

CENSIMENTI PERMANENTI
POPOLAZIONE
E ABITAZIONI



**Использование
административных данных для
подсчета численности населения
при проведении непрерывной
переписи в Италии**

Бернардини, А. Чиппа, Н. Чибелла, Дж. Галло, Д. Зиндато
Истат, Отдел переписи населения

Непрерывная перепись населения и жилого фонда (НПНЖФ) Италии

- ❖ 2018: начало первого цикла (2018-2021 гг.) Непрерывной переписи населения и жилого фонда (НПНЖФ/РРНС) = комбинированная перепись на базе регистров + *специальные* выборочные обследования — позволяет ежегодно получать подробную статистику переписи, и это заменяет межпереписной подсчет численности населения, основанный на административных данных.
- ❖ В основе Непрерывной переписи населения и жилого фонда лежит Базовый регистр населения (БРН/RBI). Вместе со Статистическим базовым регистром адресов (СБРА/RSBL) и тематическими регистрами образования и занятости он представляет собой базу для производства данных переписи в рамках комбинированной архитектуры переписи.
- ❖ **Специальные обследования применяются для измерения погрешностей охвата Базового регистра населения и для сбора данных по переменным, которые невозможно заменить (или возможна только частичная замена) данными из этих регистров.**
 - ✓ **Территориальное обследование**, проводимое по выборке адресов, взятых из Статистического базового регистра адресов (СБРА), с целью подсчета и опроса (метод CAPI) каждого домохозяйства, проживающего по своему обычному адресу.
 - ✓ **Обследование по перечневой выборке**, проводимое с использованием метода смешанного режима (CAWI, CAPI, CATI) на выборке домохозяйств, взятых из Базового регистра населения (БРН) -
 - ✓ В обоих обследованиях использовалась одна и та же анкета (она включает все переменные гиперкубов, чтобы проверить качество и охват данных, уже имеющихся в регистрах).

Подсчет численности населения в рамках Непрерывной

Данные обследования используются для исправления данных, содержащихся в Базовом регистре населения (БРН), в рамках модели оценки двойной системы, нацеленной на оценку погрешности охвата этого регистра.

- Базовый регистр населения (БРН) = первый сбор данных
- второй сбор данных = ежегодные выборочные обследования + «административные «признаки проживания»», полученные с помощью Комплексной административной базы данных (КАБД/AIDA) [т.е. чтобы исправить неполный охват опроса, отсутствующие респонденты с явными «признаками жизни» в AIDA «считаются переписанными»].
- ✓ Зональное обследование, используемое для измерения ошибки неполного охвата Базового регистра населения (БРН) по каждому муниципалитету i и профилю лиц j (лица, обычно проживающие в муниципалитете, которые не включены в БРН). →

$$p_{ij,under} = \frac{\text{Newly Enumerated}_{ij}}{\text{Total Enumerated}_{ij}}$$

- ✓ Обследование по перечневой выборке + «административные «признаки проживания»» для измерения ошибки избыточного охвата Базового регистра населения (БРН) по каждому муниципалитету i и профилю лиц j (лица, включенные в регистр, которые больше не проживают в этом муниципалитете). →

$$= \frac{\text{Expected and not found}_{ij}}{\text{Expected and not found}_{ij} + \text{Expected and Enumerated}_{ij}}$$

- ❖ Соотношение между двумя компонентами определяет «первичный поправочный коэффициент» избыточного и неполного охвата.

$$corr_{ij} = \frac{1 - p_{ij,over-coverage}}{1 - p_{ij,under-coverage}}$$

- Расчет прямых оценок, откалиброванных на предмет избыточного и неполного охвата для каждого профиля j («итальянский» или «иностранный») для муниципалитетов i , вошедших в выборку.
- Расчет косвенных оценок: модели оценки малых зон используются для сокращения вариативности прямых оценок муниципалитетов, включенных в выборку, и для расчета оценок для муниципалитетов, не вошедших в выборку.
- ❖ В конце этого процесса к каждому человеку в Базовом регистре населения (БРН) применяется «вес» (в соответствии с его/ее профилем):
 - ✓ Если на БРН данного муниципалитета не влияют ни избыточный, ни неполный охват (или если две ошибки компенсируют друг друга), то вес, применяемый к этим жителям в регистре, будет равен 1.
 - ✓ Если неполный охват БРН > избыточного охвата, → то поправочный коэффициент, применяемый к каждому человеку в БРН, будет > 1 (и общая численность населения будет выше, чем в БРН). И наоборот, если неполный охват БРН < избыточного охвата, то поправочный коэффициент, применяемый к каждой записи в БРН, будет < 1 (и общая численность населения будет ниже, чем в БРН).

