ACTEX
PUBLICATIONS
The Experts In Actuarial Career Advancement

Product Preview

For More Information: email Support@ActexMadRiver.com
or call 1(800) 282-2839

TABLE OF CONTENTS

CHAPTER 1 COMBINATORIAL PROBABILITY 1

CHAPTER 2 GENERAL RULES OF PROBABILITY 43

CHAPTER 3 DISCRETE RANDOM VARIABLES 89

CHAPTER 4 SOME DISCRETE DISTRIBUTIONS 133

CHAPTER 5 CALCULUS, PROBABILITY, AND CONTINUOUS DISTRIBUTIONS 175

CHAPTER 6 SOME CONTINUOUS DISTRIBUTIONS 233

CHAPTER 7 MULTIVARIATE DISTRIBUTIONS 305

CHAPTER 8 A PROBABILITY POTPOURRI 377

CHAPTER 9 SAMPLING DISTRIBUTIONS AND ESTIMATION 447

CHAPTER 10 HYPOTHESIS TESTING 487

1.3 SAMPLING AND DISTRIBUTION

The best way to master combinatorial (counting) type problems is to completely understand a few of the basic building-blocks, and to learn how seemingly unrelated problems are in fact structurally the same. In this section we provide a structure to formally classify certain combinatorial problems by their particular type.

One common formulation for these types of problems – as we have seen in examples from earlier sections – involves choosing subsets from a set of n distinguishable objects. In statistics, these subsets (or events) are called ***samples***. A different way of formulating problems involves assigning, or ***distributing***, markers to the n objects. It turns out that these two types of problems, *sampling* and *distribution*, are closely related. In fact every sampling problem can be recast as a distribution problem, and vice-versa.

Mathematicians often couch sampling problems in terms of *removing balls from urns*. The related assignment problem would be posed as *distributing balls into urns*. In this sense, the distribution of balls into urns is frequently referred to as *occupancy*. We have never understood the fascination with balls and urns, but the terminology pervades in the classical textbooks and is convenient to use in designing examples.

1.3.1 SAMPLING

We begin with sampling problems. Consider a set consisting of n distinguishable objects (for example, numbered balls in an urn). The set of n objects is called the *population*. Assuming all outcomes are equally likely, the probability model boils down to the question "How many distinct samples of size r can be drawn from the n objects?"

Before we can answer this big question we need the answers to two related questions:

I. Are the samples taken with or without replacement (that is, can we pick an object at most once, or can we pick the same object more than once)?

II. Does the order in which we select the items in the sample matter?

Example 1.3-1 The Four Different Types of Sampling

For each of the four types of sampling described below, calculate the number of distinct outcomes.

(1) A club consisting of 26 members needs to select an executive board consisting of a president, a vice-president, a secretary and a treasurer. Each position has different duties and responsibilities. No individual can hold more than one office.

(2) A club consisting of 26 members needs to select a delegation of 4 (distinct) members to attend a convention. The delegation of 4 wears identical goofy hats.

(3) A club consisting of 26 members requires volunteers to complete 4 distinct chores; sweeping the clubhouse, printing off raffle tickets for the drawing at the party that night, picking up the prizes, and picking up the empties after the party. The same person can volunteer for more than one job, and each job requires only one volunteer.

(4) A club consisting of 26 members requires volunteers to make 4 identical recruiting phone calls, to be chosen later from a long list. A member can volunteer for one, two, three, or all four calls.

We classify these four situations with respect to replacement/without replacement, and order matters/order doesn't matter.

Example	*n*	*r*	Replacement?	Order Matters?
#1	26	4	No	Yes
#2	26	4	No	No
#3	26	4	Yes	Yes
#4	26	4	Yes	No

Or,

	Without replacement	With replacement
Order matters	Example 1	Example 3
Order doesn't matter	Example 2	Example 4

In (1) and (2) the context makes clear we want to select four different people, so in both cases we are sampling without replacement. In (1) the order matters since if Rajulio is selected it makes a difference as to which office he holds. In (2), order doesn't matter. We are only concerned with the members of the delegation, not the order in which they were selected. In (3) and (4), since members can volunteer for multiple chores, we are sampling with replacement. In (3) the chores are different, so we need to keep track of who is volunteering for which job. In (4) the calls are essentially identical, so we are only concerned with how many calls a volunteer makes.

Each member of the club can be uniquely identified (conveniently, with a single letter of the alphabet since there are exactly 26 members). The outcomes for each of these four sampling experiments can be put into one-to-one correspondence with certain types of four letter "words." "Words" (in quotes) means, as previously, any list of four letters, not just dictionary words.

In sampling without replacement, no letter can be repeated in a word. If order matters then ABCD is different from DCBA. If order does not matter, then we agree to list the four letters chosen for the sample just once, in alphabetical order. Thus, we can rephrase our four questions this way:

(1) Calculate the number of four letter words with no duplication of letters.

(2) Calculate the number of four letter words with no duplicate letters and the letters arranged in alphabetical order.

(3) Calculate the number of four letter words, duplicate letters allowed.

(4) Calculate the number of four letter words, duplicate letters allowed, and the letters arranged in alphabetical order.

Solution
The first three are straightforward to work out using permutations, combinations and the basic multiplication principle, respectively.

(1) Answer: $\underset{\text{president}}{\underline{26}} \cdot \underset{\text{vice-pres}}{\underline{25}} \cdot \underset{\text{secretary}}{\underline{24}} \cdot \underset{\text{treasurer}}{\underline{23}} = {}_{26}P_4 = 358{,}800.$

(2) Answer: ${}_{26}C_4 = \dfrac{{}_{26}P_4}{4!} = \dbinom{26}{4} = 14{,}950.$

(3) Answer: $\underset{\text{sweep}}{\underline{26}} \cdot \underset{\text{raffle}}{\underline{26}} \cdot \underset{\text{prizes}}{\underline{26}} \cdot \underset{\text{empties}}{\underline{26}} = 26^4 = 456{,}976.$

Example (4) is a little more subtle and we will defer the solution to the next section where it will be more easily understood as a distribution problem. ❐

1.3.2 DISTRIBUTIONS

Each of the four types of sampling experiments described above has a dual formulation as a *distribution*, or occupancy, experiment. For this purpose, we let n be the number of distinguishable (fixed and labeled) urns, and let r be the number of balls to be distributed into the urns. How many different distributions are possible? Again, before we can answer the question, we need to know:

I. Can an urn hold at most one ball (exclusive) or can it hold many balls (non-exclusive)?

II. Are the balls distinguishable (for example, bearing unique numbers) or are they indistinguishable (like plain white ping-pong balls)?

