

---

# A magyar nagy fehér húsertés és magyar lapálysertés központi hízekonyságvizsgálatának (HVT) elemzése túlélés becslés (survival analysis) alkalmazásával

Nagy István<sup>1</sup> – Csató László<sup>1</sup> – Farkas János<sup>2</sup> –  
Radnóczy László<sup>3</sup> – Vígh Zsófia<sup>1</sup>

Kaposvári Egyetem, Állattudományi Kar,

<sup>1</sup>Sertés és Kisállattenyésztési Intézet, Kaposvár

<sup>2</sup>Matematikai és Informatikai Intézet, Kaposvár

<sup>3</sup>Országos Mezőgazdasági Minősítő Intézet, Budapest

## ÖSSZEFOGLALÁS

A szerzők magyar nagy fehér húsertés, illetve magyar lapálysertés hízekonysági és vágási teljesítményének vizsgálata (HVT) alapján a hizlalási napok számát elemezték a túlélés becslés (survival analysis) módszer felhasználásával. Mindkét fajta esetében ártány és emse csoportra egyaránt kiszámították az úgynevezett túlélés, illetve hazard függvényeket. A két ivar csoportjainak különbségét a log-rank teszttel ellenőrizték. A szerzők megállapították, hogy a fajtákon belül az ártány illetve emse csoportok függvényei statisztikailag igazolt módon eltérnek egymástól, nevezetesen az ártány egyedek a teljesítményvizsgálat során hamarabb érték el a vágási súlyt, illetve ha adott időpontig még egyik csoport sem érte el a 105 kg-os végsúlyt, a következő időegység alatt az ártány egyedek nagyobb valószínűséggel érték el a vágási súlyt, mint emse kortársaik.

## SUMMARY

Authors analysed the days of fattening in relation to the station test for Hungarian Large White and Hungarian Landrace pig populations using survival analysis. Survival and hazard functions were calculated in both groups separately for castrates and females. The log-rank test was applied to determine the significance of the differences across the groups. Authors established that within the breeds the functions of the castrates and females were significantly different, namely the castrates reached final weight sooner or, in cases where final weight was not reached by a certain period, castrates reached this weight during the next time unit with a higher probability than the females.

A túlélés becslés (survival analysis) egy olyan statisztikai eljárás, melynek alkalmazása abban az esetben javasolható, ha a vizsgálandó tulajdonság egy adott esemény bekövetkezéséig eltelt időtartam (a kísérlet kezdetétől mérve). Az úgynevezett esemény lehet elhullás, valamilyen betegség kialakulása, vagy bármilyen más történés, melynek bekövetkezési időpontja pontosan megállapítható. A vizsgált tulajdonság tehát az esemény jellegétől függően lehet életben maradási idő (valamilyen betegség esetében), egészséges állapot megtartásának időtartama (a betegség kialakulásáig) stb.

A kísérlet kezdete és az esemény bekövetkezése közötti időtartam mértékegysége tetszőleges, mérhető percben, órában, napban. A kísérletben résztvevő egyedekre vonatkozóan a vizsgálat kezdete és az esemény bekövetkezése közötti időtartamot túlélési

időnek (survival time) nevezzük és 't'-vel jelöljük. Fontos megemlíteni, hogy az úgynevezett túlélési idő fogalom használata az esemény jellegétől független. Ha pl. a vizsgálat célja annak elemzése, hogy adott sertéspopulációban hány nap telik el két fialás között, akkor az esemény bekövetkezése (fialás) nem kapcsolódik a biológiai értelemben vett túléléshez, illetve elhulláshoz. Ettől függetlenül a fialásig eltelt napok számát ebben az esetben is túlélési időnek nevezzük.

A túlélés becslés a rendelkezésre álló adatok alapján két függvényt állít elő, az eredmények értelmezése ezen függvények segítségével történik. Ezek a függvények az úgynevezett túlélési függvény (survival function), illetve hazard függvény (hazard function).

