Finanční Matematika – 4. přednáška

Martin Panák

15. března 2016

Odhady parametrů negativního binomického rozložení

Příklad *Předpokládejme, že máme n řidičů, každý má smlouvu na dobu d_i, za kterou nahlásil k_i škodních událostí. Zkusme modelovat počet nehod daného řidiče pomocí negativního binomického rozložení, jehož parametry a a $\lambda (= \lambda d_i$ pro i-tého řidiče) odvodíme.*

Logaritmická věrohodnostní funkce je

$$L(a,\lambda) = \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \ln(a+j) + na \ln a$$
$$- \sum_{i=1}^{n} (a + k_i) \ln(a + \lambda d_i) + \ln \lambda \sum_{i=1}^{n} k_i + konst.$$

Pro parametry a a λ pak dostáváme rovnice.

$$\frac{\partial}{\partial a} L(a, \lambda) = \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \frac{1}{a + j} + n \ln a + n - \sum_{i=1}^{n} \ln(a + \lambda d_i) -$$

$$- \sum_{i=1}^{n} \frac{a + k_i}{a + \lambda d_i} = 0$$

$$\frac{\partial}{\partial \lambda} L(a, \lambda) = - \sum_{i=1}^{n} d_i \frac{a + k_i}{a + \lambda d_i} + \frac{1}{\lambda} \sum_{i=1}^{n} k_i = 0$$

Klasifikace rizika

apriorní, aposteriorní proměnné

Klasifikace rizika

apriorní, aposteriorní proměnné
Poissonovská regrese

Klasifikace rizika

apriorní, aposteriorní proměnné
Poissonovská regrese
N_i pojistné události

$$N_i \sim Poi(d_i e^{(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij})})$$

Klasifikace rizika

apriorní, aposteriorní proměnné
Poissonovská regrese
N_i pojistné události

$$N_i \sim Poi(d_i e^{(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij})})$$

Skóre:

$$\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij})$$

Bayesovská analýza dat

Mějme Bernoulliův proces definovaný náhodnou veličinou $X \sim \mathrm{Bi}(n, \theta)$ s binomickým rozdělením pravděpodobnosti a předpokládejme, že parametr θ je přitom náhodnou veličinou s rovnoměrným rozdělením pravděpodobnosti na intervalu $(0, 1)$. Definujme *šanci na úspěch* v našem procesu jako veličinu $\gamma = \frac{\theta}{1-\theta}$. Jakou hustotu rozdělení má veličina γ?

Intuitivně asi cítíme, že nepůjde o rovnoměrné rozdělení. Označíme hledanou hustotu pravděpodobnosti $f(s)$ a ze vztahu mezi θ a γ spočteme $\theta = \frac{\gamma}{1+\gamma}$. Také okamžitě vidíme, že hustota pravděpodobnosti veličiny γ bude nenulová pouze pro kladné hodnoty proměnné. Zadání můžeme nyní zformulovat tak, že požadujeme

$$\Theta = P(\theta < \Theta) = P\left(\gamma < \frac{\Gamma}{1+\Gamma}\right) = \int_0^{\Gamma} f(s)ds, \qquad (1)$$

kde $\Gamma = \frac{\Theta}{1-\Theta}$.

Na pravé straně máme ovšem v horní mezi právě měnící se ohraničení γ a dostáváme tedy definiční vztah pro $f(s)$

$$f(s) = \left(\frac{s}{s+1}\right)' = \frac{1}{(s+1)^2}.$$

Hledaná hustota skutečně dává daleko větší pravděpodobnost malým hodnotám šance než velkým.

Jestliže pracujeme v bayesovském přístupu s binomickým modelem rozdělení pravděpodobnosti náhodné veličiny $X \sim \mathrm{Bi}(n, \theta)$, bude nás zajímat její pravděpodobnostní funkce $f_X(k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$. Na tuto funkci se ale můžeme také dívat jako na podmíněnou pravděpodobnost $P(\theta|X = k)$ při apriorním rovnoměrném rozdělení pravděpodobnosti veličiny θ na intervalu $(0, 1)$. Je to tedy právě aposteriorní rozdělení pravděpodobnosti veličiny θ odpovídající výsledku pokusu $X = k$. Následující příklad se týká obecné třídy takovýchto rozdělení pravděpodobnosti.

