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

**Лабораторная работа №4**

**Изучение законов распределения случайных величин**

Выполнил: Долматов

Дмитрий Алексеевич

Проверила: Казанова

Полина Петровна

Санкт-Петербург

2021

**Цель работы:**

Изучить законы распределения случайных величин с помощью программы Microsoft Excel.

**Задачи:**

1. Изучить и решить наиболее распространенные практические задачи по распределению случайных величин.

**Ход работы:**

**Упражнение 1:**

Биномиальное распределение определяет число успехов в n испытаний с вероятностью успешного исхода *p*. Значения параметра *p*:

Примером такого распределения является вероятность того, что трое из пяти человек, зашедших в магазин, купят товар марки *А*, при этом известно, что 85% покупателей предпочитают данный товар. Воспользуемся формулой *БИНОМРАСП*(*k*;*n*;*p*;0 или 1): 0, если нужно найти вероятность. А 1, если нужно найти вероятность того, что результат окажется меньше или равным *k*. Другими словами, интегрально-логическое значение 0 или 1 определяет, будет ли рассчитываться кумулятивная функция распределения, то есть вероятность того, что число успешных событий не больше значения аргумента *Число\_успехов*. В случае равенства 0 рассчитывается дифференциальная функция распределения, то есть вероятность того, что число успешных событий равно значению аргумента *Число\_успехов*.

Реализация представлена на рисунке 4.1.1.

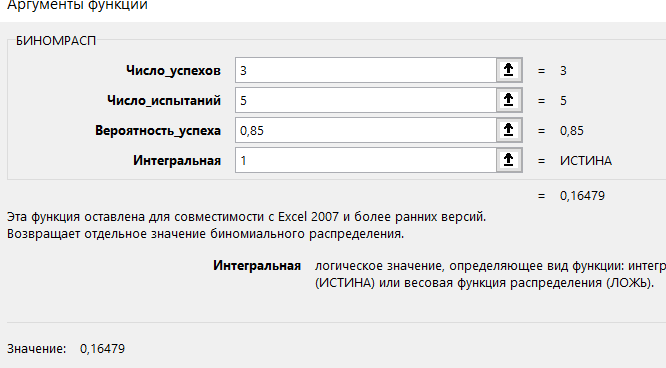




Рисунок 4.1.1 – Реализация интегрального (кумулятивного) биномиального распределения

Таким образом, вероятность того, что среди 5 покупателей вероятность того, что из них купят товар марки *А* не больше трех равно 0.16479. А вот в случае дифференциального биномиального распределения этот результат окажется равным 0.138178 (**ТРОЕ** из пяти человек купят этот товар).

**Упражнение 2:**

В данном упражнении изучим гипергеометрическое распределение. Применяется в задачах контроля за качеством продукции. Типичный пример: вероятность того, из *n* изделий, выбранных случайных образом из партии объемом *N*, ровно *k* изделий с дефектом. Формула нахождения этой случайный величины приведена на рисунке 4.2.1.

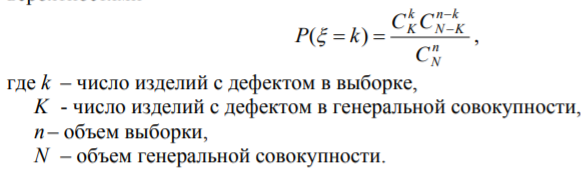


Рисунок 4.2.1 – Формула нахождения случайный величины, имеющей гипергеометрическое распределение

При этом, Задача заключается в том, что из партии, содержащей 40 изделий, случайным образом выбрали из которой 5 изделий, подвергнутых проверке на качество. Если находится брак, то снимается с продажи вся партия. Какова вероятность, что партия будет принята, если из 40 изделий 6 имеют дефекты. Используем функцию *ГИПЕРГЕОМЕТ*(*k, n, K, N*), где *k* – количество успешных событий в выборке, *n* – размер выборки, *K* – число успехов в совокупности, *N* – размер совокупности. Итого имеем ответ на рисунке 4.2.2.

