Санкт-Петербургский Национальный Исследовательский Университет Информационных технологий, механики и оптики

Лабораторная работа #4

Изучение законов распределения случайных величин

Выполнил: Канева
Тамара Игоревна
Группа № К3121
Проверила: Казанова
Полина Петровна

Санкт-Петербург
2021

Цель работы:

Изучить законы распределения случайных величин с помощью программы Microsoft Excel.

Задачи:

Изучить биномиальное распределение, гипергеометрическое распределение, нормальное распределение, экспоненциальное распределение, распределение Пуассона, распределение Пирсона, распределение Фишера и распределение Вейбулла на основе предложенных примеров и заданий для самостоятельного решения.

Ход работы:

Изучение предложенных в теоретической части примеров.

Биномиальное распределение.

Определим вероятность того, что трое из пяти человек, зашедших в магазин, купят товар марки «А», если известно, что 85% покупателей предпочитают именно товар этой марки.

Для решения этой задачи найдем вероятности $P(\xi = k)$ и $P(\xi \leq k)$, для чего воспользуемся формулой “БИНОМ.РАСП(k;n;p;i)”, где k - количество успешных испытаний, n - количество независимых испытаний, p - вероятность успеха каждого испытания, а i - логическое значение, определяющее форму функции (если 1, то возвращается вероятность того, что число успешных событий не больше значения k, а если 0, то рассчитывается вероятность того, что число успешных событий равно значению k).

В нашем случае количество успешных испытаний равно 3, количество независимых испытаний - 5, вероятность успеха каждого испытания равна 0.85, а ищем мы вероятность того, что количество успешных испытаний равно строго 3, поэтому в поле “Интегральная” вводится 0, что отражено на рисунке 1.

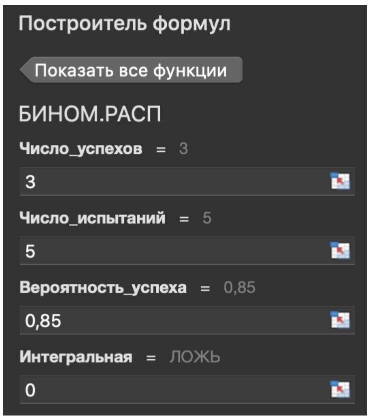


Рис. 1. Диалоговое окно “Построитель формул”, куда вводятся значения аргументов функции “БИНОМ.РАСП”.

Введя необходимые параметры, получим значение функции, равное 0.13817813 (рис. 2). Это число является ответом этой задачи.

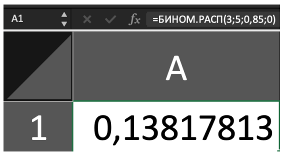


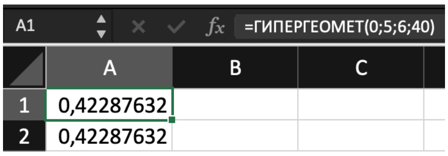
Рис. 2. Значение функции “БИНОМ.РАСП(3;5;0,85;0)”.

Гипергеометрическое распределение.

Определим вероятность того, что партия, содержащая 40 изделий, будет принята, если из 40 изделий 6 имеют дефекты. Партия принимается, если из 5 случайным образом выбранных изделий только одно оказывается бракованным.

Для решения этой задачи найдем вероятность получить *k* из выборки объемом *n*, для чего воспользуемся формулой “ГИПЕРГЕОМЕТ(k;n;K;N)”, где k - количество успешных событий в выборке, n - размер выборки, K - число успехов в совокупности, а N - размер совокупности.

В нашем случае количество успешных событий равно 0 или 1, размер выборки - 5, число успехов в совокупности равно 6, а а размер совокупности - 40. То есть нам необходимо рассмотреть два случая, получить две вероятности (для k=0 и для k=1), после чего полученные значения сложить (рис. 3 - 5).



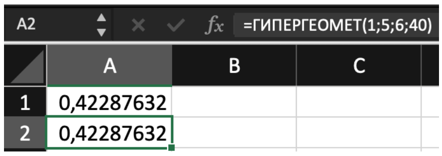


Рис. 3 - 4. Значения аргументов и самих функций “ГИПЕРГЕОМЕТ(0;5;6;40)” и “ГИПЕРГЕОМЕТ(1;5;6;40)”

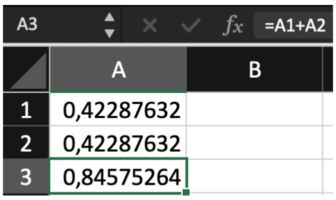


Рис. 5. Сумма вероятностей получить 0 и 1 из выборки объемом 40.

