JIRSS (2022)
Vol. 21, No. 02, pp 233-250
DOI: 10.22034/jirss.2022.706994

Characterizations of some Discrete Distributions and Upper Bounds on Discrete Residual Varentropy

Faranak Goodarzi [1]

[1] Faculty of Mathematics, Department of Statistics, University of Kashan.

Received: 26/11/2021, Accepted: 15/03/2023, Published online: 14/01/2024

Abstract. In this article, we obtain an upper bound for the variance of a function of the residual life random variable for discrete lifetime distributions. As a special case, we find an upper bound for residual varentropy. Moreover we characterize some discrete distributions by Cauchy-Schwarz inequality. We also get new expressions, bounds and stochastic comparisons involving measures in reliability and information theory.

Keywords. Characterization, Discrete aging Intensity, Discrete Cumulative Residual Entropy, Discrete Mean Residual Life Function, Discrete Residual Varentropy, Discrete Variance Residual Life Function, Hazard Rate.

MSC: 60E15, 62E10.

1 Introduction

Let T be a discrete random variable with support $N = \{1, 2, \ldots\}$, survival function $\overline{F}(k) = P(T \geq k)$ and probability mass function (pmf), $p(k) = P(T = k)$. The discrete pseudo-hazard rate and the discrete hazard rate average of T is defined as

$$r(k) = -\log \frac{\overline{F}(k+1)}{\overline{F}(k)}, \tag{1.1}$$

and

$$h(k) = \frac{1}{k} \sum_{i=1}^{k} r(i) = -\frac{1}{k} \log \overline{F}(k+1), \tag{1.2}$$

Corresponding Author: Faranak Goodarzi (f-goodarzi@kashanu.ac.ir)

respectively. By analogy with continuous distributions, the aging intensity of the discrete random variable is defined as

$$L(k) = \frac{r(k)}{h(k)} = k\left[1 - \frac{\log \overline{F}(k)}{\log \overline{F}(k+1)}\right] \text{for } k = 1, 2, \ldots. \tag{1.3}$$

It expresses the unit average aging behavior and analyzes the aging property quantitatively, the larger the aging intensity, the stronger the tendency of aging (see Szymkowiak and Iwińska (2016)).

Lemma 1.1. *(Xie et al. (2002)) Given discrete hazard rate $r(k)$, for $k = 1, 2, \ldots$, the distribution function is determined as*

$$\overline{F}(k) = \exp\left(-\sum_{i=1}^{k-1} r(i)\right), \ \ for \ \ k = 1, 2, \ldots. \tag{1.4}$$

Now, if T is a discrete random variable with support $N_0 = \{0, 1, 2, \ldots\}$, then the discrete hazard rate, the hazard rate average and the mean residual life function of T, respectively are defined as follows:

$$r_F(n) = \frac{P(T = n)}{P(T \geq n)}, \quad n = 0, 1, 2, \ldots, \tag{1.5}$$

$$A(n) = -\frac{1}{n} \log P(T \geq n), \quad n = 0, 1, 2, \ldots, \tag{1.6}$$

$$\mu_F(n) = E(T - n | T \geq n) = \frac{\sum_{x=n}^{\infty} (x - n) P(T = x)}{P(T \geq n)} = \frac{\sum_{x=n+1}^{\infty} \overline{F}(x)}{\overline{F}(n)}, \quad n = 0, 1, 2, \ldots. \tag{1.7}$$

(see Kemp (2004) and Gupta (2015)). The pseudo-hazard rate $r(n)$ and the hazard rate $r_F(n)$ are related as:

$$r(n) = -\ln(1 - r_F(n)).$$

By analogy with continuous distributions, Kemp (2004) proved that, the hazard rate, the survival function and the mean residual life are related by,

$$\begin{aligned} \overline{F}(n) &= \prod_{0 \leq i \leq n-1} [1 - r_F(i)] \\ &= \prod_{0 \leq i \leq n-1} \left[\frac{\mu_F(i)}{1 + \mu_F(i+1)}\right], \quad \mu_F(0) = E(T), \end{aligned} \tag{1.8}$$

and also the variance residual life function is defined as

$$\begin{aligned} \sigma_F^2(n) &= Var(T - n | T \geq n) \\ &= Var(T | T \geq n) \\ &= \frac{2 \sum_{x=n}^{\infty} x \overline{F}(x+1)}{\overline{F}(n)} - (2n - 1)\mu_F(n) - \mu_F^2(n). \end{aligned} \tag{1.9}$$

Definition 1.1. (Kemp , 2004)

(i) F is IFR (DFR) if $r_F(n)$ is increasing (decreasing) in n.

(ii) A lifetime distribution is said to be new better than used (NBU) or new worse than used (NWU) if $P(T \geq n + x) > (<)P(T \geq n)P(T \geq x)$.

Kemp (2004) proved that if a discrete lifetime distribution with infinite support is IFR/DFR then it is NBU/NWU and also a discrete lifetime distribution that is IFR/DFR has a decreasing/increasing mean residual life function, i.e., is DMRL/IMRL.

Moreover, for discrete nonnegative random variables X and Y with cumulative distribution functions F and G and mean residual lifetime functions $\mu_F(t)$ and $\mu_G(t)$, respectively, we say that X is smaller than Y in the

- Usual stochastic order, denoted by $X \leq_{st} Y$, if $\overline{F}(k) \leq \overline{G}(k)$ for all $k \in \{0, 1, 2, \ldots\}$,
- Mean residual lifetime order, denoted by $X \leq_{mrl} Y$, if $\mu_F(k) \leq \mu_G(k)$ for all $k \in \{0, 1, 2, \ldots\}$.

