Математички факултет

МАСТЕР РАД

Одређивање цена европских опција коришћењем непараметарске регресије

Студент:

Маријана Палибрк

Ментор:

Бојана Милошевић

Београд, 2020.

Садржај:

1 Увод

Опције су вредносни папири који представљају једну врсту финансијског деривата које инвеститори често користе да би се заштитили од различитих врста ризика у финансијском пословању. У раду ћемо се фокусирати на европске кол опције; *европске* у називу нам говоре да их можемо активирати само у уговорено време, док *кол* да имамо право, али не и обавезу куповине вредносног папира по унапред договореној цени. Наш циљ ће бити да детаљно анализирамо опције на тржишту за финансијско тржиште $S\&P$ 500 и пронађемо модел који их најбоље описује. За одређивање оцене користићемо непараметарску регресију, док ћемо за проверу квалитета те оцене користити Монте Карло метод.

2 Непараметарска регресија

У овом поглављу упознаћемо се са основним појмовима о непараметарској регресији како бисмо их у даљем раду применили у вредновању опција. За више детаља се реферишемо на књигу [10].

2.1 Локално полиномна регресија

Дато нам је n парова опсервација (X_1, Y_1), $\ldots$, (X_n, Y_n) и регресиона функција $m(x)$, која је изражена као очекивана вредност случајне промењиве Y под условом да је $X = x$, то јест

$$m(x) = E(Y|X = x).$$

За узорак $(x_1, Y_1), \ldots, (x_n, Y_n)$ везу између промењивих Y и x можемо приказати следећом једнакошћу

$$Y_i = m(x_i) + \epsilon_i,\ E(\epsilon_i) = 0,\ D(\epsilon_i) = \sigma^2,\ i = 1, \ldots, n.$$

Сада ћемо претпоставити да је регресиона функција локални полином степена p. Нека је u фиксна тачка у којој желимо да оценимо функцију $m(u)$. За вредност x у близини u, дефинишимо полином помоћу Тејлорове формуле:

$$m(u) \approx \sum_{k=0}^{p} \beta_k(x)(u - x)^k,$$

где је $\beta_k(x) = m^{(k)}(x)/k!$.

Да бисмо осигурали да оцењени коефицијенти одражавају локалну природу приказа, требало би користити регресију са тежинском функцијом $k_i = K_h(x_i - x) = K((x_i - x)/h)/h$, где је $K(\cdot)$ језгро и h параметар размере, стављајући већу тежину на тачке у близини тачке x. Оцене $\hat{\beta}_k(x)$ коефицијената $\beta_k(x)$ добијамо минимизовањем израза

$$\sum_{i=1}^{n} \left(y_i - \sum_{k=0}^{p} \beta_k(x)(x_i - x)^k \right)^2 K_h(x_i - x).$$

Ова регресија је локална у смислу да регресиони коефицијенти у једначини важе само у околини тачке x.

Оцена регресионе функције, као и њених извода, дата је са

$$\hat{m}^{(k)}(x) \equiv \hat{m}_{k,p}(x) = k!\hat{\beta}_{k,p}(x),\ k = 0, \ldots, p.$$

Специјално, када је $p = 0$ тада имамо Надараја-Вотсон ($Nadaraya - Watson$) регресију помоћу језгра, док за $p = 1$ имамо локално линеарну регресију, за $p = 2$

локално квадратну регресију, итд.
Уопштено, оцена коефицијената $\hat{\beta}_p = (\hat{\beta}_{0,p}, \hat{\beta}_{1,p}, \ldots, \hat{\beta}_{p,p})^T$ може бити записана као

$$\hat{\beta}_p = \begin{pmatrix} S_{n,0} & S_{n,1} & \cdots & S_{n,p} \\ S_{n,1} & S_{n,2} & \cdots & S_{n,p+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n,p} & S_{n,p+1} & \cdots & S_{n,2p} \end{pmatrix}^{-1} \begin{pmatrix} T_{n,0} \\ T_{n,0} \\ \vdots \\ T_{n,p} \end{pmatrix},$$

где је $S_{n,j} = \sum\limits_{i=1}^{n} (x_i - x)^j k_i$ и $T_{n,j} = \sum\limits_{i=1}^{n} (x_i - x)^j y_i k_i$.

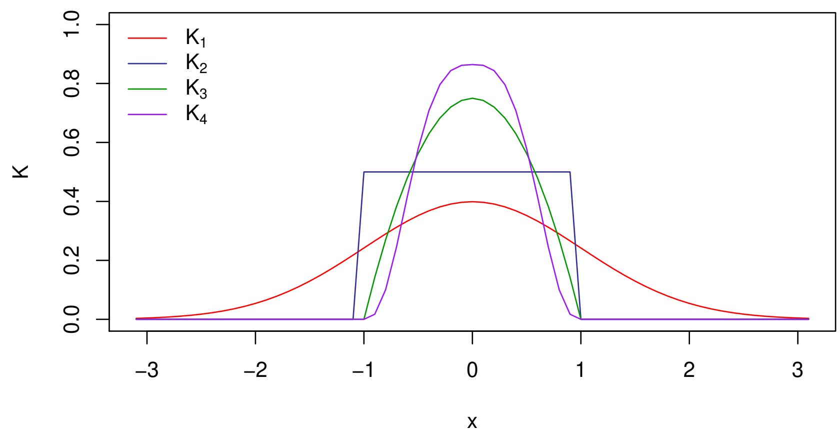
Језгро K је било која глатка функција која испуњава следеће услове:

- $K(x) \geq 0$;
- $\int\limits_{-\infty}^{\infty} K(x)dx = 1$;
- $\int\limits_{-\infty}^{\infty} xK(x)dx = 0$;
- $\sigma_K^2 \equiv \int\limits_{-\infty}^{\infty} x^2 K(x)dx > 0$.

Примери неких најчешће коришћених језгара су:

- Гаусово језгро: $K_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$;
- Униформно језгро: $K_2(x) = \frac{1}{2} I(x)$;
- Јепањешниково језгро: $K_3(x) = \frac{3}{4}(1 - x^2) I(x)$;
- Троугаоно језгро: $K_4(x) = \frac{70}{81}(1 - |x|^3)^3 I(x)$;

где је $I(x) = \begin{cases} 1, & |x| \leq 1 \\ 0, & |x| > 1 \end{cases}$.



Слика 1: Графички приказ претходно поменутих језгара.

Утицај параметра размере h на оцењивање густине показаћемо на следећем примеру. Генерисаћемо узорак из стандардне нормалне расподеле и са f означити одговарајућу густину, а затим оценити густину $\hat{f}_1$ коришћењем параметра размере h израчунатог Силвермановом методом и видети шта се дешава са оценом у случају када је параметар размере два пута мањи ($\hat{f}_2$) и два пута већи ($\hat{f}_3$) од израчунатог h. Оптимална оцена параметра размере добијена Силвермановом методом је минимизовањем средње интегралне квадратне грешке ($MISE$) ([9]):

$$MISE(\hat{f}) = E\left(\int_0^\infty (\hat{f}(x) - f(x))^2 dx\right). \tag{2.1}$$

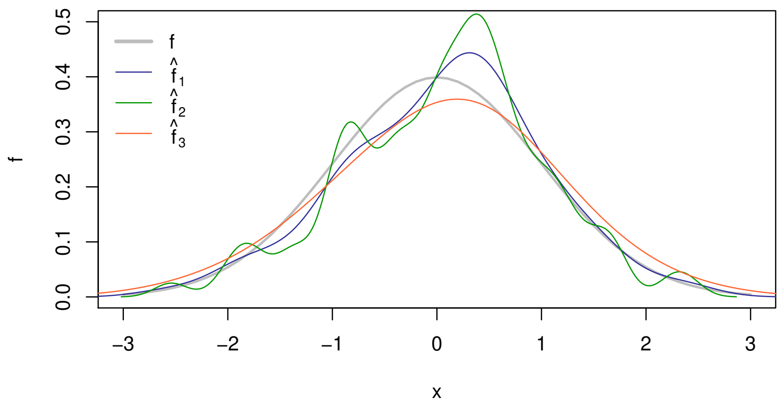
Како оцена добијена минимизовањем израза (2.1)

$$h^* = \left(\frac{\int (K(x))^2 dx}{n(\int x^2 K(x) dx)^2 \int (f''(x))^2 dx}\right)^{\frac{1}{5}} \tag{2.2}$$

