

## Statističko tumačenje rezultata u kontroli kakvoće mlijeka i mlječnih proizvoda

### (Statistical Interpretation of the Results in Quality Control of Milk and Milk's Products)

Prof. dr. Mirko FILAJDIĆ, dr. Milana RITZ, dr. Vera VOJNOVIĆ,  
Prehrambeno-biotehnološki fakultet Sveučilišta u Zagrebu

Stručni rad — Professional Paper  
Prispjelo: 10. 4. 1986.

UDK: 637.07

#### Sažetak

*U članku je na 4 primjera prikazana primjena statističkih metoda u kontroli kakvoće mlijeka i mlječnih proizvoda.*

*Primjeri ilustriraju utvrđivanje sistematske pogreške dviju metoda, provjeru kategorija kakvoće mlijeka na temelju broja živih bakterija i rješavanje spora u postupku »super-analize«.*

#### Uvod

Kontrola kakvoće u industrijskoj preradi mlijeka predstavlja zapravo ocjenu rezultata fizikalnih, kemijskih, mikrobioloških ili organoleptičkih pokazatelja proizvoda, provjeru zahtjeva normativnotehničke dokumentacije i zahtjeva zakonodavca koji se odnose na zaštitu zdravlja i ekonomskih interesa proizvođača (Potparić et. al., 1979; Trajković et. al., 1983). Djelotvornost kontrole kakvoće ovisi o pouzdanosti procjene rezultata istraživanja i o preciznosti primjenjenih metoda. Pouzdanost se može utvrditi primjenom odnosno tumačenjem metoda matematske statistike.

U praktičnoj provedbi istraživanja svi dobiveni rezultati uvijek su opterećeni tzv. slučajnom greškom, a ponekad i sistematskom greškom metode. Na rezultat provjere kakvoće mogu utjecati i izbor uzorka i veća ili manja »preciznost« upotrebene metode. Utvrđivanje pravilne procjene kakvoće nekog proizvoda moguće je uz pomoć izračunavanja disperzije date metode analize i njene pogreške i uz poznavanje vjerojatnosti da će razlika između dobivenog rezultata i stvarne vrijednosti rezultata biti samo slučajna (Hammings 1962).

Ako se potvrdi pretpostavka da doista postoji samo slučajna razlika između rezultata analize i kontrolne veličine, a u praksi se dobije nedopustivo veće odstupanje, tada treba povećati broj usporednih istraživanja. Utvrđivanje potrebnog broja ponavljanja, da bi se postigao dosta točan rezultat, moguće je samo uz pomoć statističkih metoda kontrole kakvoće (Draper i Smith 1981).

Prema tome, može se reći da je osnovna zadaća metoda statističke kontrole utvrđivanje pogreške analize i ustanovljavanje vjerodostojne razlike između rezultata analize i kontrolnih veličina kao i metrološki karakteristika konkretnih metoda analize za prihvaćanje opravdane informacije obavljenog posla.

Primjena statističkih metoda u kontroli industrijske prerade mlijeka ne pomaže samo postizavanju objektivne istine o rezultatima analize sastava i kakvoće proizvodnje, već i kontroli utroška sirovina za proizvodnju jedinice proizvoda. U proračunu bilance masti prije svega je potrebno ustanoviti pogrešku utvrđivanja masti u punomasnom i obranom mlijeku, sirutki, gotovoj prerađevini i pogrešku izmjerene količine njihove mase. Kod rješavanja takve vrste problema potrebno je utvrditi složene pogreške, a to opet zahtijeva statističke metode.

Ponekad se u industrijskoj preradi mlijeka susreću metode analize čije se metrološko testiranje nikad ne provodi. Upotreba tih metoda u načelu nije poželjna, jer se uslijed premalo informacija o pogreškama analize ne može donijeti objektivna ocjena kakvoće istraživanih uzoraka. U takvim slučajevima treba provjeriti rezultate upotrebom određene metode čija je »preciznost« i »točnost« poznata (J u r a n et. al., 1976; ISO-Standard 5725).

Pravila statističke kontrole nisu ovisna o točno definiranim objektima kontrole. Na temelju toga može se upotrijebiti primijenjeni proračun na nekom uzorku koji se istražuje određenom metodom i na druge uzorke koji se istražuju drugim metodama. S obzirom na važnost primjene statističkih metoda u kontroli kakvoće mlječnih prerađevina, ta se važnost treba prikazati sa nekoliko primjera iz prakse.

