja na način da, ako zna da tvrtka ima tri ili više puta manji PFCF neko prosjek u tom sektoru, može izračunati vjerojatnost da će postati akvizicijska meta. Koristeći (23) dobije se:

$$P(E|D) = \frac{0.75 \times 0.4}{0.75 \times 0.4 + 0.35 \times 0.6} \approx 0.59.$$

Zaključujemo da, nakon što se uzme u obzir razina PFCF, vjerojatnost da tvrtka postane akvizicijska meta raste s 40% na 59%.

Bayesov teorem i klasifikacija

Klasifikacija je jedna od bitnijih primjena diskretne verizije Bayesovog teorema. Klasifikacija se najčešće primjenjuje na području kreditnog i osigurajavajućeg rizika, kada kreditor (ili osiguravatelj) pokušava odrediti u koju kategoriju rizičnosti smjestiti potencijalnu stranku. Pretpostavimo da postoje tri kategorije kreditnog rizika: nizak (N), srednji (S) i visoki rizik (V). Za početak kreditor prikupi informacije o klijentu, koje ćemo označiti s y. Uvjetne distribucije od y ovise o navedenim kategorijama:

$$f(y|C = N),$$

$$f(y|C = S)$$

i

$$f(y|C = V)$$

gdje je C slučajna varijabla koja opisuje kategoriju. Kreditor, odnosno banka, ima uvjerenja o pripadnosti određenoj kategoriji koja ćemo označiti s π_i:

$$\pi_1 = \pi(C = N),$$

$$\pi_2 = \pi(C = S)$$

i

$$\pi_3 = \pi(C = V).$$

Sada se može primjeniti diskretna verzija Bayesovog teorema kako bi se izračunale aposteriori vjerojatnosti $\pi(C = i|y)$, i=N,S,V, odnosno vjerojatnosti da stranka pripada određenoj kategoriji.

3.3 Neprekidna verzija Bayesovog teorema

Kada se Bayesov teorem primjenjuje u financijama, pretežno se koristi u neprekidnom obliku. Kao i u diskretnom slučaju prvo ćemo izvesti neprekidnu formulu potpune vjerojatnosti.

Teorem 3.6. *Neka je B proizvoljan Borelov skup i neka su X i Y neprekidne slučajne varijable. Tada vrijedi*

$$P(X \in B) = \int_{\mathbb{R}} P(X \in B|Y = y) f_Y(y) dy. \tag{24}$$

Dokaz.

$$P(X \in B) = \int_{\mathbb{R}} \int_{B} f_{X,Y}(x,y)dxdy = \int_{\mathbb{R}} \int_{B} f_{X|Y}(x|y)f_Y(y)dxdy =$$
$$= \int_{\mathbb{R}} (\int_{B} f_{X|Y}(x|y)dx)f_Y(y)dy = \int_{\mathbb{R}} P(X \in B|Y = y)f_Y(y)dy.$$

□

Teorem 3.7 (Neprekidna verzija Bayesove formule)**.** *Neka su X i Y neprekidne slučajne varijable. Tada* $\forall x \in \mathbb{R}$ *i* $\forall y \in \mathbb{R}$ *takav da* $f_Y(y) > 0$*, vrijedi*

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{\mathbb{R}} f_{Y|X}(y|x)f_X(x)dx}. \tag{25}$$

Pretpostavimo da nam je za promatrane podatke X poznata distribucija koja ih opisuje, ali ne i njezin parametar θ. Jedna od ključnih primjena Bayesove statistike jest opisati nepoznati parametar i to kao slučajnu varijablu.

Neka je $x = (x_1, ..., x_n)$ jedna realizacija slučajnog uzorka $X = (X_1, ..., X_n)$ sa zajedničkom funkcijom gustoće $f(x|\theta)$, $x \in \mathbb{R}^n$, te neka je $\theta \in \Theta$ nepoznati parametar. Uvedimo oznake:
· $L(\theta|x)$ funkcija vjerodostojnosti nepoznatog parametra θ
· $\pi(\theta)$ apriori distribucija parametra θ, ovisi o jednom ili više parametara koje zovemo hiperparametri
· $\pi(\theta|x)$ aposteriori distribucija parametra θ
· $f_X(x)$ marginalna distribucija slučajne varijable X dana s

$$f_X(x) = \int L(\theta|x)\pi(\theta)d\theta. \tag{26}$$

Koristeći apriori uvjerenja te informacije dobivene analizom podataka x, Bayesovim teoremom dobivamo aposteriori distribuciju parametra θ

$$\pi(\theta|x) = \frac{L(\theta|x)\pi(\theta)}{f_X(x)}. \tag{27}$$

S obzirom da $f_X(x)$ ne ovisi o θ kažemo da je aposteriori distribucija proporcionalana s vjerodostojnošću i apriori distribucijom i pišemo

$$\pi(\theta|x) \propto L(\theta|x)\pi(\theta). \tag{28}$$

Primjer 3.8 (Bayesovsko zaključivanje za binomnu distribuciju)**.** *Kažemo da slučajna varijabla ima binomnu distribuciju ili da je X binomna slučajna* *varijabla, što označavamo s* $X \sim B(n,\theta)$, $\theta \in [0,1]$, $n \in \mathbb{N}$, *ako joj je funkcija gustoće dana s*

$$P(X = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, \quad k = 0, 1, 2, ..., n. \tag{29}$$

Binomna slučajna varijabla opisuje pokuse s dva moguća ishoda, pri čemu je θ *vjerojatnost pojavljivanja ishoda koji je promatraču od interesa, a* n *veličina uzorka.*

Promatramo unutardnevne promjene cijene dionice te nas zanima vjerojatnost njezinog uzastopnog porasta. Ovaj problem možemo modelirati binomnom slučajnom varijablom. Definiramo dva moguća ishoda: ,,uzastopna promjena cijene dionice rezultirala je rastom" i ,,uzastopna promjena cijene dionice rezultirala je padom ili stagnacijom". U našem slučaju θ *označava vjerojatnost uzastopnog porasta cijene dionice i cilj nam je izvući zaključak o tom nepoznatom parametru.*

Kao ilustraciju promatrimo podatke o cijenama AT&T dionica u periodu od 4.1.1993. do 26.2.1993. (55668 promjena cijene dionice). Na Slici 1. objašnjeneno je na koji način bismo definirali binomnu slučajnu varijablu u slučaju šest cijena dionice $P_1, ..., P_6$ *(odnosno pet promjena cijene). Primijetimo da A = 2 označava uzastopan rast cijene i vjerojatnost tog događaja je* $\theta = P(A = 2)$, *dok je vjerojatnost svih ostalih događaja (A = -2, -1, 0 ili 1) jednaka 1 -* θ.

S X označimo broj uzastopnih porasta cijene dionice (A=2). Tada je X binomno distribuirana slučajna varijabla s nepoznatim parametrom θ. *U promatranom periodu dogodilo se X = 176 uzastopnih porasta cijene AT&T dionice, iz čega proizlazi da funkcija vjerodostojnosti parametra* θ *glasi*

$$L(\theta|x) = \theta^{176}(1-\theta)^{55667-176}. \tag{30}$$

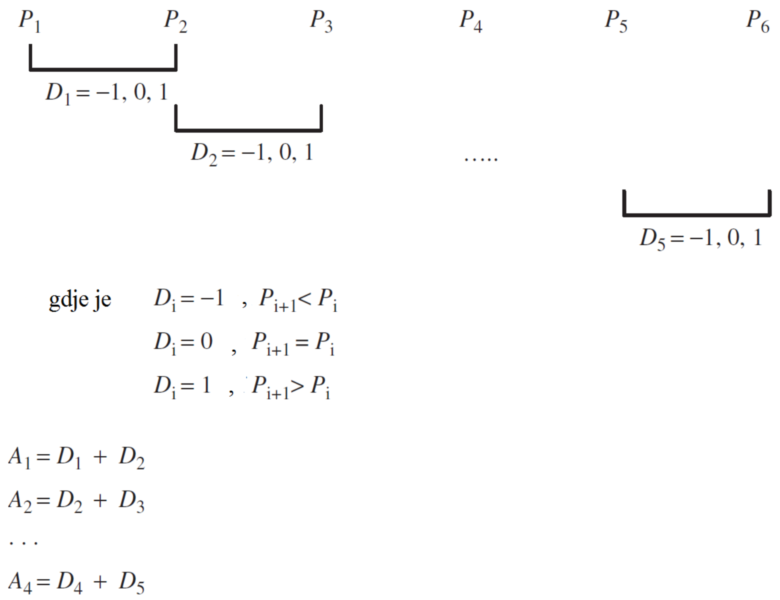
Željeli bismo ovu informaciju kombinirati s apriori uvjerenjem, s ciljem dobivanja aposteriori distribucije nepoznatog parametra θ.

