



Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Башкирский государственный аграрный университет»

Кафедра математики

Математика

Методические указания к практическим занятиям
и самостоятельной работе

**«Точечные и интервальные оценки
параметров генеральной совокупности»**

Для всех направлений бакалавриата

Уфа 2024

00УДК 51

ББК 22.1

М 33

Рекомендовано к изданию заседанием кафедры математики (протокол № 7/1 от «21 марта» 2024 года)

Составители: доценты Лукманов Р.Л., Маннанов М.М.

Ответственный за выпуск: и.о. зав. кафедрой математики
канд.психол.наук. Дик Е.Н.

Точечные оценки

Пусть многократным повторением случайного эксперимента получена выборка X_i , $i = 1, \dots, n$. Известно, что несмещенной и состоятельной оценкой математического ожидания исследуемой случайной величины X является выборочная средняя

$$x_B = \frac{1}{n} \sum_{i=1}^n X_i. \quad (1)$$

В Mathcad есть встроенная процедура $mean(x)$ для вычисления выборочной средней.

В качестве точечной оценки дисперсии используют либо выборочную дисперсию

$$D_B = \frac{1}{n} \sum_{i=1}^n (X_i - x_B)^2, \quad (2)$$

либо исправленную выборочную дисперсию

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - x_B)^2. \quad (3)$$

Сравнивая формулы (2) и (3), видим, что D_B и S^2 будут существенно отличаться лишь при малых объемах выборки n . Отличие оценок (2) и (3) в том, что D_B является смещенной оценкой дисперсии D исследуемой случайной величины X , а S^2 - несмещенной, т.е.

$$M D_B \neq D, \text{ а } M S^2 = D.$$

По этой причине при малых объемах выборки в качестве оценки дисперсии D нужно использовать исправленную выборочную дисперсию S^2 . В Mathcad D_B вычисляется с помощью процедуры $var(x)$, а S^2 - $Var(x)$.

Соответственно среднее квадратическое отклонение σ оценивается либо с помощью выборочного среднего квадратического отклонения $\sigma_B = \sqrt{D_B}$, либо исправленного среднего квадратического отклонения $S = \sqrt{S^2}$. В Mathcad для вычисления σ_B используется встроенная процедура $stdev(x)$, а для S - $Stdev(x)$. Ниже приводится пример вычисления выборочных характеристик для предварительно разыгранной случайной величины с заданными числовыми характеристиками:

ORIGIN := 1 n := 200

X := rnorm(n, 3, 2)

$$x_{\hat{a}} := \frac{1}{n} \cdot \sum_{i=1}^n X_i$$

mean(X) = 2.836 $x_{\hat{a}} = 2.836$

$$D_{\hat{a}} := \frac{1}{n} \cdot \sum_{i=1}^n (X_i - x_{\hat{a}})^2$$

var(X) = 4.104 $D_{\hat{a}} = 4.104$

$$S2 := \frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - x_{\hat{a}})^2$$

Var(X) = 4.125 $S2 = 4.125$
 stdev(X) = 2.026 $\sqrt{D_{\hat{a}}} = 2.026$

Stdev(X) = 2.031 $\sqrt{S2} = 2.031$

Как видим, значения выборочных характеристик, подсчитанные непосредственно по формулам и с помощью встроенных процедур получились одинаковыми и близкими к соответствующим значениям для генеральной совокупности.

Задания для самостоятельного выполнения.

1) Разыграть нормально распределенную случайную величину заданного объема n и с заданными значениями математического ожидания a и среднего квадратического отклонения $\sigma > 0$ (значения n, a, σ задаются преподавателем). Вычислить выборочные характеристики x_B, D_B, S^2, σ_B и S непосредственно по формулам и с помощью встроенных процедур. Убедиться в состоятельности этих оценок, т.е. в том, что с увеличением n они приближаются к соответствующим параметрам исследуемой случайной величины.

2) Выполнить задание пункта 1) для распределений «хи-квадрат», Стьюдента и Фишера с заданными (преподавателем) числами степеней свободы n и m . При этом необходимо воспользоваться следующими формулами:

$$M\chi_n^2 = n, \quad D\chi_n^2 = 2n,$$

$$Mt_n = 0, Dt_n = \frac{n}{n-2} \text{ при } n > 2,$$

$$MF_{n,m} = \frac{n}{n-2} \text{ при } n > 2,$$

$$DF_{n,m} = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \text{ при } n > 4.$$

Интервальные оценки математического ожидания и дисперсии

Точечные оценки математического ожидания и дисперсии не являются достаточно информативными, поскольку нельзя сказать, насколько они близки или далеки от соответствующего параметра исследуемой случайной величины. Именно на этот вопрос и отвечают интервальные оценки. Они представляют собой так называемые доверительные интервалы $(a; b)$, покрывающие исследуемый параметр θ с некоторой вероятностью (надежностью) γ .

