

# فهرست مندرجات

۱۵	مفاهیم و مقدمات	۱
۱۶	مقدمه	۱-۱
۱۶	داده‌های سانسور شده	۲-۱
۱۷	روش درست‌نمایی ماکسیمم	۳-۱
۱۷	تابع درست‌نمایی	۱-۳-۱
۱۸	برآوردگر درست‌نمایی ماکسیمم	۲-۳-۱
۱۹	برآورد درست‌نمایی ماکسیمم تقریبی	۴-۱
۱۹	فاصله اطمینان	۵-۱
۲۰	سری تیلور	۶-۱
۲۱	آزمون نسبت درست‌نمایی	۷-۱

۲

۲۱ . . . . . ۱-۷-۱ آزمون نسبت درست‌نمایی برای فرضیه های ساده

۲۲ . . . . . ۲-۷-۱ آزمون نسبت درست‌نمایی برای فرضیه های مرکب

۲۲ . . . . . ۸-۱ برآورد پارامتر به روش بیز

۲۳ . . . . . ۱-۸-۱ تابع چگالی پیشین

۲۳ . . . . . ۲-۸-۱ تابع چگالی پسین

۲۴ . . . . . ۳-۸-۱ تابع زیان

۲۵ . . . . . ۴-۸-۱ تابع مخاطره

۲۵ . . . . . ۵-۸-۱ مخاطره بیز

۲۵ . . . . . ۶-۸-۱ برآوردگر بیز

۲۷ . . . . . ۷-۸-۱ فاصله اطمینان بیزی

۲۸ . . . . . ۹-۱ بسندگی

۲۸ . . . . . ۱-۹-۱ آماره های بسنده

۲۹ . . . . . ۲-۹-۱ معیار تجزیه به عوامل یا فاکتورگیری

۳۰ . . . . . ۳-۹-۱ آماره های بسنده مینیمال

۳۰ . . . . . ۴-۹-۱ طرز یافتن آماره بسنده مینیمال

۳۲ . . . . . ۱۰-۱ اطلاع فیشر

۳۴ . . . . . ۱۱-۱ روش دلتا

۳۵ . . . . . ۱۲-۱ روش تکرار

۳	
۳۶	۱۳-۱ احتمال پوشش .....
۳۶	۱۴-۱ توزیع مجانبی .....
۳۷	۱-۱۴-۱ توزیع مجانبی نرمال .....
۳۸	۱۵-۱ شبیه سازی .....
۳۹	۱۶-۱ تاریخچه و مقدمه ای بر بوت استرپ .....
۴۱	۱-۱۶-۱ چه تعداد نمونه برای بوت استرپ لازم است؟ .....
۴۱	۱-۱۶-۲ انواع بوت استرپ .....
۴۲	۱-۱۶-۳ ضعف های بوت استرپ .....
۴۳	۱-۱۶-۴ فواصل اطمینان بوت استرپ .....
۴۶	۲ معرفی انواع سانسورها و کاربردهای آن
۴۷	۱-۲ مقدمه .....
۴۷	۲-۲ مقدمه ای بر تحلیل داده های بقا .....
۴۸	۳-۲ داده های سانسور شده .....
۴۹	۴-۲ انواع سانسورها .....
۴۹	۱-۴-۲ سانسور نوع اول .....
۵۰	۲-۴-۲ سانسور نوع دوم .....
۵۱	۳-۴-۲ سانسور از راست .....

