با فرستادن این مقاله به کارگاه ها و همایش های قطب علمی داده های ترتیبی و فضایی (W-OSDCE) تأیید می‌کنم که
(۱) محتوی و اصیل بودن این مقاله بر عهده من و دیگر نویسندگان مقاله است.
(۲) دیگر نویسندگان مقاله با فرستادن این مقاله به W-OSDCE موافق بوده‌اند.

مروری بر روش‌های برآوردیابی در رگرسیون زمان شکست شتابیده

امیری، ن. [1] * و فکور، و. [2]

[1,2] گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد
narjes.amiri@mail.um.ac.ir
fakoor@um.ac.ir

چکیده. بررسی تاثیر یک فاکتور به عنوان متغیر کمکی بر متغیر پاسخ از دیرباز مورد توجه بسیاری از محققان قرار گرفته است. حال اگر متغیر پاسخ زمان بقا باشد مدل‌های آماری خاصی برای آن وجود دارند که یک نوع پر کاربرد آن رگرسیون زمان شکست شتابیده است. در این مقاله مرور مختصری بر روش‌های مختلف برآورد پارامتر رگرسیونی در مدل زمان شکست شتابیده تحت داده‌های سانسور شده خواهیم داشت. روش‌های مورد بررسی عبارتند از روش کمترین مربعات -روش رتبه‌ها -روش درستنمایی ماکسیمم با هسته و روش درستنمایی تجربی . در بین این روش‌ها، روش درستنمایی تجربی یکی از بهترین روش‌های برآورد محسوب می‌شود که با استفاده از شبیه سازی و داده‌ی واقعی بیشتر مورد بررسی قرار می‌گیرد.

۰. پیش‌گفتار

گاهی در بررسی‌های آماری با داده‌هایی مواجه می‌شویم که مدت زمان لازم تا رخداد یک پیشامد هستند. منظور از زمان می‌تواند سال، ماه، هفته و یا سن فرد در زمان رخداد پیشامد مورد نظر باشد. منظور از پیشامد می تواند مرگ، بروز بیماری، جذب یک دانش آموخته دانشگاه به بازارکار و یا به طورکلی هر تجربه تعریف شده ای باشد که فرد (مولفه) با آن مواجه می‌شود. به این گونه داده‌های بر حسب زمان، داده‌های بقا[۱] گفته می‌شود. در واقع داده‌های بقا، زمان تا وقوع پیشامد خاصی را اندازه‌گیری می‌کنند که متغیری تصادفی با مقادیر مثبت هستند و معمولا دارای توزیع پیوسته فرض می شوند. این نوع ازداده‌ها در رشته‌های مختلفی از جمله پزشکی, کشاورزی و همه‌گیرشناسی پدیدار می‌شوند. یکی از ویژگی های مهم و بارز داده‌های بقا سانسور است. سانسور داده‌های بقا زمانی رخ می‌دهد که فرد (مولفه) پیشامد مورد نظر را تا پایان مطالعه تجربه نکرده باشد و زمان دقیق پیشامد مورد نظر مشخص نیست. در یک چنین حالتی زمان بقای فرد بیشتر از زمان مشاهده شده است.
نکته مهمی که در ارتباط با داده‌های بقا وجود دارد این است که چون این نوع داده‌ها اغلب در معرض سانسور قرار دارند نمی‌توان از مدل‌های رگرسیون خطی برای تحلیل آن‌ها استفاده کرد. زیرا برای داده‌های بقا فرضیات رگرسیون خطی برقرار نیست. بنابراین روش‌های رگرسیونی که برای مدل‌سازی داده‌های بقا به کار گرفته می‌شوند

2010 Mathematics Subject Classification. Primary 47A55; Secondary 39B52, 34K20, 39B82.
واژگان کلیدی. داده‌های سانسور شده، درستنمایی تجربی، رگرسیون زمان شکست شتابیده، معادله برآوردیابی، نمونه گیری طول-اریب .
* سخنران