Чем уже доверительный интервал $(a; b)$, тем лучше, поскольку при этом мы захватываем исследуемый параметр с большей точностью.

Очень часто доверительный интервал выбирается симметричным относительно несмещенной точечной оценки θ^* , т.е. выбирается интервал вида $(\theta^* - \varepsilon, \theta^* + \varepsilon)$ такой, что

$$P(|\theta - \theta^*| < \varepsilon) = \gamma. \quad (5)$$

Число $\varepsilon > 0$ характеризует точность оценки: чем меньше ε , тем оценка θ^* точнее. Надежность γ выбирается заранее, её выбор зависит от конкретно решаемой задачи. Наиболее часто задают надёжность равную 0,95; 0,99 и 0,999.

Пусть выполняется (5). Заменяя неравенство $|\theta - \theta^*| < \varepsilon$ равносильным ему двойным неравенством $\theta^* - \varepsilon < \theta < \theta^* + \varepsilon$, имеем

$$P(\theta^* - \varepsilon < \theta < \theta^* + \varepsilon) = \gamma.$$

Это означает, что вероятность того, что интервал $(\theta^* - \varepsilon, \theta^* + \varepsilon)$ заключает в себе (покрывает) неизвестный параметр θ , равна γ .

Доверительным называют интервал $(\theta^* - \varepsilon, \theta^* + \varepsilon)$, который покрывает неизвестный параметр θ с заданной надежностью γ .

Заметим, что интервал $(a, b) = (\theta^* - \varepsilon, \theta^* + \varepsilon)$ имеет случайные концы, т.к. зависит от выборки.

При построении доверительных интервалов делается предположение о нормальности исследуемого распределения. В дальнейшем будем предполагать, что производится выборка объема n из генеральной совокупности, имеющей нормальное распределение с параметрами a и σ .

Доверительный интервал для математического ожидания при известной и неизвестной дисперсии

а) Если дисперсия σ^2 известна, то доверительный интервал для $a = M(X)$ находится из следующего равенства:

$$P\left(x_B - d_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < a < x_B + d_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha,$$

где $d_{1-\frac{\alpha}{2}}$ - квантиль уровня $1 - \frac{\alpha}{2}$ стандартного нормального распределения.

Ниже приводится пример построения доверительного интервала для математического ожидания в случае известной дисперсии:

`n := 100 σ := 2`

`X := rnorm(n, 5, σ)`

`Xâ := mean(X)`

`γ := 0.95 α := 1 - γ`

`ξ := qnorm(1 - α/2, 0, 1) * σ/√n ξ = 0.392`

`L := Xâ - ξ R := Xâ + ξ`

`L = 4.678 R = 5.462`

б) Если дисперсия σ^2 неизвестна, то доверительный интервал для $a = M(X)$ находится из равенства:

$$P\left(x_B - t_{1-\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} < a < x_B + t_{1-\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}\right) = 1 - \alpha,$$

где $t_{1-\frac{\alpha}{2}}(n-1)$ - квантиль уровня $1 - \frac{\alpha}{2}$ распределения Стьюдента с $n-1$ степенью свободы.

Приведем пример построения доверительного интервала для математического ожидания в случае неизвестной дисперсии:

$$n := 100 \quad \sigma := 2$$

$$X := \text{rnorm}(n, 5, \sigma)$$

$$X_{\hat{a}} := \text{mean}(X)$$

$$\gamma := 0.95$$

$$\alpha := 1 - \gamma$$

$$\xi := \text{qnorm}\left(1 - \frac{\alpha}{2}, 0, 1\right) \cdot \frac{\text{Stdev}(X)}{\sqrt{n}}$$

$$\xi = 0.386$$

$$L := X_{\hat{a}} - \xi$$

$$R := X_{\hat{a}} + \xi$$

$$L = 4.694$$

$$R = 5.466$$

Доверительный интервал для дисперсии при известном и неизвестном математическом ожидании

а) Если $M(X) = a$ известно, то доверительный интервал для дисперсии $\sigma^2 = D(X)$ находится из равенства:

$$P\left(\frac{n\sigma_*^2}{\chi_{1-\frac{\alpha}{2}}^2(n)} < \sigma^2 < \frac{n\sigma_*^2}{\chi_{\frac{\alpha}{2}}^2(n)}\right) = 1 - \alpha,$$

