CHAPTER 2

A NEW PARAMETRIC INFORMATION GENERATING FUNCTION

2.1 INTRODUCTION

The moment generating function of a probability distribution is a convenient tool for evaluating mean, variance and other moments of a probability distribution and an effective embodiment of properties of the same for various analytical processes. The derivatives of moment generating function of a probability distribution yield the moments of a probability distribution. Similarly, the derivatives of information generating function (IGF) a probability distribution gives some statistical entities associated with the probability distribution. The IGF introduced by Golomb [102], given as follows

\[ \sum_{i \in N} p_i^\alpha, \]  

(2.1)

where \( P = (p_1, p_2, \ldots, p_i, \ldots) \) is a complete probability distribution, where \( N \) being a discrete countable sample space and \( \alpha \) is a real variable. Golomb [102] obtained simple expressions for the IGF defined by (2.1) for uniform, geometric and \( \beta \) - power distributions. The first derivative of the function (2.1) at point \( \alpha = 1 \), gives the Shannon’s entropy of the corresponding probability distribution, \( i.e., \) we have

\[ -\left( \frac{\partial}{\partial \alpha} I_\alpha(P) \right)_{\alpha=1} = -\sum_{i \in N} p_i \log p_i = H_S(P), \]

(2.2)

where \( H_S(P) \) is the well known Shannon’s entropy [1] as discussed in previous chapter 1. Further we have

\[ (-1)^r \left( \frac{\partial^r}{\partial \alpha^r} I_\alpha(P) \right)_{\alpha=1} = (-1)^r \sum_{i \in N} p_i (\log p_i)^r. \]

(2.3)

Except for factor of \((-1)^r\), the \( r^{th} \) derivative of IGF given by (2.1) gives the \( r^{th} \) moment of the self-information of the distribution.
The information generating function focuses on those aspects of a probability distribution which remains unchanged on rearranging elements of the probability space. The quantity given by (2.2) measures the average information associated with the probabilities of events but ignores the importance of these events. Thereafter, Guisau and Reischer [103] proposed the following weighted measure of information

$$H(P,U) = -\sum_{i\in N} u_i p_i \log p_i$$  \hspace{1cm} (2.4)

where $U=(u_1,u_2,\ldots,u_i,\ldots)$ and $P=(p_1,p_2,\ldots,p_i,\ldots)$ are the utility distribution and the probability distribution, respectively, $N$ being a discrete countable sample space. The function defined by (2.4) is related with the utility information scheme as follows

$$\begin{pmatrix} E_1 & E_2 & \ldots & E_n \\ p_1 & p_2 & \ldots & p_n \\ u_1 & u_2 & \ldots & u_n \end{pmatrix}; \quad 0 \leq p_i \leq 1, \sum_{i\in N} p_i = 1, \ u_i > 0. \hspace{1cm} (2.5)$$

Here $(E_1,E_2,\ldots,E_n,\ldots)$ denote a family of events with respect to some random experiment and $u_i$ denotes the utility of an event $E_i$ with probability $p_i$. In general, the utility $u_i$ is independent of its probability of occurrence.

Analogous to (2.4), Hooda and Bhaker [105] defined the following weighted IGF,

$$I(P,U,\alpha) = \sum_{i\in N} u_i p_i^\alpha, \quad \alpha \geq 1,$$ \hspace{1cm} (2.6)

$U=(u_1,u_2,\ldots,u_i,\ldots)$ and $P=(p_1,p_2,\ldots,p_i,\ldots)$ are the utility distribution and the probability distribution, respectively as defined in (2.5) and $\alpha$ is a real variable. Further, we have

$$-\left(\frac{\partial}{\partial \alpha} I(P,U,\alpha)\right)_{\alpha=1} = -\sum_{i\in N} u_i p_i \log p_i = H(P,U), \hspace{1cm} (2.7)$$

where $H(P,U)$ is the function given by (2.4).