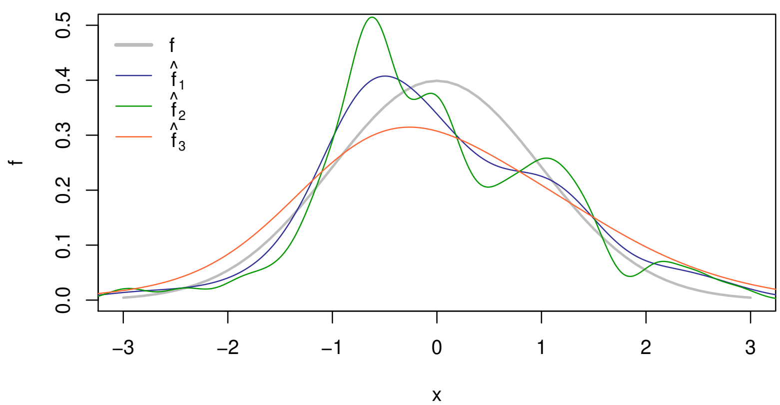
зависи од непознате функције f, ако претпоставимо да је f густина нормалне расподеле $\mathcal{N}(\mu, \sigma^2)$ и K Гаусово језгро, добијамо оптималну оцену параметра размере

$$h_{opt} = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}n^{-\frac{1}{5}}. \tag{2.3}$$

Како би побољшао квалитет оцене (2.3), Силверман је коефицијент 1.06 заменио коефицијентом 0.9 и $\hat{\sigma}$ са $\min(\hat{\sigma}, \frac{IQR}{1.34})$, при чему је $\hat{\sigma}$ стандардно одступање узорка, n величина узорка и IQR интерквартилно растојање.



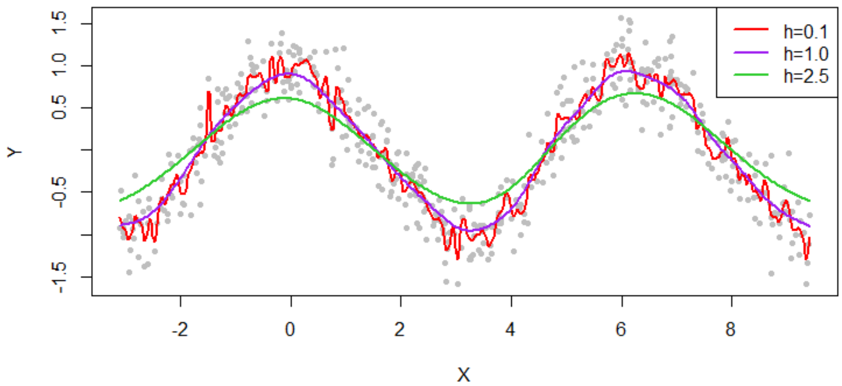
Слика 2: Оцене густина за различите вредности параметра h.



Слика 3: Оцене густина за различите вредности параметра h.

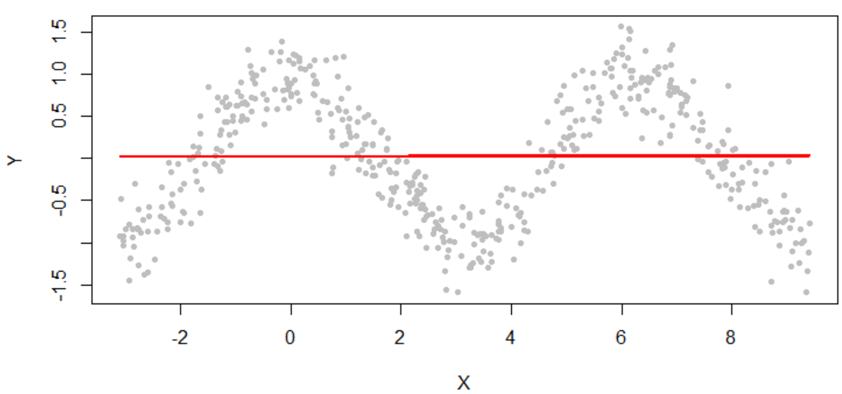
Као што видимо, одабир параметра размере је јако битан. Он утиче на то да ли ћемо превише или премало изгладити криву. Постоје разне методе одређивања параметра h као што су унакрсна валидација, минимизовање средње квадратне грешке, као и средње интегралне квадратне грешке, итд. У овом раду ћемо се фокусирати на параметар добијен из интегралне средње квадратне грешке, што ће детаљније бити објашњено у наставку.

На следећем графику можемо видети како h утиче на глаткост криве. Приликом коришћења предложене методе, треба направити компромис између пристрасности и дисперзије, јер када је h мало дисперзија је велика, али је пристрасност мала; у случају када је h велико, дисперзија је мала, али је пристрасност велика.



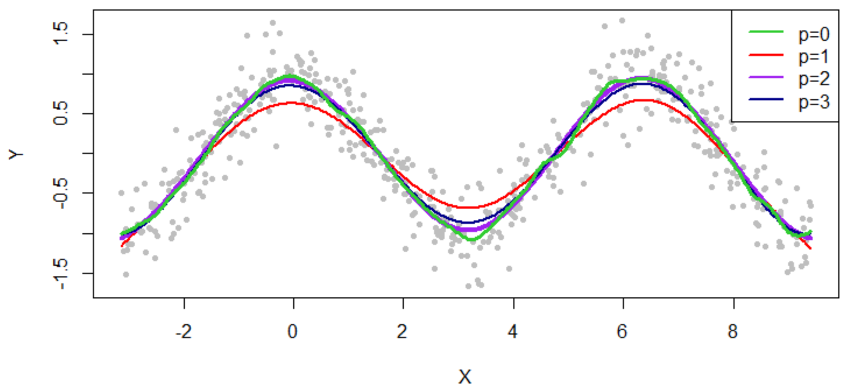
Слика 4: Симулирани подаци и оцена добијена локалном полиномном регресијом када је $p = 0$ $(Nadaraya - Watson)$. Приликом оцењивања коришћено је Гаусово језгро и три различита параметра размере.

Пре него што кренемо даље, графички ћемо представити разлику између линеарне регресије и претходно описане локалне полиномне регресије.



Слика 5: Симулирани подаци и оцена добијена линеарном регресијом.

Са графика видимо да је веза између X и Y нелинеарна и да оцена добијена линеарном регресијом не приказује добро њихову везу.



Слика 6: Оцене добијене локално полиномном регресијом за $p = 0, 1, 2, 3$ коришћењем Гаусовог језгра и глобалног параметра размере h добијеног методом која је описана у наставку.

2.2 Избор параметра размере

Параметар размере (пропусни опсег) за $h = 0$ резултира у интерполацији у свакој тачки и то је најкомплекснији модел, док бесконачан параметар размере резултира једноставним моделом у виду (једног) полинома степена p. Дакле, одабир параметра размере еквивалентан је одабиру комплексности модела. Стога је пожељно ослонити се на аутоматске процедуре које уклањају сваку потенцијалну произвољност при избору параметра размере.
За параметар размере који зависи од x кажемо да је локални. Уколико не зависи од x тада је то глобални параметар размере.

Локални параметар размере h_{local} добијамо минимизирањем условне средњеквадратне грешке у тачки x:

$$\{E[(\hat{m})^{(k)}(x)|x] - m^{(k)}(x)\}^2 + D[(\hat{m})^{(k)}(x)|x]. \tag{2.4}$$

У том случају h_{local} изражава се једнакошћу:

$$h_{local}(x) = C_{k,p}\left[\frac{\upsilon(x)}{\{m^{(p+1)}(x)\}^2\pi(x)} \times \frac{1}{n}\right]^{1/(2p+3)} \tag{2.5}$$

где је $\pi(x)$ маргинална густина регресора и $\upsilon(x)$ њихова дисперзија.

