

ФОРМИРАНЕ НА ПОНЯТИЯ ОТ ОПИСАТЕЛНАТА СТАТИСТИКА С ПОМОЩТА НА MS EXCEL

Елена Л. Каращранова

Тази статия е опит да се илюстрира използването на MS EXCEL като средство за по-добро възприемане на основни понятия в статистиката – групиране на данни, пресмятане на числови характеристики на количествени признаци, графично представяне на данните и връзката на тези понятия с емпиричното разпределение.

Статистиката като теория и като практическа дейност винаги е била свързана с информационните потребности на обществото. С течение на времето все по-осъзната става необходимостта от получаването на определени статистически знания и развитието на специфичен тип мислене, наречено статистическо.

Използването на MS EXCEL като средство за по-добро възприемане на основни понятия в статистиката е продиктувано от:

- натрупания опит в преподаването на статистика с помощта на MS EXCEL на студенти от различни специалности на ЮЗУ „Неофит Рилски“;
- възможностите за улесняване на работата на учителите и възприемането от учениците на включените в ДООИ елементи от статистиката в средното общообразователно училище;
- възможността за осъществяване на реална междупредметна връзка.

След завършване на наблюдението на изучаваното явление се разполага с огромен статистически материал. Получените данни имат ограничени познавателни възможности, тъй като се отнасят за отделните единици на наблюдението. Така получените статистически данни не дават възможност да се разкрие същността на изучаваните явления. Налага се от първичен статистически материал, отнасящ се до отделните единици, да се премине към статистически данни, характеризиращи качествено еднородни групи.

Статистическата групировка е метод с помощта на който се обособяват групи на основата на значенията на един или повече статистически признаци и се разпределят единиците на съвкупността в тези групи. На практика чрез статистическата групировка се осъществява преходът от единичното към общото.

При преподаването на статистическата групировка е важно да се акцентира върху нееднозначността на тази процедура, т. е. да се дадат примери за различни възможни групировки на статистически данни. Ако преброяването на наблюденията

се извършва ръчно, учениците бързо губят интерес и допускат грешки, особено ако обемът на извадката е голям.

MS EXCEL предоставя набор от специални инструменти за анализ и статистическа обработка на данни, наречен Analysis ToolPack. Част от функциите на това разширение позволяват:

1. Групиране на данни, получени при реални статистически изследвания;
2. Построяване на честотна таблица;
3. Построяване на хистограма на абсолютната честота;
4. Построяване на графика на кумулативните честоти на разпределението;
5. Пресмятане на мода, медиана, средна аритметична, стандартно отклонение и дисперсия.

Да разгледаме следния пример:

Пример 1. При тестово препитване на един конкурс се явили 40 кандидати. Зададени са два теста – по математика и по английски език. Резултатите от конкурса са представени в следната таблица:

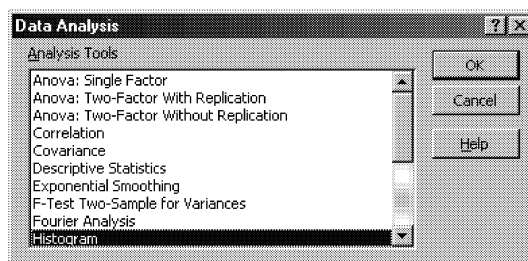
МАТЕМАТИКА	АНГЛИЙСКИ
22, 90, 86, 53, 31, 62, 75, 40	49, 38, 74, 78, 84, 88, 76, 72
55, 94, 37, 42, 76, 88, 41, 76	86, 69, 55, 42, 31, 91, 88, 72
48, 88, 76, 29, 49, 72, 64, 72	65, 66, 72, 78, 56, 99, 92, 84
50, 65, 66, 59, 85, 62, 58, 42	63, 86, 72, 67, 81, 59, 24, 77
7, 25, 66, 54, 38, 88, 76, 62	91, 86, 88, 72, 51, 84, 62, 89

а) Намерете честотното разпределение на резултатите по математика, като използвате различни ширини на интервалите – 15 и 20;

