Прежде всего, хотел бы отметить, что все, что здесь написано, появилось как результат обсуждений с В.Н.Мягковым и Н.П.Бариновым.

А. Откуда берутся интервалы.

1. Выборка. По имеющейся выборке легко находятся минимальное, максимальное, среднее арифметическое, среднее геометрическое значения.

У выборки есть аналог плотности, которую надо как то описать и измерить. Плотность – понимается как количество точек на единицу длины.

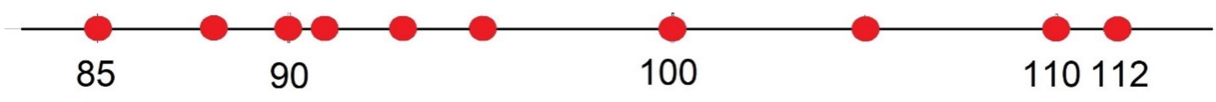


Рисунок 1. Пример расположения точек на положительной полуоси из примера В.Н.Мягкова.

Работать с плотностью выборки пактически невозможно. Обычно стоят гистограмму, а к ней подбирают непрерывную линию, описывающую плотность.

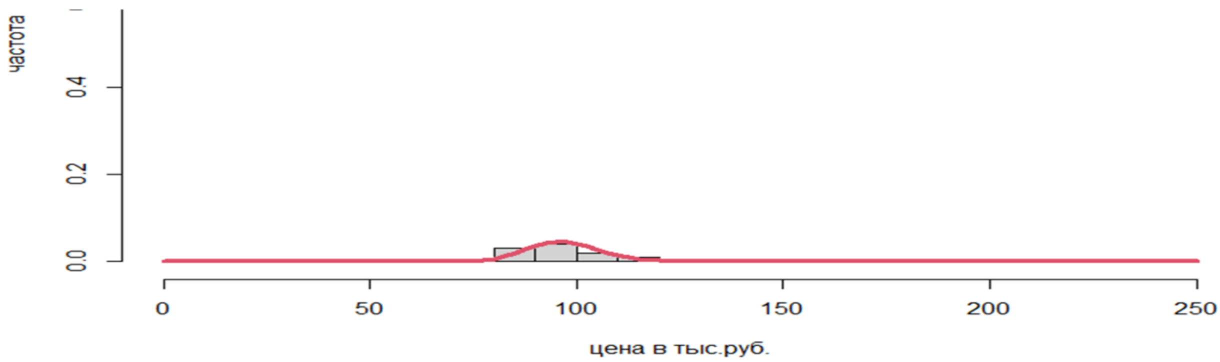


Рисунок 2. Пример гистограммы для выборки из примера В.Н.Мягкова и подобранная к ней непрерывная плотность.

2. Генеральная совокупность. Генеральная совокупность в задачах оценки – положительная числовая полуось с нулем. Плотность есть, но неизвестна. Минимум равен нулю, максимум не определен (+∞). Такой размах с точки зрения задач оценки неприемлем.

Даже если кажется, что найти дополнительльные объекты аналоги сверх имеющейся выборки невозможно, это не значит что выборка и есть генеральная совокупность. Модельная генеральная совокупность появляется немедленно, при совершении операций деления (среднее арифметическое) и взятия корня (возведение в дробную степень при расчете среднего геометрического), поскольку результаты этих операций считаются допустимыми и не обязаны совпадать с точками из выборки.

Размах от 0 до + ∞, естественно, для задач оценки не годится, поэтому возникает необходимость «обрезать» диапазон 0 до + ∞ вблизи наибольшей плотности. Для этого служат доверительные интервалы. 95% доверительного интервала вполне достаточно для целей оценки.

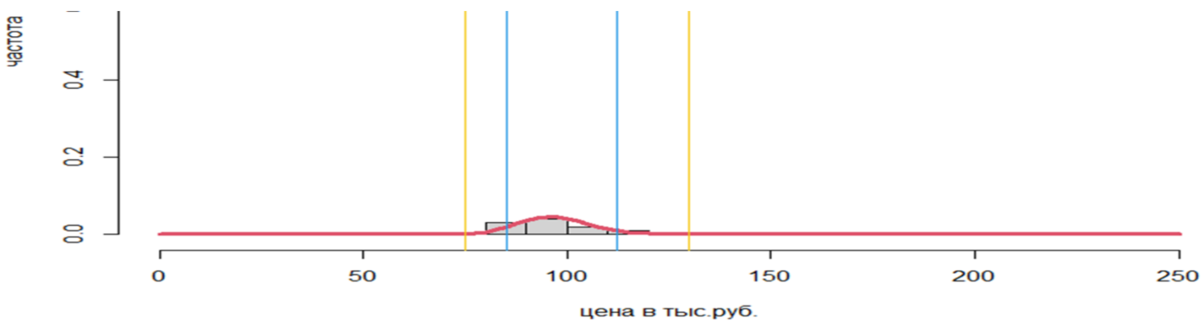


Рисунок 3. Синии линии – интервал от минимальной точки выборки до максимальной. Желтые линии – интервал охватывающий область «сгущения» плотности для данного примера.