After that, Guisau and Reischer [103] introduced the relative information generating function for any two probability distributions $P=(p_1,p_2,\ldots,p_i,\ldots)$ and $Q=(q_1,q_2,\ldots,q_i,\ldots)$, defined by

$$I_\alpha(P,Q) = \sum_{i\in N} p_i^\alpha q_i^{1-\alpha}, \hspace{1cm} (2.8)$$
where \( N \) being a discrete countable sample space. The first derivative of (2.8) at point 1 gives the Kullback and Leibler divergence [23],

\[
\left( \frac{\partial I_\alpha(P, Q)}{\partial \alpha} \right)_{\alpha=1} = \sum_{i \in N} p_i \log \left( \frac{p_i}{q_i} \right) = D_{KL}(P \mid Q),
\]

(2.9)

where \( D_{KL}(P \mid Q) \) is the divergence measure proposed by Kullback and Leibler [23], as discussed in chapter 1.

Later on, Hooda and Singh [104] developed an information improvement generating function for any three probability distributions \( P = (p_1, p_2, \ldots, p_i, \ldots), Q = (q_1, q_2, \ldots, q_i, \ldots) \) and \( R = (r_1, r_2, \ldots, r_i, \ldots) \), given by

\[
I_\alpha(P, Q, R) = \sum_{i \in N} p_i q_i^{\alpha-1} r_i^{1-\alpha},
\]

(2.10)

where \( N \) being a discrete countable sample space. The first derivative of the function defined by (2.1) at point 1 obtains Theil’s measure [8], given by

\[
\left( \frac{\partial I_\alpha(P, Q, R)}{\partial \alpha} \right)_{\alpha=1} = \sum_{i \in N} p_i \log \left( \frac{q_i}{r_i} \right)
\]

(2.11)

The particular cases of information improvement generating function give Golomb [102] generating function and Guisau and Reischer [103] relative information generating function. It also provides some expression for discrete as well as power distribution.

Now, in this chapter, we define a new weighted information generating function whose derivative at point 1 gives the measure proposed by Belis and Guisau [81]. Some particular cases of the proposed IGF function for discrete probability distributions are also discussed. Further, the characterization theorems for relative information generating function and information improvement generating function are proved and develop a relation among them.

### 2.2 NEW INFORMATION GENERATING FUNCTION

Consider the following function

\[
I_\alpha(P, U) = \sum_{i \in N} p_i^{1-u_i^{(1-\alpha)}}, \alpha \geq 1,
\]

(2.12)
where \( P = (p_1, p_2, \ldots, p_i, \ldots) \) and \( U = (u_1, u_2, \ldots, u_i, \ldots) \) are the probability and utility distributions respectively as defined in (2.5), \( N \) being a discrete countable sample space and \( \alpha \) is a real variable. Clearly, \( I_1(P,U) = 1 \) and since \( 0 \leq p_i \leq 1 \) for all \( i \), the function (2.12) is convergent for all \( u_i > 0 \). If we take \( u_i = 1 \) for all \( i \), the function (2.12) reduces to (2.1).

It further follows from (2.12) that
\[
- \left( \frac{\partial}{\partial \alpha} I_{\alpha}(P,U) \right)_{\alpha = 1} = - \sum_{i \in N} p_i \log p_i^{u_i} = - \sum_{i \in N} u_i p_i \log p_i = H(P,U).
\]

Therefore, the function defined by (2.12) can be defined as the weighted information generating function for the measure defined by (2.4). Further we have
\[
(-1)^r \left( \frac{\partial^r}{\partial r} I_{\alpha}(P,U) \right)_{\alpha = 1} = (-1)^r \sum_{i \in N} p_i (u_i \log p_i)^r.
\]

The entity \( u_i \log p_i \) can be seen as generalized (or weighted) self information for the utility information scheme given by (2.5). Therefore, except for factor of \((-1)^r\), the \( r^{th} \) derivative of weighted IGF given by (2.12) gives the \( r^{th} \) moment of the generalized self-information for the scheme defined by (2.5).

