

ОСНОВЫ МЕДИЦИНСКОЙ ИНФОРМАТИКИ

Выполнили:

Шейнина Е.М.

Сержант О.В

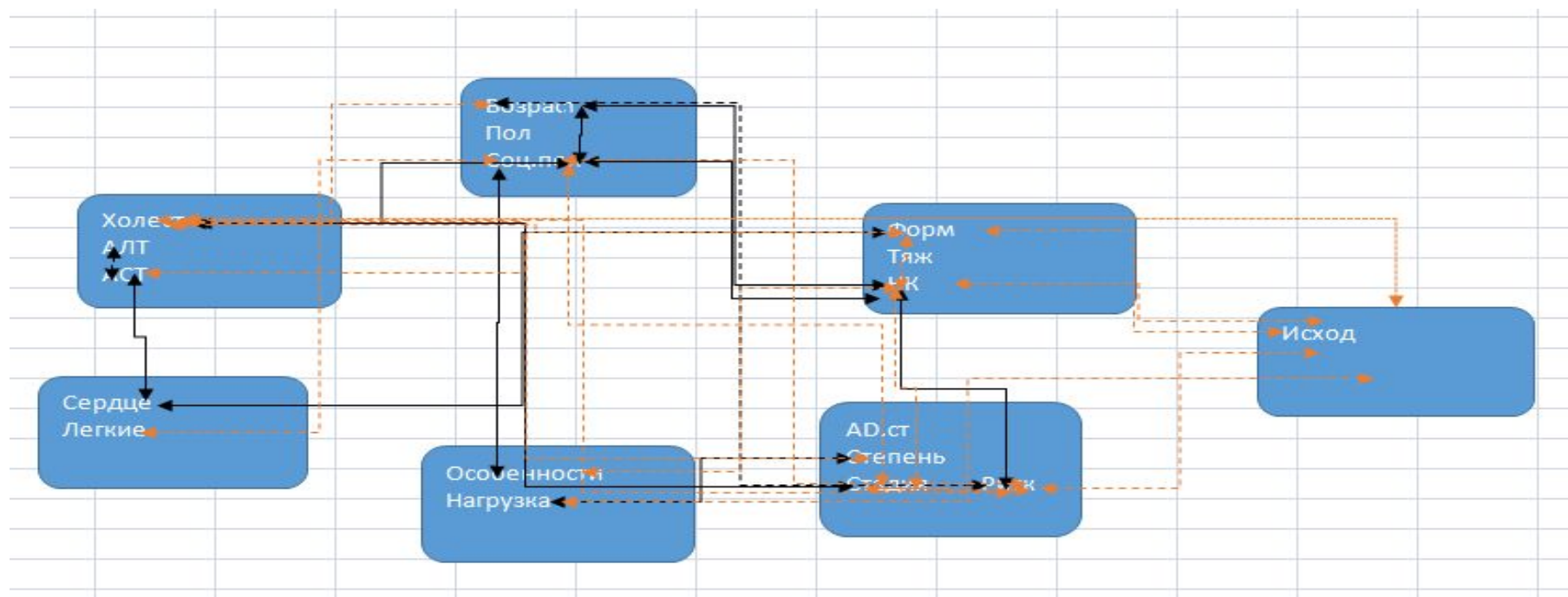


ЧТО МЫ НАУЧИЛИСЬ ДЕЛАТЬ С ПОМОЩЬЮ ПРОГРАММЫ EXCEL:

1. Составлению графических схем связей между данными.
2. Создавать таблицы
3. Рассчитывать основные параметры математической статистики на основе этих таблиц
4. Производить корреляционный анализ
5. Рассчитывать регрессию
6. Создавать модели, основанные на ней



