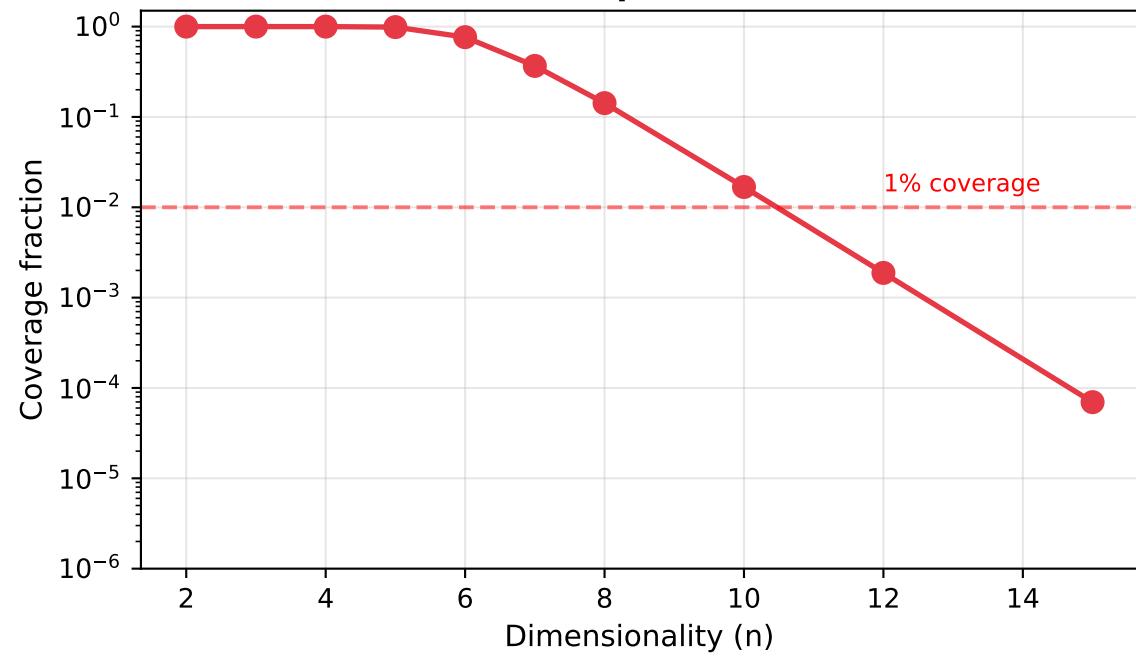
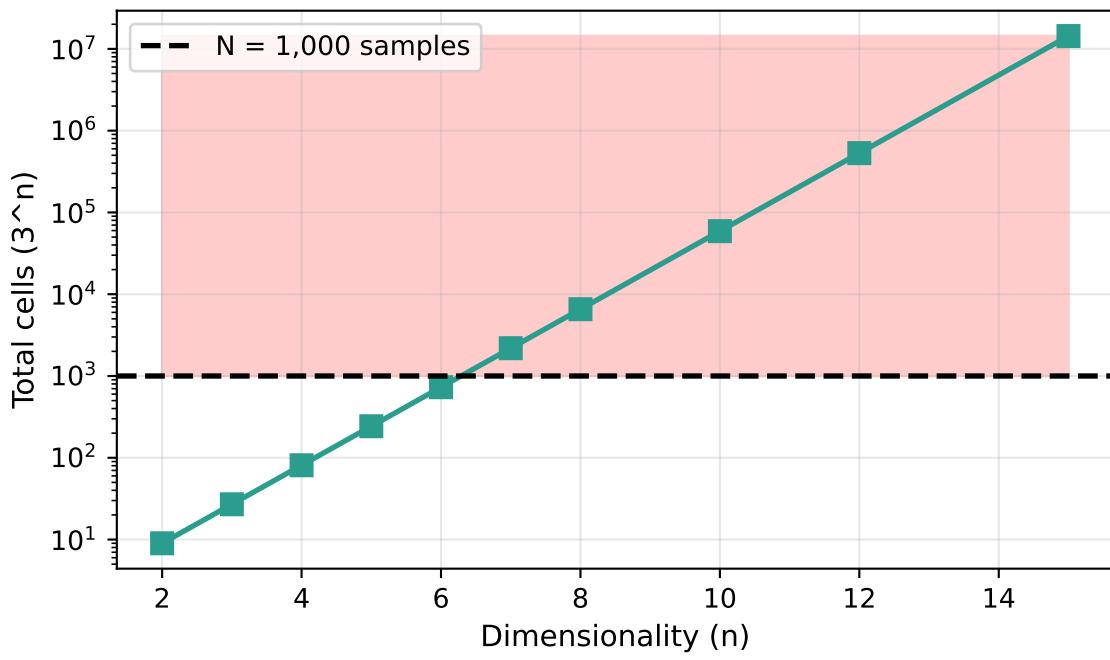