**Table 2.1: Moments for different random variables**

<table>
<thead>
<tr>
<th>( x_i = -\log p_i )</th>
<th>( x_i = -u_i \log p_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>First moment about origin</td>
<td>First moment about origin</td>
</tr>
<tr>
<td>( \sum_{i \in N} p_i x_i )</td>
<td>( \sum_{i \in N} p_i (-\log p_i) )</td>
</tr>
<tr>
<td>(Shannon entropy) [1]</td>
<td>(Belis and Guiasu) [82]</td>
</tr>
<tr>
<td>Second moment about origin</td>
<td>Second moment about origin</td>
</tr>
<tr>
<td>( \sum_{i \in N} p_i x_i^2 )</td>
<td>( \sum_{i \in N} p_i (-\log p_i)^2 )</td>
</tr>
<tr>
<td>Third moment about origin</td>
<td>Third moment about origin</td>
</tr>
<tr>
<td>( \sum_{i \in N} p_i x_i^3 )</td>
<td>( \sum_{i \in N} p_i (-\log p_i)^3 )</td>
</tr>
<tr>
<td>( r^{th} ) moment about origin</td>
<td>( r^{th} ) moment about origin</td>
</tr>
<tr>
<td>( \sum_{i \in N} p_i x_i^r )</td>
<td>( \sum_{i \in N} p_i (-\log p_i)^r )</td>
</tr>
</tbody>
</table>
The derivatives of the function proposed by Hooda and Bhaker [105] gives only the first moment for the variable \( x_i = -u_i \log p_i \), whereas the derivatives of the function (2.12) gives all the moments for the variable \( x_i = -u_i \log p_i \) as shown in Table 2.1.

In the next section, we consider three particular cases of the proposed information generating function \( I_{\alpha}(P,U) \) for discrete probability distributions.

2.3 PARTICULAR CASES OF INFORMATION GENERATING FUNCTION \( I_{\alpha}(P,U) \) FOR DIFFERENT DISCRETE PROBABILITY DISTRIBUTION

2.3.1. UNIFORM PROBABILITY DISTRIBUTION AND CONSTANT UTILITY DISTRIBUTION

A uniform random variable assumes of values with equal probabilities \( \frac{1}{n} \). Consider \( p_i = \frac{1}{n}; \ i = 1, 2, \ldots \) (uniform probability distribution) and \( u_1, u_2, \ldots = u \) (constant utility distribution), then the weighted IGF given by (2.12) yields

\[
I_{\alpha}(P,U) = \sum_{i \in \mathbb{N}} \left( \frac{1}{n} \right)^{1-u} (1-\alpha) ; \alpha \geq 1.
\]

Also, the first derivative of the above function is given by

\[
-\left( \frac{\partial}{\partial \alpha} I_{\alpha}(P,U) \right)_{\alpha=1} = \log (n)^u = u \log (n),
\]

which is same as Shannon entropy [1] presented in previous chapter 1 for uniform probability distribution and constant utility distribution.

2.3.2. GEOMETRIC PROBABILITY DISTRIBUTION AND CONSTANT UTILITY DISTRIBUTION

The geometric distribution is defined on the number of trials required for achieving the success. Consider \( p_i = qp^i \), \( p + q = 1; i = 0, 1, 2, \ldots, \infty \) (Geometric Probability Distribution) and \( u_1, u_2, \ldots = u \) (Constant utility Distribution), then the weighted IGF given by (2.12) yields
\[ I_\alpha(P,U) = \sum_{i \in N} \left( qp^i \right)^{1-u(1-\alpha)} = \frac{q^{1-u(1-\alpha)}}{1-p^{1-u(1-\alpha)}}. \]

Moreover,

\[ -\left( \frac{\partial}{\partial \alpha} I_\alpha(P,U) \right)_{\alpha=1} = -u \left( \frac{p \log p + q \log q}{q} \right) \]

This is exactly the Shannon entropy [1] for geometric probability distribution and constant utility distribution.

**2.3.3. \( \beta \) - POWER PROBABILITY DISTRIBUTION & CONSTANT UTILITY DISTRIBUTION**

The power distribution is the inverse of the Pareto distribution. If we consider \( P = (p_1, p_2, \ldots, p_i, \ldots), i \in N \) and let \( \Phi(P) = \{p^\beta : \beta \in (-\infty, +\infty)\} \) such that

\[ p^\beta = (p_1^\beta, p_2^\beta, \ldots, p_n^\beta), p_i^\beta = \frac{p_i^\beta}{\zeta(\beta)}, \]

where \( \zeta(\beta) = \sum_{i \in N} p_i^{-\beta} \) and \( N \) being a discrete countable sample space. The above distributions \( p^\beta \) are called \( \beta \)-power distribution.

