

CENSIMENTI PERMANENTI POPOLAZIONE E ABITAZIONI

Эволюция непрерывной переписи населения Италии

. Опыт реализации первого цикла и планирование непрерывной переписи после 2021 года

А. Бернардини, А. Кьеппа, Н. Чибелла,
Г. Галло, Ф. Солари, Д. Циндато
ИСТАТ, отдел переписи населения

- ❖ Первый цикл Непрерывной переписи населения и жилого фонда (НПНЖ): слабые стороны и опыт проведения в 2018-2019 годах
- ❖ 2020 и 2021 - оценка полностью основана на данных регистров
- ❖ Второй цикл НПНЖ: переосмысление процесса оценки и модели оценки качества

Первый цикл Непрерывной переписи населения и жилого фонда

Вместо переписи, которая проводилась один раз в десять лет, в 2018 году Статистическая служба Италии (ИСТАТ) начала проводить Непрерывную переписи населения и жилого фонда (НПНЖ):

- Сочетание:
 - ❖ следующих регистров: регистр населения (РН) - сокращение на итальянском языке RVI, реестр адресов (РА) - сокращение на итальянском языке RSBL, тематические регистры (уровень образования и занятость)
 - ❖ Выборочные обследования: Территориальное обследование и Списочное обследование
 - ✓ чтобы собрать данные для переменных, не входящих в регистры,
 - ✓ для оценки ошибок РН вследствие неполного или избыточного охвата
- Позволяет ежегодно иметь подробную переписную статистику (тематические переписные таблицы) и заменяет подсчеты численности населения в период между переписями на основании административных данных

Первый цикл Непрерывной переписи населения и жилого фонда

Подсчет численности населения: оценка ошибки РН вследствие неполного/избыточного охвата - Модель двойной оценки системы

- РН (первая фиксация) + данные обследования (вторая фиксация), отличается от обследования после переписи (ОПП). Цель второй фиксации - измерить и скорректировать как неполноту, так и избыточность охвата.
- Территориальное обследование призвано оценить неполноту охвата РН, а списочное обследование - его избыточность
- Для коррекции неполноты охвата Списочного обследования в процесс оценки были включены административные «признаки жизни» (ПЖ), полученные из Объединенной базы данных постоянно проживающих (AIDA), то есть не ответившие с сильными признаками жизни считались прошедшими перепись
- Модели оценки малых районов использовались для снижения разброса непосредственных оценок для муниципалитетов, попавших в выборку, и для получения оценок для муниципалитетов, не входящих в выборку
- Численность населения, полученная путем применения к каждому человеку в РН поправочных коэффициентов в соответствии с его/ее статусом (более подробно см. Gallo, G. & Zindato, D. 2021)

Слабые стороны и ограничения

- Проблемы выборки:
 - ✓ размер выборки не позволяет оценить ошибки РН вследствие неполноты/избыточности охвата с разбивкой по муниципалитетам, полу, возрастным группам и гражданству
- Проблемы, связанные с полевыми исследованиями:
 - ✓ Сильные различия между участниками полевых работ
 - ✓ Проблемы неполноты охвата в отношении территориального обследования
- Проблемы, связанные с ожиданиями пользователей:
 - ✓ Необходимость общего, подотчетного процесса оценки / необходимость производства данных о населении, которые могли бы использовать местные власти (регистр с указанием веса, то есть не просто регистр «общего количества»)
- Из-за пандемии обследования, намеченные на 2020 год, были отменены: **Для получения численности населения в 2020 году ИСТАТ решил использовать административные «признаки жизни» (ПЖ) из AIDA (архив признаков жизни), чтобы оценить ошибки вследствие неполного/избыточного охвата в регистре населения**

Подсчет численности населения в 2020 году на основе

Административные **признаки жизни (ПЖ)** означают **деятельность индивидов**, по которой можно четко определить **длительный период времени** (например, год) и **место** (какой-либо район). Примерами прямых признаков жизни являются самозанятость или работа на компанию, служба в государственных органах, наличие договора аренды жилья, посещение школы или университета. С другой стороны, косвенные признаки жизни - это **ситуации, определяющие статус или состояние, не связанное с профессией** (то есть безусловный доход, получение пенсии или наличие детей либо других родственников, указанных в качестве иждивенцев в налоговых декларациях).