Promatrat ćemo dva apriori scenarija:

1. Nemamo nikakvo posebno uvjerenje o distribuciji parametra θ. *U tom slučaju ga možemo opisati uniformnom distribucijom na intervalu [0,1], pri čemu je funkcija gustoće dana s*

$$\pi(\theta) = 1, \quad 0 \leq \theta \leq 1.$$

2. Naša intuicija kaže da je vjerojatnost uzastopnog porasta cijene dionice otprilike 2%. Jedan od mogućih izbora apriori distribucije za parametar θ *jest*



Slika 1: Uzastopne promjene cijene dionice

beta distribucija[1]*. Funkcija gustoće beta distribucije (distribucije parametra* θ*) jest*

$$\pi(\theta|\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0 \leq \theta \leq 1, \tag{31}$$

gdje su α *i* β *parametri beta distribucije, čije vrijednosti postavljamo na 1.6 odnosno 78.4.*

Slika 2 prikazuje grafove dviju apriori gustoća. Primijetimo da kod uniformne distribucije sve vrijednosti parametra θ *su jednako vjerojatne, dok to nije slučaj kod beta distribucije.*

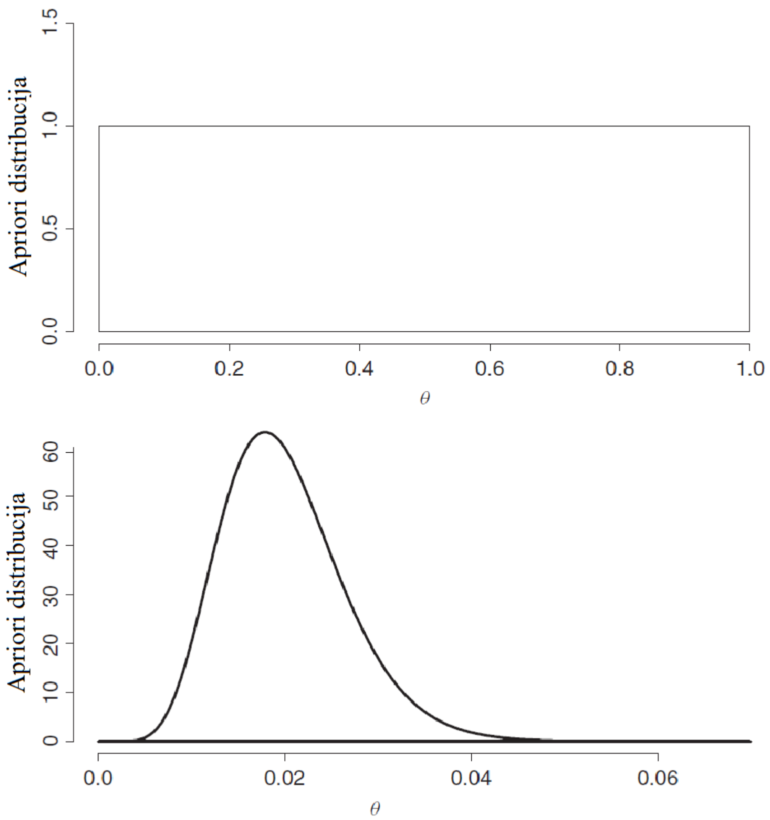
Pomoću izraza (27) odnosno (28) dobit ćemo aposteriori distribuciju parametra θ*. S obzirom da su obje apriori distribucije neprekidne, očekujemo*

[1]Kažemo da slučajna varijabla X ima beta distribuciju s parametrima $\alpha > 0$ i $\beta > 0$, što označavamo s $X \sim Beta(\alpha,\beta)$, ako joj je funkcija gustoće dana s

$$f_X(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \leq x \leq 1,$$

pri čemu je $B : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ beta funkcija definirana sa

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$



Slika 2: Krivulje funkcija gustoće dvije apriori distribucije parametra θ

da će takva biti i aposteriori pa opet promatramo dva slučaja:

1. *Aposteriori distribucija u slučaju uniformne apriori:*

$$\begin{aligned}\pi(\theta|x) &\propto L(\theta|x) \times 1 \\ &\propto \theta^{176}(1-\theta)^{55667-176} \\ &= \theta^{177-1}(1-\theta)^{55492-1}\end{aligned}$$

Izraz $\theta^{177-1}(1-\theta)^{55492-1}$ *podsjeća na funkciju gustoće beta distribucije, samo što nedostaje izraz* $Beta(177, 55492)$ *koji ne ovisi o* θ. *Zbog toga možemo odrediti jedinstvenu aposterior distribuciju parametra* θ *kao ,,krnju" beta distribuciju s parametrima* $\alpha = 177$ *i* $\beta = 55492$.

2. *Beta distribucija je konjugirana za binomnu razdiobu*[2] *pa očekujemo*

[2]Beta distribucija jest konjugirana apriori distribucija za binomnu razdiobu. O konjugiranim distribucijama pričat ćemo u sljedećem poglavlju.

da i aposteriori bude beta distribucija:

$$\begin{aligned}\pi(\theta|x) &\propto L(\theta|x)\pi(\theta)\\ &\propto \theta^{176}(1-\theta)^{55667-176}\theta^{1.6-1}(1-\theta)^{78.4-1}\\ &= \theta^{177.6-1}(1-\theta)^{55569.4-1}.\end{aligned}$$

Opet možemo zanemariti sve izraze koji ne ovise o θ pa dobivamo, kao što smo i očekivali, ,,krnju" beta distribuciju s parametrima $\alpha = 177.6$ i $\beta = 55569.4$.

Konačno, željeli bismo odrediti broj koji najbolje procjenjuje parametar θ. U teoriji klasične statistike najbolji procjenitelj od θ je procjenitelj maksimalne vjerodostojnosti (vrijednost koja maksimizira funkciju vjerodostojnosti u (30)), odnosno

$$\hat{\theta} = \frac{176}{55667} = 0.00316$$

ili 0.316%.

U Bayesovskim postavkama, jedan od mogućih procjenitelja parametra θ jest očekivanje aposteriori distribucije od θ. Očekivanje beta distribucije dano je s $\alpha/(\alpha+\beta)$ čime dobijemo procjenitelj od θ:

1. U slučaju uniformne apriori distribucije

$$\hat{\theta} = \frac{177}{177+55492} = 0.00318$$

2. U slučaju beta apriori distribucije

$$\hat{\theta} = \frac{177.6}{177.6+55569.4} = 0.00319.$$

4 Bayesovsko zaključivanje

U prošlom poglavlju predstavili smo diskretnu i neprekidnu verziju Bayesovog teorema te koncept funkcije vjerodostojnosti, apriori i aposteriori distribucija s naglaskom na parktične primjene. Sada ćemo se fokusirati na osnove Bayesovskog zaključivanja. Ključni dio takvog procesa zaključivanja jest formaliziranje znanja i intuicije kako bi se što preciznije odredila apriori distribucija, posebno kada podatci koje proučavamo nisu obilni.

4.1 Apriori distribucije

Kao što smo već spomenuli, u svijetu financijske industrije glavnu ulogu igra neprekidna verzija Bayesovog teorema dana s (27) odnosno (28). Dva su glavna faktora koja utječu na aposteriori distribuciju: snaga apriori informacija i količina dostupnih podataka. Općenito, dok god apriori nije jako informativna (u smislu koji će uskoro postati jasniji), količina podataka diktira utjecaj na aposteriori distribuciju. Vrijedi i obrat.

Jedan od osnovnih problema Bayesove statistike je kako prevesti apriori informacije o parametru u analitičku (distribucijsku) formu, $\pi(\theta)$, i koliko je aposteriori distribucija osjetljiva na njih. S obzirom da se radi o subjektivnim gledištima na dani problem, ne postoji ,,najbolje" rješenje. U nastavku ćemo iznijeti najčešće korištene metode odabira apriori distribucija.

4.1.1 Neinformativne apriori distribucije

U slučaju kada nemamo prethodnu informaciju o parametru θ, naša apriori uvjerenja su nejasna i teško ih je prevesti u informativan oblik. Stoga ne želimo da apriori nesigurnost bitno utječe na aposteriori zaključak. U tu svrhu koristimo tzv. neinformativne apriori distribucije. One nekad nisu normalizirane (niti se mogu normalizirati), no Bayesovska analiza ih formalno dopušta. Takve distribucije zovemo nepravilne apriori distribucije i za njih vrijedi

$$\int_{\Theta} \pi(\theta) d\theta = +\infty$$

ili u diskretnom slučaju, kada je Θ prebrojiv skup

$$\sum_{\theta_i \in \Theta} \pi(\theta_i) = +\infty, \quad i = 1, 2, ...$$

gdje je Θ parametarski prostor. Koristeći nepravilne apriori distribucije moguće je dobiti normaliziranu aposteriori distribuciju tako da uzmemo dobro definiranu marginalnu distribuciju.

Najčešće korištene neinformativne apriori distribucije su uniformna i Jeffreyjeva. Povijesni pristup, predvođen Laplaceom i Bayesom, zagovara uniformnu apriori distribuciju, odnosno da su sve vrijednosti u parametarskom prostoru jednako vjerojatne. Jedna od zamjerki ovakvom pristupu jest da uniformne distribucije nisu invarijantne na reparametrizaciju[3].

Jeffreyjeve apriori distribucije

Jeffreyjeve apriori distribucije bazirane su na principu invarijantnosti. Pretpostavimo da nam je poznata funkcija vjerodostojnosti $L(x|\theta)$ za podatke $x \in \mathbb{R}^n$ i nepoznati parametar θ. Jeffreyjeve apriori distribucije za nepoznati parametar mogu se dobiti iz poznate funkcije vjerodostojnosti. Primjetimo da je ovo u suprotnosti s klasičnim Bayesovskim pristupom gdje se prvo određuje apriori distribucija, nakon čega se prouče podatci i odredi vjerodostojnost. Za slučaj jednodimenzionalnog parametarskog prostora Θ Jeffreyjova apriori distribucija dana je s

$$\pi_J(\theta) \propto I(\theta)^{1/2} \tag{32}$$

gdje je I Fisherova informacija definirana kao očekivanje druge derivacije logaritma funkcije vjerodostojnosti

$$I(\theta) = -E_\theta\left[\frac{\partial^2 \ln L(\theta|x)}{\partial\theta^2}\right] \tag{33}$$

Ovakve distribucije zadovoljavaju uvjet invarijantnosti na reparametrizaciju. Ako je $h : \Theta \to \Theta$ injektivna transformacija $\phi = h(\theta)$, tada je Fisherova informacija za parametar ϕ dana s

$$I(\phi) = I(\theta)\left(\frac{\partial\theta}{\partial\phi}\right)^2, \tag{34}$$

a Jeffreyjeva apriori distribucija za ϕ s

$$\pi_J(\phi) \propto \sqrt{I(\phi)} = \sqrt{I(\theta)\left(\frac{\partial\theta}{\partial\phi}\right)^2} = \sqrt{I(\theta)}\left|\frac{\partial\theta}{\partial\phi}\right| \propto \pi_J(\theta)\left|\frac{\partial\theta}{\partial\phi}\right|. \tag{35}$$

[3]Promotrimo neinformativnu apriori distribuciju binomne razdiobe za parametar θ iz primjera 3.8. Znamo da je $\theta \in [0,1]$, pa možemo koristiti novu parametrizaciju $\rho = \ln(\frac{\theta}{1-\theta})$ koja je prirodna u slučaju da želimo da parametar θ poprimi sve realne vrijednosti. Međutim, pod novom parametrizacijom apriori distribucija $\pi(\rho)$ nije više uniformna.

