

タイトル	北海道における道東地域の鉄道に関する検討 - ベイズ法および計算ソフトJASP の活用 -
著者	上浦, 正樹; KAMIURA, Masaki
引用	開発論集(110): 97-121
発行日	2022-09-30

北海道における道東地域の鉄道に関する検討

—— ベイズ法および計算ソフト JASP の活用 ——

上 浦 正 樹*

1 はじめに

北海道における持続可能な地域社会を形成する上で鉄道は重要な位置にある。それは明治時代が始まったころから鉄道は近代化産業を牽引するとして当時の政府の考え方に通じるものがある。鉄道には都市間輸送の担い手であるとの一般的な考え方があるが、原野で鉄道を建設する北海道では成り立たない。つまり北海道の鉄道は貨物輸送を中心に炭鉱などの鉱産資源を海上輸送につなげて本州へ移送するための手段として考えられた。これは、自動車交通がまだ発達していないために馬による道路輸送、河川による船による輸送よりも鉄道がはるかに効率的な輸送であったことによる。鉄道ができると関連の産業が発達し、その結果から町ができ人が集まり、都市が形成された。その後明治時代の後半になると、国防上や効率の面から主な民間鉄道は全国で統一した基準による国有鉄道へ移行され、全国に鉄道ネットワークが構築された。その後、このネットワークに繋げる民間資本による簡易鉄道を設置するケースが増えた。一方、極寒の北海道では旧北海道庁による殖民軌道が建設された。このような国有鉄道のネットワークをメインにこれに接続するネットワーク網が出来上がった。

本研究では、道東の釧網線、花咲線、富良野線を対象に、最初に明治時代から大正時代までの鉄道建設の歴史を扱っている。道東では、明治10年(1877年)に川湯にあるアトサヌプリ(硫黄山)で硫黄の採掘が始まり、明治20年(1887年)には、アトサヌプリから標茶間の鉄道を完成した。この鉄道は硫黄山鉄道と呼ばれ、当時道内で2番目のものであった。鉄道が通るようになると、沿線は発達し明治25年(1892年)には、一般の旅客も乗せる釧路鉄道へと生まれ変わった。その後、硫黄が枯渇するようになり、明治29年(1896年)には硫黄採掘は中止され、鉄道輸送も廃止となった。だが、この鉄道が現在の釧網本線の基礎となっている。一方、旭川と釧路を結ぶ幹線鉄道として北海道官設鉄道によって旭川、釧路双方から着工された。旭川側は1899年(明治32年)に旭川—美瑛間が、釧路側は明治34年(1901年)に釧路—白糠間が開業した。大正2年(1913年)滝川—下富良野駅(現在の富良野)間の新線が開業し、起点を旭川から滝川に変更した。滝川～富良野間にある赤平炭砒は明治27年(1894年)に採掘が始まり、平成6年(1994年)に閉山となった。

* (かみうら まさき) 北海学園大学開発研究所特別研究員

一方、起点を旭川から滝川に変更したことで、線路名称は釧路本線となり、旭川―富良野間は富良野線として分離した。また花咲線は、大正10年（1921年）に根室まで延伸した。その後、殖民軌道として昭和9年（1934年）頃から道東の開拓に資するため中標別まで線路建設された。しかし自動車の普及と道路整備が進み、鉄道の利用が激減したことから現在では廃線となっている。

このように建設された鉄道が鉱産資源の枯渇や海産物などの輸送が道路交通に対抗して採算に見合う運営ができないことなどから一部が廃線となった。この結果、現状の釧網線、花咲線、富良野線の三線区での敷設位置は今から100年以上前からあまり変わっていないことが分かる。

釧網線、花咲線、富良野線の三線区はJR北海道における「単独で維持困難線区」に属する。これに対し地方鉄道のあり方を議論する国土交通省有識者会議の提言書による¹⁾と、国、自治体、鉄道事業者による「特定線区再構築協議会」（仮称）を設立することなどを提言している。このような情勢を俯瞰してみると、「単独で維持困難線区」を一律に考えることも必要だが、個別の線区ごとに分析することも重要である。そこで本研究では対象の三線区を明治時代からの経緯を踏まえ、個別の線区として定量的に分類する方法を検討する。具体的には本研究では現在の釧網線、花咲線、富良野線の沿線の人口、駅間で通過する乗車者数の分布と時間ごとの乗車者数の分布を対象にする。

本研究の統計手法には、頻度法とベイズ法の2種類である。この頻度法は大標本を必要とする理論を前提にしているもののその枠を超えて多くの場面で使用されている。これに比べ、ベイズ法はあまり一般的とは言えない。それはベイズ法が事後確率の計算で非常に面倒あることによる。しかし近年になってコンピュータの能力が飛躍的に高まったことから、心理学分野で膨大なデータ処理方法が確立され、これを簡易な計算ソフトを用いて事後確率を求めることが可能となった。

本研究で取扱うデータ数は限られており、頻度法が拠り所とする大標本理論を適するかは疑問である。そこで、ベイズ法がデータの規模が小さくても対応できる点を考慮して、両方の方法を検討することとした。また、近年のベイズ法を用いた研究が心理学分野で発達してきたが、研究事例が皆無に近い鉄道関係分野に応用できるかの検討も重要である。そこで昨年より鉄道の地域への貢献度の評価に頻度法と並行してベイズ法も解析に使用することとした。また、本研究では統計計算ソフトJASPにより頻度法とベイズ法により解析を行う。このソフトは、アムステルダム大学のワーゲンメーカー（Eric-Jan Wagenmakers）らの心理統計分野の研究者によって製作されたソフト²⁾である。またJASPにより頻度法とベイズ法の両方の分析ができる。その精度は多くの研究で検証されている³⁾。

2 道東の開拓と鉄道の歴史

1) 鉄道の国有化

欧米の鉄道では黎明期（1830年年代）を経て明治10年（1878年）頃には、鉄道が新時代に相応しい交通手段と定着しつつあった。これを担うのは民間の資本家で日本各地に私設の鉄道が建設された。一方、日清戦争（明治27年：1894年～明治28年：1895年）が勃発すると国防の観点からも鉄道を全国の交通ネットワークとして確保する必要が生じた。そこで私設鉄道買収法案が明治28年（1895年）に成立し、国が鉄道ネットワークを作るために不足する線区に対して国が鉄道の建設を行うこととなった。これを実行に移すために建設規程として鉄道敷設法が明治25年（1892年）に制定された。以上の経緯から、明治39年（1906年）に「鉄道国有法」が公布された。北海道では、私鉄の北海道鉄道（現在の函館本線の函館と小樽の間）が鉄道国有法により明治40年（1907年）に国有化された。

一方、道東の交通では最初に明治10年（1877年）に川湯にあるアトサヌプリ（硫黄山）で硫黄の採掘が始まり、この硫黄は川の船や馬によって輸送された。その後の明治20年（1887年）には、アトサヌプリから標茶間の鉄道を完成し、標茶から釧路間は蒸気船での輸送ができるようになった。この鉄道は、硫黄山鉄道と呼ばれ、当時道内で2番目のものであった。鉄道が通るようになると、沿線は発達し明治25年（1892年）には、一般の旅客も乗せる釧路鉄道へと生まれ変わった。その後、硫黄が枯渇するようになり、明治29年（1896年）には硫黄採掘は中止され、鉄道輸送も廃止となった。なお、この鉄道が現在の釧網本線の基礎となっている。

2) 簡便鉄道と殖民軌道

国有鉄道は幹線系の全国ネットワークとすれば、産業が発達するにつれて幹線系に接続する鉄道のローカルネットワークが必要となった。しかし当時は地方鉄道まで国が投資する余力がなくなった。そこで民間の鉄道のために明治33年（1900年）に私設鉄道法が制定された。この法律は国有鉄道並に制約が多いため民間資本を活用して建設するケースがあまり見られなかった。とはいえ道路交通が発達していない時代では、速度が低速で牽引用の機関車の重量が軽いなどの条件で建設規程よりも簡易な鉄道でも対応できた。そこで鉄道建設に制約に少ない簡便鉄道法が明治43年（1910年）に公布された。これにより大都市の市電や工場からの専用線など簡便鉄道として国から認められ、国内の各地で簡易鉄道の建設が進んだ。この動きは明治40年（1907年）～大正5年（1916年）代に顕著に見られた。

しかし、北海道では極寒により冬季の鉄道事業が停止するために民間主導の簡易鉄道の導入が進まなかった。そこで北海道における簡易鉄道は旧北海道庁が建設し、殖民軌道と呼ばれた。その初期段階である北海道第二期拓殖計画では昭和2年（1927年）に47線区の約800kmが計画された。しかし採算上の問題で、全ての殖民軌道は1936（昭和11）年までに廃止された。道東の昭和7年（1932年）の事例を図2-1に示す。



