Круглов А.И. Б05-202 ST 5

1 $X_1,\dots,X_n \sim N(a,\Sigma)$; найти ОМП для a и Σ. Пусть $X_i \in \mathbb{R}^d$ (d-мерная).

• Сначала решим простую одномерную задачу $(d=1)$. $L_X(\theta)=\prod\limits_{i=1}^n p_\theta(X_i)=\prod\limits_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X_i-a)^2}{2\sigma^2}} =$

$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum\limits_{i=1}^n (X_i-a)^2\right)$, $\ell_X(\theta)=\ln L_X(\theta) = -\frac n2 \ln 2\pi - \frac n2 \ln\sigma^2 - \frac{1}{2\sigma^2}\sum\limits_{i=1}^n(X_i-a)^2$

$$\begin{cases} \frac{\partial}{\partial a}\ell_X(\theta) = \frac{1}{\sigma^2}\sum\limits_{i=1}^n (X_i-a) = 0 \\ \frac{\partial}{\partial \sigma^2}\ell_X(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\limits_{i=1}^n (X_i-a)^2 = 0 \end{cases} \Rightarrow \begin{cases} \hat a = \bar X \\ \hat\sigma^2 = S^2 \end{cases}$$

(хотим получить то же) обозначение $|\Sigma| = \det\Sigma$

• Обобщим на многомерную задачу $(d>1)$. Теперь плотность $p_\theta(X_i)=\frac{1}{(2\pi)^{d/2}\sqrt{|\Sigma|}} e^{-\frac12 (X_i-a)^T\Sigma^{-1}(X_i-a)}$,

тогда $L_X(\theta)=\left((2\pi)^d|\Sigma|\right)^{-\frac n2}\cdot\exp\left(-\frac12\sum\limits_{i=1}^n\left[(X_i-a)^T\Sigma^{-1}(X_i-a)\right]\right)$, $\ell_X(\theta)=\ln L_X(\theta) = -\frac{nd}{2}\ln 2\pi -$

$- \frac n2\ln|\Sigma| - \frac12\sum\limits_{i=1}^n\left[(X_i-a)^T\Sigma^{-1}(X_i-a)\right]$. Вспомним матр. производные с теор. оптимизации:

$\frac{\partial}{\partial a}\ell_X(\theta) = -\frac12\left(\sum\limits_{i=1}^n (X_i-a)^T\Sigma^{-1}(X_i-a)\right)'_a = -\frac12\sum\limits_{i=1}^n\left[X_i^T\Sigma^{-1}X_i - a^T\Sigma^{-1}X_i - X_i^T\Sigma^{-1}a + a^T\Sigma^{-1}a\right]'_a = -\frac12\sum\limits_{i=1}^n\Big[0 +$

$\left(-a^T\Sigma^{-1}X_i - a^T(\Sigma^{-1})^T X_i + a^T\Sigma^{-1}a\right)'_a\Big] = -\frac12\sum\limits_{i=1}^n\left[-\left((\Sigma^{-1})^T+\Sigma^{-1}\right)X_i + \left((\Sigma^{-1})^T+\Sigma^{-1}\right)a\right] = 0 \Rightarrow$

$\Rightarrow (\Sigma^{-T}+\Sigma^{-1})n\bar X = (\Sigma^{-T}+\Sigma^{-1})n\hat a \Rightarrow \hat a = (\Sigma^{-T}+\Sigma^{-1})^{-1}(\Sigma^{-T}+\Sigma^{-1})\bar X = \bar X$

$\frac{\partial}{\partial\Sigma}\ell_X(\theta) = \left(-\frac n2\ln|\Sigma| - \frac12\sum\limits_{i=1}^n\left[(X_i-a)^T\Sigma^{-1}(X_i-a)\right]\right)'_\Sigma = -\frac n2\Sigma^{-T} - \frac12\sum\limits_{i=1}^n\left[(X_i-a)^T(-\Sigma^{-1}\Sigma^{-1})(X_i-a)\right]E = 0 \Rightarrow$

$\Rightarrow n\Sigma^{-T} = \sum\limits_{i=1}^n\left[(X_i-a)^T\Sigma^{-2}(X_i-a)\right]E \Rightarrow \hat\Sigma = \Sigma_{выборочная}$

Видно, что 1d случай хорошо обобщается на многомерный. Ответ: $\hat a=\bar X$, $\hat\Sigma=\Sigma_{выборочная}$

2 $X_1,\dots,X_n \sim Laplace$ $\left(p_\theta(x)=\frac12 e^{-|x-\theta|}\right)$; ОМП θ -? усл. рег. -? А.Н.О. -? А.Д. -?

