

**რეგრესიული მიდგომის გამოყენების ზოგიერთი თავისებურება სადაზღვევო პროდუქტის სამართლიანი
ტარიფიკაციის ამოცანაში**

ერეკლე ხურობე

სადაზღვევო პროდუქტის ტარიფიკაციის ამოცანაში რეგრესიული ანალიზის მეთოდების პრაქტიკაში დამკვიდრება რამოდენიმე ათწლეულს ითვლის (რაც რეგრესიული მიდგომის კლასიკურობის გათვალისწინებით ბევრი არ არის). წინამდებარე ნაშრომში განხილულია რეგრესიული მეთოდების გამოყენების ზოგიერთი თავისებურება, დაკავშირებული სადაზღვევო სპეციფიკასთან, რამაც შესაძლებელია დახმარება გაუწიოს პრაქტიკოს აქტუარებს. განსაკუთრებული ყურადღება გამახვილებულია ნორმირებულ შეფასებაზე, რომელიც, იდეის სიმარტივის მიუხედავად, ვფიქრობ არ არის საკმარისად განხილული აქტუარულ ლიტერატურაში, მაშინ, როდესაც მისი პრაქტიკული ღირებულება საკმაოდ დიდია, რაზეც ქვევით ვისაუბრებ.

სიცხადისთვის განვიხილოთ მხოლოდ ქონებრივი დაზღვევა და თავდაპირველად შევთანხმდეთ ზოგიერთ განსაზღვრებასა თუ აღნიშვნაზე: S_i -ით აღვნიშნოთ (პირობითად i -ური პოლისის შესაბამის) სადაზღვევო თანხა (როგორც წესი, დაზღვეული ობიექტის ღირებულება), ხოლო R_i -ით სადაზღვევო ტარიფი, ანუ სადაზღვევო ლიმიტის ერთი ერთეულის დაზღვევის ღირებულება. შესაბამისად, $P_i = R_i \cdot S_i$ არის სადაზღვევო პრემია, რომელიც თავის მხრივ შედგება ე.წ. რისკ-პრემიისგან და დატვირთვისგან: $P_i = P_i^{(r)} + P_i^{(l)} = R_i^{(r)} S + R_i^{(l)} S$. გარდა ამისა, L_{ij} არის შემთხვევითი სიდიდე, დამოკიდებული S_i -ზე, რომლისთვისაც $\text{supp}(L_{ij}) = (0, S]$ და რომელიც გამოხატავს i -ური პოლისზე j -ური ($j \geq 1$) სადაზღვევო შემთხვევის შედეგად დამდგარი ზარალის სიდიდეს, რომლისთვისაც:

$$\text{supp}(L_{ij}) = \left(0, S_i - \sum_{k=0}^{j-1} L_{ik} \right],$$

სადაც, $L_{i0} \equiv 0$ თითქმის აუცილებლად.

პრაქტიკული მიზნებიდან, L_{ij} -ს ნაცვლად, როგორც წესი განიხილავენ სიდიდეს

$$\bar{L}_{ij} = \min(L_{ij}, l_0(S_i)) + \mathbb{1}_{[l_0(S_i), S_i]}(L_{ij}) \cdot \left(S_i - L_{ij} - \sum_{k=0}^{j-1} \bar{L}_{ik} \right)$$

სადაც, l_0 პრაქტიკული მოსაზრებებიდან სადაზღვევო კომპანიის მიერ შერჩეული რიცხვია, როგორც წესი დამოკიდებული S_i -ზე (ტიპურად, მისი რაიმე ნაწილი) - ეს ის თანხაა, რომლის ტოლი ან მეტი ზარალის შემთხვევაში, სადაზღვევო კომპანიას ურჩევნია დამზღვევს აუნაზღაუროს დაზღვეული ობიექტის სრული ღირებულება დაზიანებული დაზღვეული ობიექტის (salvage-ის) სანაცვლოს. ამრიგად,

$$\mathbb{P}\left(\bar{L}_{ij} = S_i - \sum_{k=0}^{j-1} \bar{L}_{ik}\right) = \mathbb{P}(L_{ij} \in [l_0, S_i]) > 0$$