A túlélés függvény [S(t)], annak a valószínűségét adja meg, hogy a kísérletben résztvevők túlélési ideje nagyobb, mint a kísérlet kezdetétől számított 't' időtartam, vagyis 't' idővel a kísérlet megkezdése után az esemény még nem következett be. Az esemény jellegétől függően az egyed még nem pusztult el, a betegség még nem alakult ki, a koca még nem fialt újra stb. A függvény értékei a kísérlet megkezdésétől számítva nem növekednek (azonos szinten maradnak, vagy csökkennek), kezdeti értékük egy (a kísérlet megkezdésének pillanatában az esemény még nem következett be), végtelen idő múlva értékük nulla (elméletileg, ha a kísérlet végtelenül sokáig tartana az esemény minden résztvevő esetében bekövetkezne).

A hazard függvény [h(t)], azt az időegységre vetített valószínűséget adja meg, mely szerint 't' idővel a kísérlet megkezdése után az esemény a következő időegységben bekövetkezik. Ellentétben a túlélés függvényel, mely a túlélési időre vonatkozik (arra az időszakra, mely időtartamban az esemény nem következik be) a hazard függvény az esemény bekövetkezését jellemzi. Azt mutatja meg, mi a valószínűsége annak, hogy 't' idővel a kísérlet megkezdése után a következő időegység alatt az egyed elpusztul, a betegség kialakul, a koca ismét fial stb. (azaz az esemény bekövetkezik). Értéke nulla, illetve attól nagyobb szám lehet, melyek a kísérleti időegységtől függően is változnak (valószínűség / perc, óra, nap stb.). A két függvény tehát egymással ellentétes szemléletű, minél nagyobb adott 't' időpontban a túlélés függvény értéke, annál kisebb a hazard függvényé.

---

A túlélés becslés eljárás esetében fontos jelenség az úgynevezett „censoring”. A kísérletben résztvevők egy részénél a túlélési idő nem ismert. Ennek magyarázatát az adja, hogy a vizsgálat nem tarthat végtelen ideig, annak időtartamát valamilyen racionális időtartamban állapítják meg. Abban az esetben, ha a vizsgálati időtartamban az esemény nem következik be, akkor a túlélési idő a vizsgálatot végző számára ismeretlen. Ha tehát a kísérlet lezárása után a túlélési idő a kísérletet végző számára bármilyen oknál fogva ismeretlen, akkor beszélünk az úgynevezett „censoring” jelenségről. Ennek a jelenségnek a kezelése adja az eljárás egyik legnagyobb előnyét, mely az alapján történik, hogy a kísérlet végén a kísérleti alanyok állapotát két csoportba lehet sorolni. Ha az esemény bekövetkezett, akkor a túlélési idő ismert az állapotot jelölése „1”, míg ha az esemény nem következett be, akkor a túlélési idő ismeretlen („censored”), akkor ezt az állapotot „0”-val jelöljük. A hagyományos statisztikai eljárások, pl. ANOVA egyáltalán nem tudja kezelni azokat a megfigyeléseket, ahol a pontos túlélési idők nem állnak rendelkezésre, következésképpen az adatok egy része az értékelés számára elvesz. A túlélés becslés azonban ezeket a megfigyeléseket is beépíti az elemzésbe, vagyis könnyű belátni, hogy elemszámnövelő hatása miatt alkalmazása az eddig ismert adatstruktúra esetében előnyösebb, mint más statisztikai eljárásoknál.

