

کاربرد روش آماری تحلیل بقاء در آنالیز طول عمر تولیدی گاوهای شیری

هادی اسفندیاری^۱، علی اصغر اسلمی نژاد^۲، مجتبی طهمورث پور^۳

چکیده

ارزیابی ژنتیکی طول عمر تولیدی در دام های اهلی به دلایل زیر نیازمند روش های خاصی است:

- ۱- برخی از حیوانات در زمان ارزیابی هنوز زنده بوده و طول عمر تولیدی کامل آنها شناخته شده نیست.
- ۲- اثراتی که طول عمر تولیدی را تحت تاثیر قرار میدهند خطی نبوده و با گذشت زمان تغییر می کنند.
- ۳- توزیع داده های طول عمر تولیدی معمولاً شناخته شده نبوده و چولگی زیادی دارند. مدل هایی که بر اساس تحلیل بقاء هستند ابزار آماری مناسبی را برای آنالیز طول عمر تولیدی فراهم می کنند. چنین مدلی بر اساس تابع بقاء $\lambda_0(t)$ هستند که خطر حذف حیوان را در زمان t به شرطی که حیوان تا زمان t ماندگاری داشته باشد تعریف می کنند. تابع مخاطره برای هر حیوان به صورت حاصلضرب تابع مخاطره پایه، $\lambda_0(t)$ ، که معرف گذشت زمان و افزایش سن حیوان بوده و برداری که شامل اثرات موثر بر طول عمر تولیدی بوده و با گذشت زمان تغییر می کنند مدل بندی می شود. از آنجایی که فرم $\lambda_0(t)$ با اطمینان شناخته شده نیست $\lambda_0(t)$ را میتوان همانند مدل کاکس اختیاری در نظر گرفت. ولی در مدل های پیچیده تر همراه با تعداد زیادی داده که در ارزیابی های ژنتیکی ملی مورد استفاده قرار می گیرد برای راحتی محاسباتی $\lambda_0(t)$ را حالت پارامتریک مانند تابع مخاطره وایبول در نظر می گیرند. در چنین مدلی می توان اثرات گاو نر تصادفی یا اثر حیوان را وارد نموده و پارامتر های ژنتیکی را نیز برآورد نمود. ارزیابی های ژنتیک معمول برای طول عمر تولیدی بر اساس مدل وایبول در بسیاری از کشورها مورد استفاده قرار می گیرد.

واژه های کلیدی: ارزیابی ژنتیکی، تحلیل بقاء، طول عمر تولیدی، گاو شیری.

۱- مقدمه

که در دوره های شیردهی بالاتری قرار دارند افزایش می یابد. طول عمر بیشتر سبب پاسخ انتخاب بیشتری نیز می شود، چرا که در این حالت حیوانات کمتری جایگزین شده و شدت انتخاب بیشتری برای گاوهای ماده امکان پذیر است. از دیدگاه اصلاحی، طول عمر تولیدی^۴ مورد توجه زیادی است چرا که مستقل از میزان تولید حیوان بوده و منعکس کننده باروری، سلامت و شایستگی کلی حیوان است. اهمیت طول عمر تولیدی در طی دو ده اخیر به شدت افزایش یافته چرا که این صفت تاثیر زیادی بر هزینه های تولید داشته و از

در واحد های دامپروری بالاخص در گاوداری های شیری طول عمر دام ها از اهمیت زیادی برخوردار است چرا که این صفت سود خالص را به میزان زیادی تحت تاثیر قرار می دهد. افزایش طول عمر سبب کاهش هزینه های مرتبط با پرورش یا خرید تلیسه های جایگزین در گله می شود و از سویی دیگر میانگین تولید گله به دلیل افزایش نسبت گاوهای

^۱ دانشجوی دکتری اصلاح دام دانشگاه فردوسی مشهد

^۲ عضو هیات علمی گروه علوم دامی دانشگاه فردوسی مشهد

^۳ عضو هیات علمی گروه علوم دامی دانشگاه فردوسی مشهد

^۴ Length of Productive Life

از سویی دیگر نیاز به بهبود این صفت در گله هایی که طی سالیان دراز تنها برای افزایش تولید در آنها انتخاب انجام شده، دیده می شود.

