Porovnání síly různých testů periodicity

Daniela Jarušková, stavební fakulta Praha

1. Úvod

V časových okamžicích $t = 1,\dots,N$ pozorujeme časovou řadu $X(t)$, která je součtem deterministické periodické složky $m(t)$ a náhodného šumu $e(t)$:

$$X(t) = m(t) + e(t) .$$

Posloupnost $\{e(t)\}$ jsou nezávislé, stejně rozdělené náhodné veličiny s rozdělením $N(0,\sigma^2)$. Periodickou složku $m(t)$ lze vyjádřit jako součet kosinů s různými frekvencemi, posunutími a amplitudami :

$$m(t) = \sum_{k=1}^{R} \rho_k \cos(t\theta_k + \phi_k) .$$

Abychom mohli tuto složku z napozorovaných hodnot odhadnout, je nejprve třeba najít množinu frekvencí $\Theta = \{\theta_1,\dots,\theta_R\}$, které určují její kmitání. K vyhledávání těchto frekvencí se používá Fourierova transformace $v(\lambda) = (1/\sqrt{2\pi N}) \sum_{t=1}^{N} e^{-it\lambda} X(t)$ nebo spíše kvadrát její absolutní velikosti nazývaný periodogram $I(\lambda) = v(\lambda)v(\lambda)^*$, neboť jeho očekávaná střední hodnota nabývá velkých hodnot pro frekvence z množiny Θ a jinde je omezená.

2. Fisherův test

Nejčastějším postupem pro hledání Θ je Fisherův test. Nulová hypotéza udává, že proces vůbec neobsahuje periodickou složku, tj. $H_o : X(t) = e(t)$. Využijeme-li skutečnosti, že $I(\lambda_1),\dots, I(\lambda_n)$, kde $\lambda_p = 2\pi p/N$, $p=1,\dots,[(N-1)/2]= n$, jsou nezávislé náhodné veličiny, pro které za platnosti nulové hypotézy $4\pi I(\lambda_p)/\sigma^2 \sim \chi^2(2)$, lze jako testovou statistiku použít $\max_{i=1,\dots,n} 4\pi I(\lambda_i)/\sigma^2$. Jestliže σ^2 neznáme, lze jej odhadnout $\hat{\sigma}^2=(4\pi/N)\sum I(\lambda_p) = (1/N) \sum (X_t-\bar{X})^2$ a veličiny $4\pi I(\lambda_p)/\sigma^2$ nahradíme veličinami $Y_p = I(\lambda_p) /\sum I(\lambda_i)$. Testová statistika je pak $\max Y_p$. Její rozdělení odvodil Fisher [2]

$$P(\max Y_p > d) = \sum_{\ell=1}^{n} (-1)^{\ell-1}\binom{n}{\ell}(1- \ell d)_+^{n-1} .$$

Předpokládejme, že všechny $\theta \in \Theta$ jsou tvaru $2\pi p/N$ $(p=1,\dots,n)$, přičemž ρ_i označuje amplitudu odpovídající frekvenci $2\pi i/N$. Jestliže oproti nulové hypotéze $H_o : \rho_1=0,\dots, \rho_n=0$, stavíme jednoduché alternativy $H_i : \rho_i>0$, $\rho_j=0$, $j\neq i$, $i=1,\dots,n$, pak rozhodnutí, že platí H_i, jestliže $\max Y_p > u_P$ a $Y_i > Y_j$, $j\neq i$, a platí H_o, jestliže $\max Y_p < u_P$, kde u_P je příslušná kritická hodnota, je stejnoměrně nejsilnější symetrická invariantní rozhodovací procedura, viz Anderson [1].

3. Bølvikenův a Siegelův test

Situace je však složitější, předpokládáme-li tzv. složené alternativy, spočívající v tom, že více než jedna frekvence může mít kladnou amplitudu. V tomto případě se v součtu $\hat{\sigma}^2 = (4\pi/N)\sum I(\lambda_p)$ vyskytuje více velkých hodnot, dochází k nadhodnocení σ^2 a tím k snižování síly Fisherova testu. K odstranění tohoto nedostatku je známo několik metod. Bølviken [2] navrhl, aby se místo statistiky $\max Y_p / \sum Y_i$ použilo statistiky $\max Y_p / \sum b_j Y_{(j)}$, kde konstanty b_j jsou voleny tak, aby potlačily vliv největších pořádkových statistik $Y_{(j)}$. Při odvozování rozdělení takovéto statistiky Bølviken využil toho, že je známo, jak hledat rozdělení lineární kombinace pořádkových statistik z exponenciálního rozdělení. Nejjednodušší možná volba $\{b_j\}$ je $b_1 = \ldots = b_{n-a} = 1$, $b_{n-a+1} = \ldots = b_n = 0$, tj. jedná se o useknutí.

