

Az eseménytörténet-analízis módszertana és alkalmazása a nyerstej minőségének vizsgálataiban

Kovács Sándor

Debreceni Egyetem Agrártudományi Centrum,
Agrárgazdasági és Vidékfejlesztési Kar,
Gazdaságelemzési és Statisztikai Tanszék, Debrecen
kovacs@agr.unideb.hu

ÖSSZEFOGLALÁS

Az eseménytörténet-elemzés az egyik leggyakrabban alkalmazott matematikai módszerek közé tartozik. Ez nagyrészt annak köszönhető, hogy számos modell típus áll rendelkezésünkre. A módszer ugyancsak kiváló kockázatelemzési eszköz. Főleg az orvostudományban és az állattenyésztési kutatásokban alkalmazzák a kockázat kezelésére. Ebben a cikkben bemutatom a főbb fogalmakat, a legalapvetőbb modelleket és elméleti alapjaikat. Foglalkozom a nemparametrikus Kaplan-Meier és a parametrikus Cox proporcionális kockázati modellel. Az esettanulmányt is ezen módszerek segítségével készítettem el. Az elemzések kimutatják, hogy a 16 állásos karusszel a leghatékonyabb berendezés a tartósan extra tejminőség előállítás szempontjából. A próbacsészébe történő fejés sokkal hatékonyabb, mint a padozatra történő fejés 16 dekád elteltével.

Kulcsszavak: eseménytörténet-analízis, túlélési analízis, tejminőség-vizsgálat, fejőberendezés, Kaplan-Meier analízis, élettartam, kockázat, kockázatelemzés, Cox-modell

SUMMARY

Event-history analysis is among the most frequently adopted mathematical methods, mainly due to the numerous types of models, available in the analysis. This method is also appropriate for analysing risk. Specifically, it has been employed in medical and animal-breeding science to manage risks. In this article, I give a presentation of the basic concepts, the basic types of models and their theoretical basis. I deal with the nonparametric Kaplan Meier and the parametric Cox proportional hazard model. The case study is also analysed with these methods. The results shows, that the 16-stall carousel is the most effective facility from the viewpoint of quality. Milking into assay crucible (sample-cup) is more effective than to the floor after 16 decades.

Keywords: event-history analysis, survival analysis, qualitative acceptance test of milk, milking machine, Kaplan-Meier analysis, duration, risk, risk analysing, Cox-model

BEVEZETÉS

A versenyképesség és a piacon maradás kockázati tényezőivel már hazánk Európai Unióhoz történő csatlakozását megelőzően is számos szerző foglalkozott (Ertsey, 1990; Ertsey és Drimba, 2003; Ertsey és mtsai, 2000), ezen kategóriáknak a jelentősége azonban a csatlakozást követően is fennmaradt, és tovább erősödött. Napjaink gazdasági sajátosságai miatt kizárólag azon gazdaságok számára lehetséges a versenyképes termelés, melyek folyamatosan jó minőségű, elfogadható árú terméket

állítanak elő. Ezen megállapítás érvényes a tejtermelőkre is (Buzás és Supp, 2001). A feldolgozó üzemek azon gazdaságokkal kötnek szerződést, melyek képesek hosszabb időn át extra minőségű tejet termelni, kiváltképpen annak köszönhetően, hogy az elmúlt időszakban szigorodtak a tej minőségével szemben támasztott követelmények is. Mivel az extra minőségű és az osztályon kívüli tej átvételi ára között jelentős különbségek vannak (Nagy és Felföldi, 1999), ezért alapvetően nem közömbös egyetlen tehenészet számára sem az, hogy milyen minőségű nyerstejet állítanak elő. Természetesen mindez szoros összefüggésben áll többek között az állattenyésztő telepek műszaki állapotának tényezőivel, valamint a technológia korszerűségével, amely sok esetben kívánivalót hagy maga után a mezőgazdaság tőkeszegénységének köszönhetően. Ennek ellenére az utóbbi években jelentősen (közel 20%-kal) növekedett az extra minőségű nyerstej aránya (Szabó és Popovics, 2002). A legújabb felmérések szerint a hazánkban előállított nyerstej mintegy 90%-a extra minőségű (Popovics, 2005). A kérdés tehát az, hogy milyen vizsgálatokat végezhetünk még el, ha szinte minden gazdaság extra minőségű nyerstejet állít elő. Az eseménytörténet-analízis néven ismert matematikai módszer éppen erre a kérdésre adja meg a választ. A módszer elsősorban súlyos betegségek különböző kezeléseinek összehasonlítására alkalmazzák az orvostudományi kutatásokban (főképpen a rákkutatásban), melyekben a vizsgált esemény a beteg halála, illetve annak időpontja a kezeléstől számítva. A vizsgált jelenség pedig a beteg túlélési valószínűsége adott kezelése hatására. Ennek a cikknek a keretében azt vizsgálom, hogyan lehet az analízist az állattenyésztési ágazatban a szarvasmarhatartó telepek esetében a tejminőség vizsgálataiban alkalmazni.