We can rephrase each of the four sampling problems above as a corresponding distribution problem:

(1) In how many ways can 4 executive board positions (distinguishable balls) be distributed among 26 members (urns) with exclusion (since no member can hold more than one position)?

(2) In how many ways can 4 delegation slots (indistinguishable balls) be distributed among 26 members (urns) with exclusion?

(3) In how many ways can 4 different jobs (distinguishable balls) be distributed among 26 members (urns) without exclusion (since one member can do multiple jobs)?

(4) In how many ways can 4 identical jobs (indistinguishable balls) be distributed among 26 members (urns) without exclusion (since one member can do multiple jobs)?

Statements (1)-(3) were solved in Section 1.3.1 formulated as equivalent sampling problems. Here is a paradigm for counting in sampling or distribution problems like (4).

Imagine the 26 urns lined up in a row sharing common partitions:

A	B	C	$\cdots$	Z

We now distribute the 4 indistinguishable balls among the 26 urns. A sample outcome might be:

$\circ$	$\circ\circ$			$\circ$
A	B	C	$\cdots$	Z

Now, observe that the top row in the above schematic can be thought of as a word consisting of (26 + 1) lines and 4 circles. However, the leftmost line and the rightmost line remain fixed in every word and are therefore superfluous. A moment's thought will show that there is a one-to-one correspondence between distinct distributions into the 26 slots, and distinct words consisting of $(26-1)$ lines and 4 circles. The distribution above appears as

$$\circ|\circ\circ|||\cdots|\circ\,.$$

Thus, the question is reduced to, "How many $(26-1+4)$ letter words are there consisting of four circles and $(26-1)$ vertical lines?" Therefore, the solution to example (4) is

$$\binom{26-1+4}{4} = \binom{29}{4} = 23{,}751.$$

The above discussion leads us to the following result:

Samples with Replacement When Order Does Not Matter

The number of unordered samples of r objects, with replacement, from n distinguishable objects is

$${}_{n+r-1}C_r = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}.$$

This is equivalent to the number of ways to distribute r indistinguishable balls into n distinguishable urns without exclusion.

The same reasoning works in general, so that ${}_{n+r-1}C_r$ is the solution to the following sampling and distribution problems:

Sampling: How many distinct unordered samples of size r, with replacement, are there from n distinguishable objects?

Distribution: How many distinct ways are there to distribute r indistinguishable balls into n distinguishable urns, with multiple balls in an urn allowed?

We like to think of these types of problems in terms of distributing r dollar bills to n children. One needs $n-1$ partitions to separate the children, and then select the locations to place the r dollars.

Example 1.3-2 Samples with Replacement When Order Does Not Matter

Nicole wishes to select a dozen bagels from *Unleavened Bread Company*. Her choices include: Asiago cheese, plain, nine grain, cinnamon crunch, and very-very blueberry.

(a) How many different orders of a dozen bagels can she select?

(b) How many different orders of a dozen bagels can she select in which she has at least one of each kind?

Solution

Selecting a dozen bagels is equivalent to distributing 12 indistinguishable markers (ordering a bagel) into 5 (bagel) bins. Our answer to (a) is

$$_{(5-1+12)}C_{12} = {}_{16}C_{12} = 1{,}820.$$

For part (b), imagine that we first select one of each type of bagel. Our problem reduces to selecting the remaining seven bagels from any the five types. This can be done in

$$_{(5-1+7)}C_{7} = {}_{11}C_{7} = 330 \text{ ways.}$$ ❐

The "Boston Chicken" example, below, illustrates how identifying the particular type of sample can be very important. Boston Chicken has 16 distinct side-dishes (the population). Each dinner is served with the customer's choice of three side-dishes. Read the following article and explain the advertisement that there were "more than 3000 combinations" possible.

Thurs, Jan. 26, 1995. Rocky Mountain News, Retail & Marketing section

Catching Boston Chicken's error made Bob Swaim king for a week

Ad helps teacher prove math's important

By Lynn Bronikowski
Rocky Mountain News Staff Writer

Nearly every day, Pennsylvania high school teacher Bob Swaim harps about the importance of math skills.

Some students yawn; others chow down on chicken — the reward for Swaim's attention to detail in the world of television advertising.

The Golden-based Boston Chicken rotisserie chicken restaurant chain earlier this month aired an ad for its combo platter in which it said a customer could get more than 3,000 combinations by ordering a three-item combo.

But Swaim, who has taught math for 28 years, caught Boston Chicken with egg on its face after pointing out there are actually only 816 *combinations* but more than 3,000 *permutations*.

Boston Chicken, which corrected the ad by last weekend, won a flood of publicity by admitting the goof. Stories appeared locally and in *USA Today, The Toronto Star, Orange County Register, Chicago Sun Times, Cincinnati Post* and the Associated Press.

National Public Radio picked up the story and producers from TV network shows made calls.

Swaim, 48, became a small-town hero in Souderton, Pa. Boston Chicken donated $500 to the math department and treated 30 trigonometry students to lunch at a nearby Boston Chicken restaurant last Friday.

Swaim never actually saw the commercial starring Joe Montana, giving credit to his 15-year-old daughter, Joann, who first spotted the miscalculation.

"Everyday I'm telling my students that math is important," said Swaim, who teaches two trig classes, a geometry class and math for daily living. "I felt I had a moral responsibility to call about the mistake."

Swaim said at first he got a runaround before finally talking to the copywriter at the Integer Group in Golden.