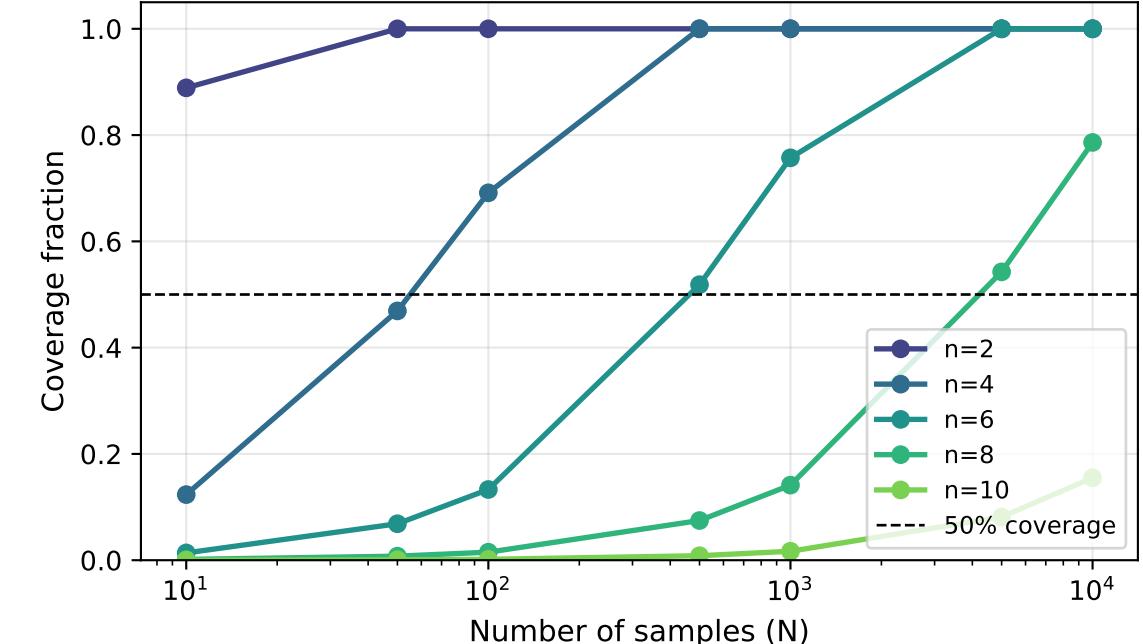
A. Space Coverage Collapse
(N=1,000 samples, 3 bins/dim)



B. Space Size Explosion
(cells = 3^n)



C. Coverage Curves by Dimension
(diminishing returns in high-D)



D. The Epistemological Point

THE CURSE OF DIMENSIONALITY

To sample a space with k bins per dimension:

$$\text{Total cells} = k^n$$

CONCRETE NUMBERS ($k=3$):

$n = 5$:	cells =	243	(doable)
$n = 10$:	cells =	59,049	(hard)
$n = 15$:	cells =	14,348,907	(impossible)
$n = 20$:	cells =	3.5 billion	(absurd)

THE BIOLOGICAL IMPLICATION:

Nonparametric inference requires "seeing" the relevant regions of state space.

But high-dimensional spaces are almost entirely empty—your samples cluster in a tiny fraction of the possible volume.

RARE EVENTS (disease subtypes, unusual cell states, outlier phenotypes) live in cells you will NEVER sample.

This is not "we need more data."
This is "the required data doesn't exist."

ESCAPE ROUTE:

Assume structure (parametric models) or find low-dimensional manifolds. But this requires knowing what you're looking for before you find it.