Теоретическое домашнее задание по теме «Сопряжённые распределения и эксепененциальный класс распределений»

Солоткий Михаил, 417 группа ВМК МГУ

28 сентября 2018 г.

1 Сопряжённое к равномерному

1.1 Оценка максимального правдоподобия θ_{ML}

Пусть дана простая выборка $x_1, \dots, x_n$ из непрервыного равномерного распределения:
$p(x_i|\theta) = \dfrac{1}{\theta} \cdot [0 \le x_i \le \theta]$. Вывести θ_{ML}.

$L(X|\theta) = \prod\limits_{i=1}^{n} \dfrac{1}{\theta} \cdot [0 \le x_i \le \theta]$

$L(X|\theta) = 0$ если $\max\limits_{i=1,\dots,n}(x_i) > \theta$ или $\min\limits_{i=1,\dots,n}(x_i) < 0$. В последнем случае любое $\theta = \theta_{ML}\ \forall\ \theta > 0$

Предположим все $x_i \ge 0$. Тогда $L(X|\theta) < L(X|\max\limits_{i=1,\dots,n} x_i)\ \forall\ \theta > \max\limits_{i=1,\dots,n} x_i$. Значит $\theta_{ML} = \max\limits_{i=1,\dots,n} x_i$.

1.2 Сопряжённое априорное и апастериорное распределение

$p(\theta|a,b) = \dfrac{b \cdot a^b}{\theta^{b+1}} \cdot [\theta \ge a]$

$p(x|\theta) = \dfrac{1}{\theta} \cdot [\theta \ge x]$

$p(\theta|X,a,b) = \dfrac{p(X|\theta) \cdot p(\theta|a,b)}{\int p(x|\theta) \cdot p(\theta|a,b) d\theta} = \dfrac{1}{z} \cdot \dfrac{b \cdot a^b}{\theta^{b+n+1}} \cdot [\theta \ge \max\limits_{i=1,\dots,n}(a, x_i)]$

$z = \int\limits_0^\infty \dfrac{b \cdot a^b}{\theta^{b+n+1}} \cdot [\theta \ge \max\limits_{i=1,\dots,n}(a,x_i)] d\theta = \int\limits_{\max\limits_{i=1,\dots,n}(a,x_i)}^{\infty} \dfrac{b \cdot a^b}{\theta^{b+n+1}} d\theta = -\dfrac{b \cdot a^b}{\theta^{b+n}}\Bigg|_{\max\limits_{i=1,\dots,n}(a,x_i)}^{\infty} = \dfrac{b \cdot a^b}{\left(\max\limits_{i=1,\dots,n}(a,x_i)\right)^{b+n} \cdot (b+n)}$

$p(\theta|X,a,b) = \dfrac{\left(\max\limits_{i=1,\dots,n}(a,x_i)\right)^{b+n} \cdot (b+n)}{\theta^{b+n+1}} = \mathrm{Pareto}(\theta|a',b')$; $a' = \max\limits_{i=1,\dots,n}(a,x_i)$; $b' = b+n$.

1.3 Вычисление статистик

$\xi \sim \mathrm{Pareto}(x|a,b)$

- $\mathsf{E}\xi = \int\limits_a^\infty \dfrac{b \cdot a^b}{x^{b+1}} \cdot x dx = \int\limits_a^\infty \dfrac{b \cdot a^b}{x^b} dx = \dfrac{b \cdot a^b}{(b-1) \cdot a^{b-1}} = \dfrac{ab}{b-1}$
- $\dfrac{1}{2} = \int\limits_a^c \dfrac{b \cdot a^b}{x^{b+1}} dx = a^b \cdot \left(\dfrac{1}{a^b} - \dfrac{1}{c^b}\right) = 1 - \left(\dfrac{a}{c}\right)^b$

  $\left(\dfrac{a}{c}\right)^b = \dfrac{1}{2}$

  $Med\xi = c = a \cdot 2^{\frac{1}{b}}$
- $Mod\xi = a$

2 Задача про автобусы

В новом городе я не знаю, какие автобусы ходят чаще каких. Поэтому предположим, что наблюдать любой номер автобуса равновероятно. Предположим, номера автобусов нумеруются подряд, начиная с первого. Обобщим дискретное равномерное распределение на непрерывное таким образом: номер автобуса - целая часть сверху от случайной величины, распределённой непрерывно на $[0, \theta]$. Параметр θ надо оценить в задаче. Понятно, что миллиарда автобусных маршрутов в городе нет, но чтобы не рассматривать ограниченные множества непонятного размера, загрубим модель: скажем, что возможно любое вещественное $\theta > 0$. Возьмём в качестве априорного распределение Парето:

$$p(\theta|a, b) = \frac{b \cdot a^b}{x^{b+1}} \cdot [\theta \geq a]$$

Параметр a можно выбрать исходя из знания от том, сколько автобусов в родном городе. Параметр b отвечает за степень доверия априорным знаниям: чем оно выше, тем больше разница менее вероятны значения, далёкие от a. Пронаблюдав автобус с номером 100, а затем автобус с номером 150 параметр а может измениться: если он раньше был больше или равен 150, то не изменится, в противном случае станет равным 150, и значения θ, большие 150 приобретут большую вероятность. При наблюдении автобуса с номером 50, большИе значения параметра a получат меньшую вероятность, так как b увеличится на единицу и хвосты станут легче.
Брать моду распределения в качестве точечной оценки значит не использовать априорной информации (если, конечно в родном горое автобусов было не больше 150).
Брать медиану можно в принципе ровно как и мат. ожидание. В любом из вариантов используется набюдаемая информация и опыт, результат будет больше 150. Насколько больше - определяется параметром b. Мат. ожидание - взвешивание с вероятностями всех допустимых значений, то есть как бы довольно много информации из распределения было взято из апастериорного распределения. В общем мат. ожидание брать адекватно. Никаких за или против медианы у меня нет, слишком неочевидно.

3 Экспоненциальный класс: распределение Парето

$$p(x|\theta) = \frac{f(x)}{g(\theta)} \cdot e^{\theta^T u(x)}$$

$$p(x|a, b) = \frac{b \cdot a^b}{x^{b+1}} \cdot [x \geq a] = b \cdot a^b \cdot e^{-(b+1) \cdot \ln x} \cdot [x \geq a] = \frac{b \cdot a^b}{x} \cdot e^{-b \cdot \ln x} \cdot [x \geq a]$$

$$g(b) = \frac{1}{b \cdot a^b}; \ f(x) = \frac{[x \geq a]}{x}$$

$$\mathsf{E}(\ln x) = g(b)' = -\frac{a^b + b \cdot a^b \cdot \ln a}{b^2 \cdot a^{2 \cdot b}} = -\frac{1 + b \cdot \ln a}{b^2 \cdot a^b}$$