Д/З №13

соотв. плотности

№1 $KL(P,Q) = E \log \frac{p(\xi)}{q(\xi)}$, где $\xi \sim P$

а) $KL(U(0,1), U(0,\theta)) =$

$= E \log p(\xi) - E \log q(\xi) =$

$\theta \geq 1 \Rightarrow E \log q(\xi) = \int_0^1 1 \cdot \log \frac{1}{\theta} dx = \log \frac{1}{\theta}$

$\theta < 1 \Rightarrow E \log q(\xi) = \int_0^\theta 1 \cdot \log \frac{1}{\theta} dx + \int_\theta^1 1 \cdot \log 0 \, dx = \theta \cdot \log \frac{1}{\theta} + (1-\theta) \cdot \log 0 =$

$= \int_0^1 1 \log 1 \, dx - \log \frac{1}{\theta} = \log(\theta)$ // при $\theta \geq 1$

б) $KL(Exp(\theta), Exp(\lambda)) =$

• Равном. распред. долго остаются "похожими", если менять парамет

$= \int_0^{+\infty} \theta e^{-\theta x} \log(\lambda e^{-\lambda x}) dx = \theta \int_0^{+\infty} e^{-\theta x} \log \lambda \, dx - \theta \lambda \int_0^{+\infty} e^{-\theta x} \cdot x \, dx$

$= \theta \log \lambda \cdot \frac{1}{-\theta} e^{-\theta x} \Big|_0^{+\infty} - \theta \lambda \int_0^{+\infty} e^{-(\theta x)} \frac{(\theta x) \theta dx}{\theta \theta} =$

$= + \log \lambda - \frac{\lambda}{\theta} \int_0^{+\infty} e^{-x} x \, dx = \log \lambda - \frac{\lambda}{\theta}$

a) $KL(Exp(\theta), Exp(\lambda)) =$

$$= \int_0^{+\infty} \theta e^{-\theta x} \log\left(\frac{\theta}{\lambda} e^{-(\theta-\lambda)x}\right) dx =$$

$$= \theta \int_0^{+\infty} e^{-\theta x} \log\frac{\theta}{\lambda}\, dx - \theta \int_0^{+\infty} e^{-\theta x} (\theta-\lambda) x\, dx =$$

$$= \theta \log\frac{\theta}{\lambda} \frac{1}{-\theta}(-1) - \theta(\theta-\lambda) \int_0^{+\infty} e^{-\theta x} \frac{\theta x \cdot \theta dx}{\theta \cdot \theta} =$$

$$= + \log\frac{\theta}{\lambda} - \left(1 - \frac{\lambda}{\theta}\right) \cdot \int_0^{+\infty} e^{-x} x\, dx =$$

$$= -\log\frac{\lambda}{\theta} - 1 + \frac{\lambda}{\theta} = \frac{\lambda}{\theta} - \log\frac{\lambda}{\theta} - 1$$

- Похожесть экспоненц. распределений определяется $k = \frac{\lambda}{\theta}$ – отнош. параметров и растёт линейно по нему.

б) $KL(Pois(\theta), Pois(\lambda)) =$

$$= \sum_{k=0}^{+\infty} \frac{e^{-\theta}\theta^k}{k!} \cdot \log\frac{e^{-\theta}\cdot\theta^k \cancel{k!}}{\cancel{k!}\, e^{-\lambda}\cdot\lambda^k} = (\lambda-\theta)\sum_{k=0}^{+\infty}\frac{e^{-\theta}\theta^k}{k!} + \log\left(\frac{\theta}{\lambda}\right)\sum_{k=1}^{+\infty}\frac{e^{-\theta}\theta^k}{(k-1)!} =$$

$$= (\lambda-\theta)\cdot 1 + \theta \log\frac{\theta}{\lambda}\cdot 1 = (\lambda-\theta) + \theta\log\frac{\theta}{\lambda}$$

- Похожесть пуас. распр. зависит от ~~[illegible]~~ порядка сравнив. параметров и растёт линейно по второму параметру.

