LaTeX Typesetting

Asako Mikami

February 13, 2019

Abstract

This is a sample document created in a LaTeX text editor. It demonstrates some of LaTeX syntax necessary to write reproducible research papers such as a) writing mathematics in LaTeX; b) inserting footnotes; c) typesetting hyperlinks to navigate the document; d) adding appendix section; e) adding bibliography at the end of the document. The document uses a `.sty` file named `sample.sty`, and a `.bib` file named `sample.bib`. Both files are stored in the same directory as this file.

Meta comments are displayed in blue. They can be hidden by

1.) opening the style file associated with this document, `sample.sty`, located in the current directory,

2.) commenting the line, `\includecomment{notes}`,

3.) uncommenting the line, `\excludecomment{notes}`.

1 Probability: random variable and cdf

Begin a section by typing `\section{}` with the section title inside the brackets.

Here is a paragraph containing *italics*, inline math delimited by `$ ...$`, and a footnote with two hyperlinks, one to the Appendix section and another to a web page.

We say $X = X(\omega)$ is a *random variable* that maps an outcome/event ω in sample space Ω to a real number.[*] For example if we have a sample space $\Omega = \{\text{success}, \text{fail}\}$, then we can map Ω to $\mathbb{R}$ by coding $X = 1$ when the outcome is a success and $X = 0$ when the outcome is a failure. In short, a random variable is a coding scheme.

The next passage contains a displayed math delimited by $$ . . . $$ and lists. For numbered lists, use the environment `\begin{enumerate}...\end{enumerate}`. For a bullet lists, use the environment `\begin{itemize}...\end{itemize}`. The lists are formatted to reduce empty space by setting the spacing parameters globally in `sample.sty` as follows:

[*] A more technical definition states that X is a measurable function mapping an *event space* $\mathcal{F}$ to $\mathbb{R}$, which implies that not all outcomes are events and not all events are outcomes. See Appendix A and this discussion on StackExchange for the motivation behind conceptualizing probability as a measure.

```
\setlist{topsep=0pt, itemsep = 0pt, leftmargin=50pt}
```

See this helpful StackExchange post for what these parameters mean.

Say we want to make a statement about the probability of an event $\{\omega \in \Omega | X(\omega) \leq x\}$ for some real number x. Such a statement is called a *cumulative distribution function*, denoted $F(x)$ or $F_X(x)$ where

$$F(x) = F_X(x) = P(\{X \leq x\}) = P(\{\omega \in \Omega, X(\omega) \leq x\}).$$

Mathematically, P is a special function called a *probability measure* and therefore has the following properties called *Kolmogorov axioms*:

1.) $0 \leq P(A) < \infty$ for all events A;

2.) for a countable sequence of disjoint events $A_1, A_2, \ldots$, $P(\bigcup\limits_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$;

3.) $P(\Omega) = 1$.

Kolmogorov axioms imply the following results.

a) $P(\varnothing) = 0$[†]

b) $0 \leq P(A) \leq 1$ for all events $A \in \omega$.

c) If $A \subseteq B$, then $P(A) \leq P(B)$.

d) For events A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

I changed the list label to appear as "a.), b.), c.)" by changing the default numeric label to alphabetic like so.

```
\begin{enumerate}[label=\alph*)]
\end{enumerate}
```

I also like my numeric labels to appear as "1.), 2.), 3.)" instead of the default, "1., 2., 3.", so I use `[label=\arabic*.)]`.

Let's prove d) for exercise.

Proof. Assume that A and B are nonempty events where $A \subseteq B$. By the first Kolmogorov axiom, we know that $0 \leq P(A), P(B) < \infty$. If $A \subset B$, then B/A and A are disjoint and their union is B. By the second axiom, $P(B) = P((B/A) \cup A) = P(B/A) + P(A)$. Therefore, $P(A) \leq P(B)$. $\square$

The `amsmath` package comes with `proof` environments:

```
\begin{proof}
\end{proof}
```

Using these properties of P, we can also derive the following three properties for cumulative distribution function, F.

[†] This is one instance where an outcome and an event are not the same because $\varnothing$ cannot an element of Ω and is not an *outcome*. However, we can still define probabilities for an *event*.

- F is nondecreasing, i.e. $F(x) \leq F(y)$ for all $x < y$.
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$
- F is right continuous, i.e. $\lim_{h \to 0} F(x+h) = F(x)$.

Let's prove the first one for exercise.