Уколико нас интересује глобални параметар размере h_{global} њега добијамо минимизацијом следећег израза:

$$\int\{\{E[(\hat{m})^{(k)}(x)|x]-m^{(k)}(x)\}^2+D[(\hat{m})^{(k)}(x)|x]\}\omega(x)dx \tag{2.6}$$

и оптимални параметар размере једнак је:

$$h_{global}=C_{k,p}\left[\frac{\int\upsilon(x)\omega(x)/\pi(x)dx}{\int\{m^{(p+1)}(x)\}^2\omega(x)dx\}}\times\frac{1}{n}\right]^{1/(2p+3)}. \tag{2.7}$$

Константе $C_{k,p}$ зависе од избора језгра. Претходно наведене формуле садрже непознате величине: $\pi(x), \upsilon(x)$ и $m^{(p+1)}(x)$, па пре него што пређемо на израчунавање параметра размере морамо прво њих оценити. Једноставан начин да се то уради је помоћу полинома степена $p+3$: $m(x)=\sum_{k=0}^{p+3}\alpha_k x^k$. Коефицијенти α_k одређују се методом најмањих квадрата, υ преко суме квадрата резидуала (тако да је оцена независна од x) а $m^{(p+1)}(x)$ као полином другог реда добијен диференцирањем полинома $m(x)$:

$$m^{(p+1)}(x)=\sum_{k=p+1}^{p+3}\alpha_k k(k-1)...(k-p+1)x^{k-(p+1)}. \tag{2.8}$$

За глобални оптимални параметар размере стандардан избор тежинских функција је $\omega(x)=\omega_0(x)\pi(x)$ где је ω_0 фиксирана функција. На пример

$$\omega_0(x)=\begin{cases}1, & E(X)-1.5\sqrt{D(X)}\leq x\leq E(X)+1.5\sqrt{D(X)}\\ 0, & \textit{иначе}\end{cases}.$$

У овом случају $\int\omega_0(x)dx=3\sqrt{D(X)}$ оцењује се заменом $D(X)$ са узорачким моментом.
Оцена глобалног параметра размере у том случају је:

$$\hat{h}_{global}=C_{k,p}\left[\frac{ssr\times\int\omega_0(x)dx}{\sum_{i=1}^{n}\{m^{(p+1)}(X_i)\}^2\omega_0(X_i)}\times\frac{1}{n}\right]^{1/(2p+3)} \tag{2.9}$$

где је ssr сума квадрата резудуала из регресије 2.8.

3 Европске кол опције

У овом поглављу упознаћемо се са основним појмовима о опцијама, за више детаља реферишемо се на [8] и [5].

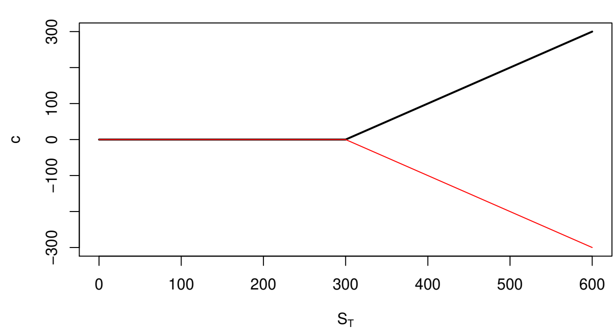
Дефиниција 3.1 *Опција је финансијски уговор којим се стиче право, али не и обавеза, да се купи или прода акција, или неки други вредносни папир, при договореним условима о уговореној цени или цени по истеку или на доспећу вредносног папира на који се односи опција и уговореном времену до истека опције.*

У раду ћемо се фокусирати на европске кол опције. *Европске* у називу нам говоре да их можемо активирати само у уговорено време, док *кол* да имамо право али не и обавезу куповине вредносног папира по унапред договореној цени.

Фактори који утичу на цену опција:

- почетна цена вредносног папира - S_0;
- уговорена цена - K;
- време до истека опције - T;
- волатилност вредносног папира - σ;
- каматна стопа - r;
- очекиване дивиденде током трајања опције - δ.

Добит кол опције у тренутку истека T износи $c = \max(S_T - K, 0)$.



Слика 7: Добит купца (црном линијом) и продавца (црвеном линијом) кол опције у тренутку истека када је уговорена цена $K = 300$.

Цена опције при процени која је неутрална од ризика је једнака садашњој вредности математичког очекивања од добити опције.

4 Оцењивање густине неутралне од ризика

4.1 Формулација проблема

Функција густине неутрална од ризика (RND) је фундаменталан концепт у финансијској математици и користи се при одређивању цене финансијских деривата. Дефинишемо је на следећи начин:

Дефиниција 4.1 (RND) *Функција густине неутрална од ризика за хартију од вредности је функција густине за коју је тренутна цена хартије од вредности једнака дисконтованој очекиваној вредности њених будућих цена.*

Циљ нам је да оценимо RND $f(\cdot)$ на основу цене опције на тржишту.

RND функција може се оцењивати помоћу параметарских и непараметарских метода. Оба приступа имају своје предности и мане. На пример, параметарске методе укључују мали број параметара и рачун који није захтеван. Са друге стране, недостаци се могу јавити уколико смо претпоставили расподелу која није довољно флексибилна да опише податке са којима радимо. Код непараметарских метода не морамо да претпоставимо расподелу RND функције тако да ту не можемо направити грешку. Ипак, непараметарски приступи су рачунски интензивнији од параметарских. Осим наведених предности и мана које су основне постоје и многе друге које су везане за конкретан метод који примењујемо, о чему ћемо касније дискутовати.

У наставку ће бити представљена и два модела, Блек-Шолс-Мертон модел и месхавина две логнормалне функције, како бисмо упоредили и одредили који модел најбоље описује цену опције и функцију неутралну од ризика (RND) .

4.2 Цена опција

Цена европске кол опције на акцију изражена је следећом једнакошћу:

$$C(K) = e^{-rt} \int_0^\infty max(0, S-K) f(S) dS$$

$$= e^{-rt} \int_K^\infty (S-K) f(S) dS \tag{4.1}$$

при чему је $f(\cdot)$ RND функција, K је уговорена цена, t је време до истека опције, а r безризична каматна стопа.

Напомена: Функција f зависи од тренутне цене акције, тј. $f(S) = f(S|S_0)$.

Диференцирањем функције $C(K)$ по K добијамо следећу једнакост:

$$\frac{\partial C(K)}{\partial K} = -e^{-rt}(S-K)f(S)\Big|_{S=K} + e^{-rt}\int_K^\infty \frac{\partial (S-K)f(S)}{\partial K}dS$$
$$= 0 + e^{-rt}\int_K^\infty -f(S)dS$$
$$= -e^{-rt}\int_K^\infty f(S)dS. \tag{4.2}$$

Уколико поново диференцирамо по K имамо:

$$\frac{\partial^2 C(K)}{\partial K^2} = e^{-rt}f(S)\Big|_{S=K} = e^{-rt}f(K). \tag{4.3}$$

Одавде видимо да функцију густине неутралну од ризика можемо изразити на следећи начин:

$$f(K) = e^{rt}\frac{\partial^2 C(K)}{\partial K^2}. \tag{4.4}$$

Претходна једнакост говори нам да уколико знамо функцију цене опције и дисконтни фактор лако можемо добити и RND функцију.

