

Probability Distributions

Gavin Cheng
GitHub: cwgavin

Wednesday, 2 December 2020

14:17

Functions:

- Gamma function
 - o factorials of complex numbers (real part > 0)
 - o $\Gamma(n) = (n - 1)!$
- Digamma function
 - o $\Psi(x) = \frac{d \log \Gamma(x)}{dx}$
 - o $\Psi(x + 1) = \Psi(x) + \frac{1}{x}$
- Beta function
 - o $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$
- Multivariate Beta function
 - o $B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^k \alpha_i)}$

Permutations and Combinations

Number of permutations
(order matters) of n things
taken r at a time:

$$P(n, r) = \frac{n!}{(n-r)!}$$

Number of combinations
(order does not matter) of n
things taken r at a time:

$$C(n, r) = \frac{n!}{(n-r)!r!}$$

Number of different permutations of n
objects where there are n_1 repeated items,
 n_2 repeated items, ... n_k repeated items

$$\frac{n!}{n_1!n_2!\dots n_k!}$$

Distributions:

- Bernoulli distribution
 - o throw a coin
 - o $P(X=1) = p$
 - o $P(X=0) = 1 - p$
 - o $E[X] = p$
 - o $\text{Var}[X] = p(1 - p)$
- Binomial distribution $B(n, p)$
 - o n Bernoulli trials, got k successes
 - o $n=1$: Bernoulli
 - o PMF: $P(X=k) = C(n, k) p^k (1 - p)^{n-k}$
 - o $E[X] = np$
 - o $\text{Var}[X] = np(1 - p)$
- Geometric distribution
 - o k Bernoulli trials until 1 success (including)
 - o PMF: $P(X=k) = (1 - p)^{k-1}p$
 - o $E[X] = \frac{1}{p}$
 - o $\text{Var}[X] = \frac{1-p}{p^2}$
- Categorical distribution
 - o roll a k -sided die
 - o PMF:
 $f(x = i | p_1, \dots, p_k) = p_i \quad \sum_{i=1}^k p_i = 1$
or
 $f(\mathbf{x} | p_1, \dots, p_k) = \prod_{i=1}^k p_i^{x_i}$

where \mathbf{x} is a k -dimensional one-hot encoding vector

- Multinomial distribution

- n trials of rolling a k -sided die
- $k=2, n=1$: Bernoulli
- $k=2, n>1$: Binomial
- PMF:

\mathbf{x} is a k -dimensional vector with counts for each possible outcome

$$f(x_1, \dots, x_k \mid n, p_1, \dots, p_k) = P(X_1 = x_1, \dots, X_k = x_k) \\ = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} \quad \sum_{i=1}^k x_i = n$$

or

$$f(x_1, \dots, x_k \mid p_1, \dots, p_k) = \frac{\Gamma(\sum_i x_i + 1)}{\prod_i \Gamma(x_i + 1)} \prod_{i=1}^k p_i^{x_i}$$

- $E[X_i] = np_i$
- $\text{Var}[X_i] = np_i(1 - p_i)$

- Beta distribution

- defined on interval $[0, 1]$
- $E[X] = \frac{a}{a+b}$
- $\text{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)}$

- Poisson distribution $\text{Poisson}(\lambda)$

- $B(n, p)$ when $np = \lambda, n \rightarrow \infty$

- Gamma distribution

- continuous version of Poisson distribution on positive real numbers
- a : shape parameter; b : rate parameter ($\frac{1}{b}$: scale parameter)
- PDF: $f(x \mid a, b) = \frac{b^a x^{a-1} e^{-bx}}{\Gamma(a)} \quad x, a, b > 0$
- $E[X] = \frac{a}{b}$
- $\text{Var}[X] = \frac{a}{b^2}$

- Dirichlet distribution $\text{Dir}(\alpha_1, \dots, \alpha_k)$

- Beta distribution in high dimensionalities
- PDF:

$$f(x_1, \dots, x_k \mid \alpha_1, \dots, \alpha_k) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^k x_i^{\alpha_i - 1} \quad \sum_{i=1}^k x_i = 1 \quad x_i \geq 0$$

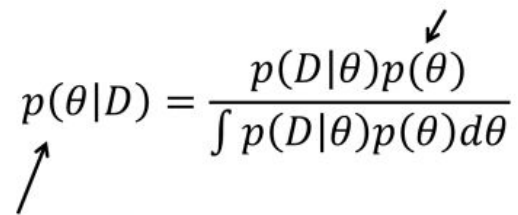
$$B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^k \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^k \alpha_i)}$$

- $E[X_i] = \frac{\alpha_i}{\sum_{i=1}^k \alpha_i}$

More about conjugate distributions:

Recall the **Bayes' Rule**:

1. Choose prior that is conjugate to likelihood

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta}$$


2. The posterior will have **same form** as conjugate prior distribution, i.e. **closed-form**.

知乎 @ailin

(E.g., Beta-Binomial distribution & Dirichlet-Multinomial distribution)