где $\sigma_*^2 = \frac{1}{n} \cdot \sum_{i=1}^n (X_i - a)^2$, а $\chi_{1-\frac{\alpha}{2}}^2(n)$ и $\chi_{\frac{\alpha}{2}}^2(n)$ - квантили распределения «хи - квадрат».

$$\text{ORIGIN} := 1$$

$$n := 100 \quad a := 5$$

$$X := \text{rnorm}(n, a, 2)$$

$$\gamma := 0.9$$

$$\alpha := 1 - \gamma$$

$$\sigma := \frac{1}{n} \cdot \sum_{i=1}^n (X_i - a)^2$$

$$L := \frac{\sigma \cdot n}{\text{qchisq}\left(1 - \frac{\alpha}{2}, n\right)}$$

$$R := \frac{\sigma \cdot n}{\text{qchisq}\left(\frac{\alpha}{2}, n\right)}$$

$$L = 3.105$$

$$R = 4.955$$

б) Если $a = M(X)$ неизвестно, то доверительный интервал для дисперсии σ^2 находится из равенства:

$$P\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) = 1 - \alpha,$$

где S^2 - исправленная выборочная дисперсия, а $\chi^2_{1-\frac{\alpha}{2}}(n-1)$ и $\chi^2_{\frac{\alpha}{2}}(n-1)$ квантили распределения «хи-квадрат» уровней $1 - \frac{\alpha}{2}$ и $\frac{\alpha}{2}$ соответственно с $n-1$ степенью свободы.

$$n := 100$$

$$X := \text{rnorm}(n, 5, 2)$$

$$\gamma := 0.9$$

$$\alpha := 1 - \gamma$$

$$L := \frac{\text{Var}(X) \cdot n}{\text{qchisq}\left(1 - \frac{\alpha}{2}, n\right)}$$

$$R := \frac{\text{Var}(X) \cdot n}{\text{qchisq}\left(\frac{\alpha}{2}, n\right)}$$

$$L = 3.226$$

$$R = 5.148$$

Задания для самостоятельного выполнения.

1. На основании 100 опытов было определено, что в среднем для производства детали требуется $\bar{t} = 5,5$ секунд. Сделав допущение, что время для производства детали распределено по нормальному закону, найти доверительный интервал для математического ожидания производства детали с надежностью 0,85, если известно, что среднее квадратическое отклонение σ равно 1,7 секундам.

2. На контрольных испытаниях 16 осветительных ламп были определены значения сроков их службы:
2992; 3029; 2998; 3018; 3028; 2994; 3008; 2999; 3013; 2999; 2996; 2986; 3011; 3025; 3004; 2968. Считая, что срок службы каждой лампы является нормально распределенной случайной величиной, определите:

а) доверительный интервал для математического ожидания с надежностью $\gamma = 0,9$;

б) вероятность, с которой можно утверждать, что абсолютная величина ошибки определения m не превысит 10 ч.

3. По 15 независимым равноточным измерениям были рассчитаны значения оценок математического ожидания и среднего квадратического отклонения максимальной скорости самолета $\bar{V} = 424,7 \text{ м/с}$ и $S = 7,7 \text{ м/с}$. Считая, что генеральная совокупность имеет нормальное распределение, определите:

а) доверительный интервал для среднего квадратического отклонения при доверительной вероятности 0,9;

б) вероятность того, что абсолютная величина случайной ошибки при определении σ_V по 15 измерениям не превзойдет 2 м/с .

4. По результатам 10 измерений емкости конденсатора прибором, получили следующие отклонения от номинального значения :

5,4; -13,9; -11; 7,2; -15,6; 29,2; 1,4; -0,3; 6,6; -9,9 (пикофарад).

Найдите 90%-ный доверительный интервал для дисперсии, предполагая, что генеральная совокупность имеет нормальное распределение.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кирьянов Д.В. Mathcad 13: книга для широкого круга пользователей.- «БХВ-Петербург», 2004.-С.590.
2. Колемаев В.А., Староверов О.В., Турундаевский В.Б. Теория вероятностей и математическая статистика: учеб. пособие для экономических специальностей вузов.- М.: Высш. шк., 1991.- С. 399.
3. Зарубина В.С., Крищенко А.П. Математическая статистика: - М.: МГТУ им. Н.Э. Баумана, 2001.- С. 421.
4. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика: учебник для специальности «Менеджмент».- М.: ИНФРА-М, 2000.- С. 300.
5. Калинина В.Н., Панкин В. Ф. Математическая статистика: учебник для средних специальных учебных заведений М.: «Высшая школа», 1998.- С. 335

