دانشگاه فردوسی مشهد

دانشکده علوم ریاضی

پایان‌نامه برای دریافت درجه کارشناسی ارشد در رشته آمار ریاضی

برآورد تابع چگالی در مدل دو نمونه‌ای نیمه‌پارامتری

استاد راهنما

دکتر وحید فکور

استاد مشاور

دکتر مهدی عمادی

نگارنده

مریم حسام عارفی

شهریورماه ۱۳۹۱

پیش‌گفتار

با توجه به اهمیت مدل‌های نیمه‌پارامتری در علوم مختلف از جمله آمار زیستی، آمار اقتصادی، پزشکی و اقتصاد بر آن شدیم که در این مجموعه به بررسی برآورد تابع چگالی در مدل دونمونه‌ای نیمه‌پارامتری بپردازیم. این پایان‌نامه مشتمل بر چهار فصل است که خلاصه مطالب هر فصل به شرح زیر می‌باشد:

• در فصل اول پایان‌نامه، تعاریف و مفاهیم مورد نیاز در فصل‌های آتی، ارائه خواهد شد. لذا نخست لازم می‌دانیم با تعاریف و ویژگی‌های اولیه‌ی تابع هسته، تابع توزیع تجربی و برخی مفاهیم دیگر که در فصل‌های آینده مورد استفاده قرار می‌گیرند، آشنا شویم. به این منظور در نخستین بخش شرح مختصری از خانواده گروه‌های مکانی، مقیاسی، مکانی-مقیاسی و خانواده نمایی ارائه می‌شود. نابرابری‌ها و قضیه‌های همگرایی که در فصل‌های آتی در اثبات قضایا به طور گسترده‌ای کاربرد خواهند داشت، در بخش ۱.۴ معرفی می‌شوند. در بخش ۱.۶ تعریف تابع هسته و انواع آن را می‌آوریم. در ادامه مروری مختصر بر مفاهیم پایه‌ای نرمال مجانبی بودن و میانگین مربعات خطا خواهیم داشت. سپس در بخش‌های ۱.۸ و ۱.۹ مروری بر مفاهیم کارایی و سازگاری داریم. در نهایت فصل اول را با تشریح روش درستنمایی ماکسیمم و قضیه‌های مربوط به آن، به پایان می‌بریم. ضمناً مطالب این فصل از منابع مختلفی گردآوری شده است که از آن جمله می‌توان به گات (۲۰۰۴)، روهاتگی (۲۰۰۱)، بیکل و داکسوم (۱۹۷۷)، برگر (۲۰۰۲) و دبارا (۱۹۸۱) اشاره کرد.

• فصل دوم این پایان‌نامه به بررسی برآورد پارامترها در مدل دو نمونه‌ای نیمه‌پارامتری با استفاده از روش برآورد درستنمایی ماکسیمم نیمه‌پارامتری ($MSLE$[۱]) اختصاص دارد. از این رو در نخستین بخش مقدمه و تاریخچه‌ای مختصر از مدل‌های نیمه‌پارامتری و کاربرد این مدل‌ها ارائه می‌شود. سپس در بخش دوم این فصل به برآورد پارامترها از روش درستنمایی ماکسیمم خواهیم پرداخت، در ادامه به بررسی کارایی نسبی برآوردگرهای بدست آمده از روش درستنمایی ماکسیمم نیمه پارامتری نسبت به برآوردگرهای درستنمایی ماکسیمم معمولی خواهیم پرداخت و در نهایت این فصل را با دو مثال کاربردی که به بررسی کارایی برآوردگرهای ذکرشده می‌پردازد، به پایان خواهیم برد. قابل ذکر است که محاسبات این مثال‌ها با استفاده از نرم افزار R و به روش شبیه‌سازی

[۱] Maximum Semi-Parametric Likelihood Estimator

مونت کارلو انجام شده است که نمونه‌ای از این برنامه را در پیوست آورده‌ایم. مطالب این فصل عمدتاً تعمیم مقاله ژنگ (۲۰۰۲) است.

• اما در فصل سوم، به این موضوعات پرداخته‌ایم: در ابتدا در مقدمه به معرفی مدل دو نمونه‌ای نیمه‌پارامتری پرداخته‌ایم. در بخش دوم، به موضوع برآورد مینیمم فاصله هلینجر، در مدل‌های دو نمونه‌ای نیمه‌پارامتری خواهیم پرداخت. در بخش سوم، به تفکیک در مورد خواص مجانبی برآوردگرهای ذکر شده که شامل سازگاری و نرمال مجانبی بودن است، بحث خواهد شد، همچنین در دو بخش بعدی فصل سوم بترتیب لم‌های مورد نیاز برای اثبات قضایا و اثبات قضایای مهم آورده شده است. مطالب این فصل تعمیم مقاله وو و همکاران (۲۰۱۰) می‌باشد.