図 2-1 道東の殖民軌道⁴⁾

その後、道東では国有鉄道によって昭和 9 年（1934 年）までに標津線厚床支線に沿って厚床～西別（別海）～中標津まで建設された。また昭和 11 年（1936 年）には東西に伸びる殖民軌道の標津線に沿って殖民軌道はレールが取り外され、釧網本線の標茶から根室標津までの線路が建設された。これらの線区は JR 北海道になって平成元年（1989 年）まで存続したが、その後廃止された。図 2-2 は標茶線、標茶線厚床支線、花咲線について参考文献⁵⁾の一部を削除し加筆したものである。

旭川と釧路を結ぶ幹線鉄道として北海道官設鉄道によって旭川、釧路双方から着工された。旭川側は明治 32 年（1899 年）に旭川―美瑛間が、釧路側は明治 34 年（1901 年）に釧路―白糠間が開業した。大正 2 年（1913 年）に、滝川―下富良野（現在の富良野）間の新線が開業し、起点を旭川から滝川に変更。線路名称は釧路本線となり、旭川―富良野間は富良野線として分離した。また花咲線は、大正 10 年（1921 年）に根室まで延伸した。

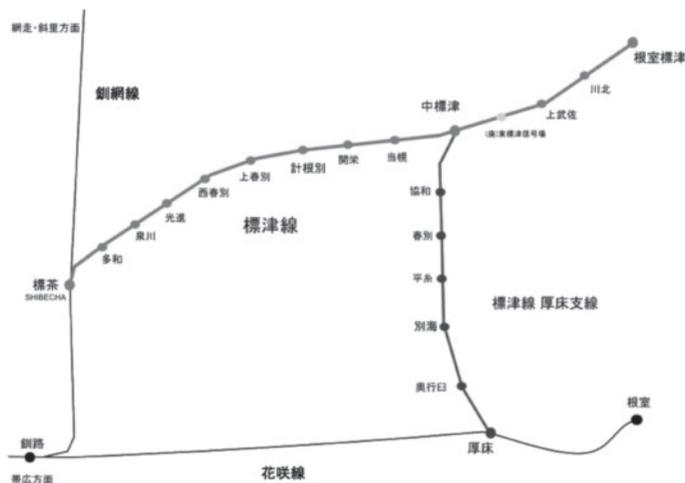


図 2-2 標茶線、標茶線厚床支線、花咲線の線路図



図 3-3 花咲線と関係する市町村⁸⁾

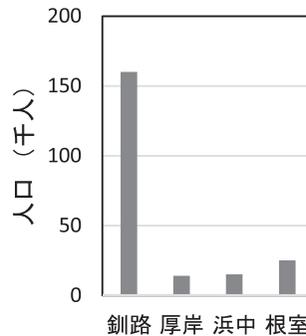


図 3-4 市町村別の人口 (花咲線)



図 3-5 富良野線と関係する市町村⁹⁾

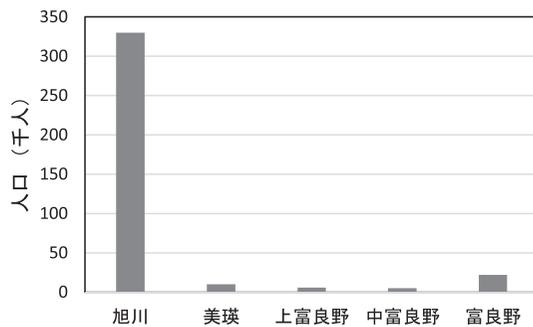


図 3-6 市町村別の人口 (富良野線)

係する市町村を人口は約 40 万人 (2020 年) である。図 3-6 に市町村別の人口を示す。この内訳は旭川 (約 33 万人) と富良野 (約 2 万人) で残りは 1 万人程度である。

(4) 貨物鉄道営業線区

図 3-7 に示すように、本研究の対象線区である釧網線、富良野線、花咲線には定期の鉄道貨物列車の運行はない。道東の農産物や海産物などで鉄道を利用する貨物はトラックなどで釧路貨物駅などに持ち込まれて、貨物列車で輸送される。

2) 各駅で通過する乗車者数

(1) 釧網線

沿線には知床国立公園、阿寒摩周国立公園、釧路湿原国立公園の 3 つの国立公園と網走国定公園があり、阿寒湖、摩周湖、釧路湿原等の豊富な自然と川湯温泉などの観光資源がある。よって釧網線はこれらの観光地へのアクセスの機能を有する。また釧路や網走・北見などの都市への近郊線区の役目を果たしている (図 3-8)。

