

Komisja Egzaminacyjna dla Aktuariuszy

LXXVIII Egzamin dla Aktuariuszy

Sesja egzaminacyjna w dniu 27 marca 2018 r.

Modelowanie

Imię i nazwisko osoby egzaminowanej:

Czas trwania egzaminu: 120 minut

Uwagi

- a) W zadaniach wartość zagrożona na poziomie ufności α jest definiowana jako kwantyl rzędu α rozkładu odpowiedniej zmiennej losowej, tzn.

$$VaR_\alpha(X) = \inf\{x: F_X(x) \geq \alpha\}.$$

- b) W zadaniach zastosowano następujące oznaczenia:

$E(X)$ – wartość oczekiwana

$D(X)$ – odchylenie standardowe

- c) Wartości $\chi_{\alpha;v}^2$ rozkładu chi-kwadrat spełniające warunek $P(\chi^2 \geq \chi_{\alpha;v}^2) = \alpha$

$v \backslash \alpha$	0,995	0,99	0,975	0,95	0,9	0,1	0,05	0,025	0,01	0,005
1	0,000	0,000	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188

- d) Kwantyle rzędu 0,995 rozkładu $Gamma(\alpha, \lambda)$ (dla wybranych parametrów α, λ)

$\alpha \backslash \lambda$	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
0,5	78,7944	39,3972	26,2648	19,6986	15,7589	13,1324	11,2563	9,8493	8,7549	7,8794
1	105,9663	52,9832	35,3221	26,4916	21,1933	17,6611	15,1380	13,2458	11,7740	10,5966
1,5	128,3816	64,1908	42,7939	32,0954	25,6763	21,3969	18,3402	16,0477	14,2646	12,8382
2	148,6026	74,3013	49,5342	37,1506	29,7205	24,7671	21,2289	18,5753	16,5114	14,8603
2,5	167,4960	83,7480	55,8320	41,8740	33,4992	27,9160	23,9280	20,9370	18,6107	16,7496
3	185,4758	92,7379	61,8253	46,3690	37,0952	30,9126	26,4965	23,1845	20,6084	18,5476
3,5	202,7774	101,3887	67,5925	50,6943	40,5555	33,7962	28,9682	25,3472	22,5308	20,2777
4	219,5495	109,7748	73,1832	54,8874	43,9099	36,5916	31,3642	27,4437	24,3944	21,9550
16	563,2811	281,6406	187,760	140,820	112,656	93,8802	80,4687	70,4101	62,5868	56,3281

- e) Wybrane kwantyle standardowego rozkładu normalnego

α	0,900	0,950	0,975	0,990	0,995	0,999
u_α	1,282	1,645	1,960	2,326	2,576	3,090

Zadanie 1.

- a) Podać definicję koherentnej (spójnej) miary ryzyka.
 b) Sprawdzić, czy określona poniższą formułą miara ryzyka jest koherentna (jest to zasada wariacji kalkulacji składki):

$$\varrho(X) = E(X) + \beta D^2(X), \quad \beta > 0,$$

gdzie X – nieujemna zmienna losowa modelująca straty.

Odpowiedzi:**Odp. a)**

Niech \mathbf{X} oznacza przestrzeń liniową zmiennych losowych określonych na odpowiedniej przestrzeni probabilistycznej. Funkcjonał $\varrho: \mathbf{X} \rightarrow \mathbb{R}$ nazywamy koherentną miarą ryzyka, gdy spełnia następujące własności:

W1: $\varrho(X + c) = \varrho(X) + c$, dla ustalonego $X \in \mathbf{X}$ i dowolnego $c \in \mathbb{R}$ (funkcjonał jest ekwiwariantny ze względu na translację)

W2: $\varrho(aX) = a\varrho(X)$, dla ustalonego $X \in \mathbf{X}$ i dowolnego $a > 0$ (funkcjonał jest dodatnio jednorodny)

W3: Jeżeli $X_1, X_2 \in \mathbf{X}$ i $X_1 \leq X_2$ prawie wszędzie, to $\varrho(X_1) \leq \varrho(X_2)$ (funkcjonał jest monotoniczny)

W4: $\varrho(X_1 + X_2) \leq \varrho(X_1) + \varrho(X_2)$, dla dowolnych $X_1, X_2 \in \mathbf{X}$ (funkcjonał jest subaddytywny)

Odp. b)

Miara ϱ nie jest koherentna. Nie spełnia warunku dodatniej jednorodności.