ГРАФИЧЕСКАЯ СХЕМА СВЯЗЕЙ МЕЖДУ ДАННЫМИ

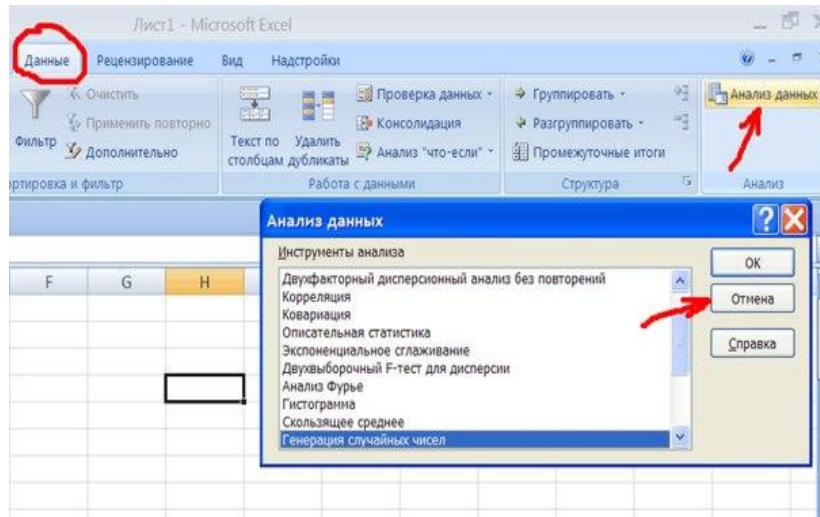


СОЗДАНИЕ ТАБЛИЦЫ

№	Пациент				ИБС			АГ				ВЭМ		RG		Лаб.обследования			Исход
Атрибут	ФИО	Возраст	Пол	Соц.пол.	Форм	Тяж	НК	AD сист.	Степень	Стадия	Риск	Особ.	Нагрузка	Сердце	Лёгкие	Холест.	АЛТ	АСТ	Исход
1	Иванов А	48	1	2	1	3	3	0	2	2	3	2	75	1	2	7,1	30	35	2
2	Вахрушев	54	0	5	1	2	2	0	0	0	2	1	50	1	1	5,2	30	38	2
3	Петров И	60	1	1	5	4	4	0	2	2	3	4	25	3	3	7,5	75	80	5
4	Сидоров	75	1	1	3	3	4	0	3	0	3	4	25	3	3	7,3	30	48	5
5	Никифор	35	1	4	3	2	2	0	3	0	0	2	50	3	1	5,7	40	50	1
6	Ломаева	80	0	1	3	2	4	0	2	2	3	4	25	3	3	7,3	42	43	1
7	Никодим	59	1	4	3	3	3	0	3	2	3	3	50	3	2	6,7	48	45	3
8	Азатов А.	46	1	4	2	4	2	0	2	0	0	3	75	3	3	5,7	75	80	1
9	Каюмов Р	50	1	5	2	2	2	0	0	0	0	0	100	1	1	5,7	35	38	1
10	Карманов	49	1	3	3	3	2	0	0	0	0	2	25	3	3	6,7	30	40	3
11	Алиева А	40	0	5	2	4	3	0	0	0	0	2	50	1	1	5	32	30	1
12	Байкузин	47	1	4	1	3	2	0	0	0	0	2	75	1	3	5	34	38	1
13	Тихомир	50	1	4	1	3	2		0	0	0	3	75	1	1	6	28	25	1
14	Куркова А	49	0	4	1	3	2	0	0	0	0	3	50	1	1	5,8	28	25	2
15	Клебанов	53	0	4	2	4	3	0	0	0	3	3	25	1	2	5,8	29	28	3



РАСЧЁТ ОСНОВНЫХ ПАРАМЕТРОВ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ НА ОСНОВЕ ТАБЛИЦ



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	№	Пациент			ИБС			АГ			ВЭМ		РГ		Лаб.исследования			Исход			
2	Атрибут	ФИО	Возраст	Пол	Соц.пол.	Форм	Тяж	НК	АД сист.	Степень	Стадия	Риск	Особ.	Нагрузка	Сердце	Лёгкие	Холест.	АЛТ	АСТ	Исход	
3	1	Иванов А.	48	1	2	1	3	3	200/120	2	2	3	2	75	1	2	7,1	30	35	2	
4	2	Вахрушев	54	0	5	1	2	2	0	0	0	2	1	50	1	1	5,2	30	38	2	
5	3	Петров И.	60	1	1	5	4	4	0	2	2	3	4	25	3	3	7,5	75	80	5	
6	4	Сидоров	75	1	1	3	3	4	0	3	0	3	4	25	3	3	7,3	30	48	5	
7	5	Никифоро	35	1	4	3	2	2	0	3	0	0	2	50	3	1	5,7	40	50	1	
8	6	Ломаяева	80	0	1	3	2	4	0	2	2	3	4	25	3	3	7,3	42	43	1	
9	7	Никодим	59	1	4	3	3	3	0	3	2	3	3	50	3	2	6,7	48	45	3	
10	8	Азатов А.	46	1	4	2	4	2	0	2	0	0	3	75	3	3	5,7	75	80	1	
11	9	Каюмов Р.	50	1	5	2	2	2	0	0	0	0	0	100	1	1	5,7	35	38	1	
12	10	Карманов	49	1	3	3	3	3	2	90/60	0	0	0	2	25	3	3	6,7	30	40	3
13	11	Алиева А.	40	0	5	2	4	3	0	0	0	0	2	50	1	1	5	32	30	1	
14	12	Байкузин	47	1	4	1	3	2	0	0	0	0	2	75	1	3	5	34	38	1	
15	13	Тихомир	50	1	4	1	3	2	0	0	0	0	3	75	1	1	6	28	25	1	
16	14	Курцова А.	49	0	4	1	3	2	0	0	0	0	3	50	1	1	5,8	28	25	2	
17	15	Клебанов	53	0	4	2	4	3	0	0	0	3	3	25	1	2	5,8	29	28	3	

