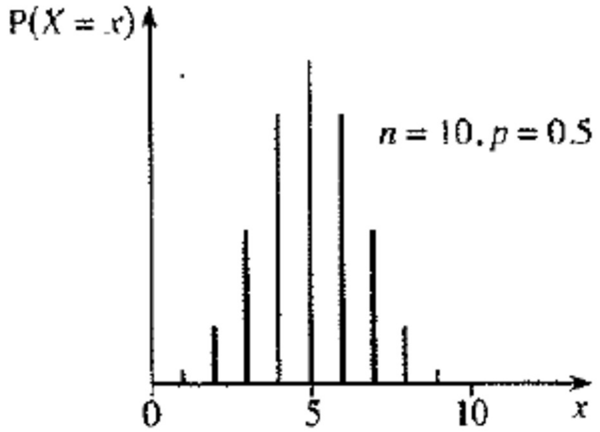
7. Проверка статистических гипотез: дискретные случайные величины
7.1. Проверка доли генеральной совокупности
Возможно, вы видели рекламу молочных спредов, в которых утверждается, что их нельзя отличить от сливочного масла. Как вы можете приступить к проверке этого утверждения? Одним из способов было бы сделать множество пар печенья: намазать одно печенье из этой пары сливочным маслом а второе печенье намазать молочным спредом. Пары печенья будут отданы нескольким дегустаторам, которых попросят идентифицировать печенье со сливочным маслом. Первая половина дегустаторов пробует сначала печенье со сливочным маслом а потом с молочным спредом. А вторая половина дегустаторов сначала пробует печенье с молочным спредом а потом со сливочным маслом.
Предположим, вы решили использовать 10 дегустаторов. Как бы вы начали делать выводы из ваших результатов? Метод проверки гипотез, описанный ранее, может быть адаптирован к этой ситуации. Сначала необходимо сформулировать нулевую гипотезу и альтернативную гипотезу. Обычно начинают с позиции сомнения: вы предполагаете, что дегустаторы не могут идентифицировать масло и что они гадают. В этой ситуации вероятность того, что выбранный случайным образом дегустатор получит правильный результат, составляет 0,5. Это можно выразить нулевой гипотезой $H_0: p = 0.5$. Если некоторые из дегустаторов действительно могут идентифицировать сливочное масло, тогда p> 0.5. Это можно выразить как альтернативную гипотезу $H_0: p > 0.5$.
Вы понимаете, почему трудно принять любую другую нулевую гипотезу?
Если H_0 верна, количество дегустаторов, которые правильно идентифицируют печенье со сливочным маслом X, является случайной величиной распределенной по биномиальному закону распределения, то есть X~B(10;0.5). На следующем рисунке показано это распределение. Высокие значения X предполагают, что H_0 следует отклонить в пользу H_1. Самый простой метод проведения проверки гипотезы для дискретной переменной - это рассчитать вероятность того, что X принимает наблюдаемое или более экстремальное значение, предполагая, что H_0 верна, и сравнить эту вероятность с указанным уровнем значимости. , Предположим, что 9 из 10 человек идентифицировали сливочное масло, а вы выбрали уровень значимости 5%.



$$P(X \geq 9) = P(X = 9) + P(X = 10) = C_{10}^{9}\left(\frac{1}{2}\right)^{9}\left(\frac{1}{2}\right) + C_{10}^{10}\left(\frac{1}{2}\right)^{10} =$$
$= 0.00976 \ldots + 0.000976 \ldots = 0.01074 \ldots = 1.07\%$, с точностью до трех значащих цифр.
Эта вероятность составляет менее 5%, поэтому результат значим на уровне 5%. H_0 отклоняется, и на уровне 5% значимости имеются доказательства того, что доля людей, способных различать сливочное масло, превышает 0.5.
Как и раньше, возможные значения X можно разделить на область принятия и область отклонения. Однако ситуация здесь осложняется тем, что X является дискретной переменной. Следующая таблица показывает закон распределения вероятностей X.

x	P(X=x)
0	0.0010
1	0.0098
2	0.0439
3	0.1172
4	0.2051
5	0.2461
6	0.2051
7	0.1172

8	0.0439
9	0.0098
10	0.0010

Для уровня значимости, скажем, 5%, вы обнаружите, что нет области отклонения, которая точно соответствует этой вероятности. Например, для области отклонения $X \geq 8$ вероятность результата в области отклонения равна 0.0439+0.0098+0.0010=0.0547, и поэтому фактический уровень значимости теста составляет 5,47%; для $X \geq 9$ фактический уровень значимости составляет 0,99%. Этот момент будет рассмотрен более подробно в дальнейшем.

