ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 206, 2007

Janusz Wywiał, Tomasz Żądło***

JACKKNIFE FORECASTS OF TIME SERIES

Abstract. In the paper we present the examples of forecasts of time series with seasonal fluctuations. Based on the jackknife method we estimate variances of seasonal factors and the *MSE* of prediction. Jackknife method has been introduced by M. Quenouille (1949) and then it has been developed among others by J. Tukey (1958) and J. Shao, D. Tu (1995).

Key words: jackknife, time series, seasonal fluctuations.

1. BASIC DEFINITIONS AND EQUATIONS

Let us consider the following model of time series:

$$Y_{t,l} = f(t, l, \boldsymbol{\theta}) + \varepsilon_{t,l}, \quad t = 1, ..., n+h, \quad l = 1, ..., r \tag{1}$$

with

$$f(t, l, \boldsymbol{\theta}) = f_*(t, \boldsymbol{\theta}_*) + C_l, \quad t = 1, ..., n+h, \quad l = 1, ..., r \tag{2}$$

or

$$f(t, l, \boldsymbol{\theta}) = f_*(t, \boldsymbol{\theta}_*) C_l, \quad t = 1, ..., n+h, \quad l = 1, ..., r \tag{3}$$

where $\boldsymbol{\theta} = [\boldsymbol{\theta}_*^T C_1 ... C_r]^T$, the function $f_*(t, \boldsymbol{\theta}_*)$ is trend function, C_l is seasonal factor in l-th among r phases of the cycle, $\varepsilon_{t,l}$ is random component. We assume that distributions of random components are identical and independent and $E(\varepsilon_{t,l}) = 0$ and $D^2(\varepsilon_{t,l}) = D^2(Y_{t,l}) = \sigma^2$ for every $t = 1, ..., n+h$.

* Professor, Department of Statistics, University of Economics in Katowice.
** Ph.D., Department of Statistics, University of Economics in Katowice.

Vector θ denotes vector of trend parameters, n – number of periods, for which realisations of Y are known. Forecasts for periods $n+h$, $h \geqslant 1$ are analysed for the version (2) of the model (1) called additive model of seasonal fluctuations, and for the version (3) of the model (1) called multiplicative model of seasonal fluctuations.

We use estimators $\mathbf{G}_n = [\mathbf{G}_{*n}^T c_1 \dots c_r]^T$ of the parameters $\theta = [\theta_*^T C_1 \dots C_r]^T$ to obtain estimates of the function $f(t, l, \theta)$, which are denoted by $F(t, l, \mathbf{G}_n)$. Estimates of the trend function $f_*(t, \theta)$ are denoted by $F_*(t, \mathbf{G}_{*n})$. The forecast of $Y_{n+h,l}$ for the period $n+h$ are given by

$$F_{n+h,l} = f(n+h, l, \mathbf{G}_n) \tag{4}$$

Let F_{*n+h} denote the forecast for the period $n+h$ based on the trend function. Hence, in the case of additive model of seasonal fluctuations, the forecast (4) is given by

$$F_{n+h,l} = F_{*n+h} + c_l \tag{5}$$

and, in the case of multiplicative model of seasonal fluctuations, it is as follows:

$$F_{n+h,l} = F_{*n+h} c_l \tag{6}$$

We denote the prediction error by

$$U_{n+h,l} = Y_{n+h,l} - F_{n+h,l} \tag{7}$$

Expected value of prediction error is given by

$$\delta(F_{n+h,l}) = \mathrm{E}(U_{n+h,l}) = \mathrm{E}(Y_{n+h,l}) - \mathrm{E}(F_{n+h,l}) \tag{8}$$

Finally, we obtain the following equation of ex ante mean square error of prediction:

$$\begin{aligned}
\mathrm{D}^2(U_{n+h,l}) &= \mathrm{E}(U_{n+h,l}^2) = \mathrm{E}(Y_{n+h,l} - F_{n+h,l})^2 = \mathrm{E}[(Y_{n+h,l} - \mathrm{E}(F_{n+h,l})) - \\
&+ (F_{n+h,l} - \mathrm{E}(F_{n+h,l}))]^2 = \mathrm{E}[(Y_{n+h,l} - \mathrm{E}(Y_{n+h,l})) + \delta(F_{n+h,l}) - \\
&+ (F_{n+h,l} - \mathrm{E}(F_{n+h,l}))]^2 = \sigma^2 + \delta^2(F_{n+h,l}) + \mathrm{D}^2(F_{n+h,l})
\end{aligned} \tag{9}$$

