Analiza przeżycia. Zagrożenie tlenową toksycznością ośrodkową cz. 3

Ryszard Kłos

Akademia Marynarki Wojennej Zakład Technologii Prac Podwodnych w Gdyni

STRESZCZENIE

Analiza przeżycia zajmuje się modelowaniem statystycznym przebiegów czasowych, dla których interesujące są czasy trwania pomiędzy wybraną chwilą a oczekiwanym zdarzeniem. Zdarzenie to nazywane jest rezultatem lub punktem końcowym.
Dane do analizy przeżycia mogą być także traktowane, jako czas do zaistnienia zdarzenia, czas przeżycia, czas do awarii, czas niezawodności, czas trwania itp. Analiza takich danych jest ważna zarówno w medycynie [1], naukach społecznych [2] jak i w inżynierii [3].
Analiza przeżycia znalazła tez zastosowanie w technice nurkowej [1]. W artykule zostały przedstawione podstawy analizy przeżycia, które posłużą do szacowania możliwości wystąpienia symptomów tlenowej toksyczności ośrodkowej przedstawione w części czwartej cyklu artykułów.
Słowa kluczowe: analiza przeżycia, szacowanie zagrożenia.

ARTICLE INFO

PolHypRes 2014 Vol. 48 Issue 3 pp. 33 – 48 **ISSN:** 1734-7009 **eISSN:** 2084-0535
DOI: HTTP://DX.DOI.ORG/10.13006/PHR.48.3
Strony: 16, rysunki: 4, tabele: 3. **page www of the periodical:** www.phr.net.pl

Typ artykułu: przeglądowy

Termin nadesłania: 21.05.2014 r
Termin zatwierdzenia do druku: 22.06.2014 r

Publisher
Polish Hyperbaric Medicine and Technology Society

ODPOWIEDŹ SYSTEMU

Przedział czasu pomiędzy wybranym punktem początkowym a zaistnieniem oczekiwanego zdarzenia można traktować, jako zmienną losową ***T*** będącą odpowiedzią systemu – nazywaną czasem przeżycia [4]. Należy zauważyć, że punkt początkowy powinien być dokładnie zdefiniowany, gdyż istnieje przeważnie kilka możliwości jego określenia [5].

Czas jest zmienną ciągłą, dlatego czas przeżycia***T*** najczęściej jest także traktowany, jako ciągła zmienna losowa. W praktyce jest on jednak odnotowywany z dokładnością do pewnego okresu [6] i często jest on wyrażany w skali dyskretnej[7].

Analiza przeżycia opiera się na wnioskowaniu statystycznym dotyczącym dystrybuanty ***F,*** czasu przeżycia ***T*** i najczęściej dotyczy prostej jej estymacji bazującej na pojedynczej homogenicznej próbie losowej, porównaniu czasu przeżycia ***T*** pomiędzy dwoma próbami[8] czy modelowaniu dystrybuanty ***F,*** jako potencjalnej funkcji kilku zmiennych objaśniających. Zagadnienia te nie różnią się od typowego wnioskowania i modelowania statystycznego, lecz przesłanki do specjalnego potraktowania analizy przeżycia[9] są następujące:

- dane do analizy przeżycia często są ucięte[10]
- standardowe rozkłady zmiennej losowej[11] często nie są adekwatne do modelowania dystrybuanty ***F,*** czasu przeżycia ***T***

DANE UCIĘTE

Badając przykładową sytuację problemową dotyczącą oceny reakcji pacjenta na zastosowany sposób leczenia, najczęściej ustala się punkt początkowy, jako włączenie pacjenta do grupy po jego przyjęciu na leczenie szpitalne. Następnie określany jest czas ***T*** do zaistnienia interesującego zjawiska – np. śmierci pacjenta[12].

W praktyce spotyka się dane pełne[13] oraz ucięte. Dla pacjentów, którzy nie umarli wiadomo, że ich czas przeżycia ***T*** musi być większy od czasu prowadzenia obserwacji t Ten rodzaj danych znany jest, jako ucięte z prawej strony – prawdziwa wartość czasu przeżycia ***T*** leży na prawo[14]. Ten typ danych jest najczęściej spotykanym wśród danych uciętych i może powstać z wielu powodów[15].

Innym typem niepełnych danych są przedziałowo lub lewostronnie ucięte. Dla danych przedziałowo uciętych czas przeżycia ***T*** nie jest dokładnie znany, lecz wiadomo, że zawiera się w określonym przedziale czasowym $T \in (t_1; t_2)$. Znaczy to, że nie zaobserwowano oczekiwanego zdarzenia dla czasu t_1, lecz nastąpiło ono przed upływem czasu t_2. Dlatego o czasie ***T*** można jedynie powiedzieć, że jest zawarty w przedziale $(t_1; t_2)$.

Dane tego typu pojawiają się często w badaniach socjologicznych prowadzonych dla konkretnego przedziału czasowego. Jeśli dla danych przedziałowo uciętych wartość początkowa wynosi zero $t_1 = 0$, to ten typ danych znany jest, jako lewostronnie ucięte. Gdy badania polegają na określeniu wystąpienia konkretnego wydarzenia w życiu człowieka to są one zawsze lewo bądź prawostronnie ucięte.

Ważną cechą odróżniającą analizę przeżycia od wielu innych metod statystyki matematycznej jest to, że dane ucięte mogą być wykorzystane i mogą nieść ze sobą istotną informację o naturze zjawiska. Nie jest to jednak regułą i należy zachować w tym względzie daleko idącą ostrożność. Jak wspomniano wcześniej, jeśli zdefiniuje się czas t, jako moment końcowy badań, to dla $T > t$ dane będą ucięte a dla $T < t$ będą pełne.