4.3 Ограничења RND функције

У овом одељку разматрамо ограничења која је потребно увести при оцени RND функције. С обзиром да говоримо о функцији густине расподеле она мора бити ненегативна и интеграл те функције по њеној области дефинисаности једнак јединици. Дакле:

$$f(K) \geq 0, K \in [0, +\infty)$$

$$\int_0^\infty f(S)dS = 1. \tag{4.5}$$

Још један услов који важи у овом случају је непостојање арбитраже односно прављења профита без ризика ([7]):

$$max(0, S_0e^{-\delta t} - Ke^{-rt}) \leq C(K) \leq S_0e^{-\delta t}$$

где је K уговорена цена, r безризична каматна стопа, δ дивиденда, S_0 садашња цена акције, а t време до истека опције. Да бисмо дискутовали о овом услову

извешћемо једнакост коју већ имамо, о цени опције на основу валуације ризик неутралним приступом.
На основу једнакости 4.1, 4.2, 4.3 примећујемо да је $f(K) \geq 0$ и $\int_K^\infty f(S)dS \geq 0$ па важи:

$$C^{'}(K) \leq 0,\ C^{''}(K) \geq 0,\ K \in [0, \infty)$$

Приметимо да $f(K)$ не може идентички бити једнака нули на $[0, \infty)$, па не могу ни $C^{'}(K)$ и $C^{''}(K)$. У том случају $C^{'}(K)$ је растућа, а $C(K)$ конвексна опадајућа функција на својој области дефинисаности. За $C^{'}(K)$ важи:

$$C^{'}(\infty) = -e^{-rt} \lim_{K \to \infty} \int_K^\infty f(S)dS = 0 \tag{4.6}$$

$$C^{'}(0) = -e^{-rt} \int_0^\infty f(S)dS = -e^{-rt} \tag{4.7}$$

$C^{'}(K)$ је растућа на $[0, \infty)$ па је:

$$-e^{-rt} \leq C^{'}(K) \leq 0 \tag{4.8}$$

За $C(K)$:

$$C(\infty) = e^{-rt} \lim_{K \to \infty} \int_K^\infty (S - K) f(S)dS = 0 \tag{4.9}$$

$$C(0) = e^{-rt} \int_0^\infty (S - 0) f(S)dS = e^{-rt} E(S) = S_0 e^{-\delta t} \tag{4.10}$$

где је $E(S)$ очекивана вредност акције у тренутку t. Последња једнакост потиче од својства мартингала у теорији цена опција:

$$e^{-(r-\delta)t} E(S) = S_0$$

где је $e^{-(r-\delta)t}$ дисконтни фактор. Како је $C(K)$ опадајућа на $[0, \infty)$:

$$0 \leq C(K) \leq S_0 e^{-\delta t}. \tag{4.11}$$

Како је $C^{'}(K)$ растућа функција, то јест $C^{'}(K) \geq C^{'}(0)$, имамо:

$$\begin{aligned} C(K) &= C(0) + \int_0^K C^{'}(S)dS \\ &\geq C(0) + \int_0^K C^{'}(0)dS \\ &= C(0) + KC^{'}(0) \\ &= S_0 e^{-\delta t} - K e^{-rt}. \end{aligned} \tag{4.12}$$

Из једнакости 4.12 видимо да за $C(K)$ важи:

$$max(0, S_0e^{-\delta t} - Ke^{-rt}) \leq C(K) \leq S_0e^{-\delta t}. \tag{4.13}$$

а то је управо услов непостојања арбитраже за цене кол опција. Дакле, да бисмо гарантовали да арбитража не постоји довољно је да важе следећи услови:

$$\begin{cases} C^{''}(K) \geq 0,\ K \in [0, \infty), \\ C^{'}(0) = -e^{-rt}, \\ C^{'}(\infty) = 0, \\ C(0) = S_0e^{-\delta t}, \\ C(\infty) = 0. \end{cases}$$

Сетимо се услова 4.5:

$$f(K) \geq 0, K \in [0, +\infty)$$
$$\int_0^\infty f(S)dS = 1$$

и једнакости 4.1, 4.2, 4.3:

$$C(K) = e^{-rt}\int_K^\infty (S-K)f(S)dS$$
$$C^{'}(K) = \frac{\partial C(K)}{\partial K} = -e^{-rt}\int_K^\infty f(S)dS$$
$$C^{''}(K) = e^{-rt}f(K).$$

На основу њих, услови $C^{'}(\infty) = 0$ и $C(\infty) = 0$ су аутоматски задовољени. Осим тога важи и:

$$f(K) \geq 0 \Longleftrightarrow C^{''}(K) \geq 0, K \in [0, \infty),$$
$$\int_0^\infty f(S)dS = 1 \Longleftrightarrow C^{'}(0) = -e^{-rt}. \tag{4.14}$$

Узимајући у обзир претходно наведено закључујемо да при оцењивању RND функције треба да важе следећи услови:

$$\begin{cases} f(K) \geq 0, K \in [0, \infty), \\ \int_0^\infty f(S)dS = 1, \\ C(0) = S_0e^{-\delta t}. \end{cases} \tag{4.15}$$

Кад се одреди тражена оцена, потребно је проверити да ли су претходно поменути услови испуњени. Уколико нису испуњени, решење тог проблема може се наћи у [2] и [4].

4.4 Мера квалитета оцене

Приликом оцењивања врло је важно испитати квалитет оцене. Када се оцењује густина једна од најчешће коришћених мера је следећа:

$$RIMSE = \left(\int_0^\infty E[(f\hat{(x)} - f(x))^2]dx \right)^{1/2}. \tag{4.16}$$

Добро је познато да се средње квадратна грешка оцене MSE може представити као збир квадрата пристрасности SB и дисперзије оцене V, то јест: $MSE = SB+V$. При оцењивању функције густине ове мере су интегрисане, а потом кореноване, па се $RIMSE$ (*Root integrated mean square error*) слично као и MSE може представити у облику: $RIMSE^2 = RISB^2 + RIV^2$, где је:

- $RISB$ (*Root integrated squared bias*):

$$RISB = \left(\int_0^\infty [(E\hat{f}(x) - f(x))^2]dx \right)^{1/2}$$

- RIV (*Root integrated variance*):

$$RIV = \left(\int_0^\infty E[(\hat{f}(x) - Ef\hat{(x)})^2]dx \right)^{1/2}$$

$\hat{f}(x)$ означава оцену RND функције а $f(x)$ праву RND функцију.
Слично као средње-квадратна грешка MSE и $RIMSE$ представља укупни квалитет оцене. $RISB$ је мера прецизности, а RIV мера стабилности.

5 Примена на реалним подацима

У овом раду подаци су преузети из базе података $Yahoo\ finance$ [1]. Посматрано финансијско тржиште је $S\&P$ 500, берзански индекс на њујоршким берзама вредносних папира $NYSE$ и $NASDAQ$. Он обједињава 500 тржишно највреднијих деоничарских друштава чијим се деоницама у САД активно тргује.

Подаци су преузети 6. априла 2020. године, са временом до доспећа 31. августа 2020. године. Цена акције у тренутку преузимања износила је $S_t = 2663.68$. На основу тих података процењујемо функцију неутралну од ризика помоћу Блек-Шолс-Мертон модела и мешавине две логнормалне расподеле, како би се открило који модел је најбоље описује. Сва оцењивања се врше локалном полиномном регресијом, док се поређење добијених резултата врши на основу $RIMSE$.

Претпоставићемо да имамо Блек-Шолсову цену кол опције. Добро је познато да у оквиру Блек-Шолс-Мертон модела цена кол опције:

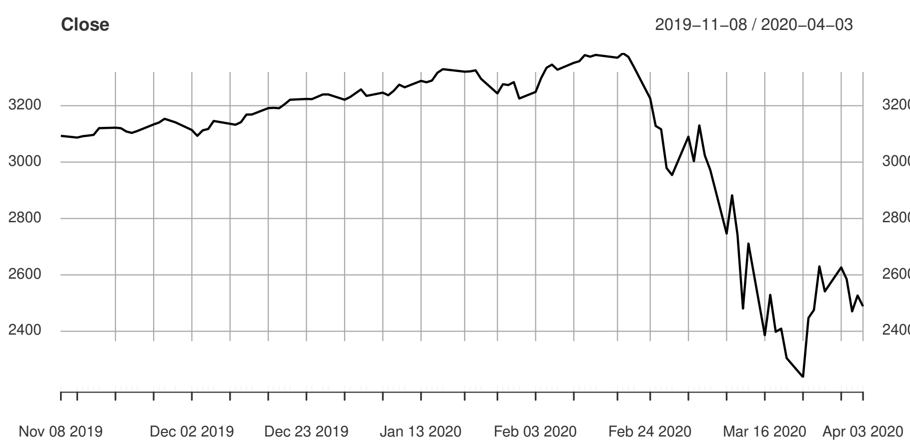
$$C(K) = C(t, T, S_t, K, r_{t,T}, \delta_{t,T}, \sigma) = e^{-\delta_{t,T}(T-t)} S_t N(d_1) - K e^{-r_{t,T}(T-t)} N(d_2),$$

где је N функција расподеле од стандарне нормалне расподеле, σ статистички оцењена волатилност и

$$d_1 = \frac{ln(S_t/K) + (r_{t,T} - d_{t,T} + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