Now, consider uniform utility distribution \( u_1, u_2, \ldots = u \) (Constant utility distribution), then the information generating function given by (2.12) reduces to

\[ I_\alpha(P,U) = \sum_{i \in N} \left( \frac{p_i^\beta}{\zeta(\beta)} \right)^{1-u(1-\alpha)}, \]

and as a result

\[ -\left( \frac{\partial}{\partial \alpha} I_\alpha(P,U) \right)_{\alpha=1} = u \left( \ln \zeta(\beta) - \frac{\beta \zeta'(\beta)}{\zeta(\beta)} \right). \]
This is exactly the Shannon entropy [1] for $\beta$ - power probability distribution and constant utility distribution. The graph of information generating function for discrete probability distribution is presented in Fig. 2.1.

We have plotted the graph of the proposed IGF for uniform and non-uniform probability and utility distribution as shown in Fig. 2.1. For uniform utility distribution, the graph represents the same behavior for uniform probability and non-uniform probability distribution. The graph of information generating function for discrete probability distribution is presented in Fig. 2.1.

![NEW INFORMATION GENERATING FUNCTION](image)

**Fig 2.1:** The variation of the new IGF given by (2.12) for uniform and non-uniform probability distributions and utility distribution
In the next section, we prove the characterization theorems for relative IGF and information improvement generating function and develop a relation among them.

### 2.4 CHARACTERIZATION OF RELATIVE INFORMATION GENERATING FUNCTION AND INFORMATION IMPROVEMENT GENERATING FUNCTION

The following theorems give the characterization of relative IGF and information improvement generating function.

**Theorem 2.1** - Let \( f : \Omega_n \times \Omega_n \rightarrow \mathbb{R} \) such that \( I_\alpha(P, Q) = \sum_{i \in \mathbb{N}} \phi(p_k, q_k) \) where \( \phi : [0,1] \times [0,1] \rightarrow \mathbb{R} \) such that \( \phi \) is continuous, \( \phi(x, x) = x \) and satisfies the following functional equation

\[
\phi(x_1, x_2, y_1, y_2) = \phi(x_1, y_1)\phi(x_2, y_2),
\]

where \( x_1, y_1, x_2, y_2 \in (0,1] \). Then \( \phi = \phi_\alpha \) and \( I = I_\alpha \), where

\[
\phi_\alpha(x, y) = x^\alpha y^{1-\alpha},
\]

and

\[
I_\alpha(P, Q) = \sum_{i \in \mathbb{N}} p_i^\alpha q_i^{1-\alpha}, \quad \alpha \in \mathbb{R}.
\]

**Proof:** Let us assume \( x = y \), equation (2.16) becomes

\[
\phi(x, x) = x
\]

Considering \( x_1 = x, x_2 = 1, y_1 = 1, y_2 = y \) in equation (2.15), we get

\[
\phi(x, y) = \phi(x, 1)\phi(1, y)
\]

Now, let \( y_1 = y_2 = 1 \) in equation (2.15), we have

\[
\phi(x_1, x_2, 1) = \phi(x_1, 1)\phi(x_2, 1)
\]

Now, again consider \( x_1 = x_2 = 1 \) in equation (2.15), we obtain
Let us assume

\[ \zeta(x) = \phi(x, 1) \]  
\[ \psi(y) = \phi(1, y). \]  

From equation (2.22), (2.19) becomes

\[ \zeta(x, y_1, y_2) = \zeta(x_1) \zeta(x_2) \]
\[ \Rightarrow \zeta(x) = x^{1-\beta} \]  

From equation (2.23), (2.21) becomes

\[ \psi(y_1, y_2) = \psi(y_1) \psi(y_2) \]
\[ \Rightarrow \psi(y) = y^{1-\alpha}, \]  

where \( \alpha, \beta \in \mathbb{R} \).