He was surprised no one else had called. "Their phones should have been ringing off the hook," said Swaim. "There should have been more people than just me (questioning) it.

"It points out some of the weaknesses in our culture. It was such a simple problem and I'm no whiz kid."

He figures a second-year algebra student could have worked the formula.

"We're not used to this happening; Souderton is such a small town," said Jacey Stroback, a junior in one of Swaim's two trig classes. "He's always talking about how we're going to use math skills later on in life, no matter what you do."

Even before the math encounter, Swaim was a Boston Chicken fan, having eaten the restaurant's food at home. And after spotting a flier in a nearby restaurant, he invited a speaker to talk to two classes about career opportunities.

"I'm always looking for people to promote positive education and I liked their attitude," said Swaim.

After his fleeting moment of fame, life has returned to normal. But not before the producers of Tom Snyder's late-night talk show called about an appearance.

"I got bumped by a guy who collects stray birds."

Permutated veggies: a reworked recipe

Boston Chicken made its mistake by confusing combinations and permutations. One combination can count as several permutations.

BOSTON ROTISSERIE CHICKEN

Corn-corn-mashed potato and corn-mashed potato-corn are two different *permutations*, but you still wind up with the same *combination* of foods.

The chain provided this corrected formula for calculating the number of possible three-item samplers that can be made from 16 Boston Chicken side items.

Add up all of the following scenarios:

■ Selecting all three of the same items. For example, corn-corn-corn. Equals 16 ways.

■ Selecting two of the same items and a third different item. For example, corn-corn-mashed potato. The first one in 16 ways; second one in 15 ways; third one in two ways (as you selected just one of the two already selected); remove combinations that are the same. Equals 240 ways.

■ Selecting all three different items. For example, corn-stuffing-mashed potato. First one in 16 ways; second one in 15 ways; third one in 14 ways; remove combinations that are the same. Equals 560 ways.

■ Total all of the above number of ways. Equals 816 ways.

Exercise 1-39 Calculate the number of possibilities under ordered samples without replacement (customers eat their side dishes in the order of selection and must choose 3 different dishes).

Exercise 1-40 Calculate the number of possibilities using unordered samples without replacement (customers can eat their dishes in any order but still must choose 3 different dishes).

Exercise 1-41 Calculate the number of possibilities using ordered samples with replacement (customers eat their side dishes in a definite order, but can order, for example, corn-corn-mashed potato, which is different from corn-mashed potato-corn.

Exercise 1.42 Calculate the number of possibilities using unordered samples with replacement (customers eat their side dishes in any order, and corn-corn-mashed potato is possible and is identical to corn-mashed potato-corn.

Exercise 1-43 Which type of sampling gives an answer closest to “more than 3000?”

Exercise 1-44 Which type of sampling provides the most realistic result?
Hint: The correct answer of 816 is given in the side-bar of the article, although it is worked out differently using multiple steps.

Exercise 1-45 Assume that Boston Chicken customers choose three side-dishes from the 16 possible in such a way that all unordered samples with replacement are equally likely. What is the probability that a customer will choose all three side-dishes the same? That is, in poker parlance, what is the probability of three-of-a-kind?

Exercise 1-46
(a) How many ways can a parent distribute five one-dollar bills to her three children?
(b) How many ways can she accomplish this if each child gets at least one dollar?

Exercise 1-47
(a) How many ways can a witch distribute ten candy bars and seven packages of gum to four trick-or-treaters?
(b) How many ways can she do this if each child receives at least one candy bar and one package of gum?

Exercise 1-48 How many ways may a parent distribute ten identical pickled beets to his five children?

Exercise 1-49 How many different 13-card bridge hands are possible?

Exercise 1-50 How many 5-card poker hands are there?

Exercise 1-51 At a local fast-food restaurant in Oregon (no sales tax), fries, soda, hamburgers, cherry pie, and sundaes cost $1 each. Chicken sandwiches cost $2 each. You have five dollars. How many different meals can you order?

Exercise 1-52 I have fifteen certificates for a free pizza and 24 cans of Coca-Cola®. How many ways may I distribute the certificates and the cans of coke to 22 students?

1.3.3 SAMPLING AND OCCUPANCY UNITED

The following diagram shows the complete set of correspondences between sampling and distribution. Although the formulation of a sampling problem may appear to be quite dissimilar to the corresponding distribution problem, the two are in fact mathematically equivalent.

Samples of size *r* from *n* distinguishable objects	**Without replacement**	**With replacement**	
Order matters	${}_nP_r$ (Example 1)	n^r (Example 3)	Distinguishable balls
Order doesn't matter	$\binom{n}{r}$ (Example 2)	$\binom{n+r-1}{r}$ (Example 4)	Indistinguishable balls
	Exclusive	Non-exclusive	Distributions of *r* balls into *n* distinguishable urns

1.4 MORE APPLICATIONS

In this section we bring together a variety of special applications of the combinatorial methods previously discussed. We begin with an illustration of some basic results involving the algebra of polynomials.

1.4.1 THE BINOMIAL AND MULTINOMIAL THEOREMS

Consider the problem of expanding the binomial expression

$$(x+y)^n = \underbrace{(x+y)}_{\text{factor 1}}\underbrace{(x+y)}_{\text{factor 2}}\underbrace{(x+y)}_{\text{factor 3}}\cdots\underbrace{(x+y)}_{\text{factor } n}.$$

Algebraically, this amounts to adding together all possible products consisting of n letters, the first selected from "factor 1," the second from "factor 2," and so forth up to "factor n." In other words, the expansion consists of the sum of all n letter "words" consisting of the letters x and y.

To simplify our final result, we group together all words that contain the same number of x's and y's. Also, a stickler for algebra would insist on proper exponent notation (and the fact that multiplication is commutative), to express words like *xxyxyyx* and *yxxxyxy* (with 4 x's and 3y's) all as x^4y^3.

