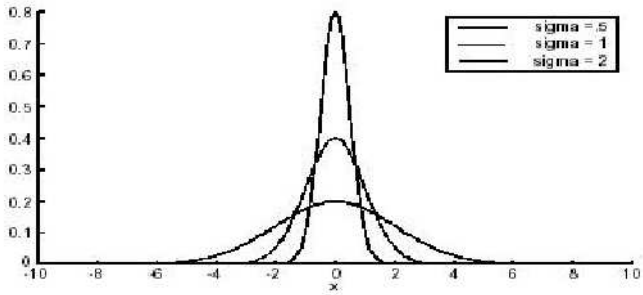
در زیر دو چگالی احتمال مهم را بررسی می کنیم:

- **چگالی احتمال یکنواخت:** یک متغییر تصادفی چگالی یکنواختی روی بازه (a, b) دارد اگر:

$$F_x(x)=\begin{cases}0, & x<a\\ (x-a)/(b-a), & a\le x\le b\\ 1, & x>b\end{cases}\quad f_x(x)=\begin{cases}1/(b-a), & a\le x\le b\\ 0, & \text{otherwise}\end{cases}\quad \sigma^2=\frac{1}{12}(b-a)^2$$

- **چگالی گووسی:** تابع چگالی گووسی یا نرمال به صورت زیر تعریف می شود (تصویر 1):

$$n(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$



تصویر 1 – تابع چگالی گووسی

تابع توزیع احتمال یک متغییر گووسی:

$$N(x;\mu,\sigma)=\int_{-\infty}^{x}n(u;\mu,\sigma)du$$

اگر تابع خطا به صورت روبرو تعریف شود:

$$erf(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-u^2/2}du$$

به فرمول زیر می رسیم:

$$N(x;\mu,\sigma)=\frac{1}{\sigma}erf(\frac{x-\mu}{\sigma})$$

4– دو متغییر تصادفی

اگر دو متغییر تصادفی X و Y با هم در نظر گرفته شوند، می توان تابع چگالی احتمال مشترک f(x,y) برای متغییرهای پیوسته یا p(x,y) برای متغییرهای گسسته.

دو متغییر تصادفی مستقلند اگر: $p(x,y)=p(x)p(y)$

با داشتن یک تابع g(x,y) امید ریاضی آن به صورت زیر تعریف می شود:

$$<g(x,y)>=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)dxdy$$

- پیوسته:
- گسسته: $< g(x,y) >= \sum_{x,y} g(x,y)p(x,y)$

ممان مشترک دو متغییر تصادفی x و y به صورت زیر محاسبه می شود: $m_{ij} = \sum_{x,y} x^i y^j p(x,y)$

ممان ها در عمل بوسیله میانگین گیری پشت سر هم آزمایش ها تخمین زده می شوند:

$$\hat{m}_{ij} = \frac{1}{N}\sum_{\delta=1}^{N} x_\delta^i y_\delta^j$$

ممان مرکزی دوم مشترک x و y کواریانس آن ها می باشد: $\sigma_{xy} =< (x-\overline{x})(y-\overline{y}) >= m_{11} - \overline{xy}$

- اگر x و y مستقل باشند کواریانس آن ها صفر است.
- ضرایب همبستگی x و y کواریانس آن ها است که به انحراف معیار نرمال شده باشد: $r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$

دو متغییر تصادفی x و y مشترکاً گووسی می باشند اگر تابع چگالی آن ها مانند روبرو باشد:

$$n(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}} \exp\left[-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2} - \frac{2rxy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}\right)\right]$$

که $r_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$

امید ریاضی جمع دو متغییر تصادفی: $< x+y >=< x > + < y >$

فرمول بالا هم برای متغییرهای مستقل و هم وابسته صادق است.

همچنین داریم: $\left\langle \sum_i x_i \right\rangle = \sum_i < x_i >$ و $< cx >= c < x >$

واریانس جمع دو متغییر تصادفی مستقل عبارت است از: $\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$

اگر دو متغییر تصادفی به هم وابسته باشند، چگالی احتمال جمع آن ها کانولوشن چگالی تک تک متغییرهاست.

- پیوسته: $f_{x+y}(z) = \int_{-\infty}^{\infty} f_x(u) f_y(z-u)du$
- گسسته: $p_{x+y}(z) = \sum_{u=-\infty}^{\infty} p_x(u) p_y(z-u)$

5– مباحث دیگر

- تئوری حد مرکزی (Central Limit Theorem)

تعریف غیررسمی: اگر تعداد زیادی متغییر تصادفی مستقل با هم جمع شوند، تابع چگالی احتمال جمع آن ها مستقل از چگالی های متغییرها به سمت یک چگالی گووسی میل می کند.

- تابع چگالی گووسی چندمتغییره

تابع چگالی نرمال را می توان به هر تعداد متغییر تصادفی عمومیت داد.

اگر X یک بردار تصادفی باشد، $N(x) = (2\pi)^{-n/2} \,|\, R \,|^{-1} \exp\left[-\frac{1}{2} Q(x-\overline{x}) \right]$

که $Q(x-\overline{x}) = (x-\overline{x})^T R^{-1} (x-\overline{x})$

ماتریس کوواریانس X با نام R: $R =< (x-\overline{x})(x-\overline{x})^T >$

- توابع تصادفی

یک تابع تصادفی تابعی است که به صورت خروجی یک آزمایش ناشی شود.