The continuity of \( \phi \) implies continuity of \( \zeta \) and \( \psi \). Therefore, we have from equation (2.19)

\[ \phi(x, y) = \zeta(x) \psi(y) \]
\[ \phi(x, y) = x^{1-\beta} y^{1-\alpha}. \]  

Using equation (2.26) in equation (2.18), we have

\[ x^{1-\beta} x^{-\alpha} = x \]
\[ 1 - \beta = \alpha. \]  

Using equation (2.27) in (2.26), we finally obtain

\[ \phi(x, y) = x^{\alpha} y^{1-\alpha} \]

and

\[ I_{\alpha}(P, Q) = \sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha}, \alpha \in \mathbb{R}, \]

which is relative information generating function proposed by Guisau and Reischer [103].
Theorem 2.2 - Let \( I : \Omega_n \times \Omega_n \times \Omega_n \rightarrow \mathbb{R} \) such that \( I_\alpha(P,Q,R) = \sum_{i\in\mathbb{N}} \phi(p_i,q_i,r_i) \) where \( \phi : (0,1] \times (0,1] \times (0,1] \rightarrow \mathbb{R} \) such that \( \phi \) is continuous, \( \phi(1,y,y) = 1 \) and satisfy the following functional equations

\[
\phi(x_1,x_2,x_3,y_1,y_2,y_3,z_1,z_2,z_3) = \phi(x_1,y_1,z_1)\phi(x_2,y_2,z_2)\phi(x_3,y_3,z_3)
\]

(2.28)

and

\[
x\phi(x,y_1,y_2,z_1,z_2) = \phi(x,y_1,z_1)\phi(x,y_2,z_2).
\]

(2.29)

Then \( \phi = \phi_\alpha \) and \( I = I_\alpha \),

where \( \phi_\alpha(x,y,z) = xy^{\alpha-1}z^{1-\alpha} \) and \( I_\alpha(P,Q,R) = \sum_{i\in\mathbb{N}} p_i q_i^{\alpha-1} r_i^{1-\alpha}, \alpha \in \mathbb{R} \).

Proof.

Let \( x_1 = x_2 = 1, x_3 = x, y_2 = y, y_1 = y_3 = 1, z_1 = z, z_2 = z_3 = 1 \) in equation (2.28), we get

\[
\phi(x,y,z) = \phi(1,1,z)\phi(1,y,1)\phi(x,1,1)
\]

(2.30)

\[
\phi(x,y,z) = m(z)h(y)l(x),
\]

(2.31)

where \( l(x) = \phi(x,1,1), h(y) = \phi(1,y,1) \) and \( m(z) = \phi(1,1,z) \).

Let \( y_1 = y_2 = y_3 = 1 \) and \( z_1 = z_2 = z_3 = 1 \) in equation (2.28), we get

\[
\phi(x_1,x_2,x_3,1,1) = \phi(x_1,1,1)\phi(x_2,1,1)\phi(x_3,1,1)
\]

\[
\Rightarrow l(x_1,x_2,x_3) = l(x_1)l(x_2)l(x_3)
\]

\[
\Rightarrow l(x) = x^{\alpha-1}
\]

(2.32)

Similarly, let \( x_1 = x_2 = x_3 = 1 \) and \( z_1 = z_2 = z_3 = 1 \) in equation (2.28), we get

\[
h(y_1,y_2,y_3) = h(y_1)h(y_2)h(y_3)
\]

\[
\Rightarrow h(y) = y^{\alpha-1}
\]

(2.33)

Similarly, let \( y_1 = y_2 = y_3 = 1 \) and \( x_1 = x_2 = x_3 = 1 \) in equation (2.28), we get

\[
m(z_1,z_2,z_3) = m(z_1)h(z_2)h(z_3)
\]
\[ m(z) = z^{\nu - 1}. \] \hspace{1cm} (2.34)

Using (2.32), (2.33), (2.34) in equation (2.31), we obtain

\[ \phi(x, y, z) = x^{\beta - 1} y^{\alpha - 1} z^{\nu - 1}. \] \hspace{1cm} (2.35)

