Research Article Open Access

Jeonghwa Lee*

Generalized Bernoulli process with long-range dependence and fractional binomial distribution

https://doi.org/10.1515/demo-2021-0100
Received October 23, 2020; accepted January 22, 2021

Abstract: Bernoulli process is a finite or infinite sequence of independent binary variables, $X_i, i = 1, 2, \cdots$, whose outcome is either 1 or 0 with probability $P(X_i = 1) = p, P(X_i = 0) = 1 - p$, for a fixed constant $p \in (0, 1)$. We will relax the independence condition of Bernoulli variables, and develop a generalized Bernoulli process that is stationary and has auto-covariance function that obeys power law with exponent $2H - 2, H \in (0, 1)$. Generalized Bernoulli process encompasses various forms of binary sequence from an independent binary sequence to a binary sequence that has long-range dependence. Fractional binomial random variable is defined as the sum of n consecutive variables in a generalized Bernoulli process, of particular interest is when its variance is proportional to n^{2H}, if $H \in (1/2, 1)$.

Keywords: Bernoulli process, Long-range dependence, Hurst exponent, over-dispersed binomial model

MSC: 60G10, 60G22

1 Introduction

Fractional process has been of interest due to its usefulness in capturing long lasting dependency in a stochastic process called long-range dependence, and has been rapidly developed for the last few decades. It has been applied to internet traffic, queueing networks, hydrology data, etc (see [3, 7, 10]). Among the most well known models are fractional Gaussian noise, fractional Brownian motion, and fractional Poisson process.

Fractional Brownian motion(fBm) $B_H(t)$ developed by Mandelbrot, B. and Van Ness, J. (1968) [11] is zero-mean, increment stationary Gaussian process that has self-similarity, i.e., for any $a > 0$,

$$\{B_H(at), t \in \mathbb{R}\} \overset{law}{=} \{a^H B_H(t), t \in \mathbb{R}\}$$

where $0 < H < 1$ is called Hurst parameter, and $\overset{law}{=}$ means equality of all the finite distributions. FBm has auto-covariance function

$$cov(B_H(s), B_H(t)) = \frac{var(B_H(1))}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right).$$

Especially, the variance of $B_H(t)$ is $var(B_H(1))|t|^{2H}$. If $H = 1/2$, then fBm becomes standard Brownian motion which has independent increment.

The one time difference of fBm, $Y_j = B_H(j) - B_H(j-1)$ for $j \in \mathbb{Z}$, is called fractional Gaussian noise(fGn) which is mean-zero, stationary Gaussian process with auto-covariance function

$$\gamma(j) := cov(Y_0, Y_j) = \frac{var(Y_0)}{2}\left(|j+1|^{2H} - 2|j|^{2H} + |j-1|^{2H}\right).$$

The auto-covariance function obeys power law with exponent $2H - 2$ for large lag,

$$\gamma(j) \sim var(Y_0)H(2H-1)j^{2H-2} \text{ as } j \to \infty. \tag{1.1}$$

***Corresponding Author: Jeonghwa Lee:** Department of Statistics, Truman State University, USA, E-mail: jlee@truman.edu

If $1/2 < H < 1$, fGn has long-range dependence,

$$\sum_{j \in \mathbb{N}} cov(Y_0, Y_j) = \infty. \tag{1.2}$$

Note that a covariance function that decreases as slowly as the power law with exponent between 0 and -1 for large lag leads to long-range dependence of the process.

The time fractional Poisson process (TFPP) developed by Laskin, N. (2003) is a fractional generalization of Poisson process. Poisson process $N(t)$ is used for counting the number of events in time-interval $[0, t]$ where inter-arrival times between events are independent and exponentially distributed with parameter λ. As a result, it has Markov property with expected number of events $E(N(t)) = \lambda t$. In [9], TFPP, $N_\mu(t), 0 < \mu \leq 1$, is defined with Mittag-Leffler distributed waiting times where the inter-arrival distribution has heavy tail that decreases as slowly as $1/t^{\mu+1}$, and this makes TFPP non-Markov process, possesses long-range dependence, and has expectation that is proportional to the fractional exponent of the time, $E(N_\mu(t)) \propto t^\mu$. For other fractional generalizations of Poisson process, such as mixed fractional Poisson process and their properties, see [2, 4, 5, 8].

In this paper, we propose a generalized Bernoulli process(GBP) that is stationary and has long-range dependence, and fractional binomial distribution as the sum of consecutive variables in a generalized Bernoulli process. Bernoulli process is a finite or infinite sequence of independent binary variables, $X_i, i = 1, 2, \cdots$, whose outcome is either 1 or 0 with probability $P(X_i = 1) = p$, $P(X_i = 0) = 1-p$, for a fixed constant $p \in (0, 1)$. If we interpret $X_i = 1$ as a success in i-th trial, and $X_i = 0$ as a failure in i-th trial, then binomial variable that is the sum of n binary variables in Bernoulli process counts the number of successes among n trials, and the binomial variable has expectation np and variance $np(1-p)$.

When applying binomial model in real data analysis, over-dispersion is frequently observed. Over-dispersion is a term referring to the phenomenon that larger variance is observed than the nominal variance under some presumed model. It occurs when assumptions in the presumed model are violated. In binomial model, it is assumed that trials are independent and each has the same probability of success. Due to dependence among trials or heterogeneous probability of success in trials, over-dispersed binomial model is widely used in application. See [6] for more account for over-dispersion in various statistical models.

In this paper, we will relax the independence condition of Bernoulli variables so that binary variables have dependence with the auto-covariance which decreases with power law with exponent $2H - 2$, for $H \in (0, 1)$. If $H \in (1/2, 1)$, the generalized Bernoulli process possesses long-range dependence, therefore, fractional binomial random variable, which is defined as the sum of n successive variables in a GBP, has larger variance than ordinary binomial random variable.

