Central Limit Theorem Note

Lizi Chen

The binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ counts the number of ways of choosing k items from n items.

Suppose $X_1, \cdots, X_n$ are n independent random variables, each equal to 1 with probability p, and 0 otherwise. Then the distribution of the random variable $Y = \sum_{i=1}^{n} X_i$ counts the total number of 1's.

This is called binomial distribution. Specifically, we have

$$\mathbf{Pr}[Y = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

The variance of a real-valued r.v. X is defined to be:

$$VarX = \mathbf{E}[(X - \mathbf{E}[X])^2]$$

The *standard deviation* of X is $\sqrt{VarX}$, is the standard measure of the "spread" of the distribution.

In one dimension, a r.v. X obeys a normal / Gaussian distribution with mean μ and standard deviation σ if its pdf is given by:

$$p(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

This means that

$$\forall z \in \mathbb{R}, \mathbf{Pr}[X \leq z] = \int_{-\infty}^{z} p(x; \mu, \sigma) dx$$

The **Central Limit Theorem** states that the sum of a large number of independent random variables, when properly standardized, <u>will converge to a normal distribution.</u>

More precisely, let $X_1, \cdots, X_n$ be iid r.v., each with mean μ and standard deviation σ. Let

$$Y_n = \sum_{i=1}^{n} X_i$$

be the sum of the first n variables, and let

$$Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

be the standardized version of the sum with mean 0 and standard deviation 1. Then the **Central Limit Theorem** states that Z_n converges in distribution to standard normal, meaning that

$$\forall z \in \mathbb{R}, \lim_{n \to \infty} \mathbf{Pr}[Z_n \leq z] = \mathbf{Pr}[Z^* \leq z]$$

where Z^* is a standard normal r.v.

Proof: Jaynes 2003, p222, or Rice 1995, p169

ToDo:

- Borel-Cantelli Lemma
- Weak Law of Large Number
- Strong Law of Large Number
- Central Limit Theorem
- Varadarajan???

References:

- Robert E. Schapire, Yoav Freund, Boosting Foundations and Algorithms.
- Ingo Steinwart, Andreas Christmann, Support Vector Machines.