Applied Statistics and Numerical Methods for Engineers
ChE 301, Fall 1998
Exam Number Four
Administered: 12:30-2:30 Saturday December 12, 1998

THE EXAM HAS 11*10=110 POINTS.
Reserve approximately 11 minutes for each of the 11 problems.

Problem 1. Probability

Consider a group of students attempting to answer a probability question on their final exam. There are 20 students in the class. 15 of those students really know how to solve the problem; the other 5 intend to rely on their skill at guessing. After the test, it is revealed that 14 of the students answered the question correctly. If the probability of answering the question correctly GIVEN that a student really knew how to solve the problem is 0.867, find the probability of answering the question correctly GIVEN that a student relied on guessing.

Solution:

Let Event A be the event that a student knows how to solve the problem, $P(A) = 15/20 = 0.75$
Let Event A' be the event that a student relies on guessing, $P(A') = 5/20 = 0.25$
Let Event B be the event that a student answers the questions correctly, $P(B) = 14/20 = 0.7$
Let Event B' be the event that a student answers the questions incorrectly, $P(B') = 6/20 = 0.3$
With these designations, we want: $P(B \mid A')$

The conditional probability that of answering the question correctly GIVEN that a student really knew how to solve the problem is 0.867, so

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

$$0.867 = \frac{P(A \cap B)}{0.75} \text{ , therefore } P(A \cap B) = 0.65$$

which is the probability that a student knew how to solve the problem AND answered it correctly. Now every student who answered the problem correctly, either knew it or guessed, so

$$P(A \cap B) + P(A' \cap B) = P(B)$$

$$0.65 + P(A' \cap B) = 0.70 \text{ , therefore } P(A' \cap B) = 0.05$$

which is the probability that a student relied on guessing AND answered it correctly. This is all we need to solve

$$P(B \mid A') = \frac{P(A' \cap B)}{P(A')} = \frac{0.05}{0.25} = 0.20$$

Problem 2. Random Variables

Given the joint PDF

$$f(x,y) = \begin{cases} \frac{2}{5}(2x + 3y) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(0 \le x \le 0.5, 0.5 \le y \le 1)$

Solution:

$$P(a < X < b, c < Y < d,) = \int_c^d \int_a^b f(x,y)\,dx\,dy$$

$$P(0 \le x \le 0.5, 0.5 \le y \le 1) = \int_{0.5}^{1} \int_{0}^{0.5} \frac{2}{5}(2x + 3y)\,dx\,dy$$

$$P = \int_{0.5}^{1} \frac{2}{5}\left(x^2 + 3yx\right)\Big|_{x=0}^{x=0.5} dy = \int_{0.5}^{1} \frac{2}{5}(0.25 + 1.5y)\,dy$$

$$P = \int_{0.5}^{1} \frac{2}{5}(0.25 + 1.5y)\,dy = \frac{2}{5}\left(0.25y + 0.75y^2\right)\Big|_{y=0.5}^{y=1}$$

$$P = \frac{2}{5}\left(\frac{1}{8} + \frac{9}{16}\right) = \frac{22}{80} = \frac{11}{40} = 0.275$$

Problem 3. Expectations
Given the PDF

$$f(x) = \frac{x^2}{9} \qquad \text{for } 0 \le x \le 3$$

find the average value of the function of the random variable x, $g(x) = (x+1)$ over the range $0 \le x \le 1$

Solution:

$$\mu_{g(x)} = E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

$$\mu_{g(x)} = \int_0^1 (x+1)\frac{x^2}{9}\,dx = \frac{1}{9}\int_0^1 (x^3 + x^2)\,dx = \frac{1}{9}\left(\frac{x^4}{4} + \frac{x^3}{3}\right)\Bigg|_{x=0}^{x=1}$$

$$\mu_{g(x)} = \frac{1}{9}\left(\frac{1}{4} + \frac{1}{3}\right) = \frac{1}{9}\left(\frac{7}{12}\right) = \frac{7}{108} = 0.0648$$

Problem 4. Discrete Distributions
The University of Tennessee has contracts with the Personal Computer manufacturers: Gateway 2000, Dell, Compaq, and IBM. A computer lab is outfitted with 10 Gateway 2000's, 8 Dell's, 4 Compaq's, and 2 IBM's. One night, three computers are stolen from the lab. The thieves randomly selected the computers. What is the probability that two Gateways and one IBM computer were stolen?