Так как выборки маленькие, гистограммы получаются наглядными при достаточно широких интервалах разбиения. Поэтому интервала между синими линиями не достаточно (поскольку допускаются значения между минимумом и максимумом, то должны допускаться и значения за пределами минимума и максимума, хотя и в значительно меньшей степени).

3. В распоряжении оценщика имеется **ограниченная по объему выборка**. Он может воспользоваться всей выборкой, а может только её частью. В таком случае оценки по всем возможным комбинациям элементов выборки будут отличаться. Далее, для определенности будем рассматривать оценку по среднему значению. Получаемые таким образом средние значения сами образуют некоторую выборку, которая может быть использована для оценки среднего средних.

Б. Доверительный интервал при известных выборочном среднем и выборочной дисперсии устанавливается как

$$\overline{X_{выб}} - t_{1-\frac{\alpha}{2},n-1} \times \frac{S}{\sqrt{N}} < \mu < \overline{X_{выб}} + t_{1-\frac{\alpha}{2},n-1} \times \frac{S}{\sqrt{N}} \qquad (1)$$

где α - доверительный уровень (в частности α =0,05), $\overline{X_{выб}}$ - выборочное среднее, S - выборочное стандартное отклонение, N - объем выборки, $t_{1-\frac{\alpha}{2},n-1}$ - квантиль распределения Стьюдента.

Размах такого интервала зависит от S , N, $t_{1-\frac{\alpha}{2},n-1}$

с ростом S и $t_{1-\frac{\alpha}{2},n-1}$ размах растет, с ростом N - размах уменьшается.

В. Определение границ интервалов.

1. Выборка $\vec{X} = (x_1, x_2, x_3, \dots, x_N)$. Оценка по среднему $\overline{X_{выб}} = \frac{1}{N}\sum_{i=1}^{n} x_i$. Интервал $[\min_i(x_i); \max_i(x_i)]$. Такая интервальная оценка неудобна из-за размаха и нестабильности границ.

2. 95% доверительный интервал для среднего генеральной совокупности

$$\overline{X_{выб}} - t_{0,975,n-1} \times \frac{S}{\sqrt{N}} < \mu < \overline{X_{выб}} + t_{0,975,n-1} \times \frac{S}{\sqrt{N}} \qquad (2)$$

Такой доверительный интервал тоже велик (за исключением большого N) и недостаточно стабилен.

3. Оценщик находится в состоянии ограниченного набора данных выборки $\vec{X} = (x_1, x_2, x_3, \dots, x_N)$. Будем считать, что оценщик может выбрать не менее трех объектов сравнения и строит оценку по среднему значению. Всего таких вариантов $C_N^k = \frac{N!}{k!(N-k)!}$, k=3,2,4,…,N.

Все возможные средние образуют новую выборку, для неё можно определить оценку среднего средних и интервал (95%). Формула (2) примет вид

$$\overline{X_{средних}} - t_{0,975,n-1} \times \frac{S_{средних}}{\sqrt{M}} < \mu < \overline{X_{средних}} + t_{0,975,n-1} \times \frac{S_{средних}}{\sqrt{M}} \qquad (3)$$

где $\overline{X_{средних}}$ - выборочное среднее средних, $S_{средних}$- выборочное стандартное отклонение средних, M=$C_N^k = \frac{N!}{k!(N-k)!}$.

M значительно больше N.

Например, при N=5 и k=3, M=10,

при N=6 и k=3, M=20,

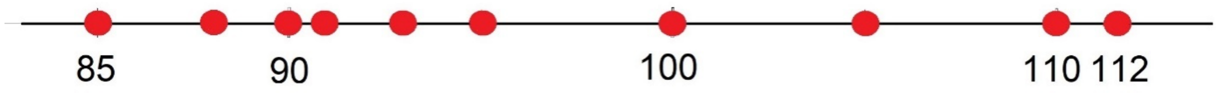
при N=7 и k=3, M=30.

Относительно среднего средних, известно, что его распределение сходится к нормальному со скоростью порядка $\sqrt{M}$ (см. формулу 3). Такая оценка будет достаточно точной для целей оценки.

Г. Покажем это на примере.

Используются данные из примера Владислава Николаевича Мягкова.

Есть выборка (рисунок 1) из 10 аналогов с ценами (1 кв.м. в тыс.руб.):



85 88 90 91 93 95 100 105 110 112

Среднее значение 96,9.

Рассмотрим все возможные оценки средних значений по выборкам по 3,4,5,6,7,8,9,10 элементов из исходной. Их всего будет

$$C_{10}^{k} = \frac{10!}{(10-k)!k!}$$

k=3 $C_{10}^{3} = \frac{10!}{(10-3)!3!} = C_{10}^{7} = 120$ (тоже при k=7)

k=4 $C_{10}^{4} = \frac{10!}{(10-4)!4!} = C_{10}^{6} = 210$ (тоже при k=6)

k=5 $C_{10}^{5} = \frac{10!}{(10-5)!5!} = C_{10}^{5} = 252$

k=8 $C_{10}^{8} = \frac{10!}{(10-8)!8!} = C_{10}^{2} = 45$

k=9 $C_{10}^{9} = \frac{10!}{(10-9)!9!} = C_{10}^{1} = 10$

всего 967 выборок. Т.е. из 10 объектов сравнения можно получить 967 выборок и 967 оценок среднего значения генеральной совокупности. Это все, что может сделать оценщик на этих данных. Естественно, результаты определения средних будут значительно отличаться при разных объемах выборок и комбинаций отобранных точек.

