5) выборочное среднее:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{15}(9.7 + \dots + 16.2) \approx 13.48$$

выборочная дисперсия:

$$\hat{\sigma}^2(\vec{x}) = \cancel{\frac{1}{n}\sum_{i=1}^{n}} \dots$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \approx 2.27$$

исправленная выборочная дисперсия:

$$s^2(\vec{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{n}{n-1}\cdot\hat{\sigma}^2(\vec{x}) \approx 2.43$$

Замечание: выборка из пред. примера смоделирована из $N(13, 2.5)$

$$\begin{cases} \bar{x} = 13.48 \\ s^2(\vec{x}) = 2.43 \end{cases}$$

07.03.24. Семинар. 5. (замена)

① $\vec{X} = (X_1, \dots, X_n)$ ген. выборка

Док-ть, что лучшая оценка величины $a = M[X]$ вида

$$\hat{a}(X) = \alpha_1 X_1 + \dots + \alpha_n X_n$$

лучшей является оценка $\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$

Док-во:

- $M[\hat{a}(X)] = M[\alpha_1 X_1 + \dots + \alpha_n X_n] = \sum_{k=1}^{n}\alpha_k M[X_k] = \left\{\begin{matrix} M[X_i] = M[X_j] \\ = M[X] = a \end{matrix}\right\}$

$= M[X]\cdot\sum_{k=1}^{n}\alpha_k = \cancel{M[X]}\sum_{k=1}^{n}\alpha_k = a\sum_{k=1}^{n}\alpha_k \Rightarrow \sum_{k=1}^{n}\alpha_k = 1$

- $D[\hat{a}(X)] = D[\alpha_1 X_1 + \dots + \alpha_n X_n] = \left\{\begin{matrix} X_i \text{ - незав,} \\ \text{расп. по} \\ \text{одному з.} \end{matrix}\right\} = \sigma^2\sum_{k=1}^{n}\alpha_k^2$

$\sigma^2\sum_{k=1}^{n}\alpha_k^2 \xrightarrow[\sum_{k=1}^{n}\alpha_k = 1]{} \min$ (т.е. лучшая)

Необх. условие условного экстремума:

$$L = \sum_{k=1}^{n}\alpha_k^2 + \lambda\left(\sum_{k=1}^{n}\alpha_k - 1\right)$$

$$L'_{\alpha_k} = 2\alpha_k + \lambda = 0^{(1)},\ k = \overline{1,n}\ ;\ L'_\lambda = 0 \Rightarrow \sum_{k=1}^{n}\alpha_k = 1^{(2)}$$

~~Подставим $\alpha_k = -\frac{\lambda}{2}$ в (2);~~

Из (1); $\lambda = -2\alpha_k$, $k=\overline{1;n} \Rightarrow \alpha_k = -\frac{\lambda}{2}$; $k=\overline{1;n}$

Подставим $\alpha_k = -\frac{\lambda}{2}$ в (2);

$$-\sum_{k=1}^{n} \frac{\lambda}{2} = 1 \Rightarrow -\frac{n\lambda}{2} = 1 \Rightarrow \lambda = -\frac{2}{n}$$

$\frac{\partial^2 L}{\partial \alpha_k^2} = 2$ — положительно определенная кв. матрица $\Rightarrow$ min.

$$\alpha_k = -\frac{\left(-\frac{2}{n}\right)}{2} = \frac{1}{n}$$

$\Rightarrow \overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ — лучшая, ч.т.д.

Несмещенность,
Несмещенная статистика

$X \sim f_X(\beta, x)$

$b(\vec{X})$ — статистика (любая ф-я от выборки)

$\vec{X} = (X_1, \dots, X_n)$

Опр: $M[b(\vec{X})] = \beta$ — несмещенная оценка β. (статистика)

Эффективность статистики

Опр: $b(\vec{X})$ — эфф. несмещенная статистика (оценка) в данном классе, если в этом классе у нее min дисперсия.

