Department of Industrial Engineering & Operations Research

IEOR 165 (Spring 2016)

Homework 4

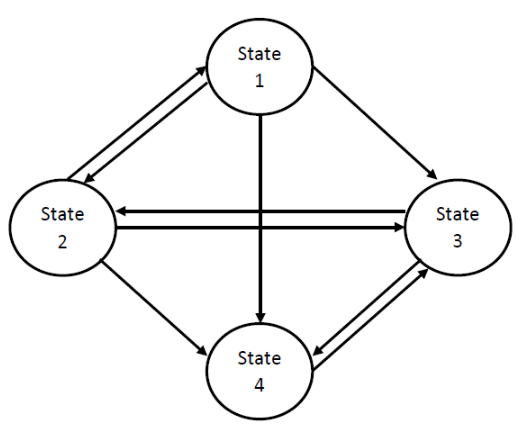
Due: Friday, Apr 8

Question 1. In most of Europe and Asia annual automobile insurance premiums are determined by use of a Bonus Malus (Latin for Good-Bad) system. Each policyholder is given a positive integer valued state and the annual premium is a function of this state (along, of course, with the type of car being insured and the level of insurance). A policyholders state changes from year to year in response to the number of claims made by that policyholder. Because lower numbered states correspond to lower annual premiums, a policyholders state will usually decrease if he or she had no claims in the preceding year, and will generally increase if he or she had at least one claim. (Thus, no claims is good and typically results in a decreased premium, while claims are bad and typically result in a higher premium.) Whereas there are usually many states (20 or so is not atypical), the following table specifies a hypothetical Bonus Malus system having four states.

		Next state if			
State	Annual Premium	0 claims	1 claim	2 claims	≥ 3 claims
1	200	1	2	3	4
2	250	1	3	4	4
3	400	2	4	4	4
4	600	3	4	4	4

Construct a graph to represent this discrete-time Markov chain.

Solution. The transition graph is shown as below (self loops are omitted). The graph is not complete, i.e. not every pair of nodes are connected by an edge. For instance, we can never go from state 3 to state 1.



Question 2. Following Question 1, suppose we are given the following data summarizing the

distribution of claims in each state, use the MLE to estimate the transition probability in the Markov chain.

State	0 claims	1 claim	2 claims	≥ 3 claims
1	226	135	89	65
2	118	190	69	73
3	99	156	187	101
4	63	132	189	199

Solution. By MLE, we have

$$
\begin{aligned}
\hat{p}_{11} &= \frac{N_{11}}{\sum_{j=1}^4 N_{1j}} = \frac{226}{515} = 0.439 \\
\hat{p}_{12} &= \frac{N_{12}}{\sum_{j=1}^4 N_{1j}} = \frac{135}{515} = 0.262 \\
\hat{p}_{13} &= \frac{N_{13}}{\sum_{j=1}^4 N_{1j}} = \frac{89}{515} = 0.173 \\
\hat{p}_{14} &= \frac{N_{14}}{\sum_{j=1}^4 N_{1j}} = \frac{65}{515} = 0.126 \\
\hat{p}_{21} &= \frac{N_{21}}{\sum_{j=1}^4 N_{2j}} = \frac{118}{450} = 0.262 \\
\hat{p}_{22} &= 0 \\
\hat{p}_{23} &= \frac{N_{23}}{\sum_{j=1}^4 N_{2j}} = \frac{190}{450} = 0.422 \\
\hat{p}_{24} &= \frac{N_{24}}{\sum_{j=1}^4 N_{2j}} = \frac{69+73}{450} = 0.316 \\
\hat{p}_{31} &= \hat{p}_{33} = 0 \\
\hat{p}_{32} &= \frac{N_{32}}{\sum_{j=1}^4 N_{3j}} = \frac{99}{543} = 0.182 \\
\hat{p}_{34} &= \frac{N_{34}}{\sum_{j=1}^4 N_{3j}} = \frac{543-99}{543} = 0.818 \\
\hat{p}_{41} &= \hat{p}_{42} = 0 \\
\hat{p}_{43} &= \frac{N_{43}}{\sum_{j=1}^4 N_{4j}} = \frac{63}{583} = 0.108 \\
\hat{p}_{44} &= \frac{N_{44}}{\sum_{j=1}^4 N_{4j}} = \frac{583-63}{583} = 0.892
\end{aligned}
$$

