Gaussian Kinematic Formula and the integral geometry of random sets

Jonathan Taylor
Stanford University

November 11, 2011

Jonathan Taylor
Stanford University

Outline

- A model for random sets.
- Some *old* integral geometry.
- Gaussian integral geometry.
- Accuracy of approximation.

Random sets

Gaussian processes: basic building blocks

Twice-differentiable Gaussian process $(f_t)_{t \in M}$ on a manifold M.

Satisfying:

- $\mathbb{E}\{f_t\} = 0$;
- $\mathbb{E}\{f_t^2\} = 1$.

Why Gaussian?

- Gaussian processes are specified by mean and covariance function.
- Finite-dimensional distributions are all simple – multivariate Gaussian.

Random sets

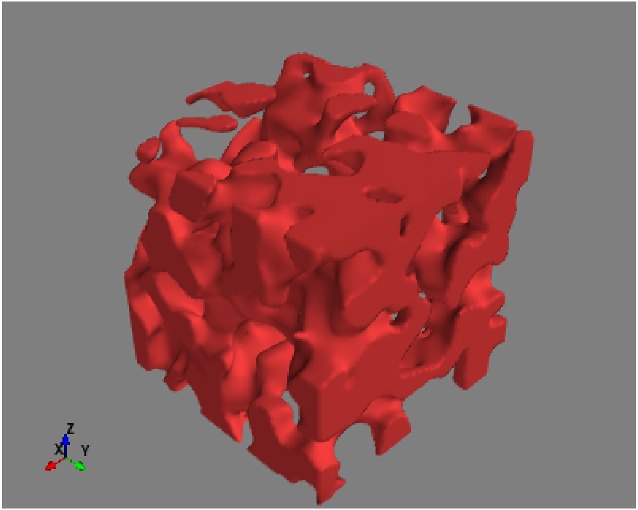
Excursion sets

Random sets we will consider are of the form:

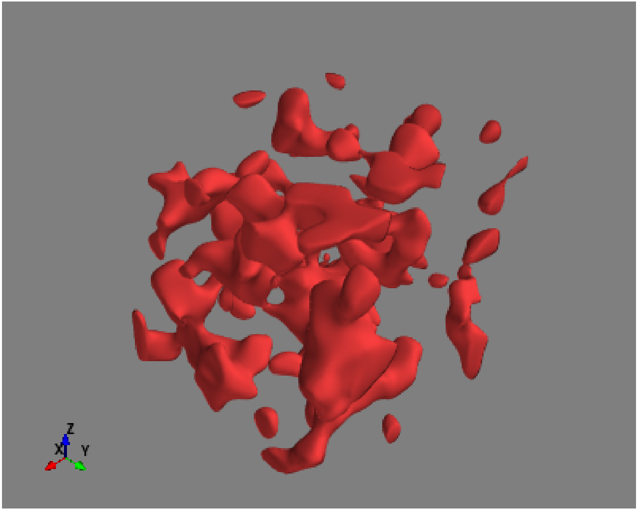
$$f^{-1}A = \{t \in M : f_t \in A\}$$

for $A \subset \mathbb{R}$. In particular, the *geometry* of these sets, and how it is determined by correlation function of f.