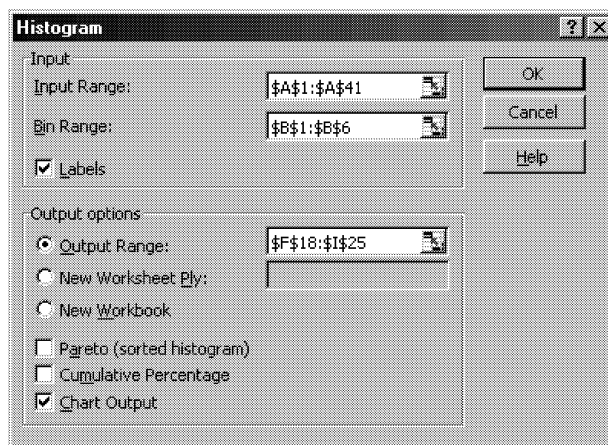
б) Постройте съответните хистограми на наблюденията

За да се използва инструментът за групиране на данни е необходимо да се въведат в колона стойностите на количествения признак, който ще бъде групиран (A). В отделни колони (B, C и D) да се въведат горните граници на интервалите, в които трябва да бъдат групирани данните. След това се избира Data Analysis от менюто Tools и от него инструмента Histogram [1]. (Фиг.1). След този избор се появява диалоговият прозорец Histogram (Фиг.2).

В този прозорец се попълват следните полета:



Фиг. 1. Диалогов прозорец Data Analysis



Фиг. 2. Диалогов прозорец Histogram

- **Input Range** – въвежда се адресът на зоната от клетки, която съдържа данните;
- **Bin Range** – въвежда се адресът на зоната от клетки, която съдържа горните граници на интервалите за групиране на данните, в случай, че това поле се пропусне, MS Excel прави автоматично групиране на данните;
- **Labels** – маркира се, ако съответните зони съдържат етикети;
- В блока **Output options** се избира мястото на извеждане на хистограмата и нейният тип.

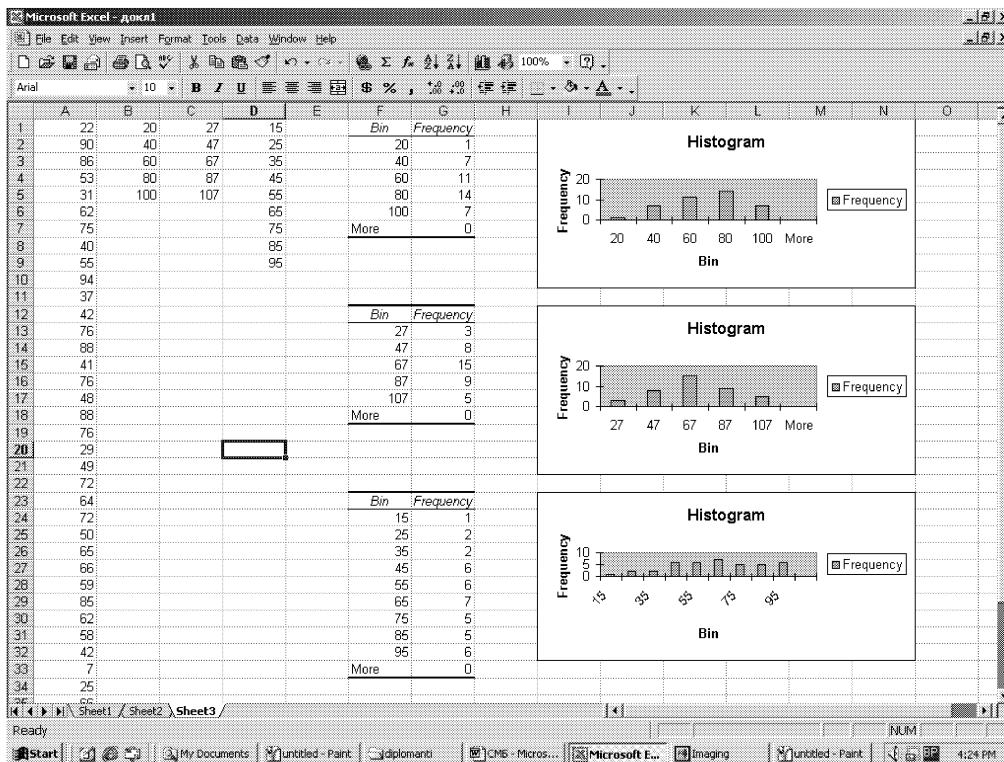
На Фиг. 3 са представени три варианта на групиране на данните (честотни таблици и хистограми) на резултатите от теста по математика, като първите два са за ширина на интервала 20, но с различна начална стойност, а третият е за ширина на интервала 15.

