Creative title: Where we give results about things

Other Author[a,1], Another Author[b], and Daniel J. McDonald[a]

[a]Department of Statistics, The University of British Columbia
[b]Department of Music, Indiana University

Last updated: August 26, 2024

Abstract

Write an abstract.

Keywords: fractals | quavers | carbonara

1   Introduction

We write some math for fun:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}|\Sigma|^{n/2}} \exp\left\{-\frac{1}{2}(y-\mu)^{\mathsf{T}}\Sigma^{-1}(y-\mu)\right\}\mathsf{d}y = 1. \tag{1}$$

We encourage the use of the various macros defined in ShorTeX, so do your best. It makes things easier to read, but also provides lots of necessary mathematics definitions that render nicely.

Be careful with things like KL divergence and conditional probability statements. I find that

$$\mathrm{KL}(q \parallel p)$$

looks much better than

$$\mathrm{KL}(q||p),$$

and I similarly prefer

$$Y \mid X \sim \mathcal{N}(X, \ \sigma^2) \quad \text{to} \quad Y|X \sim \mathcal{N}(X, \sigma^2).$$

---

[1]To whom correspondence should be addressed. E-mail: some.email@stat.ubc.ca

Note that the reals are $\mathbb{R}^p$. There is also

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2n} \|\boldsymbol{y} - \boldsymbol{X}\beta\|_2^2 + \lambda\|\beta\|_1,$$

and (with automatic sizing of the norm),

$$\widehat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \frac{1}{2n} \left\|\frac{y}{1} - X\beta\right\|_2^2 + \lambda\|\beta\|_1.$$

Note that the indicator function looks like $\mathbb{1}\{\cdot\}$, but I sometimes prefer $\mathbf{1}\{\cdot\}$.

1.1 Cleveref

We prefer to use cleveref to get nice references to things. For example, you can say that Eq. (1) was printed in Section 1. No need to write out things like "Section".

2 Some best practices

Some of these are taken from Jacob Bien. Note the use of the ShorTeX itemize environment style below.

- Section titles should be all title case or all sentence case. Don't mix and match.

- I prefer data set to dataset.

- I prefer data to be singular. There remains debate on this point. When you use the word "datum" in a sentence, then we can argue. Data is a mass noun, like "information". We don't say "How many data are enough?", we say "How much data is enough." Enough said.

- Terminology is lower case, unless it's a person's name: Nyström extension and lasso.

- Equations are parts of the sentence. Displayed equations almost always have a comma or period after. Very rarely is there a colon or comma *before* a displayed equation.

- Don't start sentences with math ("Σ is the covariance of $\boldsymbol{X}$.") or the name of a software package that's lowercase, e.g., "`glmnet` is my favourite software".

- Don't use contractions.

- No need to put dollar signs around numbers: 12 versus 12.

- DO put dollar signs around math: p not p.

- Use $x \gg y$ not $x >> y$.

- Careful with parentheticals and references. Wrong: (see, e.g., Akaike (1973)). Right: (see, e.g., Akaike, 1973).

Ingredient	Quantity
Fusili	100 g
Eggs	2
Salt	1 tsp
Guanciale	50 g
Pepper	$^1/_2$ tsp
Grated parmesan	$^1/_4$ c

Table 1: This is a nice looking table. It might make carbonara.

- Never use `eqnarray`, always use `align`. Note that ShorTeX makes \[\] into an align environment, so you can just use that always.

- For editing purposes, it is much better if the text is hard-wrapped rather than soft wrapped.

2.1 Tables

Table 1 is a nice looking table. Strive for these.

3 Discussion

We made amazing contributions to the world of musical fractal pasta (McDonald, 2017; Tibshirani, 2013). We use Natbib, so be sure to use (Stein, 1981) for parenthetical references. Or you can say, according to Hastie et al. (2009), we should strive to balance truth and lies.

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In *Proceedings of the 2nd International Symposium of Information Theory* (eds. B. N. Petrov and F. Csaki), 267–281.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Verlag.

McDonald, D. J. (2017) Minimax Density Estimation for Growing Dimension. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)* (eds. A. Singh and J. Zhu), vol. 54, 194–203. PMLR.

Stein, C. M. (1981) Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, **9**, 1135–1151.

Tibshirani, R. J. (2013) The lasso problem and uniqueness. *Electronic Journal of Statistics*, **7**, 1456–1490.