Lecture 2: Goal today: 1) derive the spacing distribution for $N \gg 1$ identically and independently distributed (iid) rv to compare with the Wigner surmise

2) derive the claimed result $P(s)$ from $N=2$ GOE matrices ($\Rightarrow$ jpdf too)

general Q: 1) Why do we choose a Gauss' distribution? (simplicity)

2) Can we choose a more general ensemble, and do we get the same answer, at least asymptotically for $N \to \infty$? (This is called universality!)

1) Only the choice of Gauss and iid leads to an invariant ensemble $g(H_s) \sim e^{-\mathrm{Tr}(H_s)^2}$

2) $\exists$ (at least) 2 ways of generalising (a "classification" comes next week)

Wigner ensembles: all H_{ij} are iid, in general jpdf not known, only probabilistic methods available, not our main focus

Invariant ensembles $\mathrm{Tr}(H_s^2) \to \mathrm{Tr}\, V(H_s)$ e.g. $V(x)$ polynomial, confining

$\to$ Jacobian in jpdf (1.15) unchanged

* for both types of ensembles universality results exist!

2.1 spacing distribution for independent random variables

(maybe show Mehta fig 1.2)

- it turns out that here we can already make a universal statement about their spacing distribution

$x_1, \dots, x_N$ iid rv $\Rightarrow$ $\boxed{p(s) = e^{-s}}$ for $N \gg 1$ Poisson distribution

without specifying $p_X(x_i)$ further!

Derivation of the Poisson distribution:

$\{X_1, \dots, X_N\}$ iid r.v. with pdf $p_X(x)$ on support σ, with cumulative distribution function (cdf) $F(x)$; $\Rightarrow F'(x) = p_X(x)$

Step 1: compute the conditional probability $p_N(s \mid X_j = x)$:

$\exists!\, j$ s.th. $X_j = x$, $\exists\, k \neq j$ s.th. $X_k = x+s$, and all other X_ℓ are less than x or larger than $x+s$. (so 2 consecutive r.v. have spacing s). It holds:

$$p_N(s \mid X_j = x) = p_X(x+s)\left[F(x) + 1 - F(x+s)\right]^{N-2}$$

, does not depend on j!

as: • 1 rv is at x, $N-1$ left.

• another rv is at $x+s$, with probab $p_X(x+s)$

• the remaining $N-2$ rv are indep., an for each of them they are with probab $F(x)$ less than x and with probability $1-F(x+s)$ larger than $x+s$.

$\Rightarrow$ total probability, irrespective of which X_j is at x reads

$$p_N(s \mid \text{any } X = x) = \sum_{j=1}^{N} p_N(s \mid X_j = x)\, \text{Prob}(X_j = x) = p_N(s \mid X_j = x) \underbrace{\sum_{j=1}^{N} \text{Prob}(X_j = x)}_{= N p_X(x)}$$

$$= N\, p_N(s \mid X_j = x)\, p_X(x)$$

$\Rightarrow$ probability that gap s is between any 2 adjacent rv, no longer conditioned on the position of one variable (i.e. X_j) is given by integration over σ:

$$\boxed{p_N(s) = \int_\sigma dx\, p_N(s \mid \text{any } X = x) = N \int_\sigma dx\, p_N(s \mid X_j = x)\, p_X(x)} \quad (2.8)$$

Exercise: check that this spacing distribution is normalized $\int_0^\infty ds\, p_N(s) = 1$

Step 2: large-N limit

- local change of variables $s \to \hat{s}$, with $\boxed{s = \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)}}$, $\hat{s} = O(1)$ when $N \to \infty$

why: the average spacing is $\sim \frac{1}{N}$ and $\sim \frac{1}{\rho_{\mathbb{X}}(x)}$

as for increasing N the same # of rv occupies the same space, with probab $\sim$ "height" $\rho_{\mathbb{X}}(x)$

$$\Rightarrow \rho_N\left(s = \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)} \,\middle|\, X_j = x\right) = \rho_{\mathbb{X}}\left(x + \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)}\right)\left[1 - \underbrace{\left(F\left(x + \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)}\right) - F(x)\right)}_{\text{Taylor} = \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)} \cdot F'(x) + O\left(\frac{1}{N^2}\right),\ F'(x) = \rho_{\mathbb{X}}(x)}\right]^{N-2}$$

So for $N \gg 1$

$$\sim \rho_{\mathbb{X}}(x) = \rho_{\mathbb{X}}(x)\left[1 - \frac{\hat{s}}{N}\right]^{N-2} \to \rho_{\mathbb{X}}(x)\, e^{-\hat{s}}$$

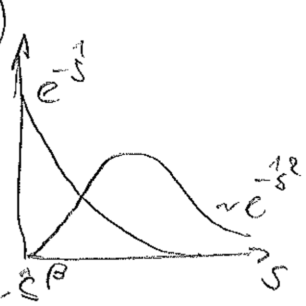
$\Rightarrow$ limiting spacing distribution $\hat{\rho}(\hat{s})$:

$$\boxed{\hat{\rho}(\hat{s}) = \lim_{N\to\infty} \rho_N\left(s = \frac{\hat{s}}{N\rho_{\mathbb{X}}(x)}\right)\frac{ds}{d\hat{s}} = \lim_{N\to\infty} \frac{\cancel{1}}{\cancel{N}} \underbrace{\int dx\,\rho_{\mathbb{X}}(x)}_{1}\, e^{-\hat{s}}\cancel{N} = e^{-\hat{s}}}$$

which is independent of $\rho_{\mathbb{X}}(x)$, the pdf of our iid rv, $\Rightarrow$ universal!