[۱] Survival data

بایستی متفاوت از رگرسیون خطی کلاسیک باشند. مدل رگرسیونی زمان شکست شتابیده[۲] (AFT) یکی از پر کاربردترین مدل‌های آماری است که برای تحلیل داده‌های بقا به کار می‌رود. رگرسیون AFT نخستین بار توسط پایک (۱۹۶۶) برای تحلیل داده‌های سرطان به کار برده شد اما نامی از این مدل در مقاله اش ذکر نشد. این مدل به دو دسته پارامتری و نیم-پارامتری تقسیم می‌شود که مدل پارامتری آن نخستین بار توسط کاکس (۱۹۷۲) و مدل نیم-پارامتری آن نخستین بار توسط کلبفلش و پرنتیس (۱۹۸۰) معرفی شدند. در این مقاله مدل نیم-پارامتری AFT مورد مطالعه قرار می‌گیرد. برآورد پارامتر رگرسیونی در این مدل تا کنون با استفاده از روش‌های کمترین مربعات، رتبهها، درستنمایی ماکسیمم با هسته و درستنمایی تجربی مورد مطالعه و تحقیق قرار گرفته است. ما در این مقاله به مرور این روش‌ها می‌پردازیم و به اختصار آنها را شرح خواهیم داد.

۲. رگرسیون زمان شکست شتابیده

فرض کنید T یک متغیر تصادفی مثبت، β بردار پارامترهای رگرسیونی، Z بردار متغیر کمکی و ε خطا باشد در اینصورت مدل رگرسیون زمان شکست شتابیده به صورت زیر تعریف می شود

$$\log(T) = \beta^\tau Z + \varepsilon, \qquad (۱.۲)$$

همانطوری که ملاحظه می‌شود در مدل AFT متغیرهای کمکی به صورت خطی روی لگاریتم متغیر پاسخ تاثیر می‌گذارند. در واقع این یکی از ویژگی‌های بارز مدل AFT محسوب می‌شود. در ادامه به مرور روش‌های برآوردیابی پارامتر β تحت داده‌های سانسور شده می پردازیم.

فرض کنید T_i طول عمر یک فرد (مولفه) باشد. تحت فرآیند سانسور راست، داده‌های مشاهده شده عبارت اند از (Y_i, δ_i, Z_i) که در آن $Y_i = \min(T_i, C_i)$ و $\delta_i = I(T_i \le C_i)$ و C_i متغیر تصادفی سانسور است. فرض کنید

$$\log(T_i) = \beta^\tau Z_i + \varepsilon_i, \quad i = 1, 2, \ldots, n. \qquad (۲.۲)$$

۱.۲. **روش کمترین مربعات.**
در غیاب سانسور یک روش کلاسیک جهت برآورد پارامتر β روش کمترین مربعات خطا است که تحت مدل (۲.۲) به فرم زیر خواهد بود

$$\int \varepsilon^2 dF = \int (\log(T_i) - \beta^\tau Z_i)^2 dF, \qquad (۳.۲)$$

که در آن F تابع توزیع متغیر تصادفی ε است.

میلر (۱۹۷۶) با استفاده از مفهوم کمترین مربعات مسئله برآورد β در حضور داده‌های سانسور راست را مورد مطالعه قرار داد. او با جایگذاری برآوردگر کاپلان-مه‌یر[۳] در (۳.۲) به معادله زیر رسید

$$\int \varepsilon^2 d\hat{F}_{KM}(\varepsilon) = \sum_{uc} \nu_i (\log(T_i) - \beta^\tau Z_i)^2, \qquad (۴.۲)$$

که در آن ν_i وزنی است که برآوردگر کاپلان-مه‌یر به ε_i ها اختصاص می‌دهد و مجموع بالا روی داده‌های سانسور نشده است. معادله (۴.۲) یک معادله ناپیوسته و برآوردگر حاصل ناسازگار است. وی با اصلاح معادله (۴.۲) برآوردگری به‌دست آورد که تحت شرطی خاص سازگار و به طور مجانبی نرمال بود اما آن شرط در عمل همیشه برقرار نبود. از طرفی برآوردگری که میلر پیشنهاد داد تنها از داده‌های غیر سانسور استفاده می‌کرد و داده‌های سانسور شده را در نظر نمی‌گرفت. این باعث می‌شود برآوردگر حاصل کارا نباشد. برای غلبه بر این مشکل باکلی و جیمز [۱]