Over-dispersion parameter $\psi_n > 0$ is defined and incorporated in binomial model as follows:

$$E(B_n) = np, \quad Var(B_n) = np(1-p)(1+\psi_n), \tag{1.3}$$

where B_n is over-dispersed binomial variable that has n number of trials and probability of success p. Among methods that incorporate over-dispersion in binomial distribution are beta-binomial model and generalized linear mixed model, [1, 6]. It turned out that the fractional binomial variable defined from a GBP shows over-dispersion (1.3) with various ψ_n. In particular, ψ_n is proportional to fractional exponent of n, $\psi_n \propto n^{2H-1}$, when $H \in (1/2, 1)$.

The contents of the paper are as follows. In section 2, we construct a generalized Bernoulli process whose auto-covariance function decreases with the power law of exponent $2H - 2$. In section 3, some interesting properties of a GBP are investigated regrading conditional probability. In section 4, fractional binomial variable is defined from a GBP, and the connection of fractional binomial distribution to the over-dispersed binomial distribution is investigated.

Throughout this paper, $\{i, i_0, i_1, \cdots\}, \{i', i'_0, i'_1, \cdots\} \subset \mathbb{N}$, and for any set $A = \{i_0, i_1, \cdots, i_n\}$, $|A| = n+1$, the number of elements in the set A, with $|A| = 0$, if $A = \emptyset$. We also define the maximum of empty set as 0, i.e., $\max \emptyset = 0$.

2 Generalized Bernoulli process

We will define stationary process, $\{X_i, i \in \mathbb{N}\}$, where each X_i takes one of two possible outcomes, 0 or 1, with $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$, and

$$cov(X_i, X_j) = c'|i-j|^{2H-2}, i \neq j,$$

for some constants $c' \in \mathbb{R}_+$, $H \in (0,1)$. If $H \in (1/2, 1)$,

$$\sum_{i=1}^{n} cov(X_1, X_i) \sim \frac{c'}{2H-1}|n|^{2H-1},$$

which diverges as n increases, and the process is said to have long-range dependence.

Below is how we will proceed to define such stationary process. Let X_i be a binary variable with

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p, \quad \text{for a fixed } p \in (0,1). \tag{2.1}$$

Define the following function P^* with constants $c > 0$, $H \in (0,1)$, and for any $i_0, i_1, \cdots, i_n$,

$$P^*(X_{i_0} = 1) := p,$$
$$P^*(X_{i_0} = 1, X_{i_1} = 1) := p\left(p + c|i_1 - i_0|^{2H-2}\right),$$
$$P^*(X_{i_0} = 1, X_{i_1} = 1, X_{i_2} = 1) := p\left(p + c|i_1 - i_0|^{2H-2}\right)\left(p + c|i_2 - i_1|^{2H-2}\right),$$

and in general,

$$P^*(X_{i_0} = 1, X_{i_1} = 1, \cdots, X_{i_n} = 1) := p\left(p + c|i_1 - i_0|^{2H-2}\right)\left(p + c|i_2 - i_1|^{2H-2}\right) \times \cdots \times \left(p + c|i_n - i_{n-1}|^{2H-2}\right). \tag{2.2}$$

Furthermore, for any disjoint sets, $A, B \subset \mathbb{N}$, $A \neq \emptyset$, $B \neq \emptyset$, define

$$\begin{aligned} P^*\left((\cap_{i' \in B}\{X_{i'} = 0\}) \cap (\cap_{i \in A}\{X_i = 1\})\right) &:= P^*(\cap_{i \in A}\{X_i = 1\}) + \sum_{k=1}^{|B|} \sum_{\substack{B' \subset B \\ |B'| = k}} (-1)^k P^*(\cap_{i \in B' \cup A}\{X_i = 1\}) \\ &= P^*(\cap_{i \in A}\{X_i = 1\}) - \sum_{j \in B} P^*(\cap_{i \in A \cup \{j\}}\{X_i = 1\}) + \sum_{\substack{j_1 < j_2 \\ j_1, j_2 \in B}} P^*(\cap_{i \in A \cup \{j_1, j_2\}}\{X_i = 1\}) \\ &+ \cdots (-1)^{|B|} P^*(\cap_{i \in A \cup B}\{X_i = 1\}), \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} P^*(\cap_{i \in B}\{X_i = 0\}) &:= 1 + \sum_{k=1}^{|B|} \sum_{\substack{B' \subset B \\ |B'| = k}} (-1)^k P^*(\cap_{i \in B'}\{X_i = 1\}) \\ &= 1 - \sum_{i \in B} P^*(\{X_i = 1\}) + \sum_{\substack{j_1 < j_2 \\ j_1, j_2 \in B}} P^*(\cap_{i \in \{j_1, j_2\}}\{X_i = 1\}) + \cdots (-1)^{|B|} P^*(\cap_{i \in B}\{X_i = 1\}). \end{aligned} \tag{2.4}$$

If we can show that P^* is non-negative in (2.3-2.4) for any disjoint sets, $A, B \subset \mathbb{N}$, then $\{X_i, i \in \mathbb{N}\}$ is stationary process with probability $P = P^*$.

2.1 Defining generalized Bernoulli process

Definition 2.1. Define the following operation on a set $A = \{i_0, i_1, \cdots, i_n\} \subset N$ with $i_0 < i_1 < \cdots < i_n$.

$$L_H(A) := \prod_{j=1,\cdots,n} \left(p + c|i_j - i_{j-1}|^{2H-2}\right).$$

If $A = \emptyset$, define $L_H(A) := 1/p$, and if $|A| = 1$, $L_H(A) := 1$.

For example, if $A = \{1, 4, 9\}$, $L_H(A) = \left(p + c|4-1|^{2H-2}\right)\left(p + c|9-4|^{2H-2}\right)$.