Proof. Assume that $x < y$. Then it must be that $\{X \leq x\} \subset \{X \leq y\}$. By the property we proved above, $P(X \leq x) \leq P(Y \leq y)$. Therefore by the definition of cdf, $F(x) \leq F(y)$. $\square$

2 Famous distributions

Notes on notation: I use $p(k)$ for probability mass function and $f(x)$ for probability density function. I use x for real numbers, and k for natural numbers. $\mathbb{I}_A$ is an indicator function where

$$\mathbb{I}_A = \mathbb{I}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

As frequently done, I have set up macros in the style file for the expectation and variance symbols.

```
\newcommand{\E}{\mathrm{E}\,}
\newcommand{\Var}{\mathrm{Var}\,}.
```

Now the expectation appears as $\mathrm{E}\,$ instead of E by typing `\E`. Similarly, the variance appears as $\mathrm{Var}\,$ instead of Var by typing `\Var`. I have `\,` added there to leave some space after the symbol.

Uniform distribution, $U(a, b)$

Begin an unnumbered subsection by typing `\subsubsection*{}` with the sub-subsection title inside the brackets. If you want it to be numbered, take out the asterik. For any environment that has the option to be unnumbered (such as lists), you can add an asterik to disable the numbering.

A random varible X has a (continuous) uniform distribution if it has constant probability over its sample space, (a, b).

$$f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}$$

$$F(x) = \begin{cases} 0 & x \leq a \\ x & a < x < b \\ 1 & x \geq b \end{cases}$$

$$\mathrm{E}\,[X] = \frac{a+b}{2}, \quad \mathrm{Var}\,(X) = \frac{(b-a)^2}{12}$$

Normal distribution, $N(\mu, \sigma^2)$

The bell curve that is observed in more places than you'd think. $X \sim N(\mu, \sigma^2)$ is a normal distribution where $\mathrm{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ -\frac{(x-\mu)^2}{2\sigma^2} \Big\} \mathbb{I}_{\mathbb{R}}$$

Exponential distribution, $\mathrm{Exp}(\beta)$

Let $X \sim \mathrm{Exp}(\beta)$ be a random variable describing the waiting time between events that occur at a constant rate. Then X is said to have an exponential distribution where $\mathrm{E}[X] = \beta$.

$$\begin{aligned} f(x) &= \frac{1}{\beta} \exp\Big\{ -\frac{x}{\beta} \Big\} \mathbb{I}_{[0,\infty)} \\ F(x) &= 1 - \exp\Big\{ -\frac{x}{\beta} \Big\} \\ \mathrm{E}[X] &= \beta, \quad \mathrm{Var}(X) = \beta^2 \end{aligned}$$

Bernoulli distribution, $\mathrm{Ber(p)}$

Let X describe a one-trial experiment whose outcome can be a success with probability p or a failure with probability $1-p$. Let $X(\text{success}) = 1$ and $X(\text{failure}) = 0$.

$$\begin{aligned} p(k) = P(X = k) &= p^k (1-p)^{1-k} \mathbb{I}_{\{0,1\}} \\ \mathrm{E}[X] = p, &\quad \mathrm{Var}(X) = p(1-p) \end{aligned}$$

Binomial distribution

A binomial distribution $X \sim \mathrm{bin(n,p)}$ models an experiment with n trials with p probability of success per trial. Alternatively, it can be modeled as a sum of n independent and identical Bernoulli trials with p probability of success.

$$\begin{aligned} P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \mathbb{I}_{\{k \leq n\}} \\ \mathrm{E}[X] &= np, \quad \mathrm{Var}(X) = np(1-p) \end{aligned}$$

Poisson distribution, $\mathrm{Poisson}(\lambda)$

Suppose we want to describe the number of plane crashes that occur in a year. Intuitively, a binomial distribution would make sense. But for rare events where there is a large number of n (number of flights per year) and a small probability of success p (the relative frequency of a plane crash in a year), a Poisson distribution is a good approximation and may be more practical as it uses the average number of successes as the parameter.

$$\begin{aligned} p(k) = P(X = k) &= \frac{\lambda^k e^{-\lambda}}{k!} \mathbb{I}_{\mathbb{N}_0} \\ \mathrm{E}[X] = \lambda, &\quad \mathrm{Var}(X) = \lambda \end{aligned}$$