Primjer 4.1 (Jeffreyjeva apriori distribucija za očekivanje normalne distribucije)**.** *Kažemo da slučajna varijabla ima normalnu distribuciju ili da je X normalna slučajna varijabla, što označavamo s $X \sim N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0$, ako joj je funkcija gustoće dana s*

$$f_X(X) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Bigg(- \frac{(x-\mu)^2}{2\sigma^2}\Bigg), \quad x \in \mathbb{R}. \tag{36}$$

Normalna slučajna varijabla najvažnija je neprekidna slučajna varijabla, pogotovo u financijskoj industriji. Parametar μ zovemo očekivanje, a σ standarndna devijacija normalne slučajne varijable.
Sada želimo, pomoću Jeffreyjevog pristupa, odrediti apriori distribuciju parametra μ (pretpostavljamo da je parametar σ^2 poznat i fiksan). Funkcija vjerodostojnosti za parametar μ normalne slučajne varijable X je oblika

$$L(\mu|X) \propto \exp\Bigg(- \frac{(X-\mu)^2}{2\sigma^2}\Bigg).$$

Slijedi da je Fisherova informacija dana s

$$I(\mu) = -E_\mu\left[\frac{\partial^2 \ln L(\mu|x)}{\partial \mu^2}\right] = -E_\mu\left[\frac{1}{\sigma^2}\right] = -\frac{1}{\sigma^2}, \tag{37}$$

iz čega dobijemo da je Jeffreyjeva apriori distribucija parametra μ oblika (s obzirom da μ ne ovisi o σ^2 koji je po pretpostavci fiksan)

$$\pi_J(\mu) \propto 1.$$

Jeffryjeve apriorine distribucije možemo generalizirati na višedimenzionalne parametre θ tako što ćemo uzeti da nam apriori distribucija bude jednaka korjenu determinante matrice Fisherove informacije, tj.

$$\pi_J(\theta) = |I(\theta)|^{1/2} \tag{38}$$

gdje je matrica Fisherove informacije dana s

$$I(\theta)_{i,j} = -E_\theta\left[\frac{\partial^2 \ln L(\theta|x)}{\partial \theta_i \partial \theta_j}\right] \tag{39}$$

Primjer 4.2 (Jeffreyeva apriori distribucija za parametre normalne distribucije)**.** *Neka je $X \sim N(\mu, \sigma^2)$ normalna slučajna varijabla. Želimo odrediti* *apriori distribuciju dvodimenzionalnog parametra* $\theta = \binom{\mu}{\sigma^2}$. *Računamo matricu Fisherove informacije*

$$\begin{aligned} I(\theta) &= -E_\theta \begin{pmatrix} \frac{1}{\sigma^2} & \frac{2(X-\mu)}{\sigma^2} \\ \frac{2(X-\mu)}{\sigma^2} & \frac{3}{\sigma^4}(X-\mu)^2 - \frac{1}{\sigma^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{\sigma^2} \end{pmatrix} \end{aligned}$$

gdje su $E_\theta(X-\mu) = 0$ *i* $E_\theta(X-\mu)^2 = \sigma^2$. *Zajednička apriori distribucija parametara* μ *i* σ *normalne slučajne varijable je*

$$\pi_J(\theta) = |I(\theta)|^{1/2} \propto \frac{1}{\sigma^2}.$$

Pokazalo se da ovakva distribucija ima jako loša konvergencijska svojstva pa je i sam Jeffrey umjesto nje uzeo $\pi_J(\theta) \propto \frac{1}{\sigma}$ *koja je ujedno i umnožak apriorinih distribucija za* μ *i* σ *(koje se dobiju zasebnim računom).*

Apriori distribucije s velikom disperzijom

Apriori neznanje može biti opisano i pravilnom distribucijom s jako velikom disperzijom. Pretostavimo da imamo podatke koji su normalno distribuirani i želimo procijeniti distribuciju parametra μ. Kako nemamo prethodno uvjerenje o njemu i ne želimo utjecati na aposteriori zaključke, možemo ga opisati normalnom distribucijom s očekivanjem centriranim oko 0 i velikom standardnom devijacijom kao npr. 10^6. Time dobijemo $\pi(\mu) = N(0, (10^6)^2)$. Takva apriori distribucija je pravilna.

4.1.2 Informativne apriori distribucije

Apriori uvjerenja su informativna ako bitno utječu na aposteriori zaključke, odnosno dominantne su u odnosu na funkcije vjerodostojnosti parametra modela. Kod odabira apriorine distribucije i hiperparametara trebamo biti pažljivi i sigurni da ona dobro reflektira naš uvjerenja. Upravo zbog toga, informativna apriorina distribucija pokazuje koliku moć ima Bayesova statistika.

Najčešći pristup „prevođenja" apriori uvjerenja je određivanje distribucije za nepoznat parametar, kao i pripadnih hiperparametara. Neka je X promatrana slučajna varijabla. Za srednju vrijednost koju poprima X radije ćemo koristiti medijan nego očekivanje (kod simetričnih distribucija te se vrijednosti podudaraju). Formaliziranje uvjerenja vezanih za disperziju distribucije od X manje je intuitivno. Najlakši način je uzimanje kvantila, kao npr.:

$$P(X < x_{0.25}) = 0.25$$

ili

$$P(X > x_{0.75}) = 0.75,$$

gdje su $x_{0.25}$ i $x_{0.75}$ prvi i treći kvartil koje određujemo subjektivno.

Primjer 4.3 (Informativna apriori distribucija za parametre normalne razdiobe)**.** *Pretpostavimo da promatramo mjesečne povrate na nekakvu financijsku imovinu koji su normalno distribuirani $N(\mu, \sigma^2)$ te da nam je poznata varijanca $\sigma^2 = \sigma^{2*}$. Želimo odrediti distribuciju za nepoznat parametar μ. Smatramo da je simetrična distribucija dobar izbor i zbog jednostavnosti odabiremo normalnu distribuciju:*

$$\mu \sim N(\eta, \tau^2),$$

gdje je η apriori očekivanje a τ^2 apriori varijanca parametra μ. Kako bi u potpunosti odredili apriorinu distribuciju trebamo odrediti vrijednosti tih hiperparametara. Vjerujemo da je najčešći mjesečni povrat oko 1% zbog čega je $\eta = 0.01$. Nadalje, subjetivno smo odredili da je 25% šanse da je prosječni mjesečni povrat manji od 0.5% ($\mu_{0.25} = 0.005$). Iz tablice za standardnu normalnu razdiobu slijedi da je varijanca $\tau^2 = 0.74^2$. Dobivena apriori distribucija glasi $\pi(\mu) \sim N(0.01, 0.74^2)$.

Vratimo se na trenutak primjeru 3.8 iz prošlog poglavlja. Jedna od apriori distribucija parametra θ je bila beta distribucija s parametrima $\alpha = 1.6$ i $\beta = 78.4$. Njih smo odredili pomoću apriori uvjerenja da je 2% uzastopnih promjena cijena rezultiralo rastom istih, te da je šansa da je uzastopnih porasta cijene manje od 1% jednaka 30%:

$$\frac{\alpha}{\alpha + \beta} = 0.02$$

i

$$P(\theta < 0.01) = 0.3.$$

Ove dvije jednadžbe u potpunosti određuju parametre α i β.

Konjugirane apriori distribucije

Ponekad je izbor apriori distribucije vođen željom da se dobije analitički poželjna aposteriori distribucija. Ako se uoči da su podatci generirani određenom familijom distribucija, tada će nam izbor tzv. konjugirane apriori distribucije garantirati da je i aposteriori iz iste familije kao i apriori. Iako apriori i aposteriori imaju istu formu, njihovi parametri se razlikuju - parametri aposteriori reflektiraju razmjenu između apriori i informacija dobivenih analizom podataka.

Sada ćemo opet promotriti slučaj jednodimenzionalne normalne distribucije, s obzirom da nam je ona najčešća kod primjena u financijama. Funkcija vjerodostojnosti normalne razdiobe za nepoznate parametre μ i σ dana je s

$$L(\mu, \sigma^2 | x_1, ..., x_n) \propto \frac{1}{\sigma^2} \exp\Bigg(- \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{2\sigma^2} \Bigg).$$

Cilj nam je naći spomenute konjugirane distribucije za μ i σ kako bi nam u konačnici aposteriori distribucija bila normalna. Pretpostavimo da je varijanca σ^2 poznata (fiksna), te da je očekivanje μ također normalno distribuirano $N(\mu_0, \sigma_0^2)$.

