

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

S01

Да се изчислят стандартизираните показатели за леталитет в две областни болници. За стандарт да се приеме структурата по отделения в болница Б. Закръгляйте до 3-тия знак след десетичната запетая!

Отделение	Болница "А"		Болница "Б"	
	Преминали	Починали	Преминали	Починали
Инфекциозно	500	10	1500	35
Хирургия	1500	50	2000	70
Онкология	1000	85	1500	120

Решение:

Стандартизация - S01 Решение

Изчисляване на нестандартизирания леталитет

Отделение	Болница "А"			Болница "Б"		
	Преминали	Починали	НС	Преминали	Починали	НС
Инфекциозни болести	500	10	0.020	1500	35	0.023
Хирургия	1500	50	0.033	2000	70	0.035
Онкология	1000	85	0.085	1500	120	0.080
Общо	3000	145	0.048	5000	225	0.045

Изчисляване на стандарт

Преминали през "Б"	Структура (S)
1500	0.3
2000	0.4
1500	0.3
5000	1.0

Стандартизация

Отделение	Болница "А"				Болница "Б"			
	Преминали	Починали	НС	С	Преминали	Починали	НС	С
Инфекциозни болести	500	10	0.020	0.006	1500	35	0.023	0.007
Хирургия	1500	50	0.033	0.013	2000	70	0.035	0.014
Онкология	1000	85	0.085	0.026	1500	120	0.080	0.024
Общо	3000	145	0.048	0.045	5000	225	0.045	0.045

Заключение

След стандартизация за болница Б - леталитетът и в двете болници е еднакъв - 4,5%

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

S02

Да се изчислят стандартизираните показатели за леталитет в две областни болници. За стандарт да се приеме структурата на преминалите болни в болница А.

Отделение	Болница "А"		Болница "Б"	
	Преминали	Починали	Преминали	Починали
Рак на стомаха	400	120	200	80
Инсулт	500	300	100	70
Пневмония	100	40	650	250

Решение:

S02 Решение

Изчисляване на нестандартизирания леталитет

Отделение	Болница "А"			Болница "Б"		
	Преминали	Починали	НС	Преминали	Починали	НС
Рак на стомаха	400	120	0.300	200	80	0.400
Инсулт	500	300	0.600	100	70	0.700
Пневмония	100	40	0.400	650	250	0.385
Общо	1000	460	0.460	950	400	0.421

Изчисляване на стандарт

Стандарт Болница А	
Преминали	Структура (S)
400	0.400
500	0.500
100	0.100
1000	1.000

Стандартизация

Отделение	Болница "А"				Болница "Б"			
	Преминали	Починали	НС	С	Преминали	Починали	НС	С
Рак на стомаха	400	120	0.300	0.120	200	80	0.400	0.16
Инсулт	500	300	0.600	0.300	100	70	0.700	0.35
Пневмония	100	40	0.400	0.040	650	250	0.385	0.039
Общо	1000	460	0.460	0.460	950	400	0.421	0.549

Заключение

След стандартизация за болница А - леталитетът в Болница Б е по-висок с 8,9 %

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

СИ1

500 пациенти с множествена склероза са включени в клинично проучване. 30 от 250 лекувани с експериментална терапия рецидивират в рамките на 6 месеца след края на терапевтичния курс, докато при конвенционалната терапия този брой е 50 от 350. Изградете 95% -овите интервали на доверителност за относителния дял на рецидивите в двете групи и проверете дали има статистически значима разлика между тях, като използвате метода на препокриването.