Now, for a given r $(r = 0,1,2,\ldots,n)$, the ***number*** of distinct n letter words with (r) x's and $(n-r)$ y's will be the coefficient of $x^r \cdot y^{n-r}$ in the expansion of $(x+y)^n$. As we have previously seen, this is just the combination ${}_nC_r = \binom{n}{r}$. This is the reason that the expression "binomial coefficient" is used to describe the coefficient $\binom{n}{r}$ of the term $x^r \cdot y^{n-r}$ of the binomial expansion for $(x+y)^n$.

The formula in the Binomial Theorem now follows from this observation.

The Binomial Theorem

For every non-negative integer n and real numbers x and y, we have

$$\begin{aligned}(x+y)^n &= \sum_{r=0}^{n} {}_nC_r \cdot x^r \cdot y^{n-r} \\ &= {}_nC_0 \cdot y^n + {}_nC_1 \cdot x^1 \cdot y^{n-1} + {}_nC_2 \cdot x^2 \cdot y^{n-2} + \cdots + {}_nC_n \cdot x^n,\end{aligned}$$

or equivalently,

$$\begin{aligned}(x+y)^n &= \sum_{r=0}^{n} {}_nC_r \cdot x^{n-r} \cdot y^r \\ &= {}_nC_0 \cdot x^n + {}_nC_1 \cdot x^{n-1} \cdot y^1 + {}_nC_2 \cdot x^{n-2} \cdot y^2 + \cdots + {}_nC_n \cdot y^n\end{aligned}$$

Note
The two forms of the theorem given are in fact equivalent since ${}_nC_r = {}_nC_{n-r}$ is the coefficient of both $x^r \cdot y^{n-r}$ and $x^{n-r} \cdot y^r$.

Example 1.4-1 The Binomial Theorem

Use the Binomial Theorem to expand $(x-2y)^3$.

Solution

To apply the binomial theorem, we rewrite $(x-2y)^3 = (x+\{-2y\})^3$.

$$\begin{aligned}(x-2y)^3 &= {}_3C_0 \cdot x^3 \cdot (-2y)^0 + {}_3C_1 \cdot x^2 \cdot (-2y)^1 + {}_3C_2 \cdot x^1 \cdot (-2y)^2 + {}_3C_3 \cdot x^0 \cdot (-2y)^3 \\ &= 1 \cdot x^3 \cdot 1 + 3 \cdot x^2(-2y) + 3 \cdot x \cdot (4y^2) + 1 \cdot 1 \cdot (-8y^3) = x^3 - 6x^2y + 12xy^2 - 8y^3\end{aligned}$$ ❒

Exercise 1-53 Use the Binomial Theorem to expand the following:

(a) $(x+3)^4$

(b) $(2x+y)^5$

(c) $(4x-5y)^3$

Exercise 1-54 Use the mathematical definition of combinations to verify the identity:

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}.$$

Note 1

For a *non*-algebraic proof, imagine that you belong to a sorority with n members, exactly r of whom get to go to a party at the fraternity next door. How many ways to select the lucky r members? Break it down into those combinations that contain you, and those combinations that do not contain you. Are you one of the party-goers? If so, it remains to choose $r-1$ members from the $n-1$ members who are not you. If you stay home completing your probability homework, then we must count the ways r members are chosen from the $n-1$ members who are not you.

Note 2

This identity is what makes the generation of Pascal's[3] triangle possible. The n^{th} row of Pascal's triangle contains the binomial coefficients for expanding $(x+y)^n$. Each new row is generated by adding the adjacent coefficients (as in the identity above) from the previous row:

Pascal's Triangle

$(x+y)^0$	1	0^{th} row
$(x+y)^1$	1 1	1^{st} row
$(x+y)^2$	1 \ / 2 1	2^{nd} row
$(x+y)^3$	1 3 3 1	3^{rd} row
$(x+y)^4$	1 4 6 4 1	4^{th} row

[3] Pascal (1623-1662), considered to be, along with his contemporary, Pierre de Fermat (1601-1665), a progenitor of modern probability theory.

Exercise 1-55 Use the binomial theorem to verify the identity $\sum_{k=0}^{n} {}_nC_k = 2^n$. For example, if $n = 4$, then $\sum_{k=0}^{4} {}_4C_k = {}_4C_0 + {}_4C_1 + {}_4C_2 + {}_4C_3 + {}_4C_4 = 16 = 2^4$.

Note
${}_nC_k$ is the number of $k-$topping pizzas that can be made if there are n toppings from which to select. You may create a zero-topping pizza, or a one-topping pizza, up to an n-topping pizza. The left-hand side of the equation is the sum of different numbers of k-topping pizzas. On the other hand, each pizza topping may either be placed on the pizza or not. There are two ways to do this (put topping on pizza or do not put topping on pizza). There are n toppings from which to select. By the multiplication principle, there are $\underbrace{2 \cdot 2 \cdot \cdots \cdot 2}_{n \text{ times}} = 2^n$ ways to do this, which is the right-hand side of the identity.

Exercise 1-56 Explain the fundamental identity $\binom{n}{r} = \binom{n}{n-r}$ in terms of pizza toppings.

Exercise 1-57 In the expansion of $(x+y)^n$ the coefficient of x^4y^{n-4} is 3,876 and the coefficient of x^5y^{n-5} is 11,628. Find the coefficient of x^5y^{n-4} in the expansion of $(x+y)^{n+1}$. What is the value of n?