Lema 4.4. *Pretpostavimo da je (z_1, z_2) normalno distribuiran slučajan vektor. Tada je $z_1|z_2$ normalno distribuirana slučajna varijabla s parametrima*

$$E(z_1|z_2) = E(z_1) + \frac{Cov(z_1, z_2)}{Var(z_2)}(z_2 - E(z_2)), \tag{40}$$

$$Var(z_1|z_2) = Var(z_1) - \frac{Cov^2(z_1, z_2)}{Var(z_2)}. \tag{41}$$

Primjenimo ovu lemu na naš slučaj tako da $\mu \to z_1$ te $x \to z_2$:

$$E(x) = \mu_0$$

$$Var(x) = \sigma^2 + \sigma_0^2$$

$$Cov(x, \mu) = E(x - \mu_0)(\mu - \mu_0) = \sigma_0^2,$$

pa prema prethodnoj lemi dobijemo

$$E(\mu|x) = \mu_0 + \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}(x - \sigma_0^2) = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2} x + \frac{\sigma^2}{\sigma_0^2 + \sigma^2} \mu_0$$

$$Var(\mu|x) = \frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2} = \frac{1}{\frac{1}{\sigma^2} + \frac{1}{\sigma_0^2}},$$

pri čemu je x iz jednadžbe za očekivanje najčešće procjenitelj metodom maksimale vjerodostojnosti (MLE). Sažmimo dobivene rezultate u sljedeću lemu.

$f(x\|\theta)$	$\pi(\theta)$	$\pi(\theta\|x)$
Normalna $N(\theta,\sigma^2)$	Normalna $N(\mu,\tau^2)$	Normalna $N(\frac{\tau^2}{\tau^2+\sigma^2}X+\frac{\sigma^2}{\tau^2+\sigma^2}\mu,\frac{\sigma^2\tau^2}{\sigma^2+\tau^2})$
Poissonova $P(\theta)$	Gamma $\Gamma(\alpha,\beta)$	Gamma $\Gamma(\alpha+X,\beta+1)$
Gamma $\Gamma(\upsilon,\theta)$	Gamma $\Gamma(\alpha,\beta)$	Gamma $\Gamma(\alpha+\upsilon,\beta+X)$
Binomna $B(n,\theta)$	Beta $Beta(\alpha,\beta)$	Beta $Beta(\alpha+X,\beta-X+n)$
Negativna binomna $Neg(m,\theta)$	Beta $Beta(\alpha,\beta)$	Beta $Beta(\alpha+m,\beta+X)$
Normlana $N(\mu,\frac{1}{\theta})$	Gamma $\Gamma(\alpha,\beta)$	Gamma $\Gamma(\alpha+0.5,\beta+\frac{(\mu+X)^2}{2})$

Tablica 1: Najčešće korištene konjugirane apriori distribucije.

Lema 4.5. *Pretpostavimo da je $X|\mu\sim N(\mu,\sigma^2)$ i $\mu\sim N(\mu_0,\sigma_0^2)$. Tada vrijedi*

$$\mu|x\sim N\left(\frac{\sigma_0^2}{\sigma_0^2+\sigma^2}x+\frac{\sigma^2}{\sigma_0^2+\sigma^2}\mu_0,\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\sigma_0^2}}\right). \tag{42}$$

Pretpostavimo sada da tražimo apriori/aposteriori distribuciju varijance normalne slučajne varijable uz poznato (fiksno) očekivanje.

Lema 4.6. *Neka su $x_1,...,x_n|\mu\sim N(\mu,\sigma^2)$ nezavisne jednako distribuirane i $\sigma^2\sim Inv\Gamma(\alpha,\beta)$.[4] Tada vrijedi*

$$\sigma^2|x_1,...,x_n\sim Inv\Gamma\left(\alpha+\frac{n}{2},\beta+\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)\right). \tag{43}$$

[4]Neka su $\alpha,\beta>0$. Za neprekidnu slučajnu varijablu X s vrijednostima u $\mathbb{R}$, kažemo da ima inverznu gama distribuciju $Inv\Gamma(\alpha,\beta)$ ako vrijedi

$$f_X(x|\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\exp(-\beta/x)}{x^{\alpha+1}},\quad x\geq 0,$$

pri čemu je $\Gamma(\alpha)=\int\limits_0^{+\infty}t^{\alpha-1}\exp(-t)dt$ gama funkcija.

4.2 Bayesovska predviđanja

Nakon što smo detaljno objasnili kako primjeniti Bayesov teorem u svrhu dobivanja apriori kao i aposteriori distribucija parametara, proučit ćemo i kako ih iskoristiti pri predviđanjima. Svrha predviđanja može biti testiranje promatranog modela (na primjer, mjerimo odstupanja rezultata dobivenih modelom od stvarnih) ili korištenje informacija dobivenih predviđanjem kod donošenja odluka. Vidjeli smo da Bayesovskim pristupom u konačnici raspolažemo s cijelom distribucijom umjesto procjeniteljima. Neka je $f(x|\theta)$ funkcija gustoće dobivena iz uzorka i $\pi(\theta|x)$ aposteriori distribucija parametra θ. Funkcija gustoće predviđanja dana je s

$$f(x_{+1}|x) = \int f(x_{+1}|\theta)\pi(\theta|x)d\theta \tag{44}$$

gdje x_{+1} označava jednu realizaciju unaprijed. Primjetimo da integriramo po svim θ što znači da funkcija gustoće predviđanja ne ovisi o θ nego samo o prošlim realizacijama slučajne varijable X (koja nam opisuje promatrane podatke).

5 Bayesovski linearni regresijski model

Regresijska analiza jedno je od najčešćih ekonometrijskih alata korištenih na području upravljanja financijskim investicijama. U ovom poglavlju iznijet ćemo Bayesovski pristup procjene univarijantnog i multivarijantnog regresijskog modela.

5.1 Univarijantni linearni regresijski model

Univarijantni linearni regresijski model pokušava objasniti utjecaj jedne ili više varijabli (neovisne varijable) na varijablu koju zovemo ovisna varijabla. Kao što samo ime modela kaže, postavlja se linearna veza među njima:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_{K-1} X_{K-1} + \epsilon, \tag{45}$$

pri čemu je
Y = ovisna varijabla
X_k = neovisna varijabla, k = 1,2,...,K-1
α = slobodni član
β_k = regresijski koeficijenti, k = 1,2,...,K-1, koji predstavljaju utjecaj jedinične promjene X_k, k = 1,2,...,K-1, na ovisnu varijablu Y
ϵ = slučajna greška ili šum.

Slučajna greška je izvor neizvjesnosti u vezi između ovisne i neovisnih varijabli. Stoga $\alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_{K-1} X_{K-1}$ predstavlja ,,objašnjeni" dio od Y, dok ϵ predstavlja varijabilni odnosno ,,neobjašnjeni" dio.

Pretpostavimo da imamo n opažanja ovisnih i neovisnih varijabli:

$$y_i = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_{K-1} x_{K-1,i} + \epsilon_i, \quad i = 1, 2, ..., n. \tag{46}$$

Kako bi opisali izvor slučajnosti, ϵ, moramo pretpostaviti njegovu distribuciju, pa zbog jednostavnosti pretpostavimo da su ϵ_i, i = 1,...,n, nezavisne i jednako distribuirane normalnom distribucijom s očekivanjem nula i poznatom varijancom σ^2. Tada je i ovisna varijabla normalno distribuirana

$$y_i \sim N(\mu_i, \sigma^2), \tag{47}$$

pri čemu je $\mu_i = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_{K-1} x_{K-1,i}$. Izraz (46) često pišemo u obliku:

$$y = X\beta + \epsilon, \tag{48}$$

gdje je y $(n \times 1)$ vektor,

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

β je $(K \times 1)$ vektor,

$$\beta = \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_{K-1} \end{pmatrix},$$

X je $(n \times K)$ matrica sa prvim stupcem jedinica,

$$X = \begin{pmatrix} 1 & x_{1,1} & ... & x_{K-1,1} \\ 1 & x_{1,2} & ... & x_{K-1,2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1,n} & ... & x_{K-1,n} \end{pmatrix},$$

i ϵ je $(n \times 1)$ vekor,

$$\epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}.$$

Distribuciju parametra ϵ pišemo kao

$$\epsilon \sim N(0, \sigma^2 I_n),$$

pri čemu je I_n $(n \times n)$ matrica identiteta. Uz ovu pretpostavku funkcija vjerodostojnosti modela je oblika

$$\begin{aligned} L(\alpha, \beta_1, \beta_2, \sigma | y, X) = &(2\pi\sigma^2)^{-n/2} \\ &\times \exp\Bigg(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \alpha - \beta_1 x_{1,i} - ... - \beta_{K-1} x_{K-1,i})^2 \Bigg) \end{aligned}$$

U vektorskoj notaciji, imamo funkciju vjerodostojnosti multivarijantne normalne distribucije,

$$L(\beta, \sigma | y, X) = (2\pi\sigma^2)^{-n/2} \exp\Bigg(-\frac{1}{2\sigma^2} (y - X\beta)'(y - X\beta) \Bigg) \tag{49}$$

Bayesovska procjena univarijantnog regresijskog modela

U klasičnim postavkama, regresijski parametri određuju se maksimiziranjem funkcije vjerodostojnosti s obzirom na β i σ^2. Ako pretpostavimo da su slučajne greške normalno distribuirane tada nam metoda najmanjih kvadrata i metoda maksimalne vjerodostojnosti daju iste procjenitelje traženih parametara:

$$\hat{\beta} = (X'X)^{-1} X y,$$

$$\hat{\sigma}^2 = \frac{1}{n-K}(y - X\hat{\beta})'(y - X\hat{\beta}).$$

Regresijska procjena parametara može se računati i Bayesovskim pristupom, gdje se primjenjuje sve prije navedeno o apriori i aposteriori distribucijama. Sada ćemo promotriti scenarij informativne konjugirane apriori distribucije.

