ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ

Самостоятельная работа № 2
по математике

Методы математической статистики

студ. ФИО
группа ...
преп.:

Санкт-Петербург
2018

Глава 1

Пример самостоятельной работы

1.1 Постановка задачи

Известно, что прибыль фирмы зависит от вложения в рекламную кампанию. Необходимо сделать прогноз значения прибыли, если в рекламную кампанию будет вложено 22 тысячи рублей в день. Дневная прибыль фирмы в зависимости от вложения в рекламу в день за первые 15 дней выписана в таблицу

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	7	4	6	6	7	5	9	3	5	6	3	8	6	7	5
y_i	13	9	11	13	13	11	17	7	11	11	6	17	11	14	12

где x_i – вложение средств в рекламную кампанию в тысячах руб. в i день, y_i – прибыль в десятках тысяч рублей в i день, $i = 1, 2, \ldots, 15$ – календарный день.

Необходимо:

1. По заданным в условии выборкам построить вариационный ряд.

2. Найти среднее выборочное, выборочную дисперсию, исправленную дисперсию, с.к.о. и исправленное с.к.о. выборок X и Y по форму-

лам:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k_x} n_i x_i \quad D_v = \frac{1}{n}\sum_{i=1}^{k_x} n_i x_i^2 - (\bar{x})^2 \tag{1.1}$$

$$S = \frac{n}{n-1}D_v, \quad \sigma = \sqrt{D_v}, \quad s = \sqrt{S}$$

где k_x – число вариант вариационного ряда.

3. Найти доверительный интервал для математического ожидания a с доверительной вероятностью $\gamma = 0.95$ по формуле.

$$a \in (\bar{x} - t_{\gamma,n}s/\sqrt{n}, \bar{x} + t_{\gamma,n}s/\sqrt{n})_\gamma \tag{1.2}$$

где постоянная $t_{\gamma,n}$ находится по таблицам распределения Стьюдента.

4. Построить гистограммы случайных величин X и Y, разделив размах выборки на 5 равных интервалов. По найденным средней выборочной и исправленной с. к. о. на гистограмме изобразить график функции плотности нормального распределения.

$$f_{n.r.}(x) = e^{-\frac{(x-a)^2}{2\sigma^2}} \tag{1.3}$$

где вместо a берётся среднее выборочное $\bar{x}$, а вместо σ^2 – исправленная выборочная дисперсия S. Для расчёта можно положить $e = 2.7$.

При построении графика функции плотности учитываем, что эта функция

– симметричная относительно своей вершины;

– в точке, удалённой от математического ожидания на расстоянии $3\sigma =$ равна практически нулю;

– в точке, удалённой от математического ожидания на расстоянии σ составляет 60% от максимума и имеет точку перегиба.

На основании гистограммы и графика функции плотности нормального распределения сделать вывод о возможном нормальном распределении случайной величины и достаточен ли объём выборки для предположения о нормальности распределения.

5. Найти выборочный r_v коэффициент корреляции между случайными величинами X и Y по формулам:

$$r_v = \frac{n\sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i \cdot \sum\limits_{i=1}^{n} y_i}{\sqrt{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}\sqrt{n\sum\limits_{i=1}^{n} y_i^2 - (\sum\limits_{i=1}^{n} y_i)^2}} \tag{1.4}$$

и проверить гипотезу о его значимости, используя распределение критических точек Стьюдента для уровня значимости $\alpha = 0.05$. (Вспомним, что данный уровень значимости означает, что мы допускаем, что не более чем в 5 процентах случаях мы можем ошибаться.)

$$T = r_v \sqrt{\frac{n-2}{1-r_v^2}} \qquad (1.5)$$

6. На плоскости XOY отметить точки $(x_i, y_i), \quad i = 1, 2, \ldots, 15$.

7. Считая, что случайная величина Y линейно зависит от случайной величины X как

$$y = ax + b$$

методом МНК найти постоянную b и коэффициент a по формулам:

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}, b = \frac{\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \qquad (1.6)$$

8. Дать прогноз прибыли фирмы при вложении в рекламу 22 тысячи рублей в день.

