Date: Tuesday, January 22, 2019

COMPUTATIONAL STATISTICS

MSDS 628-01

James Wilson

- resampling methods (bootstrap)
- MCMC
- kernel density estimation (histogram)
- artificial neural networks
- additive models

Bayesian truth serum
NLP: Latent Dirichlet Allocation overall topic w/distribution
Experimental design: multi-armed bandit gambling
Reinforcement learning

Review of Statistical Models

Random variable: Number w/ prob. distr.
what values with what probability?

Definition:

$\mathcal{X}$: Possible values for a random variable X

Pdf: $\mathbb{P}(X=x)$, $x \in \mathcal{X}$

cdf: $\mathbb{P}(X \leq x)$, $x \in \mathcal{X}$

Example:

→ 3 coin, head.

$\mathbb{P}(Y=0) = \binom{3}{0}\frac{1}{2^3}$ $\quad$ $\mathbb{P}(Y=2) = \binom{3}{2}\frac{1}{2^3}$

$\mathbb{P}(Y=1) = \binom{3}{1}\frac{1}{2^3}$ $\quad$ $\mathbb{P}(Y=3) = \binom{3}{3}\frac{1}{2^3}$

} PMF $\Sigma = 1$

Example: (Continous RVs)

$X \sim U(0,1)$

$\mathbb{P}(X = 0.5) = 0$

↳ X could take infinite values $(\frac{1}{\infty} = 0)$.
Reason why we think about intervals for continous variables.

$\mathbb{P}(X \leq 0.5) = 1/2$ PDF

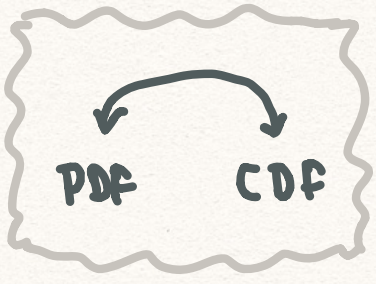
* For continous RVs, we use $\mathbb{P}(X \in A)$ (contained in some set A)

$$\mathbb{P}(X \in B) = \int_B f(x)dx$$

Example:

$\mathbb{P}(Y \leq 0) = 1/8$

$$\mathbb{P}(Y \leq y) = \begin{cases} 1/8 & y \in [0,1) \\ 4/8 & y \in [1,2) \\ 7/8 & y \in [2,3) \\ 1 & y \geq 3 \end{cases}$$



$F'(a) = \frac{d}{dx} F(a) = f(a)$.

Distributions

- Expected value
  ↳ Linear
- Standard deviation
- Variance
  ↳ Quadratic

* Jointly distributed random variables

$$F(a,b) = P(X \le a, Y \le b), \quad -\infty < a,b < \infty$$

If you have multiple rvs, knowing their joint distribution gives all information about the RV's marginal distribution. Also, having the joint pdf is equivalent to having the conditional distribution of each RV given the rest.

Having the marginal pdfs of each RVs does not give us the joint distribution or conditional distributions of all RVs.

Disjoint vs Independent

↓ same person voting twice

↓ votes of two people

joint density: product independent.

Next class: Ch1.

Clone git repository