Лабораторная работа № 2

Распределение показателей качества по качественному признаку

Качественный признак показывает, является единица продукции годной или дефектной. Качественный признак может отражать также число дефектов в единице продукции, например, на определённой площади стального листа.

При выборочном контроле по качественному признаку в выборку из партии попадает некоторое случайное число дефектных единиц продукции. Вероятности попадания в выборку того или иного количества дефектных единиц продукции составляют дифференциальную функцию распределения.

Пусть партия состоит из N изделий, D из которых бракованные. Если взять из партии случайную бесповторную выборку (какую обычно и берут в производстве) объёмом n, то вероятность того, что в выборке ровно m бракованных изделий, равна

$$P(m) = \frac{C_D^m * C_{N-D}^{n-m}}{C_N^n}$$

, где, например, $C_N^n = \frac{N!}{n!(N-n)!}$

Совокупность этих вероятностей для m=0,1,2,3,…,n при заданных N, D, n описывается дифференциальной функцией гипергеометрического распределения.

Величина *P(m)* может быть рассчитана в программе Excel при помощи статистической функции ГИПЕРГЕОМЕТ. Диалоговое окно, открывающееся при выборе этой функции, имеет четыре строки для ввода данных:

Число успехов в выборке. Необходимо ввести количество успешных испытаний в выборке. При этом под количеством успешных испытаний понимается количество элементов выборки, обладающих определённым признаком, в нашем случае – количество дефектных изделий в выборке.

Размер выборки. Вводится размер выборки.

Число успехов в совокупности Подсказка к этой строке указывает, что надо ввести количество успешных испытаний в генеральной совокупности. В нашем случае это количество дефектных изделий в партии.

Размер_совокупности. Вводится объём партии.

При очень больших значениях параметров расчёт гипергеометрического распределения может оказаться затруднительным даже при использовании компьютера. Однако, если $n \le 0{,}1N$, то гипергеометрическое распределение можно приближённо заменить биномиальным (которое имеет место при повторной случайной выборке), расчёты которого более просты. При биномиальном распределении

$$P(m) = C_n^m (1-q)^m q^{n-m},$$

где *q=D/N* – доля дефектных изделий в партии.

При биномиальном распределении величина P(m) может быть рассчитана в программе Excel при помощи статистической функции БИНОМРАСП. Диалоговое окно, открывающееся при выборе функции, имеет четыре строки для ввода данных:

Число_успехов. Необходимо ввести количество успешных испытаний. При этом под количеством успешных испытаний понимается количество элементов выборки, обладающих определённым признаком, в нашем случае – количество дефектных изделий в выборке.

Число испытаний. Предлагается ввести число независимых испытаний, т.е. объём выборки.

Вероятность успеха. Предлагается ввести вероятность успеха каждого испытания. В нашем случае это вероятность того, что случайно выбранное изделие будет бракованным, т.е. доля дефектных изделий в партии, иными словами – уровень дефектности.

Интегральный. Вводится **истина**, если рассчитывается значение интегральной функции распределения, и **ложь**, если рассчитывается значение дифференциальной функции распределения, т.е. в нашем случае – значение P(m).

Если $q \leq 0,1$ и $n \leq 0,1N$, что обычно и имеет место в практике статистического контроля, то биномиальное распределение, как и гипергеометрическое, можно приближённо заменить ещё более простым для расчётов распределением Пуассона, в котором

$P(m) = \frac{\lambda^m e^{\lambda}}{m!}$, где $\lambda = nq$ – математическое ожидание числа дефектных изделий в выборке.

При распределении Пуассона величина P(m) может быть рассчитана в программе Excel при помощи статистической функции ПУАССОН. Диалоговое окно, открывающееся при выборе функции, имеет три строки для ввода данных:

X. Количество событий, в нашем случае - количество дефектных изделий в выборке.

Среднее. Среднее ожидаемое численное значение, в нашем случае – параметр λ, т.е. математическое ожидание числа дефектных изделий в выборке.

Интегральный. Вводится **истина**, если рассчитывается значение интегральной функции распределения, и **ложь**, если рассчитывается значение дифференциальной функции распределения, т.е. в нашем случае – значение P(m).

Пример 2.1. Из партии, состоящей из 1000 изделий, 30 из которых дефектные, взята выборка объёмом 50 изделий. Построить график дифференциальной функции распределения вероятностей, используя гипергеометрическое распределение.

Открываем новую книгу Excel. В ячейку A1 вводим заголовок работы «Лаб. работа 2. Распределение показателей качества по качественному признаку». Далее вводим исходные данные (Рис. 2.1).

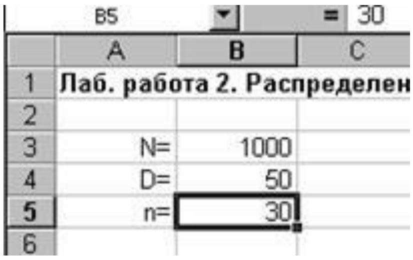

Рис.2.1. Исходные данные для расчёта распределения в примере 2.1.

Поскольку график представляет собой зависимость P(m), то для его построения понадобятся диапазоны данных **m** и **P(m)гипер.** Соответствующие заголовки вводим в ячейки A7 и B7. В диапазон A8:A38 вводим количество дефектных изделий в выборке от 0 до 40 с шагом 1.

В ячейке B8 рассчитываем вероятность для m=0 при помощи статистической функции ГИПЕРГЕОМЕТ. В первую строку диалогового окна вводим ссылку на ячейку A8. Во вторую строку вводим ссылку на ячейку B5. В третьей строке делаем ссылку на ячейку B4. В четвёртой строке делаем ссылку на ячейку B3.

В результате в ячейке B8 получаем значение 0,209681. Формулу из ячейки B8 копируем в диапазон B9:B48. Перед копированием вводим в формуле абсолютную адресацию тех ячеек, ссылки на которые не должны меняться при копировании – ячеек B3, B4, B5.

При построении графика выбираем диаграмму вида **Точечная.** Она позволяет сравнить пары значений, т.е. график будет представлять отдельные точки, не соединённые линией. Это связано с тем, что количество дефектных изделий в выборке – дискретная случайная величина, принимающая только целые значения.

На втором шаге создания диаграммы в качестве диапазона данных вводим диапазон A8:B15. Остальные значения P(m) можно на графике не использовать, поскольку они практически равны нулю, начиная с *P(7),* находящегося в ячейке B15.

После редактирования диаграммы получаем график, показанный вместе с расчётными данными на рис. 2.2.

Рис.2.2. Результаты расчётов и график дифференциальной функции

гипергеометрического распределения в примере 2.1.

Задание

1. Выполнить расчёты и построения в соответствии с примером.
2. На том же листе рабочей книги продолжить расчёты и построить графики дифференциальных функций биномиального распределения и распределения Пуассона с теми же параметрами, что и в примере. Сравнить значения вероятностей, рассчитанных по различным распределениям.
3. Измените исходные данные следующим образом:

№	**Вариант 1**		**Вариант 2**		**Вариант 3**		**Вариант 4**		**Вариант 5**	
N	20000		30000		25000		2000		15000	
D	5000		7000		7000		750		6000	
n	200		300		200		100		150	
№	**Вариант 6**		**Вариант 7**		**Вариант 8**		**Вариант 9**		**Вариант10**	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
N	15000		3000		2500		1800		13000	
D	5500		600		400		300		5400	
n	200		300		200		100		150	

4. Выполнить расчёты в соответствии с примером и построить графики плотностей вероятностей для всех распределений, выбрав оптимальные параметры диаграммы, позволяющие наглядно отразить графики.