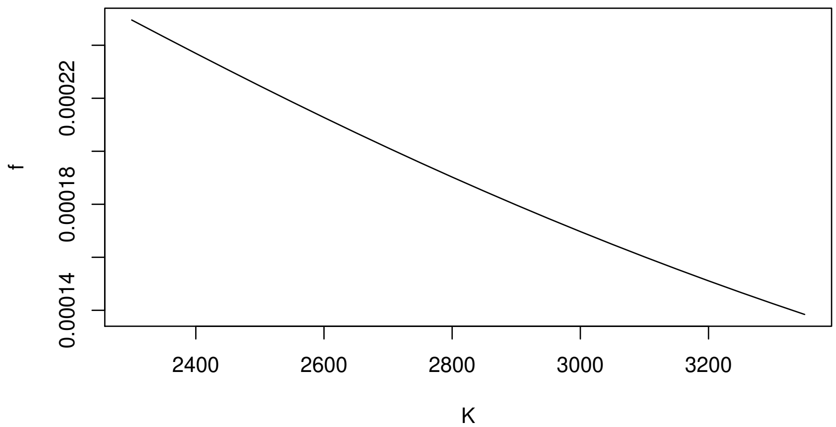
$$d_2 = d_1 - \sigma\sqrt{T-t},$$

имплицира логнормалну функцију неутралну од ризика за случајну промењиву $ln(S_T/S_t)$ са очекиваном вредношћу $\mu = (r_{t,T} - d_{t,T} + \frac{1}{2}\sigma^2)(T-t)$ и дисперзијом $\nu^2 = \sigma^2(T-t)$.
Један од начина да оценимо σ^2 је на основу историјских вредности приноса ([6]), али ту треба водити рачуна о понашању временске серије у одабраном временском интервалу.



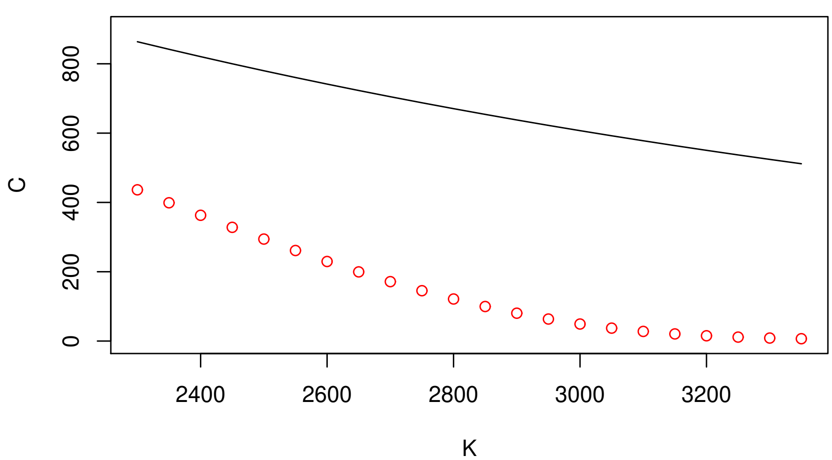
Слика 8: Цене акција за временски интервал од 8.11.2019. до 3.4.2020.

На основу графика можемо приметити да су вредности приноса поприлично нестабилне и да драстично опадају као последица короне. Из тог разлога смо се одлучили да σ^2 оценимо на основу претходна 24 дана, што је допринело не тако доброј оцени волатилности.



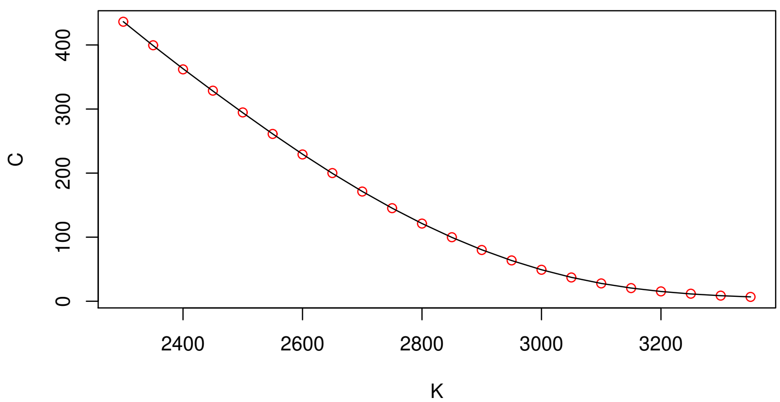
Слика 9: Оцена густине у Блек-Шолсовом моделу.

На слици изнад је приказана оцена густине у Блек-Шолсовом моделу, док су на слици испод приказане оцењене цене кол опција тим моделом и цене са тржишта. Видимо да Белк-Шолсов модел не описује најбоље кретање цена опција, што је последица оцењене волатилности.

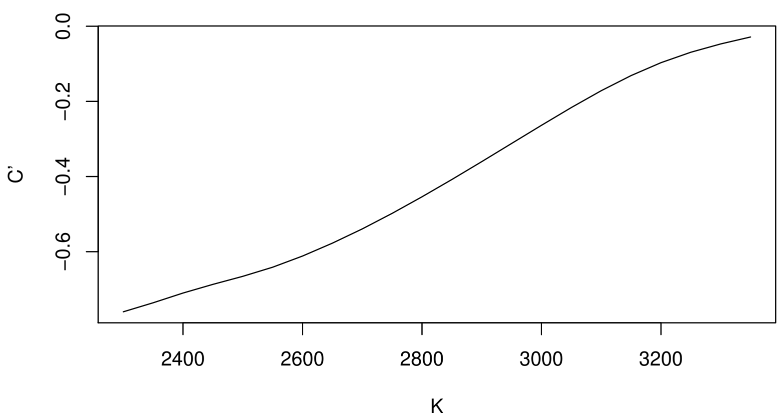


Слика 10: Оцена цена опције Блек-Шолсовим моделом (црном бојом) и цене са тржишта (црвеном бојом)

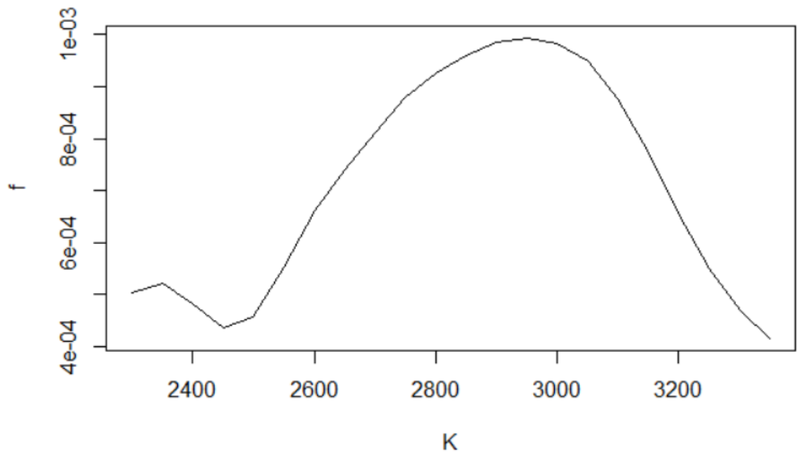
Сада ћемо анализирати реалне податке са тржишта под претпоставком да је цена кол опције просечна цена од понуђене (bid) и тражене (ask) цене. Оцене су добијене локалном квадратном регресијом, а вредност параметра размере износила је $h = 79.83$.



Слика 11: Оцена цене кол опције (пуна линија) под претпоставком да је цена кол опције (тачкице) узета као просечна цена од понуђене (bid) и тражене (ask) цене.



Слика 12: Оцена првог извода функције цене опције под претпоставком да је цена кол опције узета као просечна цена од понуђене (bid) и тражене (ask) цене.



Слика 13: Оцена функције неутралне од ризика под претпоставком да је цена кол опције узета као просечна цена од понуђене (*bid*) и тражене (*ask*) цене.

У даљем тексту испитаћемо квалитет локално полиномне регресије за оцењивање густине неутрале од ризика и то коришћењем $RIMSE$ дате изразом (4.16). У ту сврху користићемо Монте Карло метод који се заснива на понављању експеримента (вредновања опција) велики број пута ($N = 1000$), рачунањем интеграла

$$\left(\int_0^\infty E[(\hat{f}(x) - f(x))^2] dx \right)^{1/2},$$

а затим одређивањем аритметичке средине добијених израза. Како непараметарска регресија не дозвољава екстраполацију, односно, не пружа предвиђања за $E(Y|X)$ у тачкама које су изван опсега случајне променљиве X, из тог разлога ћемо у горе наведеном интегралу уместо граница 0 и ∞ одређивати интеграл на интервалу од 2300 до 3350.

