

ДЕРЖАВНА СЛУЖБА СТАТИСТИКИ УКРАЇНИ

ЗАТВЕРДЖЕНО
Наказ Державної служби
статистики України
20.08.2012 р. № 356

МЕТОДИКА
РОЗРАХУНКУ ХАРАКТЕРИСТИК НАДІЙНОСТІ ОЦІНЮВАННЯ
ПОКАЗНИКІВ ЗА РЕЗУЛЬТАТАМИ ВИБІРКОВОГО ОБСТЕЖЕННЯ
УМОВ ЖИТТЯ ДОМОГОСПОДАРСТВ

Київ-2012

Методика визначає основний зміст характеристик надійності оцінювання показників за результатами державного вибіркового обстеження умов життя домогосподарств і принципи їх розрахунку з використанням реплікаційних методів.

У методиці розглянуто інформаційне забезпечення й алгоритми розрахунків, наведено приклади визначення характеристик надійності оцінювання окремих показників на основі методу збалансованих реплікацій, реалізованого в пакеті статистичних програм „WesVarPC”.

Методика призначена для використання працівниками органів державної статистики, наукових організацій, навчальних закладів, іншими користувачами статистичної інформації, яка розробляється за підсумками обстеження умов життя домогосподарств.

Методику підготували завідувач відділу соціально-демографічної статистики Інституту демографії та соціальних досліджень НАН України ім.М.В.Птухи В.Г.Саріогло та працівники департаменту обстежень домогосподарств Державної служби статистики України:

Осипова І.І. – директор департаменту;

Пліско К.Ю. – заступник директора департаменту – начальник відділу організаційно-методологічного забезпечення обстеження умов життя домогосподарств.

Рокитько Л.Л. – начальник відділу підготовки та поширення даних обстеження умов життя домогосподарств.

Методику схвалено Комісією з питань удосконалення методології та звітної документації Державної служби статистики України (протокол від 31.07.2012р. №8).

ЗМІСТ

	стор.
Передмова.....	4
1. Визначення основних термінів	5
2. Загальні положення.....	7
3. Основні характеристики надійності оцінювання показників.....	8
4. Використання наближених методів оцінки похибки вибірки.....	11
5. Опис вхідної та вихідної інформації.....	13
Перелік використаних джерел.....	16
Додаток 1. Приклад розрахунку.....	17
Додаток 2. Інструктивні матеріали щодо розрахунку характеристик надійності оцінювання показників за результатами обстеження умов життя домогосподарств із використанням програмного пакета „WesVarPC”	19
Додаток 3. Вихідні дані та таблиці з результатами розрахунків	36
Додаток 4. Значення змінної strt_ua (страти для України по регіонах та по типах місцевості).....	38
Додаток 5. Значення змінної strt_rg (страти, побудовані у межах кожного регіону)	40

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ПТОВ	– первинна територіальна одиниця вибірки;
ВТОВ	– вторинна територіальна одиниця вибірки;
ОУЖД	– обстеження умов життя домогосподарств;
BRR	– метод збалансованих повторних реплікацій (Balanced Repeated Replications);
SPSS	– статистичний пакет для соціальних наук (Statistical Package for Social Science);
n	– обсяг вибірки;
$\hat{\theta}$	– оцінка показника за результатами вибіркового обстеження;
$\hat{\theta}_L, \hat{\theta}_R$	– верхня та нижня межі довірчого інтервалу;
$CV(\hat{\theta})$	– коефіцієнт варіації вибірових оцінок показника $\hat{\theta}$;
$deff(\hat{\theta})$	– дизайн-ефект;
$LSE(\hat{\theta})$	– гранична похибка;
$SE(\hat{\theta})$	– стандартна похибка;
$V(\hat{\theta})$	– дисперсія вибірових оцінок $\hat{\theta}$.

Передмова

При проведенні вибірових обстежень вирішальне значення для правильної інтерпретації їх результатів, контролю оптимальності дизайну обстеження та вибірки, статистичної ефективності окремих етапів обробки даних має інформація щодо надійності оцінювання показників для генеральної сукупності. Оцінка показника, надійність якої є недостатньою, не може бути використана для аналізу процесів та явищ.

Метою розробки цієї Методики є створення методичного забезпечення визначення характеристик надійності оцінювання показників, які розраховуються за результатами державного вибіркового обстеження умов життя домогосподарств (далі – ОУЖД).

Можливість визначення надійності оцінювання показників генеральної сукупності за результатами певного вибіркового обстеження є найважливішим фактором, який зумовлює ефективність вибіркового методу спостережень при зборі статистичної інформації. Ця можливість забезпечується імовірнісним характером (репрезентативністю) вибірки.

Для ймовірнісних вибірок зі складним дизайном, якою є, зокрема, вибірка ОУЖД, побудова аналітичних залежностей для розрахунку характеристик надійності є складною статистичною проблемою, оскільки такі залежності визначаються як особливостями дизайну вибірок, так і типом показників, що оцінюються. При цьому використання спрощених залежностей, які не враховують реального дизайну вибірок, неприпустимо, оскільки це може призвести до суттєвої недооцінки похибки вибірки. За таких умов найбільш ефективними на цей час є універсальні наближені розрахункові методи оцінки характеристик надійності.

З метою визначення характеристик надійності, у цій методиці використовується так званий метод збалансованих реплікацій, який дозволяє за результатами вибіркового обстеження розрахувати величину дисперсії вибірових оцінок для будь-якого дизайну вибірки та для будь-якого типу показника [5]. Цей метод дає можливість ефективно використовувати первинні дані обстеження, статистичні ваги обстежених одиниць і дизайн вибірки. Метод реалізований у статистичному програмному пакеті WesVarPC [5], який у цілому гармонізований із програмним пакетом SPSS, що є базовим програмним продуктом для обробки статистичних даних у системі Держстату України.

Методика складається з передмови, п'яти розділів, додатків і переліку використаних джерел. Документ містить визначення основних понять і термінів, опис інформаційного забезпечення та порядок розрахунків характеристик надійності оцінювання показників.

Необхідність розробки цієї методики зумовлена, насамперед, визначальною роллю інформації щодо рівня надійності результатів ОУЖД для адекватної їх інтерпретації та використання, підвищенням зацікавленості користувачів в отриманні й урахуванні характеристик надійності при використанні результатів

ОУЖД, а також рекомендаціями міжнародних статистичних служб щодо стандартів оцінювання показників і змісту публікацій за результатами зазначених обстежень.

1. Визначення основних термінів

<i>Вибіркове обстеження населення (домогосподарств)</i>	– науково обґрунтований вид несучільного спостереження, при якому обстежуються не всі одиниці сукупності, а лише певним чином відібрана їх частина, а одиницею спостереження є особа чи домогосподарство.
<i>Вибірковий метод спостереження</i>	– науково обґрунтована система правил відбору одиниць і способів характеристики сукупності досліджуваних одиниць, що їх вивчають. Вибірковий метод спостереження дає змогу розповсюдити висновки, отримані на основі вивчення частини сукупності (вибірки), на всю сукупність (генеральну).
<i>Гранична похибка</i>	– гранично припустима розбіжність між оцінкою показника за результатами вибіркового обстеження та дійсним значенням показника (невідомим) для генеральної сукупності, визначена на основі стандартної похибки вибірки та встановленого рівня довірчої ймовірності.
<i>Дизайн-ефект($deff(\hat{\theta})$)</i>	– міра відносної статистичної ефективності певного плану вибірки порівняно з простою випадковою вибіркою (відношення дисперсії вибірових оцінок показника для плану вибірки, реалізованого в обстеженні, до дисперсії оцінки цього показника, отриманої за умови використання процедури простого випадкового відбору).
<i>Дизайн (план) вибірки</i>	– організаційно-логічна модель структури вибіркової сукупності та принципи її формування. Якістю плану вибірки значною мірою визначається репрезентативність вибіркової сукупності, отриманої за визначеними у плані правилами.
<i>Дисперсія вибірових оцінок</i>	– середній квадрат відхилень оцінок показника за всіма можливими вибірками цього дизайну (реплікаціями) від середнього значення оцінки за всіма вибірками.
<i>Зміщення оцінки показника</i>	– різниця між очікуваною величиною оцінки та дійсним значенням показника для всієї сукупності. Особливість зміщення полягає в тому, що воно є постійною складовою частини похибки, яка при численній сукупності не зменшується із збільшенням

обсягу вибірки, тоді як випадкова похибка вибірки із збільшенням обсягу вибірки в середньому зменшується. Визначають зміщення як різницю середнього значення оцінок у разі повторних спостережень і оцінюваної величини.

Кластер

– сукупність географічно локалізованих одиниць генеральної сукупності (домогосподарств, сімей, осіб), яка може розглядатися як окрема територіальна одиниця відбору при формуванні територіальної вибірки для статистичних обстежень. Прикладами кластерів є населені пункти, міські та сільські ради, квартали, територіальні одиниці, сформовані для проведення переписів населення або виборчих кампаній тощо.

*Коефіцієнт варіації
(відносна стандартна похибка)*

– відносна стандартна похибка вибірки. Визначається як відношення величини стандартної похибки вибірки до оцінки показника у відсотках.

Надійність

– ступінь близькості розрахованих оцінок показників до дійсних (точних) значень показників.

Оцінка показника

– значення показника, яке отримано за неповними даними щодо сукупності, яка досліджується. Може бути отримана шляхом розповсюдження даних вибіркової сукупності на генеральну, застосування моделей, визначення значень експертами тощо.

Похибка вибірки

– частина розбіжності між істинним значенням статистичного показника та його значенням, розрахованим за вибіркою, яка зумовлена застосуванням саме несуцільного характеру обстеження.

Стандартна похибка

– середньоквадратичне відхилення вибірових оцінок показника за всіма можливими вибірками цього дизайну від дійсного значення показника для генеральної сукупності.

Стратифікація

– поділ генеральної сукупності або її частини на групи, однорідні за певною ознакою.

При формуванні переліку термінів та визначенні їх змісту використано матеріали, опубліковані у [1, 2].

2. Загальні положення

Для сучасних вибірових обстежень домогосподарств характерними є такі риси, як складний дизайн вибірки (стратифікація генеральної сукупності, застосування механізмів багатоступеневого відбору, у тому числі й відбору з

ймовірностями, що залежать від певних характеристик одиниць відбору та ін.); необхідність вимірювання значної кількості показників та проведення аналізу даних для різних типів одиниць спостереження (особи, домогосподарства), для різних підгруп домогосподарств або населення; використання оцінок різних типів показників (оцінювання сумарних, середніх значень, часток) та ін. Це значною мірою ускладнює процедури аналізу надійності оцінювання показників, оскільки зумовлює необхідність використання досить складних методів розрахунків похибок оцінювання ознак.

Із статистичної точки зору, оцінка певного показника за результатами вибіркового обстеження є надійною, якщо вибіркові оцінки, побудовані на основі всіх можливих вибірок однакового дизайну та обсягу, сконцентровані біля дійсного значення показника [3]. Висока концентрація, тобто надійність, бажана тому, що в окремому процесі побудови вибірки та розрахунку значення показника існує лише певна ймовірність наблизитися до дійсного його значення.

Близькість оцінки до реального значення показника, а саме її надійність, має дві характеристики: вибіркові оцінки мають малу дисперсію та середнє значення оцінок показника, близьке до дійсного значення, тобто зміщення відсутнє або має малий розмір (див. рис. 1).

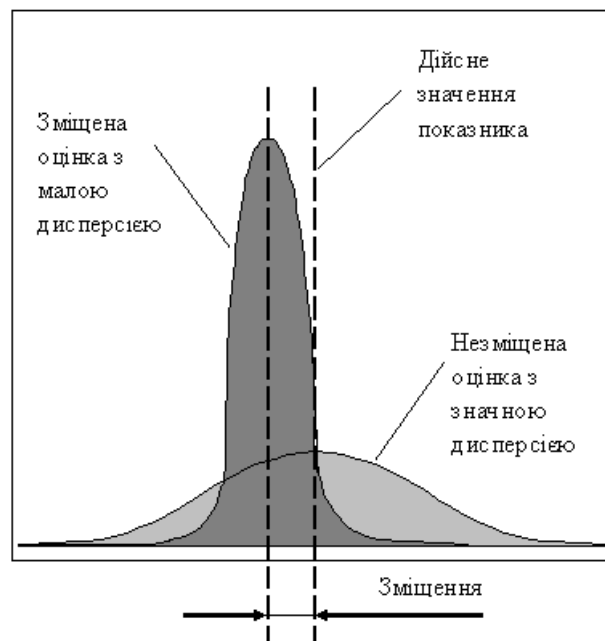


Рис. 1. Порівняння надійності оцінок при великій дисперсії без зміщення та при малій дисперсії зі зміщенням

Реальна надійність оцінок будь-якого показника, оціненого за даними вибіркового обстеження, може бути визначена лише у виключному випадку, оскільки на практиці процедура відбору реалізується лише один раз і дійсні значення показників невідомі. При обробці даних вибірових обстежень у більшості випадків надійність визначається лише наближено.

3. Основні характеристики надійності оцінювання показників

Одною з головних характеристик надійності оцінки показника, розрахованого на підставі даних вибіркового обстеження, є похибка вибірки, яка показує, наскільки оцінка може відрізнятись від дійсного значення при існуючих дизайні вибірки, обсязі вибіркової сукупності, дисперсії значень ознаки по одиницях вибіркової сукупності. За сучасними міжнародними рекомендаціями похибку вибірки, яка може визначатись як абсолютна (стандартна, гранична) або відносна, необхідно публікувати для всіх основних показників, що вимірюються або розраховуються за результатами обстеження.

Для оцінки надійності показників ОУЖД використовуються такі характеристики:

- дисперсія вибірових оцінок;
- стандартна похибка вибірки;
- гранична похибка вибірки;
- відносна стандартна похибка вибірки, або коефіцієнт варіації.

Дисперсія вибірових оцінок $V(\hat{\theta})$ показника θ визначається за формулою:

$$V(\hat{\theta}) = \frac{1}{H} \sum_{i=1}^H (\hat{\theta}_i - \bar{\hat{\theta}})^2, \quad (1)$$

де $\hat{\theta}_i$ – оцінка показника по вибірці i , $i = 1, 2, \dots, H$; $\bar{\hat{\theta}}$ – середнє значення вибірових оцінок показника, $\bar{\hat{\theta}} = \frac{1}{H} \sum_{i=1}^H \hat{\theta}_i$; H – кількість усіх можливих вибірок цього дизайну й обсягу.

Похибка вибірки показує, наскільки оцінка показника, отримана за результатами вибіркового обстеження, може відрізнятись від значення показника, яке було би отримано при ідеальному (без помилок) суцільному обстеженні одиниць генеральної сукупності [3].

При аналізі похибки вибірки розраховують стандартну похибку вибірки, граничну похибку вибірки та відносну стандартну похибку вибірки, яку частіше називають коефіцієнтом варіації CV .

