Кот ученый

Опубликовал
sobodv
Автор или источник
sobopedia
Предмет
Математическая Статистика (/Subjects/Details?id=5)
Тема
Метод максимального правдоподобия (/Topics/Details?id=31)
Раздел
Введение в ММП (/SubTopics/Details?id=109)
Дата публикации
08.01.2022
Дата последней правки
08.01.2022
Последний вносивший правки
sobodv
Рейтинг

Условие

Каждый день кот ученый мяукает до тех пор, пока его не покормят. Вероятность того, что кота покормят после очередного "мяу", не зависит от числа изданных ранее "мяу" и всегда равняется $p \in (0, 1)$. Ученый кот собрал выборку из количества мяуканий, которые ему пришлось произвести прежде, чем его покормили. Помогите ученому коту:

1. Оценить параметр p при помощи метода максимального правдоподобия.

2. Найти асимптотическое распределение ММП оценки параметра p.

3. Получить ММП оценку вероятности того, что ученого кота покормят раньше, чем он успеет трижды мяукнуть.

4. Найти асимптотическое распределение найденной в предыдущем пункте оценки, а также ее асимптотическую дисперсию и ее оценку.

5. По выборке из $n = 1000$ наблюдений оказалось, что $\overline{x}_n = 5$. Найдите, приблизительно, вероятность того, что оценка найденной в предыдущих пунктах вероятности превысит 0.35.

Решение

1. Поскольку в данном случае речь идет о выборке из геометрического распределения, то при $t \in \{1, 2, 3, \ldots\}$ получаем:

$$P(X_1 = t) = (1 - p)^{t-1} p$$

Пользуясь соответствующей функций вероятностей запишем функцию правдоподобия:

$$L(p; x) = \prod_{i=1}^{n} (1 - p)^{x_i - 1} p$$

Логарифм функции правдоподобия имеет вид:

$$\ln L(p;x) = n\ln(p) + \ln(1-p)\sum_{i=1}^{n}(x_i-1)$$

В соответствии с условиями первого порядка:

$$\frac{d\ln L(p;x)}{dp} = \frac{n}{p} - \frac{\sum_{i=1}^{n}(x_i-1)}{1-p} = 0$$

Решая данное равенство получаем точку, подозреваемую на максимум:

$$p^* = \frac{1}{\overline{x}_n}$$

Покажем, что функция правдоподобия является вогнутой, рассмотрев условия второго порядка:

$$\frac{d^2\ln L(p;x)}{d^2p} = -\frac{n}{p^2} - \frac{\sum_{i=1}^{n}(x_i-1)}{(1-p)^2} = n\left(\frac{1}{(1-p)^2} - \frac{1}{p^2}\right) - \frac{n\overline{x}_n}{(1-p)^2} \leq n\left(\frac{1}{(1-p)^2} - \frac{1}{p^2}\right) - \frac{n\times 1}{(1-p)^2} = -\frac{n}{p^2} < 0$$

Из полученного результата следует, что ММП оценка имеет следующий вид:

$$\hat{p}_n = \frac{1}{\overline{X}_n}$$

2. Сперва найдем информацию Фишера:

$$\ln L(p;X_1) = \ln(p) + \ln(1-p)(X_1-1)$$

$$\frac{d\ln L(p;X_1)}{dp} = \frac{1}{p} - \frac{X_1-1}{1-p}$$

$$\frac{d^2\ln L(p^2;X_1)}{dp} = -\frac{1}{p^2} - \frac{X_1-1}{(1-p)^2}$$

$$i(p) = -E\left(-\frac{1}{p^2} - \frac{X_1-1}{(1-p)^2}\right) = \frac{1}{p^2} + \frac{E(X_1)-1}{(1-p)^2} = \frac{1}{p^2} + \frac{\frac{1}{p}-1}{(1-p)^2} = \frac{1}{(1-p)p^2}$$

Теперь найдем асимптотическую дисперсию ММП оценки и ее оценку:

$$As.Var(\hat{p}_n) = \frac{1}{ni(p)} = \frac{(1-p)p^2}{n}$$

$$\widehat{As.Var}(\hat{p}_n) = \frac{1}{ni(\hat{p}_n)} = \frac{(1-\hat{p}_n)\hat{p}_n^2}{n} = \frac{\overline{X}_n-1}{n\left(\overline{X}_n\right)^3}$$

3. Обратим внимание, что:

$$P(X_1<3) = 1-P(X_1\geq 3) = 1-(1-p)^2$$

Поскольку данная вероятность является монотонной функцией от оцениваемого параметра, то применимо свойство инвариантности, вследствие которого получаем ММП оценку:

$$\hat{P}(X_1<3) = 1-(1-\hat{p}_n)^2 = 1-\left(1-\frac{1}{\overline{X}_n}\right)^2$$

4. Поскольку функция непрерывна, то для нахождения асимптотического распределения можно применить дельта метод. Обратим внимание, что:

$$P'(X_1 < 3) = 2(1-p)$$

Отсюда получаем, что:

$$As.Var(\hat{P}(X_1 < 3)) = P'(X_1 < 3) As.Var(\hat{p}_n) = (2(1-p))^2 \frac{(1-p)p^2}{n} = \frac{4(1-p)^3 p^2}{n}$$

$$\widehat{As.Var}(\hat{P}(X_1 < 3)) = \frac{4(1-\hat{p}_n)^3 \hat{p}_n^2}{n} = \frac{4\left(1 - \frac{1}{\overline{X}_n}\right)^3 \left(\frac{1}{\overline{X}_n}\right)^2}{n}$$

5. Поскольку $\overline{x}_n(x) = 5$, то $\hat{p}_n(x) = 0.2$, откуда:

$$\hat{P}(X_1 < 3) \dot{\sim} \mathcal{N}\left(1 - (1-0.2)^2, \frac{4(1-0.2)^3 \times 0.2^2}{1000}\right) = \mathcal{N}\left(0.36, 0.00008192\right)$$

В результате получаем:

$$P\left(\hat{P}(X_1 < 3) > 0.35\right) \approx 1 - \Phi\left(\frac{0.35 - 0.36}{\sqrt{0.00008192}}\right) \approx 0.865$$

Проверка в R:

```
options(scipen = 999)
n <- 1000
p <- 0.2
n.sim <- 10000
p.est <- rep(NA, n.sim)
prob.est <- rep(NA, n.sim)
asvar.est <- rep(NA, n.sim)
asvar.prob <- rep(NA, n.sim)
for (i in 1:n.sim)
{
 x <- rgeom(n, p) + 1
 p.est[i] <- 1 / mean(x)
 prob.est[i] <- 1 - (1 - p.est[i]) ^ 2
 asvar.est[i] <- (1 - p.est[i]) * (p.est[i] ^ 2) / n
 asvar.prob[i] <- 4 * (1 - p.est[i]) ^ 3 * p.est[i] ^ 2 / n
}
# 1
mean(p.est)
# 2
var(p.est)
mean(asvar.est)
# 3
mean(x < 3)
mean(prob.est)
1 - (1 - p) ^ 2
# 4
mean(asvar.prob)
var(prob.est)
4 * (1 - p) ^ 3  * p ^ 2 / n
# 5
mean(prob.est > 0.35)
```

```
1 - pnorm((0.35 - prob.est[1]) / sqrt(asvar.prob[1]))
```

Показать решение