სახელდობრ,

$$\mathbb{P}(\bar{L}_{i0} = S_i) > 0.$$

$\{\bar{L}_{ij} = S_i\}$ ხდომილობას ვუწოდოთ სრული ზარალი(ს შემთხვევა), ხოლო $\{\bar{L}_{ij} < S_i\} = \{\bar{L}_{ij} < l_0\}$ -ს ნაწილობრივი ზარალი(ს შემთხვევა). სადაზღვევო პერიოდში (ტიპურად 1 წელიწადი) i -ურ პოლისზე მომხდარი ზარალების რაოდენობა აღვნიშნოთ N_i ; ამრიგად, $\text{supp}(N) = \mathbb{N}_0 \equiv \{0\} \cup \mathbb{N}$. ხშირად ეს უკანასკნელი დამოკიდებულია S_i -ზე. ამრიგად, სადაზღვევო პერიოდის მანძილზე i -ურ პოლისზე

მომხდარი ზარალების მთლიანი ოდენობა აღიწერება შემდეგი (შედგენილი) შემთხვევითი სიდიდის საშუალებით:

$$C_i = \sum_{j=1}^{N_i} \bar{L}_{ij}$$

სადაც, $N_i, \bar{L}_{i1}, \bar{L}_{i2}, \dots$ ერთობლივად დამოუკიდებელია, ხოლო $\bar{L}_{i1}, \bar{L}_{i2}, \dots$ არის i.i.d. ტიპური სადაზღვევო ხელშეკრულებების დროს, აღნიშნული დაშვება არ არის ბოლომდე კორექტული, სულ მცირე იმიტომ, რომ, როგორც ზემოთ უკვე აღვნიშნეთ, $\text{supp}(L_{ij}) = (0, S_i - \sum_{k=0}^{j-1} L_{ik}]$. ამგვარი უხერხულობების თავიდან ასახილებლად, ზოგადობის შეუზღუდავად დავუშვათ შემდეგი ორი პირობა:

1. (მეტ-ნაკლებად პრაქტიკული) ნაწილობრივი ზარალის დროს, ლიმიტი (ხელშეკრულების მიხედვით) აღდგენადია (ანუ ლიმიტი ზღუდავს მხოლოდ ერთეულ და არა ჯამურ ზარალს) და ასეთი ზარალის (ქონების დაზიანების) შეკეთების შემდეგ, ქონება არ კარგავს თავდაპირველ ღირებულებას, ანუ

$$\forall j \geq 1: \text{supp}(\bar{L}_{ij} \mid \forall k < j: L_k \neq S_i) = (0, S_i]$$

2. (ტექნიკური)

$$\mathbb{P}(N_i = n > 1 \& \max(\bar{L}_{i1}, \dots, \bar{L}_{in}) = S_i) = 0$$

ამრიგად, ბოლო დაშვების შესაბამისად, N_i შეგვიძლია განვიხილოთ, როგორც შემთხვევითი ელემენტი, რომლისთვისაც:

$$\text{supp}(N_i) = \{0, 1_T, 1_P \equiv 1, 2, 3, \dots\}$$

სადაც, 1_T ნიშნავს სრულ ზარალს, ხოლო 1_P - ერთ ნაწილობრივ ზარალს. აღვნიშნოთ, $P_{Ti} \equiv \mathbb{P}(N_i = 1_T)$.

სამართლიანი ტარიფიკაციის პირობა შესაძლებელია ფორმულირდეს შემდეგნაირად: რისკ-პრემია ტოლია საშუალო მოსალოდნელი ზარალის, ანუ:

$$P_i^{(r)} = E(C_i).$$

საზოგადოდ, C_i დამოკიდებულია მთელ რიგ ფაქტორებზე, რომელთაგან ზოგიერთი დაკვირვებადია, ზოგიერთი კი არა (მაგალითად, ავტოდაზღვევის შემთხვევაში მძღოლის ტემპერამენტი, ხასიათი და ა.შ., რომელთა შეფასების სხვადასხვა გზები არსებობს, თუმცა ეს საკითხი სცდება აღნიშნული სტატიის მიზნებს). სადაზღვევო კომპანია ცდილობს გამოყოფს დაკვირვებადი ფაქტორები $(A_1, \dots, A_m)$, რომელზეც დამოკიდებულია პოტენციური ზარალიანობა და სამართლიან რისკ-პრემიას შეხედოს როგორც რეგრესიულ ფუნქციას:

$$P_i^{(r)}(a_1, \dots, a_m) = E(C_i \mid A_{i1} = a_1, \dots, A_{im} = a_m).$$

პრაქტიკაში ერთ-ერთი მნიშვნელოვანი რისკ-ფაქტორი სწორედ რომ სადაზღვევო თანხაა; ამ უკანასკნელზე შევჩერდეთ ჩვენც; ამრიგად განვიხილოთ „სამართლიანი“ (რისკ-)პრემიის ფუნქცია:

$$P_i^{(r)}(s) = E(C_i \mid S_i = s) = E\left(\sum_{j=1}^{N_i} \bar{L}_{ij} \mid S_i = s\right)$$