با این وجود، مورد توجه قرار دادن طول عمر در برنامه های اصلاحی و ارزیابی ژنتیکی حیوانات برای این صفت عموماً مشکل است. مشکل اصلی عبارت از این امر می باشد که تعدادی از حیوانات در زمان ارزیابی ژنتیکی هنوز زنده بوده و تنها کران پایین طول عمر نهایی آنها شناخته شده است. حذف چنین رکورد هایی از ارزیابی با در نظر گرفتن آنها به عنوان رکورد کامل سبب ایجاد اربیبی در نتایج می گردد. بنابراین چنین رکوردهایی بایستی به عنوان داده های سانسور شده^۱ در نظر گرفته شوند [۷].

مشکل دیگر در ارزیابی ژنتیکی طول عمر تولیدی این است که طول عمر کامل یک حیوان نتیجه حاصل ضرب عوامل موثر بر این صفت بوده و نه مجموع این اثرات [۱]. بنابراین زمانی که یکی از این اثرات ناشناخته باشند، طول عمر حیوان ناقص است [۹]. بنابراین روش های ارزیابی بر اساس مدل های خطی سنتی همانند BLUP^۲ اگرچه به طور وسیعی در اکثر کشور ها استفاده می شود با این حال این روش برای ارزیابی داده های غیر خطی طول عمر مناسب نیست. همچنین برخی از اثرات مانند میزان شیر تولیدی، اندازه گله و مدیریت که طول عمر تولیدی را تحت تاثیر قرار می دهند با گذشت زمان تغییر می کنند. از این گذشته، توزیع داده های طول عمر به شدت چولگی داشته و معمولاً مشخص نیستند. بنابراین روش هایی که بر اساس فرض نرمال بودن توزیع داده ها استوار هستند برای آنالیز طول عمر مناسب نیستند [۱۶].

تکنیک های آماری مورد استفاده در تحلیل بقاء بر اساس مدل بندی و تحلیل زمان های پاسخ است. زمان پاسخ یک فرد عبارت از متغیر تصادفی مثبتی است که فاصله بین نقطه آغاز معین و نقطه پایان را نشان می دهد. در آنالیز طول عمر تولیدی در گاوهای شیری نقطه آغاز معمولاً زمان اولین گوساله زایی و نقطه پایان (شکست) عبارت از زمانی است که حیوان از گله حذف شده و یا می میرد. این فاصله زمانی بر حسب روز، ماه و یا سال اندازه گیری می شود. عموماً، نقطه پایان می تواند هر رخداد دیگری نیز باشد به عنوان مثال بهبود بیماری و یا موفقیت پس از تلقیح مصنوعی و همچنین فاصله بین نقطه آغاز و پایان می تواند در مقیاس های دیگری غیر از زمان مانند کیلوگرم شیر تولیدی بیان شود.

ویژگی خاص تحلیل بقاء این است که با داده های سانسور شده وفق داشته و از این رو از اطلاعات دام های که در زمان ارزیابی هنوز زنده هستند استفاده می نماید. معمولترین نوع سانسور، سانسور کردن از راست^۳ است که در این حالت زمان شکست واقعی از مقدار مشاهده شده بیشتر است. از دلایل معمول سانسور کردن سمت راست عدم شکست (حذف) حیوان قبل از پایان مطالعه می باشد. در

تحلیل بقاء، به عنوان یک روش آماری که اساساً برای تحقیقات زیستی و مهندسی توسعه یافته می تواند در آنالیز داده های طول عمر مورد استفاده قرار گیرد. این روش آماری اطلاعات حاصل از دام های حذف شده (سانسور نشده) و حذف نشده (سانسور شده) را با یکدیگر ترکیب نموده و تحلیل آماری داده های سانسور شده را امکان پذیر ساخته و

۲- تکنیک های آماری تحلیل بقاء

تکنیک های آماری مورد استفاده در تحلیل بقاء بر اساس مدل بندی و تحلیل زمان های پاسخ است. زمان پاسخ یک فرد عبارت از متغیر تصادفی مثبتی است که فاصله بین نقطه آغاز معین و نقطه پایان را نشان می دهد. در آنالیز طول عمر تولیدی در گاوهای شیری نقطه آغاز معمولاً زمان اولین گوساله زایی و نقطه پایان (شکست) عبارت از زمانی است که حیوان از گله حذف شده و یا می میرد. این فاصله زمانی بر حسب روز، ماه و یا سال اندازه گیری می شود. عموماً، نقطه پایان می تواند هر رخداد دیگری نیز باشد به عنوان مثال بهبود بیماری و یا موفقیت پس از تلقیح مصنوعی و همچنین فاصله بین نقطه آغاز و پایان می تواند در مقیاس های دیگری غیر از زمان مانند کیلوگرم شیر تولیدی بیان شود.

