МС-11-12 Теоретический материал

Доверительная оценка параметров

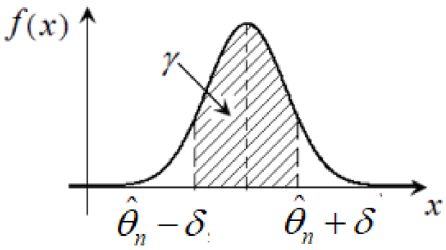
Пусть $\overrightarrow{X} = (X_1, \ldots, X_n)$ - выборка объёма n из распределения, зависящего от некоторого числового параметра $\theta \in \Theta$.

Случайный промежуток $I(\overrightarrow{X})$ называется γ-**доверительным интервалом** для параметра θ, если $\forall \theta \in \Theta$

$$P\{\theta \in I(\overrightarrow{X})\} = \gamma.$$

γ - **доверительная вероятность**.

X_1, …, X_n – независимы и из $N(m;\sigma^2)$.



Двусторонние доверительные интервалы:

a) для генерального среднего при известной дисперсии σ^2: $(\overline{X} - \delta; \overline{X} + \delta)$,

где $\delta = \dfrac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}}$, $z_{(1+\gamma)/2}$ – квантиль $N(0; 1)$ уровня $(1 + \gamma)/2$;

$\boldsymbol{\sigma}^2 = Var(X)$.

b) для генерального среднего при неизвестной дисперсии σ^2: $(\bar{X} - \delta, \bar{X} + \delta)$,

где $\delta = \dfrac{t_{n-1;(1+\gamma)/2} \cdot s}{\sqrt{n}}$, $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$,

$t_{n-1;(1+\gamma)/2}$ – квантиль распределения Стьюдента с $(n - 1)$ степенью свободы уровня $(1 + \gamma)/2$;

c) оценка дисперсии при известном генеральном среднем: $\dfrac{n \cdot s_0^2}{\chi^2_{n;(1+\gamma)/2}} < \sigma^2 < \dfrac{n \cdot s_0^2}{\chi^2_{n;(1-\gamma)/2}}$,

где $s_0^2 = \dfrac{\sum_{i=1}^{n}(X_i - \mu)^2}{n}$, $\chi^2_{k;p}$ – квантиль хи-квадрат распределения с k степенями свободы уровня p;

$\mu = E(X)$.

d) оценка дисперсии при неизвестном генеральном среднем: $\dfrac{(n-1)\cdot s^2}{\chi^2_{n-1;(1+\gamma)/2}} < \sigma^2 < \dfrac{(n-1)\cdot s^2}{\chi^2_{n-1;(1-\gamma)/2}}$,

где $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, $\chi^2_{k;p}$ – квантиль хи-квадрат распределения с k степенями свободы уровня p;

e) оценка генеральной доли признака: $(\hat{p}-\delta; \hat{p}+\delta)$,

где $\delta = z_{\frac{1+\gamma}{2}} \cdot \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$, $z_{(1+\gamma)/2}$ – квантиль $N(0;1)$ уровня $(1+\gamma)/2$;

f) интервал предсказания: $\overline{X} - \delta < X_{n+1} < \overline{X} + \delta$,

где $\delta = t_{n-1;\frac{1+\gamma}{2}} \cdot s \cdot \sqrt{1+\dfrac{1}{n}}$,

$t_{n-1;(1+\gamma)/2}$ – квантиль распределения Стьюдента с $(n-1)$ степенью свободы уровня $(1+\gamma)/2$;

g) интервальная оценка коэффициента корреляции:

$$th\left(arth\,\hat{\boldsymbol{\rho}}_n - \frac{1}{\sqrt{n-3}} z_{\frac{1+\gamma}{2}}\right) < \rho < th\left(arth\,\hat{\boldsymbol{\rho}}_n + \frac{1}{\sqrt{n-3}} z_{\frac{1+\gamma}{2}}\right),$$

где функция $th\,x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (гиперболический тангенс), $arth\,x$ – обратный гиперболический арктангенс.

Односторонние доверительные интервалы.

$\left(-\infty; \overline{X} + z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}\right)$ - **правосторонний доверительный интервал для параметра $m = E(X)$ с доверительной вероятностью γ при известной дисперсии;**

$\left(\overline{X} - z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}; +\infty\right)$ - **левосторонний доверительный интервал для параметра $m = E(X)$ с доверительной вероятностью γ при известной дисперсии;**

$\left(\overline{X} - t_{1-\alpha;\,n-1}\cdot\frac{s}{\sqrt{n}}; +\infty\right)$ - **левосторонний доверительный интервал для параметра $m = E(X)$ с доверительной вероятностью γ при неизвестной дисперсии;**

Упражнение. Записать **односторонние доверительные интервалы** для всех остальных случаев.

Поправка на конечный объем генеральной совокупности

При конечном объеме генеральной совокупности

а) **доверительный интервал для математического ожидания при известной дисперсии** имеет вид:

$$\left(\overline{x} - z_{(1+\gamma)/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}};\ \overline{x} + z_{(1+\gamma)/2}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right),$$

б) **Доверительный интервал для математического ожидания при неизвестной дисперсии:**

$(\bar{X} - \delta, \bar{X} + \delta)$, где

$$\delta = \frac{t_{n-1;(1+\gamma)/2} \cdot s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2,$$

$t_{n-1;(1+\gamma)/2}$ – квантиль распределения Стьюдента с $(n-1)$ степенью свободы уровня $(1+\gamma)/2$;

в) **Приближенный доверительный интервал для вероятности:**

$$\left(\frac{m}{n} - z_{(1+\gamma)/2} \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)(N-n)}{n(N-1)}};\ \frac{m}{n} - z_{(1+\gamma)/2} \sqrt{\frac{\frac{m}{n}\left(1 - \frac{m}{n}\right)(N-n)}{n(N-1)}} \right)$$

Допустимый объём выборки

Объем выборки, гарантирующий с вероятностью γ меньшую, чем ε, ошибку при замене математического ожидания выборочным средним:

$$n > \frac{Var(X)}{(1-\gamma)\varepsilon^2}.$$

Объем выборки, гарантирующий с вероятностью γ меньшую, чем ε, ошибку при замене вероятности р относительной частотой:

$$n > \frac{p(1-p)}{(1-\gamma)\varepsilon^2}.$$

При неизвестном значении p:

$$n > \frac{1}{4(1-\gamma)\varepsilon^2}.$$

Если X – нормально распределенная случайная величина, или объем выборки достаточно велик:

$$n > \frac{\sigma^2 z^2_{(1+\gamma)/2}}{\varepsilon^2}$$

(при замене математического ожидания выборочным средним; z_α – квантиль уровня α стандартного нормального закона);

$$n > \frac{p(1-p) \cdot z^2_{(1+\gamma)/2}}{\varepsilon^2}$$

(при замене вероятности относительной частотой).