• آخرین فصل این مجموعه به شبیه‌سازی با استفاده از روش مونت کارلو اختصاص دارد. لذا لازم دانستیم، ابتدا این روش را به طور مختصر توضیح دهیم. در این فصل به طور عددی نشان داده می‌شود برآوردگرهایی که از روش مینیمم سازی فاصله هلینجر (MHD)[۲] و روش درستنمایی ماکسیمم بدست می‌آیند، دارای چه خواصی هستند و این برآوردگرها را مورد مقایسه قرار خواهیم داد. مطالب این فصل از مقاله وو و همکاران (۲۰۱۰) برگرفته گرفته شده است. شبیه سازی‌ها را با استفاده از نرم افزار R انجام داده‌ایم و برنامه‌های شبیه‌سازی در پیوست آورده شده است.

لازم به ذکر است که در طول این پایان‌نامه همواره سعی بر آن بوده است که تمامی قضایا و نتایج مورد نیاز در همین مجموعه آورده شود تا خواننده بدون نیاز به مراجع، برهان قضایای اصلی را دنبال کند. همچنین برای جلوگیری از حجیم شدن این گردایه، از آوردن برهان این نتایج اجتناب کرده و تنها به ذکر نام مرجع بسنده کرده‌ایم.

علیرغم اینکه نمادهای مربوط به هر فصل، در همان قسمت معرفی می‌شوند، در ادامه فهرستی از نمادهای استفاده شده در کل پایان‌نامه ارائه می‌شود، تا سبب سهولت در امر آشنایی با نمادها شود.

به منظور حفظ یکپارچگی و انسجام در نوشتار پایان‌نامه و نیز جلوگیری از ترکیب کلمات انگلیسی با فارسی، واژه‌نامه‌ای را در انتها گردآوری نمودیم تا خوانندگان بتوانند به‌صورت یک جا و آنی، برگردان بیشتر واژه‌های مطرح شده در پایان‌نامه را مشاهده نمایند. در برگردان علائم، اسامی و ترجمه کلمات کلیدی از کتاب «واژه‌ها و اصطلاحات آماری» چاپ مرکز آمار ایران و سایت پژوهشکده آمار ایران استفاده گردیده است.

قطعاً این مجموعه بدون نقص نیست، لذا از خوانندگان عزیز تقاضا دارم با دقت موشکافانه خود عیب‌ها و کاستی‌های این پایان‌نامه را از طریق پست الکترونیکی[۳] به اطلاع اینجانب برسانند.

[۲] Minimum Hellinger Distance

[۳] Maryamhesam28@yahoo.com

نمادها

σ-میدان	Ω
گردایه‌ی زیرمجموعه‌های Ω	$\mathcal{A}$
ثابت عمومی	C
مستقل و هم‌توزیع	$\overset{\text{iid}}{\sim}$
مینیمم فاصله هلینجر	MHD
میانگین مربعات خطا	MSE
امید ریاضی	E_θ
توزیع نمایی با پارامتر λ	$\mathrm{E}(\lambda)$
تابع توزیع متغیر مورد بررسی	F
همگرایی در توزیع	$\xrightarrow{\mathrm{D}}$
تابع چگالی متغیر مورد بررسی	f
همگرایی در احتمال	$\xrightarrow{\mathrm{P}}$
توزیع نرمال با میانگین μ و واریانس $\sigma^۲$	$\mathrm{N}(\mu, \sigma^۲)$
نتیجه می‌دهد	$\Rightarrow$
همگرایی تقریباً مسلم	$\xrightarrow{\text{a.s.}}$
واریانس متغیر مورد بررسی	Var
بزرگترین عدد صحیح کوچکتر یا مساوی n	$[\mathrm{n}]$
مشتق مرتبه kام	$\mathrm{g}^{(\mathrm{k})}$
مجموعه‌ی اعداد حقیقی	$\mathbb{R}$
عضو، متعلق به	$\in$
وقتی که	as
برای هر	$\forall$
مشتق تابع چگالی نسبت به θ	$\frac{\partial}{\partial\theta}\mathrm{f}_\theta(\mathrm{x})$
همگراست به	$\rightarrow$
تابع نشانگر مجموعه A	$\mathrm{I}_\mathrm{A}(.)$
مجموعه نقاط درونی Θ	$\mathrm{int}(\Theta)$