Решение:

CI1 Решение

Изчисляване на относителните дялове

За експерименталната терапия:

$$\hat{p}_1 = \frac{30}{250} = 0.12 = 12\%$$

За конвенционалната терапия:

$$\hat{p}_2 = \frac{50}{350} = 0.14 = 14\%$$

Изчисляване на средната грешка

$$SE_p = \sqrt{\frac{\hat{p} \cdot (100 - \hat{p})}{n}}$$

За експерименталната терапия:

$$SE_{p1} = \sqrt{\frac{12\% \cdot (100 - 12\%)}{250}} = \sqrt{\frac{12\% \cdot (88\%)}{250}} = \sqrt{\frac{1056}{250}} = \sqrt{4,224} = 2,06\%$$

За конвенционалната терапия:

$$SE_{p2} = \sqrt{\frac{14\% \cdot (100 - 14\%)}{350}} = \sqrt{\frac{14\% \cdot (86\%)}{350}} = \sqrt{\frac{1204}{350}} = \sqrt{3,44} = 1,85\%$$

Изчисляване на 95% интервал на доверителност

$$p \pm z * SE_p$$

За експерименталната терапия:

$$12\% \pm 1,96 * 2,06\% = 12\% \pm 4,04\% = (7,96\%; 16,04\%)$$

За конвенционалната терапия:

$$14\% \pm 1,96 * 1,85\% = 14\% \pm 3,62\% = (10,38\%; 17,62\%)$$

Заклучение

Интервалите не се препокриват, следователно има статистически значима разлика между двете групи.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

ТТ1

500 пациенти с множествена склероза са включени в клинично проучване. 10 от 250 лекувани с експериментална терапия рецидивират в рамките на 6 месеца след края на терапевтичния курс, докато при конвенционалната терапия този брой е 70 от 350. Използвайте модифициран t тест за да проверите дали има статистически значима разлика между двете терапии.

Решение:

ТТ1 Решение

Изчисляване на относителните дялове

За експерименталната терапия:

$$\hat{p}_1 = \frac{10}{250} = 0.04 = 4\%$$

За конвенционалната терапия:

$$\hat{p}_2 = \frac{70}{350} = 0.20 = 20\%$$

Изчисляване на средната грешка

$$SE_p = \sqrt{\frac{\hat{p} \cdot (100 - \hat{p})}{n}}$$

За експерименталната терапия:

$$SE_{p1} = \sqrt{\frac{4\% \cdot (100 - 4\%)}{250}} = \sqrt{\frac{4\% \cdot (96\%)}{250}} = \sqrt{\frac{384}{250}} = \sqrt{1,536} = 1,24\%$$

За конвенционалната терапия:

$$SE_{p2} = \sqrt{\frac{20\% \cdot (100 - 20\%)}{350}} = \sqrt{\frac{20\% \cdot (80\%)}{350}} = \sqrt{\frac{1600}{350}} = \sqrt{4,57} = 2,14\%$$

Изчисляване на t статистиката

$$t = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{SE_{p1}^2 + SE_{p2}^2}}$$

$$t = \frac{|4\% - 20\%|}{\sqrt{1,24^2 + 2,14^2}} = \frac{|-16\%|}{\sqrt{1,54 + 4,58}} = \frac{16\%}{\sqrt{6,12}} = \frac{16\%}{2,47} = 6,48$$

Определяне на p и заключение

При $t = 6,48$; $p < 0,001$. Следователно отхвърляме нулевата хипотеза и приемаме, че има статистически значима разлика между двете терапии.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

СИ2

800 пациенти с белодробен карцином са включени в клинично проучване. 300 от 400 лекувани с експериментална терапия преживяват поне 1 година след края на терапевтичния курс, докато при конвенционалната терапия този брой е 200 от 400. Изградете 95% -овите интервали на доверителност за относителния дял на преживелите в двете групи и проверете дали има статистически значима разлика между тях, като използвате метода на препокриването.

