

СТАТИСТИЧЕСКАЯ ОБРАБОТКА И ВИЗУАЛИЗАЦИЯ МЕДИЦИНСКИХ ДАННЫХ

А.В. Носова, О.А. Маланина
(Томск, Томский Политехнический Университет)
e-mail: avn37@tpu.ru, oam7@tpu.ru

STATISTICAL PROCESSING AND VISUALIZATION OF THE MEDICAL DATA

A.V. Nosova, O.A. Malanina
(Tomsk, Tomsk Polytechnic University)
e-mail: avn37@tpu.ru, oam7@tpu.ru

Abstract. The aim of the research is to consider basic concepts of descriptive statistics and to show basic histograms for the medical data

Key words: medical data, statistic, analysis, medicine, histogram.

Статистика - это наука, изучающая сбор, анализ и обработку цифровых данных. Математический анализ данных необходим для интерпретации медицинских исследований и является важнейшим этапом изучения клинических, диагностических, лечебных и профилактических мероприятий. В медицине и здравоохранении часто используются различные статистические концепции при принятии решений по таким вопросам как оценка состояния здоровья, его прогноз, выбор стратегии и тактики профилактики и лечения, оценка отдаленных результатов и выживаемости.

Понятия из описательной статистики[1]:

- Среднее значение случайной величины представляет собой наиболее типичное, наиболее вероятное ее значение, своеобразный центр, вокруг которого разбросаны все значения признака.
- Мода — это наиболее часто встречающийся вариант ряда. Модой для дискретного ряда является варианта, обладающая наибольшей частотой.
- Медианой является такое значение случайной величины, которое разделяет все случаи выборки на две равные по численности части.
- Квартили. Предоставляют важную информацию о структуре вариационного ряда к-л признака. Вместе с медианой они делят вариационный ряд на 4 равные части. Квартилей две, их обозначают символами Q, верхняя и нижняя квартиль. 25% значений меньше, чем нижняя квартиль, 75% значений меньше, чем верхняя квартиль.
- Стандартное отклонение. Стандартное отклонение (или среднее квадратическое отклонение) является мерой изменчивости (вариации) признака. Оно показывает на какую величину в среднем отклоняются случаи от среднего значения признака.
- Дисперсия является мерой изменчивости, вариации признака и представляет собой средний квадрат отклонений случаев от среднего значения признака. В отличие от других показателей вариации дисперсия может быть разложена на составные части, что позволяет тем самым оценить влияние различных факторов на вариацию признака.
- Функция распределения — функция, характеризующая распределение случайной величины.
- **Частота** — число повторений, сколько раз за какой-то период происходило некоторое событие, проявлялось определенное свойство объекта либо наблюдаемый параметр достигал данной величины.
- Плотность вероятности — называется производная ее функции распределения.
- Кривая плотности вероятности — показывает распределение вероятностей в зависимости от величины случайной погрешности. Таким образом они являются дифференциальными кривыми.

На нашем примере мы рассмотрим как производится статистическая обработка медицинских данных. В данном примере представлены данные по такому заболеванию как гипертония в Доминиканской республике.

Данные взяты с сайта

<http://biostat.mc.vanderbilt.edu/wiki/Main/DataSets?CGISESSID=10713f6d891653ddcbb7ddb79ffb79>

В нашей выборке представлены такие данные: деревня, пол, возраст, систолическое артериальное давление и диастолическое артериальное давление. Произведем анализ данных как в RStudio, так и в Excel. Вычислим среднее значение, моду, медиану, квартили, стандартное отклонение и дисперсию (Рис. 1. и Рис. 2.).

```
> DominicanHTN <- read.csv("DominicanHTN.csv", sep = ";")
> summary(DominicanHTN)
  village      Gender      Age      SBP      DBP
Carmona   :64 Female:258 Min.   : 15.00 Min.   : 80 Min.   : 20.00
Cojobal   :59 Male  :123 1st Qu.: 38.00 1st Qu.:118 1st Qu.: 76.00
Los Gueneos :57      Median : 47.00 Median :130 Median : 82.00
Juan Sanchez :53      Mean   : 47.97 Mean   :133 Mean   : 84.23
Batey verde :41      3rd Qu.: 59.00 3rd Qu.:150 3rd Qu.: 92.00
La Altagracia:40      Max.   :100.00 Max.   :236 Max.   :152.00
(other)   :67
> var(DominicanHTN$Age)
[1] 228.3492
> var(DominicanHTN$SBP)
[1] 661.3
> var(DominicanHTN$DBP)
[1] 206.0398
> sd(DominicanHTN$Age)
[1] 15.11123
> sd(DominicanHTN$SBP)
[1] 25.71575
> sd(DominicanHTN$DBP)
[1] 14.35409
> quantile(DominicanHTN$Age)
 0% 25% 50% 75% 100%
 15  38  47  59 100
> quantile(DominicanHTN$SBP)
 0% 25% 50% 75% 100%
 80 118 130 150 236
> quantile(DominicanHTN$DBP)
 0% 25% 50% 75% 100%
 20  76  82  92 152
```