Gdy określa się czas t podczas procesu badawczego, lub decyduje wcześniej, że badania zostaną przerwane, gdy zaistnieje odpowiednia liczba zdarzeń oczekiwanych, to taki mechanizm ucinania nie wnosi istotnych danych do analizy przeżycia[16] i nie mogą być one włączone do analizy, choć czas przeżycia ***T*** nie jest w tym przypadku całkowicie niezależny od czasu ucięcia t [17]. Istotne dla analizy przeżycia mechanizmy ucinania danych pojawiają się, jeżeli są funkcjonalnie połączone z czasem przeżycia, np. wycofanie pacjenta z badań z powodu wpływającego na czas przeżycia[18].

Istotnymi są także mechanizmy ucinania, jeśli reakcja pacjenta na zastosowaną kurację jest negatywna. Nie można wykluczyć danych uciętych w analizie przeżycia, gdyż może to spowodować potencjalne poważne odchylenia przy wnioskowaniu[19], lecz obecność niektórych typów takich danych oznacza konieczność wykorzystania specjalnych metod analitycznych. Okazjonalnie można analizować sytuację, gdy ani dane pełne ani ucięte nie są możliwe do zarejestrowania.

Z taką sytuacją można mieć do czynienia, gdy trzeba podjąć wnioskowanie o czasie do uszkodzenia maszyny, która została zatrzymana na skutek wygaśnięcia certyfikatu pozwalającego na jej dalszą eksploatację lub została skierowana do obowiązkowego remontu. Takie dane nazywane są odciętymi i wymagają specjalnych metod analizy, które nie będą tutaj omawiane.

DYSTRYBUANTA CZASU PRZEŻYCIA

Dystrybuanta F czasu przeżycia, jako zmiennej losowej T, powinna być ciągła i dodatnio określona. Warunki te spełnia przykładowo dystrybuanta $rozkładu\ gamma\ \Gamma$ [20]:

Tab. 1

Uogólniony rozkład Gamma.

$$\forall_{x,a,b,c>0} L = \frac{a}{b^{c/a} \cdot \Gamma\left(\frac{c}{a}\right)} \cdot X^{c-1} \cdot exp\left(\frac{-X^a}{b}\right)$$

a	c	Distribution	Distribution density $L(X)$
1	1	Exponential	$\forall_{x,b>0} L = \frac{1}{b} \cdot exp\left(\frac{-X}{b}\right)$
1	c	Gamma	$\forall_{x,b,c>0} L = \frac{1}{b^c \cdot \Gamma(c)} \cdot X^{c-1} \cdot exp\left(\frac{-X}{b}\right)$
2	2	Reyleigh	$\forall_{x,b>0} L = \frac{2}{b} \cdot X \cdot exp\left(\frac{-X^2}{b}\right)$
a	a	Weibull	$\forall_{x,a,b>0} L = \frac{a}{b^a} \cdot X^{a-1} \cdot exp\left[-\left(\frac{X}{b}\right)^a\right]$
2	3	Maxwell	$\forall_{x,b>0} L = \frac{4}{\sqrt{\pi} \cdot b^{3/4}} \cdot X^2 \cdot exp\left(\frac{-X^2}{b}\right)$

$$\Gamma(a) = \int_0^\infty X^{a-1} \cdot e^{-X} dX; b = \frac{1}{f}$$

$$\forall_{\Gamma(a)=\int_0^\infty X^{a-1} \cdot e^{-X} dX} \quad F(t) = \frac{f^u}{\Gamma(a)} \cdot t^{a-1} \cdot \exp(-f \cdot t) \qquad (1)$$

gdzie: $F(t)$ —dystrybuanta, $L(t)$ —gęstość rozkładu prawdopodobieństwa, f —częstość, t —czas

Częściej wykorzystywane w analizie przeżycia rozkłady zebrano w tab. 2. Najbardziej popularnym w modelowaniu zależności w analizie przeżycia jest rozkład $Weibulla$, dla którego gęstość prawdopodobieństwa $L(t)$ można zapisać w postaci:

$$L(t) = a \cdot f^a \cdot t^{a-1} \cdot \exp(-f \cdot t)^a \qquad (2)$$

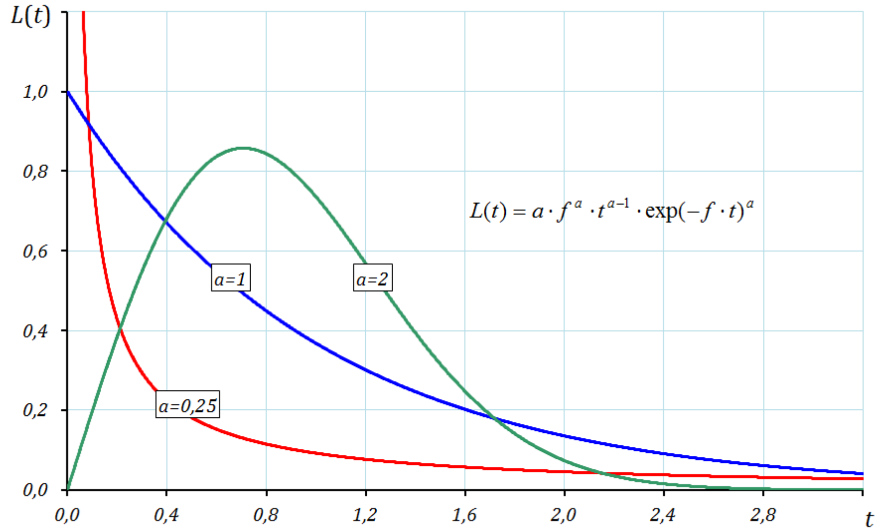
gdzie: a —stała

Parametr a z równania (2) odpowiedzialny jest za kształt zaś częstość f za skalę gęstości prawdopodobieństwa dla rozkładu $Weibulla$ – rys. 1. Średnia $\bar{x}$ dla rozkładu $Weibulla$ wynosi:

Tab. 2

Niektóre rozkłady statystyczne wykorzystywane w analizie przeżycia [2].