Definition 2.2. Define for disjoint sets, $A, B \subset \mathbb{N}$ with $|B| = m > 0$,

$$D_H(A,B) := L_H(A) - \sum_{i' \in B} L_H(A \cup \{i'\}) + \sum_{\substack{i'<i'',\\ i',i'' \in B}} L_H(A \cup \{i', i''\}) + \cdots (-1)^m L_H(A \cup B).$$

If $B = \emptyset$, $D_H(A, B) := L_H(A)$.

Note that (2.2-2.4) can be expressed as

$$P^\star\left((\cap_{i\in A}\{X_i = 1\}) \cap (\cap_{i'\in B}\{X_{i'} = 0\})\right) = pD_H(A, B), \tag{2.5}$$

for disjoint sets, $A, B \subset \mathbb{N}, A \cup B \neq \emptyset$. Now we give an assumption on constants, p, H, c, so that (2.5) is always positive.

ASSUMPTION (A1)
i) $0 < p, H < 1$.
ii)

$$0 \le c < \min\{1 - p, \frac{1}{2}(-2p + 2^{2H-2} + \sqrt{4p - p2^{2H} + 2^{4H-4}})\}. \tag{2.6}$$

Throughout this paper, it is assumed that $p, H \in (0, 1)$ and (2.6) holds.

Lemma 2.1. *For any* $i_0 < i_1 < i_2$,

$$\left(p + c|i_1 - i_0|^{2H-2}\right)\left(p + c|i_2 - i_1|^{2H-2}\right) < p + c|i_2 - i_0|^{2H-2}. \tag{2.7}$$

See proof on page 7. The Lemma 2.1 states that

$$D_H(\{i_0, i_2\}, \{i_1\}) = L_H(\{i_0, i_2\}) - L_H(\{i_0, i_1, i_2\}) > 0,$$

which is extended in the following Proposition.

Proposition 2.2. *For any disjoint sets,* $A_0, A_1 \subset \mathbb{N}, A_0 \cup A_1 \neq \emptyset$,

$$D_H(A_1, A_0) > 0. \tag{2.8}$$

See proof on page 9.
By (2.5) and Proposition 2.2, it was proved that $P^\star$ is non-negative function, and therefore, $\{X_i, i \in \mathbb{N}\}$ is a stationary process with probability $P = P^\star$ under the assumption (A1).

Theorem 2.3. *For binary variables* $\{X_i, i \in \mathbb{N}\}$ *that satisfy (2.1), if the joint probability is defined with* $P = P^\star$ *as (2.2-2.4) under the assumption (A1), then* $\{X_i, i \in \mathbb{N}\}$ *is a stationary process with covariance*

$$Cov(X_i, X_j) = pc|i - j|^{2H-2}, i \neq j. \tag{2.9}$$

Proof. By proposition 2.2, (2.3-2.4) are always positive, therefore, $\{X_i, i \in \mathbb{N}\}$ with probability $P = P^\star$ is well defined. It is a stationary process with $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$, and

$$P(X_i = 1, X_j = 1) = p^2 + pc|i - j|^{2H-2},$$

from which (2.9) is derived.

□

We call this stationary process with probability $P = P^\star$ a generalized Bernoulli process(GBP), and for the rest of the paper $\{X_i, i \in \mathbb{N}\}$ is assumed to be a GBP.

2.2 Properties of GBP

Next, we consider the conditional probability in a GBP. Following Theorem shows the result concerning the conditional probability of next observation given past observations when the last observation was 1. Recall that we define the maximum of empty set as zero, i.e., $\max \emptyset = 0$.

Theorem 2.4. *Let A_0, A_1 be disjoint subsets of $\mathbb{N}$ such that $\max A_1 > \max A_0$. Then, for $i^\star > \max A_1$, the conditional probability satisfies the following:*

$$P\Big(\{X_{i^\star} = 1\} | (\cap_{i' \in A_0} \{X_{i'} = 0\}) \cap (\cap_{i \in A_1} \{X_i = 1\})\Big) = p + c|i^\star - \max A_1|^{2H-2}. \tag{2.10}$$

If it is called "success" at time i for $X_i = 1$, and "failure" at time i for $X_i = 0$, (2.10) implies that the conditional probability of success in next observation given past observations only depends on the time difference between last observation and next observation, provided that the last observation was success. Furthermore, the conditional probability decreases as the time difference increase by power law with exponent $2H - 2$.

See the proof on page 10. Next we show the results for the conditional probability when the last observation was 0.

Theorem 2.5. *Let A_0, A_1 be disjoint subsets of $\mathbb{N}$ such that $\max A_1 < \max A_0$. Then, for $i^\star > i^{\star\star} > \max A_0$, the conditional probability satisfies the following:*

i)

$$p + c|i^\star - \max A_1|^{2H-2} > P\Big(\{X_{i^\star} = 1\} | (\cap_{i' \in A_0} X_{i'} = 0) \cap (\cap_{i \in A_1} X_i = 1)\Big). \tag{2.11}$$

ii)

$$\frac{P\Big(\{X_{i^\star} = 1\} | (\cap_{i' \in A_0} \{X_{i'} = 0\}) \cap (\cap_{i \in A_1} \{X_i = 1\})\Big)}{P\Big(\{X_{i^{\star\star}} = 1\} | (\cap_{i' \in A_0} \{X_{i'} = 0\}) \cap (\cap_{i \in A_1} \{X_i = 1\})\Big)} > \frac{p + |i^\star - \max A_1|^{2H-2}}{p + |i^{\star\star} - \max A_1|^{2H-2}}. \tag{2.12}$$

See proof on Page 10.