۴

۵۲ . . . . . ۴-۴-۲ سانسور از چپ

۵۲ . . . . . ۵-۴-۲ سانسور فاصله ای

۵۳ . . . . . ۶-۴-۲ سانسور تصادفی

۵۵ . . . . . ۷-۴-۲ سانسور دوگانه

۵۶ . . . . . ۸-۴-۲ سانسور هیبرید

۵۷ . . . . . ۹-۴-۲ سانسور پیشرونده

۵۸ . . . . . ۱۰-۴-۲ سانسور میانی

۵۹ . . . . . ۱۱-۴-۲ سانسور هیبرید پیشرونده نوع دو

۶۱ معرفی بعضی از الگوریتم‌های تکراری ۳

۶۲ . . . . . ۱-۳ مقدمه

۶۳ . . . . . ۲-۳ روش مونت کارلو

۶۴ . . . . . ۳-۳ زنجیر مارکوف

۶۶ . . . . . ۴-۳ زنجیر مارکوف مونت کارلو

۶۷ . . . . . ۱-۴-۳ Burn-In

۶۸ . . . . . ۵-۳ الگوریتم متروپلیس - هستینگ

۷۰ . . . . . ۶-۳ نمونه گیری گیبس

۷۲ . . . . . ۱-۶-۳ حالت دو متغیره

۷۳ . . . . . ۲-۶-۳ برآورد چگالی‌های حاشیه‌ای

۵

۴ استنباط براساس داده های سانسور شده هیبرید پیشرونده نوع دوم در توزیع نمایی ۷۵

۷۶ ۱-۴ مقدمه

۷۷ ۲-۴ توصیف مدل، علائم و برآوردگر درستنمایی ماکسیمم

۷۷ ۱-۲-۴ توصیف مدل

۷۸ ۲-۲-۴ برآوردگر درستنمایی ماکسیمم

۸۴ ۳-۴ فواصل اطمینان

۱-۳-۴ فاصله اطمینان مجانبی براساس برآوردگر درستنمایی

۸۴ ماکسیمم

۲-۳-۴ فاصله اطمینان مجانبی براساس لگاریتم برآوردگر

۸۶ درستنمایی ماکسیمم

۸۷ ۳-۳-۴ فاصله اطمینان براساس آزمون نسبت درستنمایی

۸۹ ۴-۳-۴ فواصل اطمینان بوت استرپ

۹۱ ۴-۴ تحلیل بیزی

۹۳ ۵-۴ مقایسه نتایج عددی

۹۶ ۶-۴ مثال عددی

۵ استنباط براساس داده های سانسور شده هیبرید و سانسور شده پیشرونده نوع دوم در

۱۰۸

توزیع وایبل

- ۱-۵ استنباط براساس داده های سانسور شده پیشرونده نوع دوم در توزیع وایبل ۱۰۹
- ۱-۱-۵ توصیف مدل . . . . . ۱۱۰
- ۲-۱-۵ برآورد درستنمایی ماکسیمم تقریبی . . . . . ۱۱۱
- ۲-۵ استنباط براساس داده های سانسور شده هیبرید در توزیع وایبل . . . . . ۱۱۷
- ۱-۲-۵ توصیف مدل . . . . . ۱۱۸
- ۲-۲-۵ برآوردگرهای درستنمایی ماکسیمم . . . . . ۱۱۹
- ۳-۲-۵ برآوردگرهای درستنمایی ماکسیمم تقریبی . . . . . ۱۲۱
- ۴-۲-۵ برآوردگرهای بیز و فواصل باور . . . . . ۱۲۵

۶ (\*\* استنباط براساس داده های سانسور شده هیبرید پیشرونده نوع دوم در توزیع

- وایبل ۱۳۰
- ۱-۶ مقدمه . . . . . ۱۳۱
- ۲-۶ توصیف مدل . . . . . ۱۳۲
- ۳-۶ برآوردگرهای درستنمایی ماکسیمم . . . . . ۱۳۳
- ۴-۶ برآوردگر درستنمایی ماکسیمم تقریبی . . . . . ۱۳۶
- ۵-۶ برآوردگرهای بیز و فواصل باور . . . . . ۱۴۴
- ۱-۵-۶ نمونه گیری گیس . . . . . ۱۴۵
- ۶-۶ شبیه سازی و تحلیل داده ها . . . . . ۱۴۸

۱۴۸	.....	شبه سازی ۱-۶-۶
۱۵۱	.....	تحلیل داده ها ۲-۶-۶
۱۶۲		برنامه‌های کامپیوتری ۷
۱۶۳	.....	مقدمه ۱-۷
۱۶۳	.....	برنامه مربوط به بخش مقایسه نتایج عددی فصل سوم ۲-۷
۱۶۳	.....	۱-۲-۷ محاسبه میزان آریبی و میانگین مربع خطای MLE
		۲-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر MLE ۱۶۴
		۳-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر فواصل اطمینان Boot-P ۱۶۵
		۴-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر فواصل اطمینان Boot-t ۱۶۷
		۵-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر فواصل باوربیز ۱۶۹
		۶-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر براساس توزیع مجانبی $ln\hat{\lambda}$ ۱۷۰
		۷-۲-۷ محاسبه متوسط طول های اطمینان و احتمال های پوشش
		متناظر براساس آزمون LR ۱۷۱
۱۷۴	.....	۸-۲-۷ محاسبه احتمالات کل $(\sum_{i=1}^m P(X_{i:m:n} < T))$
		۹-۲-۷ محاسبه احتمال خاتمه آزمایش در حالت اول
		$(P(\text{Case I terminated}))$ و مدت زمان مورد انتظار آزمایش ها ۱۷۴