Rozwiązanie:

b) $\varrho(aX) = E(aX) + \beta D^2(aX) = aE(X) + a^2\beta D^2(X),$

czyli $\varrho(aX) \neq a\varrho(X)$

Zadanie 2.

Straty dla pierwszej linii biznesu (LOB1) są modelowane za pomocą zmiennej losowej X_1 o rozkładzie $Gamma(\alpha = 2, \lambda = 0,1)$, natomiast drugiej (LOB2) – zmiennej losowej X_2 o rozkładzie $Gamma(\alpha = 1, \lambda = 0,1)$. (Uwaga! Przyjęto parametryzację, w której $E(X) = \frac{\alpha}{\lambda}$, $D^2(X) = \frac{\alpha}{\lambda^2}$.) Wymogi kapitałowe są ustalane na poziomie kapitału ekonomicznego wyznaczanego za pomocą wartości zagrożonej na poziomie ufności 0,995, tzn. $\kappa(X) = VaR_{0,995}(X) - E(X)$. Wyznaczyć efekt dywersyfikacji powstały w wyniku agregacji tych linii biznesu, jeżeli wiadomo, że:

- łączne straty są modelowane za pomocą zmiennej losowej $S = X_1 + X_2$ (funkcją agregującą jest suma),
- zmienne X_1 i X_2 są niezależne.

Odpowiedzi:**Odp.**

$D = 34,5466$ (ewentualnie: $d = 0,3551$ lub $d' = 0,6449$)

Rozwiązanie:

Efekt dywersyfikacji (możliwe sposoby wyliczenia):

- $D = \kappa(X_1) + \kappa(X_2) - \kappa(S)$
- $d = 1 - \frac{\kappa(S)}{\kappa(X_1) + \kappa(X_2)}$
- $d' = \frac{\kappa(S)}{\kappa(X_1) + \kappa(X_2)}$

1. $\kappa(X_1) = VaR_{0,995}(X_1) - E(X_1)$

$$E(X_1) = \frac{2}{0,1} = 20, VaR_{0,995}(X_1) = 74,3013$$

$$\kappa(X_1) = 54,3013$$

2. $\kappa(X_2) = VaR_{0,995}(X_2) - E(X_2)$

$$E(X_2) = \frac{1}{0,1} = 10, VaR_{0,995}(X_2) = 52,9832$$

$$\kappa(X_2) = 42,9832$$

3. Ponieważ zmienne X_1 i X_2 są niezależne, suma $S = X_1 + X_2$ ma rozkład

$Gamma(\alpha = 2 + 1, \lambda = 0,1)$. Stąd:

$$E(S) = \frac{3}{0,1} = 30, VaR_{0,995}(S) = 92,7379$$

$$\kappa(S) = 62,7379$$

Zatem: $D = 34,5466$; $d = 0,3551$; $d' = 0,6449$

Zadanie 3.

Straty w rocznym horyzoncie czasowym dla pierwszego rodzaju ryzyka są modelowane za pomocą zmiennej losowej X , która ma rozkład Pareto o gęstości: $f(x) = \frac{1}{x^2}, x > 1$, a dla drugiego – zmiennej losowej Y o rozkładzie Pareto z gęstością $f(y) = \frac{2}{y^3}, y > 1$.

Niech:

- $P_N = P(X \leq VaR_{0,95}(X), Y \leq VaR_{0,95}(Y))$ będzie prawdopodobieństwem tego, że straty z pierwszego i drugiego rodzaju ryzyka nie przekroczą odpowiednio $VaR_{0,95}(X)$ i $VaR_{0,95}(Y)$ wyznaczonym przy założeniu, że zmienne X i Y są **niezależne**.
- $P_G = P(X \leq VaR_{0,95}(X), Y \leq VaR_{0,95}(Y))$ będzie prawdopodobieństwem tego, że straty z pierwszego i drugiego rodzaju ryzyka nie przekroczą odpowiednio $VaR_{0,95}(X)$ i $VaR_{0,95}(Y)$ wyznaczonym przy założeniu, że struktura zależności między zmiennymi X i Y jest modelowana za pomocą kopuli archimedesowej o silnym generatorze $\phi(t) = (-\ln t)^2$ (jest to kopula Gumbela-Hougaard).