Let us discuss the jackknife method as the method of estimating θ based on the statistic $\mathbf{G}_n$ (see W o l t e r 1985). Let $\mathbf{G}_{n,i}$ be an estimator of θ based on the same equations as $\mathbf{G}_n$, but using observations $Y_1, \ldots, Y_n$, without the i-th one, $i = 1, \ldots, n$. Then, we compute pseudovalues

$$\hat{\mathbf{G}}_{n,i} = n\mathbf{G}_n - (n-1)\mathbf{G}_{n,i} \tag{10}$$

Finally, jackknife estimator is given by

$$\hat{\mathbf{G}} = \frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{G}}_{n,i} \tag{11}$$

If θ parameter is scalar, then variance estimators of $\hat{G}$ are as follows:

$$\hat{\mathrm{D}}_1^2(\hat{G}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(\hat{G}_{n,i} - \hat{G})^2 \tag{12}$$

$$\hat{\mathrm{D}}_2^2(\hat{G}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(\hat{G}_{n,i} - G)^2 \tag{13}$$

Jackknife method reduces the bias of estimators of parameters $\boldsymbol{\theta}$. It should also be noted that R. G. Miller (1974) discusses jackknife method for estimating parameters of nonlinear function.

We obtain jackknife forecasts similarly to jackknife estimators of $\boldsymbol{\theta}$. First, we compute forecasts (4) but based on estimators $\mathbf{G}_{n,i}$

$$F_{n+h,l,i} = f(n+h, l, \mathbf{G}_{n,i}), \quad i = 1, \ldots, n \tag{14}$$

Then, we compute pseudovalues (pseudoforecasts)

$$\hat{F}_{n+h,l,i} = nF_{n+h,l} - (n-1)F_{n+h,l,i} \tag{15}$$

Finally, jackknife forecast is as follows:

$$\hat{F}_{n+h,l} = \frac{1}{n}\sum_{i=1}^{n}\hat{F}_{n+h,l,i} \tag{16}$$

Estimators of forecasts' variances are given by (12) and (13) where $\hat{G}$, $\hat{G}_i$, G_n are replaced by $\hat{F}_{n+h,l}$, $\hat{F}_{n+h,l,i}$, $F_{n+h,l}$.

Let us assume that prediction error is given by (7). We estimate ex ante mean square error of prediction given by (9). The properties of jackknife method imply that prediction bias δ is reduced. Precisely, jackknife method eliminates $O(n^{-1})$ bias component, and hence bias is $O(n^{-2})$ (see Shao, T u 1995, p. 5). (The bias δ is $O(n^{-2})$, i.e. $\delta = O(n^{-2})$, if real positive number M exists, such that $\frac{|\delta|}{n^{-2}} \leqslant M$ for every n). Hence

$$|\delta(\hat{F}_{n+h,l})| \leqslant |\delta(F_{n+h,l})| \tag{17}$$

It should be stressed, that although jackknife estimator has smaller bias, the decrease of the bias may imply the increase of the predition variance and generally we cannot say that (see S h a o, T u 1995, p. 67) $D^2(Y_{n+h,l} - \hat{F}_{n+h,l}) \leqslant D^2(Y_{n+h,l} - F_{n+h,l})$.

Assuming, that $\delta(\hat{F}_{n+h,l})$ is enough reduced to be omitted, we obtain, based on (5), that

$$D^2(U_{n+h,l}) \approx \sigma^2 + D^2(\hat{F}_{n+h,l}) \tag{18}$$

To estimate (18), variance σ^2 should be replaced by the appropriate estimator and $D^2(\hat{F}_{n+h,l})$ is estimated based on one of the following equations:

$$\hat{D}_1^2(\hat{F}_{n+h,l}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} (\hat{F}_{n+h,l,i} - \hat{F}_{n+h,l})^2 \tag{19}$$

$$\hat{D}_2^2(\hat{F}_{n+h,l}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} (\hat{F}_{n+h,l,i} - \hat{F}_{n+h,l})^2 \tag{20}$$

In the following sections we discuss a problem of forecasting time series with seasonal fluctuations.