За теоријску (стварну вредност) функције узећемо оцену густине параметарског модела коришћеног у раду [2], односно мешавину две логнормалне расподеле описану у наставку.
Нека је $\{(x_1, c_1), (x_2, c_2), ..., (x_n, c_n)\}$ скуп података са тржишта, где x_i представља уговорену цену и c_i одговарајућу цену опције, при чему важи да је $x_i \geq 0$ и $c_i \geq 0$ за $i = 1, \ldots, n$. Претпоставимо да је

$$f(x) = p_1 \cdot logn(x; \mu_1, \sigma_1) + p_2 \cdot logn(x; \mu_2, \sigma_2),$$

где је $logn(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$.

Циљ је оценити параметре p_i, μ_i, σ_i, $i = 1, 2$ методом најмањих квадрата при чему важе услови за одсуство арбитраже:

$$\min_{p_i,\mu_i,\sigma_i,i=1,2} \sum_{i=1}^{n} (c_i - C(x_i))^2$$

$$\begin{cases} f(K) \geq 0, K \in [0,\infty), \\ \int_0^\infty f(S)dS = 1, \\ C(0) = S_0 e^{-\delta t}. \end{cases}$$

Цена опције је дата са:

$$\begin{aligned} C(K) &= e^{-rt} \int\limits_K^{+\infty} (S-K) f(S) dS \\ &= e^{-rt} \int\limits_K^{+\infty} (S-K) \sum_{i=1}^{2} p_i logn(S;\mu_i,\sigma_i) dS \\ &= e^{-rt} \sum_{i=1}^{2} p_i \int\limits_K^{+\infty} (S-K) logn(S;\mu_i,\sigma_i) dS \end{aligned}$$

Ако узмемо да је $K = 0$:

$$\begin{aligned} C(0) &= e^{-rt} \int\limits_0^{+\infty} (S-0) f(S) dS \\ &= e^{-rt} \int\limits_0^{+\infty} S \sum_{i=1}^{2} p_i \; logn(S;\mu_i,\sigma_i) dS \\ &= e^{-rt} \sum_{i=1}^{2} p_i \int\limits_0^{+\infty} S \; logn(S;\mu_i,\sigma_i) dS \\ &= e^{-rt} \sum_{i=1}^{2} p_i \; e^{\mu_i + \frac{1}{2}\sigma_i^2} \end{aligned}$$

Тада важи следеће:

$$C(0) = S_0 e^{-\delta t} \iff e^{-rt} \sum_{i=1}^{2} p_i \; e^{\mu_i + \frac{1}{2}\sigma_i^2} = S_0 e^{-\delta t}$$

Такође

$$logn(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp - \frac{(\ln x - \mu)^2}{2\sigma^2} \geq 0,$$

$$\int_0^{\infty} \frac{1}{S\sigma\sqrt{2\pi}} \exp -\frac{(\ln S-\mu)^2}{2\sigma^2} dS = 1.$$

С обзиром да захтевамо да су параметри p_1 и p_2 ненегативни, имамо да важи:

$$f(K) > 0 \iff \sum_{i=1}^{2} p_i \; logn(K; \mu_i, \sigma_i) > 0 \iff p_1 \geq 0, \; p_2 \geq 0, \; K \in [0, \infty),$$

$$\int_0^{\infty} f(S) dS = 1 \iff \int_0^{\infty} \sum_{i=1}^{2} p_i \; logn(S; \mu_i, \sigma_i) dS = 1 \iff \sum_{i=1}^{2} p_i = 1.$$

За ненегативне ε_i, $i = 1, \ldots, n$ се наш проблем одређивања параметара своди на

$$\min_{p, \mu_i, \sigma_i, i=1,2 \;\; \varepsilon_j, \;\; j=1,\ldots,n} \sum_{i=1}^{n} \varepsilon_i^2$$

$$\begin{cases} -\varepsilon_j \leq c_j - e^{-rt} \sum\limits_{i=1}^{2} p_i \int\limits_{x_j}^{+\infty} (S - x_j) logn(S; \mu_i, \sigma_i) dS \leq \varepsilon_j, \; j = 1, \ldots, n \\ p_i \geq 0, \\ \sum\limits_{i=1}^{2} p_i = 1, \\ e^{-rt} \sum\limits_{i=1}^{2} p_i \; e^{\mu_i + \frac{1}{2}\sigma_i^2} = S_0 e^{-\delta t}. \end{cases}$$

Како би се постигла већа тачност оцене биће коришћена модификација овог проблема, предложена у раду [3], где се приликом оцењивања непознатих параметара користе и одговарајуће пут опције. Параметре оцењујемо на основу података доступних на сајту $Yahoo\ finance$ [1], за већ поменуто финансијско тржиште $S\&P$ 500.

Када знамо како изгледа права RND функција можемо генерисати велики број скупова опција и користити их при решавању проблема оптимизације. Први корак је одређивање теоријске цене кол опција на основу формуле:

$$C(K) = e^{-rt} \int_K^{\infty} (S - K) f(S) dS.$$

У реалности цене опција се изводе на основу теоријских цена (из разлога несавршености маркета: несинхронизованост трговања, распон куповне и продајне цене, и тако даље). Да бисмо опонашали стварни свет на теоријске цене додајемо шум. За добијање случајног шума R_i има смисла користити распон куповне и продајне цене и ликвидност опција за различите уговорене цене, као и претпоставку је да је униформно расподељен:

$$R_i \sim U(0, \frac{A_i L_i}{2})$$

где је U ознака за униформну расподелу, A_i бид-аск ширина при уговореној цени x_i а L_i ликвидност опције при x_i.

За распон цена (куповне и продајне) претпоставља се да је 5% од тражене (ask) цене, са минималном вредношћу од 50 центи, а максималном вредношћу 2 долара:

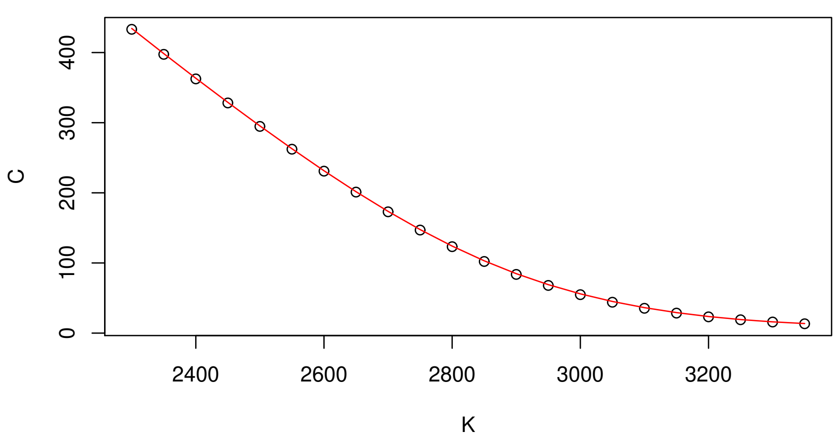
$$A_i = \begin{cases} 2, A_i > 2 \\ 5\% \times ask,\ 0.5 \leq A_i \leq 2 \\ 0.5,\ A_i < 0.5. \end{cases} \tag{5.1}$$

Фактор ликвидности L_i изражава се формулом ([2]):

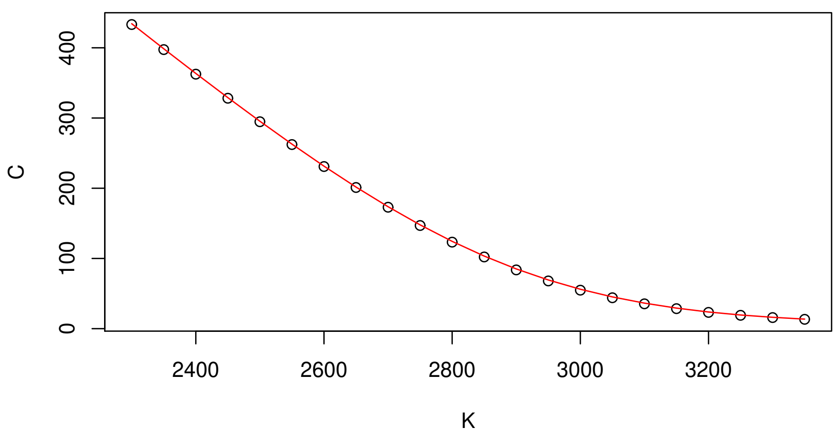
$$L_i = 1 + 10 \times |x_i/S_0 - 1|$$

где је x_i уговорена цена а S_0 тренутна цена акције.