فهرست مطالب

لیست تصاویر

لیست جداول

فصل ۱

تعاریف و مفاهیم مورد نیاز

۱.۱ مقدمه

در این فصل مفاهیمی که در فصل‌های بعد مورد استفاده قرار می‌گیرند را معرفی می‌کنیم. به این منظور در دو زیربخش ابتدایی شرح مختصری از خانواده گروه‌های مکانی، مقیاسی، مکانی-مقیاسی و خانواده نمایی ارائه می‌شود. نامساوی‌ها و قضیه‌های همگرایی که در فصل‌های بعد به طور گسترده در اثبات قضایا کاربرد دارند در زیربخش ۴.۱ معرفی می‌شوند. در زیربخش‌های ۵.۱ و ۶.۱ به‌ترتیب تابع توزیع تجربی و تابع هسته و انواع آن مورد بررسی قرار می‌گیرند و در ادامه نامساوی کرامر رائو را بیان خواهیم کرد. زیربخش‌های ۸.۱ و ۹.۱ شامل مروری بر تعاریف اولیه دو ویژگی پرکاربرد، کارایی و سازگاری می‌باشند. در ادامه مطلب، مروری مختصر بر مفاهیمی همچون نرمال مجانبی بودن، میانگین مربعات خطا و کامل بودن خواهیم داشت. در نهایت این فصل را با تشریح روش درستنمایی ماکسیمم به پایان می‌بریم.
در این پایان‌نامه با دو گروه از توزیع‌ها سر و کار داریم که یکی خانواده گروهی و دیگری خانواده نمایی از توزیع‌ها می‌باشد.

۲.۱ خانواده گروهی از توزیع‌ها

یک خانواده گروهی از توزیع‌ها به وسیله انجام یک خانواده از تبدیلات روی یک متغیر تصادفی با یک توزیع خاص حاصل می‌شود. در زیر سه خانواده‌ی مهم از خانواده‌ی گروهی مهم را معرفی می‌کنیم.

خانواده‌ی مکانی

فرض کنید U یک متغیر تصادفی با یک تابع چگالی $f(u)$ باشد. اگر μ مقداری ثابت باشد آنگاه توزیع $X = U + \mu$ عبارت است از $f_X(x) = f(x-\mu)$
توزیع حاصل برای توزیع احتمال ثابت f و $-\infty < \mu < \infty$ را یک خانواده‌ی مکانی از توزیع‌ها گویند.

مثال ۱.۱. توزیع $N(\mu, ۱)$ از توزیع $N(۰, ۱)$ با جمع با مقدار ثابت μ حاصل می‌شود، زیرا به سادگی داریم

$$Z \sim N(۰, ۱) \Rightarrow X = Z + \mu \sim N(\mu, ۱)$$

$$f_X(x) = \frac{۱}{\sqrt{۲\pi}} \exp[-\frac{۱}{۲}(x-\mu)^۲] = f(x-\mu)$$

بنابراین خانواده $N(\mu, ۱)$ با $-\infty < \mu < \infty$ یک خانواده مکانی از توزیع‌ها است. □

خانواده‌ی مقیاسی

فرض کنید U یک متغیر تصادفی با تابع چگالی $f(u)$ باشد. اگر σ مقداری ثابت باشد، آنگاه توزیع $X = \sigma U$ عبارت است از $f_X(x) = \frac{1}{\sigma} f(\frac{x}{\sigma})$، توزیع حاصل برای $\sigma > 0$ را یک خانواده‌ی مقیاسی از توزیع‌ها گویند.

مثال ۲.۱. توزیع $E(\theta)$ از توزیع ثابت $E(1)$ با ضرب در مقدار ثابت θ حاصل می‌گردد.

$$U \sim E(1) \Rightarrow X = \theta U \sim E(\theta), \quad f_X(x) = \frac{1}{\theta}\exp(-\frac{x}{\theta}) = \frac{1}{\theta} f(\frac{x}{\theta})$$

بنابراین خانواده $E(\theta)$ یک خانواده مقیاسی از توزیع‌هاست. □

خانواده‌ی مکانی-مقیاسی

فرض کنید U یک متغیر تصادفی با تابع چگالی $f(u)$ باشد. اگر σ و μ مقادیر ثابتی باشند، آنگاه توزیع $X = \mu + \sigma U$ عبارت است از: $f_X(x) = \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$.
توزیع حاصل برای $-\infty < \mu < \infty$ و $\sigma > 0$ را یک خانواده‌ی مکانی-مقیاسی از توزیع‌ها گویند.

