Record Statistics

Mild taste for beginners

written by Sejin Lim

What is record event?

1. Record

: when the new highest value appears, it is called, “A record event occurs.”.

- For example : today’s highest temperature is the new record value until 10 years.

2. Event

: all of the observable values can be events.

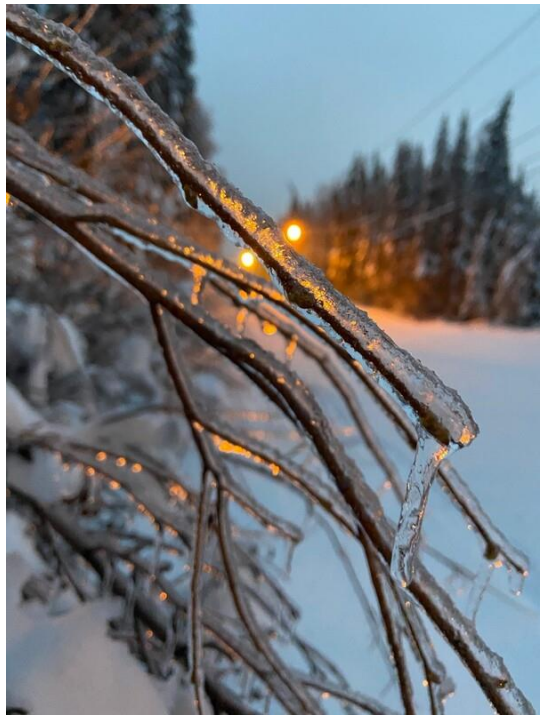
- For example : restaurant’s servings also can be events. (when a new food comes out, it can be called, “A new serving event occurs.”.)

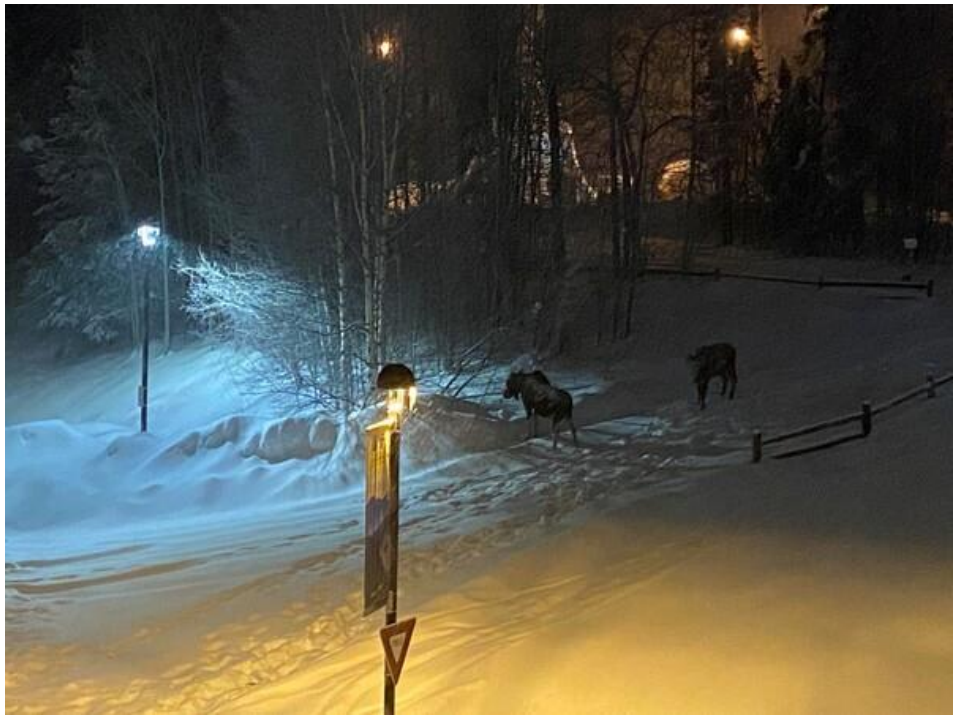
Record event

What is record event?

Record event

- Extreme weather in Alaska in December 2021.





https://www.newspenguin.com/news/articleView.html?idxno=10334

What is main concern of this study?

1. How can we analyze time series data more exactly?

2. Let's borrow "Record statistics" and try analyze time series data!

3. Analyzing Method
   - ① Prepare real time series data and lots of random walk simulation data.
   - ② Measure record statistics' variables of real data and simulation data.
   - ③ Compare the record statistics of real data with random walk simulation data.
   - ④ Find real data's parameters' regime.