პირობითი მათემატიკური ლოდინისთვის ვალდის იგივეობის ძალით:

$$P_i^{(r)}(s) = E(N_i | S_i = s) \cdot E(\bar{L}_{i1} | S_i = s).$$

ამრიგად, აქტუარის ამოცანაა $f(s) = E(N_i | S_i = s)$ და $g(s) = E(\bar{L}_{i1} | S_i = s)$ რეგრესიული წირების შეფასება.

მათ შესაფასებლად (ვუწოდოთ მათ შესაბამისად სიხშირის მოდელი და სიმძიმის მოდელი) შესაძლებელია გამოყენებულ იქნეს ზოგადი წრფივი მოდელები (GLM – generalized linear models (მაგალითად, შესაბამისად, პუასონის რეგრესია და გამა რეგრესია)); არაპარამეტრული რეგრესიული მოდელები, როგორიცაა სპლაინ-რეგრესიები, განზოგადებული ადიციური მოდელები (GAM) ან ალბათური სიმკვრივის გულოვან შეფასებებზე დაფუძნებული მოდელები, როგორებიცაა ნადარაია-უოტსონის, პრისლი-ჩაოს ან გასერ-მიულერის გულოვანი შეფასებები; უბან-უბან მუდმივი ფუნქციით შეფასებებისთვის იზოტონიკური რეგრესიები, გადაწყვეტილების ხეები და ა.შ.

პირველ რიგში შევნიშნოთ, რომ ხშირ შემთხვევაში (განსაკუთრებით, მაშინ, როდესაც სრულ ზარალებთან მიმართებაში კომპანიის პოლიტიკა არ არის კონსერვატიული; მაგალითად, როდესაც $l_0(S_i) = rS_i$ და $0 < r \ll 1$ (მაგ. $r \in [0.3, 0.5]$)) სიმძიმის მოდელის აგებისას სრული ზარალების გამოყენება დაკავშირებულია რამოდენიმე პრობლემასთან. გამოვყოთ ორი მათგანი:

1. ხშირ შემთხვევაში განსხვავებით ნაწილობრივი ზარალისგან, სრული ზარალის მოხდენა-არმოხდენა არ არის დამოკიდებული სადაზღვევო თანხაზე, ანუ

$$\mathbb{P}(N_i = 1_T | S_i = s) = \mathbb{P}(N_i = 1_T)$$

ასეთი ზარალების $g(s)$ -ის შეფასებაში გათვალისწინება ხშირად იწვევს ე.წ. გადამეტებულ მორგებას (overfitting).

შენიშვნა: გადამეტებული მორგების საფრთხე მაშინაც არსებობს, როდესაც ჩვენს მიერ მოყვანილ დაშვებას არ აქვს ადგილი, სახელდობრ მაშინ, როდესაც სრული ზარალი ხდება იშვიათად და მისი შესაბამისი ზარალის მოსალოდნელი ოდენობა მნიშვნელოვნად აღემატება ნაწილობრივი ზარალის შესაბამის მოსალოდნელ ოდენობას.

2. ზოგადი წრფივი მოდელების (რაც დღესდღეობით პრაქტიკაში ყველაზე ხშირია) გამოყენების შემთხვევაში, ირღვევა პირობა $\bar{L}_{ij}$ -ის პირობითი განაწილების შესახებ, ანუ დაშვება იმის შესახებ, რომ იგი დისპერსიული პარამეტრის შემცველ ექსპონენციალურ განაწილებათა (ED (exponential family of distributions)) ოჯახიდანაა, რაც ნიშნავს, რომ

$$f_{\bar{L}_{ij} | S_i=s}(x) = \exp\left(\frac{x\theta(s) - \alpha(\theta(s))}{\phi}\right) c(x, \phi)$$