Основания для изменения и использования административных данных

- Обследования 2020 года пришлось отменить из-за пандемии. Тем не менее, согласно плану Непрерывной переписи, данные народонаселения за 2020 год должны быть опубликованы к концу 2021 года.
- Существует необходимость дальнейшего снижения затрат и минимизации нагрузки домохозяйств, что обусловлено участием в обследованиях.
- Слабые стороны, выявленные в комбинированной архитектуре переписи по результатам первых двух волн переписи (т.е. противоречивая интерпретация некоторых результатов, зарегистрированных в системе мониторинга обследования, согласно которой домохозяйства, не принявшие участие в обследовании, классифицируются в соответствии с их «статусом охвата»; размер выборки не позволяет рассчитать поправочные коэффициенты по муниципалитетам, полу, возрасту и гражданству; из-за методологии, применяемой для коррекции Базового регистра населения (БРН), результатом является регистр с весами, а не подушевой регистр).



- Более интенсивное использование административных источников при проведении переписи после 2021 года.
- Изучение осуществимости, чтобы попытаться переписать и оценить численность населения муниципалитета с разбивкой по полу, возрасту и гражданству исключительно с использованием административных источников.

Проблемы и перспективы широкого использования административных данных

Можно ли использовать только административные источники (независимые от Регистра населения) для расчета ошибок избыточного / неполного охвата Регистра населения и, таким образом, оценить численность населения в 2020 году (т.е. получить надежные оценки по определенным территориальным единицам)?

- **анализ на микроуровне** - связь между Комплексной административной базой данных (КАБД/AIDA) и Базовым регистром населения (БРН) с целью выявления лиц, имеющих право на неполный и избыточный охват Регистра населения — изучение моделей связей между данными БРН и «признаками проживания» из КАБД на микроуровне и связей между «признаками проживания» и данными обследования целевых групп населения, т.е. в отношении «переписанных впервые» (неполный охват согласно зональному обследованию) и в отношении «предполагаемых, но не обнаруженных» (избыточный охват согласно обследованию по перечневой выборке), для определения кластеризации муниципалитетов / выявления подгрупп населения, имеющих более высокий риск ошибок охвата, что также выявляется в ходе анализа. Эти модели связей могут представлять собой области оценки, наборы определенных правил или даже полезные случайные величины в прогнозной модели.
- **макроуровневый анализ** — расчеты поправочных коэффициентов на основании только административных данных, и изучение корреляций между этими поправочными коэффициентами и поправочными коэффициентами, полученными с помощью комбинированного подхода.



- ✓ определение детерминированных критериев (критериев, устанавливаемых согласно определенным правилам) для установления неполного / избыточного охвата с помощью признаков, выявляемых на индивидуальном уровне и/или на уровне домохозяйств.
- ✓ выявление прогностических моделей переменной «обычное место жительства»; при таком подходе «признаки проживания», получаемые из административных данных, оцениваются и выбираются в соответствии с типом и надежностью источника, с географическими моделями и моделями продолжительности, а также с возможными связями (т.н. ассоциациями) с другими индивидуальными характеристиками (отношения родства в домохозяйстве) с целью определения места обычного проживания каждого лица в рамках подхода, основанного на модели.

Лица в Комплексной административной базе данных (КАБД/AIDA) по типу «признака проживания»

Виды «признаков проживания»	Абсолютные значения	% значений
Устойчивые признаки занятости/учебы	32.121.583	51,6
Признаки зачисления в ВУЗ	1.438.897	2,3
Слабые признаки занятости/учебы	1.359.072	2,2
Признаки занятости/учебы эпизодические = не применимо	3.716.311	6,0
Договор аренды	468.658	0,8
Договор аренды + действующий вид на жительство	4.634	0,0
Пенсия	13.251.073	21,3
Пенсия + действующий вид на жительство	3.121	0,0
Прочие источники дохода	2.660.642	4,3
Действующий вид на жительство	641.083	1,0
Член семьи-иждивенец	4.240.523	6,8
Косвенные «признаки проживания» — несколько источников	293.169	0,5
Признаки итальянцев, проживающих за границей (Anagrafe consolare)	112.571	0,2
Никаких «признаков проживания»	1.968.137	3,2
Общая численность населения в Комплексной административной базе данных (КАБД/AIDA)	62.279.474	100,0

Попытка остаться в рамках моделей непрерывности

Январь 2018 - декабрь 2019																								Вид присутствия по признаку работы/учебы за 24 месяца		
G	F	M	A	M	G	L	A	S	O	N	D	G	F	M	A	M	G	L	A	S	O	N	D			
																									1	Непрерывно, за 24 месяца
																									2	Непрерывно, не менее 12 месяцев
																									3	Непрерывно, не менее 12 месяцев
																									4	Непрерывно, не менее 12 месяцев
																									5	Периодически, не менее 12 месяцев
																									6	Сезонно
																									7	Периодически, < 12 месяцев
																									8	Эпизодически
																									9	Непрерывно, до декабря 2018 г.
																									10	Периодически, после декабря 2018 г.