Gupta et al. (1997) defined Glaser's function (also known as eta function) for a discrete random variable as follows:

$$\eta(k) = \frac{p(k) - p(k+1)}{p(k)}.$$

The entropy and the varentropy of a discrete random variable T taking values in the set $\{x_i; i \in I\}$ are expressed, respectively, as

$$H(T) = -\sum_{i \in I} P(T = x_i) \log P(T = x_i), \tag{1.10}$$

and

$$V(T) = \sum_{i \in I} P(T = x_i)[\log P(T = x_i)]^2 - [H(T)]^2. \tag{1.11}$$

Indeed entropy of a discrete random variable is the average number of bits of information that is obtained based on observing one symbol. In other words, if we have a sequence of symbols drawn independently according to probabilities $p_1, p_2, \ldots$, such that by observing symbol o_i, for example, we get $-\log p_i$ bits of information then entropy can be given by equation (1.10). Now if two sources have the same entropy, then for decoding, the number of digits for codeword of a symbol is nearest to entropy, when the source has the least varentropy. The varentropy measures the variability in the information content of X. For more details see Di Crescenzo and Paolillo (2021).

In this paper, we obtain upper bound for the variance of a function of random variables $T_n = (T - n | T \geq n)$ for $n = 0, 1, \ldots$. Upper and lower variance bounds of $g(X)$ for an arbitrary random variable X were considered in Cacoullos (1982) and Cacoullos and Papathanasiou (1985). Both upper and lower variance bounds may be obtained by Cauchy-Schwarz inequality. Now, in order to find the desired bounds, we use the following lemma.

Lemma 1.2. *(Cacoullos and Papathanasiou , 1985) Let X be a nonnegative integer-valued random variable with pmf $p(x)$, $E(X) = \mu$ and $g(x)$ a real-valued function defined on $\{0, 1, 2, \ldots\}$ such that $Var[g(X)] < \infty$. Then*

$$Var[g(X)] \le \sum_{x=0}^{\infty} [\Delta g(x)]^2 \left\{ \sum_{k=0}^{x} (\mu - k) p(k) \right\} = \sum_{x=0}^{\infty} [\Delta g(x)]^2 \left\{ \sum_{k=x+1}^{\infty} (k - \mu) p(k) \right\}, \qquad (1.12)$$

where $\Delta g(x) = g(x+1) - g(x)$ and equality holds if and only if g is linear.

2 Characterization by Cauchy-Schwarz Inequality

In this section, we characterize some distributions using the moments of some functions of interest in reliability theory. The following theorem characterizes the discrete Weibull (III) distribution through $E\left[\frac{T^\alpha}{r(T)}\right]$. For more details about various versions of the Weibull distribution for discrete data, see Almalki and Nadarajah (2014).

Lemma 2.1. *For a discrete random variable T, the discrete hazard rate has form $r(k) = -k^{\beta-1} \log q$ for $k = 1, 2, \ldots; 0 < q < 1$, $\beta \ge 0$ if and only if $T \sim DW(III)(q, \beta)$ with survival function $\overline{F}(k) = q^{\sum_{i=1}^{k-1} i^{\beta-1}}$.*

Theorem 2.1. *For any random variable T with support $\{1, 2, \ldots\}$,*

$$E\left[\frac{T^\alpha}{r(T)}\right] \ge \frac{1}{E\left[\frac{r(T)}{T^\alpha}\right]}, \qquad (2.1)$$

where $\alpha \ge 0$ is a real constant. The equality holds if and only if T has discrete Weibull (III) distribution.

Proof. Using Cauchy-Schwarz inequality, it follows that

$$E\left[\frac{T^\alpha}{r(T)}\right] E\left[\frac{r(T)}{T^\alpha}\right] \ge 1. \qquad (2.2)$$

The equality in (2.2) holds if and only if there exists a constant $A(> 0)$ such that, for all $k \in \{1, 2, \ldots\}$,

$$\frac{k^\alpha}{r(k)} p(k) = A \frac{r(k)}{k^\alpha} p(k), \qquad (2.3)$$

that is equivalent to $r(k) = \theta k^\alpha$ for $\theta > 0$ and by using Lemma 2.1, $T \sim DW(III)(e^{-\theta}, \alpha+1)$.

Trivially if $\alpha = 0$ then the geometric distribution is characterized. For $\alpha = 1$ we obtain the characterization of discrete Rayleigh distribution. □

Theorem 2.2. *For any random variable T with support $\{1, 2, \ldots\}$,*

$$E\left[\frac{T^{\alpha}L(T)}{r(T)}\right] \geq \frac{1}{E\left[\frac{r(T)}{T^{\alpha}L(T)}\right]}, \tag{2.4}$$

for any real constant $\alpha \geq 0$. The equality holds if and only if T has discrete Weibull (I).

Proof. Using Cauchy-Schwarz inequality, it follows that

$$E\left[\frac{r(T)}{T^{\alpha}L(T)}\right]E\left[\frac{T^{\alpha}L(T)}{r(T)}\right] \geq 1. \tag{2.5}$$

The equality in (2.5) holds if and only if there exists a constant $A(>0)$ such that, for all $k \in \{1, 2, \ldots\}$,

$$\frac{k^{\alpha}L(k)}{r(k)}p(k) = A\frac{r(k)}{k^{\alpha}L(k)}p(k), \tag{2.6}$$

that is equivalent to $h(k) = -\frac{1}{k}\log \overline{F}(k+1) = \theta k^{\alpha}$ for $\theta > 0$ and thus $\overline{F}(k+1) = e^{-\theta k^{\alpha+1}}$ for $k = 1, 2, \ldots$.

Now if we consider $e^{-\theta} = q (0 < q < 1)$, then T has discrete Weibull (I) distribution with pmf $p(k) = q^{(k-1)^{\alpha+1}} - q^{(k)^{\alpha+1}}$, $k = 1, 2, \ldots$.
It is easy to note that for $\alpha = 0$, T follows geometric distribution, with parameter $q = e^{-\theta}$ for $\theta > 0$. □

Below we obtain a lower bound for $E[T^3 r_F(T)]$ and characterize the geometric distribution in terms of it. Note that, sometimes hazard function may have a complex form and therefore finding the expected value of a function of it is not simple.

Proposition 2.1. *For any nonnegative discrete random variable T with support N_0,*

$$E[T^3 r_F(T)] \geq \frac{4(E(T^3)^2)}{E(T^2(T+1)^2)}. \tag{2.7}$$

the equality holds if and only if T is geometrically distributed.