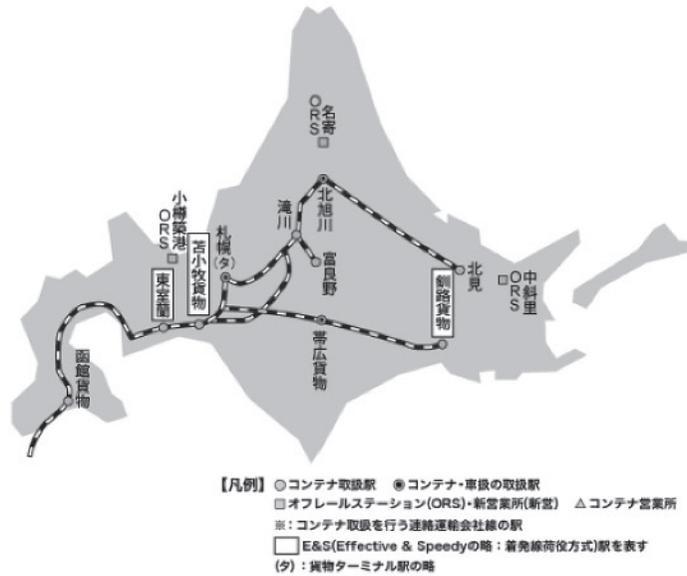


図 3-7 JR 貨物 エリア別営業案内 (北海道支社管内)¹⁰⁾

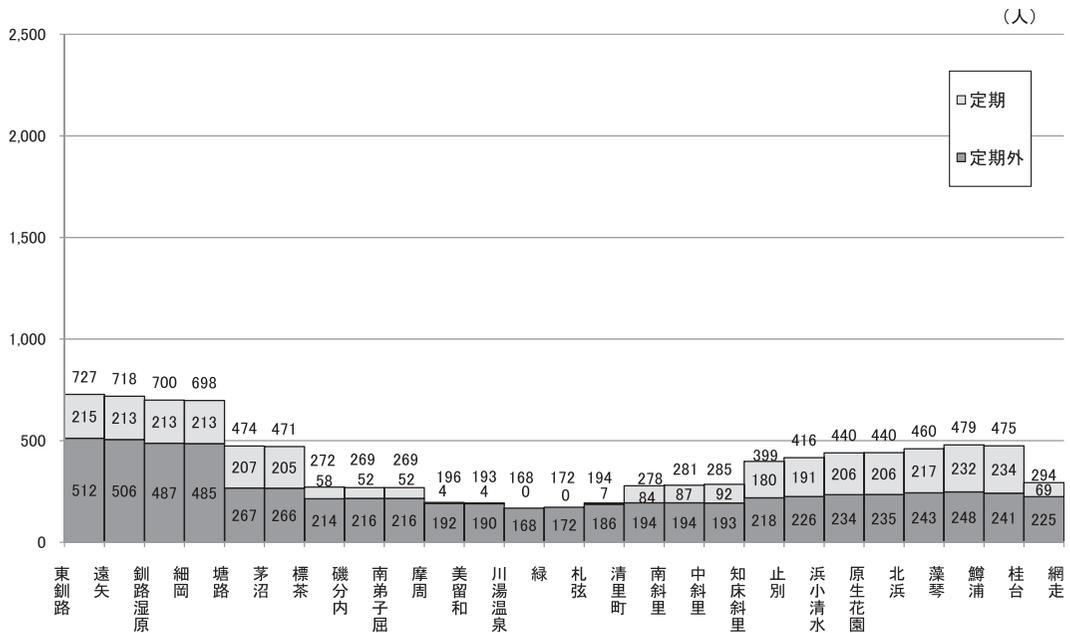


図 3-8 駅間で通過する乗車者数 (釧網線)¹¹⁾

(2) 花咲線

花咲線は、根室本線（滝川～根室）のうち釧路～根室間の通称である。沿線にはラムサール条約登録の湿地など観光地のアクセスと釧路の近郊線の役割がある（図 3-9）。

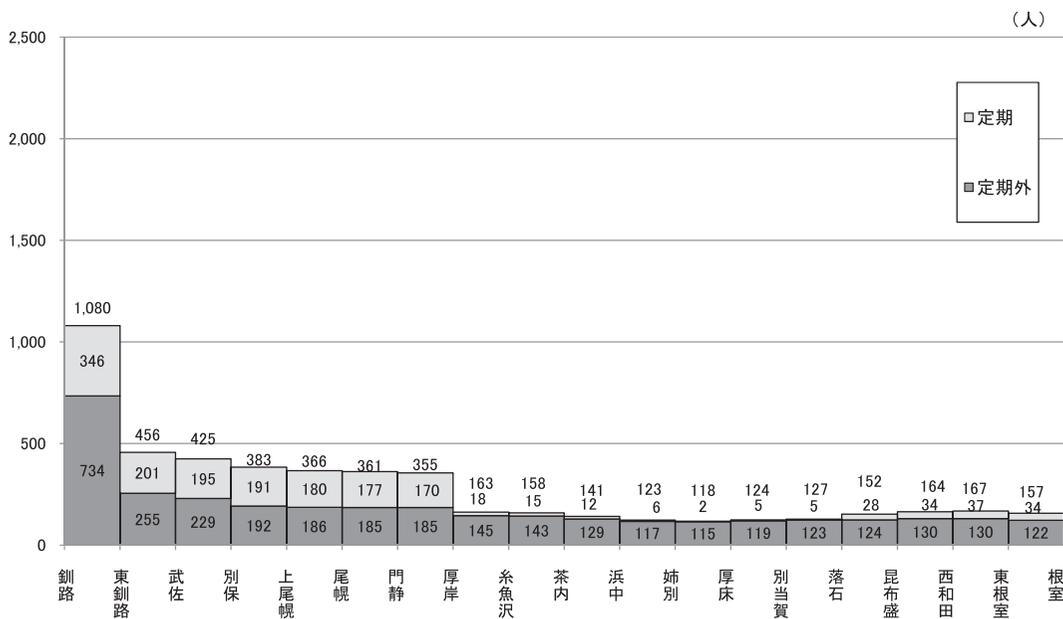


図 3-9 駅間通過の乗車者数（花咲線）¹²⁾

(3) 富良野線

沿線は美瑛、富良野などの観光地があり、その富良野線はこのアクセスに利用される。また北海道第2位の人口を有する旭川の通勤・通学のための線区や富良野の近郊線の役割もある(図 3-10)。

3) 時間ごとの乗車者数

花咲線、釧網線、富良野線における時間ごとの乗車者数の分布を図 3-11 に示す。これらの三線区では前述したように観光地へのアクセスとそれらの線区に接続する地方中核都市への近郊線区としての役割がある。これを時間による分布から推測すると、10時～16時ごろまで比較的増減が少ない乗客者数の分布から観光地へのアクセス利用が想定される。また、7時～8時が全体の中で最も多く、夕方の16時～18時ころが次に多い傾向から地方中核都市への近郊線区としては朝の通勤・通学の利用が考えられる。その結果、時間による乗車者数の分布形状の特徴はこの2つの要素の影響の割合で決まるものと思われる。

以上から本節では、現在の釧網線、花咲線、富良野線の線区の概況と沿線の人口を対比した。その結果、沿線の人口では、各線区の始発・終点の都市を除き市町村の人口は少なく、市町村間に差があまり変化はないことが明らかになった。一方、駅間で通過する乗車者数では釧網線と花咲線では始発・終点の各駅に近づくにつれて定期客の利用が多くなる傾向があった。一方、定期客外はほぼ一定に分布していた。これは三線区の人口の分布と同じ傾向を示していた。

時間ごとの乗車者数では、富良野線で朝の通勤・通学などで集中する傾向が見られる。その

(人)

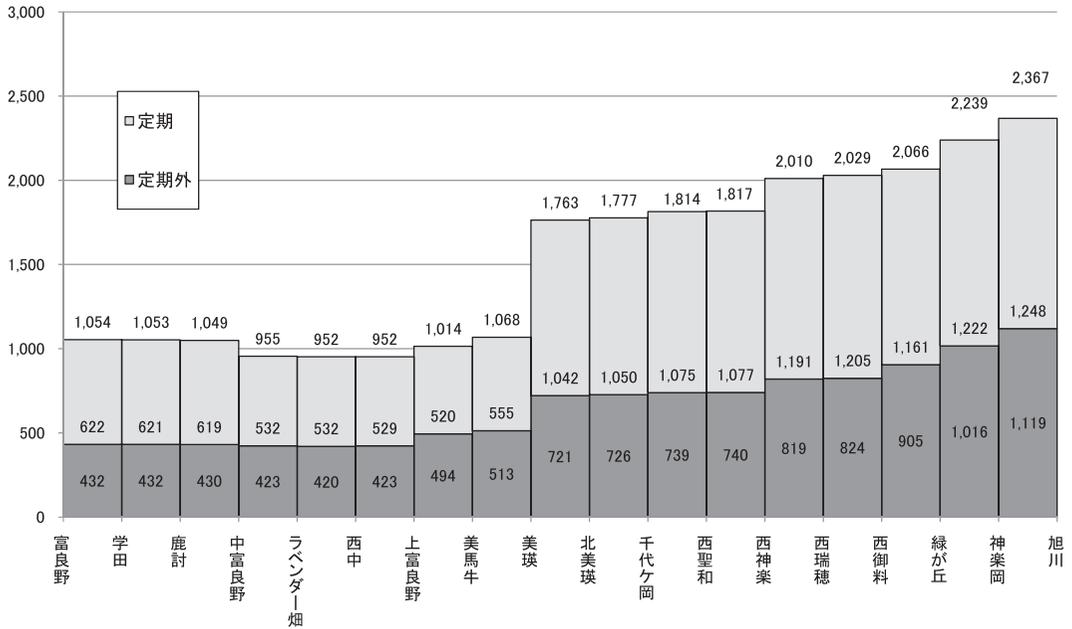


図 3-10 駅間通過の乗車者数（富良野線）¹³⁾

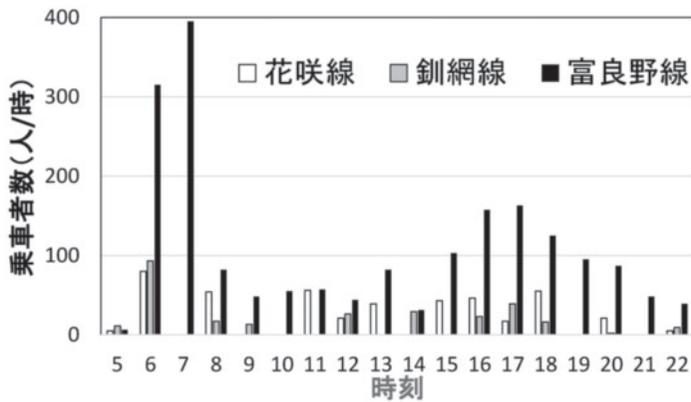


図 3-11 時間ごとの乗車者数

他の二線区ではその傾向はあまり強くなく、沿線の人口が朝の通勤・通学による定期客の影響が全体の乗降客数にあまり影響を与えていないものと推察される。

4 解析

1) 本研究における解析の目的

本研究の統計手法には、頻度法とバイズ法の2種類である。その一つは20世紀初頭にピオ

ソン・ゴセット・フィッシャーらによって確立された統計¹⁴⁾で、現在でもよく用いられている。これはベイズ法の利用拡大と相まってベイズ法と区別するために頻度法と呼ばれている。近年まで頻度法に比べ、ベイズ法はあまり一般的とは言えなかった。それはベイズ法が事後確率の計算で非常に面倒であることからあまり重要視されてこなかったことによる。しかしコンピュータの能力が飛躍的に高まったことから、心理学分野でMCMC法（マルコフ連鎖モンテカルロ法）に代表される膨大なデータ処理方法が確立され、これを簡易な計算ソフト（例えばR、Stanなど）を用いて、事前分布と尤度分布を元に乱数を無作為抽出することで、事後確率を求めることが可能となった¹⁵⁾。これにより心理学分野ではベイズ法による研究が大きく進歩し、さらに統計学分野を巻き込んで他分野に広がりつつある。

この動きに伴い頻度派とベイズ派に分かれて頻度法とベイズ法の比較論争も活発化してきた。これは頻度法が論理構成と精度が高いために20世紀初頭から頻度法が主流であり続けられた。しかし、ベイズ派からフィッシャーが確立したp値が0.05以下ではその仮説が破棄されるとの暗黙の合意に疑問がなげかられた¹⁶⁾。さらに、アメリカ統計学会の「統計的有意性とp値に関する声明」¹⁷⁾で「p値は、それだけでは統計モデルや仮説に関するエビデンスのよい指標にはならない。」と発表されたことから、ベイズ派からベイズ法の正当性に関する論文や出版物が多く出されるようになってきた。その中には「瀕死の統計学を救え！」¹⁸⁾など少し過激と思われるものもあった。この傾向は日本だけではなく、オランダの研究者の論文にも見られた¹⁹⁾。

このような動きに対し、頻度派と自認する林は「統計学は錬金術ではない」²⁰⁾において、当時の「心理学における統計革命」について、「いま起こっている「統計革命」の「革命」の旗手は、ベイズ主義（ベイズ法）である」と認めつつも、「統計革命というよりも、心理データ解析革命と呼ぶ方が穏当であるように思われる。」としている。また、「頻度法が拠り所とする大標本理論に対し、ベイズ法がデータの規模が小さくても対応できる点（著者注：事前確率分布などが採用できることによる）で過去に得られているデータの活用を目指すことができる。」に言及している。そしてp値については「誤解」や数値の悪用、「データ解析の信頼を毀損していること」があるとして「p値の誤用・悪用は学問とデータ解析の価値を墮せしめるに十分な行為である」と述べている。