Следующие примеры показывают проверку гипотез для дискретных случайных величин.

Пример 7.1.1.

Национальный опрос общественного мнения утверждает, что 40% избирателей проголосовали бы за партию R, если бы завтра были выборы. Студентка крупного колледжа подозревает, что доля молодых людей, которые будут голосовать за эту партию R, ниже. Она опросила 16 однокурсников, выбранных случайным образом из списка студентов колледжа, за какую партию они будут голосовать. Трое из 16-ти выбрали партию R. Покажите, на уровне значимости 10%, что это указывает на то, что заявленная цифра является слишком высокой для студентов колледжей.

Решение:

Нулевая и альтернативная гипотеза: $H_0: p = 0.4$ и $H_1: p < 0.4$ соответственно.

Если H_0 верна, то X~B(16;0.4).

$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) =$

$= C_{16}^0(0.4)^0(0.6)^{16} + C_{16}^1(0.4)^1(0.6)^{15} + C_{16}^2(0.4)^2(0.6)^{14} + C_{16}^3(0.4)^3(0.6)^{13}$

$= 0.00028 \dots + 0.00300 \dots + 0.01504 \dots + 0.04680 \dots = 0.06514 =$

$= 6.51\% < 10\%$

Результат является значимым для уровня значимости 10%. Это указывает на то, что указанная цифра является слишком высокой для студентов колледжа.

Пример 7.1.2.

Чтобы проверить монету на правильность, ее подбрасывают 12 раз. Результат - 9 ГЕРБов и 3 РЕШКи. Проверьте с уровнем значимости 10% что монета неправильная.

Решение:

Это двухсторонняя проверка на изменение, так как до того, как монета подброшена, нет никаких указаний, в каком направлении, если таковые имеются, вероятность может быть смещена. Если монета правильная, вероятность ГЕРБа (или РЕШКИ) равна 0.5.

Нулевая и альтернативная гипотеза: $H_0: p = 0.5$ и $H_1: p \neq 0.5$ соответственно.

Пусть X количество выпавших ГЕРБов из 12 подбрасывании. Если H_0 верна, то X~B(12; 0.5). В среднем вы ожидаете 6 ГЕРБов. Наблюдаемое значение 9 ГЕРБов выше среднего, так что:

$P(X \geq 9) = P(X = 9) + P(X = 10) + P(X = 11) + P(X = 12) =$

$= C_{12}^9\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right)^3 + C_{12}^{10}\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^2 + C_{12}^{11}\left(\frac{1}{2}\right)^{11}\left(\frac{1}{2}\right)^1 + C_{12}^{12}\left(\frac{1}{2}\right)^{12}\left(\frac{1}{2}\right)^0 =$

$= 0.05371 \dots + 0.01611 \dots + 0.00292 \dots + 0.00024 \dots = 0.07299 \dots = 7.30\%$, с точностью до трех значащих цифр.

Поскольку это двусторонняя проверка с уровнем значимости 10%, эту вероятность необходимо сравнивать с 5%. Так как 7,30%> 5%, результат незначителен, и нулевая гипотеза не отвергается: с уровнем значимости 10% недостаточно доказательств того, что монета неправильная.

Чтобы выполнить проверку гипотезы для дискретной случайной величины, рассчитайте вероятность наблюдаемого или более экстремального значения и сравните эту вероятность с уровнем значимости. Для односторонней проверки отклоните нулевую гипотезу, если эта вероятность меньше уровня значимости; для двухсторонней проверки отклоните нулевую гипотезу, если эта вероятность меньше половины уровня значимости.