---

* Под записью log подразум. натур. логарифм.

№2 $X_1 \ldots X_n$ из распр $Pois(\theta)$

1) Считаем лог. ф-ию правдоподобия и градиент.

$$l_X(\theta) = \sum_{i=1}^{n} \log\left(\frac{e^{-\theta}\theta^{x_i}}{x_i!}\right) = -\sum \log(x_i!) - n\theta + \sum x_i \cdot \log\theta$$

$$\boxed{\nabla l = \frac{\partial l_X}{\partial \theta} = -n + \frac{\sum x_i}{\theta}}$$

~~Из~~ ~~гради~~ ОМП $\hat{\theta} = \overline{X}$.

2) Считаем натур. градиент

$$\nabla_N l = i^{-1}(\theta) \cdot \nabla l$$

$$i(\theta) = -E_\theta \frac{\partial^2 l_{X_1}(\theta)}{\partial \theta^2} = -E_\theta \frac{\partial}{\partial \theta}\left(-1 + \frac{x_1}{\theta}\right) =$$

$$= -E - \frac{x_1}{\theta^2} = \theta^{-2} \cdot Ex = \frac{1}{\theta}$$

$$i^{-1}(\theta) = \theta$$

$$\boxed{\nabla_N l = -n\theta + \sum x_i}$$

Оптимиз. по градиенту:

$$\theta_{t+1} = \theta_t + \eta\left(\frac{\sum x}{\theta_t} - n\right)$$

По натур. градиенту:

$$\theta_{t+1} = \theta_t + \eta\left(\sum x_i - n\theta_t\right)$$

Метод Ньютона (запускаясь из $\theta_t \approx 0$)

$$\theta_{t+1} = \theta_t - \left(-\frac{\theta_t^2}{\sum x_i}\right)\cdot\left(\frac{\sum x_i}{\theta_t} - n\right) = 2\theta_t - \frac{n\theta_t^2}{\sum x_i}$$

Сравнение сходимости методов:

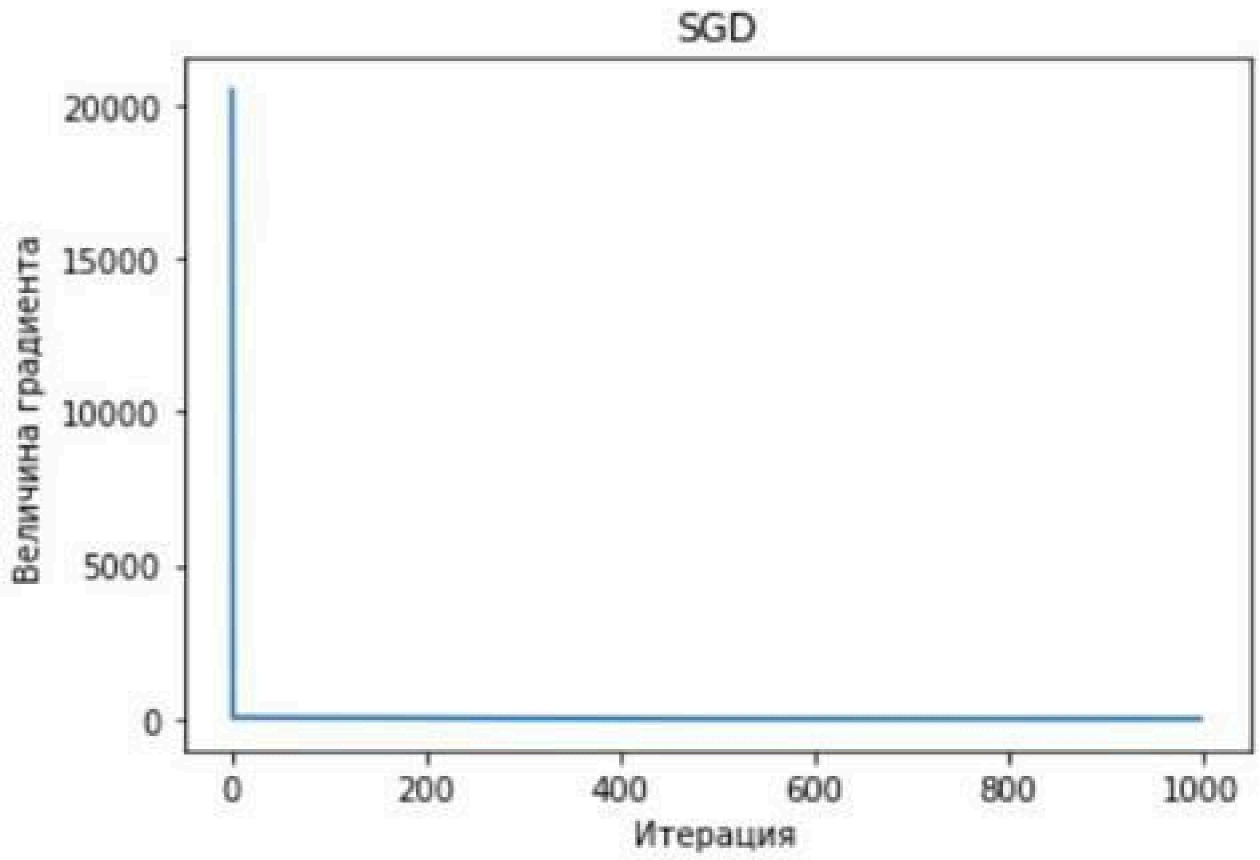
1) Натур. град. сход. быстрее обычн. град.

