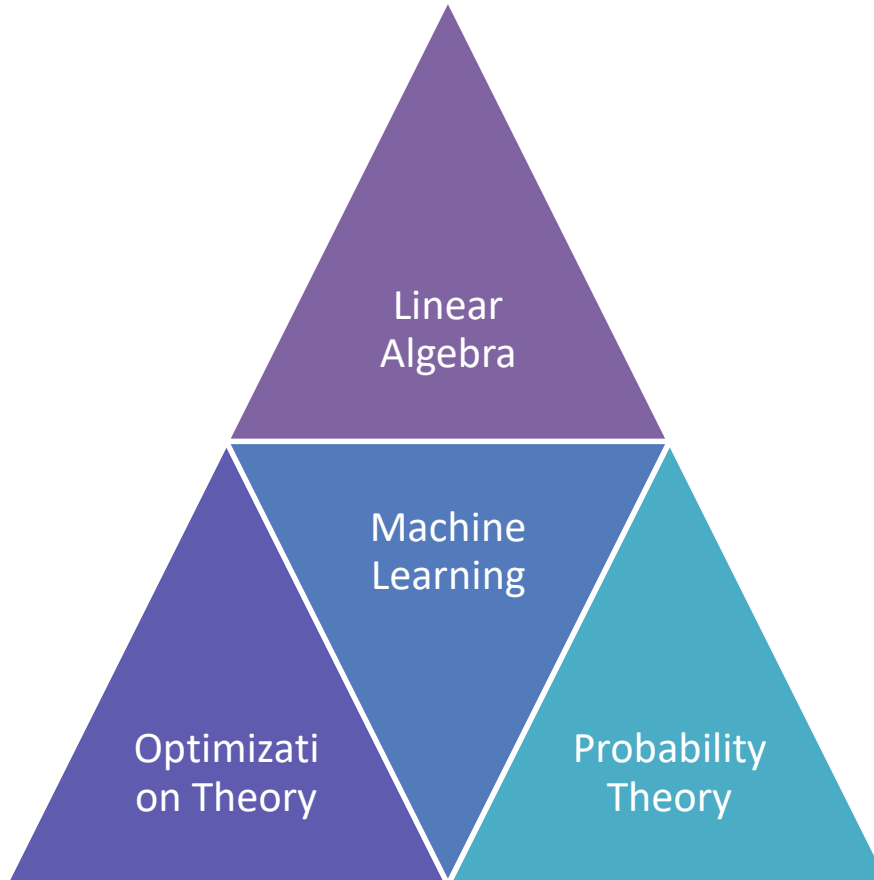


یادآوری مفاهیم پایه

گروه دایچه . dayche.com






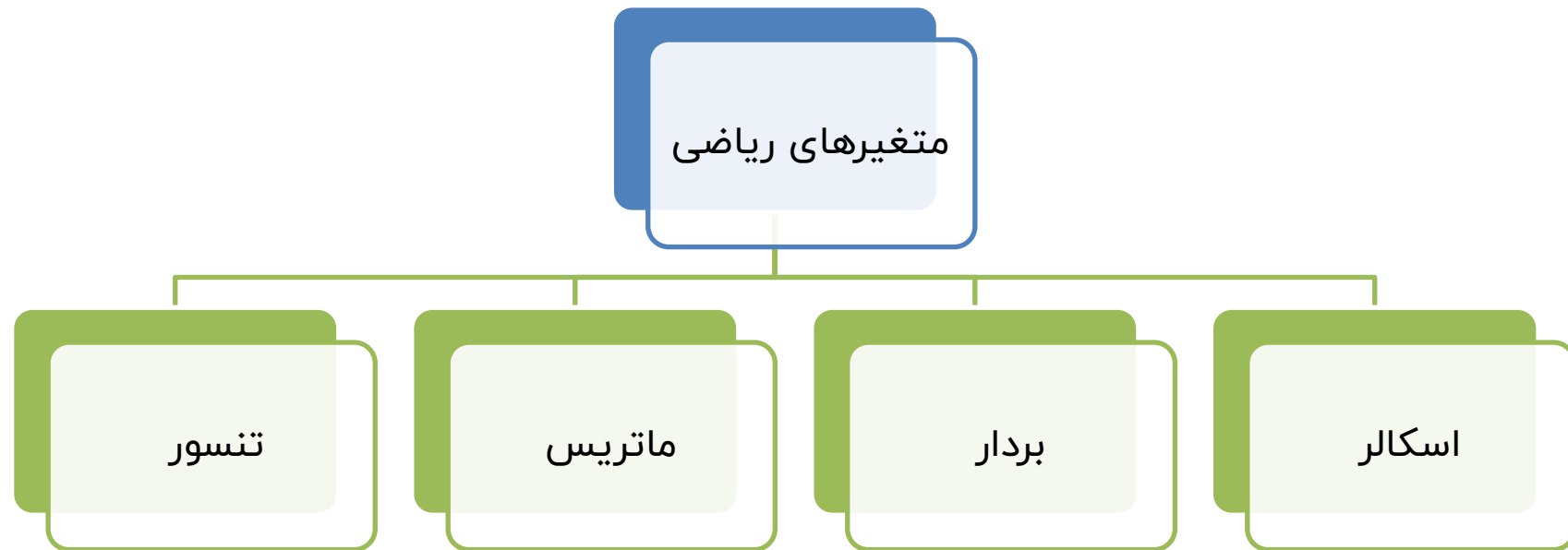
- ابزارهای مورد نیاز برای توسعه یک مدل یادگیر
- جبرخطی
- تئوری احتمالات و فرآیندهای تصادفی
- تئوری بهینه‌سازی