Remark 1. Theorem 2.4 and 2.5 imply that a GBP has what we will call "conditioned Markov property", which means that the conditional probability depends only on the time difference between the last observation and the next time of prediction, regardless of the past observations, when the last observation was success, and this is not true when the latest observation was failure.

Remark 2. Note that from (2.9) both the parameters c and H determine the strength of the dependence in the process. The larger the parameters, c, H, are, the stronger the correlation is, $Corr(X_i, X_j) = c|i-j|^{2H-2}/(1-p)$, for $i \neq j$. Especially, if $c = 0$ in (2.6), then a GBP becomes identical to Bernoulli process that has no dependence between any two variables in the process. Therefore, GBP can be considered as generalization of Bernoulli process that possesses long-range dependence if $c \neq 0$ and $H \in (1/2, 1)$.

In fact, GBP is the unique binary stationary process that has conditioned Markov property and covariance function (2.9).

Theorem 2.6. $\{X_i, i \in \mathbb{N}\}$ *is the unique stationary process that satisfies*

i)

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p, \quad p \in (0, 1),$$

ii)

$$Cov(X_i, X_j) = c|i - j|^{2H-2}, \quad i \neq j, \quad \textit{for some } c > 0,$$

iii) there is a function $h(\cdot)$ such that

$$P\Big(\{X_{i^*}=1\}|(\cap_{i'\in A_0}X_{i'}=0)\cap(\cap_{i\in A_1}X_i=1)\Big)=h(i^*-\max A_1),$$

for any $i^*\in\mathbb{N}$, and any disjoint subsets, $A_0, A_1\subset\mathbb{N}$, such that $i^*>\max A_1>A_0$.

Proof. Let $\{X_i, i\in\mathbb{N}\}$ is a stationary process that satisfies *i)-iii)*. Note that by *i)*, *ii)*,

$$P(X_{i_0}=1, X_{i_1}=1)=Cov(X_{i_0},X_{i_1})+p^2=c|i_0-i_1|^{2H-2}+p^2,$$

which results in $P(X_{i_1}=1|X_{i_0}=1)=p+c/p|i_0-i_1|^{2H-2}$. Therefore, by *iii)*, $h(i_0-i_1)=p+c/p|i_0-i_1|^{2H-2}$. By applying *iii)* repeatedly, we obtain

$$\begin{aligned}P(\cap_{i\in A_1}\{X_i=1\})&=P(X_{i_0}=1)\times P(X_{i_1}=1|X_{i_0}=1)\times P(X_{i_2}=1|X_{i_1}=1,X_{i_0}=1)\\&\times\cdots\times P(X_{i_n}=1|X_{i_{n-1}}=1,\cdots X_{i_0}=1)\\&=p\prod_{k=1,\cdots,n}h(i_k-i_{k-1}),\end{aligned}$$

which is equivalent to (2.2). □

3 Fractional binomial distribution

A GBP $\{X_i, i\in\mathbb{N}\}$ defined in Section 2 is stationary process where each X_i has two outcomes, 0 or 1, with $P(X_i=1)=p$, $P(X_i=0)=1-p$, and $cov(X_i,X_j)=c'|i-j|^{2H-2}$, $i\neq j$, for some constants $c'\in\mathbb{R}_+$, $H\in(0,1)$. Here, $c'=cp$ replaces cp in previous sections. If $H\in(1/2,1)$, the process has long-range dependence. Let $S_n=\sum_{i=1}^n X_i/n$. Then $E(S_n)=p$, and as $n\to\infty$,

$$Var(S_n)\sim\begin{cases}\left(p(1-p)+\dfrac{c'}{2H-1}\right)n^{-1} & H\in(0,1/2),\\ c'\dfrac{\ln n}{n} & H=1/2,\\ \dfrac{c'}{2H-1}|n|^{2H-2} & H\in(1/2,1).\end{cases}$$

Define $B_n=\sum_{i=1}^n X_i$. It follows that $E(B_n)=np$, and as $n\to\infty$,

$$Var(B_n)\sim\begin{cases}\left(p(1-p)+\dfrac{c'}{2H-1}\right)n & H\in(0,1/2),\\ c'n\ln n & H=1/2,\\ \dfrac{c'}{2H-1}|n|^{2H}, & H\in(1/2,1).\end{cases}$$

We call B_n fractional binomial random variable. It becomes binomial variable with parameter n, p, if $c'=0$. If $c'\neq 0$, B_n is over-dispersed binomial distribution with over-dispersion parameter,

$$\psi_n\sim\begin{cases}\dfrac{c'}{p(1-p)(2H-1)} & H\in(0,1/2),\\ \dfrac{c'\ln n}{p(1-p)}-1 & H=1/2,\\ \dfrac{c'n^{2H-1}}{p(1-p)2H-1}-1 & H\in(1/2,1),\end{cases}$$

as $n\to\infty$, by following the conventional notation of over-dispersion parameter,

$$E(B_n)=np,\quad Var(B_n)=np(1-p)(1+\psi_n).$$

If $H \in (0, 1/2)$, the over-dispersion parameter ψ_n is asymptotically constant as n increases, whereas for $H \in (1/2, 1)$, ψ_n increases with the rate of fractional exponent of n. Therefore, fractional binomial distribution can model various over-dispersed binomial random variable.

4 Conclusion

Over-dispersion in various models including binomial model is well observed in real data, and one of the reasons for over-dispersion in binomial model is explained by dependence among trials. In this paper, the independence assumption on binary sequence in Bernoulli process was relaxed, and GBP was defined as a binary sequence whose auto-covariance function decreases by power law with exponent $2H-2$, $H \in (0, 1)$. If $H \in (1/2, 1)$, the process has long-range dependence. There could be many other stationary models that feature binary sequences with dependence structure. However, GBP in this paper is the only stationary binary sequence that has the covariance function of the form,

$$Cov(X_i, X_j) = c|i-j|^{2H-2}, i \neq j, \text{ for some } c > 0, H \in (0, 1),$$

and has conditioned Markov property where the Markov property holds only when the last observation was in certain state. Also, from the conditioned Markov property and (2.2-2.4), it can be derived that the number of failures (number of 0's) between any two consecutive successes is independent of the number of failures between any other two consecutive successes and they are identically distributed.