7.2. Проверка доли генеральной совокупности для большого размера выборки

Когда размер выборки достаточной большой, вы можете рассчитать вероятности, используя тот факт, что биномиальное распределение может быть аппроксимировано нормальным распределением. Следующие примеры иллюстрируют метод.

Пример 7.2.1.

В экзаменационном тесте с несколькими вариантами ответов кандидат должен выбрать один из четырех возможных ответов на каждый вопрос. На тесте из 100 вопросов студент получает 34 правильных ответа. Проверьте с уровнем значимости 5% нулевую гипотезу о том, что ученик угадывает ответы.

Решение:

Если ученик угадывает ответы, то вероятность того, что один из ответов верен, равна 0,25. Если студент не угадывает, то доля правильных ответов должна быть больше.

Нулевая и альтернативная гипотеза: $H_0\colon p = \frac{1}{4}$ и $H_1\colon p \neq \frac{1}{4}$ соответственно.

Пусть X – количество правильных ответов. Если H_0 верна, то $X \sim B(100; \frac{1}{4})$.

Вероятность наблюдаемого или более экстремального значения равна $P(X \geq 34)$. Чтобы найти эту вероятность, X аппроксимируется нормальным распределением с параметрами: $\mu = np = 100 \cdot \frac{1}{4} = 25$, и

$\sigma^2 = 100 \cdot \frac{1}{4} \cdot \frac{3}{4} = 18.75$. Таким образом случайная величина $X \sim B(100; \frac{1}{4})$ аппроксимируется случайной величиной $V \sim N(25; 18.75)$.

$$P(X \geq 34) = P(V \geq 33.5) = P\left(Z \geq \frac{33.5 - 25}{\sqrt{18.75}}\right) = P(Z \geq 1.963) =$$
$$= 1 - \Phi(1.963) = 1 - 0.9752 = 0.0248 = 2.48\% < 5\%$$

Нулевая гипотеза может быть отклонена: на уровне значимости 5% имеются доказательства того, что учащийся не угадывает ответы.

В качестве альтернативы, нет необходимости проводить вычисления за пределами точки, где получено значение Z (= 1,963). Вместо этого вы можете использовать идею области отклонения, которая была показана ранее. Для односторонней проверки с уровнем значимости 5% область отклонения для статистики теста Z составляет $Z \geq 1.645$. Наблюдаемое значение Z находится в области отклонения, поэтому H_0 отклоняется. Вы можете использовать любой метод, который вы предпочитаете.

Ниже приведен пример двухсторонней проверки на изменение.

Пример 7.2.2.

Если в любой день недели вероятность рождения одинакова, то доля детей, рожденных в выходные, должна составлять $\frac{2}{7}$. Было установлено, что из случайной выборки из 490 детей 132 родились в выходные дни. Можно ли сказать исходя из этого, с уровнем значимости 5%, что доля детей, рожденных в выходные, отличается от $\frac{2}{7}$?

Решение:

Нулевая и альтернативная гипотеза: $H_0\colon p = \frac{2}{7}$ и $H_1\colon p \neq \frac{2}{7}$ соответственно.

Пусть X – количество детей, родившихся в выходные. Если H_0 верна, то

$X \sim N\left(490; \frac{2}{7}\right)$. Используем приближение случайной величины X случайной величиной

$V \sim N\left(490 \cdot \frac{2}{7}; 490 \cdot \frac{2}{7} \cdot \frac{5}{7}\right) = N(140; 100)$. Поскольку 132 меньше чем 140, среднее значение для H_0, находим:

$$P(X \leq 132) = P(V \leq 132.5) = P\left(Z \leq \frac{132.5 - 140}{\sqrt{100}}\right) = P(Z \leq -0.75)$$

Для двухсторонней проверки с уровнем значимости 5%, область отклонения: $|Z| \geq 1.96$. Расчетное значение $|Z| = 0.75$ и она не лежит в области отклонения. Доля детей, рожденных в выходные принимается равным $\frac{2}{7}$.

