Set 30 Let $X_1, \dots, X_n$ be independent and identically distributed random variables with the same mean μ and variance σ^2, and let $\bar{X}$ denote the sample mean. Tchebychev's inequality says that the probability $\mathbb{P}(|\bar{X} - \mu| > \delta)$ does not exceed $\mathrm{Var}(\bar{X})/\delta^2$ for every constant $\delta > 0$, where $\mathrm{Var}(\bar{X})$ denotes the variance of $\bar{X}$.

(a) Prove that the statement

$$|\bar{X} - \mu| > \sqrt{20}\,\frac{\sigma}{\sqrt{n}}$$

holds with a probability that does not exceed 0.05.

Then,

$$\mathbb{P}(|\bar{X} - \mu| > \delta) \le \frac{\mathrm{Var}(\bar{X})}{\delta^2}$$

$$\mathbb{P}\left(|\bar{X} - \mu| > \sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right) \le \frac{\mathrm{Var}(\bar{X})}{\left(\sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right)^2}$$

$$\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Finally,

$$\mathbb{P}\left(|\bar{X} - \mu| > \sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right) \le \frac{\frac{\sigma^2}{n}}{20\,\frac{\sigma^2}{n}}$$

$$\mathbb{P}\left(|\bar{X} - \mu| > \sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right) \le \frac{1}{20}$$

(b) Prove that the statement

$$\bar{X} - \sqrt{20}\,\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + \sqrt{20}\,\frac{\sigma}{\sqrt{n}}$$

holds with a probability that is not smaller than 0.95. NOTE: the previous statement gives the following confidence interval

$$\left[\bar{X} - \sqrt{20}\,\frac{\sigma}{\sqrt{n}}, \bar{X} + \sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right]$$

for the (unknown) population mean μ, with a confidence level of at least 95% and for every sample size n.

Then,

$$\bar{X} - \sqrt{20}\ \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

$$-\bar{X} + \sqrt{20}\ \frac{\sigma}{\sqrt{n}} \geq -\mu \geq -\bar{X} - \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{20}\ \frac{\sigma}{\sqrt{n}} \geq \bar{X} - \mu \geq -\sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{20}\ \frac{\sigma}{\sqrt{n}} \geq |\bar{X} - \mu|$$

$$|\bar{X} - \mu| \leq \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

However,

$$\mathbb{P}(|\bar{X} - \mu| > \delta) \leq \frac{\mathrm{Var}(\bar{X})}{\delta^2}$$

But,

$$\mathbb{P}(|\bar{X} - \mu| \leq \delta) \leq 1 - \frac{\mathrm{Var}(\bar{X})}{\delta^2}$$

Therefore,

$$\mathbb{P}\left(|\bar{X} - \mu| \leq \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \leq 1 - \frac{\frac{\sigma^2}{n}}{\left(\sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right)^2}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \leq \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \leq 1 - \frac{\frac{\sigma^2}{n}}{20\ \frac{\sigma^2}{n}}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \leq \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \leq 1 - \frac{1}{20}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \leq \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \leq \frac{19}{20}$$

(c) Prove that the statement

$$\mu - \sqrt{20}\ \frac{\sigma}{\sqrt{n}} \le \bar{X} \le \mu + \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

holds with a probability that is not smaller than 0.95.

Then,

$$\mu - \sqrt{20}\ \frac{\sigma}{\sqrt{n}} \le \bar{X} \le \mu + \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

$$-\sqrt{20}\ \frac{\sigma}{\sqrt{n}} \le \bar{X} - \mu \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

$$|\bar{X} - \mu| \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}$$

And,

$$\mathbb{P}(|\bar{X} - \mu| \le \delta) \le 1 - \frac{\mathrm{Var}(\bar{X})}{\delta^2}$$

Therefore,

$$\mathbb{P}\left(|\bar{X} - \mu| \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \le 1 - \frac{\frac{\sigma^2}{n}}{\left(\sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right)^2}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \le 1 - \frac{\frac{\sigma^2}{n}}{20\ \frac{\sigma^2}{n}}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \le 1 - \frac{1}{20}$$

$$\mathbb{P}\left(|\bar{X} - \mu| \le \sqrt{20}\ \frac{\sigma}{\sqrt{n}}\right) \le \frac{19}{20}$$

(d) Let ${\omega_1}^{act}, \ldots, {\omega_n}^{act}$ be a sample from a population Ω, and let $X: \Omega \to \mathrm{R}$ be a filter (that is, a random variable in Statistics, or a measurable function in Mathematics) that produces n outputs (called observations in Statistics) ${x_1}^{obs}, \ldots, {x_n}^{obs}$ by the formula ${x_i}^{obs} = X({\omega_i}^{act})$. Let $\bar{x}$ denote the average of ${x_1}^{obs}, \ldots, {x_n}^{obs}$. Does the confidence interval

$$\left[\bar{x} - \sqrt{20}\,\frac{\sigma}{\sqrt{n}}, \bar{x} + \sqrt{20}\,\frac{\sigma}{\sqrt{n}}\right]$$

cover the (unknown) population mean μ or not?

The confidence interval covers the population mean with a confidence percentage of 95%.