A túlélés becslési eljárások alapvetően kétfélek. Ha az  $S(t)$  eloszlása pontosan ismert (általában nem ez a helyzet), pl. Exponenciális, Weibull, Gamma stb., akkor az úgynevezett parametrikus eljárásokat kell alkalmazni, melyhez speciális szoftver áll rendelkezésre (Ducrocq és Sölkner, 1998). Abban az esetben viszont, ha az  $S(t)$  eloszlása ismeretlen, akkor az úgynevezett nem parametrikus (nonparametric) módszereket célszerű használni, melyek valamennyi túlélés függvény paramétereit jól jellemzik, függetlenül annak konkrét (számunkra ismeretlen) eloszlásától. Az úgynevezett nem parametrikus eljárásokat az általános statisztikai programcsomagok is tartalmazzák, ezért az utóbbi módszerek használata gyakoribb, mint a parametrikus eljárásoké.

Jelen dolgozatban célkitűzésünk az volt, hogy a legfontosabb nem parametrikus módszerek (Kaplan és Meier, 1958) alkalmazása alapján egyes magyar sertéspopulációk (magyar nagy fehér hússertés, magyar lapálysertés) hizlalási napjaira megadjuk az úgynevezett túlélési ( $[S(t)]$ ), illetve hazard ( $[h(t)]$ ) függvényeket, melyek alapján a hizlalási teljesítmények Magyarországon eddig még nem vizsgált szempontok szerint értékelhetők.

Az elemzéshez használt adatokat az 1996-2001 közötti hízekonysági és vágási teljesítményének vizsgálata (HVT) során rögzített adatbázis szolgáltatta. Az elemzésben 6865 magyar nagy fehér hússertés, illetve 2758 magyar lapálysertés vett részt.

Az első fajtánál az egyedek 57, a másodikonál 34 tenyészetből származtak, a HVT vizsgálatot mindkét esetben 7 állomás végezte. A vizsgált tulajdonság a hizlalási napok száma volt. Az adatbázis néhány statisztikai jellemzőjét az 1. táblázatban adtuk meg.

1. táblázat

A vizsgált adatbázis egyes statisztikai jellemzői

Tulajdonság (1)	Fajta (2)	Egyed-szám (3)	Átlag (4)	Szórás (5)
Hizlalási napok száma(6)	magyar nagy fehér hússertés(7)	6865	80.1	14.1
	magyar lapálysertés(8)	2758	76.8	13.7

Table 1: Basic statistics of the analysed database

Trait(1), Breed(2), Number of records(3), Mean(4), Standard deviation(5), Days of fattening(6), Hungarian Large White(7), Hungarian Landrace(8)

Az adatok alapján a túlélési és hazard függvényeket Kaplan és Meier (1958) módszere alapján számítottuk ki. A csoportok függvényeinek összehasonlítását a log-rank teszt segítségével végeztük. Az említett túlélés becslési eljárások elvégzéséhez az SPSS (SPSS Inc, 1999) szoftvert használtuk.

A Kaplan-Meier módszer felhasználásával kapott túlélés függvények az 1. ábrán láthatók.

Az ábra alapján jól látható, hogy mindkét vizsgált fajta esetében az ártány csoport túlélési görbéi a kísérlet (HVT) egész ideje során alatta maradtak az emse csoport egyedek által mutatott túlélési görbéknek. Az y tengelyen az úgynevezett túlélési hányad látszik, vagyis az eddigi magyarázatoknak megfelelően az állomány azon hányada, melynél az esemény még nem következett be. Jelen esetben ezek az egyedek még nem érték el a 105 kg-os testsúlyt, amikor a kísérlet (HVT) befejeződik. Az eredmény tehát azt mutatja, hogy a hizlalási napok számának emelkedésével átlagosan több ártány esetében következik be az esemény, vagyis több ártány egyed éri el a vágási súlyt, mint emse kortársa. A túlélési görbéken jól látszik, hogy azon egyedek esetében melyek a kísérlet során elérték a vágási súlyt 80 nappal a HVT megkezdése után az egyedek mintegy 50%-nál az esemény bekövetkezett. Ez az arány a kísérlet megkezdése után 100 nappal már mintegy 80%-ra nőtt. A vágási végsúlyt 120 nappal a HVT megkezdése után majdnem az összes egyed elérte. Analóg módon, ha a HVT vizsgálat kezdete után bármely 't' időpontban a vizsgálati egyedek a 105 kg-os vágási súly még nem érték el, akkor a következő időegység során az ártány egyedek nagyobb valószínűséggel érték el a 105 kg-os testsúlyt, mint emse kortársaik (2. ábra).