Najděte základní charakteristiky tzv. **Beta rozdělení** $\beta(a, b)$ s hustotou pravděpodobnosti tvaru

$$f_Y = \begin{cases} C\, y^{a-1}(1-y)^{b-1} & y \in (0,1) \\ 0 & \text{jinak.} \end{cases}$$

Najděte základní charakteristiky tzv. **Beta rozdělení** $\beta(a, b)$ s hustotou pravděpodobnosti tvaru

$$f_Y = \begin{cases} C\, y^{a-1}(1-y)^{b-1} & y \in (0,1) \\ 0 & \text{jinak.} \end{cases}$$

Konstantu C je třeba volit jako reciprokou hodnotu integrálu $\int_0^1 y^{a-1}(1-y)^{b-1}dy$, což je funkce $\mathrm{B}(a, b)$, v matematické analýze (ale také technických vědách či fyzice) známá pod názvem **Beta funkce**. Když už známe funkci Gama, která zobecňuje diskrétní hodnoty faktoriálů, vyskočí na nás např. při následujícím výpočtu:

$$\Gamma(x)\Gamma(y) = \int_0^\infty \mathrm{e}^{-t}\, t^{x-1} dt \cdot \int_0^\infty \mathrm{e}^{-s}\, s^{y-1} ds =$$
$$= \int_0^\infty \int_0^\infty \mathrm{e}^{-t-s}\, t^{x-1} s^{y-1} dt\, ds =$$
$$\text{(substituce } t = rq,\ s = r(1-q)\text{)}$$
$$= \int_{r=0}^\infty \int_{q=0}^1 \mathrm{e}^{-r} (rq)^{x-1} \big(r(1-q)\big)^{y-1} r\, dq\, dr =$$
$$= \int_{r=0}^\infty \mathrm{e}^{-r}\, r^{x+y-1} dr \cdot \int_{t=0}^1 q^{x-1}(1-q)^{y-1} dq =$$
$$= \Gamma(x+y)\mathrm{B}(x,y).$$

dostáváme tedy obecný vztah

$$\mathrm{B}(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

$$\Gamma(x)\Gamma(y) = \int_0^\infty \mathrm{e}^{-t}\, t^{x-1} dt \cdot \int_0^\infty \mathrm{e}^{-s}\, s^{y-1} ds =$$
$$= \int_0^\infty \int_0^\infty \mathrm{e}^{-t-s}\, t^{x-1} s^{y-1} dt\, ds =$$

$$= \int_{r=0}^\infty \mathrm{e}^{-r}\, r^{x+y-1} dr \cdot \int_{t=0}^1 q^{x-1}(1-q)^{y-1} dq =$$
$$= \Gamma(x+y)\mathrm{B}(x,y).$$

dostáváme tedy obecný vztah

$$\mathrm{B}(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

$$\Gamma(x)\Gamma(y) = \int_0^\infty \mathrm{e}^{-t}\, t^{x-1} dt \cdot \int_0^\infty \mathrm{e}^{-s}\, s^{y-1} ds =$$
$$= \int_0^\infty \int_0^\infty \mathrm{e}^{-t-s}\, t^{x-1} s^{y-1} dt\, ds =$$

$$= \Gamma(x+y)\mathrm{B}(x,y).$$

dostáváme tedy obecný vztah

$$\mathrm{B}(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

$$\Gamma(x)\Gamma(y) = \int_0^\infty \mathrm{e}^{-t}\, t^{x-1} dt \cdot \int_0^\infty \mathrm{e}^{-s}\, s^{y-1} ds =$$
$$= \int_0^\infty \int_0^\infty \mathrm{e}^{-t-s}\, t^{x-1} s^{y-1} dt\, ds =$$
$$\text{(substituce } t = rq,\ s = r(1-q)\text{)}$$
$$= \int_{r=0}^\infty \int_{q=0}^1 \mathrm{e}^{-r} (rq)^{x-1} \big(r(1-q)\big)^{y-1} r\, dq\, dr =$$
$$= \int_{r=0}^\infty \mathrm{e}^{-r}\, r^{x+y-1} dr \cdot \int_{t=0}^1 q^{x-1}(1-q)^{y-1} dq =$$
$$= \Gamma(x+y)\mathrm{B}(x,y).$$

dostáváme tedy obecný vztah

$$\mathrm{B}(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

a z vlastností Gamma funkce již snadno plyne, že pro přirozená kladná a, b bude platit

$$\mathrm{B}(n-k+1,k+1) = \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}\binom{n}{k}^{-1}.$$

