

EXCEL VBA ALKALMAZÁSA EGY BIOMETRIAI ESETTANULMÁNY PÉLDÁJÁN BEMUTATVA

Hampel György

Absztrakt: Az Excel táblázatkezelő program olyan szolgáltatásait alkalmazhatjuk hatékonyan, melyekkel nagymennyiségű adat egy-egy részével ismétlődően kell végrehajtani statisztikai kiértékeléseket. Szerepük alapján külön munkalapon kerültek elhelyezésre a kiértékelendő adatok, a statisztikai számítási eljárások és az eredményeket összesítő táblázatok. A munkalapok közötti adatmozgást az éppen szükséges adatok kiválasztásával Visual Basic for Application program látja el. Egy termény két termesztési módjának összehasonlítására mért 44 kémiai összetevő mennyiségének kiértékelése során csaknem félezerszer végrehajtott statisztikai számítás eredményét kapjuk meg összesítve, táblázatba rendezett formában az elkészített VBA programmal.

Abstract: Excel spreadsheet programs can be used efficiently to perform statistical evaluations repeatedly with a large amount of data. Based on their role, the data to be evaluated, the statistical calculation procedures and the tables summarizing the results were placed on separate worksheets. The data transfer between the worksheets and the selection of the required data is performed by Visual Basic for Application. The prepared VBA program creates the result of almost half a thousand statistical calculations in a cumulative, tabular form to evaluate the measured amount of 44 chemical components in order to compare two crop cultivation methods.

Kulcsszavak: Excel, VBA, statisztikai kiértékelés

Keywords: Excel, VBA, statistical evaluation

1. Bevezetés

A táblázatkezelő programokat jellemzően adatok feldolgozására, különböző számítási műveletek elvégzésére alkalmazzuk. A korai verzióikhoz képest funkcióik jelentősen kibővültek, s ezeket sokszor nem is ismerjük, valamint csak csekély hányadát alkalmazzuk mindennapi munkánk során. Pedig sokszor jelentősen segítené munkánkat, ha ismernénk és tudnánk alkalmazni őket. Ebben az esettanulmányban bemutatásra kerül, hogy az Excel Visual Basic for Application funkciója segítségével minimális, pár óra alatt elsajátítható programozói ismeret birtokában milyen nagy fegyver lehet a kezünkben az egyébként monoton ismétlődő számítási feladatok gyors elvégzésére.

2. Anyag és módszer

A statisztikai módszerekkel kiértékelendő adatok egy termény, mintánként 44 kémiai összetevőjének méréséből adódtak. A minták a termény három fajtájából, három egymást követő évben, két termesztési mód szerint álltak rendelkezésre. Azonos feltételek, mint fajta, év és termesztési mód esetében három párhuzamos mérés történt. Tehát $44 \times 3 \times 3 \times 2 \times 3 = 2376$ numerikus adatot kell statisztikai módszerekkel kiértékelni.

2.2. Az Excel és a VBA

A kiértékeléshez az Excel táblázatkezelő program biztosított minden szükséges háttérrel (Bártfai, 2016):

- Az adatok strukturált tárolása.
- A kiértékeléshez szükséges számítási támogatás.
- A kiértékelendő minták adatainak szűrése.
- Programozási környezet (VBA: Visual Basic for Application) az ismétlődő számítási eljárások automatizálásához (Walkenbach, 2013).

2.1. A Mann-Whitney-próba

A statisztikai kiértékelés során a két termesztési mód során adódó adatokat kellett összehasonlítani. Ezt a párhuzamos mérések háromelemű mintáin kellett elvégezni. A Mann-Whitney-próba ilyen kevés elemszámú minták esetén is alkalmazható (Vargha, 2002), ráadásul a kétmintás t-próbával szemben a normális eloszlás sem követelmény (Fidy–Makara, 2005; Sándor–Ádány, 2011). Az adatok egységes kiértékelésének céljából a Mann-Whitney-próbát alkalmaztuk.

3. Eredmények és értékelésük

Az adatok jól kiértékeléséhez több részfeladat megoldására volt szükség:

- Excel munkalap kialakítása az adatok olyan strukturált tárolása, mely biztosította az algoritmizálható kiértékelhetőséget.
- A Mann-Whitney-próba számításait elvégző munkalap elkészítése.
- VBA program készítése a beállított feltételeknek megfelelő adatok kiszűrésére és áthelyezésére a számításokat elvégző lapra.
- VBA program készítése, mely generálja a szűrési feltételeket és táblázatba rendezetten helyezi el az egyes szűrésekkel kapott statisztikai eredményeket.

