Tracy-Widom Distribution and its Applications

April 29, 2014

1 Introduction

Central Limit Theorem : If $X_1, X_2, \ldots$ are i.i.d. mean zero, finite variance random variables, setting $S_n = \sum_{i=1}^{n} X_i$, we have

$$\frac{S_n}{\sqrt{Var(S_n)}} \rightarrow_d \mathcal{N}(0,1) \qquad \text{as} \quad n \rightarrow \infty.$$

Several variations when X_i's are dependent given since DeMoivre's work, simplest case being the local dependence.

Example : (Local maxima) Let $\mathbf{Z} = (Z_1, ..., Z_n)$ be i.i.d. $U(0,1)$ random variables and let S_n be the number of local maxima in $\mathbf{Z}$. Then $S_n/\sqrt{Var(S_n)} \rightarrow_d \mathcal{N}(0,1)$ as $n \rightarrow \infty$.

CLT is so strong due to its universality and a natural question is whether there are other laws that own the same property. Tracy-Widom distribution and its variants form such a class and they appear in several statistics of random permutations, queue theory, random matrices, percolation, and many others.

Here is one example showing what to expect.

Example : (Longest increasing subsequences) Let $\mathbf{Z} = (Z_1, ..., Z_n)$ be i.i.d. $U(0,1)$ random variables and let S_n be the length of the longest increasing subsequence of $\mathbf{Z}$. Then

$$\frac{S_n - 2\sqrt{n}}{n^{1/6}} \longrightarrow_d TW \qquad \text{as} \quad n \rightarrow \infty.$$

where TW is the Tracy-Widom distribution.

2 Tracy-Widom distribution and basic properties

2.1 Definition

Let $u(x)$ be the solution of the Painlevé II equation (See Appendix for some info)

$$u_{xx} = 2u^3 + xu \qquad \text{with} \qquad u(x) \sim -Ai(x) \quad \text{as} \quad x \rightarrow \infty.$$

We define the Tracy-Widom distribution by giving its cumulative distribution function as

$$F(t) = \exp\left(- \int_t^\infty (x-t) u^2(x) dx \right).$$

Discussion : To check that F really defines a CDF, we need to verify (i) $F(t) \to 1$ as $t \to \infty$, (ii) $F(t) \to 0$ as $t \to -\infty$ and (iii) $F'(t) \geq 0$.

These require the following two asymptotic results for the solution of Painlevé equation:

$$u(x) = -Ai(x) + O\left(\frac{e^{-(4/3)x^{3/2}}}{x^{1/4}} \right) \qquad \text{as} \quad x \to \infty$$

and

$$u(x) = -\sqrt{\frac{-x}{2}} \left(1 + O\left(\frac{1}{x^2} \right) \right) \qquad \text{as} \quad x \to -\infty.$$

Here is an informal verification of these results (Two occurrences of C is not necessarily the same).

(i) It is enough to show $\int_t^\infty (x-t) u^2(x) dx \to 0$ as $t \to \infty$. Recall that

$$u(x) \sim -Ai(x) \quad \text{and} \quad Ai(x) \sim \frac{e^{-(2/3)x^{3/2}}}{2\sqrt{2\pi} x^{1/4}} \qquad \text{as} \quad x \to \infty.$$

Thus, for large t

$$\int_t^\infty (x-t) u^2(x) dx \leq \int_t^\infty (x-t) C Ai^2(x) dx \leq \int_t^\infty (x-t) C \frac{e^{-(4/3)x^{3/2}}}{x^{1/2}} dx \leq C \int_t^\infty e^{-(4/3)x} dx \sim 0.$$

(ii) It is enough to show $\int_t^\infty (x-t) u^2(x) dx \to \infty$ as $t \to -\infty$. For $t < 0$ and 'small', we have

$$\begin{aligned} \int_t^\infty (x-t) u^2(x) dx \geq \int_t^{t/2} (x-t) u^2(x) dx \geq \int_t^{t/2} C(x-t) \frac{-x}{2} dx \quad &\geq \quad \int_{2t/3}^{t/2} (x-t) C \frac{-x}{2} dx \\ &\geq \quad \int_{2t/3}^{t/2} \frac{-t}{3} C \frac{-x}{2} dx \sim \infty. \end{aligned}$$

(iii) Since $F(t) = \exp\left(- \int_t^\infty x u^2(x) dx + t \int_t^\infty u^2(x) dx\right)$, we have

$$F'(t) = \exp(\ldots) \exp\left(\int_t^\infty u^2(x) dx \right) > 0$$

where we do some elementary manipulations for the derivative.

Remark : Tracy-Widom distribution has several alternative representations. For example, one may define it in terms of a function of Brownian motion or as a Fredholm determinant.

