پردازش سیگنال‌های بیولوژیکی

Dr. Maleki maleki.semnan.ac.ir

مبحث سوم

فرآیندهای تصادفی

فهرست مطالب

- مقدمه
- مفاهیم نظریه‌ی احتمال
- فرآیند تصادفی
- تحلیل همبستگی
- فرآیندهای گوسی

مقدمه

عوامل تصادفی بودن سیگنال زیست‌پزشکی:

تصادفی بودن خود منبع

ناشی از سیستم اندازه‌گیری (نویز ضرب‌شونده یا جمع‌شونده‌ی خارجی)

این که سیگنال معین یا تصادفی در نظر گرفته شود به تعریف برمی‌گردد. به عنوان مثال، وقتی ویژگی‌های QRS مورد نظر است می‌توان سیگنال ECG را معین یا حتی «شبه متناوب» در نظر گرفت در حالی که وقتی تغییرات فاصله‌ی R-R مورد توجه است می‌توان آن را تصادفی دانست.

واژه‌نامه:

معین	deterministic
تصادفی	stochastic

فهرست مطالب

- مقدمه
- مفاهیم نظریه‌ی احتمال
- سیگنال تصادفی
- تحلیل همبستگی
- فرآیندهای گوسی

اهمیت طرح نظریه‌ی احتمال

نظریه‌ی احتمال نقش پایه‌ای برای تحلیل سیگنال‌های تصادفی دارد.

(event and probability) رویداد و احتمال

آزمایشی را در نظر بگیرید که خروجی آن یکی از چندین **رویداد غیر قابل پیش‌بینی** باشد.

(discrete random events)

فرض کنید آزمایش را N بار انجام دهیم که طی آن رویداد A_i به تعداد n_i دفعه رخ داده باشد:

فرکانس (بسامد) نسبی:

Relative Frequency: $f_i = \frac{n_i}{N}$

احتمال رویداد A_i :

Probability of event A_i : $P(A_i) = \lim_{N\to\infty}(f_i) = \lim_{N\to\infty}\left(\frac{n_i}{N}\right)$

با فرض این که این حد وجود داشته باشد:

$$0 \le P(A_i) \le 1$$

رویدادهای ناسازگار (mutually exclusive events)

دو رویداد را **ناسازگار** گوییم اگر رخداد یکی از آنها، بروز رویداد دیگر را غیرممکن سازد.

اگر دو رویداد A_i و A_j ناسازگار باشند:

$$P(A_i\ or A_j) = P(A_i) + P(A_j)$$

در حالت کلی‌تر:

$$P(A_1\ \ or\ A_2\ or, \dots, A_M) = \sum_{i=1}^{M} P(A_i)$$

به عنوان مثال، رویدادهای A و $not\ A$ ناسازگارند بنابراین:

$$P(A) + P(NOT\ A) = 1$$

واژه‌نامه:

رویدادهای ناسازگار	Mutually exclusive events

احتمال توأم (joint probability)

$$P(A \text{ and } B \text{ and } C \text{ and}, \dots, \text{and } J) = P(A\, B\, C, \dots, J)$$

واژه‌نامه:

Joint probability	احتمال توأم

احتمال شرطی (conditional probability)

$$P(A|B) = probability\ (A\ occurs\ given\ B\ has\ occurred)$$

واژه‌نامه:

conditional probability	احتمال شرطی

رابطه‌ی احتمال توأم و احتمال شرطی (رابطه‌ی بِیز)

$$P(AB) = \lim_{N\to\infty}\left(\frac{n_{AB}}{N}\right) = \lim_{N\to\infty}\left(\frac{n_A}{N}.\frac{n_{AB}}{n_A}\right)$$

با این فرض که N آنقدر بزرگ باشد که بتوان n_A را نیز بسیار بزرگ در نظر گرفت:

$$P(AB) = \lim_{N\to\infty}\left(\frac{n_A}{N}\right).\lim_{n_A\to\infty}\left(\frac{n_{AB}}{n_A}\right) = P(A)\,.\,P(B|A)$$

$$P(B|A) = \frac{P(AB)}{P(A)} \quad , \qquad for\ P(A) \neq 0$$

و به طور مشابه،

$$P(A|B) = \frac{P(AB)}{P(B)} \quad , \qquad for\ P(B) \neq 0$$

$$P(AB) = P(A|B)\,.\,P(B) = P(B|A).P(A)$$

$$P(B|A) = \frac{P(B).P(A|B)}{P(A)} \quad , \qquad P(A|B) = \frac{P(A).P(B|A)}{P(B)}$$

که به عنوان قاعده‌ی بِیز (**Bayes' rule**) شناخته می‌شود.

رویدادهای مستقل آماری (statistically independent events)

دو رویداد A و B را **مستقل آماری** گویند اگر:

$$P(B|A) = P(B)$$

به عبارت دیگر، دانستن احتمال رخداد A ، اطلاعاتی در مورد رویداد B فراهم نمی‌سازد.

اگر دو رویداد مستقل آماری باشند:

$$P(AB) = P(A)\,.\,P(B)$$

در حالت کلی‌تر، برای n رویداد مستقل آماری $A_i\ ,i = 1,2,\ldots,n$:

$$P(A_1A_2\ldots A_n) = \prod_{i=1}^{n} P(A_i)$$

همچنین برای هر ترکیب دلخواه با $i < j < k \ldots \leq n$ داریم:

$$P(A_iA_j) = P(A_i).P(A_j)$$

$$P(A_iA_jA_k) = P(A_i).P(A_j).P(A_k)$$

$$\vdots$$

$$P(A_1A_2\ldots A_n) = P(A_1).P(A_2).\ \ldots\ .P(A_n)$$

واژه‌نامه:

رویدادهای مستقل آماری	Statistically independent events

متغیرهای تصادفی (random variables)

متغیر تصادفی پیوسته

Continuous random variable

متغیر تصادفی : x

مقادیر متغیر تصادفی : X ; $-\infty < X < +\infty$

متغیر تصادفی گسسته

Discrete random variable

متغیر تصادفی : x

مقادیر متغیر تصادفی : $x_1 , x_2 , \dots , x_n$

متغیرهای تصادفی گسسته (discrete random variable)

x : متغیر تصادفی

$x_1 , x_2 , \dots , x_n$: مقادیر متغیر تصادفی

$P_x(x = x_i)$: احتمال اینکه متغیر تصادفی x مقدار x_i داشته باشد

در صورتی که n خروجی آزمایش ناسازگار باشند:

$$\sum_{i=1}^{n} P_x(x = x_i) = 1$$

احتمال توام (joint probability)

