Достаточные статистики

Статистика $T = T(\mathbb{X}_n)$ называется достаточной для модели $F(x,\theta), \theta \in \Theta$,если условная плотность $L(x_1, x_2, ..., x_n \,|\, T(\mathbb{X}_n) = t; \theta)$ (или условная вероятность $P\left\{X_1 = x_1, ..., X_n = x_n \,|\, T(\mathbb{X}_n) = t\right\}$ в дискретном случае) случайного вектора $\mathbb{X}_n$ при условии $T(\mathbb{X}_n) = t$ не зависит от параметра θ.

Достаточные статистики

Статистика $T = T(\mathbb{X}_n)$ называется достаточной для модели $F(x,\theta), \theta \in \Theta$,если условная плотность $L(x_1, x_2, ..., x_n \left| T\left(\mathbb{X}_n\right) = t; \theta\right)$ (или условная вероятность $P\left\{X_1 = x_1, ..., X_n = x_n \mid T\left(\mathbb{X}_n\right) = t\right\}$ в дискретном случае) случайного вектора $\mathbb{X}_n$ при условии $T(\mathbb{X}_n) = t$ не зависит от параметра θ.

Свойство достаточности статистики $T(\mathbb{X}_n)$ означает, что выборка содержит всю информацию о параметре θ, имеющуюся в выборке и поэтому все заключения, которые можно сделать при наблюдении $\mathbb{X}_n$зависит только от $T(\mathbb{X}_n)$.

Достаточные статистики

Статистика $T = T(\mathbb{X}_n)$ называется достаточной для модели $F(x,\theta), \theta \in \Theta$,если условная плотность $L(x_1, x_2, ..., x_n \,|\, T(\mathbb{X}_n) = t; \theta)$ (или условная вероятность $P\{X_1 = x_1, ..., X_n = x_n \,|\, T(\mathbb{X}_n) = t\}$ в дискретном случае) случайного вектора $\mathbb{X}_n$ при условии $T(\mathbb{X}_n) = t$ не зависит от параметра θ.

Свойство достаточности статистики $T(\mathbb{X}_n)$ означает, что выборка содержит всю информацию о параметре θ, имеющуюся в выборке и поэтому все заключения, которые можно сделать при наблюдении $\mathbb{X}_n$зависит только от $T(\mathbb{X}_n)$. Следовательно, достаточная статистика дает оптимальный в определенном смысле способ представления статистических данных, что особенно важно при обработке большого объема статистической информации.

Теорема (критерий факторизации)
Для того, чтобы статистика $T(\mathbb{X}_n)$ была достаточной для θ, необходимо и достаточно, чтобы функция правдоподобия имела вид

$$L(\mathbb{X}_n, \theta) = g(T(\mathbb{X}_n); \theta) h(\mathbb{X}_n),$$

где функция $g(t, \theta)$ зависит от выборки только через $T(\mathbb{X}_n) = t$, а функция $h(\mathbb{X}_n)$ не зависит от θ.

Теорема (критерий факторизации)
Для того, чтобы статистика $T(\mathbb{X}_n)$ была достаточной для θ, необходимо и достаточно, чтобы функция правдоподобия имела вид

$$L(\mathbb{X}_n,\theta)=g(T(\mathbb{X}_n);\theta)h(\mathbb{X}_n),$$

где функция $g(t,\theta)$ зависит от выборки только через $T(\mathbb{X}_n)=t$, а функция $h(\mathbb{X}_n)$ не зависит от θ.

Замечание 1. Всякая R-эффективная оценка является одновременно достаточной статистикой.
Однако обратное не всегда верно – достаточная статистика может существовать, но не быть эффективной.

Теорема (критерий факторизации)
Для того, чтобы статистика $T(\mathbb{X}_n)$ была достаточной для θ, необходимо и достаточно, чтобы функция правдоподобия имела вид

$$L(\mathbb{X}_n, \theta) = g(T(\mathbb{X}_n); \theta) h(\mathbb{X}_n),$$

где функция $g(t, \theta)$ зависит от выборки только через $T(\mathbb{X}_n) = t$, а функция $h(\mathbb{X}_n)$ не зависит от θ.

