INTRODUCTION TO PROBABILITY

by

Dimitri P. Bertsekas and John N. Tsitsiklis

CHAPTER 4: ADDITIONAL PROBLEMS

Last updated: July 6, 2003

Problems marked with "[D]" are from "Fundamentals of Applied Probability", by Alvin Drake, and are included here with the author's permission.

SECTION 4.1. Transforms

Problem 1. Find the transform associated with the random variable X with PMF

$$f_X(x) = \begin{cases} p\dfrac{e^{-\lambda}\lambda^k}{k!} + (1-p)\dfrac{e^{-\mu}\mu^k}{k!}, & \text{if } k = 0, 1, \ldots, \\ 0, & \text{otherwise,} \end{cases}$$

where λ and μ are positive scalars, and p satisfies $0 \le p \le 1$.

Problem 2. Let X be a random variable such that

$$M(s) = a + be^{2s} + ce^{4s}, \quad \mathbf{E}[X] = 3, \quad \text{var}(X) = 2.$$

Find a, b, and c, and the PMF of X.

Problem 3. Consider two independent random variables X and Y whose PMFs are given by

$$p_X(x) = \begin{cases} 1/2, & \text{if } x = 0, 1, \\ 0 & \text{otherwise,} \end{cases} \qquad p_Y(y) = \begin{cases} 1/2, & \text{if } y = 1, 2, \\ 0, & \text{otherwise.} \end{cases}$$

Let R be a random variable that takes, with equal probability, either the value of X or the value of Y.

(a) Evaluate $\mathbf{E}[R]$ and $\text{var}(R)$.

(b) Let G denote the sum of six independent random variables with the same distribution as R. Find the transform, the mean, and the variance associated with G.

Problem 4. The transform associated with a random variable Y has the form

$$M_Y(s) = a^6(0.1 + 2e^s + 0.1e^{4s} + 0.4e^{7s})^6.$$

Find a, $p_Y(41)$, $p_Y(11)$, the third largest possible value of Y, and its corresponding probability.

Problem 5. The transforms associated with two independent discrete random variables X and Y are

$$M_X(s) = \left(\frac{1}{2}e^{2s} + \frac{1}{2}e^{4s}\right)^7, \qquad M_Y(s) = e^{8(e^s-1)}.$$

Find $p_X(15)$, $p_Y(5)$, $\mathbf{E}[X]$, $\mathbf{E}[Y^2]$, and $\mathbf{P}(X + Y = 15)$.

Problem 6. The transform and the mean associated with a discrete random variable X are given by

$$M(s) = ae^s + be^{4(e^s-1)}, \qquad \mathbf{E}[X] = 3.$$

Find:

(a) The scalar parameters a and b.

(b) $p_X(1)$, $\mathbf{E}[X^2]$, and $\mathbf{E}[2^X]$.

(e) $\mathbf{P}(X + Y = 2)$, where Y is a random variable that is independent of X and is identically distributed with X.

Problem 7. Your section in the probability course you are attending consists of 5 students and meets for a total of 11 sessions. Each session ends with a short quiz. The score of each of the 5 students is equally likely to be 1, 2, 3, or 4, independently of other students. Let Q be the sum of the quiz scores of the students in today's session.

(a) Find the transform associated with Q.

(b) Find the range of Q and calculate $p_Q(q)$, for $q = 6$, 10, 12, 15, 18.

(c) Suppose that the time to grade all of the quizzes from a single session is a normal random variable with mean 30 minutes and variance 5 minutes, independently of other sessions. Find the PDF and the transform associated with the total amount of time spent to grade the the quizzes for all 11 sessions.

Problem 8. Suppose that

$$M_X(s) = \frac{6 - 3s}{2(1 - s)(3 - s)}.$$

Find the PDF of the associated random variable.

Problem 9. Let X_1, X_2, X_3, X_4 be independent random variables with common mean, variance, and transform, denoted by $\mathbf{E}[X]$, $\text{var}(X)$, and $M_X(s)$, respectively. Let Y be a random variable that is independent of X_1, X_2, X_3, X_4, and is associated with the transform $M_Y(s)$. Each part of this problem introduces a new random variable either as a function of X_1, X_2, X_3, X_4, and Y, or as a transform defined in terms of $M_X(s)$ and $M_Y(s)$. For each part, determine the mean and variance of the new random variable.

(a) $W = X_1 + X_2 + X_3 + X_4$.

(b) $V = 0.25(X_1 + X_2 + X_3 + X_4)$.

(c) $U = X_1 + X_2 + X_3 + X_4 + Y$.

(d) $R = 4X - Y$.

(e) $M_Q(s) = [M_X(s)]^5$.

(f) $M_H(s) = [M_X(s)]^2[M_Y(s)]^3$.

