1819-108-C1-W7-02

Kārlis Kreilis

March 2019

3.5. Axiomatization of Information Measurement

Consequently the student who is devoid of talent
will derive no more profit from this work
than barren soil from a treatise on agriculture.
Quintilian (35–95)

Information entropy (quantity of information) has an axiomatic description. Namely, there is a set of axioms, which characterize some natural properties of this information measure, and these axioms uniquely determine Shannon's entropy.

There are different systems of axioms used for characterization of the information entropy. We follow here the approach of Shannon (1948).

Building his axiomatic system based on natural properties of information measure, Shannon starts with the problem of information transmission. He analyses how to define a quantity that will measure, in some sense, how much information is "produced" by a discrete information source, or as Shannon explains, at what rate information is produced. With such a goal, the problem is specified by taking into account only information that tells the receiver whether some event has happened. To estimate this information, it is assumed that the receiver regards a set $\{E_1, E_2, \ldots, E_n\}$ of possible events with known probabilities of occurrence $p_1, p_2, \ldots, p_n$. Then it is possible to assume that the measure for the information entropy $H(p_1, p_2, \ldots, p_n)$ should have the following properties:

Axiom A1 (*Continuity*). The function $H(p_1, p_2, \ldots, p_n)$ is continuous in all arguments p_i.

Axiom A2 (*Monotonicity*). If all p_i are equal ($p_i = 1/n$), then $H(1/n, 1/n, \ldots, 1/n)$ must increase monotonically with n.

Axiom A3 (*Branching property*). If a choice is broken down into two successive choices, the original function $H(p_1, p_2, \ldots, p_n)$ should be the weighted sum of the two new values of this function and the weights are equal to the probabilities of branching.

3.5. Axiomatization of Information Measurement

Consequently the student who is devoid of talent
will derive no more profit from this work
than barren soil from a treatise on agriculture.
Quintilian (35-95)

Information entropy (quantity of information) has an axiomatic description. Namely, there is a set of axioms, which characterize some natural properties of this information measure, and these axioms uniquely determine Shannon's entropy.

There are different systems of axioms used for characterization of the information entropy. We follow here the approach of Shannon (1948).

Building his axiomatic system based on natural properties of information measure, Shannon starts with the problem of information transmission. He analyses how to define a quantity that will measure, in some sense, how much information is "produced" by a discrete information source, or as Shannon explains, at what rate information is produced. With such a goal, the problem is specified by taking into account only information that tells the receiver whether some event has happened. To estimate this information, it is assumed that the receiver regards a set $\{E_1, E_2, ..., E_n\}$ of possible events with known probabilities of occurrence $p_1, p_2, ..., p_n$. Then it is possible to assume that the measure for the information entropy $H(p_1, p_2, ..., p_n)$ should have the following properties:

Axiom A1 (*Continuity*). The function $H(p_1, p_2, ..., p_n)$ is continuous in all arguments p_i.

Axiom A2 (*Monotonicity*). If all p_i are equal ($p_i = 1/n$), then $H(1/n, 1/n, ..., 1/n\}$ must increase monotonically with n.

Axiom A3 (*Branching property*). If a choice is broken down into two successive choices, the original function $H(p_1, p_2, ..., p_n)$ should be the weighted sum of the two new values of this function and the weights are equal to the probabilities of branching.

```
\documentclass[11pt]{report}
\usepackage[utf8]{inputenc}
\usepackage{geometry}
 \geometry{paper=a4paper,
 left=50mm,
 top=50mm,
 right=47mm,
 bottom=80mm,
 }
 \usepackage{graphicx}
 \usepackage{listings}
\hyphenpenalty=100
\pagestyle{headings}
\title{1819-108-C1-W7-02}
\author{Karlis Kreilis}
\date{March 2019}
\usepackage{microtype}
\begin{document}
\maketitle
\newpage
\includegraphics[width=\textwidth,
height=\textheight]{IMG-20190313-WA0000.jpg}
\newpage
\setcounter{page}{289}
\begin{center}
    \textit{\scriptsize{Statistical Information Theory }}
\end{center}

\noindent  \scalebox{.9}[1.0]
{\textbf{3.5. Axiomatization of Information Measurement}}

\begin{flushright}
\textit{\small{Consequently the
student who is devoid of talent\\
will derive no more profit from this work\\
```

than barren soil from a
treatise on agriculture.\\}}
Quintilian (35−95)
\end{flushright}

\noindent Information entropy
(quantity~ of~ information)
~ has an axiomatic description.
Namely, there is a set of axioms,
which characterize some natural
properties of this information measure,
and these axioms uniquely
determine Shannon's entropy.

There are different systems of
axioms used for characterization
of the information entropy. We follow here
the approachof Shannon (1948).

Building his axiomatic system based
on natural properties of information measure,
Shannon starts with the problem of information
transmission. He analyses how to define a
quantity that will measure,
in some sense, how much information is
"produced" by a discrete
information source, or as Shannon explains,
at what rate information
is produced. With such a goal,
the problem is specified by taking
into account only information that tells
the receiver whether
some event has happened. To estimate
this information,
it is assumed that the receiver

```
regards a set \{$E_1, E_2, ..., E_n$\}
of possible events with known
probabilities of occurrence $p_1, p_2,
..., p_n$. Then it is possible
to assume that the measure for
the information entropy $H(p_1, p_2, ..., p_n)$
should have the following properties:

\textbf{Axiom A1} (\textit{Continuity}).
The function $H(p_1, p_2, ..., p_n)$
is continuous in all arguments $p_i$.

\textbf{Axiom A2} (\textit{Monotonicity}).
If all $p_i$ are
equal $(p_i = 1/n)$,
then $H(1/n, 1/n, ..., 1/n$\}
must increase monotonically with \textit{n}.

\textbf{Axiom A3} (\textit{Branching property}).
If a choice is
broken down into two successive choices,
the original function
$H(p_1, p_2, ..., p_n)$ should be
the weighted sum of the two new
values of this function and the weights are
equal to the probabilities of branching.
\setcounter{page}{3}
\newpage
```