

STATISTIČKA ANALIZA OSTATAKA

Dušan K. TOMKOVIĆ — Beograd*

1. UVOD

Veliki broj radova posvećen je obradi, a posebno otkrivanju grubih grešaka (anomalnih rezultata) u nizu rezultata merenja. Mereći više puta jednu istu fizičku veličinu praktično svaki geodeta se suočio sa situacijom da se jedan ili više rezultata jako razlikuju od ostalih u nizu rezultata merenja. Tradicionalni postupak otkrivanja grubih grešaka u nizu rezultata merenja je relativno prost, koristeći dozvoljeno odstupanje odbacuje se jedan ili više rezultata i vrše se nova merenja. Ovakav formalistički pristup, kada se radi o preciznim merenjima, ima nedostataka, jer anomalni rezultati mogu dati korisnu informaciju o neprimetnim promenama uslova merenja. Daleko je bolje, kada opažać ima veće teorijsko znanje i praktično iskustvo o procesu merenja koje vrši, da na osnovu očekivane varijacije rezultata uoči anomalne rezultate ili promenu uslova merenja. Kada ovo nije moguće opažać koristi statističke postupke za otkrivanje grubih grešaka u rezultatima merenja. Osnovna (nulta) hipoteza (H_0) je da su svi rezultati merenja nezavisni u statističkom smislu i imaju normalnu raspodelu $N(\mu, \sigma)$ ($H_0: l_i \in N(\mu, \sigma), i = 1, 2, \dots, n$), a alternativna hipoteza na primer, $H: l_i \in N(\mu + \Delta, \sigma), \Delta > 0$, tj. j-ti rezultat merenja ima grubu grešku. Najveći izbor statističkih kriterijuma je razrađen za najvažniji slučaj u praksi kada je μ i σ nepoznato, a mnogi su usredsređeni na proveru veličine ostataka $e_i = l_i - \bar{l}$ (u geodetskoj literaturi koriste se popravke $v_i = -e_i$) [1, 3, 4, 5, 8, 11].

2. ANALIZA OSTATAKA

2.1. Kriterijum beznačajnosti

Rezultati merenja jedne iste fizičke veličine mogu se opisati na sledeći način:

$$l_i = \mu + \varepsilon_i \quad (i = 1, 2, \dots, n); \varepsilon_i \in N(0, \sigma^2), \quad (1)$$

odnosno

$$L_{n,1} = A_{n,1}(\mu + \varepsilon_{n,1}), \quad (2)$$

* Adresa autora: Dr Dušan K. Tomković, Viša geodetska škola, Beograd, Milana Rakića 42

gde je

$$A^* = \|1, 1, \dots, 1\|.$$

Ostatak je linearna funkcija grešaka merenja ε [6]

$$e = (I - P)\varepsilon, \quad (3)$$

gde je

$$P = A(A^*A)^{-1}A^*.$$

U razvijenom obliku

$$e_i = -\frac{1}{n}\varepsilon_1 - \frac{1}{n}\varepsilon_2 - \dots - \frac{1}{n}\varepsilon_{i-1} + \frac{n-1}{n}\varepsilon_i - \frac{1}{n}\varepsilon_{i+1} - \dots - \frac{1}{n}\varepsilon_n. \quad (4)$$

Neposredno iz (4) varijansa ostsatka e_i glasi

$$V_{\varepsilon}(e_i) = \frac{n-1}{n^2}\sigma^2 + \frac{(n-1)^2}{n^2}\sigma^2, \quad (5)$$

gde prvi član sa desne strane predstavlja varijansu grešaka

$$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}, \varepsilon_{i+1}, \dots, \varepsilon_n.$$

Formula (5) je pogodna za analizu uticaja grešaka $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}, \varepsilon_{i+1}, \dots, \varepsilon_n$ na ostatak e_i . U tom cilju primenom kriterijuma beznačajnosti može se odrediti broj merenja kada je uticaj ovih grešaka beznačajan, za koeficijent beznačajnosti 0.05, odnosno

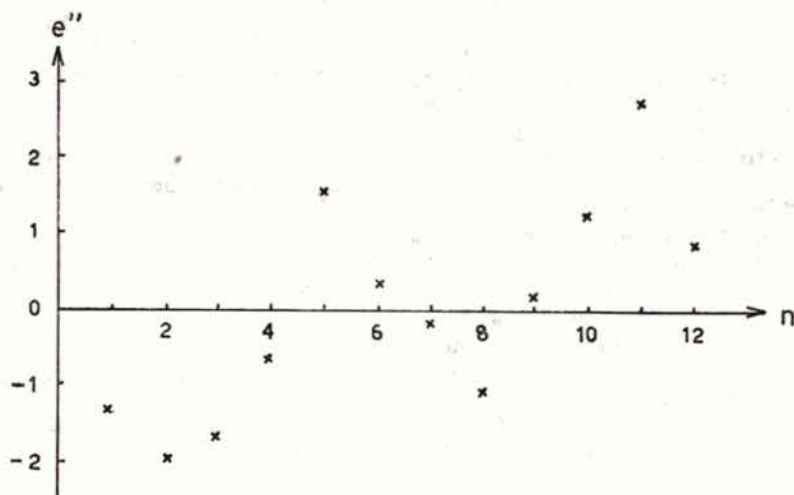
$$\frac{n-1}{n^2}\sigma^2 < 0.11 \frac{(n-1)^2}{n^2}\sigma^2, \quad (6)$$

odakle sledi da je praktično za $n \geq 10$ uticaj grešaka $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}, \varepsilon_{i+1}, \dots, \varepsilon_n$ beznačajan na ostatak e_i .

