Confidence Intervals, Conditional Expectation, LLSE, Markov Chains 12

Computer Science Mentors 70

April 23 - 25, 2017

1 Confidence Intervals

1.1 Questions

1. Define i. i. d. variables $A_k \sim$ Bern(p) where $k \in [1, n]$. Assume we can declare that $\mathrm{P}[|\frac{1}{n}\sum_k A_k - p| > 0.25] = 0.01$.

   (a) Please give a 99% confidence interval for p if given A_k.

   (b) We know that the variables X_i, for i from 1 to n, are i.i.d. random variables and have variance σ^2. We also have a value (an observation) of $A_n = \frac{X_1 + \ldots + X_n}{n}$. We want to guess the mean, μ, of each X_i.

   Prove that we have 95% confidence that μ lies in the interval $\left[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}\right]$

   That is, $\mathrm{P}\left[\mu \in \left[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}\right]\right] \geq 95\%$

(c) Give the 99% confidence interval for μ.

2. We have a die with 6 faces of values $1, 2, 3, 4, 5, 6$.

(a) Develop a 99% confidence interval for the value of n samples.

(b) Now, we say the die's face values are consecutive integers, but we do not know the starting number. The values are shifted over by some k; for example, if $k = 6$, the die faces would take on the values $7, 8, 9, 10, 11, 12$. If we observe that the average of the n samples is 15.5, develop a 99% confidence interval for the value of k.

2 Conditional Expectation

2.1 Introduction

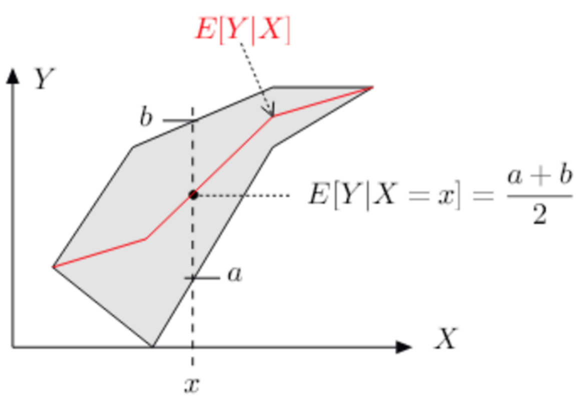
The **conditional expectation** of Y given X is defined by

$$E[Y|X=x] = \sum_{y} y \cdot P[Y=y|X=x] = \sum_{y} y \cdot \frac{P[X=x, Y=y]}{P[X=x]}$$

Properties of Conditional Expectation

$$\begin{aligned} E(a|Y) &= a \\ E(aX+bZ|Y) &= a \cdot E(X|Y) + b \cdot E(Z|Y) \\ E(X|Y) &\geq 0 \text{ if } X \geq 0 \\ E(X|Y) &= E(X) \text{ if } X, Y \text{ independent} \\ E(E(X|Y)) &= E(X) \end{aligned}$$

Here is a picture that shows that conditioning creates a new random variable with a new distribution, taken from Figure 9 of note 26.



2.2 Questions

1. Consider the random variables Y and X with the following probabilities

   This table gives the probability distribution for $P[X \cap Y]$

		X		
		0	1	2
Y	0	0	.1	.2
	1	.1	.2	.1
	2	.2	.1	0

Find:

(a) $\mathrm{E}(Y|X=0)$

(b) $\mathrm{E}(Y|X=1)$

(c) $\mathrm{E}(Y|X=2)$

(d) $\mathrm{E}(Y)$

3 Linear Least Squares Estimator

Theorem: Consider two random variables, X, Y with a given distribution $\mathrm{P}[X = x, Y = y]$. Then

$$\mathrm{L}[Y|X] = \mathrm{E}(Y) + \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)}(X - \mathrm{E}(X))$$

3.1 Questions

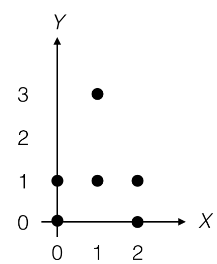
1. Assume that

$$Y = \alpha X + Z$$

where X and Z are independent and $\mathrm{E}(X) = \mathrm{E}(Z) = 0$. Find $\mathrm{L}[X|Y]$.