$b(\vec{X}) \xrightarrow{P} \beta$ — состоятельность.

$\boxed{N}$

Дано: $\hat{\sigma}^2(\vec{X}) = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X})^2$

Найти: $M[\hat{\sigma}^2(\vec{X})]$ — ?

Решение:

$$M[\hat{\sigma}^2(\vec{X})] = \frac{1}{n} M\left[\sum_{k=1}^{n}(X_k - \overline{X})^2\right] = \left\{ m \stackrel{\circ}{=} MX = M\overline{X} \right\} = \frac{1}{n} M\left[\sum_{k=1}^{n}((X_k - m) - (\overline{X} - m))^2\right] =$$

$$= \frac{1}{n} M\left[\sum_{k=1}^{n}\left((X_k - m)^2 - 2(\overline{X} - m)(X_k - m) + (\overline{X} - m)^2\right)\right] =$$

$$= \frac{1}{n} M\left[\sum_{k=1}^{n}(X_k - m)^2 - 2(\overline{X} - m)\sum_{k=1}^{n}(X_k - m) + n(\overline{X} - m)^2\right] =$$

$$= \frac{1}{n} M\left[\sum_{k=1}^{n}(X_k - m)^2 - 2n(\overline{X} - m)^2 + n(\overline{X} - m)^2\right] =$$

$$= \frac{1}{n}\left(\sum_{k=1}^{n} M[(X_k - m)^2] - n \cdot M[(\overline{X} - m)^2]\right) = \frac{1}{n}\left(\sum_{k=1}^{n} DX_k - n \cdot D\overline{X}\right) = \frac{1}{n}(n\sigma^2 - nD\overline{X}) =$$

$= \frac{1}{n}\left(n\sigma^2 - \frac{n}{n^2}\cdot D\left[\sum_{k=1}^{n} X_k\right]\right) = \left\{ \begin{array}{l} X_k \text{ - незав.} \\ k=\overline{1,n};\ X_k \text{ - распр. один.} \end{array} \right\} =$

$= \frac{1}{n}\left(n\sigma^2 - \frac{1}{n}\cdot n\sigma^2\right) = \frac{n-1}{n}\sigma^2$ — смещенная статистика, чтобы получ. несмещ., необходимо делить на $(n-1)$

<u>Ответ</u>: $M[\hat{\sigma}^2(\vec{X})] = \frac{n-1}{n}\sigma^2$.

$S^2(\vec{X}) = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X})^2$ — исправленная статистика дисперсии?

Эффективна ли в задаче? 1. Исправить. 2. Д-ть миним-ть.

! На выборочное распределение мы "смотрим" как на дискретное распределение, кот. соотв. опр. реализации.

Точечные оценки.

Э-ет два способа получения точ. оценки:
1) метод максимального правдоподобия,
2)

1) Метод максимального подобия.

Пусть
1) $\vec{X} = (X_1, \dots, X_n)$ — случ. выборка (непр. с.в.)
2) $f_X(\beta, x)$ — пл-ть распр. X_i, $i=\overline{1,n}$
3) $\vec{x}(x_1, \dots, x_n)$ — реализация (выборка)

<u>Опр.</u> Функцией правдоподобия наз. совместную пл-ть $\vec{X}$ в точке $\vec{x}$:

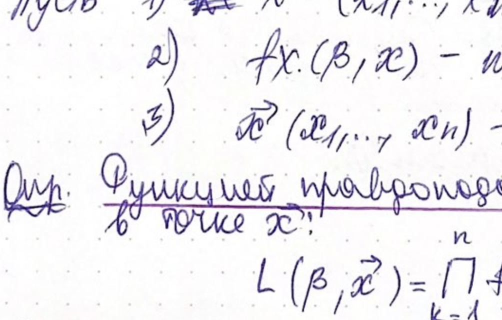
$$L(\beta, \vec{x}) = \prod_{k=1}^{n} f_X(\beta, x_k)$$

! Для дискретной с.в.:

$P_\beta\{X=x\}$ — распределение, зависящее от неизв. пар-ра β

$$L(\beta, \vec{x}) = \prod_{k=1}^{n} P_\beta\{X = x_k\}$$

Пример.



