

# Alkanoを用いた設備の異常検知

広島工業大学 工学部 電子情報工学科 前田ゼミ

高橋 秀典

---

# 目次

---

- 1.研究背景
- 2.先行研究
- 3.目的
4. 診断方法の提案
- 5.使用するデータの概要
- 6.検証と考察
- 7.まとめと今後の課題
- 8.参考文献

# 1. 研究背景

---

産業設備などのインフラ設備は消耗品であり、計画外の故障の発生は産業界や社会生活に大きな影響を及ぼす。

このため、異常に至る予兆を早期に発見し、対策を通じて、異常の発生を未然に防ぐことが必要である。

# 2. 先行研究 ( タイトル: Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati )

著者: William Gousseau, Jerome Antoni, Francois Girardin, Julien Griffaton

ベアリングを周波数領域で分析し、異常を検知する手法が報告されている。

ベアリングは図1に示す構造であり、仕様ごとに固有の周波数をもつ。  
 次式に転動体通過周波数-外輪(BPFO)を示す。

$$BPFO = \frac{n f_r}{2} \left( 1 - \frac{d}{D} \cos \phi \right)$$

$D$ …ピッチの直径  
 $d$ …転動体の直径  
 $f_r$ …シャフトの回転数  
 $n$ …転動体の数  
 $\phi$ …ベアリングの接触角

上記周波数はベアリングが故障した際に損傷特性周波数としても表れる。ただし、転動体損傷周波数は下記のBSFの2倍値となる。

$$BSF = \frac{f_r D}{2d} \left[ \left( 1 - \frac{d \cos \phi}{D} \right)^2 \right]$$

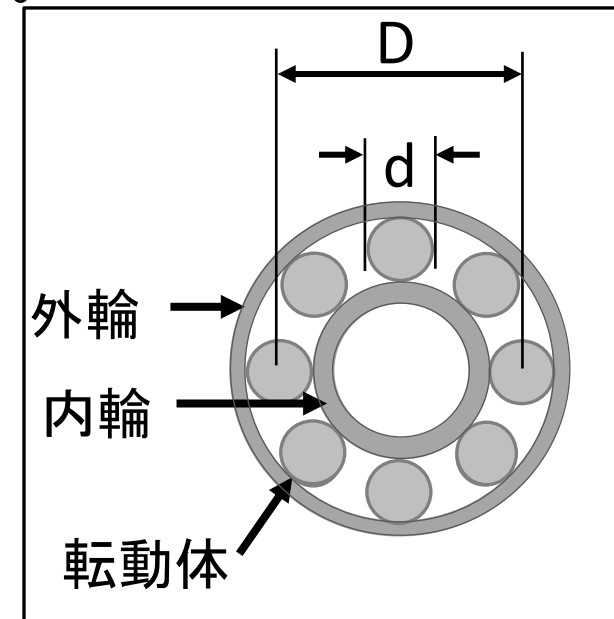


図1.ベアリング断面

## 2. 先行研究

先行研究で用いたベアリングの使用を表1に示す。  
算出されたBPFOとBSFはそれぞれ表2に示す値となった。

表1.ベアリングの特徴

ピッチの直径 $D$	71.5mm
転動体の直径 $d$	8.4mm
シャフトの回転数 $f_r$	33.3Hz
転動体の数 $n$	16
ベアリングの接触角 $\phi$	15.17°

表2.代表的な周波数

周波数特性	値
BPFO	236Hz
BSF	278Hz(2 × 139Hz)

**先行研究では上記周波数領域に着目し、ベアリングの異常の予兆を検知している。**

## 2. 先行研究

---

短時間フーリエ変換の結果、開始から3.5日後<sup>※</sup>に3倍のBPFO高調波が確認された。

(※後述するが3.5日後はファイル番号504付近に相当する)

二乗エンベロープにより尖度を最大化できる帯域で信号をフィルタリングし、カルトグラムにより可視化した結果4.8日後<sup>※</sup>に高調波が確認された。

(※後述するが4.8日後はファイル番号690付近に相当する)

### 3. 目的

---

先行研究はベアリングの型式情報から得られる周波数に着目した解析方法をとる。

→ 種々の多様な設備に適用することは困難

多様な設備に対応できる  
**汎用的な異常診断方法**を開発する。

## 4. 診断方法の提案

---

データセット2の加速度センサデータをヒストグラム化し、  
確率分布を取得する処理を行い、  
この確率分布を情報理論的観点での分析を行う。

具体的には、

エントロピー 相互情報量 カルバックライブラーダイバージェンスと  
二乗平均平方根(RMS)の比較を行う。

特に 正常・異常予兆・異常のK-meansクラスタリングによる3分類の  
可能性を評価し、考察を行った。

加えて異常より前のデータを対象に正常-異常予兆の2分類も行った。



# 4. 診断方法の提案

## エントロピー(Entropy)

情報のあいまいさを測る尺度

$$H(X) = - \sum p(x) \log p(x)$$

正規分布の式を代入すると

$$H(X) = -\frac{1}{2} (1 + \log(\pi\sigma^2)) \dots (1)$$

と変形でき、分散のみに影響を受ける。エントロピーの変化で、状態変化を確認。

## 相互情報量(MI)

2つの確率変数の従属性を測る尺度

$$I(X; Y) = \sum_{x \in X, y \in Y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)}$$

二つのBearing間の関係性が大きい程、値も大きくなる。  
異常が発生したBearingとの関係性から異常発生個所の特定に用いる。

# 4. 診断方法の提案

## カルバックライブラーダイバージェンス(KLD)

確立分布の近さを測る尺度

$$D_{\text{KL}}(p \parallel q) = \sum p(x) \log \frac{p(x)}{q(x)} \dots (2)$$

2つの確率分布が近い場合は、値が小さくなり遠いほど値は大きくなる。

信号の先頭部分とのKLDを取得することで信号の変化を検知。  
以後,KLDと省略。

## 二乗平均平方根(RMS)

振動する波形の振れ幅の平均を測る尺度

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^n x^2}{n}}$$

取得したファイルの実効値を取得し時系列変化を確認。

$$\text{RMS} = \sqrt{\text{square}(x).\text{mean}}$$

# 4. 診断方法の提案

Alkanoによるk-means法は右記条件を設定。

2分類の際は  
クラスター数を2に設定した。

基本設定

距離計算方法:	ユークリッド距離
クラスター数:	3
繰り返し最大数:	100

規格化オプション

初期クラスターの設定方法

ランダム

KMeans++

乱数シード:

# 5.使用するデータ概要

NASA Bearing Dataset(先行研究と同データ) 6000ポンドの荷重をかけた、4つのシャフト (Bearing 1,2,3,4)を、加速度センサを取り付けた状態で稼働させ、故障するまで振動データを計測した。

計測された振動信号は20.48kHzでサンプリングを行い、1つのファイルに格納している。

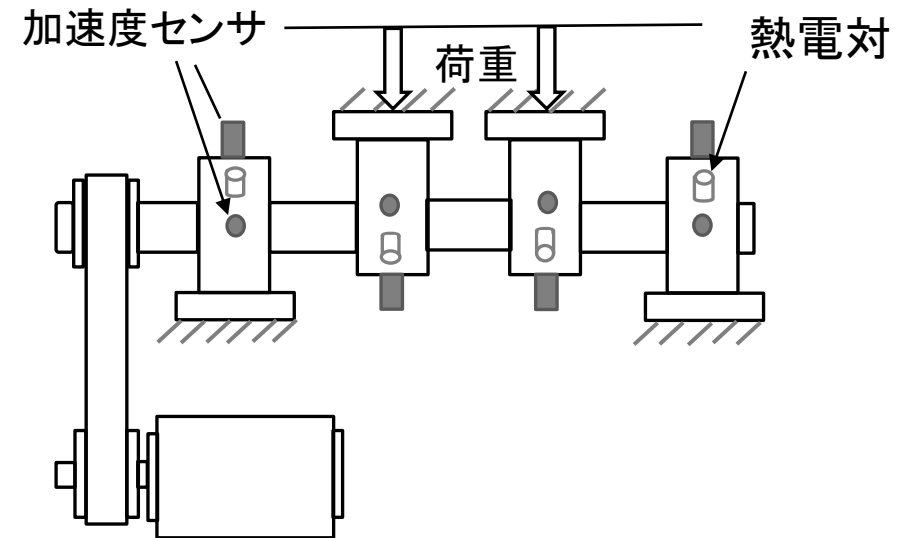
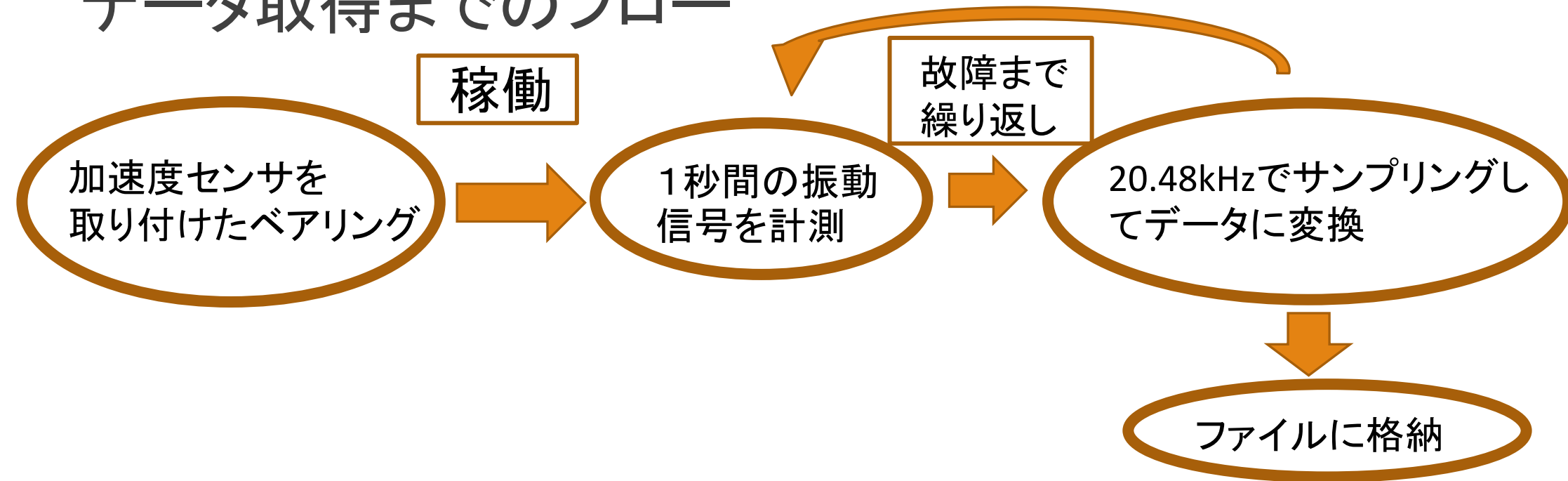


