# 1 Introduction

The modern theory of information geometry is largely based on the work of Shun-ichi Amari. He refers to information geometry as as "method of exploring the world of information by means of modern geometry."

The basic idea is to study space of models, which are probability distributions, we use these to describe the state of a system we're interested in. In statistical inference, we have data and from this data we want to infer our best guess of what that data is being generated from.

Can think o this as a best approximation problem. From a geometric perspective, the notion of distance is important. In statistics this might be likelihood and in learning it might be loss function.