KAIST 산업 및 시스템 공학과

대학원 입시(학업능력 면접) 기출 및 예시문제 공개

v1 : 2012. 5. 10

v2 : 2014. 8. 8

v3 : 2015. 6. 20

☐ 우리 학과 대학원 지원자들의 시험 준비에 도움을 주고자, 대학원 입시(학업능력 면접) 기출 및 예시문제를 출제하여 공개하오니, 시험 준비에 참고하기 바랍니다.

☐ 기출 및 예시 문제는 학사과정에서 산업공학을 전공한 학생들이 기본적으로 갖추어야 할 필수영역으로 주로 통계 및 OR 분야의 문제로 구성되어 있습니다.

☐ 학업능력면접 대상자의 경우, 개인별 면접시험 시작 전에 90분간 문제를 풀고, 그에 대한 답안지를 가지고 각 면접실에서 질문을 하는 형태로 이루어 집니다.

☐ 여기에 공개하는 기출/예시 문제 내에서만 면접 질문이 이루어지는 것은 아닙니다. 추후 문제가 추가/변경되면 version번호를 바꾸어 지속적으로 update할 예정입니다.

붙 임 : 대학원 면접 질문 예시문제

KAIST 산업 및 시스템 공학과장

<2016 Spring>

Q1. Consider the following LP problem.

$$\begin{aligned} \text{Maximize } z &= -x_1 - x_2 + 4x_3 \\ \text{subject to } \quad & x_1 + x_2 + 2x_3 \le 9 \\ & x_1 + x_2 - \;\; x_3 \le 2 \\ & -x_1 + x_2 + \;\; x_3 \le 4 \\ & x_1, \; x_2, \; x_3 \ge 0 \end{aligned}$$

(1) Let's denote by x_4, x_5, x_6 the slack variables. Fill in the table to create the initial tableau for the simplex method.

	z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
z	-1	-1	-1	4				0
x_1	0							
x_2	0							
x_3	0							

(2) Conduct one pivot operation for the initial tableau, and give the result in the following table.

	z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
z	-1							
	0							
	0							
	0							

(3) After several iterations, we've reached the following tableau. Answer the following questions.

	z	x_1	x_2	x_3	x_4	x_5	x_6	RHS
z	-1	0	-4	0	-1	0	-2	17
x_1	0	1	-1/3	0	1/3	0	-2/3	1/3
x_5	0	0	2	0	0	1	1	6
x_3	0	0	2/3	1	1/3	0	1/3	13/3

(a) Is this the optimal tableau? If yes, why? If no, identify entering and leaving variables.

For the following sub problems (b) – (e), we assume that the current tableau is optimal.

(b) Give the optimal solution (x_1^*, x_2^*, x_3^*) and the objective value.

(c) Write the optimal basis.

(d) Suppose that the RHS of the first constraint (currently 9) increases by a small amount ϵ (to 9+ ϵ). What is the optimal objective value after the change?

(e) Suppose that the objective coefficient of x_2, which is currently -1, is changed to 2, yielding the following objective function:

$$-x_1 + 2x_2 + 4x_3$$

Given this change, what are the 1) optimal objective value, and 2) optimal solution? And why?

(4) Write the dual problem

(5) Give the dual optimal solution assuming that the tableau in (3) is optimal.

Q2. Answer the following questions, which will lead to the well-known Sterling's formula via a heuristic argument:

$$n! \approx n^{n+0.5} e^{-n} \sqrt{2\pi}$$

(a) A discrete random variable X is called a Poisson random variable with parameter λ if the probability mass function of X is given by :

$$\mathrm{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},\ k = 0,1,2,\ldots$$

Find the mean and the variance of X, a Poisson random variable with parameter λ.

(b) Let $X_1, X_2, \cdots$ be independent random variables, each of which has a Poisson distribution with mean 1. What is the distribution of $S_n = \sum_{i=1}^{n} X_i$? In particular, give an explicit expression for $P(S_n = n)$.

(c) Derive the Sterling's formula by applying the central limit theorem to $P(S_n = n) = P(n-1 < S_n \le n)$ so that one can obtain an approximate formula for $P(S_n = n)$ for large n values. Then, equate this result with that of (a) in order to obtain Sterling's formula.

Q3. Consider the computer output below for fitting a simple linear regression model $y=\beta_0+\beta_1 x+\varepsilon$, where $\varepsilon \sim N(0, \sigma^2)$ given a data set. The number of observations is 13.