1. ábra: A vizsgált sertépopulációk túlélés függvényei

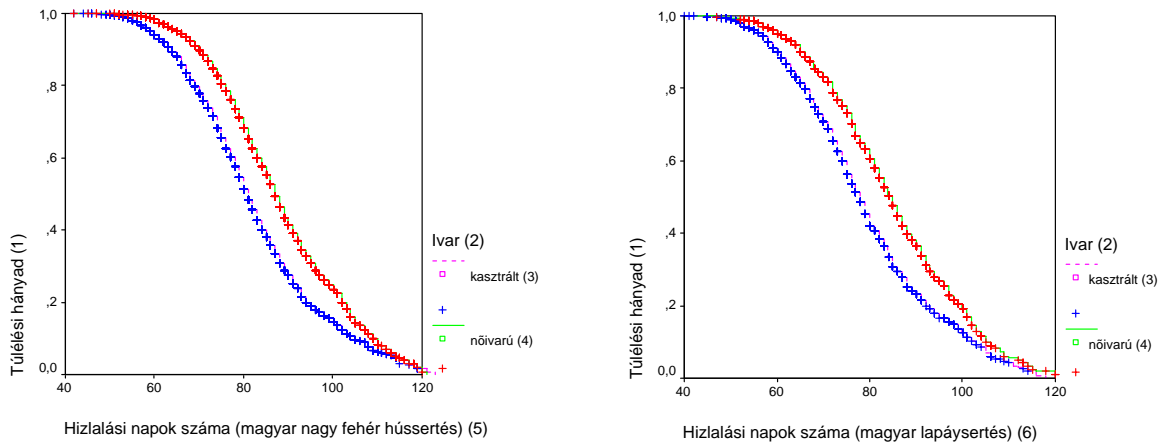


Figure 1: Survival functions of the analysed pig populations  
Cumulative survival(1), Sex(2), Castrates(3), Females(4), Days of fattening (Hungarian Large White)(5), Days of fattening (Hungarian Landrace)(6)

2. ábra: A vizsgált sertépopulációk hazard függvényei

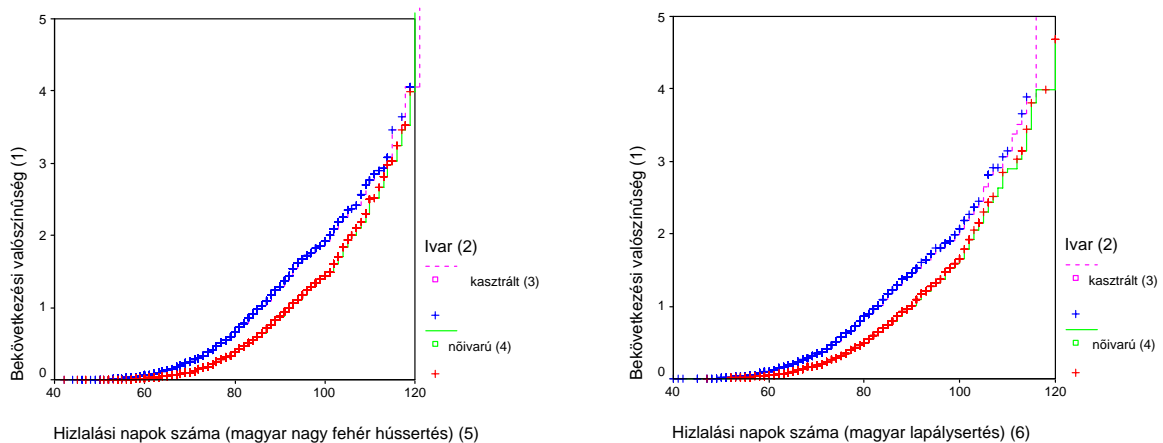


Figure 2: Hazard functions of the analysed pig populations  
Cumulative hazard(1), Sex(2), Castrates(3), Females(4), Days of fattening (Hungarian Large White)(5), Days of fattening (Hungarian Landrace)(6)