(g) $M_G(s) = e^{6s}M_X(s)$.

(h) $M_D(s) = M_X(7s)$.

Problem 10. Let X and Y be independent exponential random variables with a common parameter λ.

(a) Find the transform associated with $aX + Y$, where a is a constant.

(b) Use the result of part (a) to find the PDF of $aX + Y$, for the case where a is positive and different than 1.

(c) Use the result of part (a) to find the PDF of $X - Y$.

Problem 11. **Mean and variance of the Poisson.** Use the formula for the transform associated with a Poisson random variable X to calculate $\mathbf{E}[X]$ and $\mathbf{E}[X^2]$.

Problem 12. An integral is well-defined if the integral of the magnitude of the integrand is finite. Let us extend the definition of the transform by allowing the parameter s to be a complex number. Show that for purely imaginary values of s, the integral in the definition of the transform associated with a continuous random variable is well-defined.

Problem 13. **The z-Transform.** The z-transform associated with a discrete random variable X is a function $Q_X(z)$ of a free parameter z, defined by

$$Q_X(z) = \mathbf{E}[z^X].$$

(a) Show that the transforms $M_X(s)$ and $Q_X(z)$ are related by

$$M_X(s) = Q_X(e^s).$$

(b) Show the moment-generating properties

$$Q_X(1) = 1, \qquad \frac{d}{dz}Q_X(z)\bigg|_{z=1} = \mathbf{E}[X], \qquad \frac{d^2}{dz^2}Q_X(z)\bigg|_{z=1} = \mathbf{E}[X^2] - \mathbf{E}[X].$$

Problem 14. Let X_1, and X_2 be independent random variables. Use the moment-generating properties of transforms to verify that $\mathrm{var}(X_1 + X_2) = \mathrm{var}(X_1) + \mathrm{var}(X_2)$.

SECTION 4.2. Sums of Independent Random Variables - Convolutions

Problem 15. Let X_1 and X_2 be independent random variables with the same PMF:

$$p_{X_1}(x) = p_{X_2}(x) = \begin{cases} 1/4, & \text{if } x = 1, \\ 1/4, & \text{if } x = 2, \\ 1/2, & \text{if } x = 3, \\ 0, & \text{otherwise.} \end{cases}$$

Use convolution to obtain the PMF of $Y = X_1 + X_2$.

Problem 16. Let X be uniform on $[0, 2]$ and let Y be uniform on $[3, 4]$. Assume that X and Y independent. Find and sketch the PDF of $Z = X + Y$, using convolutions.

Problem 17. Let Y be exponentially distributed with parameter 1, and let Z be uniformly distributed over the interval $[0, 1]$. Use convolution to find the PDF of $|Y - Z|$.

Problem 18. X and Y are independent and have PDFs as shown in Fig. 4.2.1 below. Let $W = X + Y$ and find $f_W(w)$ using a graphical argument.

Problem 19. Consider two independent and identically distributed discrete random variables X and Y. Assume that their common PMF, denoted by $p(x)$, is symmetric around zero, i.e., $p(x) = p(-x)$ for all x. Show that the PMF of $X+Y$ is also symmetric around zero and is largest at zero. *Hint:* Use the Schwarz inequality: $\sum_k (a_k b_k) \le \big(\sum_k a_k^2\big)^{1/2}\big(\sum_k b_k^2\big)^{1/2}$.

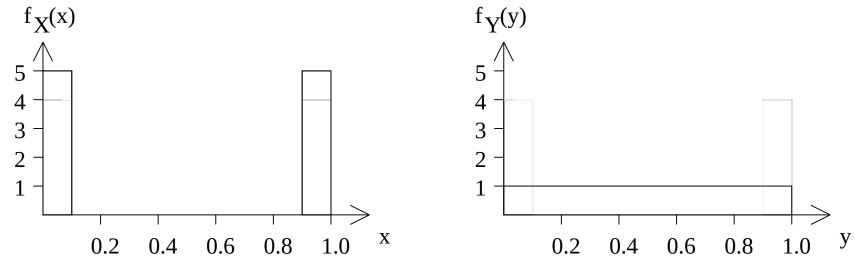

Figure 4.2.1:

Problem 20. Let X be a discrete random variable with PMF p_X and let Y be a continuous random variable, independent from X, with PMF f_Y. Derive a formula for the PDF of the random variable $X + Y$.