2. SEASONAL FACTORS' METHOD

To obtain estimates c_l of seasonal factors C_l, firstly, we eliminate trend component from the time series. For the additive model of seasonal fluctuations we have

$$z_{t,l} = y_{t,l} - F_*(t, \mathbf{G}) \tag{21}$$

For the multiplicative model of seasonal fluctuations we get

$$z_{t,l} = \frac{y_{t,l}}{F_*(t, \mathbf{G}_*)} \tag{22}$$

Let us remind that $F_*(t, \mathbf{G}_*)$ is estimate of trend function in the period t, which is received by estimating trend parameters using e.g. the Least Squares Method.

Secondly, we compute

$$z_t = \frac{1}{k} \sum_{j=1}^{k} z_{t,l} \tag{23}$$

where:

k – number of l-th phases of the cycle,

$t = l + r \cdot (j-1)$, this condition means, that we consider only periods t in the same l-th phase of the cycle,

r – number of phases in the cycle.

Thirdly, we compute estimates c_l of C_l. For additive model of seasonal fluctuations we have

$$c_l = z_t - \frac{1}{r} \sum_{t=1}^{r} z_t \tag{24}$$

For multiplicative model of seasonal fluctuations we get

$$c_l = \frac{z_t}{\frac{1}{r} \sum_{t=1}^{r} z_t} \tag{25}$$

Forecast for period $n+h$ we obtain using (5) for additive model of seasonal fluctuations or (6) for multiplicative model of seasonal fluctuations. We should stress that if we use seasonal factors' method we are not able to estimate variances of seasonal factors and MSE of prediction both for additive and multiplicative model of seasonal fluctuations. But, if we use Klein's method in the case of additive model of seasonal fluctuations we may receive estimates of variances of seasonal factors and MSE of prediction.

3. JACKKNIFE VERSION OF SEASONAL FACTORS' METHOD

We discuss the problem of variance estimation of seasonal factors based on jackknife method. We will also use the method to estimate variances of trend parameters. Let $c_{l,i}$, where $l = 1, \ldots, r$ and $i = 1, \ldots, n$, denote estimates of seasonal factors obtained using seasonal factors' method based on time series without the i-th observation. Let $\mathbf{G}_{*n,i}$ denote estimates of trend parameters obtained using e.g. the Least Squares Method based on time series without the i-th observation. Then, pseudovalues of trend parameters and pseudovalues of seasonal factors denoted by $\hat{\mathbf{G}}_{*n,i}$ and $\hat{c}_l^{\,i}$ respectively are obtained based on (10), where $\hat{\mathbf{G}}_{n,i} = [\hat{\mathbf{G}}_{*n,i}\, \hat{c}_{1,i} \dots \hat{c}_{r,i}]$, $\mathbf{G}_n = [\mathbf{G}_{*n}\, c_1 \dots c_r]$, $\mathbf{G}_{n,i} = [\mathbf{G}_{*n,i}\, c_{1,i} \dots c_{r,i}]$. Finally, jackknife estimators of trend parameters and seasonal factors denoted by $\hat{\mathbf{G}}_*$ and $\hat{c}_l$ respectively (where $l = 1, \dots, r$) are obtained based on (11), where $\hat{\mathbf{G}} = [\hat{\mathbf{G}}_*\, \hat{c}_1 \dots \hat{c}_r]$. Estimators of their variances are given by (12) or (13).

To obtain jackknife forecasts, firstly, we obtain forecasts based on time series without the i-th observation $(i = 1, \dots, n)$ using equation (14). For additive model we have

$$F_{n+h,l,i} = F_{*n+h,i} + c_{l,i} \tag{26}$$

and for multiplicative model we get

$$F_{n+h,l,i} = F_{*n+h,i}\, c_{l,i} \tag{27}$$

where $F_{*n+h,i} = f_*(n+h, \mathbf{G}_{*n,i})$ is trend forecast for period $n+h$ based on time series without the i-th observation $(i = 1, \dots, n)$. Secondly, pseudovalues (in this case pseudoforecasts) are obtained using (15). Finally, we obtain jackknife forecast based on (16) and we use the following equation to estimate *MSE* of prediction (compare with (18)):

$$\hat{\mathrm{D}}^2(U_{n+h,l}) = \hat{\sigma}^2 + \hat{\mathrm{D}}^2(\hat{F}_{n+h,l}) \tag{28}$$

where $\hat{\mathrm{D}}^2(\hat{F}_{n+h,l})$ is given by (19) or (20) and

$$\hat{\sigma}^2 = \frac{1}{n-k-r} \sum_{t=1}^{n} (Y_t - F(t, l, \mathbf{G}_n))^2 \tag{29}$$

To estimate σ^2 we can also use jackknife estimator instead of (29).