### Statističke metode u tumačenju eksperimentalnih podataka

Računske tehnike statističkih metoda sve su češće i sve se više primjenjuju u kontrolnim i istraživačkim radovima mlječne industrije. Mogu se upotrijebiti u mnogim slučajevima. Najvažniji praktični primjeri svode se na izračunavanje prosječne vrijednosti, varijance i standardne greške.

Vrlo su česti primjeri upotrebe statističkih metoda u utvrđivanju kvalitativne i kvantitativne veze između podataka dviju varijabli (korelacija i regresija) (D r a p e r i S m i t h, 1981). Na temelju nekoliko rezultata u parovima  $(x_1, y_1)$   $(x_2, y_2)$  ...  $(x_n, y_n)$  unijetih u korelacionu tablicu, izračunavaju se tzv. statističke brojčane karakteristike, kao što su:

#### Aritmetička sredina:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^m x_i \qquad \bar{y} = \frac{1}{n} \sum_{j=1}^n y_j \qquad (1)$$

#### Varijanca:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left( \sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2 \right) \qquad (2)$$

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} \left( \sum_{i=1}^n y_i^2 - n \cdot \bar{y}^2 \right) \quad (3)$$

**Kovarijanca:**

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y}) \quad (4)$$

$$s_{xy} = \frac{1}{n-1} \left( \sum_{i=1}^n x_i y_i - n \cdot \bar{x} \cdot \bar{y} \right) \quad (5)$$

**Koeficijent korelacije:**

$$\begin{aligned} r_{yx} &= \frac{s_x \cdot y}{s_x \cdot s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\left( \sum_{i=1}^n (x_i - \bar{x})^2 \right) \cdot \left( \sum_{i=1}^n (y_i - \bar{y})^2 \right)}} = \\ &= \frac{\sum_{i=1}^n x_i \cdot y_i - n\bar{x}\bar{y}}{\sqrt{\left( \sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2 \right) \left( \sum_{i=1}^n y_i^2 - n \cdot \bar{y}^2 \right)}} \end{aligned} \quad (6)$$

Koeficijent korelacije  $r_{xy}$  ima vrijednost od 0 do  $\pm 1$ , a njegov kvadrat se naziva koeficijent determinacije:  $B_{xy} = r^2_{xy}$  i iznosi  $0 \leq B_{xy} \leq 1$

Krivulja »normalne razdiobe« ima ovu jednadžbu:

$$y = \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu}{\delta} \right)^2}$$

Njeni osnovni parametri su  $\mu$  i  $\frac{\delta^2}{n}$  (aritmetička sredina i standardna greška).

Normalna neprekinuta funkcija:  $Z = (x - \mu) \sqrt{n/\delta}$  zadovoljava normalnoj razdiobi s parametrima  $\mu = 0$  i  $\delta = 1$  tj.  $Z$  ima razdiobu  $N(0,1)$ .

Prikazana uz pomoć empirijske disperzije glasi:  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ,  
a neprekinuta funkcija  $X^2 = \frac{(n-1) S^2}{\delta^2} = \frac{1}{\delta^2} \sum_{i=1}^n (X_i - \bar{X})^2$ .

Jednadžba regresije (linearna) glasi:

$y = a + bx$ , gdje je  $a$  — odrezak na ordinati  
 $b$  — koeficijent smjera ili

$$y = \bar{y} + b(x - \bar{x}); \quad b = \frac{s_{xy}}{s_x^2}$$

### Primjeri iz kontrole kakvoće mlijeka i mlječnih proizvoda

#### Primjer 1: Sistematska greška metode

Za analizu laktoze u mlječnom proizvodu upotrebljava se jodometrijska ( $M_1$ ) i polarimetrijska ( $M_2$ ) metoda. Postignuti su slijedeći rezultati:

(Tab. 1). Treba utvrditi sistematsku grešku ( $M_2$ )?