2.2 Two related distributions

The following cumulative distribution functions are the two other ones related to the Tracy-Widom distribution that also appear often in applications (some of which will be mentioned later):

- $F_1(t) = \exp\left(-\frac{1}{2}\int_t^\infty u(x)dx\right)(F_2(t))^{1/2}$ (known as 'orthogonal cdf').
- $F_4(t/\sqrt{2}) = \cosh\left(\frac{1}{2}\int_t^\infty u(x)dx\right)(F_2(t))^{1/2}$ (known as 'symplectic cfd').

These can be considered as special cases of Dyson's β distributions.

Here is a table with the means and variances of the three distributions we discussed:

β	μ_β	σ_β
1	-1.2	1.26
2	-1.77	0.9
4	-2.3	0.72

The following figure is borrowed from [7]:

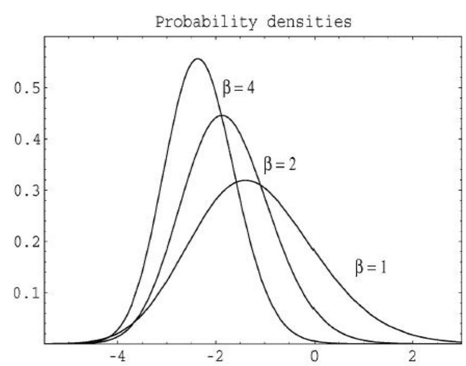

Figure 1: Density functions of Tracy-Widom distributions [7].

2.3 Tails of the Tracy-Widom distribution

Asymptotics for the tails of the Tracy-Widom distribution are well studied. Let

$$F(t) = \exp\left(-\frac{1}{2}\int_t^\infty (x-t)u^2(x)dx\right).$$

In particular, $F_2(t) = (F(t))^2$.

Then as $t \to \infty$,

$$F(t) = 1 - \frac{e^{-(4/3)x^{3/2}}}{32\pi x^{3/2}} \left(1 + O\left(\frac{1}{x^{3/2}}\right)\right).$$

The left tail had been a harder problem and been a conjecture for some time. Very recently, it has been shown that as $t \to -\infty$, we have

$$F_2(t) = \tau_2 \frac{e^{-|t|^3/12}}{|t|^{1/8}} \left(1 + \frac{3}{2^6|t|^3} + O(|t|^{-6})\right)$$

where $\tau_2 = 2^{1/2}e^{\zeta'(-1)}$ and $\zeta'(-1) = -0.165\ldots$ is the derivative of the Riemann zeta function at -1.

3 Applications on random matrices

Tracy-Widom distribution was first introduced in the context of random matrices.

3.1 Random matrix models

Real Wigner matrix : Let $X_{i,j}$ be real i.i.d. random variables with mean 0, variance 1 for $1 \leq i < j < \infty$. Set $X_{j,i} = X_{ij}$. Also let $X_{i,i}$ be real i.i.d. random variables with mean 0 and variance 1, but possibly with different distribution. Then the matrix $M = (X_{i,j})_{i,j=1}^n$ is called a *Wigner matrix*.

Gaussian orhogonal ensemble : A real Wigner matrix where $X_{i,j} \sim \mathcal{N}(0,1)$ for $i \neq j$ and $X_{i,i} \sim \sqrt{2}\mathcal{N}(0,1)$ is called a Gaussian orthogonal matrix (ensemble).

An equivalent description can be given by considering the distribution of $(A_n + A_n^t)/2$ where $A_n = (a_{i,j})_{i,j=1}^n$ and $a_{i,j}$'s are i.i.d. $\mathcal{N}(0,1)$.

Density of joint eigenvalues of a GOE is given by

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_1}\Pi_{i<j}|\lambda_j - \lambda_i|e^{-\frac{1}{4}\sum_{i=1}^n \lambda_i^2}.$$

Gaussian unitary ensemble : This is the complex version of GOE where $X_{i,i} \sim \mathcal{N}(0,1)$ and $X_{i,j} \sim \mathcal{N}(0,1/2) + i\mathcal{N}(0,1/2)$ for $i \neq j$.

Density of joint eigenvalues of a GUE is given by

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_2}\Pi_{i<j}|\lambda_j - \lambda_i|^2 e^{-\frac{1}{2}\sum_{i=1}^n \lambda_i^2}.$$

Remark : Note the similarity between densities. These are both special cases of Dyson's $\beta-$ensemble density

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_\beta}\Pi_{i<j}|\lambda_j - \lambda_i|^\beta e^{-\frac{\beta}{4}\sum_{i=1}^n \lambda_i^2}.$$

Other than the two special cases mentioned above, $\beta = 4$ is also seen often in applications – We call the matrix family as the Gaussian symplectic ensemble in that case.

3.2 Tracy-Widom in eigenvalue statistics

Theorem of Tracy-Widom : Let M a GUE matrix and λ_{max} be its largest eigenvalue. Then

$$\frac{\lambda_{max} - 2\sqrt{n}}{n^{1/6}} \longrightarrow_d F_2.$$

Remark : (1) Same result holds for GOE and GSE with F_1 and F_4, respectively.