---
[۲] Accelerated failure time
[۳] Kaplan-Meier

روش جدیدی پیشنهاد دادند. آنها دریافتند در حضور داده‌های سانسور شده تنها Y_i ها قابل مشاهده هستند لذا به جای استفاده از $\log(T_i)$ در فرمول (۳.۲)، عبارت زیر را پیشنهاد داد:

$$\hat{Y}_i(\beta) = \delta_i \log(Y_i) + (1-\delta_i)\{\frac{\int_{\varepsilon_i(\beta)}^{\infty} u d\hat{F}_{KM}(u)}{1-\hat{F}_{KM}(\varepsilon_i(\beta))} + \beta^{\tau} Z_i\}, \qquad (۵.۲)$$

بنابراین برآوردگر باکلی-جیمز پارامتر β از حل معادله زیر به دست می‌آید:

$$\sum_{i=1}^{n}(Z_i - \bar{Z})(\hat{Y}_i(\beta) - \beta^{\tau} Z_i) = 0. \qquad (۶.۲)$$

مشکلی که در رابطه با معادله باکلی- جیمز وجود دارد این است که معادله ی مذکور اتابعی یکنوا و مشتق پذیر از β نیست لذا فرآیند ریشه‌یابی دشوار خواهد بود. از این رو باکلی و جیمز بدون اثبات تئوری و با شبیه سازی رفتار حدی $\hat{\beta}$ را بررسی کردند و نشان دادند دارای توزیع تقریبی نرمال است. برای حل این مشکل جین و همکاران [۳] کلاسی از معادلات برآوردیابی یکنوا معرفی کردند که معادله برآوردیابی (۶.۲) را تقریب و به شکل زیر تعریف می‌شود

$$U(\beta, b) = \sum_{i=1}^{n}(Z_i - \bar{Z})\{(\hat{Y}_i() - \bar{Y}_i()) - \beta^{\tau}(Z_i - \bar{Z})\},$$

که در آن

$$\bar{Y}(b) = n^{-1}\sum_{i=1}^{n}\hat{Y}_i(b)$$

آنها نشان دادند $\hat{\beta}$-ای که از حل معادله $U(\beta, b) = 0$ به دست می‌آید برآوردگری سازگار و به طور مجانبی نرمال است. همچنین برای محاسبه ماتریس واریانس-کواریانس این برآوردگر از روش بازنمونه‌گیری استفاده کردند. کول و همکاران (۱۹۸۱)، ریتاو و همکاران (۱۹۹۰) و لای و همکاران (۱۹۹۱) نیز از جمله کسانی بودند که از روش کمترین مربعات خطا برای برآورد پارامتر رگرسیونی نیز استفاده کردند.

۲.۲. **روش رتبه.** پرنتیس (۱۹۷۸) با استفاده از مفهوم رتبه، برآوردگر رتبه را که بر پایه‌ی آماره لگاریتم-رتبه ساخته شد پیشنهاد داد. در حالت کلی تابع لگاریتم-رتبه [۴] وزنی برای β به صورت زیر است

$$U_{\pi}(\beta) = \sum_{i=1}^{n}\delta_i \pi(\beta, \varepsilon_i(\beta))\{Z_i - \frac{\sum_{j=1}^{n} Z_j I\{\varepsilon_j(\beta) \geq \varepsilon_i(\beta)\}}{\sum_{j=1}^{n} I\{\varepsilon_j(\beta) \geq \varepsilon_i(\beta)\}}\} \qquad (۷.۲)$$