- Установление связей между РН и архивом ПЖ (AIDA) дает в результате «Расширенный регистр населения» (РРН), который по идее должен включать все целевое население. То есть предполагается, что регистр, в котором установлены связи, может характеризоваться избыточным охватом, но не недостаточным охватом
- Административные данные использовались для расчета **избыточного охвата в регистре, объединяющем РН и архив ПЖ (РРН)**. В частности, были установлены детерминирующие критерии для определения избыточного охвата при использовании модели ПЖ на индивидуальном уровне и (или) на уровне домохозяйства

2020 год - отправная точка для активного использования административных данных

- Серьезная инновация, обеспечивающая соответствие между численностью населения согласно переписи и индивидуальными записями (больше нет весов), но
 - ✓ ИСТАТ работает над тем, чтобы улучшить использование ПЖ в рамках нового цикла НПНЖ (после 2021 года). Поворотным моментом станет использование новых источников (то есть архивов, например архивов энергопотребления - данные умных счетчиков - обеспечивающие объективные элементы оценки с точки зрения реального места жительства - Albert, A. & Rajagopal, R., 2013)
 - ✓ Ошибка неправильного присвоения адреса (то есть индивидам присваивается адрес проживания в одном из районов в соответствии с РН или ПЖ, но обычно они проживают в другом районе) пока еще не оценивалась.

Оценка численности населения в 2021 году

- В 2021 году ИСТАТ имел возможность провести переписные обследования: **Необходимо интегрировать данные регистров, административные данные и данные обследований**
- Выбор статуса согласно ПЖ и соответствующее решение установить каждый статус как «постоянно проживает» или «не проживает постоянно» могут быть:
 - ✓ Основаны на информации, предоставленной экспертами
 - ✓ Предложены путем применения статистической модели
 - ✓ Их сочетанием
- В 2021 году был сделан выбор в пользу третьего варианта (сочетание информации от экспертов и результатов статистических моделей). Статистическая модель, которая использовалась для интеграции решений экспертов, - это модель латентных классов.

Второй раунд НПНЖ

- **Более активное использование административных источников** для переписи после 2021 года
- Решающая роль **обследований** для второго раунда НПНЖ, начинающегося в 2022 году
 - ✓ Собрать данные для незаменимых (или лишь частично заменимых) переменных (списочное обследование). Более того, можно добавить новые вопросы в соответствии с новым регламентом ЕС, который будет принят
 - ✓ Усовершенствовать критерии для экспертов/модели, применяемые для ПЖ
 - ✓ Предоставить показатели качества для оценки численности населения исключительно на основе регистров
- Необходимость переосмыслить статистическую базу для общих процессов оценки планирования НПНЖ

Непрерывная перепись населения после 2021 года (1)

- ❖ Предлагаемая архитектура второго раунда НПНЖ основана на определении **Расширенного регистра населения (РРН)** по итогам процесса интеграции Регистра населения и административных архивов, содержащих ПЖ (AIDA).
- ❖ Предполагается, что на РРН может повлиять только **избыточный охват** (то есть на него не влияет недостаточный охват).
- ❖ **Статус ПЖ** определяется для того, чтобы выявить подгруппы, члены которых предположительно имеют схожее поведение избыточного охвата.
- ❖ Определяется **индикаторная функция, основанная на статусе ПЖ** (Bernardini, Cibella and Solari, 2022), согласно которой **все индивиды, имеющие один статус, классифицируются либо как включенные, либо как исключенные из подсчета численности населения.**
- ❖ Такой выбор можно рассматривать как дихотомизацию дробного подсчета, предложенного в работе Zhang (2019), где вместо статуса избыточного охвата предсказываются вероятности $[0,1]$.