სხვანაირად, რომ ვთქვათ

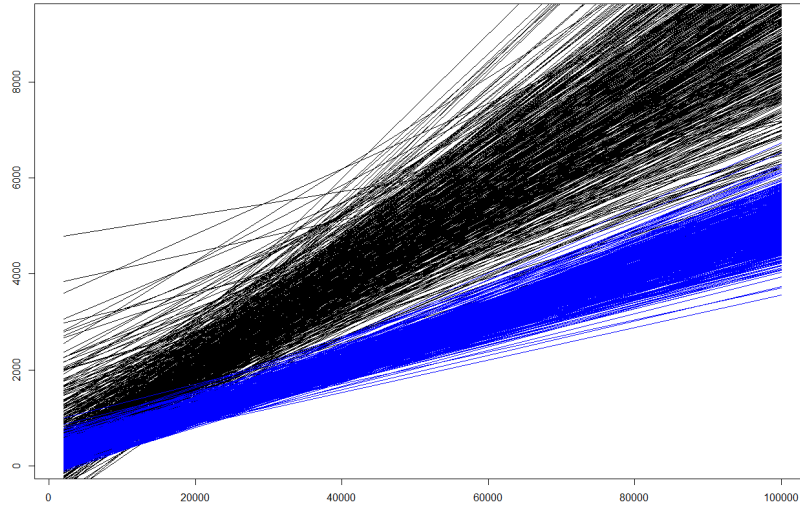
$$\bar{L}_{ij} | S_i = s \sim ED(\mu(s), \phi, \hat{\alpha}, \hat{c})$$

სადაც, $ED(\mu(s), \phi, \hat{\alpha}, \hat{c})$ აღნიშნავს, ED-ოჯახის (ფიქსირებული $\hat{\alpha}$ და $\hat{c}$ ფუნქციებისთვის) ე.წ. „საშუალო-დისპერსიული პარამეტრი“ პარამეტრიზაციას (იხ. [1], [2]).

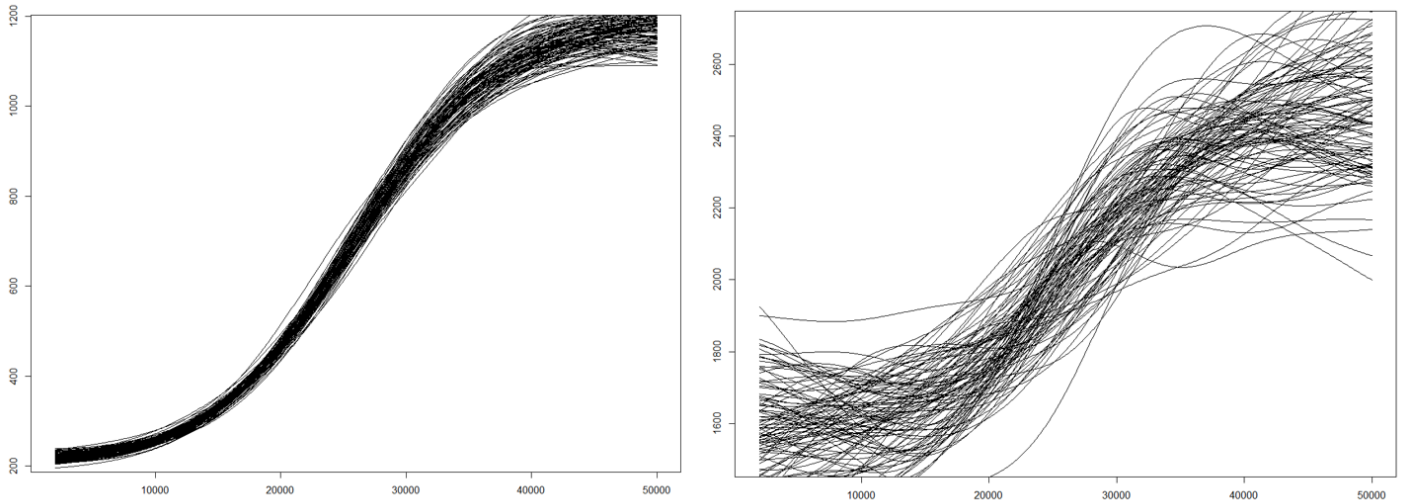
მეორეს მხრივ არაპარამეტრული რეგრესიული მოდელები ძალიან მგრძნობიარენი არიან „უხეში“ ნარეგების მიმართ.