$L_X(\theta)=\prod\limits_{i=1}^n\frac12 e^{-|X_i-\theta|} = 2^{-n}e^{-\sum\limits_{i=1}^n|X_i-\theta|}$, $\ell_X(\theta) = -n\ln2 - \sum\limits_{i=1}^n|X_i-\theta|$, $\ell_X(\theta)\to\max$ при $\underbrace{\sum\limits_{i=1}^n|X_i-\theta|}_{MAE}\to\min$

Докажем, что $\hat\theta$ - выборочная медиана: $\frac{\partial}{\partial\theta}\ell_X(\theta) = \sum\limits_{i=1}^n sign(X_i-\theta) = \sum\limits_{i:X_i>\theta}1 - \sum\limits_{j:X_j<\theta}1 = 0 \Rightarrow$

должно быть одинаковое кол-во наблюдений (X_i) меньших и больших θ $\Rightarrow$ $\hat\theta$ - выб. медиана.

Нарушается усл. рег. L6 ($p_\theta(x)$ не явл. трижды дифф. по θ для $\forall x$), разрыв из-за модуля.

$i(\theta) = E\left(sign(X_1-\theta)\right)^2 = 1$ (а.н.о.), А.Д. $= i^{-1}(\theta)=1$

Ответ: $\hat\theta$ - выборочная медиана, не все усл. рег. вып., а.н.о. и А.Д. $=1$

4 $X_1,\dots,X_n$ - выборка из распр. с $p_\theta(x) = \frac{(\ln\theta+1)(\ln\ln\ln x)^{\ln\theta}}{x\ln x\ln\ln x} I\{x\in[e^e, e^{e^e}]\}$; ОМП θ, сост.?

$L_X(\theta) = \prod\limits_{i=1}^n \frac{(\ln\theta+1)(\ln\ln\ln X_i)^{\ln\theta}}{X_i\ln X_i\ln\ln X_i} I\{X_i\in[e^e,e^{e^e}]\} = \frac{(\ln\theta+1)^n(\prod\ln\ln\ln X_i)^{\ln\theta}}{\prod X_i\prod\ln X_i\prod\ln\ln X_i} I\{e^e\le X_{(1)}\le\dots\le X_{(n)}\le e^{e^e}\}$, $\ell_X(\theta) =$

$= n\ln(\ln\theta+1) + \ln\theta\sum(\ln\ln\ln X_i) - \sum X_i - \sum\ln X_i - \sum\ln\ln X_i$, $\frac{\partial}{\partial\theta}\ell_X(\theta) = \frac{n}{(\ln\theta+1)\theta} + \frac{\sum(\ln\ln\ln X_i)}{\theta} = 0 \quad |\cdot\theta>0$

$\frac1n\sum(\ln\ln\ln X_i) = -(\ln\theta+1)^{-1} \Rightarrow \ln\theta = -1-\left(\overline{\ln\ln\ln X}\right)^{-1} \Rightarrow \hat\theta = \exp\left(-1-\left(\overline{\ln\ln\ln X}\right)^{-1}\right)$

$\boxed{5}$ $X_1,\dots,X_n \sim U[0,\theta]$; сравнить $\hat\theta$ в с/к подходе: $2\bar X$, $(n+1)X_{(1)}$, $\frac{n+1}{n}X_{(n)}$

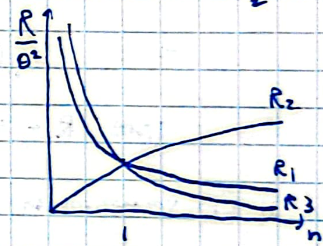
Вспомним определения: $\hat\theta_1$ не хуже ~~лучше~~ $\hat\theta_2$ в с/к подходе, если $E(\hat\theta_1-\theta)^2 \le E(\hat\theta_2-\theta)^2 \ \forall\theta\in\Theta$