Informativna apriori distribucija S obzirom na pretpostavku o normalnosti regresijske greške u (46), možemo koristiti konjugirane apriori kako bi odrazili postojeće znanje i dobili analitički pogodnu aposteriori distribuciju. Stoga pretpostavimo da je regresijski vektor koeficijenata normalno distribuiran (uvjetno na σ^2) te da σ^2 ima inverznu χ^2 [5] distribuciju:

$$\beta|\sigma \sim N(\beta_0, \sigma^2 A) \tag{50}$$

i

$$\sigma^2 \sim Inv - \chi^2(\upsilon_0, c_0^2). \tag{51}$$

Četri parametra se moraju odrediti apriori: β_0, A, υ_0 i c_0^2. Za matricu raspršenja A najčešće se uzima $\tau^{-1}(X'X)^{-1}$ kako bi je poistovjetili s matricom kovarijanci procjenitelja metodom najmanjih kvadrata (zanemarujući konstantu raspršenja). Parametar τ možemo namještati u ovisnosti koliko jako vjerujemo da je očekivanje od β upravo β_0 - manji τ znači veću nesigurnost u β_0. Najlakši način određivanja parametra β_0 jest da ga postavimo na 0 ili izjednačimo sa procjeniteljem metodom najmanjih kvadrata $\hat{\beta}$ dobivenog iz regresije provedene na apriori uzorku. Parametri inverzne χ^2 distribucije mogu se odrediti koristeći informacije dobivene iz obrađenih podataka:

$$\upsilon_0 = n_0 - K \tag{52}$$

$$c_0^2 = \frac{1}{\upsilon_0}(y_0 - X_0\hat{\beta}_0)'(y_0 - X_0\hat{\beta}_0), \tag{53}$$

gdje se indeks 0 odnosi na apriori informacije dobivene iz uzorka.

[5] Neka je $x, c, \upsilon > 0$. Slučajna varijabla X ima inverznu χ^2distribuciju sa υ stupnjeva slobode i parametrom raspršenja c ako vrijedi

$$f(x|\upsilon, c) = \frac{1}{\Gamma(\frac{\upsilon}{2})}\left(\frac{\upsilon}{2}\right)^{\upsilon/2} c^{\upsilon} x^{-\left(\frac{\upsilon}{2}+1\right)} \exp\left(-\frac{\upsilon c}{2x}\right),$$

i pišemo $X \sim Inv - \chi^2(\upsilon, c)$. Ona je poseban slučaj gama distribucije sa parametrima $\alpha = \upsilon/2$ i $\beta = \upsilon c/2$.

Aposteriori distribucije parametara β i σ^2 imaju istu formu kao i apriori, ali s ,,nadograđenim" parametrima.
Aposteriori parametra β je

$$\pi(\beta|y, X, \sigma^2) = N(\beta^*, \Sigma_\beta),$$

pri čemu su aposteriori očekivanje i kovarijacijska matrica dani s

$$\beta^* = (A^{-1} + X'X)^{-1}(A^{-1}\beta_0 + X'X\hat{\beta}) \tag{54}$$

i

$$\Sigma_\beta = \sigma^2(A^{-1} + X'X)^{-1}. \tag{55}$$

Aposteriori distribucija parametra σ^2 je

$$\pi(\sigma^2|y, X) = Inv - \chi^2(\upsilon^*, c^{2*}).$$

Parametri ove aposteriori distribucije dani su s

$$\upsilon^* = \upsilon_0 + n \tag{56}$$

i

$$\upsilon^* c^{2*} = (n - K)\hat{\sigma}^2 + (\beta_0 - \hat{\beta})'H(\beta_0 - \hat{\beta}) + \upsilon_0 c_0^2, \tag{57}$$

gdje je $H = ((X'X)^{-1} + A)^{-1}$.

Može se izračunati marginalna aposteriori distribucija parametra β integrirajući zajedničku aposteriori distribuciju danu s

$$p(\beta, \sigma^2|y, X) = p(\beta|y, X, \sigma^2)p(\sigma^2|y, X)$$

po svim σ^2. Na taj način dobijemo da je parametar β distribuiran studentovom t-distribucijom[6] $\beta \sim t(\upsilon^*, \beta^*, Q)$

$$p(\beta|y, X) \propto (\upsilon^* + (\beta - \beta^*)'Q(\beta - \beta^*))^{-\upsilon^*/2},$$

gdje je $Q = (A^{-1} + X'X)/c^{2*}$.

[6] Kažemo da slučajna varijabla X ima Studentovu t-distribuciju sa υ stupnjeva slobode ako vrijedi

$$f(x|\upsilon, \mu, \sigma) = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sigma\Gamma(\frac{\upsilon}{2})\sqrt{\upsilon\pi}}\left(1 + \frac{1}{\upsilon}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-(\upsilon+1)/2}, \quad -\infty < x < \infty$$

gdje je Γ gama funkcija, $-\infty < \mu < \infty$ očekivanje od X, i $\sigma > 0$ parametar raspršenja.

5.2 Multivarijantni linearni regresijski model

U financijama se često susrećemo sa modeliranjem podataka koji se sastoje od više imovina čiji povrati, ili neka druga značajka koju promatramo, nisu nezavisni. Jedan način prikazivanja takvih ovisnosti su multivarijantni modeli, koje ćemo sada koristiti kod regresijske procjene.

Pretpostavimo da imamo N zavisnih varijabli i za svaku T opažanja. To zapisujemo u $(T \times N)$ matricu Y,

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,N} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ y_{T,1} & y_{T,2} & \cdots & y_{T,N} \end{pmatrix}.$$

Multivarijantni linearni regresijski model dan je sa

$$Y = XB + U, \tag{58}$$

pri čemu je X $(T \times K)$ matrica opažanja K nezavisnih varijabli,

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,K} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,K} \\ \vdots & \vdots & \vdots & \vdots \\ x_{T,1} & x_{T,2} & \cdots & x_{T,K} \end{pmatrix},$$

B $(K \times N)$ matrica regresijskih koeficijenata,

$$B = \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_{K-1} \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,N} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{K-1,1} & \beta_{K-1,2} & \cdots & \beta_{K-1,N} \end{pmatrix},$$

U $(T \times N)$ matrica regresijskih grešaka,

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix} = \begin{pmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ u_{T,1} & u_{T,2} & \cdots & u_{T,N} \end{pmatrix}.$$

Prvi stupac od X obično se sastoji od jedinica kako bi se ukazalo na prisutnost slobodnog člana. Kod multivarijantnih postavki pretpostavka je da su greške nezavisne i jednako distribuirane, što znači da je svaki redak od U nezavisna realizacija iste N-dimenzionalne multivarijantne distribucije.

Multivarijantna normalna distribucija Neka je $\theta \in \mathbb{R}^n$ i Σ $(n \times n)$ simetrična pozitivno definitna matrica. Za neprekidni n-dimenzionalni slučajni vektor X s vrijednostima u $\mathbb{R}^n$ kažemo da je normalno distribuiran, $X \sim N(\theta, \Sigma)$ ako vrijedi

$$f(x|\theta,\Sigma) = (2\pi)^{-n/2}|\Sigma|^{-1/2}\exp\bigg(-\frac{1}{2}(x-\theta)'\Sigma^{-1}(x-\theta)\bigg), \tag{59}$$

pri čemu je $|\Sigma|$ determinanta matrice Σ.

Mi pretpostavljamo da je U multivarijantno normalno distribuiran s očekivanjem nula i kovarijacijskom matricom Σ,

$$u_i \sim N(0,\Sigma), \quad i = 1, ..., T.$$

Elementi izvan dijagonale matrice Σ su različiti od nula s obzirom da smo pretpostavili zavisnost varijabli, zbog čega imamo N varijanci i N(N-1)/2 različitih kovarijanci. Koristeći izraz za funkciju gustoće multivarijantne normalne distribucije dobijemo funkciju vjerodostojnosti za nepoznate parametre modela, B i Σ,

$$L(B,\Sigma|Y,X) \propto |\Sigma|^{-T/2}\exp\bigg(-\frac{1}{2}\sum_{t=1}^{T}(y_t - x_tB)\Sigma^{-1}(y_t - x_tB)'\bigg). \tag{60}$$

6 Procjena vrijednosti ulaganja

6.1 Bayesovski pristup pri odabiru portfelja

Temelje moderne teorije porfelja postavio je Markowitz (1952.) čime je uvelike utjecao na financijsku industriju. Porfelj predstavlja skup imovine koju posjeduje pojedinac ili organizacija da bi na bazi diversifikacije smanjio rizik ulaganja te tako najbolje zaštitio svoje interese. Temeljni model moderne teorije portfelja jest MV model (eng. Mean-variance model), čiji cilj nije pronaći portfelj kojim se maksimira očekivani povrat, već pronaći ravnotežu između dva osnovna parametra modela, povrata i rizika.

Pokazalo se da klasičan pristup ovom modelu ima dosta nedostataka prilikom primjene. Jedan od glavnih jest da su dobiveni udjeli porfelja jako osjetljivi na ulazne parametre, posebno na očekivani povrat i varijancu. Cilj je, dakle, problem optimizacije porfelja učiniti otpornijim. Potoji nekoliko nadogradnji MV modela koji rješavaju spomenute poteškoće, a jedan od njih je Bayesovski pristup. On parametre promatra na drugačiji način od klasičnog pristupa - opisuje ih kao slučajne varijable umjesto egzaktnih procjenitelja.

