Примеры возможных задач к семинару по лекции 3 (с решениями)

1. Пусть выборка $Z_n = (X_1, \ldots, X_n)$ порождена распределением $N(m; \theta^2)$ — нормальным распределением с известным математическим ожиданием m и неизвестной дисперсией θ^2. Найти значение C, при котором оценка

$$\hat{\theta}_n = \frac{C}{n} \sum_{k=1}^{n} |X_k - m|$$

параметра θ будет несмещённой и состоятельной.

Решение.
Сначала найдём математическое ожидание оценки $\hat{\theta}_n$:

$$\mathrm{M}\left\{\hat{\theta}_n\right\} = \mathrm{M}\left\{\frac{C}{n} \sum_{k=1}^{n} |X_k - m|\right\} = \frac{C}{n} \sum_{k=1}^{n} \mathrm{M}\left\{|X_k - m|\right\}.$$

Так как $X_k \sim N(m; \theta^2)$, то $X_k - m \sim N(0; \theta^2)$. Обозначим $Y = \frac{X_k - m}{\theta} \sim N(0; 1)$. Тогда

$$\mathrm{M}\left\{|X_k - m|\right\} = \mathrm{M}\left\{|\theta Y|\right\} = \theta \mathrm{M}\left\{|Y|\right\}.$$

Для стандартного нормального распределения известно, что $\mathrm{M}\left\{|Y|\right\} = \sqrt{\frac{2}{\pi}}$. Следовательно,

$$\mathrm{M}\left\{|X_k - m|\right\} = \theta \sqrt{\frac{2}{\pi}}.$$

Тогда

$$\mathrm{M}\left\{\hat{\theta}_n\right\} = \frac{C}{n} \cdot n \cdot \theta \sqrt{\frac{2}{\pi}} = C\theta \sqrt{\frac{2}{\pi}}.$$

Для несмещённости необходимо, чтобы $\mathrm{M}\left\{\hat{\theta}_n\right\} = \theta$, откуда

$$C\theta \sqrt{\frac{2}{\pi}} = \theta \quad \Rightarrow \quad C = \sqrt{\frac{\pi}{2}}.$$

Таким образом, при $C = \sqrt{\frac{\pi}{2}}$ оценка является несмещённой.

Теперь проверим состоятельность. По закону больших чисел для независимых одинаково распределённых случайных величин:

$$\frac{1}{n} \sum_{k=1}^{n} |X_k - m| \xrightarrow[n \to \infty]{P} \mathrm{M}\left\{|X_1 - m|\right\} = \theta \sqrt{\frac{2}{\pi}}.$$

Тогда при $C = \sqrt{\frac{\pi}{2}}$:

$$\hat{\theta}_n = \sqrt{\frac{\pi}{2}} \cdot \frac{1}{n} \sum_{k=1}^{n} |X_k - m| \xrightarrow[n \to \infty]{P} \sqrt{\frac{\pi}{2}} \cdot \theta \sqrt{\frac{2}{\pi}} = \theta.$$

Следовательно, оценка является состоятельной.

Ответ: $C = \sqrt{\frac{\pi}{2}}$.

2. Пусть выборка $Z_n = (X_1, \dots, X_n)$ порождена СВ X, имеющей биномиальное распределение $\mathrm{Bi}(N; p)$, где p — неизвестно. Пусть

$$\hat{\theta}_n = \frac{\bar{X}_n (N - \bar{X}_n)}{N},$$

где $\bar{X}_n = \dfrac{X_1 + \dots + X_n}{n}$ — выборочное среднее. Показать, что оценка $\hat{\theta}_n$ является асимптотически несмещённой оценкой для параметра $\theta = \mathrm{D}\left\{X\right\}$.

Решение.
Для биномиального распределения $\mathrm{Bi}(N; p)$ известно, что $\mathrm{M}\left\{X\right\} = Np$, $\mathrm{D}\left\{X\right\} = Np(1-p)$. Параметр $\theta = \mathrm{D}\left\{X\right\} = Np(1-p)$.

Рассмотрим математическое ожидание оценки $\hat{\theta}_n$:

$$\mathrm{M}\left\{\hat{\theta}_n\right\} = \mathrm{M}\left\{\frac{\bar{X}_n (N - \bar{X}_n)}{N}\right\} = \frac{1}{N}\left(N\mathrm{M}\left\{\bar{X}_n\right\} - \mathrm{M}\left\{\bar{X}_n^2\right\}\right).$$

Так как $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, то $\mathrm{M}\left\{\bar{X}_n\right\} = \frac{1}{n}\sum_{i=1}^n \mathrm{M}\left\{X_i\right\} = Np$. Для $\mathrm{M}\left\{\bar{X}_n^2\right\}$ воспользуемся формулой: $\mathrm{M}\left\{\bar{X}_n^2\right\} = \mathrm{D}\left\{\bar{X}_n\right\} + \left(\mathrm{M}\left\{\bar{X}_n\right\}\right)^2$. Поскольку X_i независимы и одинаково распределены, то

$$\mathrm{D}\left\{\bar{X}_n\right\} = \frac{1}{n^2}\sum_{i=1}^n \mathrm{D}\left\{X_i\right\} = \frac{1}{n^2}\cdot n \cdot Np(1-p) = \frac{Np(1-p)}{n}.$$