- $\hat\theta_1 = 2\bar X$: $E(\hat\theta_1-\theta)^2 \overset{2\bar X \text{ — несмещ.}}{=\!=\!=} E(\hat\theta_1 - E\hat\theta_1)^2 = D\hat\theta_1 = D(2\bar X) = \frac4n DX_1 \overset{\text{дисперсия } U}{=\!=\!=} \frac4n\cdot\frac{\theta^2}{12} = \frac{\theta^2}{3n}$

- $\hat\theta_2=(n+1)X_{(1)}$: $E(\hat\theta_2-\theta)^2 = E(\hat\theta_2)^2 - 2\theta E\hat\theta_2+\theta^2 = (n+1)^2 EX_{(1)}^2 - 2\theta(n+1)EX_{(1)}+\theta^2$. Изучим величину $X_{(1)}$ (возможно было в ранних ДЗ, но на всякий случай продублирую): $p_{X_{(1)}}(t) = -(P(X_{(1)}>t))' = \frac{n(\theta-t)^{n-1}}{\theta^n}$, тогда $EX_{(1)} = \int_0^\theta tn\frac{(\theta-t)^{n-1}}{\theta^n}dt = -\frac{(\theta-t)^n(nt+\theta)}{\theta^n(n+1)}\Big|_0^\theta = \frac{\theta}{n+1}$, $EX_{(1)}^2 = \int_0^\theta t^2 n\frac{(\theta-t)^{n-1}}{\theta^n}dt = \int_0^\theta (t-\theta)^2 n \frac{t^{n-1}}{\theta^n}dt = \int_0^\theta\left(\frac{nt^{n+1}}{\theta^n} - \frac{2nt^n}{\theta^{n-1}} + \frac{nt^{n-1}}{\theta^{n-2}}\right)dt = \frac{2\theta^2}{(n+1)(n+2)}$. Получим: $E(\hat\theta_2-\theta)^2 = (n+1)^2\cdot\frac{2\theta^2}{(n+1)(n+2)} - 2\theta(n+1)\cdot\frac{\theta}{n+1}+\theta^2 = \frac{n}{n+2}\theta^2$

- $\hat\theta_3 = \frac{n+1}{n}X_{(n)}$: $E(\hat\theta_3-\theta)^2 = E(\hat\theta_3)^2 - 2\theta E\hat\theta_3+\theta^2 = \left(\frac{n+1}{n}\right)^2 EX_{(n)}^2 - 2\theta\frac{n+1}{n}EX_{(n)}+\theta^2$. Изучим $X_{(n)}$: $p_{X_{(n)}}(t) = \frac{nt^{n-1}}{\theta^n}$, тогда $EX_{(n)} = \int_0^\theta t\frac{nt^{n-1}}{\theta^n}dt = \frac{n}{n+1}\theta$, $EX_{(n)}^2 = \int_0^\theta t^2\frac{nt^{n-1}}{\theta^n}dt = \frac{n}{n+2}\theta^2$. Подставим:
$E(\hat\theta_3-\theta)^2 = \left(\frac{n+1}{n}\right)^2\frac{n}{n+2}\theta^2 - 2\theta\cancel{\frac{n+1}{n}}\cancel{\frac{n}{n+1}}\theta+\theta^2 = \theta^2\left(\frac{(n+1)^2}{(n+2)n}-1\right) = \frac{\theta^2}{n(n+2)}$

Сравним оценки: $\forall\theta\in\Theta$ (т.е. $\theta>0$) сравним $R_1=\frac{\theta^2}{3n}$, $R_2=\frac{n\theta^2}{n+2}$, $R_3 = \frac{\theta^2}{n(n+2)}$. По графику ясно, что при $\theta\in(0,1)$: $R_2<R_1<R_3$, но при $\theta>1$: $R_3<R_1<R_2$.