Решение.

1.2 Первичная обработка статистических данных

1. Построение вариационного ряда.

1). Строим вариационный ряд для случайной величины X.

Таблица 1. $k_x = 7$.

x_i	3	4	5	6	7	8	9
n_i	2	1	3	4	3	1	1

2). Строим вариационный ряд для случайной величины Y.

Таблица 2. $k_y = 8$.

y_i	6	7	9	11	12	13	14	17
n_i	1	1	1	5	1	3	1	2

2. Находим среднее выборочное, выборочную дисперсию и исправленную дисперсию. Объём выборки: $n = 15$.

1). По формулам (1.1) и вариационному ряду таблицы 1 находим среднее выборочное, выборочную дисперсию и исправленную дисперсию выборки X.

$$\bar{x} = \frac{1}{15}(2 \cdot 3 + 4 + 3 \cdot 5 + 4 \cdot 6 + 3 \cdot 7 + 8 + 9) = 5.8 \tag{1.7}$$

$$DX_v = \frac{1}{15}(2 \cdot 3^2 + 3 \cdot 4^2 + 3 \cdot 5^2 + 2 \cdot 6^2 + 3 \cdot 7^2 + 8^2 + 9^2) - (5.8)^2 = 2.69 \tag{1.8}$$

$$\sigma_x = \sqrt{2.69} = 1.64, \quad S_x = \frac{15}{14} \cdot 2.69 = 2.89, \quad s_x = \sqrt{3.27} = 1.70 \tag{1.9}$$

2). Объём выборки: $n = 15$. По формулам (1.1) и вариационным рядам таблицы 2 находим среднее выборочное, выборочную дисперсию и исправленную дисперсию выборки Y.

$$\bar{y} = \frac{1}{15}(6 + 7 + 3 \cdot 9 + 4 \cdot 11 + 12 + 2 \cdot 13 + 14 + 2 \cdot 17) = 11.73 \tag{1.10}$$

$$DY_v = \frac{1}{15}(6+7+3 \cdot 9+4 \cdot 11+12+2 \cdot 13+14+2 \cdot 17) - (11.73)^2 = 8.73 \tag{1.11}$$

$$\sigma_y = \sqrt{8.73} = 2.95, \quad S_y = \frac{15}{14} \cdot 8.73 = 9.35, \quad s_y = \sqrt{9.35} = 3.06 \tag{1.12}$$

3. Нахождение доверительного интервала для математического ожидания a с доверительной вероятностью $\gamma = 0.95$ при $n = 15$.

Для заданной надёжности $\gamma = 0.95$ при $n = 15$ по таблице распределения Стьюдента находим $t_{0.95,15} = 2.15$.

1). Найдём нижнюю (α) и верхнюю (β) границы интервала математического ожидания a_X случайной величины X:

$$\alpha = \bar{x} - t_{\gamma,n} s_x/\sqrt{n} = 5.8 - 2.15 \cdot 1.81/\sqrt{15} = 4.86$$

$$\beta = \bar{x} + t_{\gamma,n} s_x/\sqrt{n} = 5.8 + 2.15 \cdot 1.81/\sqrt{15} = 6.74$$

Откуда находим, что доверительный интервал для математического ожидания a случайной величины X с доверительной вероятностью $\gamma = 0.95$ имеет вид $a_X \in (4.86; 6.74)_{\gamma=0.95}$.

2). Найдём нижнюю (α) и верхнюю (β) границы интервала математического ожидания a_Y случайной величины Y:

$$\alpha = \bar{y} - t_{\gamma,n} s_y/\sqrt{n} = 11.73 - 2.15 \cdot 3.18/\sqrt{15} = 10.04$$

$$\beta = \bar{y} + t_{\gamma,n} s_y/\sqrt{n} = 11.73 + 2.15 \cdot 3.18/\sqrt{15} = 13.43$$

Откуда находим, что доверительный интервал для математического ожидания a случайной величины Y с доверительной вероятностью $\gamma = 0.95$ имеет вид $a_Y \in (10.04; 13.43)_{\gamma=0.95}$.

