Урок 37 – 38.

Тема: Оценка параметров генеральной совокупности по ее выборке. Интервальная оценка .

Точечные оценки параметров распределения не дают информации о степени близости к соответствующему теоретическому параметру. Поэтому построение интервала, в котором с заданной степенью достоверности будет находиться оцениваемый параметр, является более информативным способом оценивания неизвестных параметров.

Интервальная оценка – это числовой интервал, который определяется двумя числами-границами интервала, содержащий неизвестный параметр генеральной совокупности.

Доверительный интервал – это интервал, в котором с той или иной заранее заданной вероятностью находится неизвестный параметр генеральной совокупности.

Доверительная вероятность p - это такая вероятность, что событие вероятности (1 - p) можно считать невозможным. $\alpha = 1 - p$ - это уровень значимости. (Обозначения могут быть любыми, часто обозначают наоборот). Обычно в качестве доверительных вероятностей используют вероятности, близкие к 1. Тогда событие, что интервал накроет характеристику, будет практически достоверным. Это $p \geq 0{,}95$, $p \geq 0{,}99$, $p \geq 0{,}999$.

Эти вероятности признаны достаточными для уверенного суждения о генеральных параметрах на основании известных выборочных показателей. Обычно указывают 95% доверительный интервал.

Для выборки малого объема (n< 30) нормально распределенного количественного признака x доверительный интервал может иметь вид:

$$\bar{x}_B - \delta(m_{\bar{x}})t \leq \mu \leq \bar{x}_B + \delta(m_{\bar{x}})t \ (p \geq 0{,}95),$$

где | μ - генеральное среднее; $\bar{x}_B$ - выборочное среднее; t - нормированный показатель распределения Стьюдента с (n - 1) степенями свободы, который определяется вероятностью попадания генерального параметра в данный интервал. Термин «степени свободы» означает, что их можно вычислить как объем выборки минус число ограничивающих условий; $\delta(m_{\bar{x}})$ — ошибка выборочной средней.

Для интерпретации доверительного интервала в клинических работах следует помнить, что ширина доверительного интервала зависит от $\delta(m_{\bar{x}})$ ошибки выборочной средней, которая в свою очередь зависит от объема выборки (n) и от изменчивости данных (S). Если выборка небольшая, то доверительный интервал более широкий, чем в случае выборки большого объема. Широкий доверительный интервал указывает на неточную оценку, а узкий – на точную оценку.

Верхний и нижние пределы доверительного интервала показывают, будут ли результаты клинически значимы.

Количественный признак x генеральной совокупности распределен нормально.

№ 1. По выборке объема n = 16 найдены выборочная средняя $\bar{x}_B$ = 20,2 и среднее квадратическое отклонение S = 0,8. Определить неизвестное математическое ожидание при помощи доверительного интервала при $p \geq 0{,}95$.

Решение:

$$\bar{x}_B - \frac{S}{\sqrt{n}}t \leq \mu \leq \bar{x}_B + \frac{S}{\sqrt{n}}t.$$

Найдем t из таблицы распределения Стьюдента при уровне значимости $\alpha \leq 0{,}95$ и числе степеней свободы f=n-1; f=16-1=15.

$$t(\alpha \leq 0{,}05, f = 15) = 2{,}13$$

Запишем:

$$20{,}2 - \frac{0{,}8}{\sqrt{16}} \cdot 2{,}13 \leq \mu \leq 20{,}2 + \frac{0{,}8}{\sqrt{16}} \cdot 2{,}13 (p \leq 0{,}95);$$

$$19{,}8 \leq \mu \leq 20{,}6 \text{ ïðè } \ p \leq 0{,}95.$$

Имеется выборка объёма n=11 - это значения систолического давления у мужчин в начальной стадии шока,

x: 127,124,155,129,77,147,65,109,145,141.

№ 2. С помощью пакета прикладных программ на ЭВМ провести статистическую обработку данных выборки и определить доверительный интервал для генеральной средней при $p \geq 0{,}95$.