Record statistics

Persistence (parameter examples : constant bias, levy exponent alpha)

Record number

Record Statistics

Mild taste(**?**) for beginners(**?**)

written by Sejin Lim

Simulation Methods

1. Simulation Method (Discrete-time Random walk)

$$x_n = x_{n-1} + \eta_n$$

- $x_0 = 0$ (origin), η : jump length
- n : each discrete time step, N : the number of total random walk steps
- N_s : the number of random walk data samples

- η_n : (identically independently distributed : i.i.d.) random variables from probability distribution $\phi(\eta)$
- $\phi(\eta)$ can be Normal distribution, or Cauchy distribution or Levy alpha-stable distribution.. And so on..

What is record event? (renewed..)

Record event's definition

: $x_n > x_0, x_1, \ldots x_{n-1}$ at discrete-time step n, x_n is a record value.

And it is called, "A record event occurs at time step n.".

Record event

What are main concerns of record statistics?

Record Statistics' variables

① Persistence : Q_n (the probability that a walker is always under the origin until step n)

② Record rate : r_n (the rate that a record event occurs at step n. By convention, $r_0 = 1$)

③ Record number : R_n (the number of record events)

④ Average record number : $\langle R_n \rangle$ (the sample average number of record events)

⑤ Variance of record number : V_n (the variance of record number)

⑥ Age statistics : $\min(\{l_i\})$, $\max(\{l_i\})$, $\quad i \in [0, \sum_{i=1}^{N_s} R_N^i]$ (inter-record time's statistics)

Definitions of record statistics' variables

① Definition of Persistence : $Q_n = Proba.[x_i > x_0 \text{ for all } i = 1,2,3, \dots n]$

② Definition of Average record number : $\langle R_n \rangle \equiv \sum_{n=1}^{N} r_n = \frac{1}{N_s} \sum_{i=1}^{N_s} R_n^i$

③ Variance of record number : $V_n = \langle R_n^2 \rangle - \langle R_n \rangle^2$

④ Age statistics : (let's skip..)

Record Statistics

Mild taste(**???**) for beginners(**???**)

written by Sejin Lim

Simulation Methods

1. Simulation Method (Discrete-time Random walk)

Strong correlation & non-biased case

① $\boldsymbol{x_n = x_{n-1} + \eta_n}$

- $x_0 = 0$ (origin), η : jump length
- n : each discrete time step, N : the number of total random walk steps
- N_s : the number of random walk data samples

- η_n : (identically independently distributed : i.i.d.) random variables from probability distribution $\phi(\eta)$
- $\phi(\eta)$ can be Normal distribution, or Cauchy distribution or Levy alpha-stable distribution.. And so on..

Simulation Methods

1. Simulation Method (Discrete-time Random walk)

Strong correlation & biased case (with a constant bias value : c)

① $x_n = x_{n-1} + \eta_n$

- $x_0 = 0$ (origin), η : jump length
- n : each discrete time step, N : the number of total random walk steps
- N_s : the number of random walk data samples

- η_n : (identically independently distributed : i.i.d.) random variables from probability distribution $\phi(\eta)$
- $\phi(\eta)$ can be Normal distribution, or Cauchy distribution or Levy alpha-stable distribution.. And so on..

② $\boldsymbol{y_n = y_{n-1} + \eta_n + c}$ (same with $y_n = x_n + cn$)

What is record event? (renewed..)

Record event's definition

: $y_n > y_0, y_1, \ldots y_{n-1}$ at discrete-time step n, y_n is a record value.

And it is called, "A record event occurs at time step n.".

Record event

What are main concerns of record statistics?

Record Statistics' variables

① Persistence : Q_n (the probability that a walker is always under the origin until step n)

② Record rate : r_n (the rate that a record event occurs at step n)

③ Record number : R_n (the number of record events)

④ Average record number : $\langle R_n \rangle$ (the sample average number of record events)

⑤ Variance of record number : V_n (the variance of record number)

⑥ Age statistics : $\min(\{l_i\})$, $\max(\{l_i\})$, $\quad i \in [0, \sum_{i=1}^{N_s} R_N^i]$ (inter-record time's statistics)

Persistence

Record number

Definitions of record statistics' variables

① Definition of Persistence : $Q_n = Proba.[y_i > y_0 \text{ for all } i = 1,2,3, \dots n]$

② Definition of Average record number : $\langle R_n \rangle \equiv \sum_{n=1}^{N} r_n = \frac{1}{N_s} \sum_{i=1}^{N_s} R_n^i$

③ Variance of record number : $V_n = \langle R_n^2 \rangle - \langle R_n \rangle^2$

④ Age statistics : (let's skip for healthy mental condition..)

How to calculate?