Solution:

multivariate hypergeometric PDF.

$$h(\{x\};N,n,\{a\}) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\binom{a_3}{x_3}\cdots\binom{a_k}{x_k}}{\binom{N}{n}}$$

$$a_1 = 10,\ a_2 = 8,\ a_3 = 4,\ a_4 = 2,\ N = \sum a_i = 24$$
$$x_1 = 2,\ x_2 = 0,\ x_3 = 0,\ x_4 = 1,\ n = \sum x_i = 3$$
$$P(\{X\} = \{2,0,0,1\}) = \frac{\binom{10}{2}\binom{8}{0}\binom{4}{0}\binom{2}{1}}{\binom{24}{3}} = \frac{90}{2024} = 0.0445$$

Problem 5. Continuous Distributions

A formulation plant produces a sulfur-containing liquid used as a fungicide. Plant exhaust is vented through the ceiling and the vent discharge is blown downwind into a near-by residential neighborhood, where the air has on average 2.0 ppm dissolved sulfur-compound with standard deviation of 1.0 ppm. The residents of a near-by neighborhood complain when they smell the noxious sulfur odor. If the odor detection level of the neighbors is 4.0 ppm, what is the probability at any given time that the residents can smell the odor?

Solution:

Normal PDF.

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\mu = 2$ and $\sigma = 1$ and $x \le 4$

$$z = \frac{x-\mu}{\sigma} = \frac{4-2}{1} = 2$$

$P(X \le 4) = P(Z \le 2) = 0.9772$ From Table A.3

so $1 - 0.9772 = 0.0228$ is the probability that the odor is detectable in the neighborhood.

Problem 6. Sampling and Estimation

A manufacturer of windshield wipers claims that her product continues to work for 18 months before requiring replacement with a standard deviation of 2 months. ($\mu = 15, \sigma = 2$) You and 12 of your friends all buy these windshield wipers and put them on your automobiles at the same time. You record the time when the windshield wipers must be replaced and find the sample mean to be $\bar{x} = 13$ and sample standard deviation $s = 1$, find a 95% confidence interval for the population mean, assuming the stated population variance is doubtful and not to be trusted.

Solution:

To estimate the mean, variance unknown, use the t-distribution.

$v = n - 1 = 12$. First, find $t_{\alpha/2}$ for $\alpha = 0.05$ from table A.4, $t_{0.025} = 2.179$

$$P(\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha$$
$$P(13 - (2.179)\frac{1}{\sqrt{13}} < \mu < 13 + (2.179)\frac{1}{\sqrt{13}}) = 0.95$$
$$P(12.40 < \mu < 13.60) = 0.95$$

Problem 7. Linear Algebra

Consider an nxn matrix, $\underline{\underline{J}}$, with determinant $\det(\underline{\underline{J}}) = -4$. Which of the following statements are true?

(a) The inverse of $\underline{\underline{J}}$ does not exist.

(b) The rows of $\underline{\underline{J}}$ are all linearly independent.

(c) The rank of $\underline{\underline{J}}$ is two.

(d) There is a unique solution to the system of linear equations $\underline{\underline{J}}\underline{x} = \underline{R}$ for any real nx1 vector, $\underline{R}$.

(e) The reduced row echelon form of $\underline{\underline{J}}$ will have at least one row completely filled with zeroes.

Solution:

(b), (c), and (d) are true. (a) and (e) are false.

Problem 8. Regression

In an isothermal, jacketed, batch reactor, the following reaction occurs:

$$A + 2B \rightarrow C$$

You know the initial concentrations of A, B, and C, A_o, B_o, and C_o; the temperature, T; and you record the concentration of C, $C(t)$, as a function of time, t. From kinetics, you know the rate of production of C is given by:

$$rate(t) = \frac{dC(t)}{dt} = A \cdot B^2 \cdot k_o \exp^{\frac{-Ea}{RT}} = (A_o - C)\cdot(B_o - 2C)^2 \cdot k_o \exp^{\frac{-Ea}{RT}}$$

You want to perform a regression to obtain the reaction rate constant, k_o and the activation energy, E_a but you couldn't measure the rate, $\frac{dC(t)}{dt}$, only $C(t)$. So, you can rearrange and analytically integrate this differential equation to yield:

$$k_o \exp^{\frac{-Ea}{RT}} t = \left(\frac{-2}{(B_o - 2A_o)^2}\right)\begin{bmatrix} 2(B_o - 2A_o)\left(\frac{1}{B_o - 2C} - \frac{1}{B_o - 2C_o}\right) \\ -2\ln\left(\frac{B_o - 2C}{B_o - 2C_o}\right) + \frac{1}{2}\ln\left(\frac{A_o - C}{A_o - C_o}\right)\end{bmatrix}$$

(a) Put this equation in a form like $y = b_o + b_1 x$ so that you could perform least squares linear regression on it to determine the reaction rate constant, k_o and the activation energy, E_a. Clearly identify all four variables in this equation, $y = b_o + b_1 x$. Clearly identify how to obtain k_o and E_a from the fit constants.