Построим диаграмму: на горизонтальной оси средние значения, на вертикальной оси стандартные отклонения.

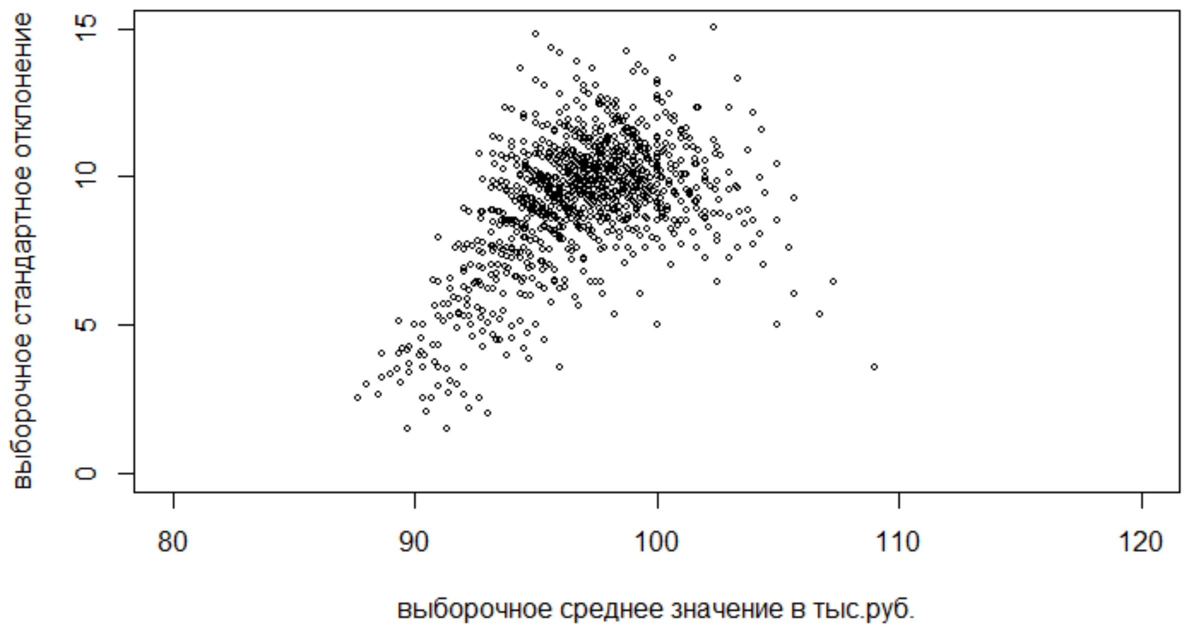


Рисунок 4. По горизонтальной оси средние значения выборок, на вертикальной оси стандартные отклонения соответствующих выборок.

Рисунок 4 показывает ожидаемый результат: распределение средних нормально, а распределение стандартных отклонений – распределение хи (не квадрат).

Для чего это показано. Горизонтальная координата каждой черной точки – один из вариантов оценки РС по среднему значению. Для каждой черной точки, являющейся оценкой РС (по среднему значению) можно построить интервал, в котором может находится РС (доверительный интервал - предлагаю его всегда фиксировать как 95% и не обсуждать).

Доверительный интервал строится по формуле (1) :

$$\overline{X_{выб}} - t_{1-\frac{\alpha}{2},n-1} \times \frac{S}{\sqrt{N}} < \mu < \overline{X_{выб}} + t_{1-\frac{\alpha}{2},n-1} \times \frac{S}{\sqrt{N}}$$

На рисунке 5 показано как сильно могут отличаться интервальные оценки для разных комбинаций исходных данных. Зеленые точки – правые границы интервальных оценок, голубые точки – левые границы интервальных оценок. Хорошо видно, что самые малые интервальные оценки – внизу, при существенно разных малых выборках и сделанных по ним оценках РС.

Наибольшая концентрация средних значений - в середине рисунка, с умеренными интервалами, но они (интервалы) все равно остаются такими, что для целей оценки их размах будет велик (+/-15%).