Proof. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}(E(T^3))^2 &= \Big[\sum_{x=0}^{\infty} x^3 p(x)\Big]^2 = \Big[\sum_{x=0}^{\infty} \sqrt{x^3 P(T \geq x)}\sqrt{\frac{x^3}{P(T \geq x)}}p(x)\Big]^2 \\ &\leq \sum_{x=0}^{\infty} x^3 P(T \geq x) \sum_{x=0}^{\infty} x^3 \frac{p^2(x)}{P(T \geq x)}. \end{aligned}\tag{2.8}$$

Since

$$\sum_{x=0}^{\infty} x^3 P(T \geq x) = \sum_{x=0}^{\infty} x^3 \sum_{y=x}^{\infty} p(y)$$
$$= \sum_{y=0}^{\infty} p(y) \sum_{x=0}^{y} x^3 = \sum_{y=0}^{\infty} \Big[\frac{y(y+1)}{2}\Big]^2 p(y)$$
$$= \frac{1}{4} E[T^2(T+1)^2], \tag{2.9}$$

and

$$\sum_{x=0}^{\infty} x^3 \frac{p^2(x)}{P(T \geq x)} = E[T^3 r_F(T)],$$

(2.8) reduces to (2.7). The equality holds if and only if there exists a constant $A > 0$ such that, for all $x \in \{0, 1, 2, \ldots\}$,

$$\sqrt{\frac{x^3 p^2(x)}{P(T \geq x)}} = A\sqrt{x^3 P(T \geq x)}.$$

This gives $r_F(x)$ =constant, which holds if and only if T is geometrically distributed. □

Theorem 2.3. *Let T be a discrete random variable with $E[\mu_F(T)/(T+1)] < \infty$ and $E[(T+1)/\mu_F(T)] < \infty$. Then*

$$E[\mu_F(T)/(T+1)] \geq \frac{1}{E[(T+1)/\mu_F(T)]}, \tag{2.10}$$

and the equality holds if and only if T follows the pmf

$$p(x) = (1 + 1/\theta)B(x+1, 2+1/\theta), \quad x = 0, 1, \ldots, \ \theta > 0,$$

where $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, a > 0,\ b > 0$.

Proof. The inequality (2.10) follows from the Cauchy-Schwarz inequality and equality holds if and only if there exists a constant $A > 0$ such that

$$\frac{(x+1)p(x)}{\mu_F(x)} = A\frac{\mu_F(x)p(x)}{x+1}, \tag{2.11}$$

which is equivalent to the fact that $\mu_F(x) = \theta(x+1)$. Applying (1.8), we have

$$\overline{F}(x) = \prod_{i=0}^{x-1} \frac{\theta(i+1)}{1+\theta(i+2)} = \prod_{i=0}^{x-1} \frac{i+1}{\frac{1}{\theta} + i + 2}, \tag{2.12}$$

and thus

$$
\begin{aligned}
P(T=x) &= P(T\ge x) - P(T\ge x+1)\\
&= \prod_{i=0}^{x-1}\frac{i+1}{\frac{1}{\theta}+i+2} - \prod_{i=0}^{x}\frac{i+1}{\frac{1}{\theta}+i+2}\\
&= \prod_{i=0}^{x-1}\frac{i+1}{\frac{1}{\theta}+(i+2)}\Big(\frac{\frac{1}{\theta}+1}{\frac{1}{\theta}+x+2}\Big)\\
&= \frac{(\frac{1}{\theta}+1)x!}{\prod_{i=0}^{x}(\frac{1}{\theta}+i+2)}\\
&= (1+\frac{1}{\theta})\frac{\Gamma(x+1)\Gamma(2+\frac{1}{\theta})}{\Gamma(x+3+\frac{1}{\theta})}\\
&= (1+\frac{1}{\theta})B(x+1,2+\frac{1}{\theta}), \qquad x=0,1,\ldots. \qquad (2.13)
\end{aligned}
$$

□

The next theorem gives a useful lower bound for $E[s^T r_F(T)], |s|\le 1$ in terms of probability generating function, and characterizes the geometric distribution.

Theorem 2.4. *Let T be a nonnegative discrete random variable with support N_0. Then*

$$E[s^T r_F(T)] \ge \frac{1-s}{1-sE(s^T)}(E[s^T])^2, \qquad (2.14)$$

for $|s|\le 1$, where the equality holds if and only if T has the geometric distribution.

Proof. Using the Cauchy-Schwarz inequality, we get

$$(E[s^T])^2 \le E[s^T r_F(T)]E\Big[\frac{s^T}{r_F(T)}\Big]. \qquad (2.15)$$

On the other hand, one can easily show that

$$
\begin{aligned}
E\Big[\frac{s^T}{r_F(T)}\Big] &= \sum_{x=0}^{\infty} s^x P(T\ge x)\\
&= \sum_{y=0}^{\infty}\sum_{x=0}^{y} s^x p(y) = \sum_{y=0}^{\infty}\frac{1-s^{y+1}}{1-s}p(y)\\
&= \frac{1}{1-s}\{1-E(s^{T+1})\}. \qquad (2.16)
\end{aligned}
$$

The equality is obtained if and only if there exists a constant $A>0$ such that, for all $x=0,1,\ldots,$

$$\sqrt{\frac{s^x p^2(x)}{P(T\ge x)}} = A\sqrt{s^x P(T\ge x)}.$$

This gives $r_F(x) = constant$, which again holds if and only if T is geometrically distributed. □

At present, we attain an upper bound for the variance of a function of residual lifetime T_n. For example, we attain an upper bound for varentropy residual lifetime, that will defined in the following literature.

