

公開月例研究会講演記録〈第 281 回 (2014. 9. 17)〉

「モバイル空間統計」

携帯電話ネットワークを活用した人口推計技術とその応用

NTT ドコモ先進技術研究所

寺田 雅之

ただ今ご紹介いただきました寺田雅之です。

本日は、ご紹介にもありましたように、モバイル空間統計という携帯電話ネットワークから作った人口統計についてお話ししたいと思います。

タイトルに、携帯電話ネットワークを活用した推計技術、なんて難しそうな言葉もあったりしますが、今日はあまり技術的にややこしいマニアックな話は抜きにして、実際にやってみた事例などのお話を中心にご紹介したいと思います。ということで、どうぞ肩の力を抜いてお聞きください。

本日の順番ですが、まずモバイル空間統計ってそもそも何でしょう、というお話から、何の役に立つの、どうやって作るの、そしてそれってどのくらい信じられるの、という順番でお話する予定です。

モバイル空間統計ってそもそも何でしょう

さて、モバイル空間統計ってそもそも何でしょう…というところのご紹介の前に、ちょっとこの写真をご覧くださいませでしょうか。これ、渋谷のスクランブル交差点の写真になるんですが、ご覧の通り沢山の人でごった返しています。人がなんだか沢山いる場所、ということではなか、海外でも結構有名らしく、海外の方々の口コミ観光スポットのようなものにもなっているそうです。

海外の有名人の方が、わざわざこの交差点をバックにご自分の写真を撮ってブログなんかアップしてみたり、というのはともかくとしまして、さて問題です。ここに何人の人がいるでしょう。

と言われても、こんなにいると、まあ数えるのは無理ですね。ひょっとして日本野鳥の会の方々ならできたりするのかもしれませんが、それ

でもたとえば24時間ずっと張りついて、というのは難しいわけで。たとえばこの例のように人が沢山来ていることはわかっていたとしても、実際のところそれがどのくらい、という定量的なことを知るのはなかなか難しいのが実状です。

モバイル空間統計ってそもそも何でしょう、というお話に戻ります。簡単にいいますと、この例のように、人がたくさん来ている、もしくは出かけて行って減っている、というときに、さて実際のところどういうところからどのくらいの人か来ているのか、もしくはどういうところにどのくらいの人が行っているのか、というのを、携帯電話ネットワークの統計データから定量的に推定してみましよう、というものです。

それをたとえば防災計画や町づくりなんかの意思決定に活かしていこう、というものになります。

ちょっと言葉だけではイメージしづらいですよ。ね、いくつか例を持ってきましたので動画でご紹介したいと思います。

モバイル空間統計の特徴の一つとしまして、日本全国の、基本的には携帯電話のつながるエリアをカバーした人口を出すことができます。

(資料1) たとえば、こちらは日本全国の時間ごとの人口の変化を表したものです。いくつか山が見えると思いますが、山の高さがそのエリアの人口の多さを表しています。東京とか大阪のあたりに高い山がある、つまり人口が集中していることがわかるかと思えます。時間を送っていきますと、朝の7時、8時と時間が進んでいくとどんどんと山が高くなる、つまり都心部に人が集中していきまして、夕方6時、7時になるとだんだんと人が少なくなっていく、みなさんが帰っていくのが分かります。このように、日本全国での人口

の移り変わりが、365日24時間、1時間ごとに分かれます。

先ほどの日本全国から、こんどは東京23区内のあたりをクローズアップしてみます。朝のラッシュアワーになると人がどんどん増えていきまして、東京、新宿、池袋などに人が集中してきます。晩になるとそれぞれだんだん減っていきます。減っているのですが、新宿とか池袋あたりには夜遅くになっても山が結構残っていて、まあこのあたり夜遅くまで人で賑っているんだな、ということが分かります。

また、これらの人口の移り変わりを、性別、年齢層別に見ることもできます。たとえばここでは秋葉原と原宿を比べてみましょう。まあ名前を聞くだけでどういう結果が出てくるかは想像がつかるところかもしれませんが、深夜は似たようなもので、どちらもあまり人がいませんが、朝になってきますとそれぞれのエリアに人が増えてきて、午前10時くらいにはくっきりと傾向に差が出てきます。やはり秋葉原は30代40代男性が多く、原宿はそれに対して20代30代女性が多めですね。このように、地域ごとの特色が、時間ごとの人口構成の移り変わりから浮き彫りになってきます。