2. The figure below shows the six equally likely values of the random pair (X, Y). Specify the functions of:

- $L[Y \mid X]$
- $E(X \mid Y)$
- $L[X \mid Y]$
- $E(Y \mid X)$



4 Markov Chains

P is a **transition probability matrix** if:

1. All of the entries are non-negative.
2. The sum of entries in each row is 1.

A **Markov chain** is defined by four things: ($\mathscr{X}$, π_0, P, $\{X_n\}_{n=0}^{\infty}$)

$\mathscr{X}$	Set of states
π_0	Initial probability distribution
P	Transition probability matrix
$\{X_n\}_{n=0}^{\infty}$	Sequence of random variables where:
	$\mathrm{P}[X_0 = i] = \pi_0(i), i \in \mathscr{X}$
	$\mathrm{P}[X_{n+1} = j \vert X_n = i, X_{n-1}, \ldots, X_0] = \mathrm{P}(i,j), \forall n \geq 0, \forall i,j \in \mathscr{X}$

A Markov chain is **irreducible** if we can go from any state to any other state, possibly in multiple steps.

Periodicity has to do with the period of occurrence of a state. If a state s has period 2, the Markov chain can be in s at every other time point. If a state has period 1, it's aperiodic; otherwise, it's periodic. More quantitatively, define value $d(i)$ for each state i as:

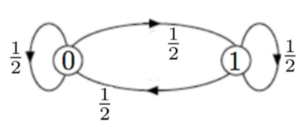
$$d(i) := g.c.d\{n > 0 | P^n(i,i) = \mathrm{P}[X_n = i | X_0 = i] > 0\}, i \in \mathscr{X}$$

If $d(i) = 1$, then the Markov chain is **aperiodic**. If $d(i) \neq 1$, then the Markov chain is periodic and its **period** is $d(i)$.

A distribution π is **invariant** for the transition probability P if it satisfies the following **balance equations**

$$\pi \cdot P = \pi.$$

Theorem 24.3: A finite irreducible Markov chain has a unique invariant distribution.
Theorem 24.4: All irreducible and aperiodic Markov chains converge to the unique invariant distribution. If a Markov chain is finite and reducible, the amount of time spent in each state approaches the invariant distribution as n grows large
Equations that model what will happen at the next step are called **first step equations**



Denote $\beta(i,j)$ as the expected amount of time it would take to move from i to j. $\beta(0,1) = 1 + \frac{1}{2} \cdot \beta(0,1)$ $\beta(1,1) = 0$

4.1 Questions

1. **Life of Alex**

   Alex is enjoying college life. She spends a day either studying, partying, or looking for housing for the next year. If she is studying, the chances of her studying the next day are 30%, the chances of her partying the next day are 50%, and the chances of her looking for housing the next day are 20%. If she is partying, the chances of her partying the next day are 10%, the chances of her studying the next day are 60%, and the chances of her looking for housing the next day are 30%. If she is looking for housing, the chances of her looking for housing the next day are 50%, the chances of her partying the next day are 30% and the chances of her studying the next day are 20%.

   (a) Draw a Markov chain to visualize Alex's life.

   (b) Write out a matrix to represent this Markov chain.

   (c) If Alex studies on Monday, what is the chance that she is partying on Friday? (Don't do the math, just write out the expression that you would use to find it.)

   (d) What percentage of her time should Alex expect to use looking for housing?

   (e) If Alex parties on Monday, what is the chance of Alex partying again before studying?

2. **Stanford Cinema**
   You have a database of an infinite number of movies. Each movie has a rating that is uniformly distributed in 0, 1, 2, 3, 4, 5 independent of all other movies. You want to find two movies such that the sum of their ratings is greater than 7.5 (7.5 is not included).

   a) A Stanford student chooses two movies each time and calculates the sum of their ratings. If is less than or equal to 7.5, the student throws away these two movies and chooses two other movies. The student stops when he/she finds two movies such that the sum of their ratings is greater than 7.5. What is the expected number of movies that this student needs to choose from the database?

   b) A Berkeley student chooses movies from the database one by one and keeps the movie with the highest rating. The student stops when he/she finds the sum of the ratings of the last movie that he/she has chosen and the movie with the highest rating among all the previous movies is greater than 7.5. What is the expected number of movies that the student will have to choose?

3. **Bet On It**
   Smith is in jail and has 3 dollars; he can get out on bail if he has 8 dollars. A guard agrees to make a series of bets with him. If Smith bets A dollars, he wins A dollars with probability 0.4 and loses A dollars with probability 0.6.

   a) Find the probability that he wins 8 dollars before losing all of his money if he bets 1 dollar each time.

   b) Find the probability that he wins 8 dollars before losing all of his money if he bets, each time, as much as possible but not more than necessary to bring his fortune up to 8 dollars.

   c) Which strategy gives Smith the better chance of getting out of jail?

4. **Tossing Coins**
   A fair coin is tossed repeatedly and independently. Find the expected number of tosses till the pattern HTH appears.