Замечание 1. Всякая R-эффективная оценка является одновременно достаточной статистикой.
Однако обратное не всегда верно – достаточная статистика может существовать, но не быть эффективной.

Замечание 2. Если статистика $T(\mathbb{X}_n)$ достаточная, то достаточной является и любая взаимно-однозначная функция $\varphi(T(\mathbb{X}_n))$.

Пример. Найдем достаточную статистику для $N(\theta,\sigma)$.

Пример. Найдем достаточную статистику для $N(\theta,\sigma)$.

$$L(\mathbb{X}_n,\theta)=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\theta)^2}=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2-2\theta X_i+\theta^2}=$$

Пример. Найдем достаточную статистику для $N(\theta,\sigma)$.

$$L(\mathbb{X}_n,\theta)=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\theta)^2}=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2-2\theta X_i+\theta^2}=$$

$$=\underbrace{\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{\frac{\theta}{\sigma^2}\underbrace{\sum_{i=1}^{n}X_i}_{T(\mathbb{X}_n)}-\frac{n\theta^2}{2\sigma^2}}}_{g(T(\mathbb{X}_n,\theta)}\cdot\underbrace{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2}}_{h(\mathbb{X}_n)}=g(T(\mathbb{X}_n,\theta)h(\mathbb{X}_n)$$

Пример. Найдем достаточную статистику для $N(\theta,\sigma)$.

$$L(\mathbb{X}_n,\theta)=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\theta)^2}=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2-2\theta X_i+\theta^2}=$$

$$=\underbrace{\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{\frac{\theta}{\sigma^2}\underbrace{\sum_{i=1}^{n}X_i}_{T(\mathbb{X}_n)}-\frac{n\theta^2}{2\sigma^2}}}_{g(T(\mathbb{X}_n,\theta)}\cdot\underbrace{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}X_i^2}}_{h(\mathbb{X}_n)}=g(T(\mathbb{X}_n,\theta)h(\mathbb{X}_n)$$

По критерию факторизации $T(\mathbb{X}_n)=\sum_{i=1}^{n}X_i$ - достаточная статистика для параметра θ.

Теорема (Рао–Блекуэлл–Колмогоров)

НОРМД, если она существует, является функцией от достаточной статистики.

Методы оценивания

Методы оценивания

Принцип подстановки: пусть параметр θ является функцией (явной или неявной) от некоторых характеристик распределения случайной величины. Подставим в эту функцию выборочные характеристики вместо истинных. Получим оценку параметра θ по принципу подстановки.

Методы оценивания

Принцип подстановки: пусть параметр θ является функцией (явной или неявной) от некоторых характеристик распределения случайной величины. Подставим в эту функцию выборочные характеристики вместо истинных. Получим оценку параметра θ по принципу подстановки.

Например, пусть известно, что $X \sim N(a,\sigma)$. Для параметров нормального распределения выполняется

$$a = EX,\ \sigma = \sqrt{DX}\ .$$

Методы оценивания

Принцип подстановки: пусть параметр θ является функцией (явной или неявной) от некоторых характеристик распределения случайной величины. Подставим в эту функцию выборочные характеристики вместо истинных. Получим оценку параметра θ по принципу подстановки.

Например, пусть известно, что $X \sim N(a,\sigma)$. Для параметров нормального распределения выполняется

$$a = EX,\ \sigma = \sqrt{DX}.$$

Значит, в качестве оценок по принципу подстановки можно использовать:

$$a^* = \bar{X},\ \sigma^* = \sqrt{S^2}.$$

Методы оценивания

Принцип подстановки: пусть параметр θ является функцией (явной или неявной) от некоторых характеристик распределения случайной величины. Подставим в эту функцию выборочные характеристики вместо истинных. Получим оценку параметра θ по принципу подстановки.

