CSE474/574 Machine Learning

1. Gradiance Quiz – Due Tuesday
2. PA2 – released
3. Need to work in groups
4. Office hours

$P(\text{heads})$ $\quad 0 \le P(\cdot) \le 1$

$P(\text{tails})$ $\quad \underline{P(\text{heads}) + P(\text{Tails}) = 1}$

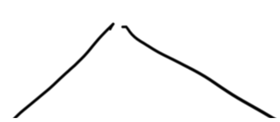
<u>Random Variable</u>

<u>Domain</u> $\{\text{Heads}, \text{Tails}\}$

$\{1, 2, 3, 4, 5, 6\}$

$\{1, 2, 3\}$

$\{1, 2, 3, 4, \ldots, 12\}$



Discrete or categorical

Continuous random variables

Domain - finite

X

x

$P(X = x) \equiv P(x)$

$\boxed{p \text{ — density}}$

Joint probability

$P(A \wedge B) \equiv P(A, B)$

$$\boxed{\begin{aligned} P(A, B) &= P(A \mid B)\, P(B) \\ &= P(B \mid A)\, P(A) \end{aligned}}$$

$$P(A \mid B) = \frac{P(A, B)}{P(B)}$$

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

$$\boxed{\begin{aligned} &P(X_1 = x_1, X_2 = x_2, \ldots, X_d = x_d) \\ &= P(X_1 = x_1)\, P(X_2 = x_2 \mid X_1 = x_1) \\ &\quad \ldots \end{aligned}}$$

Let the domain of $X = \{1, 2, 3\}$

Let domain of $Y = \{a, b\}$

$Z = (X, Y)$ $\quad \{ (1, a), (1, b), (2, a), (2, b), (3, a), (3, b) \}$

$P(X = x) = \sum P(X = x) | Y = y$

marginal dist. of

Bayes Rule or theorem

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

$$P(Y=y) = \sum_{y'} P(X=x, Y=y')$$

$$\boxed{P(X=x, Y=y) = P(Y=y \mid X=x) * P(X=x)}$$

Example:

$\boxed{P(Y=1 \mid X=1)}$

X – test is +ve or not

Y – have cancer or not

$\boxed{P(X=1 \mid Y=1)}$

$$P(Y=1 \mid X=1) = \frac{\overset{0.8}{P(X=1 \mid Y=1)}\ \overset{0.004}{P(Y=1)}}{\underset{0.8}{P(X=1 \mid Y=1)}\ \overset{0.004}{P(Y=1)} + P(X=1 \mid Y=0)\ \underset{\substack{1-0.004\\=0.996}}{P(Y=0)}}$$

$\boxed{P(X=1 \mid Y=0)} \rightarrow$ false alarm rate

0.1

X

If X is a categorical or discrete r.v

$$\text{Domain}(x) = \{1, 2, 3, 4\}$$

$P(X)$ $\quad$ $P(X=1)$

$P(X=2)$

$P(X=3)$

$P(X=4)$

Continuous r.v

Let $\text{Dom}(x) = (-\infty, \infty)$

$P(X=0)$

$P(X=0.3)$

Probability density. (PDF)

$p(\cancel{X}=x)$

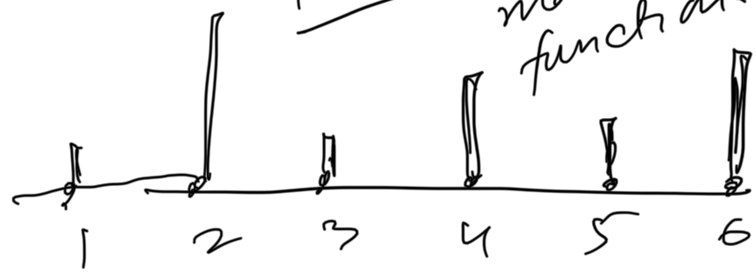
$$P(a < X \leq b)$$

$$= P(X \leq b) - P(X < a)$$

$$\boxed{\begin{array}{l}\text{CDF - cumulative distribution function.} \\ F(x) = P(X \leq x)\end{array}}$$