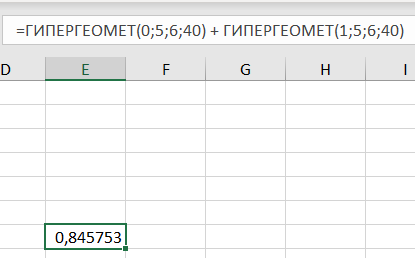


Рисунок 4.2.2 – Реализация гипергеометрического распределения

Мы используем сумму двух функций, поскольку нам нужно найти вероятность несовместимых событий (что будет найдена фальшивка, а там, где ее нет, ее собственно нет).

**Упражнение 3:**

В данном упражнении нужно изучить распределение Пуассона. Данное распределение описывает число событий, происходящих в одинаковых промежутках времени независимо друг от друга с постоянной интенсивностью. Случайная величина имеет распределение Пуассона с параметром , если она принимает значения 0, 1, 2 с вероятностями, написанными на рисунке 4.3.1.



Рисунок 4.3.1 – Уравнение случайной величины, имеющей распределение Пуассона

Решим задачу: каждый день возврат фирме составляет, в среднем, 1.5 единиц товара на ремонт. Нужно рассчитать, с какой вероятностью завтра на гарантийный ремонт не поступит ни одного изделия, или, например, поступит не более одного. Используем функцию *ПУАССОН*(со следующими параметрами:

* X – количество событий;
* Среднее – интенсивность появления событий;
* Интегральная-логическое значение (как и прошлом примере 1 – кумулятивно-интегральная, 0 – дифференциальная).

Таким образом, вероятность того, что завтра на гарантийный ремонт не поступит ни одного изделия и не более одного одного приведена на рисунке 4.3.2.

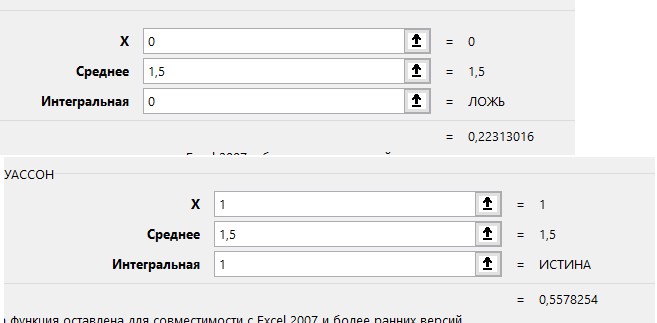


Рисунок 4.3.2 – Расчет распределения вероятности Пуассона

**Упражнение 4:**

В данном упражнении изучим нормальное распределение, которое чаще всего применяется, когда значения лежат вокруг некоего среднего значения, например, при стрельбе по мишени, измерении какого-либо параметра.

Необходимо решить задачу определения нормального распределения роста. Среди группы людей значение среднего роста получило значение 172 см и стандартного отклонения – 5 см. Требуется определить, какой процент общего числа производимых пальто должны составлять пальто 3-го роста (158-164 см). Для решения задачи используем функцию *НОРМРАСП*() со следующими параметрами:

* X – значение, для которого рассчитывается вероятность;
* Среднее – среднее значение;
* Стандартное отклонение – стандартное отклонение;
* Интегральное-логическое значение (как и в предыдущем упражнении).

Реализация нахождение ответа задачи с помощью нормального распределения представлена на рисунке 4.4.1.

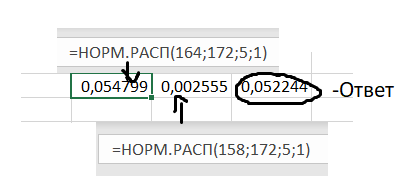


Рисунок 4.4.1 – Реализация нормального распределения

Таким образом, найдя значения того, что данная куртка больше для человека, имеющего больший рост, чем 158, но меньший, чем 164, равно разности этих отношений, представленных на рисунке 4.4.1 (в процентах). То есть пальто 3-го роста должно составлять 5.2% от общего числа.

**Упражнение 5:**

В данной упражнении изучим распределение Пирсона, в котором у нас имеются независимые случайные величины, каждая из которых имеет нормальное распределение. Закон распределения суммы квадратов независимых случайных величин называется Вероятность с помощью данного распределения можно определить с помощью функции =*ХИ2РАСП*(), в котором:

* Х – значение, для которого рассчитывается – распределение;
* Степени свободы – число степеней свободы (линейно независимых конфигураций элемента относительно системы).

