دانشگاه سیستان و بلوچستان

پایان نامهٔ کارشناسی ارشد ریاضی محض (هندسه)

عنوان:

برآورد ناپارامتری شبکه‌های عصبی به کمک نمای لیاپانوف و آزمون مستقیم برای تئوری آشوب

استاد راهنما:

دکتر حاجی‌محمد محمدی‌نژاد

تحقیق و نگارش:

سمیه جعفری‌نیا

۱۳۹۰

چکیده

این پایان‌نامه توزیع مجانبی برآوردگر ناپارامتری شبکه‌های عصبی به کمک نمای لیاپانوف در یک سیستم آشوبناک، را استنتاج می‌کند. مثبت بودن نمای لیاپانوف یک تعریف عملگری از آشوب است. در این راستا قالب‌بندی آماری برای آزمون فرضیه آشوب بر اساس نمای لیاپانوف برآورد می‌کنیم و هم‌چنین برآوردگر واریانس پایدار را معرفی می‌کنیم. یک مطالعه شبیه سازی برای ارزیابی عملکرد نمونه‌های کوچک گزارش شده است. هم‌چنین روش کارمان را برای داده‌های برگشتی سهام روزانه به‌کار می‌بریم. در بیشتر موارد، فرضیه آشوب سری‌های بازگشتی سهام در سطح یک درصد بایک استثنا در بعضی از تبدیل‌های بازگشتی مطلق، رد شده است.

پیشگفتار

نمای لیاپانوف یک اندازه مفید از پایایی یک سیستم دینامیک است. این آزمون میانگین همگرایی یا واگرایی نمایی را بین مسیرهایی که تفاوت اندکی در مقدارهای اولیه دارند، اندازه‌گیری می‌کند. برای به‌دست آوردن نمای لیاپانوف از داده‌های مشاهده شده در سال ۱۹۸۵ ایخمن و رول [۱] در [۱۰] و در سال ۱۹۸۶ ایخمن [۲] در [۹]، یک روش بر پایه رگرسیون ناپارامتری پیشنهاد دادند که‌این روش به عنوان روش ژاکوبین شناخته می‌شود. درحالی‌که هر برآوردگر رگرسیون ناپارامتری می‌تواند به عنوان روش ژاکوبین به‌کار برده شود یکی از بیشترین استفاده کاربردی این روش برآوردگر نمای لیاپانوف بر اساس شبکه‌های عصبی طرح شده در سال ۱۹۹۲ توسط نیچکا و همکارانش [۳] در [۱۸] است. برای مثال استفاده کاربردهای این رویکرد در اقتصاد شامل مطالعه روی دسته پولی در سال ۱۹۹۵ توسط سرلیتس [۴] در [۱۹]، تجزیه و تحلیل مقدار ارز خارجی در سال ۱۹۹۲ توسط دیچارت و گنسی [5] در [۷] ، تجزیه و تحلیل سری‌های بازگشتی سهام در سال ۱۹۹۷ توسط آبهانکار و همکاران [6] در [۱] می‌شود.

در این پایان‌نامه ابتدا توزیع مجانبی از برآوردگر شبکه عصبی نمای لیاپانوف را بدست می‌آوریم. یک چارچوب آماری قراردادی که علامت نمای لیاپانوف است، بر پایه برآوردگر پایدار از واریانس مجانبی معرفی می‌شود. در سال ۱۹۹۹ توسط وانگ و لینتون [۷] در [۲۱]، نرمال بودن مجانبی از برآوردگر پایه ژاکوبین که از یک رگرسیون ناپارامتری نوع هسته استفاده شده به‌دست آمده است. ایده اصلی ترکیب نتایج نظری وانگ و لینتون و نتایج اخیر مجانبی شبکه‌های عصبی بدست آمده در سال ۱۹۹۹ توسط چن و وایت [۸] در [۲۱] و دیگران است. شرایط برای نرمال مجانبی برآوردگر، از نظر تعداد واحدهای پنهان در مدل‌های شبکه عصبی