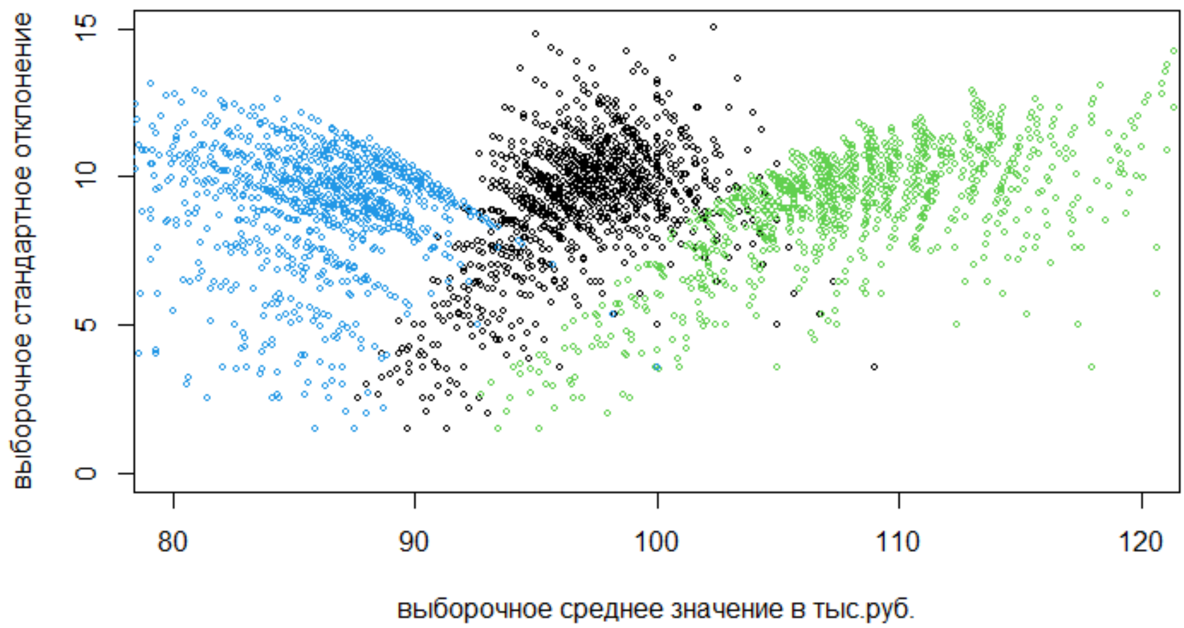


Рисунок 5. По горизонтальной оси средние значения выборок, на вертикальной оси стандартные отклонения соответствующих выборок Зеленые точки – правые границы интервальных оценок для истинного среднего, голубые точки – левые границы интервальных оценок для истинного среднего.

Однако известно, что

- среднее средних сходится к истинному значению матожидания со скоростью, порядка $\sqrt{n}$ (границы интервала сжимаются со скоростью $\frac{1}{\sqrt{n}}$)
- распределение средних с ростом n стремится к нормальному закону распределения.

Это означает, что мы можем в качестве интервальной оценки взять интервал 2σ (это чуть больше 95%) к среднему средних. НА ЭТИХ ДАННЫХ ТОЧНЕЕ УЖЕ СДЕЛАТЬ НИЧЕГО НЕЛЬЗЯ.

Это дает простой и понятный критерий сравнения двух конфликтующих оценок.

Д. Как сравнить точечные и интервальные оценки двух разных отчетов об оценке одного и того же объекта.

Значимые расхождения и незначимые расхождения.

Как признать существеность/несущественность расхождений двух оценок?

1. Предполагается, что все важные условия, отраженные в документе, присланном Е.И.Нейманом проверены и выполнены. Есть две оценки,

сделанные двумя оценщиками по двум разным выборкам. Объекты сравнения проверены, предварительная обработка данных выполнена, существенных ошибок не найдено, а оценки разные.

2. Оба оценщика получили точечную оценку РС, интервальную оценку **РС** (не цен), каждый по своей выборке **цен.**

Оба оценщика в состоянии указать

- размах выборки, минимальное, максимальное значение **выборки**

- точечную оценку **РС**, и точность её оценки по **данной, имеющейся в распоряжении** выборке

- интервальную оценку **РС,** которая **не является** интервалом цен в выборке

Рассмотрим две выборки по 5 элементов. Пусть два оценщика оценивали один и тот же объект.

Первый взял объекты сравнения с ценами: 85 88 90 91 93

Второй взял объекты сравнения с ценами: 95 100 105 110 112

И тот и другой взял в качестве оценки РС выборочное среднее.

Первый получил 89,4 тыс.руб. за 1 кв.м.

Второй получил 104,4 тыс.руб. за 1 кв.м.

Существенны эти расхождения или не существенны (независимо от того какие интервалы назвали оценщики и как они их считали)?

Рассмотрим все подвыборки по 3,4,5 элементов из выборок каждого из оценщиков.

На рисунке 6 показаны точки (выборочное среднее, выборочное стандартное отклонение) для первого оценщика и границы соответствующих доверительных интервалов.