4. EXAMPLE 1

We analyse data on quarterly production of cement in Poland in 1994–1999. We assume that the discussed data decompose into linear trend and additive seasonal fluctuations. The value of Ljung-Box test' statistic allows to accept zero hypothesis that there is no autocorrelation of random component till 16-th order (remainders are computed based on the linear trend function estimated using the Least Squares Method and seasonal factors estimated using the method described in the Section 2 of the paper). We use jackknife method to estimate parameters and their variances and to forecast time series and, finally, to estimate *MSEs* of prediction. We also use the method described in the second section of the paper with trend parameters estimated using the Least Squares Method.

The graph of the discussed data is presented on Fig. 1.

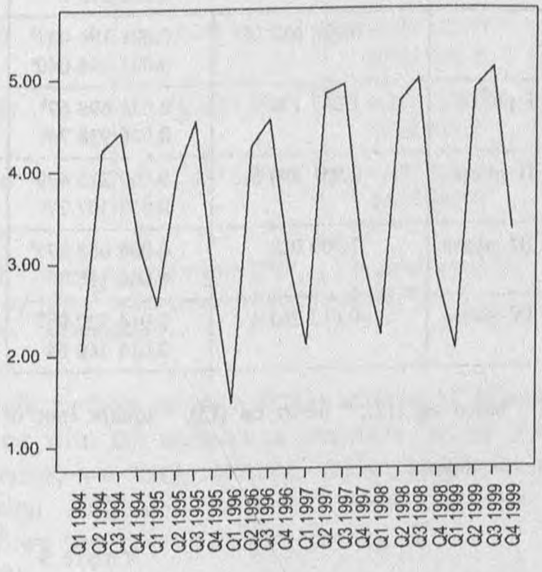

Fig. 1. Quarterly production of cement (in million tons) in Poland in 1994–1999

Source: "Rocznik Statystyczny Przemysłu" (1995–1999), GUS, Warszawa, following: Zeliaś, Pawełek, Wanat 2003, p. 115.

Table 1

Estimates of model parameters (the method of seasonal factors, trend estimated using the Least Squares Method)

Parameter	Estimate
Intercept	3.232 138
Trend parameter	0.028 395 65
Seasonal factor for I phase	–1.472 823
Seasonal factor for II phase	0.905 447 8
Seasonal factor for III phase	1.150 386
Seasonal factor for IV phase	–0.583 010 1

Source: autors' calculations.

Table 2

Jackknife estimates of model parameters and their variances

Parameters	Value of jackknife estimator[a]	Estimate of variance	Estimate of standard error[d]
Intercept	3.229 754	0.218 074 9[b] 0.218 075 1[c]	0.466 984 9 0.466 985 1
Trend parameter	0.028 680 21	0.001 194 037[b] 0.001 194 040[c]	0.034 554 84 0.034 554 89
Seasonal factor for I phase	–1.547 136	0.026 698 69[b] 0.026 938 79[c]	0.163 397 3 0.164 130 4
Seasonal factor for II phase	0.951 390 5	0.009 215 499[b] 0.009 307 27[c]	0.095 997 39 0.096 474 19
Seasonal factor for III phase	1.209 009	0.006 068 877[b] 0.006 218 3[c]	0.077 903 0.078 856 2
Seasonal factor for IV phase	–0.613 263 4	0.014 720 85[b] 0.014 760 64[c]	0.121 329 5 0.121 493 4

[a] Based on (11), [b] based on (12), [c] based on (13), [d] square root of estimated value of variance.

Source: autors' calculations.