Рис.3. Вычисления в RStudio

	Gender	Age	SBP	DBP		Age	SBP	DBP
1	Male	56	150	100	Среднее значение	47,971129	132,99475	84,228346
3	Male	42	120	90	Мода	40	120	80
4	Male	69	120	90	Медиана	47	130	82
5	Male	70	180	80	Первый квартиль	38	118	76
6	Male	62	138	78	Третий квартиль	59	150	92
7	Male	63	115	80	Стандартное отклонение	15,010808	25,582184	14,350102
8	Male	31	130	80	Дисперсия	225,32437	654,44813	205,92542

Рис.4. Вычисления в Excel

Как мы видим величины, полученные в двух разных программах, почти одинаковы. Значит все вычисления были проведены верно. Исходя из полученных данных можно сделать различные выводы о данном заболевании.

Статистика в медицине предполагает не только методы математической обработки цифрового материала, но и красивого и информативного графического или табличного его представления.

Графическим представлением является гистограмма. Гистограмма является важным инструментом статистики, позволяющим наглядно представить распределение значений анализируемой переменной. Обычная гистограмма отображает по оси ординат частоту встречаемости для каждого класса значений (в нашем случае для DBP и SBP).

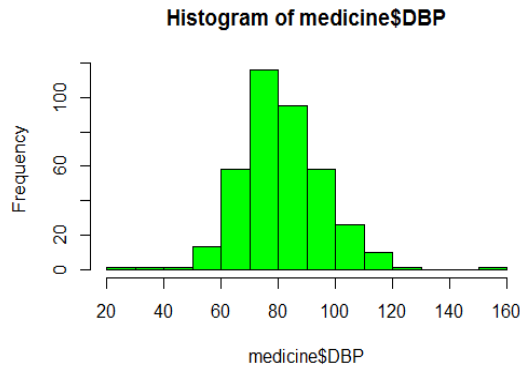


Рис.5. Гистограмма1 в RStudio

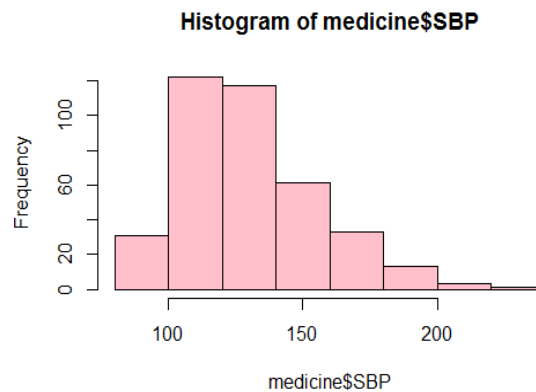


Рис.6. Гистограмма2 в RStudio

Поведение функции , задающей обычную гистограмму можно изменить, получив на оси ординат не частоту, а плотность вероятности. Так же для графического представления статистики помимо гистограммы можно использовать кривую плотности.

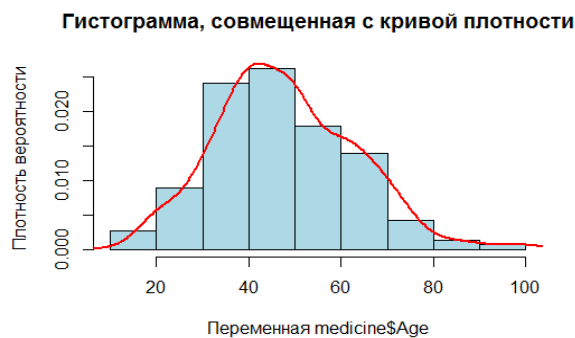


Рис.7. Гистограмма3 в RStudio

На рис.5. мы видим асимметрию влево , что означает что большинство людей, которые были описаны в статистике были представители возрастной категории меньше 50 лет.