次は、ちょっとわかりにくい画面で恐縮ですが、ある地域に人がいろいろなところから集まっているときに、それってどこに住んでいる方々が集まっているのか、ということを知ることができます。たとえば東京都千代田区、通勤などで多くの方が平日の昼にはいらっしやるかと思いますが、さて、一体どこらへんに住んでいる方が多く来られているのでしょうか。赤いところほどそこに住んでいる方が多く来ている、緑はまあそんなに多くない。そういうふうに見ていただけたらと思います。これを先ほどと同じように時間帯ごとに順々に見てみますと、深夜帯はまあほとんど千代田区に住んでいる方ばかりですが、やはり通勤時間帯になると周辺から人が集まってきて、東京23区を中心に、結構幅広いところから人が千代田区に集まっていっしやいますね。そして晩になると皆さん帰っていかれてまた千代田区にお住まいの方ばかりになります。

これまでご紹介したように、日本全国で、地域

ごとの人口の分布やその構成、そしてそれらの時間ごとの移りかわりのようなものを推計するのが、モバイル空間統計になります。というご紹介でした。

モバイル空間統計で何ができるか

さて、モバイル空間統計はどんなものかというのはいまご説明した通りなのですが、じゃあ何ができるんだ、という話に移りたいと思います。夢と目論見、なんて書いてしまいましたが、ちょっと私見を混じえまして、私の夢と目論見といたしましては、少し大げさな言い方をしてしまいますと、一言で言えばevidence based policy、日本語にすると証拠に基づく意思決定、という意味合いになりますでしょうか、これの実現に向けたお役立ちができればいいな、という思いがあります。

Evidence based policy、とか証拠に基づく意思決定、なんていうと難しそうな言葉ですが、平たく言いますと、勘と経験だけではなく、声の大きさでもなく、ましてや占いでもなく、事実と証拠やデータに基づいて政策判断や経営判断をすること、というあたりになりますでしょうか。以前、情報セキュリティ大学院大学の廣松先生という方に、何かの折にこの言葉を教えていただいた時にちょっとした感銘を受けまして、いろいろ調べてみるとこれからの日本に確かに大事なものだ、と、それで、よし、そういう世界を目指していこう、というか、こういうところにだんだん貢献していけるようになればいいな、という思いを持って研究を進めているところです。まあそういう夢と目論見です。

さて、夢とか思いのほうはともかくとして、実際問題としてどう役立つか、というのはいろいろと実証してみなければわからないわけですが、モバイル空間統計は、ちょうど1年くらい前になるんでしょうかね、2013年10月に第一段の実用化として一般向けに商用サービスを開始しました。一般のみなさまにサービスとしてご提供するからには、それが役に立つということを確認してみなくてははいけません、ということで、そのサービスを開始した2013年10月までに、2010年度秋からは東京大学の清家先生とご一緒に「まちづくりに関する共同研究」ですとか、工学院大学の村上先生とご一緒に「防災計画に関する共同研究」、