	Estimate	Standard Error	t value	p-value
β_0	-0.44	0.2	(a)	(b)
β_1	0.31	0.1	(c)	(d)

R-squared: 0.95.

(1) Fill in the missing quantities (a)～(d) of the above output. You can use the attached t-table.

(2) Predict y when x= 2.

(3) You want to test whether there is a significant linear relationship between X and Y.

(a) State the null and alternative hypothesis.

(b) Find the p-value for the above test.

(c) Based on the significance level of α =0.05, what is your conclusion?

(4) Interpret “R-squared = 0.95.”

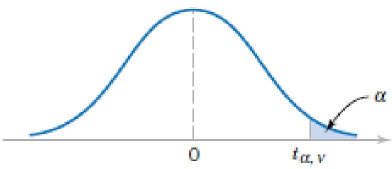


Table IV Percentage Points $t_{\alpha,\nu}$ of the t-Distribution

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.

<2015 Fall>

1. Consider the following pair of primal, dual linear programs.

(P) max $\sum_{j=1}^{n} c_j x_j$

s. t. $\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, \dots, m$

$x_j \ge 0, \quad j = 1,2, \dots, n$

and its dual

(D) min $\sum_{i=1}^{m} b_i y_i$

s. t. $\sum_{i=1}^{m} a_{ij} y_i \ge c_j, \quad j = 1, \dots, n$

$y_i \ge 0, \quad i = 1,2, \dots, m$

(a) Prove the following weak duality theorem:

If $x = (x_1, \dots, x_n)$ is a feasible solution to (P)

and $y = (y_1, \dots, y_m)$ is a feasible solution to (D),

then $\sum_{j=1}^{n} c_j x_j \le \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_j \; y_i \le \sum_{i=1}^{m} b_i y_i$

(b) State the strong duality theorem. (No need to prove it here.)

(c) Prove the following complementary slackness theorem:

Let $x^* = (x_1^*, \dots, x_n^*)$ be a feasible solution to (P) and $y^* = (y_1^*, \dots, y_m^*)$ be a feasible solution to (D). Then x^* and y^* are optimal solutions to (P) and (D) respectively if and only if

$\sum_{i=1}^{m} a_{ij} y_i^* = c_j$ or $x_j^* = 0$, for every $j = 1,2, \dots, n$

$\sum_{j=1}^{n} a_{ij} x_j^* = b_i$ or $y_i^* = 0$, for every $i = 1,2, \dots, m$

2. Suppose that you can draw independent samples $\{U_1, U_2, U_3, \dots,\}$ from uniform distribution on [0,1].

(a) Suggest a method to generate a standard normal random variable using $\{U_1, U_2, U_3, \dots,\}$. Justify your answer.

(b) How can you generate a bivariate standard normal random variable? (Note that a bivariate standard normal distribution is a 2-dimensional normal with zero mean and identity covariance matrix.)

(c) What can you suggest if you want to generate correlated normal random variables with covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$ and mean $\mu = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$?

3. Suppose that $Y_1, Y_2, \ldots, Y_n$ denote a random sample from an exponential distribution with density function given by $f(y;\theta) = \frac{1}{\theta} e^{-\frac{y}{\theta}},\ y \geq 0.$

(a) Find the mean and the variance of the exponential distribution.

(b) Find the maximum likelihood estimator (MLE) $\hat{\theta}$ for θ.

(c) Is the MLE unbiased?

4. A publisher of a newsmagazine has found through past experience that 50 % of subscribers renew their subscriptions. In a random sample of 100 subscribers, 45 indicated that they planned to renew their subscriptions.

(a) What is the p-value associated with the test that the current rate of renewals differs from the rate previously experienced? (You can use the attached Z table.)

(b) Can you conclude that the current rate of renewals differs from the rate previously experienced at α =0.1?

<2015 Spring>

1. Consider the following linear programming problem

(P) $z_{LP} = max\{c'x: Ax = b, x \geq 0, x \in R^n\}$

and its dual

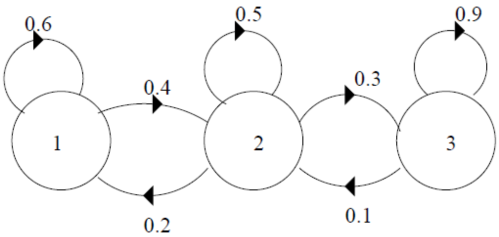
(D) $w_{LP} = min\ \{b'y: y'A \geq c', y \in R^m\}$,

where A is an $m \times n$ matrix with full row rank.