Zarówno $VaR_{0,95}(X)$ jak i $VaR_{0,95}(Y)$ odnoszą się do rocznego horyzontu czasowego.

Obliczyć: $|P_N - P_G|$.

Odpowiedzi:

Odp.

$$|P_N - P_G| = 0,0275$$

Rozwiązanie:

1. Jeżeli zmienne niezależne, to $F_N(x, y) = F_X(x) \cdot F_Y(y)$.

$$\begin{aligned} P_N &= P(X \leq VaR_{0,95}(X), Y \leq VaR_{0,95}(Y)) = F_N(VaR_{0,95}(X), VaR_{0,95}(Y)) = \\ &= F_X(VaR_{0,95}(X)) \cdot F_Y(VaR_{0,95}(Y)) = F_X(F_X^{-1}(0,95)) \cdot F_Y(F_Y^{-1}(0,95)) = \\ &= 0,95 \cdot 0,95 = 0,9025 \end{aligned}$$

2. Jeżeli zależność między zmiennymi X i Y jest modelowana za pomocą kopuli C , to $F_G(x, y) = C(F_X(x), F_Y(y))$.

$$\begin{aligned} P_G &= P(X \leq VaR_{0,95}(X), Y \leq VaR_{0,95}(Y)) = F_G(VaR_{0,95}(X), VaR_{0,95}(Y)) = \\ &= C(F_X(VaR_{0,95}(X)), F_Y(VaR_{0,95}(Y))) = C(0,95, 0,95). \end{aligned}$$

W zadaniu C jest kopulą Archimedesesa, zatem: $C = \phi^{-1}(\phi(u) + \phi(v))$.

Stąd:

$$\phi(0,95) = (-\ln(0,95))^2 = 0,002631$$

$$\phi^{-1}(z) = \exp(-\sqrt{z})$$

$$C(0,95, 0,95) = \exp(-\sqrt{2 \cdot 0,002631}) = 0,9300$$

Zadanie 4.

Wykorzystując stosowaną standardowo w Solvency II metodę agregacji (tj. metodę wariancji-kowariancji), ustalono łączny kapitałowy wymóg wypłacalności dla modułu ryzyka aktuarialnego w ubezpieczeniach innych niż ubezpieczenia na życie ($SCR_{non-life}$) na poziomie 50. Wynik otrzymano wiedząc, że:

- wymóg kapitałowy dla podmodułu ryzyka katastroficznego w ubezpieczeniach innych niż ubezpieczenia na życie jest równy: $SCR_{nlCAT} = 20$,
- wymóg kapitałowy dla podmodułu ryzyka związanego z rezygnacjami w ubezpieczeniach innych niż ubezpieczenia na życie jest równy: $SCR_{nlLAPSE} = 10$,
- współczynniki korelacji liniowej między poszczególnymi podmodułami ryzyka są następujące:

	X	SCR_{nlCAT}	$SCR_{nlLAPSE}$
X	1	0,25	0
SCR_{nlCAT}	0,25	1	0
$SCR_{nlLAPSE}$	0	0	1

- a) Jak nazywa się podmoduł oznaczony w tabeli ze współczynnikami korelacji przez X?
- b) Obliczyć dla tego podmodułu kapitałowy wymóg wypłacalności.
- c) Wyznaczyć efekt dywersyfikacji.

Odpowiedzi:**Odp. a)**

Podmoduł ryzyka składki i rezerw w ubezpieczeniach innych niż ubezpieczenia na życie.

Odp. b)

$$SCR_{nl prem res} = 40$$

Odp. c)

$$D = 20 \quad (\text{ewentualnie: } d = 0,2857 \text{ lub } d' = 0,7143)$$

Rozwiązanie:

b) Wymóg kapitałowy dla podmodułu ryzyka składki i rezerw w ubezpieczeniach innych niż ubezpieczenia na życie można wyznaczyć rozwiązując równanie:

$$[x \quad 20 \quad 10] \cdot \begin{bmatrix} 1 & 0,25 & 0 \\ 0,25 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ 20 \\ 10 \end{bmatrix} = 50^2$$

$$x^2 + 10x - 2000 = 0.$$

Stąd $x_1 = 40$ (drugie rozwiązanie jest ujemne).

$$c) \quad D = 40 + 20 + 10 - 50 = 20$$

$$d = 1 - \frac{50}{40 + 20 + 10} = 0,2857$$

$$d' = \frac{50}{40 + 20 + 10} = 0,7143$$

Zadanie 5.