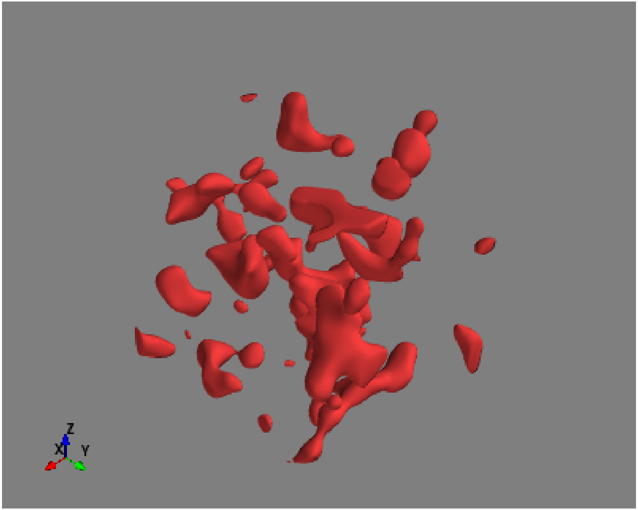
Excursion above 0



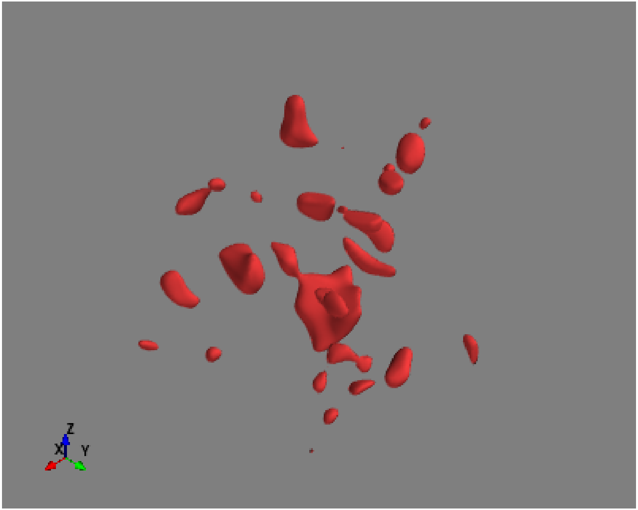
Excursion above 1



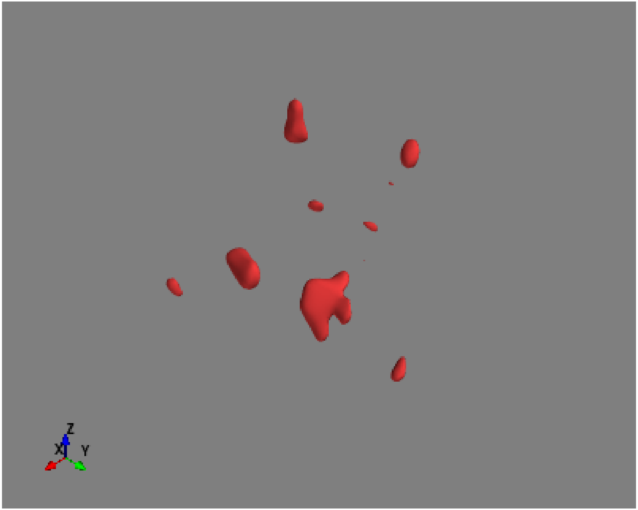
Excursion above 1.5



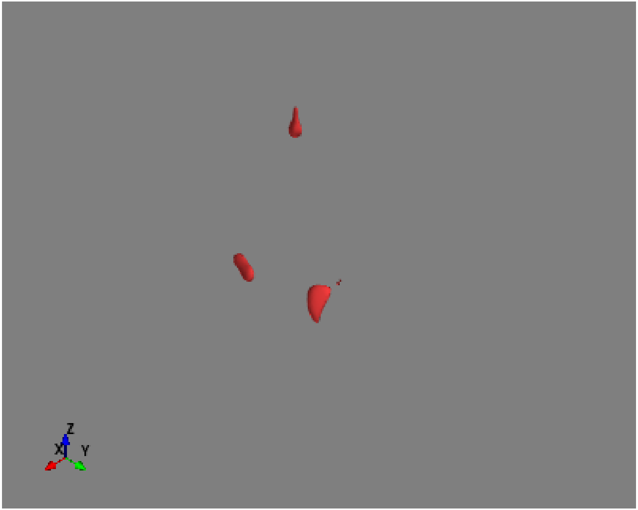
Excursion above 2



Excursion above 2.5



Excursion above 3



Excursion above 3.3

Euler characteristic

What geometric features?

We're interested in integral geometric properties.
Specifically, the (expected) Euler characteristic

$$\mathbb{E}\left\{\chi\left(f^{-1}[u,+\infty)\right)\right\} = \mathbb{E}\left\{\chi\left(M \cap f^{-1}[u,+\infty)\right)\right\}$$

EC tells you very little

Let M be a 2-manifold without boundary, then

$$\mathbb{E}\left\{\chi\left(M \cap f^{-1}[0,+\infty)\right)\right\} = \frac{1}{2} \cdot \chi(M)$$

With boundary:

$$\mathbb{E}\left\{\chi\left(M \cap f^{-1}[0,+\infty)\right)\right\} = \frac{1}{2} \cdot \chi(M) + \frac{1}{2\pi} \cdot |\partial M|$$

Euler characteristic

EC is computable

Of all quantities in the studies of Gaussian processes, the EC stands out as being *explicitly computable* in wide generality

$$\mathbb{E}\left\{\chi\left(M \cap f^{-1}[u,+\infty)\right)\right\} = \sum_{j=0}^{\dim(M)} \mathcal{L}_j(M)\rho_j(u).$$

Where do $\mathcal{L}_j$'s and ρ_j's come from?

