OVERVIEW We will go through a digest of six papers. I have broken abstracts or introductions into numbered points. I have emboldened **key terms** for **ease of reading** and I have underlined many **technical terms**. I will be collecting **feedback** from the you on **Zulip**.

INSTRUCTIONS

- If you haven't already done so, please **create** a **Zulip** account using the ***link*** on the meetup.com **meeting page**.
- We will **discuss** readings on the **Zulip** ***thread*** entitled **#learning: reading groups / Statistics reading group**.
- I will be **collecting materials** from you on the thread **#learning: reading groups / Statistics reading group—'Housekeeping'**
- I have numbered every point so that you can **quickly note** ***every*** **concept** that is **important** to you.
- For each paper, please **write** down the **author's name** followed by a **list** of **point numbers** that are most interesting you.
- For each paper, please **write at least** one sentence **relating** the **most compelling** points to your own **work** or **interests**
- Lastly, please pick the **two** papers you most want to **read right away**.
- Please **list** these two papers in the following fashion:

  'Top Papers: (1) AuthorName, (2) AuthorName'
- Submit all this information as a **post** on the thread **#learning: reading groups / Statistics reading group—'Housekeeping'**
- Next, **I will decide** which **paper to read** based on **your feedback**.
- Finally, we will **revisit** our chosen paper and **briefly** decide what **sections** to read.

Your feedback will help us chart a **roadmap** through the material. We will **chat** more on **Zulip** about that, as well as other topics such as **contriving exercises** for ourselves and **writing posts** on the **Applied Category Theory Wiki**.

What is a statistical model?
Peter McCullagh

1. This paper addresses two closely related questions,
2. **"What is a statistical model?"** and **"What is a parameter?"**
3. The notions that a model must **"make sense,"** and that a parameter must "have a well-defined meaning" are...well understood [**in *practice***], but ***absent*** from most **formal** theories...
4. In this paper, these concepts are defined in algebraic terms, using morphisms, functors and natural transformations.
5. It is argued that **inference** on the basis of a model is *not possible unless* the model admits a **natural extension** that includes the domain for which inference is required.
6. For **example**, prediction requires that the domain include all future units, subjects or time points.
7. Although it is usually not made explicit, **every sensible** statistical **model admits** such an **extension**.
8. Examples are given to show **why** such an extension is **necessary** and why a formal theory is required.
9. In the definition of a **subparameter**, it is shown that certain **parameter functions** are **natural** and others are not.
10. Inference is meaningful only for **natural parameters**.
11. This distinction has important consequences for the construction of **prior distributions**
12. and also helps to resolve a controversy concerning the **Box-Cox model**.

Statistical Isomorphism Norman Morse and Richard Sacksteder

1. A **statistical** ***problem*** consists in part of a **sample space** and a set of **probability distributions** on that space.

2. One can speak of the space and the set of probability distributions as a **"statistical system."**

3. ...different **statistical systems**...may be considered **equivalent** [if] [**stating**] a...statistical **problem** in terms of **either** system gives the statistician the **same amount** of...**information** with respect to the...problem.

4. This notion of **equivalence** has been given precise development in a number of papers dealing with the concept of **sufficiency** and with the **"comparison of experiments,"** as will be noted below.

5. .[It is] generally accepted that...Given a **sample space** and a [set of parameterized **probability distributions**] on the space,

6. if there is a map [giving] each point of [one] sample space something like a **probability distribution** on a second space,

7. and [there is] an induced set of probability distributions on the second space corresponding to those given on the first,

8. then...the second space and the induced probability distributions [are] second or induced statistical system,

9. and...the first system [is] **sufficient** for the second. Two systems are **"equivalent"** if each is **sufficient** for the other.

10. ...The formal definition of statistical isomorphism which we give is not convenient for determining whether two statistical systems are isomorphic...

11. .[We need a] complete set of **invariants** of the **isomorphism classes**. Our main result, Theorem 2, provides such a set of invariants for dominated statistical systems.

12. The invariants have a simple intuitive interpretation which we illustrate in a simple case in this section.

Probability Sheaves and the Giry Monad
Alex Simpson

1. I introduce the notion of **probability sheaf**,

2. which is a mathematical structure capturing the **relationship** between **probabilistic concepts** (such as random variable) and **sample spaces**.

3. Various **probability**-theoretic notions can be (re)formulated in terms of **category**-theoretic structure on the **category of probability sheaves**.