Fractional binomial random variable was defined as the sum of n successive variables in GBP, and it can be used for various over-dispersed binomial model whose over-dispersion parameter ranges from asymptotic constant to fractional exponent of n. As over-dispersed binomial data is widely observed, the hope is that the fractional binomial model provides another useful tool in those circumstances.

5 Proofs

Proof of Lemma 2.1. Let $c^\star = c|i_2 - i_0|^{2H-2}$. Then, (2.7) becomes

$$\left(p + c^\star|x|^{2H-2}\right)\left(p + c^\star|1-x|^{2H-2}\right) < p + c^\star, \text{ where } 0 < x = \frac{|i_1 - i_0|}{|i_2 - i_0|} < 1.$$

Since $|x(1-x)|^{2H-2}$ and $x^{2H-2} + (1-x)^{2H-2}$ are both maximized when x has the smallest fraction possible (or the largest fraction possible), (2.7) holds if

$$\left(p + c\right)\left(p + c|i_2 - 1|^{2H-2}\right) < p + c|i_2|^{2H-2}.$$

Note that

$$\min_{i_2 \in \mathbb{N}/\{1\}} \frac{p + c|i_2|^{2H-2}}{p + c|i_2 - 1|^{2H-2}} = \frac{p + c|2|^{2H-2}}{p + c},$$

the minimum occurs when $i_2 = 2$. Therefore, (2.7) holds if

$$p + c < \frac{p + c|2|^{2H-2}}{p + c},$$

that is,

$$\frac{1}{2}\left(2^{2H-2} - 2p - \sqrt{(2^{2H-2} - 2p)^2 + 4(p - p^2)}\right) < c < \frac{1}{2}\left(2^{2H-2} - 2p + \sqrt{(2^{2H-2} - 2p)^2 + 4(p - p^2)}\right).$$

□

Lemma 5.1. *For any* $\{i_0, i_1, \cdots, i_n\} \subset \mathbb{N}$,

$$D_H(\{i_0, i_n\}, \{i_1, i_2, \cdots, i_{n-1}\}) = D_H(\{i_0, i_n\}, \{i_1, i_2, \cdots, i_{n-2}\}) - D_H(\{i_0, i_{n-1}, i_n\}, \{i_1, i_2, \cdots, i_{n-2}\}).$$

In general, for any $\{a_0, a_1, \cdots, a_n, b_0, b_1, \cdots, b_m\} \subset \mathbb{N}$,

$$\begin{aligned} D_H(\{a_0, a_1, \cdots, a_n\}, \{b_0, b_1, \cdots, b_m\}) &= D_H(\{a_0, a_1, \cdots, a_n\}, \{b_0\}) \\ - D_H(\{a_0, a_1, \cdots, a_n, b_1\}, \{b_0\}) &- D_H(\{a_0, a_1, \cdots, a_n, b_2\}, \{b_0, b_1\}) \\ \cdots - D_H(\{a_0, a_1, \cdots, a_n, b_m\}, &\{b_0, b_1, \cdots, b_{m-1}\}). \end{aligned} \tag{5.1}$$

Proof. Let $A = \{a_0, a_1, \cdots, a_n\}$, $B = \{b_0, b_1, \cdots, b_m\}$, then by the definition of D_H,

$$\begin{aligned} & D_H(\{a_0, a_1, \cdots, a_n\}, \{b_0, b_1, \cdots, b_m\}) = L_H(A) - \sum_{b \in B/\{b_m\}} L_H(A \cup \{b\}) \\ & + \sum_{\substack{b' < b'', \\ b', b'' \in B/\{b_m\}}} L_H(A \cup \{b', b''\}) + \cdots (-1)^m L_H(A \cup B/\{b_m\}) \\ & - \Big(L_H(A \cup \{b_m\}) - \sum_{b \in B/\{b_m\}} L_H(A \cup \{b_m\} \cup \{b\}) \\ & + \sum_{\substack{b' < b'', \\ b', b'' \in B/\{b_m\}}} L_H(A \cup \{b_m\} \cup \{b', b''\}) + \cdots (-1)^m L_H(A \cup B)\Big) \\ & = D_H(\{a_0, a_1, \cdots, a_n\}, \{b_0, b_1, \cdots, b_{m-1}\}) - D_H(\{a_0, a_1, \cdots, a_n, b_m\}, \{b_0, b_1, \cdots, b_{m-1}\}). \end{aligned}$$

Applying the above result inductively leads to (5.1). □

Lemma 5.2. *For any* $\{a_0, a_1, \cdots, a_n, a'_0, a'_1, \cdots, a'_n\} \subset \mathbb{R}_+$ *such that* $a_0 - \sum_{i=1}^{j} a_i > 0$, $a'_0 - \sum_{i=1}^{j} a'_i > 0$, *for* $j = 1, 2, \cdots, n$,
i) if

$$\frac{a_0}{a'_0} > \frac{a_1}{a'_1} > \cdots > \frac{a_n}{a'_n},$$

then

$$\frac{a_0 - a_1 - a_2 - \cdots - a_n}{a'_0 - a'_1 - a'_2 - \cdots - a'_n} > \frac{a_0}{a'_0}.$$

ii) If

$$\frac{a_0}{a'_0} < \frac{a_1}{a'_1} < \cdots < \frac{a_n}{a'_n},$$

then

$$\frac{a_0 - a_1 - a_2 - \cdots - a_n}{a'_0 - a'_1 - a'_2 - \cdots - a'_n} < \frac{a_0}{a'_0}.$$

Proof. We will prove *i)* by mathematical induction. *ii)* follows in the same way.