Distribution	Distribution function	Density	Mean	Variance	Survival function	Hazard function	Cumulative hazard function
	$F(X)$	$L(X)$	$\acute{X}$	σ^2	$S(X) = 1 - F(X)$	$h(X) = \frac{L(X)}{S(X)}$	$H(X) = \int_0^t h(X) \cdot dt$
exponential	$1 - exp\left(\frac{-X}{b}\right)$	$\frac{1}{b} \cdot exp\left(\frac{-X}{b}\right)$	b	b^2	$exp\left(\frac{-X}{b}\right)$	$\frac{1}{b}$	$\frac{X}{b}$
logistic	$1 - \frac{1}{1+exp\left(\frac{-X-a}{b}\right)}$	$\frac{exp\left(\frac{-X-a}{b}\right)}{b\cdot\left[1-exp\left(\frac{-X-a}{b}\right)\right]^2}$	a	$\frac{\pi^2 \cdot b^2}{3}$	$\frac{1}{1+exp\left(\frac{X-a}{b}\right)}$	$\frac{1}{1+exp\left(\frac{-X-a}{b}\right)}$	$log\left[1 + exp\left(\frac{X-a}{b}\right)\right]$
Ryleigh	$1 - exp\left(\frac{-X^2}{2 \cdot b^2}\right)$	$\frac{X}{b^2} \cdot exp\left(\frac{-X^2}{2 \cdot b^2}\right)$	$b \cdot \sqrt{\frac{\pi}{2}}$	$b^2 \cdot \left(2 - \frac{\pi}{2}\right)$	$exp\left(\frac{-X^2}{2 \cdot b^2}\right)$	$\frac{X}{b^2}$	$\frac{X^2}{2 \cdot b^2}$
Weibull	$1 - exp\left[-\left(\frac{X}{b}\right)^a\right]$	$\frac{a \cdot X^{a-1}}{b^a} \cdot exp\left[-\left(\frac{X}{b}\right)^a\right]$	$b \cdot \Gamma\left(\frac{a+1}{a}\right)$	$b^2 \cdot \left[\Gamma\left(\frac{a+1}{a}\right) - \Gamma^2\left(\frac{a+1}{a}\right)\right]$	$exp\left[-\left(\frac{X}{b}\right)^a\right]$	$\frac{a \cdot X^{a-1}}{b^a}$	$\left(\frac{X}{b}\right)^a$
$\Gamma(a) = \int_0^\infty X^{a-1} \cdot e^{-X} dX; b = \frac{1}{f}$							



Rys. 1. Gęstość rozkładu prawdopodobieństwa $L(t)$ dla rozkładu $Weibulla$ dla różnych wartości parametru a i przy częstości $f = 1$.

$$\bar{x} = \frac{1}{f} \cdot \Gamma\left(1 + \frac{1}{a}\right) \tag{3}$$

a wariancja σ^2:

$$\sigma^2 = \frac{1}{f^2} \cdot \left[\Gamma\left(1 + \frac{2}{a}\right) - \Gamma^2\left(1 + \frac{1}{a}\right)\right] \tag{4}$$

Dla $a = 1$, gęstość L rozkładu $Weibulla$ przybiera formę funkcji eksponencjalnej ze średnią: $\bar{x} = \frac{1}{f}$.

ROZKŁAD WYKŁADNICZY

Dla rozkładu binominalnego, prawdopodobieństwo P zdarzenia dla braku $n = 0$ przypadku niepożądanego [21] przy N zdarzeniach[22] można zapisać, jako[23]: $\forall_{n=0}\ P(n = 0, N) = (1 - R)^N$ gdzie R oznacza prawdopodobieństwo będące ryzykiem wystąpienia niekorzystnego zjawiska. Chcąc wyrazić prawdopodobieństwo w funkcji czasu $P(n = 0, N) = f(t)$ można podzielić okres obserwacji t na szereg przedziałów Δt. Przyjmując, że wartość zagrożenia $\boldsymbol{R}$ [24] nie ulega zmianie $R(t) = idem$ dla całego czasu trwania obserwacji t, to przy dążącej do nieskończoności liczbie $N \to \infty$ przedziełów Δt, prawdopodobieństwo $P(n=0,N)$ będzie wynosiło zero [25]: $\forall_{\rho\in(0;1]}\ lim_{\Delta t\to 0}P(n=0,N)=$ $=lim_{N\to\infty}(1-R)^n\equiv 0.$

Na mocy tego, że ryzyko R jest niezmienne $R=idem$, to częstotliwość f występowania niekorzystnego zjawiska jest także niezmienna: $f=idem$. Zgodnie z częstościową definicją prawdopodobieństwa i wcześniejszym założeniem niezmienności ryzyka $R=idem$, można zapisać (3):

$$R=idem=\frac{ES_N}{N}\ \rightarrow\ R=\frac{f\cdot t}{N}=f\cdot\Delta t \qquad (5)$$

gdzie: R – ryzyko, ES_N– wartość oczekiwana E przeciętnej liczby przypadków S_N przy liczności populacji N, Δt –przedział czasowy stanowiący dokładność, z jaką liczy się liczy się upływ czasu, N –liczność populacji

Dyskretną zmienną zależną ES_N spodziewanej przeciętnej liczby przypadków n w funkcji dyskretnej zmiennej niezależnej liczby obserwacji N, można zastąpić ciągłą zmienną zależną ryzyka R w funkcji zmiennej niezależnej liczby N równych przedziałów Δt, gdzie współczynnikiem proporcjonalności będzie tutaj częstość f definiowana zgodnie z (**5**):

$$\frac{ES_N}{N}=f\cdot\frac{t}{N}\ \rightarrow\ f\equiv\frac{ES_N}{t}\Rightarrow\forall_{\Delta t=\frac{t}{N}=const}\ R=f\cdot\frac{t}{N} \qquad (6)$$