1.3 Построение гистограммы и её проверка хи-квадрат

4. Построение гистограммы для случайных величин X и Y.

1). Построим гистограмму для случайной величины X.

a). Находим размах выборки $H = x_{max} - x_{min} = 9 - 3 = 6$.

b). Находим длину интервала $h = H : 5 = 6 : 5 = 1.2$.

c). Строим интервальный ряд плотностей относительных частот, группируя варианты по интервалам.

Таблица 3.

x_i	$[3, 4.2)$	$[4.2, 5.4)$	$[5.4, 6.6)$	$[6.6, 7.8)$	$[7.8, 9]$
n_i	3	3	4	3	2
$\frac{n_i}{nh}$	0.17	0.17	0.22	0.17	0.11

d). По интервальному ряду таблицы 3 строим гистограмму относительных частот. (Площадь каждого прямоугольника равна относительной частоте $p_i = W_i h$. Полная площадь под графиком на рис. 2 равна единице.)

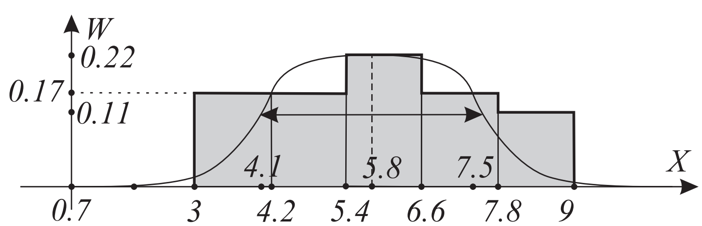

рис. 1.

На построенную гистограмму накладываем график функции плотности нормального распределения с полученными значениями: в качестве математического ожидания возьмём среднее выборочное $a_x = \bar{x} = 5.8$, а с.к.о. $= s_x = 1.7$ (исправленное выборочное с.к.о. X).

Так как $a_x = \bar{x} = 5.8$, то максимальное значение, равное 0.22 будет достигаться в точке 5.8.

Найдём его по формуле $f(0) = \dfrac{1}{\sigma\sqrt{2\pi}}$. Так как с.к.о. σ равно $s_x = 1.7$, то $f(0) = \dfrac{1}{1.7\sqrt{2\pi}} = 0.23$.

В точках $a_x - s_x = 5.8 - 1.7 = 4.1$ и $a_x + s_x = 5.8 + 1.7 = 7.5$ значение функции достигает 60% от максимума, то есть $0.22 \cdot 0.6 = 0.13$ и в этих точках возникает перегиб функции.

Из вида гистограммы на рис. 1 и наложенной на неё соответствующей функции плотности нормального распределения делаем вывод, что предположение о нормальности распределения случайной величны X вполне допустимо, но для большей надёжности желательно иметь большее число данных наблюдения.

Выдвинем гипотезу H_0 о нормальном распределении случайной величины X и проверим эту гипотезу с помощью критерия Пирсона. Для этого вычислим значение критерия ${\chi^2}_{\text{набл.}}$ случайного распределения X по формуле:

$$
{\chi^2}_{\text{набл.}} = \sum_{k=1}^{n_{int}} \frac{(n_k - n'_k)^2}{n'_k}, \tag{1.13}
$$

где $n'_k = \dfrac{nh}{\sigma}\varphi(u_k)$ – теоретические частоты, n – объём выборки (в данном случае $n = 15$), h – шаг (здесь $h = 1.2$), n_{int} – число интервалов (здесь $n_{int} = 5$):

$$
u_k = \frac{l_k - l_{v.}}{\sigma_{v.}}, \qquad \varphi(u_k) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-u_k^2/2}, \tag{1.14}
$$

$l_{v.} = \bar{x}$ – выборочное среднее и $\sigma_{v.} = s_x$ – исправленное выборочное с. к. о., l_k – середина k-го интервала (иногда в качестве l_k выбирают границу k-го интервала).