It is possible to generalize the arguments involving the binomial theorem in order to demonstrate the multinomial theorem. We wish to expand out the expression

$$(x_1 + x_2 + \cdots + x_r)^n = \underbrace{(x_1 + x_2 + \cdots + x_r)}_{\text{factor 1}}\underbrace{(x_1 + x_2 + \cdots + x_r)}_{\text{factor 2}}\cdots\underbrace{(x_1 + x_2 + \cdots + x_r)}_{\text{factor } n}.$$

The expansion now consists of all possible n-letter words consisting of the letters $x_1, x_2, \ldots, x_r$. We again group together all words with the same exponents on the letters. How many words are there with (n_1) x_1's, (n_2) x_2's, (n_3) x_3's, …, (n_r) x_r's (think MISSISSIPPI)? The answer is the number of partitions of n objects into subsets of sizes $n_1, n_2, n_3, \ldots, n_r$, that is $\binom{n}{n_1, n_2, \cdots, n_r}$. Therefore, the multinomial expansion can be expressed as follows:

The Multinomial Theorem

$$\left(x_1 + x_2 + \cdots + x_r\right)^n = \sum_{n_1 + \cdots + n_r = n} \binom{n}{n_1, n_2, \cdots, n_r} \cdot x_1^{n_1} \cdot x_2^{n_2} \cdot \cdots \cdot x_r^{n_r}.$$

The sum runs over all possible partitions $n_1, n_2, n_3, \ldots, n_r$ such that $n_1 + n_2 + n_3 + \cdots + n_r = n$. Finding all of these partitions is the challenging part.

Note
The number of such partitions is the number of ways of distributing n indistinguishable balls into r urns non-exclusively, that is, $\binom{n+r-1}{n}$.

Example 1.4-2 The Multinomial Theorem

Use the multinomial theorem to expand $(x+y+z)^4$.

Solution
You should take the time to find all of the ways to combine x, y, and z so that the total powers sum to 4.

$$\begin{aligned}(x+y+z)^4 &= \binom{4}{4,0,0}x^4y^0z^0 + \binom{4}{3,1,0}x^3y^1z^0 + \cdots + \binom{4}{0,0,4}x^0y^0z^4 \\ &= x^4 + 4x^3y + 4x^3z + 6x^2y^2 + 6x^2z^2 + 12x^2yz + 4xy^3 + 12xy^2z \\ &\qquad + 12xyz^2 + 4xz^3 + y^4 + 4y^3z + 6y^2z^2 + 4yz^3 + z^4\end{aligned}$$

Note that the number of terms is 15, the same as $\binom{n+r-1}{n} = \binom{4+3-1}{4}$. ❐

Exercise 1-58 Use the multinomial theorem to expand the following:

(a) $(x-2y+5z)^3$

(b) $(w+x-y+2z)^2$.

1.4.2 POKER HANDS

Our simple version of ***poker*** is played with a standard four-suit, 52-card deck. We are serious players, so there are neither jokers nor wild cards in our deck. A poker hand consists of 5 cards dealt from a standard deck. In other words, a poker hand is an unordered random sample of size 5 chosen from a population of size 52, without replacement (you wouldn't want to be caught in Dodge City with 2 Queens of Hearts in your hand). The ace can be played as either high or low, as explained below. We present the definitions of the various types of poker hands.

Straight flush: Five cards of the same suit in sequence, such as 7♥6♥5♥4♥3♥. The Ace-King-Queen-Jack-Ten (A♣K♣Q♣J♣T♣) is called a ***royal flush***. The ace can also play low so that 5♣4♣3♣2♣A♣ is another straight flush.

***Four-of-a-kind*:** Four cards of the same denomination accompanied by another card, like 7♣7♦7♥7♠9♥.

***Full house* (*a.k.a.* a *boat*):** Three cards of one denomination accompanied by two of another, such as Q♣Q♦Q♥4♠4♥.

***Flush*:** Five cards of the same suit, such as K♠Q♠9♠6♠4♠. Straight flushes are excluded (they form their own category above).

***Straight*:** Five cards in sequence, such as J♥T♦9♣8♦7♠. The ace plays either high or low, but a collection like 32AKQ is not allowed. That is, there are no wrap-around straights. Again, straight flushes are excluded, since they have been counted separately.

***Three-of-a-kind*:** Three cards of the same denomination and two other cards of different denominations, such as 7♠7♣7♦K♣2♦.

***Two Pair*:** Two cards of one denomination, two cards of another denomination and a fifth card of a third denomination, such as K♠K♣8♦8♠7♥.

***One Pair*:** Two cards of one denomination accompanied by three cards of different denominations, such as T♠T♣Q♦8♠7♥.

***High Card* (*a.k.a.* Nothing):** Any hand that does not qualify as one of the hands above, such as A♠Q♣9♦8♠7♥.

Various combinatorial techniques are employed to calculate the probabilities of being dealt these hands on an initial deal from a standard deck of cards. We illustrate several of these calculations and provide a summary of the probabilities of all types of poker hands.

Since poker hands consist of five cards (any order) selected without replacement from a population of size 52 cards, the size of the sample space is $\binom{52}{5} = 2{,}598{,}960.$

Example 1.4-3 Straight Flush

Compute the probability of being dealt a straight flush.

Solution

The highest ranked straight flush is AKQJT in one of the four suits (clubs, diamonds, hearts, and spades), the lowest ranked straight flush is 5432A. There are 10 of these rankings. There are four suit choices. By the multiplication rule, there are $10 \cdot 4 = 40$ possible straight flushes.

$$\begin{aligned}\Pr(\text{straight flush}) &= \frac{\text{number of straight flushes}}{\text{number of five card poker hands}} \\ &= \frac{10\cdot 4}{\binom{52}{5}} = \frac{40}{2{,}598{,}960} = .00001539.\end{aligned}$$ ❒

Example 1.4-4 Full House

Compute the probability of being dealt a full house on the initial deal of 5 cards.

Solution

We begin by selecting a card denomination (Ace, King,..., 2) for the three-of-a-kind. One has 13 ways to do this. We then select three of the four cards from the selected denomination to create the three-of-a-kind. For the pair, there are only 12 remaining denominations to select, since the card value chosen for the three-of-a-kind is no longer available. We need to choose two cards with this second denomination.

$$\begin{aligned}\Pr(\text{full house}) &= \frac{\text{number of full houses}}{\text{number of five card poker hands}} \\ &= \frac{\binom{13}{1}\cdot\binom{4}{3}\cdot\binom{12}{1}\cdot\binom{4}{2}}{\binom{52}{5}} = \frac{13\cdot 4\cdot 12\cdot 6}{2{,}598{,}960} = .00144058.\end{aligned}$$ ❒

Exercise 1-59 Compute the probabilities for all nine poker hand types. Do this by yourself prior to looking at the following summary. It is an important exercise, even if you struggle.