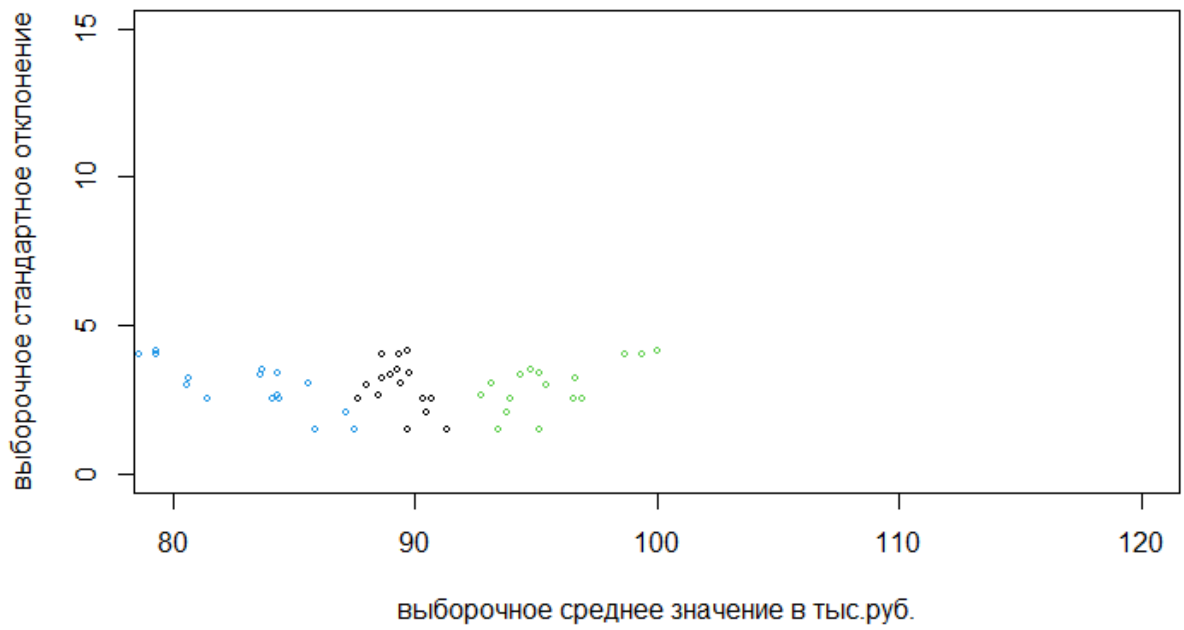


Рисунок 6. Все возможные средние (и соответствующие им границы интервала истинного среднего) для выборки первого оценщика.

Всего на рисунке 6 показаны 16 черных точек.

У них среднее средних: **89.4**

Стандартное отклонение средних от среднего средних **0.9907909**

Квантиль распределения Стьюдента **qt**(0.975,15)= **2.13145**

$$t_{1-\frac{0.05}{2},16-1} \times \frac{0.99}{\sqrt{16}} \sim 2.13 \times 0.25 \sim \mathbf{0,527}$$

1) Точечная оценка РС **89.4,** точность определения среднего средних $\mathbf{88.873 < \mu < 89.927}$

2) Интервал **РС**

$$\mathbf{89.4 - 2 \times 0.99 < \mu <} 89.\mathbf{4 + 2 \times 0.99}$$

окончательно интервал РС

$$\mathbf{87.42 < \mu <} 91.\mathbf{38}$$

На рисунке 7 показаны точки (выборочное среднее, выборочное стандартное отклонение) для второго оценщика и границы соответствующих доверительных интервалов.

.

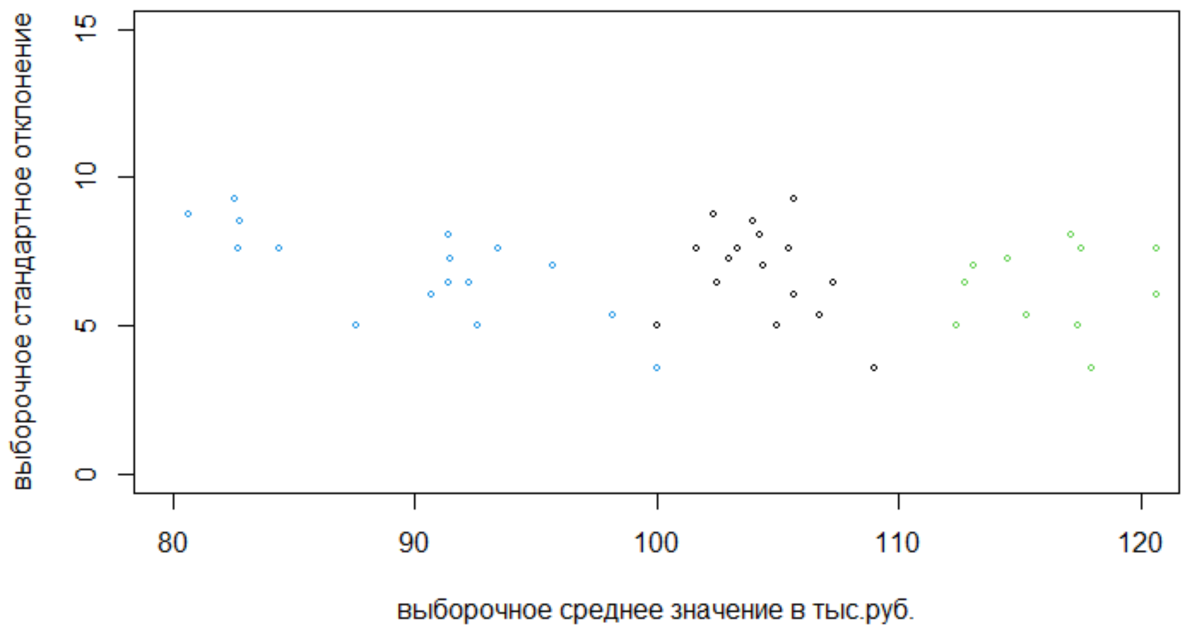


Рисунок 7. Все возможные средние (и соответствующие им границы интервала истинного среднего) для выборки второго оценщика.