Proposition 2.2. *Let T be a nonnegative discrete random variable with pmf $p(x)$ and survival function $\overline{F}(x)$. If g is a function with forward difference $\Delta g(x) = g(x+1) - g(x)$ then*

$$Var[g(T_n)] \leq E\left[\Big(\frac{1}{r_F(T_n+n)} - 1\Big)\Big(\mu_F(n+T_n+1) - \mu_F(n) + T_n + 1\Big)\Delta g^2(T_n)\right]. \tag{2.17}$$

Proof. We know that

$$P(T_n = x) = P(T - n = x | T \geq n) = \frac{P(T = x+n)}{P(T \geq n)}.$$

Using Lemma 1.2, and since $E[T_n] = \mu_F(n)$, we can write that

$$\begin{aligned}\sum_{k=x+1}^{\infty}(k - \mu_F(n))\frac{P\{T = k+n\}}{P\{T \geq n\}} &= \frac{1}{P(T \geq n)}\left\{\sum_{k=x+1+n}^{\infty}(k-n)p(k) - \mu_F(n)\sum_{k=x+n+1}^{\infty}p(k)\right\} \\ &= \frac{P(T \geq n+x+1)}{P(T \geq n)}\left\{\sum_{k=x+1+n}^{\infty}\frac{(k-(n+x+1))p(k)}{P(T \geq n+x+1)} + (x+1) - \mu_F(n)\right\} \\ &= \frac{P(T \geq n+x+1)}{P(T \geq n)}[\mu_F(n+x+1) - \mu_F(n) + (x+1)], \end{aligned} \tag{2.18}$$

and further using Lemma 1.2 and substituting the right hand side of (2.18) in inequality (1.12), we have

$$\begin{aligned} Var[g(T_n)] &\leq \sum_{x=0}^{\infty}[\Delta g(x)]^2\frac{\overline{F}(x+n+1)}{\overline{F}(n)}\{\mu_F(n+x+1) - \mu_F(n) + x + 1\} \\ &= \sum_{x=0}^{\infty}[\Delta g(x)]^2\frac{\overline{F}(x+n+1)}{P(T = x+n)}\{\mu_F(n+x+1) - \mu_F(n) + x + 1\}\frac{P(T = x+n)}{\overline{F}(n)} \\ &= \sum_{x=0}^{\infty}[\Delta g(x)]^2\frac{P(T \geq x+n) - P(T = x+n)}{P(T = x+n)}\{\mu_F(n+x+1) - \mu_F(n) + x + 1\}\frac{P(T = x+n)}{P(T \geq n)} \\ &= \sum_{x=0}^{\infty}[\Delta g(x)]^2(\frac{1}{r_F(x+n)} - 1)\{\mu_F(n+x+1) - \mu_F(n) + x + 1\}P(T_n = x) \\ &= E\Big[(\Delta g(T_n))^2(\frac{1}{r_F(T_n+n)} - 1)\{\mu_F(n+T_n+1) - \mu_F(n) + T_n + 1\}\Big]. \end{aligned}$$

□

Corollary 2.1. *In Proposition 2.2, if let* $g(x) = -\log \frac{p(n+x)}{\overline{F}(n)}$ *then* $\Delta g(x) = -\log \frac{p(n+x+1)}{p(n+x)} = -\log(1-\eta_{n+x})$ *and hence*

$$Var[-\log p(T_n)] \leq E\Big\{[\log(1-\eta_{n+T_n})]^2(\frac{1}{r_F(T_n+n)} - 1)[\mu_F(n+T_n+1) - \mu_F(n) + T_n + 1]\Big\}. \tag{2.19}$$

According to Lemma 1.2, the equality in (2.19) *holds if and only if* $g(x) = -\log p(n+x) + \log \overline{F}(n)$ *is linear in x. Thus,* $\log p(n+x) = ax + b$ *for constants a and b, and therefore in inequality* (2.19) *the equality holds if and only if T has a geometric and discrete uniform distribution. In fact if* $a \neq 0$ *then* $p(n+x) = e^{ax+b}$ *and thus* $p(x) = (e^a)^x e^{b-n}$ *and simplifyig we can show that T has a geometric distribution. On the other hand if* $a = 0$ *then* $p(n+x) = e^b$ *and hence* $p(x) = c$ *for a finite set of values of random variable T and this shows that T has a discrete uniform distribution.*

Now, we consider the generalization of the entropy to the residual lifetime distributions, that is given by

$$H(T_n) = -\sum_{x=n}^{\infty} \frac{p(x)}{\overline{F}(n)} \log \frac{p(x)}{\overline{F}(n)} = \log \overline{F}(n) - \frac{1}{\overline{F}(n)} \sum_{x=n}^{\infty} p(x) \log p(x), \tag{2.20}$$

which is called residual entropy, in short. Based on Equation (1.11), the residual varentropy, varentropy of the residual lifetime T_n, is defined as

$$\begin{aligned} V(T_n) &= Var[-\log p(T_n)] \\ &= \sum_{x=n}^{\infty} \frac{p(x)}{\overline{F}(n)} \Big(\log \frac{p(x)}{\overline{F}(n)}\Big)^2 - \bigg(\sum_{x=n}^{\infty} \frac{p(x)}{\overline{F}(n)} \Big(\log \frac{p(x)}{\overline{F}(n)}\Big)\bigg)^2 \\ &= \frac{1}{\overline{F}(n)} \sum_{x=n}^{\infty} p(x)(\log p(x))^2 - (\log \overline{F}(n) - H(T_n))^2. \end{aligned} \tag{2.21}$$

It can be easily seen that

$$\lim_{n \to 0} V(T_n) = V(T),$$

where $V(T)$ is the varentropy of T in Equation (1.11). Computing the residual varentropy using the definition in Equation (2.21) is presented in the following example for the geometric distribution.

Example 2.1. Let T be a random variable that follows discrete uniform distribution on the set $\{0, 1, \ldots, \theta\}$, then

$$H(T_n) = \log\Big(1 - \frac{n}{\theta+1}\Big) - \Big(\frac{\theta+1}{\theta+1-n}\Big) \sum_{x=n}^{\theta} \frac{1}{\theta+1} \log \frac{1}{\theta+1} = \log\Big(1 - \frac{n}{\theta+1}\Big) + \log(1+\theta), \tag{2.22}$$

and thus

$$V(T_n)=\frac{\theta+1}{\theta+1-n}\sum_{x=n}^{\theta}\frac{1}{\theta+1}\Big(\frac{1}{\theta+1}\Big)^2-\Big(\log\Big(\frac{\theta+1-n}{\theta+1}\Big)-\log\Big(\frac{\theta+1-n}{\theta+1}\Big)-\log(1+\theta)\Big)^2=0. \tag{2.23}$$

On the other hand, since $-\log(1-\eta_{n+x})=-\log\frac{p(n+x+1)}{p(n+x)}$=0, hence the equality holds in inequality (2.19).