Решение:

CI2 Решение

Изчисляване на относителните дялове

За експерименталната терапия:

$$\hat{p}_1 = \frac{300}{400} = 0.75 = 75\%$$

За конвенционалната терапия:

$$\hat{p}_2 = \frac{200}{400} = 0.50 = 50\%$$

Изчисляване на средната грешка

$$SE_p = \sqrt{\frac{\hat{p} \cdot (100 - \hat{p})}{n}}$$

За експерименталната терапия:

$$SE_{p1} = \sqrt{\frac{75\% \cdot (100 - 75\%)}{400}} = \sqrt{\frac{75\% \cdot (25\%)}{400}} = \sqrt{\frac{1875}{400}} = \sqrt{4,6875} = 2,16\%$$

За конвенционалната терапия:

$$SE_{p2} = \sqrt{\frac{50\% \cdot (100 - 50\%)}{400}} = \sqrt{\frac{50\% \cdot (50\%)}{400}} = \sqrt{\frac{2500}{400}} = \sqrt{6,25} = 2,5\%$$

Изчисляване на 95% интервал на доверителност

$$p \pm z * SE_p$$

За експерименталната терапия:

$$75\% \pm 1,96 * 2,16\% = 75\% \pm 4,24\% = (70,76\%; 79,24\%)$$

За конвенционалната терапия:

$$50\% \pm 1,96 * 2,5\% = 50\% \pm 4,9\% = (45,1\%; 54,9\%)$$

Заклучение

Интервалите не се препокриват, следователно има статистически значима разлика между двете групи.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

ТТ2

800 пациенти с белодробен карцином са включени в клинично проучване. 300 от 400 лекувани с експериментална терапия преживяват поне 1 година след края на терапевтичния курс, докато при конвенционалната терапия този брой е 200 от 400. Използвайте модифициран t тест за да проверите дали има статистически значима разлика между двете терапии.

Решение:

ТТ2 Решение

Изчисляване на относителните дялове

За експерименталната терапия:

$$\hat{p}_1 = \frac{300}{400} = 0.75 = 75\%$$

За конвенционалната терапия:

$$\hat{p}_2 = \frac{200}{400} = 0.50 = 50\%$$

Изчисляване на средната грешка

$$SE_p = \sqrt{\frac{\hat{p} \cdot (100 - \hat{p})}{n}}$$

За експерименталната терапия:

$$SE_{p1} = \sqrt{\frac{75\% \cdot (100 - 75\%)}{400}} = \sqrt{\frac{75\% \cdot (25\%)}{400}} = \sqrt{\frac{1875}{400}} = \sqrt{4,6875} = 2,16\%$$

За конвенционалната терапия:

$$SE_{p2} = \sqrt{\frac{50\% \cdot (100 - 50\%)}{400}} = \sqrt{\frac{50\% \cdot (50\%)}{400}} = \sqrt{\frac{2500}{400}} = \sqrt{6,25} = 2,5\%$$

Изчисляване на t статистиката

$$t = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{SE_{p1}^2 + SE_{p2}^2}}$$

$$t = \frac{|75\% - 50\%|}{\sqrt{2,16^2 + 2,5^2}} = \frac{|25\%|}{\sqrt{4,68 + 6,25}} = \frac{25\%}{\sqrt{10,93}} = \frac{25\%}{3,31} = 7,55$$

Определяне на p и заключение

При $t = 7,55$; $p < 0,001$. Следователно отхвърляме нулевата хипотеза и приемаме, че има статистически значима разлика между двете терапии.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

СИЗ

Да се изчисли 95% доверителен интервал за средния популационен инкубационен период при пациенти с хепатит А.

Инкубационен период [дни]	Честота f [брой пациенти]
1 – 15	80
16 – 30	240
31 – 45	150
46 – 60	28
61 – 75	10
Общо	250

Решение:

СІЗ Решение

Изчисляване на средната аритметична

Използваме формулата за средна аритметична при групирани данни в интервален ред:

$$\bar{x} = \frac{\sum x_u \cdot f}{\sum f}$$

Инкубационен период [дни]	Честота f [брой пациенти]	x_u	$x_u \cdot f$
1 – 15	80	8	640
16 – 30	240	23	5520
31 – 45	150	38	5700
46 – 60	28	53	1484
61 – 75	10	68	680
Общо	250		14024