Если вы хотите работать с вероятностями, вы можете продолжить вычисление, $P(Z \leq -0.75) = 0.2266 = 22.66\%$. Поскольку 22.66%> 2.5% (помните, что это двухсторонняя проверка с уровнем значимости 5%), нулевая гипотеза не отвергается.

7.3. Проверка среднего значения генеральной совокупности по распределению Пуассона.

Идеи, разработанные ранее, могут также применяться, когда распределение Пуассона является подходящей моделью для генеральной совокупности, из которой взята выборка.

Пример 7.3.1.

Старый офисный копировальный аппарат выходил из строя, в среднем, три раза каждые две недели. Испытывается новый, более дорогой копировальный аппарат, который, как утверждают производители, является более надежным. В первые четыре недели использования этот новый копировальный аппарат выходил из строя всего один раз. Предполагая, что сбои копировального аппарата происходят независимо и случайным образом, проверьте с уровнем значимости 5%, есть ли доказательства того, что новый копировальный аппарат более надежен, чем старый.

Решение:

Если сбои копировального аппарата происходят независимо и случайным образом, то число сбоев в заданном интервале времени может быть смоделировано распределением Пуассона. Нулевой гипотезой будет гипотеза «без изменений»; старый копировальный аппарат выходил из строя в среднем 3 раза каждые две недели. Итак, λ, среднее количество сбоев в неделю, равно 1.5.

Нулевая и альтернативная гипотеза: $H_0: \lambda = 1.5$ и $H_1: \lambda \neq 1.5$ соответственно.

Пусть X – количество сбоев за 4 недели. Тогда если H_0 верна, то

$X \sim Po(4 \cdot 1.5) = Po(6)$

Наблюдаемое значение X равно 1.

$P(X \leq 1) = e^{-6} + e^{-6}\frac{6}{1!} = 0.002478 \ldots + 0.014872 \ldots = 0.01735 \ldots = 1.74\%$, с точностью до трех значащих цифр.

Поскольку 1,74% <5%, результат значим на уровне значимости 5%, и есть свидетельства того, что новый копировальный аппарат более надежен, чем старый.

Воодушевленный этим, офис решает продолжать использовать новый копировальный аппарат до конца года. Результаты анализируются в следующем примере.

Пример 7.3.2.

При дальнейшей эксплуатации нового копировального аппарата в оставшейся части первого года эксплуатации количество сбоев составила 57 раз. Используя результаты за весь год, проверьте с уровнем значимости 5%, является ли новый копировальный аппарат более надежным, чем старый.

Решение:

Нулевая и альтернативная гипотеза: $H_0: \lambda = 1.5$ и $H_1: \lambda \neq 1.5$ соответственно.

λ – среднее количество сбоев за неделю. Пусть X – количество сбоев за год, то есть 52 недели. Если H_0 верна, то $X \sim Po(52 \cdot 1.5) = Po(78)$.

Наблюдаемое значение X равна 1+57, то есть 58 сбоев за год. Так как λ>15, вероятность $P(X \leq 58)$ можно посчитать, используя подходящее приближение нормально распределенной случайной величиной.

Случайная величина $X \sim Po(78)$ может быть аппроксимирована случайной величиной $Y \sim N(78; 78)$ с использованием коррекции непрерывности.

$P(X \leq 58) = P(Y \leq 58.5) = P\left(Z \leq \frac{58.5-78}{\sqrt{78}}\right) = P(Z \leq -2.208)$

Для односторонней проверки с уровнем значимости 5% область отклонения $Z \leq -1.645$. Вычисленное значение статистики теста Z равна -2.208. И она лежит в области отклонения нулевой гипотезы. Данная проверка подтвердила прежний результат, то есть новый копировальный аппарат более надежен чем старый.

Если вы хотите дальше работать с вероятностями то продолжив вычисление, находим, что $P(Z \leq -2.208) = 0.0135 = 1.35\% < 5\%$. Поскольку 1.35%<5%, то нулевая гипотеза отклоняется.

Задачи.