Example 2.2. Let T be a geometric random variable with probability mass function $p(x)=pq^x, x=0,1,2,\ldots$. Then

$$H(T_n)=\log\frac{1-p}{p}-\frac{1}{p}\log(1-p),$$

and thus

$$\begin{aligned}V(T_n)&=\frac{1}{(1-p)^n}\sum_{x=n}^{\infty}p(1-p)^x(\log p(1-p)^x)^2-\Big(\log(1-p)^n+\log\frac{1-p}{p}-\frac{1}{p}\log(1-p)\Big)^2\\&=(\log(1-p))^2\frac{1-p}{p^2}=V(T).\end{aligned} \tag{2.24}$$

On the other hand, since $r_F(n)=p$, $\mu_F(n)=\frac{q}{p}$ and hence $\eta_{n+x}=1-\frac{p(1-p)^{k+1}}{p(1-p)^k}=p$, so by using (2.19) the upper bound for residual varentropy of geometric distribution is obtained as follows:

$$E[(\log(1-p))^2(\frac{1}{p}-1)(T_n+1)]=(\log(1-p))^2\frac{1-p}{p^2}, \tag{2.25}$$

which agrees with the result of Proposition 2.2.

Remark 1. Let X and Y be related by $Y=aX+b$. Hence the residual varentropy of Y is obtained as

$$V(Y_n)=\frac{1}{\overline{F}_X(\frac{n-b}{a})}\sum_{x=\frac{n-b}{a}}^{\infty}p_X(x)(\log p_X(x))^2-\Big(\frac{1}{\overline{F}_X(\frac{n-b}{a})}\sum_{x=\frac{n-b}{a}}^{\infty}p_X(x)(\log p_X(x))\Big)^2. \tag{2.26}$$

Proposition 2.3. *For all $n\in\mathbb{N}\cup\{0\}$, the difference of the residual varentropy is*

$$\Delta V(T_n)=\frac{p(n)}{\overline{F}(n+1)}\Big\{V(T_n)+\Big(H(T_n)+\log r_F(n)\Big)\Big(\log\frac{\overline{F}(n+1)}{\overline{F}(n)}-H(T_{n+1})-\log r_F(n)\Big)\Big\}. \tag{2.27}$$

Proof. By differencing of Equation (2.21) and recalling (1.5), we have

$$\Delta V(T_n) = \frac{p(n)}{\overline{F}(n+1)}\left\{-(\log p(n))^2 + \sum_{x=n}^{\infty}\frac{p(x)}{\overline{F}(n)}(\log p(x))^2\right\} - (\log \overline{F}(n+1) - H(T_{n+1}))^2 + (\log \overline{F}(n) - H(T_n))^2$$
$$= \frac{p(n)}{\overline{F}(n+1)}\left\{V(T_n) + (\log \overline{F}(n) - H(T_n))^2 - (\log p(n))^2\right.$$
$$\left. + (\log r_F(n) + H(T_n))(\log[\overline{F}(n)\overline{F}(n+1)] - H(T_n) - H(T_{n+1}))\right\}$$
$$= \frac{p(n)}{\overline{F}(n+1)}\left\{V(T_n) + \Big(H(T_n) + \log r_F(n)\Big)\Big(\log \frac{\overline{F}(n+1)}{\overline{F}(n)} - H(T_{n+1}) - \log r_F(n)\Big)\right\}. \tag{2.28}$$

□

In the previous theorem, for all n, if $H(T_n) = H(T_{n+1})$, then

$$\Delta V(T_n) = \frac{p(n)}{\overline{F}(n+1)}\Big(V(T_n) + (H(T_n) + \log r_F(n))\log \frac{\overline{F}(n+1)}{\overline{F}(n)} - (H(T_n) + \log r_F(n))^2\Big). \tag{2.29}$$

Proposition 2.4. *Let T have a pmf such that $p(n) > 0$ for all $n = 0, 1, 2, \ldots$.*

Let $c \in \mathbb{R}$; if $H(T_n) = H(T_{n+1})$ and $\log \frac{\overline{F}(n+1)}{\overline{F}(n)} - H(T_{n+1}) - \log r_F(n) = c$, then

$$V(T_n) = \frac{1}{\overline{F}(n)}(V(T) - c^2 F(n-1)) + \frac{c}{\overline{F}(n)}\sum_{i=0}^{n-1} p(i)\log \frac{\overline{F}(i+1)}{\overline{F}(i)}. \tag{2.30}$$

Proof. If the conditions given in the theorem are satisfied, then Equation (2.27) becomes

$$\Delta V(T_n) = \frac{p(n)}{\overline{F}(n+1)}[V(T_n) + c(\log \frac{\overline{F}(n+1)}{\overline{F}(n)} - c)],$$

with initial condition $\lim_{n\to 0} V(T_n) = V(T)$. Now, since

$$V(T_{n+1}) - V(T_n)\Big\{\frac{p(n)}{\overline{F}(n+1)} + 1\Big\} = \frac{cp(n)}{\overline{F}(n+1)}\Big\{\log \frac{\overline{F}(n+1)}{\overline{F}(n)} + c\Big\}, \tag{2.31}$$

is a first order inhomogeneous difference equation, then

$$V(T_n) = \frac{V(T)}{\overline{F}(n)} + \frac{c}{\overline{F}(n)}\sum_{i=0}^{n-1} p(i)\Big\{\log \frac{\overline{F}(i+1)}{\overline{F}(i)} - c\Big\}, \tag{2.32}$$

and the desired result is obtained. □

In the following, we want to find a bound for $V(T_n)$ in terms of the weighted residual entropy of T, which is a weighted version of the residual entropy in (2.20) and is given by

$$\begin{aligned} H^w(T_n) &= -\sum_{x=n}^{\infty} x\frac{p(x)}{\overline{F}(n)}\log\frac{p(x)}{\overline{F}(n)} \\ &= -\frac{1}{\overline{F}(n)}\sum_{x=n}^{\infty} xp(x)\log p(x) + \frac{\log\overline{F}(n)}{\overline{F}(n)}\sum_{x=n}^{\infty} xp(x), \qquad n\in\{0,1,\ldots\}. \end{aligned} \tag{2.33}$$