Столбец1		Столбец1	
Среднее	53	Среднее	6,166667
Стандартн	3,048809	Стандартн	0,222468
Медиана	50	Медиана	5,8
Мода	50	Мода	5,7
Стандартн	11,80799	Стандартн	0,861615
Дисперси	139,4286	Дисперси	0,742381
Экссесс	1,363003	Экссесс	-1,34895
Асимметр	1,101123	Асимметр	0,246215
Интервал	45	Интервал	2,5
Минимум	35	Минимум	5
Максимум	80	Максимум	7,5
Сумма	795	Сумма	92,5
Счет	15	Счет	15
Уровень н	6,539046	Уровень н	0,477147

С помощью функции анализ данных мы нашли такие параметры как среднее, стандартная ошибка, медиана, мода, стандартное отклонение, дисперсия выборки, эксцесс, асимметричность, интервал, минимум, максимум, сумма, счёт.



КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

-это метод, позволяющий
обнаружить зависимость между
несколькими случайными
величинами.



Допустим, проводится независимое измерение различных параметров у одного типа объектов. Из этих данных можно получить качественно новую информацию - о взаимосвязи этих параметров. Для этого вводится коэффициент корреляции. Это величина, характеризующая направление и силу связи между признаками. Одним числом дает представление о направлении и силе связи между признаками (явлениями), пределы его колебаний от 0 до ± 1 .



Сила корреляционной связи:

сильная: $\pm 0,7$ до ± 1

средняя: $\pm 0,3$ до $\pm 0,699$

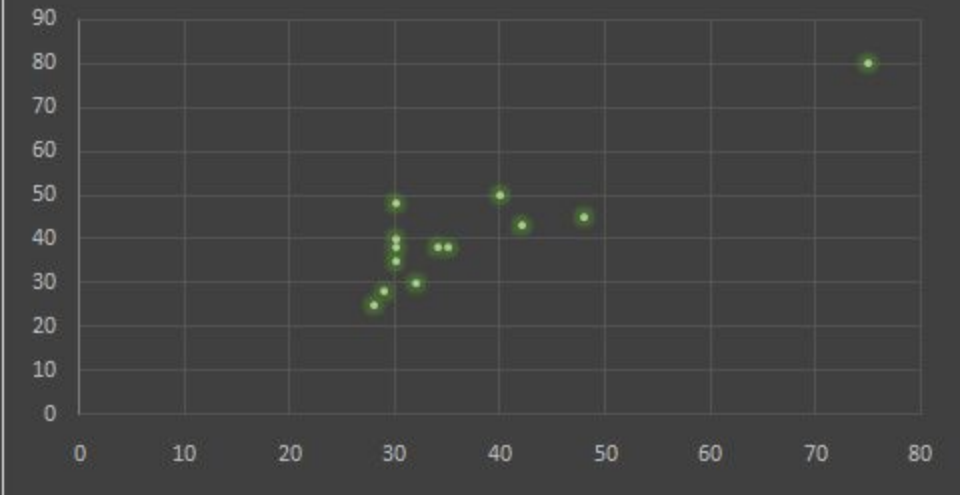
слабая: 0 до $\pm 0,299$



С ПОМОЩЬЮ ФУНКЦИИ АНАЛИЗ ДАННЫХ,
КОРРЕЛЯЦИЯ, МЫ СОЗДАЛИ
КОРРЕЛЯЦИОННУЮ МАТРИЦУ, НАШЛИ
КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ МЕЖДУ
ОТДЕЛЬНЫМИ ПАРАМЕТРАМИ И
ПРЕДСТАВИЛИ ЭТО В ВИДЕ ДИАГРАММЫ
РАССЕИВАНИЯ. ТЕМ САМЫМ НАЙДЯ
ЗАВИСИМОСТЬ МЕЖДУ НА ПЕРВЫЙ ВЗГЛЯД
ОБОСОБЛЕННЫМИ ПОКАЗАТЕЛЯМИ
ЗДОРОВЬЯ, СОЦИАЛЬНЫМ ПОЛОЖЕНИЕМ
И ИСХОДОМ ЗАБОЛЕВАНИЯ БОЛЬНЫХ