احتمال توام : $P_{xy}(x = x_i\ , y = y_j)$

$$\sum_i \sum_j P_{xy}(x = x_i\ , y = y_j) = 1$$

احتمال شرطی (conditional probability)

احتمال شرطی : $P_x(x = x_i \mid y = y_j)$

$$\sum_i P_x\left(x = x_i \middle| y = y_j\right) = 1$$

$$\sum_j P_y\left(y = y_j \middle| x = x_i\right) = 1$$

قاعده‌ی بِیز (Bayes' rule)

$$P_x\left(x = x_i \,\middle|\, y = y_j\right) = \frac{P_{xy}\left(x = x_i, y = y_j\right)}{P\left(y = y_j\right)}$$

$$P_y\left(y = y_j \middle| x = x_i\right) = \frac{P_{xy}\left(x = x_i, y = y_j\right)}{P(x = x_i)}$$

$$P_x\left(x = x_i \,\middle|\, y = y_j\right) = \frac{P_y\left(y = y_j \middle| x = x_i\right) P_x(x = x_i)}{P_y\left(y = y_j\right)}$$

متغیرهای تصادفی پیوسته (continuous random variable)

x : متغیر تصادفی

$X; \quad -\infty < X < +\infty$: مقادیر متغیر تصادفی

در اینجا لازم است مفاهیم **تابع توزیع احتمال** و **تابع چگالی احتمال** معرفی شوند.

تابع توزیع احتمال (probability distribution function)

تابع توزیع احتمال : $P(x \leq X)$

احتمال اینکه متغیر تصادفی x مقداری کوچکتر مساوی مقدار ثابت X داشته باشد.

تابع توزیع احتمال را تابع توزیع تجمعی نیز می‌نامند.

$$P(x \geq \infty) = P(x \leq -\infty) = 0$$

$$P(x \leq \infty) = 1$$

$$P(x = X_0) = 0$$

$$P(X_2 \leq x \leq X_1) = P(x \leq X_1) - P(x \leq X_2)$$

$$if\ X_1 > X_2 \qquad P(x \leq X_1) \geq P(x \leq X_2)$$

بنابراین تابع توزیع احتمال، تابعی غیرمنفی و غیرکاهشی است که مقداری در محدوده‌ی [0,1] دارد.

تابع توزیع احتمال توام (joint probability distribution func.)

با فرض دو متغیر تصادفی x و y ، $(-\infty \leq X \leq +\infty \;\; , -\infty \leq Y \leq +\infty)$ ،

تابع توزیع احتمال توام : $P(x \leq X \;\; , y \leq Y)$

بر این اساس داریم:

$$P(x \leq X\,, y \leq \infty) = P(x \leq X)$$

$$P(x \leq \infty\,, y \leq Y) = P(y \leq Y)$$

$$P(x \leq -\infty\,, y \leq Y) = P(x \leq X\,, y \leq -\infty) = 0$$

$$P(x \leq \infty\,, y \leq \infty) = 1$$

تابع چگالی احتمال (probability density function)

x : متغیر تصادفی

X : مقدار متغیر تصادفی

ΔX : همسایگی بسیار کوچکی حول X

$p(x) = \lim_{\Delta X \to 0} \frac{P(X - \Delta X \le x \le X)}{\Delta X}$: تابع چگالی احتمال

$$\Rightarrow p(x) = \lim_{\Delta X \to 0} \frac{P(x \le X) - P(x \le X - \Delta X)}{\Delta X}$$

$$\Rightarrow p(X) = \frac{d}{dx} P(x \le X)$$

$$\Rightarrow P(x \le X) = \int_{-\infty}^{X} p(x)dx$$

همچنین داریم:

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

$$P(X_1 \le x \le X_2) = \int_{X_1}^{X_2} p(x)dx$$

$$P(x = X_0) = 0$$

تابع توزیع احتمال و تابع چگالی احتمال

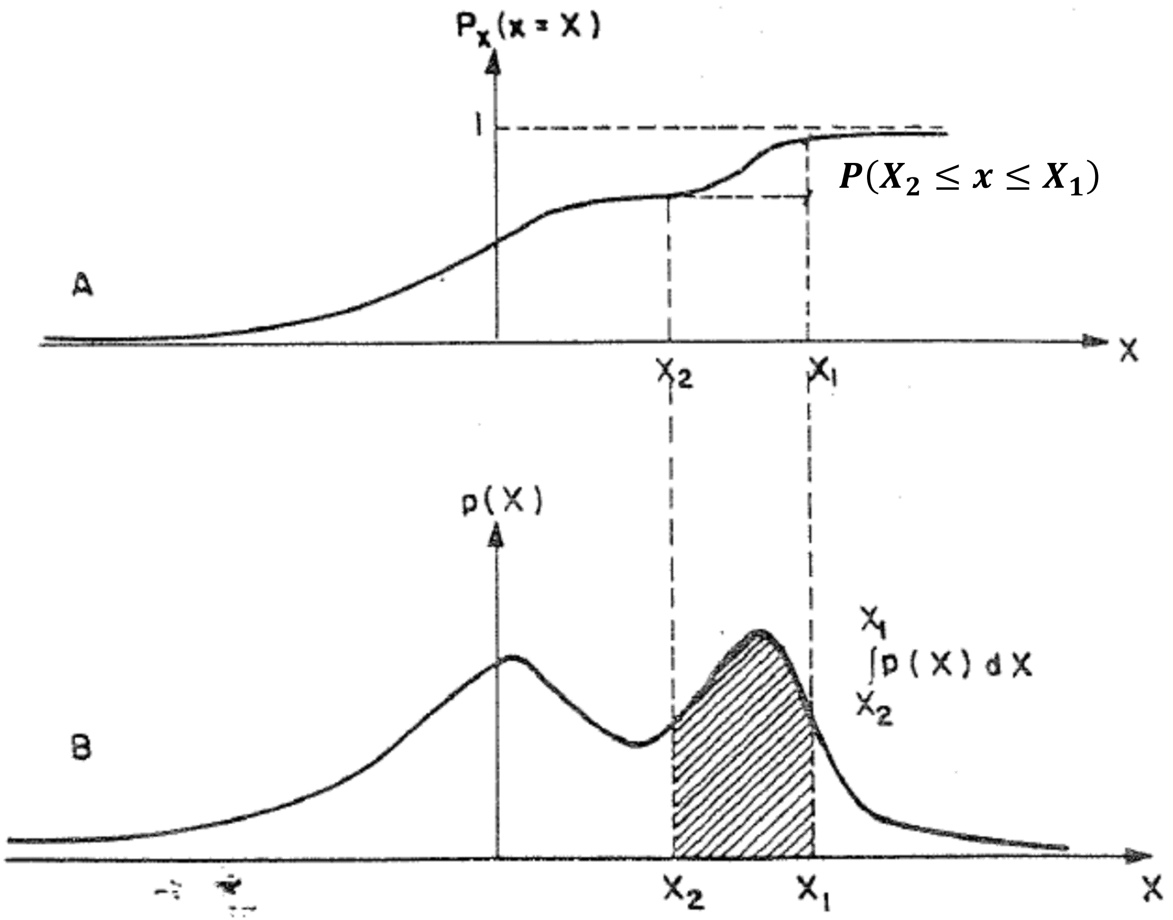


FIGURE 1. Probability functions. (A) Probability distribution function; (B) probability density function (PDF).

تابع چگالی احتمال توام (joint probability density func.)

با فرض وجود مشتق جزئی:

$$p(X,Y) = \frac{\partial^2}{\partial x \partial y} P(x \le X \, , y \le Y)$$

$$P(x \le X \, , y \le Y) = \int_{-\infty}^{Y} \int_{-\infty}^{X} p(x,y) dxdy$$

تابع توزیع احتمال شرطی و تابع چگالی احتمال شرطی

احتمال اینکه متغیر تصادفی y کوچکتر مساوی مقدار Y باشد با فرض اینکه متغیر تصادفی x در محدوده‌ی $X - \Delta X \leq x \leq X$ باشد:

$$P(y \leq Y | x = X) \triangleq \frac{\int_{-\infty}^{Y} p(X, y) dy}{P(X)}$$

فهرست مطالب

- مقدمه
- مفاهیم نظریه‌ی احتمال
- فرآیند تصادفی
- تحلیل همبستگی
- فرآیندهای گوسی

عناوین این بخش:

- معرفی اصطلاحات و مفاهیم فرآیندهای تصادفی
- فرآیندهای تصادفی مستقل آماری
- فرآیند ایستان
- میانگین‌گیری آماری (امیدها)

پیشگفتار

در قیاس با آنچه تا کنون مطرح شده است، در اینجا به جای متغیرهای تصادفی حاصل از آزمایش‌ها، می‌خواهیم به توابعی زمانی بپردازیم. توابع زمانی از آن جهت مورد توجه قرار گرفته‌اند که در اغلب موارد اینچنین است؛ هرچند در حالت کلی می‌تواند توابع مکانی و ... نیز مطرح گردد.

معرفی اصطلاحات

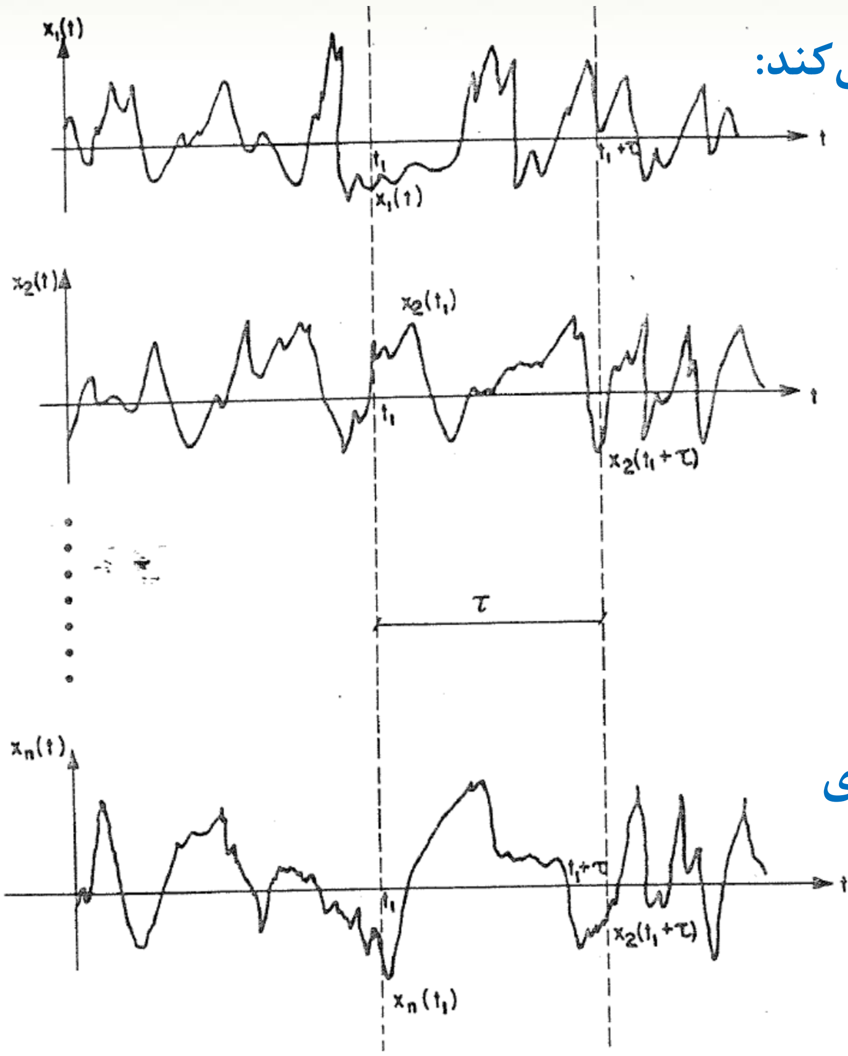


FIGURE 2. The ensamble of the random process x(t).

منبعی را در نظر بگیرید که تابع تصادفی $x(t)$ را ایجاد می‌کند:

تابع نمونه (sample function): هر یک از توابع $x_1(t)$، $x_2(t)$ و ...

ensamble : مجموعه‌ای شامل تمام توابع نمونه.

فرآیند تصادفی (random process): فرآیند تصادفی x که ensamble را ایجاد نموده است.

متغیر تصادفی (random variable): مقادیر توابع نمونه‌ی یک ensamble در یک زمان خاص نظیر t_1.