- obviously it is normalised on $\mathbb{R}_+$ and has 1st moment $= 1$

($\to$ exercise) (this is why we imposed this condition on the local spacing distribution.)



$\ast$ comparison with the RMT spacing distribution

Wigner surmise $\boxed{\rho_\beta(\hat{s}) = a_\beta\, e^{-b_\beta \hat{s}^2}\, \hat{s}^\beta}$

eigenvalues repell each other, due to the log-interaction to power β

whereas iid random variables have a finite probability to

have spacing $\hat{s} = 0$, no gap between consecutive variables.

2.2. spacing distribution and jpdf for 2x2 GOE matrices

* this calculation is an illustration only, this will not be the right way to proceed for general N
* we will see why different variances for diagonal and offdiagonal H_{ij} matter

2x2 GOE matrix $H_s = \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix}$, $x_1, x_2 \in \mathcal{N}(0,1)$, $x_3 \sim \mathcal{N}(0,\frac{1}{2})$

What are the eigenvalues of H_s?

$$0 = \det(\lambda \mathbb{1}_2 - H_s) = \begin{vmatrix} \lambda - x_1 & -x_3 \\ -x_3 & \lambda - x_2 \end{vmatrix} = (\lambda - x_1)(\lambda - x_2) - x_3^2$$

$$= \lambda^2 - (x_1+x_2)\lambda + x_1 x_2 - x_3^2 \quad \left(= \lambda^2 - \mathrm{Tr}(H_s)\lambda + \det(H_s) \right)$$

so $\lambda_{1,2} = \frac{(x_1+x_2)}{2} \pm \sqrt{\frac{(x_1+x_2)^2}{4} - x_1 x_2 + x_3^2} = \frac{1}{2}\left(x_1 + x_2 \pm \sqrt{\underbrace{(x_1-x_2)^2 + 4x_3^2}_{>0}} \right)$

where we chose $\lambda_1 > \lambda_2$

$\Rightarrow |\lambda_1 - \lambda_2| = \lambda_1 - \lambda_2 = \sqrt{(x_1-x_2)^2 + 4x_3^2} \equiv S'$ spacing

How do we obtain the jpdf $\rho(\lambda_1, \lambda_2)$ and the spacing distrib $p(s)$?

- $$\rho(\lambda_1,\lambda_2) = \int_{\mathbb{R}^3} dx_1 dx_2 dx_3 \frac{e^{-\frac{1}{2}(x_1^2+x_2^2) - x_3^2}}{\sqrt{2\pi}^2 \sqrt{\pi}} \delta\left(\lambda_1 - \left(\frac{x_1+x_2+S}{2}\right)\right) \delta\left(\lambda_2 - \left(\frac{x_1+x_2-S}{2}\right)\right)$$

$\Rightarrow$ jpdf (2.15) for $\beta = 1$ with ordered ev $\lambda_1 > \lambda_2$ (Exercise)

- $$p(s) = \int_{\mathbb{R}^3} dx_1 dx_2 dx_3 \frac{e^{-\frac{1}{2}(x_1^2+x_2^2) - x_3^2}}{2\pi\sqrt{\pi}} \delta(s - S')$$

change variables:

$x_1 - x_2 = r\cos\theta$, $\quad r \in \mathbb{R}_+$

$2x_3 = r\sin\theta$, $\quad \theta \in [0, 2\pi]$

$x_1 + x_2 = \psi$, $\quad \psi \in \mathbb{R}$

$\Leftrightarrow \begin{cases} x_1 = \frac{1}{2}(\psi + r\cos\theta) \\ x_2 = \frac{1}{2}(\psi - r\cos\theta) \\ x_3 = \frac{1}{2} r\sin\theta \end{cases}$

$$\Rightarrow \text{Jacobian} = \det\left(\frac{\partial x_1, x_2, x_3}{\partial(r,\theta,\varphi)}\right) = \begin{vmatrix} \frac{\cos\theta}{2} & -\frac{r\sin\theta}{2} & \frac{1}{2} \\ -\frac{\cos\theta}{2} & \frac{r\sin\theta}{2} & \frac{1}{2} \\ \frac{\sin\theta}{2} & \frac{r\cos\theta}{2} & 0 \end{vmatrix} = -\frac{r}{4}$$

$$\Rightarrow p(s) = \frac{1}{2\pi^{\frac{3}{2}}} \int_0^\infty dr \int_0^{2\pi} d\theta \int_{-\infty}^{\infty} d\varphi \; e^{-\frac{1}{2}\left(\frac{(\varphi + r\cos\theta)^2}{4} + \frac{(\varphi - r\cos\theta)^2}{4} + \frac{r^2\sin^2\theta}{4}\right)} \left|\frac{r}{4}\right| \delta(s-r)$$

as $s = \sqrt{(x_1-x_2)^2 + 4x_3^2} = \sqrt{r^2\cos^2\theta + r^2\sin^2\theta} = r$

with the exponent reading $-\frac{1}{4}(\varphi^2 + r^2)$

we finally obtain $p(s) = \frac{1}{8\pi^{\frac{3}{2}}} \int_0^\infty dr\, r\, e^{-\frac{r^2}{4}} \delta(s-r) \underbrace{\int_0^{2\pi} d\theta}_{2\pi} \underbrace{\int_{-\infty}^{\infty} d\varphi\, e^{-\frac{1}{4}\varphi^2}}_{\sqrt{4\pi}}$

$$= \frac{s}{2} e^{-\frac{s^2}{4}}$$

as was claimed on the basis of jpdf (2.15) for $N=2$ and exercise 1.2.

* it turns out that the result based on 2×2 matrices is an extremely good approximation for $N\times N$ matrices when $N \to \infty$!