Значит, данные оценки нельзя сравнить в с/к подходе

Ответ: сравнить оценки в с/к подходе нельзя

$\boxed{6}$ $X_1,\dots,X_8 \sim N(\theta,1)$, $\hat\theta_1=\bar X$, $\hat\theta_2=\frac{\bar X}{2}$, $\hat\theta_3=0$, MSE-loss; сравнить в равн., байес., минимакс. подходах, если а) $\Theta=[-1,1]$ б) $\Theta=[-0.5,0.5]$ в) $\Theta=[-0.1,0.1]$

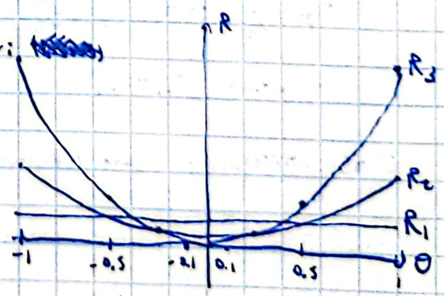
взять Q — равномерное на Θ

- равном. (с/к): $E(\hat\theta-\theta)^2 = E\hat\theta^2 - 2\theta E\hat\theta+\theta^2$; $E\bar X \overset{\text{несмещ.}}{=} \theta$, $E\bar X^2 - (E\bar X)^2 = D\bar X = \frac{1}{n^2}n\cdot 1 = \frac1n \Rightarrow E\bar X^2 = \frac1n+\theta^2$

θ_1: $E(\hat\theta_1-\theta)^2 = E(\bar X)^2 - 2\theta E\bar X+\theta^2 = \frac{n\theta^2+1}{n} - 2\theta^2+\theta^2 = \frac1n \overset{n=8}{=\!=} \frac18 = R_1$

θ_2: $E(\hat\theta_2-\theta)^2 = \frac14 E(\bar X)^2 - \theta E\bar X+\theta^2 = \frac14\left(\theta^2+\frac1n\right) - \theta^2+\theta^2 = \frac{\theta^2}{4}+\frac{1}{4n} \overset{n=8}{=\!=} \frac14\left(\theta^2+\frac18\right) = R_2$

θ_3: $E(\hat\theta_3-\theta)^2 = E(0^2) - \cancel{2\theta E(0)} + \theta^2 = \theta^2 = R_3$. Сравним по графику:



а) $\Theta=[-1,-1]$: оценки сравнить нельзя (все графики пересекаются попарно)

б) $\Theta=[-0.5,0.5]$: $R_2<R_1 \to \hat\theta_2$ лучше $\hat\theta_1$, но $\hat\theta_3$ сравнить с ними нельзя

в) $\Theta=[-0.1,0.1]$: $\forall\theta\in\Theta$ $R_3<R_2<R_1 \to \hat\theta_3$ лучше $\hat\theta_2$ лучше $\hat\theta_1$

- байес: $\hat\theta_1$ не хуже $\hat\theta_2 \overset{def}{\Leftrightarrow} E_Q R_1(\theta) \le E_Q R_2(\theta)$, где Q возьмем $U[-\lambda,\lambda]$, $\lambda=0.1, 0.5, 1$

θ_1: $\int_{-\lambda}^{\lambda} R_1 d\theta = \frac18\cdot 2\lambda = \frac\lambda4$, θ_2: $\int_{-\lambda}^{\lambda} R_2 d\theta = \int_{-\lambda}^{\lambda}\left(\frac14\theta^2+\frac1{32}\right)d\theta = \frac{\lambda^3}{6}+\frac{\lambda}{16}$, θ_3: $\int_{-\lambda}^{\lambda} R_3 d\theta = \int_{-\lambda}^{\lambda} \theta^2 d\theta = \frac23\lambda^3$

а) $\lambda=1$: $\frac{\lambda^3}{6}+\frac{\lambda}{16} < \frac\lambda4 < \frac23\lambda^3 \to \hat\theta_2$ лучше $\hat\theta_1$ лучше $\hat\theta_3$

б) $\lambda=0.5$: $\frac{\lambda^3}{6}+\frac\lambda{16} < \frac23\lambda^3 < \frac\lambda4 \to \hat\theta_2$ лучше $\hat\theta_3$ лучше $\hat\theta_1$