مثال ۳.۱. توزیع $N(\mu, \sigma^2)$ از توزیع $N(0, 1)$ به صورت زیر حاصل می‌گردد.

$$U \sim N(0, 1) \Rightarrow X = \mu + \sigma U \sim N(\mu, \sigma^2),$$

پس

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\} = \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$$

بنابراین $N(\mu, \sigma^2)$ به خانواده‌ی مکانی-مقیاسی از توزیع‌ها تعلق دارد. □

۳.۱ خانواده نمایی از توزیع‌ها

تعریف ۱.۱. (خانواده نمایی) یک خانواده $\{f_\theta : \theta \in \Theta\}$ از توزیع‌ها را یک خانواده‌ی نمایی k پارامتری گویند، اگر توزیع f_θ برای هر $\theta \in \Theta$ و هر $x \in S_X$ (تکیه‌گاه X) دارای تابع چگالی بفرم زیر نسبت به اندازه‌ی

لبگ (μ) باشد

$$f_\theta(x) = \Big\{\exp[\sum_{i=1}^{k} c_i(\theta)T_i(x) - B(\theta)]\Big\}h(x)$$

که در آن

۱) S_X بستگی به مقادیر θ نداشته باشد.

۲) $c_i(\theta)$، توابعی غیر صفر و پیوسته از θ باشند.

۳) در حالت پیوسته $T_i(x)$، توابعی پیوسته روی S_X و در حالت گسسته $T_i(x), i = 1, \ldots, k$، توابعی غیر صفر روی S_X باشند.

۴) در حالت پیوسته $h(x)$ تابعی پیوسته روی S_X باشد.

نکته ۱.۱. اگر X_1، ...، X_n یک نمونه‌ی تصادفی مستقل و هم توزیع از خانواده‌ی نمایی k پارامتری باشند، آنگاه توزیع توأم X_1، ...،X_n نیز به خانواده‌ی نمایی k پارامتری متعلق است.

اثبات.

$$\begin{aligned}
f_\theta(x_1, \ldots, x_n) &= \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n}\Big\{\exp\sum_{j=1}^{k}\Big[c_j(\theta)T_j(x_i) - B(\theta)\Big]\Big\}h(x_i) \\
&= \exp\Big[\sum_{i=1}^{n} c_j(\theta)\sum_{i=1}^{n} T_j(x^i)_j(x_i) - nB(\theta)\Big]\prod_{i=1}^{n} h(x_i) \\
&= \exp\Big[\sum_{i=1}^{n} c_j(\theta)T_j^*(x_i) - nB(\theta)\Big]h^*(x_i)
\end{aligned}$$

که در آن $h^*(x) = \prod_{i=1}^{n} h(x_i)$ و $T_j^*(x) = \sum_{i=1}^{n} T_j(x_i)$.

□

نکته ۲.۱. اگر X_1، ...، X_n یک نمونه‌ی تصادفی مستقل و هم توزیع از خانواده‌ی نمایی با تابع چگالی $f_\theta(x) = \exp[c(\theta)T(x) - B(\theta)]h(x)$ باشند، در این صورت

$$E_\theta(T(x)) = \frac{B'(\theta)}{c'(\theta)}$$

اثبات. با توجه به خاصیت تابع چگالی که داریم

$$\int \exp[c(\theta)T(x) - B(\theta)]h(x)d\mu(x) \;=\; ۱$$

با مشتق‌گیری از طرفین تساوی بالا نسبت به θ داریم

$$\begin{aligned}\int [c'(\theta)T(x) - B'(\theta)]f_\theta(x)d\mu(x) &= \circ \Rightarrow \int c'(\theta)T(x)f_\theta(x)d\mu(x) - \int B'(\theta)f_\theta(x)d\mu(x) = \circ \\ \Rightarrow c'(\theta)E_\theta(T(x)) - B'(\theta) &= \circ \\ \Rightarrow E_\theta(T(x)) = \frac{B'(\theta)}{C'(\theta)}\end{aligned}$$

□

۴.۱ تعاریف و قضایای مورد نیاز

تعریف ۲.۱. (حد) گوییم تابع $f(x)$ در نقطه $x = a$ دارای حدی معادل L است، هرگاه برای هر $\varepsilon > \circ$، $\delta > \circ$ وجود داشته باشد به قسمی که

$$\circ < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon \Leftrightarrow \lim_{x\to a} f(x) = L$$

به همین ترتیب، می‌توان حدود چپ و راست را تعریف نمود.

برای حد راست داریم: برای هر $\varepsilon > \circ$، $\delta > \circ$ وجود داشته باشد به قسمی که

$$\circ < x - a < \delta \Rightarrow |f(x) - L| < \varepsilon \Leftrightarrow \lim_{x\to a^+} f(x) = L$$

برای حد چپ داریم: برای هر $\varepsilon > \circ$، $\delta > \circ$ وجود داشته باشد به قسمی که

$$\circ < a - x < \delta \Rightarrow |f(x) - L| < \varepsilon \Leftrightarrow \lim_{x\to a^-} f(x) = L$$

تعریف ۳.۱. گوییم تابع $f(x)$ در نقطه $x_\circ$ پیوسته است هرگاه:

$$\lim_{x\to x_\circ} f(x) = f(x_\circ)$$

تعریف ۴.۱. برآوردگر $T(X)$ را برای $\gamma(\theta)$ نااریب گویند، هرگاه برای هر $\theta \in \Theta$،

$$E_\theta[T(X)] = \gamma(\theta)$$

و اگر برآوردگر $T(X)$ برای $\gamma(\theta)$ نااریب نباشد، مقدار $b(\theta)=E_\theta[T(X)]-\gamma(\theta)$ را اریبی برآوردگر $T(X)$ در برآورد $\gamma(\theta)$ گویند.

