

مقایسه مشاهدات شکست و بقا در برآورد پارامتر و پیش‌بینی

مریم آقای رشتی، دانشجوی کارشناسی ارشد گروه آمار، دانشگاه اصفهان، maryam.ay@gmail.com

چکیده: چکیده: برای پیش‌بینی طول عمر یک قطعه جدید، نیاز به انجام آزمون طول عمر n قطعه مشابه است. اما سوالی که در این‌جا مطرح می‌شود این است که در حین انجام آزمون چه چیزهایی را مشاهده کنیم؟ زمان‌های شکست یا زمان‌های بقا؟ به بیان دیگر برای تحلیل، در طول آزمون اطلاعات قطعاتی که از بین رفته‌اند مفیدتر است یا اطلاعات قطعاتی که هنوز در حال کار هستند (سانسور شده‌اند)؟ بسیاری از مهندسين به شکست اهمیت بیشتر می‌دهند و معتقدند شکست اطلاع بیشتری برای مدل فراهم می‌کند. ما نشان خواهیم داد که برای کلاس خاصی از توزیع‌ها، شکست اطلاعات بیشتری نسبت به بقا در پیش‌بینی طول عمر دارد و در مقابل برای کلاس بزرگ تری از مدل‌های طول عمر، شکست اطلاعات کمتری دارد. در این مقاله، برای مقایسه از اندازه‌های اطلاع استفاده شده است و پیش‌بینی تحت دیدگاه بیزی بررسی می‌شود.

کلمات کلیدی: آنتروپی، توزیع پیشین، توزیع پسین، توزیع پیش‌بین بیزی، اطلاع متقابل

مقدمه

توزیع‌های فراوانی، تابعی منحصر بفردی از توزیع‌ها را معرفی می‌کند که مقدار اطلاع توسط آن اندازه‌گیری می‌شود که آنتروپی شانون می‌نامند و که به صورت

$$H(Q) = - \int f(q) \log f(q) dq$$

تعریف می‌شود. فرض کنید $f(y|\theta)$ تابع چگالی توزیع طول عمر Y با پارامتر θ باشد. هدف مقایسه اطلاع فراهم شده توسط بردار داده‌های $y = (y_1, y_2, \dots, y_n)$ در مورد پارامتر θ یا تابعی از پارامترمانند $\frac{1}{\theta}$ و یا پیش‌بینی طول عمر یک قطعه جدید y_v ، تحت دو نمونه داده زیر می‌باشد.

$$D_k \equiv D(y, k) : k = n \quad (1)$$

$$D_k \equiv D(y, k) : k < n \quad (2)$$

که در آن

$$D(y, k) = (y_1, \dots, y_k, y_{k+1}, \dots, y_n)$$

k تعداد شکست‌ها و $n - k$ تعداد بقاها می‌باشد. در حالت (۱) تمام n مولفه زمان‌های شکست و در (۲) $y_1, \dots, y_k$ زمان‌های شکست و $y_{k+1}, \dots, y_n$ زمان‌های بقا می‌باشند. با فرض استقلال و هم‌توزیعی داده‌های موجود، تابع درستنمایی برای پارامتر θ به صورت

هر قطعه‌ای از یک مدل طول عمر پیروی می‌کند و هر مدل شامل پارامترهای مجهولی است، بنابراین در پیش‌بینی طول عمر یک قطعه جدید نیاز به اطلاع در مورد مدل طول عمر است، به همین ترتیب در مقایسه مشاهدات شکست و بقا در پیش‌بینی طول عمر از این حقیقت استفاده می‌کنیم که کدام یک اطلاع بیشتری در برآورد پارامترهای مدل فراهم می‌کنند. برای به دست آوردن مشاهدات مورد نیاز می‌توان وضعیتی را در نظر گرفت که n قطعه غیر قابل تعمیر، به طور تصادفی انتخاب و وارد آزمون می‌کنیم و سپس می‌توان زمان شکست n قطعه را ثبت نمود اما در شرایطی تحت محدودیت‌هایی تنها می‌توان تا زمان T که می‌تواند تصادفی یا ثابت انتخاب شود، منتظر بمانیم. در این صورت ممکن است r تا از قطعات تا T شکست خورده باشند و $n - r$ تا هنوز در حال کار باشند که با سانسور نوع اول مواجهیم و یا داده‌ها را به نحوی بخواهیم که تنها r تا شکست خورده و بقیه در حال کار باشند، که سانسور نوع دو نامیده می‌شود. هدف، مقایسه مشاهده شکست در زمان y_i در مقابل زمان بقا y_i ، برای پیش‌بینی طول عمر و برآورد پارامترهای مدل می‌باشد.

همچنین سوالات دیگری نیز مطرح می‌شود از جمله نقش توزیع پیشین در این نتایج چیست؟ آیا حجم نمونه در پاسخ گویی تاثیرگذار است؟ ما نتایج را از خانواده نمایی به خانواده بزرگی از مدل‌های طول عمر به نام تبدیلات زمان نمایی TTE^1 تعمیم می‌دهیم.

اندازه اطلاع

اندازه اطلاع به معنای میزان اطلاع فراهم شده توسط یک آزمایش است و به منظور مقایسه بعضی از جفت آزمایش‌ها استفاده می‌شود. شانون دو ایده‌ی مهم و اساسی در قضایای اطلاع در ارتباطات مهندسی بیان کرده‌است. ایده‌ی اول، این است که اطلاع یک مفهوم آماری است و دومین ایده که از ایده اول سرچشمه می‌گیرد، بر اساس

¹time-transformed exponential