که در آن π یک تابع وزنی است که می‌تواند بر حسب داده‌ها بیان شود و $I(\cdot)$ تابع نشانگر است. در صورتی که $\pi = 1$ آنگاه معادله (۷.۲) آماره لگاریتم-رتبه است که توسط مانتل و همکاران (۱۹۷۷) معرفی شد و اگر $\pi = \sum_{j=1}^{n} I\{\varepsilon_j(\beta) \geq t\}$ آنگاه معادله حاصل , آماره گهان (گهان،(۱۹۶۵)) خواهد بود تیسیاتیس (۱۹۹۰) لای و یینگ (۱۹۹۱b) و یینگ (۱۹۹۳) نشان دادند $\hat{\beta}$ -ای که از حل $U_{\pi}(\beta) = 0$ به دست می آید دارای توزیع مجانبی نرمال است. در اینجا نیز واریانس برآوردگر حاصل فرم پیچیده ای دارد و برآورد آن مستلزم برآورد تابع خطر است. مانند روش کمترین مربعات، $U_{\phi}(\beta)$ تابعی است که پیوسته و یکنوا نیست لذا پیدا کردن $\hat{\beta}$ کاری دشوار خواهد بود. جین و همکاران [۲] کلاسی از معادلات برآوردیابی یکنوا معرفی کردند که معادله (۷.۲) را تقریب می‌زند. این معادله عبارت است از

$$\tilde{U}(\beta; \hat{\beta}) = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\psi(\hat{\beta}; \varepsilon_i(\hat{\beta}))\delta_i(Z_i - Z_j) I\{\varepsilon_i(\beta) \leq \varepsilon_j(\beta)\} \qquad (۸.۲)$$

---

[۴] Log-rank

که در آن
$\psi(b;x) = \pi(b;x)/\sum_{j=1}^{n} I\{\varepsilon_j(\beta) \geq t\}$.
آنها نشان دادند برآوردگر حاصل از حل معادله $\tilde{U}(\beta;\hat{\beta}) = 0$ برآوردگری سازگار و به طور مجانبی نرمال است. همچنین برای محاسبه ماتریس واریانس–کواریانس این برآوردگر از روش بازنمونه‌گیری نیز استفاده کردند.

۳.۲. **روش درستنمایی ماکسیمم با هسته.** زنگ و لین [۵] با به کارگیری تابع درستنمایی و تابع هسته برآوردگری برای پارامتر های رگرسیونی ارائه دادند که این برآوردگر های حاصل سازگار و به طور مجانبی دارای توزیع نرمال هستند. بر خلاف بسیاری از روش ها که واریانس پیچیده ای دارند او نشان داد که ماتریس کواریانس به طور مجانبی برابر با باند کارا efficiency می‌bound باشد و به راحتی برآورد می شود.

۴.۲. **روش درستنمایی تجربی.** روش درستنمایی تجربی یک روش ناپارامتری در استنباط آماری است که به تحلیل‌گر این اجازه را می‌دهد تا بدون دانستن توزیع داده‌ها و از طریق روش‌های درستنمایی به استنباط آماری بپردازد. درستنمایی تجربی نخستین بار توسط توماس و گرانک مه‌یر (۱۹۷۵) به منظور ساخت فواصل اطمینان بهتری برای تابع بقای جامعه معرفی شد اما آنها نامی از این روش در مقاله ذکر نکردند. بر اساس ایده‌ی توماس و گرانک مه‌یر, آون (۱۹۸۸) چارچوب کلی از درستنمایی تجربی را در استنباط ناپارامتری برای داده‌های کامل ارائه نمود. از آن زمان استفاده از درستنمایی تجربی در بسیاری از شاخه‌های آمار گسترش یافت. درستنمایی تجربی دارای ویژگی‌های آماری مطلوبی است که موجب برتری این روش نسبت به روش‌های موجود می‌شود. یکی از ویژگی‌های بارز درستنمایی تجربی ساخت فواصل اطمینان بدون نیاز به محاسبه و برآورد واریانس است. حال از آنجاییکه در تحلیل بقا با داده های ناتمام مواجه هستیم و برآورد واریانس‌های نامعلوم با وجود این داده‌ها بسیار دشوار بوده و اغلب به صورت محدود در عمل مورد استفاده قرار می‌گیرند این روش می تواند راهکار بسیار مناسبی برای حل این مشکل باشد. بنابراین با استفاده از ابزار درستنمایی تجربی در حضور داده‌های سانسور شده, ژو [۶] و ژو و لی [۷] در مدل AFT نتایجی برای پارامتر رگرسیونی β ارائه دادند.
در حضور داده‌های سانسور شده, درستنمایی تجربی به صورت زیر خواهد بود:

$$EL(\beta,F) = \prod_{i=1}^{n} p_i^{\delta_i}(1 - \sum_{\varepsilon_j(\beta)\leq\varepsilon_i(\beta)} p_j)^{1-\delta_i} \tag{۹.۲}$$

که در آن $p_i \geq 0, \sum_{i=1}^{n} p_i = 1$.

۱.۴.۲. *روش درستنمایی تجربی بر پایه‌ی معادله برآوردیابی رتبه.* ژو با به کارگیری معادله برآوردیابی بر پایه رتبه (۷.۲) محدودیتی به فرم زیر به دست آورد

$$0 = \sum_{i=1}^{n} \delta_i p_i \pi(\varepsilon_i(\beta)) \frac{Z_i - \bar{Z}}{n w_i} \tag{۱۰.۲}$$

که در آن

$$\bar{Z} = \frac{\sum_{j=1}^{n} Z_j I\{\varepsilon_j(\beta) \geq \varepsilon_i(\beta)\}}{\sum_{j=1}^{n} I\{\varepsilon_j(\beta) \geq \varepsilon_i(\beta)\}},$$

و w_i اندازه جهش برآوردگر کاپلان–مه‌یر در نقطه $\varepsilon_i(b)$ است. وی نشان داد تحت فرض $\beta = \beta_0$

$$-2\log \frac{\sup_F EL(\beta_0, F)}{EL(\hat{\beta}, \hat{F}_{KM})}$$

برای اندازه نمونه‌های بزرگ به سمت توزیع کی دو با درجه آزادی p میل می‌کند. قابل ذکر است صورت کسر تحت محدودیت (۱۰.۲) ماکسیمم می شود. ژو با استفاده از شبیه سازی و تحلیل داده‌های واقعی به بررسی صحت روش پیشنهادی خود پرداخت.

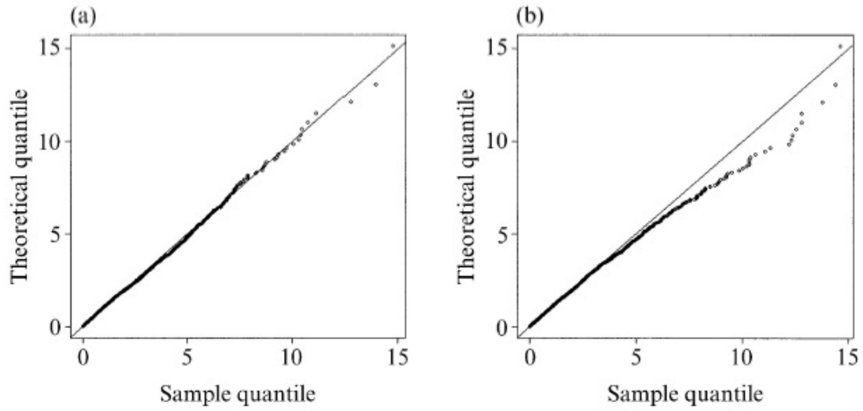
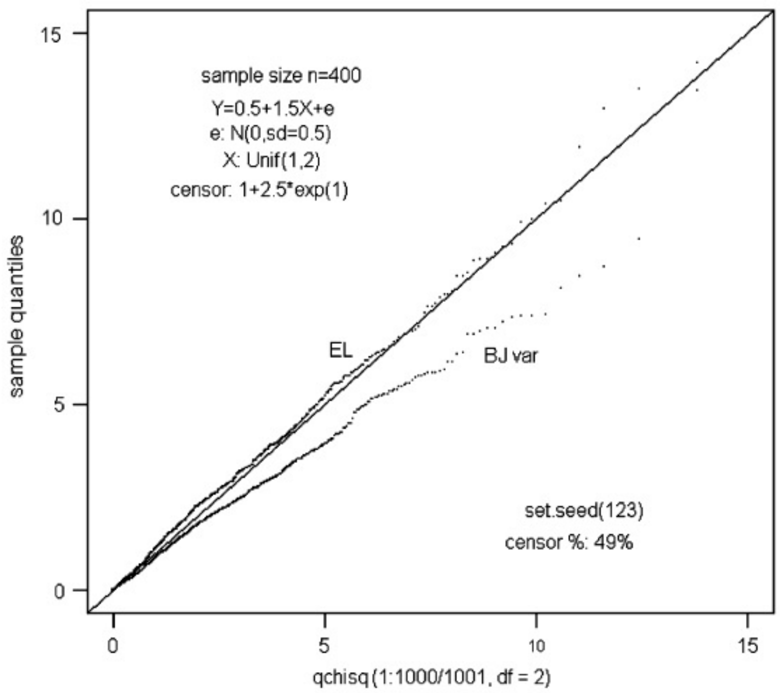