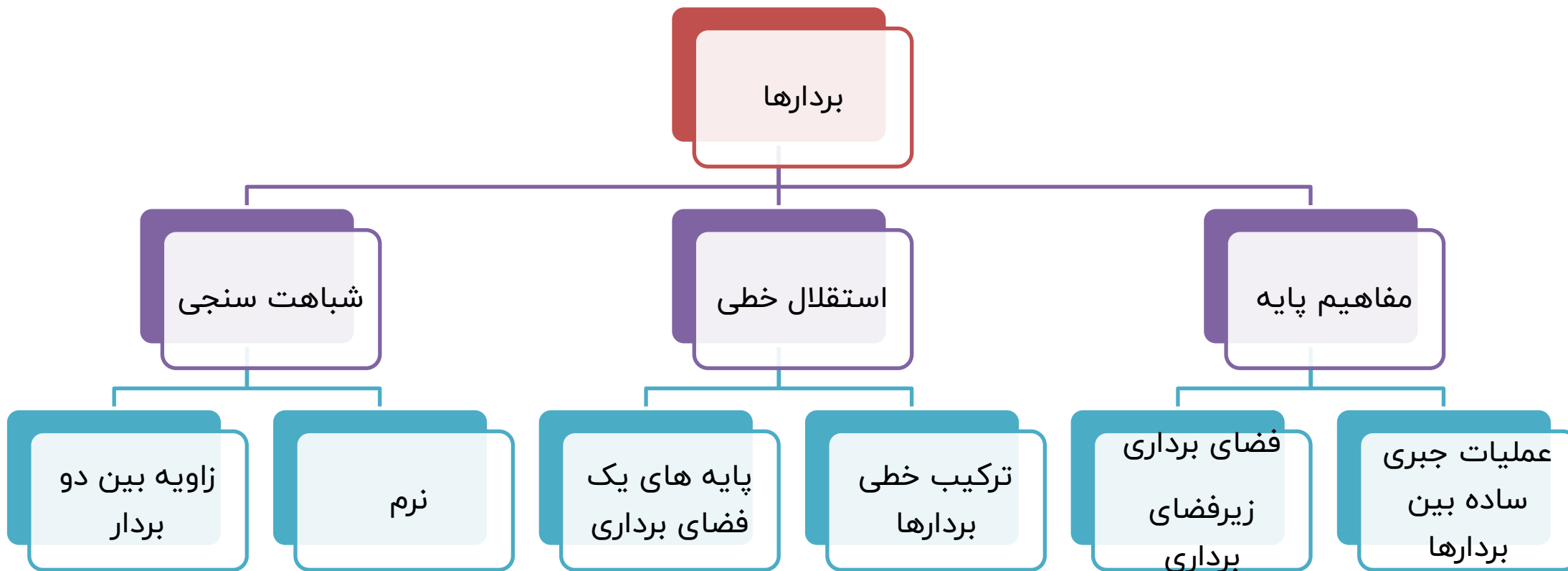
تولید محتوا: وحید محمدزاده ایوقی

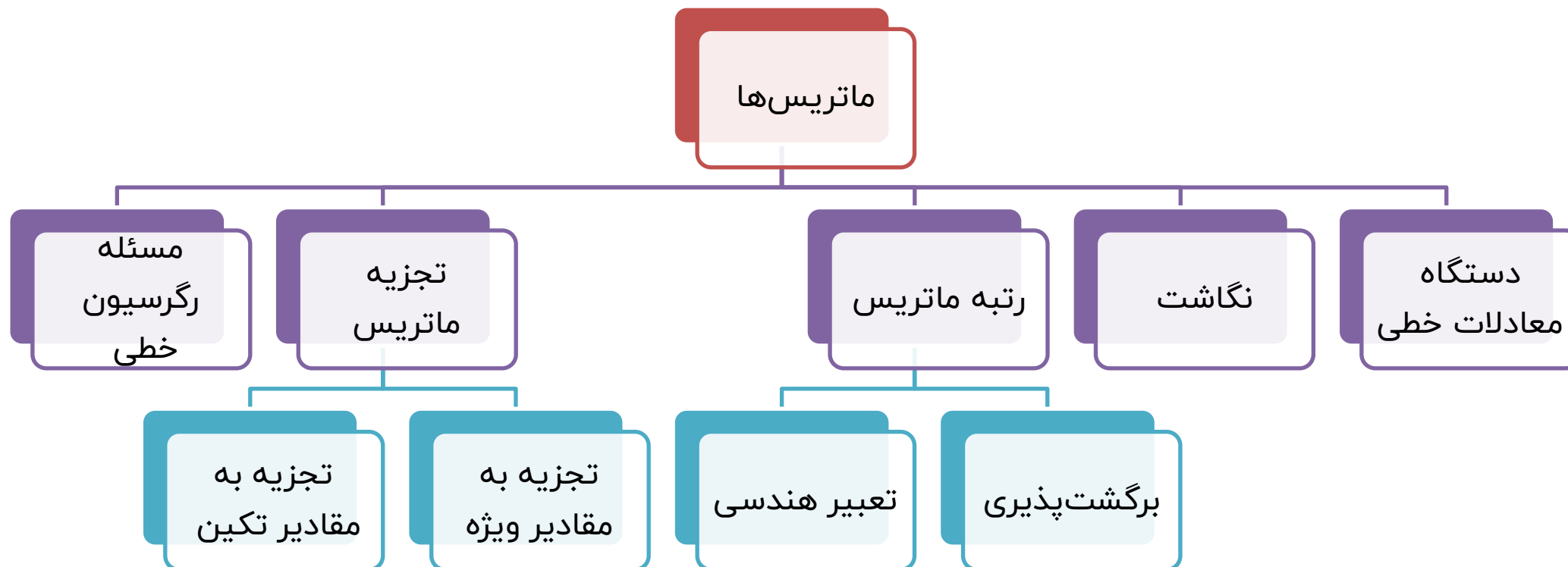
daychegroup 

daychegroup 

dayche.com | گروه دایکه 







ترکیب خطی بردارها



- ترکیب خطی بردارها
- منظور از ترکیب خطی بردارها، جمع وزن دار بردارهاست (هر وزن یک اسکالر است)


$$y = c_1x_1 + c_2x_2 + \dots + c_nx_n \qquad \begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 50 \\ 0 \\ 12 \end{bmatrix}$$