Всего на рисунке 7 показаны 16 черных точек.

У них среднее средних: **104.4**

Стандартное отклонение средних от среднего средних: **2.281203**

Квантиль распределения Стьюдента **qt**(0.975,15)= **2.13145**

$$t_{1-\frac{0.05}{2},16-1} \times \frac{2.281203}{\sqrt{16}} \sim 2.13 \times 0.57 \sim 1{,}215$$

1) Точечная оценка РС **104.4**, точность определения среднего средних $\mathbf{103.185} < \boldsymbol{\mu} < 105.\mathbf{615}$

2) Интервал **РС**

$$\mathbf{104.4 - 2 \times 2.28} < \boldsymbol{\mu} < 104.\mathbf{4 + 2 \times 2.28}$$

окончательно интервал РС

$$\mathbf{99.84 < \boldsymbol{\mu} < 108.96}$$

ОБЪЕДИНЯЕМ ВЫБОРКИ ДВУХ ОЦЕНЩИКОВ В ОДНУ

При сравнении результатов оценки мы бы выполнили аналогичные расчеты, рассматривая все возможные подвыборки из 3,4,5,6,7,8,9,10 элементов из 10.

На рисунке 8 еще раз (такой же как рисунок 5) показаны точки выборочных средних и выборочных стандартных отклонений и соответствующие им интервальные оценки для объединенной выборки (всего 967 точек – выборочных средних.

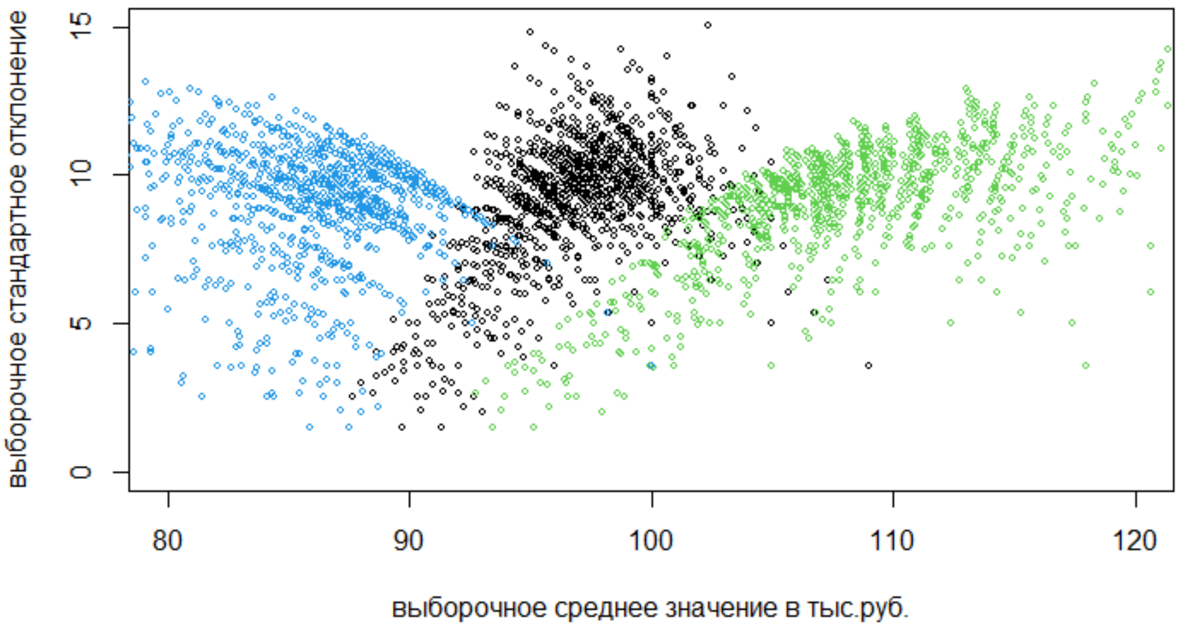


Рисунок 8. (такой же как рисунок 5)

Что получится:

Среднее средних: **96.89**

Стандартное отклонение средних от среднего средних: **3.104731**

Квантиль распределения Стьюдента **qt**(0.975,965)= **1.962425**

$$t_{1-\frac{0.05}{2},966-1} \times \frac{3.104731}{\sqrt{966}} \sim 1.96 \times 0.1 \sim 0{,}196$$

1) Точечная оценка РС **96.89**, точность определения среднего средних $\mathbf{96.694 < \mu < 97.086}$