Решение: Пусть расчет на ЭВМ дал: выборочное среднее $\bar{x}_B = 122{,}01; m_x = 8{,}59$

По таблице распределения Стьюдента найдем:

$t(\alpha \leq 0{,}05, f = 11 - 1 = 10) = 2{,}23.$

$\mu = \bar{x}_B \pm \delta t;$

$\mu = 122{,}01 \pm 8{,}59 \cdot 2{,}23 \qquad (p \geq 0{,}95);$

$\mu = 122 \pm 19; \qquad (p \geq 0{,}95).$

№ 3. Из обследованных 430 машин 37 оказались с наличием неисправности. Каковы 95%-ные доверительные границы процента неисправности данной партии машин?

Решение: **Выборочный средний процент пораженности составляет:**

$$p = \frac{37}{430} = 0{,}086 = 8{,}6.$$

Теперь по формуле находим:

$$\sigma_p = \sqrt{\frac{8{,}6 \cdot 91{,}4}{430}} \approx 1{,}35\%$$

.

Тогда для 95%-ного доверительного уровня имеем доверительные границы:

$$p \pm 1{,}96\sigma_p = (8{,}6 \pm 2{,}6)\%$$

,

т. е . 95%-ный доверительный интервал есть (6,0 $\pm 11,$ 2)%.

№ 4. При рентгеновском облучении 10 мышей дозой в 550 Р погибло 5 мышей. Каковы 99%-ные доверительные границы для доли мышей, погибающих пол действием данной дозы облучения?

Решение: **Имеем:**

$$p = \frac{5}{10} = 0{,}5; \qquad \sigma_p = \sqrt{\frac{0{,}5 \cdot 0{,}5}{10}} \approx 0{,}158;$$

.

поэтому при Р=99% (и $u_p = 2{,}58$) доверительные границы будут:

0,5-2,58*=0,5-0,408=0,092=9,2%,

0,5+2,58*0,158=0,5+0,408=0,908=90,8% .

Так как найденный доверительный интервал перекрывает почти весь возможный диапазон расположения истинной доли погибающих мышей (от 0 до 100%), то следует заключить, что опыт вообще не дал почти никакого результата (кроме указания, что при данной дозе облучения выборка из 10 мышей недостаточно велика для нахождения ответа на поставленный вопрос).

№ 5. При изучении в 10 опытах образования у собаки условного рефлекса под действием ранее индифферентного раздражителя были получены результаты (время между моментом включения условного раздражителя и моментом начала слюноотделения): $\bar{x} = 8{,}47$ с, $s_{\bar{x}} = 1{.}19$ с. Надо найти 95% -ный доверительный интервал для μ , характеризующий данное животное.

Решение: Для Р = 95% и f = n - 1 = 9 (число степеней свободы дисперсии) находим в приложении значение t = 2,26. Поэтому границы доверительного интервала будут:

$8{,}47 - 2{,}26 \cdot 1{,}19 = 8{,}47 - 2{,}69 = 5{,}78,$

$8{,}47 + 2{,}26 \cdot 1{,}19 = 8{,}47 + 2{,}69 = 11{.}16$

Результаты обычно записываются в одной из следующих двух форм:

$5{,}78 \div 11{,}16,$ или $8{,}47 \pm 2{,}69.$

Из табл. приложений видно, что значения t_P зависят особенно резко от f при малых f. Поэтому увеличение малых n приводит к сужению доверительного интервала определяемого величиной $t_P s_{\bar{x}} = \frac{t_P s}{\sqrt{n}}$ не только за счет уменьшения множителя $1/\sqrt{n}$, но в еще большей степени за счет уменьшения t_P . Так, при $P = 95\%$ изменении n с двух опытов до трех уменьшает множитель $t_P/\sqrt{n}$ с $12{,}71/\sqrt{2} = 9{,}0$ до $4{,}30/\sqrt{3} = 2{,}5$, т.е. доверительный интервал сужается в 9,0 : 2,5 = 3,6 раза; при P=99% ширина доверительного интервала уменьшается даже примерно в 8 раз ($63{,}66/\sqrt{2} = 45{,}0$; $9{,}93/\sqrt{3} = 5{,}7$; $45{,}0/5{,}7 = 7{,}9$). При больших значениях n увеличение n на одну единицу сказывается на ширине доверительного интервала гораздо меньше.