---

[۱] Ruelle

[۲] Eckmann

[۳] Nychka

[۴] Serletis

[5] Dechart,Gencay

[6] Abhyankar

[۷] wang, Linton

[۸] Chen,White

همانند طول بلوک، برای هر دو حالت چند بعدی و یک بعدی بدست آمده است، مقدار صعود لازم از طول بلوک و مقدار همگرایی از برآوردگر شبکه عصبی از نمای لیاپانوف با آنهایی که بر پایه برآوردگر هسته هستند مقایسه می‌شود. مثبت بودن نمای لیاپانوف در یک سیستم غیر خطی تلف کننده کراندار، یک تعریف قراردادی عمومی به‌کار رفته از آشوب است(ایخمن و رول در [۱۰]در سال ۱۹۸۵). چون وجود آشوب تصادفی در یک سیستم را پذیرفتیم، برآوردگر پایدار واریانس مجانبی از برآوردگر نمای لیاپانوف یک چهارچوب آماری قراردادی برای آزمون فرضیه مثبت بودن نمای لیاپانوف در یک محیط تصادفی پیشنهاد می‌کند، به عبارت دیگر می‌توانیم یک آزمون مستقیم برای استفاده آشوب با خطای استاندارد پایدار پیشنهاد شده در این کار بسازیم. بعد از مطالعات نظری سابق روی خواص آماری از برآوردگر شبکه عصبی نمای لیاپانوف در سال ۱۹۹۲ توسط مکافی [۹] در [۱۵]، روی یک رده پیش خور تک لایه پنهان از شبکه‌های عصبی مصنوعی تمرکز می‌کنیم. به نظر می رسد بیشترین مزیت نظری قابل توجه از استفاده شبکه‌های عصبی خواص تقریبی عمومی آنها باشد، به طور نظری انتظار می‌رود شبکه‌های عصبی نسبت به دیگر روش‌های تقریب حداقل در محدوده رده خاصی از توابع در نظر گرفته شده بهتر عمل می‌کنند. به طور کلی نتایج شبیه سازی موجود در مطالعات پیشین به روش شبکه‌های عصبی توجه می‌کنند، با توجه به انعطاف پذیری شبکه‌های عصبی خواص تقریبی نزدیک به دقیق شبکه‌های عصبی را نشان می‌دهد، حتی اگر تابع غیرخطی به اندازه کافی پیچیده برای تولید آشوب باشد، نیرومندی شبکه‌های عصبی به انتخاب تعداد واحدهای پنهان است.

[۹] McCaffrey

فهرست مندرجات

فصل ۱

تعاریف و مفاهیم مقدماتی

مقدمه

در این فصل تعاریف و قضایا و مفاهیم مقدماتی جهت آشنایی با موضوع پایان‌نامه آورده شده است. این فصل شامل چهار بخش می‌باشد. در بخش اول به تعاریف مقدماتی برای معرفی نرم سوبولوف و نامساوی الحاق و مفاهیم مقدماتی آماری می‌پردازیم. در بخش دوم مفاهیم مقدماتی نظریه آشوب را بیان می‌کنیم. بخش سوم به معرفی سیستم‌های دینامیکی اختصاص دارد. در بخش چهارم مفاهیم شبکه‌های عصبی را بیان می‌کنیم.

۱ — ۱ تعاریف و مفاهیم مقدماتی

در این بخش تعاریف و مفاهیم مقدماتی ریاضی و آماری را بیان می کنیم.

تعریف ۱.۱.۱. فرض کنید X یک مجموعه ناتهی باشد، گردایه S از زیر مجموعه‌های X را یک نیم‌حلقه گویند، اگر در سه خاصیت زیر صدق کند:

(۱) مجموعه تهی متعلق به S باشد؛

(۲) هر گاه $A, B \in S$، آنگاه $A \cap B \in S$؛

(۳) به ازای هر دو مجموعه از S، تفاضل آنها را به توان به صورت اجتماعی متناهی از اعضای از هم جدای S نوشت، یعنی، به ازای هر $A, B \in S$، مجموعه‌هایی چون $C_1, ..., C_n$ در S وجود داشته باشند که $A - B = \bigcup_{i=1}^{n} C_i$ و $i \neq j$، $C_i \cap C_j = \emptyset$.