Fig. 1. Simulation study for (a) Gehan estimator, (b) log-rank estimator: Q–Q plot of $-2 \log \mathrm{ELR}$, based on 5000 simulation runs, with sample size 100.

شکل ۱: نمودار Q–Q برای نسبت درستنمایی تجربی با محدودیت بر پایه رتبه

- **شبیه سازی** وی با در نظر گرفتن مدل $\log(T) = 2x + \varepsilon$ و با رسم نمودار Q-Qمطابق شکل ۱ نشان داد توزیع حدی نسبت درستنمایی تجربی با آماره های گهان و لگاریتم–رتبه، توزیع کای–دو با یک درجه آزادی است.

- **داده‌ی واقعی** ژو همچنین با استفاده از داده‌های پیوند قلب استنفورد، مدل $\log(y) = \beta age + \varepsilon$ و آماره گهان برآورد نقطه ای و فاصله ای پارامتر β را به‌دست آورد. برآوردگر رتبه ای نقطه‌ای پارامتر β برای این داده‌ها برابر است با ۰/۰۲۵۳-. علاوه بر آن فاصله اطمینان‌های ۹۵ در صدی درستنمایی تجربی و والد به ترتیب برابرند با (۰/۰۰۳۰۱-و۰/۰۴۴۶-) (۰/۰۰۴۴-و۰/۰۴۶۲۳-). همانطوری که مشاهده می‌شود هر دو فاصله اطمینان با هم برابرند. قابل ذکر است برای ساختن فاصله اطمینان والد نیاز به برآورد واریانس $\hat{\beta}$ است که با ۱۰۰۰۰ بار باز نمونه‌گیری به‌دست آمده است. در نتیجه مطالعات شبیه سازی و داده‌ی واقعی نشان می دهند روش پیشنهادی توسط ژو روش خوبی است زیرا نیاز به برآورد واریانس ندارد و لذا محاسبات در زمان کوتاهتری انجام می‌شوند.



شکل ۲: نمودار Q–Q برای نسبت درستنمایی تجربی با استفاده از محدودیت بر پایه باکلی–جیمز