ALKALMAZOTT MÓDSZEREK BEMUTATÁSA

Az eseménytörténet-analízis alkalmazásával célom megmagyarázni azt, hogy az egyes technológiák (fejőberendezés típusa, első tejsugarak kifejtésének módja) esetében mennyire nagy a kockázata annak, hogy a nyerstej minősége egy-egy tartósan extra szakasz után osztályon kívüli lesz. Mindez speciális modellekkel végezhető el, mint például az esemény fennállásának ideje (élettartam modell), azaz a túlélési idő analízis (Vermunt és Moors, 2005). Az eseménytörténet és túlélési analízis, valamint a túlélési és kockázati modell

szakkifejezéseket többnyire egymás szinonimájaként használja számos szerző (Pötter és Rohver, 1999; Vermunt és Moors, 2005). Az alapvető modellek bemutatása előtt fontosnak tartom tisztázni a lényegesebb fogalmakat, és bemutatni a főbb képleteket a tej minőségének romlása példáján keresztül.

Az analízis megkezdése előtt definiálnunk kell a megkülönböztetendő állapotokat, melyeket vizsgálni szeretnénk. Az egyes állapotok a magyarázandó változó kategóriái. A tejminőség, mint megmagyarázandó változónak 2 kategóriája lehetséges: extra, illetve nem extra minőség. A lehetséges állapotok együttesen alkotják az állapotteret. Az esemény az egyik állapotból a másikba történő átmenetként értelmezhető. Egy lehetséges megfigyelhető esemény a tartósan extra minőség termelésének az ideje, azaz a tej minőségének a romlásáig eltelt idő. Ezt a sorozatos „extra” állapotból a „nem extra” állapotba történő átmenetként definiálhatjuk. A kezdőállapot az „extra”, a végállapot a „nem extra” kategória.

Amennyiben a tej minősége változó nem csak két kategóriából állna, akkor megfigyelhetnénk más eseményeket is. Például az „extra” minőségű tejmintát követhetné 1-es, 2-es, 3-as, vagy éppen „osztályon kívüli” minősítésű minta is. A saját vizsgálataimban csak két kategóriára (extra, nem extra) szorítkozom, mivel 2003. március hó 2. dekájától már csak ez a minősítési rendszer van érvényben, a mintában szereplő korábbi évekre vonatkozó minősítéseket pedig ennek megfelelően módosítottam.

A fent leírtak alapján az állapotter meghatározza a lehetséges események halmazát. Az is érthető, hogy ahhoz, hogy az esemény bekövetkezzen, a kezdőállapotnak fenn kell állnia. A kezdőállapot fennállása azonban még nem feltétlenül vonja maga után a végállapotot, vagyis fennáll a lehetősége annak, hogy az esemény nem következik be a megfigyelés ideje alatt. Ekkor csak részinformációval rendelkezünk, és az esemény bekövetkeztét csonkoljuk az adott egyeddel kapcsolatban. A kockázati időszakról vagy periódusról az indulóállapottól kezdve beszélünk, mert ebben az időszakban bármikor bekövetkezhet az esemény. A minőség romlásának kockázatát vizsgálva a kockázati időszak az „extra” minőség előállításával kezdődik meg, hiszen azután bármikor romolhat a tej minősége.

A csonkolás, vagy csonkítás azt jelenti, hogy nem történt minőségi romlás a vizsgálat ideje alatt, így csak részinformációkkal rendelkezünk. Az eseménytörténeti analízis ezeket a részinformációkat is fel tudja használni, s ezzel a túlélési idők alsó, de igen pontos becslését megadni. A szakirodalom ezen eseteket csonkított eseteknek nevezi (McGrady, 2005; Bolla és Krámlí, 2005).

