**II. Sampling scheme**

1. **Sample survey on yields and production of cereals**

The aim of the survey was to collect data on yields and sown area, as well as on production of cereals by voivodships. The surveyed population was about 819 thousand farms with the area of agricultural land above 1 ha, and assumed sample size was about 18000.

* 1. **Sampling fr**

The sampling frame was based on the results of “June agricultural survey” conducted in June and in July 2015 as of 1 June 2015 (reference day). As additional condition we selected only those farms which showed greater than zero sown area of cereals or sown area of rape.

For each farm in the frame the following information was recorded:

* ID of a farm,
* address data,
* agricultural land area of the farm,
* sown area of cereals,
* sown area of rape.
  1. **Sampling scheme**

In order to draw the sample, a stratified sampling scheme was used, with strata from “June agricultural survey”. To increase the efficiency of the sample, within the existing strata, an additional strata comprising units with the large values ​​of the selected attributes were created.

The boundaries of these strata were established using algorithm from the paper: Hidiroglou, M.A. (1986), “The construction of a self-representing stratum of large units in survey design,” *The American Statistician*, 40(1), 27-31. As a result of applying this algorithm 326 strata were established. Delimitation of additional strata on the basis of sown area is presented in table 1.

After determining the strata the problem of allocation of sample units among voivodships and strata was solved by means of numerical optimization method, so that the expected relative error for the selected variable does not exceed the set level for all voivodships. As a key variable sown area of cereals was used.

The overall sample contained 18000 farms, including 4095 from upper strata.

* 1. **Extrapolation method and precision evaluation**

The basic parameter estimated in this survey is the yield of a given crop. This parameter is a quotient of random variables, i.e.:

(1) 

where:

X – production of a given crop,

Y – sown area for a given crop.

Estimation value X for the w-th voivodship is counted according to formula:

(2)  ( i = 1, 2, ... , nwh; h = 1, 2, ... , 7)

where:

xwhi – value of *X*  variable in i-th farm (sampling unit) drawn from the h-th stratum in w-th voivodship,

Nwh –  number of sampling units in h-th stratum of w-th voivodship,

nwh - number of sampling units drawn for the sample from h-th stratum of w-th voivodship.

The sum of values of *Y* variable for the w-th voivodship is calculated analogically, and then the rw value is estimated according to the following formula:

(3) 

Estimation of sum of variables *X* and *Y* for Poland is constituted by the sum of the values estimated for voivodships, i.e.

(4) 

(5)  ( w = 1, 2, ... , 16)

(6) 

For selected important variables estimations were made (as precision measures) of variation coefficient related to yields, production and crops area. Calculation of precision involved formulas appropriate for stratified sampling scheme. Table 3 includes some estimated coefficients of variation (relative standard errors).

Table 1. Delimitation of additional strata in particular voivodships (in hectares) in the 2015 survey.

|  |  |  |  |
| --- | --- | --- | --- |
| **voivodship** | **sown area of cereals** | **sown area of wheat** | **sown area of rape** |
| 02 | 65,4 | 49,1 | 32,0 |
| 04 | 80,7 | 37,8 | 27,6 |
| 06 | 71,1 | 40,2 | 15,9 |
| 08 | 42,3 | 15,4 | 14,5 |
| 10 | 62,3 | 18,1 | 4,9 |
| 12 | 26,4 | 15,9 | 1,4 |
| 14 | 75,2 | 21,6 | 9,6 |
| 16 | 60,7 | 36,5 | 24,0 |
| 18 | 27,5 | 17,3 | 4,9 |
| 20 | 61,5 | 10,0 | 2,2 |
| 22 | 66,6 | 32,5 | 21,6 |
| 24 | 33,5 | 14,1 | 4,9 |
| 26 | 33,7 | 16,2 | 3,3 |
| 28 | 74,1 | 37,0 | 25,3 |
| 30 | 104,8 | 28,4 | 17,3 |
| 32 | 69,5 | 34,6 | 34,5 |

1. **Sample survey of yields of selected agricultural crops**

The aim of the survey was to collect data on yields and sown area as well as production of selected crops, i.e. potatoes, sugar beets, edible pulses, maze, as well as on area of meadow. The surveyed population was about 774 thousand farms, and the sample size – 18  thousands farms.