ویژگی خاص تحلیل بقاء این است که با داده های سانسور شده وفق داشته و از این رو از اطلاعات دام های که در زمان ارزیابی هنوز زنده هستند استفاده می نماید. معمولترین نوع سانسور، سانسور کردن از راست^۳ است که در این حالت زمان شکست واقعی از مقدار مشاهده شده بیشتر است. از دلایل معمول سانسور کردن سمت راست عدم شکست (حذف) حیوان قبل از پایان مطالعه می باشد. در

^۱ Censoring data

^۲ Best Linear Unbiased Production

^۳ Right censoring

بیانگر احتمالی است که شکست در فاصله زمانی t و Δt اتفاق بیافتد.

تابع مخاطره^۳:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob} [t \leq T < t + \Delta t \mid T \geq t]}{\Delta t}$$

$$= \frac{f(t)}{S(t)} = -\frac{d \log S(t)}{dt},$$

بیانگر احتمال شرطی است که حیوان در فاصله زمانی $(t, t + \Delta t]$ با این فرض که تا زمان t ماندگاری داشته است دچار شکست شود. تمام روابط فوق بهم وابسته هستند.

۲-۲- برآورد تجربی تابع ماندگاری

تابع ماندگاری تجربی ما را از توزیع زمان های ماندگاری آگاه می سازد. تابع توزیع تجربی را می توان از طریق فرمول کاپلان - مایر^۴ محاسبه کرد:

$$\hat{S}_{KM}(t) = \prod_{k | T_{[k]} \leq t} \left(\frac{n_k - d_k}{n_k} \right).$$

در این فرمول $\hat{S}_{KM}(t)$ مقدار تابع ماندگاری در زمان t و $T_{[k]}$ نمایانگر زمان های شکست به ترتیب از کوچکترین به بزرگترین و d_k تعداد حیواناتی است که در زمان T_k دچار شکست شده اند. $\hat{S}_{KM}(t)$ برآورد حد حاصلضرب یا برآورد کاپلان - مایر تابع بقاء نامیده می شود. همچنین $\hat{S}_{KM}(t)$ برآورد درست نمایی ماکسیمم $S(t)$ شامل همه توزیع های ممکن است. در ضمن از $\hat{S}_{KM}(t)$ تابع مخاطره تجربی یعنی $\lambda(t)$ را می توان محاسبه نمود. برآورد های تجربی λ و S در مطالعات مقدماتی حائز اهمیت هستند چرا که می توان توزیع داده ها را با دانستن آنها به دست آورد و همچنین انتخاب مدل برای آنالیزهای بیشتر را فراهم کرد و از سویی امکان اعتبار سنجی این مدل ها را میسر می سازند [۵].

ارزیابی طول عمر در گله های گاو شیری داده سانسور شده به دلیل یکی از عوامل زیر بوجود می آید: حیوان در انتهای دوره مطالعه و جمع آوری داده هنوز زنده باشد، هنگامی که حیوان از یک گله به گله دیگر که تحت بررسی نمی باشد فروخته می شود و یا زمانی که کل حیوانات موجود در یک گله از برنامه ارزیابی حذف می شوند.

تحلیل بقاء همچنین دارای توانایی استفاده از داده های بریده شده است. در این سری داده ها نقطه آغاز خارج از زمان شروع جمع آوری داده است. یک مشاهده، زمانی گفته می شود بریده شده است که تاریخ اولین گوساله زایی دام زودتر از زمان آغاز جمع آوری داده ها باشد. از آنجایی که هیچ اطلاعاتی در زمان قبل از جمع آوری داده ها در دسترس نمی باشد، چنین دامی فرض می شود که تنها پس از آغاز جمع آوری داده ها در معرض خطر حذف باشد. رکوردهای بریده شده برخلاف رکوردهای سانسور شده که دارای اطلاعات جزئی هستند رکوردهای کاملی در تحلیل ها به حساب می آیند.