Egy adott időpontban a kockázati halmazt azon egyedek alkotják, amelyeknél fennáll az esemény megtörténtének a veszélye. A fent definiált

fogalmakkal az eseménytörténet-analízis keretében azt az időtartamot vizsgáljuk, amelyben az esemény nem fordult elő a kockázati periódus alatt, illetve a két bekövetkezés között eltelt időtartam hosszát, azaz az élethosszát. A matematikai képletekben az élettartamot a T időváltozó jelöli, amely csak pozitív valós számot vehet fel. Az eseménytörténet-analízis definiálási módja nagymértékben függ attól is, hogy a T időváltozót folytonosnak vagy egész értékűnek vesszük. Habár a legtöbb alkalmazásban természetes módon a T változót folytonosként kezelik, ennek a feltételezésnek néha nincs valóságalapja.

Mivel a tejminőségre vonatkozó vizsgálatokban a mintát dekádonként veszik le, így a T időváltozót egészértékűnek kell tekinteni. A módszer az élettartamok statisztikai leírására törekszik, ezért a T változó valószínűségi változót jelöl, amely egyúttal azt is jelenti, hogy az összes, számunkra lényeges információt annak $F(t)$ eloszlásfüggvénye hordozza:

$$F(t) = P(T \leq t),$$

ahol t a vizsgálat tartamán belüli időt jelöli, kutatásaimban értéke 1-től 162 dekádig terjed.

Legyen $S(t)$ ennek komplementer-függvénye:

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t)$$

Ezt a függvényt – amely azt a valószínűséget adja meg, hogy az esemény nem fordul elő a t időpontig, azaz a vizsgált időtartam hossza legalább t – számos irodalom túlélési függvénynek nevezi. A T változó sűrűségfüggvényét jelölje $f(t)$, számítása a következő:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t},$$

feltéve, ha ez a határérték mindenütt létezik. Ennek teljesülése esetén az eloszlás folytonos. A sűrűségfüggvény – ha létezik – az eloszlásfüggvényénél sokkal rugalmasabban használható a valószínűségek leírására, kifejezésére.

Az eloszlásfüggvény, s ennek megfelelően a túlélési függvény a sűrűségfüggvény segítségével előállítható a következőképpen:

$$F(t) = P(T \leq t) = \int_0^t f(u) du$$

és

$$S(t) = P(T > t) = \int_t^{\infty} f(u) du$$

Az $S(t)$ függvény szigorúan monoton csökkenő alakú (1. ábra). A folyamat 1 valószínűséggel megéri a 0. időpontot, ez a valószínűség csökken az idő függvényében. Mértéke az utolsó időpillanatban 0 is lehet, ha sosem éri meg a folyamat az adott időpontot, ellenkező esetben 0-hoz közeli érték adódik.

1. ábra: A túlélési függvény alakja

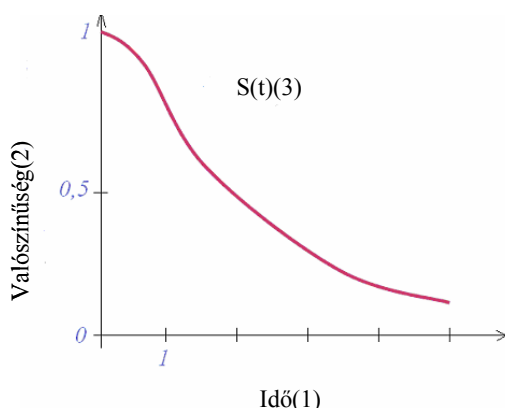


Figure 1: The plot of the survival function time(1), probability(2), survival function at time t(3)

Megjegyzendő, hogy $F(t)$ függvényt szokás halálzási valószínűségnek is nevezni (Ágoston és Kovács, 2000).

Olyan problémák is előállhatnak vizsgálataink során, amelyekre már az eloszlás függvények, sűrűségfüggvények nem képesek választ adni, vagy pedig részletesebb információra van szükség. Ezért kell olyan további központi fogalmakat bevezetni, mint például a halálzási intenzitás, vagy kockázati ráta fogalmak. Az arány fogalma a rátától eltérően azt fejezi ki, hogy az adott esemény összes bekövetkezéseiből mennyi hányadot tesznek ki a kedvező bekövetkezések. Rátát akkor kapunk, ha az arányt egy adott időszakra számoljuk ki, vagyis adott időszak kedvező bekövetkezéseit viszonyítjuk az időszaki összes bekövetkezéshez. Diszkrét esetben a halálzási intenzitás az alábbi formában adható meg:

$$h(t_i) = P(T = t_i | T \geq t_i) = \frac{f(t_i)}{S(t_i)}$$

Ez azt a feltételes valószínűséget jelenti, hogy az esemény bekövetkezett a t_i időpontban feltéve, hogy az nem következett be azelőtt. Itt az a koncepció jelenik meg, hogy a megfigyelő rendelkezik már egy bizonyos információval, vagyis azzal, hogy a t_i időpont előtt mi történt. Az intenzitásba feltételként be van építve ez az információ. A folytonos modellben a megfelelő definíció határértékkel történik az alábbi módon (Heien és mtsai, 2003):

$$h(t_i) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(T \in [t, t + \Delta t | T \geq t]) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}$$