Přímým výpočtem vidíme, že střední hodnota veličiny $X \sim \beta(a, b)$ s beta rozdělením je (využíváme vztah $\Gamma(z+1) = z\Gamma(z)$)

$$\mathrm{E}\, X = \frac{\mathrm{B}(a+1, b)}{\mathrm{B}(a, b)} = \frac{a}{a+b}.$$

Je-li $a = b$ vyjde střední hodnota i medián $\frac{1}{2}$.
Přímo se také spočte rozptyl

$$\operatorname{var} X = \mathrm{E}(X - \mathrm{E}\, X)^2 = \frac{ab}{(a+b)^2(a+b+1)}.$$

Pro $a = b$ tedy dostáváme $\operatorname{var} X = \frac{1}{8a+4}$, což ukazuje, že pro rostoucí $a = b$ klesá rozptyl. V případě $a = b = 1$ dostáváme obyčejné rovnoměrné rozdělení na intervalu $(0, 1)$.

Předpokládejme, že v Bernoulliho procesu je šance zdaru θ náhodná veličina s rozdělením pravděpodobnosti $\beta(a, b)$. Jak bude vypadat rozdělení pravděpodobnosti veličiny $\gamma = \frac{\theta}{1-\theta}$?

Předpokládejme, že v Bernoulliho procesu je šance zdaru θ náhodná veličina s rozdělením pravděpodobnosti $\beta(a, b)$. Jak bude vypadat rozdělení pravděpodobnosti veličiny $\gamma = \frac{\theta}{1-\theta}$?
V již řešeném příkladě jsme měli speciální případ s rovnoměrným rozdělením $\beta(1, 1)$. Můžeme tedy pokračovat v řešení v rovnosti $\|1\|$, kdy jsme tvar tohoto rozdělení použili. Dostáváme nyní na levé straně místo Θ výraz

