

Рыбак

Опубликовал

sobodv

Автор или источник

sobopedia

Предмет

Математическая Статистика (/Subjects/Details?id=5)

Тема

Теория проверки статистических гипотез (/Topics/Details?id=35)

Раздел

Критерий Хи-Квадрат Пирсона (/SubTopics/Details?id=128)

Дата публикации

28.05.2019

Дата последней правки

20.02.2022

Последний вносивший правки

sobodv

Рейтинг



Условие

Лаврентий каждое утро отправляется на рыбалку. На протяжении месяца он вел счет пойманных рыб. Его итоговая таблица получилась следующей:

Рыба	Карась	Окунь	Щука	Сом	Лещ
Количество	10	20	30	40	50

1. На уровне значимости 0.01 проверьте гипотезу о том, что вероятности поймать Карася, Окуня, Щуку, Сом и Леща составляют 0.1, 0.15, 0.2, 0.25 и 0.3 соответственно. Для этого воспользуйтесь тестом Хи-Квадрат Пирсона.
2. Найдите минимальный уровень значимости, при котором нулевая гипотеза будет отвергнута, то есть p-value.
3. Повторите предыдущие пункты при помощи теста отношения правдоподобия.

Решение

1. В данном случае мы имеем выборку размера $n = 10 + 20 + 30 + 40 + 50 = 150$ из мультиномиального распределения $M(1, p_1, p_2, p_3, p_4, p_5)$. Нулевая гипотеза заключается в том, что $H_0 : p_1^0 = 0.1, p_2^0 = 0.15, p_3^0 = 0.2, p_4^0 = 0.25, p_5^0 = 0.3$. Для проверки соответствующей гипотезы воспользуемся тестом Хи-квадрат Пирсона.

Рассмотрим реализацию тестовой статистики:

$$T(x) = \frac{(10 - 150 * 0.1)^2}{150 * 0.1} + \frac{(20 - 150 * 0.15)^2}{150 * 0.15} + \frac{(30 - 150 * 0.2)^2}{150 * 0.2} + \frac{(40 - 150 * 0.25)^2}{150 * 0.25} + \frac{(50 - 150 * 0.3)^2}{150 * 0.3} = \frac{8}{3} \approx 2.667$$

Критическое значение составляет $\chi_{5-1}^{2,0.99} = 13.2767$. Поскольку $2.667 \leq 13.2767$, то нулевая гипотеза не отвергается.

2. Очевидно, что поскольку используемый тест является правосторонним, то:

$$p\text{-value} = 1 - F_{\chi_4^2}(2.667) \approx 0.615$$

3. Для начала рассмотрим модель без ограничений, предварительно заменив p_5 на $(1 - p_1 - p_2 - p_3 - p_4)$. Нетрудно показать, что оценки метода максимального правдоподобия для выборки из мультиномиального распределения принимают следующий вид: $(\hat{p}_F^{ML} | X = x) = (\frac{10}{150}, \frac{20}{150}, \frac{30}{150}, \frac{40}{150})$. При условии рассмотренных ограничений оценки метода максимального правдоподобия будут $(\hat{p}_R^{ML} | X = x) = (0.1, 0.15, 0.2, 0.25)$

Рассчитаем значение тестовой статистики:

$$\begin{aligned} 2 \ln \left(\frac{L(\hat{p}_F^{ML}; x)}{L(\hat{p}_R^{ML}; x)} \right) &= 2 \ln \left(\frac{L(\frac{10}{150}, \frac{20}{150}, \frac{30}{150}, \frac{40}{150}; x)}{L(0.1, 0.15, 0.2, 0.25; x)} \right) = \\ &= 2 \ln \left(\left(\frac{10}{150} \right)^{10} \left(\frac{20}{150} \right)^{20} \left(\frac{30}{150} \right)^{30} \left(\frac{40}{150} \right)^{40} \left(1 - \frac{10}{150} - \frac{20}{150} - \frac{30}{150} - \frac{40}{150} \right)^{50} \right) - \\ &\quad - 2 \ln \left((0.1)^{10} (0.15)^{20} (0.2)^{30} (0.25)^{40} (1 - 0.1 - 0.15 - 0.2 - 0.25)^{50} \right) = \\ &= 2 * (-223.461 + 224.902) = 2.882 \end{aligned}$$

Поскольку на параметры были наложены четыре ограничения, то критическое значение составляет $X_4^{2,0.99} = 13.2767$. Поскольку $2.882 \leq 13.2767$, то нулевая гипотеза не отвергается.

Рассчитаем p-value по аналогии с предыдущим пунктом:

$$\text{p-value} = 1 - F_{\chi_4^2}(2.882) \approx 0.578$$

Показать решение

Пожалуйста, войдите или зарегистрируйтесь, чтобы оценивать задачи, добавлять их в избранные и совершать некоторые другие, дополнительные действия.