- ۳-۷ برنامه مربوط به بخش مقایسه نتایج عددی فصل ششم . . . . . ۱۷۵
- ۱-۳-۷ محاسبه  $\hat{\alpha}$  و  $\hat{\lambda}$  از روش تکرار عددی . . . . . ۱۷۶
- ۲-۳-۷ محاسبه اریبی و میانگین مربع خطاهای  $\hat{\alpha}$  و  $\hat{\lambda}$  از روش تکرار عددی . . . . . ۱۷۷
- ۳-۳-۷ محاسبه متوسط طول های اطمینان پارامتر  $\lambda$  و احتمال پوشش متناظر آنها از روش تکرار عددی . . . . . ۱۷۸
- ۴-۳-۷ محاسبه ی متوسط طول اطمینان و احتمال پوشش  $\hat{\alpha}$  به روش تکرار عددی . . . . . ۱۸۱
- ۵-۳-۷ محاسبه  $\hat{\mu}$  و  $\hat{\sigma}$  توسط AMLE . . . . . ۱۸۳
- ۶-۳-۷ محاسبه اریبی و میانگین مربع خطاهای  $\hat{\alpha}$  و  $\hat{\lambda}$  توسط AMLE . . . . . ۱۸۶
- ۷-۳-۷ محاسبه متوسط طولهای اطمینان و احتمال پوشش  $\hat{\alpha}$  توسط AMLE . . . . . ۱۸۹
- ۸-۳-۷ محاسبه متوسط طولهای اطمینان و احتمال پوشش  $\lambda$  توسط AMLE . . . . . ۱۹۲
- ۹-۳-۷ محاسبه  $\hat{\alpha}$  و  $\hat{\lambda}$  بیزی از روش نمونه گیری گیبس . . . . . ۱۹۵
- ۱۰-۳-۷ محاسبه اریبی و میانگین مربع خطاهای  $\hat{\alpha}$  و  $\hat{\lambda}$  بیزی از روش نمونه گیری گیبس . . . . . ۱۹۸
- ۱۱-۳-۷ محاسبه متوسط طول اطمینان و احتمال های پوشش  $\hat{\alpha}$  و  $\hat{\lambda}$  بیزی از روش نمونه گیری گیبس . . . . . ۲۰۰
- ۱۲-۳-۷ رسم دو نمودار برای مشاهده عمل Burn-In . . . . . ۲۰۳
- ۴-۷ برنامه مربوط به بخش مثال عددی . . . . . ۲۰۴
- ۱-۴-۷ ورود داده های مثال عددی . . . . . ۲۰۴



- ۲-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\lambda}$  و متوسط طول اطمینان  $\hat{\lambda}$  از طریق روش تکرار عددی وقتی  $D = 10$  ..... ۲۰۴
- ۳-۴-۷ محاسبه  $\hat{\alpha}$ ، اریبی  $\hat{\alpha}$ ، میانگین مربع خطای  $\hat{\alpha}$  و متوسط طول اطمینان  $\hat{\alpha}$  از طریق روش تکرار عددی وقتی  $D = 10$  ..... ۲۰۶
- ۴-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\lambda}$  و متوسط طول اطمینان  $\hat{\lambda}$  از طریق روش تکرار عددی وقتی  $D = 7$  ..... ۲۰۸
- ۵-۴-۷ محاسبه  $\hat{\alpha}$ ، اریبی  $\hat{\alpha}$ ، میانگین مربع خطای  $\hat{\alpha}$  و متوسط طول اطمینان  $\hat{\alpha}$  از طریق روش تکرار عددی وقتی  $D = 7$  ..... ۲۱۰
- ۶-۴-۷ محاسبه  $\hat{\alpha}$ ، اریبی  $\hat{\alpha}$ ، میانگین مربع خطای  $\hat{\alpha}$  و متوسط طول اطمینان  $\hat{\alpha}$  از طریق روش AMLE وقتی  $D = 10$  ..... ۲۱۱
- ۷-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\lambda}$  و متوسط طول اطمینان  $\hat{\lambda}$  از طریق روش AMLE وقتی  $D = 10$  ..... ۲۱۳
- ۸-۴-۷ محاسبه  $\hat{\alpha}$ ، اریبی  $\hat{\alpha}$ ، میانگین مربع خطای  $\hat{\alpha}$  و متوسط طول اطمینان  $\hat{\alpha}$  از طریق روش AMLE وقتی  $D = 7$  ..... ۲۱۵
- ۹-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\lambda}$  و متوسط طول اطمینان  $\hat{\lambda}$  از طریق روش AMLE وقتی  $D = 7$  ..... ۲۱۸
- ۱۰-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\alpha}$  و  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\alpha}$  و  $\hat{\lambda}$ ، متوسط طول اطمینان  $\hat{\alpha}$  و  $\hat{\lambda}$  بیزی وقتی  $D = 10$  ..... ۲۲۲
- ۱۱-۴-۷ محاسبه  $\hat{\lambda}$ ، اریبی  $\hat{\alpha}$  و  $\hat{\lambda}$ ، میانگین مربع خطای  $\hat{\alpha}$  و  $\hat{\lambda}$ ، متوسط طول اطمینان  $\hat{\alpha}$  و  $\hat{\lambda}$  بیزی وقتی  $D = 7$  ..... ۲۲۴
- ۱۲-۴-۷ برنامه مربوط به مقایسه توزیع  $\frac{\hat{\alpha}}{\alpha}$  به ازاء  $\alpha$  و  $T$  های مختلف در MLE ..... ۲۲۷
- ۱۳-۴-۷ برنامه مربوط به مقایسه توزیع  $\frac{\hat{\alpha}}{\alpha}$  به ازاء  $\alpha$  و  $T$  های مختلف در AMLE ..... ۲۲۹