تعریف ۲.۱.۱. یک حلقه گردایه‌ای ناتهی از مجموعه‌ها، مانند R از مجموعه X است که از دو خاصیت زیر برخوردار است:

(۱) هر گاه $A, B \in R$، آنگاه $A \cup B \in R$؛

(۲) هر گاه $A, B \in R$**، آنگاه** $A - B \in R$.

تعریف ۳.۱.۱. فرض کنید S **نیم‌حلقه‌ای از زیر مجموعه‌های** X **باشد. تابع مجموعه‌ای** $\mu : S \to [\circ, \infty]$ **را یک اندازه روی** S **گویند، اگر دارای خواص زیر باشد:**

(۱) $\mu(\emptyset) = \circ$؛

(۲) هر گاه $\{A_n\}$ **دنباله‌ای از هم‌جدا از** S **که** $\bigcup_{n=1}^{\infty} A_n \in S$ **باشد، آنگاه**

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right).$$

سه تایی (X, S, μ) **که در آن** X **یک مجموعه ناتهی،** S **زیرحلقه‌ای از زیر مجموعه‌های** X **و** μ **یک اندازه روی** S **است، یک فضای اندازه نام دارد.**

تعریف ۴.۱.۱. فرض کنید S **نیم‌حلقه مرکب از مجموعه تهی و تمام مجموعه‌ها به شکل** $A = \prod_{i=1}^{n} [a_i, b_i)$ **که به ازای هر** i**،** $-\infty < a_i < b_i < \infty$ **باشد. تابع مجموعه‌ای** $\lambda : S \to [\circ, \infty)$ **با تعریف** $\lambda(\emptyset) = \circ$**،** $\lambda\left[\prod_{i=1}^{n} [a_i, b_i)\right] = \prod_{i=1}^{n} (b_i - a_i)$ **یک اندازه است که آن را اندازه لبگ روی** S **می نامند.**

تعریف ۵.۱.۱. تابع توسعه یافته حقیقی غیر منفی مجموعه‌ای μ^* **را اندازه خارجی گویند، هر گاه دارای خواص زیر باشد:**

(۱) $\mu^*(\emptyset) = \circ$؛

(۲)$A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$؛

(۳) اگر E_n **یک دنباله دلخواه از زیرمجموعه‌های** X **باشد، آنگاه**

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} \mu^*\left(E_n\right).$$

تعریف ۶.۱.۱. دنباله $\{f_n\}$ از توابع اندازه‌پذیر همگرا به $f \in m$ در اندازه (یا در احتمال) است اگر به ازای هر $\varepsilon > ۰$،

$$\lim \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = ۰$$

و آن را با $f_n \stackrel{\mu}{\to} f$ نمایش می‌دهند.

گزاره ۷.۱.۱. فرض کنید (X, S, μ) یک فضای اندازه باشد. اگر f یک تابع اندازه‌پذیر باشد، $|f|^p$ نیز به ازای هر $p > ۰$ اندازه‌پذیر است.

برهان. به [۲۲] مراجعه شود.

تعریف ۸.۱.۱. فرض کنید $۰ < p < \infty$ در این صورت گردایه تمام توابع اندازه‌پذیر f که $|f|^p$ انتگرال‌پذیر است با $L_p(\mu)$ نشان داده می‌شود.

گزاره ۹.۱.۱. $L_p(\mu)$ یک فضای برداری است.

برهان. به [۲۲] مراجعه شود.

تعریف ۱۰.۱.۱. فرض کنید X یک مجموعه ناتهی باشد. تابع حقیقی d تعریف شده روی $X \times X$ را یک متر روی X نامند، هرگاه برای هر $x, y, z \in X$ شرایط زیر برقرار باشد:

(۱) $d(x, y) \geq ۰$ و $d(x, y) = ۰$ اگر و تنها اگر $x = y$؛

(۲) $d(x, y) = d(y, x)$؛

(۳) $d(x, y) \leq d(x, z) + d(z, y)$.

تعریف ۱.۱.۱۱. نقطه p یک نقطه حدی مجموعه E است، هر گاه هر همسایگی p شامل نقطه‌ای چون $q \in E$ غیر از p باشد.