Poker Probability Summary:

1. **Straight Flush**

$$\begin{aligned}\Pr(\text{Straight Flush}) &= \frac{\text{number of straight flushes}}{\text{number of five card poker hands}} \\ &= \frac{4\text{ suits}\cdot 10\text{ possible straights}}{\binom{52}{5}} = \frac{40}{2{,}598{,}960} = .00001538.\end{aligned}$$

2. **Four-of-a-Kind**

$$\Pr(\text{Four-of-a-Kind}) = \frac{\binom{13}{1}\cdot\binom{4}{4}\cdot\binom{12}{1}\cdot\binom{4}{1}}{\binom{52}{5}} = \frac{13\cdot 1\cdot 12\cdot 4}{2{,}598{,}960} = \frac{624}{2{,}598{,}960} = .00024010.$$

3. **Full House**

$$\Pr(\text{Full House}) = \frac{\binom{13}{1}\cdot\binom{4}{3}\cdot\binom{12}{1}\cdot\binom{4}{2}}{\binom{52}{5}} = \frac{13\cdot 4\cdot 12\cdot 6}{2{,}598{,}960} = \frac{3{,}744}{2{,}598{,}960} = .00144058.$$

4. **Flush**

$$\Pr(\text{Flush}) = \frac{\overbrace{\binom{4}{1}}^{\text{suits}}\cdot\overbrace{\binom{13}{5}}^{\text{5 cards selected from the suit}} - \overbrace{40}^{\text{straight flushes already accounted}}}{\binom{52}{5}}$$

$$= \frac{4\cdot 1{,}287 - 40}{2{,}598{,}960} = \frac{5{,}108}{2{,}598{,}960} = .00196540.$$

5. **Straight**

$$\Pr(\text{Straight}) = \frac{10\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1} - 40}{\binom{52}{5}}$$

$$= \frac{10\cdot 4^5 - 40}{2{,}598{,}960} = \frac{10{,}200}{2{,}598{,}960} = .00392465.$$

6. **Three-of-a-Kind**

$$\Pr(\text{Three-of-a-Kind}) = \frac{\binom{13}{1}\cdot\binom{4}{3}\cdot\binom{12}{2}\cdot\binom{4}{1}\cdot\binom{4}{1}}{\binom{52}{5}}$$

$$= \frac{13\cdot 4\cdot 66\cdot 4\cdot 4}{2{,}598{,}960} = \frac{54{,}912}{2{,}598{,}960} = .02112846.$$

7. **Two Pair**

$$\Pr(\text{Two Pair}) = \frac{\binom{13}{2}\cdot\binom{4}{2}\cdot\binom{4}{2}\cdot\binom{11}{1}\cdot\binom{4}{1}}{\binom{52}{5}}$$

$$= \frac{78\cdot 6\cdot 6\cdot 11\cdot 4}{2{,}598{,}960} = \frac{123{,}552}{2{,}598{,}960} = .04753902.$$

8. **One Pair**

$$\Pr(\text{One Pair}) = \frac{\binom{13}{1}\cdot\binom{4}{2}\cdot\binom{12}{3}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}}{\binom{52}{5}}$$

$$= \frac{13\cdot 6\cdot 220\cdot 4^3}{2{,}598{,}960} = \frac{1{,}098{,}240}{2{,}598{,}960} = .42256903$$

9. **Nothing**

$$\Pr(\text{Nothing}) = 1 - \Pr(\text{other possibilities})$$

$$= \frac{\binom{52}{5} - (40 + 624 + 3{,}744 + 5{,}108 + 10{,}200 + 54{,}912 + 123{,}552 + 1{,}098{,}240)}{\binom{52}{5}}$$

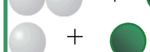
$$= \frac{2{,}598{,}960 - 1{,}296{,}420}{2{,}598{,}960} = \frac{1{,}302{,}540}{2{,}598{,}960} = .50117739.$$

Exercise 1-60 (Poker Dice): Play poker using 5 fair dice rather than a deck of cards. Roll the five dice onto the table. The possible *hands* are: five-of-a-kind, four-of-a-kind, a full house, three-of-a-kind, two pair, one pair, a straight, and nothing. Find the probabilities of these events on a single roll of the dice. Where should the straight ***rank***? The less likely (lower probability) an event is to happen, the higher it should rank.

1.4.3 THE POWERBALL® LOTTERY

The following rules, prizes, and odds were once found at www.musl.com. To play the game, we draw five balls out of a drum with 53 numbered white balls, and one power ball out of a drum with 42 numbered green balls.

Powerball® Prizes and Odds

Match	Prize	Odds
	Grand Prize	1 in 120,526,770.00
	$100,000	1 in 2,939,677.32
	$5,000	1 in 502,194.88
	$100	1 in 12,248.66
	$100	1 in 10,685.00
	$7	1 in 260.61
	$7	1 in 1696.85
	$4	1 in 123.88
	$3	1 in 70.39

The overall odds of winning a prize are 1 in 36.06.
The odds presented here are based on a $1 play and are rounded to two decimal places.

CHAPTER 7

MULTIVARIATE DISTRIBUTIONS

Often we wish to study several random variables simultaneously. For example, throughout the book we have been making use of sums of two or more random variables. We first encountered this in Section 3.6, where we worked with several different random variables arising from the same probability experiment. In this chapter we present a more complete framework for studying these, so-called, ***multivariate distributions*****,** alternatively called ***jointly distributed*** random variables.

Here, we will take a more general approach than in Chapter 3, allowing us to transcend our earlier reliance on a common underlying probability experiment. An important consideration in our previous encounters with multiple random variables involved independence. We might, for example, be looking at summing random variables arising from observing independent repetitions of the same experiment. An important subtext for our studies here involves the interaction between the random variables, which can range from independence to total dependence. We will be presenting the machinery for quantifying and calculating this.