2) Ньютон сходится быстрее всех

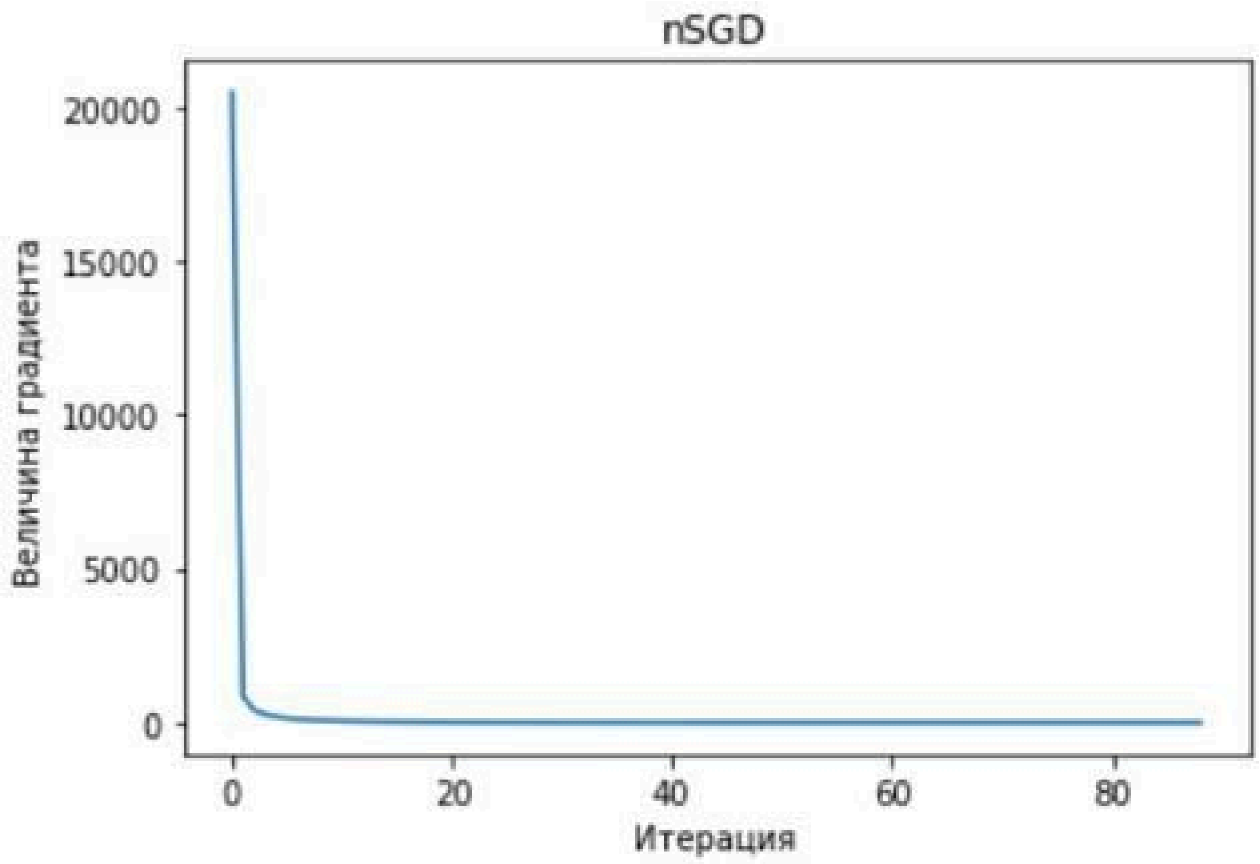
In [176]:
```python
real_theta = 10

X = sps.poisson(real_theta).rvs(size=100)
```