١٥

٢٣٣

٨ مراجع

## فهرست علائم

MLE	برآورد درست‌نمایی ماکسیمم
AMLE	برآورد درست‌نمایی ماکسیمم تقریبی
pdf	تابع چگالی
MCMC	زنجیر مارکف مونت کارلو
MSE	میانگین توان دوم خطا
IId	نمونه‌های مستقل و هم‌توزیع

## چکیده

کوندا و جواردر (۲۰۰۵) طرح سانسور هیبرید پیشرونده نوع دوم که در آزمون های طول عمر و قابلیت اعتماد بسیار مرسوم می باشد، معرفی و به بررسی استنباطی و اثبات برخی ویژگی ها در این سانسور پرداختند. در این رساله سانسور هیبرید پیشرونده نوع دوم هنگامی که توزیع طول عمرها نمایی می باشند، بررسی می شود. از روش برآورد درستنمایی ماکسیمم برای برآورد پارامتر مجهول استفاده کرده و سپس فواصل اطمینان مختلف را محاسبه می کنیم. با استفاده از شبیه سازی، نتایج حاصل را با یکدیگر مقایسه می نماییم. همچنین به بررسی داده های سانسور هیبرید و سانسور پیشرونده نوع دوم زمانی که توزیع طول عمرها وایبل می باشند، پرداخته و برآوردگرهای ماکسیمم درستنمایی و ماکسیمم تقریبی و بیز را به دست می آوریم. در ادامه به عنوان کار جدید، به بررسی سانسور هیبرید پیشرونده نوع دوم با توزیع طول عمرهای وایبل پرداخته و از روش های برآوردگرهای ماکسیمم درستنمایی، برآوردگرهای ماکسیمم تقریبی برای برآورد پارامترهای مجهول استفاده کرده و سپس برآورد بیز را با توزیع پیشین مناسب با استفاده از نمونه گیری گیبس محاسبه می کنیم. با استفاده از شبیه سازی، به مقایسه نتایج حاصل از برآوردهای معرفی شده می پردازیم. برای درک بهتر مطالب مثال های عددی از داده های واقعی را همراه با جداول حاصل ارائه می کنیم و در پایان برنامه های کامپیوتری نوشته شده توسط نرم افزار R در فصل هفتم قرار داده شده است.

**واژه های کلیدی:** برآورد بیز، برآورد درستنمایی ماکسیمم، برآورد درستنمایی ماکسیمم تقریبی، توزیع نمایی، توزیع وایبل، سانسور هیبرید، سانسور پیشرونده نوع دوم، سانسور هیبرید پیشرونده نوع دوم، نمونه گیری گیبس.

