Практикум 17
Раздел 17. «*Основы математической статистики*»

Требования к выполнению заданий

Решение задач необходимо располагать в порядке, указанном в заданиях, сохраняя номера задач. Перед решением каждой задачи желательно кратко записать имеющиеся данные из условия. Решение же следует излагать подробно и аккуратно, объясняя все действия по ходу решения и делая необходимые чертежи. Записывайте или выделяйте ответы.

Критерии оценки заданий

Правильное полное решение всех заданий оценивается в 10 баллов (1 задача=2 балла). При наличии неточностей или ошибок выставляемый за задачу балл может быть уменьшен до 0 либо 1.

Разбор типовых задач

Пример 1.

По заданной выборке:

3; 6; 11; 3; 14; 4; 6; 1; 2; 1; 6; 11; 12; 12; 1; 2; 12; 1; 2; 4

1. составить вариационный ряд;
2. построить полигон относительных частот;
3. найти $\overline{X}$, $S_x^{\ 2}$, S_x;
4. построить доверительные интервалы при $\gamma = 0{,}95$ для

а) математического ожидания при известной дисперсии $\sigma = s$;
б) математического ожидания при неизвестной дисперсии;
в) среднего квадратического отклонения.

Решение.

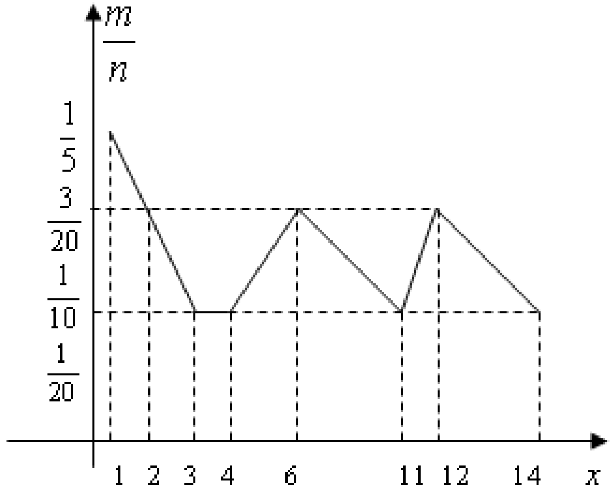
1. Составим вариационный ряд. Объем выборки $n = 20$.

x_i	1	2	3	4	6	11	12	14
m_i	4	3	2	2	3	2	3	1
$\frac{m_i}{n}$	$\frac{1}{5}$	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{1}{10}$

2. Построим полигон относительных частот. На плоскости O_{xy} построим точки:

$$\left(1; \frac{1}{5}\right), \left(2; \frac{3}{20}\right), \left(3; \frac{1}{10}\right), \left(4; \frac{1}{10}\right), \left(6; \frac{3}{20}\right), \left(11; \frac{1}{10}\right), \left(12; \frac{3}{20}\right), \left(14; \frac{1}{10}\right).$$

Соединим их. Полученная ломаная – полигон относительных частот.



3. Вычислим выборочную среднюю $\overline{X}$:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{k} x_i m_i = \frac{1}{20}(1\times4+2\times3+3\times2+4\times2+6\times3+11\times2+12\times3+14\times1) =$$

$$\frac{1}{20}(4+6+6+8+18+22+36+14) = \frac{114}{20} = 5{,}7.$$

Вычислим исправленную выборочную дисперсию:

$$S_x{}^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} x^2{}_i m_i - n(\overline{X})\right) = \frac{1}{19}(1^2\times4+2^2\times3+3^2\times2+4^2\times2+6^2\times3+$$

$$+11^2\times2+12^2\times3+14^2\times1-20\times(5{,}7)^2) \approx 20{,}75.$$

Следовательно, $S_x \approx 4{,}56$.

При построении доверительного интервала для математического ожидания $M(x)=m$ с известной дисперсией $D(x)=S^2=20{,}75$ воспользуемся формулой:

$$\overline{X} - Z_\gamma \frac{\sigma}{\sqrt{n}} < m < \overline{X} + Z_\gamma \frac{\sigma}{\sqrt{n}},$$

где $\overline{X}=5{,}7$, $\sigma = S = 4{,}56$, $n=20$.