Вычислим наблюдаемое значение ${\chi^2}_{\text{набл.}}$ случайной величины X. Для этого составим таблицу 4, подставляя в формулы (1.14) $l_{v.} = 5.8$ и $\sigma_{v.} = 1.7$:

Таблица 4.

k	l_k	$u_k = \dfrac{l_k - l_{v.}}{\sigma_v}$	$\varphi(u_k)$	n'_k	n_k	$(n_k - n'_k)^2/n'_k$
1	3.6	-1.29412	0.1727	1.828	3	0.75
2	4.8	-0.58824	0.3356	3.553	3	0.09
3	6.0	0.11765	0.3962	4.195	4	0.01
4	7.2	0.82353	0.2842	3.009	3	0.0
5	8.4	1.52941	0.1239	1.312	2	0.36

Откуда находим, что $\chi^2_{\text{набл.}} = 0.75 + 0.09 + 0.01 + 0.0 + 0.36 = 1.21$.

Зададим уровень значимости $\alpha = 0.05$. По таблице критических точек распределения, по заданному уровню значимости α и числу степеней свободы $k_0 = n_{int.} - 3$ ($n_{int.}$ – число групп выборки или число интервалов, здесь $n_{int.} = 5$) найдём критическую точку правосторонней критической области: $\chi^2_{kr.}(\alpha = 0.05, k_0 = n_{int.} - 3 = 5 - 3 = 2) = 6.0$.

Так как $\chi^2_{\text{набл.}} = 1.21 < \chi^2_{kr.} = 6.0$, то эмпирические и теоретические частоты, различаются незначимо (случайно) и нет оснований отвергать гипотезу о нормальном распределении генеральной совокупности случайной величины X.

2). Построим гистограмму для случайной величины Y.

a). Находим размах выборки $H = y_{max} - y_{min} = 17 - 6 = 11$.

b). Находим длину интервала $h = H : 5 = 11 : 5 = 2.2$.

c). Строим интервальный ряд плотностей относительных частот, группируя варианты по интервалам.

Таблица 5.

x_i	$[6, 8.2)$	$[8.2, 10.4)$	$[10.4, 12.6)$	$[12.6, 14.8)$	$[14.8, 17]$
n_i	2	1	6	4	2
$\dfrac{n_i}{nh}$	0.06	0.03	0.18	0.12	0.06

d). По интервальному ряду таблицы 5 строим гистограмму относительных частот.

На построенную гистограмму накладываем график функции плотности нормального распределения с полученными значениями: в качестве математического ожидания возьмём среднее выборочное $a_y = \bar{y} = 11.73$, а с. к. о. $\sigma = s_y = 3.06$ (исправленное выборочное с.к.о. Y).

Так как $a_y = \bar{y} = 11.73$, а с.к.о. $s_y = 3.06$, то максимальное значение, равное 0.13 будет достигаться в точке 11.73. В точках $\sigma = a_y - s_y = 11.73 - 3.06 = 8.7$ и $a_y + s_y = 11.73 + 3.06 = 14.8$ значение функции достигает 60% от максимума, то есть $0.13 \cdot 0.6 = 0.08$ и в этих точках возникает перегиб функции.

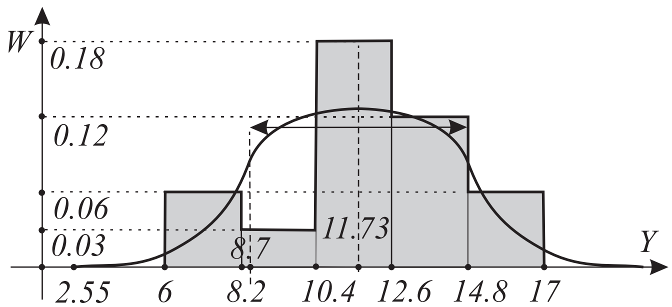

рис. 2.

Из вида гистограммы на рис. 2 и наложенной на неё соответствующей функции плотности нормального распределения делаем вывод, что для предположения о нормальности распределения случайной величны Y необходимо иметь большее число данных наблюдения.