Из генеральной совокупности извлечена выборка объема n=10:

- вариант xi -2 1 2 3 4 5

- частота ni 2 1 2 2 2 1

№ 6. Оценить с надежностью 0,95 математическое ожидание а нормально распределенного признака генеральной совокупности по выборочной средней при помощи доверительного интеграла.

Решение:

Выборочную среднюю и исправленное среднее квадратическое отклонение найдем соответственно по формулам:

$$\overline{X}_B = \frac{\sum n_i x_i}{n}, \qquad S = \sqrt{\frac{\sum n_i (x_i - \overline{x}_B)^2}{n-1}}.$$

Подставим в эти формулы данные задачи, получим

$$\overline{X}_B = 2, \quad S = 2{,}4.$$

Найдем t_γ. Пользуясь таблицей, по g=0,95 и n=10 находим t_γ =2,26.

Найдем искомый доверительный интервал

$$\overline{X}_B - t\gamma\frac{S}{\sqrt{n}} < \alpha < \overline{X}_B + t\gamma\frac{S}{\sqrt{n}}.$$

Подставляя $\overline{X}_B = 2,\ t_\gamma = 2{,}26,\ S = 2{,}4,\ n = 10,$ получим искомый доверительный интервал 0,3<a<3,7, покрывающий неизвестное математическое ожидание а с надежностью 0,95.

№ 7. По данным 9 независимых равноточных измерений некоторые физические величины найдены среднее арифметическое результатов измерений $\overline{X}_B = 30{,}1$ и исправленное квадратическое отклонение S=6. Оценить истинное значение измеряемой величины с помощью доверительного интервала с надежностью g=0,99.

Решение:

Истинное значение измеряемой величины равно ее математическому ожиданию а. Поэтому задача сводится к оценке математического ожидания (при неизвестном s) при помощи доверительного интервала $\overline{X}_B - t\gamma\frac{S}{\sqrt{n}} < \alpha < \overline{X}_B + t\gamma\frac{S}{\sqrt{n}}$

Здесь все величины, кроме tg, известны. Найдем tg. По таблице g=0,95 и n=9 находим tg=3,36.

Подставим $\overline{X}$ =30,1, tg=3,36, S=6, n=9, получим искомый интервал

23,38<a<36,82

№ 8. По данным выборки объема n=16 из генеральной совокупности найдено исправленное среднее квадратическое отклонение s=1 нормально распределенного количественного признака. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение s с надежностью 0,95.

Решение:

Задача сводится к отысканию доверительного интервала

S(1-q)<s<s(1+q), (если q<1) (*)

или 0<s<s(1+q) (если q>1).

По данным g=0,95 и n=16 найдем q=0,44. Так как q<1, то, подставив s=1, q=0,44 в соотношение (*), получим искомый доверительный интервал 0,56<s<1,44.

ЗАДАНИЕ ДЛЯ САМОКОНТРОЛЯ.

№ 1. При исследовании частоты дыхания по выборке объема n = 15 были получены выборочная средняя $\overline{x}_B$ =18,5 и среднее квадратическое отклонение S = 0,6. Определить интервальную оценку математического ожидания с вероятностью $p \ge 0{,}95$.

№ 2. Найти доверительный интервал для оценки с уровнем доверительной вероятности $p \geq 0{,}95$ неизвестного математического ожидания нормального распределения признака x – диаметра эритроцита – генеральной совокупности, если выборочная средняя $\bar{x}_B$ =10,2 мкм; исправленное выборочное среднее квадратическое отклонение S = 4 и объем выборки п = 16. 16. При исследовании эффективности диуретика было установлено, что среднее увеличение диуреза в группе из n=10 человек составило $\bar{x}_B = 220$ìë . Найти доверительный интервал ($p \geq 0{,}95$) для среднего изменения диуреза, если ошибка выборочной средней $\delta(m_{\bar{x}}) = 8{,}9$.

Выполненные задания высылаем мне на почту, при этом не забываем подписывать. Удачи вам в выполнении работы.