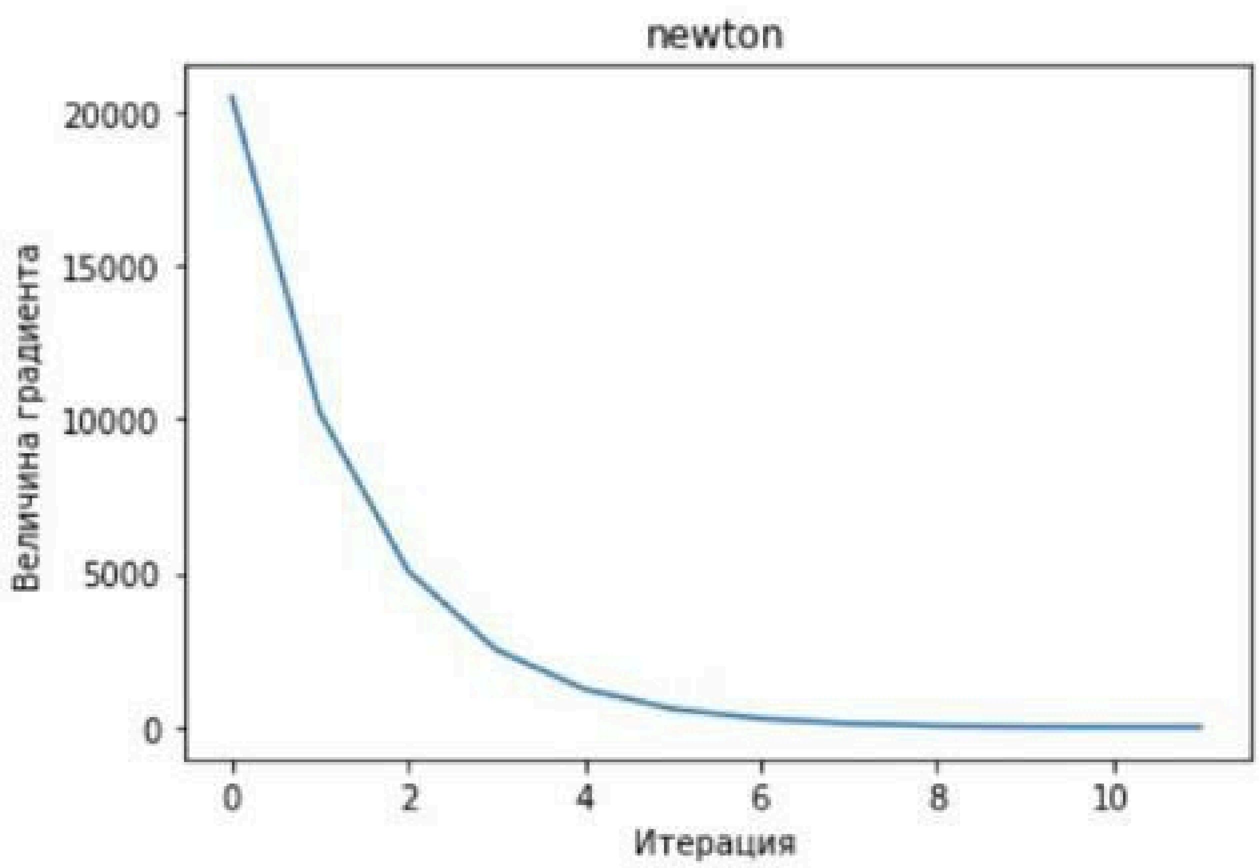
In [177]:
```python
run_method(X, SGD, 0.05, 1e-2)
```