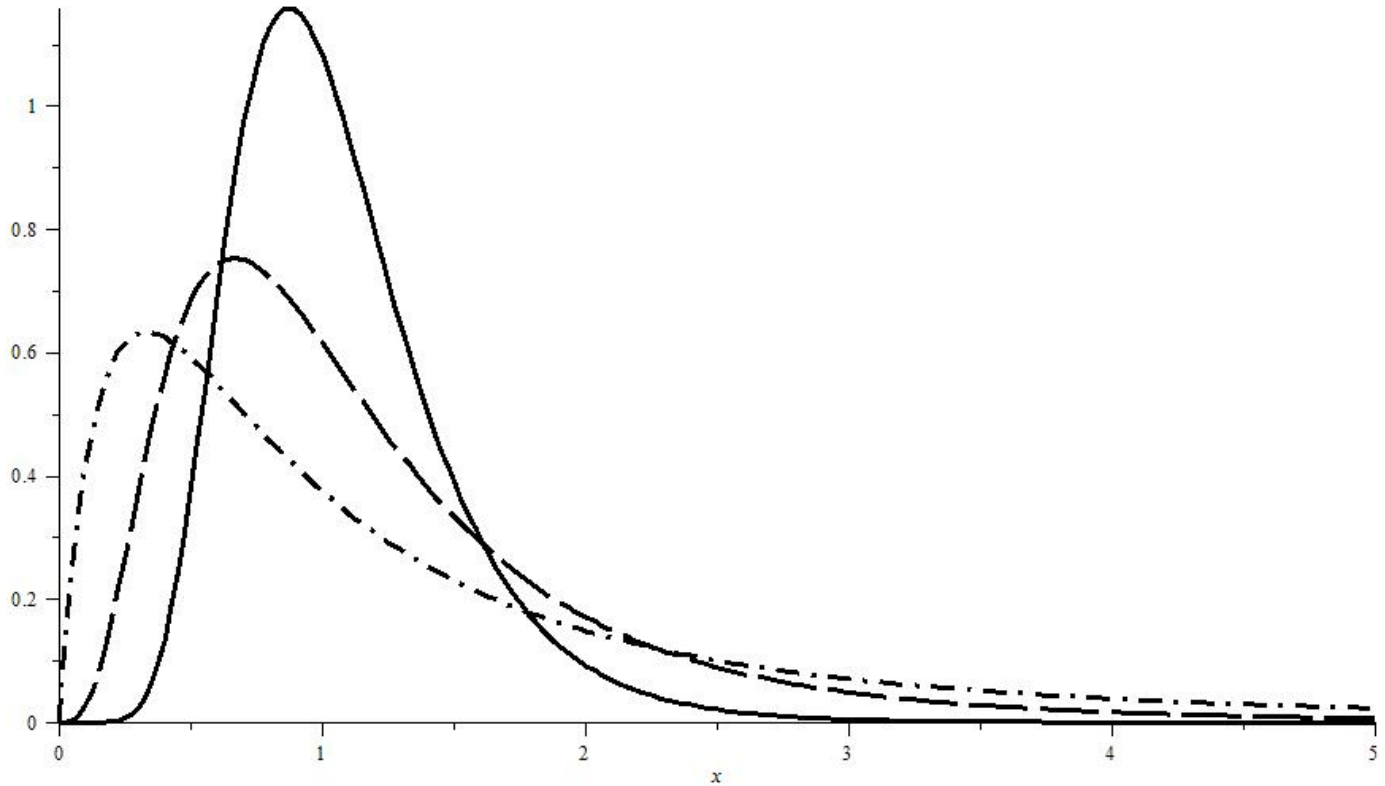
$$\frac{1}{\mathrm{B}(a, b)} \int_0^{\Theta} t^{a-1}(1-t)^{b-1} dt$$

a při derivování musíme použít pravidlo pro derivování integrálu s proměnnou horní mezí. Dostáváme proto pro hledanou hustotu

$$\mathrm{B}(a, b) f(s) = \left(\frac{s}{s+1}\right)^{a-1} \left(1 - \frac{s}{s+1}\right)^{b-1} \frac{1}{(s+1)^2} =$$

$$= \left(\frac{s^{a-1}}{s+1}\right)^{a+b}.$$

Na obrázku jsou vyneseny hustoty pro hodnoty $a = b = p = 2, 5, 15$.



V případě Bernoulliho procesu popsaného náhodnou veličinou $X \sim \mathrm{Bi}(n, \theta)$ a apriorní pravděpodobnosti náhodné veličiny θ s beta rozdělením, má i aposteriorní pravděpodobnost opět beta rozdělení s vhodnými parametry závislými na výsledku pokusu. Jaká bude aposteriorní střední hodnota veličiny θ (tj. bayesovský bodový odhad této náhodné veličiny)?

Jak je zdůvodněno v odstavci **??** teoretického sloupce, bude aposteriorní hustota pravděpodobnosti, až na násobek vhodnou konstantou, dána jako součin apriorní hustoty pravděpodobnosti

$$g(\theta) = \frac{1}{\mathrm{B}(a,b)}\theta^{a-1}(1-\theta)^{b-1}$$

a pravděpodobnosti sledované veličiny X za podmínky, že nastala hodnota θ. Dostáváme tedy za předpokladu, že v Bernoulliho procesu nastalo k zdarů, aposteriorní hustotu (použitý znak místo rovnosti značí „proporcionální“)

$$\begin{aligned} g(\theta|X=k) &\propto P(X=k|\theta)g(\theta) \propto \\ &\propto \theta^k(1-\theta)^{n-k}\theta^{a-1}(1-\theta)^{b-1} = \\ &= \theta^{a+k-1}(1-\theta)^{b+n-k-1}. \end{aligned}$$

Dostali jsme tedy, až na konstantu, kterou nemusíme vůbec vyčíslovat, skutečně hustotu aposteriorního rozdělení pro veličinu θ s rozdělením $\mathrm{B}(a+k, b+n-k)$.