Table 3

Forecasts based on the seasonal factors' method with trend estimated using the Least Squares Method

Date	Forecast
I quarter of 2000	2.469 206
II quarter of 2000	4.875 872
III quarter of 2000	5.149 206
IV quarter of 2000	3.444 206

Source: autors' calculations.

Table 4

Jackknife forecasts and estimates of *MSE* of prediction

Date	Jackknife forecast[a]	Value of estimated MSE	Value of estimated *RMSE*[d]
I quarter of 2000	2.399 623	0.464 553 7[b] 0.464 764 2[c]	0.681 581 8 0.681 736 2
II quarter of 2000	4.926 83	0.403 780 3[b] 0.403 893 2[c]	0.635 437 0.635 525 9
III quarter of 2000	5.213 129	0.399 860 3[b] 0.400 037 9[c]	0.632 345 1 0.632 485 5
IV quarter of 2000	3.419 537	0.407 515 2[b] 0.407 541 7[c]	0.638 369 2 0.638 389 9

[a] based on (16), [b] based on (28) where $D^2(\hat{F}_{n+h,l})$ is given by (19), [c] based on (28) where $D^2(\hat{F}_{n+h,l})$ is given by (20), [d] square root of estimated *MSE*.
Source: autors' calculations.

Using jackknife method we may obtain additional information (see Tab. 2 and 4) comparing with the method of seasonal factors (see Tab. 1 and 3). Being more precise, we may estimate variances of seasonal factors and *MSE* of prediction. Although, we should stress that these information are also available if we use Klein's model (see e.g. Zeliaś, Pawełek, Wanat 2003, p. 88).

5. EXAMPLE 2

We analyse data on quarterly incomes of Polifarb Cieszyn company in 1995–1999. We assume that the discussed data decompose into linear trend and multiplicative seasonal fluctuations. The value of Ljung-Box test' statistics allows to accept zero hypothesis that there is no autocorrelation of random component till 14-th order (remainders are computed based on the linear trend function estimated using the Least Squares Method and seasonal factors estimated using the method described in the Section 2 of the paper). Similarly to the first example, we use jackknife method to estimates parameters and their variances and to forecast time series and, finally, to estimate *MSES* of prediction. We also use the method described in the Section 2 of the paper with trend parameters estimated using the Least Squares Method.

The graph of the discussed data is presented on Fig. 2.

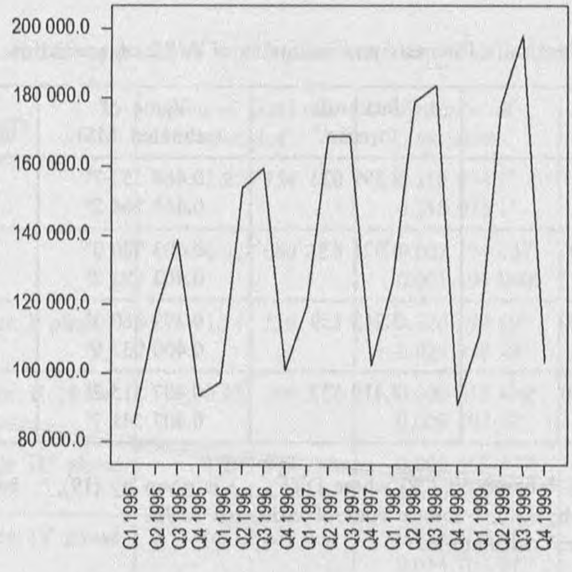

Fig. 2. Quarterly incomes of Polifarb Cieszyn (in thousand zł) company in 1995–1999

Source: www.bossa.pl.

Table 5

Estimates of model parameters (the method of seasonal factors, trend estimated using the Least Squares Method)

Parameter	Estimate
Intercept	110 616.7
Trend parameter	2 206.496
Seasonal factor for I phase	0.826 776 4
Seasonal factor for II phase	1.197 014
Seasonal factor for III phase	1.255 150
Seasonal factor for IV phase	0.721 059 1

Source: authors' calculations.