ქვემოთ მოცემული მაგალითები (სიმულაციის შედეგები) გვამღევეს აღნიშნულის თვალსაჩინო ვიზუალურ ილუსტრაციებს (სტატისტიკის მიზნიდან გამომდინარე შემოვიფარგლე განსხვავების ვიზუალური წარმოჩენის).

მაგალითი 1. დავუშვათ $\bar{L}_{ij} | S_i = s \text{ \& } N_i \neq 1_T \sim \text{Gamma}(\mu(s) = 0.1s + 50, \phi = 0.7)$ და $\mathbb{P}(N = 1_T | S_i = s) = 0.1$. გრაფიკზე მოცემულია სიმულირებულ მონაცემებზე (შერჩევის რაოდენობა - 100) აგებული 1000 (GLM-) შეფასების გრაფიკი (შავი - სრულ ზარალებთან ერთად, ლურჯი - სრული ზარალების გარეშე).



მაგალითი 2: დავუშვათ $\bar{L}_{ij} | S_i = s \text{ \& } N_i \neq 1_T \sim \text{Gamma}\left(\mu(s) = 200 + \frac{1000}{1+e^{-0.000208s+5.408}}, \phi = 0.7\right)$ და $\mathbb{P}(N = 1_T | S_i = s) = 0.05$. ქვემოთ მოცემულია სიმულირებულ მონაცემებზე (შერჩევის რაოდენობა - 100) აგებული ნადარაია-უოტსონის 1000 შეფასება, მარცხნივ სრული ზარალების გარეშე, ხოლო მარჯვნივ სრულ ზარალებთან ერთად.



ამრიგად, უმჯობესია, თუ ჩავწერთ:

$$C_i | S_i = s \cong \sum_{j=1}^{N_{Pi}} \hat{L}_{ij} + D_T s$$

სადაც, $N_{Pi} \cong N_i | N_i = 1_T$ აღწერს ნაწილობრივი ზარალების რაოდენობას, $\hat{L}_{ij}$ - ამ შემთხვევაში უკვე ნაწილობრივი ზარალის სიდიდეს, ხოლო $D_T \sim \text{Bernoulli}(p = P_{Ti})$. ამრიგად, თუ დავუშვებთ, რომ სრული ზარალის მოხდენა-არმოხდენის საკითხი სადაზღვევო თანხისგან დამოუკიდებელია, გვექნება:

$$E(C_i | S_i = s) = P_i^{(r)}(s) = E(N_{Pi} | S_i = s) \cdot E(\hat{L}_{i1} | S_i = s) + P_{Ti} s \equiv P_i^{(r,p)}(s) + P_i^{(r,t)}(s)$$

ხოლო:

$$R_i^{(r)}(s) = \frac{E(N_{Pi} | S_i = s) \cdot E(\hat{L}_{i1} | S_i = s)}{s} + P_{Ti} \equiv R_i^{(r,p)}(s) + R_i^{(r,t)}.$$

სადაც ტოლობის მარჯვენა მხარის პირველი შესაკრებები გვიჩვენებს ნაწილობრივი ზარალების შესაბამის, ხოლო მეორე შესაკრები სრული ზარალების შესაბამის რისკ-პრემიასა (რომელიც s -ის წრფივი ფუნქციაა) და რისკ-ტარიფს (რომელიც არ არის დამოკიდებული s -ზე). $P_i^{(r,p)}(s)$ -ის სტატისტიკური შეფასება აღვნიშნოთ $\hat{P}_i^{(r,p)}(s)$ -ით. იგი უნდა წარმოვიდგინოთ, როგორც ნამრავლი: $\hat{P}_i^{(r,p)}(s) = \phi(s) \cdot \psi(s)$. დავუშვათ, რომ შეფასებები აიგება შემდეგი ორი ამოკრეფის საფუძველზე (ცხადია ამოკრეფები არ არიან დამოუკიდებლები): $((S_i, N_i))_{i=1}^m$ და $((S_{j(l)}, \hat{L}_l))_{l=1}^{\sum_{i=1}^m N_i \equiv K}$ (შემთხვევას, როდესაც ამოკრეფა მოიცავს ინფორმაციას სხვადასხვა სიგრძის პერიოდის განმავლობაში მოქმედი პოლისების შესახებ, აქ არ ვიხილავთ).