- بردارهای x_1, x_2, \dots, x_n فضای برداری را اسپن می کنند هرگاه هر بردار دلخواه از این فضا را بتوان بر حسب ترکیب خطی این بردارها نوشت

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

پایه‌های فضای برداری

- استقلال خطی

– مجموعه بردارهای x_1, x_2, \dots, x_n دارای استقلال خطی هستند هرگاه:

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = 0 \longrightarrow \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- پایه‌های یک فضای برداری

$$\begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix} = 5 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} + 0 \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} + 2 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}$$

بردارهای پایه فضای سه بعدی


$$\begin{bmatrix} 4 \\ -1 \end{bmatrix} = 4 \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} - 1 \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}$$

بردارهای پایه فضای دو بعدی

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

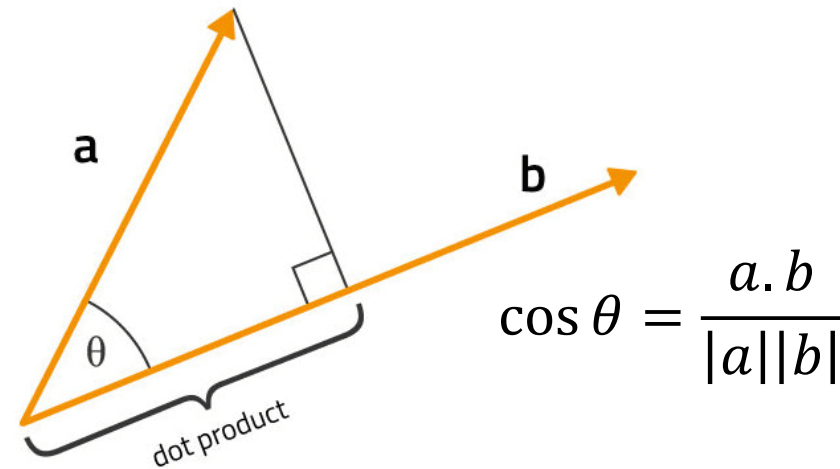
daychegroup 

dayche.com | گروه دایچه 

• ضرب داخلی

- شباهت دو بردار بر اساس این معیار، بر اساس زاویه بین دو بردار تعریف می شود Cosine similarity
- از این معیار در سنجش میزان شباهت بین دو رشته متنی در زمینه پردازش زبان طبیعی استفاده می شود

$$x^T y = \sum_{i=1}^n x_i y_i$$
$$x^T x = \sum_{i=1}^n x_i^2$$





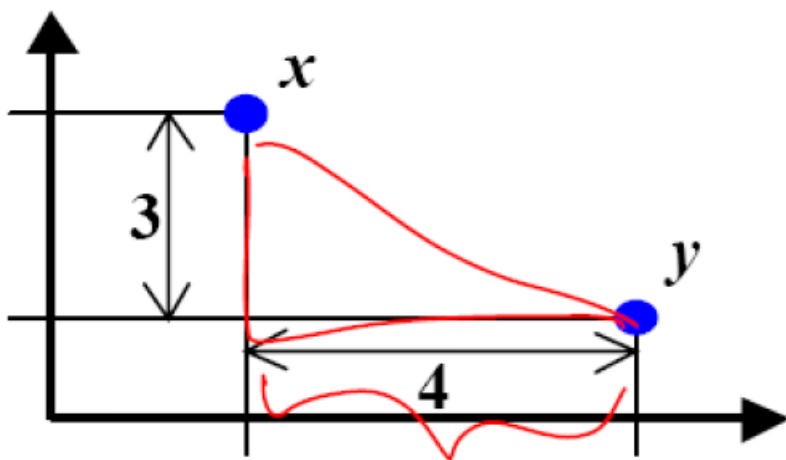
- مفهوم فاصله بین دو بردار را به شکل زیر می‌توان داد:

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

- مسئله اصلی خوشه بندی
- دسته بندی اشیا به نحوی که اشیا هر دسته دارای کمترین فاصله با خود و بیشترین فاصله با اشیا دیگر دسته ها باشد

$$d(x, y) = \left(\sum_{i=1}^n w_i |x_i - y_i|^p \right)^{\frac{1}{p}}$$

- فاصله اقلیدسی $p = 2$ (نرم-2)
- فاصله منهتن $p = 1$ (نرم-1)
- ماکزیمم فاصله $p = +\infty$ (نرم-بی نهایت)
- مینیمم فاصله $p = -\infty$



فاصله اقلیدسی: $\sqrt{4^2 + 3^2} = 5$

فاصله منهتن: $4 + 3 = 7$

سوپریمم (ماکزیمم فاصله): 4

دستگاه معادلات خطی

• دستگاه معادلات خطی


$$y = c_1x_1 + c_2x_2 + \dots + c_nx_n = \underbrace{[x_1 \quad x_2 \quad x_3 \quad \dots \quad x_n]}_{\text{ماتریس ضرایب } (M)} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix}}_{\text{بردار پارامترها } X} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}}_{\text{بردار هدف } b}$$

- چه ترکیب خطی ای از ستون های ماتریس M منجر به تولید بردار هدف b خواهد شد؟
- اگر بردار پارامتر مشخص باشد - رویکرد مستقیم (نگاشت)
- اگر بردار پارامتر مشخص نباشد - رویکرد معکوس (تخمین پارامتر)