$$\bar{x} = \frac{\sum x_u \cdot f}{\sum f} = \frac{14024}{250} = 56,1$$

Изчисляване на стандартното отклонение

$$SD = \sqrt{\frac{\sum (x_u - \bar{x})^2 \cdot f}{\sum f}}$$

Инкубационен период [дни]	Честота f [брой пациенти]	x_u	$x_u \cdot f$	$x_u - \bar{x}$	$(x_u - \bar{x})^2$	$(x_u - \bar{x})^2 \cdot f$
1 – 15	80	8	640	-48.1	2313.23	185058.4
16 – 30	240	23	5520	-33.1	1095.35	262884
31 – 45	150	38	5700	-18.1	327.47	49120.5
46 – 60	28	53	1484	-3.1	9.59	268.52
61 – 75	10	68	680	11.9	141.71	1417.1
Общо	250		14024			498748.52

$$SD = \sqrt{\frac{\sum (x_u - \bar{x})^2 \cdot f}{\sum f}} = \sqrt{\frac{498748.52}{250}} = \sqrt{1994.99} = 44.67$$

Изчисляване на стандартната грешка

$$SE_m = \frac{44.67}{\sqrt{250}} = \frac{44.67}{15.81} = 2.83$$

Изчисляване на 95% доверителен интервал

$$\bar{x} \pm z * SE_m = 56.1 \pm 1.96 * 2.83 = 56.1 \pm 5.55 = (50.55; 61.65)$$

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

СИ4

Да се изчисли 95% доверителен интервал за средната популационна преживяемост при пациенти с белодробен карцином, лекувани с експериментална терапия.

Преживяемост	Честота f
1 – 10	30
11-20	60
21-30	90
31-40	70
41-50	30
Общо	280

Решение:

СИ4 Решение

Изчисляване на средната аритметична

Използваме формулата за средна аритметична при групирани данни в интервален ред:

$$\bar{x} = \frac{\sum x_u \cdot f}{\sum f}$$

Преживяемост	Честота f	x_u	$x_u \cdot f$
1 – 10	30	5,5	165
11-20	60	15,5	930
21-30	90	25,5	2295
31-40	70	35,5	2485
41-50	30	45,5	1365
Общо	280		7240

$$\bar{x} = \frac{\sum x_u \cdot f}{\sum f} = \frac{7240}{280} = 25,86$$

Изчисляване на стандартното отклонение

$$SD = \sqrt{\frac{\sum (x_u - \bar{x})^2 \cdot f}{\sum f}}$$

Преживяемост	Честота f	x_u	$x_u \cdot f$	$x_u - \bar{x}$	$(x_u - \bar{x})^2$	$(x_u - \bar{x})^2 \cdot f$
1 – 10	30	5,5	165	-20.36	414.02	12420.6
11-20	60	15,5	930	-10.36	107.39	6443.4
21-30	90	25,5	2295	-0.36	0.13	11.7
31-40	70	35,5	2485	9.64	93.02	6511.4
41-50	30	45,5	1365	19.64	385.6	11568
Общо	280		7240			36955,1

$$SD = \sqrt{\frac{\sum (x_u - \bar{x})^2 \cdot f}{\sum f}} = \sqrt{\frac{36955,1}{280}} = \sqrt{132} = 11.5$$

Изчисляване на стандартната грешка

$$SE_m = \frac{11.5}{\sqrt{280}} = \frac{11.5}{16.73} = 0.69$$

Изчисляване на 95% доверителен интервал

$$\bar{x} \pm z * SE_m = 25.86 \pm 1.96 * 0.69 = 25.86 \pm 1.35 = (24.51; 27.21)$$

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

ХТ1

Ретроспективно проучване анализа наблюдаваните усложнения след аборт, извършен по два различни метода – класически и аспирационен. Има ли зависимост между метода за аборт и вида на усложненията след това?