Для определения квантиля – распределения следует использовать функцию *=ХИ2ОБР*(*уровень значимости*; *число степеней свободы*), которое возвращает обратное значение вероятности. Квантиль – это значение, которое заданная случайная величина не превышает с фикс.вероятностью.

\**Упражнение 6:**

Следующим распределением будет *F-распределение* (Фишера), использующееся при проверке адекватности уравнений регрессии. Используем функцию =FРАСП() со следующими параметрами:

* Х – значение, для которого рассчитывается функция;
* Число степеней свободы первого набор данных;
* Число степеней свободы второго набора данных.

Для определения квантиля *F-распределения* используем функцию *=FРАСПОБР*(*уровень значимости; число степеней свободы1; число степеней свободы2*), возвращает обратное значение вероятности.

**Упражнение 7:**

В данном упражнении разберем распределение *Вейбулла*, которое применяют в теории надежности при описании этапа старения аппаратуры. Для вычисления вероятности применяется функция *=ВЕЙБУЛЛ* с параметрами:

* X – значение, для которого выполняется функция;
* Альфа – параметр распределения;
* Бета – параметр распределения;
* Кумулятивный или дифференцированный параметр (см.прошлые упражнения).

**Упражнение 8:**

Завершающим распределением будет экспоненциальное. Его применяют в теории надежности при описании режима нормальной эксплуатации техники. Для вычисления пользуются функцией ЭКСПРАСП с параметрами:

* X – значение, для которого вычисляется функция;
* Лямбда – параметр распределения;
* Кумулятивный или дифференциальный параметр.

**Упражнение 9:**

Нужно найти вероятность того, что завтра позвонят не менее 3-х клиентов, если всего клиентов 8, а вероятность, что позвонит каждый из них – 0.25.

На мой взгляд, наилучшим распределением здесь будет биномиальное, поскольку у нас есть число успехов в *n* испытаний. Таким образом, поскольку функция *БИНОМРАСП* работает в тех вариантах, когда нам нужно найти “не более”, то решим задачу от противного. Скажем, что вероятность того, что человек не позвонит равна 75%. А количество людей, которые не позвонят, не более 5. Вероятность того, что не более 5-ти человек не позвонят показана на рисунке 4.9.1.

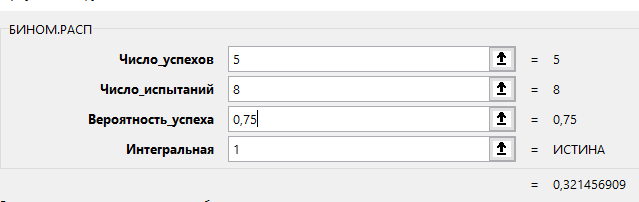


Рисунок 4.9.1 – Расчет вероятности от противного

Вероятность того, что не менее 3 человек позвонят равна 1 - 0.32 = 0.68 = 68%

**Упражнение 10:**

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

Задача достаточно тривиальная. Воспользуемся распределением Пуассона. Задача решена на рисунке 4.10.1.

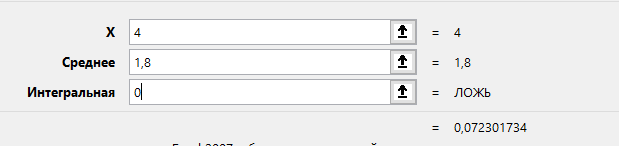


Рисунок 4.10.1 – Решение задачи

Таким образом, вероятность, что на ремонт отдадут 4 детали равна 7%.

**Упражнение 11:**

Из опрошенных 30% выразили симпатию к мужским пальто серого цвета, 26% – черное пальто, 22% – синего, 22% – коричневого. Проверим гипотезу о принадлежности данных опроса к равномерному распределению, используя критерий (при n = 100). Вычислим меру отклонения Пирсона.

Среднее количество человек в группе 100/4 = 25;

Отклонение для каждой группы равно соответственно В результате сложения получаем критерий отклонения Пирсона, равный 1.76.

Значение вычисление квантиля приведено на рисунке 4.11.1.

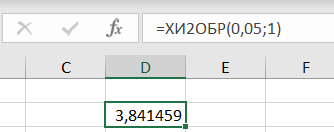


Рисунок 4.11.1 – Значение квантиля

Разность между данными значениями не существенно, поэтому разницы при выборе цвета нет (3.84 > 1.76).