Стандартна похибка вибірки SE при оцінюванні показника θ розраховується як корінь квадратний від дисперсії вибірових оцінок $V(\hat{\theta})$ за формулою:

$$SE(\hat{\theta}) = \sqrt{V(\hat{\theta})}. \quad (2)$$

На практиці розрахувати стандартну похибку з використанням формули (2) неможливо, оскільки у дослідника є лише одна вибірка з усіх можливих. Тому для визначення SE користуються спеціальними підходами.

Для вибірок із простим дизайном найчастіше використовують зв'язок між величинами $V(\hat{\theta})$ і дисперсією значень ознаки, за якою оцінюється показник θ , по одиницях вибірки σ^2 . Наприклад, для безповторної простої випадкової вибірки з малою часткою відбору f (звичайно при $f < 0,05$) цей зв'язок має вид:

$$V(\hat{\theta}) = \frac{\sigma^2}{n}, \quad (3)$$

де n - обсяг вибірки, $\hat{\theta}$ - оцінка показника θ за результатами вибіркового обстеження;

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}, \quad (4)$$

де y - ознака (характеристика), за якою оцінюється показник $\hat{\theta}$; y_i - значення ознаки для i -го домогосподарства (наприклад, якщо θ - грошовий дохід, то y_i - грошовий дохід певного домогосподарства з вибірки); \bar{y} - середнє значення ознаки по всіх одиницях вибірки.

Для вибірок зі складним дизайном, якою є, зокрема, вибірка для ОУЖД, величина $V(\hat{\theta})$ може бути розрахована на основі формули, аналогічній (3), але з урахуванням додаткового параметра $deff(\hat{\theta})$:

$$V(\hat{\theta}) = deff(\hat{\theta}) \cdot \frac{\sigma^2}{n}, \quad (5)$$

де $deff(\hat{\theta})$ - параметр, який відображає вплив дизайну вибірки на величину дисперсії вибірових оцінок показника (дизайн-ефект). При побудові вибірки за допомогою методу простого випадкового відбору $deff(\hat{\theta}) = 1$.

Визначення величини $deff(\hat{\theta})$ для конкретного показника та конкретного дизайну вибірки потребує застосування спеціальних методів.

Гранична похибка вибірки $\Delta(\hat{\theta})$ є максимально можливою похибкою вибірки для цієї довірчої ймовірності ρ і визначається за формулою:

$$\Delta(\hat{\theta}) = t \cdot SE(\hat{\theta}), \quad (6)$$

де t - довірче число (квантиль нормального розподілу або квантиль розподілу ймовірностей), визначає співвідношення граничної та стандартної похибки при цій ймовірності ρ (ρ - ймовірність того, що похибка вибірки для оцінки показника не перевищить величину граничної похибки $\Delta(\hat{\theta})$). Типові значення t наведено в табл.1.

Таблиця 1. Взаємовідповідність між величинами довірчої ймовірності ρ і довірчого числа t

p	0,50	0,80	0,90	0,95	0,99
t	0,67	1,28	1,64	1,96	2,58

Гранична похибка вибірки використовується для побудови довірчих границь оцінок (границь довірчих інтервалів). Визначення нижньої $\hat{\theta}_L$ та верхньої $\hat{\theta}_R$ довірчих границь інтервальної оцінки показника $\hat{\theta}$ здійснюється за формулами:

$$\hat{\theta}_L = \hat{\theta} - \Delta(\hat{\theta}); \quad \hat{\theta}_R = \hat{\theta} + \Delta(\hat{\theta}). \quad (7)$$

Відносна стандартна похибка $RSE(\hat{\theta})$ (коефіцієнт варіації $CV(\hat{\theta})$) розраховується за формулою:

$$RSE(\hat{\theta}) = CV(\hat{\theta}) = \frac{SE(\hat{\theta})}{\hat{\theta}} \cdot 100\%. \quad (8)$$

Використання відносних стандартних похибок (коефіцієнтів варіації) при аналізі даних обстежень особливо доцільно, коли ознаки, що розглядаються, є додатними величинами. У цьому випадку для порівняння ступеня варіювання (і, як наслідок, точності оцінювання) різних ознак доцільно використовувати саме відносні величини, оскільки вони не залежать від рівня ознаки. Крім того, відносні стандартні похибки використовуються для порівняння якості оцінок різних за змістом ознак.

Величина відносної стандартної похибки часто використовується як показник придатності даних для аналізу. Так, якщо $RSE < 5\%$, то оцінка вважається надійною, якщо $5\% < RSE \leq 10\%$ - оцінка є придатною для кількісного аналізу, але її надійність недостатньо висока, якщо $10\% < RSE \leq 25\%$ - оцінка придатна лише для якісного аналізу і її слід використовувати обережно (іноді публікують дані, для яких RSE досягає 30% і навіть 40%).

Для визначення дисперсій, стандартних похибок і величини дизайн-ефектів для вибірок зі складним дизайном найчастіше використовуються спеціалізовані розрахункові методи.

4. Використання наближених методів оцінки похибки вибірки

Як витікає з методології розрахунку характеристик надійності оцінювання показників за результатами вибірових обстежень (див. п. 3), такі характеристики надійності, як $SE(\hat{\theta})$, $\Delta(\hat{\theta})$, $CV(\hat{\theta})$, розраховуються на основі величини дисперсії вибірових оцінок $V(\hat{\theta})$. Таким чином, визначення цієї величини є ключовою проблемою при їх розрахунках. Слід зазначити, що наведені у п. 3 залежності відповідають випадку, коли при оцінюванні показників відсутнє зміщення.

Один із найбільш поширених методів розрахунку похибки вибірки базується на вимірюванні розбіжностей оцінок показників, побудованих на

спеціальних незалежних підвбірках із повної вибірки (реплікаціях). Реплікації повторюють дизайн основної вибірки і відрізняються від неї лише обсягом.

Основним реплікаційним методом, який використовується для розрахунків дисперсії оцінок показників по вибірках зі складним дизайном, є метод збалансованих повторних реплікацій - *BRR*.

Метод *BRR* можна використовувати, коли структура вибірки представлена набором страт, у кожній з яких міститься лише два кластери. Саме така структура вибірки була реалізована в ОУЖД.

Метод *BRR* передбачає, що повна вибірка складається з H страт, у кожній з котрих міститься $m_h = 2$ кластери (наприклад, дві первинні територіальні одиниці вибірки - ПТОВ). Кожна реплікація – напіввибірка, що будується шляхом випадкового відбору по одній ПТОВ із кожної страти. Таким чином, може бути сформовано $T = 2^H$ реплікацій.

Розрахунок дисперсії вибірових оцінок показника здійснюють за формулою:

$$V(\hat{\theta}) = \frac{1}{T} \cdot \sum (\hat{\theta}_t - \hat{\theta})^2, \quad (9)$$

де $\hat{\theta}_t$ - оцінка показника θ по реплікації t ; $\hat{\theta}$ - оцінка показника θ по всій вибірці; t – поточний номер реплікації, $t = 1, 2, \dots, T$; T - кількість реплікацій.

На практиці необхідна кількість реплікацій T визначається з використанням методу балансування. Це досягається спеціальною процедурою формування напіввбірок на основі ортогональних матриць Адамара. Такі матриці мають порядок $4z$, де z - будь-яке натуральне число [6]. Елементи матриць Адамара приймають значення 1 або -1.

При використанні матриць Адамара для формування реплікацій «1» означає, що з відповідної страти включається перша ПТОВ, «-1» - друга. Приклад матриці Адамара для $z = 2$ наведено в табл. 2. Як видно із представлених даних, якщо всі ПТОВ можуть бути згруповані попарно у вісім страт, то при формуванні першої реплікації ($t = 1$) з кожної страти ($h = 1, 2, \dots, 8$) включається перша ПТОВ, а друга ПТОВ виключається. Статистичні ваги одиниць вибірки з кожної першої ПТОВ у кожній страті подвоюються для забезпечення наближеної загальної суми ваг по кожній страті. Таки скориговані статистичні ваги називаються реплікаційними вагами. Реплікаційні ваги для першої реплікації використовуються для розрахунку оцінки показника $\hat{\theta}_1$ (див. формулу (9)). Доцільно зауважити, що такий підхід до розрахунку реплікаційних ваг передбачає необхідність об'єднання у страти схожих ПТОВ як за змістом явища, що вивчається, так і за розміром.

При формуванні другої реплікації ($t = 2$) зі страт 1, 3, 5, 7 ($h = 1, 3, 5, 7$) включається перша ПТОВ, а зі страт 2, 4, 6, 8 - друга ПТОВ. Для ПТОВ, що включені у реплікацію, статистичні ваги одиниць вибірки подвоюються і формуються реплікаційні ваги для другої реплікації. Далі на основі цих ваг розраховується оцінка показника $\hat{\theta}_2$.

Після визначення оцінок $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_8$ за формулою (9) визначається дисперсія вибіркових оцінок $V(\hat{\theta})$.

При цьому, якщо вибірка складається з меншого числа страт, ніж число $4z$, (наприклад, $T = 9$ при $z = 3$ і, відповідно, порядку повної матриці Адамара, що дорівнює 12), то для побудови реплікаційних ваг слід обрати будь-які T рядків повної матриці Адамара (при $T = 9$ - будь-які 9 рядків з 12 рядків повної матриці).

Таблиця 2. Зв'язок елементів матриці Адамара при $z = 2$ зі стратами для вибірки, яка складається із 8 страт (умовний приклад)

Напіввибірка	Страта, h							
t	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	1	-1	1	-1	1	-1	1	-1
3	1	1	-1	-1	1	1	-1	-1
4	1	-1	-1	1	1	-1	-1	1
5	1	1	1	1	-1	-1	-1	-1
6	1	1	1	-1	1	1	-1	1
7	1	1	-1	-1	-1	-1	1	1
8	1	-1	-1	1	-1	1	1	-1

Доцільно зазначити, що, як видно з розглянутого алгоритму розрахунків, застосування реплікаційного методу дає можливість визначити наближені характеристики надійності оцінювання будь-якого показника, що розраховується за результатами вибіркового обстеження. Це можуть бути як сумарні або відносні значення, так і характеристики нерівності, наприклад індекс Джині, або показники, розраховані на основі моделі регресії.

Реплікаційний метод *BRR* реалізований у стандартному пакеті програм *WesVarPC*.

5. Опис вхідної та вихідної інформації.

Технологія розрахунку похибки вибірки для результатів ОУЖД орієнтована на використання стандартних пакетів *SPSS*, *WesVarPC*, *EXCEL*.

5.1. Вхідна інформація.

Вхідною інформацією для розрахунку характеристик надійності оцінювання показників за результатами ОУЖД є файл первинних даних обстеження у форматі програми *SPSS*, що містить ознаки (по домогосподарствах або по членах домогосподарств), по яких будуються оцінки та додаткові змінні, необхідні для розрахунків.

Для розрахунку характеристик надійності основних показників (на прикладі доходів і витрат) за результатами ОУЖД у файлі аналізу даних обстеження, що містить інформацію по домогосподарствах або по членах домогосподарств, формуються необхідні для розрахунків змінні:

Змінні характеристики (приклад для показників основних груп доходів і витрат):

- грошові доходи (cashinc);
- грошові витрати (cashexp);
- сукупні витрати (totalexpr);
- загальні доходи (totalinc);
- сукупні ресурси (resurs).

Класифікаційні змінні, для яких розраховуються основні показники доходів і витрат:

- тип населеного пункту (tp_ns_p)
- код територій (регіонів) за КОАТУУ (cod_obl);
- код економічного району (rn).

Спеціальні змінні, що використовуються програмою при розрахунках:

- порядковий номер запису (id) ;
- статистична вага (w_q);
- страти для України (strt_ua);
- страти для регіонів (strt_rg);
- страти для економічних районів (strt_rn);
- код страти первинної одиниці вибірки (psu_ua).

У діалогових вікнах програми WesVarPC задаються:

- набір необхідних для розрахунків змінних;
- метод розрахунку;
- змінні страт;
- вихідні таблиці з характеристиками надійності.

5.2. Вихідна інформація.

Результати роботи програми зберігаються у виводі WesVarPC (*.lst), який завантажується через середовище Excel для подальшого аналізу та формування таблиць у заданих форматах:

1). У верхньому лівому куті листа міститься інформація про програму, дату розрахунку, місцезнаходження вхідної та вихідної інформації та загальні параметри для розрахунку.

2). Нижче послідовно виводяться таблиці з результатами розрахунків, які містять таку інформацію:

- класифікаційні змінні, для яких побудовані таблиці;
- інформація про характер і тип оцінки (графи «STATISTIC» та «EST_TYPE»);
- оцінка показника (графа «ESTIMATE»);
- стандартна похибка оцінки (графа «STDERROR»);
- коефіцієнт варіації (відносна стандартна похибка) оцінки (графа «CV»);
- розмір групи, для якої розраховується оцінка показника (графа «N»);

–оцінка величини дизайн-ефекту (графа «DEFF»).

Структуру вихідного файлу наведено в додатку 3.

5.3. Порядок проведення розрахунків основних характеристик надійності оцінювання показників.

Основні етапи роботи щодо розрахунку характеристик надійності основних показників доходів і витрат за даними ОУЖД:

1). У програмному середовищі SPSS готується робочий файл аналізу даних по домогосподарствах, який містить цільові показники, ключові змінні (код домогосподарства, загальнодержавний код ПТОВ), класифікаційні змінні (тип населеного пункту) та статистичні ваги. Цей масив не повинен містити ніяких зайвих змінних. Для забезпечення розрахунку необхідно перевірити наявність пропущених значень у будь-якій змінній з масиву, які призводять до відкидання всіх спостережень із пропусками при аналізі похибки вибірки програмою WesVarPC. Тому здійснюється перевірка наявності пропущених значень, змінні з пропусками аналізуються окремо.

2). У робочому файлі аналізу даних формуються спеціальні змінні, необхідні для розрахунків. Це здійснюється автоматизовано шляхом використання функції SPSS злиття робочого файла аналізу даних зі спеціальним масивом, який містить спеціальні змінні. Ключовою змінною для злиття масивів є загальнодержавний номер ПТОВ, ВТОВ. Крім того, для файлів, створених у версіях „SPSS”, які перевищують восьму (наприклад, 10-12 версії), слід здійснити конвертацію робочого файла у формат *.dbf. Це пов’язано з тим, що програма "WesVarPC" не працює із зазначеними версіями.

3). Робочий файл аналізу даних імпортується у формат програми WesVarPC (створюється файл із розширенням *.var) та подальші розрахунки проводяться в середовищі цієї програми.

4). Розраховуються реплікаційні ваги домогосподарств.

5). Формуються таблиці для аналізу.

6). Здійснюється розрахунок оцінок показників і характеристик їх надійності.

7). Файл із результатами розрахунків відкривається в середовищі EXCEL, де проводиться формування остаточних таблиць для публікацій результатів ОУЖД.

Порядок побудови спеціальних змінних.