A kockázati ráta vagy halálzási intenzitás egy olyan mérték, amely az adott esemény előfordulásának intenzitását adja meg. Ez folytonos esetben nem egy valószínűség, mivel egynél nagyobb értéket is felvehet (Vermount és Moors, 2005). A definícióból az is kiderül, hogy akkor és csak akkor létezik, ha a sűrűségfüggvény is létezik. Az eloszlás függvény, a túlélési függvény, a sűrűségfüggvény, a halálzási intenzitás ekvivalens leírási módjai egy pozitív valószínűségi változó eloszlásának (Vermount és Moors, 2005). Az egyik függvény

ismeretében a többi függvény azonnal leszámaztatható analitikusan. Diszkrét esetben a túlélési függvényt az alábbiak szerint származtatjuk:

$$S(t) = \prod_{j=1}^i (1 - h(t_j))$$

Saját vizsgálataimban a $h(t)$ függvény jelentése megadja annak a kockázatát, hogy a tej minősége éppen a t . dekádban romlik. Azt, hogy éppen a tej minőségének romlását vizsgálom, a fentiekben tárgyaltak indokolják. Ugyanis két minőségi romlás között eltelt idő éppen a folyamatosan extra tej előállításának hosszát adja meg. Ennek az időtartamnak a hosszát befolyásoló tényezők vizsgálata pedig rávilágíthat arra, hogyan javíthatjuk a nyerstej minőségét.

PARAMETRIKUS BECSLÉSEK

Az eseménytörténet-analízis során az $S(t)$ túlélési függvényt, illetve a $h(t)$ kockázati rátát állítjuk elő. A modellek két csoportra oszthatóak előállításuk módja alapján. Az egyik csoport a parametrikus modellek, a másik a nemparametrikus modellek csoportja. A parametrizálás tágabb értelemben véve magában foglalja az eloszlások tág halmazát, regressziót, néhány valós érték kölcsönös összefüggését. A parametrizálás számos célra szolgál: összefoglalja az adathalmaz jellemzőit, és a speciális vonásokra tereli a figyelmet. Egyszerű összehasonlításokat tesz lehetővé különböző helyzetek között, és még szimulációra is alkalmas.

Proporcionális kockázat modellek

A proporcionális modellek legfontosabb sajátosságát az adja, hogy az időtől való függés és a magyarázó változók között nem tételezünk fel kapcsolatot, ez a két tényező független egymástól. A magyarázó változóinkat valamilyen alapvető függvényrel (például exponenciális) transzformáljuk. Az alapvető kockázati függvényt a magyarázó változók transzformáltjával szorozzuk. Az alapvető kockázati függvény annak a kockázatnak a mértékét adja, amely akkor áll fenn, ha a magyarázó változóktól eltekintünk. Ennek jele: $h_0(t)$. Mivel a $h(t)$ kockázati függvény pozitív valós szám, ezért a transzformáltjának is pozitívnak kell lennie. Így a legkézenfekvőbb, ha a magyarázó változók exponenciális függvényével transzformálunk (Pötter és Rohver, 1999):

$$h(t | x, \beta) = e^{x\beta} h_0(t)$$

A szakirodalom ezt a modellt Cox proporcionális modellnek nevezi. A képletben x a magyarázó változók vektora, β a paraméterek vektora. Könnyen belátható, hogy $x\beta$ pozitív értékeire magasabb intenzitás adódik, mint az $x\beta = 0$ helyzetre. Ebből az következik, hogy minden t értéknél magasabb intenzitással az események egyre korábban következnek be, és az élethosszak így rövidülnek, míg alacsonyabb intenzitással (ha $x\beta$ negatív) az

események egyre később következnek be, hosszabbodó élettartammal. Tehát $x\beta$ pozitív értékeiből magasabb kockázati mértékek adódnak, míg negatív értékekre alacsonyabb kockázati mértékek adódnak. A modell ezen tulajdonságának szerepe lesz az esettanulmány paraméterbecslésénél.