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

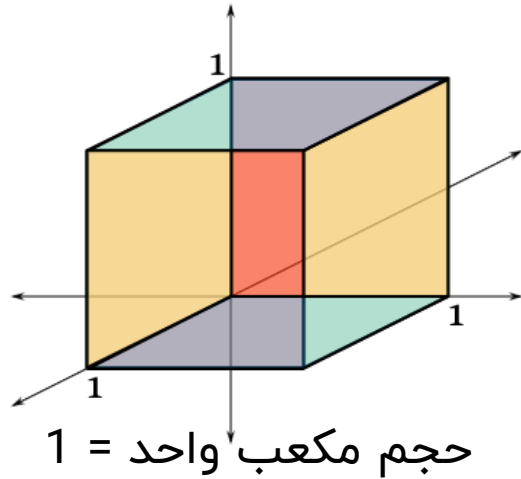


• ساده‌ترین مسئله نگاشت

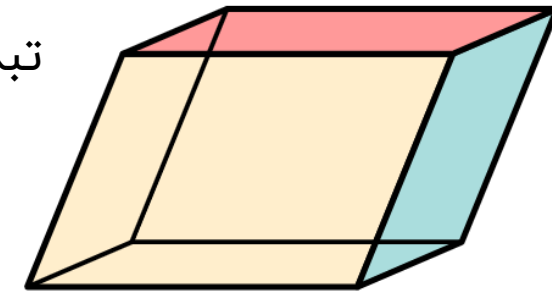
- ضرب ماتریس در یک بردار ستونی (قواعد ضرب ماتریسی و تصویر ستون)
- تبدیل بردار X تحت نگاشت A برابر با بردار b است

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$



تبدیل ماتریس واحد تحت نگاشت A



حجم مکعب انتقالی = دترمینان ماتریس A

• تبدیل خطی

ماتریس A

- تحت چه شرایطی این تبدیل برگشت پذیر است؟



- دستگاه معادلات خطی
- مسئله اصلی جبر خطی

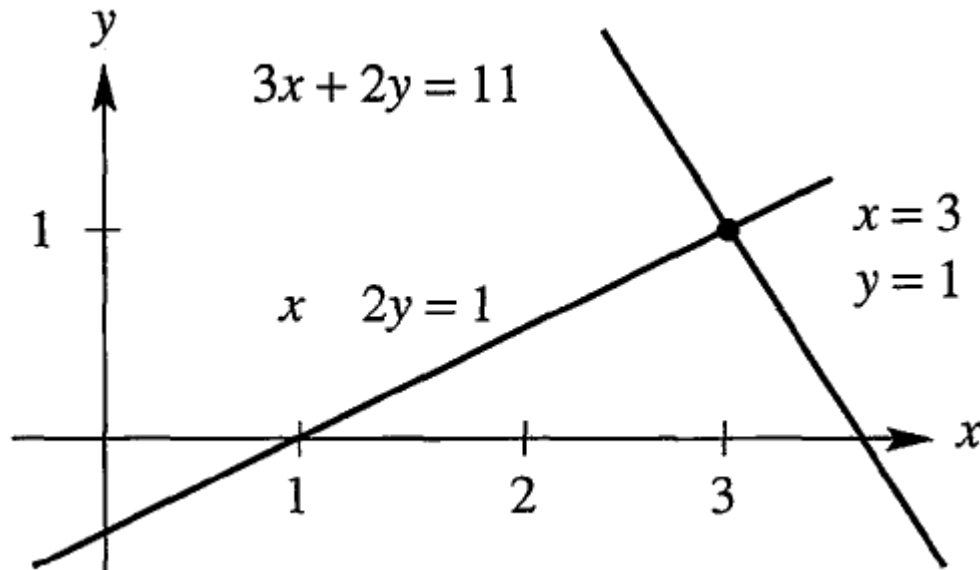
$$y = AX$$

$$\begin{cases} x - 2y = 1 \\ 3x + 2y = 11 \end{cases} \rightarrow \underbrace{\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}}_{\text{ماتریس ضرایب}} \begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 11 \end{bmatrix}}_{\text{بردار هدف}}$$

- روش اول: تصویر بردارهای ستون
- چه ترکیب خطی از بردارهای ستونی ماتریس ضرایب منجر به تولید بردار هدف خواهد شد؟

$$x \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$$

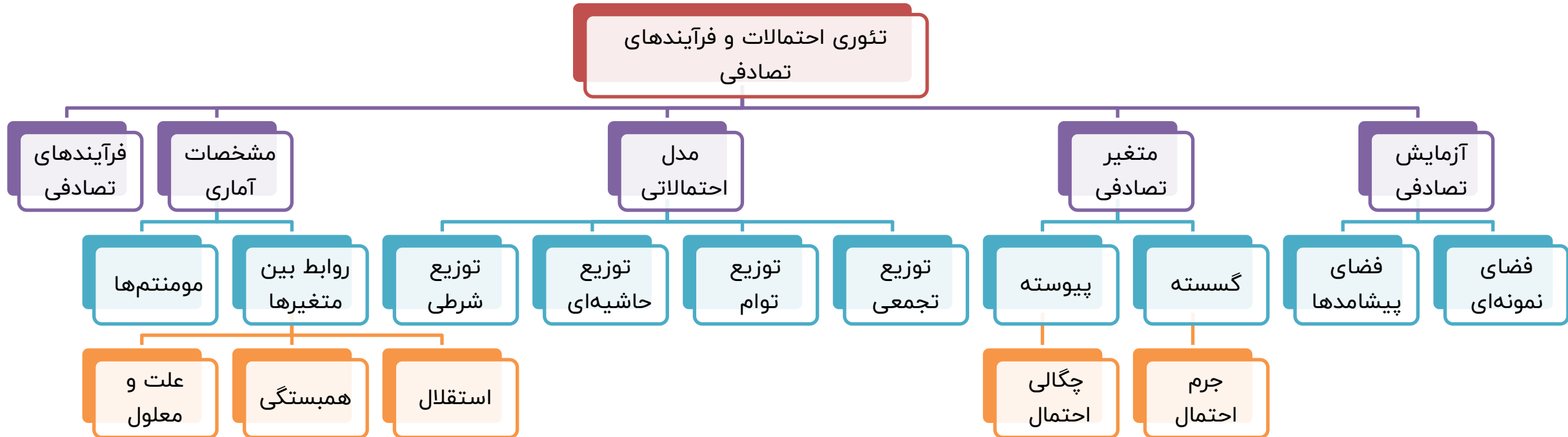
$$\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$$