۴ تعریف زیر از کتاب واسرمن (۲۰۰۶) انتخاب شده است.

تعریف ۵.۱. فرض کنید $\{a_n\}$ و $\{b_n\}$ دنباله‌هایی از اعداد حقیقی باشند. گوییم $a_n=O(b_n)$، هرگاه ثابت مثبت و متناهی C و عدد صحیح مثبت $n(C)$ موجود باشند به قسمی که برای هر $n\geq n(C)$ داشته باشیم

$$|\frac{a_n}{b_n}|\leq C.$$

مثال ۴.۱. فرض کنید $a_n=\gamma^2n^{-1}+kn^{-2}+n^{-3}$ که γ و k متناهی و نامعلوم هستند. در اینصورت رابطه‌ی $a_n=O(n^{-1})$ در این مثال برقرار است. حال اگر $\gamma=0$ باشد، در این‌صورت $a_n=O(n^{-2})$ است و اگر $\gamma=k=0$ قرار دهیم رابطه $a_n=O(n^{-3})$ برقرار خواهد بود. □

مثال ۵.۱. در نظر بگیرید $a_n=\mu^2+n\sigma^2$ که در اینجا σ و μ متناهی هستند رابطه $a_n=O(n)$ را داریم و هنگامی‌که $\sigma=0$ باشد $a_n=O(1)$ خواهد بود. □

تعریف ۶.۱. فرض کنید $\{a_n\}$ و $\{b_n\}$ دنباله‌هایی از اعداد حقیقی باشند. گوییم $a_n=o(b_n)$، هرگاه برای هر $\varepsilon>0$، عدد صحیح مثبت $n(\varepsilon)$ موجود باشد به قسمی که برای هر $n\geq n(\varepsilon)$ داشته باشیم

$$|\frac{a_n}{b_n}|\leq \varepsilon.$$

مثال ۶.۱. در نظر بگیرید $a_n=\gamma^2n^{-1}+kn^{-2}+n^{-3}$ که در اینجا γ و k متناهی هستند رابطه $a_n=o(n^{-\frac{1}{2}})$ و $a_n=O(n^{-1})$ را داریم. اما رابطه $a_n=o(n^{-1})$ تنها در صورتی برقرار خواهد بود که $\gamma=0$ باشد.

□

تعریف ۷.۱. فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی و $\{b_n\}$ یک دنباله دلخواه باشد، $X_n=O_p(b_n)$، هرگاه به ازای هر $\eta>0$، ثابت مثبت و متناهی C، عدد صحیح مثبت $n(\eta)$ موجود باشد، به قسمی که برای هر $n\geq n(\eta)$ داشته باشیم

$$P(|\frac{X_n}{b_n}|\leq C)\geqslant 1-\eta.$$

مثال ۷.۱. فرض کنید Y_n دارای توزیع نمایی با میانگین ۱ باشد آنگاه $Y_n = O_p(1)$ اگر $Z_n = n^{-1} Y_n$باشد، در اینصورت $Z_n = O_p(n^{-1})$. زیرا با فرض $\eta > 0$ داده شده، ثابت مثبت C را طوری انتخاب می‌کنیم که

$$P(\frac{Z_n}{n^{-1}} \leq C) = P(Y_n \leq C) = 1 - e^{-C} \geqslant 1 - \eta \quad \forall \eta > 0.$$

اکنون با فرض

$$e^{-C} < \eta \Rightarrow -C < \ln \eta \Rightarrow C > -\ln \eta$$

داریم $Z_n = O_p(n^{-1})$. □

تعریف ۸.۱. فرض کنید $\{X_n\}$ دنباله‌ای از متغیرهای تصادفی و $\{b_n\}$ یک دنباله دلخواه باشند. گوییم $X_n = o_p(b_n)$، هرگاه به ازای هر $\eta > 0$و $\varepsilon > 0$، عدد صحیح مثبت $n(\varepsilon, \eta)$، موجود باشد، به قسمی که برای هر $n \geqslant n(\varepsilon, \eta)$ داشته باشیم

$$P\big(|\frac{X_n}{b_n}| > \varepsilon \quad for\ some \quad N \geq n\big) < \eta.$$

مثال ۸.۱. فرض کنید Y_n دارای توزیع نمایی با میانگین ۱ باشد و برای $\delta > 0$ رابطه $X_n = \dfrac{Y_n}{n^{1+\delta}}$ برقرار باشد. در این حالت $X_n = o_p(n^{-1})$ است. □

در ادامه چند قضیه در مورد نامساوی‌ها بدون اثبات ذکر می‌شود، که خواننده برای مشاهده اثبات قضایا می‌تواند به کتاب دبارا (۱۹۸۱) مراجعه کند.