Let $n = 1$. Since $\frac{a_0}{a'_0} > \frac{a_1}{a'_1}$, $a_0 - a_1 > \frac{a_0}{a'_0}(a'_0 - a'_1)$, which leads to

$$\frac{a_0 - a_1}{a'_0 - a'_1} > \frac{a_0}{a'_0}.$$

Assume *i)* holds for $n = k$. Let $n = k + 1$. Since

$$\frac{a_0 - a_1 - \cdots - a_k}{a'_0 - a'_1 - \cdots - a'_k} > \frac{a_0}{a'_0} > \frac{a_{k+1}}{a'_{k+1}},$$

$(a_0 - a_1 - \cdots - a_k) - a_{k+1} > \frac{a_0}{a'_0}\{(a'_0 - a'_1 - \cdots - a'_k) - a'_{k+1}\}$, therefore, *i)* holds for $n = k + 1$. By mathematical induction, *i)* follows.

□

Proof of Proposition 2.2. If $A_0 = \emptyset$, $D_H(A_1, A_0) = L_H(A_1) > 0$.
Assume $A_0 \neq \emptyset$. We will prove (2.8) by mathematical induction. If $|A_0| = 1$, $D_H(A_1, A_0) = L_H(A_1) - L_H(A_1 \cup \{A_0\}) > 0$, by Lemma 2.1. Assume (2.8) holds with $|A_0| \leq m$. We will show that (2.8) holds when $|A_0| = m+1$. Let $A_0 = \{i'_0, i'_1, i'_2, \cdots, i'_m\}$ with $i'_0 < i'_1 < i'_2 < \cdots < i'_m$. By Lemma 5.1, it is enough to show

$$D_H(A_1, A_0/\{i'_m\}) > D_H(A_1 \cup \{i'_m\}, A_0/\{i'_m\}).$$

By (5.1),

$$\frac{D_H(A_1, A_0/\{i'_m\})}{D_H(A_1 \cup \{i'_m\}, A_0/\{i'_m\})} = \frac{a_0 - a_1 - a_2 - a_3 \cdots - a_{m-1}}{a'_0 - a'_1 - a'_2 - a'_3 \cdots - a'_{m-1}}, \tag{5.2}$$

where

$$a_j = \begin{cases} D_H(A_1 \cup \{i'_j\}, \{i'_0, i'_1, \cdots, i'_{j-1}\}) & \text{for } j = 1, \cdots, m-1, \\ D_H(A_1, \{i'_0\}) & \text{for } j = 0, \end{cases}$$

and

$$a'_j = \begin{cases} D_H(A_1 \cup \{i'_j, i'_m\}, \{i'_0, i'_1, \cdots, i'_{j-1}\}) & \text{for } j = 1, \cdots, m-1, \\ D_H(A_1 \cup \{i'_m\}, \{i'_0\}), & \text{for } j = 0. \end{cases}$$

We will apply Lemma 5.2 *i)* to show that (5.2) > 1. Note that $a_0 - a_1 - \cdots - a_j > 0$ and $a'_0 - a'_1 - \cdots - a'_j > 0$ for $j = 0, 1, \cdots, m-1$, by the earlier assumption that (2.8) holds for $|A_0| \leq m$.

Define $i^*_j = \max\{i'_j, i : i \in A_1, i < i'_m\}$, and $i^* = \min\{i : i \in A_1, i > i'_m\}$ if $\max A_1 > i'_m$. Note that i^*_j is non-decreasing as j increases from 1 to $m-1$. For $j = 1, \cdots, m-1$,

$$\frac{a_j}{a'_j} = \frac{D_H(A_1 \cup \{i'_j\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})}{D_H(A_1 \cup \{i'_j, i'_m\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})}$$
$$= \begin{cases} \dfrac{p + c|i^* - i^*_j|^{2H-2}}{(p + c|i'_m - i^*_j|^{2H-2})(p + c|i'_m - i^*|^{2H-2})} & \text{if } \max A_1 > i'_m, \\ \dfrac{1}{p + c|i'_m - i^*_j|^{2H-2}} & \text{if } \max A_1 < i'_m, \\ \dfrac{1}{p + c|i'_m - i'_j|^{2H-2}} & \text{if } A_1 = \emptyset, \end{cases} \tag{5.3}$$

which non-increases as j goes from 1 to $m-1$. Also,

$$\frac{a_0}{a'_0} = \frac{D_H(A_1, \{i'_0\})}{D_H(A_1 \cup \{i'_m\}, \{i'_0\})} = \frac{L_H(A_1) - L_H(A_1 \cup \{i'_0\})}{L_H(A_1 \cup \{i'_m\}) - L_H(A_1 \cup \{i'_0, i'_m\})}$$

with

$$\frac{L_H(A_1)}{L_H(A_1 \cup \{i'_m\})} = \begin{cases} \dfrac{1}{p + c|i'_m - i^*|^{2H-2}} & \text{if } \min A_1 > i'_m, \\ \dfrac{1}{p + c|i'_m - i_\star|^{2H-2}} & \text{if } \max A_1 < i'_m, \\ \dfrac{p + c|i^* - i_\star|^{2H-2}}{(p + c|i'_m - i^*|^{2H-2})(p + c|i'_m - i_\star|^{2H-2})} & \text{if } \min A_1 < i'_m < \max A_1, \\ 1/p & \text{if } A_1 = \emptyset, \end{cases}$$

where $i_\star = \max\{i : i \in A_1, i < i'_m\}$ if $\min A_1 < i'_m$, and

$$\frac{L_H(A_1 \cup \{i'_0\})}{L_H(A_1 \cup \{i'_0, i'_m\})} = (5.3)$$

with $j = 0$. Since $i_\star \leq i^*_0 \leq i^*_1 \leq i^*_2 \leq \cdots \leq i^*_{m-1}$,