Выдвинем гипотезу H_0 о нормальном распределении случайной величины Y и проверим эту гипотезу с помощью критерия Пирсона. Для этого вычислим значение критерия $\chi^2_{\text{набл.}}$ случайного распределения Y по формулам 1.13 и 1.14.

Вычислим наблюдаемое значение $\chi^2_{\text{набл.}}$ случайной величины Y. Для этого составим таблицу 6, подставляя в формулы (1.14) $l_{v.} = 11.73$ и $\sigma_{v.} = 3.06$:

Таблица 6.

k	l_k	$u_k = \dfrac{l_k - l_{v.}}{\sigma_v}$	$\varphi(u_k)$	n'_k	n_k	$(n_k - n'_k)^2/n'_k$
1	7.1	-1.5131	0.127	1.370	3	0.29
2	9.3	-0.7941	0.291	3.139	3	1.46
3	11.5	-0.07516	0.398	4.290	4	0.68
4	13.7	0.6438	0.324	3.497	3	0.07
5	15.9	1.3628	0.158	1.700	2	0.53

Откуда находим, что $\chi^2_{\text{набл.}} = 0.29 + 1.46 + 0.68 + 0.07 + 0.53 = 2.55$.

Так как $\chi^2_{\text{набл.}} = 2.55 < \chi^2_{kr.} = 6.0$, то эмпирические и теоретические частоты, различаются незначимо (случайно) и нет оснований отвергать гипотезу о нормальном распределении генеральной совокупности случайной величины X.

1.4 Построение линейной регрессии

5. Найдём коэффициент корреляции случайных величин X и Y по формуле (1.4), которую с учётом предыдущих рассчётов запишем несколько иначе:

$$r_v = \frac{\overline{xy} - \bar{x}\bar{y}_v}{\sigma_x \sigma_y} \tag{1.15}$$

где $\overline{xy} = \dfrac{1}{n}\sum_{i=1}^{n} x_i y_i$.

$\overline{xy} = (7\cdot 13 + 4\cdot 9 + 6\cdot 11 + 6\cdot 13 + 7\cdot 13 + 5\cdot 11 + 9\cdot 17 + 3\cdot 7 + 5\cdot 11 + 6\cdot 1 + 3\cdot 8 + 8\cdot 17 + 6\cdot 11 + 7\cdot 14 + 5\cdot 12)/15 = 72.67$

$$r_v = \frac{72.67 - 5.8\cdot 11.73}{1.64\cdot 2.95} = 0.95$$

Проверим гипотезу о значимости коэффициента корреляции.

Для этого выдвигаем гипотезу $H_0 : r_v = 0$ с конкурирующей гипотезой $H_1 : r_v \neq 0$ (именно так применяется этот критерий) и найдём

$$T = r_v\sqrt{\frac{n-2}{1-r_v^2}} = 0.95\cdot\sqrt{\frac{15-2}{1-0.95^2}} = 11.15 \tag{1.16}$$

По условию конкурирующая гипотеза $H_1 : r_v \neq 0$, поэтому область двусторонняя (допускаются как значения $r_v < 0$, так и $r_v > 0$). По таблицам распределения критических точек Стьюдента двусторонней области находим, что $t_{kr}(\alpha = 0.05, k = 15 - 2 = 13) = 2.16$. Видим, что $t_{kr} = 2.16 < T = 11.15$, поэтому основную гипотезу о равенстве нулю коэффициента корреляции отвергаем и принимаем конкурирующую гипотезу, полагая коэффициент корреляции между случайными величинами X и Y равным выборочному коэффициенту корреляции $r_v = 0.95$.

6. С учётом ранее проведённых расчётов найдём линейное уравнение регрессии $y = ax + b$ преобразуя формулу (1.6) к виду:

$$y = r_v\frac{\sigma_y}{\sigma_x}x + \bar{y}_v - r_v\frac{\sigma_y}{\sigma_x}\bar{x} \tag{1.17}$$

$a = r_v\dfrac{\sigma_y}{\sigma_x} = 0.95\cdot\dfrac{2.95}{1.64} = 1.71, \quad b = \bar{y}_v - a\bar{x} = 11.73 - 1.71\cdot 5.8 = 1.80$

Следовательно

$$y = 1.71x + 1.8$$

7. Отметим на плоскости пары (x_i, y_i) и прямую $y = 1.71x + 1.8$. (Повторяющиеся пары выделим как точки, большие диаметром.)

