

## KOMPLETT SZTOCHASZTIKUS MODELLEK AZ AGRÁRÁGAZATBAN

*Ferenczy Antal, toni@hoya.kee.hu*

*Vargha András, vargha@ludens.elte.hu*

*Kertészeti és Élelmiszeripari Egyetem Matematika és Informatika Tanszék  
Eötvös Lóránd Tudomány Egyetem Általános Pszichológiai Tanszék*

### Stochastic linear models in agriculture

Antal Ferenczy,  
Department of Mathematics and Informatics  
Andras Vargha  
Eotvos Lorand University, Department of General Psychology

In the last couple of years we have encountered numerous empirical research in agriculture where the main goal was to assess the combined effect of one or more discrete grouping factors (e.g. plant species or types, level of temperature, lightening conditions) and a continuous concomitant variable (e.g. harvest/gathering time or number of flowers) upon the general level of a continuous dependent variable (e.g. flower characteristics, morfological measures, amount of crop, content of  $\beta$ -glucanas, number of tubers/bulbs developed).

Assuming that the concomitant variable exerts only a linear effect on the dependent variable these problems could always been adequately handled by means of a complex analysis of covariance design with the following general model:

$$\eta = \mu + \rho + \varepsilon,$$

where  $\eta$  is the dependent variable,  $\mu$  represents the complex linear ANOVA effect structure,  $\rho$  is the continuous regression effect and  $\varepsilon$  is the error term.

To perform an appropriate statistical analysis we applied several statistical programs (SPSS, Statgraf, Kovarian, Ministat). Our experience was that the simpler models could very easily and effectively handled by the programs Kovarian (written by the first author) and Ministat (developed by the second author).

### 1. Bevezetés

Az utóbbi néhány évben olyan komplett sztochasztikus modellek lehetőségével találkoztunk (értékelhető mérési adatsorokkal együtt), hogy felmerült bennünk egy komplett modellezési rendszer kidolgozása és irodalmi vagy saját forrásból megoldási módszer, számítógépes algoritmus kidolgozása.

A modell általános alakban:

$$\eta = \mu + \rho + \varepsilon \quad (1)$$

ahol  $\eta$  = eredményváltozó,  $\mu$  = varianciamodell,  $\rho$  = regressziós modell és  $\varepsilon$  = hibtag

## 2. Előzményeink

A bevezetésben vázlatosan ismertetett komplett sztochasztikus modell legegyszerűbb változatával Sváb János (1981) foglalkozik. Ez a modell egytényezős teljes véletlen elrendezésű varianciamodellt és egyváltozós lineáris regressziós modellt kapcsol össze. A Sváb János (1981) 407-424. oldalakon szereplő algoritmus alapján Ferenczy Antal készített Basic nyelvű programot Wellisch Péter instrukciói alapján. A programot 1993-ban alkalmaztuk először Komonyi Melitta (1993) diplomamunkája során alkalmaztuk publikáltan sikeresen (Élelmiszeripari szakkonzulensek: . Rezessyné dr. Szabó Judit és dr. Seres Gáborné dr. Mekis Erika.)

Az 1995/96-os tanévben a Kertészeti és Élelmiszeripari Egyetemen felmerülő következő feladatok indítottak bennünket a komplett sztochasztikus modellek további vizsgálatára:

- Hogyan befolyásolja az idő és a fajta a bársonyvirág virágzásdinamikáját? (Pluhár Zsuzsanna)
- Hogyan befolyásolja az idő és a taxon a bazsalikom morfológiai jellemzőinek változását? (Szabó Krisztina).
- Mi befolyásolja az alma termésmennyiségét? Milyen előrejelzési módszereket tudunk javasolni a termelőknek? (Kókai Zoltán)
- Mitől függ a nagyzezerjófű csirázási üteme? (Neumayer Éva- Tóth Andrea)
- Hogyan függ a Valeriana taxonok viselkedése a termőhelytől? (Neumayer Éva-Petheő Ágnes Flóra),

Vizsgálatainkhoz a következő software háttérrel fogtunk hozzá:

- Statgraf 5.1 : kovariáns hatást vizsgál a többváltozós varianciaanalízis módosított modelljében,
- SPSS 4.0: van kovariáns változatot elemző sztochasztikus modellje,
- Ministat 2.2: az egyszempontos csoportátlagvizsgálatnál együttesen és csoportonként vizsgálja a kovariáns változó lineáris regressziós hatását minősített korrelációs együtthatók segítségével (Vargha András software-fejlesztése,.)
- Kovarian: az 1989-ben Ferenczy Antal által megírt Basic-software és annak 1993-ban továbbfejlesztett változata.