EC tells you a lot

For "nice enough" M

$$\mathbb{E}\left\{\chi\left(M \cap f^{-1}[u,+\infty)\right)\right\} \overset{u\to\infty}{\simeq} \mathbb{P}\left\{\sup_{t\in M} f_t \geq u\right\}$$

Euler characteristic

Accuracy

- If the parameter set is *locally convex*, then

$$\left| \mathbb{P}\left\{ \sup_{t \in M} f_t \geq u \right\} - \mathbb{E}\left\{ \chi\left(M \cap f^{-1}[u, +\infty) \right) \right\} \right| \overset{u \to \infty}{=} O_{\exp}\left(e^{-u^2/2 \cdot (1 + \frac{1}{\sigma_c^2(f)})} \right)$$

- Since, $\rho_j(u) = O_{\exp}(e^{-u^2/2})$, the approximation has *exponential* relative accuracy!

Integral geometry

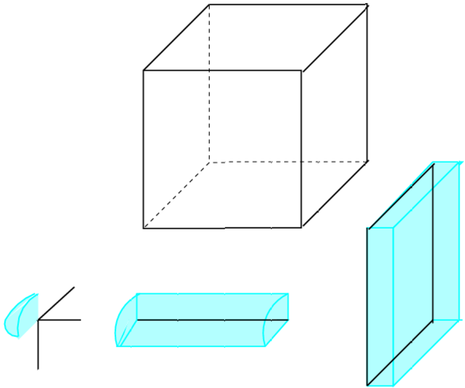
Jonathan Taylor
Stanford University

Tube formulae

- Suppose f is (the restriction to M) of an isotropic process $\tilde{f}$, with $\text{Var}\left\{d\tilde{f}/dx_i\right\} = 1$.
- For small r, the functionals $\mathcal{L}_j(M)$ are implicitly defined by Steiner-Weyl formula

$$\mathcal{H}_k\left(x \in \mathbb{R}^k : d(x, M) \leq r\right) = \sum_{j=0}^{k} \omega_{k-j} r^{k-j} \mathcal{L}_j(M; M)$$

Integral geometry: tubes



$$\mathcal{H}_3\left(\text{Tube}([0,a]\times[0,b]\times[0,c],r)\right)$$

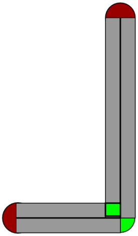
$$= abc + 2r\cdot(ab+bc+ac) + (\pi r^2)\cdot(a+b+c) + \frac{4\pi r^3}{3}$$

Integral geometry: tubes

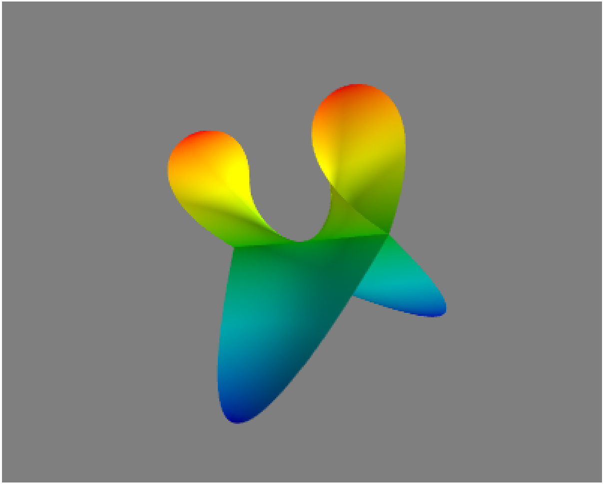
Integral geometry: tubes

Local convexity is important!

- Singularity means implicit definition is invalid, BUT $\mathcal{L}_j(\cdot)'s$ are still well defined ...
- They are defined for a large class of sets ...

