Вор должен сидеть в тюрьме

Опубликовал
sobodv
Автор или источник
sobopedia
Предмет
Математическая Статистика (/Subjects/Details?id=5)
Тема
Доверительные интервалы (/Topics/Details?id=33)
Раздел
Введение в доверительные интервалы (/SubTopics/Details?id=114)
Дата публикации
16.04.2020
Дата последней правки
14.02.2023
Последний вносивший правки
sobodv
Рейтинг
★

Условие

Рабочее время (в часах), которое требуется Жеглову на раскрытие очередного преступления, подчиняется экспоненциальному распределению с параметром λ и не зависит от времени, потраченного на раскрытие других преступлений. За год Жеглов раскрыл 146 преступлений, потратив на это 4380 часов.

Через $X=(X_1,\cdots,X_{146})$ обозначим выборку, наблюдения в которой отражают время, затраченное Жегловым на раскрытие преступления. Через $x=(x_1,\cdots,x_n)$ будем обозначать реализацию данной выборки.

Подсказка: Во-первых, если $a>0$ и $V\sim EXP(\lambda)$, то $aV\sim EXP\left(\frac{\lambda}{a}\right)$. Во-вторых, если $V\sim EXP(0.5)$, то $V\sim\chi^2(2)$.

1. Найдите распределение случайной величины $2\lambda\sum\limits_{i=1}^{146}X_i$. Проверьте, является ли данная случайная величина центральной статистикой.

2. Постройте симметричный 95%-й доверительный интервала для параметра λ, а также найдите реализацию этого доверительного интервала.

3. Найдите реализацию симметричного 95%-го асимптотический доверительный интервал для параметра λ используя распределение оценки, полученное с помощью метода максимального правдоподобия.

4. Найдите реализацию симметричного 95%-го асимптотический доверительный интервал для параметра λ альтернативным способом (не прибегая к ММП оценке).

5. Сравните реализации обычного и асимптотических доверительных интервалов, посчитанных в предыдущих пунктах.

6. Найдите реализации границ обыкновенного и асимптотического 95%-ых доверительных интервалов для ожидаемого времени на раскрытие одного преступления, используя реализации соответствующих границ доверительных интервалов для λ.

Решение

1. Пусть $i, j \in \{1, \cdots, 146\}$. Обратим внимание, что в силу первой подсказки $2\lambda X_i \sim EXP(0.5)$, а значит, согласно второй подсказке $2\lambda X_i \sim \chi^2(2)$. Используя свойство суммы независимых Хи-квадрат случайных величин получаем:

$$\sum_{i=1}^{146} 2\lambda X_i \sim \chi^2(2 * 146) = \chi^2(292)$$

Данная функция является центральной статистикой для параметра λ, поскольку ее распределение не зависит от соответствующего параметра и $2\lambda \sum\limits_{i=1}^{146} x_i$ строго монотонна по λ. Следовательно, ее можно взять за основу для построения доверительного интервала.

2. Через $\chi^2_{k,\alpha}$ обозначим квантиль уровня α случайной величины с распределением $\chi^2\,(k)$, где $k \in N$.

Используя результат предыдущего пункта получаем:

$$P\left(\chi^2_{292,0.025} \leq 2\lambda \sum_{i=1}^{146} X_i \leq \chi^2_{292,0.975}\right) = 0.95$$

Перепишем данный результат в следующем виде:

$$P\left(\frac{\chi^2_{292,0.025}}{2\sum\limits_{i=1}^{n} X_i} \leq \lambda \leq \frac{\chi^2_{292,0.975}}{2\sum\limits_{i=1}^{n} X_i}\right) = 0.95$$

В итоге получаем следующие границы 95%-го доверительного интервала для λ:

$$T_1(X) = \frac{\chi^2_{292,0.025}}{2\sum\limits_{i=1}^{n} X_i}$$

$$T_2(X) = \frac{\chi^2_{292,0.975}}{2\sum\limits_{i=1}^{n} X_i}$$

Учитывая, что $\chi^2_{292,0.025} = 246.5571$ и $\chi^2_{292,0.975} = 341.2296$, сам доверительный интервал будет иметь вид:

$$\left(\frac{123.2785}{\sum\limits_{i=1}^{n} X_i}, \frac{170.6148}{\sum\limits_{i=1}^{n} X_i}\right)$$

Теперь найдем реализации границ данного доверительного интервала:

$$T_1(x) = (T_1(X)|X = x) = \frac{123.2785}{4380} \approx 0.02814578$$

$$T_2(x) = (T_1(X)|X = x) = \frac{170.6148}{4380} \approx 0.03895315$$

В итоге реализацией данного доверительного интервала будет:

$$[0.02814578, 0.03895315]$$

3. Запишем ММП оценку и оценку информации Фишера:

$$\hat{\lambda}_{146} = \frac{1}{\overline{X}_n}, \qquad i(\hat{\lambda}_{146}) = \frac{1}{\hat{\lambda}_n^2}$$