Змінні *strt_ua*, *strt_rg*, *strt_rn*, *psu_ua*, *psu_rg*, *psu_rn* будуються окремо для кожного дизайну вибірки та для конкретного методу розрахунку похибки вибірки. Якщо дизайн вибірки не змінюється (не змінюється кількість та нумерація ПТОВ по стратах) і не змінюється метод розрахунку похибки вибірки, то ці змінні також використовуються без змін.

Побудова змінної strt_ua

Змінна *strt_ua* – страти для України по регіонах та по типах місцевості, приймає значення, які наведено в додатку 4.

Побудова змінної strt_rg

У межах кожного регіону будується своя система страт із нумерацією від 1 до *S*. Для цього:

а) усі ВТОВ по міських поселеннях і ПТОВ по сільській місцевості в межах регіону розташовуються в порядку зростання номерів;

б) кожні дві ВТОВ по міських поселеннях або ПТОВ по сільській місцевості, починаючи з першої у ранжованому ряді, зараховуються до окремої страти. За необхідністю декілька територіальних одиниць можуть об'єднуватися в один кластер. Страти, побудовані для міських поселень, не можуть включати територіальні одиниці сільської місцевості, а страти сільської місцевості – територіальні одиниці міських поселень;

в) страти нумеруються в порядку зростання номерів – $1, 2, 3, \dots, S$.

Якщо кількість ПТОВ непарна, то окрему ПТОВ умовно можна розділити на дві. Значення змінної *strt_rg* для всіх регіонів України та номери ПТОВ, які входять до кожної страти, наведені у додатку 5.

Побудова змінної psu_ua

У межах кожної страти для України (*strt_ua*) ПТОВ розташовуються у порядку зростання номерів. Для першої по порядку ПТОВ змінна *psu_ua* приймає значення 1, для другої – 2.

Побудова змінної psu_rg

У межах кожної страти для регіонів (*strt_rg*) ПТОВ розташовуються у порядку зростання номерів. Для першої по порядку ПТОВ змінна *psu_rg* приймає значення 1, для другої – 2.

Інструктивні матеріали розрахунку характеристик надійності оцінювання показників за допомогою програми „WesVarPC” наведено в додатку 2.

Перелік використаних джерел

1. Вибіркове спостереження: Термінологічний словник. – К.: НТК статистичних досліджень, 2004.
2. Методологічні основи формування вибірових сукупностей для проведення органами державної статистики України базових державних вибірових обстежень населення (домогосподарств) / Затверджені наказом Держкомстату України від 2 серпня 2005 р. № 223.
3. Kish L. Survey sampling. – Wiley Classics Library Edition Published 1995. – 643 p.
4. Кокрен У. Методы выборочного исследования. – М.: Статистика, 1976. – 440 с.
5. A User's Guide to WesVarPC // J.M. Brick, P. Broene, P. James, J. Severynse. – Rockville: Westat, Inc., 1997. – 214 p.
6. Лидл Р., Пильц Г. Прикладная абстрактная алгебра: Учебн. пособ. – Екатеринбург: Изд-во Уральского ун-та, 1996. – 744с.
7. Методика розрахунку характеристик надійності оцінювання показників за результатами вибірового обстеження населення (домогосподарств) з питань економічної активності / Затверджена наказом Держкомстату України від 29 грудня 2006 р. № 639.

**Приклад розрахунку середньомісячного значення показника
"Загальні доходи" за результатами обстеження умов
життя домогосподарств у 2010 році.**

Розглянемо приклад розрахунку характеристик надійності оцінювання середньомісячного значення показника "Загальні доходи" за результатами обстеження умов життя домогосподарств у 2010 році по Рівненській області. У 2010 році по Рівненській області в обстеженні взяли участь 377,2 тис. домогосподарств. Пряма оцінка середньомісячних загальних доходів пересічного домогосподарства (надалі – загальні доходи) у генеральній сукупності (побудована з урахуванням системи статистичних ваг осіб) дорівнює 3342,54 грн.

Для визначення характеристик надійності прямої оцінки на основі реплікаційного методу *BRR* побудовано одинадцять страт, у кожній з яких виділено по два кластери.

За принципами побудови збалансованих реплікацій сформовано матрицю Адамара, яка має такий вигляд (матриця побудована в програмному пакеті „WesVarPC”):

Таблиця 3

Система реплікацій для вибірки, яка складається з 11 страт

Реплікація, <i>t</i>	Страта, <i>h</i>										
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>
<i>1</i>	1	1	1	1	1	1	1	1	1	1	1
<i>2</i>	-1	1	-1	1	1	1	-1	-1	-1	1	-1
<i>3</i>	-1	-1	1	-1	1	1	1	-1	-1	-1	1
<i>4</i>	1	-1	-1	1	-1	1	1	1	-1	-1	-1
<i>5</i>	-1	1	-1	-1	1	-1	1	1	1	-1	-1
<i>6</i>	-1	-1	1	-1	-1	1	-1	1	1	1	-1
<i>7</i>	-1	-1	-1	1	-1	-1	1	-1	1	1	1
<i>8</i>	1	-1	-1	-1	1	-1	-1	1	-1	1	1
<i>9</i>	1	1	-1	-1	-1	1	-1	-1	1	-1	1
<i>10</i>	1	1	1	-1	-1	-1	1	-1	-1	1	-1
<i>11</i>	-1	1	1	1	-1	-1	-1	1	-1	-1	1
<i>12</i>	1	-1	1	1	1	-1	-1	-1	1	-1	-1

Оцінки показника "Загальні доходи" та квадрата різниці прямої та реплікаційних оцінок для кожної реплікації наведені у табл. 4 (тут і далі для спрощення сприйняття результати наводяться після округлення до двох знаків після коми).

**Оцінки середньомісячного значення показника "Загальні доходи" за
реплікаціями, побудованими для визначення
характеристик надійності цього показника по Рівненській області**

Номер реплікації, t	$\hat{\theta}_t$, за місяць, грн.	$(\hat{\theta}_t - \hat{\theta})^2$
1	3322,73	392,00
2	3307,88	1201,06
3	3436,93	8911,06
4	3405,31	3940,72
5	3422,82	6445,66
6	3395,85	2843,04
7	3149,86	37123,28
8	3509,81	27980,59
9	3359,95	303,39
10	3174,03	28392,95
11	3423,91	6622,66
12	3200,48	20178,32
		$\sum_{t=1}^{12} (\hat{\theta}_t - \hat{\theta})^2 = 144334,73$

Пряма оцінка загальних доходів $\hat{\theta} = 3342,54$ грн.

Дисперсія прямої оцінки визначається за формулою (9):

$$V(\hat{\theta})_T = \frac{1}{T} \cdot \sum_{t=1}^{12} (\hat{\theta}_t - \hat{\theta})^2 = \frac{1}{12} \cdot 144334,73 = 12027,89.$$

Стандартна похибка визначається за формулою (2):

$$SE(\hat{\theta}) = \sqrt{V(\hat{\theta})_T} = \sqrt{12027,89} = 109,67 \text{ грн.}$$

Коефіцієнт варіації визначається за формулою (4):

$$CV = \frac{SE(\hat{\theta})}{\hat{\theta}} \cdot 100\% = \frac{109,67}{3342,54} \cdot 100\% = 3,28\%.$$

Гранична похибка оцінки загальних доходів визначається за формулою (6):

$$\Delta(\hat{\theta}) = t \cdot SE(\hat{\theta}) = 1,96 \cdot 109,67 = 214,96 \text{ грн.}$$

Інструктивні матеріали щодо розрахунку характеристик надійності оцінювання показників за результатами вибіркового обстеження умов життя домогосподарств із використанням програмного пакета „WesVarPC”

Перед використанням програми "WesVarPC" для розрахунків характеристик надійності показників необхідно підготувати вхідний файл даних у форматі програмного пакета „SPSS” (наприклад, з іменем "for_wesvar_analisis_ua.sav"). Цей файл формується на основі квартальних (річних) масивів даних ОУЖД. Після завершення формування вхідного файла даних його слід зберегти у форматі „dbf” за допомогою програмного засобу „dBase”.

1. Запуск програми

Для інсталяції програми „WesVarPC” необхідно знайти місцезнаходження файла „wesvar32.exe” і запустити його. Після інсталяції програми її ярлик буде виведено на робочий стіл „Windows” (див. рис. 1), потім програму слід запустити безпосередньо з робочого столу.



Рис. 1. Ярлик для запуску програми „WesVarPC”

Після запуску програми на екрані з’являється вікно програми (див. рис. 2).

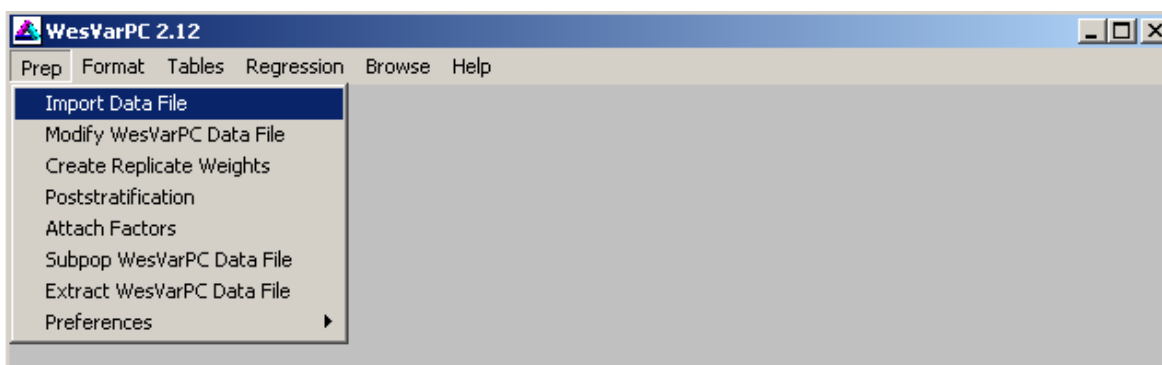


Рис. 2. Вікно програми „WesVarPC” та опції меню “Prep”

2. Конвертація вхідного файла із формату програми „dBase” у формат програмного пакета „WesVarPC”

У пункті меню **“Prep”** вибрати опцію **“Import Data File”** для відкриття вхідного файла формату dBase. На екрані з’являється діалогове вікно **“Input Data Files”** (див. рис. 3).

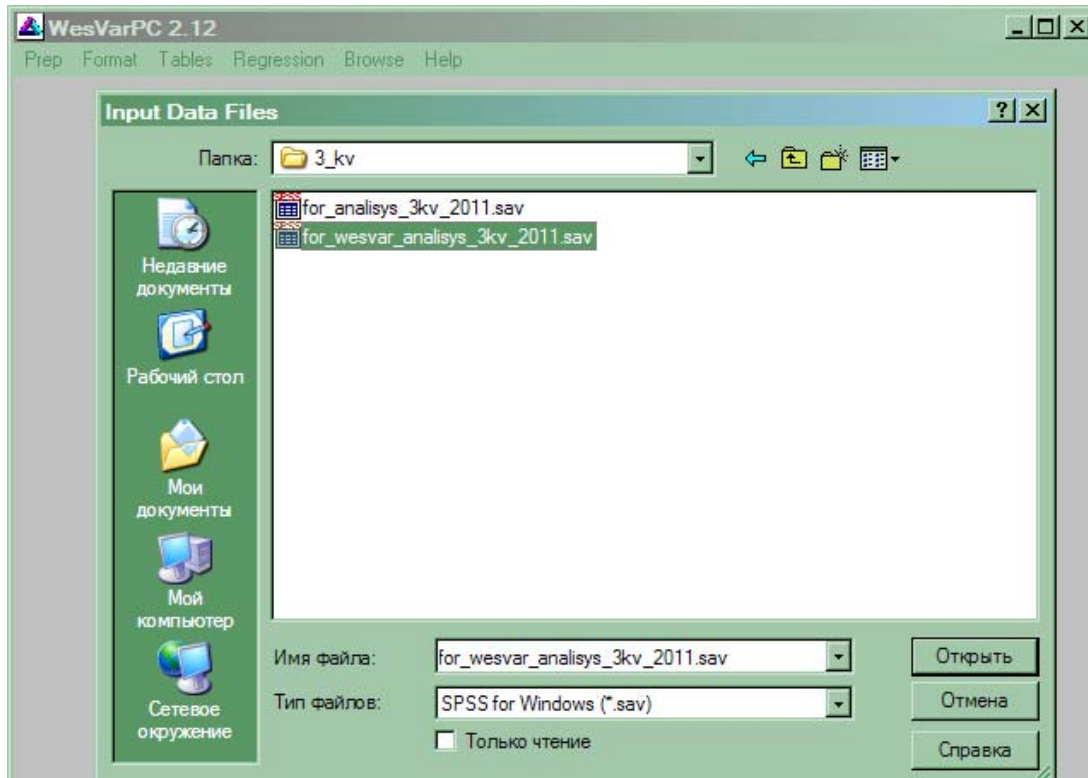


Рис. 3. Діалогове вікно “Input Data Files”

Вибираємо папку, у якій збережено вхідний файл, у полі **“Тип файлов”** слід вибрати тип файла **“dBase(.dbf)”**, у полі **“Имя файла”** вказати ім’я вхідного файла (наприклад, „for_wesvar_analysys_3kv_2011.sav”). Після натискання кнопки **“Открыть”** з’являється вікно **“Import dBase Data File”** з іменем файла, що імпортується (див. рис. 4).

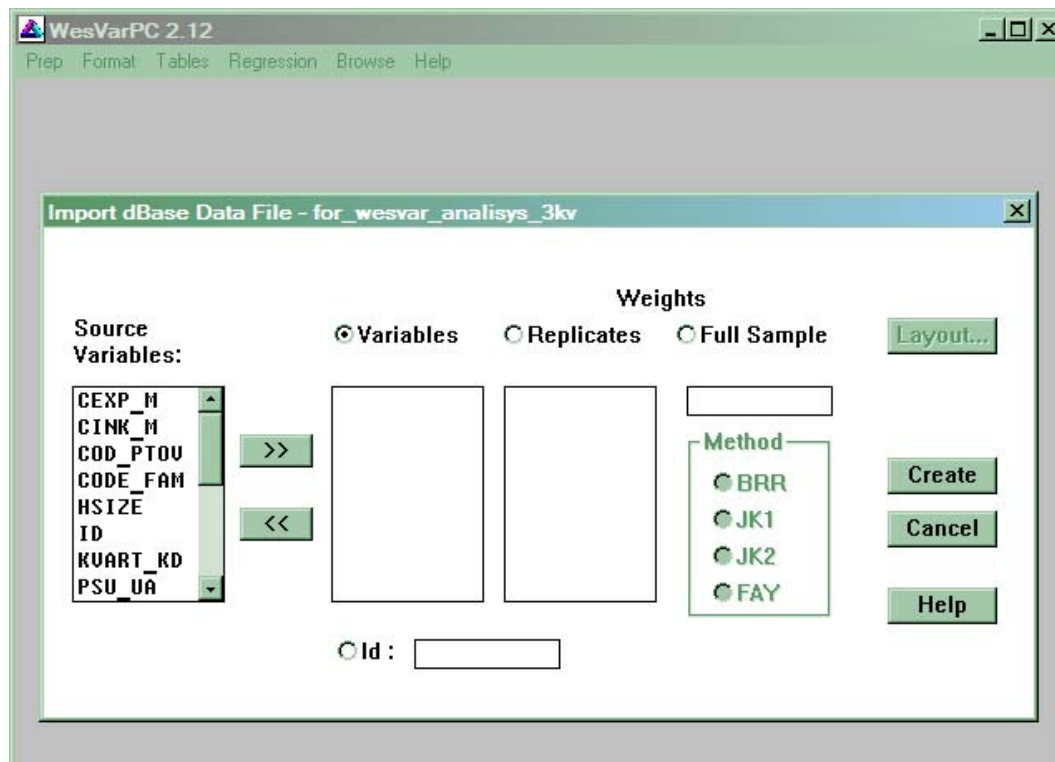


Рис. 4. Початковий вигляд вікна “Import dBase Data File”

У полі “Source Variables” великими літерами вказані імена всіх змінних вхідного файлу.