تعریف ۱.۱.۱۲. فرض کنید X یک فضای متری و E زیر مجموعه‌ای از X باشد. E در X چگال است، هر گاه هر نقطه X یک نقطه حدی E یا یک نقطه E و یا هر دو باشد.

تعریف ۱.۱.۱۳. فرض کنید $\{f_n\}$ دنباله‌ای از توابع باشد که روی مجموعه E تعریف شده‌اند. $\{f_n\}$ روی E به طوریکنواخت کراندار است، هر گاه عددی مانند M باشد به‌طوری‌که

$$|f_n(x)| \leq M \qquad (x \in E, \quad n = ۱, ۲, ۳, ...).$$

تعریف ۱.۱.۱۴. دنباله‌ای از توابع مانند $\{f_n\}$ که $n = ۱, ۲, ۳, ...$، روی E به طور یکنواخت به تابع f همگراست، هر گاه به ازای هر $\varepsilon > ۰$ عدد صحیحی مانند N باشد به‌طوری‌که $n \geq N$ نامساوی $|f_n(x) - f(x)| \leq \varepsilon$ را به ازای هر $x \in E$ ایجاب کند.

تعریف ۱.۱.۱۵. فرض کنید (X, d) یک فضای متری و I یک بازه روی خط حقیقی $\mathbb{R}$ باشد. تابع $f : I \to R$ روی I پیوسته مطلق است، اگر برای هر ε مثبت $\delta > ۰$ ای و یک دنباله متناهی دوبه دو مجزا (x_k, y_k) از I وجود داشته باشند که در رابطه زیر صدق کنند:

$$\sum_k |y_k - x_k| < \delta \Rightarrow \sum_k |f(y_k) - f(x_k)| < \varepsilon.$$

تعریف ۱.۱.۱۶. فرض کنید X فضای برداری روی میدان اسکالر F باشد. نرم بردار $x \in X$ را با $\|x\|$ نمایش می‌دهند. در واقع یک نرم تعریف شده روی X یک تابعی چون $\|.\| : X \to R^+ \cup \{۰\}$ که $x \longrightarrow \|x\|$ واجد خواص زیر است:

(۱) به ازای هر $x \in X$؛ $\|x\| \geq ۰$ و $\|x\| = ۰$ اگر و تنها اگر $x = ۰$؛

(۲) به ازای هر $x \in X$ و هر $\alpha \in \mathbb{F}$، $\|\alpha x\| = |\alpha| . \|x\|$؛

(۳) به ازای هر $x, y \in X$، $\|x+y\| \leq \|x\| + \|y\|$.
در این صورت زوج مرتب $(X, \|.\|)$ یک فضای برداری نرم‌دار است.

تعریف ۱.۱.۱۷. فرض کنید که U و V مجموعه‌های بازی به ترتیب در $\mathbb{R}^n$ و $\mathbb{R}^p$ باشند. نگاشت $f : U \subseteq \mathbb{R}^n \to V \subseteq \mathbb{R}^p$ در نقطه $x_\circ \in U$ مشتق‌پذیر است، هر گاه یک نگاشت خطی چون $(Df)_{x_\circ} : \mathbb{R}^n \to \mathbb{R}^p$ موجود باشد به‌طوری‌که $\lim\limits_{\|\vec{h}\| \to \circ} \frac{f(x_\circ + \vec{h}) - f(x_\circ) - (Df)_{x_\circ} \vec{h}}{\vec{h}} = \circ$.
اگر f در نقطه $x_\circ \in U$ مشتق پذیر باشد، آنگاه $(Df)_{x_\circ}$ به طور یکتا توسط ماتریس زیر تعیین می‌شود:

$$(Df)_{x_\circ} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \frac{\partial f_p}{\partial x_2} & \cdots & \frac{\partial fp}{\partial x_n} \end{bmatrix}_{(x_\circ)}$$

نگاشت $(Df)_{x_\circ}$ را نگاشت مشتق f در نقطه $x_\circ$ و ماتریس $(Df)_{x_\circ}$ را ماتریس ژاکوبین f در $x_\circ$ گویند.

حال نرم سوبولف به صورت زیر تعریف می‌شود.