Platí

$$P\left(\max Y_p \Big/ \sum_{j=1}^{n-a} Y_{(j)} > d\right) = \sum_{k=1}^{n} (-1)^{k-1}\binom{n}{k} Q_k(d)$$

$$Q_k(d) = \begin{cases} \prod_{j=1}^{n-a} \left(1 + \dfrac{jkd}{(a-k+j)}\right)^{-1} & k \leq a \\ (1-(k-a)d)^{n-1}(1+ad)^{-(n-a)} \prod_{j=1}^{a-1} \left(1 + \dfrac{(k-a)jd}{k-j}\right)^{-1} & a < k < \frac{1}{d}+a \\ 0 & k > \frac{1}{d} + a \end{cases}$$

Jiný přístup navrhl Siegel [4]. Jeho testová statistika se neopírá pouze o $\max Y_p$ nýbrž o všechny velké hodnoty periodogramu, které jsou větší než prahová hodnota a. Navrhovaná statistika má tvar $\sum (Y_p - a)_+$, kde $a = \lambda u_p$, přičemž u_p je kritická hodnota Fisherova testu a λ je konstanta z intervalu $\langle 0,1 \rangle$.

Platí

$$P\left(\sum (Y_p - a)_+ > t\right) = \sum_{\ell=1}^{n} \sum_{k=0}^{\ell-1} (-1)^{k+\ell+1} \binom{n}{\ell}\binom{\ell-1}{k}\binom{n-1}{k} t^k (1-\ell a - t)_+^{n-k-1} .$$

4. Simulace

Oba testy obsahují subjektivní prvek. V Bølvikenově testu je třeba se rozhodnout, jak velký počet a největších hodnot periodogramu vypustíme pro zlepšení odhadu σ^2, v Siegelově testu je třeba předem zvolit konstantu λ. Pro vytvoření představy, jak se mění síla testu v závislosti na volbě a, resp. λ a k jejich vzájemnému porovnání jsme provedli simulace.

Nesimulovali jsme přímo časové řady, nýbrž již odvozené hodnoty periodogramu, přičemž jsme využili toho, že jestliže pro amplitudu ρ_p odpovídající frekvenci $\lambda_p = 2\pi p/N$ platí

$\rho_p = 0 \Longrightarrow (4\pi/\sigma^2)\, I(\lambda_p) \sim \chi^2(2)$ a

$\rho_p > 0 \Longrightarrow (4\pi/\sigma^2)\ I(\lambda_p) \sim \chi^2(2,$ par.necentr. $\Delta = \frac{N}{2}\ \frac{\rho_p^2}{\sigma^2})$.

Simulovali jsme tedy vektory o n=20 složkách, což odpovídá časové řadě s N = 41 nebo 42 členy, resp. o n=30, což odpovídá časové řadě o N=61 nebo 62 členy. Náhodné vektory byly vytvořeny 1,2 nebo 3 necentrálními χ^2 rozdělenými náhodnými veličinami, zbytek pak centrálními χ^2 rozdělenými náhodnými veličinami, což postupně odpovídá tomu, že se v periodické složce vyskytují 1, 2, 3 frekvence s kladnými amplitudami. Parametr necentrality jsme volili Δ = 4, 9, 16, 25. Hladina významnosti α = 0.05. Celkově jsme pro n = 20 i 30 provedli 5000 simulací, z kterých jsme odhadovali pravděpodobnost zamítnutí nulové hypotézy, to znamená, že směrodatná odchylka odhadu je menší než 0.007. V následujících tabulkách jsou uvedeny relativní četnosti zamítnutí H_0 pro Fisherův test (f), Siegelův test s λ=0.4 (s0.4) a s λ=0.6 (s0.6), Bølvikenův test při vypuštění 10% (b10), 20%(b20) a 30% (b30) největších hodnot periodogramu.