2011年に東京都と「帰宅困難者数調査」、柏市で「商業実態調査」、2012年度は埼玉県で「帰宅困難者数調査」、沖縄県で「観光客の実態調査」という感じで検証を進めてまいりました。今回、その中で最後の三つについてもう少し詳しくご説明しようかと思います。まず埼玉県の帰宅困難者調査、これはどういう調査をしたかといいますと、たとえば首都直下型地震ですとか、もう交通機関が全部止まってしまうような大きな災害が起きてしまったようなときに、埼玉県内でどのくらいの数の帰宅困難者の方々が出てくるんだろう、というのを推定したものになります。帰宅困難者って耳慣れない言葉かもしれませんが、まだおそらくみなさんの記憶にも新しい、東日本大震災の後にちよくちよく話題に上ったりもしましたが、覚えていらっしゃるでしょうか。電車が止まったりして家に帰れない人が都内に滞留して大変だという話がありましたよね。その帰宅困難になる人がどのくらいになるんだろうというのを推計したものが帰宅困難者数になります。ただ、帰宅困難者数の定義なんですけれども、人によって、「おれは30キロでも40キロでも歩いて帰るぞ」という人もいれば、「私なんか5キロも歩けないわ」みたいな方も、いろいろいらっしゃるんで、一律に決めるのってなかなか難しいんですよ。ただ、一つのモデルとしまして、10キロ以内に住んでいる人は皆さん帰れるでしょう、20キロ超えると帰れないんじゃないか。10キロから20キロの間は、1キロごとに帰宅困難な方の割合が10%ずつ増える。つまり、11キロ離れたところに住んでいる人は、そのうち10%の方が帰宅困難になって、19キロだと90%の方が帰宅困難になります。そういうふうな帰宅困難者の割合を推定するモデルがあります。内閣府なんかで使われているモデルなのですが、このモデルにモバイル空間統計の居住地別人口をあてはめて帰宅困難者数を推計してみました。その帰宅困難者数をどういう切り口で、どういう単位で推計するか、というのはいろいろ考えられるわけですが、ここでは、まず埼玉県にやってきて帰れなくなってしまった方々の数字として、大宮駅、浦和駅、川越駅、川口駅、南越谷駅という埼玉県内の主要5駅と言われるところの周辺にいらっしゃる方の帰宅困難者数を推計しました。これとその

逆、つまり埼玉県外、東京とか神奈川とかの他県に出かけて行って埼玉県に帰れなくなってしまった埼玉県民の方の人数、これを出かけた先の市区町村別に推計する。この二つで調査いたしました。結果としては、ここに書いてある数字の通りですけれども、まず埼玉県内の各駅で帰れなくなってしまふ人のほうで言うと、やはり大宮駅が一番多くて3.4万人、続いて浦和駅2.1万人、川越駅1万人という数字が出ています。つまり、たとえば大宮駅周辺では3万4000人の方が帰れなくなってしまふことを想定して、食料・飲料水などをどのくらい備蓄しなくちゃいけないかとか、その備蓄場所とか考えなくちゃいけないかもしれませんね、ということがわかるわけです。また、それらの方々の住んでいる都道府県別の内訳を見ますと、南のほうの方がやはり中心的で、多い順に東京、千葉、群馬、神奈川、栃木と続いています。こういうことがわかれば、初期対応が一段落して、さあ、みなさんおうちに帰りましょう、というときに、どういう方面にどのくらいの足が必要になるか、というのが予測できるんじゃないか、というわけです。こんどはその逆、埼玉県外の出先で帰宅困難になってしまう埼玉県民の方々の数ですけれども、これは千代田区にいらっしゃる方が一番多くて、約11万人の埼玉県民の方々が帰宅困難になるのではないかと。あと新宿区、中央区と続きます。なので、他県で帰宅困難になってしまう県民の方々への支援を考えるときには、こういうところの自治体とか、もしくはそのあたりから帰ってくる経路上にある自治体とか、というあたりと連携していく必要があるかもしれませんね、ということが想定できます。次に、沖縄県で「観光客実態調査」というのをやりました、というお話です。これは何を調べたかといいますと、沖縄県の外、県外から観光とかにいらっしゃる方々が、実際のところ観光かビジネスかはわからないので出張で来ている可能性もあつたりしますが、まあそれには目を瞑って、その県外から来た方々が、沖縄県内でどの市町村にどのくらい集まっていられるかというあたりについて、平日と休日との違いや、その時間変化を調べてみました。あともう一つ、県内で開催されるイベント、具体的に言いますとプロ野球のキャンプになりますが、その周辺にどれだけの観光客の方々が集