в) $\lambda=0.1$: $\frac23\lambda^3 < \frac{\lambda^3}{6}+\frac{\lambda}{16} < \frac\lambda4 \to \hat\theta_3$ лучше $\hat\theta_2$ лучше $\hat\theta_1$

продолжение →

6 (продолжение)

• минимакс: $\hat\theta_1$ не хуже $\hat\theta_2$ $\overset{def}{\Leftrightarrow}$ $\sup_{\Theta} R_1 \leq \sup_{\Theta} R_2$. Удобно, что R_1 - прямая (не зависит от θ), а R_2 и R_3 имеют ~~…~~ max на краях интервала $(\theta=\lambda)$ для $\forall\,\Theta$.

$\sup_{\Theta} R_1 = R_1 = \frac{1}{8}$, $\sup_{\Theta} R_2 = \sup_{[-\lambda,\lambda]} R_2 = \frac{\lambda^2}{4} + \frac{1}{32}$, $\sup_{\Theta} R_3 = \sup_{[-\lambda,\lambda]} R_3 = \lambda^2$

а) $\lambda=1$: $\sup R_1 < \sup R_2 < \sup R_3$ $\rightarrow$ $\hat\theta_1$ лучше $\hat\theta_2$ лучше $\hat\theta_3$

б) $\lambda=0{,}5$: —//— $\rightarrow$ $\hat\theta_1$ лучше $\hat\theta_2$ лучше $\hat\theta_3$

в) $\lambda=0{,}1$: $\sup R_3 < \sup R_2 < \sup R_1$ $\rightarrow$ $\hat\theta_3$ лучше $\hat\theta_2$ лучше $\hat\theta_1$

7 формулировки

а) оценка – статистика, которую применяют для оценивания параметра распределения

ф-ция потерь: $L: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$, показывающая отклонение оценки от параметра. Примеры: MSE, MAE

сравнение оценок (с/к): $\hat\theta_1$ не хуже $\hat\theta_2$ $\overset{def}{\Leftrightarrow}$ $R_1 \leq R_2$ $\forall\theta\in\Theta$

—//— (байес): $\hat\theta_1$ не хуже $\hat\theta_2$ $\overset{def}{\Leftrightarrow}$ $E_Q R_1 \leq E_Q R_2$, Q - распределение на Θ

—//— (минимакс): $\hat\theta_1$ не хуже $\hat\theta_2$ $\overset{def}{\Leftrightarrow}$ $\sup_{\Theta} R_1 \leq \sup_{\Theta} R_2$

(где R - ф-ция риска, $R = E_\theta L$)

б) $\hat\theta$ - несмещенная оценка $\Leftrightarrow$ $E\hat\theta = \theta$

линейность матожидания: X_i - норм, $\alpha_i \in \mathbb{R}$ $\Rightarrow$ $E(\sum \alpha_i X_i) = \sum \alpha_i EX_i$

8 $X_1,\dots,X_n \sim N(a,\sigma^2)$, a,σ^2 - неизвестны, $K = \{\frac{1}{c_n}\sum_{i=1}^n (X_i-\overline{X})^2, c_n\in\mathbb{R}\}$ - для σ^2; в K найти все несмещ. оценки θ, ОМП, bias-variance, лучшая оценка в с/к подходе

Помним, что несмещенная оценка σ^2 - это $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\overline{X})^2$, откуда $c_n = n-1$.

ОМП: $L(\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i-a)^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-a)^2\right)$, $l(\theta) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-a)^2$, $\frac{\partial}{\partial(\sigma^2)} l(\theta) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (X_i-\overline{X})^2 = 0$ $\Rightarrow$ $\sum_{i=1}^n (X_i-\overline{X})^2 = n\sigma^2$, $\hat\sigma^2 = \frac{1}{n}\sum_{i=1}^n (X_i-\overline{X})^2$, откуда $c_n = n$

bias-variance: для несмещенной, понятно, bias = 0, вычисляли когда-то variance $= \frac{2\sigma^4}{n-1}$, то есть MSE $= \frac{2\sigma^4}{n-1}$; для ОМП bias $= -\frac{\sigma^2}{n}$, variance $= \frac{2\sigma^4}{n}$, т.е. MSE $= \left(\frac{\sigma^2}{n}\right)^2 + \frac{2\sigma^4}{n} = \sigma^4\left(\frac{1}{n^2}+\frac{2}{n}\right)$