### 3. Elvégzett vizsgálataink és eredményeink

A továbbiakban először felvázoljuk az egyes feladatok elemzésénél követett stratégiát, majd ezután néhány tipikus elemzés tapasztalatait, következtetéseit összegezzük.

#### 3.1 Követett módszertani lépések

A következő alfejezetekben az alábbi szempontok szerint tárgyaljuk az egyes modelleket:

- ismertetjük a feladatot az agrár szakember szavaival,
- leírjuk a matematikai modellt,
- felsoroljuk azokat a software-eket, amelyek legalább valamely részeredmény elérésére alkalmasak,
- ismertetjük konkrét alkalmazási tapasztalatainkat,
- javaslatot teszünk a szakmailag előremutató megoldás elérésére.

#### 3.2 Alma (*Malus domestica*) termésmennyiségének becslése a virágzatok száma alapján

Hevesen egy karcsúorsós ültetvényben figyeltük meg 5 almafajta esetén a virágzatok számát és ennek a termésmennyiségre gyakorolt hatását kerestük (A feladat Kókai Zoltán agrárinformatika szakos V. évfolyamos hallgató diplomamunkájából való. Konzulensei: dr. Edda Henze és Ferenczy Antal).

A feladat matematikai modellje:

$$\eta_{ij} = \mu + \alpha_i + \beta_i \varpi + \varepsilon_{ij} \quad (2)$$

ahol  $\eta_{ij}$  a becsült almatermés az  $i$ -ik fajta  $j$ -ik parcelláján,  $\mu$  a várható átlagtermés,  $\alpha_i$  az  $i$ -ik fajtától várható terméseltérés,  $\beta_i$  egy virágra jutó terméstöbblet az  $i$ -ik fajtánál,  $\varpi$  a virágszámot szimbolizálja a modellben,  $\varepsilon_{ij}$  az  $i$ -ik fajta  $j$ -ik parcelláján tapasztalt véletlen hiba.

Mivel modellünk a Sváb János (1981)-ben leírt modell kertészeti megvalósítása, ezért a Kovarian program egyszerű alkalmazásával oldottuk meg a feladatot. Sikeresen alkalmazhattuk volna még Vargha András Ministat-ját is.

Az illeszkedés erősségének hiányában nem kaptunk említésre méltó szakmai eredményt. A kísérletet elprelátóbb tervezéssel megismételjük.

#### 3.3 $\beta$ -glükánáz tartalom változásának vizsgálata

A sör egy igen fontos tulajdonsága, hogy a  $\beta$ -glükánáz tartalom minél hamarabb elérje a 250 mg/l értéket. Két hőmérsékleten 3 féle malátaalapanyaggal próbáltuk ki a FIA módszert. Az eljárás során percenként jegyeztük fel a  $\beta$ -glükánáz értékeket. Arra vagyunk kíváncsiak, hogy van-e különbség a különböző hőmérsékletek és alapanyagok hatására a folyamat lejátszódásában.

A feladat matematikai modellje a következő meglátások alapján építhető fel:

- a regressziós kapcsolat alapján véve hiperbolikus jelleget mutat,
- a második és harmadik maláta esetén a nevezőben felmerül harmadfokú polinom szerepeltetése,
- a variancia modellrészt kéttényezős teljes véletlen elrendezésű varianciaanalízissel javasoljuk modellezni.

Ha feladatunkat csak az első megjegyzés felhasználásával óhajtjuk megoldani, akkor a  $\beta$ -glükánáz értékek reciprokát képezve alkalmazhatjuk az előző alfejezetben javasolt megoldási módszert. Ha figyelembe

vesszük a második és harmadik megjegyzést, akkor már csak a később ismertetésre kerülő komplex ún. Statgraph-os elemzési módszert tudjuk javasolni.

A feladatban a következő szereposztást alkalmaztuk:

- eredményváltozó a  $\beta$ -glükánáz tartalom,
- kéttényezős teljes véletlen varianciamodell, ahol az egyik tényező a maláta 3-féle megválasztása, a B-tényező pedig a folyamatban alkalmazott hőmérséklet (40 és 50 °C),
- hiperbolikus regresszió a percben mért idő függvényében.

### 3.4 Bazsalikom (*Ocimum basilicum*) taxonok morfológiai jellemzőinek összehasonlítása

A nemesítési alapanyagok felmérése során a taxonok fejlődési ütemeinek különbségein keresztül azok koraiságára, elérhető magasságára és hozamára következtethetünk.