2) Интервал **РС**

$$\mathbf{96.89 - 2 \times 3.1 < \mu <} \, 96.89 + \mathbf{2 \times 3.1}$$

Окончательно интервал РС на объединенной выборке

$$\mathbf{90.69 < \mu < 103.09}$$

При этом точечная оценка первого оценщика 89,4

точечная оценка второго оценщика 104,4

В интервал $90.69 < \mu < 103.09$ НЕ ПОПАЛИ ТОЧЕЧНЫЕ ОЦЕНКИ НИ ПЕРВОГО **(89,4)** НИ ВТОРОГО ОЦЕНЩИКА **(104,4)**

ОЦЕНКИ РАЗЛИЧАЮТСЯ СУЩЕСТВЕННО

При оценке существенности несущественности расхождения использованы результаты обеих оценщиков. Ни одного наблюдения не удалено. Естественно, перед таким исследованием, каждый из объектов оценки должен быть рассмотрен по всем возможным ценообразующим факторам на предмет подходит/не подходит как объект сравнения (т.е. выполнены условия Е.И.Неймана). Рассматриваемая процедура строится в предположении, что это уже сделано и все объекты сравнения приемлемы.

Теперь рассмотрим другой случай.

Первый оценщик взял объекты сравнения с ценами: 85 95 90 105 93

Второй оценщик взял объекты сравнения с ценами: 88 100 91 110 112

И тот и другой взяли в качестве оценки РС выборочное среднее.

Первый получил 93,6 тыс.руб. за 1 кв.м.

Второй получил 100,2 тыс.руб. за 1 кв.м.

Существенны эти расхождения или не существенны (независимо от того какие интервалы назвали оценщики и как они их считали)?

Рассмотрим все подвыборки по 3,4,5 элементов из выборок каждого из оценщиков.

На рисунке 9 показаны точки (выборочные средние, выборочные стандартные отклонения) для первого оценщика и границы доверительных интервалов для них.

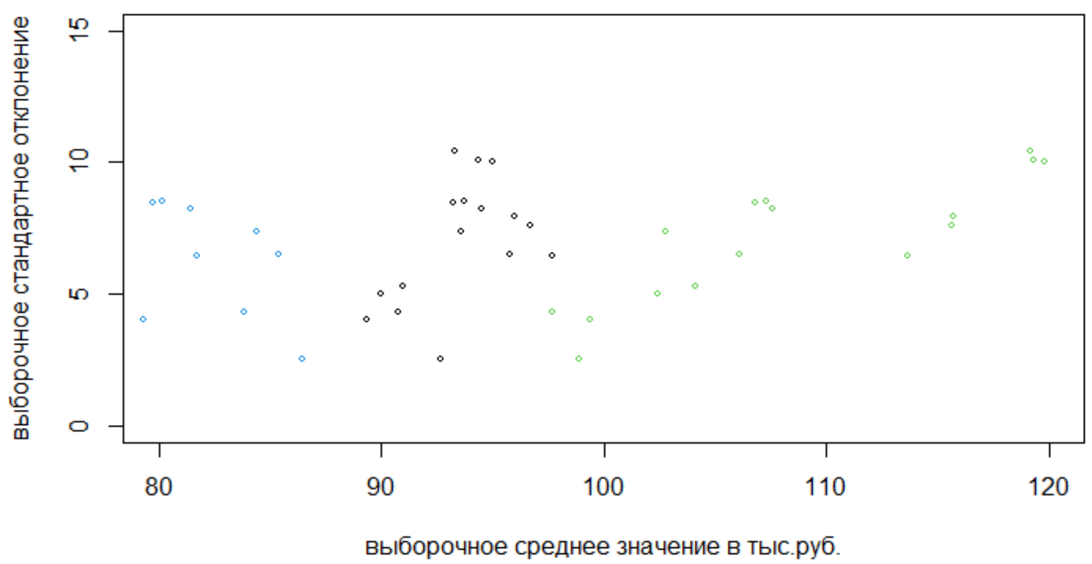


Рисунок 9. Все возможные средние (и соответствующие им границы интервала истинного среднего) для выборки первого оценщика.

Всего на рисунке 9 показаны 16 точек.

У них среднее средних: **93.6**

Стандартное отклонение средних от среднего средних: **2.405087**

Квантиль распределения Стьюдента **qt**(0.975,15)= **2.13145**

$$t_{1-\frac{0.05}{2},16-1} \times \frac{2.405087}{\sqrt{16}} \sim 2.13 \times 0.6 \sim 1{,}278$$

1) Точечная оценка РС **93.6**, точность определения среднего средних $92.\mathbf{32} < \boldsymbol{\mu} < 94.\mathbf{88}$

2) Интервал **РС**

$$\mathbf{93.6 - 2 \times 2.4} < \boldsymbol{\mu} < 93.\mathbf{6 + 2 \times 2.4}$$

Т.е**.**

$$\mathbf{88.8} < \boldsymbol{\mu} < 98.\mathbf{4}$$

На рисунке 10 показаны точки (выборочные средние, выборочные стандартные отклонения) для второго оценщика и границы доверительных интервалов для них.