まっているか、ということ調べてみました。

まず市町村別のほうですが、ちょっと分析していて、おやっ、と思ったところとして、県南部の糸満市での推計結果について少しご紹介します。糸満市に来た観光客の方々の時間帯ごとの増減を平日と休日で見比べてみると、なぜだか平日のほうが休日よりあからさまに多い。これ、都内なんかのビジネス街ではまあ当然そうなるものなんですけど、なんで糸満でそんなことになるんだ、何か間違えたんちゃうか、とも思ったりもしましたが、もう少し詳しく見てみると、平日に増えているほとんどは10代の方ばかりですね、ということがわかりまして、ああ、なるほど、と。修学旅行だったんですね。ちなみに、この修学旅行の影響というのは、人数で言いますと、シーズンオフだから、というところも当然ありますが、少なくともその期間では、糸満市だけじゃなくて沖縄県全体で無視できないボリュームになっていたりもします。

プロ野球キャンプのほうは、まあやはりと言いますか、各チームの地元にお住まいの方々が多くいらっしゃるってですね、という結果が出ました。たとえば、北海道日本ハムが阪神と練習試合をした時で言いますと、その周辺の観光客のうち1/3くらいが北海道の方のようだ、とか。このあたりの数字を具体的に押さえていけば、今後の観光プロモーションなんかにも役立つ可能性がありますね、という話になります。

最後に、街づくりなんですけど、千葉県柏市で「商業実態調査」をいたしました。これは、柏市の中心市街地、JRの柏駅前あたりですかね、その柏駅周辺のエリアについて、人口が時間ごとにどう増えたり減ったりしているのかというのを曜日ごとに比較したりですとか、それを年齢、男女、住所別にわけて比較したりですとかをやりました。その結果、たとえば平日の時間ごとの変化を年齢・男女別に比較してみると、女性で30代、40代の方は昼に多くて、20代の方は、まあ昼もそれなりに多いのですが、昼よりも晩のほうが多い。こういうことがわかりますと、まあ物にもよるかもしれませんが、たとえば20代の方を対象にしたイベントとかの施策なんかは昼にやるんじゃなくて晩にやったほうが効果的かもしれないね、なんて推測ができたりします。という定性的

なお話だけですと、地元の方にとっては、なんだそんなこと知ってるよ、という話もあったりしますが、こういうのが定量的に客観的な数字としてわかりますよ、というのが大事なのかな、と思っています。

さて、ここまでモバイル空間統計で何ができるの、というところについていくつか事例をご紹介しました。次に、じゃあモバイル空間統計ってどうやってつくるの、というところにつきまして、このあたりあまり細かく話し出すと朝までかかりかねないので、なるべく簡単にご紹介したいと思っています。まずその前に、モバイル空間統計を作るにあたって、一番大事なことは何ですか、ということにつきまして。これは、お客さまのプライバシーが第一。つまり、プライバシーファーストというかたちでモバイル空間統計はつくっています。少なくともそのつもりです。別にこれはええ格好しいでこんなことを言っているというわけじゃなくて、実際に本当に大事なものですし、もしそれを軽んじるようなことがあってそれで本業に影響が出てしまったらどうすんの、というところもあるので、これは本気で考えているところです。それで、じゃあお客さまのプライバシーが第一はいいけど、そのためにモバイル空間統計って具体的に何をやっているかといいますと、携帯電話ネットワークの運用データというものに対して、運用データってなに、ということについては後で具体的にご説明しますが、この運用データに対して、非識別化処理、集計処理、秘匿処理という3段階の処理をしてモバイル空間統計というのをつくっています。順を追ってそれぞれ簡単に説明します。まず携帯電話ネットワークの運用データって何ですか、というところにつきまして。一言で言うと携帯電話ネットワークを運用するためのデータです、ということになってしまうんですが、そのあたりもうちょっと具体的に、というところから。携帯電話、皆さんお持ちだと思いますけど、電話がかかってくるたら当然着信しますよね。圏外じゃなければ、ですが。でも、なんで着信できるんでしょう。これ、不思議に思ったことはないでしょうか。どうやっているんだって、その仕組みを簡単に言いますと、たとえば携帯電話どうして電話をするようなときには、まず携帯電話は

最寄りの基地局、基地局というのはビルのとっぺんとかにアンテナが立っているようなあれですが、基地局と電波で繋がって、そこから携帯電話ネットワークをぐるっと回って、また相手の最寄りの基地局から電波でもう片方の携帯に、という感じで繋がります。いろいろはしりましたが、まあ基本はそんなところかな、という感じでなんとなく想像ください。