**2.1 Sampling frame**

The sampling frame was based on the results of “June agricultural survey” conducted in June and in July 2015 as of 1 June 2015 (reference day). As additional condition we selected only those farms which showed greater than zero at least one of the following variables: sown area of meadow, sown area of edible pulses, sown area of maize, sown area of potatoes or sugar beet.

For each farm the following characteristics were recorded:

* address data,
* agricultural land area of the farm,
* area of meadow,
* area of potatoes or sugar beet,
* area of maize,
* area of edible pulses.

**2.2 Sampling scheme**

In order to draw the sample, a stratified sampling scheme was used, using strata from farm structure survey. To increase the efficiency of the sample, within the existing strata, an additional strata comprising units with the large values ​​of the selected attributes were created.

The boundaries of these strata were established using algorithm from the paper: Hidiroglou, M.A. (1986), “The construction of a self-representing stratum of large units in survey design,” *The American Statistician*, 40(1), 27-31. As a result of applying this algorithm 337 strata were established. Delimitation of additional strata on the basis of sown area is presented in table 2.

After determining the strata the problem of allocation of sample units among voivodships and strata was solved by means of numerical optimization method, so that the expected relative error for the selected variable does not exceed the set level for all voivodships. As a key variable area of meadow was used.

The overall sample contained 18000 farms, including 4069 from upper strata.

# Table 2. Upper boundaries for additional strata (in hectares) in the survey of yields of selected agricultural crops in 2015.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **voivodship** | **area of meadow** | **area of edible pulses** | **area of potatoes**  **or sugar beet** | **area of maize** |
| 02 | 23,8 | 1,0 | 7,3 | 18,9 |
| 04 | 16,8 | 1,0 | 7,3 | 18,9 |
| 06 | 27,1 | 1,0 | 7,3 | 18,9 |
| 08 | 28,9 | 1,0 | 7,3 | 18,9 |
| 10 | 22,0 | 1,0 | 7,3 | 18,9 |
| 12 | 24,6 | 1,0 | 7,3 | 18,9 |
| 14 | 50,9 | 1,0 | 7,3 | 18,9 |
| 16 | 13,2 | 1,0 | 7,3 | 18,9 |
| 18 | 23,5 | 1,0 | 7,3 | 18,9 |
| 20 | 60,5 | 1,0 | 7,3 | 18,9 |
| 22 | 28,6 | 1,0 | 7,3 | 18,9 |
| 24 | 15,7 | 1,0 | 7,3 | 18,9 |
| 26 | 17,7 | 1,0 | 7,3 | 18,9 |
| 28 | 62,2 | 1,0 | 7,3 | 18,9 |
| 30 | 33,0 | 1,0 | 7,3 | 18,9 |
| 32 | 38,0 | 1,0 | 7,3 | 18,9 |

**2.3 Extrapolation method and precision evaluation**

The results of survey were generalized the same manner as results of yields of cereals survey. Analogous way was used for accuracy of the results assessment.

Table 3. Relative standard error for yields in Poland

|  |  |  |
| --- | --- | --- |
| no. of  characteristic | name of the characteristic | Relative standard  error  cv(r) in % |
| 1 | winter wheat | 0,5 |
| 2 | spring wheat | 1,4 |
| 3 | rye | 1,1 |
| 4 | winter barley | 1,5 |
| 5 | spring barley | 0,9 |
| 6 | oats | 1,2 |
| 7 | winter triticale | 0,7 |
| 8 | spring triticale | 1,8 |
| 9 | winter cereal mixed | 2,0 |
| 10 | spring cereal mixed | 0,9 |
| 11 | maize for grain | 0,6 |
| 12 | potatoes | 1,3 |