نوع دیگر سانسور، سانسور شدن از سمت چپ^۱ است که در این حالت حذف حیوان (شکست) قبل از نقطه شروع اتفاق می افتد. این نوع سانسور شدن در اصلاح دام مورد توجه قرار نمی گیرد.

۲-۱- توزیع زمان های شکست

تحلیل داده های بقاء بر اساس استفاده از توزیع و تابع های خاصی می باشد [۱۳]. تابع ماندگاری^۱:

$$S(t) = \text{Prob}(T \geq t) = 1 - \text{Prob}(T < t) = 1 - F(t).$$

بیانگر این احتمال است که حیوان حداقل تا زمان T ماندگاری داشته باشد. $S(t)$ نسبت حیواناتی است که در زمان t زنده هستند. $F(t)$ تابع چگالی احتمال تجمعی می باشد.

تابع چگالی احتمال $f(t)$:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{Prob} [t \leq T < t + \Delta t]}{\Delta t} = -\frac{d S(t)}{dt},$$

³ Hazard Function

⁴ Kaplan - Meier

¹ Left censoring

² Survival function

۲-۳- مدل های رگرسیونی

در بسیاری از حالات، خصوصیات اصلی تابع ماندگاری یا تابع چگالی شناخته شده نبوده اما برخی اطلاعات در مورد تغییرات میزان شکست $\lambda(t)$ در دسترس می باشد. بنابراین، مدل های تحلیل بقاء معمولاً از روی تابع مخاطره که ریسک حذف حیوان در زمان t را نشان می دهد ساخته می شوند. فرض کلی در این حالت این است که تابع مخاطره برای هر حیوان دارای فرم پایه ای مشترکی است که برای تمامی حیوانات یکسان بوده و مفهوم میانگین کل را دارد. تابع مخاطره پایه هر حیوان به وسیله اثراتی که حذف حیوان را تحت تاثیر قرار داده و فاکتورهای استرس نامیده می شوند تغییر می کند. مدل ماندگاری در فرم عمومی را می توان به شکل زیر نوشت:

$$\lambda(t, x_i) = \lambda_0(t) \exp(x_i \beta).$$

که $\lambda(t, x_i)$ بیانگر مخاطره برای حیوان i ام بوده، $\lambda_0(t)$ عبارت از تابع مخاطره پایه برای هر حیوان بوده و تنها به زمان وابسته است. در مدل های پارامتریک فرض می شود که تابع مخاطره پایه از توزیع نمایی خاصی پیروی می کند که این توزیع معمولاً یا توزیع نمایی است (مدل نمایی) و یا توزیع وایبول (مدل وایبول) می باشد. از سویی دیگر تابع مخاطره پایه در مدل نیمه پارامتری یا مدل کاکس اختیاری در نظر گرفته می شود. x_i عبارت از بردار متغیرها (فاکتورهای استرس) بوده که حذف حیوان را تحت تاثیر قرار می دهند و β عبارت از بردار ضرایب رگرسیون می باشد.

۲-۴- روش برآورد

برآورد اثرات متغیرها بر منحنی ماندگاری بر اساس روش های درست نمایی ماکسیم صورت می گیرد. با این حال ساخت تابع درست نمایی در مدل های نیمه پارامتری پارامتری نیازمند روش های متفاوتی است. در مدل های پارامتریک می توان نشان داد که سهم داده سانسور نشده در درستنمایی برای برابر با مقدار تابع چگالی در زمان شکست می باشد و سهم داده سانسور شده در

درست نمایی برابر با مقدار تابع ماندگاری در زمان سانسور شدن می باشد [۴]. در نتیجه تابع درست نمایی کامل از حاصلضرب سهم مستقل هر یک از رکورد ها به دست می آید. سپس برآورد پارامترها از طریق ماکسیم کردن این تابع یا لگاریتم آن به دست می آید [۷]. پارامترهای تابع مخاطره پایه نیز همراه با ضرایب رگرسیون برآورد می شوند.