Исходя из графиков мы можем сделать вывод о заболевании, а точнее о том, в каком возрасте больше распространяется данное заболевание или какое давление встречается чаще всего среди представителей данной статистики.

Таким образом, нами были рассмотрены основные понятия из описательной статистики, проанализированы данные по такому заболеванию как гипертония в Доминиканской республике и сделаны различные выводы по поводу данного заболевания.

ЛИТЕРАТУРА

1. Елисеева И.И., Юзбашев М.М. Общая теория статистики: Учебник / Под ред. И.И. Елисеевой. — 5-е изд., перераб. и доп. — М.: Финансы и статистика, 2004. — 656 с.

РЫНОК ИЗДЕЛИЙ ДЛЯ ЭНДОВАСКУЛЯРНОЙ ХИРУРГИИ. МЕСТО СОСУДОВ МАЛОГО ДИАМЕТРА

П. А. Пономарев, Е. Е. Пономарева
(г. Новосибирск, Новосибирский государственный университет,
Сибирский государственный университет путей сообщения)
Email: ponomarevpavel1995@gmail.com, iee-1995@mail.ru

MARKET OF PRODUCTS FOR ENDOVASCULAR SURGERY. PLACE OF SMALL DIAMETER VESSELS

P. A. Ponomarev, E. E. Ponomareva
(Novosibirsk, Novosibirsk state university, Siberian transport university)

Abstract. In this article authors analyses local market of products for endovascular surgery and small diameter vascular grafts. There are 4 main types of small diameter vascular grafts on the market today, but they are not suitable for all operations and show insufficient results in the long run. Due to demographic factors and grow of local healthcare market new generation of small diameter vascular grafts is required. The best way for solving this problem is electrospun vascular grafts made from different non-biodegradable polymers. Technology of this type are already in development stage in local scientific organizations.

Key words: healthcare market, endovascular surgery, vascular grafts, shunting, prosthetics.

Актуальность В настоящее время существует большое количество демографических явлений, значимых для национальной экономики таких, как ускоряющееся старение населения, низкая физическая активность отдельных социальных и возрастных групп, обострение некоторых новых групп хронических и возрастных заболеваний, приводящих к возникновению заболеваний сердечно-сосудистой системы (например, атеросклероз) при которых необходима реваскуляризация (восстановление утраченной функции кровоснабжения участка ткани) пораженных артерий [1]. Широкое распространение получили процедуры баллонной ангиопластики и процедуры с применением стентов различной конфигурации, причем процент проходимости спустя три года после проведения процедуры составляет лишь немногим более 19% [2]. Это является показанием для процедуры замены пораженных сосудов, которая является открытым хирургическим вмешательством, сопряженным с большим комплексом рисков как объективного, так и субъективного порядка. Как правило, в современной эндоваскулярной хирургической практике используются аутологичные (свои собственные) сосуды или синтетические протезы сосудов из различных полимерных материалов. Синтетические протезы сосудов являются стандартом при замене сосудов большого диаметром, однако, сосуды диаметром меньше 6 мм часто подвержены стенозу из-за недостаточной эластичности и биосовместимости. Следствием этого является снижение сечения просвета в районе анастомозов за счет тромбозом в постоперационном периоде [3]. Что касается применения аутологичных сосудов, то они не всегда бывают доступны по медицинским показаниям. Ксеногенные протезы различного происхождения, как правило приводят в отдаленной перспективе к возникновению аневризм. Таким образом, разработка и внедрение в клиническую практику протезов сосудов малого диаметра из синтетических полимеров, которые будут более механически, био- и гемосовместимы является актуальной задачей.

Потребность рынка/рынок сбыта. В Российской Федерации на протяжении последних лет наблюдается рост числа артериальных реконструкций. Однако, число подобных оперативных вмешательств должно быть многократно возрастать, т. к. освоение данной хирургической технологии делает это вмешательство все более доступным для многих пациентов. По