• روش دوم: تصویر بردارهای ردیف

روش های ذکر شده (مصورسازی معادلات و روش حذفی) در شرایطی که تعداد متغیرها بالا باشند (ابعاد ماتریس ضرایب بیشتر از 2 باشد) قابل پیاده سازی نیستند - راه حل استفاده از روش های مبتنی بر بهینه سازی است


نئوری احتمالات و فرآیندهای تصادفی



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

گروه دایچه | dayche.com 



مسیر حل یک پروژه داده محور

جمع آوری داده



پیش پردازش داده



انتخاب مدل



ارزیابی مدل

- اندازه گیری
- برچسب گذاری

• بسته به نوع تسک، چه پردازشی نیاز است تا نمایش داده را طوری تغییر دهد تا توسعه مدل تصمیم ساده تر صورت بگیرد؟

- مدل انتخاب شده تا چه اندازه به واقعیت مکانیزم تولید داده نزدیک است؟
- مدل انتخاب شده تا چه اندازه در تصمیم گیری های آتی کمک می کند؟

تولید محتوا: وحید محمدزاده ایوقی

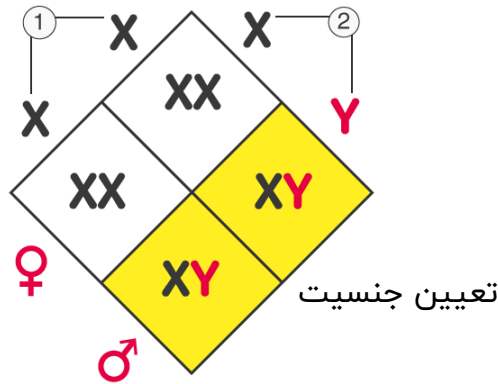
daychegroup

daychegroup

dayche.com | گروه دایچه

آزمایش تصادفی

- آزمایشی که نتیجه آن از قبل مشخص نیست.



انداختن تاس




انداختن سکه

- فضای نمونه‌ای - مجموعه تمام نتایج ممکن از یک آزمایش تصادفی
- برای یک تاس شش وجهی، فضای نمونه برابر است با مجموعه {۱، ۲، ۳، ۴، ۵، ۶}
- مجموعه اعداد حقیقی بین صفر و یک $[0,1]$
- فضای پیشامدها - مجموعه تمام نتایج مطلوب

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

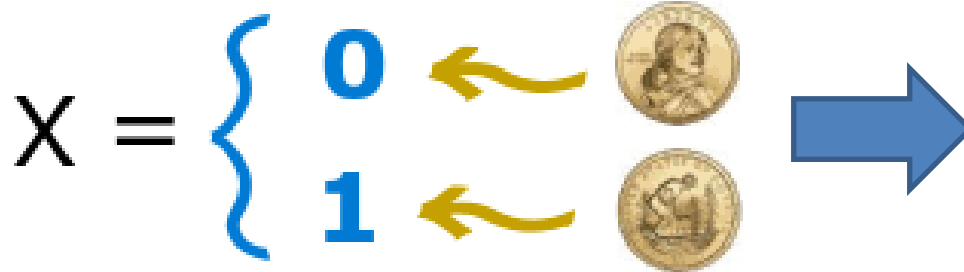
متغیر تصادفی

- نگاهی از فضای نمونه‌ای به مجموعه اعداد
- تغییری که چندین مقدار مختلف را با احتمالات مختلف اختیار می‌کند.
- اصل اساسی در تئوری احتمالات محاسبه این احتمالات است.

Random Variable

Possible Values

Random Events




- محاسبه احتمال بسته به نوع متغیر تصادفی متفاوت است
 - متغیر تصادفی گسسته
 - متغیر تصادفی پیوسته

تولید محتوا: وحید محمدزاده ایوقی

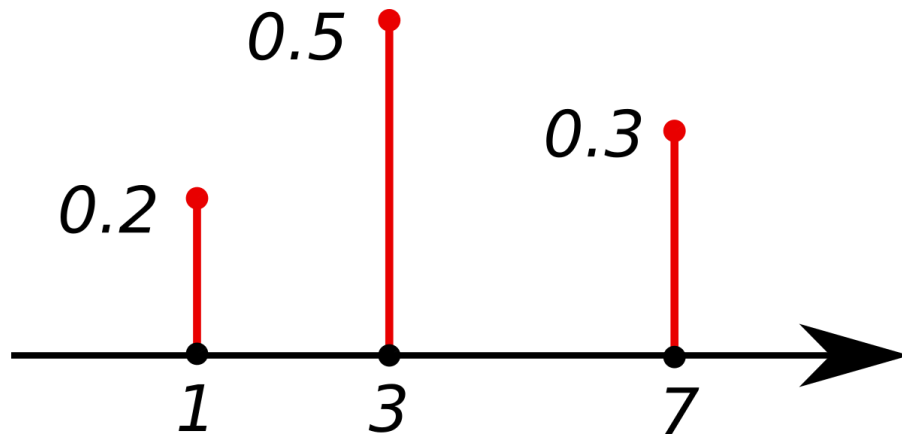
daychegroup 

daychegroup 

dayche.com | گروه دایکه 

محاسبه جرم احتمال

- منظور از احتمال، جرم احتمال است.