Straty dla pierwszej linii biznesu (LOB1) są modelowane za pomocą zmiennej losowej X_1 o rozkładzie Pareto z gęstością $f(x_1) = \frac{2}{x_1^3}, x_1 > 1$, dla drugiej (LOB2) – zmiennej losowej X_2 o rozkładzie $Gamma(\alpha = 3, \lambda = 0,5)$. (Uwaga! Przyjęto parametryzację, w której $E(X_2) = \frac{\alpha}{\lambda}, D^2(X_2) = \frac{\alpha}{\lambda^2}$). Struktura zależności między stratami X_1 i X_2 (wektora losowego (X_1, X_2)) jest modelowana za pomocą kopuli Gumbela z parametrem $\theta = 2$. W celu wyznaczenia wartości zagrożonej na poziomie ufności 0,95 dla łącznych strat, tzn. $VaR_{0,95}(S)$, gdzie $S = X_1 + X_2$ zastosowano symulacyjną metodę agregacji za pomocą kopuli (Gumbela z parametrem $\theta = 2$).

- a) Wiadomo, że z kopuli wygenerowano następującą parę (obserwację): $(u_1, v_1) = (0,717, 0,995)$. Wyznaczyć odpowiadającą tej parze wartość sumy $S = X_1 + X_2$ (realizację zmiennej S).
- b) Poniższa tabela przedstawia wyniki symulacji:

Uzyskane podczas symulacji realizacje zmiennej S s_i	Liczba realizacji z danego przedziału n_i
(0, 5]	300
(5, 10]	450
(10, 15]	180
(15, 20]	45
(20, 30]	20
(30, 60]	5
powyżej 60	0

Oszacować $VaR_{0,95}(S)$, przyjmując jednostajny rozkład realizacji s_i w poszczególnych przedziałach klasowych (czyli na podstawie krzywej ogiwalnej).

Odpowiedzi:**Odp. a)**

$$s_1 = 20,427379$$

Odp. b)

$$VaR_{0,95}(S) = 17,2222$$

Rozwiązanie:**a)**

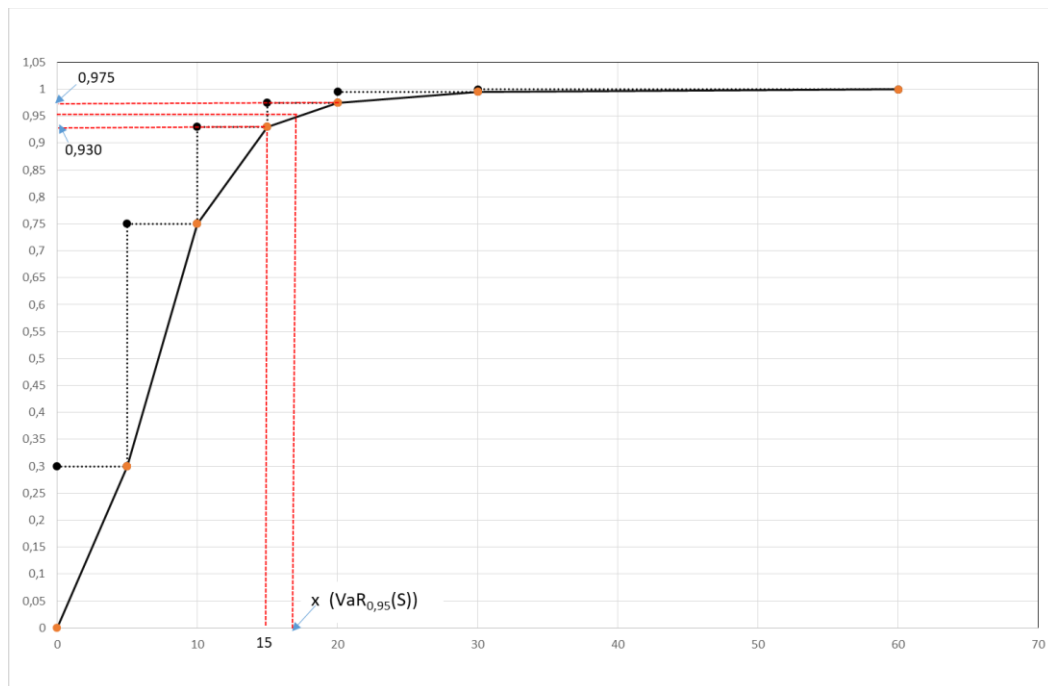
$$s_1 = F_1^{-1}(u_1) + F_2^{-1}(v_1)$$

$$F_1(x) = 1 - \frac{1}{x^2}, \text{ stąd } F_1^{-1}(u) = \sqrt{\frac{1}{1-u}}, \text{ zatem } F_1^{-1}(0,717) = 1,879779$$