Table 6

Jackknife estimates of model parameters and their variances

Parameters	Value of jackknife estimator[a]	Estimate of variance	Estimate of standard error[d]
Intercept	110 209.1	207 536 621[b] 207 545 366[c]	14 406.13 14 406.43
Trend parameter	2 209.446	2 396 577[b] 2 396 578[c]	1 548.088 1 548.088
Sasonal factor for I phase	0.814 139 3	0.001 353 306[b] 0.001 361 711[c]	0.036 787 31 0.036 901 37
Seasonal factor for II phase	1.205 322	0.001 578 348 0.001 581 980[c]	0.039 728 43 0.03977411
Seasonal factor for III phase	1.273 929	0.000 948 983 6[b] 0.000 967 543 4[c]	0.030 805 58 0.031 105 36
Seasonal factor for IV phase	0.706 610 2	0.000 959 539 1[b] 0.000 970 527[c]	0.030 976 43 0.031 153 28

[a] based on (11), [b] based on (12), [c] based on (13), [d] square root of estimated value of variance.

Source: authors' calculations.

Table 7

Forecasts based on the seasonal factors' method with trend estimated using the Least Squares Method

Date	Forecast
I quarter of 2000	129 765.1
II quarter of 2000	190 516.3
III quarter of 2000	202 538.7
IV quarter of 2000	117 945.5

Source: authors' calculations.

Table 8

Jackknife forecasts and estimates of MSE of prediction

Date	Jackknife forecast[a]	Value of estimated MSE	Value of estimated *RMSE*[d]
I quarter of 2000	127 388.9	490 167 378[b] 490 464 572[c]	22 139.72 22 146.43
II quarter of 2000	191 510.4	863 236 925[b] 863 288 939[c]	29 380.89 29 381.78
III quarter of 2000	205 334.5	975 641 641[b] 976 053 027[c]	31 235.26 31 241.85
IV quarter of 2000	115 163.5	536 695 761[b] 537 103 111[c]	23 166.70 23 175.49

[a] based on (16), [b] based on (28) where $\hat{D}^2(\hat{F}_{n+h,l})$ is given by (19), [c] based on (28), where $\hat{D}^2(\hat{F}_{n+h,l})$ is given by (20), [d] square root of estimated *MSE*.

Source: authors' calculations.

Similarly to the case of the additive model, if we use jackknife method in the second example we may obtain additional information (see Tab. 6 and 8) comparing with the method of seasonal factors (see Tab. 5 and 7) including estimates of variances of seasonal factors and *MSE* of prediction.

6. SUMMARY

In the paper we propose the jackknife method to forecast time series with seasonal fluctuations. The problem of estimating *MSE* of prediction is also taken into consideration. Two examples of forecasting real time series are presented.

REFERENCES

Miller R. G. (1974), *An unbalanced jackknife*, "Annals of Statistics", **2**, 880–891.

Quenouille M. (1949), *Approximations tests of correlations in time series*, "Journal of the Royal Statistical Society", Ser. B, **11**, 533–538.

"Rocznik Statystyczny Rzeczypospolitej Polskiej" (1998 i 1999), GUS, Warszawa.

Shao J., Tu D. (1995), *The jackknife and bootstrap*, Springer-Verlag, New York.

Tukey J. (1958), *Bias and confidence in not quite large samples*, "Annals of Mathematical Statistics", **29**, 614.

Wolter K. (1985), *Introduction to variance estimations*, Springer-Verlag, New York.

Zeliaś A., Pawełek B., Wanat S. (2003), *Prognozowanie ekonomiczne. Teoria, przykłady, zadania*, Wydawnictwo Naukowe PWN, Warszawa.

Janusz Wywiał, Tomasz Żądło

WYKORZYSTANIE METODY *JACKKNIFE* DO PROGNOZOWANIA SZEREGÓW CZASOWYCH

W pracy zaproponowano wykorzystanie metody *jackknife* do prognozowania szeregów czasowych. Oprócz problemu prognozowania tą metodą, podjęto także problem oceny średniego błędu tak wyznaczanych prognoz. W oparciu o rzeczywiste dane zaprezentowane zostały przykłady prognozowania szeregów czasowych z wahaniami sezonowymi przy wykorzystaniu wersji *jackknife* metody wskaźników sezonowości. Oprócz wyznaczenia wartości prognozowanej w rozważanym przypadku będzie możliwa ocena wariancji błędu predykcji. Metodę *jackknife* wprowadził M. Quenouille (1949), a była rozwijana m. in. przez J. Tukey'a (1958) oraz J. Shao i D. Tu (1995).