Явные «признаки проживания»

Слабые «признаки проживания»

«Признаки проживания» непригодны

Предварительные критерии домохозяйств для оценки избыточного / неполного национального охвата

Тип домохозяйства	Критерии для лиц в Базовом регистре населения (БРН), имеющих право на избыточный охват	Результат
Домохозяйство, состоящее из одного человека	Домохозяйство, состоящее из одного человека, без признаков проживания на протяжении последних 5 лет, не проживающее в приграничных муниципалитетах.	<i>национальный избыточный охват</i>
Член семьи, состоящей из нескольких человек	Все члены семьи без «признаков проживания»	<i>национальный избыточный охват</i>
	Лицо младше 14 лет, посещающее школу в том же муниципалитете Базового регистра населения (БРН)	<i>отсутствие национального избыточно охвата</i>
	Если хотя бы один член семьи владеет жилищем (или имеет договор аренды) в том же муниципалитете БРН.	<i>отсутствие национального избыточно охвата</i>
	Муж / жена (даже если не является иждивенцем) работающего / учащегося / получателя пенсии	<i>отсутствие национального избыточно охвата</i>

Лица, не включенные в Базовый регистр населения (БРН), имеющие право на неполный охват	Если все критерии подтверждены
Признаки занятости/учебы не менее 12 месяцев за соответствующий период	<i>национальный неполный охват</i>
Собственник жилой недвижимости (или арендатор)	
Муниципалитет, не находящийся на приграничной территории	

Выявление подгрупп населения, имеющих право на избыточный / неполный охват

Предварительная классификация после применения детерминированных критериев (следующих определенным правилам)		Абсолютные значения	% значений
0	Лица, присутствующие как в Базовом регистре населения (БРН), так и в Комплексной административной базе данных (КАБД/AIDA) в одном и том же муниципалитете ---> подтверждаются в БРН.	58.309.572	93,6
1	Неполный охват на уровне страны ---> подлежит добавлению к полученной численности	416.807	0,7
2	Избыточный охват на национальном уровне (на уровне страны) ---> подлежит исключению из подсчетов	368.185	0,6
3	Серая зона (потенциальный избыточный / неполный охват, но без четких сигналов ---> необходимо изучение этого вопроса	281.679	0,5
4	Избыточный / неполный охват на местном уровне (другая провинция, взятая в Базовом регистре населения (БРН)) ---> подлежит подтверждению в другом муниципалитете.	12.626	0,0
5	Лица, не включенные в БРН, но не имеющие неполного охвата (слабые «признаки проживания») ---> не должны включаться в показатель численности населения.	895.208	1,4
6	Лица без «признаков проживания» в ИКР ---> требуется дальнейшее расследование	1.599.952	2,6
7	В настоящее время не поддается классификации - ---> требуется дальнейшее изучение	395.445	0,6
Общая численность населения в Комплексной административной базе данных (КАБД/AIDA)		62.279.474	100,0

Два возможных сценария расчета ошибок избыточного / неполного охвата Регистра населения и оценки численности населения в 2020 г.

А - Использование административных данных (детерминированный и/или модельный подход)

- ✓ **определение критериев прогнозирования «обычного места жительства»**
- ✓ **индивидуация «серых зон»** (подгруппы населения или муниципалитеты с особыми характеристиками) → — что делать с подгруппами населения, у которых «признаки проживания» НЕ являются достаточно предсказуемыми?

В — Применение поправочного коэффициента (вес, применяемый к каждой записи БРН) = применить поправочный коэффициент 2018/2019 года избыточного и неполного охвата к БРН 2020 г.

СПИСОК ЛИТЕРАТУРЫ

Bernardini A., Cibella N., Gallo G., & Al. (2019), “Empirical evidence for population counting: the combined use of administrative sources and survey data”. Paper presented at ESS Workshop on the use of administrative data and social statistics, Valencia, June 4th 2019

Chieppa A., Gallo G., Tomeo V. & Al. (2018), “Knowledge discovery for inferring the usually resident population from administrative registers”. In Mathematical Population Studies. International Journal of Mathematical Demography, Published online: 27 Jul 2018. <http://www.tandfonline.com/loi/gmps20>

Chieppa A., Ferrara R., Gallo G., Tomeo V. (2017), «Towards The Register-Based Statistical System: A New Valuable Source For Population Studies”. Presentazione al Convegno Statistics and Data Science: New Challenges, New Generations. University of Florence, June 2017

Kim J.K., Rao J.N.K. (2012), Combining data from two independent surveys: a model assisted approach, Biometrika, Vol. 99(1), 85-100.

Nirel, R. , Glickman, H. (2009), Ch. 21 - Sample Surveys and Censuses. In: Rao, C.R. (ed.) Handbook of Statistic, Elsevier.

Pfeffermann D. (2015), METHODOLOGICAL ISSUES AND CHALLENGES IN THE PRODUCTION OF OFFICIAL STATISTICS. Journal of Survey Statistics and Methodology 3, 425–483.



GRAZIE

PER

L'ATTENZIONE

Бернардини, А. Чиппа, Н. Чибелла, Дж. Галло, Д.
Зиндато

anberna@istat.it, chieppa@istat.it, cibella@istat.it,
gegallo@istat.it, zindato@istat.it