Za početak ćemo općenito opisati problem odabira porfelja, a kasnije ga nadograditi Bayesovskim pristupom.

6.1.1 Problem odabira porfelja

MV model pretpostavlja da su povrat i rizik jedino što investitior promatra prilikom odabira porfelja. Stoga je logično da preferira porfelj sa većim očekivanim povratom za danu razinu rizika. Pokazalo se da postoje tri formulacije spomenutog problema koje daju isto (optimalno) rješenje.

Pretpostavimo da postoji N imovina u koje imamo mogućnost ulagati. Označimo s R_t njihov vektor povrata u trenutku t

$$R_t = (R_{1,t}, ..., R_{N,t}),$$

te pretpostavimo da imaju multivarijantnu normalnu distribuciju

$$p(R_t|\mu, \Sigma) = N(\mu, \Sigma),$$

pri čemu je μ $(N \times 1)$ vektor očekivanja, a Σ $(N \times N)$ matrica kovarijanci. Udjeli u porfelju su omjeri sredstava uloženih u imovinu i, i = 1,...,N, i ukupnih sredstava. Označimo ih sa $\omega = (\omega_1, ..., \omega_N)'$.

Definicija 6.1. *Uz gornje oznake povrat portfelja dan je s*

$$R_p = \sum_{i=1}^{N} \omega_i R_{i,t} = \omega' \mathbb{R}_t,$$

dok su očekivani povrat i varijanca porfelja definirani kao

$$\mu_p = \sum_{i=1}^{N} \omega_i \mu_i = \omega' \mu,$$

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j cov(R_i, R_j) = \omega' \Sigma \omega,$$

pri čemu su μ_i očekivani povrati na imovinu i, a $cov(R_i, R_j)$ kovarijanca između povrata na imovinu i i j.

Pretpostavimo da je cilj investitora maksimizirati bogatstvo na kraju perioda držanja određenog porfelja $T + \tau$, gdje je T zadnji period za koji su dostupne informacije o porfelju, a τ vrijeme koje ga investitor planira držati. Već smo spomenuli da postoje tri načina formulacije MV modela. Sljedeća dva mogu se iskazati kao dualni problemi:

$$\begin{aligned} \min_{\omega} \sigma_p^2 &= \min_{\omega} \omega' \Sigma_{T+\tau} \omega \\ uz\ uvjet \quad & \omega' \mu_{T+\tau} \geq \mu^* \\ & \omega' 1 = 1, \end{aligned} \tag{61}$$

pri čemu je 1 $(N \times 1)$ vektor jedinica, a μ^* minimalan prihvatljivi povrat, te

$$\begin{aligned} \max_{\omega} \mu_p &= \max_{\omega} \omega' \mu_{T+\tau} \\ uz\ uvjet \quad & \omega' \Sigma_{T+\tau} \omega \leq \sigma^{2*} \\ & \omega' 1 = 1, \end{aligned} \tag{62}$$

gdje je σ^{2*} maksimalno prihvatljiv rizik. U trećoj formulaciji, optimizacija portfelja izražena je pomoću maksimizacije funkcije korisnosti investitora:

$$\begin{aligned} \max_{\omega} E[U(\omega' R_{T+\tau})] &= \max_{\omega} \int U(\omega' R_{T+\tau}) p_{T+\tau}(R_{T+\tau} | \mu, \Sigma) dR_{T+\tau} \\ uz\ uvjet \quad & \omega' 1 = 1. \end{aligned} \tag{63}$$

Rješavajući navedene probleme optimizacije dolazimo do optimalnog porfelja (vektora optimalnih udjela) danog s

$$\omega^* = \frac{\Sigma_{T+\tau}^{-1} \mu_{T+\tau}}{1' \Sigma_{T+\tau}^{-1} \mu_{T+\tau}}. \tag{64}$$

Naveli smo osnovne smjernice pri konstrukciji portfelja pod pretpostavkom da je vektor povrata normalno multivarijantno distribuiran. Preostaje nam odrediti parametre te distribucije. Klasičnim MV pristupom nepoznate parametre odredili bismo egzaktno na osnovu povijesnih podataka koje imamo i pretpostavili bi da se neće mijenjati tokom vremena, odnosno perioda držanja portfelja.

6.1.2 Bayesovska nadogradnja MV modela

Za razliku od klasičnog pristupa, koji se nije pokazao najpraktičnijim, sada ćemo u nepoznate parametre ugraditi neizvjesnost - promatrat ćemo ih kao slučajne varijable.

Prisjetimo se diskusije iz 4.2 koja se odnosila na predviđanja. Ako s R_{T+1} označimo vektor povrata sljedećeg perioda, tada nam je funkcija gustoće predviđanja dana s

$$p(R_{T+1}|R) \propto \int p(R_{T+1}|\mu,\Sigma)\pi(\mu,\Sigma|R)d\mu d\Sigma, \tag{65}$$

gdje je

$$\begin{aligned} \text{R} &= \text{podatci o povratima dostupni do vremena T} \\ & \quad\ \text{- } (T \times N) \text{ matrica} \\ \pi(\mu,\Sigma|R) &= \text{zajednička aposteriori gustoća parametara} \\ & \quad\ \text{multivarijantne normalne distribucije} \\ p(R_{t+1}|\mu,\Sigma) &= \text{multivarijantna normalna gustoća.} \end{aligned}$$

Primijetimo da zbog integriranja po svim mogućim parametrima, predviđanja ovise samo o povijesnim podatcima. Zbog pretpostavke o normalnosti povrata, može se pokazati da je distribucija predviđanja poznata distribucija (kao što ćemo vidjeti kasnije). Uzimajući u obzir navedenu funkciju gustoće predviđanja, probleme (61) i (63) možemo pisati na sljedeći način

$$\begin{aligned} \min_{\omega} \sigma_p^2 &= \min_{\omega} \omega'\widetilde{\Sigma}\omega \\ uz\ uvjet \quad & \omega'\widetilde{\mu} \geq \mu^* \\ & \omega' 1 = 1, \end{aligned} \tag{66}$$

$$\begin{aligned} \max_{\omega} E[U(\omega' R_{T+1})] &= \max_{\omega} \int U(\omega' R_{T+1}) p(R_{T+1}|R) dR_{T+1} \\ uz\ uvjet \quad & \omega' 1 = 1, \end{aligned} \tag{67}$$

pri čemu su $\widetilde{\mu}$ očekivanje, a $\widetilde{\Sigma}$ kovarijacijska matrica vektora povrata R_{T+1}. U ovom slučaju izraz za optimalan portfelj (64) dan je s

$$\omega^* = \frac{\widetilde{\Sigma}^{-1}\widetilde{\mu}}{1'\widetilde{\Sigma}^{-1}\widetilde{\mu}}. \tag{68}$$

6.2 Apriori uvjerenja i modeli vrednovanja imovine

Postoji mnoštvo financijskih modela koji opisuju ponašanje povrata na razne vrste imovine. Međutim, treba uzeti u obzir da niti jedan od njih nije u potpunosti precizan. Znači li to da ih ne treba koristiti? U ovom poglavlju pokazat ćemo kako investitor može kombinirati svoja apriori uvjerenja s određenim modelom, pri čemu je također u mogućnosti iskazati stupanj vjerovanja u valjanost tog modela. Nadovezat ćemo se i na prošlo poglavlje optimizacije porfelja te obogatiti apriori uvjerenja s implikacijama promatranog modela.

Ekonometrijski model opisuje cijene ili povrate kao funkcije koje ovise o egzogenim varijablama - faktorima ili portfelj faktorima. One su obično mjere rizika. Ako investitor vjeruje u valjanost modela, tada će se njegov portfelj sastojati od kombinacije faktora koji ga izlažu samo onim izvorima rizika koji su utkani u model. Ako je pak investitor skeptičan u svezi modela i želi optimizirati svoj porfelj samo na osnovu rizika koje obuhvaća procjena parametara, tada će koristiti pristup objašnjen u 6.1.

Od sada pretpostavljamo da su modeli CAPM (Capital asset pricing model) odnosno općeniti faktorski model moguće alternative pri optimizaciji portfelja.

CAPM model CAPM model je ekonometrijski model baziran na dvije vrste pretpostavki:
(1) Investitori imaju averziju prema riziku te odluke o investiranju donose isključivo na osnovu očekivanog povrata i varijance promatranih povrata.
(2) Kapitalna tržišta su savršeno konkurentna. Postoji bezrizična imovina.
S obzirom na navedene pretpostavke CAPM model možemo zapisati u obliku

$$E(R_i) - R_f = \beta_i(E(R_M) - R_f), \tag{69}$$

pri čemu su
$E(R_i)$ = očekivani povrat na rizičnu imovinu i
R_f = povrat na bezrizičnu imovinu (konstanta)
$E(R_M)$ = očekivani povrat na tržišni portfelj
β_i = mjera tržišnog rizika na imovinu i.

CAPM tvrdi da je očekivani povrat na imovinu linearna funkcija koja ovisi isključivo o mjeri tržišnog rizika (beta). Rizik koji dolazi iz ostalih izvora možemo zanemariti. Model možemo zapisati u empirijski analognom obliku linearne regresije

$$R_{i,t} - R_f = \alpha + \beta(R_{M,t} - R_f) + \epsilon_{i,t}, \quad i = 1, ..., K, \tag{70}$$

gdje je:
$R_{i,t}$ = povrat na imovinu i u trenutku t
$R_{M,t}$ = povrat na tržišni portfelj u trenutku t
$\epsilon_{i,t}$ = specifični povrat na imovinu i u trenutku t.