مثالی برای آشنایی با اصطلاحات

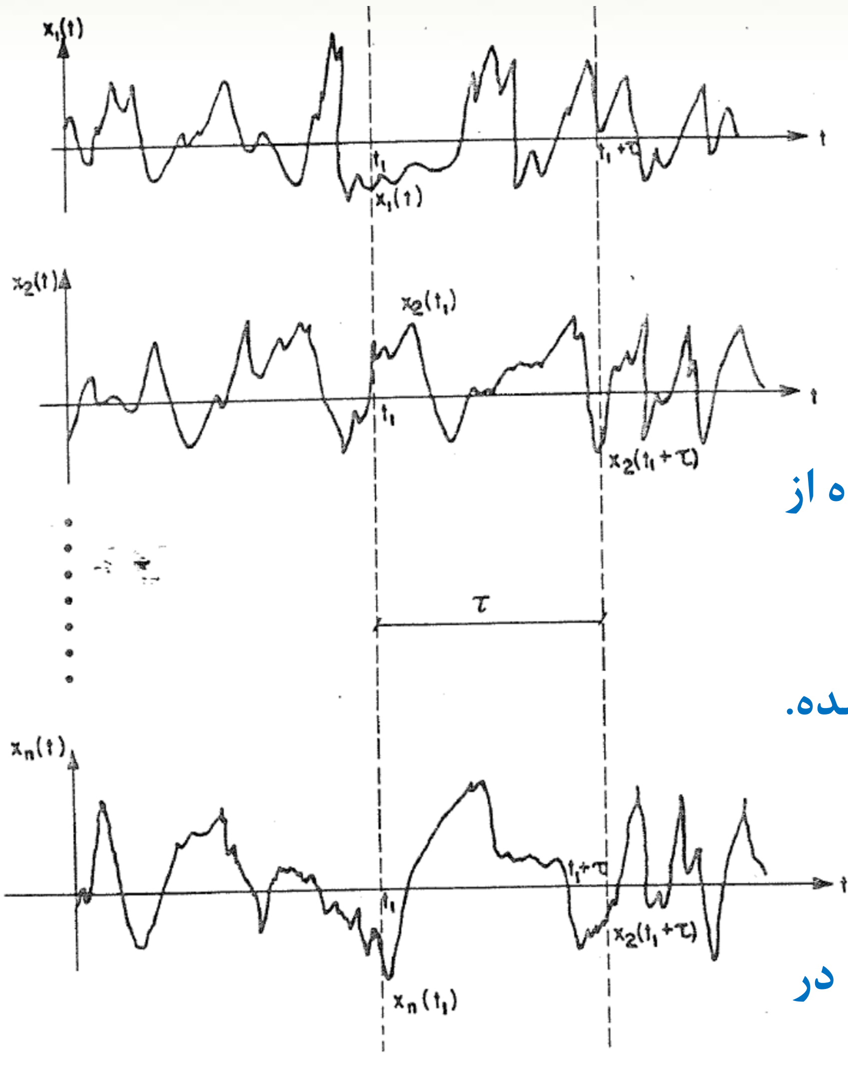


FIGURE 2. The ensamble of the random process x(t).

فرض کنید می‌خواهیم سیگنال EEG نقطه‌ی خاصی از سطح جمجمه را در مجموعه‌ای مشخص از افراد (مثلا ۳۰ کودک سالم ۱۰ ساله) بررسی کنیم. می‌توان این را یک فرآیند تصادفی در نظر گرفت که:

تابع نمونه (**sample function**): سیگنال EEG ثبت شده از هر فرد.

ensamble : مجموعه‌ی تمامی سیگنال‌های EEG ثبت شده.

فرآیند تصادفی (**random process**): ....

متغیر تصادفی (**random variable**): مقدار n تابع نمونه در زمان t_i.

تابع چگالی احتمال و تابع توزیع احتمال فرآیند تصادفی

فرآیند تصادفی x با ensamble ی شامل n تابع نمونه $x_i(t); i = 1,2, \ldots, n$ را در نظر بگیرید. برای مقادیر بزرگ n،

احتمال : $P(x(t) \leq X)$

احتمال توام : $P(x(t_1) \leq X_1 , x(t_2) \leq X_2 , \cdots , x(t_N) \leq X_N)$
$\triangleq P(x_1 , x_2 , \cdots , x_N)$

تابع چگالی احتمال توام فرآیند تصادفی : $p(x_1, x_2, \cdots , x_N)$

و بر این اساس،

$$p(x_1, x_2, \ldots , x_N) \geq 0$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, x_2, \ldots , x_N) dx_1, \ldots , dx_N = 1$$

$$p\left(x_{j+1}, x_{j+2}, \ldots , x_N \middle| x_1, x_2, \ldots , x_j\right) = \frac{p(x_1, x_2, \ldots , x_N)}{p(x_1, x_2, \ldots , x_j)}$$

فرآیند تصادفی مستقل آماری

دو فرآیند تصادفی $x(t)$ و $y(t)$ را مستقل آماری گویند اگر،

$$p(x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_M) = p(x_1, x_2, \dots, x_N)\, p(y_1, y_2, \dots, y_M)$$

فرآیند تصادفی ایستان (stationary)