図2.Bearing装置とセンサの配置

# 5.使用するデータ概要

## データ取得までのフロー



# 5.使用するデータ概要

データセット2(ファイル数984)  
→試験終了時にBearing1外輪破損

Bearing1-ch1 (破損)

Bearing2-ch2

Bearing3-ch3

Bearing4-ch4

- 2004.02.12.10.32.39
- 2004.02.12.10.42.39
- 2004.02.12.10.52.39
- 2004.02.12.11.02.39
- 2004.02.12.11.12.39
- 2004.02.12.11.22.39
- 2004.02.12.11.32.39
- 2004.02.12.11.42.39
- 2004.02.12.11.52.39
- 2004.02.12.12.02.39
- 2004.02.12.12.12.39
- 2004.02.12.12.22.39
- 2004.02.12.12.32.39
- 2004.02.12.12.42.39
- 2004.02.12.12.52.39
- 2004.02.12.13.02.39

Bearing	1	2	3	4
1	-0.049	-0.071	-0.132	-0.01
2	-0.042	-0.073	-0.007	-0.105
3	0.015	0	0.007	0
4	-0.051	0.02	-0.002	0.1
5	-0.107	0.01	0.127	0.054
6	-0.078	-0.212	0.042	-0.044
7	-0.02	-0.01	-0.144	-0.007
8	-0.046	0.112	0.034	0.034
9	-0.063	-0.154	0.071	0.076
		■		
		■		
		■		
20474	-0.029	-0.046	0.005	0.061
20475	-0.017	-0.063	0.002	0.007
20476	0.049	-0.051	-0.039	-0.044
20477	0.037	0.061	0.115	0.007
20478	-0.012	0.007	0.056	-0.007
20479	-0.012	0.093	0.017	-0.044
20480	0.02	0.076	-0.042	-0.029

# 5.使用するデータ概要(確率分布の取得)

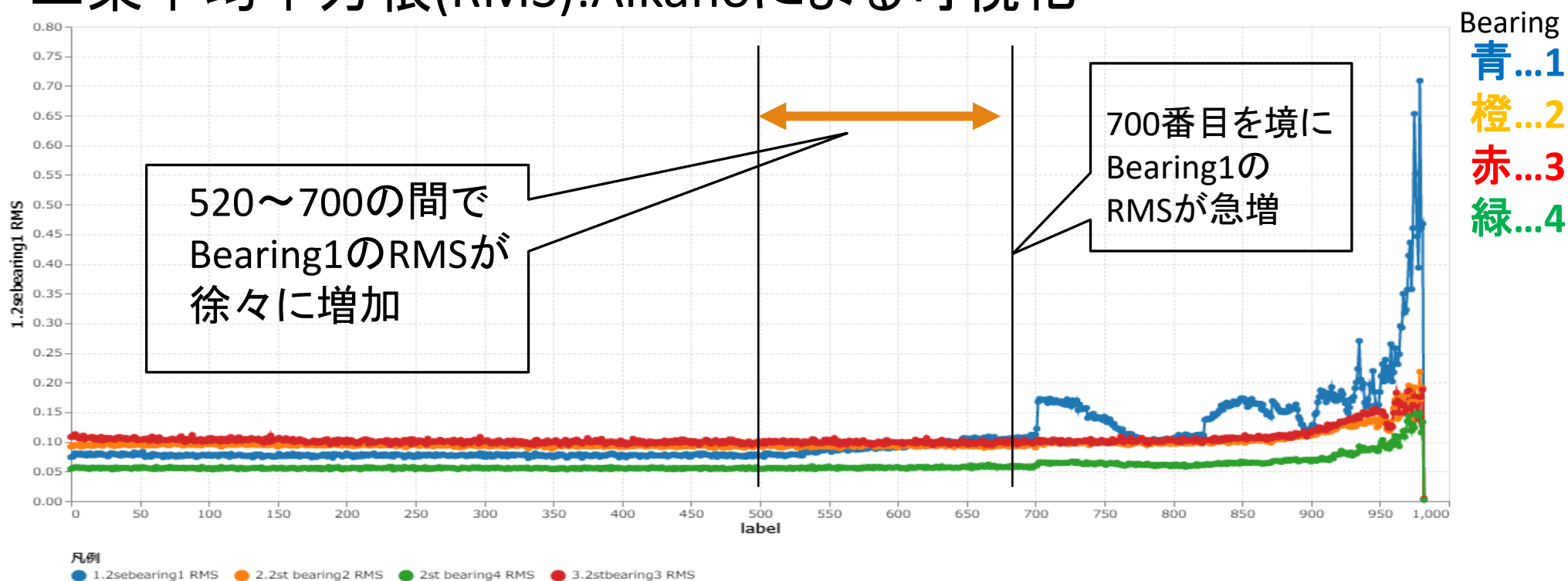
Bearing	1	2	3	4
1	-0.049	-0.071	-0.132	-0.01
2	-0.042	-0.073	-0.007	-0.105
3	0.015	0	0.007	0
4	-0.051	0.02	-0.002	0.1
5	-0.107	0.01	0.127	0.054
6	-0.078	-0.212	0.042	-0.044
7	-0.02	-0.01	-0.144	-0.007
8	-0.046	0.112	0.034	0.034
9	-0.063	-0.154	0.071	0.076
⋮				
20474	-0.029	-0.046	0.005	0.061
20475	-0.017	-0.063	0.002	0.007
20476	0.049	-0.051	-0.039	-0.044
20477	0.037	0.061	0.115	0.007
20478	-0.012	0.007	0.056	-0.007
20479	-0.012	0.093	0.017	-0.044
20480	0.02	0.076	-0.042	-0.029

1ファイル20,480個のデータにおいて、ヒストグラムに変換することでそれぞれを確率分布とみなす。  
ヒストグラム化する際のビン数は300  
範囲は-5~5に設定。

後の説明に出てくる  
 $P(x)$ ,  $P(y)$ は対象のベアリングの確率分布  
 $P(x, y)$ は同時分布である。

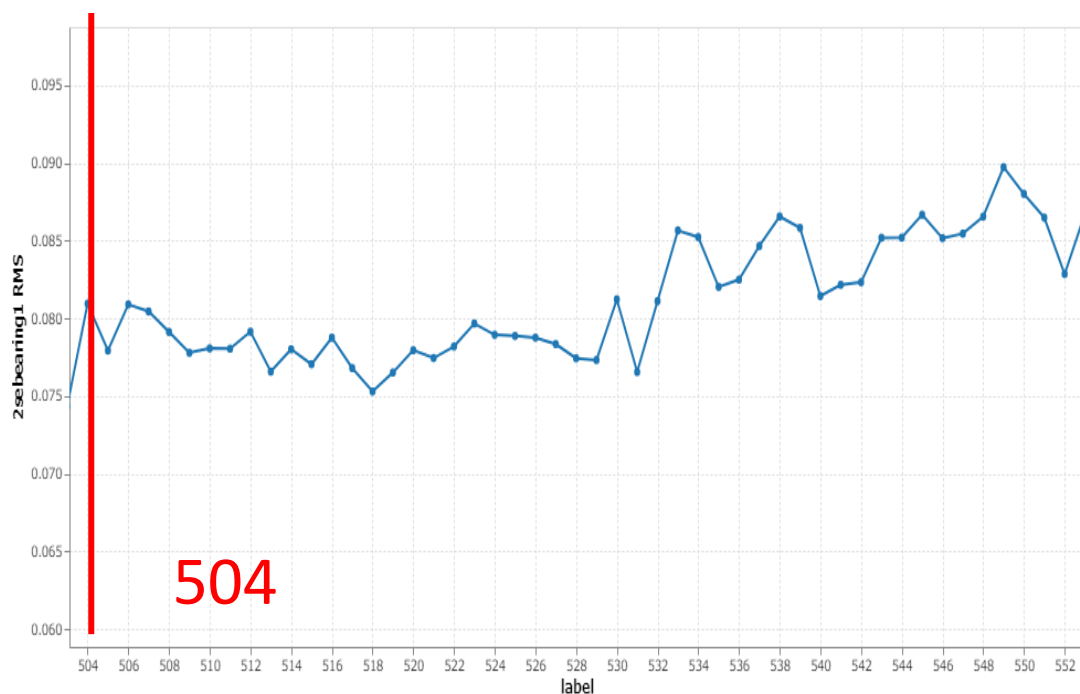
# 6. 検証と考察(k-meansによる3分類)