$$\begin{aligned} P(X = 1) &= 0.2 \\ P(X = 3) &= 0.5 \\ P(X = 7) &= 0.3 \end{aligned} \quad \rightarrow \quad \sum_x P(X = x) = 1$$

شرط نرمال بودن


- جرم احتمال چگونه محاسبه می‌شود؟

- بسامد نسبی
- آیا می‌توان همین دیدگاه را برای متغیرهای پیوسته تعمیم داد؟

تولید محتوا: وحید محمدزاده ایوقی

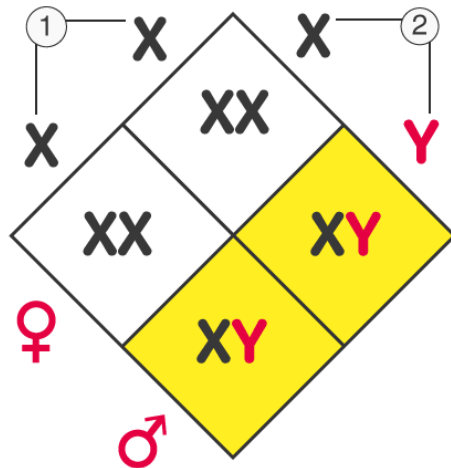
daychegroup 

daychegroup 

dayche.com | گروه دایچه 



• احتمال مبتنی بر تکرار



• احتمال بیزین

تولید محتوا: وحید محمدزاده ایوقی

daychegroup

daychegroup

dayche.com | گروه دایکه

تابع توزیع تجمعی

• تعریف

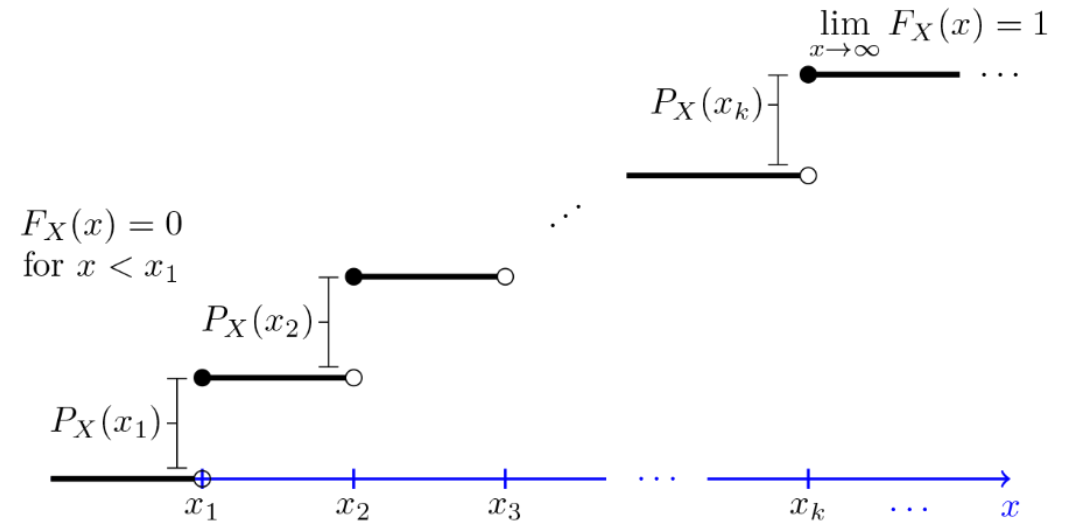
• برای یک متغیر تصادفی X تابع توزیع تجمعی به صورت زیر قابل تعریف است:

$$F_x(x) = Pr\{X < x\}$$

$$F_x(-\infty) = 0, F_x(+\infty) = 1$$

$$x_2 \geq x_1 \rightarrow F_x(x_2) \geq F_x(x_1)$$


$$P(x_1 \leq x \leq x_2) = F_x(x_2) - F_x(x_1)$$



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

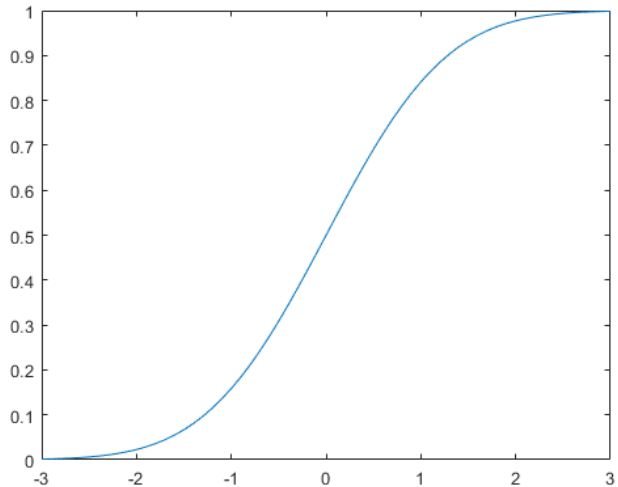
dayche.com | گروه دایچه 

• تابع چگالی احتمال

$$P(x_1 \leq X \leq x_2) = F_x(x_2) - F_x(x_1)$$

$$P(x \leq X \leq x + \Delta x) = F_x(x + \Delta x) - F_x(x) \quad \Rightarrow \quad f(x) = \frac{F_x(x + \Delta x) - F_x(x)}{\Delta x} = F'_x(x)$$

$$f(x)\Delta x = F_x(x + \Delta x) - F_x(x)$$




• تعریف

- متغیر تصادفی پیوسته - تابع توزیع تجمعی پیوسته
- متغیر تصادفی گسسته - تابع توزیع تجمعی گسسته
- متغیر تصادفی مخلوط - تابع توزیع تجمعی تکه‌ای پیوسته