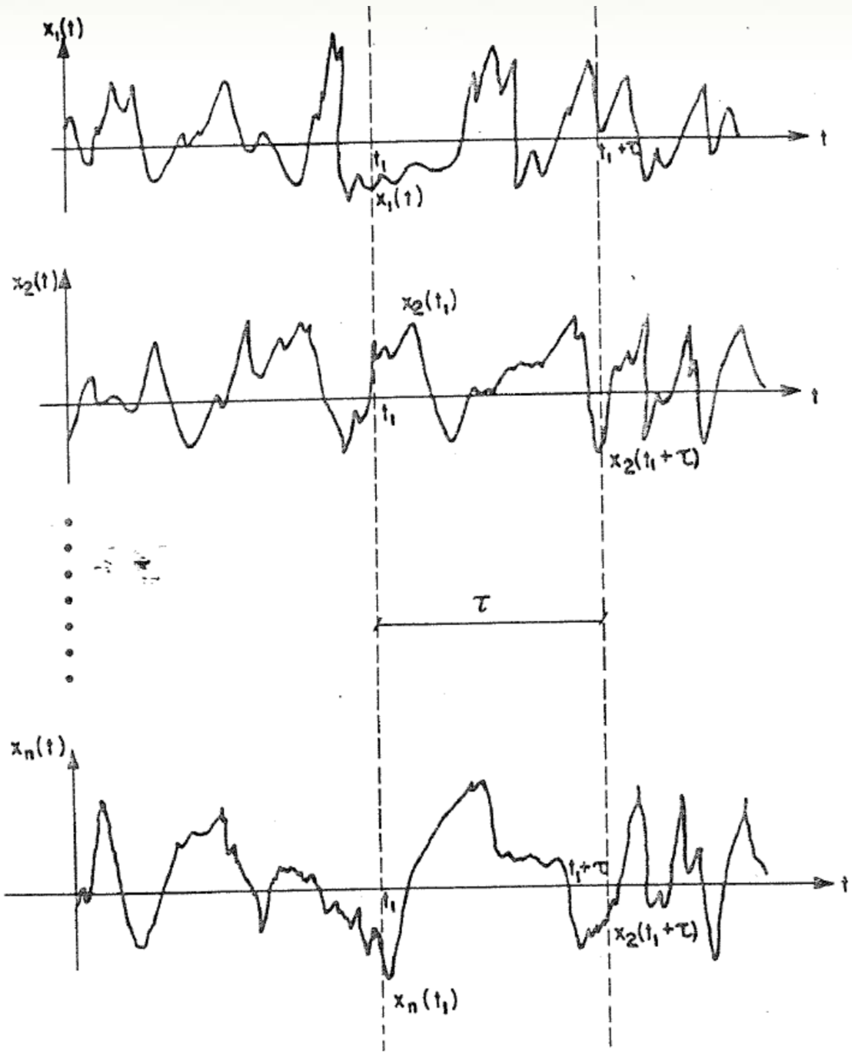


FIGURE 2. The ensamble of the random process x(t).

اگر تابع چگالی احتمال فرآیند $x(t)$ (که در زمان‌های مختلف نظیر t_1 و $t_1+\tau$ در شکل روبرو، تخمین زده می‌شود) در تمامی زمان‌ها و به ازای تمام N یکسان باشد فرآیند را ایستان (**stationary**) گویند.

میانگین‌های آماری (امیدها)

دو متغیر تصادفی x و y با تابع چگالی احتمال توام $p(x,y)$ را فرض کنید. همچنین متغیر تصادفی جدید $z = f(x,y)$ که تابعی تک مقداری از این دو متغیر است را در نظر بگیرید.

$$E\{z\} = E\{f(x,y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)p(x,y)\,dxdy$$: تعریف امید (**expectation**)

به امید، نام‌های متوسط آماری (**statistical average**) و میانگین (**mean**) هم اطلاق می‌گردد.

چنانچه x و y گسسته باشند به نحوی که:

مقادیر x : $= x_i \,; i = 1, 2, \cdots, I$

مقادیر y : $= y_j \,; j = 1, 2, \cdots, J$

$$E\{z\} = E\{f(x,y)\} = \sum_{i=1}^{I}\sum_{j=1}^{J} f(x,y)p(x,y)$$: تعریف امید (**expectation**)

به سادگی می‌توان نشان داد که امید تابعی خطی است!

ممان (moment)

تابع $z = x^n$ را در نظر بگیرید:

(nth moment of x) x ام n ممان : $E\{x^n\} = \int_{-\infty}^{\infty} x^n p(x)dx$

ممان اول : $E\{x\} = \int_{-\infty}^{\infty} x\, p(x)dx = m$ میانگین (mean)

n ام ممان مرکزی : $\mu_n = E\{(x-m)^n\} = \int_{-\infty}^{\infty} (x-m)^n p(x)dx$

(nth central moment)

ممان مرکزی دوم : $\mu_2 = E\{(x-m)^2\} = \int_{-\infty}^{\infty} (x-m)^2 p(x)dx = {\sigma_x}^2$

واریانس (variance) که با ${\sigma_x}^2$ نیز نشان داده می‌شود.

σ_x نشانگر انحراف معیار (standard deviation) است.

ممان توام مرتبه $(n+m)$ام : $E\{x^n y^m\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^m p(x,y)dxdy$

ممان مرکزی توام مرتبه $(n+m)$ام : $\mu_{nm} = E\{(x-m_x)^n (y-m_y)^m\}$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-m_x)^n (y-m_y)^m p(x,y)dxdy$$

ممان مرکزی توام μ_{11} از اهمیت ویژه‌ای برخوردار بوده و کوواریانس (covariance) نامیده می‌شود.

$$\mu_{11} = E\{ (x - m_x) \left(y - m_y\right)\}$$

مثال:

$E\{x^2\}$ و $E\{xy\}$ را بر حسب میانگین، واریانس و کوواریانس توصیف نمایید.

$$\sigma_x{}^2 = E\{(x-m_x)^2\}$$
$$= E\{x^2 - 2xm_x + m_x{}^2\}$$
$$= E\{x^2\} - 2m_x E\{x\} + m_x{}^2$$
$$= E\{x^2\} - 2m_x\, m_x + m_x{}^2$$
$$= E\{x^2\} - m_x{}^2$$

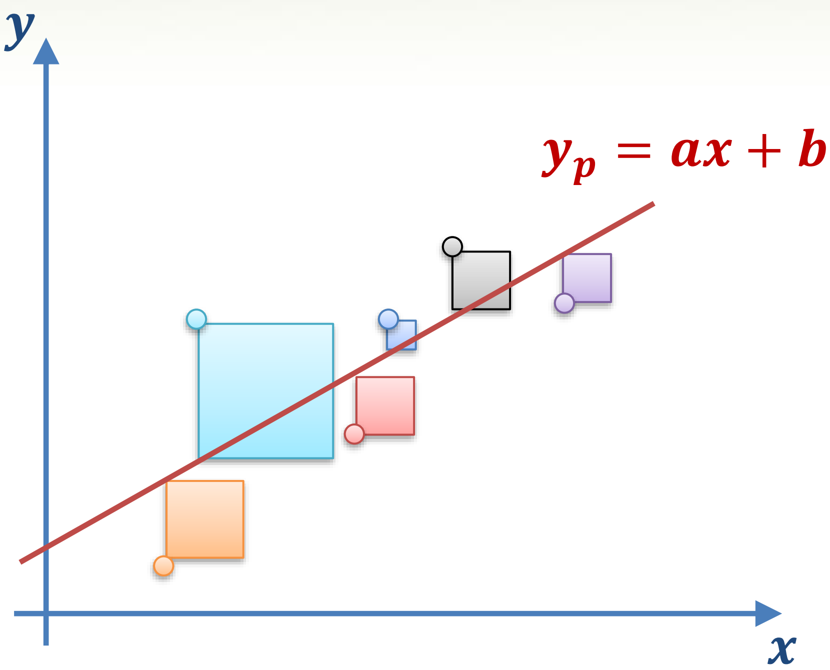
$$\Rightarrow \boxed{E\{x^2\} = \sigma_x{}^2 + m_x{}^2}$$

$$\mu_{11} = E\{(x-m_x)(y-m_y)\}$$
$$= E\{xy - xm_y - ym_x + m_x m_y\}$$
$$= E\{xy\} - m_x m_y - m_x m_y + m_x m_y$$
$$= E\{xy\} - m_x m_y$$

$$\Rightarrow \boxed{E\{xy\} = \mu_{11} + m_x m_y}$$

فهرست مطالب

خط رگرسیون (regression line)