## 二乗平均平方根(RMS):Alkanoによる可視化





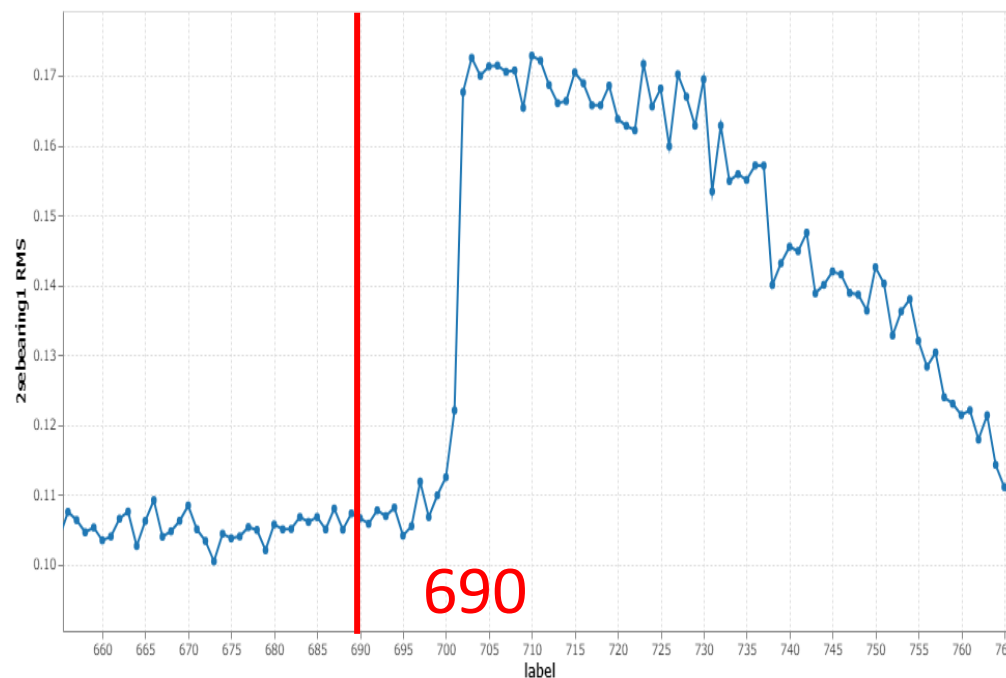
# 6. 検証と考察(k-meansによる3分類)



504

凡例  
● 2sebearing1 RMS

Bearing1 RMS 504~550



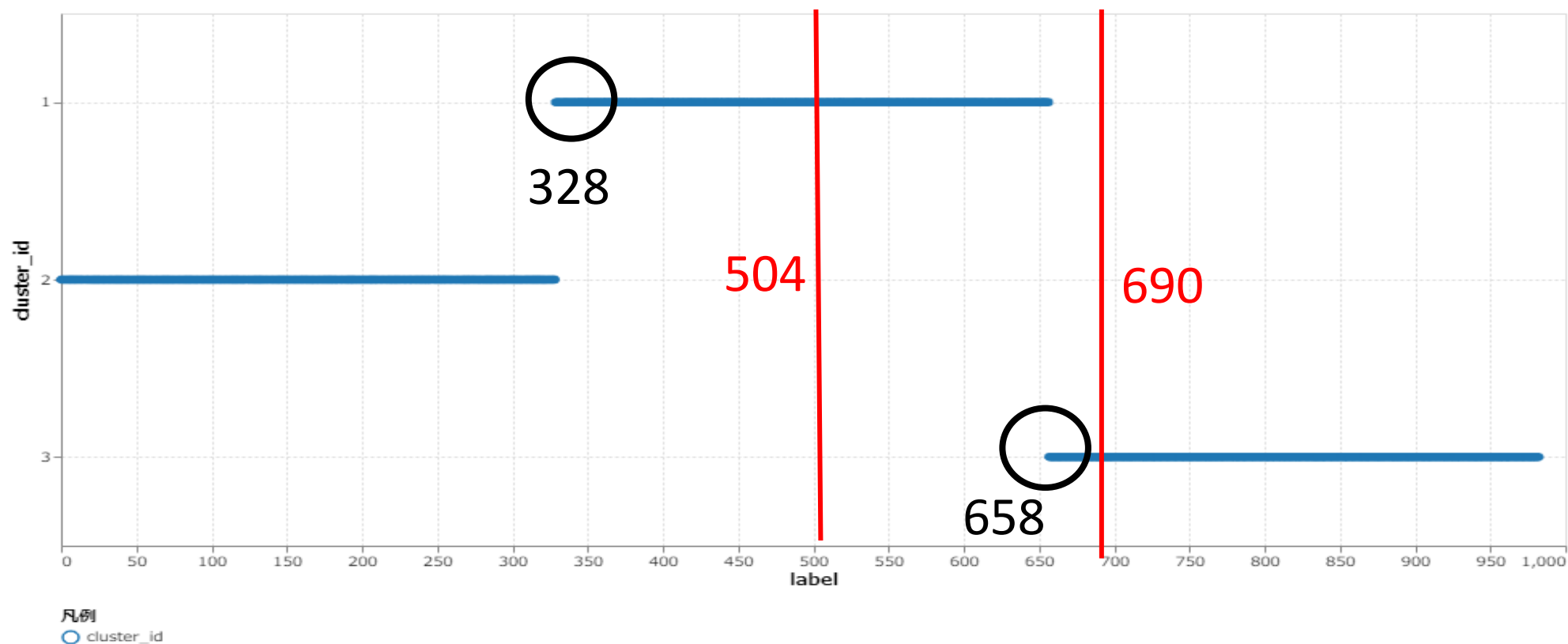
690

凡例  
● 2sebearing1 RMS

Bearing1 RMS 660~760

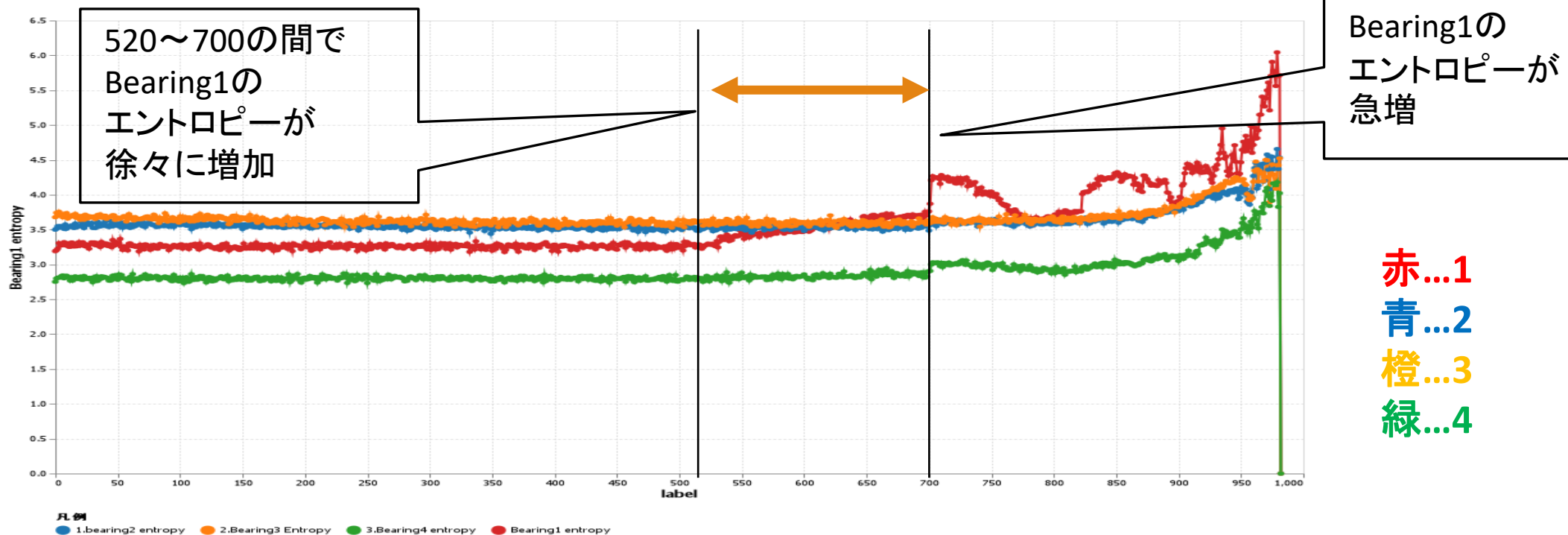
# 6. 検証と考察(k-meansによる3分類)

二乗平均平方根(RMS):K-means法(ユークリッド距離,クラスター数3)