Например, пусть известно, что $X \sim N(a, \sigma)$. Для параметров нормального распределения выполняется

$$a = EX, \ \sigma = \sqrt{DX}.$$

Значит, в качестве оценок по принципу подстановки можно использовать:

$$a^* = \bar{X}, \ \sigma^* = \sqrt{S^2}.$$

Достоинства: простота; все другие известные способы основаны на этом принципе.

Методы оценивания

Принцип подстановки: пусть параметр θ является функцией (явной или неявной) от некоторых характеристик распределения случайной величины. Подставим в эту функцию выборочные характеристики вместо истинных. Получим оценку параметра θ по принципу подстановки.

Например, пусть известно, что $X \sim N(a,\sigma)$. Для параметров нормального распределения выполняется

$$a = EX, \ \sigma = \sqrt{DX} .$$

Значит, в качестве оценок по принципу подстановки можно использовать:

$$a^* = \bar{X}, \ \sigma^* = \sqrt{S^2} .$$

Достоинства: простота; все другие известные способы основаны на этом принципе.

Недостаток: оценка не всегда однозначна.

На принципе подстановки основаны

- метод моментов;

На принципе подстановки основаны

- метод моментов;
- метод максимального правдоподобия.

На принципе подстановки основаны

- метод моментов;
- метод максимального правдоподобия.

Метод моментов

- вместо моментов теоретического распределения подставляются моменты эмпирического распределения.

На принципе подстановки основаны

- метод моментов;
- метод максимального правдоподобия.

Метод моментов

- вместо моментов теоретического распределения подставляются моменты эмпирического распределения.

Пусть для параметра θ известно, что k-й момент $EX^k = h(\theta)$, где h - некоторая функция. Тогда оценка по методу моментов (ММ-оценка) является решением уравнения

$$EX^k = h(\theta)\Big|_{\theta=\theta^*} = \overline{X^k} = \frac{1}{n}\sum X_i^k .$$

На принципе подстановки основаны

- метод моментов;
- метод максимального правдоподобия.

Метод моментов

- вместо моментов теоретического распределения подставляются моменты эмпирического распределения.

Пусть для параметра θ известно, что k-й момент $EX^k = h(\theta)$, где h - некоторая функция. Тогда оценка по методу моментов (ММ-оценка) является решением уравнения

$$EX^k = h(\theta)\Big|_{\theta = \theta^*} = \overline{X^k} = \frac{1}{n}\sum X_i^k .$$

Замечания.

1. Оценка может быть не единственная (т.к. параметр может быть выражен через различные моменты).

На принципе подстановки основаны

- метод моментов;
- метод максимального правдоподобия.

Метод моментов

- вместо моментов теоретического распределения подставляются моменты эмпирического распределения.

Пусть для параметра θ известно, что k-й момент $EX^k = h(\theta)$, где h - некоторая функция. Тогда оценка по методу моментов (ММ-оценка) является решением уравнения

$$EX^k = h(\theta)\Big|_{\theta=\theta^*} = \overline{X^k} = \frac{1}{n}\sum X_i^k .$$

Замечания.

1. Оценка может быть не единственная (т.к. параметр может быть выражен через различные моменты).

2. Если $\theta = (\theta_1, ..., \theta_k)$, то необходимо решить k уравнений для k различных моментов.

3. Можно рассмотреть более общий случай:
Пусть параметр θ удовлетворяет уравнению: $Eg(X) = h(\theta)$,
где g - некоторая функция (ранее было $g(X) = X^k$).

3. Можно рассмотреть более общий случай:
Пусть параметр θ удовлетворяет уравнению: $Eg(X) = h(\theta)$, где g - некоторая функция (ранее было $g(X) = X^k$). Тогда ММ-оценка θ^* - решение уравнения $\overline{g(X)} = h(\theta^*)$:

$g(X)$ называется пробной функцией.

3. Можно рассмотреть более общий случай:
Пусть параметр θ удовлетворяет уравнению: $Eg(X) = h(\theta)$, где g - некоторая функция (ранее было $g(X) = X^k$). Тогда ММ-оценка θ^* - решение уравнения $\overline{g(X)} = h(\theta^*)$:

$g(X)$ называется пробной функцией.