(a) Let x^0 be a feasible solution to P and y^0 be a feasible solution to D. Show that $c'x^0 \leq z_{LP} \leq w_{LP} \leq b'y^0$.

(b) Suppose P has finite optimal value and x^* is an optimal solution. We assume that the simplex method can find an optimal solution if one exists, possibly using the two-phase method. Prove that the dual problem D also has an optimal solution y^* and $c'x^* = b'y^*$, i.e. $z_{LP} = w_{LP}$.

2. Consider a Markov chain $\{X_n; n = 0,1, \dots\}$, specified by the following transition diagram.



(a) Let $Y_n = X_n - X_{n-1}$. Thus, $Y_n = 1$ indicates that the *n*th trnaisition was to the right, $Y_n = 0$ indicates it was a self-transition, and $Y_n = -1$ indicates it was a transition to the left. Find $\lim_{n\to\infty} \mathbf{P}(Y_n = 1)$.

(b) Given that the *n*th transition was a transition to the right ($Y_n = 1$), find the probability that the previsous state was state 1. (You can assume that *n* is large.)

3. Each of 150 newly manufactured items is examined and the number of scratches per item is recorded, yielding the following data:

Number of scratches per item	0	1	2	3	4	5	6	7
Observed frequency	18	37	30	42	13	7	2	1

Let *X* be the number of scratches in a randomly chosen item, and assume that *X* has the Poisson distribution with parameter λ :

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(a) Find the maximum likelihood estimator of λ and then compute the estimate for the given data.

(b) Derive the standard error of your estimator found in part (1) and then compute the estimated standard error of your estimate.

4. Let $X \sim Uniform(0; 1)$ and $Y = -\log(X)$. Find $E(Y)$ and $Var(Y)$.

<2014 Fall>

1. Consider the following LP problem.

$$\begin{aligned} \text{maximize} \quad & 10x_1 + 12x_2 + 12x_3 \\ \text{subject to} \quad & x_1 + 2x_2 + 2x_3 + x_4 = 20 \\ & 2x_1 + x_2 + 2x_3 + x_5 = 20 \\ & 2x_1 + 2x_2 + x_3 + x_6 = 20 \\ & x_1, \dots, x_6 \geq 0 \end{aligned}$$

(a) Suppose we are solving the problem using the simplex method and $\{x_4, x_1, x_6\}$ are in the basis in the current basic feasible solution. Construct the tableau corresponding to the current basis. Identify the current basic feasible solution and its objective value. (Do not find the tableau by performing a pivot. You have to find the tableau using information about the current basis.)

(b) Perform a simplex iteration on the tableau found in (1-1). Use smallest subscript rule to find the entering and the leaving variable. (If you cannot solve problem (1-1), you may find the tableau for part (1-1) by performing a pivot on the original problem.). What can be observed on the solutions? Explain the meaning of this phenomenon.

2. Suppose that you are given with two set of samples F = $\{x_1, x_2, \dots, x_{100}\}$ and M = $\{y_1, y_2, \dots, y_{100}\}$. x_i (or y_i) represents the height of a female (or male) individual. Let μ_F and μ_M denote the mean heights of female and male population, respectively.

Sample mean of F : $m_F = 169.1$, Sample standard deviation of F : $S_F = 5$
Sample mean of M : $m_M = 170.5$, Sample standard deviation of M : $S_M = 6$

(a) Provide a 95% confidence interval of the mean of total (male+female) population.

(b) We would like to test the following hypotheses :

$H_o : \mu_F = \mu_M$ vs $H_1 : \mu_F \neq \mu_M$

What is the p-value for this test?

(c) Now, consider the following hypotheses :

$H_o : \mu_F = \mu_M$ vs $H_1' : \mu_F < \mu_M$

Provide the rejection region when the significance level $\alpha = 5\%$.

Can we reject H_o with significance level $\alpha = 5\%$?

(d) If the answer to (2-2) is different from that of (2-3), please explain why.
Note that H_1' is a stronger argument than H_1.

(e) Now, further suppose that x_i is the spouse of y_i for each i = 1~100. The Pearson correlation coefficient between F and M is $\rho = 0.7$. Using this extra information, can we tell that husbands are taller than their wives, with the significance level $\alpha = 1\%$?