№1. Ведущая газета сообщила, что 2 из каждых 3 болельщиц женского футбольного клуба смогли правильно объяснить правило офсайда. Сторонники Gaussian Rovers верили, что больше их поклонниц смогут правильно объяснить это правило. Чтобы доказать свою точку зрения, случайная выборка из 20 женщин-фанаток Gaussian Rovers была опрошена за пределами Белл-парка после матча, и 17 из них смогли правильно объяснить правило офсайда.

а) Проведите тест на уровне значимости 10% гипотезы о том, что доля женщин-фанаток Gaussian Rovers, способных правильно объяснить закон офсайда, превышает 2/3.

б) Сторонники "доказали свою точку зрения"?

Ответ №1: а) $H_0: p = \frac{2}{3}$, $H_1: p > \frac{2}{3}$, $X \sim B\left(20; \frac{2}{3}\right)$, $P(X \geq 17) = 0{,}0604$, поэтому отклоняется нулевая гипотеза и принимается альтернативная гипотеза, что доля больше чем $\frac{2}{3}$.

б) Статистические тесты не являются доказательством: они указывают на вероятность правильности гипотезы.

№2. Вероятность приземления чертежной булавки «лицом вверх» при падении на горизонтальный пол с высоты одного метра обозначается p. Когда чертежная булавка падала 25 раз, и она приземлялась «лицом вверх» 5 раз. Проверьте нулевую гипотезу – $p = 0{,}4$ и альтернативную гипотезу $p < 0{,}4$ с уровнем значимости 2,5%.

Ответ №2: $X \sim B(25; 0{,}4)$; $P(X \leq 5) = 0{,}0294 > 0{,}025$, поэтому принимаем нулевую гипотезу $p = 0{,}4$

№3. Предполагается, что треть всех математиков-левши. По результатам опроса из 174 математиков, 51 из них ответили что они левши. Предположив, что выборка была случайной, проведите проверку с уровнем значимости 5%, подтверждает ли выборка предположение (что треть математиков-левши).

Ответ №3: $H_0: \mu = \frac{1}{3}$, $H_1: \mu \neq \frac{1}{3}$, $X \sim B\left(174; \frac{1}{3}\right) \approx N(58; 38{,}67)$, $P(X \leq 51) = \Phi(-1{,}045)$, так как $-1{,}96 < -1{,}045 < 1{,}96$ (или так как $0{,}148 > 0{,}025$), принимаем нулевую гипотезу, что треть математиков-левши.

№4. При продвижении Doggo, нового корма для животных, утверждалось, что все больше собак предпочитают новый корм нынешнему лидеру бренда. 40 собакам был предоставлен выбор между Doggo и нынешним лидером бренда.

а) Найдите наименьшее количество собак, которые должны были бы предпочесть Doggo, чтобы утверждение промоутера было принято на уровне значимости 5%.

б) Каков же тогда уровень значимости теста?

Ответ №4: а) $H_0: p = \frac{1}{2}$, $H_1: p > \frac{1}{2}$, (p – доля собак предпочитающих Doggo), $X \sim B\left(40; \frac{1}{2}\right) \approx N(20; 10)$, $n = 26$. б) 4,1%

№5. Супермаркет покупает большую партию пластиковых пакетов у производителя для использования в магазине. В предыдущих партиях 7% пакетов были бракованными. Менеджер по контролю качества хочет проверить, имеет ли партия более высокий процент брака, чем 7%, и в этом случае партия будет возвращена производителю. Он осматривает 125 случайно выбранных пакетов и обнаруживает, что 14 из них бракованы. Проведите тест на уровне значимости 3% и укажите, следует ли возвращать партию.

Ответ №5: $H_0: p = 0{,}07$, $H_1: p > 0{,}07$; $X \sim B(125; 0{,}07) \approx N(8{,}75; 8{,}1375)$, $P(X \geq 14) = 1 - \Phi(1{,}665)$, так как $1{,}665 < 1{,}881$ (или так как $0{,}048 > 0{,}03$), принимаем нулевую гипотезу и оставляем партию