本研究にあたり地域活性化における鉄道の果たす役割を研究する上で用いるデータ数は限られており、頻度法が拠り所とする大標本理論に適するかは疑問である。そこで、ベイズ法がデータの規模が小さくても対応できる点を考慮して、両方の方法を検討することとした。また、心理学分野ではない鉄道関係分野にベイズ法に基づいた統計処理を行うことも価値があると考えた。そこで昨年より鉄道の地域への貢献度の評価に頻度法と並行してベイズ法も解析に使用することとした。

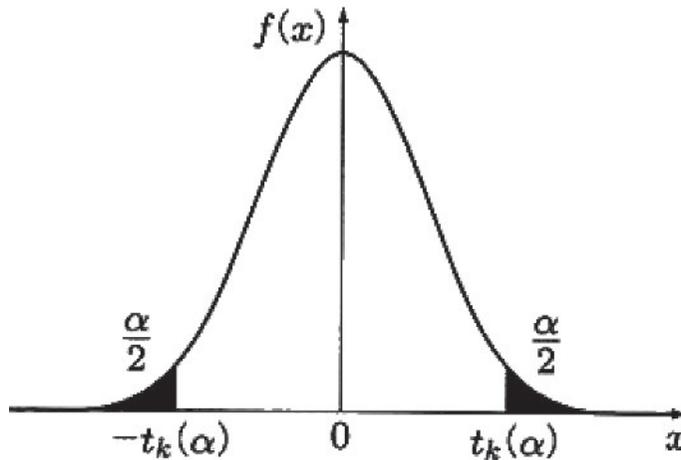


図 4-1 未知のパラメータの範囲 (正規分布関数)

2) 頻度法から見たベイズ法の特徴

(1) 未知のパラメータの扱い方

統計を用いて2線区での1日あたりの乗客者数の分布を比較する場合、平均値と標準偏差を対象とした。ここで平均値を未知のパラメータとする。頻度法では(1)未知のパラメータ(ここでは平均値 μ)を一つ定まった値(真値)が存在すると仮定し(2)未知のパラメータ μ は、正規分布関数などの尤度関数を用いて推定するとしている。図4-1ではp値(図では $\alpha=0.05$)において未知のパラメータは0.95の確率で次の範囲に存在する。 $(-t_k(\alpha) \leq \mu \leq t_k(\alpha))$

一方、ベイズ法では(1)未知のパラメータを、正規分布関数などの確率分布による変数として考え、(2)ベイズの定理によってパラメータの推測は、事後分布(事後確率分布)に基づくとしている。これらの関係を具体的に述べる。ここで、 $P(\theta|D)$:事後確率分布、 $P(D|\theta)$:尤度関数(ここでは正規分布関数)、 $P(\theta)$:事前確率関数(ここでは正規分布関数)とすると、既往の文献²¹⁾により4-1式となり、事後確率分布は尤度関数と事前確率の積に比例する。

$$P(\theta|D) \propto P(D|\theta)P(\theta) \quad (4-1)$$

文献²¹⁾ではこれに基づいた汎用ソフトRによるプログラム電算プログラム(図4-2)を作成している。

このプログラムでは①~④までが入力条件と入力値である。⑤には2ケース(③と④)のデータの平均値と標準偏差を $\theta_1 \sim \theta_4$ が定義している。⑥では尤度関数(確率)の値を計算している。

その例として10,000回繰り返し設定された正規分布関数の一部を図4-3に示す。なお、図4-2の⑦は4-1式を用いてMCMC法によって事後確率分布を求めているプログラムである。

```

#対象：北見市 月ごとの入れ込み観光客数の差 (H28とH31)

#マルコフ連鎖モンテロ法のパッケージ
① library(MCMCpack)
#乱数のセット
② setseed(1)
#H28のデータ(百万人/月)
③ treat<-c(0.33,0.88,0.68,1.20,1.43,0.95,0.78,0.37,0.27,0.39,0.68,0.43)
#H31のデータ(百万人/月)
④ control<-c(0.47,0.96,0.81,1.20,1.57,0.86,0.73,0.36,0.30,0.41,0.72,0.37)

⑤ #尤度関数 theta：推定する4つの母数 (H28の平均θ1・標準偏差θ2
    およびH31の平均θ3・標準偏差θ4)
#分布には正規分布を仮定
lhorm fun <- function(theta, x) {
  a <- x[[1:12]]
  b <- x[[13:24]]
  if(theta[2] < 0.0 || treat[4] < 0){
    l <- -hf
  }
  else {
    l <- sum(log(dnorm(a, m.ean = theta[1], sd = theta[2]))) +
      sum(log(dnorm(b, m.ean = theta[3], sd = theta[4])))
  }
  return(l)
}

⑥
#事後分布f(theta|x)
postdiff <- MCMCmetropIR(lhorm fun, theta.in = c(1,0.5,1,0.5),
  x = c(treat,control),
  thin = 1, mcmc = 10000, burnin = 500,
  tune = 1.0,
  verbose = 500, logfun = TRUE)

⑦
#事後分布のプロット
plot(postdiff)

```

図 4-2 事後確率 (汎用ソフト R による)

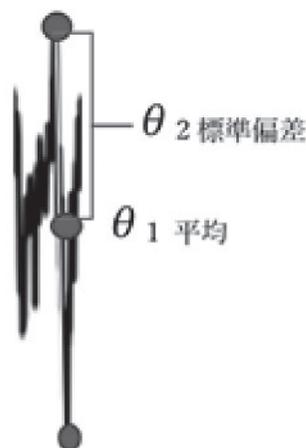


図 4-3 繰り返し計算に用いられる平均値と標準偏差の例

このように、事後分布の平均値を未知のパラメータとして正規分布関数などの確率分布による変数として考え、ベイズの定理によって事後分布の平均値の推測は、事後確率分布に基づくとしている。

(2) ベイズファクターと頻度法で用いる効果量

ベイズファクターとは、帰無仮説と対立仮説のいずれかが可能性が高いかを表す指標¹⁹⁾である。ここで帰無仮説を H_0 の尤度と対立仮説 H_1 の尤度とするとベイズファクター BF_{10} は 4-2 式で定義される。

$$BF_{10} = \frac{P(D \vee H_1)}{P(D \vee H_0)} \quad (4-2)$$

ベイズファクター BF_{10} の基準の例を表 4-1 に示す²²⁾

また、次の 4-3 式はベイズファクター BF_{10} の逆数としてベイズファクター BF_{01} を示している。

表 4-1 ベイズファクター BF_{10}

BF_{10}	対立仮説の評価	
1-3	Anecdotal	裏付けが乏しい
3-10	Moderate	中程度
10-30	Strong	強い
30-100	Very strong	かなり強い
>100	Extreme	非常に強い

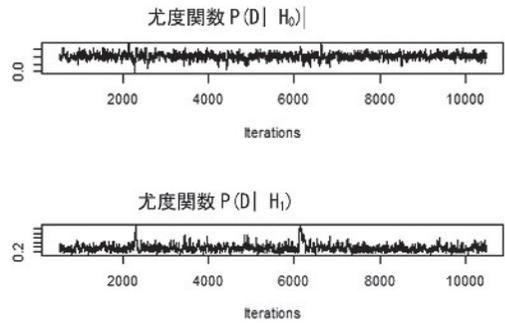


図 4-4 尤度関数 (MCMC 法)

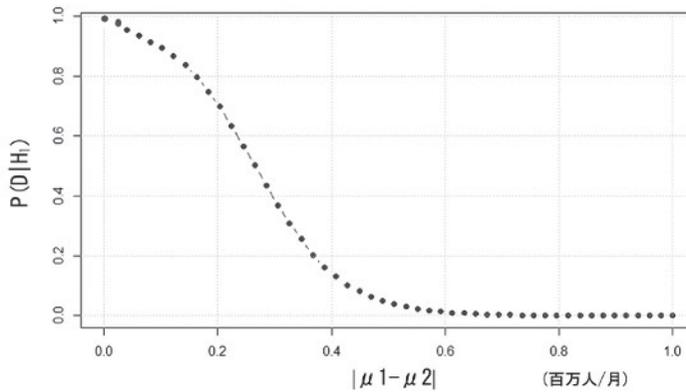


図 4-5 $P(D|H_1)$ の分布

$$BF_{01} = \frac{P(D \vee H_0)}{P(D \vee H_1)} \quad (4-3)$$

ここである線区における 2 年間の乗降者数の平均値を μ_1 と μ_2 とする。これらを用いて帰無仮説 $H_0: \mu_1 \neq \mu_2$ とし、対立仮説 $H_1: \mu_1 = \mu_2$ とする。正規分布関数である尤度関数 $P(D|H_0)$ と $P(D|H_1)$ を MCMC 法によって図 4-2 のプログラムから 10000 回のケースを求めた (図 4-4)。