Furthermore, it is based on the so-called vitality function of T, i.e.

$$\nu(n) = E[X|X\geq n] = \mu_F(n) + n. \tag{2.34}$$

Theorem 2.5. *If T is a random lifetime such that its probability mass function satisfies*

$$e^{-\alpha x-\beta} \leq p(x) \leq 1, \tag{2.35}$$

for all x, with $\alpha > 0$ and $\beta \geq 0$, then for all $n \in \{0,1,\ldots\}$,

$$V(T_n) \leq \alpha[-\log\overline{F}(n)\nu(n) + H^w(T_n)] + \beta[-\log\overline{F}(n) + H(T_n)] - (\log\overline{F}(n) - H(T_n))^2. \tag{2.36}$$

Proof. From Equation (2.21), due to (2.35) we have

$$V(T_n) \leq \frac{-1}{\overline{F}(n)}\sum_{x=n}^{\infty} p(x)\log p(x)(\alpha x+\beta) - (\log\overline{F}(n) - H(T_n))^2. \tag{2.37}$$

We note that Equations (1.7) and (2.34) give

$$\sum_{x=n}^{\infty} xp(x) = \overline{F}(n)\nu(n),$$

hence, recalling (2.34), Equation (2.33) implies:

$$\sum_{x=n}^{\infty} xp(x)\log p(x) = -\overline{F}(n)\{H^w(T_n) - \log\overline{F}(n)\nu(n)\}. \tag{2.38}$$

Moreover, from (2.20), we have

$$\sum_{x=n}^{\infty} p(x)\log p(x) = -\overline{F}(n)\{H(T_n) - \log\overline{F}(n)\}. \tag{2.39}$$

Finally, substituting (2.38) and (2.39) in (2.37), we obtain the inequality (2.36). □

Remark 2. In Proposition 2.2, if T is a nonnegative discrete random variable and F is *IFR* then

$$Var[g(T_n)] \leq E\Big[\Big(\frac{1}{r_F(T)} - 1\Big)(T+1)\Delta g^2(T)\Big]. \tag{2.40}$$

The upper bound of (2.40) is equal to (2.17) if and only if F has geometric distribution with pmf $P(T = x) = pq^x$, $x = 0,1,\ldots$, since T having constant hazard rate, having constant mean residual life and having a geometric distribution are all equivalent.

3 Results on Discrete Cumulative Residual Entropy and Variance

Let X be a nonnegative discrete random variable with survival function $\overline{F}(k) = P(T \geq k)$ and pmf $p(k)$. The discrete cumulative residual entropy of T was defined by Baratpour and Bami (2012) as follows,

$$d\mathscr{E}(X) = -\sum_{k=1}^{\infty} \overline{F}(k) \log \overline{F}(k). \tag{3.1}$$

Theorem 3.1. *If X is a nonnegative discrete random variable, then*

$$d\mathscr{E}(X) \leq 2E(X)\left(\frac{E(X^2)}{E(X^2) - E(X)} - 1\right) + \frac{1}{2}\left(\frac{E(X^2)}{E(X)} - 1\right). \tag{3.2}$$

Proof. By log-sum inequality (see Cover and Thomas (2006)), for $0 < p < 1$ and $q = 1-p$, we have

$$\sum_{k=0}^{\infty} P(X > k) \log \frac{P(X > k)}{q^{k+1}} \geq E(X)\left(\log \frac{E(X)}{\frac{q}{p}}\right). \tag{3.3}$$

Expanding the LHS of (3.3) we get

$$\sum_{k=0}^{\infty} P(X > k) \log P(X > k) - \log q \sum_{k=0}^{\infty} (k+1) P(X > k) \geq E(X) \log\left(\frac{p}{q} E(X)\right).$$

On the other hand, since $\sum_{k=0}^{\infty}(k+1)P(X > k) = \frac{1}{2}(E(X^2) + E(X))$ thus

$$\sum_{k=0}^{\infty} P(X > k) \log P(X > k) \geq \log q \left[\frac{1}{2}(E(X^2) + E(X))\right] + E(X) \log\left[\frac{1-q}{q} E(X)\right]. \tag{3.4}$$

This is valid for all $0 < q < 1$. The maximum of the RHS of (3.4) is attained when $q = \frac{E(X^2) - E(X)}{E(X^2) + E(X)}$.

Substituting this value of q into (3.4) we get

$$\begin{aligned}\sum_{k=0}^{\infty} P(X > k) \log P(X > k) &\geq \log\left(\frac{E(X^2) - E(X)}{E(X^2) + E(X)}\right)\left[\frac{1}{2}(E(X^2) + E(X))\right] + E(X) \log\left[\frac{2E^2(X)}{E(X^2) - E(X)}\right] \\ &\geq \left(1 - \frac{E(X^2) + E(X)}{E(X^2) - E(X)}\right)\left[\frac{1}{2}(E(X^2) + E(X))\right] + E(X)\left(1 - \frac{E(X^2) - E(X)}{2E^2(X)}\right).\end{aligned}$$

Here we used $\log x \geq 1 - \frac{1}{x}$. At last, straightforward computations yield that

$$d\mathscr{E}(X) \leq 2E(X)\left(\frac{E(X^2)}{E(X^2) - E(X)} - 1\right) + \frac{1}{2}\left(\frac{E(X^2)}{E(X)} - 1\right). \tag{3.5}$$

□

In the following, we want to find the introduced formula for $d\mathscr{E}(X)$. Let X be a nonnegative discrete random variable with probability mass function $p(x)$ and survival function $\overline{F}(x)$. Assume that $0 \equiv X_0 \leq X_1 \leq X_2 \leq \ldots$ denote the epoch times, where $X_n, n \geq 1$, denote the time until the nth event of a discrete-time stochastic process and X_1 has the same distribution as X. In this case, $X_{n+1} - X_n$, describes the duration of the interepoch intervals or the interoccurrence times. By denoting the survival function of X_2 as $\overline{F}_2(t)$, it follows that

$$\overline{F}_2(t) = \sum_{k=0}^{1} P(k \text{ failures in } [0,t)) = \overline{F}(t) + \sum_{i=1}^{t-1} r(i)\overline{F}(t) = \overline{F}(t)(1 - \log \overline{F}(t)). \tag{3.6}$$