3.1. Az adatok elrendezése munkalapon

Az adatok algoritmizálható visszakereshetőségét, szűrését kellett biztosítani az adatok megfelelő elrendezésével. Azaz, ha egy-egy sorszám jellegű adat jelzi a szűrendő adatsor esetén:

- a termény termesztési évét (1, 2 vagy 3),
- a termény fajtáját (1, 2 vagy 3),
- a kémiai összetevőt (1, 2, ..., 44),

akkor VBA programmal kiszámítható legyen, hogy a kétfajta termesztési mód esetén ismert három-három darab adat hol található a táblázatban. Az adatok megfelelő elrendezését az *1. táblázat* mutatja.

1. táblázat: Az adatok elrendezése

Év	Fajta	Termesztési mód	Párhuzamos mérés sorszáma	Összetevő1	Összetevő2	...	Összetevő44
2014	fajta1	mód1	1	3,230	4,575	...	16,992
2014	fajta1	mód1	2	2,869	6,561	...	16,614
2014	fajta1	mód1	3	2,632	7,931	...	19,182
2014	fajta1	mód2	1	5,749	9,183	...	16,614
2014	fajta1	mód2	2	5,874	9,851	...	15,557
2014	fajta1	mód2	3	5,076	9,517	...	16,086
2014	fajta2	mód1	1	5,537	8,849	...	10,799
2014	fajta2	mód1	2	7,034	8,849	...	11,403
2014	fajta2	mód1	3	5,500	9,266	...	12,008
2014	fajta2	mód2	1	2,295	10,602	...	19,333
2014	fajta2	mód2	2	1,571	10,602	...	18,125
2014	fajta2	mód2	3	1,609	10,518	...	18,200
...
2016	fajta3	mód1	1	5,924	11,019	...	19,408
2016	fajta3	mód1	2	4,328	8,598	...	12,234
2016	fajta3	mód1	3	5,076	10,685	...	15,783
2016	fajta3	mód2	1	2,819	10,685	...	23,260
2016	fajta3	mód2	2	2,856	10,936	...	21,447
2016	fajta3	mód2	3	3,642	12,271	...	22,505

Forrás: A szerző saját szerkesztése.

Hat sort igényel, hogy egy adott év és fajta esetén a két termesztési mód 3-3 párhuzamos méréséből származó 44 féle kémiai összetevő értékét tároljuk. Így 54 sorban soronként 44 adat található, azaz 2376 adat.

3.2. A Mann-Whitney-próba számításainak munkalapja

A statisztikai kiértékelés lapját úgy kellett kialakítani, biztosítsa:

- a szűrési feltételek beállíthatóságát,
- a szűrés végrehajtását,
- a szűrt adatok tárolását, megjelenítését,
- a szűrt adatoktól függő statisztikai számítás végrehajtását,
- a statisztikai számítás eredményének tárolását, megjelenítését.

A szűrési feltételek (év, fajta, mód) beállítását három cella biztosítja. Ezek a cellák felhasználóbarát módon legördülő listában jelenítik meg a választható lehetőségeket.

Amikor beállítottuk a feltételeket, egy parancsgombra kattintva indíthatjuk el azt a VBA programot, mely az adatokat tartalmazó munkalapról kigyűjti a

számítások munkalapjára a feltételeknek megfelelő adatokat. Az így megváltozott adatok miatt újraértékelődnek a számítások munkalap további cellái, azaz megkapjuk a statisztikai kiértékelés eredményét egy cellában.

A kialakított munkalap univerzális módon bármely két adatsor összehasonlítására alkalmas a Mann-Whitney-próba alapján, tehát bármikor alkalmazható, újrarahasznosítható.

3.2.1. A szűrés VBA programja

A szűrést elvégző, adatok() nevű szubrutin utasításait láthatjuk a következő sorokban (1. program).

1. program: A szűrést elvégző adatok() szubrutin

```
Public Sub adatok()
    ev = Worksheets("Számítások").Cells(1, 3)
    fajta = Worksheets("Számítások").Cells(2, 3)
    osszetevo = Worksheets("Számítások").Cells(3, 3)
    oszlop = osszetevo + 4
    kezdosor = (ev - 1) * 6 * 3 + (fajta - 1) * 6 + 2
    For i = 0 To 2
        Worksheets("Számítások").Cells(i+7,2) =_
            Worksheets("Adatok").Cells(kezdosor+i,oszlop)
        Worksheets("Számítások").Cells(i+7,3) =_
            Worksheets("Adatok").Cells(kezdosor+i+3,oszlop)
    Next i
End Sub
```

Forrás: A szerző saját szerkesztése.