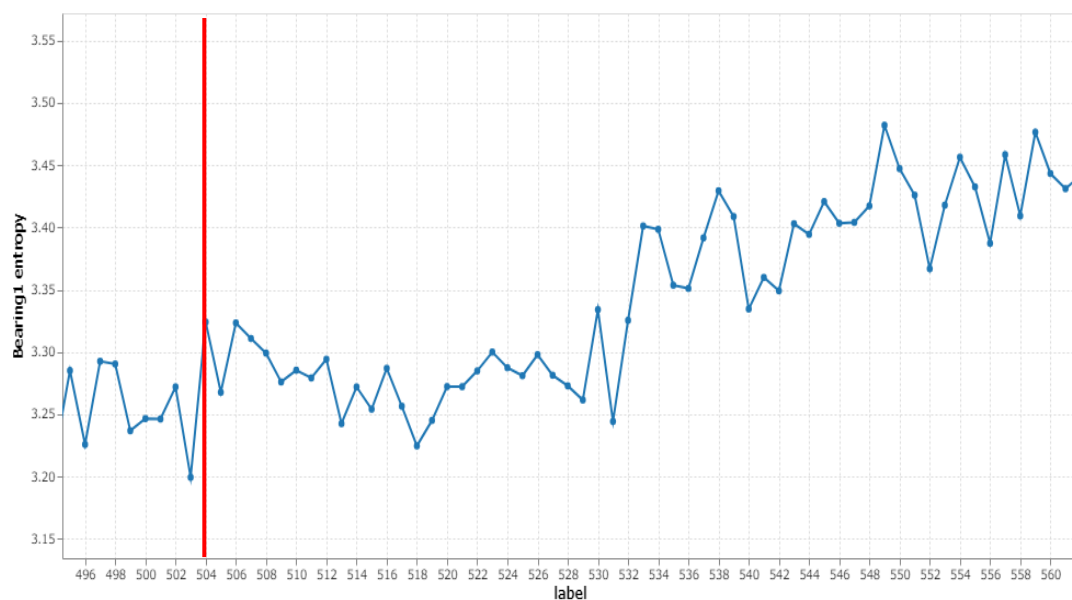


# 6. 検証と考察(k-meansによる3分類)

## エントロピー:Alkanoによる可視化



# 6. 検証と考察(k-meansによる3分類)

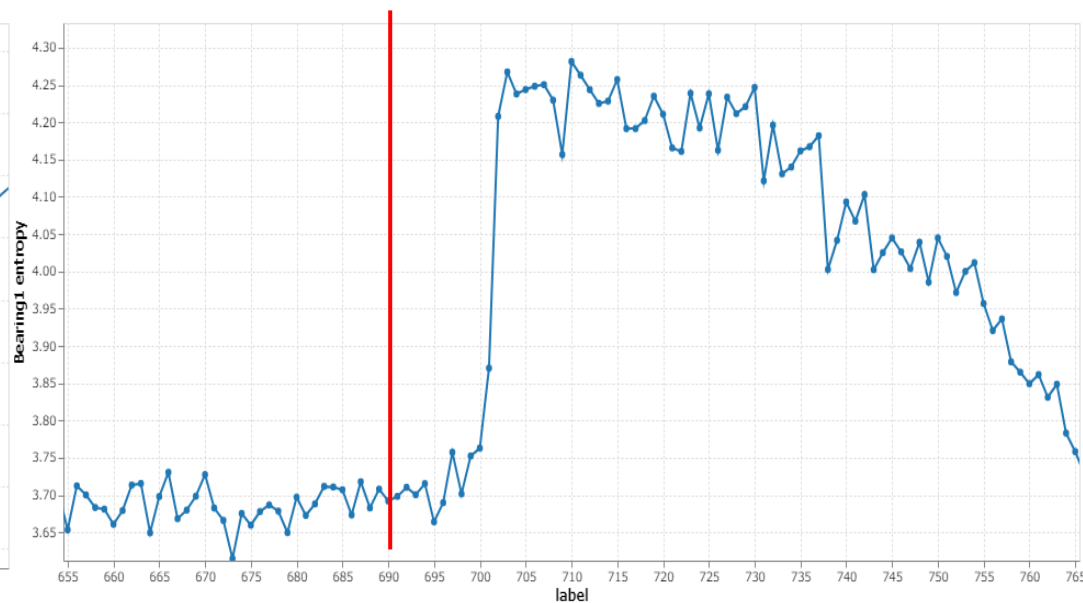


504

凡例

● Bearing1 entropy

Bearing1 RMS 496~560



690

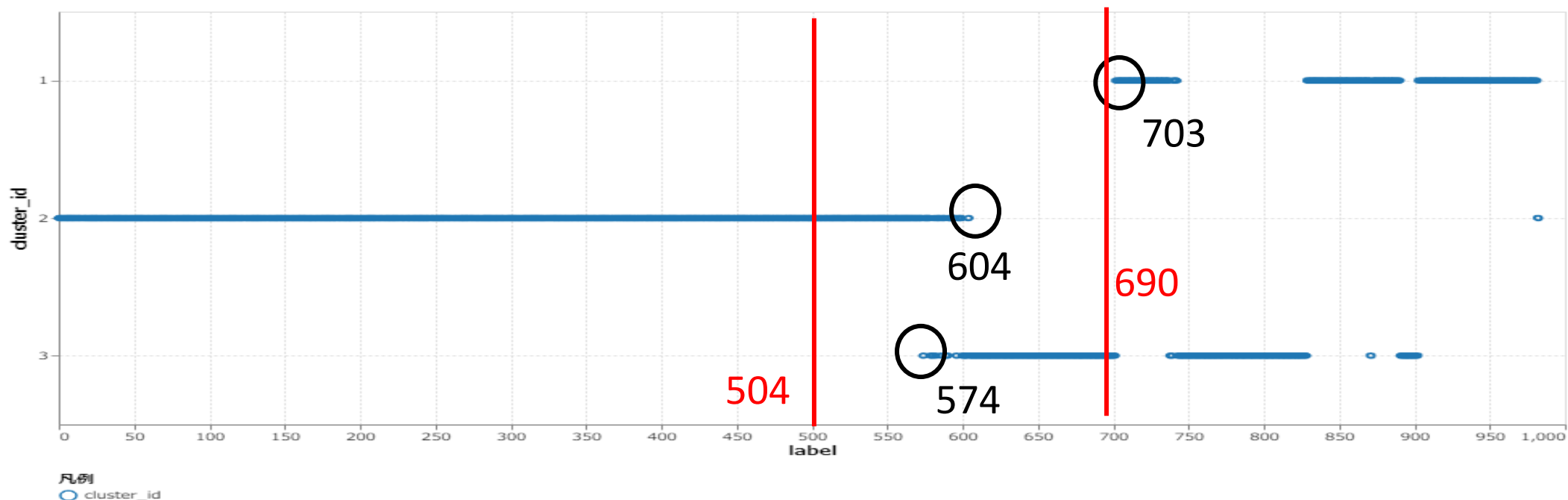
凡例

● Bearing1 entropy

Bearing1 RMS 655~765

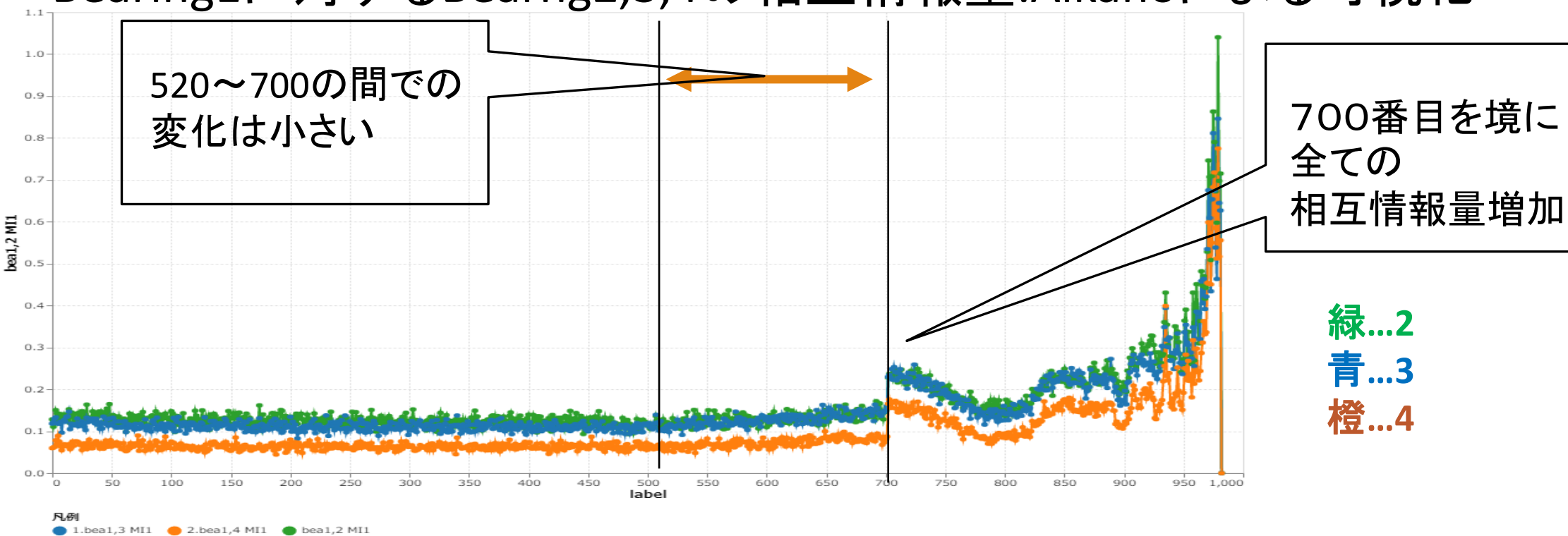
# 6. 検証と考察(k-meansによる3分類)

エントロピー: K-means法(ユークリッド距離, クラスタ数3)



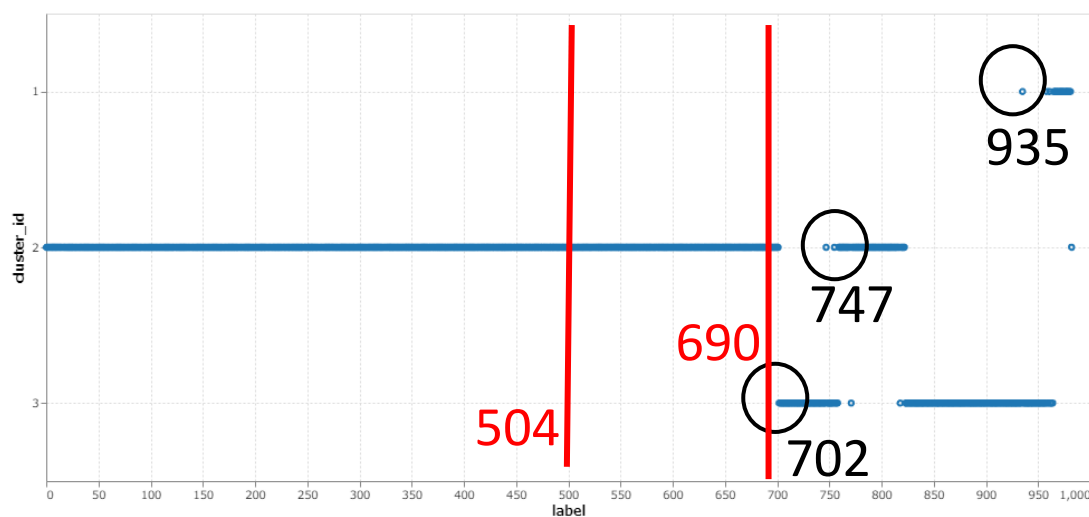
# 6. 検証と考察(k-meansによる3分類)

Bearing1に対するBearing2,3,4の相互情報量:Alkanoによる可視化

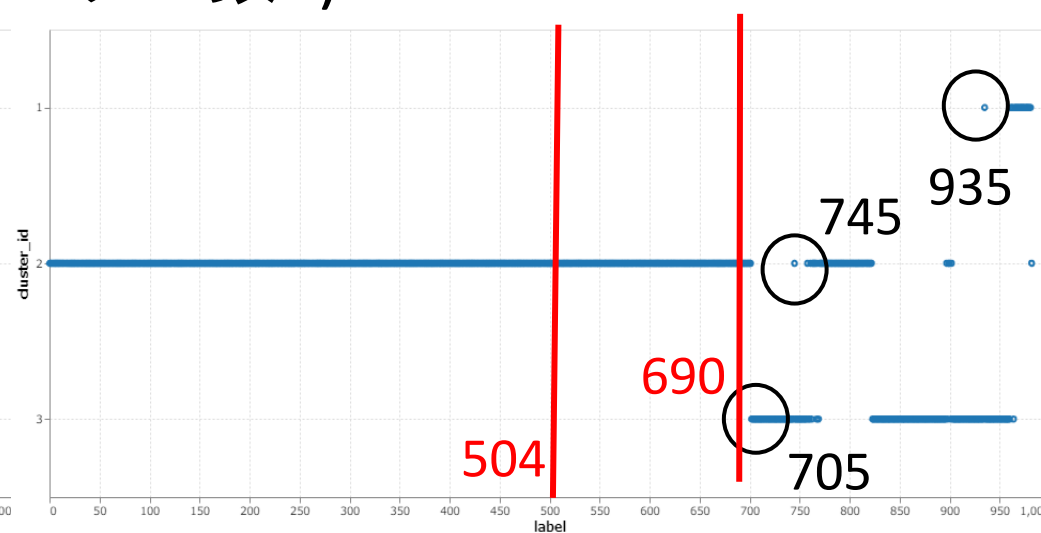


# 6. 検証と考察(k-meansによる3分類)

Bearing1に対するBearing2,3の相互情報量:  
K-means法(ユークリッド距離, クラスタ数3)