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



- اگر بیش از یک متغیر تصادفی در دست باشد:

$$X = \begin{bmatrix} x_a \\ x_b \end{bmatrix}, x_a \in R^{m_1}, x_b \in R^{m_2}$$

$$F_x(x_a, x_b) = P(X_a \leq x_a, X_b \leq x_b)$$

$$P(a_1 \leq x_a \leq b_1, a_2 \leq x_b \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_a, x_b) dx_a dx_b$$

$$f(x_a) = \int_{-\infty}^{+\infty} f(x_a, x_b) dx_b$$

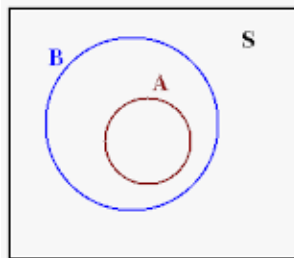
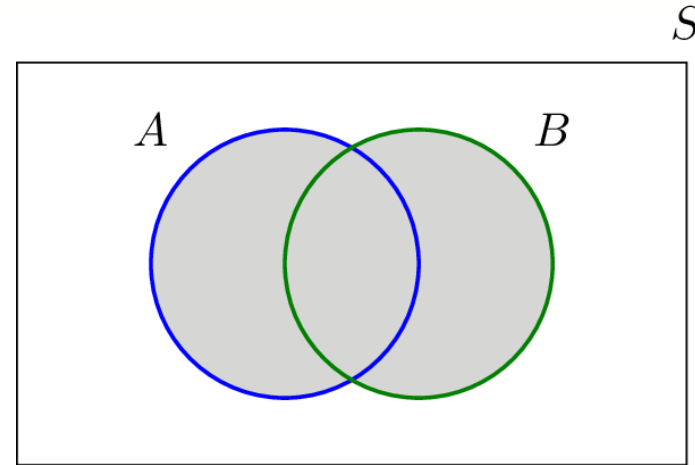
- توزیع حاشیه‌ای

• تعریف

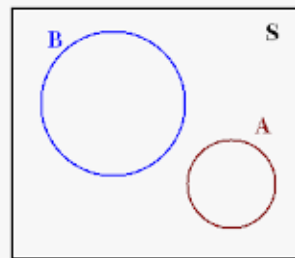
$$f(A|B) = \frac{f(A, B)}{f(B)}$$

$$f(A, B) = f(A|B)f(B)$$

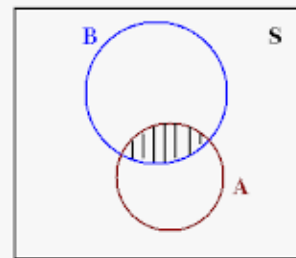
$$f(A) = \sum f(A|B)f(B)$$



(1)



(2)



(3)

(1): $P(B|A) = 1$


(2): $P(B|A) = P(A|B) = 0$

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

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 


توزیع شرطی



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



• تعریف

• در صورت تکرار یک آزمایش تصادفی به تعداد نامتناهی چه اتفاقی رخ می‌دهد؟

$$E(x) = \int_{-\infty}^{+\infty} xf(x)dx \quad \text{گشتاور مرتبه اول}$$

$$E(x^n) = \int_{-\infty}^{+\infty} x^n f(x)dx \quad \text{گشتاور مرتبه } n$$


$$E(x^n y^m) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^n y^m f(x, y) dx dy \quad \text{گشتاور توام مرتبه } mn$$

• اهمیت گشتاورها

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 


dayche.com | گروه دایچه 



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

گروه دایکه | dayche.com 



- **تعریف**

- تنها تعریف استقلال آماری بین دو متغیر به صورت زیر وجود دارد:

$$f(x, y) = f(x)f(y)$$

- **تفسیر**

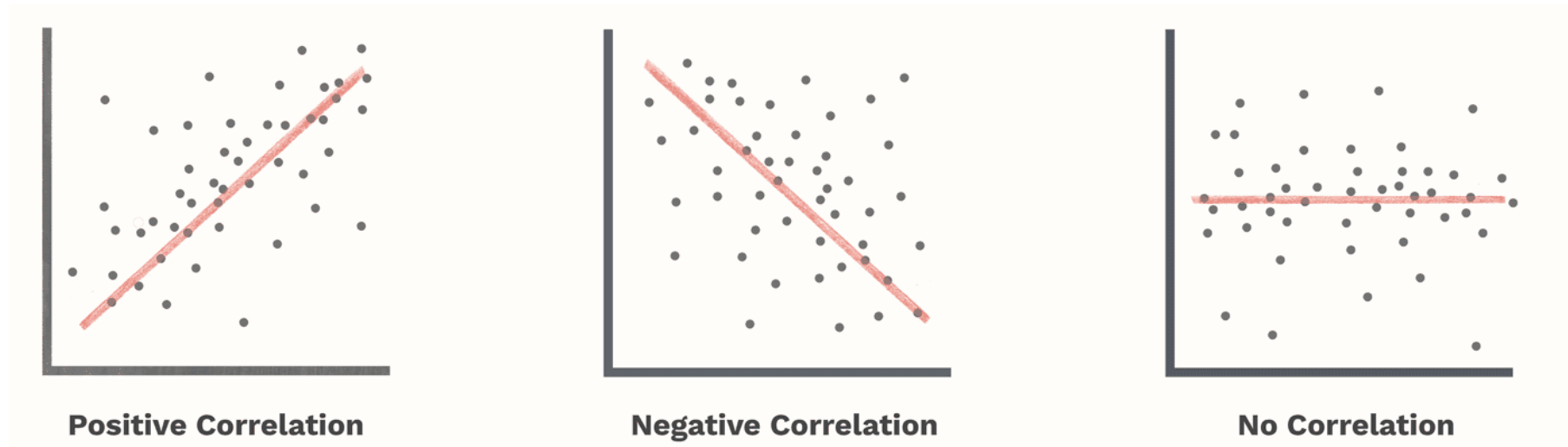
- داشتن اطلاعاتی درباره یک متغیر، چه اطلاعاتی از سایر متغیرها به دست می‌دهد؟

- **علیت**

- استقلال یک مفهوم متقارن است.
- ارتباط بین دو متغیر که توسط متغیرهای زبانی بیان می‌شود، استقلال آماری را عنوان نمی‌کند.

