پیشگفتار

به دلیل فراوانی سری‌های زمانی در آمار، اقتصاد، الکترونیک، هواشناسی، صنعت و غیره تحلیل سریهای زمانی یکی از کاربردی‌ترین و پرطرفدارترین شاخه‌های علم آمار است به طوری که هر روزه به علاقه‌مندان یادگیری روشهای تحلیل سری‌های زمانی افزوده می‌شود. از طرفی علیرغم کتب متنوعی و ارزشمندی که در این مبحث تألیف یا ترجمه شده است این علاقه‌مندان بخصوص دانشجویان در آغاز راه به سختی می‌توانند از منابع موجود بهرهٔ لازم را ببرند. این ادعایی نیست که فقط اینجانب با تجربهٔ چند سال تدریس درس سری‌های زمانی مطرح کرده باشم بلکه اغلب مدرسین محترم و بخصوص دانشجویان و پژوهشگران گرامی بر آن اذعان دارند. هدف اصلی از تألیف کتاب حاضر نیز اقدامی در راستای رفع این مشکل و ارائهٔ منبعی بوده که با حفظ نظم و دقت مناسب در ارائهٔ مطالب، بستر آموزشی مستمری برای خوانندهٔ علاقه‌مند بوجود آورد. مفاهیم اولیهٔ آمار و ریاضی مورد نیاز برای مطالعهٔ این کتاب در سطح متوسط دانشجویان کارشناسی آمار و ریاضی در نظر گرفته شده تا این کتاب برای مطالعهٔ دانشجویان سایر رشته‌ها نیز مناسب باشد. ارائهٔ برنامه‌های رایانه‌ای برای انجام کلیهٔ محاسبات این امکان را برای خوانندهٔ محترم فراهم کرده که نتایج مباحث تئوری را به سادگی برای مدل‌ها و داده‌های خود استخراج کند. با توجه به اینکه کتاب براکول و دیویس (۱۹۹۱) کتابی مرجع برای دانشجویان تحصیلات تکمیلی است و

فصل ۱

مقدمه

سریهای زمانی مشاهداتی‌اند که در طول زمان جمع آوری می‌شوند. فراوانی چنین مشاهداتی تحلیل سریهای زمانی را به یکی از کاربردی‌ترین شاخه‌های علم آمار تبدیل کرده است.

به عنوان مثال، متوسط درجه حرارت، رطوبت یا سرعت باد که بسته به نیاز به طور روزانه، هفتگی یا ماهیانه ثبت می‌شوند، قیمت سهامی خاص یا شاخصی کلی در بازار بورس، مقدار تقاضا، تولید یا فروش محصولات یک شرکت، درآمد این شرکت و مبلغی که این شرکت بابت تبلیغات محصولات خود صرف می‌کند، تعداد توریستها و درآمد حاصل از این صنعت، مشخصه‌هایی از کیفیت کالای تولیدی مثل تعداد نقصها، طول عمر، میزان چسبندگی یا غلظت و غیره که در طی فرآیند ساخت اندازه‌گیری می‌شوند، نمونه‌هایی از سریهای زمانی هستند. غالباً زمان ثبت مشاهدات گسسته است (مثالهای فوق) و در بیشتر

مواردی فاصلهٔ زمانی بین مشاهدات متوالی ثابت است. اما این زمان می‌تواند پیوسته نیز باشد. به‌عنوان مثال وضعیتی را در نظر بگیرید که اشغال (1) یا عدم اشغال (0) شبکهٔ رایانه‌ای به‌طور پیوسته ثبت می‌شود، یا دبی آب خروجی از یک منبع به‌طور پیوسته اندازه‌گیری می‌شود. در این کتاب سریهایی را مطالعه می‌کنیم که زمان ثبت گسسته با فواصل زمانی ثابت دارند.