We will begin first with joint distribution tables for discrete distributions - first introduced in Section 3.6 - defining the concepts of marginal and conditional distributions. We then proceed to adapt these concepts to the case of jointly distributed continuous random variables. While the concepts we encounter can be applied to any number of jointly distributed random variables, we will concentrate mainly on pairs of random variables. Similar results for higher dimensions can be derived by extension of the two-dimensional case.

7.1 JOINT DISTRIBUTIONS FOR DISCRETE RANDOM VARIABLES

In Chapter 3 we defined discrete random variables and generally portrayed their distributions in a tabular format. In Section 3.6 we saw that two discrete random variables arising from the same experiment could be displayed in a 2-dimensional table. For our purposes here we will take the 2-dimensional table, with suitable restrictions, as the definition for a discrete joint distribution of two random variables.

Joint Probability Distribution

Let R be a set of ordered pairs in the plane of the form (x_i, y_j); $i = 1, \ldots, m$, $j = 1, \ldots, n$. Let p be a function on R satisfying:

(1) $p(x_i, y_j) \geq 0$ for all pairs (x_i, y_j).

(2) $\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i, y_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} p(x_i, y_j) = 1.$

Let X be the random variable taking on values (x_i); $i = 1, \ldots, m$, with probability distribution given by $p_X(x_i) = \sum_{j=1}^{n} p(x_i, y_j)$ and let Y be the random variable taking values (y_j); $j = 1, \ldots, n$, with probability distribution given by $p_Y(y_j) = \sum_{i=1}^{m} p(x_i, y_j)$.

Then X and Y are said to be ***jointly distributed*** on R with probability distribution given by

$$\Pr[X = x_i,\ Y = y_j] = p(x_i, y_j).$$

The individual random variables X and Y are referred to as the ***marginal distributions***

Notes

(1) Properties (1) and (2) assure that the joint distribution is in fact a probability distribution.

(2) Property (2) also assures that the individual probability distributions for X and Y each sum to one. This is because, in the case of X,

$$\sum_{i=1}^{m} p_X(x_i) = \sum_{i=1}^{m}\left[\sum_{j=1}^{n} p(x_i, y_j)\right] = \sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i, y_j) = 1.$$

A similar argument holds for Y.

(3) It is possible for m and/or n to be infinite, as might be the case in dealing with, for example, Poisson or geometric random variables.

The information for a discrete joint distribution can be neatly summarized in tabular form as follows:

		Y: y_1	$\cdots$	y_n	$p_X(x)$
X	x_1	$p(x_1, y_1)$		$p(x_1, y_n)$	$p_X(x_1)$
	$\vdots$				
	x_m	$p(x_m, y_1)$		$p(x_m, y_n)$	$p_X(x_m)$
	$p_Y(y)$	$p_Y(y_1)$		$p_Y(y_n)$	1

The pairs (x_i, y_j) in R constitute the sample space, or outcomes, for a probability experiment. The dark green cells in the table contain the individual probabilities of the outcomes and constitute the joint probability distribution. The marginal distribution $p_X(x_i) = \sum_{j=1}^{n} p(x_i, y_j)$ for X is tabulated along the right margin, as the sums of the probabilities in the corresponding rows. Similarly, the marginal distribution for Y is displayed along the bottom margin, with $p_Y(y_j) = \sum_{i=1}^{m}(x_i, y_j)$, the sum along the corresponding column.

Example 7.1-1 Discrete Joint Distributions

At Simple Choices Insurance Company, customers choose from 3 levels of automobile liability coverage: 1, 3, or 5 million (state law requires at least 1 million). They may also choose from 3 levels of personal liability coverage: 0, 1, or 4 million. Let X be the level of auto liability selected (in millions) and let Y be the level of personal liability selected (again, in millions). Then (X, Y) is a joint probability distribution with $p(x, y)$ representing the probability that a customer chooses the levels (x, y) of coverage. The joint distribution is shown in the following table:

		Y		
		0	**1**	**4**
X	**1**	.10	.05	.15
	3	.05	.20	.25
	5	.15	.00	.05

(a) Verify that this is in fact a joint distribution table.
(b) Calculate $\Pr(X = 3, Y = 4)$.
(c) Write the marginal distributions for X and Y.

Solution

		Y			
		0	**1**	**4**	$p_Y(y)$
X	**1**	.10	.05	.15	**.30**
	3	.05	.20	.25	**.50**
	5	.15	.00	.05	**.20**
	$p_Y(y)$	.30	.25	.45	1.00

X	$p_Y(y)$
1	**.30**
3	**.50**
5	**.20**
	1.00

Y	$p_Y(y)$
0	.30
1	.25
4	.45
	1.00

(a) We note that all the $p(x_i, y_j) \geq 0$ and sum to 1.00 (lower right hand corner).

(b) $p(3,4) = \Pr(X = 3, Y = 4) = .25$. The marginal distribution for X is obtained by summing the rows, while the marginal for Y is obtained by summing the columns. Marginal distributions can be displayed individually as well, as shown alongside the joint table. ❒

The calculations for expected values, moments, and more general functions of X and Y are straightforward, stemming from the following procedure:

Calculating Expected Values

Let X and Y be jointly distributed with probability distribution

$$p(x_i, y_j);\ i = 1, \ldots\ldots, m,\ \ j = 1, \ldots, n.$$

If $f(x, y)$ is a real-valued function, then

$$E[f(X,Y)] = \sum_{i=1}^{m}\sum_{j=1}^{n} f(x_i, y_j)\,p(x_i, y_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} f(x_i, y_j)\,p(x_i, y_j).$$

It is easy to show that if $f(x, y)$ is a function of either X alone or Y alone, then the expected value can be calculated using the appropriate marginal with the same result.

Example 7.1-2 Expectations with Discrete Joint Distributions

A fair coin is tossed twice. Let X be the number of heads on the first toss. Let Y be the number of heads on the first two tosses.

