تئوری اطلاعات : جلسه دوم 99/7/5

مروری بر احتمالات

متغیر تصادفی : $X(\omega): \Omega \to \mathbb{R}$

داشتن فضای نمونه و تابع احتمال و تابع $X:\Omega\to\mathbb{R}$ $\longleftrightarrow$ $\boxed{F_X(x) = \mathbb{P}[X(\omega) \leq x] \quad \forall x\in\mathbb{R}}$

CDF

↓

همه نکات لازم درباره متغیر تصادفی X را بیان می کند.

امید ریاضی :

- متغیر گسسته : $\mathbb{E}_{P_X}[X] \triangleq \sum_{x\in\mathcal{X}} x\, P_X(x)$
- متغیر های پیوسته : $\mathbb{E}_{f_X}[X] \triangleq \int_{\mathbb{R}} x\, f_X(x)\,.dx$
- برای هر متغیر تصادفی : $\mathbb{E}_{F_X}[X] = -\int_{-\infty}^{0} F_X(x)\,dx + \int_{0}^{+\infty} (1-F_X(x))\,dx$

* تابعی از یک متغیر تصادفی :

$$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$$

$Y = g(X) \Rightarrow$ Y هم یک متغیر تصادفی

با فرض Y گسسته :

$$\mathbb{E}_{P_Y}[Y] = ? = \sum_y y\, P_Y(y)$$

$$= \mathbb{E}_{P_X}[g(X)] = \sum_x g(x)\, P_X(x)$$

به طور مشابه برای پیوسته :

$$\mathbb{E}_{f_Y}[Y] = \int_{\mathbb{R}} g(x)\, f_X(x)\, d$$

- واقعه $A\in\mathcal{F}$ ، و تابع نشانگر های واقعه A

Indicator function

$$1_A : \begin{cases} 1 & : \omega\in A \\ 0 & : o.w. \end{cases}$$

↳ R.V.

همیشه برقرار است

$$\mathbb{E}[1_A] = \sum_{\omega\in\Omega} 1_A(\omega)\cdot\mathbb{P}(\omega) = \sum_{\omega\in A} + \cancel{\sum_{\omega\in A^c}}^{\,0}$$

$$= \sum_{\omega\in A} \mathbb{P}(\omega) = \mathbb{P}[A]$$

با فرض گسسته بودن فضای نمونه

خواص امید ریاضی :

• متغیر متناهی یک

اگر در نقاط x_i جرم p_i قرار دهیم ← مرکز جرم جرم‌ها $= \mathbb{E}[X]$

• عملگر خطی :

$$\mathbb{E}[\alpha_1 X_1 + \cdots + \alpha_n X_n + \beta] = \sum_{i=1}^{n} \alpha_i \, \mathbb{E}[X_i] + \beta$$

• فرض کنید متغیر X را می‌خواهیم با یک عدد ثابت تقریب بزنیم :

$X \longrightarrow C$: عدد ثابت

خطای متوسط تقریب کمینه شود :

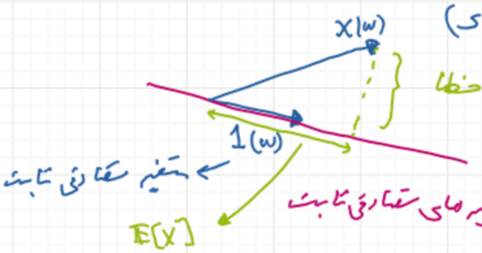
$$\min_{C} \mathbb{E}[f(X - C)]$$

f یک تابع درجه ۲ :

بهترین تقریب با عدد ثابت :
با معیار خطای میانگین مربعات

$$\arg\min_{C} \mathbb{E}[(X - C)^2] = \mathbb{E}[X]$$

• در حالت کلی تر می‌شود به متغیرهای تصادفی به شکل یک بردار نگاه کرد.
(یک فضای برداری هم‌بعدی)



خطا : $\sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]} = \sqrt{Var[X]}$

متغیر تصادفی ثابت :

$$1(\omega) : \Omega \longrightarrow \mathbb{R}$$
$$1(\omega) = 1 \quad \forall \omega \in \Omega$$

واریانس :

$$Var[X] = \mathbb{E}[(X - \mu_X)^2]$$

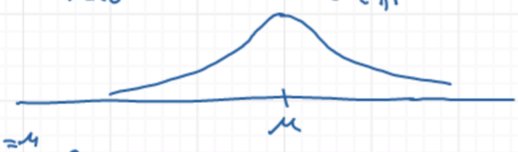
$$Var[\alpha X + \beta] = \alpha^2 Var[X]$$

* گشتاورهای یک متغیر تصادفی (Moments) :

گشتاور iام : $\mathbb{E}[X^i]$
$i \in \mathbb{N} \cup \{0\}$

* توزیع گاوسی تک متغیره:

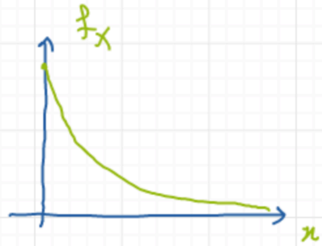
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x \in \mathbb{R}$$



$\mathbb{E}[X] = \mu$

$Var(X) = \sigma^2$

توزیع نمایی (پیوسته):

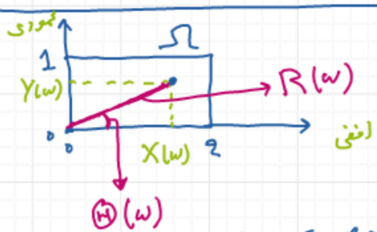


$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

توزیع نمایی بدون حافظه است.