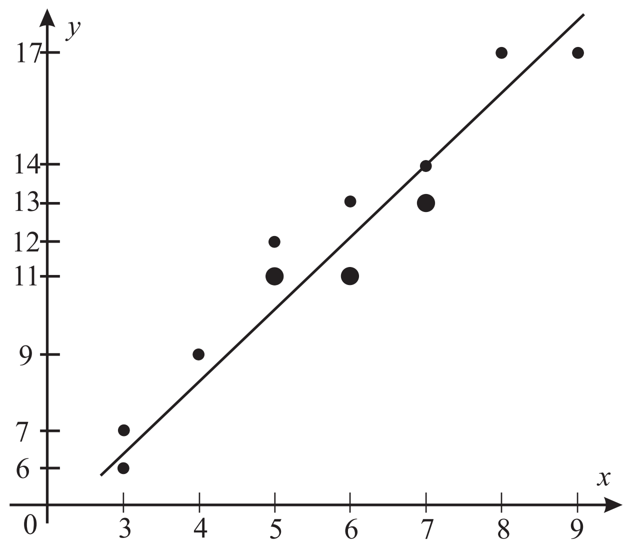

рис. 3.

8. Если вкладывать в рекламную кампанию $x = 22$ тыс. рублей в день, то прибыль окажется равной $y = 1.71 \cdot 22 + 1.8 = 39.48$ или 394.8 тыс. рублей в день.

1.5 Линейная множественная регрессия

Известно, что прибыль фирмы зависит от вложения в рекламную кампанию на ТВ и радио. Необходимо сделать прогноз значения прибыли, если в рекламную кампанию будет вложено 50 тысячи рублей в день на ТВ и 20 тыс. руб. в день на радио. Дневная прибыль фирмы y в десятках тысяч рублей зависимости от суточного вложения в рекламу на ТВ (x_1) и на радио (x_2) за первые 15 дней выписана в таблицу

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_1^i	7	4	6	6	7	5	9	3	5	6	3	8	6	7	5
x_2^i	6	8	5	2	15	10	20	4	8	5	4	13	10	13	8
y^i	20	2	18	24	30	17	26	21	16	5	1	34	32	26	20

Необходимо:

1. Найти выборочный r_v коэффициенты парной корреляции между случайными величинами X^1, X^2 и Y по формулам (1.4).

2. Считая, что случайная величина Y линейно зависит от случайных величин X_1 и X_2 как

$$Y = a_1 X_1 + a_2 X_2 + b$$

найти a_1, a_2, b по формулам:

$$a_1 = \beta_1 \frac{\sigma_y}{\sigma_{x_1}}, \quad a_2 = \beta_2 \frac{\sigma_y}{\sigma_{x_2}}, \quad b = \bar{y} - a_1 \bar{x}_1 - a_2 \bar{x}_2 \tag{1.18}$$

где нормализованные коэффициенты β_1 и β_2 находятся из системы линейных уравнений

$$\begin{cases} \beta_1 + \beta_2 r_{x_1 x_2} = r_{y x_1} \\ \beta_1 r_{x_2 x_1} + \beta_2 = r_{y x_2} \end{cases} \tag{1.19}$$

Решение.

1). Найдём среднии выборочные каждой из выборок.

$\bar{x}_1 = (7+4+6+4+7+5+9+3+5+4+3+8+6+7+5)/15 = 5.8$

$\bar{x}_2 = (6+8+5+2+15+10+20+4+8+5+4+13+10+13+8)/15 = 8.73$

$\bar{y} = (2\cdot 20+2+18+24+30+17+2\cdot 26+21+16+5+1+34+32)/15 = 19.47$

2). Найдём выборочные дисперсии и с. к. о. каждой из выборок.