Podstawiając równanie (**6**) określające ryzyko R do zależności $\forall_{\rho\in(0;1]}\ \lim_{\Delta t\to 0}P(n=0,N)=\lim_{N\to\infty}(1-R)^N\equiv 0$ i korzystając z definicji funkcji eksponencjalnej[26] $\forall_{k\in\mathbb{N}}\ \lim_{k\to\infty}(1+\frac{-a}{k})^k\stackrel{\text{def}}{=}\exp(-a)$, można otrzymać:

$$\forall_{\rho\in(0;1]}\ \lim_{\Delta t\to 0}P(n=0,N)=\lim_{N\to\infty}(1-f\cdot\frac{t}{N})^N\stackrel{\text{def}}{=}\exp(-f\cdot t) \qquad (7)$$

gdzie: P –prawdopodobieństwo, n–liczba zdarzeń w próbie o liczności N

Dystrybuantę $F(t,f)$ dla prawdopodobieństwa (**7**), można zapisać korzystając z definicji prawdopodobieństwa odwrotnego:

$$\forall_{f>0}\ F(t,f)=P(0\le T\le t\mid f)=1-\exp(-f\cdot t) \qquad (8)$$

a gęstość L rozkładu wykładniczego można będzie znaleźć różniczkując dystrybuantę F z zależności (8): $\forall_{f>0}\ L(t,f)=\frac{dF(t,f)}{dt}=f\cdot\exp(-f\cdot t)$. Całkując gęstość L w granicach od zera do nieskończoności można pokazać, że rozkład wykładniczy jest unormowany[27].

ROZKŁAD WEIBULLA

Prawdopodobieństwo dodatkowego przeżycia czasu Δt, gdy do chwili obecnej czas życia wynosi t jest prawdopodobieństwem warunkowym:

$$P(\Delta t\mid t)=\frac{P(\Delta t\cap t)}{P(t)} \qquad (9)$$

Licznik ułamka z równania (**9**) to prawdopodobieństwo przeżycia czasu łącznego $t+\Delta t$, dlatego zależność (**9**) można przepisać do postaci:

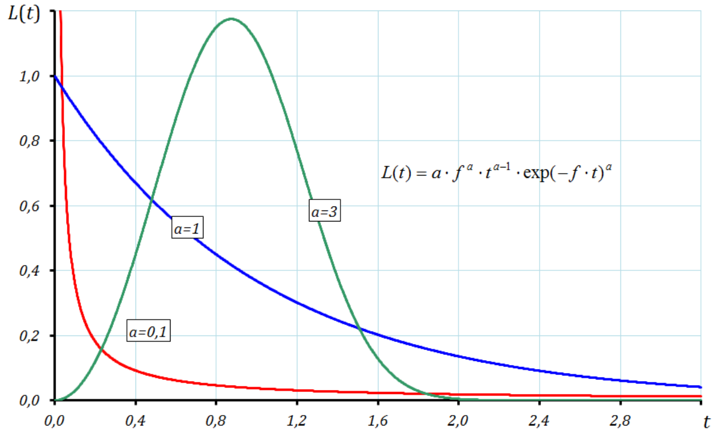
$$P(\Delta t|t)=\frac{P(\Delta t+t)}{P(t)}=\frac{\exp[-f(\Delta t+t)]}{\exp(-f\cdot t)}=\exp(-f\cdot\Delta t) \qquad (10)$$

Z zależności (10) wynika, że dla rozkładu wykładniczego prawdopodobieństwo warunkowe przeżycia dodatkowego czasu Δt [28] nie jest funkcją dotychczasowego czasu t $P(\Delta t|t)\neq f(t)$przeżycia.

Określa się to, jako niezależność rozkładu wykładniczego od obecnego wieku [29]: $\forall_{P=exp(-f\cdot\Delta t)}\ P(\Delta t+t)=P(t)\cdot P(\Delta t)$, lecz właściwość ta stoi w sprzeczności z doświadczeniem [30].

Modyfikując sposób podejścia przy wyprowadzeniu formuły na gęstość rozkładu prawdopodobieństwa L można poprawić właściwości rozkładu wykładniczego $\forall_{f>0}\ L(t,f)=f\cdot\exp(-f\cdot t)$ zakładając, że prawdopodobieństwo wystąpienia zdarzenia dla pojedynczej próby Bernoulliego [31] jest proporcjonalne do czasu trwania próby Δt i jest niezależne od numeru tej próby – rys. 2.

Przyjmując, że prawdopodobieństwo wystąpienia zdarzenia w i–tej próbie można zapisać, jako: $\forall_{f_i<f_{i+1};i>1}\ P_i=f_i\cdot\Delta t$, gdzie częstotliwość f_i występowania zdarzenia ulega zwiększeniu $f_i<f_{i+1}$ wraz ze starzeniem się organizmu czy urządzenia. Stąd przyjmując niezależność prób, prawdopodobieństwo braku $n=0$ zdarzeń niekorzystnych w N próbach będzie wyrażało się iloczynem:



Rys. 2. Gęstość rozkładu prawdopodobieństwa $L(t)$ rozkładu $Weibulla$ dla różnych wartości parametru a i przy częstości $f = 1$.

$$P(\Delta t \mid t) = \frac{P(\Delta t \cap t)}{P(t)} \qquad (11)$$

Podobnie jak poprzednio można obliczyć granicę dla $N \to \infty \Leftrightarrow \Delta t \to 0$. Wygodnie jest uczynić to po uprzednim zlogarytmowaniu wyrażenia (11): $\ln P(n = 0, N) = \sum_{i=1}^{N} \ln(1 - f_i \cdot \Delta t)$, skąd prawdopodo-bieństwo to można przyjąć w przybliżeniu, jako: $\forall_{\Delta t \to 0} \ln P(n = 0, N) \cong -\sum_{i=1}^{N} f_i \cdot \Delta t$. Dla granicznie małych Δt można policzyć granicę: $-\lim_{\substack{N \to \infty \\ \Delta t \to 0}} \sum_{i=1}^{N} f_i \cdot \Delta t = -\int_0^t f(t)dt \equiv -\Lambda(t)$.