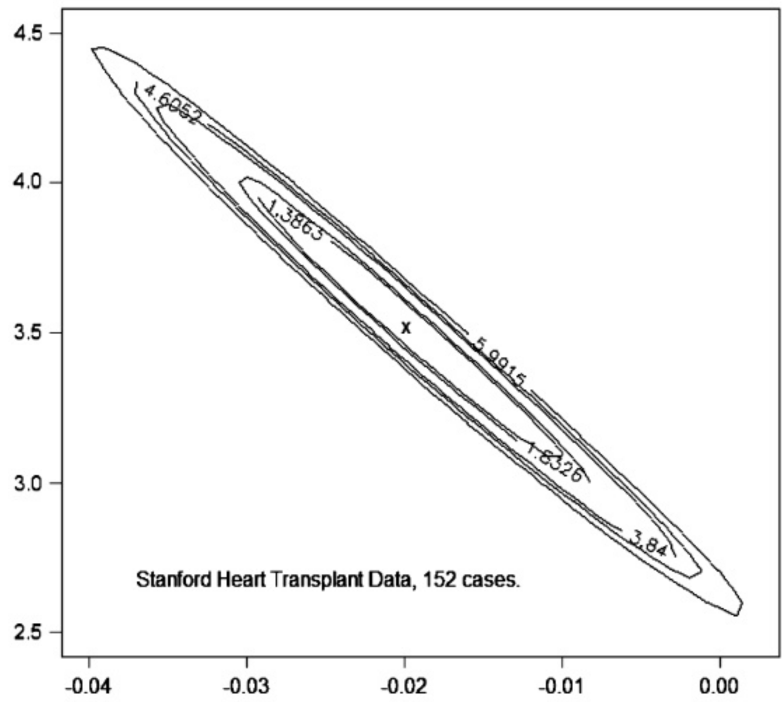
۲.۴.۲. *روش درستنمایی تجربی بر پایه‌ی معادله برآوردیابی باکلی–جیمز.* ژو و لی با استفاده از معادله برآوردیابی باکلی–جیمز (۶.۲) محدودیتی به فرم زیر ساختند که تحت آن صورت کسر درستنمایی تجربی ماکسیمم می شود

$$0 = \sum_{i=1}^{n} \delta_i p_i \varepsilon_i(b)\{Z_i + \sum_{\delta_j=0, j:\varepsilon_j \leq \varepsilon_i} \frac{Z_j \Delta \hat{F}_{KM}(\varepsilon_i)}{1-\hat{F}_K M(\varepsilon_j)}\}, \qquad (۱۱.۲)$$

آنها ثابت کردند توزیع حدی لگاریتم نسبت درستنمایی تجربی برای n های بزرگ کی دو با یک درجه آزادی است و برای ارزیابی روش پیشنهادی خود به تحلیل نتایج شبیه سازی و داده‌های پیوند قلب استنفورد پرداختند.

- **شبیه سازی** آنها در ابتدا با درنظر گرفتن مدل $y = 0.5 + 1.5x + \varepsilon$ با انجام ۱۰۰۰ بار شبیه سازی مقدار آماره‌های نسبت درستنمایی تجربی (EL) و والد (BJvar) را که با استفاده از واریانس برآوردگر باکلی–جیمز به دست می‌آید را محاسبه کردند و نمودار Q–Q را مطابق شکل ۲ ارایه دادند. شکل ۲ نشان دهنده‌ی این است که برازش توزیع کای–دو برای آماره ی نسبت درستنمایی تجربی خوب بوده است و همچنین مقدارهای آماره والد[5] محاسبه شده کمتر از مقدارهای کای–دو متناظرش است به طوریکه حتی با افزایش اندازه نمونه این اختلاف کاهش نمی‌یابد. شکل ۲ در واقع برتری روش درستنمایی تجربی را

[5]Wald statistics



شکل ۳: کانتر پلات مربوط به آماره نسبت درستنمایی تجربی

نسبت به روش نرمال مجانبی به نمایش می‌گذارد. آنها همچنین روش پیشنهادی خود ELBJ)) را با روش درستنمایی تجربی بر پایه داده‌های مصنوعی[6] (ELSD) (کین و جینگ، (۲۰۰۱) و لی و ونگ، (۲۰۰۳)) تحت مدل $y = x + \varepsilon$ مقایسه کردند . نتایج شبیه سازی نشان دهنده‌ی این واقعیت بود که برای اندازه نمونه‌های بزرگ هر دو روش خوب عمل می کنند اما برای اندازه‌های نمونه‌های کوچک بخصوص با درصد سانسور بالا ELBJ بهتر از ELSD عمل می‌کند.