まあ繋がったあとはそんな感じで良いのですが、そもそも最初に電話をかけるとき、どうしましょう。最寄りの基地局まで電波を飛ばすのは良いのとして、相手の電話がどこにいるかわからないと、どこの基地局から電波を飛ばせば良いかわかりません。日本全国の基地局で電波を飛ばしくって探すというのもそれはそれで大変です。というわけで、それぞれの携帯電話は、最寄りの基地局との間で、ほくはここにいるよ、あなたの電波が届く圏内にいるんだよ、ということを定期的にやりとりしています、それをネットワーク側で受けて、この携帯電話はこの基地局のあたりにいるんですね、ということ进行管理しています。

そうしておく、わざわざ日本全国で探しまくなくても、電話がかかってきたときに、その基地局のあたりだけにピンポイントで、電話がかかってきたよ、ということを知らせる信号を送れば良いわけです。つまり、この仕組みを使えば、それぞれの基地局は、自分の電波が届く範囲に、どういう携帯電話がどのくらいいるのか、というのを知ることができるんですね。これがたとえば携帯電話ネットワークの運用データの一つになります。これと、もう一つの運用データとして、契約時にいただく加入者情報のデータ、これはそれぞれの携帯電話をどういう方がご契約されたのかというデータになりますが、これらの運用データ、いわゆる携帯電話が繋がる仕組みを支えるためのデータをもとにしてモバイル空間統計は作られます。

お気付きの方もいるかと思いますが、この運用データ、つまりどの基地局の圏内、つまり電波が届くエリアにどのくらいの携帯電話があるか、それはどういう方がお使いなのか、というデータがあれば、携帯がたくさんあれば人が多い、少なければ人が少ないというところがわかります。これ

が、言ってみればモバイル空間統計の基本的な考え方になります。

ただ、考え方としてはそうなのですが、このデータは、お話の通り携帯電話を繋ぐためのものですので、当然ながら電話番号ですとかも含んでいる、まごうことなき個人情報を含んだデータになります。なので、そこからモバイル空間計という人口統計を作るにあたって、先ほどの三段階の処理、非識別化、集計、秘匿の各処理をしてプライバシーを保護しながら作ります、というお話があることになるわけです。

まず非識別化処理ですが、これは簡単に言えば統計を作るのには不要な、個人ごとの細かい情報を捨てたり丸めたりします。モバイル空間統計をつくるに当たりまして、こんな携帯電話番号ですとか生年月日ですとか、そんなものは扱いが怖いだけで必要ありませんので、もともとの運用データをたとえば「女性、40代、埼玉県A市B町」、もしくは「男性、50代、千葉県C市D町」というかたちに、人口推計に必要なデータのみにしていきます。これが非識別化処理になります。

次に、統計を作るというところ、集計処理になります。さっきの非識別化処理で出てきた、基地局エリアごとの「男性、50代、千葉県なんとか町」というデータを単純に集計すれば、基地局エリアごとの携帯電話の数はわかりますが、基地局エリアごとにわかっても不便なので、たとえば500mメッシュとかに変換してやる。また、そもそも知りたいのは携帯電話の数じゃなくて人口なので、これは携帯電話利用率みたいなのをを使ってその数字を割り戻してやります。単に言えば、二人に一人がドコモの携帯電話を使っていて、台数が1000台であれば、2000人、ということになるわけです。これ、地域とか年代とかによってこの利用率というのはずいぶん違いますので、各地域ごと、年代ごとに割合を調整しながら各エリアの推計人口というのをつくっていきます。つまり、携帯電話の利用率をサンプル率として、層別にサンプル調査をするようなものになります。これが集計処理です。そうすると、たとえば50代男性というのは、このメッシュに75人、このメッシュに130人、このメッシュに30人、という感じで推計されます。という統計化をすれば、これでプライバシー的には、まあほぼほぼ大丈夫にはなる