NEMPARAMETRIKUS MÓDSZEREK AZ ESEMÉNYTÖRTÉNET ANALÍZISBEN

A Kaplan-Meier becslés

Ezt a módszert, melyet termék-ciklus becslésnek is neveznek, Kaplan és Meier (1958) vezette be. Arra az esetre fejlesztették ki, amikor adott egy sokaság, de semmiféle magyarázó változó nem áll rendelkezésünkre.

Megfigyeléseink egy adott esemény bekövetkeztéig eltelt időtartamot ölelnek fel minden egyéb információ nélkül. Anélkül, hogy felosztanánk a túlélési időket intervallumokra, közvetlenül is megbecsülhetjük a túlélés valószínűségeit a túlélési idők alapján.

A Kaplan-Meier becslés diszkrét időpontok esetére nyújt megoldást. Használhatjuk a túlélési idő mediánjának meghatározására, vagy pedig egyes időszakokra vonatkozó túlélési arány meghatározására. A számítás során feltételes valószínűségeket szoroztatni kell vennünk:

$$P(T \geq t_i) = P(T \geq t_i | T \geq t_{i-1}) = P(T \geq t_i | T \geq t_{i-1}) \cdot P(T \geq t_{i-1} | T \geq t_{i-2}) \cdot \dots \cdot P(T \geq t_0) = \prod_{j=1}^i \left(1 - \frac{d_j}{n_j} \right),$$

ahol T a túlélési idő, a t_1, t_2, \dots, t_i jelenti azokat az időpontokat, ahol a vizsgált esemény bekövetkezett, a d_j mutatja a t_j időpontban bekövetkezett események számát, az n_j pedig a t_j időpontban azon egyedek számát tükrözi, amelyeknél az adott esemény még bekövetkezhet (Wais, 2004). Az n értékek számításánál vesszük figyelembe a csonkított eseteket:

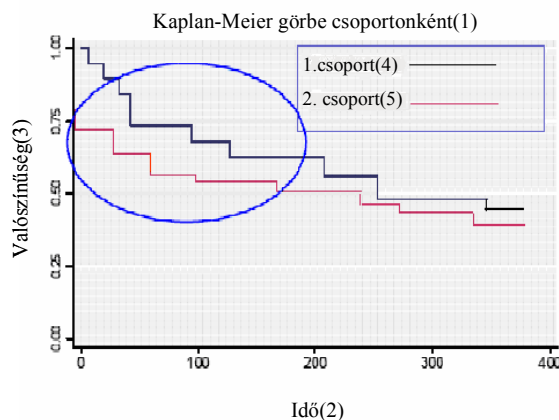
$$n_i = n_{i-1} - d_{i-1} - c_{i-1},$$

ahol c_{i-1} a t_{i-1} időpontban csonkított esetek számát jelenti.

A túlélési függvényre adott Kaplan-Meier görbék lépcsős alakúak. Két Kaplan-Meier görbe összehasonlításánál a közöttük lévő távolságot vizsgáljuk. A függőleges irányú rés azt szemlélteti, hogy egy adott pillanatban az egyik csoportnál mennyivel nagyobb a „túlélés” aránya a másik csoporthoz viszonyítva. A vízszintes távolság megfigyelésével azt olvashatjuk le az ábráról, hogy az egyik csoportnál mennyivel később következik be, hogy a túlélés aránya megegyezzen. Egzakt statisztikai tesztek is a rendelkezésünkre állnak ezen becslt függvények eltéréseinek vizsgálatára. A két leggyakrabban alkalmazott próba az általánosított Wilcoxon próba, vagy más néven Gehan teszt (2. ábra), és a log-rank próba (3. ábra).

Az első a tekintett időtartam elején lévő különbségekre, míg a második a folyamat végén lévőkre érzékenyebb.

2. ábra: A Gehan-próba alkalmazása

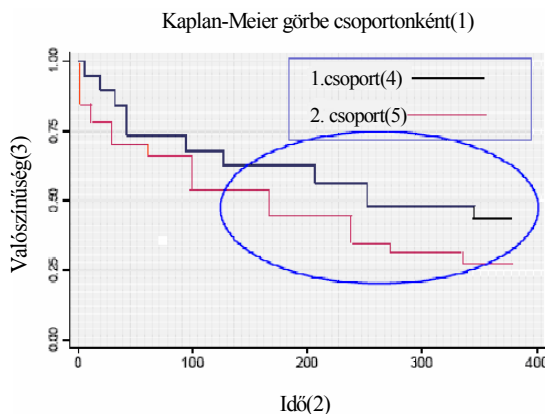


Forrás: McGrady (2005)

Figure 2: Application of the Gehan-test
Kaplan-Meier curve by group(1), time(2), probability(3), group 1(4), group 2(5)

A 2. ábrán látható, hogy a különbségek főképpen a vizsgált időszak elején mutatkoznak. Ezek kimutatására ez a próba a legalkalmasabb.