A log-rank teszt valamennyi bemutatott függvény esetében igazolta az ártány, illetve emse csoportok közötti különbségeket. Az eltérés minden esetben 0,1%-os szinten is szignifikáns volt. A magyar nagy fehér hússertés esetében a túlélési idő az egyedek 33,7%-ban, míg a magyar lapálysertések esetében 32,9%-ban ismeretlen („censored”), volt, melynek jelentése, hogy a vizsgálat tartamát 120 napra korlátozva az összes egyedszámuk ezen hányada nem érte el a 105 kg-ot. Meg kell azonban jegyeznünk, hogy a Sertés Teljesítményvizsgáló Kódex (OMMI, 2000) előírásai alapján a vágási végsúly  $105 \pm 2$  kg, vagyis az egyedeket már akkor is le lehet vágni, ha a 105 kg-os testsúlyt megközelítik. Ezek az egyedek nyilvánvalóan ismeretlen túlélési időt mutatnak, mely magyarázza az ismeretlen túlélési időt mutató egyedek nagy hányadát.

Sajnos az olyan közleményeket, melyek a sertések hízekonysági és vágási teljesítményének vizsgálatát a túlélés becslés módszerével elemezték volna, nem találtunk. Ugyanakkor megállapításunk, mely szerint az ártány csoport egyedei hamarabb érik el a vágási súlyt, vagyis jobb gyarapodást mutatnak, mint az emse csoport egyedei teljes mértékben megegyezik Horn és mtsai (1988) megállapításaival, akik részletesen elemezték a HVT eljárást és javaslatokat tettek az eltérő ivari teljesítmény alapján történő korrekcióra, amennyiben a vizsgálat pontosságát a különböző ivararányú elhullások torzítják.

A kísérlet alapján látható, hogy a túlélés becslés eljárás alkalmazása ideális az időskálán mérhető tulajdonságok esetében. A kapott függvények rendkívül szemléletesen mutatják a növekedés folyamatát. Hagyományos statisztikai módszerek

---

alkalmazásával (pl. t-próba, ANOVA) a jelenlegi adatbázis kevesebb mint 70%-át lehetett volna felhasználni a vizsgálati csoportok eredményeinek elemzése céljára, hiszen mindkét vizsgált fajta esetében az egyedek több mint 30%-a ismeretlen túlélési időt mutatott. Nem nehéz belátni, hogy a

jelen esetben 32-33%-kal több adat elemzése olyan előny, mely adott esetben szignifikancia kimutatásának lehetőségét jelentheti, vagyis az ismertetett eljárás az eddigiektől szélesebb körben történő használata mindenképpen javasolható.

#### *IRODALOM*

Duqrocq, V.-Sölkner, J. (1998): The Survival Kit-V3.0. A package for large analyses of survival data. Proc. 6-th WCGALP, Armidale, Australia 27. 447-448.

Horn, P.-Csató, L.-Baltay, L.-Farkas, J. (1988): Differences between the performance of sows and barrows and its influence on progeny test result. Acta Agronomica Hungarica, 37. 93-100.

Kaplan, E. L.-Meier, P. (1958): Nonparametric Estimation from Incomplete Observations. J. Am. Stat. Assoc. 53. 457-481.

OMMI (Országos Mezőgazdasági Minősítő Intézet) (2000): Sertés Teljesítményvizsgálati Kódex. 3. Budapest.

SPSS Inc. (1999): SPSS for Windows. Version 10.