Faktorski model Općeniti faktorski model tvrdi da je očekivani povrat na imovinu proporcionalan linearnoj kombinaciji faktora koji opisuju rizik, što zapisujemo kao

$$E(R_i) - R_f = \beta_{i,1}(E(f_1) - R_f) + ... + \beta_{i,k}(E(f_k) - R_f), \tag{71}$$

gdje je:
$E(f_j)$ = očekivani povrat faktora j
$\beta_{i,j}$ = osjetljivost očekivanog povrata imovine i na faktor j.

Kako bismo odredili ,,osjetljivosti" $\beta_{i,j}$ model pišemo u empirijskom obliku

$$R_{i,t} - R_f = \alpha + \beta_{i,1}(f_{1,t} - R_f) + ... + \beta_{i,K}(f_{K,t} - R_f) + \epsilon_{i,t}. \tag{72}$$

Koeficijent α u (70) i (72) zovemo ex post, a označava mjeru sposobnosti procjene vrijednosti imovine aktivnog portfelj menadžera. Pozitivan (negativan) α znači da je vrijednost imovine podcijenjena (precijenjena).

6.2.1 Perturbirani model

U slučaju da je promatrani model valjan, α je jednaka 0. Ekvivalentno rečeno, valjani model ne dopušta greške pri procjeni vrijednosti. Zamislimo investitora koji je skeptičan oko preciznosti procjene u (69) i (71). Taj skepticizam možemo izraziti pomoću λ:

$$E(R_i) - R_f = \lambda + \beta_i(E(R_M) - R_f)$$

ili

$$E(R_i) - R_f = \lambda + \beta_{i,1}(E(f_1) - R_f) + ... + \beta_{i,k}(E(f_k) - R_f).$$

Naš cilj je, pomoću Bayesovske teorije, odrediti perturbirani model kako bi reflektirali investitorovu nesigurnost u snagu procjene modela. U konačnici će distribucija predviđanja biti rezultat ne samo određivanja parametara, već i investitorove apriori nesigurnosti ,,nadograđene" informacijama iz povijesnih podataka.

Zbog jednostavnosti, od sada ćemo povrat na imovinu R_t promatrati u ovisnosti o K benchmark portfelja (egzogene varijable - faktori). U slučaju CAPM modela K = 1, s obzirom da je samo jedan benchmark portfelj - tržišni porfelj. Podaci su dostupni za T perioda, t = 1, 2, ..., T. Model koji želimo procjeniti dan je s:

$$R_t = \alpha + \beta_1 f_{t,1} + ... + \beta_K f_{t,K} + \epsilon_t, \tag{73}$$

što možemo zapisati u matričnoj formi:

$$R = Xb + \epsilon, \tag{74}$$

pri čemu je:

	R	=	$(T \times 1)$ vektor povrata na rizičnu imovinu
	X	=	(1F) matrica
	F	=	$(T \times K)$ matrica benchmark povrata
	ϵ	=	$(T \times 1)$ vektor slučajnih pogrešaka
	b	=	$((K+1) \times 1)$ vektor regresijskih koeficijenata.

Kao što znamo iz poglavlja 5, problem Bayesovske linearne regresije obuhvaća sljedeće korake: određivanje funkcije vjerodostojnosti parametara modela, subjektivno odabiranje apriori distribucija te računanje aposteriori distribucija. Sada ćemo proći kroz sve korake.

Funkcija vjerodostojnosti

Pretpostavljamo da slučajne pogreške u (74) nisu korelirane s benchmark povratima F, te da su nezavisne i jednako normalno distribuirane sa očekivanjem 0 i varijancom σ^2. Slijedi da su povrati R također normalno distribuirani s očekivanjem Xb i varijancom σ^2

$$R \sim N(Xb, \sigma^2 I_T),$$

pri čemu je I_T $(T \times T)$ matrica identiteta. Funkcija vjerodostojnosti parametara b i σ^2 dana je s (kao u poglavlju 5)

$$L(b, \sigma | R, X) \propto (\sigma^2)^{-T/2} \exp\Big(- \frac{1}{2\sigma^2}(R - Xb)'(R - Xb)\Big). \tag{75}$$

Postavlja se pitanje na koji način tretirati benchmark povrate F, kao konstantne ili kao stohastičke varijable. Kako bismo uzeli u obzir nesigurnost pri određivanju komponenti od F, pretpostavit ćemo da prate normalnu multivarijantnu distribuciju

$$F_t \sim N(E, V),$$

pri čemu je: F_t = $(1 \times K)$ vektor benchmark povrata u trenutku t
E = $(1 \times K)$ vektor očekivanih povrata
V = $(K \times K)$ kovarijacijska matrica benchmark povrata.

Funkcija vjerodostojnosti parametara E i V dana je s

$$L(E, V|F) \propto |V|^{-T/2} \exp\left(-\frac{1}{2} \sum_{t=1}^{T} (F_t - E)' V^{-1} (F_t - E) \right). \tag{76}$$

Apriori distribucije

Kao što smo spomenuli, perturbaciju idealnog modela izrazit ćemo pomoću apriori distribucije parametra α. Očekujemo da je model valjan što znači da očekivanje od α postavljamo na 0. Odabiranjem vrijednosti standardne devijacije σ_α investitor reflektira stupanj nesigurnosti u promatrani model vrednovanja imovine. Što je manji σ_α veća je sigurnost u implikacije modela.

Pretpostavljamo da parametri modela prate konjugirane apriori distribucije (kao u Poglavlju 5)

$$b|\sigma \sim N(b_0, \sigma^2 \Sigma_0) \tag{77}$$

$$\sigma^2 \sim Inv - \chi^2(v_0, c_0^2). \tag{78}$$

Parametri v_0 i c_0 određeni su s (52) odnosno (53). Apriori očekivanje regresijskih parametara je

$$b_0 = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix},$$

pri čemu je $\alpha_0 = 0$ kao što je objašnjeno ranije. Kovarijacijska matrica Σ_0 je pozitivno definitna matrica definirana na sljedeći način:

$$\Omega_0 = \begin{pmatrix} \sigma_\alpha^2 \frac{1}{\sigma^2} & 0 \\ 0 & \Sigma_\beta \end{pmatrix}. \tag{79}$$

Elemente izvan dijagonale postavili smo na 0 s obzirom da nemamo apriori razloga vjerovati kako su α i regresijski koeficijenti u korelaciji. Nadalje, pretpostavljamo da je Σ_β dijagonalna matrica s velikim vrijednostima na dijagonali, kako nebi previše utjecali na zaključak o koeficijentima β. Cilj je odrediti prvi element, odnosno σ_α^2. Uskoro ćemo istražiti kakav utjecaj imaju različite veličine od σ_α^2 na odabir optimalnog portfelja.

Pretpostavljamo da je apriori distribucija vektora očekivanja E i kovarijacijske matrice V benchmark povrata Jeffreyjeva[7] apriori distribucija:

$$E, V \propto |V|^{-(T+1)/2}.$$

Aposteriori distribucije

Aposteriori distribucija parametra b uvjetno na σ^2 Kako smo pretpostavili konjugirane apriori distribucije, očekujemo da će aposteriori biti istog oblika s nadograđenim parametrima.

Aposteriori distribucija parametra b uvjetno na σ^2 je multivarijantna normalna distribucija s očekivanjem b^* i kovarijacijskom matricom $\sigma^2\Omega^*$,

$$b|\sigma^2, R, X \propto N(b^*, \sigma^2\Omega^*), \tag{80}$$

pri čemu pomoću (54) i (55) dobijemo izraze za hiperparametre

$$\Omega^* = (\Omega_0^{-1} + X'X)^{-1}, \tag{81}$$

$$b^* = \begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \Omega^*(\Omega_0^{-1}b_0 + X'X\hat{b}). \tag{82}$$

U (82) $\hat{b}$ označava procjenitelj od b metodom najmanjih kvadrata (koji je jednak kao i procjenitelj metodom maksimalne vjerodostojnosti).

Aposteriori distribucija parametra σ^2 Aposteriori distribucija parametra σ^2 je inverzna χ^2 distribucija,

$$\sigma^2 \sim Inv - \chi^2(\upsilon^*, c^{2*}), \tag{83}$$

s aposteriori parametrima dobivenim pomoću (56) i (57):

$$\upsilon^* = \upsilon_0 + T \tag{84}$$

$$\upsilon^* c^* = \upsilon_0 c_0^2 + (R - X\hat{b})'(R - X\hat{b}) + (b_0 - \hat{b})'K(b_0 - \hat{b}), \tag{85}$$

gdje je $K = ((X'X)^{-1} + \Omega_0)^{-1}$.

[7] Pretpostavljamo neovisnost između elemenata vektora E i V.

Aposteriori distribucije benchmarh parametara Aposteriori distribucije parametara benchmark povrata su normalna i inverzna Wishartova[8]

$$E|V, F \sim N\big(\hat{E}, \frac{V}{T}\big) \tag{86}$$

$$V|F \sim IW(\Psi, T-1), \tag{87}$$

pri čemu je

$$\hat{E} = \frac{1}{T}\sum_{t=1}^{T} F_t$$

i

$$\Psi = \sum_{t=1}^{T}(F_t - \hat{E})'(F_t - \hat{E}).$$

6.2.2 Distribucije predviđanja i Odabir portfelja

U prvom dijelu ovog poglavlja obradili smo problem odabira portfelja te ga nadogradili Bayesovskim pristupom. Pokazalo se da je krajnji rezultat, ujedno i naš cilj, funkcija gustoće predviđanja dana s (65), kao i formula za optimalan portfelj dana s (68).

