SecondExam-PartTwo

Anna Bogachova

April 2019

about the outcome of the experiment H or about the event E in the message m that informs about an event C is determined by the following formula

$$I(X; Y) = H(X) - H(X \mid Y)$$

Here are some basic properties of statistical information measures (cf. (Aczél and Daróczy, 1975)):

$$H(X, Y) = H(X) + H(Y)$$
$$I(X; Y) = H(X) + H(Y) - H(X, Y)$$
$$I(X; Y) = H(Y) - H(Y \mid X)$$
$$H(X) \geq H(X \mid Y)$$
$$H(X) - H(Y) = H(X \mid Y) - H(Y \mid X)$$

Statistical information theory has been successfully applied to many important problems in communication practice, allowing one to theoretically estimate different characteristics of communication systems, such as channel capacity and amount of information for discrete and continuous, noisy and noiseless systems.

One of the basic results of the statistical information theory is the noisy-channel coding theorem, also called the fundamental theorem of information theory, or just Shannon's theorem because it was first presented with an outline of the proof in the classical paper (Shannon, 1948). A rigorous proof was given later. Here we consider this theorem only as an example of proved in statistical information theory mathematical results with important practical applications. There are many books and papers with dozens similar results about communication channels.

Let us take the channel capacity defined as

$$C = \max \{I(X, Y); p_X\},$$

and the binary entropy function

$$H(p_b) = -[p_b \cdot \log_2 p_b + (1 - p_b) \log_2 (1 - p_b)]$$

Theorem 3.2.1 (Shannon, 1948). For every discrete memoryless channel, the channel capacity C has the following properties:

1. For any $\varepsilon > 0$ and $R < C$, for large enough N, there exists a code of length N and a rate larger than or equal to R and a decoding algorithm, such that the maximal probability of block error is $\leq \varepsilon$.

about the outcome of the experiment H or about the event E in the message m that informs about an event C is determined by the following formula

$$I(X;Y) = H(X) - H(X|Y)$$

Here are some basic properties of statistical information measures (cf. (Aczél and Daróczy, 1975)):

$$H(X;Y) = H(X) + H(Y)$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$
$$I(X;Y) = H(Y) - H(Y|X)$$
$$H(X)/geqH(X|Y)$$
$$H(X) - H(Y) = H(X|Y) - H(Y|X)$$

Statistical information theory has been successfully applied to many important problems in communication practice, allowing one to theoretically estimate different characteristics of communication systems, such as channel capacity and amount of information for discrete and continuous, noisy and noiseless systems.

One of the basic results of the statistical information theory is the noisy-channel coding theorem, also called the fundamental theorem of information theory, or just Shannon's theorem because it was first presented with an outline of the proof in the classical paper (Shannon, 1948). A rigorous proof was given later. Here we consider this theorem only as an example of proved in statistical information theory mathematical results with important practical applications. There are many books and papers with dozens similar results about communication channels.

Let us take the channel capacity defined as

$$C = max\{I(X,Y); p_x\},$$

and the binary entropy function

$$H(p_b) = -[p_b \cdot log_2 p_b + (1 - p_b) log_2 (1 - p_b)]$$

Theorem 3.2.1 (Shannon, 1948). For every discrete memoryless channel, the channel capacity C has the following properties:

1. For any $\varepsilon > 0$ and $R < C$, for large enough N, there exists a code of length N and a rate larger than or equal to R and a decoding algorithm, such that the maximal probability of block errors is $\leq \varepsilon$.

LaTeX code

```
 \documentclass{report}
\usepackage[utf8]{inputenc}
\usepackage[paperheight=247mm,paperwidth=178mm,margin=34mm,heightrounded]
{geometry}
\usepackage{nopageno}
\usepackage{graphicx}

\title{SecondExam-PartTwo}
\author{Anna Bogachova}
\date{April 2019}

\begin{document}
\thispagestyle{empty}
\maketitle

\includegraphics[width=\textwidth]{photo.jpg}

\pagebreak


\noindent 274   \hspace{2.9cm} \textit{Theory of Information}

\hfill

\noindent about the outcome of the experiment \textit{H} or about
the event \text{E} in the message \textit{m} that informs about an
event \textit{C} is determined by the following formula

$$I(X;Y) = H(X) - H(X|Y)$$

Here are some basic properties of statistical information measures
(cf. (Aczél and Daróczy, 1975)):
    $${H(X;Y) = H(X) + H(Y)}$$
   $$ I(X;Y) = H(X) + H(Y) - H(X,Y)$$
   $$I(X;Y) = H(Y) - H(Y|X)$$
   $$H(X) /geq H(X|Y)$$
   $$H(X) - H(Y) = H(X|Y) - H(Y|X)$$

Statistical information theory has been successfully applied to many
important problems in communication practice, allowing one to theoretically
estimate different characteristics of communication systems, such as channel
```

```
capacity and amount of information
for discrete and continuous, noisy and noiseless systems.

One of the basic results of the statistical information theory is the
noisy-channel coding theorem, also called the fundamental theorem of
information theory, or just Shannon’s theorem because it was first presented
with an outline of the proof in the classical paper (Shannon, 1948).
A rigorous proof was given later. Here we consider this theorem only as an
example of proved in statistical information theory mathematical results with
important practical applications. There are many books and papers with dozens

Let us take the channel capacity defined as
$$C = max  \{I(X,Y);p_x\},$$

and the binary entropy function
$$H(p_b) = -[p_b \cdot log_2p_b + (1-p_b)log_2(1-p_b)]$$

\textbf{Theorem 3.2.1} (Shannon, 1948). For every discrete memoryless
channel, the channel capacity \textit{C} has the following properties:

\begin{enumerate}
    \item For any $\varepsilon > 0$ and $R < C$, for large enough
    \textit{N}, there exists a code
\end{enumerate}
\vspace*{-.8em}
of length \textit{N} and a rate larger than or equal to \textit{R} and a
decoding algorithm, such that the maximal probability of block errors
is $\leq \varepsilon $.
```