んですが、実はまだ若干問題がありまして、たとえばここに6人というエリアがあります。こういう少数の方しかいらっしやらないエリアというのはどうしても個人を特定しやすくなるという懸念がありますので、それに対して秘匿処理という処理をしまして、たとえば6人という数字を消してしまって、この部分のエリアに関してはデータありませんよ、というふうにしてしまいます。こうして秘匿までできたものが、モバイル空間統計として世に出ていくことになります。このあたりの話につきましては、技術的なお話もさることながら、社会への影響ですとか、公的なコンプライアンスとかいうところもいろいろ多方面に考えなくちゃいけません。というあたりまで、われわれだけで考えます、というのもなかなか難しいところもありますので、2009年9月24日から12月10日まで、「モバイル空間統計による社会と産業の発展に関する研究会」という社外有識者の方々による研究会を開催いたしました。

こうした有識者研究会の中で、モバイル空間統計について、どういうやり方をしたら有用なのか、法的にどうか、技術的にどうか、社会的に安心して使ってもらえるためにはどういうふうにしていったらいいのかということをご議論いただきまして、それに従って、モバイル空間統計ガイドラインというものを作って公表しました。

このガイドラインは、モバイル空間統計を作ったり提供したりするにあたってはこういうふうにしていますよ、もしくはこういうふうにしなくてはいけませんよ、ということを決めています。たとえば、さっきの3段階の処理をやったものだけがモバイル空間統計と呼ばれるものですよ、というのもそのガイドラインに書いてあって、たとえばそのあたりの処理をすっ飛ばしてしまって、個人情報に相当するようなものが残ってしまっているものをモバイル空間統計という名前でしれっと出してしまふ、なんてことができないようにルールとして決めたりしているわけです。

このあたりのお話は、やりだすといくら時間があっても足りなくなってしまうんですが、日本全体でも、こういうことをルールとして整備していこうね、という流れにもなっています。ビッグデータの有効な活用による経済の活性化、という流れの中で、とても重要な位置付けを占めてくるもの

かと考えています。

モバイル空間統計はどういうふうにするのかということがわかりました。プライバシーについてこういうふうを考えています。というところまでお話ししました。次は最後のお話といたしまして、じゃあそれってどのくらい信じられるの、というところになります。この部分のお話ですが、独立行政法人の統計センターさん。ご存じの方いらっしゃるか。国勢調査とかの統計、公的統計と呼ばれたりしますが、その統計を作成する、製表というところを担っていらっしゃるんです。その統計センターさんと共同研究をさせていただきまして、そこでの内容につきまして、統計センターの方々と共著で *Statistical Journal of IAOS* という論文誌に採録された論文の一部を引用しながらご紹介させていただきます。

まず、どのくらい信じられるの、というのがなんで大事なの、というところにつきまして、自戒を込めて、というのも含みつつ少しお話しします。これまで、モバイル空間統計を使ってこういうのができます、ああいうのができます、そういうお話をさせていただきまして、いただいたんですけれども、実際のところ、数字を出すだけだったら、べつに誰でもできるお話なんですよ。極端な話、たとえば私、双子座生まれなんですけど、双子座のきょうのラッキーナンバーは3なので、じゃあきょうの水道橋の人口は3000人ということにしよう。というのでも、人口を推計していると言えば推計しているわけです。根拠はともかく、つまり、占いでいいのであれば、数字を出すだけなら誰でも作れてしまうんですね。でも、先ほど夢と目論見としての、というところで、*evidence based policy* ということについてお話をさせていただいたんですけれども、*evidence*、つまり証拠に基づいて、なんて言う限りは、ちゃんと事実に基づいたデータでないという意味がないということです。じゃなきゃ、双子座だとか占いと一緒だよ、というお話にもなりかねません。ただ、注意なくちゃいけないのは、事実に基づかないと、ということを杓子定規に考えて、統計に100パーセント正確性を求めるというのも、つまり100%正確じゃないと使い物にならない、なんていうのもまたナンセンスなお話になります。すべての統計には、と言ってしまってもいいと思