Таким образом, получаем искомую вероятность, равную 0.84575264.

Распределение Пуассона.

Определим вероятности того, что на гарантийный ремонт не поступит ни одного изделия и поступит не более одного, если каждый день ожидается возврат на гарантийный ремонт в среднем только 1,5 единиц товара.

Для решения этой задачи найдем вероятности $P(\xi = k)$ и $P(\xi \leq k)$, для чего воспользуемся формулой “ПУАССОН(*k*,λ,i)”, где k - количество событий, λ - интенсивность появления событий, а i - логическое значение, определяющее форму функции (если 1, то возвращается вероятность того, что число успешных событий не больше количества событий, а если 0, то возвращается вероятность того, что число успешных событий равно количеству событий).

В случае, когда мы ищем вероятность того, что на гарантийный ремонт не поступит ни одного изделия, количество событий равно 0, интенсивность появления событий - 1.5, а поскольку мы ищем вероятность того, что число успешных событий равно количеству событий, то значение i равно 0 (рис. 6).

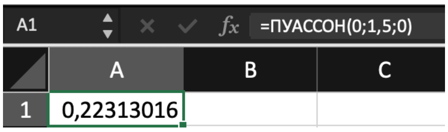


Рис. 6. Вероятность того, что на гарантийный ремонт не поступит ни одного изделия, равная значению функции “ПУАССОН(0;1,5;0).

Таким образом, вероятность того, что на гарантийный ремонт не поступит ни одного изделия равна 0.22313016.

В случае, когда мы ищем вероятность того, что на гарантийный ремонт поступит не более одного изделия, количество событий равно 1, интенсивность появления событий - 1.5, а поскольку мы ищем вероятность того, что число успешных событий не больше количества событий, то значение i равно 1 (рис. 7).

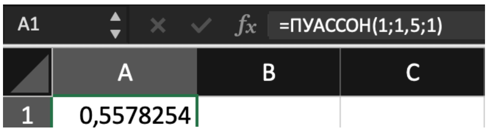


Рис. 7. Вероятность того, что на гарантийный ремонт поступит не более одного изделия, равная значению функции “ПУАССОН(1;1,5;1).

Таким образом, вероятность того, что на гарантийный ремонт не поступит не более одного изделия равна 0.5578254.

Нормальное распределение.

Определим процент общего числа производимых пальто, который должны составлять пальто 3-го роста (158 - 164 см), если среднее значение роста группы людей равно 172 см и стандартное отклонение равно 5 см.

Для решения этой задачи воспользуемся функцией “НОРМРАСП(k;a;d;i), где k - значение, для которого рассчитывается вероятность, a - среднее значение, d - стандартное отклонение, i - логическое значение, определяющее форму функции (если 1, то возвращается вероятность того, что результат не больше значения, для которого рассчитывается вероятность, а если 0, то возвращается значение k).

В нашем случае нужно найти разность значений функции “НОРМРАСП” для k=164 (рис. 8) и k=158 (рис. 9) в каждом отдельном случае и a = 172, d = 5, i = 1 в обоих случаях. Полученное значение разности (рис. 10), умноженное на 100% и будет ответом этой задачи.

Рис. 8 - 9. Значения функций “НОРМРАСП(164;172;5;1)” и “НОРМРАСП(158;172;5;1).

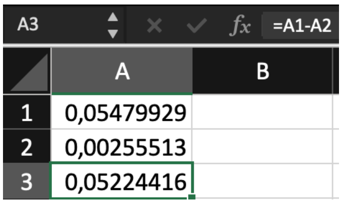


Рис. 10. Разность значений функций “НОРМРАСП(164;172;5;1)” и “НОРМРАСП(158;172;5;1).

Таким образом, искомый процент будет равен (0.05479929 - 0.00255513)·100%=0.05224416·100%=5.224416%.

Выполнение заданий для самостоятельного решения.

Задание 2.1.

Найдем вероятность того, что завтра позвонят не менее 3-х клиентов, если всего фирма располагает 8 клиентами, а вероятность с которой может позвонить каждый из них равна 0.25.