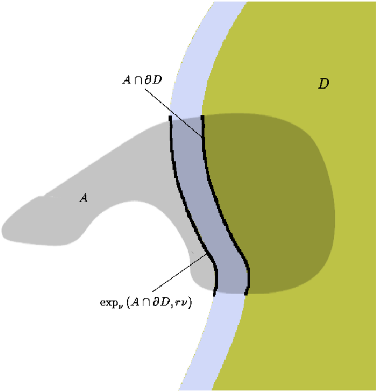


Gaussian processes



Integral geometry: curvature measures

Jonathan Taylor
Stanford University

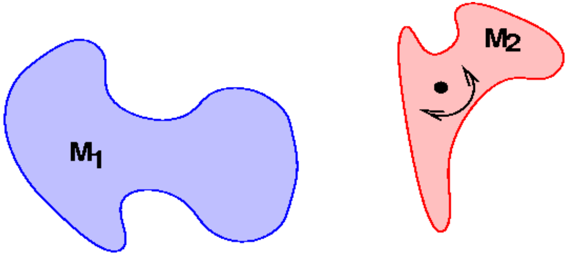


$$\mathcal{H}_{k-1}\left(\exp_\nu\left(A \cap \partial D, r\nu\right)\right) = \sum_{j=1}^{k} r^{j-1} \cdot \omega_j \int_A \mathcal{L}_{k-j}(D; dp)$$

Integral geometry

Kinematic Fundamental Formula

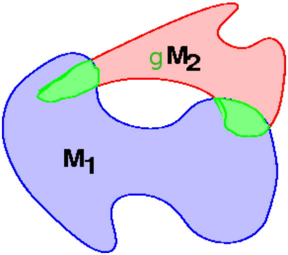
- Where else do we see $\mathcal{L}_j(\cdot)$'s?



Integral geometry

Kinematic Fundamental Formula

- Considers the "average" curvature measures of $M_1 \cap gM_2$, i.e.



Integral geometry: KFF

The KFF on $\mathbb{R}^N$

- Isometry group G_N of rigid motions of $\mathbb{R}^N$,

$$G_N \sim \mathbb{R}^N \rtimes O(N)$$

- Fix a Haar measure:

$$\nu_N\left(\{g_N \in G_N : g_N x \in A\}\right) = \mathcal{H}_N(A)$$

-

$$\int_{G_N} \mathcal{L}_i\left(M_1 \cap g_N M_2\right) \, d\nu_N(g_N) = \sum_{j=0}^{N-i} \begin{bmatrix} i+j \\ i \end{bmatrix} \begin{bmatrix} N \\ j \end{bmatrix}^{-1} \mathcal{L}_{i+j}(M_1)\mathcal{L}_{N-j}(M_2)$$