ბუნებრივია, რომ შეფასების მიმართ გვქონდეს შემდეგი მოთხოვნა:

$$\sum_{i=1}^m \hat{P}_i^{(r,p)}(s_i) = \sum_{i=1}^m \phi(s_i) \cdot \psi(s_i) = \sum_{l=1}^K \hat{L}_l$$

რაც ტოლფასია შემდეგის:

$$\frac{\sum_{i=1}^m \hat{P}_i^{(r,p)}(s_i)}{\sum_{i=1}^m s_i} = \frac{\sum_{i=1}^m s_i \hat{R}_i^{(r,p)}(s_i)}{\sum_{i=1}^m s_i} = \left\langle \hat{R}_i^{(r,p)}(s_i) \right\rangle_{\{s_i\}} = \frac{\sum_{l=1}^K \hat{L}_l}{\sum_{i=1}^m s_i} = BE$$

სადაც, BE აღნიშნავს ე.წ. წაუგებლობის პრემიას (Break-Even Rate), ხოლო $\langle a_i \rangle_{\{b_i\}}$ კი a_i -ების შეწონილ საშუალოს b_i -ების მიმართ.

ეს მოთხოვნა ნიშნავს, იმას, რომ რეგრესიულმა შეფასებებმა მხოლოდ რისკის შესაბამისად უნდა „გადაანაწილოს“ პრემია, საშუალო ტარიფი ან პორტფელის უცვლელობის პირობაში ჯამური პრემია კი, უნდა იყოს იგივე, რაც მუდმივი (s -გან დამოუკიდებელი) შეფასების შემთხვევაში იქნებოდა.

პრაქტიკოსებს შორის გავრცელებული არასწორი წარმოდგენა იმის შესახებ, რომ GLM მოდელები ამას თავისით უზრუნველყოფენ, ეფუძნება შემდეგ ცნობილ შედეგს:

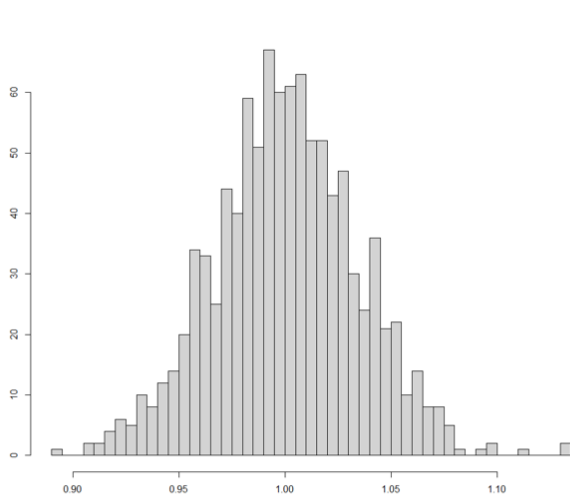
წინადადება: კანონიკური მაკავშირებელი ფუნქციის (link function) შემთხვევაში (მაკავშირებელ h ფუნქციას უწოდებენ კანონიკურს, თუ $h^{-1} = \alpha$ (მაგალითად, გამა რეგრესიითვის კანონიკური მაკავშირებელი ფუნქციაა შებრუნებული ფუნქცია, პუასონისთვის - ლოგარითმული ფუნქცია, გაუსურისთვის - იგივე და ა.შ.)) სამართლიანია:

$$\sum_i y_i = \sum_i \hat{\mu}_i$$

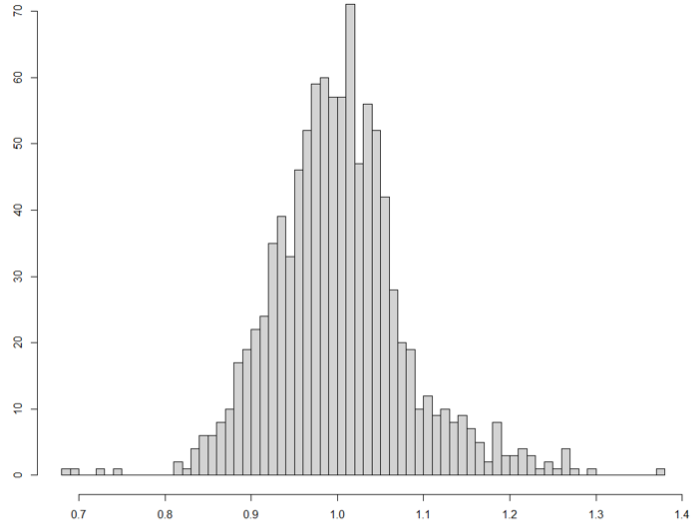
სადაც, $\hat{\mu}_i$ - რეგრესიით შეფასებული მნიშვნელობებია, y_i -ები კი ასახსნელი ცვლადის მნიშვნელობები ამოკრეფაში.