Az első három utasítás beállítja az ev, fajta és osszetevo változók értékét a Számítások munkalap egyes celláiból kiolvasott értékekre. A következő utasítások kiszámítják a szűrendő adatok elhelyezkedését, s ez lesz az oszlop és a kezdosor változók értéke. Az adatok kiolvasását az Adatok munkalapról, és elhelyezésüket a Számítások lapra egy For ... Next ciklus végzi el. A ciklus biztosítja, hogy az utasítások ciklikusan többször végrehajtsanak. Erre azért van szükség, mert a kétféle termesztési mód 3-3 adatát kell kiolvasni egymást követő sorokból.

3.3. A számítások automatizálása

Kihaszználhatjuk, hogy a „Számítások” munkalapon újraértékelődik az eredmény amikor megváltoznak rajta az adatok. Így lehetőségünk van egy VBA szubrutinnal (2. program), hogy generálva egymás után a szükséges szűréseket, az ehhez tartozó eredményeket egy táblázatba gyűjtsük össze. A táblázatban évenként, fajtánként és összetevőnként megkapjuk a kétféle termesztési mód összehasonlításának eredményeit. Mindez 3 év, 3 fajta és 44 összetevő esetén $3 \times 3 \times 44 = 396$ ismétlése a Mann-Whitney-próbának. Az eredményeket a program egy 44 sorból és 9 oszlopból álló táblázatban helyezi el.

3.3.1. A számítássorozatot generáló VBA program

Az előző szubrutinhoz (1. program) képest az eltérés, hogy ez a szubrutin (2. program) az év, fajta és összetevo változók értékét nem kiolvassa egy munkalapról, hanem egy-egy For ciklusban állítja be, hogy milyen kezdőérték és végérték között változzanak. Így nem csak egy adott értékük mellett értékelődik ki a Mann-Whitney-próba, hanem a változók minden értékvariációja mellett. Ezért kellett úgy kialakítani a ciklusokat, hogy egymásba ágyazottak legyenek. Ez azt jelenti, hogy a külső ciklus változójában beállított egyes érték mellett a belső ciklus változójának minden egyes értéke esetén végrehajtnak az utasítások. Így a számítások minden természetési év, fajta és kémiai összetevő esetén végrehajtnak. Természetesen minden kiértékelést követően az eredmény a Számítások munkalapról átkerül az Eredmények munkalap táblázatába. Azaz automatikusan elkészül az eredményeket összesítő táblázat a VBA program segítségével.

2. program: A számítások automatizálását elvégző szamitas() szubrutin

```
Public Sub szamitas()
    For ev = 1 To 3
        For fajta = 1 To 3
            For osszetevo = 1 To 44
                oszlop = osszetevo + 4
                kezdosor = (ev - 1) * 6 * 3 + (fajta - 1) * 6 + 2
                For i = 0 To 2
                    Worksheets("Számítások").Cells(i + 7, 2) = _
                        Worksheets("Adatok").Cells(kezdosor+i, oszlop)
                    Worksheets("Számítások").Cells(i + 7, 3) = _
                        Worksheets("Adatok").Cells(kezdosor+i+3, oszlop)
                Next i
                Worksheets("Eredmények")._
                    Cells(osszetevo+9, (ev - 1) * 3 + fajta + 1) = _
                        Worksheets("Számítások").Cells(2, 9)
            Next osszetevo
        Next fajta
    Next ev
End Sub
```

Forrás: A szerző saját szerkesztése.

4. Következtetések, összegzés, záró megjegyzések, záró gondolatok

Az Excel táblázatkezelő nagy lehetőséget biztosít számunkra munkánk automatizálásához a Visual Basic for Application fejlesztőeszközeinek segítségével. Ezt ismétlődő számítási feladatok esetén tudjuk hatékonyan kihasználni.

Irodalomjegyzék

- Bártfai B. (2016): *Office 2016 – Word, Excel, Access, Outlook, PowerPoint*. BBS-Info Könyvkiadó, Budapest.
- Fidy J., Makara G. (2005): *Biostatisztika*. InforMed 2002 Kft., Budapest.
- Sándor J., Ádány R. (2011): *Biostatisztika*. Medicina Könyvkiadó, Budapest.

- Vargha A. (2002): Független minták összehasonlítása új rangsorolási eljárásokkal. *Statisztikai Szemle*, 80 (4): 354–370.
- Walkenbach, J. (2013): *Excel VBA Programming for Dummies*. 3rd edition. John Wiley & Sons Inc., New Jersey.