これに基づき、 $|\mu_1 - \mu_2|$ に対する $P(D|H_1)$ の分布を図 4-4 に示す。この図から $\mu_1 - \mu_2 = 0$ の対立仮説 $P(D|H_1)$ は 1 であり、帰無仮説 $P(D|H_1)$ は 0 となる。この結果からベイズファクター BF_{10} は「>100 以上」となり非常につよい仮説となる。一方、 $|\mu_1 - \mu_2| = 0.1$ では、 $\mu_1 = \mu_2$ とする誤差が 0.1 (百万人/月) を示している。その場合図 4-5 より対立仮説 $P(D|H_1)$ は 0.9 であるので帰無仮説 $P(D|H_1)$ は 0.1 となる。よってベイズファクター $BF_{10} = 9$ となり、表 4-1 より対立仮説は中程度に支持される仮説となる。

頻度法で用いる効果量 d とは 2 種類
の母集団の平均値 μ_1 と平均値 μ_2 につ
いて、差 $\mu_1 - \mu_2$ と差の標準偏差 σ との
比により 4-4 式で定義している²³⁾。

$$d = \frac{\mu_1 - \mu_2}{\sigma} \quad (4-4)$$

表 4-2 効果量 d

d	効果	
<0.2	Trivial	わずか
0.2-0.5	Small	小さい
0.5-0.8	medium	中程度
>0.8	Large	大きい

(3) バイズ法のバイズファクターと頻度法の情報量規準

従前は帰無仮説が棄却されるか否かとの判別に用いられていたが、バイズファクターの導入により、Eric-Jan Wagenmakers が「バイズファクターの値が 3 より小さいときには何か結論を与えるには弱い結果である」²⁴⁾ ということを示唆しているように、仮説の可能性の程度をバイズファクターで評価することができる。一方で、頻度法においても赤池情報量規準などに代表される情報量規準によって複数の数量化モデルにおける適性の相対評価をする方法がある。しかし、林が「二者のどちらがより良い方法かという議論にはあまり意味がない」¹⁸⁾ としている。つまり「重要なことは、研究上の関心がどのように量的に表現され、その量がどの程度であるかを研究デザイン時に把握しておくことであろう」²⁵⁾ と指摘しているように、数量化モデル自体によってバイズ法と頻度法を使い分けるのが重要と考えられる。

3) 本研究における仮説検定の条件設定

釧網線と花咲線では列車間隔が 1 時間や 2 時間が混在している。一方富良野線は列車間隔の最大でも 1 時間以内である。そこでこれらの 3 線区を比較する上で、1 時間に間隔としてその間に列車がない場合は乗客者数を 0 とした。また時間ごとの乗車者数分布の傾向を検討するために、各線区の時間ごとの乗車者数を一日の総計の乗車者数により標準化した。

仮説検定における頻度法については詳しくは文献²¹⁾ によるが、ここではその一部を引用する。「ネイマン・ピアソンの共同研究 (1928 年) では、検定されるべき仮説 (帰無仮説) とそれ以外の仮説 (対立仮説) の 2 つの仮説を設けて p 値を用いて検定する理論を生み出した。」また、「フィッシャー流を拡大して 2 つのケースでは $SE = \sqrt{\sigma_A^2/m + \sigma_B^2/n}$ としている。ここで帰無仮説 $\mu_A = \mu_B$ より $\mu_A - \mu_B = 0$ の仮説となる。標本集団の平均値の差 $x_A - x_B$ ($= \mu_A - \mu_B$) のばらつきの程度を調べる。このばらつきに係する確率変数 $Z = \frac{x_A - x_B}{SE}$ を求める」としている。これから p 値を求める。前述のバイズ法ではこの確率変数 Z を効果量 d としている。

例えば、釧網線と花咲線における時間ごとの乗車者数分布のそれぞれの平均値 (μ_A , μ_B) の差 $\delta 1$ ($= \mu_A - \mu_B$) する。ここで帰無仮説 $H_0: \delta 1 = 0$ とし対立仮説 $H_1: \delta 1 \neq 0$ とする。ここで有意水準 $\alpha = 0.05$ とする。ここで p 値とは、「帰無仮説が正しいとしたときに、観測データから得られる確率」として、帰無仮説を棄却する条件は、 p 値が有意水準 α より小さいことを意

味する。ここで、各線区で列車間隔を1時間としたことから、6時～22時までの17個のデータを対象とする。

ベイズ法において標準化を取り入れると4-1式から4-5式として事後確率分布 $f(\theta|D)$ を求めることができる。

$$f(\theta|D) = \frac{f(D|\theta)f(\theta)}{f(D)} \quad (4-5)$$

ここで $f(D)$ は正規化定数と呼ばれる。また $f(\theta)$ は事前分布で一般的には正規分布を用いる²⁶⁾。なお $f(\theta|D)$ は尤度関数である。

5 解析結果

1) 概要

本研究では、統計計算ソフト JASP¹⁾ により頻度法とベイズ法により解析を行う。その精度は多くの研究で検証されている²⁾。JASP では頻度法では一般的な統計解析である t 検定、分散分析、 χ^2 検定、回帰分析に加え、主成分分析、因子分析などの多変量分析にも対応している。また、ノンパラメトリック検定、階層線形モデルやネットワーク分析といった高度な統計分析も可能としている。一方で、ベイズ法では上記の頻度法でのそれぞれの検定や分析ができるように構成されている。なお、操作方法などの詳細は本論文の補足で説明する。このように、前回の論文¹⁹⁾ では計算過程を検証するために汎用ソフト R や Stan によるプログラム作成と操作によって図 4-2、図 4-3、図 4-4、図 4-5 を求めたが、JASP では計算過程はシステム内部で行われ、出力として頻度法の t 値やベイズ法のベイズファクター BF_{10} が出力される。

次に、対象の線区の略称を釧網線：「SL」、花咲線：「HL」、富良野線：「FL」とする。これらの線区のうち二線区を選択し、帰無仮説と対立仮説の仮説検定により平均値の差の有無を検討することとする。よって帰無仮説 H_0 では、「2線区の時間ごとの乗車者数分布に差がない： $\delta_1 = 0$ 」とし、対立仮説は「二線区の時間ごとの乗車者数分布に差がある： $\delta_1 \neq 0$ 」となる。

また JASP の出力様式にある Raincloud Plots を示す。ここの2データの対応、箱ひげ図、分布状態を全体で把握できる表示方法である。ここで箱ひげ図では箱部分は全体の50%に範囲を示し、その内部の線は中央値である。ひげの2つの端部は最大値と最小値を表している。

2) 解析結果

(2-1) 頻度法

頻度法ではその分布に真値が存在していると仮定する。ここでは真値を平均値としている。また、分布の形状を比較するために時間ごとの乗降客数をその一日の総計により標準化している。標準化した二線区の平均の差を効果量 d とし、表 4-2 により効果量 d を判定する。

表 5-1 仮説検定表（花咲線（HL）と釧網線（SL））

Paired Samples T-Test

Measure 1	Measure 2	t	df	p	VS-MPR*	Mean Difference	SE Difference	95% CI for Mean Difference		Cohen's d
								Lower	Upper	
HL	- SL	0.090	16	0.930	1.000	0.002	0.019	-0.038	0.041	0.022

Note. Student's t-test.

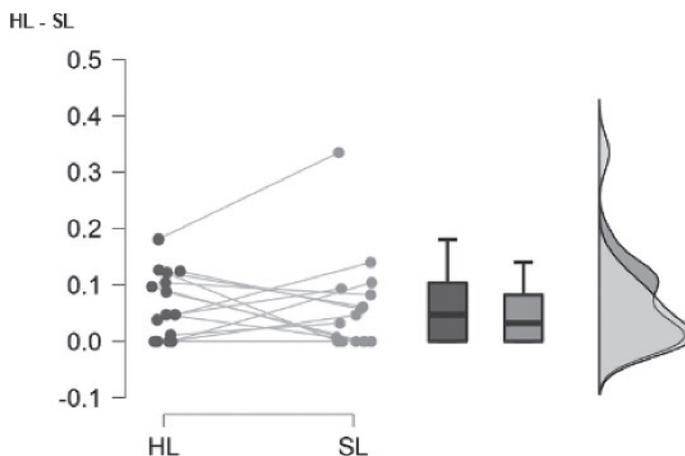


図 5-1 Raincloud Plots（花咲線（HL）と釧網線（SL））

①花咲線（HL）と釧網線（SL）

解析した結果を表 5-1 に示す。この表から p 値 (0.930) > t 値 (0.090) であるので、帰無仮説 $H_0: \delta 1 = 0$ が成り立つ。つまり、花咲線（HL）と釧網線（SL）の時間ごとの乗車者数分布のそれぞれの平均値の差 $\delta 1$ はないと認められる。次に 95% 以内に入る範囲は -0.038 ~ 0.041 である。その効果量 d は、0.022 であり、表 4.2 から、0.2 以下であるのでその効果はわずかである。