Tabulka 1 (n = 20)

počet necentr. χ^2 rozděl.náh.vel = 1 (odpovídá 1 kladné amplitudě)

	amplituda pro N=41	f	s0.4	s0.6	b10	b20	b 30
4	0.441	0.089	0.089	0.089	0.088	0.089	0.086
9	0.662	0.318	0.288	0.315	0.317	0.312	0.295
16	0.883	0.635	0.580	0.630	0.633	0.620	0.597
25	1.104	0.892	0.851	0.887	0.892	0.886	0.873

počet necentr. χ^2 rozděl.náh.vel = 2 (odpovídá 2 kladným amplitudám)

	amplituda pro N=41	f	s0.4	s0.6	b10	b20	b30
4	0.441	0.120	0.132	0.127	0.125	0.128	0.131
9	0.662	0.359	0.409	0.399	0.391	0.402	0.391
16	0.883	0.702	0.806	0.767	0.765	0.768	0.754
25	1.104	0.924	0.973	0.963	0.961	0.959	0.928

počet necentr. χ^2 rozděl.náh.vel = 3 (odpovídá 3 kladným amplitudám)

	amplituda pro N=41	f	s0.4	s0.6	b10	b20	b30
4	0.441	0.125	0.138	0.138	0.131	0.139	0.138
9	0.662	0.310	0.468	0.390	0.366	0.428	0.422
16	0.883	0.553	0.851	0.755	0.693	0.788	0.736
25	1.104	0.768	0.991	0.956	0.914	0.979	0.975

Tabulka 2 (n = 30)

počet necentr. χ^2 rozděl.náh.vel.= 1 (odpovídá 1 kladné amplitudě)

	amplituda pro N=61	f	s0.4	s0.6	b10	b20	b30
4	0.362	0.094	0.088	0.089	0.098	0.090	0.090
9	0.543	0.315	0.280	0.319	0.320	0.316	0.295
16	0.724	0.639	0.567	0.629	0.634	0.627	0.612
25	0.905	0.891	0.831	0.882	0.891	0.881	0.872

počet necentr. χ^2 rozděl.náh.vel.= 2 (odpovídá 2 kladným amplitudám)

	amplituda pro N=61	f	s0.4	s0.6	b10	b20	b30
4	0.362	0.115	0.125	0.122	0.121	0.123	0.121
9	0.543	0.370	0.401	0.396	0.395	0.393	0.390
16	0.724	0.800	0.837	0.843	0.830	0.831	0.815
25	0.905	0.957	0.967	0.971	0.972	0.966	0.964

počet necentr. χ^2 rozděl.náh.vel.=3 (odpovídá 3 kladným amplitudám)

	amplituda pro N=61	f	s0.4	s0.6	b10	b20	b30
4	0.362	0.143	0.168	0.162	0.155	0.161	0.157
9	0.543	0.398	0.550	0.468	0.455	0.471	0.475
16	0.724	0.752	0.918	0.879	0.868	0.885	0.878
25	0.905	0.950	0.996	0.991	0.989	0.992	0.991

K největšímu rozdílu v síle testu dochází kolem hodnoty $\Delta=16$. Všimněme si Siegelova testu s $\lambda=$ 0.6 (s0.6) a Bølvikenova testu (b10) s vynecháním 10% pozorování. Pro jednu periodu s kladnou amplitudou dochází u nefisherovských testů s0.6 a b10 pouze k velmi malému snížení síly testu ≤ 0.005(pro n=20) a ≤ 0.01(pro n=30). Zvýšení síly testu pro složenou amlitudu je daleko podstatnější. Pro n=20, $\Delta=16$ a 2 kladné amplitudy činí rozdíl vzhledem k Fisherově testu 0.065 (pro s0.6), 0.063 (pro b10). Pro n=20, $\Delta=16$, 3 kladné amplitudy činí rozdíl 0.202 (pro s0.4) a 0.140 (pro b10). Pro Siegelův test s $\lambda=0.4$ (s0.4) a Bølvikenův test s vypuštěním většího počtu pozorování jsou rozdíly v sílách testu větší, jak pro jednu kladnou amplitudu, kdy nefisherovské testy jsou výrazněji horší, tak pro více kladných amplitud, kdy naopak dávají podstatně lepší výsledky.

Je tedy patrné, že jakmile máme sebemenší pochybnost o tom, že proces obsahuje pouze jednu frekvenci s kladnou amplitudou, je vhodnější použít buď Siegelův nebo Bølvikenův test. Abychom příliš nesnížili sílu testu v případě, že by proces přece jen kmital s jednou kladnou amplitudou, doporučila bych použít s0.6 nebo b10 nebo b20.

Literatura

[1] Anderson, T.W. (1971). The statistical analysis of time series. Wiley

[2] Bølviken, E. (1983). New Tests of Significance in Periodogram Analysis, Scand. J.Statist. 10, 1-9

[3] Fisher, R.A. (1929) Tests of significance in harmonic analysis. Proc.Roy.Soc., London, Ser.A, 125, 54-59

[4] Siegel, A. (1980) Testing for periodicity in a time series. J.Amer.Statist. Assoc. 75, 345-348