- **داده‌ی واقعی** ژو و لی مدل زیر را برای تحلیل داده‌های پیوند قلب در نظر گرفتند

$$\log(T_i) = \beta_1 + \beta_2 age + \varepsilon_i,$$

تحت مدل بالا برآوردگر باکلی-جیمز $(\hat{\beta}_1, \hat{\beta}_2)$ عبارت است از (۳/۵۲۷،۰/۰۱۹-). همچنین با رسم کانتر پلات (شکل ۳) توانستند فاصله اطمینان درستنمایی تجربی برای β_1 و β_2 و رابطه بین آن دو را بررسی کنند. با مشاهده شکل ۳ به راحتی دیده می شود که فاصله اطمینان تجربی تقریبی در سطح ۹۵ درصد برای پارامتر β_1 و β_2 به ترتیب برابر است با (۲/۷۵و۴/۲۵) و (۰/۰۳۶-و۰/۰۰۲۶-). در حالی که فاصله اطمینان نرمال مجانبی برای این دو پارامتر که با استفاده از واریانس برآوردگر باکلی-جیمز به‌دست می آید به ترتیب برابر است با (۳/۴۸و۳/۵۷) و (۰/۰۲۰۹-و۰/۰۱۸۹-). آنها متوجه شدند فاصله اطمینان درستنمایی تجربی پهن تر از فاصله اطمینان نرمال مجانبی است زیرا برای محاسبه فاصله

[6]Synthetic data

اطمینان نرمال مجانبی از واریانس برآوردگر باکلی–جیمز استفاده کردند که هیچ توجیه نظری ندارد. علاوه بر آن، از آنجایی که نقاط کانتورها حالت نزولی (از بالا سمت چپ به پایین سمت راست) دارند برآوردگر $\hat{\beta}_1$ با برآوردگر $\hat{\beta}_2$ رابطه‌ی معکوس دارد.

۳. نتیجه گیری

مدل رگرسیونی زمان شکست شتابیده یکی از مدل‌های مفید و کاربردی برای تحلیل داده های بقا محسوب می شود که مورد توجه بسیاری از محققان قرار گرفته است. در این مقاله به اختصار روش‌های برآورد پارامتر رگرسیونی تحت این مدل را معرفی کردیم. خواننده با مطالعه‌ی این مقاله به این موضوع پی می‌برد در بین این روش‌ها، روش درستنمایی تجربی بهترین روش است. نتایج شبیه سازی و داده‌ی واقعی اهمیت این موضوع را می رسانند. از طرفی مسئله برآورد درستنمایی تجربی با استفاده از معادله‌ی برآوردیابی درستنمایی ماکسیمم با هسته (زنگ و لین، (۲۰۰۷)) می تواند برای پژوهش‌های آتی در نظر گرفته شود.

مراجع

1. Buckley, J., James, I. (1979). Linear regression with censored data. Biometrika, 66(3), 429-436.
2. Jin, Z., Lin, D. Y., Wei, L. J., Ying, Z. (2003). Rank-based inference for the accelerated failure time model. Biometrika, 90(2), 341-353.
3. Jin, Z., Lin, D. Y., Ying, Z. (2006). On least-squares regression with censored data. Biometrika, 93(1), 147-161.
4. Ryabenkii V.S. and Tsynkov, S. (2007), *A theoretical introduction to numerical analysis*, Chapman & Hall/CRC, Boca Raton, FL.
5. Zeng D and Lin DY (2007). Efficient estimation for the accelerated failure time model, Journal of theAmerican Statistical Association, 102, 1387–1396.
6. Zhou, M. (2005). Empirical likelihood analysis of the rank estimator for the censored accelerated failure time model. Biometrika, 92(2), 492-498.
7. Zhou, M., Li, G. (2008). Empirical likelihood analysis of the Buckley–James estimator. Journal of multivariate analysis, 99(4), 649-664.