Její aposteriorní střední hodnota je

$$\hat{\theta} = \frac{a + k}{a + b + n}.$$

Pro n a k jdoucí do nekonečna, tak aby $k/n \to p$, bude i pro náš aposteriorní odhad platit $\hat{\theta} \to p$. Je tedy vidět, že při velkých hodnotách n a k bude převažovat pozorovaný podíl úspěšných pokusů nad apriorním předpokladem. Nicméně pro menší hodnoty je apriorní předpoklad naopak velice významný.

Příklad odhadu nehodovosti

Máme data o nehodovosti $N = 20$ řidičů za posledních $n = 10$ let (k–tá položka označuje počet roků, ve kterých došlo k nehodě u k–tého řidiče):

$$0, 0, 2, 0, 0, 2, 2, 0, 6, 4, 3, 1, 1, 1, 0, 0, 5, 1, 1, 0.$$

Předpokládáme, že pravděpodobnosti nehod u jednotlivých řidičů jsou konstanty p_j, $j = 1, \ldots, N$.
Odhadněte pro každého řidiče pravděpodobnost, že bude mít nehodu v následujícím roce (např. pro určení jeho individuálního pojistného). [1]

[1] Tento příklad je převzat z příspěvku M. Friesl, Bayesovské odhady v některých modelech, publikováno v: Analýza dat 2004/II (K. Kupka, ed.), Trilobyte Statistical Software, Pardubice, 2005, pp. 21-33.

Zavedeme si náhodné veličiny X_{ij} s hodnotami 0, když i-tý řidič v j-tém roce neměl žádnou nehodu, a hodnotami 1 pokud nehodu měl. Jednotlivé roky považujeme za nezávislé, můžeme proto předpokládat, že náhodné veličiny $S_j = \sum_{i=1}^{n} X_{ji}$ udávající počet nehod za všech $n = 10$ let mají rozdělení $\mathrm{Bi}(n, p_j)$.

Zavedeme si náhodné veličiny X_{ij} s hodnotami 0, když i-tý řidič v j-tém roce neměl žádnou nehodu, a hodnotami 1 pokud nehodu měl. Jednotlivé roky považujeme za nezávislé, můžeme proto předpokládat, že náhodné veličiny $S_j = \sum_{i=1}^{n} X_{ji}$ udávající počet nehod za všech $n = 10$ let mají rozdělení $\mathrm{Bi}(n, p_j)$.
Samozřejmě bychom mohli odhadnout pravděpodobnosti pro všechny řidiče společně, tj. pomocí aritmetického průměru

$$\hat{p} = \frac{1}{N} \sum_{j=1}^{n} S_j \frac{1}{n} = \frac{1}{20} \frac{29}{10} = 0{,}145.$$

Když ale uvážíme homogennost rozdělení veličin X_j, těžko je lze považovat za shodné, proto bude takovýto odhad jistě zavádějící.

Zavedeme si náhodné veličiny X_{ij} s hodnotami 0, když i-tý řidič v j-tém roce neměl žádnou nehodu, a hodnotami 1 pokud nehodu měl. Jednotlivé roky považujeme za nezávislé, můžeme proto předpokládat, že náhodné veličiny $S_j = \sum_{i=1}^{n} X_{ji}$ udávající počet nehod za všech $n = 10$ let mají rozdělení $\mathrm{Bi}(n, p_j)$.
Samozřejmě bychom mohli odhadnout pravděpodobnosti pro všechny řidiče společně, tj. pomocí aritmetického průměru

$$\hat{p} = \frac{1}{N} \sum_{j=1}^{n} S_j \frac{1}{n} = \frac{1}{20} \frac{29}{10} = 0{,}145.$$

Když ale uvážíme homogennost rozdělení veličin X_j, těžko je lze považovat za shodné, proto bude takovýto odhad jistě zavádějící. Opačný extrém, tj. zcela nezávislý a individuální odhad

$$\hat{p}_j = \frac{1}{n} S_j$$

je samozřejmě také nevhodný, protože jistě nechceme předepisovat nulové pojistné, dokud nedojde k první nehodě.