A feladat matematikai modellje a következő

$$\eta_{ij} = \alpha_i + \beta_i(1 - e^{-\gamma_i \tau}) + \varepsilon_{ij} \quad (3)$$

ahol  $\eta_{ij}$  az egyes bazsalikom növények magassága a  $\tau$ -ik időpontban az  $i$ -ik taxon a  $j$ -ik parcellán,  $\alpha_i$  az  $i$ -ik taxon induló magassága,  $\beta_i$  az  $i$ -ik taxon elérhető maximális magassága (de ezt élettani okok miatt soha sem éri el),  $\gamma_i$  az ún. simulási együttható,  $\varepsilon_{ij}$  pedig a hibtag.

A feladatot a Statgraph nonlinear regression nevű menüpontjával telítődési függvények ún. illesztési sorozatával oldottuk meg. Ez a programcsomag közli a taxononkénti paraméterbecslés 3 értékét és a becslés hibáját. Ebből eldönthetjük, hogy az egyes paraméterek becslései lényegesen különböznek-e 0-tól illetve a különböző taxonok azonos paraméterbecslései egymástól.

Két telítődési görbét akkor tekintünk statisztikailag azonosnak, ha mindhárom paraméterpárjának becslése statisztikailag azonos. Ez a paraméterek becsléséből, a hozzátartozó standard hibából és a megfelelő  $t$ -értékekből meghatározható. Nyilván egyetlen igazolt eltérés a fejlődési tendenciák különböző voltát igazolja.

Valamennyi taxon esetében statisztikailag igazolt telítődési görbét kaptunk. A teljes adathalmazra az összevont illesztést nem tudtuk elvégezni, mivel alapadatokból dolgoztunk s ez azt jelentette, hogy több mint 2000 adatsorunk volt.

### 3.5 Bársonyvirágok (*Tagetes erecta*, *T. patula*) fejlődési folyamatai

A bársonyvirág fajták tövenkénti virágszámának változását követtük nyomon az idő függvényében. A virágzásdinamika ismeretében meghatározhatjuk az egyes fajták optimális betakarítási időpontját és következtethetünk a fajták droghozamára.

Az előző alfejezet modelljéhez hasonlóan itt is egytényezős varianciaanalízist kapcsoltunk össze telítődési függvény illesztésével. Itt a tövenkénti virágszám volt, a fajta volt a tényező. Mivel itt parcellaátlagokkal számoltunk. Így a komplex modellt is realizáltuk tudtuk a Statgraph-fal. Mivel a megfigyelések túlmutattak a virágzási csúcson, ezért a teljes időszakra két szakaszban kellett volna illesztést végeznünk. Mi csak az első csúcsig terjedő szakaszt elemeztük.

### 3.6 Nagyezerjófű (*Dictamnus albus*) csírázási üteme

Napjaink nagyon fontos feladata a megbízható szaporítóanyag előállítása. Ezt mutatjuk be egy gyógyyn9vény csiráztatási adatsorának elemzésével. 3 giberelin koncentráció és az idő hatását vizsgáljuk a nagyzezerjófű csirázási eredményeire.

A matematikai modell az előző két feladattal azonos.

Ha az előzőhöz hasonló módszerrel, akár az alacsonyabb koncentrációk magasabbal ekvivalens voltát, akár a csiráztatási időszak rövidíthetőségét, akár a csirázási eredmény

#### 4. Általános következtetések

Minden időben lejátszódó folyamat a variancia- és a regressziós modell összekapcsolásával és az említett software-k felhasználásával az eddigi elemzéseknél lényegesen hatékonyabban elemezhető.

#### Irodalom:

- Ferenczy Antal(1991): Kísérletelemző információs rendszer (magyar nyelvű Quattro, Stagraph 3.0)
- Komonyi Melitta(1993) Maláta eredetű  $\beta$ -glükánáz enzimek vizsgálata flow injection (FIA) módszerrel. Diplomamunka KÉE.
- Marija J. Norusis(1990): SPSS-PC + Statistics 4.0 for the IBM PC/XT/AT and PS/2
- Sváb János(1981) Biometriai módszerek a kutatásban, Mezőgazdasági Kiadó.
- Szabó, Krisztina(1996): Morphological and chemical variability of basil genotypes. International Symposium, Breeding Research on Medicinal and Aromatic Plants June 30- July 4. 1996 Quedlinburg, Germany. in press.
- Petheő Ágnes Flóra (1996): Vadon előforduló és termesztett Valeriana taxonok összehasonlítása. Diplomamunka KÉE.
- Tóth Andrea (1996): Dictamnus albus környezeti igényeinek feltárása és szaporításának lehetőségei. Diplomamunka KÉE.
- Vargha András (1995) Ministat Felhasználói füzet