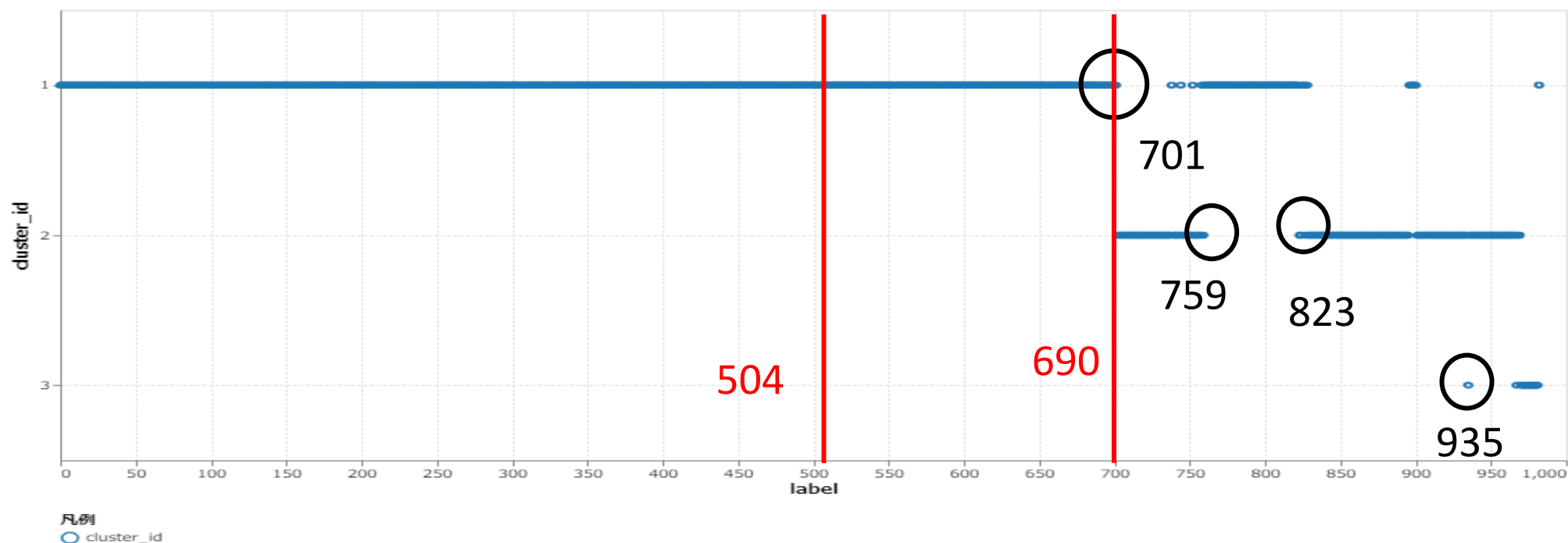
Bearing1-2



Bearing1-3

# 6. 検証と考察(k-meansによる3分類)

Bearing1に対するBearing4の相互情報量:  
K-means法(ユークリッド距離, クラスタ数3)

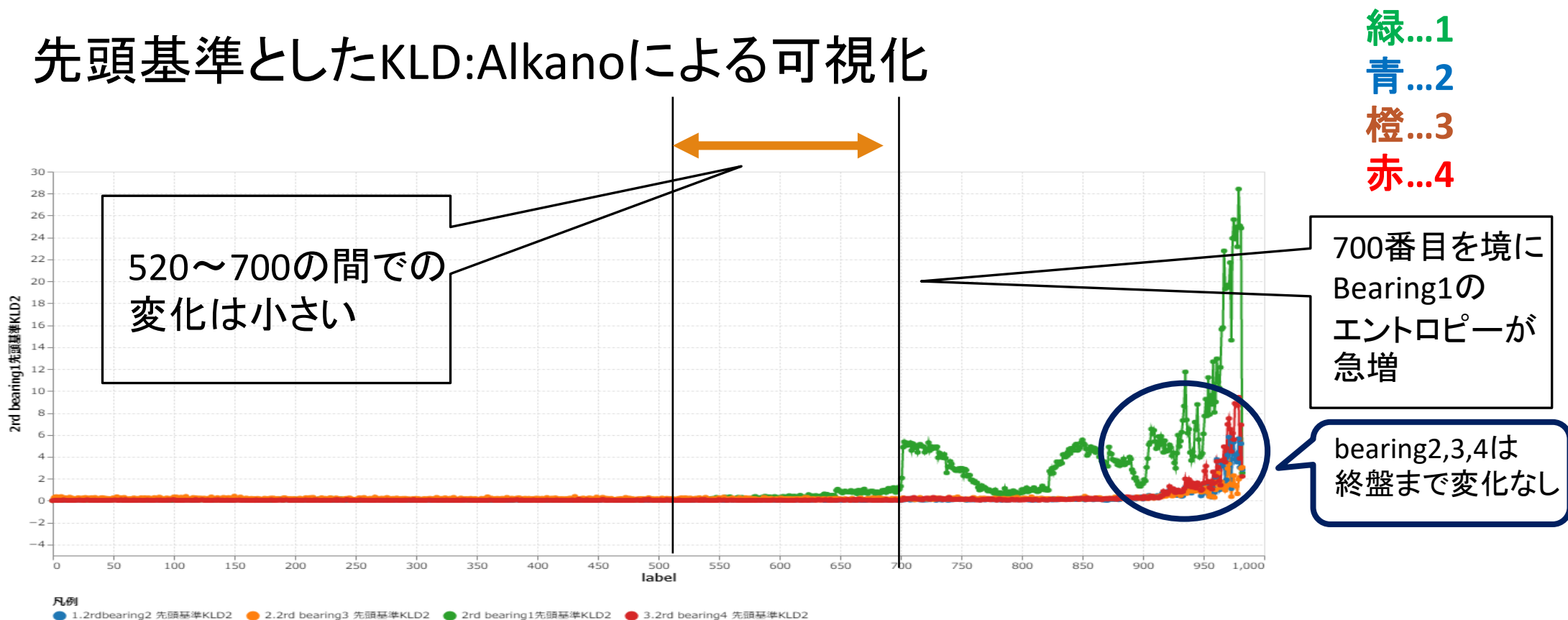


Bearing1-4



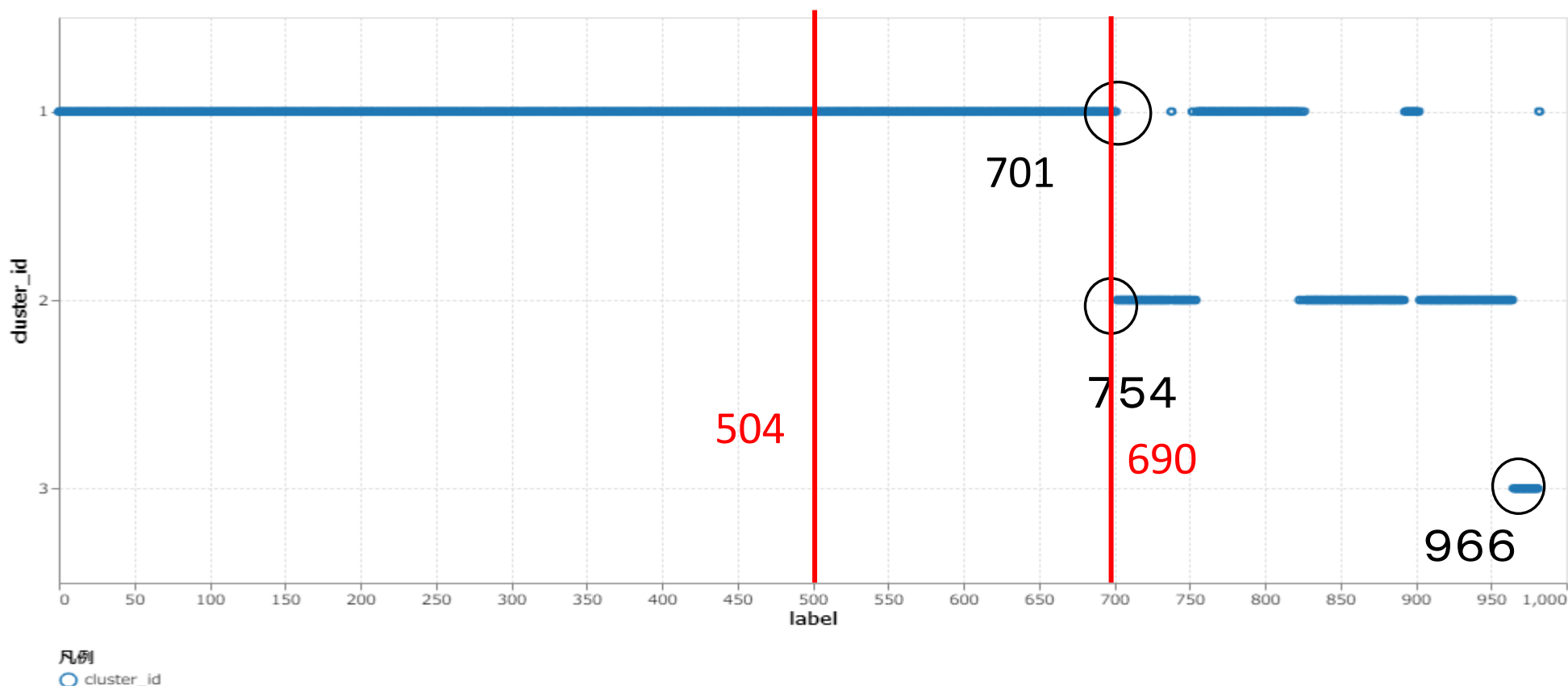
# 6. 検証と考察(k-meansによる3分類)