• تعریف

• تغییرات دو متغیر، چه ارتباطی با هم دارند؟




$$E(xy) = \int_{-\infty}^{+\infty} xyf(x, y)dxdy$$

• همبستگی خطی

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 


همبستگی آماری



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

گروه دایکه | dayche.com 

فرآیندهای تصادفی




- تعریف
- تابعی از دو متغیر - یک متغیر تصادفی و یک متغیر معین
- ایستایی
- ارگادیسیتة
- نمونه‌گیری
- تئوری تخمین

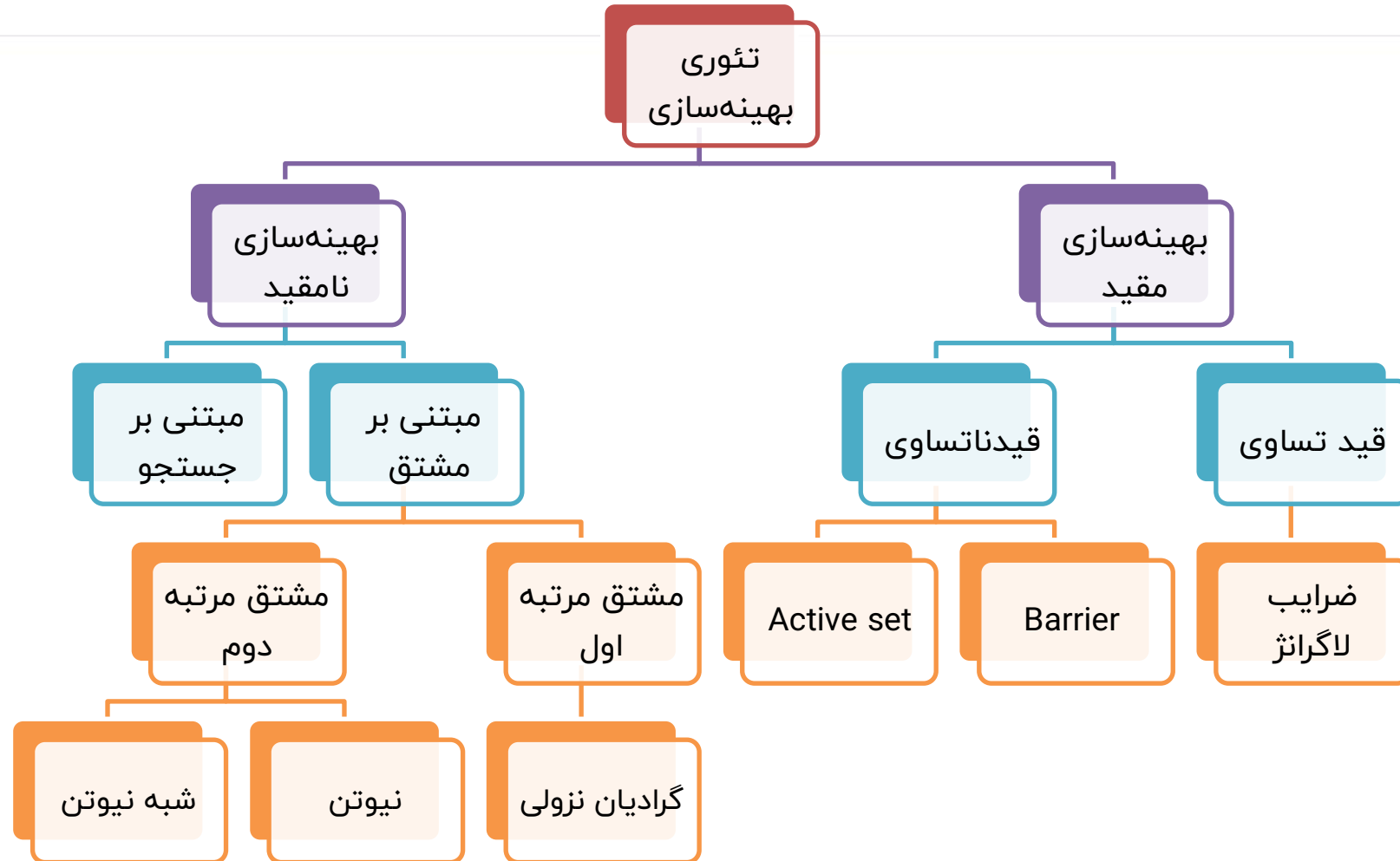
تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایکه 


تئوری بهینه‌سازی



تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

مسئله استاندارد بهینه‌سازی

➤ مسئله استاندارد بهینه‌سازی

➤ یک مسئله بهینه‌سازی استاندارد همواره به صورت زیر نوشته می‌شود.

$$x^* = \arg \min f(x)$$

$$h_j(x) = 0, j = 1, 2, \dots, k$$

$$g_i(x) < 0, i = 1, 2, \dots, m$$

Linear programming

$$x^* = \arg \min C^T x$$

$$A_j x = 0, j = 1, 2, \dots, k$$

$$B_i x - b < 0, i = 1, 2, \dots, m$$



$$x^* = \arg \max f(x) = - \arg \min f(x)$$

Quadratic programming

$$x^* = \arg \min x^T C x$$

$$A_j x = 0, j = 1, 2, \dots, k$$

$$B_i x - b < 0, i = 1, 2, \dots, m$$

تولید محتوا: وحید محمدزاده ایوقی

daychegroup 

daychegroup 

dayche.com | گروه دایچه 