$$\frac{L_H(A_1)}{L_H(A_1 \cup \{i'_m\})} \geq \frac{L_H(A_1 \cup \{i'_0\})}{L_H(A_1 \cup \{i'_0, i'_m\})} \geq \frac{D_H(A_1 \cup \{i'_1\}, \{i'_0\})}{D_H(A_1 \cup \{i'_1, i'_m\}, \{i'_0\})}$$

$$\geq \frac{D_H(A_1 \cup \{i'_2\}, \{i'_0, i'_1\})}{D_H(A_1 \cup \{i'_2, i'_m\}, \{i'_0, i'_1\})} \geq \cdots \geq \frac{D_H(A_1 \cup \{i'_{m-1}\}, \{i'_0, i'_1, \cdots, i'_{m-2}\})}{D_H(A_1 \cup \{i'_{m-1}, i'_m\}, \{i'_0, i'_1, \cdots, i'_{m-2}\})},$$

therefore, by Lemma 5.2 *i*),

$$(5.2) \geq \frac{L_H(A_1)}{L_H(A_1 \cup \{i'_m\})}.$$

The result follows as $\frac{L_H(A_1)}{L_H(A_1 \cup \{i'_m\})} > 1$ by Lemma 2.1. □

Proof of Theorem 2.4. (2.10) is derived from the fact that

$$P\Big((\cap_{i' \in A_0}\{X_{i'} = 0\}) \cap (\cap_{i \in A_1 \cup \{i^\star\}}\{X_i = 1\})\Big) =$$
$$P\Big((\cap_{i' \in A_0}\{X_{i'} = 0\}) \cap (\cap_{i \in A_1}\{X_i = 1\})\Big) \times \Big(p + c|i^\star - \max A_1|^{2H-2}\Big),$$

since there is no element i' such that $i' \in A_0$ and $\max A_1 < i' < i^\star$. □

Proof of Theorem 2.5. Let $A_0 = \{i'_0, i'_1, \cdots, i'_m\}$, $A_1 = \{i_0, i_1, \cdots, i_n\}$ with $\max A_1 = i_n < \max A_0 = i'_m$.
i) Note

$$\frac{P\Big((\cap_{i' \in A_0}\{X_{i'} = 0\}) \cap (\cap_{i \in A_1 \cup \{i^\star\}}\{X_i = 1\})\Big)}{P\Big((\cap_{i' \in A_0}\{X_{i'} = 0\}) \cap (\cap_{i \in A_1}\{X_i = 1\})\Big)} = \frac{D_H(A_1 \cup \{i^\star\}, A_0)}{D_H(A_1, A_0)}, \tag{5.4}$$

and by (5.1),

$$\frac{D_H(A_1 \cup \{i^\star\}, A_0)}{D_H(A_1, A_0)} = \frac{a_0 - a_1 - a_2 - a_3 \cdots - a_m}{a'_0 - a'_1 - a'_2 - a'_3 \cdots - a'_m}, \tag{5.5}$$

where

$$\frac{a_j}{a'_j} = \begin{cases} \dfrac{D_H(A_1 \cup \{i'_j, i^\star\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})}{D_H(A_1 \cup \{i'_j\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})} & \text{for } j = 1, \cdots, m, \\ \dfrac{D_H(A_1 \cup \{i^\star\}, \{i'_0\})}{D_H(A_1, \{i'_0\})} & \text{for } j = 0. \end{cases}$$

We obtain

$$\frac{a_0}{a'_0} = \frac{D_H(A_1 \cup \{i^\star\}, \{i'_0\})}{D_H(A_1, \{i'_0\})} = \frac{L_H(A_1 \cup \{i^\star\}) - L_H(A_1 \cup \{i^\star, i'_0\})}{L_H(A_1) - L_H(A_1 \cup \{i'_0\})} < p + c|i^\star - i_n|^{2H-2},$$

by

$$\begin{aligned} p + c|i^\star - i_n|^{2H-2} &= \frac{L_H(A_1 \cup \{i^\star\})}{L_H(A_1)} \\ &\leq \frac{L_H(A_1 \cup \{i^\star, i'_0\})}{L_H(A_1 \cup \{i'_0\})} = p + c|i^\star - \max\{i_n, i'_0\}|^{2H-2} \end{aligned}$$

and Lemma 5.2 *ii*). Note also that

$$\frac{D_H(A_1 \cup \{i^\star\}, \{i'_0\})}{D_H(A_1, \{i'_0\})} \leq \frac{D_H(A_1 \cup \{i^\star, i'_1\}, \{i'_0\})}{D_H(A_1 \cup \{i'_1\}, \{i'_0\})} = p + c|i^\star - \max\{i_n, i'_1\}|^{2H-2}.$$

As $i'_0 < i'_1 < \cdots < i'_j$, we have

$$\frac{D_H(A_1 \cup \{i^\star, i'_j\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})}{D_H(A_1 \cup \{i'_j\}, \{i'_0, i'_1, \cdots, i'_{j-1}\})} = p + c|i^\star - \max\{i_n, i'_j\}|^{2H-2},$$

which is non-decreasing as j increases from 1 to m. Applying Lemma 5.2 *ii*) leads to

$$\frac{D_H(A_1 \cup \{i^\star\}, A_0)}{D_H(A_1, A_0)} < p + c|i^\star - i_n|^{2H-2}.$$