Число Z_γ находим с помощью таблицы приложения из уравнения

$$\Phi(Z\gamma) = \frac{\gamma}{2} => \Phi(Z\gamma) = 0{,}475 => Z\gamma = 1{,}96.$$

Следовательно, искомый доверительный интервал имеет вид:

$$5{,}7 - 1{,}96\frac{4{,}56}{\sqrt{20}} < m < 5{,}7 + 1{,}96\frac{4{,}56}{\sqrt{20}},$$

т.е. $3{,}7 < m < 7{,}7$.

4. При построении доверительного интервала для математического ожидания $M(x)=m$ с неизвестной дисперсией воспользуемся формулой:

$$\overline{X} - t_\gamma \frac{S}{\sqrt{n}} < m < \overline{X} + t_\gamma \frac{S}{\sqrt{n}}$$

Число t_γ находим с помощью таблицы приложения при $n=20$ и $\gamma = 0{,}95$:

$$t_\gamma(0{,}95,\ 20) = 2{,}093.$$

Следовательно, искомый доверительный интервал имеет вид:

$$5{,}7 - 2{,}093 \times \frac{4{,}56}{\sqrt{20}} < m < 5{,}7 + 2{,}093 \times \frac{4{,}56}{\sqrt{20}},$$

$$3{,}565 < m < 7{,}835.$$

5. При построении доверительного интервала для среднего квадратического отклонения σ воспользуемся формулой:

$$S(1-q) < \sigma < S(1+q),$$

где $S = 4{,}56$, q - число, которое находим с помощью таблицы приложения при $n = 20$ и $\gamma = 0{,}95$, т.е. $q(0{,}95,\ 20) = 0{,}37$.

Следовательно, искомый доверительный интервал имеет вид:

$$4{,}56(1-0{,}37) < \sigma < 4{,}56(1+0{,}37),$$

т.е. $2{,}87 < \sigma < 6{,}25$.

Ответ:

1. $\overline{X} = 5{,}7$,
2. $S_x^{\ 2} = 20{,}75$,
3. $S_x \approx 4{,}56$,
4. а) $3{,}7 < m < 7{,}7$, б) $3{,}565 < m < 7{,}835$, в) $2{,}87 < \sigma < 6{,}25$.

Пример 2.

Исследуется зависимость между количеством (N) покупателей в ювелирном магазине и количеством (Q) проданных товаров в один день. За 10 дней наблюдений получена таблица пар $(N;\ Q)$:

$X = N$	50	61	72	43	60	65	76	55	62	40
$Y = Q$	10	12	20	9	15	15	21	14	18	7

Найти выборочный коэффициент корреляции и проверить гипотезу о его значимости.

Решение.

1. Найдем выборочные характеристики.

$$\overline{N} = \overline{X} = \frac{1}{10}\left(50+61+72+43+60+65+76+5+62+40\right) = 58{,}4;$$

$$\overline{Q} = \overline{Y} = \frac{1}{10}(10+12+20+9+15+15+21+14+18+7) = 14{,}1;$$

$$D_{bx} = \frac{1}{10}\left(50^2+61^2+72^2+43^2+60^2+65^2+76^2+5^2+62^2+40^2\right) - (58{,}4)^2 = 121{,}84;$$

$$D_{by} = \frac{1}{10}(10^2+12^2+20^2+9^2+15^2+15^2+21^2+14^2+18^2+7^2) - (14{,}1)^2 = 19{,}69;$$

$$\sigma_{bx} = \sqrt{121{,}84} = 11{,}04;$$

$$\sigma_{by} = \sqrt{19{,}69} = 4{,}44;$$

$$\widetilde{M}_{xy} = \frac{1}{10}\left(50\times10+61\times12+72\times20+43\times9+60\times15+65\times15+76\times21+55\times18+62\times18+40\times7\right) -$$
$$-58{,}4\times14{,}1 = 46{,}3;$$

$$\widetilde{r}_{xy} = \frac{\widetilde{M}_{xy}}{\sigma_{bx}\times\sigma_{by}};$$

$$\widetilde{r}_{xy} = \frac{46{,}3}{11{,}04\times4{,}44} = \frac{46{,}3}{49{,}02} = 0{,}94.$$