Для решения этой задачи воспользуемся законами биномиального распределения. С помощью функции “БИНОМРАСП” найдем искомую вероятность. В нашем случае количество успешных испытаний равно 3, количество независимых испытаний - 8, вероятность успеха каждого испытания равна 0.25, а ищем мы вероятность того, что количество успешных испытаний равно не менее 3, поэтому значение последнего аргумента равно 1.

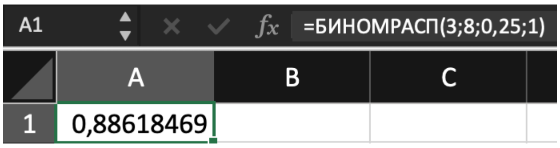


Рис. 11. Значение функции “БИНОМРАСП(3;8;0,25;1)”.

Таким образом, искомая вероятность равна 0.88618469 (рис. 11).

Задание 2.2.

Определим вероятность того, что завтра на ремонт поступит 4 детали, если в среднем возвращается 1.8 единиц.

Для решения этой задачи воспользуемся законами распределения Пуассона. С помощью функции “ПУАССОН” найдем искомую вероятность. В нашем случае количество событий равно 4, интенсивность появления событий - 1.8, а поскольку мы ищем вероятность того, что число успешных событий равно количеству событий, то значение i равно 0 (рис. 12).

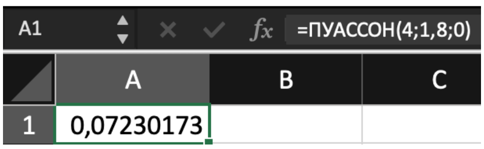


Рис. 12. Значение функции “ПУАССОН(4;1,8;0)”.

Таким образом, значение искомой вероятности равно 0.07230173.

Задание 2.3.

Выясним имеется ли существенная разница между цветами пальто с точки зрения покупателей, если считать, что результаты получены при *n* = 100, критерий χ^2 -квадрат при уровне значимости *α* = 0.05, 30% опрошенных предпочитают пальто серого цвета, 26% - чёрного, 22% - синего и 22% - коричневого.

Для решения этой задачи проверим гипотезу о принадлежности имеющихся данных к равномерному распределению, число интервалов K=4. Тогда при n=100 $n_1 = 0.3 \cdot n =$ $= 30$, $n_2 = 26$, $n_3 = 22$, $n_4 = 22$. Вероятности p_1=p_2=p_3=p_4=0.25, а теоретические частоты np_1=np_2=np_3=np_4=25. Критическое значение $\chi^2_{1-a} = \chi^2_{0.95} = 3.8$ с числом степеней свободы r =(4-2-1)=1.

$$\chi^2_n = \frac{(n_1 - np_1) + (n_2 - np_2) + (n_3 - np_3) + (n_4 - np_4)}{np_1} = \frac{(30-25)^2 + (26-25)^2 + (22-25)^2 + (22-25)^2}{25} = 1.76 < \chi^2_{0.95}$$

Значение критерия χ^2_n меньше критического значения $\chi^2_{0.95}$. Таким образом, гипотеза верна. Существенной разницы между пальто разного цвета нет.

Задание 2.4.

Решим задачу из задания 2.3 для n=1000. Решение будет полностью аналогично, но $n_1 = 300$, $n_2 = 260$, $n_3 = 220$, $n_4 = 220$, а теоретические частоты np_1=np_2=np_3=np_4=250. Тогда $\chi^2_n = \frac{(300-250)^2 + (260-250)^2 + (220-250)^2 + (220-250)^2}{250} = 17.6 > \chi^2_{0.95}$. Значение критерия χ^2_n больше критического значения $\chi^2_{0.95}$. Таким образом, гипотеза неверна. Разница между пальто разного цвета существенна.

Вывод:

В ходе этой лабораторной работы были получены сведения о нескольких законах распределения случайных величин и навыки практической работы с этим законами. Были решены конкретные задачи, связанные с различными видами распределения случайных величин, таких как биномиальное, гипергеометрическое или распределение Пирсона.

Ответы на контрольные вопросы:

1. Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.
2. Экспоненциальное распределение применяют в теории надежности при описании режима нормальной эксплуатации техники, например, срока службы радиоэлектронной аппаратуры. Также это распределение используется в теории массового обслуживания, например, с его помощью можно описать время ожидания при техническом обслуживании или продолжительность телефонных разговоров.