Using equation (2.35) in equation (2.29), we have

\[ x \cdot x^{\beta - 1} (y_1 y_2)^{\alpha - 1} (z_1 z_2)^{\nu - 1} = x^{\beta - 1} y_1^{\alpha - 1} z_1^{\nu - 1} x^{\beta - 1} y_2^{\alpha - 1} z_2^{\nu - 1} \]

\[ \Rightarrow \beta = 2. \]

From equation (2.35), we have

\[ \phi(x, y, z) = xy^{\alpha - 1} z^{\nu - 1}. \] \hspace{1cm} (2.36)

Now, considering

\[ \phi(1, y, y) = 1 \]

\[ \Rightarrow y^{\alpha - 1} y^{\nu - 1} = y^0 \]

\[ \Rightarrow y^{\alpha + \nu - 2} = y^0 \]

\[ \Rightarrow \alpha + \nu - 2 = 0 \]

\[ \Rightarrow \nu - 1 = 1 - \alpha. \]

From equation (2.36), we have

\[ \phi(x, y, z) = xy^{\alpha - 1} z^{1 - \alpha}. \]

This completes the proof.

**Theorem 2.3** – For \( P = (p_1, p_2, \ldots, p_i, \ldots), \ Q = (q_1, q_2, \ldots, q_i, \ldots), i \in N, \) we have

\[ I_\alpha(P) \leq I_\alpha(P, Q). \]

**Proof.** Since we know that

\[ q^\alpha \leq q \]

\[ \Rightarrow 1 \leq q^{1 - \alpha} \]
\[ \Rightarrow \sum_{i \in N} p_i^\alpha \leq \sum_{i \in N} p_i^\alpha q_i^{1-\alpha} \]

\[ I_\alpha(P) \leq I_\alpha(P, Q). \]

This develop a relation between the proposed IGF and the function proposed by Golomb [102].

### 2.5 PROPERTIES AND RELATION AMONG INFORMATION GENERATING FUNCTIONS

The properties and relations among existing and proposed generating functions are as follows

i. \( I_\delta(P) = n, i.e., \) does not depend on the elements of \( P \).

ii. \( I_\delta(P) = n I_\delta(P, Q). \)

iii. \( I_\alpha(P, Q) \) is monotonic decreasing function of \( \alpha \).

iv. If we consider \( \beta \)-power distribution, then

\[
I_\alpha(P, P^\beta) = \sum_{i \in N} p_i^\alpha \left( \frac{p_i^\beta}{\sum_{i=1}^{n} p_i^\beta} \right)^{1-\alpha}
\]

\[
= \sum_{i \in N} \frac{p_i^{\alpha+\beta-\alpha\beta}}{\sum_{i \in N} p_i^\beta(1-\alpha)}
\]

\[
= \frac{I_{\alpha+\beta-\alpha\beta}(P)}{I_{(1-\alpha)\beta}(P)}.
\]

v. \( I(P, U, \alpha) = \sum_{i \in N} p_i^\alpha \left( \frac{1}{n} \right)^{1-\alpha} = n^{\alpha-1} \sum_{i \in N} p_i^\alpha = n^{\alpha-1} I_\alpha(P). \)

vi. \( I_\delta(P, Q, R) = \sum_{i=1}^{n} \frac{r_i}{q_i}, i.e., \) depends on the elements of \( Q \) and \( R \).

vii. \( I_\alpha(P, Q, U) = n^{\alpha-1} I_\alpha(P, Q). \)

viii. \( I_\alpha(P, Q, R) \neq I_\alpha(P, R, Q), \) except where \( \alpha = 1 \).
ix. \( I_1(P) = I_1(P, Q) = I_0(P, Q) = I(P, U, 1) = I_1(P, U) = I_1(P, Q, R) = 1 \), i.e., do not depends on the elements of \( P, Q \) and \( R \).

x. \( I_a(P, P, P) = I_a(P, Q, Q) = I_a(P) \).

### 2.6 CONCLUSION

In this chapter, we have introduced a new weighted IGF for discrete probability distribution, which is a generalization of IGF proposed by Golomb and also defined the generalizations of the proposed measure and developed the relation between them. Characterization theorems for relative information generating function and information improvement generating function are also developed. After discussed the generating functions for various entropy, divergence and other information measure, we now turn our attention towards proposing new information theoretic measures and discuss their applications in multiple disciplines.