(2) There exist Painleve type representations for distributions of second largest, etc. eigenvalues of random variables.

3.3 Universality comment of Tracy

An immediate question is whether we can relax the Gaussian assumption on the random matrices. The following discussion on universality is quoted from [7]:

"A natural question is to ask whether the above limit laws depend upon the underlying Gaussian assumption on the probability measure. To investigate this for unitarily invariant measures ($\beta = 2$) one uses a density of the form

$$c_n \exp(-tr(V(A)))dA$$

where $dA = \Pi_i dA_{ii} \Pi_{i<j} d_{A_{ij}}$. Bleher and Its choose

$$V(A) = gA^4 - A^2,\ g > 0,$$

and subsequently a large class of potentials V was analyzed by Deift et al.. These analyses require proving new Plancherel-Rotach type formulas for nonclassical orthogonal polynomials. The proofs use Riemann-Hilbert methods. It was shown that the generic behavior is GUE; and hence, the limit law for the largest eigenvalue is F_2. However, by finely tuning the potential new universality classes will emerge at the edge of the spectrum. For $\beta = 1, 4$ a universality theorem was proved by Stojanovic for the quartic potential. In the case of noninvariant measures, Soshnikov proved that for real symmetric Wigner matrices (complex hermitian Wigner matrices) the limiting distribution of the largest eigenvalue is F_1 (respectively, F_2). The significance of this result is that nongaussian Wigner measures lie outside the "integrable class" (e.g. there are no Fredholm determinant representations for the distribution functions) yet the limit laws are the same as in the integrable cases."

4 Other applications

Longest increasing subsequences : Letting L be the length of the longest increasing subsequence of a uniformly random permutation, scaled and centered version of L converges to Tracy-Widom distribution. This is a fundamental result due to Baik, Deift and Johannson.

Largest principal component : The largest principal component of the covariance matrix $X^t X$ where X is an $n \times p$ data matrix all of whose entries are independent standard normal has F_1 as the limiting distribution after proper centering and scaling.

Queuing theory : Consider a series of n single server queues each with unlimited waiting space with a first-in first-out service. Service times are i.i.d. with distribution V, with mean 1 and variance σ^2. Quantity of interest is $D(k,n)$ which is the departure time of customer k from the last queue. For a fixed number of customers, k, it was first shown that $(D(k,n) - n)/(\sigma\sqrt{n})$ converges in distribution to a functional $\hat{D}_k$ of a $k-$dimensional Brownian motion. Later it was realized that $\hat{D}_k$ had the same distribution as the largest eigenvalue of a $k \times k$ GUE matrix.
More recently, Johannson proves when V is Poisson,

$$\frac{D(\lfloor xn \rfloor, n) - c_1 n}{c_2 n^{1/3}} \longrightarrow_d F_2(t)$$

as $n \to \infty$ for some c_1, c_2 depending on x.

Yet others :

- Growth processes
- Random tilings
- Superconductors

Appendix : Painleve equation

Consider the integral equation

$$K(x,\xi) = kAi\left(\frac{x+\xi}{2}\right) + \frac{k^2}{4}\int_x^\infty \int_x^\infty K(x,s)Ai\left(\frac{s+t}{2}\right)Ai\left(\frac{t+\xi}{2}\right)dsdt.$$

Then $u(x) = K(x,x)$ satisfies $u_{xx} = 2u^3 + xu$ which is the Painleve II equation, with the boundary condition

$$u(x) \sim kAi(x), \qquad \text{as} \quad x \to \infty.$$

Details about the asymptotics of the solutions of Painleve equations can be found in [3]. This reference also gives several historical notes, applications and various details about Painleve equations.

References

[1] Edelman, A. and Rao, N. R., *Random matrix theory*, Acta Numerica, pp. 1-65, 2005.

[2] Baik, J., Deift, P. and Johansson, K., *On the distribution of the length of the longest increasing subsequence of random permutations*, Journal of the American Mathematical Society 12 (4): 1119-1178, 1999.

[3] Clarkson, P. A., *Painlevé equations-nonlinear special functions, Orthogonal polynomials and special functions*, Vol. 1883 of Lecture Notes in Math., Springer, Berlin, pp. 331411, 2006.

[4] Johnstone, I. M., *Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy-Widom limits and rates of convergence*, Annals of Statistics 36 (6): 2638-2716, arXiv:0803.3408, 2008.

[5] Tracy, C. A. and Widom, H., *Level-spacing distributions and the Airy kernel*, Physics Letters B 305 (1-2): 115-118, 1993.

[6] Tracy, C. A. and Widom, H., *Level-spacing distributions and the Airy kernel*, Communications in Mathematical Physics 159 (1): 151-174, 1994.

[7] Tracy, C. A. and Widom, H., *Distribution Functions for Largest Eigenvalues and Their Applications*, ICM 2012, Vol. I, 587-596, 2012.