Broj pokusa	Laktoza ‰		d	$(d_i - \bar{d})$	$(d_i - \bar{d})^2 \cdot 10^{-4}$
	M-1	M-2			
1	2,20	2,25	- 0,05	0,06	36
2	2,87	2,82	0,05	0,04	16
3	3,12	3,20	- 0,08	0,09	81
4	3,62	3,48	0,14	0,13	169
5	3,90	3,89	0,01	0,00	0
6	4,30	4,35	- 0,05	0,06	36
7	4,95	4,90	0,05	0,04	16
Ukupno			+ 0,07		354

$$\bar{d} = \frac{0,07}{7} = 0,010 \quad s^2 = \frac{354 \cdot 10^{-4}}{6} = 5,9 \cdot 10^{-3}$$

$$s = 7,68 \cdot 10^{-2}$$

$$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}} = \frac{0,01}{\frac{7,68 \cdot 10^{-2}}{\sqrt{7}}} = \frac{0,01 \cdot \sqrt{7}}{7,68 \cdot 10^{-2}} = 3,445$$

Vrijednost iz statističkih tablica iznosi  $t_{0,05}(6) = 2,447$ .

Metoda (Trajković et. al., 1983) u odnosu na M-1 ima sistematsku grešku. Ponavljanjem mjerenja sa M-2 dobiveno je

	M-1	M-2	d	$d_i - \bar{d}$	$(d_i - \bar{d})^2 \cdot 10^{-4}$
1	2,20	2,26	- 0,06	0,23	529
2	2,87	2,87	0,20	0,03	9
3	3,12	2,82	0,30	0,13	169
4	3,62	3,42	0,20	0,03	9
5	3,90	3,95	- 0,05	0,22	484
6	4,30	4,05	0,25	0,08	64
7	4,85	4,50	0,35	0,18	324
			1,19		1588

$$\bar{d} = \frac{1,19}{7} = 0,17 \quad s^2 = \frac{0,1588}{6} = 0,0265$$

$$s = 0,163$$

$$t = \frac{0,17}{0,163} \sqrt{7} = 2,76; \quad t_{0,05}(6) = 2,447$$

Statističko tumačenje rezultata u ...

S obzirom da je eksperimentalna vrijednost  $t = 2,76$  veća od granične vrijednosti  $t = 2,447$ , zaključujemo da polarimetrijska (M-2) metoda ima sistematsku grešku u odnosu na jodometrijsku (M-1) i ona iznosi  $+ 0,17\%$  apsolutno (smanjivati treba rezultat).

Slučajna pogreška te veličine računa se iz formule:

$$\varepsilon = s \cdot \frac{t}{\sqrt{n} \cdot M}$$

$$t_{0,05} (7) = 2,365$$

$$\sqrt{n} = \sqrt{7} = 2,64575$$

$$s = \sqrt{0,0265} = 0,163$$

$$M \text{ za } f = 7 - 1 = 6 \text{ iznosi } 0,959$$

$$\varepsilon = \frac{0,163 \cdot 2,365}{2,64575 \cdot 0,959} = 0,1519 = 0,15$$

Prema tome, sistematska pogreška polarimetrijske (M-2) metode iznosi  $+ (0,2 \pm 0,2) \%$  apsolutno.

## Primjer 2

Analiziran je uzorak Na-kazeinata kod proizvođača i kupca i postignuti su slijedeći rezultati:

Pokazatelji	Rezultati istraživanja	
	Proizvođač	Kupac
Bjelančevine (dostavni uzorak) %	88,4	87,0
Broj ponavljanja	6	6
Varijance	0,295	0,690
Suha tvar %	95,0	94,8
Broj ponavljanja	4	4
Varijance	0,0088	0,0220
Bjelančevine na suhu tvar %	93,0	91,8
Klasa kakvoće	I	II

Prvo se uspoređuju rezultati varijanci suhe tvari:

$$F = \frac{0,0220}{0,0088} = 2,5 \quad f_1 = f_2 = 4 - 1 = 3; \quad F_{0,05} (3/3) = 9,28$$

Budući da je eksperimentalna vrijednost F- kvocijenta (2,5) manja od vrijednosti F- kvocijenta iz tablice, zaključujemo da su varijance homogene. Prema tome, može se izračunati prosječna varijanica koja će biti:

$$\bar{s}^2 = \frac{0,0220 \cdot 3 + 0,0088 \cdot 3}{3 + 3} = 0,0154$$

$\bar{s} = 0,1241$       $d = 95,0 - 94,8 = 0,2$  u suhoj tvari

$$t = \frac{d}{\bar{s}} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = \frac{0,2}{0,1241} \sqrt{\frac{4 \cdot 4}{4 + 4}} = 2,28$$

$$t_{0,05}(6) = 2,447$$

**Zaključak:** Razlika prosjeka suhe tvari nije statistički izrazita (nalazi se unutar pogreške mjerenja).

Usporedbom rezultata bjelančevina dobivamo:

$$F = \frac{0,690}{0,295} = 2,34 \quad F_{0,05}(5/5) = 5,05 \text{ (iz tablica)}$$