$Dx_1 = (7^2+4^2+6^2+4^2+7^2+5^2+9^2+3^2+5^2+4^2+3^2+8^2+6^2+7^2+5^2)/15 - \bar{x}_1^2 = 2.69, \quad \sigma_{x_1} = \sqrt{3.05} = 1.64$

$Dx_2 = (6^2+8^2+5^2+2^2+15^2+10^2+20^2+4^2+8^2+5^2+4^2+13^2+10^2+13^2+8^2)/15 - \bar{x}_2 = 22.20, \quad \sigma_{x_2} = \sqrt{22.20} = 4.71$

$Dy = (20^2+2^2+18^2+24^2+30^2+17^2+26^2+21^2+16^2+5^2+1^2+34^2+32^2+26^2+20^2)/15 - \bar{y}^2 = 97.58, \quad \sigma_y = \sqrt{97.58} = 9.88$

3). Найдём коэффициенты парной корреляции.

$\overline{x_1 x_2} = (7\cdot 6+4\cdot 8+6\cdot 5+6\cdot 2+7\cdot 15+5\cdot 10+9\cdot 20+3\cdot 4+5\cdot 8+6\cdot 5+3\cdot 4+8\cdot 13+6\cdot 10+7\cdot 13+5\cdot 8)/15 = 56$

$r_{x_1x_2} = \dfrac{\overline{x_1x_2} - \bar{x}_1 \cdot \bar{x}_2}{\sigma_{x_1} \cdot \sigma_{x_2}} = \dfrac{56 - 5.53 \cdot 8.73}{1.75 \cdot 4.71} = 0.69$

$\overline{yx_1} = (7 \cdot 6 + 4 \cdot 8 + 6 \cdot 5 + 6 \cdot 2 + 7 \cdot 15 + 5 \cdot 10 + 9 \cdot 20 + 3 \cdot 4 + 5 \cdot 8 + 6 \cdot 5 + 3 \cdot 4 + 8 \cdot 13 + 6 \cdot 10 + 7 \cdot 13 + 5 \cdot 8)/15 = 123.4$

$r_{yx_1} = \dfrac{\overline{yx_1} - \bar{y} \cdot \bar{x}_1}{\sigma_y \cdot \sigma_{x_1}} = \dfrac{123.4 - 19.47 \cdot 5.53}{9.88 \cdot 1.75} = 0.65$

$\overline{yx_2} = (7 \cdot 6 + 4 \cdot 8 + 6 \cdot 5 + 4 \cdot 2 + 7 \cdot 15 + 5 \cdot 10 + 9 \cdot 20 + 3 \cdot 4 + 5 \cdot 8 + 4 \cdot 5 + 3 \cdot 4 + 8 \cdot 13 + 6 \cdot 10 + 7 \cdot 13 + 5 \cdot 8)/15 = 194.33$

$r_{yx_2} = \dfrac{\overline{yx_2} - \bar{y} \cdot \bar{x}_2}{\sigma_y \cdot \sigma_{x_2}} = \dfrac{194.33 - 19.47 \cdot 8.73}{9.88 \cdot 4.71} = 0.52$

4). Найдём нормализованные коэффициенты. Для этого составим нормализованную систему уравнений и решим её методом подстановки.

$\begin{cases} \beta_1 + 0.69\beta_2 & = 0.65 \\ 0.69\beta_1 + \beta_2 & = 0.52 \end{cases} \begin{cases} \beta_1 = 0.65 - 0.69\beta_2 \\ 0.69(0.65 - 0.69\beta_2) + \beta_2 = 0.52 \end{cases}$

$\begin{cases} \beta_2 = 0.14 \\ \beta_1 = 0.65 - 0.69 \cdot 0.14 \end{cases} \quad \begin{cases} \beta_2 = 0.14 \\ \beta_1 = 0.55 \end{cases}$

Найдём коэффициенты a_1 и a_2:

$$a_1 = \frac{0.55 \cdot 9.88}{1.75} = 3.3, \quad a_2 = \frac{0.14 \cdot 9.88}{4.71} = 0.3$$