Back to Gaussian processes

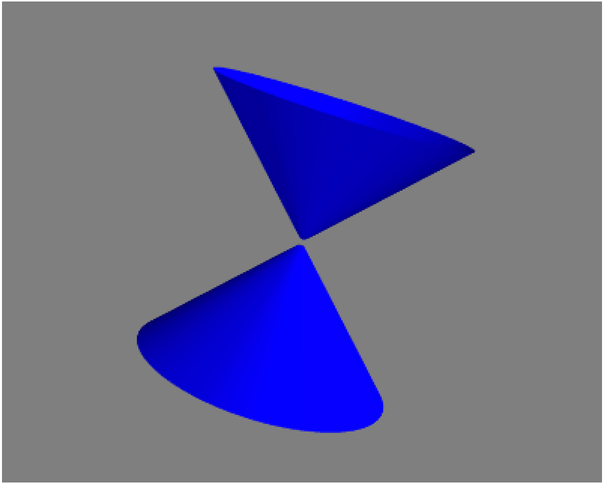
Analogy with KFF

- Recall what we were trying to study

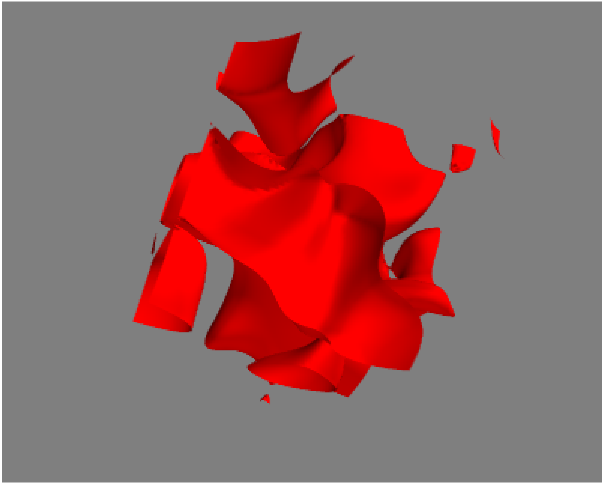
$$\begin{aligned}
&\mathbb{E}\left\{\chi\left(M \cap f^{-1}[u,+\infty)\right)\right\} \\
&\quad = \int_\Omega \mathcal{L}_0(M \cap f(\omega)^{-1}[u,+\infty))\, \mathbb{P}(d\omega) \\
&\quad = \sum_{j=0}^{\dim(M)} \mathcal{L}_j(M)\rho_j(u)
\end{aligned}$$

- This *looks like* KFF on $\mathbb{R}^N$ where $g_N M_2$ is replaced by $f^{-1}[u,+\infty) = f^{-1}D$.
- Can replace f with $f = (f_1, \dots, f_j)$. Let's look at some examples . . .

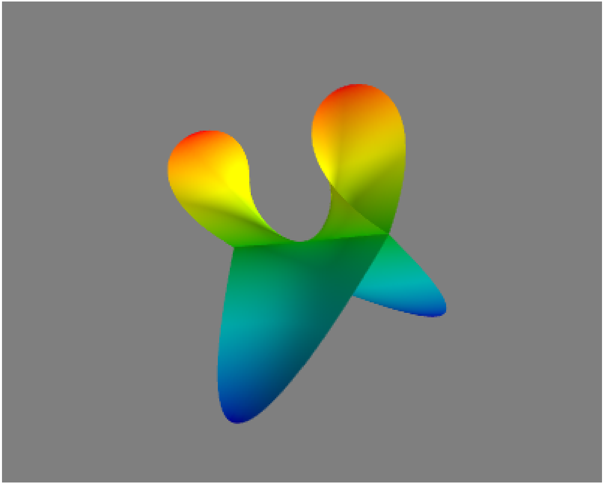
Gaussian processes: D a cone



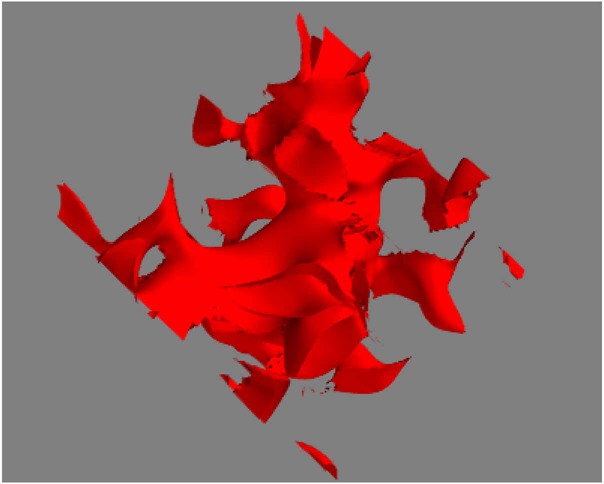
Gaussian processes: $f^{-1}D$



Gaussian processes: D a variety



Gaussian processes: $f^{-1}D$



Gaussian processes

Gaussian Kinematic Formula

- Let $f = (f_1, \dots, f_k)$ be made of IID copies of a Gaussian field
- Consider the additive functional on $\mathbb{R}^k$ that takes a "rejection region"

$$D \mapsto \mathbb{E}\left\{\chi\left(M \cap f^{-1}D\right)\right\}.$$

- **Questions:** how do the $\mathcal{L}_j$'s enter into this functional? How about D?