Jako realistický se jeví postup, ve kterém využijeme stejný předpoklad apriorního rozdělení pravděpodobnosti p_j nehodovosti u jednotlivých řidičů. V praxi se zpravidla používá model s Poissonovým rozdělením $\mathrm{Po}(\lambda_j)$ u j-tého řidiče s dalšími předpoklady o rozdělení parametru λ mezi řidiči. Docela dobře (a hlavně jednoduše) můžeme také předpokládat, že v našem případě půjde o rozdělení

Jako realistický se jeví postup, ve kterém využijeme stejný předpoklad apriorního rozdělení pravděpodobnosti p_j nehodovosti u jednotlivých řidičů. V praxi se zpravidla používá model s Poissonovým rozdělením $\mathrm{Po}(\lambda_j)$ u j-tého řidiče s dalšími předpoklady o rozdělení parametru λ mezi řidiči. Docela dobře (a hlavně jednoduše) můžeme také předpokládat, že v našem případě půjde o rozdělení $p_j \sim \beta(a, b)$ s vhodnými parametry a, b, které by měly odrážet kumulované výsledky všech řidičů. Pojďme tedy touto cestou.

Víme, že aposteriorní rozdělení pravděpodobností bude $(p_j|S_j = k) = \beta(a + k, b + n - k)$, takže příslušná střední hodnota bude

$$\hat{p}_j^b = \frac{a + k}{a + b + n}.$$

Víme, že aposteriorní rozdělení pravděpodobností bude $(p_j|S_j = k) = \beta(a + k, b + n - k)$, takže příslušná střední hodnota bude

$$\hat{p}_j^b = \frac{a + k}{a + b + n}.$$

Srovnejme si tento odhad s výše uvedeným společným odhadem $\hat{p}$ a individuálním $\hat{p}_j$. Zaveďme si k tomu hodnoty $p_0 = \frac{a}{a+b}$, tj. střední hodnotu apriorního společného rozdělení pro všechny řidiče, a $n_0 = a + b$. Dostáváme

$$\hat{p}_j^b = \frac{(a + b)a}{(a + b + n)(a + b)} + \frac{nk}{(a + b + n)n} = \frac{n_0}{n_0 + n} p_0 + \frac{n}{n_0 + n} \hat{p}_j,$$

což je lineární kombinace střední hodnoty p_0 a individuálního odhadu $\hat{p}_j$.

Zbývá nám tedy už jen smysluplně odhadnout neznámé parametry a, b. Víme přitom

$$\mathrm{E}\, X_{ji} = \mathrm{E}\, \mathrm{E}(X_{ji}|p) = \mathrm{E}\, p = p_0$$
$$\frac{\mathrm{E}\, \mathrm{var}(X_{ji}|p)}{\mathrm{var}\, \mathrm{E}(X_{ji}|p)} = \frac{\mathrm{E}(p(1-p))}{\mathrm{var}\, p} = a + b = n_0$$

a přitom veličiny na levých stranách můžeme přímo odhadnout.

Zbývá nám tedy už jen smysluplně odhadnout neznámé parametry a, b. Víme přitom

$$\mathrm{E}\, X_{ji} = \mathrm{E}\, \mathrm{E}(X_{ji}|p) = \mathrm{E}\, p = p_0$$
$$\frac{\mathrm{E}\, \mathrm{var}(X_{ji}|p)}{\mathrm{var}\, \mathrm{E}(X_{ji}|p)} = \frac{\mathrm{E}(p(1-p))}{\mathrm{var}\, p} = a + b = n_0$$

a přitom veličiny na levých stranách můžeme přímo odhadnout.

