

ANALIZA VARIANCE (ANOVA)

Z ANOVO (**A**nalysis of **V**ariance) lahko ločimo in ovrednotimo različne vire variacij in primerjamo več srednjih vrednosti ter presodimo ali je med njimi statistično pomembna razlika.

Variacije znotraj vzorcev (within-sample estimate)

$$\sigma_0^2 = \frac{\sum_i s_i^2}{h} = \frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)^2}{h \cdot (n-1)} \quad \text{Prostostne stopnje: } \nu = h \cdot (n-1)$$

h je število skupin rezultatov, n pa število rezultatov znotraj skupine.

Variacije med vzorci (between-sample estimate)

$$\sigma_0^2 = \frac{n \sum_i (\bar{x}_i - \bar{x})^2}{h-1} \quad \text{Prostostne stopnje: } \nu = h-1$$

Srednjo vrednost vseh rezultatov izračunamo po formuli $\bar{x} = \sum \bar{x}_i / h$.

Izračunamo $F_{h-1/(h(n-1))}$ in ga primerjamo s tabelarično vrednostjo za enostranski F -test pri 95 % ravni zanesljivosti. Če je tabelarična vrednost presežena, je med srednjimi vrednostmi statistično pomembna razlika, kar pomeni, da ima nadzorovani faktor, ki je v našem primeru način shranjevanja fluorescenčnega reagenta, statistično pomemben vpliv.

Vzroka sta lahko, da ena srednja vrednost izrazito odstopa od drugih, da se ena podskupina vrednosti izrazito razlikujeta od drugih, ali pa se vse srednje vrednosti statistično pomembno razlikujejo med seboj. To lahko presodimo, če srednje vrednosti uredimo po velikosti v naraščajočem vrstnem redu in za vsak par sosednjih vrednosti izračunamo razliko med njima. Te vrednosti primerjamo z najmanjšo statistično pomembno razliko (least significant difference; LSD), ki jo izračunamo po formuli:

$$LSD = t_{h(n-1)} \cdot s \cdot \sqrt{2/n}$$

Standardni odklon s , ki ga uporabimo v tej zvezi je izračunan iz variance za variacije znotraj vzorcev.

1. Presoditi želimo, če je razlika, če hranimo fluorescenčni reagent pri različnih pogojih. Podatki so v tabeli 1.

Tabela 1. Rezultati merjenja fluorescence reagenta hranjenega pri različnih pogojih

	Ponovitve			\bar{x}	s^2
A) Sveže pripravljen	102	100	101		
B) Po eni uri v temi	101	101	104		
C) Po eni uri pri blede svetlobi	97	95	99		
D) Po eni uri na svetlobi	90	92	94		

Izračunajte LSD in ugotovite, katere srednje vrednosti se statistično pomembno razlikujejo med seboj in to označite v preglednici z rezultati!

2. Radi bi ugotovili, če je v porah peska na različnih globinah različen masni delež vode. Podatki so v tabeli 2. Označite, katere srednje vrednosti se statistično pomembno razlikujejo med seboj!

Tabela 2. Delež vode v porah peska na različnih globinah

h (m)	w (%)						\bar{x}	s^2
7	33,3	33,3	35,7	38,1	31,0	33,3		
8	43,6	45,2	47,7	45,4	43,8	46,5		
16	73,2	68,7	73,6	70,9	72,5	74,5		
23	72,5	70,4	65,2	66,7	77,6	69,8		

3. Šest analitikov je šestkrat določilo paracetamol v tabletah (tabela 3). Ali je med njihovimi srednjimi vrednostmi statistično pomembna razlika?

Tabela 3. Rezultati določitve paracetamola v tabletah

Oseba	w (%)						\bar{x}	s^2
A	84,32	84,51	84,63	84,61	84,46	84,51		
B	84,24	84,25	84,41	84,13	84,00	84,30		
C	84,29	84,40	84,68	84,28	84,40	84,36		
D	84,14	84,22	84,02	84,48	84,27	84,33		
E	84,50	83,88	84,49	83,91	84,11	84,06		
F	84,70	84,17	84,11	84,36	84,61	83,81		

4. ANOVA je tudi ena od statističnih možnosti obravnave podatkov v Excelu. Uporabite jo za vsaj enega od obravnavanih primerov, da boste znali uporabiti podatke, ki jih dobimo v izpisu.