3. Чтобы ввести формулу в Microsoft Excel, нужно перейти во вкладку “Формулы” и нажать в ней на кнопку “Вставить функцию”. Откроется окно “Построитель формул”, в котором нужно сначала выбрать нужную функцию, а после ввести значения параметров.

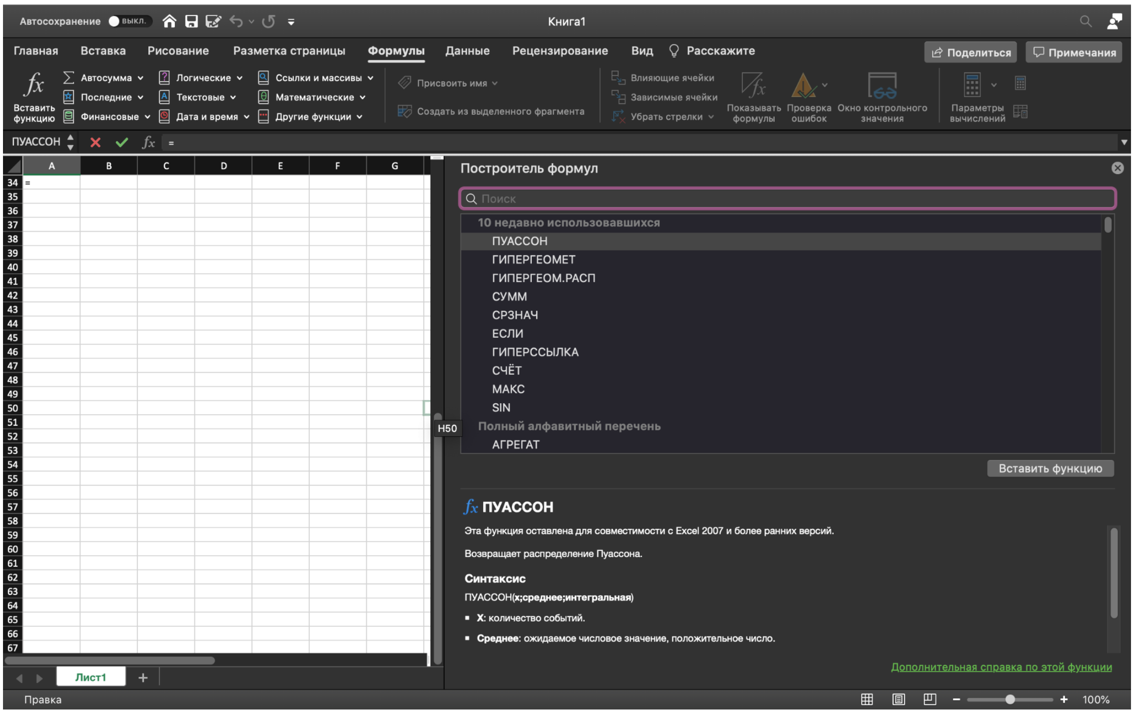


Рис. 13. Ввод формул в Microsoft Excel.

4. Существуют следующие параметры нормального распределения: k - значение, для которого рассчитывается вероятность, a - среднее значение, d - стандартное отклонение, i - логическое значение, определяющее форму функции (если 1, то возвращается вероятность того, что результат не больше значения, для которого рассчитывается вероятность, а если 0, то возвращается значение k).

5. Нормальное распределение часто применяется в тех случаях, когда возможные значения лежат вокруг некого среднего значения, например, при стрельбе по мишени, измерении какого-либо параметра и т.д.

6. Дисперсия - квадрат отклонения случайной величины от её математического ожидания.

7. Квантиль - значение, которое заданная случайная величина не превышает с фиксированной вероятностью.

8. Количество степеней свободы — это количество значений в итоговом вычислении статистики, способных изменяться. Например, при определении критерия согласия Пирсона число степеней свободы определяется как разность числа исследуемых независимых случайных величин и числа независимых условий, наложенных на частоты p_i.

9. При применении критерия согласия строится таблица результатов. Подсчитываются вероятности попадания в j-й интервал. Проверяется выполнение условия все np_i не меньше 5, объединяются те интервалы, в которых np_j меньше 5. Перестраивается таблица и подсчитывается статистика

Пирсона. На последнем этапе сравнивается полученная величина с квантилем распределения Пирсона на заданном уровне значимости.