先頭基準としたKLD:Alkanolによる可視化



# 6. 検証と考察(k-meansによる3分類)

Bearing1先頭基準KLD :K-means法(ユークリッド距離,クラスター数3)



# 6. 検証と考察

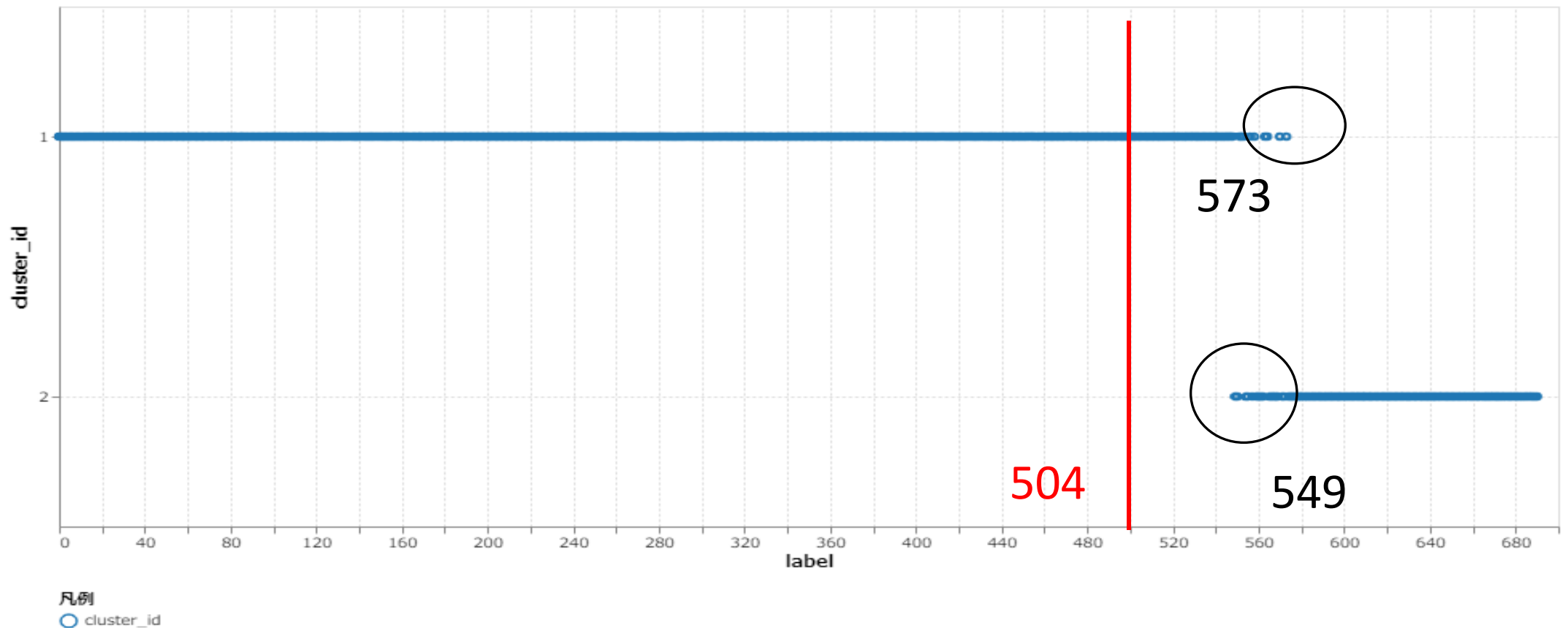
3分類の結果と先行研究から690-700番目以降は**異常である**と推測できる。

そのため取得したエントロピー・相互情報量・KLD・RMSの690番目より前のデータを対象に、正常-異常予兆の2分類を再びK-means法で行い考察する。



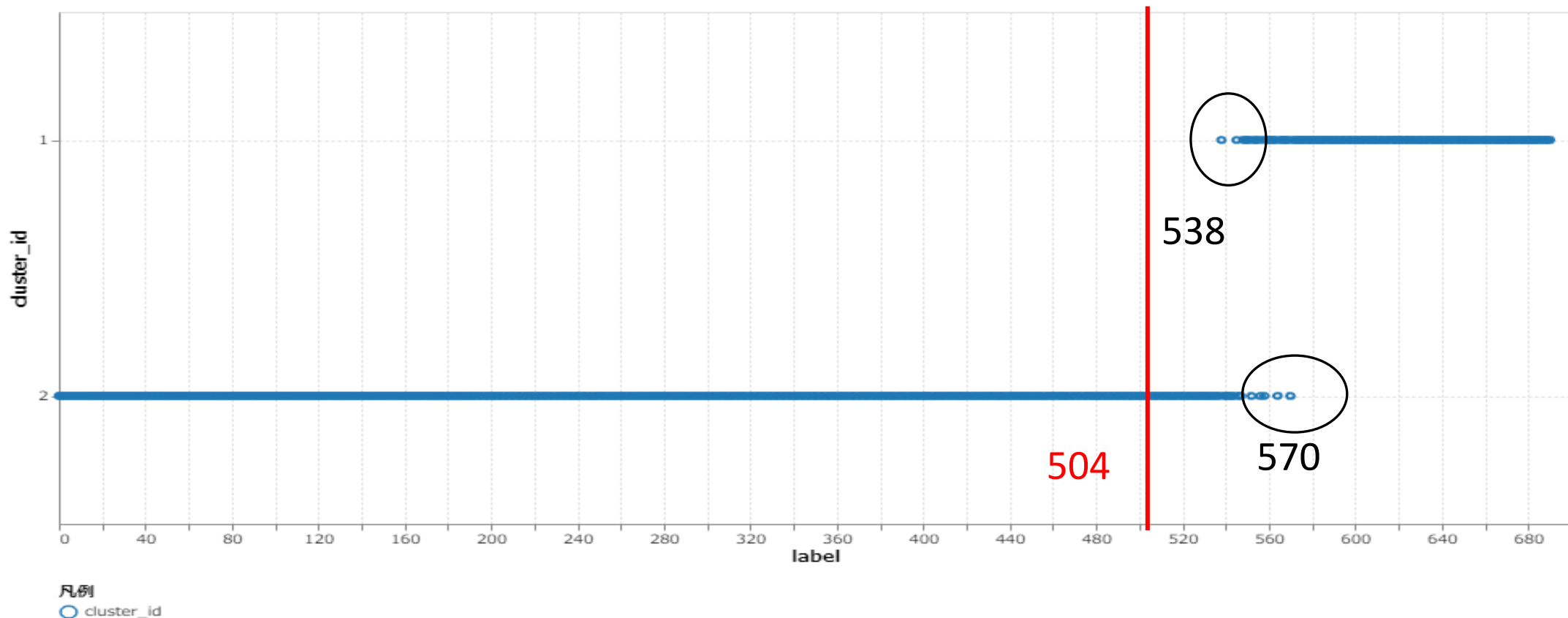
# 6. 検証と考察(k-meansによる2分類)

0~690 RMS:K-means法(ユークリッド距離, クラスタ数2)



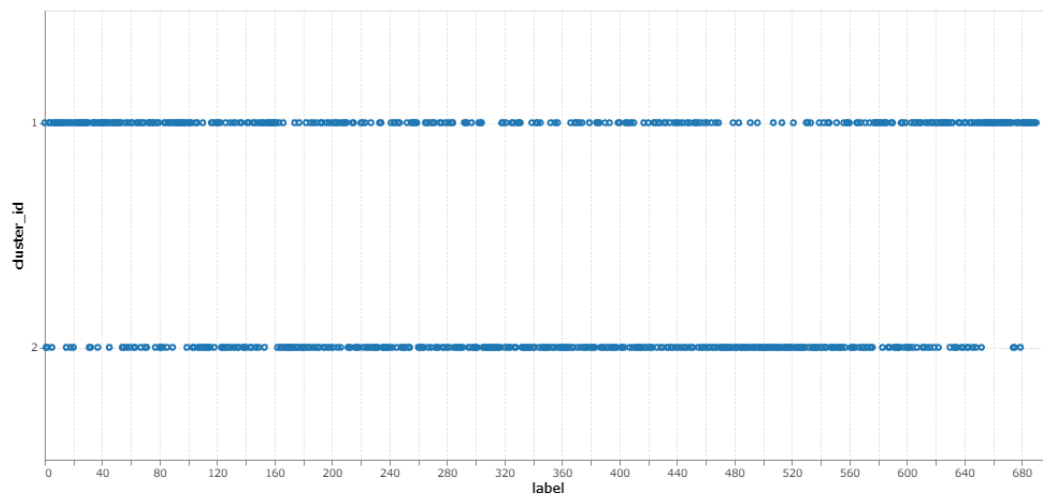
# 6. 検証と考察(k-meansによる2分類)

0~690 エントロピー: K-means法(ユークリッド距離, クラスタ数2)