$$F_2^{-1}(0,995) = 18,5476$$

$$s_1 = 1,879779 + 18,5476 = 20,427379$$

b)



Wartość zagrożoną $VaR_{0,95}(S)$ można wyznaczyć rozwiązując następujące równanie:

$$\frac{20-15}{0,975-0,930} = \frac{x-15}{0,950-0,930}.$$

Stąd $x = 17,2222$.

Zadanie 6.

Wysokość szkód (Y) modelowano wykorzystując uogólniony model liniowy (GLM). Przyjęto rozkład gamma dla zmiennej zależnej Y oraz logarytmiczną funkcję wiążącą (*link function*). Uwzględniono dwie zmienne niezależne (regresory):

- wiek samochodu (*wiek.samochodu*), przyjmującą 4 kategorie wiekowe: *wiek1*, *wiek2*, *wiek3*, *wiek4*,
- miejsce rejestracji (*m.rejestracji*), przyjmującą 6 kategorii związanych z obszarem rejestracji: A, B, C, D, E, F.

Uzyskano następujące wyniki:

Zmienna	Oszacowanie
wyraz wolny (Intercept)	7,317033
<i>wiek.samochodu</i> : <i>wiek2</i>	0,095588
<i>wiek.samochodu</i> : <i>wiek3</i>	0,132231
<i>wiek.samochodu</i> : <i>wiek4</i>	0,210605
<i>m.rejestracji</i> : B	0,043452
<i>m.rejestracji</i> : C	0,139058
<i>m.rejestracji</i> : D	-0,009905
<i>m.rejestracji</i> : E	0,239069
<i>m.rejestracji</i> : F	0,356198

Wiadomo, że parametr dyspersji wynosi: 2,853176.

Niech:

$\hat{\mu}_{(wiek1,A)}$ - oznacza prognozę wysokości szkody dla samochodu z pierwszej kategorii wiekowej (*wiek1*) i zarejestrowanego w obszarze A,

$\hat{\mu}_{(wiek1,F)}$ - oznacza prognozę wysokości szkody dla samochodu z pierwszej kategorii wiekowej (*wiek1*) i zarejestrowanego w obszarze F.

- a. Obliczyć: $\hat{\mu}_{(wiek1,F)} - \hat{\mu}_{(wiek1,A)}$.
- b. Obliczyć wariancję i odchylenie standardowe prognozy $\hat{\mu}_{(wiek1,A)}$.
- c. Dla jakiej grupy samochodów (tzn. w jakim wieku i miejscu rejestracji) prognoza wysokości szkód będzie najniższa?

Odpowiedzi:

Odp. a)

$$\hat{\mu}_{(wiek1,F)} - \hat{\mu}_{(wiek1,A)} = 644,2871$$

.....

Odp. b)

$$D^2(\hat{\mu}_{(wiek1,A)}) = 6468784,356, \quad D(\hat{\mu}_{(wiek1,A)}) = 2543,380$$

.....

Odp. c)

Dla samochodu z grupy wiekowej *wiek1* i obszaru rejestracji D.

Rozwiązanie:

a)

$$\hat{\mu}_{(wiek1,F)} = \exp(7,317033 + 0,353198) = 2150,01694$$

$$\hat{\mu}_{(wiek1,A)} = \exp(7,317033) = 1505,729835$$

b)

W przypadku rozkładu gamma wariancja $D^2(y) = \phi\mu^2$,

gdzie:

ϕ - parametr dyspersji,

μ - średnia.

Zatem:

$$D^2(\hat{\mu}_{(wiek1,A)}) = 2,853176 \cdot 1505,729835^2 = 6468784,356$$

$$D(\hat{\mu}_{(wiek1,A)}) = 2543,380$$

Zadanie 7.