3. ábra: A log-rank teszt alkalmazása



Forrás: McGrady (2005)

Figure 3: Application of the log-rank test
Kaplan-Meier curve by group(1), time(2), probability(3), group 1(4), group 2(5)

A 3. ábra azt az esetet ábrázolja, amikor az idő előre haladásával várhatóak a lényeges eltérések a valószínűségekből. Ekkor alkalmazzuk a log-rank tesztet. Amennyiben a két görbe metszené egymást, úgy akármelyik teszt alkalmazható a szignifikáns különbség kimutatására.

A Kaplan-Meier módszer alkalmazhatóságának főbb feltételei, hogy a csonkolt és a nem csonkolt eseteknek függetleneknek kell lenniük, nem tartalmazhatnak rejtett magyarázó faktorokat, nem

lehet túl sok a csonkolt esetek száma, valamint, hogy az információ hiányában csonkolt eseteknek az időtől függetleneknek kell lenniük (Anonym, Prophet Statguide, 2005).

SAJÁT VIZSGÁLATOK

A Cox modell és a Kaplan-Meier becslés alkalmazása

Az elemzésekhez az adatokat a Tejlabor szolgáltatotta a 2000-2004 éveket átfogóan, 162 dekádon keresztül. A vizsgálatba 33 Hajdú-Bihar megyei telep került bevonásra. Mivel a minta a megye tejminőségi vizsgálat alatt álló telepeinek közel 50%-ából áll, ezért az reprezentatívnak tekinthető. A vizsgált telepeken megfigyeléseket végeztem a technológiára vonatkozóan (fejőberendezés típusa, korszerűsége, első tejsugarak kifejtésének módja). A fejőberendezések esetén a kevésbé korszerű berendezések közé soroltam Magda és Marsalek (2000) munkája alapján a sajtáros, tejvezetékes berendezéseket, s ezen berendezéseket kihagytam vizsgálataimból. Másrészt pedig a telepek jelentős részén nem ilyen típusú berendezésekkel dolgoznak. A korszerűbb fejőberendezések közé soroltam a fejőházi mobil, 16 állásos karusszal, a fejőházi nem mobil, index, poligon, halszállás berendezéseket. Az első tejsugarak kifejtése történhet próbacsészébe vagy padozatra. Szakmai szempontból ennek jelentősége lehet, ugyanis a próbacsészébe fejtéskor a tej minősége jobban kontrollálható egy esetleges fertőzés esetén.

Az eseménytörténet analízis nem parametrikus módszerét, vagyis a Kaplan-Meier becslést választottam a túlélési függvény becsléséhez a fejőkehely leválasztási módjának elemzésére. Az ábrát a Statistika 6.0 program segítségével

készítettem el. Kutatásaim során az alábbiakban leírt kérdésekre kerestem a becslés alapján választ az első tejsugarak kifejtésére vonatkozóan.

Az első tejsugarak kifejtésének technikáját tekintve, mi a valószínűsége, hogy az extra minőséget adott dekádban biztosítani tudják? Az esetleges eltérések a valószínűségekből hányadik dekádtól kezdve jelentkeznek? A fejőberendezésekre vonatkozóan az alábbi kérdéseket fogalmaztam meg.

Melyek a legkorszerűbb fejőházi berendezések a nyerstej minősége alapján? Mennyivel nagyobb a nem-extra minőség előállításának kockázata, illetve azonos dekádokat tekintve mennyi az esélye az extra minőségű tej nyerésének?

Az adatokra Cox proporcionális kockázati modellt illesztettem, melyet a LEM (Log-linear and event-history modelling with Expectation Maximization algorithm) program segítségével valósítottam meg (Van Tilburg, 2005).

Egy eseménynek ebben a tanulmányban az egy dekádon belüli, extra illetve nem extra minőségű tejminta számít. A túlélés pedig azt jelenti, hogy mennyi időn át (hány dekádon keresztül) maradt fenn az extra állapot, vagyis amint bekövetkezett egy nem extra minőségű minta vétele, akkor megszakad a túlélés ideje. Mindezekből következően az esemény (extra minőségű minta vétele) fennállási idő modelljével dolgoztam.