Można zatem zapisać, że $\ln P(n = 0, N) = -\int_0^t f(t)dt$ a stąd przekształcając, można otrzymać zależność na prawdopodobieństwo $P(n = 0, N)$ braku niekorzystnych przypadków w funkcji czasu t: $P(n = 0, N) = \exp[-\int_0^t f(t)dt]$. Stąd podobnie jak dla zależności (**8**) dystrybuantę $F(t)$ można zapisać korzystając z definicji prawdopodobieństwa odwrotnego: $F(t, f) = 1 - \exp[-\int_0^t f(t)dt]$. Gęstość L rozkładu w tym przypadku wyniesie:

$$\forall_{f>0} \; L(t, f) = \frac{dF(t)}{dt} = f(t) \cdot exp[-\int_0^t f(t)dt] \equiv f(t) \cdot exp[-\Lambda(t)] \qquad (12)$$

W szczególności, dla gęstości rozkładu prawdopodobieństwa L z formuły (12), gdy częstość $f = \frac{n}{N}$ występowania niekorzystnego zjawiska nie jest funkcją czasu $f \neq f(t)$, to zależność (12) wyraża gęstość dla rozkładu wykładniczego.
Dla $\Lambda(t) \triangleq t^{\alpha}$ zależność (12) wyraża gęstość rozkładu $Weibulla$ [4]:

$$\forall_{f>0} \; L(t, f) = f(t) \cdot exp[-t^{\alpha}] \qquad (13)$$

GĘSTOŚĆ ROZKŁADU CZASU PRZEŻYCIA

Dla ciągłej zmiennej losowej T oznaczającej czas przeżycia, z funkcji gęstości rozkładu prawdopodobieństwa L można określić rozkład prawdopodobieństwa wystąpienia czasu T w przedziale $(t_1; t_2)$:

$P(t_1 \leq T \leq t_2) = \int_{t_1}^{t_2} L(t)dt$. Dystrybuanta F czasu T określona jest formułą:

$$F(t) = P(T \leq t) = \int_0^t L(t)dt \qquad (14)$$

W analizie przeżycia często preferowane jest wykorzystanie trzech alternatywnych funkcji prawdopodobieństwa definiujących rozkład zmiennej losowej T: $funkcji\ przeżycia\ S(t)$, $hazardu\ h(t)$ i $skumulowanego\ hazardu\ H(t)$.

FUNKCJA PRZEŻYCIA

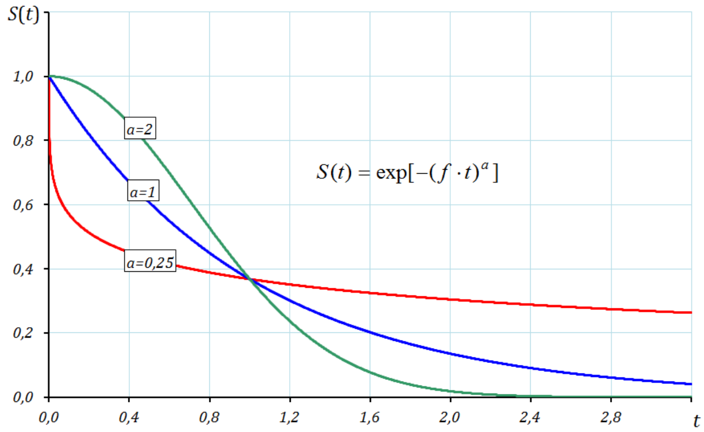
Funkcja przeżycia $S(t)$ określa prawdopodobieństwo, przetrwania[32] dłużej niż pewien przeciętny czas t[33]:

$$\forall_{t \geq 0} \; S(t) \stackrel{\text{def}}{=} P(T > t) = 1 - F(t) \qquad (15)$$

gdzie: $S(t)$ —funkcja przeżycia

Funkcja przeżycia[34] $S(t)$ jest nierosnącą funkcją ciągłą, dla której $S(0) = 1$. Dla rozkładu $Weibulla$ funkcja przeżycia $S(t)$ dana jest zależnością: $S(t) == \exp[-(f \cdot t)^{\alpha}]$.
Na rys. 3 pokazano funkcje przeżycia $S(t)$ dla rozkładu $Weibulla$ dla różnych wartości parametrów a i częstości $f = 1$.

Wartość oczekiwana czasu przeżycia ET powiązana jest z funkcją przeżycia $S(t)$ następująca formułą: $ET = \int_0^{\infty} S(t)dt$, stąd wartość ta reprezentowana jest przez pole pod funkcją przeżycia $S(t)$.



Rys. 3. Funkcja przeżycia $S(t)$ dla rozkładu $Weibulla$ dla różnych wartości parametru a i częstości $f = 1$.