Када на теоријске цене додамо случајан шум који смо генерисали, добијемо један скуп уговорених цена и њима одговарајућих цена опција. Уколико поновимо процес генерисања скупова случајних шумова 1000 пута добићемо 1000 скупова уговорених цена и њима одговарајућих цена опција. За сваки добијени скуп решавањем проблема оптимизације добијамо по једну оцену за праву RND фунцкију. На следећим графицима приказане су средње вредности оцена добијене локалном квадратном ($p = 2$) и кубном ($p = 3$) регресијом.

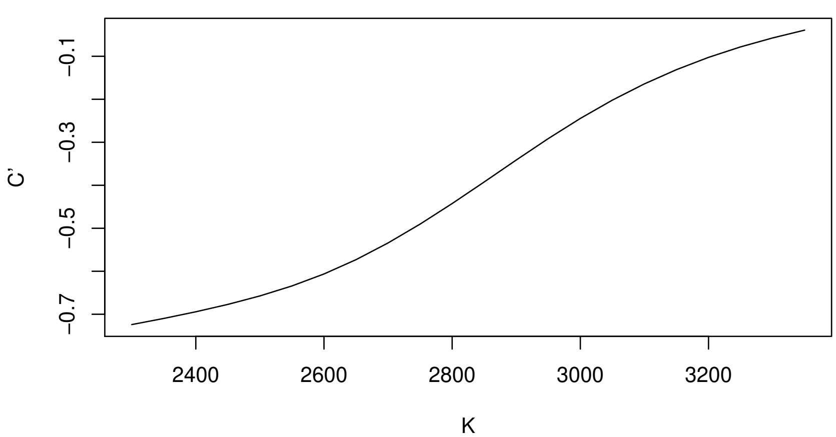


Слика 14: Средња вредност оцена цена опција добијених локалном квадратном регресијом (црвена линија).

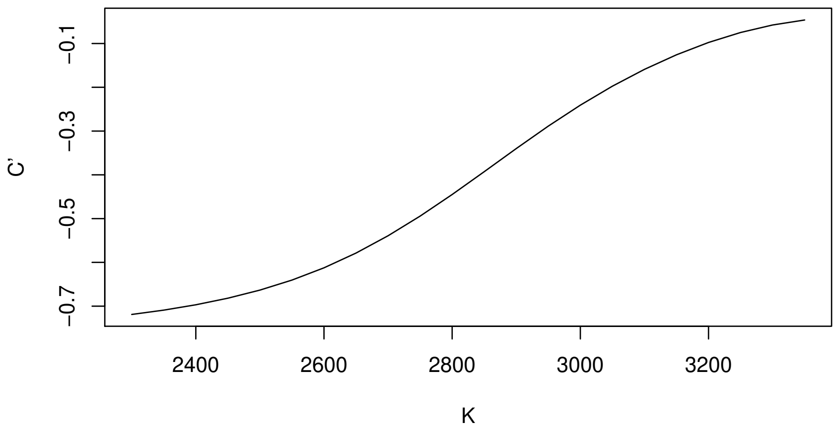


Слика 15: Средња вредност оцена цена опција добијених локалном кубном регресијом (црвена линија).

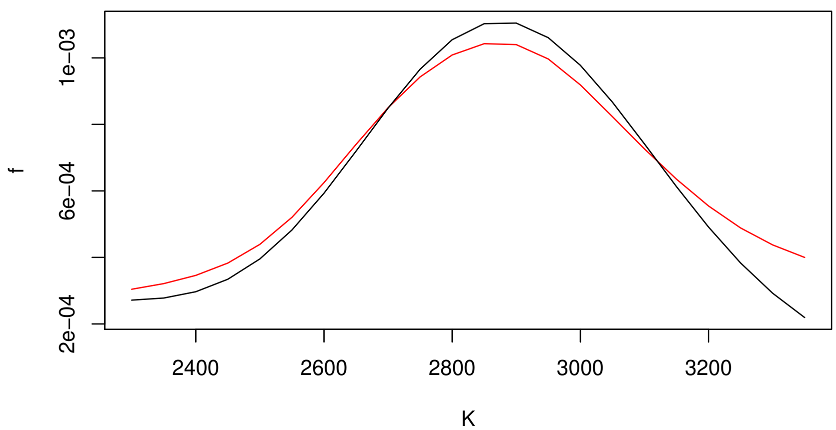
Као што видимо из претходног, оба метода су погодна за оцењивање цена опција.



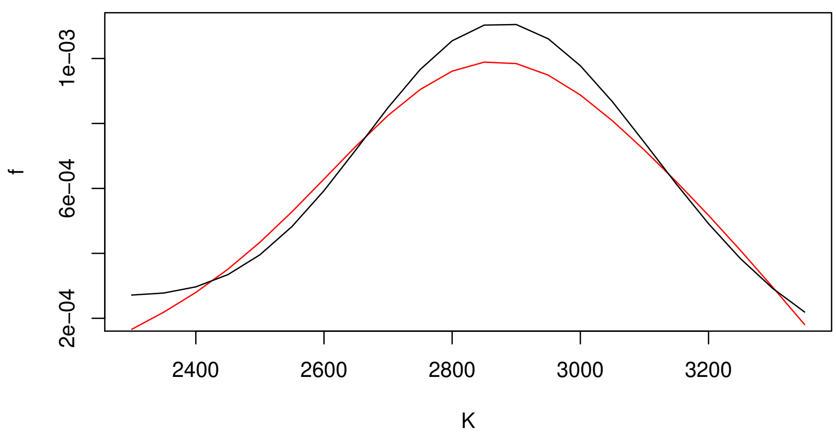
Слика 16: Средња вредност оцена првих извода цена опција добијених локалном квадратном регресијом.



Слика 17: Средња вредност оцена првих извода цена опција добијених локалном кубном регресијом.



Слика 18: Средња вредност оцена функција неутралних од ризика добијених локалном квадратном регресијом (црвена линија) и густина мешавине две логнормалне расподеле (црна линија).



Слика 19: Средња вредност оцена функција неутралних од ризика добијених локалном кубном регресијом (црвена линија) и густина мешавине две логнормалне расподеле (црна линија).

Циљ ове симулације је одредити квалитет оцене и упоредити који метод је бољи за оцењивање функције неутралне од ризика. На основу мере (4.16) и Монте Карло симулације добили смо да је мера квалитета за оцену добијену локално квадратном регресијом 0.00252, док је 0.00292 за оцену добијену локално кубном регресијом. Из приложеног можемо закључити да је оцена добијена локалном квадратном регресијом за нијансу боља од оцене добијене локалном кубном регресијом.