<More questions: basic concepts in OR and Statistics >

1) Linear algebra: vector space, linear independence, basis, rank, linear transformation, range space-null space of linear transformation, nonsingular matrix, determinant, inverse, solving linear equations, eigenvalues, eigenvectors

2) Linear programing: convex sets, convex functions, formulations, basic (feasible) solutions, extreme points, algebraic representation of extreme points, simplex method in tableau form, revised simplex method, duality theory, weak duality, strong duality, complementary slackness conditions, dual simplex method, sensitivity analysis, Farkas' lemma (theorem of alternatives)

3) Integer programing: formulations, branch-and-bound algorithm

4) Statistics: population, sample, disjoint (mutually exclusive), independent/dependent, Bayes' rule, random variable, (cumulative) distribution function, probability density (mass) function, expectation, variance, discrete random variables (e.g. binomial, geometric, negative binomial, hypergeometric, Poisson), continuous random variables (e.g. uniform, normal, gamma, chi-square, exponential, beta, t, F), moment, moment generating function, Tchebysheff′s Theorem, quantile, joint distribution function, joint probability density (mass) function, marginal distribution, conditional distribution, covariance, correlation, estimator, estimate, order statistics, sampling distribution, central limit theorem, parameter, bias, mean square error, unbiased estimator, standard error, consistent estimator, confidence interval, method of moments, maximum likelihood estimator, hypothesis testing, null hypothesis, alternative hypothesis, test statistics, rejection region, type I error, type II error, significance level, p-value, t-test, F-test, linear regression, least squares estimation, dependent/independent variable (response/predictor), residual, R-square, ANOVA, chi-square test

Example)

1. Convex function의 정의를 기술하시오.

2. Convex optimization problem의 정의를 기술하시오.

3. Simplex에서 shadow price의 의미를 설명하고 duality의 개념과 연관하여 설명하시오.

4. Simplex에 degeneracy가 발생할 수 있는 상황의 예를 드시오.

5. Branch and bound algorithm을 개념적으로 설명하시오.

6. Network model의 해가 항상 integer로 나오기 위한 조건을 설명하고, 그 이유를 개념적으로 설명하시오

7. 무기억속성(memoryless property)과 마코프속성(Markovian property)의 비교 : 차이점은 무엇이고 공통점 (또는 유사점)은 무엇인가?

8. likelihood function(우도함수)의 의미는 무엇이며, 어디에 어떻게 쓰이는가?

9. MLE (Maximum likelihood estimator, 최우추정)이란 무엇인가? MLE는 불편추정 (Unbiased estimator) 인가? MLE가 bias가 되는 예는? MLE는 consistent estimator인가?

10. 샘플 y1, ... , yn이 관찰되었다. 확률변수 Y의 모수를 maximum likelihood estimation을 통해 추정하고자 한다. 그 과정을 기술하시오.

11. 사건(events) A와 B가 상호배반(mutually exclusive)이면 A와 B는 독립인가?

12. 통계적 가설검정은 귀무가설(null hypothesis) H0 와 대립가설(alternative hypothesis) H1, 두 가설을 주어진 샘플을 바탕으로 분석하여 한 가설을 선택하는 과정이다. 이 때, 일어날 수 있는 두 가지 오류가 있는데 이들은 무엇인가?

13. 통계적 가설검정 (statistical test)에서 p-value가 의미하는 바는 무엇인가?

14. [Bayes Rule] Suppose that a test for AIDS shows the accuracy of 90%. That is, Prob(Positive | AIDS) = 0.9 and Prob(Negative | Normal) = 0.9. Assume 1% of the entire population is AIDS patients. Now, if someone's test result is positive, what is the probability that this person is indeed an AIDS patient?

15. For many simulation models, arrivals are modeled to follow a Poisson process.

(a) List the three properties that define a Poisson process.

(b) When arrivals follow a Poisson process, what can you say about the interarrival times of those arrivals?

16. (a) Let X1 and X2 be a uniformly distributed random variable, independent of each other. What would be the distribution of Z = X1 + X2?

(b) Repeat the above question when we have X1, ..., Xn, where X1, ..., Xn are iid uniform RV's & n is a very large number.

(c) Explain how, if possible, you would derive the distribution for Z = X1/X2.

17. Suppose buses arrive at the bus stop every 30 min on average.

(a) If there the interarrival time is a constant (i.e. the interarrival time is exactly 30 min), what would be the expected waiting time for a randomly arriving passenger?

(b) What if the interarrival time is random? Is the waiting time shorter than, same as, or longer than your answer to part (a)?