Непрерывная перепись населения после 2021 года (2)

- ❖ Чтобы **обеспечить объективные оценки на основе регистров (с использованием модели)**, вместо установления значения индикаторной функции на основе ПЖ для каждого статуса ПЖ на постоянное значение, равное 0 или 1, можно генерировать случайные значения 0 или 1, начиная с некоторых указанных предсказанных вероятностей θ_h . Более того, для получения стабильного списка лиц, для которых за какое-то время характерен избыточный охват, вместо генерирования независимых случайных чисел каждый год, каждому лицу могут присваиваться постоянные случайные числа.

Показатели качества и проверочное обследование (1)

- ❖ Показатели качества, связанные с оценкой численности населения, можно оценить путем сравнения оценок на основе регистров с оценками обследований (подробно см. Zhang, 2022b). **Это означает, что обследования проводятся для оценки показателей качества, а не для производства оценок. Такой принцип называется проверочным обследованием.**
- ❖ Согласно **принципу проверочного обследования** низкая эффективность обследования, например отмеченная в 2018 и 2019 годах, сможет повлиять только на оценку показателей качества, а не на оценку численности населения.
- ❖ Принцип **проверочного обследования** позволяет сократить затраты по сравнению со стандартными территориальными обследованиями, то есть размер выборки для проверочного обследования может быть меньше, чем для обследования, направленного на оценку численности населения.

Показатели качества и проверочное обследование (2)

- ❖ Помимо ошибки вследствие избыточного охвата, на РРН может оказать влияние неправильное присвоение адреса, то есть лица могут быть корректно включены в РРН, но присвоенный им адрес будет неверен.
- ❖ Следует использовать продвинутую версию РРН, **в котором записи относятся к парам индивид-адрес, а не только к индивидам** (см. Zhang, 2021 and Bernardini, Cibella and Solari, 2022).
- ❖ При таком подходе в оценки по итогам обследований не нужно добавлять компоненты смещения выборки и соответственно показатели качества, связанные с оценкой численности населения по данным регистров (см. Bernardini, Cibella and Solari, 2022).

Показатели качества и проверочное обследование (3)

- ❖ Наиболее адекватный подход к созданию выборки для проверочного обследования состоит в выборке индивидов из РРН или пар индивид-адрес из усовершенствованной версии РРН.
- ❖ Для некоторых наборов индивидов информация о их потенциальном месте постоянного жительства может быть неполной или недостоверной. Возможные решения:
 - Территориальная выборка (если предполагается, что они сосредоточены на отдельных частях территории)
 - Косвенная выборка (то есть путем выбора адресов их ПЖ, связанных с работой или учебой, для того, чтобы попробовать собрать полезную информацию о реальном месте их жительства).

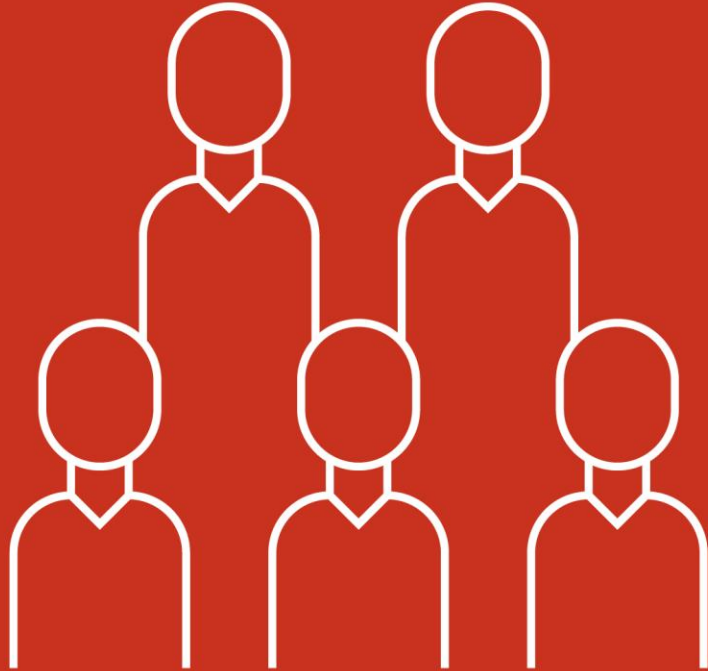
Показатели качества и проверочное обследование (4)