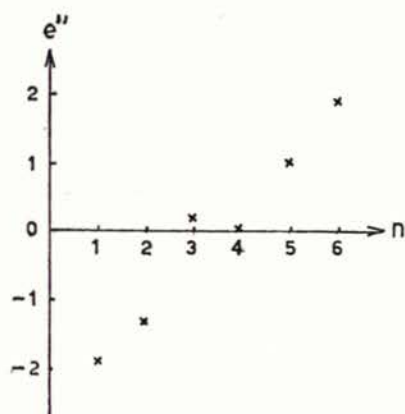
2.2. Autokorelacija

U mnogim slučajevima provera veličine ostataka daje rezultat da nema grubih grešaka u rezultatima merenja. Međutim, dalje ispitivanje ostataka može dati odgovor o adekvatnosti modela (1), odnosno o strukturi greške ε . Grafički prikaz ostataka u zavisnosti od broja merenja u nekim primerima može ukazati na postojanje autokorelacije ostataka. U svojstvu primera korišćeni su rezultati merenja iz [9(str. 118)] i [10]), sa napomenom da se u [9] mogu naći korisni podaci i rezultati ispitivanja merenja iz [9]. Na sl. 1 [9], sl. 2 i sl. 3 [10] dati su grafici ostataka u funkciji broja merenja.

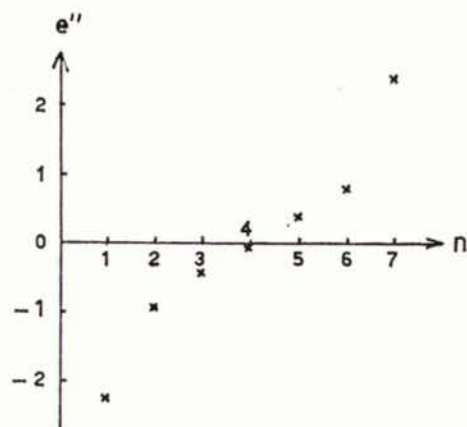
Kada veličine e_i isuviše retko ili često menjaju znak može se sumnjati na prisustvo autokorelacije ostataka. U datim primerima veličine e_i retko menjaju znak te se može pretpostaviti prisustvo pozitivne autokorelacije ostataka, odnosno strukturom greške ε nije obuhvaćena promenljiva sistematska greška (periodična (sl. 1.) i linearna (sl. 2 i 3.)).



Sl. 1



Sl. 2



Sl. 3

Metodom razlaganja niza ostataka [2,7] na serije sa jednim istim znakom može se utvrditi na jednostavan način prisustvo autokorelacije ostataka. Verovatnoće pojave broja serija istog znaka (u_r), $p(u \leq u_r)$, pozitivnih (n_+) i negativnih (n_-) ostataka za date primere su: $p_1 = 0.067$, $p_2 = 0.133$, $p_3 = 0.057$. Sračunate verovatnoće nisu značajne (na nivou $\alpha = 0.05$), ali su dovoljno male te ukazuju na pozitivnu autokorelaciju ostataka u ovim primerima. Postoje i

druge metode za otkrivanje autokorelacije ostataka (na primer metod Durbin-Watsona [6,7]), ali s obzirom na mali broj merenja koji se vrši pri merenju jedne iste fizičke veličine u geodeziji teško je pouzdano otkriti njeno prisustvo.

LITERATURA

- [1] Dixon, W. J.: Processing data for outliers, *Biometrics*, 1953, 22, 74—80.
- [2] Drejper, N., Smit, G.: *Prikladnoj regressionnyj analiz*, Moskva, Statistika 1973.
- [3] Dejvid, G.: *Porjadkovye statistiki*, Moskva 1979.
- [4] Ellenberg, J. H.: The Joint Distribution of the Standardized Least Squares Residuals from a General Linear Regression, *JASA*, 1973., vol 68, No. 344, 941—943.
- [5] Grabbs, F. E.: Sample criteria for testing outlying observations, *Anal. of Mathematical Statistics*, 1950., vol. 21, 27—58
- [6] Kendel, M.: *Vremennye rjady*, Moskva 1981.
- [7] Kejn, E.: *Ekonomičeskaja statistika i ekonometrija*, Moskva, Statistika, 1977.
- [8] Lund, R. E.: Tables for Approximate Test for Outliers in Linear Models, *Technometrics*, 1975., vol. 17, No 4, 473—476.
- [9] Kemnic, Ju. V.: *Teorija ošibok izmerenij*, Moskva, Nedra, 1977.
- [10] Tomković, D. K.: Otkrivanje sistematskih grešaka u uglovnim merenjima, *Geodetska služba*, 1972., br. 5, 15—17.
- [11] Tietjen, G. L. Moore, R. H.: Some Grubbs-Type Statistics for the Detection of Several Outliers, *Technometrics*, 1972., vol. 14, No. 3, 583—597.

REZIME

U radu je primenom kriterijuma beznačajnosti utvrđen broj merenja kada je uticaj grešaka $\epsilon_1, \epsilon_2, \dots, \epsilon_{i-1}, \epsilon_{i+1}, \dots, \epsilon_n$ beznačajan na ostatak e_i i ispitivanjem autokorelacije ostataka analizira se struktura grešaka merenja.

Primljeno: 1985-10-01