## پیشگفتار

بسیاری از اوقات در آزمون‌های طول عمر، آزمایش‌های کلینیکی، مطالعات تأثیر دوز (Dose) سم‌ها، تحقیقات زیست‌شناسی و دیگر زمینه‌های علم آمار، این امکان وجود دارد که زمان شکست کامل بعضی از واحدها مشاهده نشود، به این معنی که در برخی وضعیت‌ها، برای برخی از واحدها در طول آزمایش شکست اتفاق نمی‌افتد و یا از ادامه آزمایش باز نمی‌مانند و به جای دانستن زمان شکست، تمام آنچه می‌دانیم این است که این واحدها طول عمری متجاوز از مقداری مانند  $y$  دارند. این محدودیت‌های پیش آمده در نمونه‌ها را سانسور گوئیم. در مطالعات طبی و یا تحلیل‌های قابلیت اعتماد، بسیار معمول است که شکست هر واحد به بیش از یک علت مربوط باشد. به علاوه، داده‌های مشاهده شده اغلب سانسور شده‌اند. طرح سانسور هیبرید که ترکیبی از طرح‌های سانسور نوع یک و دو می‌باشد، در آزمون‌های طول عمر و یا آزمایشات قابلیت اعتماد بسیار سودمند است. اخیراً طرح سانسور پیشرونده نوع دوم برای تحلیل داده‌های با قابلیت اعتماد بسیار بالا معروف شده است. اما در این حالت ممکن است طول زمان آزمایش بسیار طولانی باشد. بنابراین طرح سانسور هیبرید پیشرونده نوع دوم معرفی می‌شود که ترکیبی از طرح‌های سانسور پیشرونده نوع دوم و سانسور هیبرید می‌باشد.

این رساله مشتمل بر هفت فصل است که ابتدای هر فصل شامل مقدمه ای است که جزئیات مطالب مندرج در آن فصل را توضیح می‌دهد، در فصل اول ابتدا مفاهیم و مقدمات مورد نیاز برای سایر فصل‌ها را بیان می‌کنیم. در فصل دوم مفاهیم و کلیاتی درباره داده‌های سانسور شده و انواع سانسورها بیان می‌کنیم. در فصل سوم توضیح مختصری در رابطه با برخی از الگوریتم‌های تکرار عددی داده می‌شود. در فصل چهارم به معرفی سانسور جدیدی به نام سانسور هیبرید پیشرونده نوع دوم پرداخته و برآوردگر درست‌نمایی ماکسیمم پارامتر مجهول و فواصل اطمینان مختلف بر اساس این سانسور را در توزیع نمایی به دست می‌آوریم و با استفاده از شبیه سازی به مقایسه نتایج عددی می‌پردازیم. در فصل پنجم به تحلیل داده‌های سانسور شده هیبرید و داده‌های سانسور شده پیشرونده نوع دوم در توزیع وایبل پرداخته و AMLE، MLE و برآوردگرهای بیزی و فواصل باور را به دست می‌آوریم. در فصل ششم به کمک فصل‌های چهارم و پنجم به بررسی داده‌های نمونه سانسور شده هیبرید پیشرونده نوع دوم در توزیع وایبل پرداخته و AMLE، MLE و برآوردگرهای بیزی و فواصل باور را با استفاده از نمونه گیری گیبس که در فصل سوم معرفی شد بررسی کرده و با استفاده از شبیه سازی نتایج حاصل را اعلام می‌نماییم. در انتها در فصل هفتم، برنامه‌های کامپیوتری مربوط به فصل‌های چهارم و ششم که توسط نرم افزار R

نوشته شده را قرار داده ایم.

بسیار کوشش شده است تا این رساله را با حداقل کاستی ارائه دهم اما نمی دانم تا چه اندازه در این کار موفق بوده ام. قضاوت را به خوانندگان واگذار می کنم.

در این رساله مطالبی که توسط اینجانب اثبات شده است با علامت (\*) و مطالبی که تماماً ایده جدید می باشد با (\*\*\*) مشخص شده اند.

در پایان لازم می دانم تشکر و قدردانی خالصانه خود را نسبت به استاد عزیز و ارجمندم، سرکار خانم دکتر حبیبی راد که زحمت راهنمایی این رساله را بر عهده داشتند و همچنین جناب آقای دکتر اراقامی که سمت مشاوره این رساله را عهده دار بودند، اعلام نمایم. از داوران محترم، جناب آقای دکتر دوست پرست و سرکار خانم دکتر برات پور که حضور ایشان باعث دلگرمی اینجانب بوده تشکر می نمایم. از سرکار خانم دکتر یوسف زاده که با صبر در امر برنامه نویسی با نرم افزار R اینجانب را یاری و راهنمایی نمودند، کمال تشکر و قدردانی را دارم. از همکاریهای صمیمانه مسئولین کتابخانه که در تهیه مراجع و کتب مورد نیاز اینجانب را یاری رساندند کمال تشکر و احترام را دارم. و نیز زحمات مسئولین آموزش از چشم اینجانب پوشیده نخواهد ماند. در نهایت، از کمک ها، صبر و شکیبایی بی دریغ مادر نازنینم و پدر عزیزم که نهایت عشق و ایثار را به من ارزانی کردند و مرا از تکیه کردن بی نیاز گردانیدند و در تمام مراحل زندگی مرا یاری رساندند، سپاسگزارم.