برای مدل های نیمه پارامتری (مدل کاکس) روش متفاوتی مورد استفاده قرار می گیرد. برآورد β از طریق ماکسیم کردن تابع درست نمایی و بدون ایجاد هیچ فرضی در مورد $\lambda(t)$ صورت می گیرد. کاکس اصول روش درستنمایی جزئی را برای اولین بار ارائه داد [۲]. تابع درست نمایی جزئی بخشی از تابع درست نمایی کامل بوده و به $\lambda_0(t)$ وابسته نیست. این تابع از طریق درست نمایی حاشیه ای رتبه ای زمان های شکست بدست می آید.

پس از حداکثر نمودن تابع درستنمایی جزئی، تابع مخاطره پایه و یا تابع ماندگاری پایه را می توان با استفاده از برآوردهای انجام شده برای ضرایب رگرسیون همانند روشی مشابه با روش حد حاصلضرب (کاپلان - مایر) محاسبه کرد.

۲-۵- مدل مخاطره نسبی و بسط آن

یکی از مدل های که در تحلیل طول عمر به طور وسیعی مورد استفاده قرار می گیرد مدل مخاطره نسبی است که در این مدل، مخاطره دو حیوان A و B در طی زمان با میزان مخاطره $\exp(x_A - x_B)\beta$ ثابت می ماند. با این وجود در برخی حالات فرض مخاطره نسبی همواره برقرار نیست چرا که مخاطره برای هر یک از حیوانات در طی زمان به صورت متفاوتی بسته به شاخص های حذف در گله تغییر می نماید.

یک راه برای حل این مشکل که امکان استفاده از مدل مخاطره نسبی را امکان پذیر می باشد لایه بندی^۱ نام دارد. در لایه بندی برای لایه های مختلف (به عنوان مثال حیواناتی که در یک سال در یک گله متولد شده اند و سطح تولید یکسانی

^۱ Stratification

دارند می توانند یک لایه باشند) تابع های مخاطره متفاوتی تعریف می شود.

خصوصیت دیگر مدل های مخاطره نسبی این است که در این مدل ها می توان متغیرهای که مسئول تغییر در سیاست حذف در طی زمان هستند همانند اثر گله در سال^۱ یا تعداد شیر دهی را در مدل وارد کرد. در این حالت، محور زمان به فواصل زمانی متعددی تفکیک می شود و به طوری که در هر فاصله زمانی فرض مخاطره نسبی برقرار بوده اما میزان مخاطره در هر فاصله زمانی متفاوت است. با متغیرهای وابسته به زمان می توان سطوح متفاوت اثرات ثابت را به یک رکورد منفرد طول عمر تولیدی مرتبط ساخت.

۲-۶- تفاوت مدل کاکس و وایبول در اصلاح دام

از مدل کاکس در اصلاح دام به دلیل اینکه نیازی به داشتن اطلاعات در مورد توزیع زمان های ماندگاری که در اغلب حالات شناخته شده نیستند، نمی باشد به طور وسیعی استفاده می شود. با این وجود زمانی که مدل کاکس در برگیرنده کواریت های وابسته به زمان متعددی با سطوح زیاد باشند، برآورد ضرایب رگرسیون، β ، نیازمند محاسبات پیچیده ای می باشد. بنابراین چنین مدلی در ارزیابی های ملی با تعداد زیادی داده کاربرد ندارد.

در صورتی که تابع مخاطره پایه فرم پارامتری مشخصی داشته باشد، برآورد β و $\lambda_0(t)$ راحت تر است [۳]. به طور کلی مناسبترین فرم پارامتری که برای تابع مخاطره پایه در نظر گرفته می شود توزیع وایبول می باشد. در مدل رگرسیون وایبول تابع مخاطره پایه، $\lambda_0(t) = \lambda \rho (\lambda t)^{\rho-1}$ با دو پارامتر λ و ρ تعریف می شود که به ترتیب نشان دهنده پارامتر مقیاس و شکل تابع مخاطره پایه می باشند.