قضیه ۱.۱. (همگرایی کراندار) اگر $\{f_n\}$ دنباله‌ای از توابع اندازه‌پذیر باشند، به طوری که

$$lim_{n \to \infty} f_n(x) = f(x)$$

و یک تابع انتگرال‌پذیر g چنان باشد که $|f_n(x)| \leq g(x)$ آنگاه f_n و f انتگرال‌پذیر هستند و

$$\int f d\mu = lim_{n \to \infty} \int f_n d\mu$$

قضیه ۲.۱. (نامساوی هولدر) فرض کنید $1 < p < \infty$ ، $1 < q < \infty$ ، $1/p + 1/q = 1$، و فرض کنید $f \in L^p(\mu)$، $g \in L^q(\mu)$. در این صورت $fg \in L^1(\mu)$ و

$$\int |fg| d\mu \leqslant (\int |f|^p d\mu)^{1/p} . (\int |g|^q d\mu)^{1/q}$$

قضیه ۳.۱. (نامساوی کوشی-شوارتز) نامساوی هولدر در حالتی که $p = q = ۲$ باشد، نامساوی کوشی-شوارتز نامیده می‌شود.

قضیه ۴.۱. (نامساوی مینکوفسکی) فرض کنید $p \geqslant ۱$ و $f, g \in L^p(\mu)$ در این صورت

$$\Big(\int |f+g|^p d\mu \Big)^{۱/p} \leqslant \Big(\int |f|^p d\mu \Big)^{۱/p} + \Big(\int |g|^p d\mu \Big)^{۱/p}$$

تعریف ۹.۱. (بسط تیلور) بسط تیلور تابع $f(x)$ حول نقطه $x_\circ$ به صورت زیر نوشته می‌شود:

$$f(x) \ = \ \sum_{n=\circ}^{\infty} a_n (x - x_\circ)^n$$

که در آن ضرایب این سری عبارتند از:

$$a_\circ = f(x_\circ) \quad a_۱ = \frac{f'(x_\circ)}{۱!} \quad a_۲ = \frac{f''(x_\circ)}{۲!} \dots \quad a_n = \frac{f^n(x_\circ)}{n!}$$

تعریف ۱۰.۱. (σ-میدان) فرض کنید Ω یک مجموعه باشد و $\mathcal{A}$ مجموعه‌ای غیرتهی از زیر مجموعه‌های Ω باشد. $\mathcal{A}$ را یک σ-میدان گویند، به شرط آن که شرایط زیر برای آن برقرار باشد

الف) اگر $A \in \mathcal{A}$ آنگاه $A^C \in \mathcal{A}$

ب) اگر $A_۱, A_۲, \dots,$ مجموعه‌هایی در $\mathcal{A}$ باشند، آنگاه $\bigcup_{n=۱}^{\infty} A_n \in \mathcal{A}$

از تعریف فوق نتیجه می‌شود اگر $\mathcal{A}$ یک σ-میدان باشد، آنگاه $\Omega \in \mathcal{A}$، $\varnothing \in \mathcal{A}$ و $\bigcap_{n=۱}^{\infty} A_n \in \mathcal{A}$.

تعریف ۱۱.۱. (اندازه) یک اندازه μ یک تابع مجموعه‌ای روی σ-میدان $\mathcal{A}$ از زیر مجموعه‌های Ω است یعنی $\mu : A \to [\circ, \infty)$ هرگاه در شرایط زیر صدق کند

الف) اگر $A \in \mathcal{A}$ آنگاه $\mu(A) \geq \circ$

ب) اگر $A_۱, A_۲, \dots, \in \mathcal{A}$ و A_iها مجزا باشند، یعنی $\forall i \neq j$, $A_i \bigcap A_j = \varnothing$ آنگاه

$$\mu \left[\bigcup_{n=۱}^{\infty} A_n \right] = \sum_{n=۱}^{\infty} \mu(A_n)$$

اگر $\mathcal{A}$ یک σ-میدان از زیرمجموعه‌های Ω باشد آنگاه $(\Omega, \mathcal{A})$ را یک فضای اندازه‌پذیر گویند و مجموعه‌ی $A \in \mathcal{A}$ را اندازه‌پذیر گویند. همچنین اگر μ یک اندازه باشد که روی $(\Omega, \mathcal{A})$ تعریف شده باشد، آنگاه سه تایی $(\Omega, \mathcal{A}, \mu)$ را یک فضای اندازه می‌نامند.