---

* چند متغیر تصادفی همزمان:



$$\left\{ \begin{array}{l} X: \Omega \to \mathbb{R} \\ Y \text{ -- -- --} \\ \Theta \text{ -- -- --} \\ R \text{ -- -- --} \end{array} \right.$$

به صورت همزمان بر روی این آزمایش تصادفی تعریف شده اند.

راه‌های توصیف همزمان متغیرهای تصادفی:

تابع توزیع تجمعی مشترک: Joint Cumulative Dist. Func.

$$F_{X_1 \cdots X_n}(x_1, \ldots, x_n) = \mathbb{P}[X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n]$$

$\hookrightarrow \mathbb{R}^n \to [0, 1]$

* حاشیه‌سازی: Marginalization

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y) = \mathbb{P}[X \leq x, Y \leq \underset{\infty}{y}]$$

---

* اگر $X_1 \cdots X_n$ پیوسته باشند: $\leftarrow$ $f_{X_1 \cdots X_n}(x_1, \ldots, x_n)$

$\hookrightarrow$ تابع چگالی احتمال مشترک

$$f_{X_1 \cdots X_n}(x_1, \ldots, x_n) = \frac{\partial F_{X_1 \cdots X_n}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

• حاشیه‌سازی

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dy$$

---

* برای متغیرهای گسسته: تابع جرم احتمال

$$P_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \mathbb{P}[X_1 = x_1, \ldots, X_n = x_n]$$

* توزیع شرطی :

مثلا: Y گسسته ← $F_{X|Y}(x|y) \triangleq \mathbb{P}[X \le x \mid Y=y]$

مثال: X و Y هر دو پیوسته :

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

↳ $\forall y: f_Y(y) > 0$

---

* استقلال متغیرهای تصادفی :

متغیرهای $X_1 \cdots X_n$ را (مشترکا) مستقل گوییم اگر که :

$$F_{X_1\cdots X_n}(x_1,\dots,x_n) = F_{X_1}(x_1)\cdots F_{X_n}(x_n) \quad \forall x_1,\dots,x_n \in \mathbb{R}$$

$$\Longleftrightarrow$$

برای گسسته :

$$P_{X_1\cdots X_n}(x_1,\dots,x_n) = P_{X_1}(x_1)\cdots P_{X_n}(x_n)$$

پیوسته

$$f_{X_1\cdots X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n)$$

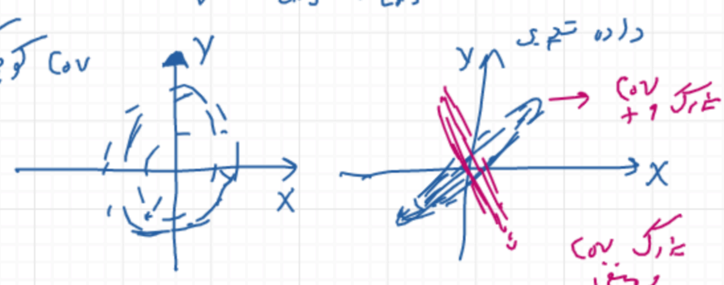
---

* کواریانس بین دو متغیر تصادفی :

$$\mathrm{Cov}[X,Y] = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]$$

$$-1 \le \mathrm{Cor}[X,Y] = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\cdot\mathrm{Var}[Y]}} \le 1$$

(X,Y) Cov کوچک



بردارهای تصادفی :

متغیر تصادفی → $X: \Omega \to \mathbb{R}$

بردار تصادفی → $\underline{X}: \Omega \to \mathbb{R}^k$

↳ $= (X_1,\dots,X_k) \longrightarrow \mathbb{E}[\underline{X}] = (\mathbb{E}[X_1],\dots,\mathbb{E}[X_k])$

Covariance Matrix :

$$\Sigma = \begin{bmatrix} \mathrm{Cov}(X_1,X_1) & \cdots & \mathrm{Cov}(X_1,X_k) \\ & & \\ \mathrm{Cov}(X_k,X_1) & \cdots & \mathrm{Cov}(X_k,X_k) \end{bmatrix} \in \mathbb{R}^{k\times k}$$

Whiteboard link: https://r2.whiteboardfox.com/266764-2288-4112