Следовательно,

$$\mathrm{M}\left\{\bar{X}_n^2\right\} = \frac{Np(1-p)}{n} + (Np)^2.$$

Подставим в выражение для $\mathrm{M}\left\{\hat{\theta}_n\right\}$:

$$\begin{aligned}
\mathrm{M}\left\{\hat{\theta}_n\right\} &= \frac{1}{N}\left(N\cdot Np - \left(\frac{Np(1-p)}{n} + (Np)^2\right)\right)\\
&= \frac{1}{N}\left(N^2 p - \frac{Np(1-p)}{n} - N^2p^2\right)\\
&= Np - \frac{p(1-p)}{n} - Np^2\\
&= Np(1-p) - \frac{p(1-p)}{n}\\
&= \theta - \frac{p(1-p)}{n}.
\end{aligned}$$

Отсюда видно, что $\mathrm{M}\left\{\hat{\theta}_n\right\} = \theta - \frac{p(1-p)}{n}$. Смещение оценки равно $-\frac{p(1-p)}{n}$. При $n \to \infty$ смещение стремится к нулю, поэтому оценка является асимптотически несмещённой.

3. Показать, что оценка $\hat{\theta}$ из задачи 2 будет также состоятельной оценкой для θ.

Решение.
Для доказательства состоятельности покажем, что $\hat{\theta}_n \xrightarrow[n\to\infty]{P} \theta$. Имеем:

$$\hat{\theta}_n = \frac{\bar{X}_n (N - \bar{X}_n)}{N}.$$

По закону больших чисел, $\bar{X}_n \xrightarrow[n\to\infty]{P} \mathrm{M}\left\{X\right\} = Np$. Так как функция $g(x) = \frac{x(N-x)}{N}$ непрерывна, то по теореме о наследовании сходимости:

$$\hat{\theta}_n = g(\bar{X}_n) \xrightarrow[n\to\infty]{P} g(Np) = \frac{Np(N-Np)}{N} = Np(1-p) = \theta.$$

Таким образом, оценка является состоятельной.

4. Пусть выборка $Z_n = (X_1, \ldots, X_n)$ порождена СВ X, имеющей равномерное распределение на отрезке $[0;\theta]$, где $\theta > 0$ — неизвестный параметр. Показать, что оценка $\hat{\theta}_n = X_{(n)}$ будет состоятельной оценкой для θ. ($X_{(n)} = max(X_1, \ldots, X_n)$ - n-я порядковая статистика).

   Решение.
   Найдём функцию распределения $X_{(n)}$:

$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P(X_1 \le x, \ldots, X_n \le x) = \prod_{i=1}^{n} P(X_i \le x) = \left(\frac{x}{\theta}\right)^n, \quad 0 \le x \le \theta.$$

   Плотность распределения:

$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n}, \quad 0 \le x \le \theta.$$

   Вычислим математическое ожидание и дисперсию $X_{(n)}$:

$$\mathrm{M}\left\{X_{(n)}\right\} = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n} \int_0^\theta x^n\, dx = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta.$$

$$\mathrm{M}\left\{X_{(n)}^2\right\} = \int_0^\theta x^2 \cdot \frac{nx^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n} \int_0^\theta x^{n+1}\, dx = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2.$$

   Тогда дисперсия:

$$\begin{aligned}
\mathrm{D}\left\{X_{(n)}\right\} &= \mathrm{M}\left\{X_{(n)}^2\right\} - \left(\mathrm{M}\left\{X_{(n)}\right\}\right)^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \theta^2\left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right) \\
&= \theta^2 \cdot \frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2} = \theta^2 \cdot \frac{n(n^2+2n+1) - n^3 - 2n^2}{(n+2)(n+1)^2} \\
&= \theta^2 \cdot \frac{n^3 + 2n^2 + n - n^3 - 2n^2}{(n+2)(n+1)^2} = \theta^2 \cdot \frac{n}{(n+2)(n+1)^2}.
\end{aligned}$$

   Таким образом,

$$\mathrm{M}\left\{\hat{\theta}_n\right\} = \frac{n}{n+1}\theta, \quad \mathrm{D}\left\{\hat{\theta}_n\right\} = \theta^2 \cdot \frac{n}{(n+2)(n+1)^2}.$$

   Оценка смещённая, но асимптотически несмещённая, так как $\mathrm{M}\left\{\hat{\theta}_n\right\} \to \theta$ при $n \to \infty$. Для состоятельности достаточно показать, что $\hat{\theta}_n \xrightarrow[n\to\infty]{P} \theta$. Воспользуемся критерием состоятельности: если $\mathrm{M}\left\{\hat{\theta}_n\right\} \to \theta$ и $\mathrm{D}\left\{\hat{\theta}_n\right\} \to 0$ при $n \to \infty$, то оценка состоятельна. Мы уже видели, что $\mathrm{M}\left\{\hat{\theta}_n\right\} \to \theta$. Теперь проверим дисперсию:

$$\mathrm{D}\left\{\hat{\theta}_n\right\} = \theta^2 \cdot \frac{n}{(n+2)(n+1)^2} \sim \theta^2 \cdot \frac{n}{n^3} = \frac{\theta^2}{n^2} \to 0 \quad \text{при } n \to \infty.$$

   Следовательно, оценка $\hat{\theta}_n = X_{(n)}$ является состоятельной оценкой параметра θ.