1 Introduction

Stein’s method was invented in early 70s by Charles Stein as a method for proving central limit theorems.

1.1 Convergence in distribution

The cumulative distribution function of a random variable \( X \) is defined as

\[
F(t) := \mathbb{P}(X \leq t).
\]

Suppose we have a sequence of r.v. \( \{X_n\} \) with c.d.f. \( \{F_n\} \) and a r.v. \( X \) with c.d.f. \( F \). We say that \( X_n \) converges in distribution (or converges in law, or converges weakly) to \( X \) if

\[
\lim_{n \to \infty} F_n(t) = F(t) \quad \text{for all continuity points } t \text{ of } F.
\]

This is denoted by \( X_n \Rightarrow X \) or \( F_n \Rightarrow F \) or \( X_n \Rightarrow F \). The following theorem is standard.

**Theorem 1** The following are equivalent:

1. \( X_n \Rightarrow X \)
2. \( \mathbb{E} f(X_n) \rightarrow \mathbb{E} f(X) \) for all bounded continuous \( f \)
3. \( \mathbb{E} f(X_n) \rightarrow \mathbb{E} f(X) \) for all bounded Lipschitz \( f \)
4. \( \mathbb{E}(e^{itX_n}) \rightarrow \mathbb{E} e^{itX} \) for all \( t \)

Recall that the function \( \varphi(t) := \mathbb{E}(e^{itX}) \) is known as the characteristic function of \( X \).
1.2 Central limit theorems

Recall: The standard gaussian distribution $N(0,1)$ has density
\[
\varphi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}.
\]
We will usually denote standard gaussian r.v. by $Z$.

Basic central limit theorem: If $X_1, X_2, \ldots$ iid r.v. with mean 0 variance 1 then
\[
\frac{\sum_{i=1}^{n} X_i}{\sqrt{n}} \Rightarrow N(0,1).
\]
One standard method of proof uses characteristic functions.

\[
E \left[ e^{it\sum_{i=1}^{n} X_i/\sqrt{n}} \right] = E \left[ e^{it\sum_{i=1}^{n} X_i/\sqrt{n}} \right]^n = \left[ 1 + \frac{it}{\sqrt{n}} E X + \frac{(it)^2}{2n} E X^2 + \cdots \right]^n 
\approx \left( 1 - \frac{t^2}{2n} \right)^n \rightarrow e^{-t^2/2} = E \left( e^{itZ} \right).
\]

1.3 Some examples that we will cover

We will apply Stein’s method to situations where it’s hard to apply standard arguments. Some examples are as follows.

Hoeffding’s combinatorial CLT

Suppose $\pi$ is a random (uniform) permutation of $\{1, \ldots, n\}$, and consider the following distance from identity:
\[
W_n = \sum_{i=1}^{n} |i - \pi(i)|
\]
This is known as Spearman’s footrule.

Known result: As $n$ becomes large,
\[
\frac{W_n - EW_n}{\sqrt{\text{Var}(W_n)}} \Rightarrow N(0,1).
\]

More generally, we have Hoeffding’s combinatorial CLT.
• Array of numbers \((a_{ij})_{i,j \leq n}\) satisfying certain conditions.

• \(\pi\) a random permutation

• \(W_n = \sum_i a_{i\pi(i)}\) (for spearman, \(a_{ij} = |i - j|\))

How close is \(\frac{W_n - EW_n}{\sqrt{\text{Var}(W_n)}}\) to \(N(0, 1)\)? This was Stein’s original motivation.

**Linear statistics of eigenvalues**

Suppose \((X_{ij})_{1 \leq i,j \leq n}\) are iid rv’s with mean 0 and variance 1, \(X_{ji} = X_{ij}\). Then

\[A_n = \frac{1}{\sqrt{n}}(X_{ij})\]

is known as a *Wigner matrix*. Let \(\lambda_1, \ldots, \lambda_n\) be the eigenvalues of \(A_n\). Then it is known that

\[\frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \implies \text{semicircle law}\]

which has density

\[\frac{1}{2\pi} \sqrt{4 - x^2}\]

on \([-2, 2]\).

We may want to look at fluctuations of random distribution about a fixed distribution.

Look at \(W_n = \sum_{i=1}^{n} f(\lambda_i)\).

Then \(W_n - EW_n \implies N(0, \sigma^2(f))\). Main restrictions needed:

1. \(\mathbb{E}\left(X_{ij}^{2m}\right) \leq (Cm)^m\) for all \(m\).

2. \(X_{ij}\)’s have symmetric distribution around zero; not needed to be iid.

Original proof is by method of moments.

**Curie-Weiss Model**

• \(N\) magnetic particles, each with spin +1 or -1.

• Spins denoted by \(\sigma_1, \ldots, \sigma_N\).
• The particles try to align themselves together with the same spin.

• Simplest model [mean-field model]:
  \[ P(\sigma) = \frac{1}{Z} \exp \left( \frac{\beta}{2} \sum_{1 \leq i < j \leq n} \sigma_i \sigma_j \right) \]
  where \( \beta = 1/kT \) with \( T \) being the temperature and \( k \) the Boltzmann constant.

• The magnetization of the system
  \[ m(\sigma) = \frac{1}{n} \sum \sigma_i \]
  If \( \beta = 0 \), then we have iid, and magnetization is close to 0.

• Known that for \( \beta \leq 1 \), \( m(\sigma) \to 0 \) in probability as \( n \to \infty \).

• For \( \beta > 1 \), the equation \( x = \tanh(\beta x) \) has two solutions, \( m^*(\beta) \) and \( -m^*(\beta) \), say \( m^* > 0 \) and
  \[ m(\sigma) \Rightarrow \frac{1}{2} (\delta_{m^*(\beta)} + \delta_{-m^*(\beta)}) . \]

• If \( \beta < 1 \), then
  \[ \sqrt{n} m(\sigma) \Rightarrow N(0, \beta^2) \]

• The model has a phase transition at \( \beta = 1 \). If \( \beta = 1 \), then
  \[ n^{1/4} m(\sigma) \Rightarrow \text{the distribution with density } \propto e^{-x^4/12}. \]

**Sherrington-Kirkpatrick Model**

Spin glass model for \( N \) spins.

\[ P(\sigma) = \frac{1}{Z} \exp \left( \frac{\beta}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{ij} \sigma_i \sigma_j + h \sum \sigma_i \right) \]

where \( g_{ij} \) is a fixed realization of iid \( N(0,1) \). The idea is that some particles try to align in the same direction, and some repel each other.

We will prove various results about this model using Stein’s method.

If \( h = 0 \), it is known that \( \beta = 1 \) is the critical temperature.

Overlap: generate two vectors, \( \sigma^1 \) and \( \sigma^2 \) independently from the Gibbs measure.

\[ R_{1,2} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^1 \sigma_i^2 \]

It is known that \( R_{1,2} = O\left( \frac{1}{\sqrt{N}} \right) \) if \( \beta < 1 \).

Open question: What is the magnitude of \( R_{1,2} \) at \( \beta = 1 \)?