Теорема 5 (состоятельность метода моментов). Пусть $\theta^* = h^{-1}(\overline{X^k})$ - ММ-оценка параметра θ, причем функция h^{-1} непрерывна. Тогда θ^* состоятельна.

3. Можно рассмотреть более общий случай:
Пусть параметр θ удовлетворяет уравнению: $Eg(X) = h(\theta)$, где g - некоторая функция (ранее было $g(X) = X^k$). Тогда ММ-оценка θ^* - решение уравнения $\overline{g(X)} = h(\theta^*)$:

$g(X)$ называется пробной функцией.

Теорема 5 (состоятельность метода моментов). Пусть $\theta^* = h^{-1}(\overline{X^k})$ - ММ-оценка параметра θ, причем функция h^{-1} непрерывна. Тогда θ^* состоятельна.

Доказательство. По закону больших чисел, $\overline{X^k} \xrightarrow{p} EX^k$ при $n \to \infty$.

Так как h^{-1} непрерывна, $\theta^* = h^{-1}(\overline{X^k}) \xrightarrow{p} h^{-1}(EX^k) = \theta$.

Пример.

Пусть известно, что $X \sim U(a,b)$ (равномерное распределение). По свойству распределения,

$$EX = \frac{b+a}{2},\ DX = \frac{(b-a)^2}{12}.$$

Пример.

Пусть известно, что $X \sim U(a,b)$ (равномерное распределение). По свойству распределения,

$$EX = \frac{b+a}{2},\ DX = \frac{(b-a)^2}{12}.$$

Система уравнений:

$$\left.\begin{aligned} \bar{X} &= \frac{b+a}{2} \\ S &= \frac{b-a}{2\sqrt{3}} \end{aligned}\right\}$$

Пример.

Пусть известно, что $X \sim U(a,b)$ (равномерное распределение). По свойству распределения,

$$EX = \frac{b+a}{2},\ DX = \frac{\left(b-a\right)^2}{12}.$$

Система уравнений:

$$\left.\begin{aligned} \overline{X} &= \frac{b+a}{2} \\ S &= \frac{b-a}{2\sqrt{3}} \end{aligned}\right\} \Rightarrow \begin{cases} b^* + a^* = 2\overline{X} \\ b^* - a^* = 2S\sqrt{3} \end{cases}$$

Пример.

Пусть известно, что $X \sim U(a,b)$ (равномерное распределение). По свойству распределения,

$$EX = \frac{b+a}{2},\ DX = \frac{(b-a)^2}{12}.$$

Система уравнений:

$$\left.\begin{aligned} \bar{X} &= \frac{b+a}{2} \\ S &= \frac{b-a}{2\sqrt{3}} \end{aligned}\right\} \Rightarrow \begin{cases} b^* + a^* = 2\bar{X} \\ b^* - a^* = 2S\sqrt{3} \end{cases} \Rightarrow \begin{aligned} a^* &= \bar{X} - S\sqrt{3} \\ b^* &= \bar{X} + S\sqrt{3}. \end{aligned}$$

Метод максимального правдоподобия

Пусть X – дискретная случайная величина,

$$P(X = x_i) = p_i(\theta),\ i = 1, \dots, k$$

вид функций $p_i(\theta)$ известен.

Метод максимального правдоподобия

Пусть X – дискретная случайная величина,

$$P(X = x_i) = p_i(\theta),\ i = 1,\dots,k$$

вид функций $p_i(\theta)$ известен. Имеется выборка v наблюдений над X объема n. Пусть значение x_i наблюдалось с частотой n_i.

Метод максимального правдоподобия

Пусть X – дискретная случайная величина,

$$P(X = x_i) = p_i(\theta),\ i = 1,...,k$$

вид функций $p_i(\theta)$ известен. Имеется выборка v наблюдений над X объема n. Пусть значение x_i наблюдалось с частотой n_i.

