

Notation

Contents

1	Notation	1
1.1	Information matrices	1
1.2	Percent sign (“%”)	2
1.3	Proofs	2
1.4	Why is notation in probability and statistics so inconsistent and disorganized?	2

1 Notation

Table 1: Notation used in this book

symbol	meaning	LaTeX
\neg	not	<code>\neg</code>
\forall	all	<code>\forall</code>
\exists	some	<code>\exists</code>
\cup	union, “or”	<code>\cup</code>
\cap	intersection, “and”	<code>\cap</code>
$ $	given, conditional on	<code>\mid</code> , <code> </code>
\sum	sum	<code>\sum</code>
\prod	product	<code>\prod</code>
μ	mean	<code>\mu</code>
E	expectation ¹	<code>\mathbb{E}</code>
x^\top	transpose of x	<code>x^{\top}</code>
$'$	transpose or derivative ²	<code>'</code>
\perp	independent ³	
\therefore	therefore, thus	<code>\therefore</code>
η	linear component of a GLM ⁴	<code>\eta</code>
$\lfloor x \rfloor$	floor of x : largest integer smaller than x	<code>\lfloor x \rfloor</code>
$\lceil x \rceil$	ceiling of x : smallest integer larger than x	<code>\lceil x \rceil</code>

1.1 Information matrices

There is no consistency in the notation for observed and expected information matrices (see Table 2).

¹[probability.qmd#def-expectation](#)

²depending on whether it is applied to a matrix or a function

³[probability.qmd#def-indpt](#)

⁴https://en.wikipedia.org/wiki/Generalized_linear_model#:~:text=The%20linear%20predictor%20is%20the,data%20through%20the%20link%20function

Table 2: notation for information matrices

book	observed information	expected information
Dobson and Barnett (2018)	U'	\mathfrak{I}
Dunn and Smyth (2018)	\mathfrak{I}	\mathcal{I}
McLachlan and Krishnan (2007)	I	\mathcal{I}
Wood (2017)	\hat{I}	\mathcal{I}

These notes currently have a mixture of notations, depending on my whims and what reference I had last looked at. Eventually, I will try to standardize my notation to I for observed information and \mathcal{I} for expected information.

1.2 Percent sign (“%”)

The percent sign “%” is just a shorthand for “/100”. The word “percent” comes from the Latin “per centum”; “centum” is Latin for 100, so “percent” means “per hundred” (c.f., <https://en.wikipedia.org/wiki/Percentage>)

So, contrary to what you may have learned previously, $10\% = 0.1$ is a true and correct equality, just as $10\text{kg} = 10,000\text{g}$ is true and correct.

Proof.

$$\begin{aligned}
 10\% &= 10/100 \\
 &= \frac{10}{100} \\
 &= 0.1
 \end{aligned}$$

□

You are welcome to switch between decimal and percent notation freely; just make sure you execute it correctly.

1.3 Proofs

We can use any of:

- \therefore (`\therefore` in LaTeX),
- \Rightarrow (`\Rightarrow`),
- \models (`\models`)

to denote logical entailments (deductive consequences).

Let’s save \rightarrow (`\rightarrow`) for convergence results.

1.4 Why is notation in probability and statistics so inconsistent and disorganized?

In grad school, we are asked to learn from increasingly disorganized materials and lectures. Not coincidentally, as the amount of organization decreases, the amount of complexity increases, the amount of difficulty increases, the number of reliable references decreases, and the amount of inconsistency in notation and content increases (both between multiple references and within single references!). In other words, as you approach the cutting-edge of most fields, you start to encounter into content that hasn’t been fully thought through or standardized. This lack of clarity is unfortunate and undesirable, but it is understandable and inevitable.

It’s worth noting that calculus was formalized in the 1600s⁵, elementary algebra was formalized around 820⁶, and arithmetic even earlier⁷. And calculus still has several competing notation

⁵https://en.wikipedia.org/wiki/Leibniz%27s_notation

⁶<https://en.wikipedia.org/wiki/Al-Jabr>

⁷<https://en.wikipedia.org/wiki/Arithmetic#History>

systems⁸. In contrast, the field of statistics only emerged in the late 1800s and early 1900s⁹, so it's not surprising that the notation and terminology is still developing. Generalized linear models were only formalized in 1972 (Nelder and Wedderburn (1972)), which is very recent in terms of the pace of scientific development¹⁰.

- Dobson, Annette J, and Adrian G Barnett. 2018. *An Introduction to Generalized Linear Models*. 4th ed. CRC press. <https://doi.org/10.1201/9781315182780>.
- Dunn, Peter K, and Gordon K Smyth. 2018. *Generalized Linear Models with Examples in R*. Vol. 53. Springer. <https://link.springer.com/book/10.1007/978-1-4419-0118-7>.
- McLachlan, Geoffrey J, and Thiriyambakam Krishnan. 2007. *The EM Algorithm and Extensions*. 2nd ed. John Wiley & Sons. <https://doi.org/10.1002/9780470191613>.
- Nelder, John Ashworth, and Robert WM Wedderburn. 1972. "Generalized Linear Models." *Journal of the Royal Statistical Society Series A: Statistics in Society* 135 (3): 370–84.
- Wood, Simon N. 2017. *Generalized Additive Models: An Introduction with r*. chapman; hall/CRC.

⁸https://en.wikipedia.org/wiki/Notation_for_differentiation

⁹https://en.wikipedia.org/wiki/History_of_statistics#Development_of_modern_statistics

¹⁰https://en.wikipedia.org/wiki/The_Structure_of_Scientific_Revolutions