Следующая часть упражнения имеет измененное условие. Количество человек, участвующих в выборке, равна не 100, а 1000. Среднее значение количества человек на группы равно 250. А отклонения, соответственно, равны 10, 0.4, 3.6, 3.6. Сумма отклонений равна 17.6, что значительно превышает значение квантиля в 3.84. Поэтому при большой выборке цвет оказывает бОльшее значение, чем при относительно маленькой.

**Вывод:**

Таким образом, в результате выполнения лабораторной работы, мы познакомились и изучили законы распределения случайных величин. Довольно подробно усвоили, в каких случаях нужно использовать определенное распределение, применимое на практике. Самое главное, научились находить распределения и вероятности с помощью функций *Excel*.

**Контрольные вопросы**:

1. Закон распределения случайной величины – это математическая инструкция, которая определяет соотношение связи между возможными значениями случайной величины и соответствующими им вероятности;
2. Экспоненциальный закон распределения применяется для определения надежности техники при нормальной ее эксплуатации, например, при определении срока службы радиоэлектронной аппаратуры;
3. Порядок ввода формул в *Excel* таков, что сначала мы определяем клетку, на которой будет выводится значение этой функции. Дальше выбираем вкладку “Формулы” → “Ввод функции”. Выбираем из списка функцию, которая нам необходима, а затем вводим аргументы для специфической функции. Это показано на рисунке 4.12.1;

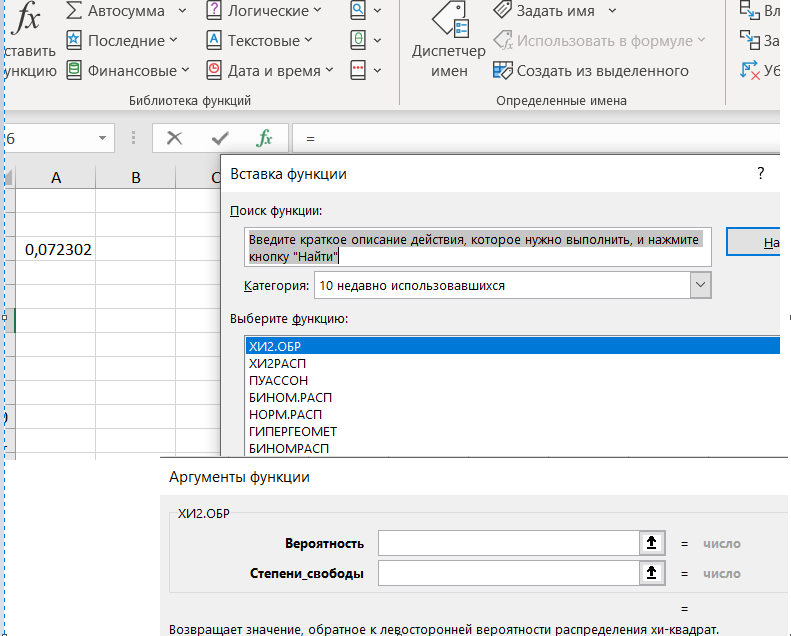


Рисунок 4.12.1 – Ввод функции

1. Параметры нормального распределения представляют собой значение, для которого рассчитывается вероятность, среднее значение, стандартное отклонение и интегрально-логическое значение (кумулятивная или дифференциальная);
2. Нормальный закон распределения описывает события, которые связана со сравнением с каким-то идеальным значением или средним (будь то попадание дротиков в тире, стрельба из лука, или рост людей, для которых шьется одежда разных размеров);
3. Дисперсия – это мера разброса значений случайной величины относительно ее математического ожидания;
4. Квантиль – это значение, которое заданная случайная величина не превышает с фиксированной вероятностью. Также ее называют процентилем, если вероятность задана в процентах;
5. Вообще, количество степеней свободы – это количество значений, которые способны варьироваться (причем линейно независимый каждый элемент степени свободы);
6. Критерий согласия Пирсона подразумевает наличие независимых случайных величин, каждая из которых имеет нормальное распределение и количество степеней свободы для определения квантиля;
7. Отличие в решении задач связана с измененным количеством измерений. Чем они больше, тем результат точнее.