3. Формування робочого файлу у програмному пакеті „WesVarPC”

Для формування робочого файлу необхідно виконати таке:

1) за допомогою курсора „миші” слід активізувати індикатор “Id” (у віконці індикатора з’явиться чорна крапка). Потім із списку змінних у полі “Source Variables” вибрати змінну „IDD”. Після натискання кнопки >> або подвійного натискання на змінну „IDD” у списку змінних зазначена змінна буде перенесена в поле індикатора “Id”;

2) аналогічно в полі індикатора “Full Sample” задається змінна, яка відповідає статистичним вагам – „W_q”;

3) у полі індикатора “Variables” із списку змінних вхідного файлу переносяться всі змінні, необхідні для подальших розрахунків:

CEXP_M – змінна - середньомісячні грошові витрати домогосподарства;

TEXP_M – змінна - середньомісячні сукупні витрати домогосподарства;

CINK_M – змінна - середньомісячні грошові доходи домогосподарства;

TINK_M – змінна - середньомісячні загальні доходи домогосподарства;

TRES_M – змінна - середньомісячні сукупні ресурси домогосподарства;

Tr_ns_p – змінна-ознака тип місцевості;

PSU_UA – код територіальних одиниць вибірки для України (шари для України);

RG – код регіону за КОАТУУ;

RN – код економічного району;

STRS_RG – страти для регіонів;
STRT-UA – страти для України;
STRT_RN – страти для економічних районів.

Для того, щоб видалити непотрібну змінну із будь-якого поля, слід натиснути кнопку **<<** або двічі – на певну змінну.

Кінцевий вигляд вікна **“Import dBase Data File”** наведено на рис. 5.

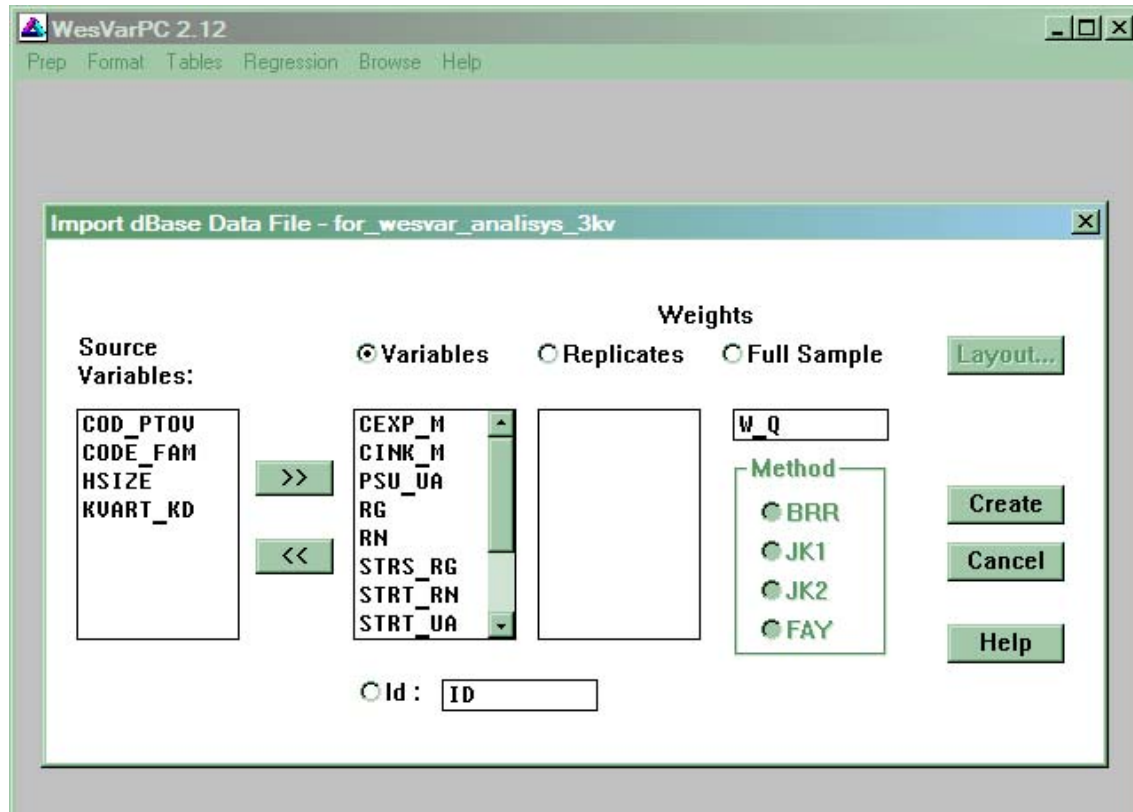


Рис. 5. Кінцевий вигляд вікна “Import dBase Data File”

Для збереження заданих параметрів робочого файла потрібно натиснути кнопку **“Create”**. З’являється вікно **“Notice”** (див. рис. 6) з інформацією про відсутність реплікаційних ваг, які необхідні при розрахунках похибки вибірки. Для продовження роботи програми слід натиснути кнопку **“Да”**.

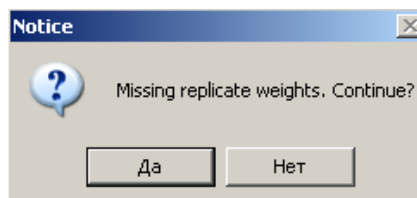


Рис. 6. Вікно з інформацією про відсутність реплікаційних ваг

Після цього з’являється діалогове вікно **“WESVAR Files”** (див. рис. 7), у якому вибирається папка та задається ім’я робочого файла формату „WesVarPC” (*.var). Для збереження файла необхідно натиснути кнопку **“Сохранить”**. При цьому програма автоматично залишає ім’я вхідного файла та змінює тільки розширення файла з **„*.dbf”** на **„*.var”**.

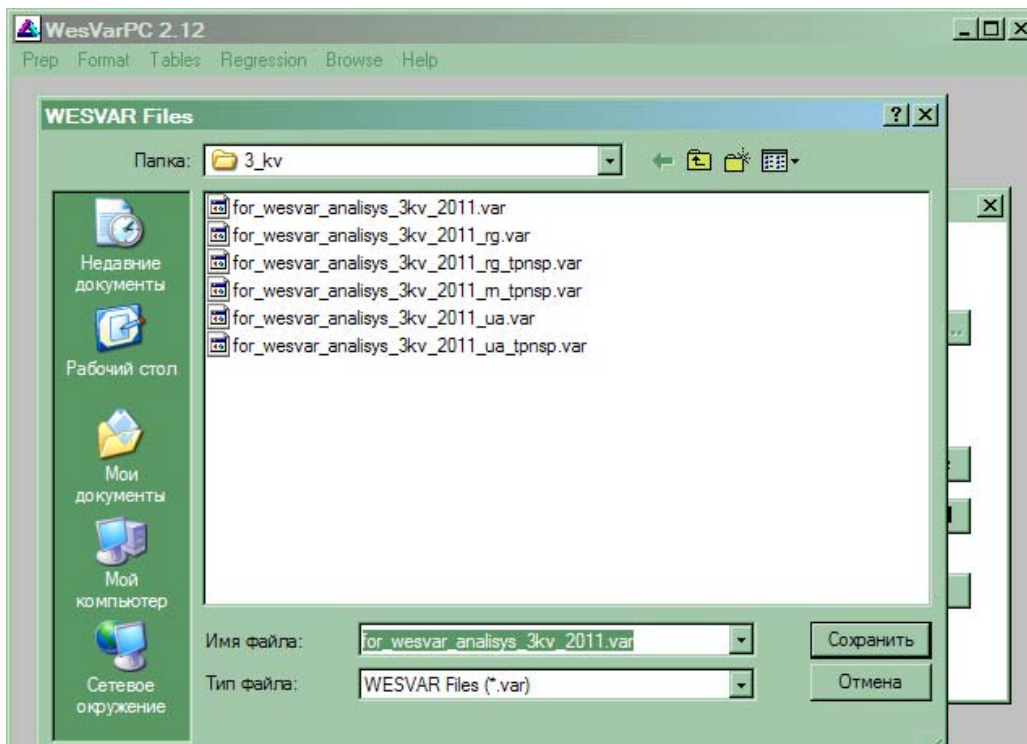


Рис. 7. Вікно для збереження робочого файла WesVar

У вікні програми, яке відображає динаміку процесу збереження робочого файла „WesVarPC”, у разі необхідності, можливо перервати цей процес за допомогою кнопки “Cancel” (див. рис. 8).

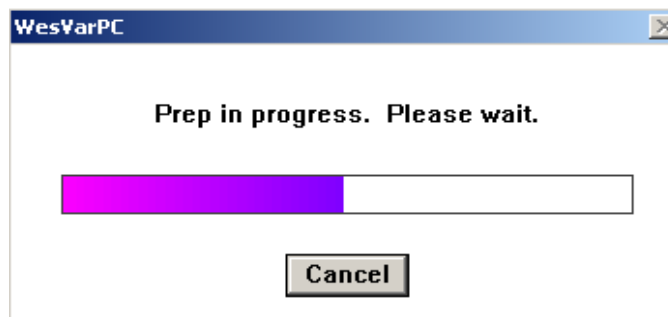


Рис. 8. Вікно відображення процесу збереження робочого файла „WesVarPC”

Після завершення формування робочого файла „WesVarPC” з’являється вікно “Information” з відповідною інформацією (див. рис. 9).

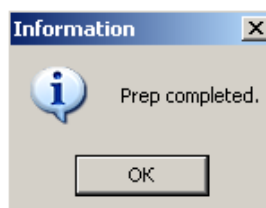


Рис. 9. Вікно з інформацією про завершення процесу збереження

Після натиснення кнопки “OK” з’являється вікно “Notice” з пропозицією створити реплікаційні ваги (див. рис. 10). Натиснути кнопку “Да”.

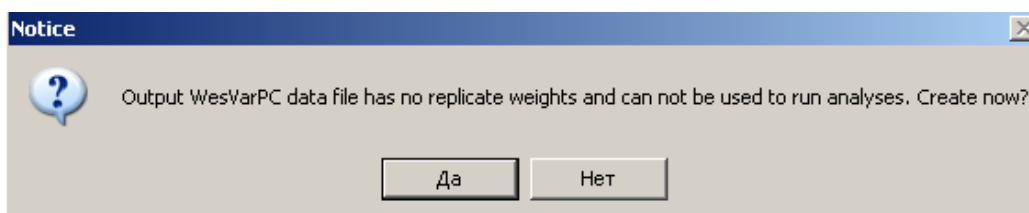


Рис. 10. Вікно з пропозицією створити реплікаційні ваги

4. Формування реплікаційних ваг

Після погодження з пропозицією про створення реплікаційних ваг з’явиться діалогове вікно “Creating Weights for WesVarPC Data File”, яке представлено на рис. 11.

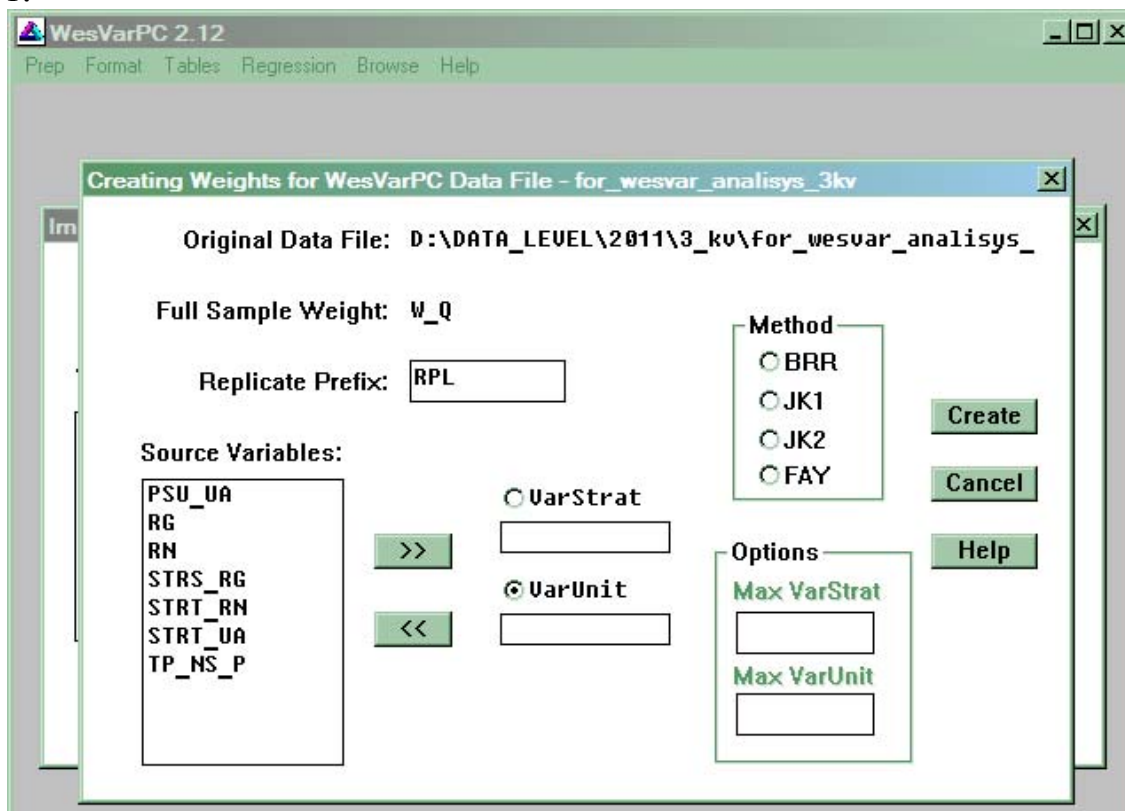


Рис. 11. Початковий вигляд вікна “Creating Weights for WesVarPC Data File”

У полі “Source Variables” вказані імена всіх змінних сформованого робочого файла „WesVarPC”.