تعریف ۱.۱.۱۸. فرض کنید $\mu = (\mu_1, ..., \mu_d)^t$ برداری از ثابت‌های صحیح غیرمنفی و $x^\mu = \prod\limits_{i=1}^{d} x_i^{\mu_i}$، که در آن x یک بردار به شکل $x = (x_1, ..., x_d)^t$ باشد، که برای هر تابع حقیقی $g(x)$ در $\mathbb{R}^d$ و برای هر اسکالر μ رابطه $D^\mu g(x) = \frac{\partial^{|\mu|} g(x)}{\partial x_1^{\mu_1}, ..., \partial x_d^{\mu_d}}$ برقرار باشد. در این صورت $D^\mu g(x)$ را مشتق مرتبه μام g در x می نامیم.
از قرارداد $D^\circ g(x) = g(x)$ و $D^1 g(x) = Dg(x)$ استفاده می‌کنیم. از β_d^m برای نمایش فضای سوبولف همه توابع وزنی در R^d که مشتق به‌طوریکنواخت کراندار و پیوسته تا مرتبه m دارند استفاده

می‌کنیم ونرم سوبولف به صورت زیر تعریف می‌شود:

$$\|g\|_{\beta_d^m} \equiv \max_{0 \leq |\mu| < m} \sup_{x \in R^d} |D^\mu g(x)| < \infty \qquad \forall g \in \beta_d^m.$$

تعریف ۱۹.۱.۱ (نامساوی کوشی شوارتز). فرض کنید $f, g \in L_2(\mu)$. آنگاه $fg \in L_1(\mu)$ و

$$\int |fg| d\mu \leq (\int |f|^2 d\mu)^{1/2} . (\int |g|^2 d\mu)^{1/2} = \|f\|_2 . \|g\|_2 .$$

برهان. به [۲۲] مراجعه شود.

تعریف ۲۰.۱.۱ (نامساوی الحاق). فرض کنید S پاره‌خطی از خط حقیقی و $f(x)$ تابع اندازه‌پذیر تعریف شده روی S باشد، $1 \leq r, p, q$، اعداد حقیقی و k, l که در آن $0 \leq k < l$ اعداد صحیح باشند، هم‌چنین

$$\|f\|_{L_{p(s)}} = \begin{cases} \left(\int_s |f(x)|^p dx\right)^{1/p} & 1 \leq p < \infty \\ ess \sup_{x \in s} |f(x)| & p = \infty \end{cases}$$

در این صورت گوییم تابع $f(x)$ به رده $W_{p,r}^l(s)$ تعلق دارد، اگر

(۱) $f(x)$ و $f^{(l-1)}(x)$ روی هر پاره خط $[a, b] \subset S$ پیوسته مطلق باشد،

(۲) $\left\|f^{(l)}\right\|_{L_r(s)} < \infty$، $\|f\|_{L_p(s)} < \infty$.

قضایای زیرتعیین می کند که برای مجموعه داده شده از اعداد $p, q, r, k, l, \alpha, \beta$ و هر تابع $f(x) \in W_{p,r}^l(s)$ آیا نامساوی،

$$\left\|f^k\right\|_{L_q(s)} \leq k \|f\|_{L_p(s)}^\alpha \left\|f^l\right\|_{L_r(s)}^\beta$$

برقرار است، که در آن ثابت k به f بستگی ندارد این نامساوی را نامساوی برای توابع $W_{p,r}^l(s)$ می‌نامیم. ملاحظه می شود که اگر $p, q, r \geq 1$ و $0 \leq k < l$ و $(l-k)p^{-l} + kr^{-l} \geq lq^{-l}$، آنگاه

برای رده ای از توابع $W^l_{p,r}(s)$ و برای هر δ که $0 < \delta \leq s$ نامساوی زیر بر قرار است:

$$\left\|f^k\right\|_{L_q(s)} \leq A(\delta^{-k-p^{-1}+q^{-1}} \left\|f\right\|_{L_p(s)} + \delta^{l-k-r^{-1}+q^{-1}} \left\|f^{(l)}\right\|_{L_r(s)}),$$

که در آن A به δ و S بستگی ندارد.

برهان . به [۲۰] مراجعه شود.