წინადადება სამართლიანია მხოლოდ კანონიკური მაკავშირებლის შემთხვევაში; სხვა შემთხვევაში მაგალითად $\frac{\sum_{l=1}^m \psi(s_l)}{\sum_{l=1}^K \hat{L}_l}$ შეფარდება შესაძლებელია მნიშვნელოვნად დამორდეს 1-ს. ქვემოთ მოცემულია შესაბამისი მაგალითი (1000 სიმულაციის შედეგად მიღებული შეფარდებების შესაბამისი ჰისტოგრამები და დეცილთა ცხრილი):

მაგალითი 3: მარჯვენა შემთხვევაში $\hat{L}_l | S_l = s \sim \text{Gamma}(\mu(s) = 0.1s + 50, \phi = 0.7)$, ხოლო მეორე შემთხვევაში $\phi = 3$.



10%	20%	30%	40%	50%	60%	70%	80%	90%
0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.05



10%	20%	30%	40%	50%	60%	70%	80%	90%
0.91	0.94	0.97	0.98	1.00	1.02	1.03	1.05	1.09

შევნიშნოთ, რომ ჩვენს კონკრეტულ ამოცანაში, ვინაიდან ჩვენი შეფასება წარმოადგენს ორი რეგრესიული მოდელის ნამრავლს, $\left\langle \hat{R}_i^{(r,p)}(s_i) \right\rangle_{\{s_i\}} = BE$ პირობა იმ შემთხვევაშიც არ სრულდება, როცა ორივე რეგრესიაში ვიყენებთ მის შესაბამის კანონიკურ მაკავშირებელ ფუნქციას, თუნდაც მხოლოდ იმიტომ, რომ $\psi(s)$ ფუნქციის შესაფასებლად გამოყენებული ამოკრეფის პირველი კოორდინატებისგან შემდგარი ვექტორი არ არის $(S_1, \dots, S_m)$ (ანუ იგივე, რაც $\phi(s)$ -ის შესაფასებლად გამოყენებული ამოკრეფის შემთხვევაში).

განვიხილოთ რისკ-პრემიის და რისკ-ტარიფის ე.წ. ნორმირებული შეფასებები:

$$\hat{P}_i^{(r,p)}(s) = \frac{\sum_{l=1}^K \hat{L}_l}{\sum_{i=1}^m \phi(s_i) \cdot \psi(s_i)} \phi(s) \cdot \psi(s) \equiv J\phi(s) \cdot \psi(s) \equiv J\hat{P}_i^{(r,p)}(s)$$

და

$$\hat{R}_i^{(r,p)}(s) = \frac{J\phi(s) \cdot \psi(s)}{s} = \frac{\sum_{l=1}^K \hat{L}_l / \sum_{i=1}^m s_i}{\sum_{i=1}^m \phi(s_i) \cdot \psi(s_i) / \sum_{i=1}^m s_i} \frac{\phi(s) \cdot \psi(s)}{s} \equiv \frac{BE \phi(s) \cdot \psi(s)}{\bar{b} s}$$

$\hat{R}_i^{(r,p)}(s)$ -ს შემდეგი მოსაზრებით შეგვიძლია ვუწოდოთ წაუგებლობის (brake-even) ფუნქცია:

$$\begin{aligned}
BE &= \frac{\sum_{l=1}^K \hat{L}_l}{\sum_{i=1}^m s_i} = \frac{\sum_{l=1}^K \hat{L}_l}{\sum_{i=1}^m \phi(s_i) \cdot \psi(s_i)} \cdot \frac{\sum_{i=1}^m \phi(s_i) \cdot \psi(s_i)}{\sum_{i=1}^m s_i} = \frac{\sum_{i=1}^m \frac{\phi(s_i) \cdot \psi(s_i)}{s_i} s_i}{\sum_{i=1}^m s_i} = \frac{\sum_{i=1}^m \hat{R}_i^{(r,p)}(s_i) s_i}{\sum_{i=1}^m s_i} = \\
&= \frac{\int_0^\infty \hat{R}_i^{(r,p)}(s) s d\hat{F}_{S,n}(s)}{\int_0^\infty s d\hat{F}_{S,n}(s)} \approx \frac{\int_0^\infty \hat{R}_i^{(r,p)}(s) s dF_S(s)}{\int_0^\infty s dF_S(s)} = \frac{\int_0^\infty \hat{R}_i^{(r,p)}(s) s f_S(s) ds}{\int_0^\infty s f_S(s) ds}
\end{aligned}$$