FUNKCJA HAZARDU

Zgodnie z definicją prawdopodobieństwa warunkowego można zapisać:

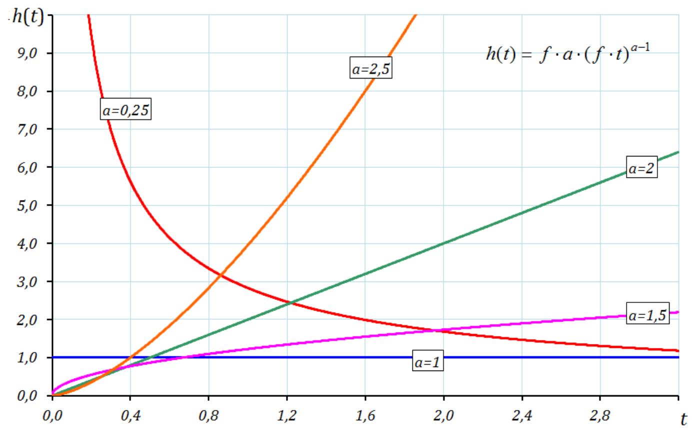
$$P(T \leq t + \Delta t \mid T > t) \equiv \frac{P(t<T\leq t+\Delta t)}{P(T>t)} = \frac{P(T>t \cap T \leq t+\Delta t)}{P(T>t)} = \frac{L(t) \cdot \Delta t}{S(t)} = h(t) \cdot \Delta t \qquad (16)$$

gdzie: $h(t)$ –funkcja hazardu

W zależności (16) funkcja $h(t)$ określana jest, jako *funkcja hazardu*: $h(t) = \frac{L(t)}{S(t)}$. Wychodząc z zależności (16) można przedstawić funkcję hazardu $h(t)$, jako granicę prawdopodobieństwa warunkowego na jednostkę czasu:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(T \leq t+\Delta t \mid T>t)}{\Delta t} \qquad (17)$$

Funkcja hazardu *h(t)* reprezentuje prawdopodobieństwo tego, że czas przeżycia $\boldsymbol{T}$ będzie znajdować się w otoczeniu wybranego czasu t, lecz nie pojawi się przed tym czasem[35]. Opisuje ona intensywność niepowodzeń dla wybranego czasu[36] t.



Rys. 4. Funkcja hazardu $h(t)$ dla rozkładu $Weibulla$ dla różnych wartości parametru a i przy częstości $f = 1$.

Funkcja hazardu $h(t)$podaje wartość prawdopodobieństwa na jednostkę czasu (17) może więc wystąpić sytuacja[37] dla której jej wartość będzie większa od . Dla rozkładu1 $Weibulla$ funkcja ta wyraża się równaniem: $h(t) = f \cdot a \cdot (f \cdot t)^{a-1}$. Na rys. 4 pokazano wybrane kształty funkcji hazardu $h(t)$ dla rozkładu $Weibulla$, dla różnych wartości parametru a przy częstości $f = 1$.

SKUMULOWANA FUNKCJA HAZARDU

Skumulowaną funkcję hazardu $H(t)$ można zdefiniować, jako całkę z funkcji hazardu $h(t)$:

$$H(t) = \int_0^t h(t)dt \tag{18}$$

gdzie: $H(t)$ —skumulowana funkcja hazardu

Dla rozkładu $Weibulla$ wyrazi się one zależnością: $H(t) = (f \cdot t)^{\alpha}$

FUNKCJA RYZYKA

W technice, funkcja hazardu $h(t)$ nazywana jest czasami *funkcją ryzyka* $R(t)$ lub *intensywności uszkodzeń* $\lambda(t)$ i definiowana jest, jako iloraz gęstości prawdopodobieństwa czasu pracy elementu $L(t)$ w punkcie t przez prawdopodobieństwo, dla którego czas pracy elementu jest co najmniej równy t:

$$\forall_{t \geq 0;\, S(t) \geq 0;\, \frac{dF}{dt} = \dot{F}(t) \equiv L(t)} \quad h(t) \equiv R(t) \stackrel{\text{def}}{=} \frac{L(t)}{S(t)} \equiv \frac{\dot{F}(t)}{S(t)} \tag{19}$$

gdzie: $R(t)$ —funkcja ryzyka

Dla dyskretnego rozkładu prawdopodobieństwa intensywności uszkodzeń $R(t)$ można zapisać zależność: $\exists_{P_k = P(k \leftarrow X)}\ R(t) = \frac{P_k}{\sum_{i=k}^{\infty} P_i}$. Różniczkując funkcję przeżycia $\forall_{t \geq 0}\ S(t) = P(T > t) = 1 - F(t)$ można pokazać ciekawą zależność: $\dot{F}(t) = \frac{d[1 - S(t)]}{dt} = -\dot{S}(t)$, co wobec zależności (**19**) daje: $\forall_{S(t)>0}\ R(t) = -\frac{\dot{S}(t)}{S(t)} =$ $= -\frac{1}{S(t)} \cdot \frac{dS(t)}{dt} \Rightarrow \int_0^t R(t)dt = -\int_0^{S(t)} \frac{dS(t)}{S(t)} = -\ln|S(t)| = -\ln S(t)$ ·

Co w konsekwencji skutkuje zależnościami:

$$\forall_{S(t)>0}\ S(t) = exp\left[-\int_0^t R(t)dt\right] = \exp[-H(t)] \tag{20}$$

Równanie (20) nosi nazwę *związku Wienera*. Funkcja skumulowanego hazardu $H(t)$ nazywana jest w technice bezpieczeństwa *funkcją rozkładu zawodności bezpieczeństwa.*

ZALEŻNOŚCI MIĘDZYFUNKCYJNE

Podsumowując przegląd podstawowych funkcji stosowanych w analizie przeżycia można stwierdzić, że wystarcza podać jedną z nich aby opisać inne, gdyż są one pomiędzy sobą powiązane zależnościami zebranymi w tab. W analizie przeżycia najczęściej operuje się funkcjami przeżycia ***S (t)*** i hazardu ***h (t)*** [38]

Z praktyki wiadomo, że rzetelnego oszacowania specyficznych funkcji analizy przeżycia można dokonać, jeśli liczność próby do wnioskowania jest większa niż $N = 30$, w przeciwnym przypadku wyniki estymacji są obciążone.