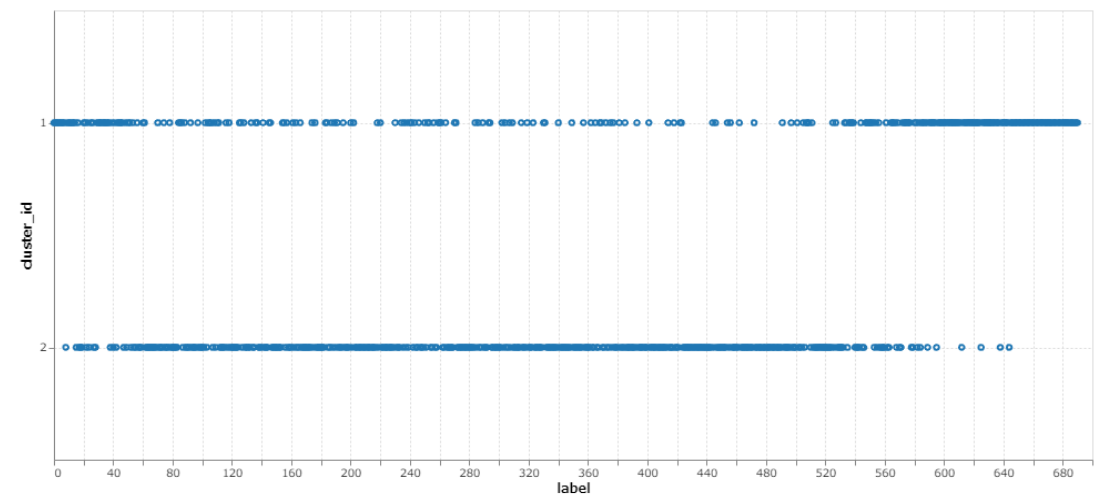


# 6. 検証と考察(k-meansによる2分類)

0~690 Bearing1に対するBearing2,3の相互情報量:K-means法  
(ユークリッド距離, クラスタ数2)



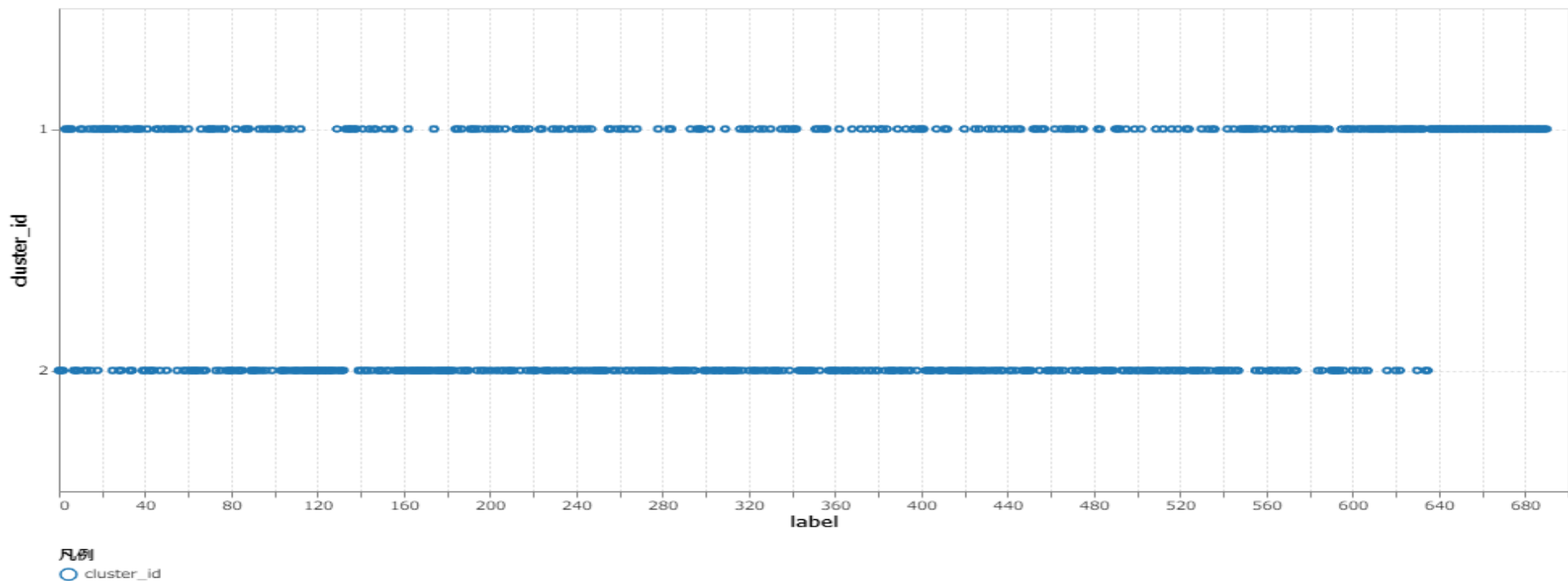
Bearing1-2



Bearing1-3

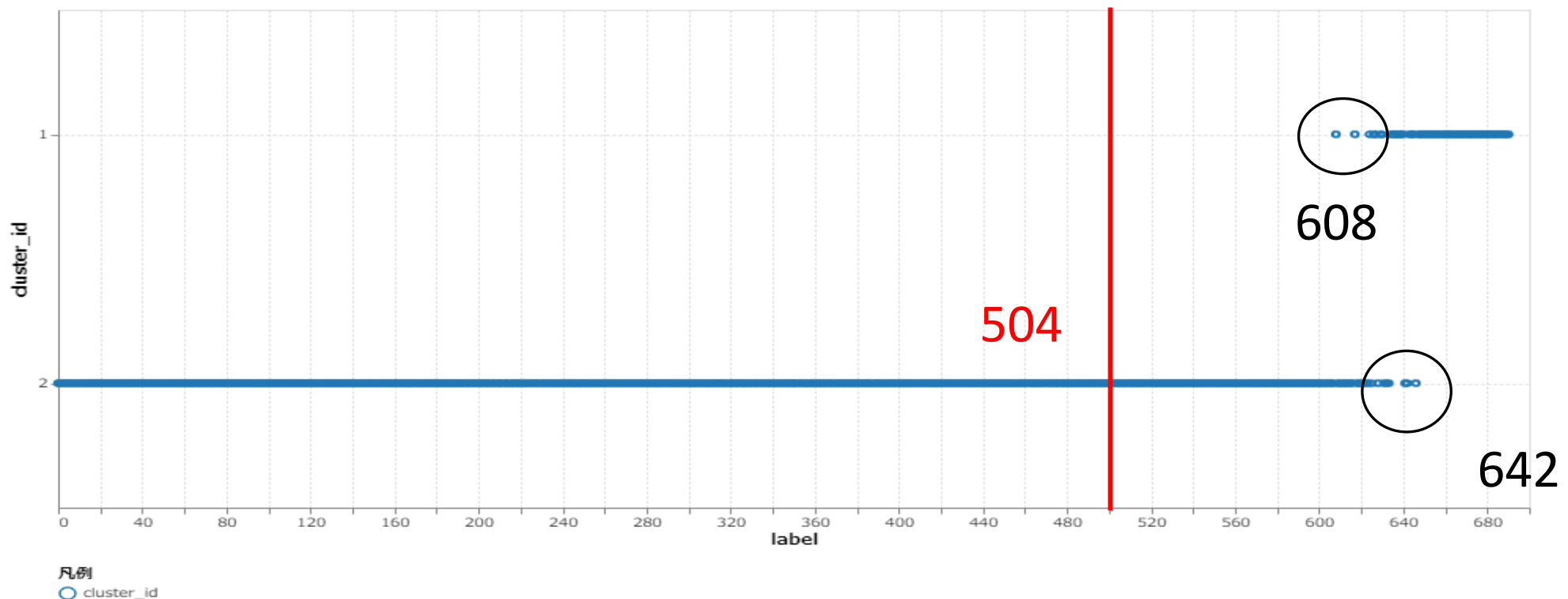
# 6. 検証と考察(k-meansによる2分類)

0~690 Bearing1に対するBearing4の相互情報量:K-means法  
(ユークリッド距離, クラスタ数2)



# 6. 検証と考察(k-meansによる2分類)

0~690 Bearing1先頭基準KLD:K-means法  
(ユークリッド距離, クラスタ数2)





## 6. 考察 (k-meansによる3分類)

・Alkanoによるエントロピーの可視化(スライド.20)により  
先行研究で高周波を確認した(504,690)番付近で値の変化が確認できた。

➡ 値の分散の大きさに依存するため異常点検知に活用可能。

・相互情報量の可視化(スライド.23,26)により  
Bearing2,3,4において500付近の変化は小さかったが690付近での変化が  
確認できた。

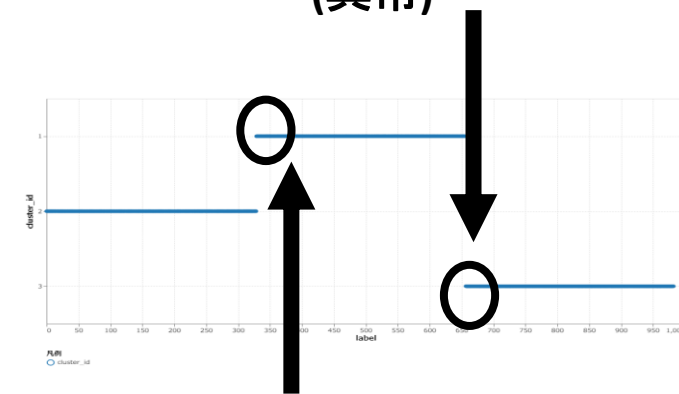
➡ ベアリング1の500付近のわずかな変化には影響されないが、  
変化の大きい690付近でBearing1の影響を強く受けたと考えられる。  
よって、異常個所の特定に利用可能。

# 6. 考察 (k-meansによる3分類)

3分類のうちクラスター2,3の先頭番号を表にまとめる。  
 クラスター2を異常予兆・クラスター3を異常とみなす

クラスター3先頭  
(異常)

評価項目		クラスター2先頭 (異常予兆)	クラスター3先頭 (異常)
RMS		328(-176)	658(-32)
エントロピー		574(+70)	703(+13)
相互情報量	1-2	702 (+198)	935 (+245)
	1-3	705 (+198)	935 (+245)
	1-4	702 (+198)	935 (+245)
KLD		703(+199)	906(+216)



クラスター2先頭  
(異常予兆)

※()は先行研究で観測した番号との差

# 6. 考察 (k-meansによる3分類)

3分類の結果RMSは異常予兆、異常の番号よりも早い段階で分類された。

これに対するエントロピー、相互情報量 KLDの分類結果を比較する。

## エントロピー

異常予兆は遅れて検知されたが、RMSと比べると異常はほぼ的確に検知された。

これは、RMSは値の平方根、エントロピーは分散の対数を取るため、**小さな変化でも比較的大きな値が取得**され、異常の検知で良い結果となったと考える。

評価項目		クラスター 2先頭 (異常予兆)	クラスター3 先頭 (異常)
RMS		328(-176)	658(-32)
エントロピー		574(+70)	<b>703(+13)</b>
相互情報量	1-2	702 (+198)	935 (+245)
	1-3	705 (+198)	935 (+245)
	1-4	702 (+198)	935 (+245)
KLD		703(+199)	906(+216)