تعریف ۱.۱۲. (اندازه لبگ)

فرض کنید Ω فضای اقلیدسی n بعدی باشد، $\Omega = \{(w_1, ..., w_n) \mid w_i \in \mathbb{R}, \forall i\}$ و $\mathcal{A}$ کوچکترین σ-میدانی باشد که شامل تمام ابر مستطیل‌های باز زیر باشد:

$$A = \{(w_1, ..., w_n) \mid a_i < w_i < b_i, -\infty < a_i < b_i < \infty\}$$

به اعضای $\mathcal{A}$ مجموعه‌های بورل گویند. تنها یک اندازه μ را می‌توان روی این $\mathcal{A}$ معرفی کرد که به مجموعه A فوق اندازه

$$\mu(A) = (b_1 - a_1)(b_2 - a_2), \ldots, (b_n - a_n)$$

را نسبت می‌دهد (که حجم مجموعه A است) و به آن اندازه لبگ گویند.

تعریف ۱.۱۳. (قابل شناسایی)

فرض کنید $\varphi = \{f_\theta, \theta \in \Theta\}$ یک مدل آماری باشد که در اینجا فضای پارامتر Θ دارای بعد متناهی یا نامتناهی است. گوییم φ قابل شناسایی است اگر $\theta \to f_\theta$ یک به یک باشد. بدین معنی که برای هر $\theta_1, \theta_2 \in \Theta$ داشته باشیم

$$f_{\theta_1} = f_{\theta_2} \Rightarrow \theta_1 = \theta_2$$

مثال ۱.۹. فرض کنید φ یک خانواده مکانی-مقیاسی نرمال باشد:

$$\varphi = \{f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} \mid \theta = (\mu, \sigma) : \mu \in \mathbb{R}, \sigma > 0\}$$

در این صورت

$$\begin{aligned}
f_{\theta_1} = f_{\theta_2} \quad &\Leftrightarrow \quad \frac{1}{\sqrt{2\pi}\sigma_1} \exp \frac{-1}{2\sigma_1^2}(x-\mu_1)^2 = \frac{1}{\sqrt{2\pi}\sigma_2} \exp \frac{-1}{2\sigma_2^2}(x-\mu_2)^2 \\
&\Leftrightarrow \quad \frac{1}{\sigma_1^2}(x-\mu_1)^2 + \ln \sigma_1^2 = \frac{1}{\sigma_2^2}(x-\mu_2)^2 + \ln \sigma_2^2 \\
&\Leftrightarrow \quad x^2(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}) - 2x(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2}) + (\ln \sigma_1^2 - \ln \sigma_2^2) = 0
\end{aligned}$$

این عبارت برای همه مقادیر x زمانی مساوی صفر است که همه ضرایب صفر باشند و این زمانی میسر است که $\mu_1 = \mu_2$ و $|\sigma_1| = |\sigma_2|$ باشد. چون پارامتر مکانی به مقادیر بزرگتر از صفر محدود شده است مدل زمانی قابل تشخیص است که $f_{\theta_1} = f_{\theta_2} \Rightarrow \theta_1 = \theta_2$. □

قضیه ۵.۱. (روش دلتا[۱]) فرض کنید $\{Y_n\}$ دنباله‌ای از متغیرهای تصادفی باشند

$$\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2)$$

در این صورت برای تابع g و یک مقدار θ، با این فرض که $g'(\theta)$ وجود داشته و مخالف صفر باشد، داریم

$$\sqrt{n}\,[g(Y_n) - g(\theta)] \xrightarrow{D} N(0, \sigma^2 (g'(\theta))^2)$$

برای مشاهده اثبات این قضیه می‌توانید به کتاب کسلا و برگر (۲۰۰۲) مراجعه کنید.

۵.۱ تابع توزیع تجربی

در این قسمت به بررسی برآورد تابع توزیع می‌پردازیم، شاید بتوان گفت که این مسئله به تنهایی چندان درخور توجه نمی‌باشد بلکه آنچه مهم است، این است که این مسئله، نخستین مسئله در برآورد توابع آماری، انجام آزمون فرض، بدست آوردن فاصله اطمینان و غیره می‌باشد.

تعریف ۱۴.۱. (تابع توزیع تجربی) فرض کنید $\{X_i, i \geq 1\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع، با توزیع مشترک F باشد. تابع توزیع تجربی این متغیرها عبارت است از

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x) \quad x \in \mathbb{R},$$

که در آن $I(.)$ تابع نشانگر می‌باشد.

در قضیه زیر خواص تابع توزیع تجربی بدون اثبات آورده شده است، برای مشاهده اثبات این قضیه می‌توانید به کتاب واسرمن (۲۰۰۶) مراجعه کنید.