4. As a main example, I consider the **Giry monad**, which, in its original formulation, constructs spaces of probability measures.

5. I show that the **Giry monad generalises to** the category of **probability sheaves**, where it turns out to have a simple, purely category-theoretic definition.

The Algebra and Machine Representation of Statistical Models
Evan Patterson

1. This dissertation takes steps toward **digitizing** and **systematizing**... **statistical models** and **data analyses**.

2. Using tools from...categorical logic, a **precise analogy** is drawn between [**statistical models**] and [**logical models**]...

3. **Statistical theories**, being algebraic structures, are amenable to **machine representation** and are equipped with **morphisms** that formalize the relations **between** different **statistical methods**.

4. ...a **software system** for creating **machine representations** of **data analyses, in...Python or R** programs, is designed and implemented.

5. The representations aim to capture the ***semantics of data analyses,*** **independent of the programming language** and libraries in which they are implemented.

6. ...The necessary background in category theory is presented in Chapter 2.

7. ...In Chapters 3 and 4, I develop the **algebra** of statistical theories, [models], and their morphisms.

8. ...In the second major part, I describe the design and **implementation** of a **software** system for creating semantic **representations** of data science **workflows**.

9. Chapter 5 [is] on the **analysis** of data science **code**...

10. In Chapter 6, I present the **Data Science Ontology** and a **procedure** for the **semantic enrichment** of idealized **computer programs**.

11. The concluding Chapter 7 describes limitations of the work, suggests directions for future work, and offers a general outlook on how the **structuralist approach** to data analysis might transform the scientific process.

Categorical Probability and Stochastic Dominance in Metric Spaces
Paolo Perrone

Disclaimer: *There is a lot of overlap between this thesis and papers co-authored by Perrone and Fritz. We can select readings from this dissertation, or corresponding papers, depending on interest and other factors. Think of this outline as a multi-source summary.*

1. In this work we introduce...category-theoretical concepts...to study **probability distributions** on **metric spaces** and ***ordered*** **metric spaces**.

2. The leading themes in this work are **Kantorovich duality** [Vil09, Chapter 5],

3. **Choquet theory** [Win85, Chapter 1],

4. and the categorical theory of **monads** and their algebras [Mac00, Chapter VI].

5. ...**Probability monads**[, discussed in Chapter 1,] can be interpreted as a categorical tool to talk about **random elements** of a space [and their **convex combinations**].

6. ...to every **monad** corresponds an **adjunction**.

7. For **probability monads**, this **adjunction** can be interpreted in terms of **Choquet theory**.

8. In Chapter 2 we define a **probability monad** on the category of **complete metric spaces** and **1-Lipschitz maps** called the **Kantorovich monad**...

9. This monad assigns to each complete metric space X its **Wasserstein space** PX, which is itself a **complete metric space**.

10. In Chapter 3 we **extend** the **Kantorovich monad** of Chapter 2 to metric spaces equipped with a **partial order**. The **order is inherited** by the **Wasserstein space**, and is called the **stochastic order**.

11. ...we define a **compatibility condition** of the **order** with the **metric**...We call the spaces with this property **L-ordered spaces**

12. ...In Chapter 4 we study a **different order** between probability measures, which [points] in the direction of **increasing randomness**.

13. ...we develop a **new** categorical formalism to describe **operations evaluated *partially***.

14. ...the **partial evaluation order** is equivalent to the order known in the literature as the **convex [order]** or **Choquet order**.

15. ...we study the relation between these **partial evaluation orders** and **convex functions**

16. ...[we] derive a [new] **general duality** result valid on **all ordered Banach spaces**

17. ...for every two probability measures p and q over A, $\int f dp \leq \int f dq$ for all convex monotone functions f if and only if $p \preceq_l q$ for the lax partial evaluation order.

A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics
Tobias Fritz

1. We develop **Markov categories** as a framework for synthetic probability and statistics...
2. ...we treat the following concepts in purely abstract categorical terms:
3. **conditioning** and **disintegration**;
4. various versions of **conditional independence** and its standard properties;
5. **conditional products**;
6. **almost surely**;
7. **sufficient statistics**;
8. versions of **theorems** on **sufficient statistics** due to Fisher-Neyman, Basu, and Bahadur.
9. ...[This approach] provides a uniform treatment of...
10. **discrete probability theory**,
11. **measure-theoretic probability** with **general** measurable spaces,
12. **Gaussian probability**,
13. **Markov processes** of either of these kinds, and many others.