انتخاب توزیع وایبول به دلیل سادگی تابع ماندگاری وایبول، $S_0(t) = \exp(-(\lambda t)^\rho)$ ، و قابلیت انعطاف آن می باشد. در مدل رگرسیون وایبول بسته به مقدار پارامتر ρ ، تابع مخاطره می تواند فرم ثابت ($\rho = 1$)، صعودی ($\rho > 1$) و یا نزولی ($\rho < 1$) داشته باشد. مزیت دیگر مدل وایبول امکان

۲-۷- مدل های شکنندگی^۲

بسط آن به مدل های ماندگاری مختلط است که می توانند اثرات تصادفی همچون رابطه خویشاوندی بین نرها را شامل شوند [۸].

در مدل های مخاطره نسبی چه پارامتری و چه نیمه پارامتری می توان اثرات تصادفی (ژنتیکی) را وارد نمود. چنین مدل های ماندگاری مختلطی عموماً مدل های شکنندگی نامیده می شوند. عبارت V_m بیانگر اثرات ژنتیکی در مدل های شکنندگی بوده که به صورت حاصلضربی تابع مخاطره هر حیوان یا گروهی از حیوانات را تحت تاثیر قرار می دهد. برای مثال، مخاطره برای حیوان m را می توان به صورت زیر نوشت:

$$\lambda(t, x_m) = \lambda_0(t) v_m \exp(x_m \beta),$$

جزء شکنندگی نشان دهنده بخشی از واریانس است که در مدلی که تنها اثرات ثابت را دارد، نمی توان محاسبه نمود. تبدیل ساده $s_m = \log v_m$ امکان قرار دادن جزء V_m در عبارت $\exp(x \beta)$ مدل رگرسیونی را ممکن می سازد. در صورتی که Z نشان دهنده وکتور طرح برای اثرات تصادفی $S = \{s_m\}$ باشد مدل ماندگاری مختلط را می توان به صورت زیر نوشت:

$$\lambda(t, x, z) = \lambda_0(t) \exp(x \beta + z' s).$$

به دلیل قابلیت انعطاف و راحتی محاسباتی توزیع گاما فرض می شود که جزء V_m در مدل های ماندگاری مختلط دارای این توزیع می باشد [۱۱]. اما با این حال این توزیع با توزیع نرمال چند متغیره ای که در اصلاح دام تحت مدل پلی ژنتیک نام محدود فرض می شود مطابقت ندارد [۸]. با این وجود، می توان نشان داد که برآوردهای بدست آمده برای توزیع گامای جزء v در جمعیت های گاوهای شیری نسبتاً بزرگ بوده و این بدین معنی است که v تقریباً داری توزیع لگ نرمال بوده و در نتیجه S دارای توزیع نرمال است [۱۱].

چندین روش برای برآورد پارامترهای توزیع v که هایپر پارامتر نامیده می شوند پیشنهاد شده است که از جمله آنها می

۲Frailty

^۱ Herd by year

^۲Frailty

در ارزیابی های ملی و بین المللی استفاده کرد. توسعه قدرتهای محاسباتی و نرم افزار های مرتبطی چون Survival kit استفاده از این روش آماری را برای ارزیابی های ژنتیکی در اصلاح دام بسیار مفید ساخته لیکن برای استفاده کارآمدتر از این روش آماری تحقیقات بیشتری لازم است.

منابع

- [1]. Beilharz, R.G., Luxford, B.G. and Wilkinson, J.I. (1993). Quantitative genetics and evolution: Is our understanding of genetics sufficient to explain evolution? *J. Anim. Breed. Genet.* 110: 161 – 170.
- [2]. Cox, D. R. (1972). Regression Models and life tables (with discussion). *Journal of the Royal Statistical Society. Series B.* 34(2): 187-220.
- [3]. Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data.* Chapman and Hall, London, England.
- [4]. Ducrocq, V. (1987). An analysis of length of productive life in dairy cattle. Ph.D. diss., Cornell Univ., Ithaca, NY.
- [5]. Ducrocq, V. (1992). Survival analysis. Workshop on advanced biometrical methods in animal breeding. Flawil, Switzerland.
- [6]. Ducrocq, V. (1994). Statistical Analysis of Length of Productive Life for Dairy Cows of Normande Breeds. *J. Dairy Sci.* 77: 855-866.
- [7]. Ducrocq, V. (1997). Survival analysis, a statistical tool for longevity data. 48th Annual Meeting of the EAAP, Vienna, Austria.
- [8]. Ducrocq, V. and Casella, G. (1996). A Bayesian analysis of mixed survival models. *Genetics Selection Evolution.* 28: 505-529.
- [9]. Ducrocq, V. and Solkner, J. (1998). "The survival Kit-V3.0"—a package for large analysis of survival data. *Proc. 6th World Cong. Genet. Appl. Livest. Prod., Armidale, NSW, Australia, vol. 27, pp. 447–448.*
- [10]. Ducrocq, V. and Solkner, J. (1998). Implementation of a routine breeding value