АЛТ от АСТ



АЛТ	АСТ
30	35
30	38
75	80
30	48
40	50
42	43
48	45
75	80
35	38
30	40
32	30
34	38
28	25
28	25
29	28

Корреляция

Входные данные: \$B\$17:\$C\$31

Выходной интервал: \$D\$21

Группирование: по столбцам по строкам

Метки в первой строке

Параметры вывода: Выходной интервал Новый рабочий лист Новая рабочая книга

Buttons: OK, Отмена, Справка

Лист1 - Microsoft Excel

Меню: Данные, Рецензирование, Вид, Настройки

Панель инструментов: Фильтр, Дополнительно, Проверка данных, Группировать, Анализ данных

Панель задач: Работа с данными, Структура

Анализ данных

Инструменты анализа:

- Двухфакторный дисперсионный анализ без повторений
- Корреляция
- Ковариация
- Описательная статистика
- Экспоненциальное сглаживание
- Двухвыборочный F-тест для дисперсии
- Анализ Фурье
- Гистограмма
- Скользящее среднее
- Генерация случайных чисел

Buttons: OK, Отмена, Справка

	Возраст	Пол	Соц.пол.	Форм	Тяж	НК	АД сист.	Степень	Стадия	Риск	Особ.	Нагрузка	Сердце	Лёгкие	Холест.	АЛТ	АСТ	Исход
Возраст	1																	
Пол	-0,13637	1																
Соц.пол.	-0,70307	-0,20135	1															
Форм	0,374636	0,255377	-0,56561	1														
Тяж	-0,16005	-4,3E-17	-0,06498	0,164845	1													
НК	0,740869	-0,11952	-0,78213	0,610468	0,231455	1												
АД сист	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	#ДЕЛ/0!	1											
Степень	0,343805	0,412242	-0,55839	0,555098	-0,07257	0,51512	#ДЕЛ/0!	1										
Стадия	0,462538	0,1066	-0,60098	0,435572	-2,3E-17	0,637059	#ДЕЛ/0!	0,535332	1									
Риск	0,699524	-0,13047	-0,62388	0,374827	0,063161	0,799173	#ДЕЛ/0!	0,488941	0,695384	1								
Особ	0,585857	-0,04336	-0,70709	0,465024	0,419788	0,673657	#ДЕЛ/0!	0,484195	0,397467	0,480794	1							
Нагрузка	-0,4972	0,355371	0,541757	-0,59638	-0,19662	-0,60679	#ДЕЛ/0!	-0,23592	-0,20565	-0,51331	-0,6295	1						
Сердце	0,386568	0,377964	-0,55174	0,796318	0	0,395285	#ДЕЛ/0!	0,750735	0,342475	0,246557	0,524379	-0,49889	1					
Лёгкие	0,516175	0,316228	-0,68977	0,471084	0,306186	0,472456	#ДЕЛ/0!	0,355534	0,3371	0,360997	0,548408	-0,40135	0,597614	1				
Холест	0,684522	0,294486	-0,90081	0,621888	0,010967	0,724264	#ДЕЛ/0!	0,602754	0,712365	0,674203	0,608917	-0,48013	0,599337	0,546213	1			
АЛТ	0,077579	0,320675	-0,25827	0,587562	0,385854	0,236294	#ДЕЛ/0!	0,468577	0,385649	0,123886	0,362288	-0,03825	0,58719	0,457805	0,262533	1		
АСТ	0,136968	0,435587	-0,38102	0,649754	0,273722	0,265485	#ДЕЛ/0!	0,571676	0,290903	0,157127	0,330427	-0,12243	0,702704	0,561014	0,341761	0,934961	1	
Исход	0,481368	0,173344	-0,58636	0,602045	0,335679	0,600838	#ДЕЛ/0!	0,340408	0,273482	0,621924	0,493011	-0,64065	0,399659	0,438529	0,640056	0,187377	0,321827	1

Важную роль при исследовании взаимосвязей между статистическими выборками кроме корреляционного и дисперсионного анализа играет регрессионный анализ. Регрессия позволяет проанализировать воздействие на какую-либо зависимую переменную одной или более независимых переменных и позволяет установить модель этой зависимости.