Метод за аборт	Кръвотечение	Инфекция	Инфертелитет
Класически	26	32	42
Аспирационен	27	55	18

Решение:

ХТ1 Решение

Изчисляване на очакваните честоти

Използваме формулата за очаквана честота:

$$f_{tcell} = \frac{\sum_{columns} \cdot \sum_{rows}}{\sum_{total}}$$

Метод за аборт	Кръвотечение	Инфекция	Инфертелитет	по редове
Класически	27	44	30	100
Аспирационен	27	44	30	100
по колони	53	87	60	200

Изчисляване на статистиката на Пиърсн

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

$$\begin{aligned}\chi^2 &= \frac{(26 - 27)^2}{27} + \frac{(32 - 44)^2}{44} + \frac{(42 - 30)^2}{30} + \frac{(27 - 27)^2}{27} + \frac{(55 - 44)^2}{44} + \frac{(18 - 30)^2}{30} = \\ &= \frac{(1)^2}{27} + \frac{(-12)^2}{44} + \frac{(12)^2}{30} + \frac{(0)^2}{27} + \frac{(11)^2}{44} + \frac{(-12)^2}{30} = \\ &= \frac{1}{27} + \frac{144}{44} + \frac{144}{30} + \frac{0}{27} + \frac{121}{44} + \frac{144}{30} = 0,003 + 3,27 + 4,8 + 0 + 2,75 + 4,8 = 15,63\end{aligned}$$

Определяне на степените на свобода

Степените на свобода се определят по формулата:

$$df = (r - 1)(c - 1)$$

където r е броят на редовете, а c е броят на колоните.

В нашия случай имаме 2 реда и 3 колони, следователно:

$$df = (2 - 1)(3 - 1) = 2$$

Определяне на p и заключение

При $\chi^2 = 15,63$ и степен на свобода = 2. От таблицата за разпределението на χ^2 се определя $p < 0,05$. Следователно отхвърляме нулевата хипотеза и приемаме, че има статистически значима зависимост между метода за аборт и вида на усложненията след това.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

ХТ2

Ретроспективно проучване анализира честотата на неврологични проблеми при IT специалисти. Има ли зависимост между продължителността на стажа при IT специалисти и наличието на неврологични проблеми?

Продължителност на стажа	С проблеми	Без проблеми
До 5 г.	10	35
6 – 10 г.	14	30
11 – 15 г.	36	25
Над 15 г.	40	10

Решение:

ХТ2 Решение

Изчисляване на очакваните честоти

Използваме формулата за очаквана честота:

$$f_{tcell} = \frac{\sum_{columns} \cdot \sum_{rows}}{\sum_{total}}$$

Продължителност на стажа	С проблеми	Без проблеми	по редове
До 5 г.	23	23	45
6 – 10 г.	22	22	44
11 – 15 г.	31	31	61
Над 15 г.	25	25	50
по колонии	100	100	200

Изчисляване на статистиката на Пиърсън

$$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t}$$

$$\begin{aligned}\chi^2 &= \frac{(10 - 23)^2}{23} + \frac{(14 - 22)^2}{22} + \frac{(36 - 31)^2}{31} + \frac{(40 - 25)^2}{25} + \frac{(35 - 23)^2}{23} \\ &\quad + \frac{(30 - 22)^2}{22} + \frac{(25 - 31)^2}{31} + \frac{(10 - 25)^2}{25} = \\ &= \frac{(-13)^2}{23} + \frac{(-8)^2}{22} + \frac{(5)^2}{31} + \frac{(15)^2}{25} + \frac{(12)^2}{23} + \frac{(8)^2}{22} + \frac{(-6)^2}{31} + \frac{(-15)^2}{25} =\end{aligned}$$

$$= \frac{169}{23} + \frac{64}{22} + \frac{25}{31} + \frac{225}{25} + \frac{144}{23} + \frac{64}{22} + \frac{36}{31} + \frac{225}{25} = 7,35 + 2,91 + 0,81 + 9 + 6,26 + 2,91 + 1,16 + 9 = 39,5$$

Определяне на степените на свобода

Степените на свобода се определят по формулата:

$$df = (r - 1)(c - 1)$$

където r е броят на редовете, а c е броят на колоните.