Csonkított eset akkor fordulhat elő, amikor folyamatosan extra minőségű volt a minta egészen a vizsgálat végéig. Azt azonban nem lehet tudni, hogy a vizsgálat utolsó dekádja után milyen volt a tej minősége, vagyis pontosan meddig tartott az extra minőség hossza. Ez természetesen azért történhet meg, mert a Tejlabor adatai csak egy bizonyos időszakra vonatkoznak. A próbacsészébe történő fejtés hatékonyabb, ha több mint 16 dekádon keresztül fenn akarjuk tartani az „extra” tejminőséget (4. ábra).

4. ábra: A tartósan extra minőségű tej időtartama az első tejsugarak kifejtésének módszerétől függően

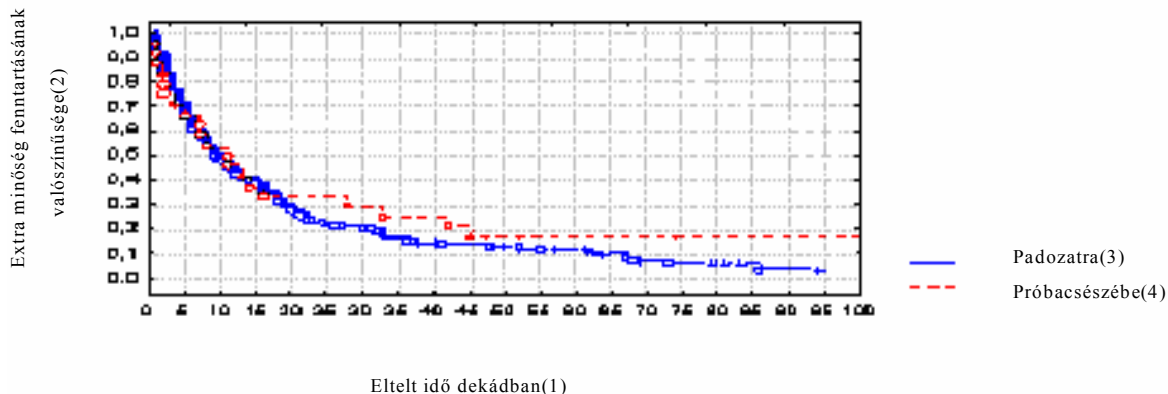


Figure 4: Durations of the steadily extra quality milk depending of the milking technique of the primary milk produced Durations (units are given in decades)(1), cumulative surviving proportion(2), to the floor(3), to assay crucible (milking cup)(4)

Az adatokat a 16. dekádtól tekintve a Gehan teszt 94% megbízhatósággal (p=0,06 szignifikanciával) kimutatja a különbségeket, melyek 10%-os szignifikancia szinten statisztikailag is jelentősek. A

próbacsésze használata hosszútávon statisztikailag igazolva előnyösebb lehet, hiszen segítségével kiszűrhető a beteg állatok tőgyéből kifejt tej. A 4. ábra alapján az is elmondható, hogy mindkét eljárást

tekintve 50% valószínűséggel a 10. dekádtól várható minőségi romlás. Tehát a folyamatosan „extra” minőség termelésének a mediánja 10 dekád.

A β paraméterek a proporcionális kockázati modelleknél bemutatott képletben szereplő paraméterek, amelyek alapján képezhetők az e^{β} relatív kockázati értékek. A kockázati függvény értékeinél azt az eseményt figyeljük, amikor minőségi romlás következik be. Az eredmények alapján a legkorszerűbbnek a 16 állásos karusszeles berendezés mutatkozott, mivel a paraméter alapján látható, hogy csökken a minőségi romlás kockázata (1. táblázat).

Ezzel a berendezéssel például közel kétszer akkora az esély arra, hogy „extra” minőségű tejet állítanak elő, mint az indexes berendezéssel. A szakemberek által korszerűnek vélt poligon, index, halszállka, berendezések a karusszelhez képest a nyerstej minősége szempontjából kevésbé hatékonyak bizonyultak. A karusszeles berendezés után a poligonos berendezéssel is igen jó eséllyel állíthatunk elő folyamatosan „extra” tejet, a halszállás és

indexes berendezésekkel közel ugyanolyan eséllyel állíthatunk elő „extra” tejet.