Označimo s F_{T+1} $(1 \times K)$ vektor benchmark povrata, te s R_{T+1} vektor povrata rizične imovine sljedećeg perioda. Obzirom da investitor analizira vezu između rizične imovine i benchmark portfelja, prediktivni parametri $\widetilde{\mu}$ i $\widetilde{\Sigma}$ su zapravo zajedničko očekivanje i varijanca vektora (R_{T+1}, F_{T+1}).

Kako smo pretpostavili da su benchmark predviđanja slučajne varijable, prvo moramo predvidjeti F_{T+1} kako bi bili u mogućnosti odrediti R_{T+1}.

Može se pokazati da predviđanje F_{T+1} ima multivarijantnu Studentovu t-distribuciju[9] s T - K stupnjeva slobode. Očekivanje i varijanca od F_{T+1} su dani s:

$$\widetilde{E}_F = \hat{E} \quad i \quad \widetilde{V}_F = \frac{T+1}{T(T-K-2)}\Psi. \tag{88}$$

[8] Inverzna Wishartova distribucija jest konjugirana distribucija za kovarijacijsku matricu normalne distribucije. Neka je Q pozitivno definitna matrica i sa S označimo njezin inverz, $S = Q^{-1}$. Kažemo da S ima inverznu Wishartovu distribuciju i pišemo $S \sim IW(\Psi, v)$ ako joj je funkcija gustoće dana s

$$f(S|\Psi, v) = \frac{|\Psi|^{v/2}}{2^{vN/2}\pi^{N(N-1)/4}\prod_{i=1}^{N}\Gamma(v-i+1)/2}\frac{\exp(-\frac{1}{2}trS^{-1}\Psi)}{|S|^{(v+n+1)/2}},$$

pri čemu je $\Psi = \Sigma^{-1}$, a v stupanjevi slobode takav da $v \geq N$.

[9] Kažemo da p-dimenzionalni slučajni vektor X s vrijednostima u $\mathbb{R}^p$ ima multivari-

Distribucija predviđanja od R_{T+1} je također multivarijantna Studentova t-distribucija ali s $T+\upsilon_0$ stupnjeva slobode, čije su očekivanje i varijanca dani s:

$$\widetilde{E}_R = \widetilde{E}_X b^* \quad i \quad \widetilde{V}_R = \frac{T+\upsilon_0}{T+\upsilon_0-2} c^{2*}(1-\widetilde{E}_X \Omega^* \widetilde{E}_X), \tag{89}$$

pri čemu je $\widetilde{E}_X = (1\widetilde{E}_F)$. Konačno kovarijanca predviđanja između povrata na rizičnu imovinu R_{T+1} i povrata na benchmark portfelj j, $F_{j,T+1}$ dana je s

$$\widetilde{V}_{R,F} = \beta_j^* \widetilde{V}_{F,jj}, \tag{90}$$

pri čemu je β_j^* aposteriori očekivanje j-tog koeficijenta, a $\widetilde{V}_{F,jj}$ j-ta dijagonalna komponenta od $\widetilde{V}_F$.

Objedinjujući (88), (89) i (90) dobivamo zajedničko očekivanje i kovarijancu predviđanja korištenu pri optimizaciji portfelja

$$\widetilde{\mu} = \begin{pmatrix} \widetilde{E}_R \\ \widetilde{E}_F \end{pmatrix}$$

i

$$\widetilde{\Sigma} = \begin{pmatrix} \widetilde{V}_R & \widetilde{V}_{R,F} \\ \widetilde{V}_{R,F} & \widetilde{V}_F \end{pmatrix}.$$

Koristeći formulu (68) možemo izračunati optimalan portfelj. Naglašavamo da nam ne treba analitička forma distribucije predviđanja, sve dok imamo njezino očekivanje i kovarijacijsku matricu.

---

jantnu Studentovu t-distribuciju sa υ stupnjeva slobode ako vrijedi

$$f(x|\upsilon,\theta,\Sigma) = \frac{\Gamma(\frac{\upsilon+p}{2})}{|\Sigma|^{1/2}\Gamma(\frac{\upsilon}{2})(\upsilon\pi)^{p/2}} \left(1+\frac{1}{\upsilon}(x-\theta)'\Sigma^{-1}(x-\theta)\right)^{-(\upsilon+p)/2}, \quad -\infty < x < \infty$$

gdje je Γ gama funkcija, $\upsilon > 0$, $\theta \in \mathbb{R}^p$, a Σ $(p \times p)$ simetrična pozitivno definitna matrica.

Literatura

[1] P. Congdon, *Bayesian Statistical Modelling*, Wiley, 2001.

[2] M. Huzak, *Matematička statistika*, predavanja, PMF-MO, 2014.

[3] M. I. Jordan, *Bayesian Modeling and Inference*, predavanja, UC Berkeley, 2010.

[4] T. Khun, *The Structure of Scientific Revolutions*, University of Chicago Press, 1996.

[5] S. T. Rachev et al., *Bayesian Methods in Finance*, Wiley, 2008.

[6] N. Sarapa, *Teorija vjerojatnosti*, Školska knjiga, Zagreb, 2002.

[7] H. Šikić, *Mjera i integral*, predavanja, PMF-MO, 2014.

Sažetak

Prvi cilj bio je predstavljanje teorije utkane u Bayesovske metode. Kako bismo to učinili, započeli smo s teorijom vjerojatnosti čija je opsežnost podređena metodološkim zahtjevima Bayesovskih okvira obrađenih u radu. Spomenuti uvodni dio zaključili smo s diskretnom verzijom Bayesovog teorema koji je kamen temeljac za daljnju raspravu.

Prvi korak bio je postaviti temelje Bayesovske analize, s naglaskom na praktične aspekte. Cilj je bio iskoristiti njezine alate u svrhu kombiniranja informacija dobivenih iz različitih izvora - obrađenih podataka i apriori uvjerenja. Uveli smo pojmove funkcije vjerodostojnosti, apriori distribucija, aposteriori distribucija, kao i neprekidnu verziju Bayesovog teorema.

Zatim smo analizirali prirodu neinformativnih i informativnih apriori distribucija. Oprezan odabir parametara apriori distribucije nužan je kako bi osigurali da što preciznije odražavaju našu intuiciju. Također smo objasnili kako koristiti Bayesovske aposteriori zaključke u svrhu predviđanja budućih realizacija slučajnih varijabli.

Bayesovsko zaključivanje prošireno je univarijantnim i multivarijantnim regresijskim modelima. U normalnim postavkama i pod konjugiranim pretpostavkama, aposteriori zaključci su standardni.

Drugi cilj bio je predstaviti općeniti okvir pomoću kojega izražavamo stupanj vjerovanja u modele vrednovanja imovine prilikom odabira optimalnog portfelja. Za početak ovog dijela analizirali smo Bayesovski pristup MV (mean-variance) modelu. Klasični pristup koristi egzaktne procjenitelje za očekivanje i kovarijancu promatranih povrata i smatra ih točnima. Nedostatak ovakvog razmišljanja očituje se u tome što su težine optimalnog portfelja osjetljive da male promjene ulaznih parametara. Rješavanje problema Bayesovskim alatima pomaže nositi se s takvim preprekama.

Konačno, kombinirali smo apriori uvjerenja s modelima vrednovanja imovine kako bi predstavili strukturu i ekonomsku opravdanost Bayesovskog odabira portfelja.

Summary

Our first main goal was to provide an overview of the theory of Bayesian methods. To do so, we started with probability theory, whose depth and scope are subordinated to the methodological requirements of the Bayesian framework discussed throughout this thesis. We closed that introduction theory with discrete version of Bayes' theorem which is our cornerstone for following discussion.

First step was to lay foundations of Bayesian analysis, emphasizing its practical rather then philosophical aspects. The objective was to employ its framework for combining information derived from different sources - the observed data and prior beliefs. We introduced likelihood functions, prior distributions, posterior distributions, and continuous version of Bayes' theorem.

We next turned to discussing the nature of uninformative and informative priors. Careful selection of the parameters of the prior distributions is necessary to ensure that they accurately reflect our prior intuition. We also presented how to use Bayesian posterior inference to predict realizations of the random variable ahead in time.

Bayesian inference is extended with the univariate and multivariate linear regression models. In a normal setting and under conjugate priors, the posterior and predictive results are standard.

Our second main goal was to introduce a general framework for reflecting degrees of belief in an asset pricing model when selecting the optimal portfolio. First we discuss the Bayesian approach to mean-variance portfolio selection. The classical framework uses the sample estimates of the mean and the covariance of returns as if they were the true parameters. The failure to account for the parameter uncertainty leads to optimal porfolio weights that are too sensitive to small changes in the inputs of the portfolio optimization problems. Casting the problem in a Bayesian framework helps deal with this sensitivity.

Finally, we were combining prior beliefs with asset pricing models to introduce structure and economic justification into Bayesian portfolio selection.

Životopis

Tončica Topić rođena je 25. kolovoza 1992. godine u Splitu. Nakon završene Osnovne škole ,,Manuš" opredjeljuje se za Prirodoslovno - matematičku gimnaziju. Akademske godine 2011./2012. upisuje preddiplomski studij Matematika na Prirodoslovno - matematičkom fakultetu Sveučilišta u Zagrebu, na kojem, 2014. godine, stječe akademski naziv sveučilišne prvostupnice matematike (bacc. univ. math.). Iste godine je upisala diplomski studij Financijska i poslovna matematika.

Paralelno s obrazovanjem, Tončica se aktivno bavi sportom. Godine 2007. postaje član hrvatske mačevalačke reprezentacije, od kada je sedam puta odnijela titulu državne prvakinje u kategoriji mač seniorke. Postigla je i zapažene međunarodne rezultate u svim dobnim uzrastima.