دو متغیر تصادفی x و y را در نظر بگیرید. فرض کنید این دو متغیر ارتباطی خطی داشته باشند به نحوی که اگر نمونه‌های این دو متغیر تصادفی در صفحه‌ی $x - y$ ترسیم گردد نقاطی حوالی خط مستقیم زیر حاصل گردد:

$$y_p = ax + b$$

$\epsilon = E\{(y - y_p)^2\} = E\{(y - ax - b)^2\}$: میانگین مربعات خطا بین نقاط تصادفی و این خط

برای این نقاط، بهترین خطی که ϵ را کمینه می‌نماید **خط رگرسیون (regression line)** نامیده می‌شود.

تعیین پارامترهای a و b خط رگرسیون

یادآوری : $\epsilon = E\{(y-y_p)^2\} = E\{(y-ax-b)^2\}$

$$\frac{d\epsilon}{db} = 0 \Rightarrow -2E\{y-ax-b\} = 0$$

$$E\{y-ax-b\} = 0$$

$$E\{y\} - aE\{x\} - b = 0$$

$$m_y - am_x - b = 0 \Rightarrow \boxed{b = m_y - am_x}$$

$$\frac{d\epsilon}{da} = 0$$

$$-2E\{(y-ax-b)x\} = 0$$

$$E\{(y-ax-b)x\} = 0$$

$$E\{xy\} - a\,E\{x^2\} - bE\{x\} = 0$$

$$(\mu_{11} + m_x m_y) - a(\sigma_x^2 + m_x^2) - (m_y - am_x)m_x = 0$$

$$\mu_{11} + m_x m_y - a\,\sigma_x^2 - a\,m_x^2 - m_x m_y + a\,m_x^2 = 0$$

$$\mu_{11} - a\,\sigma_x^2 = 0 \Rightarrow \boxed{a = \frac{\mu_{11}}{\sigma_x^2}} \Rightarrow \boxed{b = m_y - \frac{\mu_{11}}{\sigma_x^2} m_x}$$

خط رگرسیون : $y_p = ax + b = \frac{\mu_{11}}{\sigma_x^2} x + m_y - \frac{\mu_{11}}{\sigma_x^2} m_x = \frac{\mu_{11}}{\sigma_x^2}(x - m_x) + m_y$

توجه کنید که خط رگرسیون از نقطه‌ی (m_x, m_y) می‌گذرد.

تعیین ضریب همبستگی (correlation coefficient)

با نرمالیزه کردن متغیرهای تصادفی داریم:

$$\xi = \frac{x - m_x}{\sigma_x} \quad ; \quad \Phi = \frac{y - m_y}{\sigma_y} \quad ; \Phi_p = \frac{y_p - m_y}{\sigma_y};$$

بازنویسی رابطه‌ی خط رگرسیون با متغیرهای جدید:

یادآوری : $y_p = \frac{\mu_{11}}{{\sigma_x}^2}(x - m_x) + m_y$

$$\Rightarrow y_p - m_y = \frac{\mu_{11}}{{\sigma_x}^2}(x - m_x) \quad \Rightarrow \frac{y_p - m_y}{\sigma_y} = \frac{\mu_{11}}{\sigma_x\,\sigma_y}\,\frac{x - m_x}{\sigma_x}$$

$$\Rightarrow \Phi_p = \frac{\mu_{11}}{\sigma_x\,\sigma_y}\xi \quad \Rightarrow \Phi_p = \rho\xi$$

$$\rho = E\{\xi\Phi\} = \frac{\mu_{11}}{\sigma_x\,\sigma_y}$$

ضریب همبستگی (correlation coefficient) یا

کوواریانس نرمالیزه شده (normalized covariance)

ρ نشانگر شیب خط رگرسیون متغیرهای نرمالیزه شده است.

تعیین محدوده‌ی مقدار ضریب همبستگی

با توجه به مربع بودن عبارت $E\{(\xi-\Phi)^2\}\geq 0$: میانگین مربعات تفاضل

$$\Rightarrow E\{\xi^2\}-2E\{\xi\Phi\}+E\{\Phi^2\}\geq 0$$

$$\Rightarrow 1-2\rho+1\geq 0 \Rightarrow \rho\leq 1$$

با توجه به مربع بودن عبارت $E\{(\xi+\Phi)^2\}\geq 0$: میانگین مربعات مجموع

$$\Rightarrow E\{\xi^2\}+2E\{\xi\Phi\}+E\{\Phi^2\}\geq 0$$

$$\Rightarrow 1+2\rho+1\geq 0 \Rightarrow \rho\geq -1$$

$$\Rightarrow -1\leq\rho\leq 1$$

متغیرهای ناهمبسته (uncorrelated variables)

متغیرهای x و y را ناهمبسته (uncorrelated) یا مستقل خطی (linearly independent) گویند اگر ضریب همبستگی آنها صفر باشد ($\rho = 0$).

$$\rho = 0 \Rightarrow E\{\xi\Phi\} = 0 \Rightarrow E\left\{\frac{x-m_x}{\sigma_x} \cdot \frac{y-m_y}{\sigma_y}\right\} = 0$$

$$\Rightarrow E\{(x-m_x).(y-m_y)\} = 0$$

$$\Rightarrow E\{x\,y\} - E\{x\,m_y\} - E\{y\,m_x\} + m_x m_y = 0$$

$$\Rightarrow E\{x\,y\} - m_x m_y - m_x m_y + m_x m_y = 0$$

$$\Rightarrow E\{x\,y\} - m_x m_y = 0 \Rightarrow E\{x\,y\} - E\{x\}\,E\{y\} = 0$$

$$\Rightarrow E\{x\,y\} = E\{x\}\,E\{y\}$$

بنابراین متغیرهای تصادفی مستقل آماری، ناهمبسته هم می‌باشند.