また Raincloud Plots を図 5-1 に示す。この図において左の図は時間ごとの標準化したデータ比較では 1 点を除き対応関係が 0 ~ 0.15 の範囲内に収まっている。次に、この図の中央にある箱ひげ図では 50% の範囲を示す箱の範囲と中央値はほぼ同じと見なせる。また右の図は 2 データの分布の重なり度合いも同様に同じと考えられる。

以上から p 値、効果量 d、Raincloud Plots の 3 点から総合すると、花咲線（HL）と釧網線（SL）の分布にあまり差がないことが分かる。

②富良野線（FL）と釧網線（SL）

解析した結果を表 5-2 に示す。この結果から p 値 (0.912) > t 値 (0.112) であるので、帰無仮説 $H_0: \delta 1 = 0$ が成り立つ。よって①と同様に富良野線（FL）と釧網線（SL）の時間ごと

表 5-2 仮説検定表 (富良野線 (FL) と釧網線 (SL))

Paired Samples T-Test

Measure 1	Measure 2	t	df	p	VS-MPR*	Mean Difference	SE Difference	95% CI for Mean Difference		Cohen's d
								Lower	Upper	
FL	- SL	0.112	16	0.912	1.000	0.002	0.019	-0.038	0.043	0.027

Note. Student's t-test.

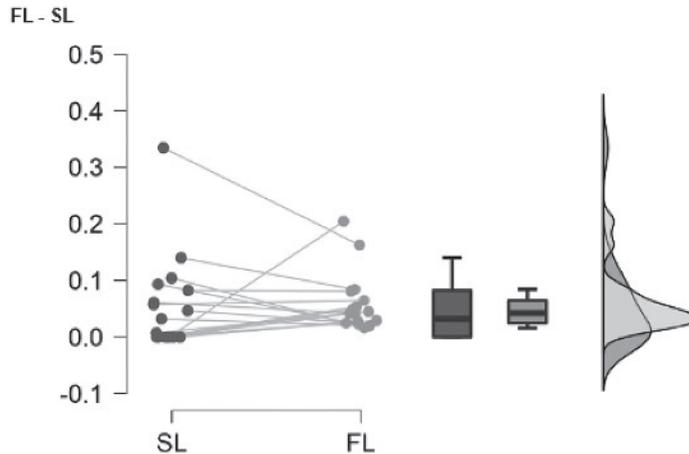


図 5-2 Raincloud Plots (富良野線 (FL) と釧網線 (SL))

の乗車者数分布のそれぞれの平均値の差 $\delta 1$ はないと認められる。次に 95% 以内に入る範囲は $-0.038 \sim 0.043$ である。その効果量 d は、0.027 であり、表 4.2 から、0.2 以下であるのでその効果はわずかである。

また Raincloud Plots を図 5-2 に示す。この図において左の図は時間ごとの標準化したデータ比較では 1 点を除き対応関係が $0 \sim 0.2$ の範囲内に収まっている。次に、この図の中央にある箱ひげ図では 50% の範囲を示す箱の範囲は富良野線 (FL) の範囲は小さいが中央値はほぼ同じと見なせる。また右の図では富良野線 (HL) で近郊線の分布傾向が見られ、ピーク値は異なるが時間ごとの分布は同じような傾向が見られた。

以上から、①と同じく p 値などの 3 点から総合すると、富良野線 (FL) と釧網線 (SL) の分布にあまり差がないことが分かる。

③富良野線 (FL) と花咲線 (HL)

解析した結果を表 5-3 に示す。この結果から p 値 (0.977) $>$ t 値 (0.029) であるので、帰無仮説 $H_0: \delta 1 = 0$ が成り立つ。よって①と同様に富良野線 (FL) と花咲線 (HL) の時間ごとの乗車者数分布のそれぞれの平均値の差 $\delta 1$ はないと認められる。次に 95% 以内に入る範囲は

表 5-3 仮説検定表（富良野線（FL）と花咲線（HL））

Paired Samples T-Test

Measure 1	Measure 2	t	df	p	VS-MPR'	Mean Difference	SE Difference	95% CI for Mean Difference		Cohen's d
								Lower	Upper	
FL	- HL	0.029	16	0.977	1.000	4.827e-4	0.016	-0.034	0.035	0.007

Note. Student's t-test.

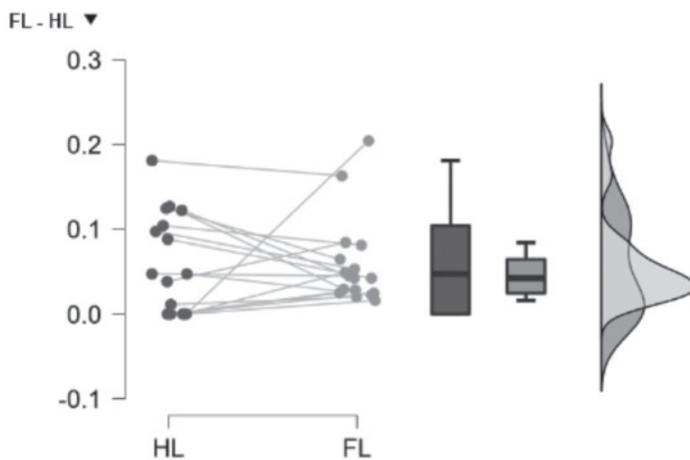


図 5-3 Raincloud Plots（富良野線（FL）と花咲線（HL））

-0.034~0.035である。その効果量 d は、0.007であり、表 4.2 から、0.2 以下であるのでその効果はわずかである。

また Raincloud Plots を図 5-3 に示す。この図において左の図は時間ごとの標準化したデータ比較では 2 点を除き対応関係が 0~0.15 の範囲内に収まっている。次に、この図の中央にある箱ひげ図では 50% の範囲を示す箱の範囲は富良野線（FL）の範囲は小さいが、中央値はほぼ同じと見なせる。また右の図では富良野線（HL）で近郊線の分布傾向が見られ、2 データの分布の重なり度合いはピークの形状が異なっている。

以上から、①と同じく p 値などの 3 点から総合すると、富良野線（FL）と花咲線（HL）の分布にあまり差がないことが分かる。

(2-2) バイズ法

頻度法では、二線区の時間ごとの乗車者数分布の平均値の差 δ_1 が対象となったが、バイズ法では時間ごとの乗車者数分布を対象に二線区の差である効果量 δ の検定を行う。ここで帰無仮説 H_0 を $\delta = 0$ とし対立仮説 H_1 を $\delta \neq 0$ とする。また 4-5 式において事前分布を正規分布関数とする。

表 5-4 バイズファクター BF_{10}
(釧網線 (SL) と花咲線 (HL))

Measure 1		Measure 2	BF_{10}
SL	-	HL	0.250

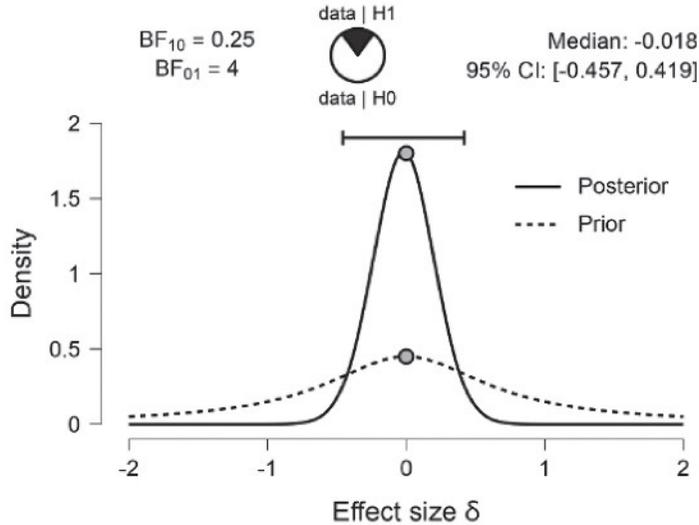


図 5-4 バイズ法による検定結果 (釧網線 (SL) と花咲線 (HL))

① 釧網線 (SL) と花咲線 (HL)

解析した結果を表 5-4 に示す。この結果からバイズファクター $BF_{10}=0.250$ である。これは表 4-1 の BF_{10} が 1 以下であり、対立仮説の裏付けは乏しいことを示している。つまり帰無仮説「釧網線 (SL) と花咲線 (HL) の時間ごとの乗車者数分布に差がない」が支持されることとなる。

図 5-4 は釧網線 (SL) と花咲線 (HL) の時間ごとの乗車者数分布の中央値の差 δ (Median) と 95% 信頼区間 (-0.457, 0.419) が示されている。また図の実線は事前分布で正規分布関数であり、破線は事後分布である。この結果は図からも δ がほぼ 0 であり、釧網線 (SL) と花咲線 (HL) の差が小さいと考えられる。