W poniższej tabeli przedstawiono dane dotyczące liczby wypadków (odnotowanych w ciągu jednego roku) oraz wieku (w latach) dla 6 kierowców.

Nr kierowcy (i)	Wiek (x_i)	Liczba wypadków (y_i)
1	30	3
2	40	1
3	35	0
4	35	2
5	20	4
6	40	0
Suma	200	10

Zakładamy, że y_1, y_2, \dots, y_6 są realizacjami niezależnych zmiennych losowych Y_1, Y_2, \dots, Y_6 o rozkładach Poissona ze średnią $\lambda_i = \beta x_i, i = 1, 2, \dots, 6$.

- Wykorzystując Metodę Największej Wiarygodności oszacować parametr β .
- Wykorzystując oszacowany model wyznaczyć prawdopodobieństwo, że 50-cio letni kierowca spowoduje więcej niż jeden wypadek.

Odpowiedzi:**Odp. a)**

$$\hat{\beta} = 0,05$$

Odp. b)

$$P(Y > 1) = 0,7127$$

Rozwiązanie:

$$\text{a) } L(\beta) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \frac{\exp(-\beta x_i) \cdot (\beta x_i)^{y_i}}{y_i!}$$

$$l(\beta) = \ln(L(\beta)) = \sum_{i=1}^n [-\beta x_i + y_i \ln(\beta) + y_i \ln(x_i) - \ln(y_i!)]$$

$$l'(\beta) = -\sum_{i=1}^n x_i + \frac{1}{\beta} \sum_{i=1}^n y_i$$

$$\text{W zadaniu: } l'(\beta) = -200 + \frac{10}{\beta}.$$

$$l'(\beta) = 0, \text{ gdy } \beta = \frac{1}{20}, \text{ zatem } \hat{\beta} = 0,05.$$

b) W przypadku 50-cio letniego kierowcy parametr rozkładu Poissona wynosi:

$$\lambda = 0,05 \cdot 50 = 2,5$$

$$\text{Stąd } P(Y > 1) = 1 - (P(Y = 0) + P(Y = 1)) = 1 - 3,5 \cdot \exp(-2,5) = 0,7127$$

Zadanie 8.

W poniższej tabeli przedstawiono dane 12 losowo wybranych kierowców dotyczące:

- liczby wypadków (zmienna *liczba*),
- wieku (zmienna *wiek* [w latach]),
- płci (zmienna *plec*, gdzie: K - kobieta, M - mężczyzna),
- użytkownika samochodu (zmienna *uzytkowanie*, gdzie: P – w celach prywatnych, S- w celach służbowych).

Nr obserwacji	1	2	3	4	5	6	7	8	9	10	11	12
<i>liczba</i>	0	0	1	1	0	2	0	3	0	1	1	0
<i>wiek</i>	45	50	25	30	45	20	30	22	50	30	30	50
<i>plec</i>	M	K	M	K	K	K	M	M	M	K	M	K
<i>uzytkowanie</i>	P	P	S	S	P	S	P	S	S	P	P	P

Na podstawie tych danych oszacowano uogólniony model liniowy (GLM) dla liczby wypadków. W modelu tym dla zmiennej zależnej przyjęto rozkład Poissona z logarytmiczną funkcją wiążącą (*link function*) oraz uwzględniono wszystkie zmienne objaśniające (tj. *wiek*, *plec* i *uzytkowanie*).