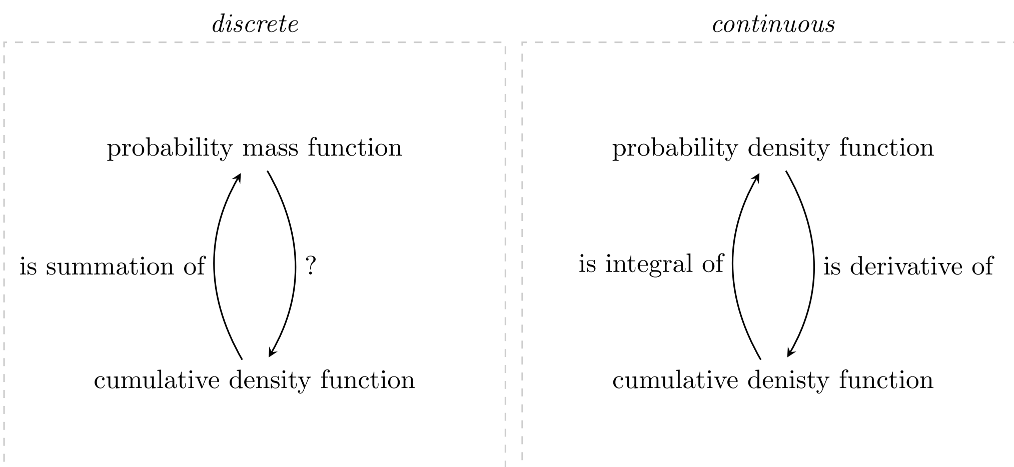
Use `\appendix` command to begin appendix.

A Motivation for a measure theory-based probability

I have created a hyperlink to this appendix section by adding `\label{measure}` next to the section tag. I can refer to this section by writing `\ref{measure}`.

The goal of a probability measure theory is to unify the discrete and continous cases as the two cases do not parallel each other nicely. In an introductory probability textbook such as Ross 2014, students first learn that a random variable X may be associated with either a discrete sample space (i.e. the sample space Ω is countable) or a continuous sample space (i.e. the sample space Ω is uncountable). Next, they define a probability mass function (pmf) for a discrete r.v. and a probability density function (pdf) for a continuous random variable. At this point, only the pmf actually expresses probability. In order to make a probability statement for a continuous r.v., we need to get to the cumulative distribution function (cdf). This is the first instance where the two cases are not congruent. Second, unlike the relationship between pdf and cdf, where the former is the derivative of the latter and the latter is the integral of the former, the relationship between the pmf and the cdf is not one of a derivative and an antiderivative. The following figure attempts to illustrate the situation.

Figure 1: The discrete case and the continuous case do not parallel each other.



A measure theory-based probability solves these inconsistencies. By defining probability as a measure on a measurable set, regardless of whether our random variable is discrete or continuous, it becomes mathematically possible to state that the pdf (pmf) is the derivative of the cdf and that the cdf is the integral of the pdf (pmf).

B How to get help on LaTeX

Google it.

There are many ebooks, books, and websites out there illustrating the A to Z of LaTeX syntax. None of them are short, nor is reading through any of them actually helpful for learning to write in LaTeX. If you have any question about how to do something in LaTeX, or in any programming language for that matter, then google it. Chances are that somebody else has already asked the question on Tex StackExchange and that one of the answers will be the right fit for you. Even when you encounter errors, copy-paste the error message and you'll find a solution on Tex StackExchange.

C How to get better at LaTeX

Practice.

The best way to get better at LaTeX is to stop using Microsoft Word and do more writing on a LaTeX text editor. Write your notes in LaTeX.[‡] Write your paper in LaTeX. Write your CV in LaTeX. Write your presentation slides in LaTeX. Look up how you can do something on LaTeX by googling it and try it out.

[‡]Here are some examples of what I practiced: my probability notes, my game theory notes.

Use `\printbibliography` command to insert a bibliography. In the style file, I have the following lines specifying that I want the `biblatex` package to write my bibliography in author-year format with the `biber` engine as my `.bib` processor.

```
\usepackage[backend=biber, style=authoryear, citestyle=authoryear]{biblatex}
\addbibresource{sample}
```

The `.bib` file stores all my bibliography entries. You can find a `bibtex` formatted citation by googling the work you want to cite on Google Scholar. For example, the entry for Ross 2014 looks like this:

```
@book{ross,
   title={A first course in probability},
   author={Ross, Sheldon},
   year={2014},
   publisher={Pearson}
}
```

References

Ross, Sheldon (2014). *A first course in probability*. Pearson.