Представим выборку как набор частот $v = (n_1,...,n_k)$

Метод максимального правдоподобия

Пусть X – дискретная случайная величина,

$$P(X = x_i) = p_i(\theta),\ i = 1,\dots,k$$

вид функций $p_i(\theta)$ известен. Имеется выборка v наблюдений над X объема n. Пусть значение x_i наблюдалось с частотой n_i.

Представим выборку как набор частот $v = (n_1,\dots,n_k)$

Вероятность получить выборку:

$$\boxed{L(\theta,v) = \prod_{i=1}^{k}\left(p_i(\theta)\right)^{n_i}}.$$

Метод максимального правдоподобия

Пусть X – дискретная случайная величина,

$$P(X = x_i) = p_i(\theta),\ i = 1,\ldots,k$$

вид функций $p_i(\theta)$ известен. Имеется выборка v наблюдений над X объема n. Пусть значение x_i наблюдалось с частотой n_i.

Представим выборку как набор частот $v = (n_1,\ldots,n_k)$

Вероятность получить выборку:

$$\boxed{L(\theta,v) = \prod_{i=1}^{k}\left(p_i(\theta)\right)^{n_i}}.$$

Функция $L(\theta,v)$ - функция правдоподобия; зависит от неизвестного параметра θ и выборки.

Метод максимального правдоподобия: в качестве оценки θ^* брать то значение, при котором вероятность $L(\theta, v)$ для данной выборки максимальна.

θ^* - оценка максимального правдоподобия параметра θ.

Пусть теперь X – непрерывная случайная величина с плотностью вероятности $f(x,\theta)$. Функция правдоподобия:

$$\boxed{L(\mathbf{X}_n,\theta)=\prod_{i=1}^{n} f(X_i,\theta)}\,.$$

Пусть теперь X – непрерывная случайная величина с плотностью вероятности $f(x,\theta)$. Функция правдоподобия:

$$\boxed{L(\mathbf{X}_n,\theta)=\prod_{i=1}^{n} f(X_i,\theta)}\,.$$

Пусть теперь X – непрерывная случайная величина с плотностью вероятности $f(x,\theta)$. Функция правдоподобия:

$$\boxed{L(\mathbf{X}_n,\theta)=\prod_{i=1}^{n} f(X_i,\theta)}\,.$$

Оценка максимального правдоподобия – (МП-оценка) в качестве параметра θ берется значение θ^*, в которой $L(\mathbf{X}_n,\theta)$ достигает максимума.

Пусть теперь X – непрерывная случайная величина с плотностью вероятности $f(x,\theta)$. Функция правдоподобия:

$$\boxed{L(\mathbf{X}_n,\theta)=\prod_{i=1}^{n} f(X_i,\theta)}.$$

Оценка максимального правдоподобия – (МП-оценка) в качестве параметра θ берется значение θ^*, в которой $L(\mathbf{X}_n,\theta)$ достигает максимума.

Замечание. Удобно рассматривать логарифмическую функцию правдоподобия (точки в которых достигается максимум совпадают):

$$\ln L(\mathbf{X}_n,\theta)=\sum_{i=1}^{k} n_i \ln p_i(\theta)$$ (для дискретных величин)

Пусть теперь X – непрерывная случайная величина с плотностью вероятности $f(x,\theta)$. Функция правдоподобия:

$$\boxed{L(\mathbf{X}_n,\theta)=\prod_{i=1}^{n} f(X_i,\theta)}.$$

Оценка максимального правдоподобия – (МП-оценка) в качестве параметра θ берется значение θ^*, в которой $L(\mathbf{X}_n,\theta)$ достигает максимума.