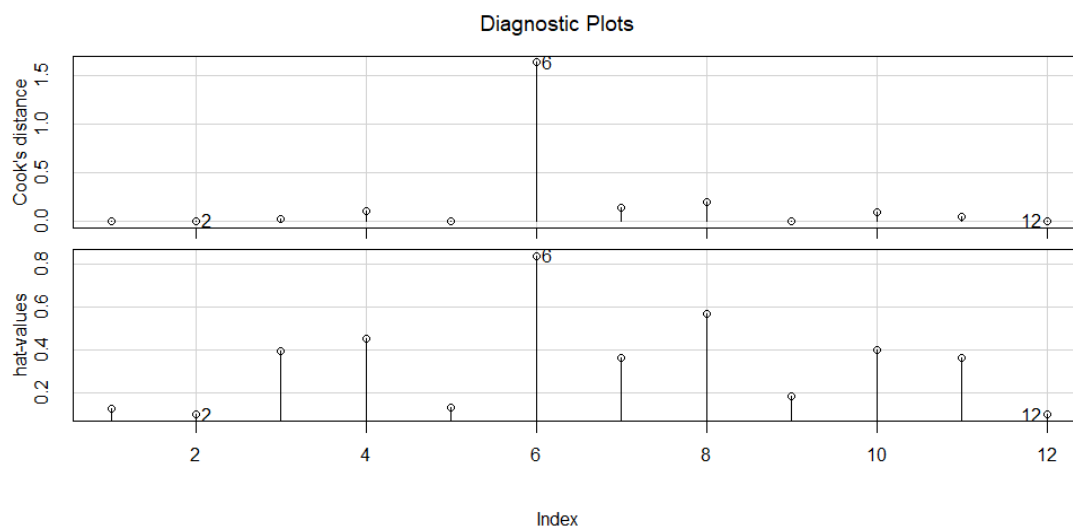
Otrzymano następujące parametry modelu:

(Intercept)	<i>wiek</i>	<i>plec</i> M	<i>uzytkowanie</i> S
4.0816990	-0,1549239	0,1067970	0,0257586

oraz elementy leżące na głównej przekątnej macierzy „daszkowej” H (*hat matrix*):

0,1235 0,0948 0,3909 0,4494 0,1301 0,8394 0,3633 0,5691 0,1808 0,4007 0,3633 0,0948

Ponadto wyznaczono następujący wykres diagnostyczny (na osi poziomej są numery obserwacji):



- a) Zinterpretować wykres diagnostyczny.
- b) Dla szóstej obserwacji wyznaczyć resztę Pearsona i miarę Cooka (*Cook's distance*).

Odpowiedzi:**Odp. a)**

1. *Cook's distance*: Można uznać, że obserwacja nr 6 miała istotny wpływ na obciążenie parametrów modelu.
2. *hat-values*: Obserwację nr 6 można uznać za wpływową, tzn. mającą istotny wpływ na wartość teoretyczną $\hat{\mu}_6$

Odp. b)Resztę Pearsona: $e_{P,6} = -0,44851$ Miara Cooka: $D_6 = 1,63669$ **Rozwiązanie:****b)**

$$1. \text{ Reszta Pearsona: } e_{P,i} = \frac{y_i - \hat{\mu}_i}{\sqrt{D^2(y_i|\mathbf{x})/\phi}}$$

W modelu Poissona parametr dyspersji $\phi = 1$ oraz $D^2(y_i|\mathbf{x}) = \mu_i$.

Wartość teoretyczna dla obserwacji nr 6 wynosi:

$$\hat{\mu}_6 = \exp(4,081699 - 0,1549239 \cdot 20 + 0,0257586) = 2,7428.$$

Stąd:

$$e_{P,6} = \frac{2 - 2,7428}{\sqrt{2,7428}} = -0,44851.$$

$$2. \text{ Miara Cooka: } D_i = \frac{e_{PS,i}^2}{k} \cdot \frac{h_i}{1-h_i},$$

gdzie:

$$e_{PS,i} = \frac{y_i - \hat{\mu}_i}{\sqrt{D^2(y_i|\mathbf{x})(1-h_i)}} - \text{standaryzowana reszta Pearsona;}$$

 k – liczba parametrów modelu; h_i – i -ty element leżący na głównej przekątnej macierzy „daszkowej” H .

Zatem dla szóstej obserwacji:

$$e_{PS,6} = \frac{2 - 2,7428}{\sqrt{2,7428 \cdot (1 - 0,8394)}} = -1,119185,$$

$$\text{skąd } D_6 = \frac{(-1,119185)^2}{4} \cdot \frac{0,8394}{1 - 0,8394} = 1,63669.$$

Zadanie 9.

Dla pewnego portfela ubezpieczeń, wykorzystując model ryzyka kolektywnego oszacowano następujące parametry rozkładu zmiennej łącznych szkód Z :

- wartość oczekiwaną: $\mu_Z = 100$,
- wariancję: $\sigma_Z^2 = 1600$,
- współczynnik skośności: $\gamma_Z = 0,5$.

Rozkład zmiennej Z aproksymowano wykorzystując:

1. rozkład normalny,
2. przesunięty rozkład gamma, tzn. przyjęto, że $Z \approx Y + c$, gdzie Y ma rozkład gamma, a c jest parametrem przesunięcia.