Tab. 3

Użyteczne zależności dla częściej spotykanych funkcji w analizie przeżycia.

$$S(t) = 1 - F(t) = \int_t^{\infty} L(t)dt$$

$$L(t) = \frac{-d}{dt} \cdot S(t)$$

$$R(t) \equiv h(t) = \frac{-d}{dt} \cdot lnS(t)$$

$$H(t) = -lnS(t) = \int_0^t R(t) \cdot dt \equiv \int_0^t h(t) \cdot dt$$

$$S(t) = exp[-H(t)] = exp\left[-\int_0^t h(t) \cdot dt\right]$$

F —distribution function	L —probability density
h —hazard function	R —risk function
H —cumulative hazard function	S —survival function

Pomimo zbieżności formuł do obliczania różnych charakterystycznych parametrów rozpatrywanych tutaj przebiegów czasowych istnieją różnice ich interpretacji pomiędzy teorią niezawodności, analizą bezpieczeństwa czy w innych zastosowaniach analizy przeżycia[39].

ZAGROŻENIE

Funkcja ryzyka $R(t)$ może reprezentować prawdopodobieństwo wystąpienia objawów $CNSyn$ lub DCS w funkcji czasu t. Korzystając z zależności $\forall_{t \geq 0}\ S(t) = P(T > t) = 1 - F(t)$ i $\forall_{S(t)>0}\ S(t) = exp\left[-\int_0^t R(t)dt\right]$ można dystrybuantę

$F(t)$ prawdopodobieństwa wystąpienia objawów DCS czy $CNSyn$ w funkcji czasu wyrazić poprzez wartość funkcji ryzyka $R(t)$:

$$\forall_{\xi(t)\triangleq F(t)} \quad F(t) \equiv 1 - S(t) = 1 - exp\left[-\int_0^t R(t) \cdot dt\right] \qquad (21)$$

gdzie: $\xi(t) \triangleq F(t)$ – funkcja zagrożenia DCS lub $CNSyn$ tożsamościowo równa dystrybuancie czasu przeżycia $F(t)$

Całka funkcji ryzyka $R(t)$ od momentu $t = 0$ do t określa integralne ryzyko wystąpienia w tym okresie czasu przypadku DCS lub $CNSyn$. Czyli wartość funkcji ryzyka $R(t)$ z wyrażenia (21) określa wartość funkcja zagrożenia $\xi(t)$ wystąpieniem objawów DCS lub $CNSyn$. Wartości parametrów funkcji ryzyka $R(t)$ można określić, poprzez dopasowanie jej do danych eksperymentalnych. Granice całkowania mogą rozciągać się także na kilka godzin po zakończeniu nurkowania.

Przy wykorzystaniu analizy przeżycia do modelowania matematycznego funkcji ryzyka $R(t)$ i zagrożenia $F(t)$ wystąpieniem objawów DCS i $CNSyn$ należy rozróżniać te pojęcia. Ryzyko $R(t)$ jest tutaj utożsamiane z prawdopodobieństwem wystąpienia DCS lub $CNSyn$, zaś zagrożenie wystąpieniem DCS jest utożsamiane z dopełnieniem funkcji przeżycia $S(t)$ będące dystrybuantą czasu przeżycia: $F(t) = 1 - S(t)$. Zagrożenie $F(t)$ jest prawdopodobieństwem wystąpienia DCS lub $CNSyn$ pod warunkiem zaakceptowania poziomu ryzyka $R(t)$ wystąpieniem DCS lub $CNSyn$.

PODSUMOWANIE

Metody analizy przeżycia zostały wprowadzone do problematyki nurkowej przez Weathersby'ego i Thalmanna [1]. Zaproponowany przez $US\ Navy$ model predykcji zagrożenia $F(t)$ możliwością wystąpienia $CNSyn$ pochodzący z tej teorii wydaje się być dostatecznie precyzyjny. Będzie on przedstawiony w następnym artykule tej serii.

Artykuł jest trzecim z serii zawierających wyniki badań przeprowadzonych w Akademii Marynarki Wojennej w Gdyni a finansowanych ze środków na naukę w latach 2009 – 2011 w ramach projektu rozwojowego Nr O R00 0001 08 p.t.: Projektowanie dekompresji w misjach bojowych.

BIBLIOGRAFIA

1. Gerth W.A. Overview of survival functions and methodoology . [author of the book] Gerth W.A. Weathersby P.K. *Survival analysis and maximum likelihood techniques as applied to physiological modeling.* Kesington : Undersea and Hyperbaric Medical Society Inc., 2002
2. Evans M., Hastings N., Peacock B. *Statistical Distributions.* New York: John Willey & Sons, Inc. , 1993. ISBN 0-47 1-55951-2
3. Kłos R. *Application of statistical method in diving technique - Skrypt.* Gdynia: Polish Society of Hyperbaric Medicine and Technique, 2007b. ISBN 978-83-924989-26
4. Nowak R. *Statistics for physicists.* Warsaw: Wydawnictwo Naukowe PWN, 2002. ISBN 83-01-13702-9
5. Jaźwiński J. i Ważyńska-Fiok K. *System security.* Warsaw: Wydawnictwo Naukowe PWN, 1993

dr hab. inż. Ryszard Kłos, prof. nadzw. AMW
Akademia Marynarki Wojennej im. Bohaterów Westerplatte
Zakład Technologii Prac Podwodnych
81 – 103 Gdynia 3, ul. Śmidowicza 69
TEL.: +58 626 27 46, FAX.: +58 626 27 61

[1] np. analiza czasu od rozpoczęcia kuracji do nawrotu choroby, śmierci itp.,

[2] np. czas pozostawania bezrobotnym, wiek pierworództwa itp.,

[3] np. czas do uszkodzenia elementu wyposażenia, czas bezawaryjnej pracy itd.,

[4] zależnie od rozpatrywanego systemu może być nazywany czasem bezawaryjnej pracy, trwania, oczekiwania czy odpowiedzi,

[5] przykładowo, przy określeniu przeżywalności ataku serca punkt początkowy może być odniesiony do czasu pojawienia się objawów, przyjęcia do hospitalizacji, rozpoczęcia określonej kuracji itp.,

[6] jak dzień, godzina, minuta itp.

