1– مقدمه

اهداف درس:

آشنایی با مفاهیم پایه ای تئوری احتمال

خیلی از تکنیک های پردازش گفتار نیاز به کار با تئوری احتمال و آمار دارد.

دو دو کاربرد اصلی که برخورد خواهیم کرد عبارتند از:

- بازشناسی الگوری آماری
- مدل کردن سیستم های خطی

2– رخداد ها (Events)

معمول است که به یک احتمال، رخداد گفته شود.

یک رخداد، یک مجموعه مشخص از خروجی (outcome) های یک آزمایش می باشد.

فرض می شود خروجی ها دوبدو با هم اشتراک ندارند و اجتماع آن ها کل حالات را پوشش می دهد.

به هر رخداد A می توان عددی P(A) اختصاص داد که از قواعد زیر پیروی می کند:

- $P(A) \geq 0$
- $P(S)=1$
- اگر A و B دوبدو غیرمشترک باشند آنگاه $P(A+B)=P(A)+P(B)$
- عدد P(A) را احتمال A می گویند.

از قواعد بالا، قضایای زیر به دست می آید:

- اگر $\overline{A}$ مکمل A باشد آنگاه
  - $(A+\overline{A})=S$
  - $P(\overline{A})=1-P(A)$
- P(0) احتمال رخداد غیرممکن صفر است.
- $P(A) \leq 1.$

- اگر دو رخداد اشتراک داشته باشند، می توان نشان داد که *P(A+B)=P(A)+P(B)-P(AB).*

احتمال شرطی

احتمال شرطی یک رخداد A با دانستن اینکه رخداد B رخ می دهد به صورت زیر تعریف می شود:

$$P(A|B)=\frac{P(AB)}{P(B)}$$

می توان P(B|A) را بوسیله قانون بیزین استنباط کرد:

$$P(B|A)=P(A|B)\frac{P(B)}{P(A)}$$

استقلال

اگر رخدادهای A و B هیچ ربطی به هم نداشته باشند، می توان گفت که آن ها مستقل اند.

دو رخداد مستقل اند از هم اگر *P(AB)=P(A)P(B)*.

از تعریف استقلال قواعد زیر به دست می آید:

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

سه رخداد A و B و C مستقل اند اگر و فقط اگر:

$$\begin{cases} P(AB)=P(A)P(B) \\ P(AC)=P(A)P(C) \\ P(BC)=P(B)P(C) \\ P(ABC)=P(A)P(B)P(C) \end{cases}$$

3– متغییرهای تصادفی

یک متغییر تصادفی عددی است که به صورت تصادفی به عنوان خروجی آزمایش انتخاب شده است.

متغییرهای تصادفی ممکن است حقیقی و یا مختلط باشند و ممکن است گسسته و یا پیوسته باشند.

البته معمولاً متغییرهای تصادفی گسسته و خقیقی هستند.

می توان یک متغییر تصادفی را با توزیع احتمال آن یا با تابع توزیع احتمال (probability distribution function) آن توصیف کرد.

تابع توزیع یک متغییر تصادفی y احتمال این است که y از یک مقدار u بیشتر نشود:

$$F_y(u)=P(y\le u)$$

همچنین داریم:

$$P(u<y\le v)=F_y(v)-F_y(u)$$

$$f_y(u)=\frac{d}{du}F_y(u)$$

تابع چگالی احتمال مشتق تابع توزیع احتمال می باشد:

- همچنین: $P(u < y \le v) = \int_u^v f_y(y)dy$
- $F_y(\infty) = 1$
- $\int_{-\infty}^{+\infty} f_y(y)dy = 1$

امید ریاضی

می توان یک متغییر تصادفی را علاوه بر تابع توزیع احتمالش با شاخص های آماریش نیز توصیف کرد.

یکی از این شاخص های آماری امید ریاضی (Expected Value) می باشد.

امید ریاضی برای g(x) به صورت E{g(x)} یا <g(x)> نمایش داده می شود و به صورت زیر تعریف می شود:

- متغییر تصادفی پیوسته: $< g(x) >= \int_{-\infty}^{+\infty} g(x) f(x)dx$
- متغییر تصادفی گسسته: $< g(x) >= \sum_x g(x) p(x)$

ممان های متغییر تصادفی

یکی از شاخص های آماری مهم ممان ها (moments) p(x) می باشد.

K امین ممان p(x) برابر امید ریاضی x^k می باشد.

برای یک متغییر تصادفی گسسته: $m_k =< x^k >= \sum_x x^k p(x)$

میانگین و واریانس

ممان اول m_1، همان میانگین (mean) متغییر تصادفی x می باشد.

- پیوسته: $\overline{x} = \int_{-\infty}^{+\infty} xf(x)dx$
- گسسته: $\mu = \overline{x} =< x >= \sum_x xp(x)$

ممان مرکزی دوم، همان واریانس p(x) می باشد:

$$\sigma^2 = \sum_x (x - \overline{x})^2 p(x) = m_2 - \overline{x}^2$$

برای تخمین شاخص ها آماری یک متغییر تصادفی، آزمایشات بسیار زیادی را انجام می دهیم که متغییر را برای دفعات زیادی تولید کند.

در صورتی که آزمایش را Nبار انجام دهیم، هر مقدار x ، Np(x) بار اتفاق می افتد:

- تخمین ممان kام: $\hat{m}_k = \frac{1}{N}\sum_{i=1}^{N} x_i^k$
- تخمین میانگین: $\hat{\mu}_x = \frac{1}{N}\sum_{i=1}^{N} x_i$

دانشگاه امام رضا (علیه السلام)