Для формування реплікаційних ваг необхідно:

1) обрати метод розрахунку в полі “Method”, поставивши індикатор біля позначки “BRR”;

2) у полі індикатора “VarUnit” задати змінну “PSU_UA” із списку змінних робочого файла;

3) у полі індикатора “VarStrat” поставити змінну “STRT_UA” із списку змінних робочого файла (якщо розрахунки будуть проводитися для рівня

України) або змінну “STRS_RG”(для рівня регіонів) та “STRS_RN” (для рівня економічних районів).

Кінцевий вигляд вікна “Creating Weights for WesVarPC Data File” наведено на рис.12.

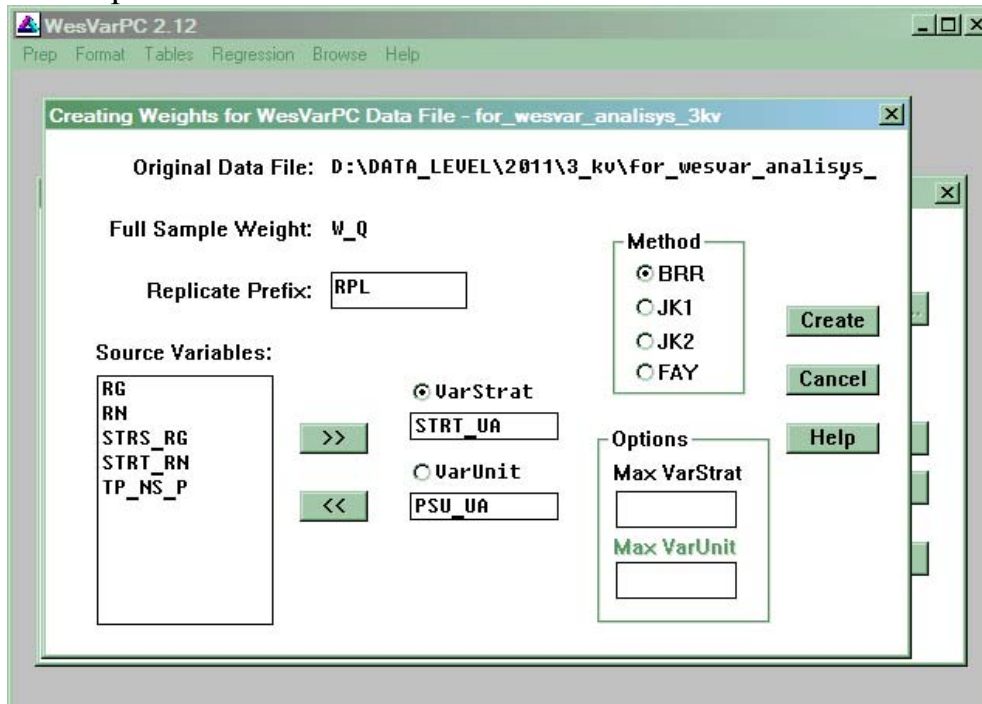


Рис. 12. Кінцевий вигляд діалогового вікна “Creating Weights for WesVarPC Data File”

Для збереження заданих параметрів формування реплікаційних ваг слід натиснути кнопку “Create”. У вікні, що з’явиться (див. рис. 7), необхідно задати ім’я файла та зробити до запропонованого імені букви „_UA” або „_RG” у залежності від рівня, для якого здійснюється обробка даних (по Україні чи регіонах, економічних районах). Наприклад, назва файла буде відповідно „for_wesvar_analisys_3kv_2011_UA.var”, „for_wesvar_analisys_3kv_2011_RG.var”, „for_wesvar_analisys_3kv_2011_RN.var”.

Після виконання зазначеної процедури слід натиснути кнопку “Сохранить”.

У вікні “WesVarPC” буде відображено динаміку процесу створення робочого файла програми з реплікаційними вагами та є можливість зупинити цей процес за допомогою кнопки “Cancel” (див. рис. 13).

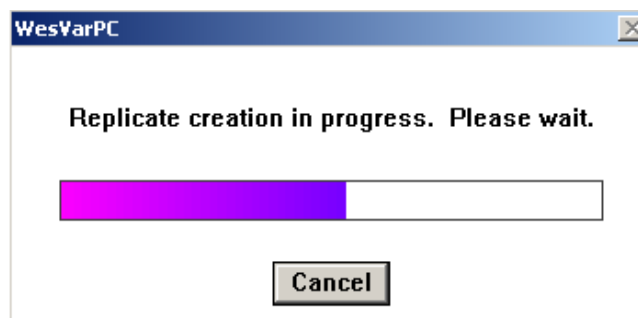


Рис. 13. Вікно відображення процесу створення файлу з реплікаційними вагами

Після завершення процесу формування файлу з'являється вікно **“Information”** з інформацією про те, що реплікаційні ваги сформовано успішно (див. рис. 14).

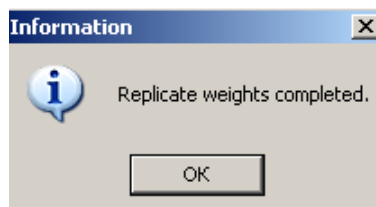


Рис. 14. Вікно з інформацією про успішне формування реплікаційних ваг

Після закінчення вищезазначеної процедури слід повернутися до основного діалогового вікна програми **„WesVarPC”** (див. рис. 2).

5. Повторне формування реплікаційних ваг

Якщо розрахунки здійснюються для різних рівнів агрегації (по Україні чи регіонах), то необхідно виконати таке:

1) сформувати робочий файл із реплікаційними вагами (**„for_wesvar_analisys_3kv_2011_UA.var”**) для рівня України, як описано в п. 3 та 4;

2) для формування реплікаційних ваг у пункті меню **“Prep”** основного діалогового вікна програми вибрати опцію **“Create Replicate Weights”** (див. рис. 15);

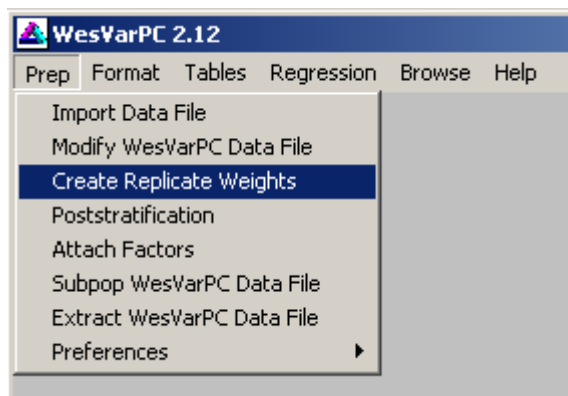


Рис. 15. Вибір опції формування реплікаційних ваг

3) вибрати робочий файл без реплікаційних ваг (**„for_wesvar_analisys_3kv_2011_RG.var”**) та відкрити його, натиснувши кнопку **“Открыть”** (див. рис. 16);

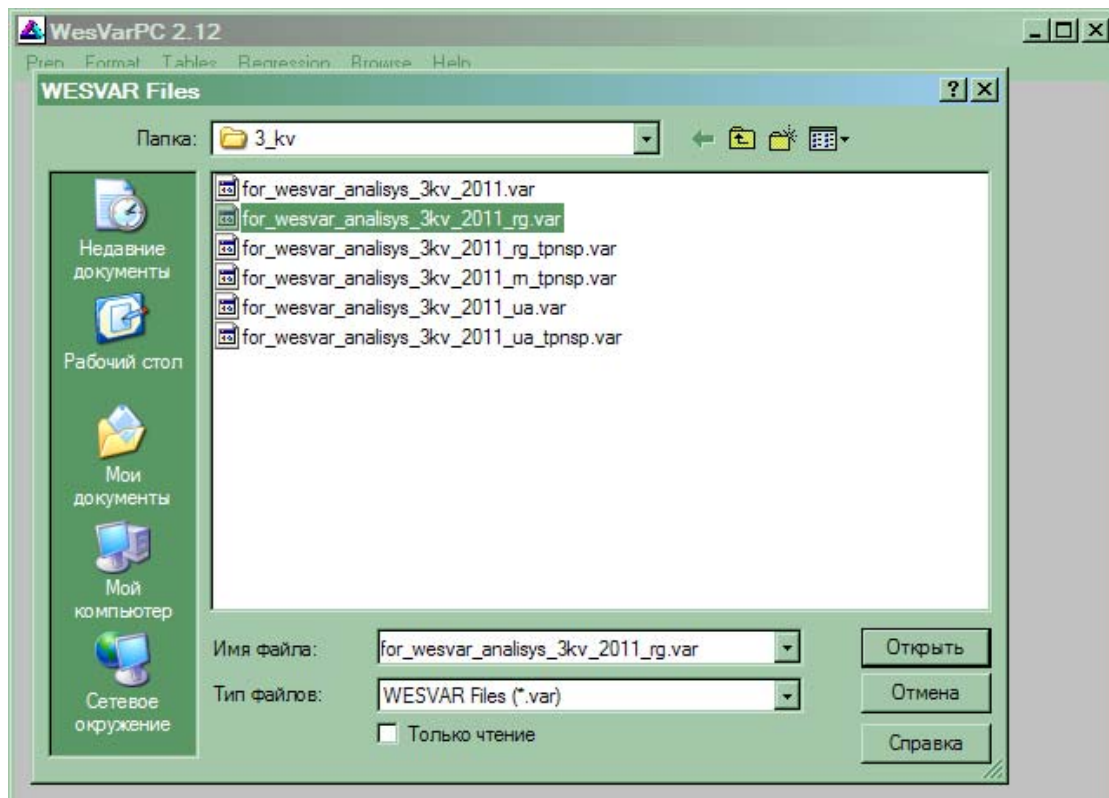


Рис. 16. Вікно вибору робочого файлу

4) після появи діалогового вікна **“Creating Weights for WesVarPC Data File”** (див. рис. 11) виконати операції, описані в п. 4. У результаті буде отримано робочий файл із реплікаційними вагами для подальших розрахунків для рівня регіонів (**for_wesvar_analisys_3kv_2011_RG.var**);

5) повернутися в основне вікно програми.

6. Побудова таблиць з характеристиками надійності

Для побудови таблиць необхідно обрати у пункті меню **“Tables”** основного діалогового вікна опцію **“New”** (див. рис. 17).

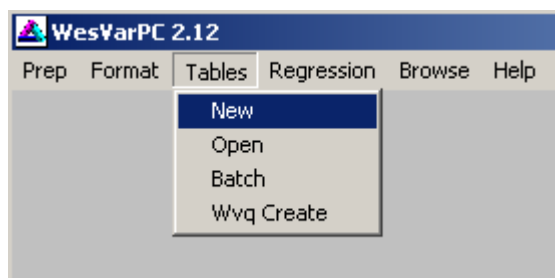


Рис. 17. Вибір опції для формування таблиць із характеристиками надійності

У діалоговому вікні **“WESVAR Files”** (див. рис. 16) необхідно вибрати робочий файл із реплікаційними вагами **for_wesvar_analisys_3kv_2011_UA.var** та відкрити його, натиснувши кнопку **“Открыть”**.

Далі з’являється діалогове вікно **“Table Request”**, яке представлене на рис. 18.

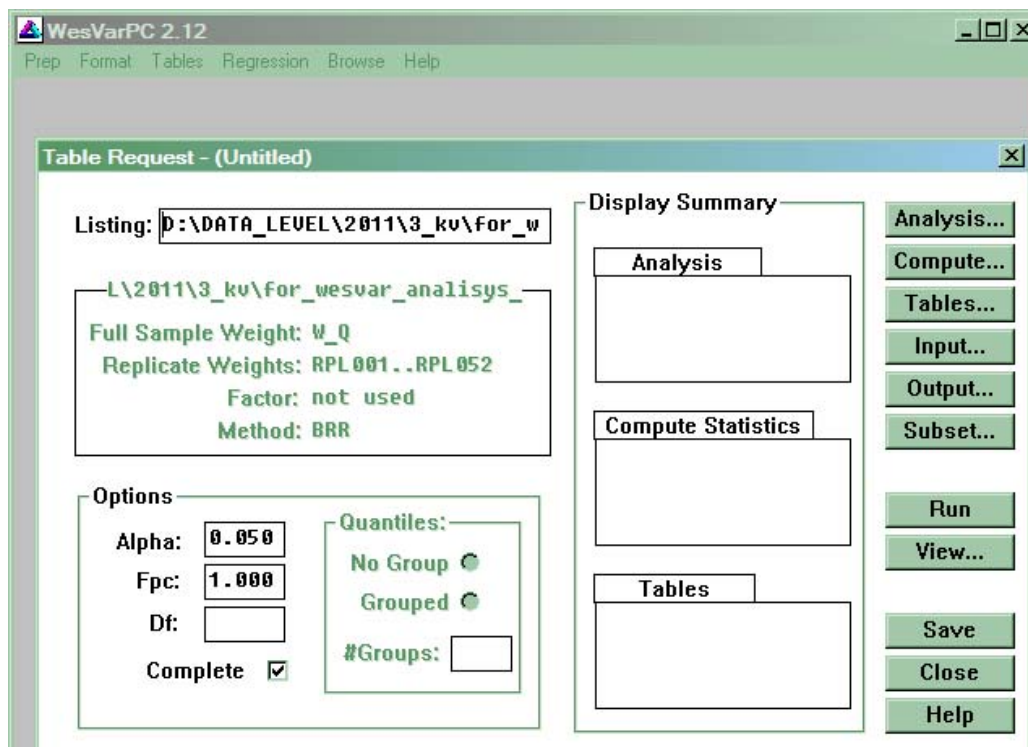


Рис. 18. Вікно “Table Request”

Зазначене вікно можна умовно розділити на чотири зони:

- **інформаційна** зона, яка містить інформацію про робочий файл і файл виводу результатів розрахунків;
- зона **опцій** (“Options”), що задає певні параметри для розрахунків;
- зона **описання виводу** (“Display Summary”), що зосереджує інформацію про основні операції розрахунків і формат виводу результатів;
- зона **функціональних кнопок**, за допомогою яких формуються основні параметри розрахунків, здійснюється управління процесом розрахунку та робочим файлом.

6.1. Формування таблиць

Для формування таблиць із характеристиками надійності необхідно натиснути кнопку “Tables...” у зоні функціональних кнопок, після чого з’явиться діалогове вікно “Tables” (див. рис. 19).