Решението на задачата с MS EXCEL дава възможност ясно да се види, че групирането на данните на практика е от същесвено значение за вида на честотното разпределение на статистическите признаци. Лесно може да се направи връзката между групирането на данните и емпиричното разпределение, което на практика моделира изследваното явление. Доброто познаване на същността на статистическите признаци и техниката за статистическа групировка са от съществено значение за построяването на подходящ модел.

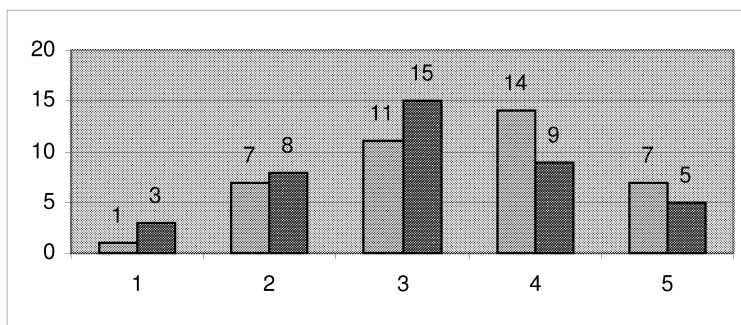
Данните от честотните таблици могат да се използват за получаване на различни графични изображения.

Ако към предложения пример се поставят допълнителни условия – хистограмите да се представят на една графика, това може лесно да се реализира с помощта на **Chart Wizard** (Фиг 4.).

Използването на стандартните графични възможности на MS Excel не би трябвало да затрудни учениците, тъй като графичното представяне на данни се изучава в курса по информационни технологии в 9 клас. Този факт подпомага учениците



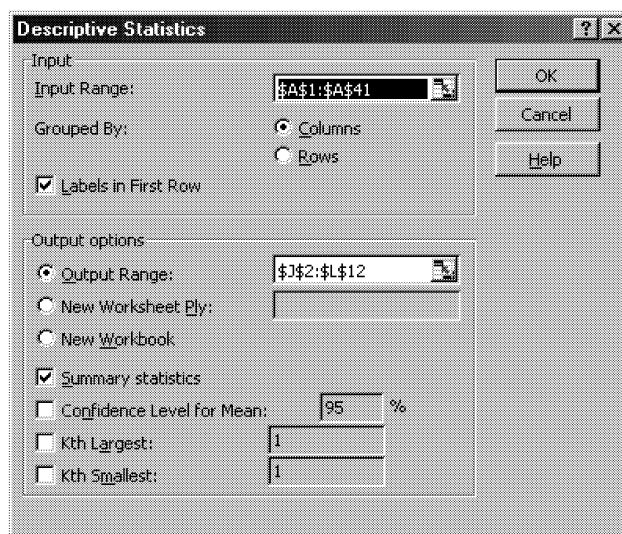
Фиг. 3. Решение на Пример 1



Фиг. 4.

да интерпретират и обработват лесно статистическата информация, нещо повече в часовете по математика се създава възможност да се обогати представата им за приложенията на MS EXCEL при създаване на графични изображения.