قضیه۲۱.۱.۱. فرض کنید $p, q, r \geq 1$ اعداد حقیقی و k, l که $0 \leq k < l$ صحیح و $E = [0, \infty)$ باشند. آنگاه نامساوی زیر برای رده ای از توابع $W^l_{p,r}(E)$ برقرار است:

$$\left\|f^{(k)}\right\|_{L_q(E)} \leq K \left\|f\right\|^{\alpha}_{L_p(E)} \left\|f^{(l)}\right\|^{\beta}_{L_r(E)},$$

که در آن شرایط زیر برای وجود این نامساوی شرایط لازم و کافی هستند:

(۱) $\alpha = \frac{l-k-r^{-1}+q^{-1}}{l-r^{-1}+p^{-1}}$ و $\beta = \frac{k-q^{-1}+p^{-1}}{l-r^{-1}+p^{-1}}$؛

(۲) $(l-k)p^{-1} + kr^{-1} \geq lq^{-1}$؛

برهان. به [۱۹] مراجعه شود.

تعریف ۲۲.۱.۱. صفت یا ویژگی نامعلوم یک جامعه را پارامتر آن جامعه گویند، که معمولا نامعلوم است و آنرا بدست می‌آورند یا برآورد می‌کنند.

تعریف ۲۳.۱.۱. اگر S یک فضای نمونه‌ای با یک اندازه احتمال و X یک تابع با مقادیر حقیقی باشد که روی عناصر S تعریف شده است، آنگاه X متغیر تصادفی نامیده می‌شود.

تعریف ۲۴.۱.۱. دو متغیر تصادفی X و Y را هم‌توزیع می نامند، هرگاه برای هر عدد حقیقی z،

$$P(X \leq z) = P(Y \leq z).$$

تعریف ۲۵.۱.۱. متغیرهای تصادفی $X_n, ..., X_2, X_1$ یک نمونه تصادفی n تائی از X هستند، هرگاه مستقل از یکدیگر و با X هم‌توزیع باشند.

تعریف ۲۶.۱.۱. تابع $U = g(X_1, X_2, ..., X_n)$ از نمونه تصادفی $X_n, ..., X_2, X_1$ را که در آن پارامتر مجهول نباشد، یک آماره می‌نامند.

تعریف ۲۷.۱.۱. هرگاه به جای نمونه تصادفی یافته‌های آن را در آماره $U = g(X_1, X_2, ..., X_n)$ بگذاریم، یافته این آماره عدد $u = g(x_1, x_2, ..., x_n)$ می‌شود. اگر این عدد را به عنوان مقدار تقریبی پارامتر به‌کار بریم، آنگاه آماره U را یک بر آوردگر و عدد u را یک برآورد برای پارامتر θ می‌گویند.

تعریف ۲۸.۱.۱. فرایند تصادفی با این خاصیت که در یک سلسله از پیشامدهای تصادفی، احتمال وقوع هر پیشامد تنها به نتیجه ماقبل آن بستگی دارد را فرایند مارکف[1] می‌نامند.

تعریف ۲۹.۱.۱. فرض کنید X یک متغیر تصادفی روی یک احتمال باشد. تابع حقیقی $F_x(x) = P(X \leq x)$ را که در آن x یک عدد حقیقی است، تابع توزیع X می‌نامند.

تعریف ۳۰.۱.۱. تابعی که احتمال را در هر نقطه برای متغیر تصادفی گسسته تعیین می‌کند تابع چگالی احتمال گویند.

تعریف ۳۱.۱.۱. فرض کنید A و B دو پیشامد باشند به گونه‌ای که $P(B) > 0$. آنگاه احتمال شرطی پیش آمد A هنگامی که B رخ داده است را با $P(A|B)$نمایش داده و به صورت زیر تعریف

[1] Markov

می کنیم:

$$P(A\,|B) = \frac{P(A\cap B)}{P(B)}, \qquad P(B) \neq \circ.$$

تعریف ۳۲.۱.۱. متغیر تصادفی X پیوسته مطلق است، اگر تابع چگالی احتمال $f_x : R \to [\circ, \text{۱}]$ **برای X وجود داشته باشد به‌طوری‌که برای هر بازه** $[a,b] \subset \mathbb{R}$، $P\left(X \in [a,b]\right) = \int_a^b f_X(x)dx$.