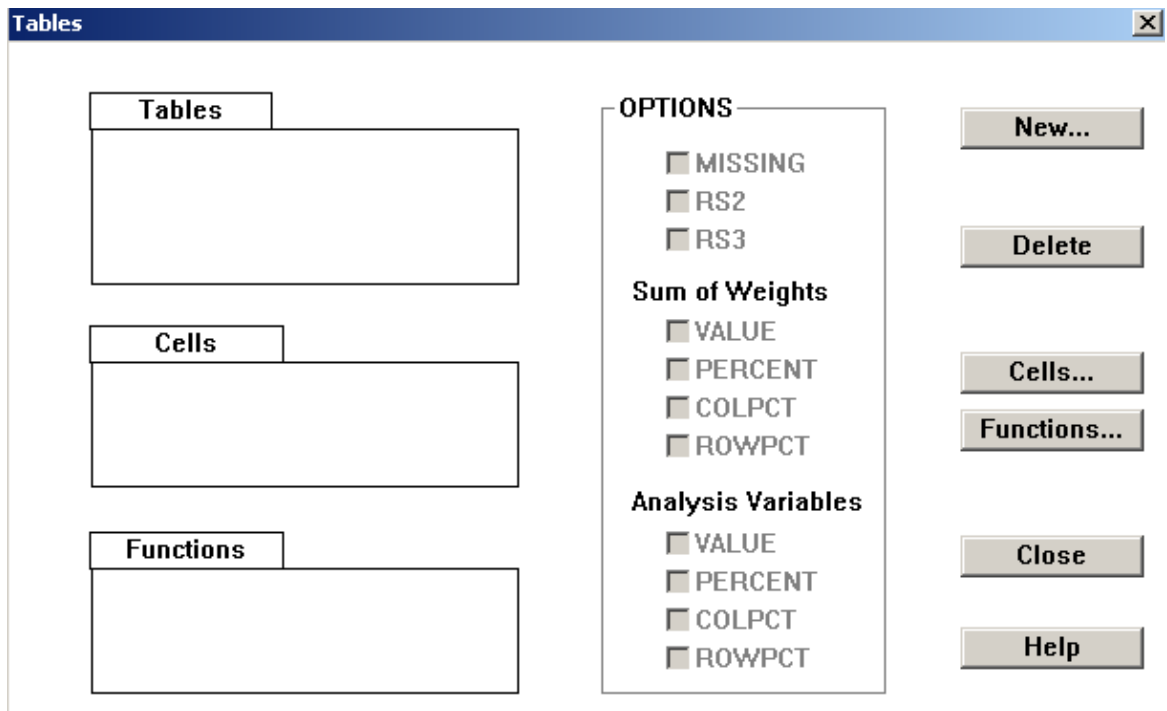


Рис. 19. Вікно “Tables”

Після натискання кнопки “New...” з’явиться вікно “Define Table” визначення таблиці (див. рис. 20).

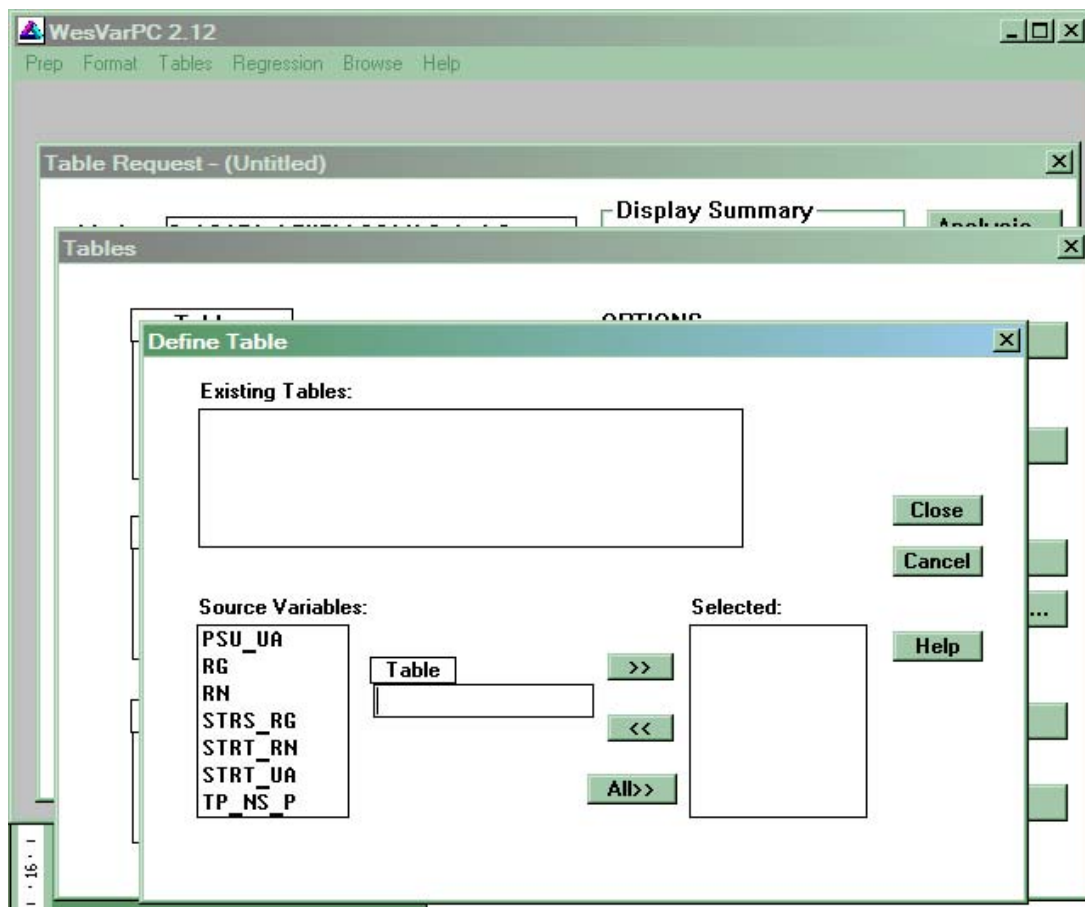


Рис. 20. Вікно “Define Table”

У діалоговому вікні **“Define Table”** визначаються необхідні таблиці на основі існуючих змінних зі списку **“Source Variables”**. Для кожної таблиці розраховуються необхідні оцінки показників і характеристики її надійності.

6.1.1. Побудова одновимірних таблиць

У полі **“Source Variables”** слід двічі натиснути „мишею” по змінній, за якою будується одновимірна таблиця.

Після перенесення змінної у поле **“Table”** слід натиснути кнопку **>>**. У результаті чого змінна перенесеться у поле **“Selected”**, бо лише після цього по ній буде побудовано одновимірну таблицю.

6.1.2. Побудова двовимірних таблиць

Для побудови двовимірних таблиць необхідно вибрати першу змінну (двічі натиснувши „мишею”), за якою будується двовимірна таблиця (основний розподіл), а потім двічі натиснути другу змінну, по якій будується двовимірна таблиця (внутрішній розподіл). Після того, як обидві змінні будуть розділені символом **“*”** та з’являться у полі **“Table”**, слід натиснути кнопку **>>**, що дозволить перенести їх у поле **“Selected”**. Так буде побудовано двовимірну таблицю.

Аналогічно задаються будь-які інші таблиці.

Для прикладу на рис.21 представлено вигляд вікна **“Define Table”** у процесі формування одновимірної (розподіл населення за статтю) та двовимірної таблиць (розподіл зайнятих за статтю).

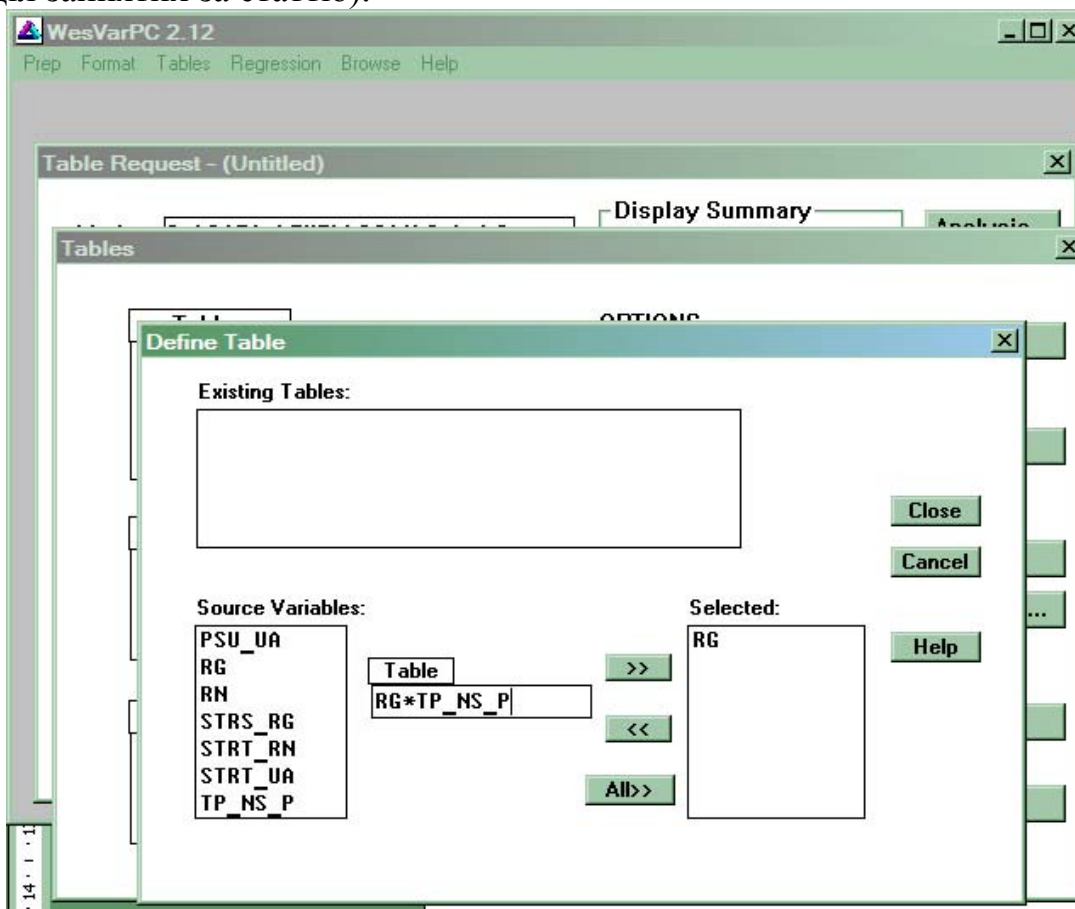


Рис. 21. Вигляд вікна **“Define Table”** у процесі формування таблиць

Після того, як у полі **“Selected”** будуть знаходитись усі необхідні таблиці (за розрізами), слід повернутися до попереднього діалогового вікна **“Tables”**, натиснувши на кнопку **“Close”**. У полі **“Tables”** цього вікна будуть зазначені всі сформовані таблиці, а в полі **“Options”** – перелік оцінок, для яких розраховуються характеристики надійності (див. рис.22). При цьому перелік оцінок може бути збільшений або зменшений шляхом проставлення або зняття галочки біля певної оцінки. Для розрахунків використовується перелік оцінок, що задається автоматично.

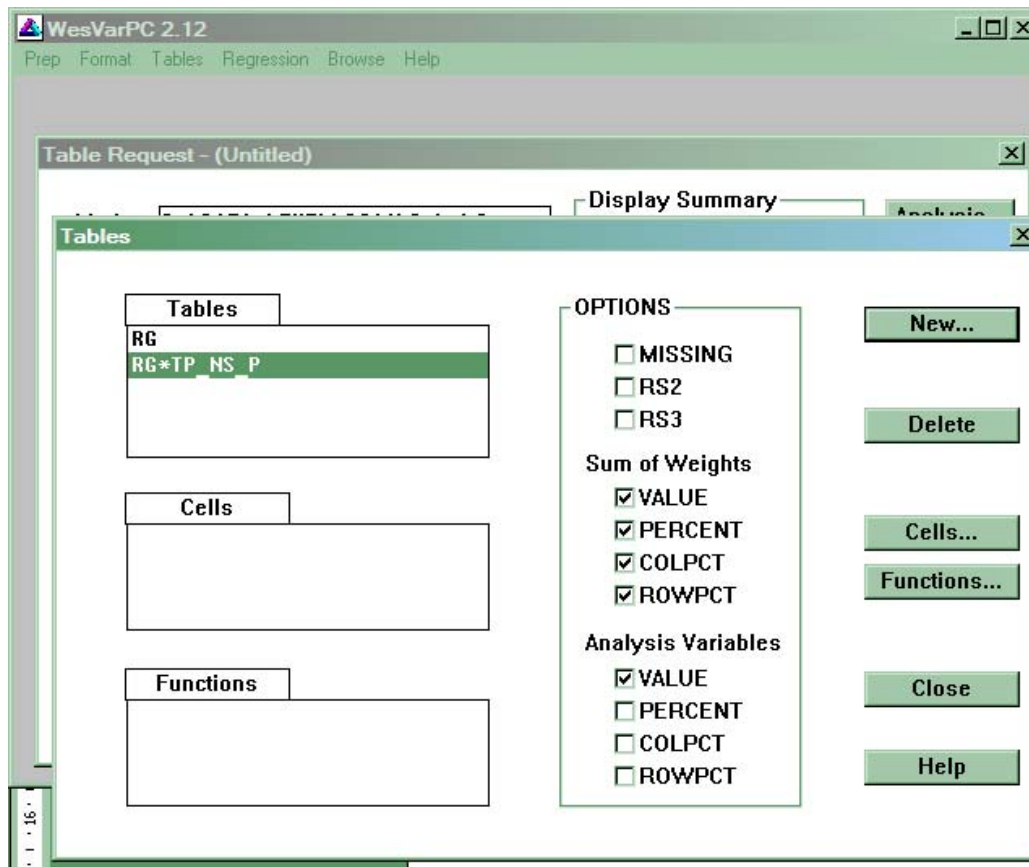


Рис. 22. Вигляд вікна **“Tables”** після формування таблиць

Натиснувши кнопку **“Close”**, повертаємося до вікна **“Table Request”**. У полі **“Tables”** цього вікна перелічені всі сформовані таблиці (див. рис. 23).

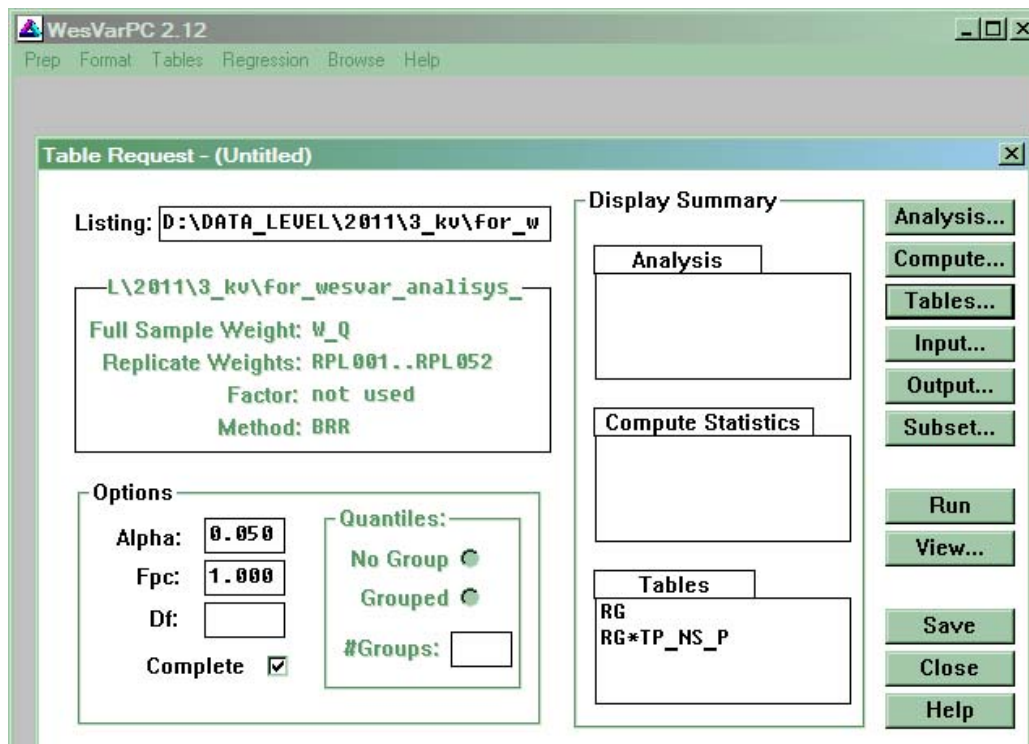


Рис. 23. Вигляд вікна “Table Request” після формування таблиць
У полі “Compute” визначаються типи оцінок показників (рис. 24 та рис.25).