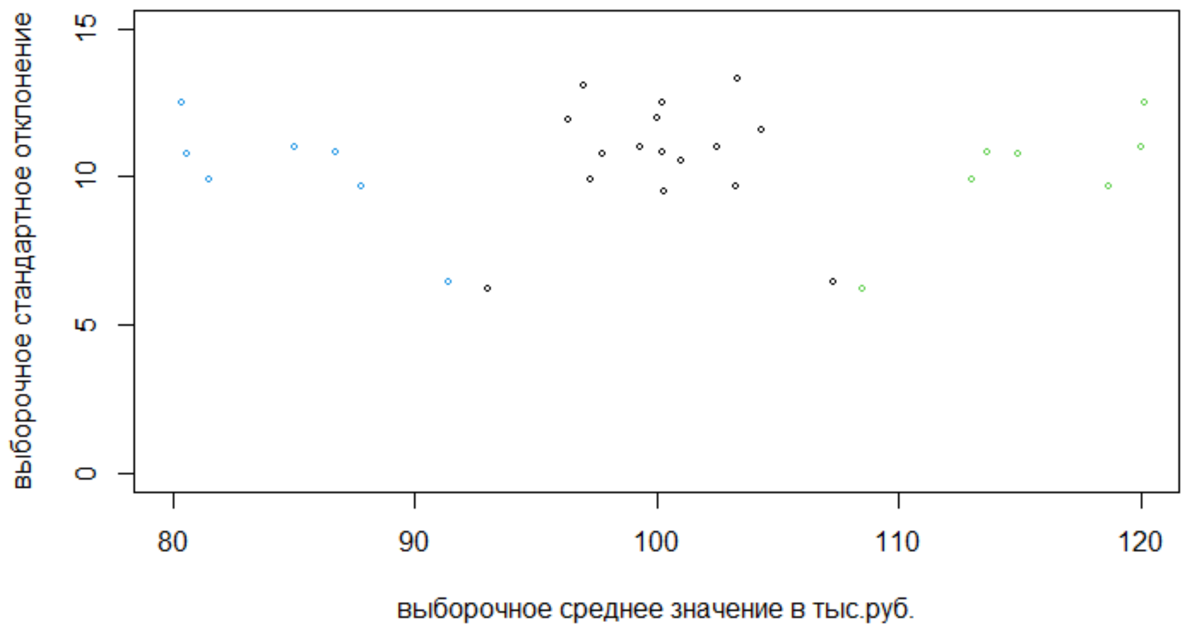


Рисунок 10. Все возможные средние (и соответствующие им границы интервала истинного среднего) для выборки второго оценщика.

Всего на рисунке 10 показаны 16 точек.

У них среднее средних: **100.2**

Стандартное отклонение средних от среднего средних: **3.517259**

Квантиль распределения Стьюдента **qt**(0.975,15)= **2.13145**

$$t_{1-\frac{0.05}{2},16-1} \times \frac{3.517259}{\sqrt{16}} \sim 2.13 \times 0.88 \sim 1{,}874$$

1) Точечная оценка РС **100.2**, точность определения среднего средних $\mathbf{98.326} < \boldsymbol{\mu} < 102.\mathbf{074}$

2) Интервал **РС**

$$\mathbf{100.2} - \mathbf{2} \times \mathbf{3.51} < \boldsymbol{\mu} < 100.2 + \mathbf{2} \times \mathbf{3.51}$$

Т.е.

$$\mathbf{93.18} < \boldsymbol{\mu} < 107.\mathbf{22}$$

А на объединенной выборке было получено

Точечная оценка РС **96.89**, Интервал **РС** $\mathbf{90.69} < \boldsymbol{\mu} < 103.\mathbf{09}$

Точечная оценка первого оценщика **93,6** попала в интервал объединенной выборки (интервальная у неё $\mathbf{93.18} < \boldsymbol{\mu} < 107.\mathbf{22}$)

Точечная оценка второго оценщика **100,2** тоже попала в интервал объединенной выборки (интервальная у неё $\mathbf{93.18} < \boldsymbol{\mu} < 107.\mathbf{22}$)

Оценки отличаются несущественно.

Здесь важными являются два момента:

- оценщики получили несущественное различие в оценках, нельзя говорить о том, что один из них прав, а другой нет – правы оба

- оценка на объединенной выборке является уточнением к результатам двух оценщиков и может быть выбрана как компромиссная.

Три комментария.

1. Я отдаю себе отчет в том, что приведенная процедура – это некоторая подмена устоявшихся методов статистики по проверке гипотез. Но так как я уже неоднократно слышал и имел возможность убедится, что оценочное сообщество не очень склонно опираться на устоявшиеся статистические методы, мне такой подход представляется более простым для понимания

2. На первый взгляд, может показаться, что был показан трудоемкий процесс (967 выборок проверили!). Это не так. Подобные расчеты в статистических пакетах занимают 1-2 секунды.

Больше времени уходит на описание в Ворде. Т.е. эксперт, пользующийся таким методом потратит секунды на расчет, и 1-2 часа на описание результата. И на то и на другое просто будут созданы шаблоны.

3. Работа оценщика со статистическими пакетами – это новый вызов времени.