Varijance su homogene te se može računati prosječna varijanca  $\bar{s}^2$ :

$$\bar{s}^2 = \frac{s_1^2 f_1 + s_2^2 f_2}{f_1 + f_2} = \frac{0,690 \cdot 5 + 0,295 \cdot 5}{5 + 5} = 0,4925$$

$$\bar{s} = 0,702$$

$$d = 88,4 - 87,0 = 1,4 \quad t = \frac{1,4}{0,702} \sqrt{\frac{6 \cdot 6}{6 + 6}} = 3,454$$

$$t_{0,05}(5 + 5) = 2,228 \text{ (iz statističkih tablica)}$$

Razlika rezultata količine bjelančevina statistički je izrazita.

Za rješenje spora o klasi kakvoće kazeina nužna je SUPERANALIZA.

Rezultati superanalize bile su slijedeći:

Količina bjelančevina (dostavni uzorak)	88,7%
Broj određivanja	12
Disperzija (varijanca) ispitivanja	1,76
Količina suhe tvari	94,9%
Broj određivanja	4
Disperzija	0,0165

$$\text{Usporedbom rezultata superanalize i proizvođača: } F = \frac{1,76}{0,295} = 5,97$$

$F_{0,05}(11/5) = 4,72$  varijance **nisu homogene.**

$$d = 88,7 - 88,4 = 0,3$$

$$t = \frac{d}{\sqrt{\frac{s_s^2}{n_s} + \frac{s_p^2}{n_p}}} = \frac{0,3}{\sqrt{\frac{1,76}{12} + \frac{0,295}{6}}} = 678$$

$$c = \frac{\frac{s_s^2}{n_s}}{\frac{s_s^2}{n_s} + \frac{s_p^2}{n_p}} = \frac{\frac{1,76}{12}}{\frac{1,76}{12} + \frac{0,295}{6}} = 0,749$$

$$f = \frac{(n_s - 1)(n_p - 1)}{[n_s(1-C) + (n_p - 1)C]^2} = \frac{11 \cdot 5}{11(1 - 0,749) + 5 \cdot 0,749^2} = 16$$

$$t_{0,05}(16) = 2,120$$

Dakle, razlika prosjeka bjelančevina superanalize i proizvođača NIJE STATISTIČKI IZRAZITA.

Usporedbom rezultata superanalize s rezultatima kupca dobivamo:

$$F = \frac{1,76}{0,690} = 2,55; \quad F_{0,05}(11/5) = 4,72 \text{ (vrijednost iz statističkih tablica).}$$

varijance su **homogene**, te se može izračunati prosječna varijanća.

$$\text{Približna vrijednost prosječne varijance: } \bar{s}^2 = \frac{1,76 \cdot 11 + 0,690 \cdot 5}{11 + 5} = 1,42;$$

$$\bar{s} = 1,19 \quad d = 88,7 - 87 = 1,7$$

$$t = \frac{1,7}{1,19} \cdot \sqrt{\frac{12 \cdot 6}{12 + 6}} = 2,857 \text{ i } t_{0,05}(16) = 2,120 \text{ (vrijednost iz statističkih}$$

tablica)

Razlika prosjeka rezultata bjelančevina superanalize i kupca STATISTIČKI JE IZRAZITA. (rizik znač.  $p < 0,05$ ).

Prema tome, kupac je napravio grešku u analizi. DATI PROIZVOD ima I KLASU, kako je to ustanovio proizvođač.

## Primjer 3.

Trebna utvrditi ukupni broj živih bakterija u konzumnom mlijeku

Broj Petrij.	Razrjeđenje (faktor)	Broj bakt. ( $v_i$ )	Transfor. varijabla ( $x_i - \bar{x}$ ) $10^{-4}$	( $x_i - \bar{x}$ ) $\cdot 10^{-5}$	( $x_i - \bar{x}$ ) $Y_i$	$\hat{Y}$	( $Y_i - \hat{Y}$ )	( $Y_i - \hat{Y}$ ) $^2$
1	50	200	—	—	—	—	—	—
2	100	193	0,010	1,85	0,83	186	7	49
3	200	84	0,005	0,05	—0,06	92	8	64
4	500	46	0,002	1,37	—0,17	37	9	81
5	1000	18	—	—	—	—	—	—
Ukupno		323	0,017	3,27	0,60	—	—	194