تعریف ۳۳.۱.۱. فرض کنید X یک متغیر تصادفی با چگالی $f(x)$ باشد. امید ریاضی X عددی به صورت زیرتعریف می شود:

$$E\left(x\right) = \begin{cases} \sum\limits_x xf(x) \\ \int_{-\infty}^{+\infty} xf(x)dx. \end{cases}$$

عبارت اول برای حالت گسسته و عبارت دوم برای حالت پیوسته است. امید ریاضی با μ نشان داده می شود.

تعریف ۳۴.۱.۱. اگر X یک متغیر تصادفی با چگالی $f(x)$ و $y = g(x)$ یک تابع حقیقی باشد، امید ریاضی $g(x)$ به صورت زیر است:

$$E\left[g(x)\right] = \begin{cases} \sum\limits_x g(x)f(x) \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, \end{cases}$$

که عبارت اول برای حالت گسسته و عبارت دوم برای حالت پیوسته است.

تعریف ۳۵.۱.۱. امید شرطی متغیر تصادفی گسسته X هنگامی که متغیر تصادفی گسسته Y مقدار y را دارد به صورت زیر تعریف می‌شود:

$$E\left(X\,|Y = y\right) = \sum_x xf_{x|y}\left(x\,|y\right).$$

تعریف ۳۶.۱.۱. میانگین $(X-\mu)^{\text{۲}}$ **را واریانس X می‌نامند و با** $var(x)$ **یا** $\sigma^{\text{۲}}$ **نشان می‌دهند، به عبارت دیگر** $var(X) = \sigma^{\text{۲}} = E((X-\mu)^{\text{۲}}) \geq \circ$.

تعریف ۳۷.۱.۱. فرض کنید X و Y دو متغیر تصادفی با میانگین‌های μ_X و μ_Y باشند. کواریانس این دو متغیر به صورت زیر تعریف می‌شود:

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$

تعریف ۳۸.۱.۱. چگالی زیر را چگالی نرمال با پارامترهای μ و $\sigma^۲$ می‌نامند:

$$f(x;\mu,\sigma^۲) = \frac{۱}{\sqrt{۲\pi\sigma^۲}} e^{-\frac{۱}{۲\sigma^۲}(x-\mu)^۲} \qquad -\infty < x < \infty.$$

هر گاه متغیر تصادفی X دارای این چگالی باشد، می‌گویند X دارای توزیع نرمال با پارامترهای μ و $\sigma^۲$ می‌باشد و با $X \sim N(\mu,\sigma^۲)$ نمایش داده می‌شود.

قضیه ۳۹.۱.۱(قضیه حد مرکزی). فرض کنید $X_۱, X_۲, ..., X_n$ یک نمونه تصادفی از توزیع $F(x)$ با میانگین و واریانس μ و $\sigma^۲$ باشد، به طوری‌که $\sigma^۲ < \infty$. اگر n به اندازه کافی بزرگ باشد، دو نتیجه تقریبی زیر را داریم:

(۱) $T_n = X_۱ + X_۲ + ... + X_n$ تقریبا دارای توزیع $N(n\mu, n\sigma^۲)$ است؛

(۲) $\bar{X}_n = \frac{X_۱+X_۲+...+X_n}{n}$ تقریبا دارای توزیع $N(\mu, \frac{\sigma^۲}{n})$ است.

برهان. به [۲۲] مراجعه شود.

این توزیع‌های تقریبی را به صورت زیرنشان می‌دهند:

$$T_n \approx N(n\mu, n\sigma^۲) \qquad , \qquad \bar{X}_n \approx N(\mu, \frac{\sigma^۲}{n}).$$

اگر Φ توزیع نرمال استاندارد باشد، به صورت حدی برای هر z داریم:

$$\lim_{n\to\infty} P(\frac{T_n - n\mu}{\sqrt{n\sigma}} \leq z) = \Phi(z),$$

$$\lim_{n\to\infty} P(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \le z) = \Phi(z).$$

اگر z دارای توزیع نرمال استاندارد باشد، آنگاه $\Phi(z) = P(Z \le z)$.