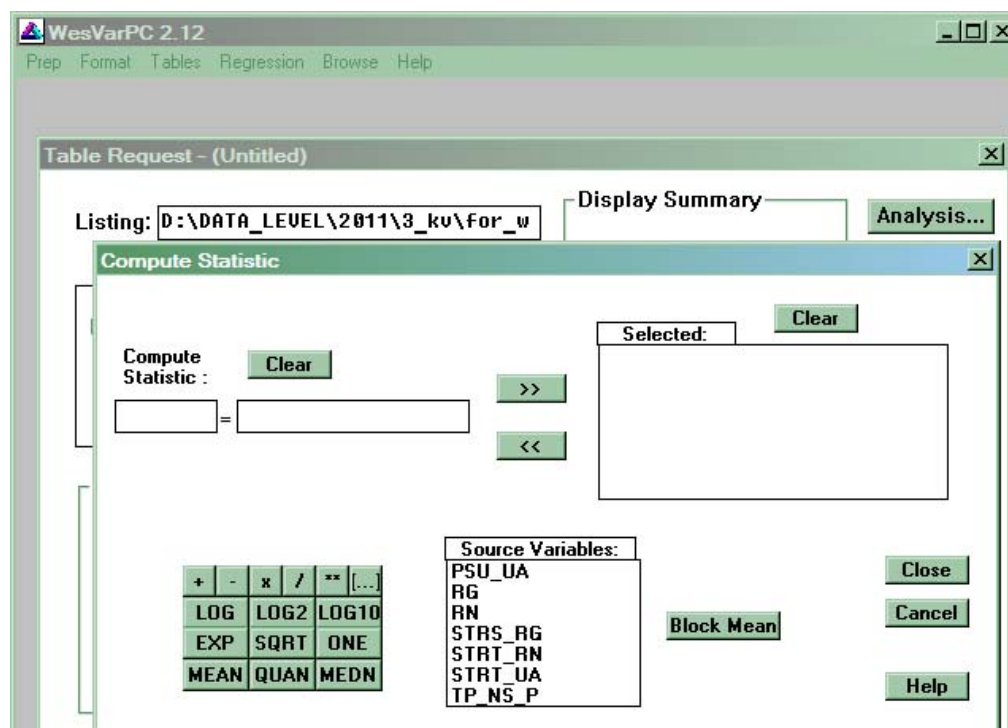


Рис. 24. Вигляд вікна “Compute Statistic” для вибору типу оцінок показників.

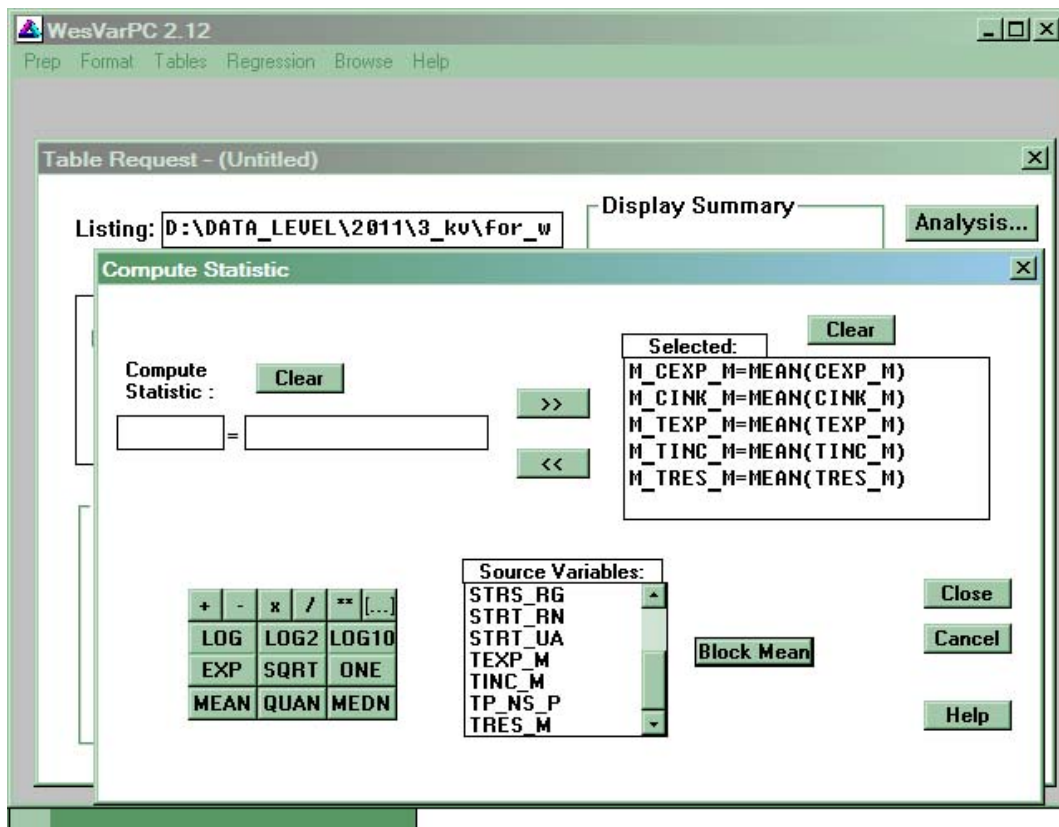


Рис. 25. Вікно “Compute Statistic” із завданням на розрахунок середніх значень змінних.

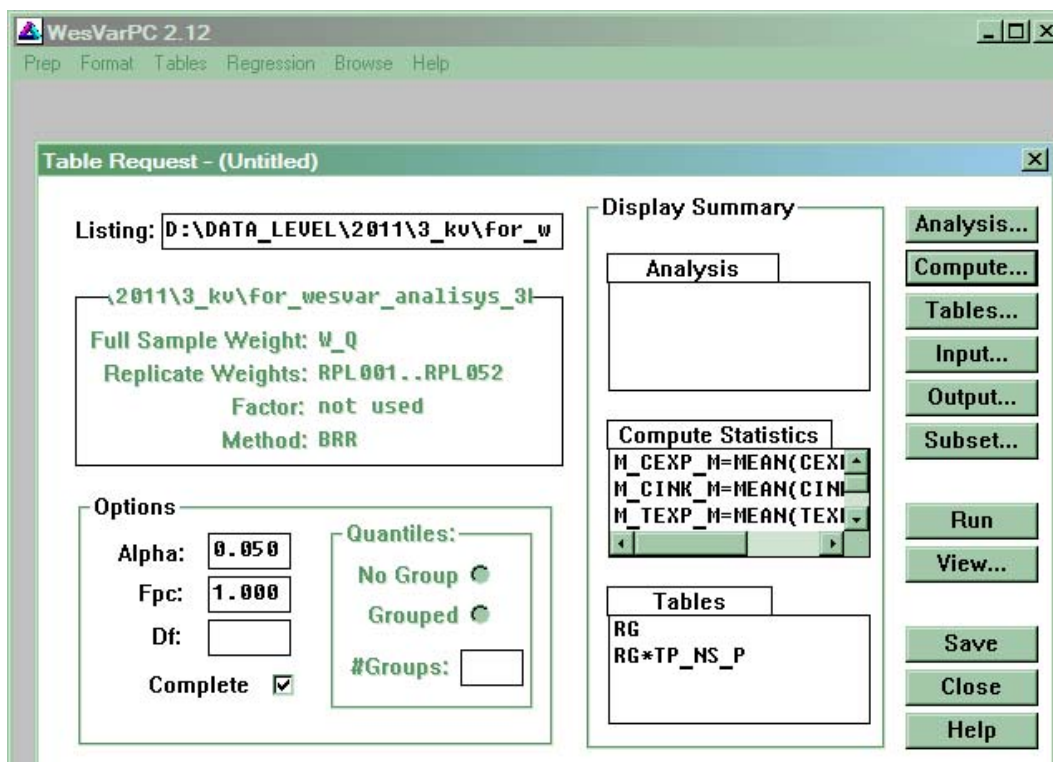


Рис. 26. Вікно “Table Request” програми WesVarPC після формування таблиць для аналізу та вибору типу оцінок.

Для виконання цієї процедури слід натиснути кнопку **“Run”**. З’являється вікно, яке відображає динаміку процесу розрахунку оцінок і побудови таблиць (див. рис. 27).

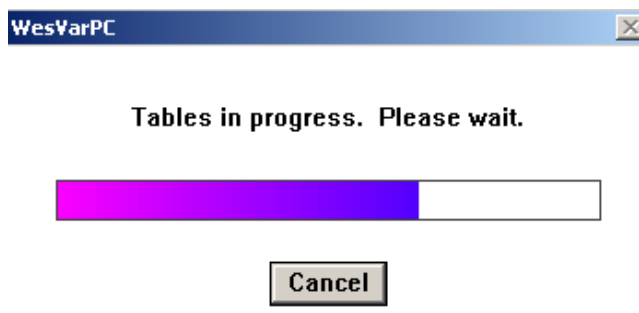


Рис. 27. Вікно відображення процесу побудови таблиць

Після завершення процесу побудови таблиць з’являється відповідне повідомлення (див. рис. 28). За допомогою кнопки **“OK”** слід повернутися до вікна **“Table Request”** (див. рис. 23).

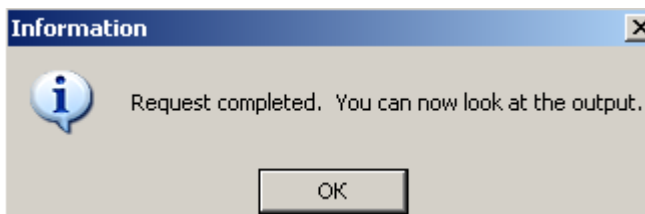


Рис. 28. Повідомлення про завершення процесу побудови таблиць

Результати розрахунків зберігаються у відповідному файлі з розширенням **„*.LST”**. Ознайомитись із результатами роботи програми можна, натиснувши на кнопку **“View”** у вікні **“Table Request”** (див. рис. 23). У результаті цієї маніпуляції з’являється вікно виводу **“Browse”** у тестовому форматі. Також результати можна переглянути, відкривши файл із результатами розрахунків у програмі **„Excel”**.

Розрахунок показників якості даних обстеження для рівня регіонів та економічних районів здійснюється аналогічно як і для всіх домогосподарств по Україні. При розрахунку показників для рівня регіонів використовується файл **„for_wesvar_analisys_3kv_2011_RG.var”**, для рівня економічних районів - **„for_wesvar_analisys_3kv_2011_RN.var”**.

Після завершення роботи програми WesVarPC у директорії, де знаходився вихідний робочий файл у форматі SPSS, створюються нові файли з такими ж самими іменами і розширенням *.var, *.wvq, *.lst. Результати здійснених розрахунків знаходяться у текстовому файлі *.lst. Приклад такого файла наведено в додатку 2.

7. Завершення роботи програми „WesVarPC”

Для завершення роботи програми „WesVarPC” необхідно вийти з головного вікна (рис. 2) одним з таких способів:

- натиснути на кнопку “**x**” у правому верхньому куті вікна;
- застосувати комбінації клавіш “**ALT+F4**”;
- натиснути на емблему програми у лівому верхньому куті вікна й обрати опцію “**Закреть**”.

**Вихідні дані та таблиці з результатами розрахунків
для прикладу розрахунків основних показників доходів і витрат**

PC	WESVAR	VERSION	NUMBER:	02.Гпр			
TIME	THE	JOB	EXECUTED:	16:09:53	01/19/11		
INPUT	DATASET	NAME:					
\\Invest1\D\DATA_LEVEL\2011\3_kv\for_wesvar_analisys_3kv_2011_ua_tpnsp.var							
OUTPUT	LISTING:						
\\Invest1\D\DATA_LEVEL\2011\3_kv\for_wesvar_analisys_3kv_2011_ua.LST							
NOSUMMAR							
OPTION	Y	IS:	OFF				
OPTION	FUNCTION	LOG	IS:	OFF			
OPTION	ALIGNMENT	IS:	OFF				
OPTION	EXPORT	IS:	OFF				
VARIANC							
E	ESTIMATION	METHOD:	BRR				
FINITE	N	CORRECTION	FACTOR:	1.00000	INTERVA		
VALUE	OF	ALPHA	(CONFIDENCE	L	%)		
0.05000	(95.00000						
DEGREES	OF	FREEDOM:	INFINITE				
t	VALUE:	1.960					
OPTION	COMPLETE	IS:	ON				
FULL	SAMPLE	WEIGHT:	W_Q				
REPLICAT							
E	WEIGHTS:	RPL001...RPL052					
ANALYSIS	VARIABLES:	None	Specified.				
COMPUTE	STATISTIC:	M_CEXP_M=MEAN(CEXP_M)					
M_CINK_M=MEAN(CINK_M)							
M_TEXP_M=MEAN(TEXP_M)							
M_TRES_M=MEAN(TRES_M)							
M_TINC_M=MEAN(TINC_M)							
TABLE	REQUESTS:	TP_NS_P					
FACTOR(S							
):	1.00						
NUMBER	OF	REPLICATES:	52				
NUMBER	OF	S	READ:	10663			
WEIGHTE							
D	NUMBER	OF	S	READ:	17022668.000		
TABLE	REQUEST	:	TP_NS_P				
TP_NS_P	STATISTIC	EST_TYPE	ESTIMATE	STDERRO	CV(%)	N	DEFF
1	SUM_WTS	VALUE	7370635.00	127432.014	1.729	4388	N/A
2	SUM_WTS	VALUE	4404012.00	68187.907	1.548	2868	N/A
3	SUM_WTS	VALUE	5248021.00	103815.205	1.978	3407	N/A
MARGINA							
L	SUM_WTS	VALUE	17022668.00	169015.413	0.993	10663	N/A
1	M_CEXP_M	VALUE	3761.24	66.889	1.778	4388	0.813
2	M_CEXP_M	VALUE	2948.12	36.843	1.250	2868	0.906
3	M_CEXP_M	VALUE	2651.41	75.932	2.864	3407	3.305
MARGINA	M_CEXP_M	VALUE	3208.72	39.423	1.229	10663	1.215