② 釧網線 (SL) と富良野線 (FL)

解析した結果を表 5-5 に示す。この結果からバイズファクター $BF_{10}=0.250$ である。これは①と同様に BF_{10} が 1 以下であり、対立仮説の裏付けは乏しいことを示している。つまり帰無仮説「釧網線 (SL) と富良野線 (FL) の時間ごとの乗車者数分布に差がないこと」が支持されることとなる。

図 5-5 は釧網線 (SL) と富良野線 (FL) の時間ごとの乗車者数分布の中央値の差 δ

表 5-5 バイズファクター BF_{10}
(鈎網線 (SL) と富良野線 (FL))

Measure 1		Measure 2	BF_{10}
SL	-	FL	0.250

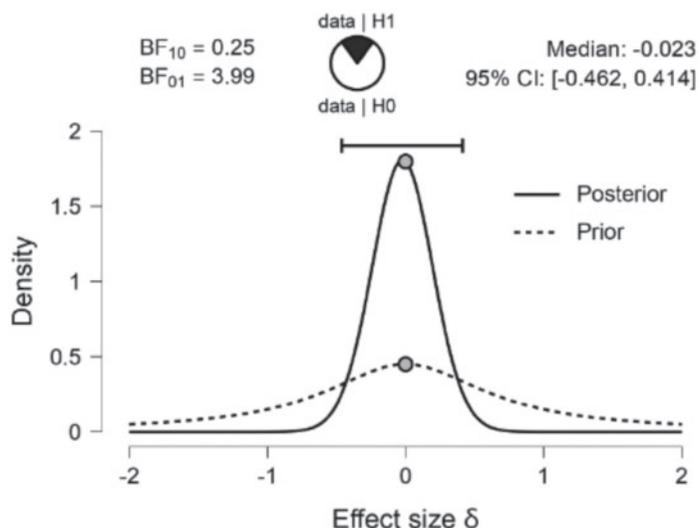


図 5-5 バイズ法による検定結果 (鈎網線 (SL) と富良野線 (FL))

(Median) と 95%信頼区間 (-0.462, 0.414) が示されている。また図の実線は①と同様に事前分布で正規分布関数であり、破線は事後分布である。この結果は図からも δ がほぼ 0 である。この結果は図からも①と同様に鈎網線 (SL) と富良野線 (FL) の差が小さいと考えられる。

③富良野線 (FL) と花咲線 (HL)

解析した結果を表 5-6 に示す。この結果からバイズファクター $BF_{10}=0.249$ である。これは①と同様に BF_{10} が 1 以下であり、対立仮説の裏付けは乏しいことを示している。つまり帰無仮説「鈎網線 (SL) と富良野線 (FL) の時間ごとの乗車者数分布に差がないこと」が支持されることとなる。

図 5-6 は富良野線 (FL) と花咲線 (HL) の時間ごとの乗車者数分布の中央値の差 δ (Median) と 95%信頼区間 (-0.431, 0.444) が示されている。また図の実線は①と同様に事前分布で正規分布関数であり、破線は事後分布である。この結果は図からも δ がほぼ 0 である。

この結果は図からも①と同様に鈎網線 (SL) と花咲線 (HL) の差が小さいと考えられる。

表 5-6 バイズファクター BF_{10}
富良野線 (FL) と花咲線 (HL)

Measure 1		Measure 2	BF_{10}
FL	-	HL	0.249

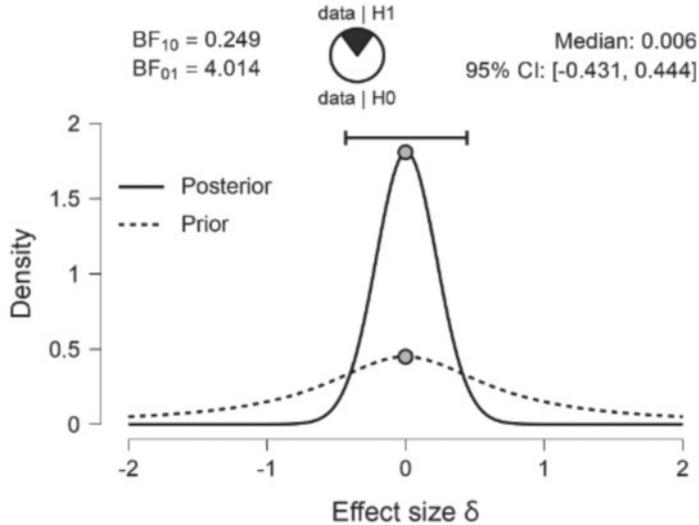


図 5-6 バイズ法による検定結果 (富良野線 (FL) と花咲線 (HL))

3) 解析結果のまとめ

3章では定性的な考察を行ったが、本章では定頻度法とバイズ法の統計手法により定量的に検討した。その結果を表 5-7 に示す。なお、この表の枠内で左下側は頻度法、右上側はバイズ法を記載している。また、釧網線、花咲線、富良野線のそれぞれ二線を選択して全部を組み合わせで解析した。

解析の前提として平均値 (μ_A, μ_B) の差 $\delta (= \mu_A - \mu_B)$ とし、帰無仮説 $H_0 : \delta = 0$ とし対立仮説 $H_1 : \delta \neq 0$ とした。頻度法では、平均値の差 δ において帰無仮説と対立仮説の仮説検定により p 値を求めた。ここで、 p 値が 0.05 より大きいときは帰無仮説 H_0 が有意ではないとして帰無仮説は棄却されない。一方、バイズ法では頻度法と同様に帰無仮説と対立仮説の仮説検定においてバイズファクター BF_{10} を求め、表 4-1 を用いて対立仮説の評価を行った。この場合に BF_{10} が 3 以下であれば裏付けが乏しく、100 以上であれば裏付けが非常に強いことになる。

以上より表 5-7 によると、頻度法とバイズ法ともに時間別の釧網線、花咲線、富良野線では平均の差が認められなかった。これは時間ごとの乗車者数分布において、これらの三線区で大きな差が認められないこと明らかになった。

表 5-7 解析結果のまとめ

		ベイズ法		
		釧網線 SL	花咲線 HL	富良野線 FL
頻度法	釧網線 SL		0.250 0.930	0.250 0.912
	花咲線 HL	0.250 0.930		0.249 0.977
	富良野線 FL	0.250 0.912	0.249 0.977	

6 ま と め

本研究の対象である道東の釧網線、花咲線、富良野線の三線区では北海道の他の線区と同様に、貨物輸送を中心に炭鉱などの鉱産資源や海産物を海上輸送につなげて本州へ移送するための手段として建設されてきた。しかし、鉱産資源の枯渇や石炭から石油などのエネルギー革命によってその目的が無くなり、これに道路交通網の整備や船舶の大型化の影響を受けると、これらの線区が JR 北海道における「単独で維持困難線区」に属することとなった。

本研究では、「単独で維持困難線区」として全体と見ること必要だが、個別の線区ごとに分析することも重要であるとの視点で、各線区の建設の歴史と現状の利用実態や線区が関係している市町村の人口の対比、乗降客数の分布の比較を試みた。その結果以下の点が明らかになった。

①釧網線の一部は、当時道内で2番の鉄道でアトサヌプリ（硫黄山）から標茶間に敷設された。その後硫黄の枯渇で明治29年（1896年）には硫黄採掘は中止され、鉄道輸送も廃止となった。この路線は釧網線に繰り入れられた。花咲線は海産物の輸送のために1921年（大正10年）に根室まで延伸し富良野線では旭川と富良野・帯広・釧路を結ぶ路線の一部として1899年（明治32年）に建設された。しかし、滝川と富良野間の赤平炭鉱の輸送のために滝川・富良野から釧路までの路線（根室線）が1913年（大正2年）に完成したことにより、旭川・富良野間が分離されて富良野線として発足した。

以上から現状の釧網線、花咲線、富良野線の三線区において敷設位置は今から100年以上前からあまり変わっていないことを確認した。

②現在の釧網線、花咲線、富良野線の各線区では、それぞれの始発・終点の都市を除き市町村の人口は少なく、市町村間に差があまり変化はないことが明らかになった。一方、これらの線区を通過する乗車者数では釧網線と花咲線では始発・終点の各駅に近づくにつれて定期客の利用が多くなる傾向があった。特に富良野線ではこの傾向が大きく見られた。一方、定期客外はほぼ一定に分布していた。これは三線区の人口の分布と同じ傾向を示していた。

③本研究では統計ソフト JASP を用いて仮説検定による釧網線、花咲線、富良野線における時間ごとの乗車者数の分布を検討した。その結果、頻度法とベイズ法ともに時間別の釧網線、

花咲線、富良野線では差が認められなかった。

以上から、釧網線、花咲線、富良野線の時間ごとの乗車者数の分布などからその差はあまりないことが明らかになった。また、統計ソフト JASP は本研究で扱った鉄道分野でも有効な手法であることを確認した。