Srednja vrijednost ovisne varijable:  $\frac{\sum y_i}{n \cdot \bar{x}} = \frac{323}{3 \cdot 0,0057} = 18889/\text{ml}$ Regresija  $\hat{Y} = b_0 + b_1 x_i$ 

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{0,60}{3,27 \cdot 10^{-5}} = 18349$$

$$b_0 = \frac{\sum_{i=1}^n Y_i}{n} - (b_1 \cdot \bar{x}) = \frac{323}{3} - (18349 \cdot 0,0057) = 3,07 = 3$$

$$\hat{Y} = 3 + 18349 x_i$$

$$[\bar{x} = 0,017 : 3 = 0,0057.]$$



$$\text{Varijanca adekvatnosti regresije (»oko regresije«): } s_{\text{ad}}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n - 2} =$$

$$= \frac{194}{3 - 2} = 194$$

$$s_{\text{ad}} = \sqrt{194} \doteq 14$$

$t_{0,10}(1) = 6,314$  (vrijednost iz statističkih tablica)

$$\text{Pogreška proračuna: } \frac{14 \cdot 6,314}{3 \cdot 0,0057} \doteq 5169$$

$$\left( \varepsilon = \frac{s \cdot t}{n \cdot \bar{x}} \right)$$

Prema tome, broj bakterija po 1 ml našeg istraživanog proizvoda iznosi od 19 000 do 24 000/ml. Rezultat je dobiven iz slijedećih podataka:

$$(18889 \doteq 19000) \left( \begin{array}{c} 18889 \\ + 5169 \\ \hline 24058 \doteq 24000 \end{array} \right)$$

Za precizniji rezultat trebalo bi povećati pokuse tj. raditi s razrjeđenjima 1 : 100, 150, 200, 300 i 500.

#### Primjer 4:

U klasi A-mlijeka dopušteno je najviše 75 000 ukupnih živih bakterija/l ml.

U klasi B-mlijeka broj bakterija može biti do 150 000/l ml.

Koristeći podatke primjera 3, istraženi uzorak treba uvrstiti u jednu od klasa i utvrditi svrsishodnost povećanja broja pokusa u datoj analizi. Broj ukupnih bakterija iznosio je 19000 — 24000/ml. Gornja granica intervala pouzdanosti za klasu A iznosi 75000/ml, budući da je rezultat istraženog uzorka daleko manji od 75000/ml, proizvod odgovara zahtjevima klase A te stoga povećanje točnosti analize nije svrsishodno.

#### Zaključak

U članku je na 4 primjera prikazana primjena statističkih metoda:

- (1) utvrđivanje sistematske pogreške,
- (2) rješavanje spora između rezultata analize proizvođača i kupca u postupku superanalize,
- (3) utvrđivanje ukupnog broja živih bakterija po mililitri u konzumnom mlijeku,
- (4) provjera kategorija kakvoće u dva uzorka mlijeka na temelju ukupnog broja živih bakterija.

Na temelju primjene jednostavnih statističkih formula u kontroli kakvoće mlijeka i nekih njegovih proizvoda, prikazana je veća pouzdanost i sigurnost u pravilnom tumačenju rezultata uobičajenih analiza u mljekarskoj industriji.

**Literatura**

- DRAPER, N. R., SMITH, H.: »Applied Regression Analysis« 2nd Ed. Wiley J., New York — Singapore, 70—136, 1981.
- HAMMING, R. W.: »Numerical Methods for Scientists and Engineers«, Mc. Graw-Hill Book Co., New York — Toronto — London — Tokyo, 1962.
- JURAN, J. M., GYNA, F. i BINGHAM, R. S.: »Quality Control Handbook«, 3th Ed. Mc. Graw-Hill Book Co., New York, 27—47, 1976.
- ISO — Standard 5725 »Precision of Test Methods — Determination of Repeatability and Reproducibility«.  
International Organization for Standardization. 1980.
- POTPARIĆ, M., MIJUŠKOVIĆ, M., MRVOŠ, N., SANCIN, B.: »Kontrola namirnica« Privredni pregled, Beograd, str. 431—493, 1979.
- TRAJKOVIĆ, J., MIRIĆ, M., BARAS, J., ŠILER, S.: »Analiza životnih namirnica« Tehnološko-metalurški fakultet, Beograd, str. 359—402, 1983.