Question 3. A normal population distribution is known to have standard deviation 20. Determine the p-value of a test of the hypothesis that the population mean is equal to 50, if the average of a sample of 64 observations is
(a) 52.5; **(b)** 55.0; **(c)** 57.5.

Solution. $\sigma = 20$, $n = 64$ and $\mu_0 = 50$. Using the formula of two-tailed test

$$p = \mathbb{P}\Big(|Z| > \sqrt{n}\left|\frac{\overline{X} - \mu_0}{\sigma}\right|\Big) = 2(1 - \mathbb{P}(Z \leq \sqrt{n}\left|\frac{\overline{X} - \mu_0}{\sigma}\right|))$$

(a) $\overline{X} = 52.5$, $p = 2(1 - \mathbb{P}(Z \leq 1)) = 0.317$.
(b) $\overline{X} = 55$, $p = 2(1 - \mathbb{P}(Z \leq 2)) = 0.0455$.
(c) $\overline{X} = 57.5$, $p = 2(1 - \mathbb{P}(Z \leq 3)) = 0.0027$.

Question 4. The mean breaking strength of a certain type of fiber is required to be at least 200 psi. Past experience indicates that the standard deviation of breaking strength is 5 psi. If a sample of 8 pieces of fiber yielded breakage at the following pressures,

$$210, 198, 195, 202, 197.4, 196, 199, 195.5$$

would you conclude, at the 5 percent level of significance, that the fiber is unacceptable? What about at the 10 percent level of significance?

Solution. Let's say, you are the manufacturer and you are confident about your product, then your null hypothesis is

$$H_0 : \ \mu \geq 200$$

So we will conduct a one-tailed test:

$$p = \mathbb{P}\Big(Z < \sqrt{n}(\frac{\overline{X} - \mu_0}{\sigma})\Big) = \mathbb{P}\Big(Z < -0.502\Big) = 0.308 > 0.1 > 0.05$$

Thus we should accept H_0, and so accept the fiber.
Some may use the H_0 as $\mu \leq 200$, which is valid if you are suspicious about the quality of this fiber. Then the p-value should be 0.692, for which you should also accept the null hypothesis . Note that the choice of null hypothesis depends on the real scenario and it will make a difference in our final conclusion.

Question 5. Twenty years ago, entering male high school students of Central High could do an average of 24 pushups in 60 seconds. To see whether this remains true today, a random sample of 36 freshmen was chosen. If their average was 22.5 with a sample standard deviation of 3.1, can we conclude that the mean is no longer equal to 24? Use the 5 percent level of significance.

Solution. This is a two-tailed test and the null hypothesis is

$$H_0 : \ \mu = 24$$

So we calculate the p-value as ($n = 36$, $s = 3.1$, $\overline{X} = 22.5$)

$$p = \mathbb{P}\Big(|t| > \sqrt{n}\left|\frac{\overline{X} - \mu_0}{s}\right|\Big) = 2(1 - \mathbb{P}(t \leq \sqrt{n}\left|\frac{\overline{X} - \mu_0}{s}\right|)) = 2(1 - P(t_{35} \leq 2.90)) = 0.0064 \leq 0.05$$

Note that we need to check the t-table with $(n - 1) = 35$ degrees of freedom. So we should reject the null hypothesis and we can conclude that the mean is no longer equal to 24.

Another method is to compare 2.90 with $t_{35}(0.975) = 2.03$. Since $2.90 > 2.03$, we should reject the null hypothesis at the significance level of 0.05.