補足

JASP の使用方法

1) 解析の内容

JASP は心理統計学のアムステルダム大学のワーゲンメーカー (Eric-Jan Wagenmakers) らによって製作されたソフトで Windows や他の OS も対応可能である。JASP (free) download Windows version (freedownload-manager.org) によりダウンロードできる。

JASP では、頻度法で使われる分析とベイズ論で使われる分析ができる。この例を図 1 に示す。JASP の操作で t 検定を行うと、この図が最初に表示される。このように頻度法 (ここでは Classical) とベイズ法 (ここでは Bayesian) がセットになっている。

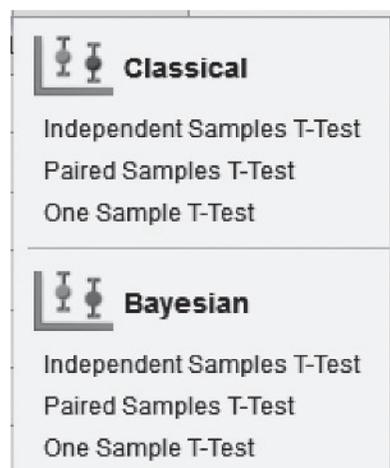


図 1 頻度法 (Classical) とベイズ法 (Bayesian) の選択

2) 立ち上げの操作方法

このソフトを作動させて、左端をクリックすると図 2 となる。ここで “Open” を開くと “Recent”, “Computer”, “OSF”, “Data Library” の 4 つの選択肢が表示される。解析に必要なファイルのアクセスでは、最近使用したファイルを用いるには “Recent” を開く。また、コンピュータの中にあるデータを選択して読み込むには “Computer” を使用する。“OSF” は、Open Science Framework の略であり、フリーのオープンソースウェブアプリケーションである。“Data Library” には、今まで履歴が示される。

3) データの読み込み

最初に excel のデータを csv 形式で作成する。“Computer” を用いてファイルを開き、作成した csv 形式のデータを取り込む。例を図 3 に示す。HL は花咲線を示し SL は釧網線を示している。表の数字は時間に対応するその線区の一日の総計で標準化した値である。

4) 記述統計量 (Descriptives)

“Descriptives” をクリックし、HL (花咲線) と SL (釧網線) のデータを variables (変数)

へ移動すると図4が表示される。ここで図の右にある統計量である記述統計量（Descriptive statistic）が計算される。そのなかで Valid とは有効変数の数， Missing とは値の表示のない数， Mean は平均値， Std. Deviation は標準偏差， Minimum は最小値， Maximum は最大値である。

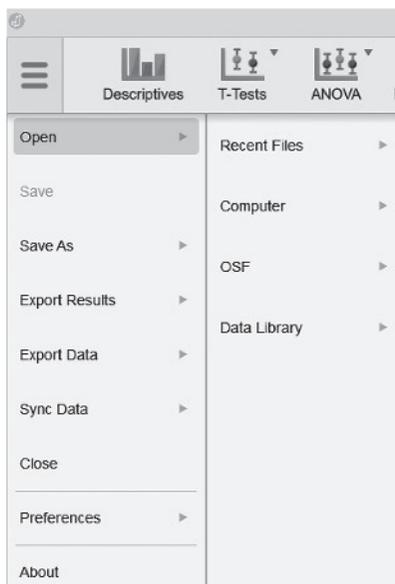


図2 立ち上げの操作画面

	V1	HL	SL
1	1	0.180995475	0.334532374
2	2	0	0
3	3	0.122171946	0.061151079
4	4	0	0.04676259
5	5	0	0
6	6	0.126696833	0
7	7	0.047511312	0.09352518
8	8	0.088235294	0
9	9	0	0.104316547
10	10	0.097285068	0
11	11	0.104072398	0.082733813
12	12	0.038461538	0.14028777
13	13	0.124434389	0.057553957
14	14	0	0
15	15	0.047511312	0.007194245
16	16	0	0
17	17	0.011312217	0.032374101

図3 データの入力画面

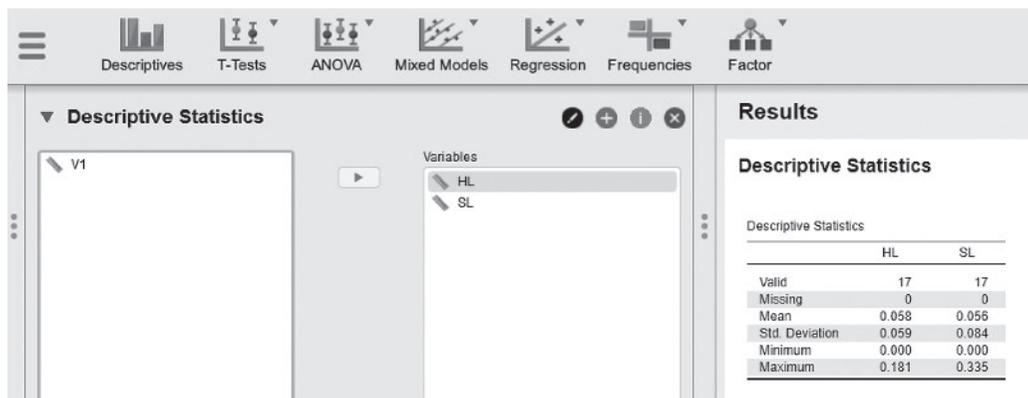


図4 記述統計量の出力

参考文献

- 1) 北海道新聞：フォーカス「赤字線区 協議会設置へ」, 2022年7月26日
- 2) 難波修史・由比 匠・神原広平・宮谷真人・中尾 敬：JASPによる心理学者のためのベイズ統計, 広島大学心理学研究第16号, pp.97-108, 2016
- 3) JASP Team. (2016). JASP (Version 0.8.0.1) [Computer software]. Retrieved from <https://jasp-stats.org/>
- 4) 開拓を支えた小さな鉄道・簡易軌道：http://kai-hokkaido.com/feature_vol45_lightrail/
- 5) 廃線鉄道寮：<http://haisentetsudouryou.travel.coocan.jp/134shibetsu1.html>
- 6) 令和2年住民基本台帳人口・世帯数：<https://www.pref.hokkaido.lg.jp/ss/tuk/900brr/index2.htm>
- 7) JR北海道 釧網線第2期アクションプラン：別添資料 p4, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_01_actionplan_2ki.pdf
- 8) JR北海道 花咲線第2期アクションプラン：別添資料 p4, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_02_actionplan_2ki.pdf
- 9) JR北海道 富良野線第2期アクションプラン：別添資料 p4, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_05_actionplan_2ki.pdf
- 10) JR貨物サービス向上の取り組み：<https://www.jrfreight.co.jp/service/area.html#hokkaido>
- 11) JR北海道 釧網線第2期アクションプラン：別添資料 p11, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_01_actionplan_2ki.pdf
- 12) JR北海道 花咲線第2期アクションプラン：別添資料 p11, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_02_actionplan_2ki.pdf
- 13) JR北海道 富良野線第2期アクションプラン：別添資料 p12, https://www.jrhokkaido.co.jp/corporate/region/pdf/8senku/8_05_actionplan_2ki.pdf
- 14) デイヴィット・ザルツブルグ（竹内恵行・熊谷悦生訳）：統計学を拓いた異才たち, 日本経済新聞, 2006.3
- 15) 朝野熙彦, 土田尚弘, 小野滋：ベイズ統計学, 朝倉書店, p95, 2017.2
- 16) 豊田秀樹：瀕死の統計学を救え!, 朝倉書店, pp.6-13, 2020.3
- 17) 清水優菜, 山本光：JASPによるデータ解析, コロナ社, p64, 2020.3
- 18) 豊田秀樹：瀕死の統計学を救え!, 朝倉書店, 2020.3
- 19) Thomas J. Faulkenberry, Alexander Ly, Eric-Jan Wagenmakers: Bayesian Inference in Numerical Cognition: A Tutorial Using JASP, Journal of Numerical Cognition, Vol. 6(2), pp. 231-259, 2020
- 20) 林賢一：統計学は錬金術ではない, Japanese Psychological Review, Vol. 61, No. 1, pp. 147-155, 2018
- 21) 上浦正樹：北海道主要都市の観光入込客数に関する基礎的研究, 開発論集 第108号, pp. 1-17, 2021.9
- 22) Eric-Jan Wagenmakers 他：Bayesian inference for psychology. Part II: Example applications with JASP, Psychon Bull Rev, 2017
- 23) Mark Goss-Sampson: Statistical-Analysis-in-JASP-A-Students-Guide-v14-Nov, 2020
- 24) Eric-Jan Wagenmakers: Bayesian Hypothesis Testing, Using JASP, 2019
- 25) 清水優菜, 山本光：JASPによるデータ解析, コロナ社, p74, 2020.3
- 26) 清水優菜, 山本光：JASPによるデータ解析, コロナ社, p69, 2020.3