الهام بیات مختاری

شهریور ۱۳۸۸

## فصل ۱

# مفاهيم و مقدمات

## ۱-۱ مقدمه

در این فصل به تعریف مفاهیم و اصطلاحات مورد نیاز در فصل‌های بعدی می‌پردازیم. مفاهیمی نظیر داده‌های سانسور شده، روش درستی‌نمایی ماکسیمم، فاصله اطمینان، آمار بیز و مفاهیمی دیگر که در فصول بعدی مورد استفاده قرار خواهند گرفت.

## ۱-۲ داده‌های سانسور شده

یکی از مشکلات در مطالعه داده‌های زمان تا پیشامد و در واقع آنچه تحلیل بقاء را از سایر مباحث آماری جدا می‌کند عمل سانسور است. در تحلیل داده‌های بقاء مانند آزمون‌های طول عمر، آزمایشهای کلینیکی، مطالعات تأثیر دوز سم‌ها و دیگر زمینه‌های علم آمار، این امکان وجود دارد که زمان شکست کامل برخی از واحدها مشاهده نشود. به این معنی که در برخی وضعیت‌ها، برای برخی از واحدها در طول آزمایش شکست اتفاق نمی‌افتد و یا از ادامه آزمایش باز نمی‌مانند و به جای دانستن زمان شکست  $t$ ، تمام آنچه می‌دانیم این است که این واحدها طول عمری متجاوز از مقداری مانند  $y$  دارند، در آمار سانسور زمانی رخ می‌دهد که تنها قسمتی از مقدار یک مشاهده معلوم باشد یا هنگامی که یک مقدار خارج از حیطه اندازه‌گیری روی دهد.

این محدودیت، ممکن است از روی اجبار و یا به صورت اختیاری توسط آمارگرا اعمال شود. بعضی از این محدودیت‌ها عبارتند از: فرصت کم برای اعلام نتایج، طولانی شدن مدت آزمایش، عدم دسترسی به همه واحدها و یا مایوس شدن از نتیجه دادن همه واحدها. این محدودیت‌های پیش آمده در نمونه‌ها را سانسور<sup>۱</sup> می‌نامند. سانسورها معمولاً به صورت‌های مختلفی اعمال می‌شوند. بعضی از انواع سانسورها عبارتند از: سانسور نوع اول<sup>۲</sup>، سانسور نوع دوم<sup>۳</sup>، سانسور تصادفی<sup>۴</sup>، سانسور دوگانه<sup>۵</sup>، سانسور هیبرید<sup>۶</sup> و دیگر سانسورها که در فصل بعدی به طور کامل معرفی خواهند شد. برای جزئیات بیشتر درباره

Censoring<sup>۱</sup>Type-I censoring<sup>۲</sup>Type-II censoring<sup>۳</sup>Random censoring<sup>۴</sup>Double censoring<sup>۵</sup>Hybrid censoring<sup>۶</sup>



انواع سانسورها و کاربرد آنها می توان به کتاب های آرنولد و همکاران <sup>۷</sup> (۱۹۹۲)، لاولس <sup>۸</sup> (۱۹۸۲) و نلسن <sup>۹</sup> (۲۰۰۴) مراجعه کرد.

یکی از اولین تلاشها در جهت تحلیل مسائل آماری که شامل داده های سانسور شده بود توسط دانیل برنولی <sup>۱۰</sup> (۱۷۶۶) انجام گرفت که در تحلیل شیوع بیماری آبله و اطلاعات مربوط به مرگ و میر در اثر این بیماری بود.

### ۳-۱ روش درستنمایی ماکسیمم

روش درستنمایی ماکسیمم <sup>۱۱</sup> (ML) یکی از قدیمی ترین و پر اهمیت ترین روش ها در نظریه برآوردهاست. برآورد ML که متداول ترین روش در بین استفاده کنندگان کاربردی از علم آمار است، اولین بار توسط گوس <sup>۱۲</sup> (۱۸۲۱) به کار گرفته شده است و پس از آن به صورت گسترده ای توسط فیشر <sup>۱۳</sup> (۱۹۲۵) در انتقاد از روش "برآورد گشتاوری" مورد استفاده قرار گرفت. روش  $MLE$  عبارت است از دستورالعملی برای به دست آوردن برآوردگری به نام "برآوردگر درستنمایی ماکسیمم" و مبتنی بر تابع آماری به نام تابع درستنمایی است.

