Домашнее задание 3

Юрасов Никита Андреевич

Обновлено 24 октября 2019 г.

Содержание

Эта работа представляет собой отчет к Домашнему Заданию №3.
Так как в моделировании используется пакет `numpy.random`, то, пусть, для него будет выставлено по умолчанию стартовое значение генератора 12345678.

1 Нахождение выборочного среднего и выборочной дисперсии

Определение 1.1. *Пусть $X_1, X_2, \cdots, X_n$ – выборка из какого-то распределения вероятности. Тогда ее выборочным средним называется случайная величина*

$$\overline{X} = \frac{1}{n}\sum_{i=0}^{n} X_i$$

Определение 1.2. *Пусть $X_1, X_2, \cdots, X_n$ – выборка из какого-то распределения вероятности. Тогда выборочная дисперсия – это случайная величина*

$$S_n^2 = \frac{1}{n}\sum_{i=0}^{n} \left(X_i - \overline{X}\right)^2,$$

где $\overline{X}$ – выборочное среднее.

Для нахождения этих двух значений можно воспользоваться методами `mean` и `var` библиотеки `numpy`, но для реализации была написана собственная функция `sample_mean` и `sample_variance` соответственно. Время работы на больших выборках почти одинаковое.

1.1 Распределение Пуассона

```
Выборочное среднее для выборки [2. 1. 1. 3. 0.] = 1.4
Выборочное среднее для выборки [2. 1. 2. 3. 7. 2. 3. 1. 2. 0.] = 2.3
Выборочная дисперсия для выборки [2. 1. 1. 3. 0.] = 1.04
Выборочная дисперсия для выборки [2. 1. 2. 3. 7. 2. 3. 1. 2. 0.] = 3.21
```

1.2 Распределение Эрланга

```
Выборочное среднее для выборки [ 9.43737905 11.94755981 1.6335522 11.63186523
1.95757948] = 7.321587156076815
Выборочное среднее для выборки [ 8.30297653 17.47684737 5.71182291 2.67860603
19.66877258 7.92660288 5.52776384 8.11891813 9.22277337 16.35132395] = 10.098640758
Выборочная дисперсия для выборки [ 9.43737905 11.94755981 1.6335522 11.63186523
1.95757948] = 21.11620307310885
Выборочная дисперсия для выборки [ 8.30297653 17.47684737 5.71182291 2.67860603
19.66877258 7.92660288 5.52776384 8.11891813 9.22277337 16.35132395] = 29.294400894
```

2 Нахождение параметров распределений событий

Для каждого из двух распределений будем строить оценку максимального правдоподобия.

2.1 Распределение Пуассона

Пусть функция распределения будет выглядеть следующим образом:

$$f(x, \theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x \geq 0$$

Тогда функция правдоподобия:

$$L(x, \theta) = \prod_{i=0}^{n} f(x, \theta) = \prod_{i=0}^{n} \left(\frac{\theta^{x_i} e^{-\theta}}{x_i!} \right) = e^{\theta n} \frac{\theta^{\sum_{i=0}^{n} x_i}}{\prod_{i=0}^{n} x_i!}$$

Возьмем от функции правдоподобия натуральный логарифм:

$$lnL(x, \theta) = -\theta n + \sum_{i=0}^{n} x_i \cdot ln\theta - ln \prod_{i=0}^{n} x_i!$$

Продифференцируем полученное выражение по θ и приравняем к нулю:

$$\frac{\partial lnL(x, \theta)}{\partial \theta} = -n + \frac{\sum_{i=0}^{n} x_i}{\theta} = 0$$

Будем решать это уравнение относительно θ:

$$\hat{\theta} = \frac{1}{n} \sum_{i=0}^{n} x_i$$

В итоге, получается, что оценка максимального правдоподобия параметра θ распределения Пуассона имеет вид выборочного среднего (см. определение 1.1).

Предложенная оценка $\hat{\theta}$ является <u>несмещенной</u>, так как выборочное среднее является в свою очередь несмещенной оценкой.

Состоятельность можно проверить по утверждению, что выборочные моменты k-го порядка сходятся к k-ым моментам X, то есть:

$$\hat{\alpha_k} \xrightarrow{P} MX^k$$

В нашем случае k=1:

$$\hat{\alpha_k} \xrightarrow{P} MX$$

Следовательно оценка $\hat{\theta} = \hat{\alpha_k}$, которая является <u>состоятельной</u>.

Эффективность представленной оценки также подтверждается, так как $\hat{\theta}$ – оценка максимального правдоподобия, а такая оценка <u>эффективна</u>.

2.2 Распределение Эрланга

Рассмотрим функцию распределения, которая зависит от двух параметров:

$$f(x,m,\lambda)=\frac{\lambda^m x^{m-1}}{\Gamma(m)}e^{-\lambda x},\quad m\in\mathbb{N},\lambda>0,x>0$$

Построим функцию правдоподобия:

$$L(x,m,\lambda)=\prod_{i=0}^{n}f(x,m,\lambda)=\left(\frac{\lambda^m}{\Gamma(m)}\right)^n\prod_{i=0}^{n}x_i^{m-1}e^{-\lambda x_i}=\left(\frac{\lambda^m}{\Gamma(m)}\right)^n e^{-\lambda\sum_{i=0}^{n}x_i}\prod_{i=0}^{n}x_i^{m-1}$$

Возьмем натуральный логарифм от $L(x,m,\lambda)$:

$$lnL(x,m,\lambda)=mnln\lambda+(m-1)\prod_{i=0}^{n}lnx_i-\lambda\sum_{i=0}^{n}x_i-nln\Gamma(m)$$

Продифференцируем полученное по λ и приравняем к нулю:

$$\frac{\partial lnL(x,m,\lambda)}{\partial\lambda}=\frac{mn}{\lambda}-\sum_{i=0}^{n}x_i=0$$

Решая относительно λ, получим:

$$\hat{\lambda}=\frac{mn}{\sum\limits_{i=0}^{n}x_i}$$

Также можно предложить оценку параметра λ, используя метод моментов.