Materials for calculation…

① Sparre-Andersen Theorem (Q_n from $\phi(\eta)$)

② First-passage probability f_n from survival probability Q_n (=persistence)

③ Using generating functions of Q_n and f_n, we can get average record number $\langle R_n \rangle$

generating function $\tilde{Q}_z$ of Q_n

: $\tilde{Q}_z = \sum_{n=0}^{\infty} Q_n z^n$. $(|z| < 1.)$

Sparre-Andersen Theorem (non-biased case)

$$\tilde{Q}_z = \sum_{n=0}^{\infty} Q_n z^n = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} p_n z^n\right)$$

- $p_n \equiv Proba.[x_i < x_0 \quad \text{for any } i = 1,2,\dots,n]$
- $\tilde{\phi}(k) \equiv \int_{-\infty}^{\infty} \phi(\eta) e^{ik\eta} d\eta$
- $P_n(x) \equiv Proba.[x_n = x] = \int_{-\infty}^{\infty} P_{n-1}(x')\phi(x - x')dx' = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \left[\tilde{\phi}(k)\right]^n e^{-ikx}$
- For the Normal jump length distribution, (with $\langle \eta \rangle = 0$)
  - $\phi(\eta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\eta^2}{2\sigma^2}\right), \quad \tilde{\phi}(k) = \exp\left(-\frac{k^2}{2\sigma^2}\right)$
- $p_n = \int_0^{\infty} P_n(x)dx = \frac{1}{2}$

Record statistics

Persistence

Record number

Sparre-Andersen Theorem (non-biased case)

$$\tilde{Q}_z = \sum_{n=0}^{\infty} Q_n z^n = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} p_n z^n\right)$$

- $p_n = \int_0^{\infty} P_n(x)dx = \frac{1}{2}$
- $\sum_{n=1}^{\infty} \frac{1}{2n} z^n = -\frac{1}{2}\ln(1-z)$ (from generating function table by *Ziff*.)
- $\tilde{Q}_z = \exp\left(-\frac{1}{2}\ln(1-z)\right) = \frac{1}{\sqrt{1-z}}$
- $Q_n = \binom{2n}{n} 2^{-2n}$ (from generating function table by *Ziff*.)

First-passage probability (non-biased case)

- $f_n = -\frac{\partial}{\partial n} Q_n$ (from *A Guide to First-Passage Processes* by *Sidney Redner*)
- For discrete time step n, $f_n = Q_{n-1} - Q_n$.

- $\tilde{f}_z = \sum_{n=0}^{\infty} f_n z^n = Q_0 + \sum_{n=1}^{\infty} (Q_{n-1} - Q_n) z^n = Q_0 + z \sum_{n=0}^{\infty} Q_n z^n - \sum_{n=0}^{\infty} Q_n z^n = 1 - (1-z)\tilde{Q}_z$
  by definition of persistence (of survival probability), $Q_0 = 1$.

- $\tilde{f}_z = 1 - \sqrt{1-z}$

Average record number (non-biased case)

- $\tilde{Q}_z = \frac{1}{\sqrt{1-z}}$
- $\tilde{f}_z = 1 - \sqrt{1-z}$

- $P(R|n) \equiv Proba.[R_n = R]$
- $\tilde{P}(z,n) = \sum_{n=m-1}^{\infty} P(m|n) z^n = \tilde{f}_z^{m-1} \tilde{Q}_z$ (skip so many details..)

- $\langle R_n \rangle = RP(R|n)$
- $\tilde{R}_z = \sum_{n=0}^{\infty} \langle R_n \rangle z^n = \sum_{n=0}^{\infty} RP(R|n) z^n = \sum_{m=1}^{\infty} m \sum_{n=m-1}^{\infty} z^n P(m|n) = \sum_{m=1}^{\infty} m \tilde{f}_z^{m-1} \tilde{Q}_z = \frac{\tilde{Q}_z}{(1-\tilde{f}_z)^2}$
- $\tilde{R}_z = (1-z)^{-1.5}$

- $R_n = \binom{n+\frac{1}{2}}{n}$ (from generating function table by *Ziff*.)

GG…

자세한 설명은 생략…

참고로 biased case 에서는 generating function table으로 해결 안 됨…

Keywords : Cauchy Integral (Cauchy Inversion Formula), Inverse Laplace transform (Bromwich Integral) …

References

Non-biased case : Satya N. Majumdar and Robert M. Ziff 2008 Phys.Rev.Lett. **101** 050601

Biased case : Satya N Majumdar *et al* 2012 J. Phys. A: Math. Theor. **45** 355002

So many papers studied by Satya N Majumdar…

Generating Function Table by *Robert Ziff.*