В нашия случай имаме 4 реда и 2 колони, следователно:

$$df = (4 - 1)(2 - 1) = 3$$

Определяне на p и заключение

При $\chi^2 = 39,5$ и степен на свобода = 3. От таблицата за разпределението на χ^2 се определя $p < 0,001$. Следователно отхвърляме нулевата хипотеза и приемаме, че има статистически значима зависимост между продължителността на стажа при ИТ специалисти и наличието на неврологични проблеми.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

КОРП1

Използвайки данните от таблицата по-долу, да се измери силата и посоката на зависимостта между възрастта при поставяне на диагноза и преживяемостта при пациенти с белодробен карцином.

Пациент №	Възраст при диагноза X [години]	Преживяемост Y [месеци]
1	60	112
2	70	95
3	75	72
4	78	45
5	77	15
6	80	20
7	85	12

Решение:

КОРП1 Решение

Изчисляване на средната аритметична

Използваме формулата за средна аритметична при негрупирани данни:

За възрастта при диагноза:

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{60 + 70 + 75 + 78 + 77 + 80 + 85}{7} = \frac{525}{7} = 75$$

За преживяемостта:

$$\bar{y} = \frac{\sum y}{n}$$

$$\bar{y} = \frac{112 + 95 + 72 + 45 + 15 + 20 + 12}{7} = \frac{371}{7} = 53$$

Изчисляваме отклоненията dx и dy , техните произведения и квадратите им

Пациент №	Възраст	Преживяемост Y	dx	dy	$dx \cdot dy$	dx^2	dy^2
1	60	112	-15	59	-885	225	3481
2	70	95	-5	42	-210	25	1764
3	75	72	0	19	0	0	361
4	78	45	3	-8	-24	9	64
5	77	15	2	-38	-76	4	1444
6	80	20	5	-33	-165	25	1089
7	85	12	10	-41	-410	100	1681
средна	75	53	суми:		-1770	388	9884

Изчисляване на коефициента на корелация

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$\rho = \frac{-1770}{\sqrt{388 \cdot 9884}} = \frac{-1770}{\sqrt{3834032}} = \frac{-1770}{1958,3} = -0.904$$

Съществува силна негативна корелация между възрастта при поставяне на диагноза и преживяемостта при пациенти с белодробен карцином.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

КОРП2

В таблицата по-долу са отчетени резултатите от изминатото разстояние при 6-минутен тест с ходене (6MWT) в километри и нивото на лабораторния показател NT-proBNP (pg/ul) сред 7 пациенти с начална сърдечна недостатъчност. Да се измери силата и посоката на зависимостта между изминатото разстояние при 6-минутен тест с ходене и нивото на лабораторния показател NT-proBNP

Пациент №	Изминато разстояние X [метри]	Ниво на NT-proBNP Y [pg/ul]
1	120	26
2	220	35
3	300	35
4	450	42
5	500	56
6	610	74
7	642	82

Решение:

КОРП2 Решение

Изчисляване на средната аритметична

Използваме формулата за средна аритметична при негрупиран данни:

За изминатото разстояние при 6-минутен тест с ходене:

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{120 + 220 + 300 + 450 + 500 + 610 + 642}{7} = \frac{2842}{7} = 406$$

За нивото на лабораторния показател NT-proBNP:

$$\bar{y} = \frac{\sum y}{n}$$

$$\bar{y} = \frac{26 + 35 + 35 + 42 + 56 + 74 + 82}{7} = \frac{350}{7} = 50$$

Изчисляваме отклоненията dx и dy , техните произведения и квадратите им

Пациент №	Разстояние X	NT-proBNP Y	dx	dy	$dx \cdot dy$	dx^2	dy^2
1	120	26	-286	-24	6864	81796	576
2	220	35	-186	-15	2790	34596	225
3	300	35	-106	-15	1590	11236	225
4	450	42	44	-8	-352	1936	64
5	500	56	94	6	564	8836	36
6	610	74	204	24	4896	41616	576
7	642	82	236	32	7552	55696	1024
средна	406	50	суми:		23904	235712	2726

Изчисляване на коефициента на корелация

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$\rho = \frac{23904}{\sqrt{235712 \cdot 2726}} = \frac{23904}{\sqrt{642982912}} = \frac{23904}{25338,5} = 0.943$$

Съществува силна позитивна корелация между изминатото разстояние при 6-минутен тест с ходене и нивото на лабораторния показател NT-proBNP.