ますが、統計にはすべからず誤差がつきものです。たとえば推計してみたところ渋谷で10080人、新宿で10100人だったから、新宿の方が人が多い、なんていうのは、ほとんどの場合あまり意味がない議論で、一万人のうちの10人20人の違いは、まあほとんどの場合は、誤差、という言葉で片付けられるものですよね。そのくらいの違いだったらほとんど一緒、と結論づけるべきだったりすることが多かったです。

このように、統計の信頼性というのは少し難しいお話もはらみますが、それだけに、モバイル空間統計という人口統計について、統計を名乗る以上は、それがどの程度信頼できるか、ということについて継続的に評価と検証をしていく、ということが重要になってきます。

統計センターの方々とはいろんな観点で検証させていただいたのですが、その中の一つについて簡単に結果を紹介させていただきます。2010年の10月1日に行われました国勢調査の人口と、同じ日のモバイル空間統計の人口を比較してみました。この比較ですが、実は同じ「人口」と言っても指しているものが違う、ということに注意しなくちゃいけません。ということ、まず最初にお話しします。

国勢調査の人口というのは、常住人口とか常住地主義に基づく人口とか言われたりしますが、たとえば10月1日が国勢調査の実施日だといって、10月1日にどこにいたかということを出すのではなく、10月1日時点でどこに住んでいましたか、ということ、を推計するものになります。それに対して、モバイル空間統計というのは携帯電話の場所に基づいて作るもので、こちらのほうは、その時間にどこにいたかということ、をベースに推計します、という現在人口とか現在地主義に基づく人口、と呼ばれるものに近いものになります。ということで、そもそも定義が違うものを比較するということになってしまうんですけれども、まあ深夜の時間帯でしたら、ほとんどの人が家にいるので、どちらも自宅でカウントされるよね、国勢調査とモバイル空間統計でだいたい同じような結果が出るはずだね、ということで、国勢調査の結果とモバイル空間統計の深夜時間帯人口との比較をしました。その結果がこのグラフで、若干分かりにくいかもしれないんですけれども、横軸

がモバイル空間統計による推計人口で、縦軸が国勢調査による人口とどのくらい違うかの比率になります。点々がちらばっているのは、ある場所のメッシュを表します。つまり、もし仮に、国勢調査とモバイル空間統計とが完全に一致するのであれば、全ての点が、このx軸の線の上にびしっと乗ることになります。この線の上側に点があれば、その点が表すメッシュはモバイル空間統計が多めに推計された、下側であれば少なめに推計された。そういうふうにごらんいただければと思います。グラフを見ればわかりますように、メッシュ人口がたとえば100人とか1000人、2000人というあたりまでは、上と下のどちらにも結構点がちらばってますね。つまり人口が少ない地域ではモバイル空間統計はあまり精度よくないということを表しています。なぜかといいますと、これは理由がありまして、モバイル空間統計は基地局のエリアごとに人口推計したものをメッシュに直すという話をさせていただきましたが、東京都心とかの都市部ですと、基地局の間隔ってものすごく密に置かれているんですが、郊外とか山奥のほうに行きますと、何kmごとにしか置かれてないということがありますので、人口が少ないところで狭いエリアの人口を出そうとすると、どうしても推計の精度が悪くなります。人口が多いところ、人口密度が高いところになってくると、ちらばり方は少なくなっていきまして、ざっくりした言い方をしてしまいますと、1kmメッシュ人口、500mメッシュ人口とも、人口が2000人から3000人以上のエリアでは、80パーセントのメッシュがプラマイ2割くらいの範囲におさまる、というような結果が出ています。大ざっぱに言って、1kmメッシュごとの推計ですと、いわゆる人口集中エリア、DIDとか呼ばれていますが、そういうところだったらまあ信頼できそうだよ、というあたりが結論の一つになっています。

総じて言うとそんな感じになりますが、では地域ごとにもう少し細かく見るとどうでしょう、ということで、たとえば東京周辺を見てみましょう。この地図の上で緑色になっているところは、さっきの深夜で国勢調査とモバイル空間統計があまり変わらないところで、赤いところほどモバイル空間統計の方が多く、青いところが少ないというものになります。ぱっと見てもわかるように、