2. Проверим гипотезу о значимости выборочного коэффициента корреляции

$$\begin{matrix} H_0: & r_{xy} = 0 \\ H_1: & r_{xy} \neq 0 \end{matrix}.$$

Зададим уровень значимости $\alpha = 0{,}05$. Вычислим $T_{набл} = \frac{\widetilde{r}_{xy}\sqrt{n-2}}{\sqrt{1-\widetilde{r}^{\,2}_{xy}}} = \frac{0{,}94\times\sqrt{10-2}}{\sqrt{1-(0{,}94)^2}} = 7{,}82$.

По таблице для $n-2=8$ степеней свободы и заданного уровня значимости $\alpha = 0{,}05$ находим $T_{крит}(0{,}05;\, 8) = 2{,}3$.

Так как $T_{набл} > T_{крит} \Rightarrow H_0$ - отвергаем, т.е. $\widetilde{r}_{xy} \neq 0$.

Таким образом, величины X и Y признаны коррелированными, т.е. между ними есть корреляционная зависимость.

Ответ: $\begin{matrix} \widetilde{r}_{xy} = 0{,}94 \\ r_{xy} \neq 0 \end{matrix}$

Пример 3.

Дана корреляционная зависимость:

Y \ X	0	1	2	3
0	1	5	3	-
1	-	2	8	1

Написать выборочные уравнения линейной регрессии Y на X и X на Y.

Решение.

Найдем выборочные характеристики.

- выборочные средние:

$$\overline{X} = \frac{1}{20}(0 \times 1 + 1 \times 7 + 2 \times 11 + 3 \times 1) = \frac{32}{20} = \frac{8}{5} = 1{,}6$$

$$\overline{Y} = \frac{1}{20}(0{,}9 + 1{,}11) = \frac{11}{20} = 0{,}55.$$

- выборочные дисперсии:

$$D_{bx} = \frac{1}{10}\left(0^2 \times 1 + 1^2 \times 7 + 2^2 \times 11 + 3^2 \times 1\right) - \left(\frac{8}{5}\right)^2 = 3 - \frac{64}{25} = 0{,}44;$$

$$D_{by} = \frac{1}{20}(0^2 \times 9 + 1^2 \times 11) - \left(\frac{11}{20}\right)^2 = \frac{99}{400} = 0{,}2475.$$

- выборочные среднеквадратичные отклонения:

$$\sigma_{bx} = \sqrt{0{,}44} = 0{,}663, \ \sigma_{by} = \sqrt{0{,}2475} = 0{,}497;$$

- выборочный корреляционный момент (нулевые слагаемые не вписаны):

$$\tilde{M}_{xy} = \frac{1}{20}(1 \times 1 \times 2 + 1 \times 2 \times 8 + 1 \times 3 \times 1) - \frac{11}{20} \times \frac{8}{5} = \frac{21}{20} - \frac{22}{25} = 0{,}17;$$

- выборочный коэффициент корреляции:

$$\tilde{r}_{xy} = \frac{0{,}17}{0{,}497 \times 0{,}663} = 0{,}516.$$

Выборочные уравнения линейной регрессии имеют вид:

$$y_x = \overline{y} + \tilde{r}_{xy}\frac{\sigma_{by}}{\sigma_{bx}}(x - \overline{x}) \ (Y \text{ на } X)$$

$$x_y = \overline{x} + \tilde{r}_{xy}\frac{\sigma_{bx}}{\sigma_{by}}(y - \overline{y}) \ (X \text{ на } Y).$$

Следовательно,

$$y_x = 0{,}55 + 0{,}516\frac{0{,}497}{0{,}663}(x - 1{,}6) = 0{,}387x - 0{,}07$$

$$x_y = 1{,}6 + 0{,}516\frac{0{,}663}{0{,}497}(y - 0{,}55) = 0{,}688y + 1{,}22$$

Ответ: $y_x = 0{,}387x - 0{,}07$; $x_y = 0{,}688y + 1{,}22$.