تعریف ۴۰.۱.۱. نمونه تصادفی $X_1, X_2, ..., X_n$ را با چگالی $f(x;\theta)$ در نظر بگیرید. فرض کنید مقدار واقعی پارامتر θ به مجموعه Θ ها برابر با حاصل‌ضرب X_i و چگالی توام $f(x_i;\theta)$ برابر با X_iها هم‌توزیع و مستقل هستند، بنابراین چگالی هر X_iبه نام مجموعه پارامتر، تعلق داشته باشد. چون چگالی‌های آنها می‌باشد. این تابع چگالی توام را با$L(\theta)$ نشان می‌دهند و آن را تابع راست‌نمایی برای نمونه تصادفی می‌نامند. به طور خلاصه داریم:

$$L(\theta) = \prod_{i=1}^{n} f(x;\theta).$$

تعریف ۴۱.۱.۱. مقداری از θ در فضای پارامتر که $L(\theta)$ را ماکزیمم می‌کند، برآورد راست‌نمایی ماکزیمم نامیده می‌شود و آن را به $\hat{\theta}$ نشان می‌دهیم، یعنی برای θ، $\hat{\theta} \in \Theta$ دارای برآورد راست‌نمایی ماکزیمم است، هر گاه داشته باشیم

$$L(\hat{\theta}) \ge L(\theta) \qquad \forall\theta \in A.$$

۱—۲ سیستم های پویا

بررسی پیشینه سیستم‌های پویا نشان می‌دهد که مطالعه این سیستم‌ها از اواسط سال‌های ۱۶۰۰ میلادی با کشف قوانین جاذبه و حرکت و معرفی معادلات دیفرانسیل توسط نیوتن[۲] و توجیه قوانین

[۲] Newton

کپلر[۳] در مورد حرکت سیارات بر پایه آنها شکل گرفته است. بدین ترتیب نیوتن قادر به حل مساله حرکت زمین به دور خورشید گردید. تلاش ریاضی‌دانان و و فیزیک‌دانان برای تعمیم مساله به سه جسم (خورشید، زمین، ماه) منجر به فهم این نکته شد که حل مساله سه جسم اساسا غیرممکن است، در اواخر ۱۸۰۰ میلادی هانری پوانکاره[۴] با دید جدیدی به مساله سه جسم نگریست. او به جای اینکه بخواهد مکان دقیق سیارات را در تمام زمان‌ها بدست آورد، مساله پایداری یا ناپایداری سیستم خورشیدی را مورد توجه قرار داده و امکان بروز آشوب را مطرح ساخت.
از آن‌جا که عنوان سیستم پویا به سیستم‌هایی داده می‌شود که در گذر زمان دستخوش تحول می‌شوند، لذا یک سیستم پویا را می‌توان توسط سه پارامتر زمان، حالت‌ها و قاعده‌هایی که بیانگر نحوه تحول این سیستم‌هاست، شکل داد. برای درک سیستم پویا بایستی بر شرایط اولیه حاکم بر سیستم احاطه داشت. اگر سیستم به طور گسسته با زمان تحول یابد سیستم در قالب نگاشت‌های تکرار مانند نگاشت لجستیک [۵] مورد مطالعه قرار می‌گیرد که یک نگاشت درجه دوم یک بعدی است.

تعریف ۱.۲.۱. سیستم‌هایی که در آنها یک رابطه خطی میان سرعت و موقعیت برقرار باشد، را سیستم‌های پویای خطی می‌نامند. اگر دو جواب برای سیستم خطی داشته باشیم، مجموع آنها نیز یک جواب برای سیستم است. هم‌چنین سیستم‌های خطی از این قابلیت برخوردارند که آنها را می‌توان با تجزیه مسأله، به اجزا کوچک‌تر، مورد بررسی قرار داده و سپس با جمع‌بندی نتایج، به تحلیل کلی آنها اقدام کرد. این از جمله مواردی است که تحلیل سیستم‌های خطی را آسان می‌سازد.

تعریف ۲.۲.۱. اگر رابطه میان سرعت و موقعیت غیرخطی باشد، سیستم پویای غیر خطی است. در

[۳] kepler
[۴] Hanry poincare
[۵] lgistic