С/К: сравним $MSE_{несм.} = \frac{2\sigma^4}{n-1}$ и $MSE_{омп} = \sigma^4\left(\frac{1}{n^2}+\frac{2}{n}\right)$. По графику было видно, что $MSE_{омп} < MSE_{несм.}$ $\forall\theta$ (т.е. $\forall\sigma$) $\Rightarrow$ $MSE_{омп}$ лучше в с/к подходе

$\boxed{9}$ $X_1, \dots, X_n \sim Bern(\theta)$, $\tilde{\theta}$ - оценка Ходжеса-Лемана; $MSE_{\tilde{\theta}}(\theta)$ - ?

$$MSE = E(\tilde{\theta} - \theta)^2 = E\tilde{\theta}^2 - 2\theta E\tilde{\theta} + \theta^2$$

$$E\tilde{\theta} = E\left(\bar{X} + \frac{1}{1+\sqrt{n}}\left(\frac{1}{2} - \bar{X}\right)\right) = E\bar{X} + \frac{1}{2(1+\sqrt{n})} - \frac{1}{1+\sqrt{n}} E\bar{X} = \frac{1}{2(1+\sqrt{n})} + \frac{\sqrt{n}}{1+\sqrt{n}} E\bar{X} = \frac{1}{2(1+\sqrt{n})} + \frac{\sqrt{n}}{1+\sqrt{n}}\theta$$

$$E\tilde{\theta}^2 = E\left(\bar{X} + \frac{1}{1+\sqrt{n}}\left(\frac{1}{2} - \bar{X}\right)\right) = E\left(\bar{X}^2 + 2\bar{X}\frac{1}{1+\sqrt{n}}\left(\frac{1}{2} - \bar{X}\right) + \frac{1}{(1+\sqrt{n})^2}\left(\frac{1}{2} - \bar{X}\right)^2\right) = E\bar{X}^2 + \frac{1}{1+\sqrt{n}} E\bar{X} - \frac{2}{1+\sqrt{n}} E\bar{X}^2 + \frac{1}{(1+\sqrt{n})^2} E\left(\frac{1}{2} - \bar{X}\right)^2 =$$

$$= \frac{\sqrt{n}-1}{\sqrt{n}+1} E\bar{X}^2 + \frac{1}{1+\sqrt{n}} E\bar{X} + \frac{1}{4(1+\sqrt{n})^2} - \frac{1}{(1+\sqrt{n})^2} E\bar{X} + \frac{1}{(1+\sqrt{n})^2} E\bar{X}^2 = \frac{n}{(1+\sqrt{n})^2} E\bar{X}^2 + \frac{\sqrt{n}}{(1+\sqrt{n})^2} E\bar{X} + \frac{1}{4(1+\sqrt{n})^2} =$$

$$= \frac{1}{(1+\sqrt{n})^2}\left(n\left(D\bar{X} + (E\bar{X})^2\right) + \sqrt{n}\,\theta + \frac{1}{4}\right) = \frac{1}{(1+\sqrt{n})^2}\left(n\left(\frac{\theta(1-\theta)}{n} + \theta^2\right) + \sqrt{n}\,\theta + \frac{1}{4}\right) = \frac{1}{(1+\sqrt{n})^2}\left(\theta(1-\theta) + n\theta^2 + \sqrt{n}\,\theta + \frac{1}{4}\right)$$

$$MSE = \frac{1}{(1+\sqrt{n})^2}\left(\theta(1-\theta) + n\theta^2 + \sqrt{n}\,\theta + \frac{1}{4}\right) - 2\theta\frac{1}{1+\sqrt{n}}\left(\frac{1}{2} + \sqrt{n}\,\theta\right) + \theta^2 = \theta^2 \cdot (0) + \theta \cdot (0) + \frac{1}{4(1+\sqrt{n})^2} =$$

$$= \frac{1}{4(1+\sqrt{n})^2}$$

Ответ: $\frac{1}{4(1+\sqrt{n})^2}$