- ❖ Наконец, для оценки возможной неполноты охвата РРН или продвинутой версии РРН:
 - ✓ **малое территориальное обследование**
 - ✓ Внедрение в проверочное обследование **компонента обратной проверки записей** (то есть перекрестной проверки записей, относящихся к не ответившим).
- ❖ Хотя объединение разных обследований может создать организационные проблемы и сложности при проведении полевых работ, с чисто статистической точки зрения их можно связать путем выборки графиков (см. Zhang, 2022a)

- ❖ На пути к более широкому применению административных данных
 - ✓ Переосмысление в целом процессов оценки, используемых в рамках НПНЖ (данные обследований используются для определения качества оценки численности населения исключительно на основе данных регистров)
 - ✓ Необходимость усовершенствовать использование признаков жизни (усовершенствовать критерии классификации, получить новые источники)
 - ✓ Потребность в статистической базе для оценки качества
- ❖ Огромное значение будет иметь обработка данных переписи 2021 года, которая сейчас ведется, учитывая наличие и данных обследований, и административных данных. Сравнение различных моделей оценки, интеграция административных данных и данных обследований, оценка качества полевых работ - все это важные направления работы для того, чтобы лучше планировать предстоящие циклы НПНЖ.

- Albert, A. & Rajagopal, R. (2013), Smart Meter Drive Segmentation: What Your Consumption Says About You. IEEE Transactions on power systems: 4019-4030.
- Bernardini A., Cibella N. & Solari F. (2022). A Statistical Framework for Register Based Population Size Estimation, Technical Report, Istat Advisory Committee on Statistical Methods 2022 Springtime Meeting, Roma, 19-20 May 2022.
- Falorsi, S. (2017), The Italian experience on the Population and Housing Census: the Master Sample, UNECE Meeting, October 4-6 (2017), https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2017/Meeting-Geneva-Oct/Day2_1130_Italy_falorsi_presentation.ppt__1_.pdf
- Gallo, G. & Zindato, D. (2018). Annex H. Italy case study, in UNECE, Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses, Geneva, <https://unece.org/guidelines-use-registers-and-administrative-data-population-and-housing-censuses-0>.
- Gallo, G. & Zindato, D. (2021). Italy: The combined use of survey and register data for the Italian Permanent Population Census count in UNECE, Guidelines for Assessing the Quality of Administrative Sources for Use in Censuses (endorsed by the 69th plenary session of the Conference of European Statisticians), <https://unece.org/statistics/publications/CensusAdminQuality>.
- Istat (2020). Nota tecnica sulla produzione dei dati del Censimento Permanente: la stima della popolazione residente per sesso, età cittadinanza, grado di istruzione e condizione professionale per gli anni 2018 e 2019, <https://www.istat.it/it/files/2020/12/NOTA-TECNICA-CENSIPOP.pdf>
- Istat (2020). Nota tecnica sulla produzione dei dati del Censimento Permanente: la popolazione residente per genere, età, cittadinanza e grado di istruzione al 31.12.2020, https://www.istat.it/it/files//2021/12/NOTA-TECNICA-CENSIMENTO-POPOLAZIONE_2020.pdf
- Zhang, L.-C. (2019). On provision of UK neighbourhood population statistics beyond 2021, Report for ONS, <https://arxiv.org/pdf/2111.03100.pdf>.
- Zhang, L.-C. (2021). Discussion at ISI session 'Population statistics using administrative data instead of census', Virtual 63rd ISI World Statistics Congress, 11–16 July 2021.
- Zhang, L.-C. (2022a). Graph Sampling, Chapman and Hall/CRC.
- Zhang, L.-C. (2022b). Complementarities of survey and population registers, in N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri and J. L. Teugels (eds.), Wiley StatsRef: Statistics Reference.

CENSIMENTI PERMANENTI
POPOLAZIONE
E ABITAZIONI



**БЛАГОДАРЮ ЗА
ВНИМАНИЕ**