The Basics of Mathematical Statistics

Problem set 1

1. A coin is tossed three times and the sequence of heads and tails is recorded.

    (a) List the sample space.

    (b) List the elements that make up the following events:

    - A = at least two heads
    - B = the first two tosses are heads
    - C = the last toss is a tail.

    (c) List the elements of the following events:

    - $\bar{A}$
    - $A \cap B$
    - $A \cup C$

2. The first three digits of a university telephone exchange are 452. If all the sequences of the remaining four digits are equally likely, what is the probability that a randomly selected university phone number contains seven distinct digits?

3. The four players in a bridge game are each dealt 13 cards. How many ways are there to do this?

4. If you throw a die twice, what is the probability that a 1 will be followed by a 3? What will be the chance of throwing a 3 followed by any other number but a three?

5. If a five-letter word is formed at random (meaning that all sequences of five letters are equally likely), what is the probability that no letter occurs more than once?

6. What is the coefficient of $x^2y^2z^3$ in the expansion of $(x+y+z)^7$?

7. A software development company has three jobs to do. Two of the jobs require three programmers, and the other requires four. If the company employs ten programmers, how many different ways are there to assign them to the jobs?

Problem set 2

1. A family that owns two automobiles is selected at random. Let $A_1 = \{\text{The older car is American}\}$ and Let $A_2 = \{\text{The newer car is American}\}$. If $P(A_1) = 0.7$, $P(A_2) = 0.5$ and $P(A_1 \cap A_2) = 0.4$, compute the following:

   (a) The probability that at least one car is American.

   (b) The probability that neither car is American.

   (c) The probability that exactly one of the two cars is American.

2. Let A denote the event that the next item checked out at a public library is a nonfiction book, and let B be the event that the next item checked out is a work of fiction. Suppose that $P(A) = 0.35$ and $P(B) = 0.50$.

   (a) Why is it not the case that $P(A) + P(B) = 1$?

   (b) Calculate $P(\bar{A})$.

   (c) Calculate $P(A \cup B)$.

   (c) Calculate $P(\bar{A} \cap \bar{B})$.

3. One box contains six red balls and four green balls, and a second box contains seven red balls and three green balls. A ball is randomly chosen from the first box and placed in the second box. Then a ball is randomly selected from the second box and placed in the first box.

   (a) What is the probability that a red ball is selected from the first box and a red ball is selected from the second box?

   (b) At the conclusion of the selection process, what is the probability that the numbers of red and green balls in the first box are identical to the numbers at the beginning?

4. The blood type distribution in the United States is type A, 41%; type B, 9%; type AB, 4%; and type O, 46%. It is estimated that during World War II, 4% of inductees with type O blood were typed having type A; 88% of those with type A were correctly typed; 4% with type B blood were typed as A; and 10% with type AB were typed as A. A soldier was wounded and brought to surgery. He was typed as having a type A blood. What is the probability that this is his true blood type?

Problem set 3

1. A certain company sends 40% of its overnight mail parcels via express mail service E_1. Of these parcels, 2% arrive after the guaranteed delivery time (denote by L the event "late delivery"). Suppose that 50% of the overnight parcels are sent via express mail service E_2 and the remaining 10% are sent via E_3. Of those sent via E_2, only 1% arrive late, whereas 5% of the parcels handled by E_3 arrive late.

   (a) What is the probability that a randomly selected parcel arrived late?

   (b) If a randomly selected parcel has arrived on time, what is the probability that it was not sent via E_1?

2. For each random variable defined here, describe the set of possible values for the variable and state whether the variable is discrete.

   (a) X = The number of unbroken eggs in a randomly chosen standard egg carton

   (b) Y = The number of students on a class list for a particular course who are absent on the first day of classes

   (c) Z = The length of a randomly selected ratlesnake

   (d) U = The amount of royalties earned from the sale of a first edition of 10.000 textbooks

   (e) K = The pH of a randomly chosen soil sample

3. A mail-order computer business has six telephone lines. Let X denote the number of lines in use at a specified time. Suppose the pmf (*probability mass function*) of X is as given in the accompanying table.

x	0	1	2	3	4	5	6
$p(x)$	0.10	0.15	0.20	0.25	0.20	0.06	0.04

   Calculate the probability of each of the following events.

   (a) at most 3 lines are in use

(b) fewer than 3 lines are in use

(c) at least 3 lines are in use

(d) between 2 and 5 lines, inclusive, are in use

4. Which one is more likely: 9 heads in 10 tosses of a fair coin or 18 heads in 20 tosses?

5. The time X (min) for a lab assistant to prepare the equipment for a certain experiment is believed to have a uniform distribution with $A = 25$ and $B = 35$.

    (a) Write the pdf (*probability density function*) of X

    (b) What is the probability that preparation time exceeds 33 min?

    (c) What is the probability that preparation time is within 2 min of the mean time?

Problem set 4

1. A fair coin is thrown once; if it lands heads up, it is thrown a second time. Find the frequency function of the total number of heads.

