Check Your Understanding of the Lecture Material

Finger Exercises with Solutions

Instructor: David Dobor

April 15, 2017

While the solutions to these questions are available, you are strongly encouraged to do your best to answer them on your own and to only peek into the solutions if you get stuck or to verify your answers.

Understanding this material is your key to success on your final examination.

Check Your Understanding: 1. Markov inequality

Let Z be a nonnegative random variable that satisfies $\mathbf{E}[Z^4] = 4$. Apply the Markov inequality to the random variable Z^4 to find the tightest possible (given the available information) upper bound on $\mathbf{P}(Z \geq 2)$.

$\mathbf{P}(Z \geq 2) \leq$

Check Your Understanding: 2. Chebyshev inequality

Let Z be normal with zero mean and variance equal to 4. For this case, the Chebyshev inequality yields:

$\mathbf{P}\big(|Z| \geq 4\big) \leq$

Check Your Understanding: 3. Chebyshev versus Markov

Let X be a random variable with zero mean and finite variance. The Markov inequality applied to $|X|$ yields

$$\mathbf{P}\big(\,|X| \geq a\big) \leq \frac{\mathbf{E}\big[\,|X|\,\big]}{a},$$

whereas the Chebyshev inequality yields

$$\mathbf{P}\big(\,|X| \geq a\big) \leq \frac{\mathbf{E}[X^2]}{a^2}.$$

a) Is it true that the Chebyshev inequality is stronger (i.e., the upper bound is smaller) than the Markov inequality, when a is very large?

Yes

No

b) Is it true that the Chebyshev inequality is always stronger (i.e., the upper bound is smaller) than the Markov inequality?

Yes

No

Check Your Understanding: 4. Sample mean bounds

We saw during the lecture that if X_i are i.i.d. with mean μ and variance σ^2, and if $M_n = (X_1 + \cdots + X_n)/n$, then we have an inequality of the form

$$\mathbf{P}\big(|M_n - \mu| \geq \epsilon\big) \leq \frac{a\sigma^2}{n},$$

for a suitable value of a.

a) If $\epsilon = 0.1$, then the value of a is: ...

b) If we change $\epsilon = 0.1$ to $\epsilon = 0.1/k$ for $k \geq 1$ (i.e., if we are interested in k times higher accuracy), how should we change n so that the value of the upper bound does not change from the value calculated in part a)?

   n should:

   - stay the same
   - increase by a factor of k
   - increase by a factor of k^2
   - decrease by a factor of k
   - none of the above

Check Your Understanding: 5. Polling

Consider the polling example we discussed in class. We saw that if we want to have a probability of at least 95% that the poll results are within 1 percentage point of the truth, Chebyshev's inequality recommends a sample size of $n = 50,000$. This is very large compared to what is done in practice. Newspaper polls use smaller sample sizes for various reasons. For each of the following, decide whether it is a valid reason.

In the real world,

a) the accuracy requirements are looser.

Yes

No

b) the Chebyshev bound is too conservative.

Yes

No

c) the people sampled are all different, so their answers are not identically distributed.

Yes

No

d) the people sampled do not have independent opinions.

Yes

No

Check Your Understanding: 6. Convergence in probability

a) Suppose that X_n is an exponential random variable with parameter $\lambda = n$. Does the sequence $\{X_n\}$ converge in probability?

Yes

No

b) Suppose that X_n is an exponential random variable with parameter $\lambda = 1/n$. Does the sequence $\{X_n\}$ converge in probability?

Yes

No

c) Suppose that the random variables in the sequence $\{X_n\}$ are independent, and that the sequence converges to some number a, in probability. Let $\{Y_n\}$ be another sequence of random variables that are dependent, but where each Y_n has the same distribution (CDF) as X_n. Is it necessarily true that the sequence $\{Y_n\}$ converges to a in probability?

Yes

No

Answers

QUESTION 1

We have

$$\mathbf{P}(Z \geq 2) = \mathbf{P}(Z^4 \geq 16) \leq \frac{\mathbf{E}[Z^4]}{16} = \frac{4}{16} = \frac{1}{4}.$$

QUESTION 2

We have

$$\mathbf{P}\big(|Z| \geq 4\big) \leq \frac{\mathrm{var}(Z)}{4^2} = \frac{4}{4^2} = \frac{1}{4}.$$

QUESTION 3

a) Yes, because for very large a, the term $1/a^2$ will be much smaller than $1/a$.

b) No. For example, suppose that $a = 1$. It is certainly possible to have $\mathbf{E}[X^2] > \mathbf{E}\big[\,|X|\,\big]$, in which case the Markov inequality provides a stronger bound.

QUESTION 4

a) Chebyshev's inequality yields

$$\mathbf{P}\big(\,|M_n - \mu| \geq \epsilon\big) \leq \frac{\sigma^2}{n\epsilon^2},$$

so that $a = 1/\epsilon^2 = 1/0.1^2 = 100$.

b) In order to keep the same upper bound, the term $n\epsilon^2$ in the denominator needs to stay constant. If we reduce ? by a factor of k, then ϵ^2 gets reduced by a factor of k^2. Thus, n will have to be increased by a factor of k^2.

QUESTION 5

a) Requiring the accuracy to be within one percentage point is too strict for most real world situations.

b) The Chebyshev bound is conservative as discussed in class

c), d) No matter how opinions get formed, as long as we choose who to ask at random, independently and uniformly, the opinions reported will be i.i.d. random variables, so that the last two considerations do not apply.

QUESTION 6

a) In the first case, for any $\epsilon > 0$, we have $\mathbf{P}(X_n \geq \epsilon) = e^{-n\epsilon}$, which converges to zero. Therefore, we have convergence in probability.

b) In the second case, for any $\epsilon > 0$, we have $\mathbf{P}(X_n \geq \epsilon) = e^{-\epsilon/n}$, which converges to one. Therefore, we do not have convergence in probability.

c) Dependence will not make a difference because the definition of convergence in probability involves probabilities of the form $\mathbf{P}\big(|Y_n - a| \geq \epsilon\big)$. These probabilities are completely determined by the marginal distributions of the random variables Y_n, and these marginal distributions are the same as for the sequence X_n.