Пресмятането на числови характеристики на количествен признак може да се направи чрез вградените в MS EXCEL статистически функции
AVERAGE(number1,number2,...) – за средно аритметична стойност,
STDEV(number1,number2,...) – за стандартно отклонение,
SKEW(number1,number2,...) – коефициент на асиметрия,
VAR(number1,number2,...) – за дисперсия,
MEDIAN(number1,number2,...) – за медиана,
MODE(number1,number2,...) – за мода и
KURT(number1,number2,...) – коефициент за ексцес
или от менюто DESCRIPTIVE STATISTICS на DATA ANALYSIS. (Фиг.1). На фиг. 5 е представен диалоговият прозорец DESCRIPTIVE STATISTICS. В полето Input Range се въвежда адресът на зоната от данни, за които се изчисляват дескриптивните статистики. Избира се начинът на групиране на данните: по колони или по редове в полето Grouped By. В блока Output Options се избира мястото на извеждане на резултата. Важно е да се маркира полето Summary Statistics, чрез което се избира извеждането на числовите характеристики. Изходните резултати са представени в Табл. 1.



Фиг. 5. Диалогов прозорец DESCRIPTIVE STATISTICS

Mean	59.225
Standard Error	3.352092
Median	62
Mode	76
Standard Deviation	21.20049
Sample Variance	449.4609
Kurtosis	-0.44479
Skewness	-0.35378
Range	87
Minimum	7
Maximum	94
Sum	2369
Count	40

Таблица 1. Числови характеристики за резултатите по математика

Използването на DATA ANALYSIS за пресмятането на описателни статистики на количествен признак е по-подходящо в сравнение с използването на вградените статистически функции, тъй като едновременно се получава таблица със съответните описателни статистики – средна стойност, стандартна грешка, медиана, мода, стандартно отклонение, коефициент на ексцес, коефициент на асиметрия, размах, минимална и максимална стойности, сумата от стойностите и броя на наблюденията.

Предложеният подход е възможно да бъде реализиран без допълнително да натоварва учениците, защото в учебната програма по Информационни технологии за IX клас (задължителна подготовка) е включен модулът Електронни таблици.

Използването на MS EXCEL в обучението по статистика позволява:

1. Решаване на реални задачи, т. е. резултатите могат да са от реално статистическо изследване;
2. Да се илюстрира приложението на електронните таблици за обработка на статистически данни;
3. Да подпомогне учителя от една страна и да мотивира учениците към преподаваната тема;
4. Да се акцентира върху моделирането на честотното разпределение;
5. Да се групират данни и анализират честотни разпределения;
6. Да се избира средна стойност (мода, медиана, средна аритмитична), характеризираща най-добре статистическата съвкупност.

В продължение на няколко години в ЮЗУ „Неофит Рилски“ в курсовете по статистика със студенти от различни специалности се използва MS EXCEL. Този подход обогатява тяхната статистическа грамотност и им позволява за в бъдеще да го прилагат в различни области. През учебната 2001/2002 студенти от курсовете за следдипломна квалификация на учители по информатика приложиха експериментално дискутирания по-горе подход в 11 клас на ЕГ „Акад. Людмил Стоянов“ гр. Благоевград. Получиха се много добри резултати при усвояването на статистическите понятия, като в това число се повиши и интересът на учениците към възможностите за статистически анализ на данни чрез средствата на MS EXCEL.

Опитът показва, че обучението по статистика включително и в средното училище ще спечели от прилагането на информационните технологии в учебния процес. Необходимо е по-широко популяризиране на възможностите на MS EXCEL както сред учениците, така и сред учителите по математика.

ЛИТЕРАТУРА

[1] Ръководство за потребителя – Microsoft Excel., Microsoft Corporation.

Елена Лазарова Караштранова
ЮЗУ „Неофит Рилски“
бул. Иван Михайлов 66
2700 Благоевград
e-mail: helen@aix.swu.bg

FORMING CONCEPTS OF THE DESCRIPTIVE STATISTICS BY THE USE OF MS EXCEL

Elena L. Karashtranova

In this paper we try to illustrate the use of information technologies as a tool for better perception of basic concepts in statistics – data grouping, graphical representation of data and relation between these concepts and empiric distribution.