Рассчитаем реализации соответствующий оценок:

$$\hat{\lambda}_{146}(x) = \frac{146}{4380} = \frac{1}{30}, \qquad i(\hat{\lambda}_{146}(x)) = \frac{1}{\left(\frac{1}{30}\right)^2} = 900$$

Используя надлежащую квантиль стандартного нормального распределения получаем искомую реализацию:

$$\left[\frac{1}{30} - 1.96\sqrt{\frac{1}{146 \times 900}}, \frac{1}{30} + 1.96\sqrt{\frac{1}{146 \times 900}}\right] \approx [0.02792631, 0.03874036]$$

4. Обратим внимание, что $E(X_i) = \frac{1}{\lambda}$ и $Var(X_i) = \frac{1}{\lambda^2}$ для любого $i \in \{1, \cdots, n\}$. Пользуясь данным обстоятельством и используя ЦПТ получаем, что:

$$\frac{\overline{X} - \frac{1}{\lambda}}{\sqrt{\frac{1}{n\lambda^2}}} = \sqrt{n}\left(\lambda\overline{X} - 1\right) \xrightarrow{d} \mathcal{N}(0, 1)$$

Обозначим через z_α квантиль стандартного нормального распределения уровня α. Обратим внимание, что:

$$P\left(z_{0.025} \leq \sqrt{n}\left(\lambda\overline{X} - 1\right) \leq z_{0.975}\right) = 0.95$$

Подставляя $z_{0.025} = -1.959964$ и $z_{0.975} = 1.959964$ имеем:

$$P\left(\frac{\sqrt{n} - 1.959964}{\sqrt{n}\overline{X}} \leq \lambda \leq \frac{\sqrt{n} + 1.959964}{\sqrt{n}\overline{X}}\right)$$

В итоге, полагая $n = 146$, получаем асимптотический 95%-й доверительный интервал:

$$\left[\frac{\sqrt{146}-1.959964}{\sqrt{146}\overline{X}}, \frac{\sqrt{146}+1.959964}{\sqrt{146}\overline{X}}\right]$$

Реализация данного доверительного интервала имеет вид:

$$\left[\frac{\sqrt{146}-1.959964}{\sqrt{146}*\frac{4380}{146}}, \frac{\sqrt{146}+1.959964}{\sqrt{146}*\frac{4380}{146}}\right] = (0.027926408, 0.038740259]$$

5. Нетрудно заметить, что различия в реализациях доверительных интервалов обоих типов оказались несущественными:

$$\text{Обычный: } [0.02814578, 0.03895315]$$

$$\text{Асимптотические с ММП: } [0.027926408, 0.03874036]$$

$$\text{Асимптотические без ММП: } [0.02792631, 0.038740259]$$

6. Поскольку $E(X_i) = \frac{1}{\lambda}$ является непрерывной и строго убывающей функцией от λ, то нетрудно показать, что:

$$P\left(\frac{\chi^2_{292,0.025}}{2\sum\limits_{i=1}^{n} X_i} \le \lambda \le \frac{\chi^2_{292,0.975}}{2\sum\limits_{i=1}^{n} X_i}\right) = P\left(\frac{2\sum\limits_{i=1}^{n} X_i}{\chi^2_{292,0.975}} \le \frac{1}{\lambda} \le \frac{2\sum\limits_{i=1}^{n} X_i}{\chi^2_{292,0.025}}\right) = 0.95$$

Поэтому реализация обычного доверительного интервала примет вид:

$$\left[\frac{1}{0.03895315}, \frac{1}{0.02814578}\right] = (25.67186, 35.52930)$$

Для нахождения реализации асимптотического доверительного интервала проще всего воспользоваться дельта методом. Обозначим $g(\lambda) = \frac{1}{\lambda}$, а значит $g'(\lambda) = -\frac{1}{\lambda^2}$, откуда получаем искомую реализацию:

$$\text{Асимптотический: } \left[\frac{1}{\frac{1}{30}} - 1.96\sqrt{\frac{\left(-\frac{1}{\left(\frac{1}{30}\right)^2}\right)^2}{146\times 900}}, \frac{1}{\frac{1}{30}} + 1.96\sqrt{\frac{\left(-\frac{1}{\left(\frac{1}{30}\right)^2}\right)^2}{146\times 900}}\right] \approx [25.13368, 34.86632]$$

Если бы нам была также известная реализация выборочной дисперсии, то также можно было бы воспользоваться формулой для асимптотического доверительного интервала для математического ожидания.

Показать решение