## 6. 考察 (k-meansによる3分類)

---

### 相互情報量

500付近の異常予兆の分類はされず、690付近の異常で初めて検知された。

➡ ベアリング1の異常予兆部分の微量な変化は ベアリング2, 3, 4への影響が少なかったため相互情報量の変化が乏しかったが原因。これに対し異常部分ではベアリング1の影響を受け相互情報量が大きくなり、RMSと比べて正確な検知が行われた。  
よって、相互情報量は**異常個所の特定**に活用できると考える。

## 6. 考察 (k-meansによる3分類)

---

### KLD

500付近の異常予兆の分類はされず、690付近の異常で初めて検出された。

➡ データの可視化では500付近の値の変化は確認することが出来たが式(1)(2)より、エントロピーと比較して変化値が小さいことが、K-meansの分類に影響したと考える。

# 6. 考察 (k-meansによる2分類)

2分類の結果は右の表のようになった。  
相互情報量は異常予兆での分類が不可  
と判断し省略した。

## RMSとの比較

エントロピーはRMSと比べて先行研究の番号近い番号で  
検知された。

➡ 要因は3分類の時と同様と考えられる。

KLDはRMSと比較すると先行研究より遅い番号で検知  
された

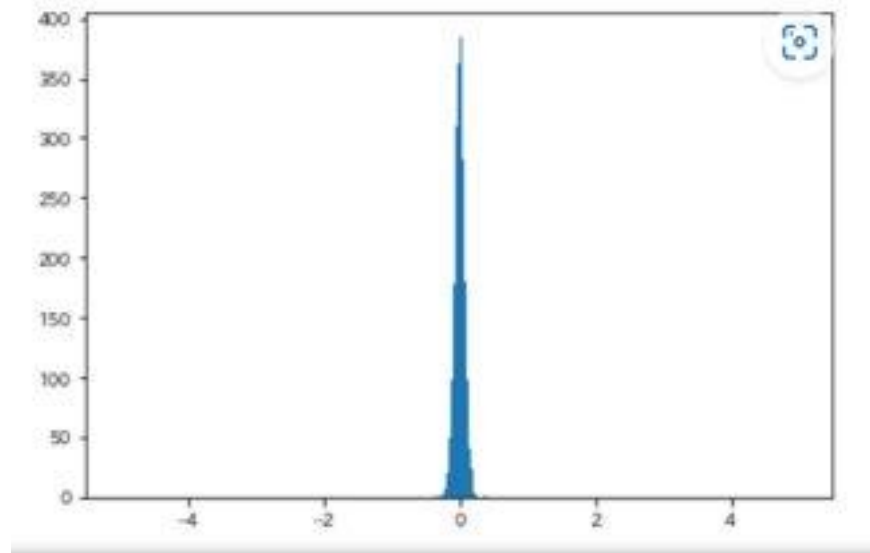
➡ 変化の割合が小さいことが要因と考えられるが  
しっかり2つに分かれていることから、クラスタリング  
法の工夫により精度向上の可能性が期待できる。

評価項目		クラスター2 先頭 (異常予兆)
RMS		549(+45)
エントロピー		538(+34)
相互情報量	1-2	
	1-3	
	1-4	
KLD		608(+104)

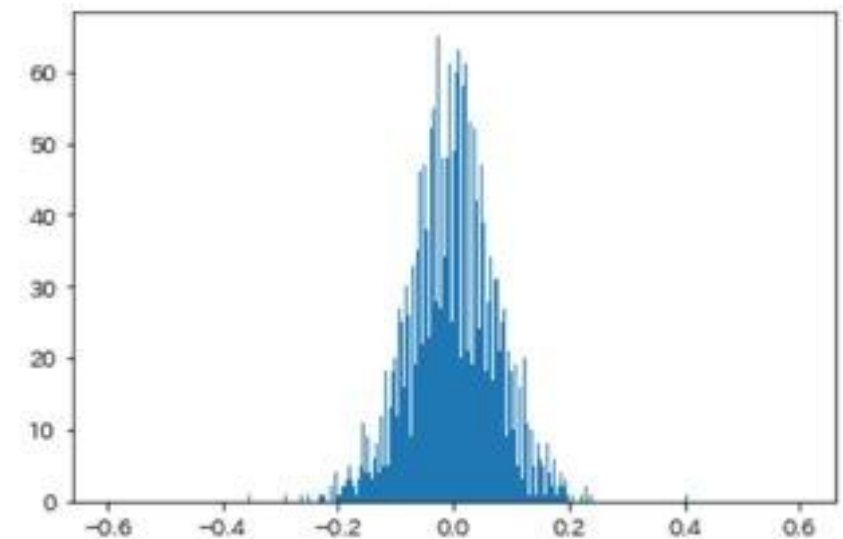
※()は先行研究で観測した番号との差

## 6. 検証と考察(さらなる検討)

分類の結果をさらに改善するためデータの存在範囲を考慮し、**ヒストグラム化の範囲を  $[-0.6 \sim 0.6]$**  に設定した上で、エントロピーとKLDを対象に検討した。



範囲 $[-5 \sim 5]$



範囲 $[-0.6 \sim 0.6]$

## 6. 考察(エントロピーを対象とした3分類)

再検討した結果、エントロピーはヒストグラム範囲の変更前と比べて**異常予兆の検知が改善された**。

➡ データの存在範囲を考慮したヒストグラム化を行ったため、**小さな変化を捉えることが出来た**と考える。

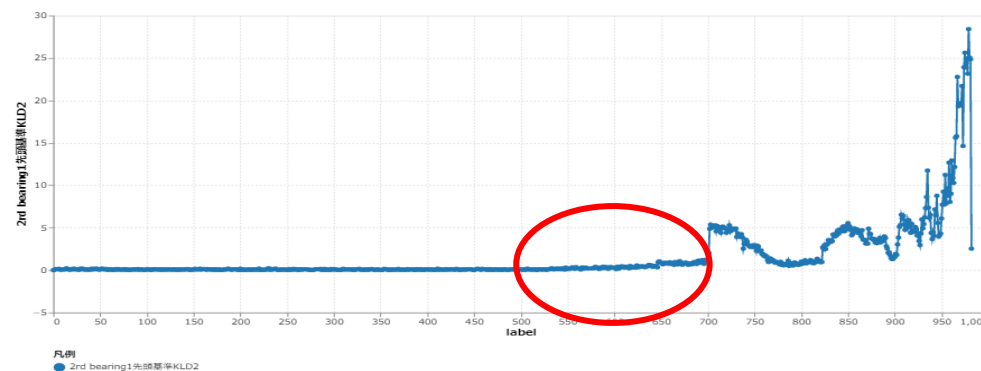
エントロピー	クラスター2先頭 (異常予兆)	クラスター3先頭 (異常)
範囲[-5~5]	574(+70)	703(+13)
範囲[-0.6~0.6]	<b>549(+45)</b>	703(+13)



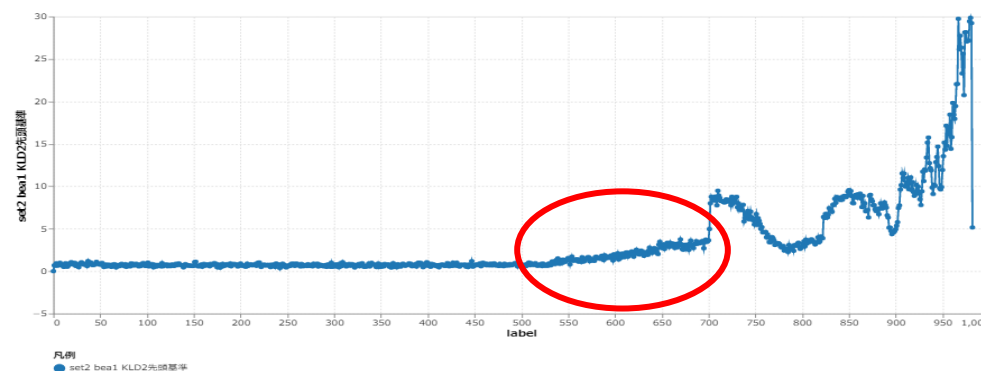
# 6. 考察(KLDを対象とした2分類)

KLD	クラスター2先頭 (異常予兆)
範囲[-5~5]	608 (+104)
範囲[-0.6~0.6]	492 (-6)

KLDが改善され、異常予兆の変化を的確に表現することに成功し、検知性能が大幅に改善された。  
なお、異常予兆のみを対象とした。



変更前: 範囲[-5~5]



変更後: 範囲[-0.6~0.6]

# 7.まとめと今後の課題

---

- 様々な設備に対応できる汎用的な異常診断方法を立案した。
- エントロピーをクラスタリングに活用することで、正常・異常予兆・異常の3分類は可能であることが分かった。
- 相互情報量により他の設備との関係性を数値化することができ、異常個所の特定が可能である。
- データの値を考慮したヒストグラム化により分類の精度を上げることが出来ると分かった。

## 7.まとめと今後の課題

---

- 今回はK-means法のみでの検証のため、他のクラスタリング法で検証を行い、3分類の精度を上げる必要がある。
- 2つの分布の入れ替えによって得られるKLDの双対性を利用して新たな視点から異常検知を検討する。

# 8.参考文献

---

- **NASA Bearing Dataset**

<https://www.kaggle.com/datasets/vinayak123tyagi/bearing-dataset>

- **転動体ベアリングの故障診断**

<https://jp.mathworks.com/help/predmaint/ug/Rolling-Element-Bearing-Fault-Diagnosis.html>

- 居森 章, 前田俊二, 小松 堯, 谷口哲至, 外田 脩, エントロピーがもつ統一性, 双対性, 従属性による異常検知 ~ エントロピー・カルバック-ライブラーダイバージェンス・相互情報量の有用性 ~, 電子情報通信学会 情報理論研究会, 2022-01-RCS-SIP-IT(2022.1)

- [William Gousseau](#), [Jérôme Antoni](#), [François Girardin](#), [Julien Griffaton](#). **Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati**  
CM2016, Oct 2016, Charenton, France