هدف از تحلیل یک سری زمانی چیست؟ هر چند توصیف رفتار یک سری زمانی از لحاظ تغییرات موضعی و درازمدت در آن، یا مطالعهٔ وابستگی‌های موجود بین عناصر سری از بررسی‌های متداولی است که روی سریهای زمانی انجام می‌شود اما می‌توان گفت که مهمترین هدف از تحلیل یک سری زمانی پیش‌بینی مقادیر آیندهٔ آن است. در صورتی که با بیش از یک سری زمانی سر و کار داشته باشیم به این هدف مطالعهٔ روابط بین این سریها را نیز می‌توان اضافه نمود. این روابط در صورت وجود علاوه بر اینکه به ما در رسیدن به پیش‌بینی‌های دقیق‌تر کمک می‌کنند در بحث کنترل رفتار یک سری با کنترلی که بر سریهای مرتبط با آن داریم نیز کاربرد دارند. در این کتاب با روشهایی برای تحلیل سریهای زمانی یک متغیره آشنا می‌شویم. خوانندهٔ علاقمند به تحلیل سریهای زمانی چند متغیره و توابع انتقال و کنترل می‌تواند به وی[۱] (۱۹۹۰)، براکول[۲] و دیویس[۳] (۱۹۹۶) و باکس[۴] و جنکینز[۵] (۱۹۷۶) مراجعه کند.

---

۱ Wei

۲ Brockwell

۳ Davis

۴ Box

۵ Jenkins

برای بیان تغییرات منظم در سطح سری اضافه می‌شد. به چنین مدلهایی مدلهای کلاسیک می‌گوییم، به عنوان مثال از مدلهای کلاسیک،

$$X_t = \beta_0 + \epsilon_t \tag{۱.۱}$$

$$X_t = \beta_0 + \beta_1 t + \epsilon_t \tag{۲.۱}$$

$$X_t = \beta_0 + \beta_1 t + \sum_{j=1}^{6}[\beta_{j1}\cos(\frac{2\pi jt}{12}) + \beta_{j2}\sin(\frac{2\pi jt}{12})] + \epsilon_t, \quad \beta_{62} = 0 \tag{۳.۱}$$

که ϵ_t ها متغیرهای تصادفی ناهمبسته با میانگین صفر و واریانس σ^2 هستند، به ترتیب در مدل کردن سریهای (i)، (ii) و (iv) در شکل (۱.۱) می‌توان استفاده کرد. در مدل (۳.۱) با توجه به اینکه برای $j = 6$، $\sin(\frac{2\pi jt}{12}) = 0$ است می‌توان عبارت متناظر آن را حذف کرد، برای حفظ تقارن در نگارش مدل، اینکار را با اضافه کردن $\beta_{62} = 0$ انجام داده‌ایم. عدم همبستگی ϵ_t ها به هم که عدم همبستگی X_t را به همراه دارد از نقاط ضعف این مدلها است. استفاده کنندگان از این مدلها این ضعف را در پیش‌بین‌های حاصل از این مدلها جبران کرده و پیش‌بین‌هایی مثل

$$\hat{X}_n(l) = \hat{\beta}_0(n) \tag{۴.۱}$$

$$\hat{X}_n(l) = \hat{\beta}_0(n) + \hat{\beta}_1(n)(n+l) \tag{۵.۱}$$

$$\hat{X}_n(l) = \hat{\beta}_0(n) + \hat{\beta}_1(n)(n+l) + \sum_{j=1}^{6}\{\hat{\beta}_{j1}(n)\cos[\frac{2\pi j(n+l)}{12}] + \hat{\beta}_{j2}(n)\sin[\frac{2\pi j(n+l)}{12}]\}, \ \hat{\beta}_{62}(n) = 0 \tag{۶.۱}$$

را برای پیش‌بینی مقدار X_{n+l} پیشنهاد کرده‌اند که در آنها برآوردهای $\hat{\beta}(n)$ در مبدأ پیش‌بینی n حساب