Gaussian processes

Gaussian Kinematic Formula

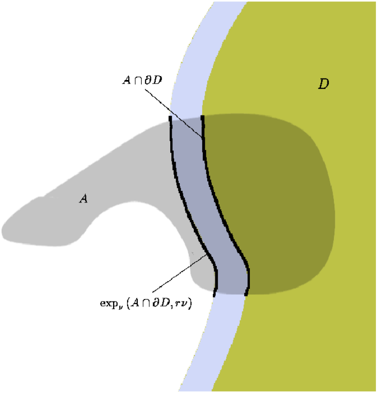
- Define the functionals $\mathcal{M}_j^{\gamma_k}(\cdot)$ implicitly by

$$\gamma_k\left(y \in \mathbb{R}^k : d(y, D) \leq r\right) = \sum_{j \geq 0} \frac{(2\pi)^{j/2} r^j}{j!} \mathcal{M}_j^{\gamma_k}(D).$$

- Then:

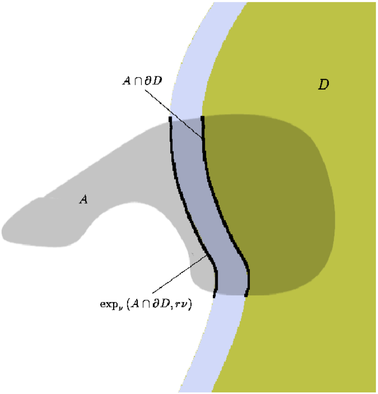
$$\mathbb{E}\left\{\chi\left(M \cap f^{-1} D\right)\right\} = \sum_j \mathcal{L}_j(M) \cdot \mathcal{M}_j^{\gamma_k}(D)$$

Integral geometry: curvature measures



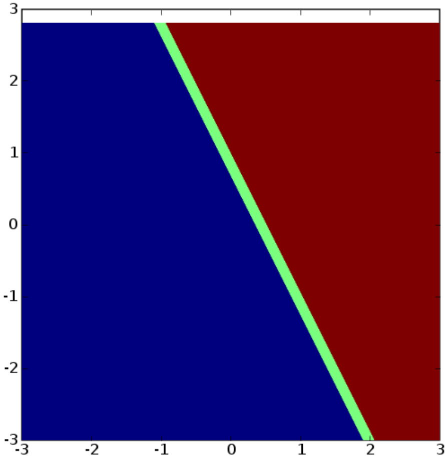
Q: How do we compute $\mathcal{M}_j^{\gamma_k}(\cdot)$?

Integral geometry: curvature measures

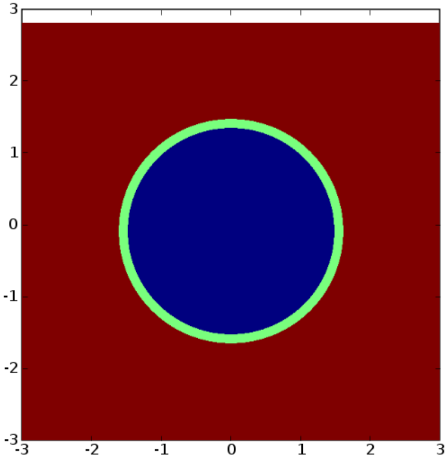


A: Integrate power series expansion for density over hypersurface at distance r ...