توان به استفاده از الگوریتم EM، روش بیزی و تکنیک مونت کارلو اشاره کرد.

۳- نرم افزار Survival kit

سورویوال کیت در اصل یک بسته از برنامه های فرتن بوده که توسط دکروک و یوهان سولکنر [۶] با کاربری در اصلاح دام نوشته شده ولی با این حال در اپیدمیولوژی نیز کاربرد دارد. ویژگی برجسته این کیت قابلیت کاربرد آن برای مدل هایی است که دارای کواریت های وابسته به زمان و یا اثرات تصادفی می باشد. این نرم افزار متشکل از سه برنامه می باشد. برنامه اول به نام Prepare بوده و همانطور که از نام آن مشخص است داده ها را برای تحلیل آماده می سازد به عبارتی متغیرهای دارای سطوح متفاوت ثبت شده و در صورت وجود کواریت های وابسته به زمان رکوردها به صورت رکوردهایی تبدیل می شوند که فاصله زمانی از یک تغییر در هر متغیر وابسته به زمان نسبت به بعدی را پوشش می دهند. سپس داده های ثبت شده با استفاده از برنامه COX بر اساس مدل نیمه پارامتری کاکس یا برنامه WEIBULL با فرض توزیع پایه ای وایبول آنالیز می شوند. مدل ها می توانند در برگزیده متغیرهای پیوسته و یا گسسته (وابسته به زمان یا غیر وابسته به زمان) و همچنین تابع های مخاطره پایه متفاوتی باشند (لایه بندی). اثرات تصادفی می توانند توزیع نرمال، نرمال چند متغیره (با ماتریس روابط خویشاوندی) یا لگ نرمال داشته باشند. با استفاده از این بسته نرم افزاری منحنی های ماندگاری مورد انتظار نیز محاسبه می شوند. نسخه جدید این نرم افزار با نام «The Survival Kit – V5.0» کارآمدتر بوده و امکان برآورد های پارامترهایی همچون واریانس ژنتیکی اثرات تصادفی با استفاده از روش بیزی امکان پذیر است.

۴- نتیجه گیری

تحلیل بقاء یک روش آماری مناسب برای تحلیل داده های طول عمر می باشد. همچنین این روش آماری را می توان برای ارزیابی ژنتیکی طول عمر تولیدی گاوهای شیری

evaluation for longevity of dairy cows using survival analysis techniques. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, Australia. 23: 359 – 362.

- [11]. Ducrocq, V. Quass, R.L. Pollak, and E.J. Casella, G. (1988). Length of productive life in dairy cows. 2. Variance component estimation and sire evaluation. *J. Dairy Sci.*, 71: 3071-3079.
- [12]. Famula, T.R. (1981). Exponential stability model with censoring and covariates. *J. Dairy Sci.* 64:538.
- [13]. Kalbfleish, J.D. and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York, NY.
- [14]. Smith, S.P. (1983). The extension of failure time analysis to problems of animal breeding. Ph.D. Diss., Cornell Univ., Ithaca, NY.
- [15]. Smith, S.P. and Quaas, R.L. (1984). Productive life span of bull progeny groups: failure time analysis. *J. Dairy Sci.* 67:299-3007.
- [16]. Vukasinovic, N. (2001). Application of survival analysis in breeding for longevity. Proc. Int. Workshop on EU Concerted Action Genetic Improvement of Functional Traits in Cattle (GIFT); Longevity, Jouy-en-Josas, France. *Interbull Bulletin*, 21:3-10.