Now, from (3.6) we obtain

$$d\mathscr{E}(X) = E[X_2 - X_1] = \sum_{x=0}^{\infty}[\overline{F}_2(x) - \overline{F}_1(x)] = -\sum_{x=0}^{\infty} \overline{F}(x) \log \overline{F}(x), \tag{3.7}$$

and also $p_2(x) = p(x)(1-\log \overline{F}(x)) - \overline{F}(x+1)r(x)$. The equation (3.6) corresponds to the case $n = 1$ of Corollary 4.1 of Kapodistria and Psarrakos (2012) for continuous distribution function.

Hereafter, we obtain a lower bound for the variance of the random variable X, in terms of the mean residual life function.

Theorem 3.2. *Let X be a discrete random variable with support N_0, then*

$$E[\mu_F^2(X)] \leq Var(X). \tag{3.8}$$

Proof. By using (1.7), we have

$$E[\mu_F^2(X)] = \sum_{x=0}^{\infty}\Big(\sum_{y=x+1}^{\infty}\frac{\overline{F}(y)}{\overline{F}(x)}\Big)^2 p(x) \leq \sum_{x=0}^{\infty}\frac{(\sum_{y=x+1}^{\infty}\overline{F}(y))^2}{\overline{F}(x)\overline{F}(x+1)}p(x). \tag{3.9}$$

Applying summation by parts

$$\sum_{a}^{b} u\Delta v = uv|_a^b - \sum_{a}^{b} Sv\Delta u, \tag{3.10}$$

where Sv is a shift operator, i.e. $Sv(x) = v(x+1)$, we have, with considering $(\sum_{y=x+1}^{\infty}\overline{F}(y))^2 = u$ and $\frac{1}{\overline{F}(x)} = v$, and therefore $\Delta u = \overline{F}^2(x+1) - 2\overline{F}(x+1)\sum_{y=x+1}^{\infty}\overline{F}(y)$, and also

$$\Delta v = \frac{p(x)}{\overline{F}(x)\overline{F}(x+1)}, \tag{3.11}$$

$$
\begin{aligned}
E[\mu_F^2(X)] &\le \frac{(\sum_{y=x+1}^{\infty}\overline{F}(y))^2}{\overline{F}(x)}\Big|_0^{\infty} - \sum_{x=1}^{\infty}\frac{\overline{F}^2(x) - 2\overline{F}(x)\sum_{y=x}^{\infty}\overline{F}(y)}{\overline{F}(x)} \\
&= -E^2(X) - \sum_{x=1}^{\infty}\overline{F}(x) + 2\sum_{x=1}^{\infty}\sum_{y=x}^{\infty}\overline{F}(y) \\
&= -E^2(X) - E(X) + 2\sum_{y=1}^{\infty}\sum_{x=1}^{y}\overline{F}(y) \\
&= -E^2(X) - E(X) + 2(\frac{1}{2}\left[E(X^2) + E(X)\right]) \\
&= Var[X].
\end{aligned}
$$

□

For an absolutely continuous nonnegative random variable, Toomaj and Di Crescenzo (2020) proved that $E[\mu_F^2(X)] = Var(X)$, and by giving an example, they stated that the equality does not hold for a discrete random variable.

Now we must notice that, since $E^2(\mu_F(X)) \le E(\mu_F^2(X)) \le Var(X)$, consequently $E(\mu_F(X)) \le \sqrt{Var(X)}$.

Szymkowiak and Iwińska (2016) obtained a lower bound for $E[\mu_F(X)]$. Also Baratpour and Bami (2012) proved $E[\mu_F(X)] \le d\mathscr{E}(X)$ where $d\mathscr{E}(X)$ is defined in equation (3.1). In the following proposition, we obtain the other lower bound on $E[\mu_F(X)]$ in terms of cumulative residual Tsallis entropy of order 2 as $\xi_2(X) = \sum_{x=1}^{\infty}(\overline{F}(x) - \overline{F}^2(x))$, (see Rajesh and Sunoj (2019)).

Proposition 3.1. *Let X be a nonnegative discrete random variable, then*

$$E[\mu_F(X)] \ge \xi_2(X). \tag{3.12}$$

Proof. By using (1.7), we have

$$
\begin{aligned}
E[\mu_F(X)] &= \sum_{k=0}^{\infty}\sum_{t=k+1}^{\infty}\frac{\overline{F}(t)}{\overline{F}(k)}p(k) = \sum_{t=1}^{\infty}\overline{F}(t)\sum_{k=0}^{t-1}\frac{p(k)}{\overline{F}(k)} \\
&\ge \sum_{t=1}^{\infty}\overline{F}(t)\sum_{k=0}^{t-1}p(k) = \sum_{t=1}^{\infty}\overline{F}(t)(1 - \overline{F}(t)) = \xi_2(X).
\end{aligned}
$$

□

Example 3.1. Let X have a three point discrete distribution with the pmf $p(0) = \frac{1}{2}$, $p(1) = \frac{1}{4}$ and $p(2) = \frac{1}{4}$. Then from (1.7), we obtain $\mu_F(0) = \frac{3}{4}$ and $\mu_F(1) = \frac{1}{2}$ while $\mu_F(2) = 0$. It is obvious that $\sigma_F^2(X) = \frac{11}{16}$. However, we get

$$E[\mu_F^2(X)] = 11/32 \le \sigma_F^2(X).$$

On the other hand, since $d\mathscr{E}(X) = 0.693$, we can show that $E[\mu_F(X)] \leq d\mathscr{E}(X) \leq \sigma_F(X)$.

Also we can find lower bounds for $E[\mu_F(X)]$ using the lower bound given in Szymkowiak and Iwińska (2016) and inequality (3.12). Simple algebraic calculations show that the lower bounds are 0 and 7/16 respectively, and thus the bound (3.12) is sharper.