Принцип правдоподобия закл. в том, что проверяет оценка вер-тей попарно точек

Маловероятно, что в случае для картинки $x_1, \dots, x_n$ явл. реализацией сплошной.

Более вероятно, что они реализуют пунктир.

Т.е. ищем $\max L(\beta, \vec{x})$ по β.

$L(\beta, \vec{x}) \to max$ - принцип.

$\boxed{N1}$

Найти МП-оценку (макс.-правдопод. оценку) для $X \sim N(a, \sigma^2)$ для обоих параметров: $\hat{a}_{ОМП}$, $\hat{\sigma}^2_{ОМП}$. Реализация: $\vec{x} = (x_1, ..., x_n)$ - дано.

Решение:

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \frac{(x-a)^2}{\sigma^2}}, \quad x \in R$$

$$\cancel{L(\hat{a}_{ОМП}, \vec{x}) = f_X(\beta, x_1) \cdot ... \cdot f_X(\beta, x_n) =}$$

$$\cancel{= \frac{1}{(\sqrt{2\pi})^n \cdot \sigma^n} \cdot e^{-\sum_{k=1}^{n} \frac{(x-a)^2}{2\sigma^2}}}$$

! Необх. избавиться от экспоненты:

$y'(x) = 0$

Нужно найти $x : y'(x) = 0$ вне везде так можно...

$g(x) = \ln y(x)$

$g'(x) = \frac{1}{y(x)} \cdot y'(x)$

$P\{x \mid g'(x) = 0\} = P\{x \mid y'(x) = 0\}$

$$L(a, \sigma^2, \vec{x}) = \prod_{k=1}^{n} f_X(a, \sigma^2, x_k)$$

$$L(a, \sigma^2, \vec{x}) = f_X(a, \sigma^2, x_1) \cdot ... \cdot f_X(a, \sigma^2, x_n) =$$

$$= \frac{1}{(\sqrt{2\pi})^n \cdot \sigma^n} \cdot e^{-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (x_k - a)^2};$$

$$\ln[L(a, \sigma^2, \vec{x})] = C - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{k=1}^{n} (x_k - a)^2;$$

1) $\hat{a}_{ОМП}$.

$$\left. \frac{\partial \ln[L(a, \sigma^2, \vec{x})]}{\partial a} \right|_{a = \hat{a}_{ОМП}} = 0$$

$$\left. + \frac{\cancel{2}}{\cancel{2\sigma^2}} \sum_{k=1}^{n} (x_k - a) \right|_{a = \hat{a}_{ОМП}} = 0$$

$$\sum_{k=1}^{n} (x_k - \hat{a}_{ОМП}) = 0 \Rightarrow \hat{a}_{ОМП} = \frac{1}{n} \sum_{k=1}^{n} x_k$$ - совпадает с выборочной пск.

☆

2) $\hat{\sigma}^2_{омп}$:

$$\left.\frac{\partial \ln[L(a,\sigma^2;\vec{x})]}{\partial \sigma}\right|_{\sigma=\hat{\sigma}_{омп}} = 0$$

$$\left.-\frac{n}{\sigma} - \frac{1}{2}(-2)\frac{1}{\sigma^3}\cdot\sum_{k=1}^{n}(x_k-a)^2\right|_{\sigma=\hat{\sigma}_{омп}} = 0$$

$$n = \frac{1}{\hat{\sigma}^2_{омп}}\cdot\sum_{k=1}^{n}(x_k-a)^2$$

$$\hat{\sigma}^2_{омп} = \frac{1}{n}\sum_{k=1}^{n}(x_k-a)^2$$ – смещенная