

Анализ данных

Семинар 2

Задачи

2. В городе проживает 47500 человек. Для определения средней заработной платы горожан был произведен опрос 1970 человек об их уровне дохода в долларах. Получены следующие значения:

$$\sum_{i=1}^{1970} y_i = 559037; \quad \sum_{i=1}^{1970} y_i^2 = 163533026.$$

Найдите доверительные границы, в которых, с вероятностью 95%, находится среднее значение заработной платы горожан в долларах.

Из 483 двухгодичных колледжей была отобрана простая случайная выборка размером 60 единиц с целью определения среднего числа студентов в каждом колледже. Получены следующие значения:

$$\sum_{i=1}^{60} y_i = 81\,796; \quad \sum_{i=1}^{60} y_i^2 = 112\,805\,447.$$

Найдите доверительные границы для среднего количества студентов на один колледж. Доверительная вероятность равна 95%.

2.4. Простая случайная выборка объемом в 30 домохозяйств была отобрана из городского района, содержащего 14 848 домохозяйств. Числа лиц для каждого домохозяйства в выборке следующие:

5, 6, 3, 3, 2, 3, 3, 3, 4, 4, 3, 2, 7, 4, 3, 5, 4, 4, 3, 3, 4, 3, 3, 1, 2, 4, 3, 4, 2, 4.
Оцените общее число людей в районе и вычислите вероятность того, что эта оценка находится в пределах $\pm 10\%$ истинного значения.

2.5. При изучении возможности применения выборочного метода для учета товаров на складе произвели подсчет стоимости товаров на каждой из 36 полок складского помещения. Получены следующие округленные до целых долларов значения:

29, 38, 42, 44, 45, 47, 51, 53, 53, 54, 56, 56, 56, 58, 58, 59, 60, 60,
60, 60, 61, 61, 61, 62, 64, 65, 65, 67, 67, 68, 69, 71, 74, 77, 82, 85.

Оценка суммарного значения, сделанная по выборке, должна быть точной в пределах 200 долл., за исключением одного случая из 20. Эксперт считает, что для этого достаточно взять простую случайную выборку объемом в 12 полок. Согласны ли вы с ним?

$$\Sigma y = 2138; \quad \Sigma y^2 = 131\,682.$$

2.6. После того как выборка, приведенная в табл. 2.1 (с. 42), была получена, подсчитали число полностью заполненных листов (каждый с 42 подписями), оно оказалось равным 326. Пользуясь этими сведениями, получите улучшенную оценку общего числа подписей и найдите стандартную ошибку вашей оценки.

В микрорайоне проживает 14 000 семей. Был произведен опрос на предмет наличия детей в этих семьях и получены следующие результаты:

| | | | | | | |
|------------------|-----|-----|-----|----|----|----|
| Количество детей | 0 | 1 | 2 | 3 | 4 | 5 |
| Количество семей | 100 | 200 | 120 | 40 | 20 | 20 |

Используя данные центра социального обеспечения района о зарегистрированном количестве многодетных семей, найдите доверительные границы, в которых, с вероятностью 95%, находится значение общего количества детей в микрорайоне.

Данные центра социального обеспечения:

| | | | |
|--------------------------------------|-----|-----|---|
| Количество детей | 4 | 5 | 6 |
| Количество семей (в общем, в районе) | 501 | 497 | 2 |

2.14. Два зубных врача *A* и *B* проводят обследование состояния зубов

у 200 детей некоторой деревни. Д-р *A* отобрал простую случайную выборку, включающую 20 детей, и подсчитал число испорченных зубов у каждого ребенка, получив следующие результаты:

| | | | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|---|---|----|
| Число испорченных зубов на одного ребенка | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Число детей | 8 | 4 | 2 | 2 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |

Д-р *B*, пользуясь теми же стоматологическими методами, осмотрел всех 200 детей, просто отмечая тех, у кого не было испорченных зубов. Число таких детей оказалось равным 60.

Оцените суммарное число испорченных зубов у детей этой деревни: (а) пользуясь только результатами *A*; (б) пользуясь результатами *A* и *B*; (в) будут ли