[7] przykładowo, w technice niezawodności może być wyrażany w ilości wykonanych przez maszynę cykli do chwili zaistnienia awarii,

[8] np. dla dwóch alternatywnych sposobów leczenia,

[9] różnicującymi analizę przeżycia od podobnych, wspomnianych wcześniej sytuacji problemowych,

[10] dane nazywane są uciętymi, gdy nie można ich wartość dokładnie określić – będzie o tym mowa dalej,

[11] np. binominalny, normalny, F, t itp.,

[12] przy typowym wnioskowaniu statystycznym należałoby czekać do chwili śmierci pacjenta, lecz to może trwać wiele lat czy dekad, czyli czasami nierealistycznym jest oczekiwanie na zaistnienie ustalonego punktu końcowego badania,

[13] np. dla zmarłych pacjentów można odnotować czas przeżycia ***T,***

[14] tzn. czas przeżycia $\boldsymbol{T}$ jest większy niż punkt końcowy badań t - wiadomo, że czas przeżycia $\boldsymbol{T}$ leży w przedziale $(t; \infty)$,

[15] pacjent może być wycofany z programu badań, można stracić z nim kontakt po zakończeniu cyklu badań, nie ma uzasadnienia ekonomicznego lub praktycznego do dalszego monitorowania pacjenta do czasu pojawienia się zaplanowanego zdarzenia końcowego itp.,

[16] bardzo często występujące nieprawidłowości w praktyce,

[17] wystarczającym warunkiem, aby dane nie wnosiły istotnej informacji o naturze zjawiska jest fakt niezależności czasu przeżycia T od momentu końcowego t,

[18] na skutek choroby lub utraty łączności z wyleczonym pacjentem czy wskutek jego niestawienia się na umówione wizyty,

[19] podobnie, uwzględnienie danych „nieprawidłowo" uciętych może spowodować nieadekwatne wyniki wnioskowania,

[20] uogólnionego rozkładu Γ – tab.1,

[21] np. wystąpienia objawów $CNSyn$,

[22] przykładowo, przy N okresach czasowych – patrz dalej,

[23] jak to będzie pokazane dalej rozpoczęcie rozumowania od zdarzenia odwrotnego upraszcza wyprowadzenie dystrybuanty zagrożenia,

[24] np. zagrożenia wystąpieniem objawów $CNSyn$,

[25] zawsze istnieje jakieś choćby niewielkie ryzyko, stąd postulowanie wartości ryzyka na poziomie zerowym $R \equiv 0$ stoi w sprzeczności z obserwowaną rzeczywistością,

[26] wykładniczej,

[27] $\forall_{f>0} \int_0^\infty L(t,f) \cdot dt = f \cdot \int_0^\infty e^{-f\cdot t} \cdot dt = f \cdot \frac{1}{-f}\Big|_0^\infty = 0 - \frac{f}{-f} = 1$,

[28] gdy czas życia wyniósł dotąd t,

[29] pozostały czas życia Δt nie zależy od przeszłości i ma ten sam wykładniczy rozkład, co dotychczasowy czas przeżycia t,

[30] z reguły, ludzie umierają a maszyny zaczynają się psuć osiągnąwszy pewien wiek,

[31] np. podczas pojedynczego cyklu nurkowego,

[32] prawdopodobieństwo bezawaryjnej pracy, przeżycia, zajścia innego zdefiniowanego zdarzenia itp.,

[33] przykładowo, $S(t)$ to prawdopodobieństwo, że dana osoba dożyje aż do czasu t,

[34] w technice odpowiednik funkcji przeżycia $S(t)$ służy do określania niezawodności i jest nazywany *funkcją niezawodności bezpieczeństwa* [5],

[35] wartość funkcji hazardu $h(t)$ należy traktować, jako potencjał do wystąpienia spodziewanego zdarzenia (najczęściej niepowodzenia) przedstawiający dopełnienie analizy sytuacji problemowej charakteryzowanej przez funkcję przeżycia $S(t)$ – gdy funkcja $S(t)$ maleje to $h(t)$ rośnie. Funkcja $h(t)$ można obrazowo porównać do działania prędkościomierza samochodowego. Na podstawie jego stałego wskazania można wnioskować, jaki dystans zostanie przebyty po upływie wybranego czasu – na podstawie stałej wartości funkcji $h(t)$ można wnioskować o liczbie oczekiwanych zdarzeń w ciągu wybranego czasu.,

[36] w technice bezpieczeństwa funkcję hazardu $h(t)$ określa się jako *intensywność zawodności bezpieczeństwa* i oznacza często jako $\lambda_B(t)$ [5],

[37] w zależności od przyjętych jednostek czasu,

[38] jedną z ważnych przyczyn wykorzystania funkcji hazardu $h(t)$ jest fakt, że warunkowy rozkład przewidywanego przeżycia poza moment czasu t_0 może być bezpośrednio z niej obliczony dla $h(t > t_0)$,

[39] przykładowo, miarą niezawodności jest prawdopodobieństwo spełnienia wymagań dotyczących systemu w jednostce czasu, zaś miarą niebezpieczeństwa jest prawdopodobieństwo zaistnienia sytuacji niekorzystnej dla otaczającej system czasoprzestrzeni.