なんか東京の都心部が真っ赤ですよ。あんな、人口がたくさんいて基地局もたくさんあるはずなのに、これはほっておいたらあまりよろしくないお話なので、一体何が起きているんだろう、というところでこのあたりを少し詳しく見てみました。

こちらのグラフですが、それぞれの点、つまりメッシュの中で、特徴的なところについて地名をつけてみました。そうすると、月島、恵比寿、入船、というあたりはほとんど差がない。その逆に、このずいぶんと派手に上にはみだしているあたり、つまりモバイル空間統計のほうが国勢調査より多いところを見ると、六本木、銀座、新橋という地名が出てきます。なんだこれ、みんな夜の街じゃん、みたいな。

先ほどのお話で、国勢調査とモバイル空間統計では定義に違いがあります、というお話をしました。10月1日の、この時間帯に夜の街で飲み歩いている人は、まあ飲んでいるとは限らないかもしれませんが、モバイル空間統計では実際にそこで飲み歩いているので人口としてカウントしますが、国勢調査の場合は、別に住んでいるわけではないのでカウントしません。おうちにいるものとしてカウントします。その分、モバイル空間統計のほうが、こういうところで多く人口をカウントしているのは、そりゃ当たり前のお話ですよ、ということで、そのあたりの常住人口ですとか現在人口ですとかの定義の違いによる影響だろう、という説明がつけられることがわかりました。実際に酔っ払いの人をカウントしたわけではないので、あくまで仮説ということにはなるとは思いますが、まあそういうところなんじゃないかな、とそう思うように思っています。

このあたりの話につきましては、国勢調査とモバイル空間統計を混ぜて使うようなときに注意しなくてはいけませんよ、ということを示唆する一方で、逆に言えば国勢調査と組み合わせると、その差を見ることによって夜の街度みたいなものがわかるかも、というのは、それはそれで面白い結果だったりするところもあるかな、とも感じたところです。

というところで、お時間的にはちょうどいいので、ご紹介はこれで終わりになります。きょうお伝えしたことを順番に、モバイル空間統計って

何、何の役に立つの、どうやって作るの、どのくらい信じられるの、ということをお話しさせていただきました。モバイル空間統計に限らず、データとか統計ってまあそれだけ見ると数字の羅列ですけど、こうやっていろいろな見方をしていくことによって、楽しい、面白い、役に立つんだということ、少しでもお伝えできていたらうれしいかなと思います。

ただ、その一方で、データって手に入ったら何でもしていいの、とか、面白ければどんなものを出してもオッケーかという、そうではなくて、やっぱりプライバシーの保護についてきちんと考えるのは大事、これからますます大事になってきますし、占いじゃない以上は信頼性が大切なんだよねというところ、まあこのあたりややこしいところではありますが、大切なところですので、ぜひご理解のほどいただけましたらと思っております。これは私自身への自戒も込めまして、そのように考えています。

最後に、今後につきまして、これまで何度か *evidence based policy* という言葉でお話しをさせていただきましたけれども、いままで定量的にはあまりよくわからなかった地域ごと時間ごとの人口の変動というものを踏まえると、必要な場所、必要な人に対して、必要なものを必要なだけ提供する、ということがやりやすくなるのかな、ということを期待しています。まだまだそのあたり、何にどういうふうに使っていったらいいのか、というのはこれからいろいろと実例ですとか実績を重ねたり考えたりしなくては、ということもあると思いますが、モバイル空間統計を通じまして、そのようにしてより快適な安心・安全社会の実現、というものに貢献していけるといいな、と思っております。

学術的にも、ある時間にある場所にいた人数という、いろいろなことに対する説明変数となりそうなものについて、これまでは定量的に知るのが大変だったものを簡単に知ることができるようになりました。ということで、これを使った社会分析・最適化とかそういうものを通して、何か新しい分野の開拓など、学術的なところにつきましても、その発展への貢献に繋がれば嬉しく思っております。みなさまも、もしご興味がございましたらこれを機会にご検討のほどいただけましたら、

というお願いを最後にさせていただきまして、本日の私からのお話を終わらせていただきたいと思いを思います。

つたない説明ではございましたが、ご清聴ありがとうございました（拍手）。