#### ۱-۳-۱ تابع درستنمایی

فرض کنید  $\underline{X} = (X_1, \dots, X_n)$  بردار  $n$  متغیر تصادفی با تابع چگالی احتمال توأم  $f_{\theta}(\underline{x})$ ،  $\theta \in \Theta \subseteq R^k$ ، باشد. آنگاه برای هر مقدار داده شده  $\underline{X} = \underline{x}$ ، تابع درستنمایی  $\underline{X}$  را تابع چگالی احتمال توأم  $\underline{x}$ ، یعنی  $f_{\theta}(\underline{x})$ ، تعریف می کنیم که بصورت تابعی از  $\theta$  در نظر گرفته می شود و آن را با نماد  $L(\theta)$  نمایش می

---

Arnold et al<sup>۷</sup>

Lawless<sup>۸</sup>

Nelson<sup>۹</sup>

Daniel Bernoulli<sup>۱۰</sup>

Maximum likelihood<sup>۱۱</sup>

Guse<sup>۱۲</sup>

Fisher<sup>۱۳</sup>

دهیم:

$$\begin{aligned} L(\theta) &= f_{\theta}(x_1, \dots, x_n), \\ &= L(\theta; x_1, \dots, x_n), \\ &= L(\cdot; x_1, \dots, x_n). \end{aligned}$$

لازم به ذکر است که

الف) تابع درست‌نمایی  $L(\theta)$  لزوماً نسبت به  $\theta$  مشتق پذیر نیست.

ب) تابع درست‌نمایی  $L(\theta)$  لزوماً یک تابع چگالی احتمال نیست.

ج) اگر  $X_1, \dots, X_n$  یک نمونه تصادفی  $n$  تایی از توزیعی از خانواده چگالیهای  $\{f_{\theta}(x) : \theta \in \Theta\}$  باشد، آنگاه:

$$L(\theta) = \prod_{i=1}^n f_{\theta}(x_i)$$

### ۱-۳-۲ برآوردگر درست‌نمایی ماکسیمم

بطور کلی، مقداری از  $\theta$  را که تابع درست‌نمایی را ماکسیمم می‌کند برآوردگر درست‌نمایی ماکسیمم آن تابع گویند.

به عبارت دیگر باید داشته باشیم

$$L(\hat{\theta}, x_1, \dots, x_n) = \max_{\theta \in \Theta} L(\theta; x_1, \dots, x_n)$$

بنابراین اگر  $\delta(\underline{X})$  برآوردگری برای  $\theta$  باشد، بطوریکه:

$$i) P_{\theta}(\delta(\underline{X}) \in \Theta) = 1$$

$$ii) L(\delta(\underline{x})) \geq L(\theta); \quad \forall \theta \in \Theta$$

آنگاه  $\delta(X)$  یک برآوردگر درست‌نمای ماکسیمم  $\theta$  تعریف می‌شود، که معمولاً آن را با  $\hat{\theta}$  نشان می‌دهند. در اغلب مواقع به خصوص برای مشتق‌گیری آسانتر آن است که بجای  $L(\theta)$  از  $l(\theta) = \ln L(\theta)$  استفاده کنیم.

## ۴-۱ برآورد درست‌نمایی ماکسیمم تقریبی

برآورد درست‌نمایی ماکسیمم پارامترها (مکانی، مقیاسی) که بر اساس نمونه‌های کامل و یا سانسور شده می‌باشند، همواره مورد بحث و بررسی آماردانان واقع شده است. در برخی از توزیع‌ها روش درست‌نمایی ماکسیمم، برآوردگرهای واضح و روشنی را ایجاد نمی‌کند. از این رو، مطلوب است (در صورت امکان) تقریب‌هایی را برای این روش فراهم آوریم به طوری که برآوردگرهایی را برای پارامترهای مکانی و مقیاسی که به صورت توابع واضحی از آماره‌های مرتب هستند، پدید آورد که دارای خواص بهینه باشند. به طور کلی، برآورد درست‌نمایی ماکسیمم تقریبی<sup>۱۴</sup> زمانی استفاده می‌شود که نمی‌توان  $\theta$  را مستقیماً و بطور دقیق از روش درست‌نمایی ماکسیمم به دست آورد.