$$p(a < X \le b) = \int_a^b p(x)\,dx$$

pmf probability mass function



$p(X=x)$

$p(x)$

pdf

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

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

DRV: $\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x\,P(X=x)$

domain of X

CRV $\mathbb{E}[X] = \int_{\mathcal{X}} x\, p(x)\, d$

Also known as the mean (μ)

$$f(X) = \begin{cases} -100 & \text{if } x = \text{tails} \\ +900 & \text{if } x = \text{heads} \end{cases}$$

$$p(X = h) = 0.5$$
$$p(X = t) = 0.5$$

$$\mathbb{E}[f(X)] \sum_{x \in \mathcal{X}} f(x)\, P(X = x)$$

$$= 0.5 \times (-100) + 0.5 \times 900$$
$$= \underline{400}$$

Annoucements

(i) Gradiance 5 released
– Due next Thesday midnight

Random Variable

X — Domain $\mathcal{X}$ (Support)

Discrete

Continuous

$\boxed{P(X=x)}$ $x \in \mathcal{X}$

PMF

$p(X=x)$

min $\leq x \leq$ m

PDF $p(x)$

$\boxed{P_A(X=x)}$

$P_B(X=x)$

A probability distribution

A PDF or a PMF

1. Domain ✓
2. PMF or PDF ✓

Bernoulli

Domain — $\{0, 1\}$

$\{\text{Yes}, \text{No}\}$

$\{\text{Heads}, \text{Tails}\}$

PMF $\boxed{\text{Parameter} - \theta}$

Bernoulli parameter — p.

$0 \leq p \leq 1$

$$\boxed{\begin{aligned} P(X = \text{heads}) &= p \\ P(X = \text{tails}) &= 1 - p \end{aligned}} \longrightarrow \text{PMF}$$

Binomial Distribution

$\underline{n, \theta}$ (Same as p)

Domain? $\{0, 1, 2, \ldots, n\}$

PMF $p(X=k) = Bin(k|n,\theta)$

$$= \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

$$\boxed{{}^nC_k = \frac{n!}{k!(n-k)!}}$$

Multinoulli → Generalization of Bernoulli

$$\theta = [p_1, p_2, p_3 \ldots p_K] \rightarrow \boxed{\begin{array}{l} 0 \le p_i \le 1 \\ \sum p_i = 1 \end{array}}$$

PMF $p(X=1) = p_1$

$\vdots$

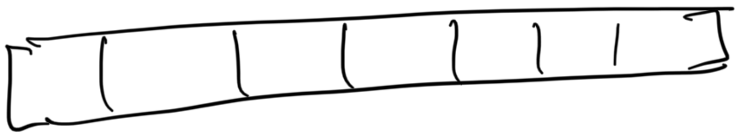
$p(X=k) = p_k$

0	0	0	1	0	0	0

Multinomial — Generalization of Binomial

n, $\theta \rightarrow$ vector

$X \quad n)$



Poisson

$$\boxed{P(X=k) = \frac{\lambda^{k} e^{-\lambda}}{k!}}$$

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^{k} e^{-\lambda}}{k!}$$

$$= \lambda$$

Gaussian (or Normal) Distribution

Domain: $-\infty, \infty$

$$\boxed{\text{pdf: } p(X=x) = \mathcal{N}(x \mid \mu, \sigma^2)}$$

$$\boxed{= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}}$$

$$\mathbb{E}[X] = \int^{\infty} x\, p(X=x)\, dx$$

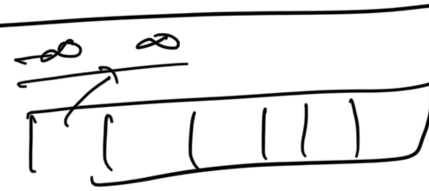
$-\infty$

$= \mu$

first moment.

$$\boxed{E[X - E[X]]^2 = \sigma^2}$$

If $X \in \mathbb{R}^D$



MVN → Multivariate Gaussian or Normal

1D: μ, σ^2

$\underline{\mu}$ → a vector with D values

$\underline{\Sigma}$ → a matrix $(D \times D)$

X → a vector with D values

$$\text{pdf}(X) = \mathcal{N}(X \mid \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right]$$

Σ – Σ^{-1} should exist