Example: linear statistic



Example: χ^2 statistic



Examples

EC densities

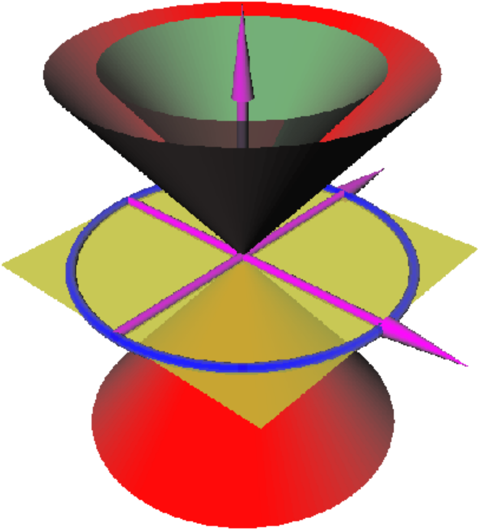
- Gaussian EC densities

$$\rho_j(u) = (-1)^j \frac{d^j}{du^j} \mathbb{P}\left\{N(0,1) > u\right\}$$

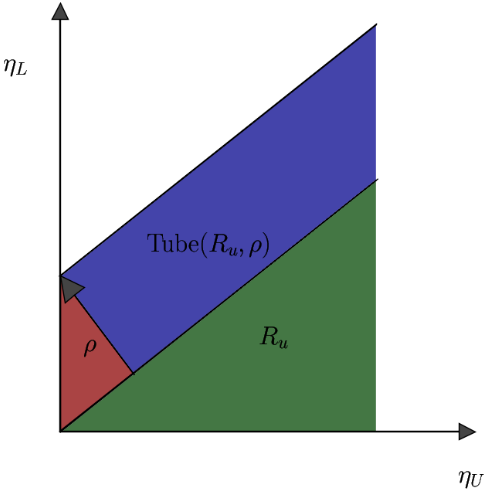
- χ^2_k EC densities

$$\rho_{j,\chi^2_k}(u) = (-1)^j \frac{d^j}{dx^j} \mathbb{P}\left\{\sqrt{\chi^2_k} > x\right\}\Bigg|_{x=\sqrt{u}}$$

t or F statistic



t or F statistic



Ideas behind the proof

Hidden embedding

- A Gaussian process is just a mapping

$$t \mapsto f_t \in L^2(\Omega, \mathcal{F}, \mathbb{P})$$

- Our assumptions about mean and variance implies the image is in the unit sphere in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, and it is an embedding.
- This suggests that the relevant "geometry" to prove this result is spherical.
- Short answer: *yes*.

Proof: spherical KFF

Spherical KFF

- For $\kappa \in \mathbb{R}$ define

$$\mathcal{L}^{\kappa}(\cdot) = \sum_{n=0}^{\infty} \frac{(-\kappa)^n}{(4\pi)^n} \frac{(i+2n)!}{n!\, i!} \mathcal{L}_{i+2n}(\cdot).$$

- For $M_1, M_2 \subset S_{n^{1/2}}(\mathbb{R}^n)$

$$\int_{G_n} \mathcal{L}_i^{n^{-1}} \left(M_1 \cap g_n M_2\right) \; d\nu_{n,\lambda}(g_n) = \sum_{j=0}^{n-1-i} \begin{bmatrix} i+j \\ i \end{bmatrix} \begin{bmatrix} n-1 \\ j \end{bmatrix}^{-1} \mathcal{L}_{i+j}^{n^{-1}}(M_1) \mathcal{L}_{n-1-j}^{n^{-1}}(M_2)$$

where $G_n = O(n)$, appropriately normalized.

Proof: Poincaré's limit

Model process

- For $M \subset S(\mathbb{R}^j)$ define a $\mathbb{R}^k$ valued process

$$f^n(t, g_n) = \pi_k(n^{1/2} g_n t)$$

  where $g_n \in O(n)$ is a Haar-distributed random matrix and $\pi_k : S_{n^{1/2}}(\mathbb{R}^n) \to \mathbb{R}^k$ is projection onto the first k coordinates.

- Poincaré's limit (and generalizations) ensures that the process $f^n = (f_1^n, \dots, f_k^n)$ converges in variation to a vector of IID zero mean, unit variance Gaussian processes $f = (f_1, \dots, f_k)$.

Proof: connecting Gaussian processes with KFF

Expected EC for model process

- For $D \subset \mathbb{R}^k$

$$\int_{G_n} \mathcal{L}_i^1 \left(M \cap (f^n)^{-1}(D)\right) \; d\nu_{n,\lambda}(g_n)$$

$$= n^{-i/2} \int_{G_n} \mathcal{L}_i^{n^{-1}} \left(n^{1/2} M \cap \pi_k^{-1} D\right) \; d\nu_{n,\lambda}(g_n)$$

$$= c_n \sum_{j=0}^{n-1-i} \begin{bmatrix} i+j \\ i \end{bmatrix} \begin{bmatrix} n-1 \\ j \end{bmatrix}^{-1} \mathcal{L}_{i+j}^1(M) \mathcal{L}_{n-1-j}^{n^{-1}}(\pi_k^{-1} D)$$

- The set $\pi_k^{-1} D$ is the disjoint union of a warped product and $D \cap S_{n^{1/2}}(\mathbb{R}^k)$. Need to analyse curvatures of warped product asymptotically.
- The rest is combinatorics … almost.