این روش توسط بالاکریشنان<sup>۱۵</sup> (۱۹۹۰b, c, d, ۱۹۸۹a, b)، بالاکریشنان و وارادان<sup>۱۶</sup> (۱۹۹۰)، بالاکریشنان و وانگ<sup>۱۷</sup> (۱۹۹۰a, b) و چان<sup>۱۸</sup> (۱۹۸۹) معرفی و گسترش یافت. لازم به ذکر است که این روش اولین بار توسط بالاکریشنان در زمینه سانسورها مورد استفاده قرار گرفته است. علت نامگذاری آن به عنوان برآورد درست‌نمایی ماکسیمم تقریبی استفاده از تقریب سری تیلور می‌باشد، که نحوه استفاده از این روش به طور کامل در فصل‌های بعدی بیان شده است.

## ۵-۱ فاصله اطمینان

اگر  $X_1, \dots, X_n$  یک نمونه تصادفی از چگالی  $f(\cdot; \theta)$  باشد و  $T_1$  و  $T_2$  دو آماره باشند بطوریکه  $T_1 \leq T_2$  و  $P_\theta[T_1 < \theta < T_2] = \gamma$  که به  $\theta$  بستگی نداشته و برابر با  $1 - \alpha$  باشد، در این صورت بازه تصادفی  $[T_1, T_2]$  را یک فاصله اطمینان<sup>۱۹</sup>  $100\gamma$  درصدی برای  $\theta$  و  $\gamma$  را ضریب اطمینان می‌نامیم که بصورت دلخواه توسط پژوهشگر انتخاب می‌شود. اغلب مقادیر ۵% و ۱% که متناظر با سطوح اطمینان ۹۵% و

<sup>۱۴</sup> Approximate Maximum Likelihood Estimation

<sup>۱۵</sup> Balakrishnan

<sup>۱۶</sup> Varadan

<sup>۱۷</sup> Wong

<sup>۱۸</sup> Chan

<sup>۱۹</sup> Confidence Interval

۹۹٪ می باشد برای  $\alpha$  در نظر گرفته می شود.

لازم به ذکر است که تعداد کمی فاصله اطمینان دقیق شناخته شده است. بنابراین، در اکثر مواقع، نیازمند جستجوی راهی برای دستیابی به فواصل اطمینان تقریبی با دقتی در سطح قابل قبول می باشیم. یک راه برای رسیدن به فواصل اطمینان تقریبی، فاصله اطمینان مجانبی<sup>۲۰</sup> است که بصورت زیر تعریف می شود. تعریف ۱-۱.۵ فاصله اطمینان مجانبی: بازه  $[T_1, T_2]$  را فاصله اطمینان مجانبی گویند هرگاه گزاره

$$P(T_1 < \theta < T_2) = \gamma$$

به صورت حدی وقتی  $n$  به سمت بینهایت میل کند، درست باشد.

## ۱-۶ سری تیلور

توابعی نظیر توابع لگاریتمی، نمایی و مثلثاتی وجود دارند که به آسانی قابل محاسبه نیستند. لیکن، چند جمله‌ایها را می توان به جای تابع اصلی، وقتی تفاضل تابع و تقریب چند جمله‌ای به قدر کافی کوچک است، در محاسبات به کاربرد.

برای تقریب یک تابع بوسیله چند جمله‌ایها روش‌های متعددی وجود دارند. یکی از این روش‌ها، روشی ناشی از فرمول تیلور است، که اولین بار توسط یک ریاضیدان انگلیسی به نام بروک تیلور<sup>۲۱</sup> معرفی و به افتخار این ریاضیدان به سری تیلور نام گرفت. در ریاضی، سری تیلور نماینده‌ای از یک تابع است که بصورت مجموع نامتناهی از جملات که توسط مشتقات آن در یک نقطه محاسبه می شود، می باشد. اگر سری تیلور حول نقطه صفر مرکزی شود، به آن سری مک لورن گویند که به افتخار ریاضیدان اسکاتلندی بنام کالین مک لورن<sup>۲۲</sup> نامگذاری شده است.

فرض کنید  $f$  تابع حقیقی یا مختلط مشتق‌پذیر از تمام مراتب بر بازه  $I$  شامل نقطه  $a$  باشد که  $a$  میتواند حقیقی یا مختلط باشد، در این صورت به ازای هر  $x$  در  $I$  سری توانی

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n = f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f''(a)}{2!} (x-a)^2 + \dots$$

<sup>۲۰</sup>Asymptotic Confidence Interval

<sup>۲۱</sup>Brook Taylor

<sup>۲۲</sup>Colin Maclaurin