Если рассматривается зависимость между одной зависимой переменной Y и несколькими независимыми X_1, X_2, \dots, X_n , то речь идет о множественной линейной регрессии. В этом случае уравнение регрессии имеет вид: $Y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n$, где a_1, a_2, \dots, a_n - коэффициенты при независимых переменных, которые нужно вычислить (коэффициенты регрессии), a_0 - константа.



При построении регрессионной модели важнейшими моментами являются оценка ее адекватности (эффективности) и значимости, на основании которых можно судить о возможности применения в практике полученной модели. Мерой оценки адекватности регрессионной модели является коэффициент детерминации R^2 (R-квадрат), который определяет, с какой степенью точности полученное уравнение регрессии аппроксимирует исходные данные. Значимость регрессионной модели оценивается с помощью критерия Фишера (F – критерия). Если величина F – критерия значима ($p < 0,05$), то регрессионная модель является значимой.

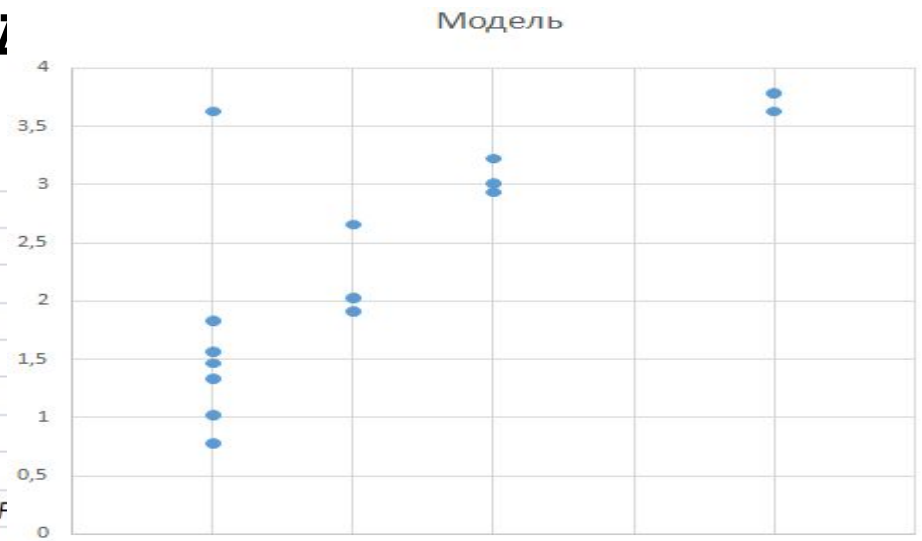


С помощью функции анализ данных, регрессия, мы научились строить модель основанную на регрессии, нашли коэффициент Фишера и R-квадрат. Тем самым получив предсказание исхода забот

Регрессионная статистика	
Множественный R	0,764711
R-квадрат	0,584783
Нормированный R-квадрат	0,418697
Стандартная ошибка	1,073095
Наблюдения	15

Дисперсионный анализ					
	df	SS	MS	F	Значимость F
Регрессия	4	16,21799	4,054499	3,520955	0,048428
Остаток	10	11,51534	1,151534		
Итого	14	27,73333			

Коэффициентная статистика - Значения нижние 95%, верхние 95%, средние 95%, нижние 95%								
Y-пересечение	-2,4389	6,285996	-0,38799	0,706155	-16,445	11,56718	-16,445	11,56718
Соц.пол.	0,197661	0,474465	0,416597	0,685774	-0,85951	1,254834	-0,85951	1,254834
Риск	0,200224	0,271627	0,737128	0,477983	-0,405	0,805447	-0,405	0,805447
Нагрузка	-0,024	0,014843	-1,617	0,136949	-0,05707	0,009071	-0,05707	0,009071
Холест.	0,790263	0,81826	0,965784	0,356928	-1,03294	2,61346	-1,03294	2,61346



Соц.пол.	Риск	Нагрузка	Холест.	Исход	Модель
2	3	75	7,1	2	2,659
5	2	50	5,2	2	2,028
1	3	25	7,5	5	3,785
1	3	25	7,3	5	3,627
4	0	50	5,7	1	1,833
1	3	25	7,3	1	3,627
4	3	50	6,7	3	3,223
4	0	75	5,7	1	1,333
5	0	100	5,7	1	1,023
3	0	25	6,7	3	2,933
5	0	50	5	1	1,47
4	0	75	5	1	0,78
4	0	75	6	1	1,57
4	0	50	5,8	2	1,912
4	3	25	5,8	3	3,012





Спасибо за внимание!