متحرک توسط اسلاتسکی[۱] (۱۹۲۷) مطالعه و مورد استفاده قرار گرفته است اما بواسطهٔ کارهای جامع باکس و جنکینز در راستای تعمیم این مدلها، چگونگی محاسبهٔ پیش‌بینی‌های حاصل از آنها و مطالعهٔ رفتار این پیش‌بینی‌ها به خصوص ارائهٔ راه‌کارهایی نسبتاً کامل در استفادهٔ عملی از این مدلها، آنها را به عنوان مدلهای باکس و جنکینز می‌شناسیم. از مهمترین این تعمیم‌ها که توسط باکس و جنکینز صورت گرفته ارائهٔ مدلهای اتورگرسیو – میانگین متحرک جمع بسته برای سریهای نامانا و مدلهای اتورگرسیو – میانگین متحرک جمع بستهٔ فصلی برای سریهای نامانای فصلی است. در دستهٔ دوم از این مدلها که منشأ آن مدل معروف به مدل خطوط هوایی باکس و جنکینز است، استفاده از تفاضل‌گیری فصلی این امکان را می‌دهد که به عنوان مثال مدل خطوط هوایی با تنها دو پارامتر منجر به پیش‌بینی‌هایی مشابه (۶.۱) با سیزده پارامتر شود (فصل نهم را ببینید).

در فرآیند مدل کردن یک سری چه به کمک مدلهای کلاسیک چه مدلهای باکس و جنکینز چرخهٔ زیر بایستی تکرار شود،

الف) شناسایی مدل ب) برازش مدل ج) ارزیابی مدل برازش شده

تا نهایتاً مدلی برای تبیین رفتار سری مشاهده شده و پیش‌بینی آیندهٔ آن مناسب تشخیص داده شود.

هدف اصلی در این کتاب این است که در ابتدا خواننده به شناخت نسبتاً جامعی از مدلهای باکس و جنکینز (فصول چهارم، پنجم و نهم) برسد یعنی ساختار این مدلها، ساختار همبستگی در این مدلها، رفتار مصادیق حاصل از این مدلها، چگونگی محاسبهٔ بهترین پیش‌بینی‌ها توسط این مدلها و رفتار این

[۱] Slutsky

مواردی چون اشتباه در اندازه‌گیری یا ثبت مشاهدات، وقوع پدیده‌هایی چون خشکسالی، سیل، جنگ، صلح، انقلاب و غیره منجر به مشاهدات دورافتاده در سریهای زمانی می‌شوند. وجود چنین مشاهداتی مشکلاتی در روشهای معمول تجزیه و تحلیل سریهای زمانی بوجود می‌آورد. در فصل هشتم ضمن آشنایی با انواع دورافتاده‌ها با روشهایی برای مدل کردن و تحلیل سریهای زمانی در حضور مشاهدات دورافتاده آشنا می‌شویم. فصل نهم را به سریهای فصلی اختصاص داده و ضمن معرفی مدلهای اتورگرسیو - میانگین متحرک جمع بسته فصلی، با کاربرد این مدلها در مدل‌سازی و پیش‌بینی سریهای زمانی فصلی آشنا می‌شویم.

برای انجام محاسبات در این کتاب از نرم‌افزار S-PLUS استفاده می‌کنیم. S-PLUS یک زبان و محیط برنامه‌نویسی قوی برای تحلیل داده‌ها و رسم نمودارهای آماری است. توابع موجود در S-PLUS امکان تحلیل داده‌های آماری به روشهای معمول و پیشرفته را فراهم کرده‌اند. علاوه بر این، قابلیت برنامه‌نویسی ساده و اجرای سریع برنامه‌ها در این محیط، S-PLUS را به یکی از برنامه‌های جذاب و مورد علاقهٔ آمارشناسان تبدیل کرده است. توابع داخلی S-PLUS برای انجام بعضی از محاسبات در این کتاب کفایت می‌کنند، اما در بسیاری از موارد که اینطور نبوده، از قابلیت برنامه‌نویسی در محیط S-PLUS استفاده کرده توابعی برای انجام اینگونه محاسبات تهیه کرده‌ایم. در نوشتن این برنامه‌ها که در پیوست ۴ آمده‌اند تا حد امکان سعی شده از روابط و الگوریتم‌های نظیرشان همان‌گونه که در متن کتاب آمده استفاده شود تا ضمن اینکه هدف اصلی یعنی انجام محاسبات را برآورده می‌کنند مثال‌هایی نیز برای خوانندهٔ علاقمند به یادگیری برنامه‌نویسی در محیط S-PLUS باشند. فرمت باز این برنامه‌ها این امکان