Замечание. Удобно рассматривать логарифмическую функцию правдоподобия (точки в которых достигается максимум совпадают):

$$\ln L(\mathbf{X}_n,\theta)=\sum_{i=1}^{k} n_i \ln p_i(\theta) \quad \text{(для дискретных величин)}$$

или

$$\ln L(\mathbf{X}_n,\theta)=\sum_{i=1}^{n} \ln f(X_i,\theta) \quad \text{(для непрерывных величин).}$$

Свойства ОМП:

1. Если существует R-эффективная оценка $T\left(\mathbb{X}_n\right)$, то $\hat{\theta}=T(\mathbb{X}_n)$, так как по критерию эффективности Рао-Крамера

$$\frac{\partial \ln L(\mathbb{X}_n,\theta)}{\partial\theta}=\frac{1}{a\left(\theta\right)}\left[T(\mathbb{X}_n)-\theta\right]=0\Rightarrow T(\mathbb{X}_n)=\theta.$$

2. Если существует достаточная статистика $T=T(\mathbb{X}_n)$ и *ОМП* существует и единственная, то она является функцией от T. Так как по теореме факторизации $L(\mathbb{X}_n,\theta)=g(T(\mathbb{X}_n),\theta)h(\mathbb{X}_n)$ и максимизация функции правдоподобия эквивалентна максимизации $g(T(\mathbb{X}_n);\theta)$ по $\theta\Rightarrow\hat{\theta}$ зависит от $T(\mathbb{X}_n)$.

3. *ОМП* является инвариантной относительно преобразования параметров, то есть $\hat{\tau}\left(\theta\right)=\tau\left(\hat{\theta}\right)$ если $\tau\left(\theta\right)$ - взаимнооднозначное преобразование.

Свойства ОМП:

4. *ОМП* является асимптотически несмещенной, то есть $M\hat{\theta} \to \theta, n \to \infty$.

5. Если модель является регулярной, а функция правдоподобия $\forall n \geq 1$ и $\mathbb{X}_n \in \mathbb{X}$ имеет один локальный максимум, то $\hat{\theta}$ является состоятельной оценкой параметра θ то есть $\hat{\theta} \xrightarrow{p} \theta, \forall \theta \in \Theta$.

6. Если F регулярная, а $L(\mathbb{X}_n, \theta)$ имеет один максимум, $f(x,\theta)$ трижды дифференцируема и $\exists M(x): \forall \theta \in \Theta \left| \frac{\partial^3 \ln f(x,\theta)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| < H(x)$, то при $n \to \infty$

$\sqrt{n}(\widehat{\theta_n} - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))$, где $I(\theta)$ - информационная матрица Фишера.

Свойства ОМП:

7. Если $\tau(\theta)$ непрерывно дифференцируемая функция от θ и $\hat{\tau}_n = \tau(\hat{\theta})$ - *ОМП* $\tau(\theta)$, то

$$\sqrt{n}(\widehat{\tau_n} - \tau(\theta)) \xrightarrow{d} N(0, \sigma_\tau{}^2)$$

$$\sigma_\tau{}^2 = b^T(\theta) I^{-1}(\theta) b(\theta), b(\theta) = \left(\frac{\partial \tau(\theta)}{\partial \theta_1}, \dots \frac{\partial \tau(\theta)}{\partial \theta_r} \right).$$

.

Алгоритм оценивания методом МП

Предполагается:

1. Вид закона распределения известен;

Алгоритм оценивания методом МП

Предполагается:

1. Вид закона распределения известен;
2. Плотность вероятности(в случае непрерывного распределения) - гладкая функция.

Последовательность решения:

Алгоритм оценивания методом МП

Предполагается:

1. Вид закона распределения известен;
2. Плотность вероятности(в случае непрерывного распределения) - гладкая функция.

Последовательность решения:

1. Составляется функция правдоподобия
2. Вычисляется логарифм функции правдоподобия

Алгоритм оценивания методом МП

Предполагается:

1. Вид закона распределения известен;
2. Плотность вероятности(в случае непрерывного распределения) - гладкая функция.

Последовательность решения:

1. Составляется функция правдоподобия

2. Вычисляется логарифм функции правдоподобия

3. Оценки параметров получаются в результате решения системы уравнений вида:

$$\frac{\partial \ln L(\mathbf{X}_n, \theta)}{\partial \theta} = 0$$

Алгоритм оценивания методом МП

Предполагается:

1. Вид закона распределения известен;
2. Плотность вероятности(в случае непрерывного распределения) - гладкая функция.