ii) The proof is similar to *i*). First, note that

$$\frac{P\Big(\{X_{i^*}=1\}|(\cap_{i'\in A_0}\{X_{i'}=0\})\cap(\cap_{i\in A_1}\{X_i=1\})\Big)}{P\Big(\{X_{i^{**}}=1\}|(\cap_{i'\in A_0}\{X_{i'}=0\})\cap(\cap_{i\in A_1}\{X_i=1\})\Big)}=\frac{D_H(A_1\cup\{i^*\},A_0)}{D_H(A_1\cup\{i^{**}\},A_0)},$$

and by (5.1),

$$\frac{D_H(A_1\cup\{i^*\},A_0)}{D_H(A_1\cup\{i^{**}\},A_0)}=\frac{a_0-a_1-a_2-a_3\cdots-a_m}{a'_0-a'_1-a'_2-a'_3\cdots-a'_m},\tag{5.6}$$

where

$$\frac{a_j}{a'_j}=\begin{cases}\dfrac{D_H(A_1\cup\{i'_j,i^*\},\{i'_0,i'_1,\cdots,i'_{j-1}\})}{D_H(A_1\cup\{i'_j,i^{**}\},\{i'_0,i'_1,\cdots,i'_{j-1}\})} & \text{for } j=1,\cdots,m,\\ \dfrac{D_H(A_1\cup\{i^*\},\{i'_0\})}{D_H(A_1\cup\{i^{**}\},\{i'_0\})} & \text{for } j=0.\end{cases}$$

It is derived that

$$\begin{aligned}\frac{D_H(A_1\cup\{i^*\},\{i'_0\})}{D_H(A_1\cup\{i^{**}\},\{i'_0\})}&=\frac{L_H(A_1\cup\{i^*\})-L_H(A_1\cup\{i^*,i'_0\})}{L_H(A_1\cup\{i^{**}\})-L_H(A_1\cup\{i^{**},i'_0\})}\\&>\frac{p+c|i^*-i_n|^{2H-2}}{p+c|i^{**}-i_n|^{2H-2}},\end{aligned}$$

by Lemma 5.2 *i*) and the fact that

$$\begin{aligned}\frac{p+|i^*-i_n|^{2H-2}}{p+c|i^{**}-i_n|^{2H-2}}&=\frac{L_H(A_1\cup\{i^*\})}{L_H(A_1\cup\{i^{**}\})}\\&\geq\frac{L_H(A_1\cup\{i^*,i'_0\})}{L_H(A_1\cup\{i^{**},i'_0\})}=\frac{p+c|i^*-\max\{i_n,i'_0\}|^{2H-2}}{p+c|i^{**}-\max\{i_n,i'_0\}|^{2H-2}}.\end{aligned}$$

Note also that

$$\frac{D_H(A_1\cup\{i^*\},\{i'_0\})}{D_H(A_1\cup\{i^{**}\},\{i'_0\})}\geq\frac{D_H(A_1\cup\{i^*,i'_1\},\{i'_0\})}{D_H(A_1\cup\{i^{**},i'_1\},\{i'_0\})}=\frac{p+c|i^*-\max\{i_n,i'_1\}|^{2H-2}}{p+c|i^{**}-\max\{i_n,i'_1\}|^{2H-2}}.$$

As $i'_0<i'_1<\cdots<i'_j$, we have

$$\frac{D_H(A_1\cup\{i^*,i'_j\},\{i'_0,i'_1,\cdots,i'_{j-1}\})}{D_H(A_1\cup\{i^{**},i'_j\},\{i'_0,i'_1,\cdots,i'_{j-1}\})}=\frac{p+c|i^*-\max\{i_n,i'_j\}|^{2H-2}}{p+c|i^{**}-\max\{i_n,i'_j\}|^{2H-2}},$$

which is non-increasing as j increases from 1 to m. Applying Lemma 5.2 *i*) leads to

$$\frac{D_H(A_1\cup\{i^*\},A_0)}{D_H(A_1\cup\{i^{**}\},A_0)}>\frac{p+c|i^*-i_n|^{2H-2}}{p+c|i^{**}-i_n|^{2H-2}}.$$

□

References

[1] Ahn, H. and J. J. Chen (1995). Generation of over-dispersed and under-dispersed binomial variates. *J. Comput. Graph. Statist. 4*(1), 55–64.

[2] Aletti, G., N., Leonenko and E. Merzbach (2018). Fractional Poisson fields and martingales. *J. Stat. Phys. 170*(4), 700–730.

[3] Benson, D. A., Meerschaert, M. M., Baeumer, B. and Scheffler, H. P. (2006). Aquifer operator-scaling and the effect on solute mixing and dispersion. *Water Resour. Res. 42*(1), 1–18.

[4] Beghin, L. and E. Orsingher (2009). Fractional Poisson processes and related planar random motions. *Electron. J. Probab. 14*(61), 1790–1827.

[5] Beghin, L. (2012). Random-time processes governed by differential equations of fractional distributed order. *Chaos Solitons Fractals*. *45*(11), 1314–1327.
[6] Dean, C.B. and E. R., Lundy (2016). Overdispersion. *Wiley StatsRef: Statistics Reference Online*. Available at https://doi.org/10.1002/9781118445112.stat06788.pub2.
[7] Delgado, R. (2007). A reflected fBm limit for fluid models with ON/OFF sources under heavy traffic. *Stochastic Process. Appl*. *117*(2), 188–201.
[8] Kataria, K. K. and M. Khandakar (2020). On the long-range dependence of Mixed fractional Poisson process. *J. Theoret. Probab*. Available at https://doi.org/10.1007/s10959-020-01015-y.
[9] Laskin, N. (2003). Fractional Poisson process. *Commun. Nonlinear Sci. Numer. Simul*. *8*(3–4), 201–213.
[10] Majewski, K. (2005). Fractional Brownian heavy traffic approximations of multiclass feedforward queueing networks. *Queueing Syst*. *50*(2–3), 199–230.
[11] Mandelbrot, B. and J. Van Ness (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev*. *10*(4), 422–437.