SECTION 4.3. More on Conditional Expectation and Variance

Problem 21. The random variables X and Y have the joint PDF shown below:

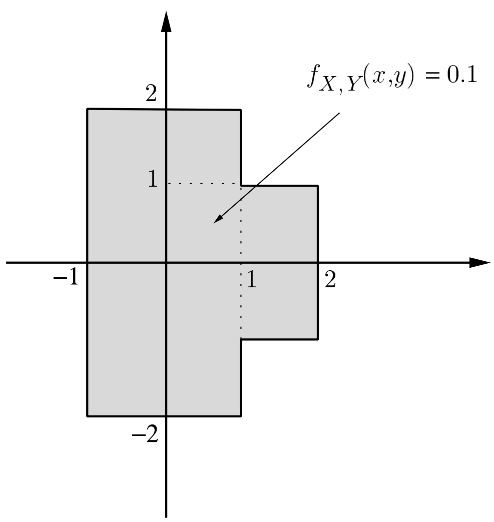

Figure 4.3.1:

(a) Find the conditional PDFs $f_{Y|X}(y \mid x)$ and $f_{X|Y}(x \mid y)$, for various values of x and y, respectively.

(b) Find $\mathbf{E}[X \mid Y]$, $\mathbf{E}[X]$, and $\text{var}(X \mid Y)$. Use these to calculate $\text{var}(X)$.

(c) Find $\mathbf{E}[Y \mid X]$, $\mathbf{E}[Y]$, and $\text{var}(Y \mid X)$. Use these to calculate $\text{var}(Y)$.

Problem 22. Oscar is an engineer who is equally likely to work between zero and one hundred hours each week (i.e., the time he works is uniformly distributed between zero and one hundred). He gets paid one dollar an hour. If Oscar works more than fifty hours during a week, there is a probability of $1/2$ that he will actually be paid overtime, which means he will receive two dollars an hour for each hour he works longer than fifty hours. Otherwise, he will just get his normal pay for all hours he worked that week. Independently of receiving overtime pay, if Oscar works more than seventy five hours in a week, there is a probability of $1/2$ that he will receive a one hundred dollar bonus, in addition to whatever else he earns. Find the expected value and variance of Oscar's weekly pay.

Problem 23. Let X be a geometric random variable with parameter P, where P is itself random and uniformly distributed from 0 to $(n-1)/n$. Let $Z = \mathbf{E}[X \mid P]$. Find $\mathbf{E}[Z]$ and $\lim_{n\to\infty} \mathbf{E}[Z]$.

Problem 24. The random variables X and Y are described by a joint PDF which is constant within the unit area quadrilateral with vertices $(0,0)$, $(0,1)$, $(1,2)$, and $(1,1)$. Use the law of total variance to find the variance of $X+Y$.

Problem 25. Let X, Y, and Z be discrete random variables. Show the following generalizations of the law of iterated expectations.

(a) $\mathbf{E}[Z] = \mathbf{E}\big[\mathbf{E}[Z \mid X, Y]\big]$.

(b) $\mathbf{E}[Z \mid X] = \mathbf{E}\big[\mathbf{E}[Z \mid X, Y] \mid X\big]$.

(c) $\mathbf{E}[Z] = \mathbf{E}\big[\mathbf{E}\big[\mathbf{E}[Z \mid X, Y] \mid X\big]\big]$.

SECTION 4.4. Sum of a Random Number of Independent Random Variables

Problem 26.

(a) You roll a fair six-sided die, and then you flip a fair coin the number of times shown by the die. Find the expected value and the variance of the number of heads obtained.

(b) Repeat part (a) for the case where you roll two dice, instead of one.

Problem 27. A fair coin is flipped independently until the first head is obtained. For each tail observed before the first head, the value of a continuous random variable with uniform PDF over the interval $[0,3]$ is generated. Let the random variable X be defined as the sum of all the values obtained before the first head. Find the mean and variance of X.

Problem 28. Widgets are stored in boxes, and boxes are assembled in crates. Let X be the number of widgets in any particular box, let N be the number of boxes in any particular crate, and let K be the number of crates in any particular shipment. Assume that X, N, and K are independent integer-valued random variables, with expected value equal to 10, and variance equal to 16. Evaluate the expected value and variance of the following random variables.

(a) The number of widgets in any particular crate.

(b) The total number of widgets in any particular shipment.

Problem 29. The transform associated with N, the total number of living groups contacted about the MIT blood drive, is

$$M_N(s) = \left(\frac{1}{3} + \frac{2}{3}e^s\right)^{10}.$$

(a) Determine the PMF of N.

(b) Let the number K of people in any particular living group, be an independent random variable with associated transform

$$M_K(s) = \frac{\frac{1}{5}e^{4s}}{1 - \frac{4}{5}e^s}.$$

Find $p_K(k)$, $\mathbf{E}[K]$, and $\mathrm{var}(K)$.

(c) Let L be the total number of people whose living groups are contacted about the blood drive. Determine the transform, the mean, and the variance associated with L.

(d) Suppose that any particular person, whose living group is contacted, donates blood with probability $1/4$, and that all such individuals make their decisions independently. Let D denote the total number of blood donors from the contacted living groups. Calculate the transform and mean associated with D, and the probability that there will be no donors at all.