и свободный член b:

$$b = 19.47 - 3.3 \cdot 5.8 - 0.3 \cdot 8.73 = -2.29$$

Таким образом линейное уравнение множественной регрессии имеет вид:

$$y = 3.3x_1 + 0.3x_2 - 2.29$$

Если в рекламу на ТВ вкладывать 50 тыс. рублей в день, а в рекламу на радио 20 – тысяч рублей в день, то предполагаемая выручка составит $y = 3.3 \cdot 50 + 0.3 \cdot 20 - 2.29 = 168.60$ или 1 млн 686 тыс. рублей в день.

Глава 2

Таблицы данных по вариантам

Таблица № 1 данных.

№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1.	6	6	5	6	8	1	4	3
2.	6	7	7	2	6	4	3	7
3.	1	8	2	5	7	7	8	4
4.	9	3	3	7	7	4	4	2
5.	8	6	5	4	1	5	5	5
6.	5	6	8	1	6	5	5	7
7.	3	9	6	8	6	8	1	7
8.	3	7	3	1	2	2	7	8
9.	4	6	9	2	4	1	3	5
10.	1	4	5	6	9	7	8	6
11.	9	7	2	1	4	1	5	4
12.	5	3	9	9	4	3	5	7
13.	8	9	1	6	5	6	4	5
14.	1	2	6	3	9	8	5	9
15.	5	4	9	6	3	3	8	1

Таблица № 2 соответствия данных Таблицы № 1 и номера варианта по балльно-рейтинговой системе

1	2	3	4	5
X_1, X_3	X_1, X_4	X_1, X_5	X_1, X_6	X_1, X_7
6	7	8	9	10
X_1, X_8	X_3, X_4	X_3, X_5	X_3, X_6	X_3, X_7
11	12	13	14	15
X_2, X_4	X_2, X_5	X_2, X_6	X_2, X_7	X_2, X_8
16	17	18	19	20
X_4, X_5	X_4, X_6	X_4, X_7	X_4, X_8	X_5, X_6
21	22	23	24	25
X_5, X_7	X_5, X_8	X_6, X_7	X_6, X_8	X_3, X_8
26	27	28	29	30
X_6, X_7	X_5, X_8	X_7, X_8	X_4, X_7	X_2, X_6

Таблица № 3 соответствия данных Таблицы № 1 и номера варианта по балльно-рейтинговой системе

1	2	3	4	5
X_1, X_2, X_3	X_1, X_2, X_4	X_1, X_2, X_5	X_1, X_2, X_6	X_1, X_2, X_7
6	7	8	9	10
X_1, X_2, X_8	X_1, X_3, X_4	X_1, X_3, X_5	X_1, X_3, X_6	X_1, X_3, X_7
11	12	13	14	15
X_2, X_3, X_4	X_2, X_3, X_5	X_2, X_3, X_6	X_2, X_3, X_7	X_2, X_3, X_8
16	17	18	19	20
X_3, X_4, X_5	X_3, X_4, X_6	X_3, X_4, X_7	X_3, X_4, X_8	X_4, X_5, X_6
21	22	23	24	25
X_4, X_5, X_7	X_4, X_5, X_8	X_4, X_6, X_7	X_4, X_6, X_8	X_1, X_3, X_8
26	27	28	29	30
X_5, X_6, X_7	X_5, X_6, X_8	X_6, X_7, X_8	X_4, X_7, X_8	X_2, X_4, X_6

Глава 3

Обязательные требования к самостоятельной работе

1. Работа начинается с титульного листа.

2. Номер варианта задания №1 и №2 выбираются в соответствии с номером в балльно-рейтинговой системе при разбивке списка группы по алфавиту.

3. К каждой задаче необходимо представить текстовую формулировку на профессиональную тему, сочинённую самостоятельно.

4. В решении должны присутствовать все пункты исследования, приведённые в данной показательной работе Главы 1.

5. Решение должно содержать развёрнутый ответ.

Оглавление