$$\mathrm{E}\, X_{ij} = \mathrm{E}\ \mathrm{E}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N} \hat{p}_j$$

$$\mathrm{E}\,\mathrm{var}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big)$$

$$\mathrm{var}\,\mathrm{E}(X_{ji}|p) \simeq s^2_{\hat{p}_j} - \frac{1}{nN}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big),$$

kde $s^2_{\hat{p}_j}$ označuje výběrový rozptyl mezi individuálními odhady (čtenář si může promyslet, že odečtením posledního výrazu vpravo zajišťujeme, aby i poslední odhad byl nestranný).
Protože pro uvedená data takto dostáváme $n_0 \simeq 3{,}8643$ a $p_0 \simeq 0{,}1450$, vyjde nám bayesovský odhad individuální pravděpodobnosti nehod

$$\hat{p}^b_j = 0{,}154 \cdot 0{,}145 + 0{,}846 \cdot \hat{p}_j.$$

$$\mathrm{E}\, X_{ij} = \mathrm{E}$$

$$\mathrm{E}\,\mathrm{var}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big)$$

$$\mathrm{var}\,\mathrm{E}(X_{ji}|p) \simeq s^2_{\hat{p}_j} - \frac{1}{nN}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big),$$

kde $s^2_{\hat{p}_j}$ označuje výběrový rozptyl mezi individuálními odhady (čtenář si může promyslet, že odečtením posledního výrazu vpravo zajišťujeme, aby i poslední odhad byl nestranný).
Protože pro uvedená data takto dostáváme $n_0 \simeq 3{,}8643$ a $p_0 \simeq 0{,}1450$, vyjde nám bayesovský odhad individuální pravděpodobnosti nehod

$$\hat{p}_j^b = 0{,}154 \cdot 0{,}145 + 0{,}846 \cdot \hat{p}_j.$$

$$\mathrm{E}\, X_{ij} = \mathrm{E}$$

$$\operatorname{var} \mathrm{E}(X_{ji}|p) \simeq s^2_{\hat{p}_j} - \frac{1}{nN} \sum_{j=1}^{N} \Big(\frac{n}{n-1} \hat{p}_j (1 - \hat{p}_j) \Big),$$

kde $s^2_{\hat{p}_j}$ označuje výběrový rozptyl mezi individuálními odhady (čtenář si může promyslet, že odečtením posledního výrazu vpravo zajišťujeme, aby i poslední odhad byl nestranný).
Protože pro uvedená data takto dostáváme $n_0 \simeq 3{,}8643$ a $p_0 \simeq 0{,}1450$, vyjde nám bayesovský odhad individuální pravděpodobnosti nehod

$$\hat{p}^b_j = 0{,}154 \cdot 0{,}145 + 0{,}846 \cdot \hat{p}_j.$$

$$\mathrm{E}\, X_{ij} = \mathrm{E}\ \mathrm{E}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N} \hat{p}_j$$

$$\mathrm{E}\,\mathrm{var}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big)$$

$$\mathrm{var}\,\mathrm{E}(X_{ji}|p) \simeq s^2_{\hat{p}_j} - \frac{1}{nN}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big),$$

$$\mathrm{E}\, X_{ij} = \mathrm{E}\ \mathrm{E}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N} \hat{p}_j$$

$$\mathrm{E}\,\mathrm{var}(X_{ji}|p) \simeq \frac{1}{N}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big)$$

$$\mathrm{var}\,\mathrm{E}(X_{ji}|p) \simeq s^2_{\hat{p}_j} - \frac{1}{nN}\sum_{j=1}^{N}\Big(\frac{n}{n-1}\hat{p}_j(1-\hat{p}_j)\Big),$$

kde $s^2_{\hat{p}_j}$ označuje výběrový rozptyl mezi individuálními odhady (čtenář si může promyslet, že odečtením posledního výrazu vpravo zajišťujeme, aby i poslední odhad byl nestranný).
Protože pro uvedená data takto dostáváme $n_0 \simeq 3{,}8643$ a $p_0 \simeq 0{,}1450$, vyjde nám bayesovský odhad individuální pravděpodobnosti nehod

$$\hat{p}^b_j = 0{,}154 \cdot 0{,}145 + 0{,}846 \cdot \hat{p}_j.$$

Jde tedy o kombinaci spolehlivého odhadu $\hat{p} = 0{,}145$ kolektivní pravděpodobnosti p_0 s individuálním (frekvenčním) odhadem $\hat{p}_j$, který je pořízen z malého počtu pozorování $n = 10$ u jediného řidiče.