تابع تصادفی لزوماً توابعی از زمان نیستند، ولی در مطالعه ما معمولاً تابعی از زمانند.

یک فرآیند تصادفی گسسته بوسیله تعداد زیادی چگالی احتمال توصیف می شود.

$$p(x_1, x_2, x_3, ..., x_n, t_1, t_2, t_3, ..., t_n)$$

در صورتی که سیگنال های تصادفی از هم مستقل باشند،

$$p(x_1, x_2, ..., x_n, t_1, t_2, ..., t_n) = p(x_1, t_1) p(x_2, t_2) ... p(x_n, t_n)$$

اگر همه این چگالی های احتمال یکسان باشند، نتیجتاً دنباله ای از نمونه های مستقل و یکسان توزیع شده (i.i.d) خواهیم داشت.

- میانگین و خودهمبستگی

میانگین امید ریاضی x(t) می باشد: $\overline{x}(t) =< x(t) >= \sum_x x p(x,t)$

تابع خود همبستگی امید ریاضی x(t1)x(t2) می باشد: $r(t_1, t_2) =< x(t_1) x(t_2) >= \sum_{x_1, x_2} x_1 x_2 p(x_1, x_2, t_1, t_2)$

- میانگین زمانی و میانگین کلی (ensemble)

میانگین و خودهمبستگی را می توان به دو صورت مشخص کرد:

1. آزمایش می تواند به تعداد خیلی زیاد انجام شود و میانگین روی همه این توابع گرفته شود. به این میانگین، «میانگین ensemble» می گویند.
2. یکی از توابع را در نظر گرفته و آن را نمایانگر کل در نظر بگیریم. میانگین را از یک سری نمونه های این تابع محاسبه نماییم. به این مورد، «میانگین زمانی» می گویند.

- Ergodic بودن و ایستا بودن

اگر میانگین زمانی و کلی یک تابع تصادفی یکسان باشد، به آن ergodic می گویند.

یک تابع تصادفی ایستا است اگر شاخص آماری آن با تغییر زمان تغییر نکند.

همه توابع ergodic، ایستا هستند.

در یک سیگنال ایستا داریم: $\overline{x}(t) \equiv \overline{x}$ $\quad p(x_1, x_2, t_1, t_2) \equiv p(x_1, x_2, \tau)$ که $\tau = t_2 - t_1$

تابع خودهمبستگی عبارت است از: $r(\tau) = \sum_{x_1, x_2} x_1 x_2 p(x_1, x_2, \tau)$

یک x(t)ergodic است که میانگین و خودهمبستگی آن:

$$\overline{x} = \lim_{N\to\infty} \frac{1}{N} \sum_{t=-N}^{N} x(t) \qquad r(\tau) = < x(t)x(t-\tau) > = \lim_{N\to\infty} \frac{1}{N} \sum_{t=-N}^{N} x(t)x(t-\tau)$$

- همبستگی تقاطعی

همبستگی تقاطعی دو تابع تصادفی ergodicعبارت است از:

$$r_{xy}(\tau) = < x(t)y(t-\tau) > = \lim_{N\to\infty} \frac{1}{N} \sum_{t=-N}^{N} x(t)y(t-\tau)$$

اندیس xy نشان دهنده تقاطعی بودن است.

- چگالی توانی: تبدیل فوریه $r(\tau)$ را چگالی توانی طیف x(t) می گویند $S(\omega) = \sum_{\tau=-\infty}^{\infty} r(\tau) e^{-j\omega\tau}$
- چگالی طیفی تقاطعی دو تابع تصادفی عبارت است از:

$$S_{xy}(\omega) = \sum_{\tau=-\infty}^{\infty} r_{xy}(\tau) e^{-j\omega\tau}$$

برای سیگنال های ergodic می توان به صورت زیر نوشت: $r(\tau) = x(\tau) * x(-\tau)$

و بر اساس تبدیل فوریه:

$$\begin{aligned} S(\omega) &= X(\omega)X(-\omega) \\ &= X(\omega)X^*(\omega) \\ &= |X(\omega)|^2 \end{aligned}$$

در صورتی که همه مقادیر یک سیگنال تصادفی غیرهمبسته باشند (یعنی سیگنال نویز سفید باشد)

$$r(\tau) = \sigma^2 \delta(\tau)$$

طیف توانی نویز سفید مقدار ثابت است: $S(\omega)=\sigma^2$

نویز سفید مخلوطی از تمامی فرکانس ها می باشد.

اگر T[] عملیات خطی باشد، $<T[x(t)]>=T[<x(t)>]$

با داشتن یک پاسخ فرکانسی h(n): $<y(n)>=<x(n)*h(n)>=<x(n)>*h(n)$

یک سیگنال ایستا که به یک سیستم خطی اعمال می شود یک خر.جی ایستا می دهد.

$$r_{yy}(\tau)=r_{xx}(\tau)*h(\tau)*h(-\tau)$$

$$S_{yy}(\omega)=S_{xx}(\omega)\,|H(\omega)|^2$$

6 – خلاصه و نتیجه گیری

در این فصل مروری بر بحث احتمال انجام دادیم.

7 – منابع درس

1- Rabiner, “Fundumentals of Speech Recognition”
2- Huang, Acero, “Spoken Language Processing”
3- Deller, “Discrete-time processing of speech signals”

دانشگاه امام رضا (علیه السلام)