სადაც, $\hat{F}_{S,n}(s)$ არის $(S_1, \dots, S_m)$ ამოკრეფის შესაბამისი ემპირიული განაწილების ფუნქცია, ხოლო F_S - S შემთხვევითი სიდიდის თეორიული განაწილების ფუნქცია.

მოყვანილი ნორმირებული შეფასებები რამოდენიმე პრაქტიკული მოსაზრებით არის მნიშვნელოვანი:

1. GLM-ის გამოყენებისას მაკავშირებელი ფუნქციის არჩევაში თავისუფალი არჩევანის გაჩენა.
2. არაპარამეტრული რეგრესიული მეთოდების (რომლებსაც ზემოთ მოყვანილი სასურველი თვისება თავისთავად არ გააჩნიათ) გამოყენების საშუალება.
3. ტარიფების მართვის მოქნილი პრაქტიკული ინსტრუმენტი (შედეგად, ბიზნეს გარემოს ცვლილებებზე მყისიერი (სწრაფი) რეაგირების საშუალება).
4. ტარიფის შეფასებები სხვადასხვა სანდოობით.

მეოთხე პუნქტთან დაკავშირებული მოკლე შენიშვნა: რეგრესიებისთვის, როგორც წესი, განიხილავენ ორი ტიპის სანდოობის ინტერვალებს: 1. წერტილოვანს, ანუ სანდოობის ინტერვალებს $E(Y | X = x_0)$ ტიპის თეორიული სიდიდეებისთვის (პარამეტრულ და არაპარამეტრულ რეგრესიებში) და 2. სანდოობის ინტერვალებს რეგრესიული კოეფიციენტებისთვის (ტიპურად, პარამეტრულ შემთხვევებში). აქტუარისთვის კი მნიშვნელოვანია სანდოობის შეფასება/განსაზღვრა მთლიანი ტარიფიკაციისათვის. ამისთვის ბუნებრივია, თუ განვსაზღვრავთ ფუნქციის რისკ-ნორმას, ანუ შემოვიტანთ რისკ-დალაგების მიმართებას. რისკ-ნორმის ერთ-ერთ ყველაზე ბუნებრივ ფორმად შეგვიძლია განვიხილოთ შემდეგი:

$$\|R\|_{risk} = \int_S |R(s)| s f_S(s) ds$$

ამრიგად, ჩვენს შემთხვევაში გვექნება:

$$\begin{aligned}
\|\hat{R}_i^{(r,p)}\|_{risk} &= \int_S \frac{\sum_{l=1}^K \hat{L}_l}{\sum_{i=1}^m \phi(s_i) \cdot \psi(s_i)} \frac{\phi(s) \cdot \psi(s)}{s} s f_S(s) ds = \\
&= \int_S \frac{\sum_{l=1}^K \hat{L}_l}{\int_S \hat{P}_i^{(r,p)}(s) d\hat{F}_{S,n}(s)} \frac{\phi(s) \cdot \psi(s)}{s} s f_S(s) ds = \int_S \frac{\sum_{l=1}^K \hat{L}_l}{\int_S \hat{P}_i^{(r,p)}(s) d\hat{F}_{S,n}(s)} \hat{P}_i^{(r,p)}(s) F_S(s) \approx \sum_{l=1}^K \hat{L}_l = m
\end{aligned}$$

ხოლო, m -ის შეფასება სხვადასხვა სანდოობით უკვე სტანდარტული სტატისტიკური ამოცანაა, რომლისთვისაც შეგვიძლია გამოვიყენოთ, მაგალითად, ბუტსტრაპის მიდგომები.

გამოყენებული ლიტერატურა:

1. M. Denuit, D. Hainaut, J. Trufin – Effective Statistical Learning Methods for Actuaries I (GLM and Extensions)
2. P. Dunn, G. Smyth – Generalized Linear Models with Examples in R
3. J. Faraway – Extending the Linear Model with R