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

РЕГ1

Използвайки данните от таблицата по-долу, да се състави модел на зависимостта между възрастта при поставяне на диагноза и преживяемостта при пациенти с белодробен карцином.

Пациент №	Възраст при диагноза X [години]	Преживяемост Y [месеци]
1	60	112
2	70	95
3	75	72
4	78	45
5	77	15
6	80	20
7	85	12

Решение:

РЕГ1 Решение

Изчисляване на $\sum X$, $\sum Y$, $\sum XY$

Пациент №	Възраст X	Преживяемост Y	XY	X ²
1	60	112	6720	3600
2	70	95	6650	4900
3	75	72	5400	5625
4	78	45	3510	6084
5	77	15	1155	5929
6	80	20	1600	6400
7	85	12	1020	7225
сума	525	371	26055	39763
средни	75	53		

Изчисляване на регресионния коефициент b

$$\beta = \frac{(n \cdot \sum (X \cdot Y)) - (\sum X \cdot \sum Y)}{(n \cdot \sum X^2) - (\sum X)^2}$$

$$\beta = \frac{(7 \cdot 26055) - (525 \cdot 371)}{(7 \cdot 39763) - (525)^2} = \frac{182385 - 194775}{278341 - 275625} = \frac{-12390}{2716} = -4,56$$

Изчисляване на константата a

$$a = \bar{y} - (\beta \cdot \bar{x})$$

$$a = 53 - (-4,56 \cdot 75) = 53 + 342,00 = 395,00$$

Регресионният модел е във вида:

$$y = 395,00 - 4,56x$$

Колоквиум

Медицинска статистика

Име _____, Специалност: _____

Дата: _____, Група: _____, Фак.номер _____

РЕГ2

В таблицата по-долу са отчетени резултатите от изминатото разстояние при 6-минутен тест с ходене (6MWT) в километри и нивото на лабораторния показател NT-proBNP (pg/ul) сред 7 пациенти с начална сърдечна недостатъчност. Да се състави модел на зависимостта между изминатото разстояние при 6-минутен тест с ходене и нивото на лабораторния показател NT-proBNP.

Пациент №	Изминато разстояние X [метри]	Ниво на NT-proBNP Y [pg/ul]
1	120	26
2	220	35
3	300	35
4	450	42
5	500	56
6	610	74
7	642	82

Решение:

РЕГ2 Решение

Изчисляване на $\sum X$, $\sum Y$, $\sum XY$

Пациент №	Разстояние X	NT-proBNP Y	XY	X^2
1	120	26	3120	14400
2	220	35	7700	48400
3	300	35	10500	90000
4	450	42	18900	202500
5	500	56	28000	250000
6	610	74	45140	372100
7	642	82	52644	412164
сума	2842	350	166004	1389564
средни	406	50		

Изчисляване на регресионния коефициент b

$$\beta = \frac{(n \cdot \sum (X \cdot Y)) - (\sum X \cdot \sum Y)}{(n \cdot \sum X^2) - (\sum X)^2}$$

$$\beta = \frac{(7 \cdot 166004) - (2842 \cdot 350)}{(7 \cdot 1389564) - (2842)^2} = \frac{1162028 - 994700}{9736948 - 8085764} = \frac{167328}{1651184} = 0,101$$

Изчисляване на константата a

$$a = \bar{y} - (\beta \cdot \bar{x})$$

$$a = 50 - (0,101 \cdot 406) = 50 - 41 = 9$$

Регресионният модел е във вида:

$$y = 9 + 0,101x$$