تابع همبستگی (correlation function)

متغیرهای تصادفی $x(t)$ و $y(t)$ را در نظر بگیرید. با فرض $x(t_1) = x_1$ و $y(t_2) = y_2$:

تابع همبستگی متقابل : $E\{x_1 . y_2\}$ (تابعی از t_1 و t_2 است)

(cross correlation function)

تابع خود همبستگی : $E\{x_1 . x_2\}$

(auto-correlation function)

تابع همبستگی : $r_{xy}(t_1, t_2) = E\{x(t_1). y(t_2)\}$

(correlation function)

تابع همبستگی نرمالیزه شده :

(normalized correlation function)

$$\rho_{xy}(t_1, t_2) = E\left\{\frac{x_1 - m_{x1}}{\sigma(x_1)} . \frac{y_2 - m_{y2}}{\sigma(y_2)}\right\}$$

$$= \frac{E\{(x_1 - m_{x1})(y_2 - m_{y2})\}}{\sigma(x_1)\sigma(y_2)}$$

$$= \frac{r_{xy}(t_1, t_2) - m_{x1}m_{y2}}{\sigma(x_1)\sigma(y_2)}$$

ماتریس همبستگی (correlation matrix)

همبستگی دو فرآیند با خود همبستگی هر کدام از آنها و همبستگی متقابل بین آنها توصیف می‌شود:

ماتریس همبستگی :
(correlation matrix)

$$R_{xy}(t_1\ t_2) = \begin{bmatrix} r_x(t_1\ t_2) & r_{xy}(t_1\ t_2) \\ r_{yx}(t_1\ t_2) & r_y(t_1\ t_2) \end{bmatrix}$$

کواریانس (covariance)

کواریانس دو فرآیند با انحراف استاندارد هر کدام از آنها و کواریانس توام آنها توصیف می‌شود:

کواریانس توام دو فرآیند : $\mu_{xy} = E\{(x(t_1) - m_{x1})(y(t_2) - m_{y2})\}$

(joint covariance)

ماتریس کواریانس : $W_{xy}(t_1\ t_2) = \begin{bmatrix} \mu_x(t_1, t_2) & \mu_{xy}(t_1, t_2) \\ \mu_{yx}(t_1, t_2) & \mu_y(t_1, t_2) \end{bmatrix}$

(covariance matrix)

همبستگی فرآیندهای ایستان

اگر فرآیندهای تصادفی **ایستان** باشند امیدها تابعی از زمان‌های واقعی t_1 و t_2 نخواهد بود بلکه تنها تابعی از تفاوت زمانی $\tau = t_1 - t_2$ می‌باشند.

بنابراین برای فرآیندهای **ایستان** :

همبستگی : $$r_{xy}(t_1, t_2) = r_{xy}(t\,, t-\tau) = r_{xy}(\tau) = E\{x(t)y(t-\tau)\}$$

همبستگی نرمالیزه : $$\rho_{xy}(\tau) = \frac{r_{xy}(\tau) - m_x m_y}{\sigma_x \sigma_y}$$

$$r_{xy}(t_1\, t_2) = r_{xy}(\tau) = r_{yx}(-\tau)$$

$$r_x(\tau\,) = \; r_x(\,-\tau)$$

$$|r_{xy}(\tau)| \le \sqrt{r_x(0\,).\, r_y(0\,)}$$

$$|r_x(\tau)| \le r_x(0\,)$$

فرآیندهای تصادفی ایستان

Wide Sense Stationary (WSS)

اگر همبستگی متقابل تنها تابعی از اختلاف زمانی باشد.

$$r_{xy}(t_1, t_2) = r_{xy}(t, t-\tau) = r_{xy}(\tau)$$

Strict Sense Stationary (SSS)

اگر تابع چگالی احتمال توام فرآیندها، مستقل از زمان باشند.

مفهوم تابع خودهمبستگی

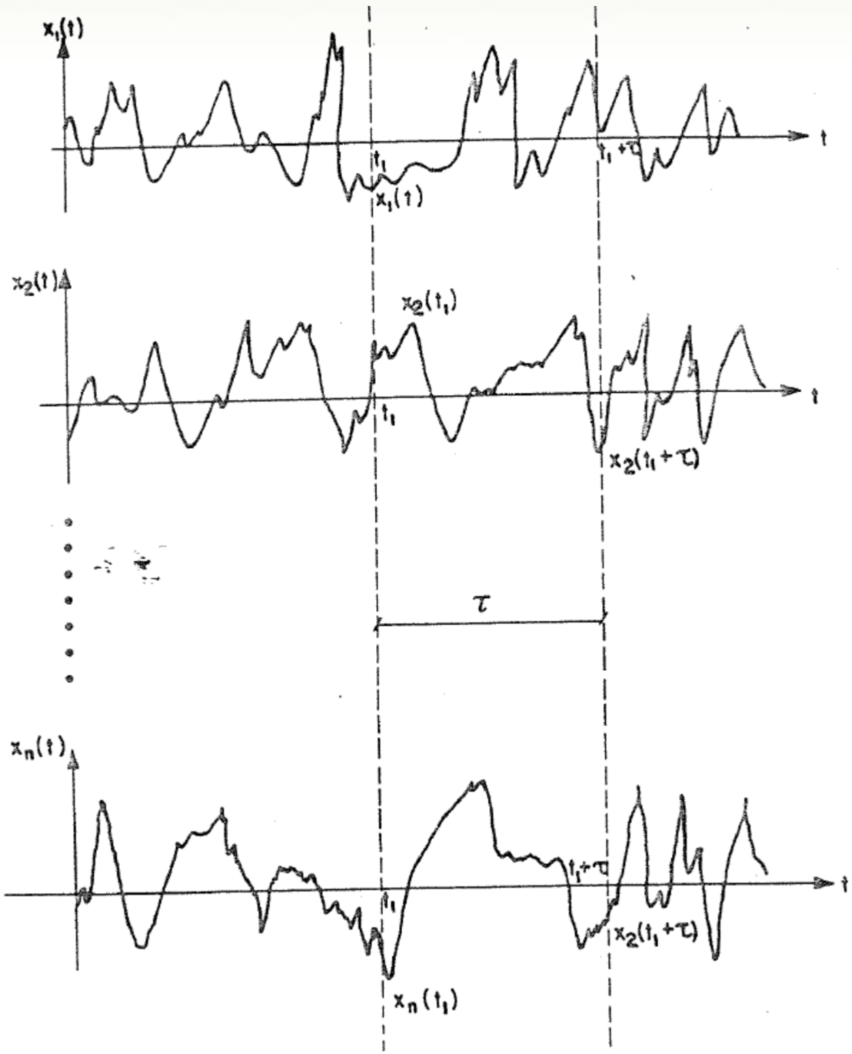


FIGURE 2. The ensamble of the random process x(t).