Wyznaczyć różnicę między oszacowaniem składki kwantylowej na poziomie ufności 0,995 uzyskanym za pomocą pierwszego i drugiego przybliżenia, tzn.:

$$|F_1^{-1}(0,995) - F_2^{-1}(0,995)|,$$

gdzie F_1 i F_2 oznaczają dystrybuanty odpowiednio dla pierwszej i drugiej metody aproksymacji.

Uwaga! Współczynnik skośności dla rozkładu gamma jest równy: $\gamma = \frac{2}{\sqrt{\alpha}}$ (w parametryzacji, w której $E(X) = \frac{\alpha}{\lambda}$, $D^2(X) = \frac{\alpha}{\lambda^2}$).

Odpowiedzi:

Odp.

$$|F_1^{-1}(0,995) - F_2^{-1}(0,995)| = 18,6006$$

Rozwiązanie:

1. Rozkład normalny: $Z \sim N(100; 40)$,

$$\text{Stąd } F_1^{-1}(0,995) = 100 + \Phi^{-1}(0,995) \cdot 40 = 100 + 2,576 \cdot 40 = 203,04.$$

2. Przesunięty gamma.

Parametry wyznaczamy rozwiązując układ równań:

$$\begin{cases} \mu_Z = c + \frac{\alpha}{\lambda} \\ \sigma_Z^2 = \frac{\alpha}{\lambda^2} \\ \gamma_Z = \frac{2}{\sqrt{\alpha}} \end{cases}$$

Otrzymujemy:

$$c = \mu_Z - \frac{2\sigma_Z}{\gamma_Z} = -60; \quad \alpha = \frac{4}{\gamma_Z^2} = 16; \quad \lambda = \frac{2}{\gamma_Z \cdot \sigma_Z} = 0,1.$$

Korzystamy z następującej własności funkcji odwrotnej do dystrybuanty:

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(p)), \text{ gdy } g \text{ jest funkcją rosnącą i ciągłą,}$$

otrzymując $F_2^{-1}(0,995) = c + F_Y^{-1}(0,995) = -60 + 281,6406 = 221,6406$.

Zadanie 10.

Na podstawie tego samego zbioru danych oszacowano dwa uogólnione modele liniowe M1 i M2. Wiadomo, że w modelu M2 uwzględniono wszystkie zmienne objaśniające z modelu M1 i dodano jedną zmienną jakościową, przyjmującą trzy wartości (kategorie): K1, K2 i K3. Dla modelu M1 wartość kryterium informacyjnego AIC wynosi: 22,548, a dla M2: 25,833. Sprawdzić na poziomie istotności 0,05, czy model M2 ma lepszą zdolność predykcyjną w porównaniu z modelem M1 (tzn. czy dodanie do M1 dodatkowej zmiennej istotnie wpływa na polepszenie zdolności predykcyjnej).

Odpowiedzi:**Odp.**

Stosując test oparty na ilorazie wiarygodności, na poziomie istotności 0,05, nie ma podstaw do odrzucenia hipotezy, że model M1 jest tak samo dobry jak M2. Zatem można przyjąć, że M2 nie ma lepszych zdolności predykcyjnych w porównaniu z M1.

Rozwiązanie:

Statystyka testowa: $T = 2(l_2 - l_1)$, gdzie l_2, l_1 to logarytmy wiarygodności odpowiednio modelu M2 i M1.

Korzystając z kryterium informacyjnego AIC:

$$M1: AIC_1 = 2(-l_1 + k_1)$$

$$M2: AIC_2 = 2(-l_2 + k_1 + 2) .$$

Po przekształceniach:

$$\begin{cases} AIC_1 = -2l_1 + 2k_1 \\ AIC_2 = -2l_2 + 2k_1 + 4 \end{cases}$$

skąd

$$2(l_2 - l_1) = AIC_1 - AIC_2 + 4.$$

Zatem

$$T = 22,548 - 25,833 + 4 = 0,715.$$

Wartość krytyczna:

$$\chi_{0,05;2}^2 = 5,991$$

Egzamin dla Aktuariuszy
Sesja egzaminacyjna w dniu 27 marca 2018r.

Modelowanie

Arkusz ocen

Zadanie nr	Punktacja
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	