6 Кодови

Код 1: Слике 2 и 3

```
x=seq(-3, 3, 0.1)
plot(x,dnorm(x),"l",col="grey",lwd=2,ylab="f")
rx=rnorm(100)
lines(density(rx),col="#333399",lwd=1)
lines(density(rx,adjust=0.5),col="#009900",lwd=1)
lines(density(rx,adj=2),col="#FF6633",lwd=1)
exp1=as.expression(bquote(hat(f)[1]))
exp2=as.expression(bquote(hat(f)[2]))
exp3=as.expression(bquote(hat(f)[3]))
legend=c("f",exp1,exp2,exp3)
color=c("grey","#333399","#009900","#FF6633")
lwd=c(3,1,1,1)
legend("topleft",legend,bty="n",col=color,lwd=lwd)
```

Код 2: Слике 4,5 и 6

```
X = runif(500, min=-pi, max=3*pi)
Y = cos(X) + rnorm(500, sd=0.3)

NW1=ksmooth(x=X,y=Y,kernel="normal",bandwidth = 0.1)
NW2=ksmooth(x=X,y=Y,kernel="normal",bandwidth = 1)
NW3=ksmooth(x=X,y=Y,kernel="normal",bandwidth = 2.5)

plot(X,Y,pch=20,col= "grey")
lines(NW1, lwd=2, col="red")
lines(NW2, lwd=2, col="purple")
lines(NW3, lwd=2, col="limegreen")
color=c("red","purple","limegreen")
legend=c("h=0.1","h=1.0","h=2.5")
legend("topright", legend=legend , lwd=2, col=color)

fit = lm(Y~X)
plot(X,Y, pch=20,col="grey")
lines(X,fit$fitted.values, lwd=1, col="red")
```

Код 3: Слике 4, 5 и 6

```
D=data.frame(X=X)
D$Y=Y

h=lpbwselect(Y,X,kernel="gau",p=0,bwselect="imse-rot")
h1=lpbwselect(Y,X,kernel="gau",p=1,bwselect="imse-rot")
h2=lpbwselect(Y,X,kernel="gau",p=2,bwselect="imse-rot")
h3=lpbwselect(Y,X,kernel="gau",p=3,bwselect="imse-rot")

nw=ksmooth(x=X,y=Y,kernel="normal",bandwidth=h$bws[1,2])
llr=locpol(Y~X,data=D,bw=h1$bws[1,2],kernel=gaussK,deg=1)
lqr=locpol(Y~X,data=D,bw=h2$bws[1,2],kernel=gaussK,deg=2)
lcr=locpol(Y~X,data=D,bw=h3$bws[1,2],kernel=gaussK,deg=3)

plot(X,Y, pch=20,col="grey")
lines(nw$x,nw$y,lwd=3,col="limegreen")
lines(llr$lpFit$X,llr$lpFit$Y,type="l",col="red",lwd=2)
lines(lqr$lpFit$X,lqr$lpFit$Y,type = "l",col="purple",lwd=4)
lines(lcr$lpFit$X,lcr$lpFit$Y,type = "l",col="darkblue",lwd=2)
color=c("limegreen","red","purple","darkblue")
legend= c("p=0","p=1","p=2","p=3")
legend("topright",legend=legend, lwd=2, col=color)
```

Код 4: Оцена волатилности

```
d=getSymbols("^GSPC",from='2020-03-01',to="2020-04-06",
                     warnings=FALSE,auto.assign=FALSE)
Close=d[,"GSPC.Close"]
n=length(Close)
X=rep(0,n)
for(i in 1:(n-1)){
  j=i+1
  X[i]=log(Close$GSPC.Close[j]/as.numeric(Close$GSPC.Close[i]))
  }
X=na.omit(X)
sigma=sqrt(var(X)*252)
sigma=as.numeric(sigma)
```

Код 5: Слике 11, 12 и 13

```
df=exp(baza$`T`*baza$`r`)
xeval=seq(2300,3350,50)
x=baza$`K`
y=baza$`c`
h=lpbwselect(y,x,kernel="gau",p=2,bwselect="imse-rot",deriv=2)
locPol=locPolSmootherC(x,y,xeval= xeval,h$bws[1,2],deg=2,
                       kernel=gaussK,DET = FALSE)

plot(x,y,type="p", ylab="C", xlab="K")
lines(locPol$x,locPol$beta0,col="red")

plot(locPol$x,locPol$beta1, ylab="C'", xlab="K",type="l")

plot(xeval, bsm.density,col="red",ylab="f",xlab="K")
lines(locPol$x,locPol$beta2*df)
```

Код 6: Шум

```
L=function(K,S){
  l=1+10*abs(K/S-1)
  return(l)
}

A=function(ask){
  s=0.05*ask
  if(s<0.5)
    return(0.5)
  if(s>2)
    return(2)
  else
    return(s)
 }
#za odgovarajuce vrednosti K, S i ask imamo shum
shum=runif(1,0,L(K,S)*A(ask)/2)
```

Код 7: Слике 14, 16 и 18

```
beta20=rep(0,length(xeval))
beta21=rep(0,length(xeval))
beta22=rep(0,length(xeval))
broj_simulacija=1000
ise=rep(0,broj_simulacija)

for (i in 1:broj_simulacija){
  y=rep(0,length(mln.kol))

  #vraca nam pravu cenu opcije (mln.kol) plus shum
  y=yShum(mln.kol,baza$shum)

  h=lpbwselect(y,x,kernel = "gau",p=2, bwselect = "imse-rot",
                         deriv = 2)
  locPol=locPolSmootherC(x, y, xeval= xeval, h$bws[1,2],
                              deg=2, kernel=gaussK, DET = FALSE)
  h=as.numeric(h$bws[1,2])

  #"razlika" predstavlja razliku izmedju "prave" RND
  #i ocene za RND dobijene iz lokalno kvadratne regresije

  integral=integrate(razlika,2300,3350)
  ise[i]=integral$value

  beta20=beta20+locPol$beta0
  beta21=beta21+locPol$beta1
  beta22=beta22+locPol$beta2
  }
#srednje vrednosti ocena
sv20=beta20/broj_simulacija
sv21=beta21/broj_simulacija
sv22=beta22/broj_simulacija

rimse=sqrt(mean(ise))
```

Код 8: Слике 15, 17 и 19

```
beta30=rep(0,length(xeval))
beta31=rep(0,length(xeval))
beta32=rep(0,length(xeval))
broj_simulacija=1000
ise=rep(0,broj_simulacija)

for (i in 1:broj_simulacija){
  y=rep(0,length(mln.kol))

  #vraca nam pravu cenu opcije (mln.kol) plus shum
  y=yShum(mln.kol,baza$shum)

  h=lpbwselect(y,x,kernel = "gau",p=3, bwselect = "imse-rot",
                        deriv = 3)
  locPol=locPolSmootherC(x, y, xeval= xeval, h$bws[1,2],
                              deg=3, kernel=gaussK, DET = FALSE)
  h=as.numeric(h$bws[1,2])

  #"razlika" predstavlja razliku izmedju "prave" RND
  #i ocene za RND dobijene iz lokalno kubne regresije

  integral=integrate(razlika,2300,3350)
  ise[i]=integral$value

  beta30=beta30+locPol$beta0
  beta31=beta31+locPol$beta1
  beta32=beta32+locPol$beta2
  beta33=beta33+locPol$beta3
  }
#srednje vrednosti ocena
sv30=beta30/broj_simulacija
sv31=beta31/broj_simulacija
sv32=beta32/broj_simulacija
sv33=beta33/broj_simulacija

rimse=sqrt(mean(ise))
```

Литература

[1] Yahoo finance (https://finance.yahoo.com/).

[2] Y. Aıt-Sahalia and J. Duarte. Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116(1-2):9–47, 2003.

[3] B. Bahra. Probability distributions of future asset prices implied by option prices. *Bank of England Quarterly Bulletin*, 36(3):299–311, 1996.

[4] M. Birke and K. F. Pilz. Nonparametric option pricing with no-arbitrage constraints. *Journal of Financial Econometrics*, 7(2):53–76, 2009.

[5] J. Hull. *Options, futures and other derivatives.* Upper Saddle River, NJ: Prentice Hall, 2009.

[6] S. M. Iacus. *Option pricing and estimation of financial models with R.* Wiley Online Library, 2011.

[7] J. C. Jackwerth and M. Rubinstein. Recovering probability distributions from option prices. *The Journal of Finance*, 51(5):1611–1631, 1996.

[8] S. Janković and B. Milošević. *Elementi finansijske matematike.* Matematički fakultet, Beograd, 2017.

[9] S. J. Sheather. Density estimation. *Statistical science*, pages 588–597, 2004.

[10] L. Wasserman. *All of nonparametric statistics.* Springer Science & Business Media, 2006.

Биографија

Маријана Палибрк, рођена 21.09.1992. године у Чачку. Основну школу "Свети Сава" завршила је 2007. године и уписала Гимназију у Чачку, природно-математички смер. Године 2011. уписала Математички факултет, смер Статистика, актуарска и финансијска математика. Након завршених основних академских студија 2016. године уписује мастер академске студије на истом факултету и истом смеру. Претходне три године била је запослена у Републичком заводу за статистику, сектор за Националне рачуне, цене и пољопривреду.