SECTION 4.5. Covariance and Correlation

Problem 30. The random variables $X_1, \ldots, X_n$ have common mean μ, common variance σ^2 and, furthermore, $\mathbf{E}[X_i X_j] = c$ for every pair of distinct i and j. Derive a formula for the variance of $X_1 + \cdots + X_n$, in terms of μ, σ^2, c, and n.

Problem 31. Show that

$$\mathrm{cov}(X_1 + \cdots X_m, Y_1 + \cdots + Y_n) = \sum_{i=1}^{m} \sum_{i=1}^{n} \mathrm{cov}(X_i Y_j).$$

Problem 32. Let $X_1, \ldots, X_n$ be some random variables and let $c_{ij} = \mathrm{cov}(X_i, X_j)$. Show that for any numbers $a_1, \ldots, a_n$, we have

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i c_{ij} a_j \geq 0.$$

Problem 33. Consider n independent tosses of a die. Each toss has probability p_i of resulting in i. Let X_i be the number of tosses that result in i. Show that X_1 and X_2 are negatively correlated (i.e., a large number of ones suggests a smaller number of twos).

Problem 34. Let $X = Y - Z$ where Y and Z are nonnegative random variables such that $YZ = 0$.

(a) Show that $\text{cov}(Y, Z) \leq 0$.

(b) Show that $\text{var}(X) \geq \text{var}(Y) + \text{var}(Z)$.

(c) Use the result of part (b) to show that

$$\text{var}(X) \geq \text{var}\big(\max\{0, X\}\big) + \text{var}\big(\max\{0, -X\}\big).$$

Problem 35. Consider two random variables X and Y. Assume for simplicity that they both have zero mean.

(a) Show that X and $\mathbf{E}[X \mid Y]$ are positively correlated.

(b) Show that the correlation coefficient of Y and $\mathbf{E}[X \mid Y]$ has the same sign as the correlation coefficient of X and Y.

SECTION 4.6. Least Squares Estimation

Problem 36. The continuous random variables X and Y have a joint PDF given by

$$f_{X,Y}(x, y) = \begin{cases} c, & \text{if } (x, y) \text{ belongs to the shaded region} \\ 0, & \text{otherwise.} \end{cases}$$

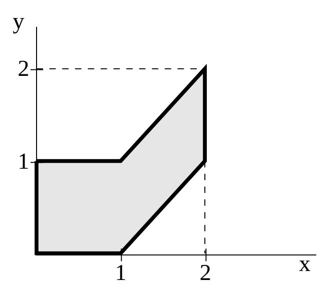

Figure 4.6.1:

(a) Find the least squares estimate of Y given that $X = x$, for all possible vaalues x.

(b) Let $g^*(x)$ be the estimate from part (a), as a function of x. Find $\mathbf{E}[g^*(X)]$ and $\text{var}(g^*(X))$.

(c) Find the mean square error $\mathbf{E}[(Y - g^*(X))^2]$. Is it the same as $\mathbf{E}[\text{var}(Y \mid X)]$?

(d) Find $\text{var}(Y)$.

Problem 37. We are given that $\mathbf{E}[X] = 1$, $\mathbf{E}[Y] = 2$, $\mathbf{E}[X^2] = 5$, $\mathbf{E}[Y^2] = 8$, and $\mathbf{E}[XY] = 1$. Find the linear least squares estimator of Y given X.

Problem 38. In a communication system, the value of a random variable X is transmitted, but what is received (denoted by Y) is the value of X corrupted by some additive noise W; that is, $Y = X + W$. We know the distribution of X and W, and let us assume that these two random variables are independent and have the same PDF. Calculate the least squares estimate of X given Y. What happens if X and W are dependent?

Problem 39. Consider three zero-mean random variables X, Y, and Z, with known variances and covariances. Give a formula for the linear least squares estimator of X based on Y and Z, that is, find a and b that minimize

$$\mathbf{E}\big[(X - aY - bZ)^2\big].$$

For simplicity, assume that Y and Z are uncorrelated.

Problem 40. Provide a new derivation of the formula

$$\mathrm{var}(X) = \mathrm{var}(\hat{X}) + \mathrm{var}(\tilde{X})$$

using the law of total variance. Here, as in the text, $\hat{X} = \mathbf{E}[X \mid Y]$, and $\tilde{X} = X - \hat{X}$.

SECTION 4.7. The Bivariate Normal Distribution

Problem 41. Let U and V be independent standard normal random variables. Let

$$X = U + V, \qquad Y = U - 2V.$$

(a) Do X and Y have a bivariate normal distribution?

(b) Provide a formula for $\mathbf{E}[X \mid Y]$.

(c) Write down the joint PDF of X and Y.