قضیه ۶.۱. (خواص تابع توزیع تجربی) فرض کنید $X_1, \ldots, X_n$ نمونه‌ای تصادفی از توزیع F با تابع توزیع تجربی F_n باشد، خواص زیر برقرارند:

الف) نااریبی $E(F_n(x)) = F(x)$

ب) سازگاری ضعیف $F_n(x) \xrightarrow{P} F(x) \quad \forall\, x \in \mathbb{R}$

[۱] Delta Method

ج) سازگاری قوی $F_n(x) \xrightarrow{a.s.} F(x) \quad \forall\, x \in \mathbb{R}$

د) سازگاری قوی به طور یکنواخت

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \xrightarrow{a.s.} ۰ \qquad (۱.۱)$$

مثال ۱۰.۱. فرض کنید مقادیر ۳, ۱/۵, ۳/۵, ۶ داده شده‌اند، $F_n(x)$ را محاسبه کنید.

حل: در اینجا $n = ۴$ می‌باشد، لذا $P(X_i = x) = \frac{۱}{۴}$

$$F_n(x) = \begin{cases} ۰ & x < ۱/۵, \\ ۱/۴ & ۱/۵ \le x < ۳ \\ ۲/۴ & ۳ \le x < ۳/۵ \\ ۳/۴ & ۳/۵ \le x < ۶ \\ ۱ & x \ge ۶ \end{cases}$$

□

تعریف ۱۵.۱. (برآوردگر جایگزینی) یک برآوردگر طبیعی و منطقی $T(F)$، برآوردگر جایگزینی می‌باشد که به صورت $T(F_n)$ تعریف می‌شود .

مثال ۱۱.۱. فرض کنید $X_۱$، ...، X_n نمونه‌ای تصادفی از توزیع F باشند، حال می‌خواهیم برآوردگر جایگزینی برای $T(F) = \int xdF(x)$ بدست آوریم.

حل:

$$\begin{aligned} T(F_n) &= \int xdF_n(x) = \int xd(\frac{۱}{n}\sum_{i=۱}^{n} I(X_i \le x)) \\ &= \frac{۱}{n}\sum_{i=۱}^{n}\int xdI(X_i \le x) \\ &= \frac{۱}{n}\sum_{i=۱}^{n} X_i \\ &= \bar{X} \end{aligned}$$

□

در این قسمت به یادآوری تعاریف همگرایی از کتاب گات (۲۰۰۴) می‌پردازیم.

تعریف ۱.۱۶. (همگرایی تقریباً مسلم) دنباله‌ی متغیرهای تصادفی X_n به متغیر تصادفی X همگرای تقریباً مسلم($a.s.$) است، اگر و تنها اگر وقتی که n به سمت بی‌نهایت میل می‌کند

$$P\big(\omega : X_n(\omega) \longrightarrow X(\omega)\big) = ۱$$

و می‌نویسیم $X_n \xrightarrow{a.s} X$.

تعریف ۱.۱۷. (همگرایی در احتمال) دنباله‌ی متغیرهای تصادفی X_n را وقتی که n به سمت بی‌نهایت میل می‌کند در احتمال به متغیر تصادفی X همگرا گویند هرگاه برای هر $\epsilon > ۰$

$$P\Big(\big|X_n - X\big| > \epsilon\Big) \longrightarrow ۰$$

و می‌نویسیم $X_n \xrightarrow{P} X$.

تعریف ۱.۱۸. (همگرایی در توزیع) مجموعه نقاط پیوستگی تابع $F_X(.)$ را بصورت

$$C(F_X) = \Big\{x : \text{در آن پیوسته است } F_X\Big\}$$

در نظر بگیرید. X_n در توزیع به X همگراست اگر و تنها اگر برای هر $x \in C(F_X)$ وقتی‌که n به سمت بی‌نهایت میل می‌کند، داشته باشیم

$$F_{X_n}(x) \longrightarrow F_X(x)$$

و می‌نویسیم $X_n \xrightarrow{D} X$.

در انتها به تعریف همگرایی مغلوب می‌پردازیم.

تعریف ۱.۱۹. فرض کنید برای هر n، $\big|X_n\big| < Y$ باشد که در آن $E(Y) < \infty$ و داشته باشیم $X_n \xrightarrow{a.s} X$ در اینصورت وقتی که $n \longrightarrow \infty$، آنگاه $E\big|X_n - X\big| \longrightarrow ۰$ و در نتیجه $E(X_n) \longrightarrow E(X)$.

در قسمت بعد به تعریف تابع هسته و انواع آن پرداخته‌ایم که این مطالب از واسرمن (۲۰۰۶) انتخاب شده است.