(b) If you only have $C(t)$, can you independently identify k_o and E_a, if you perform this experiment only once? If not, what variable would you have to change (and run the experiment again) in order to independently identify k_o and E_a?

Solution:

Rewrite the equation as:

$$k_o \exp^{\frac{-Ea}{RT}} t = RHS$$

$$\ln(k_o) + \frac{-E_a}{RT} + \ln(t) = \ln(RHS)$$

$$\ln(RHS) = \left[\ln(k_o) + \frac{-E_a}{RT}\right] + \ln(t)$$

so $y = \ln(RHS)$ $b_o = \left[\ln(k_o) + \frac{-E_a}{RT}\right]$, $b_1 = 1$, and $x = \ln(t)$

The intercept of the best-fit line would give us a function of k_o and E_a. This experiment would have to be repeated at a different temperature to independently determine k_o and E_a.

Problem 9. Solutions of Systems of Nonlinear Algebraic Equations

Consider the system of equations

$$x + y = 1$$

$$3x - y = -2$$

The Jacobian for this system is $\underline{\underline{J}} = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$ with inverse $\underline{\underline{J}}^{-1} = \begin{bmatrix} 0.25 & 0.25 \\ 0.75 & -0.25 \end{bmatrix}$

Perform one complete Newton-Raphson step to find the roots of the system of equations, starting with the initial guess $\underline{x_0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Solution:

$$\underline{R} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\underline{\underline{J}}\underline{\delta x} = \underline{R}$$

$$\underline{\delta x} = \underline{\underline{J}}^{-1}\underline{R} = \begin{bmatrix} -0.25 \\ 1.25 \end{bmatrix}$$

$$\underline{x}_1 = \underline{x}_0 + \underline{\delta x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -0.25 \\ 1.25 \end{bmatrix} = \begin{bmatrix} -0.25 \\ 1.25 \end{bmatrix}$$

Problem 10. Numerical Integration

Perform a numerical using Simpson's 1/3 rule on the function, $f(x) = x^2$ over the interval $2 \le x \le 4$ using $n = 2$ intervals.

Solution:

$$\int_a^b f(x)dx \approx \frac{h}{3}\left(f(a) + 4\sum_{i=2,4,6}^{n-1} f(x_i) + 2\sum_{i=3,5,7}^{n-2} f(x_i) + f(b) \right)$$

a = 2, b=4, h = (b-a)/n= 1

$$\int_2^4 x^2 dx \approx \frac{1}{3}(f(2) + 4f(3) + f(4)) = \frac{1}{3}(4 + 36 + 16) = 18.667$$

$$\int_2^4 x^2 dx = \left. \frac{x^3}{3} \right|_{x=2}^{x=4} = \frac{56}{3} = 18.667$$

Problem 11. Solutions of Ordinary Differential Equations

Perform one Euler method step on the differential equation: $\frac{dy}{dx} = \frac{xy}{2}$ subject to the initial condition, $y(x = 1) = 2$, using a step size $\Delta x = 0.1$

Solution:

$$y(x_{i+1}) = y(x_i) + \Delta x * f(y(x_i), x_i)$$

$$y(x_1) = y(x_0) + \Delta x * f(y(x_0), x_0)$$

$$y(x_1) = 2.0 + 0.1 * f(2,1) = 2.0 + 0.1 * 1 = 2.1$$

analytical solution:

$$\frac{dy}{y} = \frac{xdx}{2}, \qquad \ln\left(\frac{y}{y_o}\right) = \frac{x^2 - x_o^2}{4}, \qquad y = y_o \exp\left[\frac{x^2 - x_o^2}{4}\right]$$

$$y(x = 1.1) = 2\exp\left[\frac{1.1^2 - 1^2}{4}\right] = 2.107805$$