1. táblázat

A fejőberendezések kockázati függvényeinek paraméterbecslése a Cox-modell alapján

Fejő-berendezés(1)	Cox modell β paraméterei(2)	Relatív kockázati érték(3)	paraméterek a karusszelhez képest(4)
16 állásos karusszel(5)	-0,4784	0,6198	1
Poligon(6)	-0,0016	0,9984	1,611
Index(7)	0,2547	1,2900	2,08
halszállás(8)	0,2253	1,2527	2,021

Table 1: parameter estimates of the hazard function of the milking-machines by the Cox-model
milking-machine(1), parameters of the Cox regression model(2), relative risk value(3), parameters in the ratio of the caroussel(4), 16 stall caroussel(5), polygon(6), index(7), fishbone(8)

IRODALOM

Anonym, Prophet StatGuide (2005): http://www.quality-control-plan.com/StatGuide/survival_tests_ass_viol.htm (2005.08.18.)
 Ágoston K.-Kovács E. (2000): Halandósági modellek. Aula Könyvkiadó, Budapest, 12.
 Bolla M.-Krámlí A. (2005): Statisztikai következtetések elmélete. Typotex Könyvkiadó, Budapest, 110-111.
 Buzás, F.E.-Supp, Gy. (2001): How can small dairy farms cope with the EU requirements? Prospects for the 3rd millennium Agriculture konferencia, Kolozsvár, 153-156.
 Ertsey I. (1990): A kockázat mérésének módszertani kérdései a növénytermesztésben. Tiszántúli Mezőgazdasági Tudományos Napok előadásai-kivonatai, Debrecen.
 Ertsey I.-Drimba P. (2003): A kukorica terméseredményeinek elemzése a műtrágyázás függvényében, a kockázat figyelembevételével. In: Nagy J: Kukorica hibridek adaptációs képességének és terméshatóságának javítása. Civis-Copy Kft., Debrecen, 149-163.
 Ertsey, I.-Drimba, P.-Petró, Zs. (2000): Risk programming models for planning plant production, EURO XVII. 17th European Confer. on Operational Research, Bp., Hungary, July 16-19.
 Heien, H.C.-Baumann, W.A.-Rahman, M. (2003): Inferences in Log-Rate Models. http://www.mnsu.edu/research/URC/OnlinePublications/URC2003OnlinePublication/Heien_Baumann.doc (2006.03.14)
 Kaplan, E.L.-Meier, P. (1958): Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association, 53. 282. 457-481.
 McGrady, J. (2005): When Time is of Interest: The Case for Survival Analysis. John Hopkins University.

<http://www.twocw.net/jhsp/courses/StatisticalReasoning1/PDFs/Lecture7.pdf> (2005. 08. 18.)
 Magda S.-Marselek S. (2000): A tehenészet technológiája. In: Állattenyésztés (Szerk.: Magda S.-Marselek S.) Mezőgazdasági Szaktudás Könyvkiadó, Budapest, 85.
 Nagy, T.-Felföldi, J. (1999): Connections between cattle keeping and quality of milk. Debreceni Agrártudományi Egyetem Tudományos Közleményei, XXXIV. Debrecen, 175-180.
 Popovics P. (2005): A tejtermelés jelene és jövője az Európai Unió csatlakozását követően, Agrárgazdaság, Vidékfejlesztés, Agrárinformatika AVA-2 Nemzetközi konferencia CD-kiadványa, Debrecen
 Pötter, U.-Rohver, G. (1999): Introduction to Event History Analysis, <http://www.stat.ruhr-uni-bochum.de/scrip.html> (2005. 08.12.)
 Szabó G.-Popovics P. (2002): A tehenészeti ágazat helyzete a '90-es években, különös tekintettel az EU-integrációra, Agrártermelés-Életminőség, XXIX. Óvári Tudományos Napok CD-kiadványa, Mosonmagyaróvár
 Van Tilburg Egyetem statisztikai szoftvere (LEM) (2005): <http://www.uvt.nl> (2005. augusztus)
 Vermunt, J.K.-Moors, G. (2005): Event history analysis. In: B. Everitt and D. Howell (eds.): Encyclopedia of Statistics in the Behavioral Science, Wiley: Chichester, <http://arno.uvt.nl/show.cgi?fid=13313> (2005. 08.21.)
 Wais, R. (2004): Algorithmen für SPSS 12. <http://www.rz.uni-hamburg.de/RRZ/Software/SPSS/Algorithm120/km.pdf> (2005.08.08)