In [178]:
```python
run_method(X, nSGD, 0.05, 1e-2)
```



In [179]:
```python
run_method(X, newton, 0.05, 1e-2)
```



(№3)

Теор. (Колм-а, Блекуэлл, Рао)

$\hat{\theta}$ - несм. оценка $\tau(\theta)$ прич. $E_\theta \hat{\theta}^2 < +\infty$

Пусть $S(x)$ - дост. статист.

Тогда $\theta^* = E_\theta(\hat{\theta} \mid S(x))$ - несм оценка $\tau(\theta)$

$D_\theta \theta^* \leq D_\theta \hat{\theta} \quad \forall \theta \in \Theta$

Задача: имея $T(X) = \sum\limits_{i=1}^{n} x_i$ - дост. статист.

на выборке $X_1 \dots X_n$ из $Bern(\theta)$

улучшить несм оценку $\hat{\theta} = X_1$.

• Найдём $E_\theta(X_1 \mid \sum\limits_{i=1}^{n} x_i) = P(X_1 = 1 \mid \sum x_i) =$

$$= \frac{C_{n-1}^{\sum x_i - 1}}{C_n^{\sum x_i}} = \frac{C_{n-1}^{k-1}}{C_n^{k}} = \frac{\frac{(n-1)!}{(k-1)!\cancel{(n-k)!}}}{\frac{n!}{k!\cancel{(n-k)!}}} = \frac{k}{n} = \frac{\sum x_i}{n}$$

улучшенная оценка.

№4 Найти оптим. оценку пар-ра θ.

Опр. Оптимальная оценка – несмещ. оценка,

кот. имеет равном. мин. дисперс.

т.е. $\underbrace{}_{\text{такая } T(X), \text{ что}}$ $\forall T_1(X)$ – др. оценка $D_\theta(T(X)) \le D_\theta(T_1(X))$ $\forall \theta$

~~Теор. (Рао-Блекуэлла-Колмог.)~~

~~Пусть $T(X)$ – дост. стат.~~

~~Если $\exists$ опт. оц. $T_1(X)$~~

Утв. Измеримая функция от полной дост. статистики явл. оптим. оценкой своего мат. ожидания.

~~Дост. статистика $\left(\sum X_i, \sum X_i^2\right)$~~

а) Дано $X_1 \dots X_n$ из $N(\theta, 1)$

$$p_\theta(x_1 \dots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_i-\theta)^2}{2}} = (\sqrt{2\pi})^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \theta)^2}$$

$$= (\sqrt{2\pi})^{-\frac{n}{2}} e^{-\frac{1}{2}\left(\sum(x_i-\bar{x})^2 - 2\sum(x_i - \bar{x})(\theta - \bar{x}) + n(\theta-\bar{x})^2\right)} =$$

$$= \underbrace{(\sqrt{2\pi})^{-\frac{n}{2}} e^{-\frac{1}{2}(\sum x_i - \bar{x})}}_{h(x)} \cdot \underbrace{e^{-\frac{n}{2}(\theta - \bar{x})}}_{\psi(\sum x_i, \theta)}$$

$\boxed{T(x) = \sum x_i}$ – дост. статистика по крит. Неймана-Фишера.