2. A drug is used to maintain a steady heart rate in patients who have suffered a mild heart attack. Let X denote the number of heartbeats per minute obtained per patient. Consider the hypothetical frequency function given in table below.

x	40	60	68	70	72	80	100
$p(x)$	0.01	0.04	0.05	0.80	0.05	0.04	0.01

   What is the average heart rate obtained by all patients receiving this drug?

3. A white blood cell count of healthy individual can average as low as 6000 per cubic millimeter of blood. To detect a white-cell deficiency, a 0.001 cubic millimeter drop of blood is taken and the number of white cells X is found. How many white cells are expected in a healthy individual? If at most two are found, is there evidence of a white cell deficiency?

4. The distribution of IQ scores in a particular occupation is assumed normal with mean 100 and a standard deviation 10. Find the probabilities that a randomly selected person in this occupation has an IQ

   (a) below 95

   (b) between 105 and 112.

Problem set 5

1. The joint frequency function of two random variables X and Y is given with one omission:

	$x=1$	$x=2$	$x=3$
$y=1$	0.20	0.15	?
$y=2$	0.05	0.15	0.10
$y=3$	0.00	0.10	0.25

   (a) Find the marginal frequency functions of X and Y.

   (b) Compute $E(X)$, $E(X^2)$, $Var(X)$, and σ_X.

   (c) Find the conditional expectations $E(X|Y=1)$, $E(X|Y=2)$, $E(X|Y=3)$.

   (d) Verify the Law of Total Expectation using your answers to a-c.

2. In a certain city, Channel 3 has 50% of the viewing audience on Saturday nights, Channel 12 has 30%, and Channel 10 has 20%. What is the probability that among eight television viewers randomly selected in that city on a Saturday night, five will be watching Channel 3, two will be watching Channel 12, and one will be watching Channel 10?

3. Material strength investigations provide a rich area of application for statistical methods. The article "Effects of Aggregates and Microfillers on the Flexural Properties of Concrete" (*Magazine of Concrete Research*, 1997:81-83) reported on a study of strength properties of high-performance concrete obtained by using superplasticizers and certain binders. The compressive strength of such concrete had previously been investigated, but not much was known about flexural strength (a measure of ability to resist failure in bending). The accompanying data on flexural strength (in MegaPascal, MPa) appeared in the article cited:

   5.9, 7.2, 7.3, 6.3, 8.1, 6.8, 7.0, 7.6, 6.8, 6.5, 7.0, 6.3, 7.9, 9.0,

   8.2, 8.7, 7.8, 9.7, 7.4, 7.7, 9.7, 7.8, 7.7, 11.6, 11.3, 11.8, 10.7

   (a) Calculate a point estimate of the mean value of strength for the conceptual population of all beams manufactured in this fashion. *Hint:* $\sum x_i = 219.8$.

(b) Calculate a point estimate of the population standard deviation. *Hint:* $\sum x_i^2 = 1860.94$.

(c) Calculate a point estimate of the proportion of all such beams whose flexural strength exceeds 10 MPa *Hint:* Think of an observation as a "success" if it exceeds 10.

Problem set 6

1. Suppose $H_0 : p = p_0$ is tested against $H_1 : p > p_0$.

    (a) If the null hypothesis is rejected at the $\alpha = 0.05$ level of significance, will it necessarily be rejected at the $\alpha = 0.01$ level of significance?

    (b) If the null hypothesis is rejected at the $\alpha = 0.01$ level of significance, will it necessarily be rejected at the $\alpha = 0.05$ level of significance?

2. True or false:

    (a) The significance level of a statistical test is equal to the probability that the null hypothesis is true.

    (b) If a test is rejected at the significance level α, the probability that the null hypothesis is true equals α.

    (c) The probability that the null hypothesis is falsely rejected is equal to the power of the test.

    (d) A type II error is more serious than a type I error.

3. Suppose that a 99% confidence interval for the mean μ of a normal distribution is found to be (-2.0, 3.0). Would a test of $H_0 : \mu = -3$ versus $H_1 : \mu \neq= -3$ be rejected at the 0.01 significance level?

4. What values of a chi-square test statistic with 7 degrees of freedom yield a p-value less than or equal to 0.10?

5. Suppose that a test statistic T has a standard normal null distribution.

    (a) If the test rejects for large values of $|T|$, what is the p-value corresponding to $T = 1.50$?

    (b) Answer the same question if the test rejects for large T.

Problem set 7

1. An urn contains two balls. The prior information about the colors is such that three possibilities are considered to be equally likely

$$\begin{aligned} A_0 &= \text{no white balls} \\ A_1 &= \text{one ball is white, the other ball is not white} \\ A_2 &= \text{two white balls} \end{aligned}$$

   Random experiment: place an additional white ball in the urn and choose one ball out of three at random. Suppose the experiment resulted in a white ball = event W.

   (a) Compute the conditional probability $P(W|A_2)$.

   (b) Compute the unconditional probability of the event W.

   (c) Compute the posterior probability $P(A_0|W)$. Explain the meaning of the posterior probability.

2. Suppose that you have a prior distribution for the probability π of success in a certain kind of gambling game which has mean 0.4, and that you regard your prior information as equivalent to 12 trials. You then play the game 25 times and win 12 times. What is your posterior distribution for π?