Последовательность решения:

1. Составляется функция правдоподобия
2. Вычисляется логарифм функции правдоподобия
3. Оценки параметров получаются в результате решения системы уравнений вида:

$$\frac{\partial \ln L(\mathbf{X}_n, \theta)}{\partial \theta} = 0$$

4. Проверяется условие максимума функции правдоподобия

Пример. Пусть $X \sim N(a, \sigma)$, т.е. $\theta = (a, \sigma)$.

Пример. Пусть $X \sim N(a, \sigma)$, т.е. $\theta = (a, \sigma)$.

$$L(a,\sigma) = \frac{1}{(\sigma\sqrt{2\pi})^{n}} \exp\left(\sum_{i=1}^{n} -\frac{(x_i - a)^2}{2\sigma^2}\right)$$

Пример. Пусть $X \sim N(a, \sigma)$, т.е. $\theta = (a, \sigma)$.

$$L(a, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left(\sum_{i=1}^{n} -\frac{(x_i - a)^2}{2\sigma^2}\right)$$

$$\ln L(a, \sigma) = -n \ln \sigma - n \ln \sqrt{2\pi} - \frac{1}{2\sigma^2} \sum_{i} (x_i - a)^2 .$$

Пример. Пусть $X \sim N(a, \sigma)$, т.е. $\theta = (a, \sigma)$.

$$L(a,\sigma) = \frac{1}{(\sigma\sqrt{2\pi})^{n}} \exp\left(\sum_{i=1}^{n} -\frac{(x_i - a)^2}{2\sigma^2}\right)$$

$$\ln L(a,\sigma) = -n\ln\sigma - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_i (x_i - a)^2 .$$

Уравнения:

$$\frac{\partial}{\partial a}\ln L(a,\sigma) = \frac{2}{2\sigma^2}\sum_i (x_i - a) = 0$$

Пример. Пусть $X \sim N(a, \sigma)$, т.е. $\theta = (a, \sigma)$.

$$L(a,\sigma) = \frac{1}{(\sigma\sqrt{2\pi})^{n}} \exp\left(\sum_{i=1}^{n} -\frac{(x_i - a)^2}{2\sigma^2}\right)$$

$$\ln L(a,\sigma) = -n\ln\sigma - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i}(x_i - a)^2 .$$

Уравнения:

$$\frac{\partial}{\partial a}\ln L(a,\sigma) = \frac{2}{2\sigma^2}\sum_{i}(x_i - a) = 0$$

$$\frac{\partial}{\partial \sigma}\ln L(a,\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i}(x_i - a)^2 = 0$$

Пример 2. Пусть $X \sim U(a, b)$.

Плотность равномерного распределения

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & else. \end{cases} \Rightarrow$$

Пример 2. Пусть $X \sim U(a, b)$.

Плотность равномерного распределения

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & else. \end{cases} \Rightarrow$$

$$L(\mathbf{X}_n, a, b) = \begin{cases} \dfrac{1}{(b-a)^n}, \forall X_i \in [a,b] \\ 0, \quad else. \end{cases}$$

Пример 2. Пусть $X \sim U(a, b)$.

Плотность равномерного распределения

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & else. \end{cases} \Rightarrow$$

$$L(\mathbf{X}_n, a, b) = \begin{cases} \dfrac{1}{(b-a)^n}, \forall X_i \in [a,b] \\ 0, \quad else. \end{cases}$$

$L \ = \max \Leftrightarrow \ \forall X_i \in [a,b]$ и $b - a = \min \Leftrightarrow$

Пример 2. Пусть $X \sim U(a, b)$.

Плотность равномерного распределения

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0, & else. \end{cases} \Rightarrow$$

$$L(\mathbf{X}_n, a, b) = \begin{cases} \dfrac{1}{(b-a)^n}, \forall X_i \in [a,b] \\ 0, \quad else. \end{cases}$$

$L \ = \max \Leftrightarrow \ \forall X_i \in [a,b]$ и $b - a = \min \Leftrightarrow$

$$a^* = \min(x_i), \ b^* = \max(x_i)$$