L								
	1	M_CINK_M	VALUE	4092.88	44.229	1.081	4388	0.439
	2	M_CINK_M	VALUE	3373.93	41.342	1.225	2868	0.959
	3	M_CINK_M	VALUE	3305.82	75.469	2.283	3407	2.588
MARGINA								
L		M_CINK_M	VALUE	3664.23	33.232	0.907	10663	0.962
	1	M_TEXP_M	VALUE	3991.32	73.584	1.844	4388	0.904
	2	M_TEXP_M	VALUE	3232.25	39.510	1.222	2868	0.837
	3	M_TEXP_M	VALUE	3122.48	80.451	2.577	3407	3.238
MARGINA								
L		M_TEXP_M	VALUE	3527.07	41.680	1.182	10663	1.232
	1	M_TRES_M	VALUE	4479.85	125.705	2.806	4388	1.135
	2	M_TRES_M	VALUE	3655.52	42.686	1.168	2868	0.866
	3	M_TRES_M	VALUE	3815.29	81.718	2.142	3407	2.672
MARGINA								
L		M_TRES_M	VALUE	4061.70	62.116	1.529	10663	1.337
	1	M_TINC_M	VALUE	4198.06	46.052	1.097	4388	0.472
	2	M_TINC_M	VALUE	3566.44	43.310	1.214	2868	0.979
	3	M_TINC_M	VALUE	3716.14	80.455	2.165	3407	2.733
MARGINA								
L		M_TINC_M	VALUE	3886.08	34.822	0.896	10663	1.033

Значення змінної strt_ua
(страти для України по регіонах та по типах місцевості)

Регіон	Тип місцевості	strt_ua
Автономна Республіка Крим	міська	1
	сільська	2
Вінницька	міська	3
	сільська	4
Волинська	міська	5
	сільська	6
Дніпропетровська	міська	7
	сільська	8
Донецька	міська	9
	сільська	10
Житомирська	міська	11
	сільська	12
Закарпатська	міська	13
	сільська	14
Запорізька	міська	15
	сільська	16
Івано-Франківська	міська	17
	сільська	18
Київська	міська	19
	сільська	20
Кіровоградська	міська	21
	сільська	22
Луганська	міська	23
	сільська	24
Львівська	міська	25
	сільська	26
Миколаївська	міська	27
	сільська	28
Одеська	міська	29
	сільська	30
Полтавська	міська	31
	сільська	32
Рівненська	міська	33
	сільська	34
Сумська	міська	35
	сільська	36
Тернопільська	міська	37
	сільська	38

Регіон	Тип місцевості	strt ua
Харківська	міська	39
	сільська	40
Херсонська	міська	41
	сільська	42
Хмельницька	міська	43
	сільська	44
Черкаська	міська	45
	сільська	46
Чернівецька	міська	47
	сільська	48
Чернігівська	міська	49
	сільська	50
м. Київ	міська	51
м. Севастополь	міська	52

Значення змінної strt_rg
(страти, побудовані у межах кожного регіону)

Регіон	strt_rg	Номери ВТОВ, ПТОВ
Автономна Республіка Крим	1	2, 4
	2	6, 8
	3	10, 12
	4	14, 16
	5	17, 19
	6	21, 23
	7	25, 27
	8	30, 32
	9	33, 35
	10	38, 40
	11	43, 45
	12	45, 47
	13	49, 51
	14	54, 56
	15	58, 59
	16	61, 63
	17	2001-2003, 2008
	18	2004-2007
	19	2009-2012
	20	2013-2017
Вінницька	1	65, 67
	2	69, 71
	3	73, 75
	4	77, 79
	5	81, 83
	6	86, 88
	7	90, 92
	8	94, 96
	9	98, 100
	10	101, 103
	11	105, 107
	12	110, 112

	13	113, 115
	14	2018 - 2021
	15	2022-2025
	16	2026-2029
	17	2030-2033
	18	2034-2037
	19	2038-204, 2043
	20	2041, 2042, 2044-2046
Волинська	1	118, 120
	2	122, 124
	3	126, 128
	4	130, 132
	5	133, 135
	6	137, 139
	7	141, 143
	8	146, 148
	9	150, 152
	10	154, 156
	11	158, 160
	12	2047-2050
	13	2051-2053, 2056
	14	2054, 2055, 2057, 2058
	15	2059-2062
	16	2063-2065
Дніпропетровська	1	162, 164
	2	166, 168
	3	170, 172
	4	174, 176
	5	178, 180
	6	182, 184
	7	186, 188
	8	190, 192
	9	194, 196
	10	198, 200
	11	201, 203
	12	205, 207
	13	209, 211
	14	213, 215

	15	217, 219
	16	221, 223
	17	225, 227
	18	230, 232
	19	234, 236
	20	238, 240
	21	242, 244
	22	246, 248
	23	250, 252
	24	253, 255
	25	258, 260
	26	262, 264
	27	265, 267
	28	270, 272
	29	274, 276
	30	278, 280
	31	2066-2069
	32	207-2073
	33	2074-2077
	34	2078-2081
	35	2082-2085
	36	2086-2089
Донецька	1	282, 284
	2	286, 288
	3	290, 292
	4	294, 296
	5	298, 300
	6	302, 304
	7	306, 308
	8	310, 312
	9	314, 316
	10	318, 320
	11	322, 324
	12	326, 328
	13	330, 332
	14	335, 337
	15	339, 340
	16	342, 344

	17	346, 348
	18	350, 351
	19	354, 356
	20	357, 359
	21	362, 363
	22	366, 368
	23	370, 372
	24	373, 375
	25	377, 379
	26	381, 384
	27	386, 388
	28	390, 392
	29	393, 394
	30	398, 400
	31	401, 403
	32	406, 408
	33	409, 411
	34	413, 415
	35	418, 419
	36	2090-2093
	37	2094-2096, 2099
	38	2097, 2098, 2100, 2101
	39	2102-2104
	40	2106-2108, 2111
	41	2109, 2110, 2112, 2113
Житомирська	1	421, 423
	2	425, 427
	3	429, 431
	4	433, 435
	5	438, 440
	6	442, 444
	7	445, 447
	8	450, 452
	9	453, 455
	10	2114- 2117
	11	2118-2121
	12	2122-2125

	13	2126-2129
	14	2130-2133
	15	2134-2137
Закарпатська	1	457, 459
	2	461, 463
	3	466, 468
	4	469, 471
	5	474, 476
	6	478, 480
	7	481, 483
	8	486, 488
	9	489, 491
	10	2138-2140, 2144
	11	2141-2143, 2147
	12	2145, 2146, 2148, 2149
	13	2150-2154
	14	2155-2158
Запорізька	1	493, 495
	2	497, 499
	3	501, 503
	4	505, 507
	5	509, 511
	6	513, 515
	7	517, 519
	8	521, 523
	9	525, 527
	10	530, 532
	11	534, 536
	12	538, 540
	13	542, 544
	14	546, 548
	15	550, 552
	16	554, 556
	17	557, 559
	18	2159 - 2162
	19	2163-2166
	20	2167-2170

	21	2171-2174
	22	2175-2177, 2182
	23	2178-2181
	24	2183-2185
Івано-Франківська	1	561, 563
	2	565, 567
	3	569, 571
	4	573, 575
	5	577, 579
	6	582, 583
	7	585, 587
	8	589, 591
	9	593, 595
	10	2186-2188, 2191
	11	2189, 2190, 2192, 2193
	12	2194-2197
	13	2198-2201
	14	2202-2206
Київська (без м. Києва)	1	598, 600
	2	602, 604
	3	606, 608
	4	609, 611
	5	614, 616
	6	618, 620
	7	623, 624
	8	625, 627
	9	630, 632
	10	634, 636
	11	637, 639
	12	642, 644
	13	646, 648
	14	649, 651
	15	653, 655
	16	2207-2209, 2212
	17	2210, 2211, 2213, 2214
	18	2215-2218
	19	2219-2222

	20	2223-2227
	21	2228-2232
Кіровоградська	1	658, 660
	2	662, 664
	3	666, 668
	4	670, 672
	5	674, 676
	6	677, 679
	7	681, 683
	8	686, 688
	9	689, 691
	10	694, 696
	11	697, 699
	12	702, 704
	13	2233-2236
	14	2237-2240
	15	2241-2244
	16	2245-2249
	17	2250-2254
Луганська	1	706, 708
	2	710, 712
	3	714, 716
	4	718, 720
	5	722, 724
	6	725, 727
	7	729, 731
	8	734, 736
	9	737, 739
	10	741, 744
	11	745, 747
	12	749, 751
	13	754, 756
	14	757, 759
	15	761, 763
	16	766, 768
	17	770, 772
	18	773, 775
	19	778, 780

	20	2255-2258
	21	2259-2262
	22	2263-2266
	23	2267-2270
	24	2271- 2273
	25	2274-2277
Львівська	1	782, 784
	2	786, 788
	3	790, 792
	4	794, 796
	5	798, 800
	6	802, 804
	7	806, 808
	8	810, 812
	9	814, 816
	10	818, 820
	11	821, 823
	12	825, 827
	13	830, 832
	14	833, 835
	15	837, 839
	16	842, 844
	17	845, 847
	18	2278-2281
	19	2282-2284, 2291
	20	2285-2287, 2297
	21	2288-2290, 2299
	22	2292, 2293
	23	2294-2296, 2304
	24	2300-2303
	25	2305-2308
Миколаївська	1	850, 852
	2	854, 856
	3	858, 860
	4	862, 864
	5	866, 868
	6	870, 872
	7	874, 876

	8	878, 880
	9	881, 883
	10	885, 887
	11	889, 891
	12	893, 894
	13	2309-2312
	14	2313-2316
	15	2317-2320
	16	2321-2324
	17	2325-2329
Одеська	1	898, 900
	2	902, 904
	3	906, 908
	4	910, 912
	5	914, 916
	6	918, 920
	7	922, 924
	8	926, 928
	9	930, 932
	10	934, 936
	11	938, 940
	12	941, 943
	13	946, 948
	14	949, 951
	15	953, 955
	16	958, 960
	17	962, 964
	18	965, 967
	19	2330-2333
	20	2334-2337
	21	2338-2341
	22	2342-2345
	23	2346-2349
	24	2350-2353
	25	2354-2357
Полтавська	1	969, 971
	2	973, 975
	3	977, 979

	4	982, 984
	5	986, 988
	6	989, 991
	7	993, 995
	8	997, 999
	9	1002, 1004
	10	1005, 1007
	11	2358-2361
	12	2362-2365
	13	2366-2369
	14	2370-2373
	15	2374-2377
	16	2378-2381
	17	2382-2384
Рівненська	1	1010, 1012
	2	1014, 1016
	3	1018, 1020
	4	1022, 1024
	5	1025, 1027
	6	1030, 1032
	7	1034, 1036
	8	2385-2388
	9	2389-2392
	10	2393-2397
	11	2398-2402
Сумська	1	1037, 1039
	2	1041, 1043
	3	1045, 1047
	4	1049, 1051
	5	1054, 1056
	6	1057, 1059
	7	1062, 1064
	8	1066, 1068
	9	1070, 1072
	10	1074, 1076
	11	2403-2406
	12	2407-2410
	13	2411-2414

	14	2415-2418
	15	2419-2422
Тернопільська	1	1077, 1079
	2	1081, 1083
	3	1085, 1087
	4	1089, 1091
	5	1093, 1095
	6	1097, 1099
	7	1102, 1104
	8	1106, 1108
	9	2423-2426
	10	2427-2430
	11	2431-2434
	12	2435-2437,2440
	13	2438, 2439, 2441-2443
Харківська	1	1109, 1111
	2	1113, 1115
	3	1117, 1119
	4	1121, 1123
	5	1125, 1127
	6	1129, 1131
	7	1133, 1135
	8	1137, 1139
	9	1141, 1143
	10	1145, 1147
	11	1149, 1151
	12	1153, 1155
	13	1157, 1159
	14	1161, 1163
	15	1165, 1167
	16	1169, 1171
	17	1174, 1176
	18	1178, 1180
	19	1182, 1184
	20	1185, 1187
	21	1190,1192
	22	1194, 1196
	23	2444-2447

	24	2448-2451
	25	2452-24555
	26	2456-2459
	27	2460-2463
	28	2462-2467
	29	2468-2472
Херсонська	1	1197, 1199
	2	1201, 1203
	3	1205, 1207
	4	1209, 1211
	5	1213, 1215
	6	1217, 1219
	7	1222, 1223
	8	1225, 1227
	9	1230, 1232
	10	1234, 1236
	11	1238, 1240
	12	1242, 1244
	13	1246, 1248
	14	2473-2476
	15	2477-2481
	16	2482-2485
	17	2486-2489
	18	2490-2493
	19	2494-2497
	20	2498-2501
Хмельницька	1	1249, 1251
	2	1253, 1255
	3	1257, 1259
	4	1261, 1263
	5	1265, 1267
	6	1269, 1271
	7	1274, 1276
	8	1277, 1279
	9	1281, 1283
	10	1285, 1287
	11	2502-2505
	12	2506-2508, 2511

	13	2509, 2510, 2512, 2513
	14	2514-2518
	15	2519-2523
Черкаська	1	1289, 1291
	2	1292, 1293
	3	1295, 1296
	4	1298, 1301
	5	1306, 1308
	6	1310, 1312
	7	1313, 1315
	8	1317, 1319
	9	1322, 1324
	10	1326, 1328
	11	1329, 1331
	12	2524-2527
	13	2528-2531
	14	2532-2535
	15	2536-2539
	16	2540-2544
Чернівецька	1	1333, 1335
	2	1337, 1339
	3	1341, 1343
	4	1345, 1347
	5	1349, 1351
	6	1354, 1356
	7	1358, 1360
	8	1362, 1364
	9	2545-2548
	10	2549-2552
	11	2553-2557
	12	2558-2561
Чернігівська	1	1366, 1368
	2	1370, 1372
	3	1374, 1376
	4	1378, 1380
	5	1382, 1384
	6	1386, 1388

	7	1390, 1392
	8	1393, 1395
	9	1397, 1399
	10	2562-2565
	11	2566-2569
	12	2570-2573
	13	2574-2577
	14	2578-2581
	15	2582-2584
м.Київ	1	1401, 1403
	2	1405, 1407
	3	1409, 1411
	4	1413, 1415
	5	1417, 1419
	6	1421, 1423
	7	1425, 1431
	8	1429, 1431
	9	1433, 1435
	10	1437, 1439
	11	1441, 1443
	12	1445, 1447
	13	1449, 1451
	14	1453, 1455
	15	1457, 1459
	16	1461, 1463
	17	1465, 1467
	18	1469, 1471
	19	1473, 1475
	20	1477, 1479
	21	1481, 1483
	22	1485, 1487
	23	1489, 1491
	24	1493, 1495
	25	1497, 1499
	26	1501, 1503
	27	1505, 1507
	28	1509, 1511
м. Севастополь	1	1514, 1516

	2	1517, 1519
	3	1521, 1523
	4	1525, 1527
	5	1529, 1531
	6	1533, 1535
	7	1538, 1540
	8	1542, 1544