- По теор. с лекции $p_\theta(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi} e^{\frac{\theta^2}{2}}} \cdot e^{x \cdot \theta} \Rightarrow \sum x_i$ – полн. и дост. статист.
(экспон. семейство)

- $\varphi(t) = \frac{t}{n}$ – изм. функц

$$E_\theta(\varphi(T(x))) = E_\theta \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum E x_i = \theta$$

Сл-но $\frac{\sum x_i}{n} = \bar{X}$ – оптим. оценка θ

б) $X_1 \dots X_n = Pois(\theta)$

- $p_\theta(x_1 \dots x_n) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta} = \underbrace{\frac{1}{\prod_{i=1}^{n} x_i!}}_{h(x)} \underbrace{\theta^{\Sigma x_i} \cdot e^{-\theta n}}_{\psi(\Sigma x_i, n)}$

$T(x) = \underbrace{\sum_{i=1}^{n} x_i}_{\sim Pois(n\theta)}$ – дост. статист.

- Покажем полноту: Пусть $E_\theta \varphi(T(x)) = 0 \quad \forall \theta > 0$

$$E_\theta \varphi(T(x)) = \sum_{k=0}^{\infty} \varphi(k) e^{-n\theta} \frac{(n\theta)^k}{k!} = 0 \quad \forall \theta > 0.$$

$$\Leftrightarrow \sum_{k=0}^{\infty} \varphi(k) \frac{(n\theta)^k}{k!} = 0 \quad \forall \theta > 0$$

$$= \varphi(0) + \theta n \sum_{k=0}^{\infty} \varphi(k+1) \frac{(n\theta)^k}{(k+1)!} = 0 \quad \forall \theta > 0.$$

Устремив $\theta \to 0$, получим $\varphi(0) = 0$.
Далее разделим на θn оставш. часть и получим $\sum_{k=0}^{\infty} \varphi(k+1) \frac{(n\theta)^k}{(k+1)!} = 0$. Аналогично $\varphi(k) = 0 \quad \forall k \geq 0, k \in \mathbb{Z}$

т.е $\varphi(T(x)) = 0 \quad \forall \theta \Rightarrow$ полная.

- Рассм $\psi(t) = \frac{t}{n}$

$$E(\psi(T(x))) = E_\theta \frac{\Sigma x_i}{n} = \frac{1}{n} \sum_{i=1}^{n} E x_i = \theta$$

$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{X}$ – опт. оценка

№5 $X_1 \dots X_n \sim Bin(m, \theta)$, m – изв.

Опр Эффективная оценка – кот. меет $\mathcal{D}_\theta \hat{\theta} = \frac{(\tau'(\theta))^2}{I_X(\theta)}$ // из Рао-Крамера

По Теор. (Крит. эфф. оценки):

$\hat{\theta}$ – эфф. оценка $\tau(\theta) \Leftrightarrow \hat{\theta} - \tau(\theta) = c(\theta) \cdot U_X(\theta)$

причём $c(\theta) = \frac{\tau'(\theta)}{I_X(\theta)}$

$$U_X(\theta) = \frac{\partial l_X}{\partial \theta}(\theta) = \left(\sum_{i=1}^n \log\left(C_m^{x_i} \cdot \theta^{x_i} (1-\theta)^{m-x_i}\right)\right)' =$$

$$= \frac{\sum x_i}{\theta} + (nm - \sum x_i) \cdot (-1) \cdot \frac{1}{1-\theta} =$$

$$= \frac{(\sum x_i)(\theta - 1) + \theta(nm - \sum x_i)}{\theta(\theta - 1)} = \frac{\theta \cdot n \cdot m - \sum x_i}{\theta(\theta - 1)} =$$

$$= \frac{nm\left(\theta - \frac{\overline{X}}{m}\right)}{\theta(\theta - 1)} = \frac{nm\left(\frac{\overline{X}}{m} - \theta\right)}{\theta(1-\theta)}$$

$$\frac{\overline{X}}{m} - \theta = U_X(\theta) \cdot \frac{\theta(1-\theta)}{nm} \rightarrow \frac{\overline{X}}{m} \text{ – эфф. оценка}$$