Example 3.2. Let X be a geometric random variable with probability mass function $p(x) = pq^x$, $x = 0, 1, \ldots$. Then $E[\mu_F^2(X)] = (\frac{q}{p})^2 \leq \frac{q}{p^2} = Var[X]$, also $E[\mu_F(X)] = \frac{q}{p} \leq -\frac{q}{p^2}\log q = d\mathscr{E}(X)$ because $\log q \leq q - 1$ and hence $-\frac{\log q}{p} \geq \frac{1}{\sqrt{q}}$.

Theorem 3.3. *Let X and Y be nonnegative discrete random variables with mean residual life functions $\mu_X(t)$ and $\mu_Y(t)$, respectively. Let $X \leq_{st} Y$ and $X \leq_{mrl} Y$. If either X or Y is IMRL; then $E[\mu_X^2(X)] \leq \sigma^2(Y)$*

Proof. Let Y be IMRL. From (3.8), we get

$$E[\mu_X^2(X)] \leq E[\mu_Y^2(X)] \leq E[\mu_Y^2(Y)] \leq \sigma^2(Y). \tag{3.13}$$

The first inequality is obtained from the assumption $X \leq_{mrl} Y$ while the last inequality is obtained by the fact that $X \leq_{st} Y$, which implies that $E(g(X)) \leq E(g(Y))$ for all increasing functions $g(\cdot)$. Now let X be IMRL. Then, we similarly have

$$E[\mu_X^2(X)] \leq E[\mu_X^2(Y)] \leq E[\mu_Y^2(Y)] \leq \sigma^2(Y), \tag{3.14}$$

and hence the result stated is obtained. □

The proof of Theorem 3.3 proceeds similarly as the proof of Theorem 2 of Toomaj and Di Crescenzo (2020).

Example 3.3. Let X and Y have discrete Weibull distributions with distribution functions $\overline{F}(x) = (0.2)^{x^2}$ and $\overline{G}(x) = (0.2)^x$ for $x = 0, 1, \ldots$, respectively. Then $E[\mu_X(X)^2] \leq \sigma^2(Y) = 0.2/(0.8^2) = 0.3125$

4 Conclusion

In this article, we obtained an upper bound for variance of a function of discrete residual life random variable T_n. In view of the importance of varentropy in the information theory, we investigated the discrete residual varentropy, that is the varentropy of the residual lifetime distribution for discrete random variables. We obtained an upper bound for it in terms of the mean residual life and eta function. Moreover, we presented the other upper bound for the discrete residual varentropy that involves the weighted residual entropy. We analyzed the effect of linear transformations on the discrete residual varentropy.

We also characterized the geometric distribution in terms of $E[T^3 r_F(T)]$ and $E[s^T r_F(T)]$. The geometric distribution can be characterized by $E[T^m r_F(T)]; m \leq 7$. Moreover, we characterized discrete Weibull distribution via Cauchy-Schwarz inequality.

Furthermore some results have been obtained for discrete cumulative residual entropy and some stochastic orders. Also we showed that, the variance of a discrete random variable is an upper bound for the expectation of the square of the mean residual life. We gave a lower bound for the expectation of the mean residual life in terms of cumulative residual Tsallis entropy of order 2.

References

Almalki, S.J., Nadarajah, S. (2014). Modifications of the Weibull distribution: A review, *Reliability Engineering and System Safety*, **124**, 32–55.

Baratpour, S. and Bami, Z. (2012). On the Discrete Cumulative Residual Entropy. *Journal of the iranian statistical society*, **11**, 203-215.

Cacoullos, T. (1982). On upper and lower bounds for the variance of the function of a random variable. *Annals of Probability*, **10**, 799-809.

Cacoullos, T. and Papathanasiou, V. (1985). On upper bounds for the variance of functions of random variables. *Statistics and Probability Letters*, **3**, 175-184.

Cover, T. M. and Thomas J. A. (2006). elements of information theory. Second Edition. Wiley Interscience.

Di Crescenzo, A. and Paolillo, L. (2021). Analysis and applications of the residual varentropy of random lifetimes. *Probability in the Engineering and Informational Sciences*, **35**, 680-698.

Gupta, P. L. (2015). Properties of reliability functions of discrete distributions. *Communicatons in Statistics-Theorey and Methods*, **44(19)**, 4114-4131.

Gupta, R. C. (2006). Variance residual life function in reliability studies. *Metron-International Journal of Statistics LXIV*, **3**, 343-355.

Gupta, R. C. and Kirmani, S.N.U.A. (1998). *Residual Life Function in Reliability Studies*. In: Fronteriors of Reliability, Basu A.P., Basu S.K. and Mukhopadhyay S. (eds.), World Scientific, New Jersey.

Gupta, R. C., Kirmani, S.N.U.A. and Launer, R. L. (1987). On life distributions having monotone residual variance. *Probability in the Engineering and Informational Sciences*, **1**, 299-307.

Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (1997). On the monotonic properties of discrete failure rates. *Journal of Statistical Planning and Inference*, **65(2)**, 255-268.

Kapodistria, S., and Psarrakos, G. (2012). *Probability in the Engineering and Informational Sciences*, **26(1)**, 129-146.

Kemp, A.W. (2004). Classes of discrete lifetime distributions. *Communication in Statistic Theory and Methods*, **33**, 3069-3093.

Rajesh, G. and Sunoj, S. M. (2019). Some properties of cumulative Tsallis entropy of order {*alpha*}. *Statistical Papers*, **60(3)**, 933–943.

Szymkowiak, M. and Iwińska, M. (2016). Characterizations of Discrete Weibull related distributions. *Statistics and Probability Letters*, **111**, 41-48.

Toomaj, A., and Di Crescenzo, A. (2020). Generalized entropies, variance and applications. *Entropy*, **22**, 709.

Xie, M., Gaudoin, O., Bracqemond, C. (2002). Redefining failure rate function for discrete distribution. *International Journal of Reliability, Quality Saftey Enginearing*, **9(3)**, 275–285.