فرآیند تصادفی شکل روبرو را در نظر بگیرید:

برای هر یک از تابع‌های نمونه‌ی فرآیند تصادفی، مقدار در زمان‌های t_1 و $t_1 + \tau$ را در هم ضرب می‌کنیم. این حاصل‌ضرب برای تمام تابع‌های نمونه محاسبه می‌گردد.

با میانگین‌گیری مقادیر حاصل‌ضرب روی کل ensamble ، مقدار خودهمبستگی در زمان‌های t_1 و $t_1 + \tau$ حاصل می‌گردد.

چنانچه فرآیند ایستان باشد (WSS کفایت می‌کند) نتیجه مستقل از t_1 خواهد بود.

WSS: Wide Sense Stationary

میانگین زمانی (time averaging)

در محاسبه‌ی **میانگین زمانی**، برخلاف میانگین، میانگین‌گیری روی تابع نمونه انجام می‌گردد نه کل ensamble.

تابع نمونه تصادفی $z(t) = f\big(x(t)\big)$ را در نظر بگیرید،

میانگین زمانی $z(t)$: $$\langle z(t)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} z(t)dt = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} f(x(t))dt$$

میانگین زمانی اول $x(t)$: $$\langle x(t)\rangle = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)dt$$

میانگین زمانی دو فرآیند تصادفی $x(t)$ و $y(t)$: $$\langle r_{xy}\rangle(t,\tau) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t+\tau)y(t)dt$$

همبستگی متقابل زمانی (time cross correlation)

برای فرآیندهای **ایستان**، این رابطه تنها تابعی از τ خواهد بود.

قضیه‌ی ارگودیسیتی (ergodicity theorem)

قضیه‌ی ارگودیسیتی:

در شرایط خاصی، میانگین‌های زمانی (time averages) یک فرآیند تصادفی ایستان با احتمال یک برابر امیدهای آماری (statistical expectation) آن است.

اثبات و همچنین پرداختن به شرایط این قضیه فراتر از مباحث این کتاب است.

توجه کنید که برای این که فرآیند ارگودیک باشد لازم است فرآیند تصادفی ایستان باشد.

اهمیت قضیه‌ی ارگودیسیتی: تحلیل‌های آماری نیازمند تخمین امیدهای مختلف است. در عمل دستیابی به تعداد کافی از توابع نمونه برای این منظور، کاری مشکل است. این در حالی است که چنانچه فرآیند ارگودیک باشد امید را می‌توان به استفاده از تنها یک تابع نمونه و با میانگین‌گیری زمانی تخمین زد.

فهرست مطالب

- مقدمه
- مفاهیم نظریه‌ی احتمال
- فرآیند تصادفی
- تحلیل همبستگی
- فرآیندهای گوسی

توزیع گوسی یا نرمال (Gaussian or Normal distribution)

در عمل، تابع چگالی احتمال فرآیندهای مورد بررسی در دسترس نیست از این رو لازم است با استفاده از توابع نمونه به تخمین آنها بپردازیم.

از مهمترین توزیع‌ها می‌توان به **توزیع گوسی یا نرمال** اشاره نمود.

متغیر تصادفی دارای توزیع گوسی یا نرمال است اگر دارای تابع چگالی احتمال و تابع توزیع احتمال زیر باشد:

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} exp\left(-\frac{(x-m_x)^2}{2{\sigma_x}^2}\right)$$

$$P(x \le X) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \int_{-\infty}^{X} exp\left(-\frac{(x-m_x)^2}{2{\sigma_x}^2}\right) dx$$

قابل توجه است که توزیع گوسی تنها با ممان‌های اول و دوم به طور کامل توصیف می‌گردد. این توزیع نرمال اغلب به صورت $N(m_x, \sigma_x)$ نشان داده می‌شود.

اغلب با نرمالیزه کردن متغیر تصادفی x ، به متغیر تصادفی جدیدی (ξ) با توزیع $N(0,1)$ دست می‌یابیم.

قضیه‌ی حد مرکزی (central limit theorem)

قضیه‌ی حد مرکزی:

برای نمونه‌های مستقل آماری، چنانچه تعداد نمونه‌ها به اندازه‌ی کافی افزایش یابد، تابع توزیع احتمال میانگین زمانی محدودی از نمونه‌ها، مستقل از توزیع احتمال متغیر تصادفی، به گوسی میل می‌کند.

به همین دلیل، در پردازش سیگنال‌های واقعی معمولا از توزیع گوسی استفاده می‌گردد.

از آنجا که مجموع متغیرهای تصادفی مستقل دارای توزیع گوسی هستند می‌توان انتظار داشت که به عنوان نمونه، EEG دارای توزیع تقریبا گوسی باشد؛ زیرا در برخی حالت‌ها می‌توان EEG را در مجموع میدان‌های تولید شده توسط فعالیت تعداد زیادی نورون که مستقل از یکدیگر عمل می‌کنند دانست.

فرآیند گوسی چندمتغیره (multivariate Gaussian process)

$$\underline{x}^T = [x_1, x_2, \dots, x_n]$$

$$\underline{m}^T{}_x = E\{\underline{x}^T\} = [E\{x_1\}, E\{x_2\}, \dots, E\{x_n\}]$$

$$W_x = E\{(\underline{x} - \underline{m}_x)(\underline{x} - \underline{m}_x)^T\}$$

$$p(x_1, x_2, \dots, x_n) = p(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|W|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(\underline{x} - \underline{m}_x)^T W^{-1}(\underline{x} - \underline{m}_x)\right]$$

$$p(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Omega|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\underline{x}^T\Omega^{-1}\underline{x}\right]$$

فهرست مطالب

- مقدمه
- مفاهیم نظریه‌ی احتمال
- فرآیند تصادفی
- تحلیل همبستگی
- فرآیندهای گوسی

همت بلند دار که مردان روزگار

از همت بلند به جایی رسیده اند