(a) Calculate the joint distribution table for X and Y.
(b) Using the joint table, calculate the quantities, $E[X]$, $Var[X]$, $E[Y]$, and $Var[Y]$.
(c) Calculate the marginal distributions for X and Y. Use them to recalculate the quantities in (b).
(d) Calculate $E[X+Y]$ and $Var[X+Y]$.
(e) Show that $E[X+Y] = E[X] + E[Y]$, but that $Var[X+Y] \neq Var[X] + Var[Y]$.

Solution

(a) The joint distribution table is easily calculated based on the four equally likely outcomes to the underlying experiment of tossing a coin twice:

		Y		
		0	**1**	**2**
X	**0**	.25	.25	0
	1	0	.25	.25

(b) From the above definition, we take $f(x, y) = x$ and evaluate,

$$E[X] = \sum_{i=1}^{2}\sum_{j=1}^{3} x_i\,p(x_i, y_j) = \left[0\cdot.25 + 0\cdot.25 + 0\cdot 0\right] + \left[1\cdot 0 + 1\cdot.25 + 1\cdot.25\right] = .50.$$

Similarly, with $f(x, y) = x^2$, we obtain

$$E[X^2] = \left[(0)^2(.25) + (0)^2(.25) + (0)^2(0)\right] + \left[(1^2)(0) + (1^2)(.25) + (1^2)(.25)\right] = 0.50.$$

It follows that $Var[X] = 0.50-(0.50)^2 = 0.25$.

Again, using the definition above with $f(x,y)=y$ and $f(x,y)=y^2$ we find

$$\begin{aligned} E[Y] &= \sum_{i=1}^{2}\sum_{j=1}^{3} y_j p(x_i,y_j) \\ &= \big[(0)(.25)+(1)(.25)+(2)(0)\big]+\big[(0)(0)+(1)(.25)+(2)(.25)\big] = 1.0. \end{aligned}$$

$$\begin{aligned} E[Y^2] &= \sum_{i=1}^{2}\sum_{j=1}^{3} y_j^2 p(x_i,y_j) \\ &= \big[(0^2)(.25)+(1^2)(.25)+(2^2)(0)\big]+\big[(0^2)(0)+(1^2)(.25)+(2^2)(.25)\big]=1.5, \end{aligned}$$

so that $Var[Y]=1.5-1^2=0.5$.

We chose to sum horizontally first (the inner sum, with respect to *j*), then vertically (the outer sum, with respect to *i*). Clearly, reversing the order of summation between *i* and *j* would produce the same results.

(c) We calculate the marginal distributions of *X* and *Y* and display the individual tables for the first and second moments:

		Y			
		0	**1**	**2**	$p_X(x)$
X	**0**	.25	.25	0	0.50
	1	0	.25	.25	0.50
$p_Y(y)$		0.25	0.50	0.25	1.00

X	$p_X(x)$	$E[X]$	$E[X^2]$
0	.5	0	0
1	.5	.5	.5
Total	1.0	.5	.5

Y	$p_Y(y)$	$E[Y]$	$E[Y^2]$
0	.25	0	0
1	.50	.50	.50
2	.25	.50	1.00
Total	1.00	1.00	1.50

Thus, we see that using the marginal distributions produces the same moments for *X* and *Y* as the joint distribution, and therefore, of course, the same expectations and variances.

(d) We use $f(x,y)=x+y$ and $f(x,y)=(x+y)^2$ to calculate $E[X+Y]$ and $E\big[(X+Y)^2\big]$.

$$E[X+Y] = \sum_{i=1}^{2}\sum_{j=1}^{3}(x_i+y_j)p(x_i,y_j)$$
$$=\left[(0+0)(.25)+(0+1)(.25)+(0+2)(0)\right]$$
$$+\left[(1+0)(0)+(1+1)(.25)+(1+2)(.25)\right] = 1.5.$$
$$E\left[(X+Y)^2\right] = \sum_{i=1}^{2}\sum_{j=1}^{3}(x_i+y_j)^2 p(x_i,y_j)$$
$$= \left[(0+0)^2(.25)+(0+1)^2(.25)+(0+2)^2(0)\right]$$
$$+\left[(1+0)^2(0)+(1+1)^2(.25)+(1+2)^2(.25)\right] = 3.5.$$

It follows from the variance formula that $Var[X+Y] = 3.5-1.5^2 = 1.25$.

For (e) we see that $E[X+Y] = 1.5 = .5+1.0 = E[X]+E[Y]$. On the other hand, $Var[X+Y] = 1.25 \neq .75 = .25+.50 = Var[X]+Var[Y]$.

Note

We see that the means and variances of the marginal distributions could be calculated either from the joint distribution (as in (b)), or from the individual marginal distributions (as in (c)). We also make the observation, from the original statement of the problem, that $X \sim Binomial(p=.5, n=1)$ and $Y \sim Binomial(p=.5, n=2)$. Thus, we can calculate the means and variances directly from the standard formulas. Of course, each of these methods is valid; that is to say, all roads may lead to Rome, but some routes are shorter than others.

❒

Example 7.1-3 Expectations

Compute $E\left[2X-\sqrt{Y}\right]$ if the random variables X and Y are discrete with joint probability function as follows:

		Y		
		0	**1**	**4**
	1	.10	.05	.15
X	**3**	.05	.20	.25
	5	.15	.00	.05

Solution

$$E\left[2X-\sqrt{Y}\right] = \left(2\cdot1-\sqrt{0}\right)\cdot.1+\left(2\cdot1-\sqrt{1}\right)\cdot.05+\left(2\cdot1-\sqrt{4}\right)\cdot.15$$
$$+\cdots+\left(2\cdot5-\sqrt{4}\right)\cdot.05 = 4.45.$$

❒

Example 7.1-4 Infinite Discrete Joint Distribution

Consider random variables X and Y with joint probability distribution $p(x,y)=\frac{e^{-4}3^x}{x!\cdot y!}$ for $x=0,1,2,\dots$ and $y=0,1,2,\dots$.

(a) Show that $p(x,y)$ is a joint probability distribution.

(b) Find the marginal probabilities $p_Y(y)=\sum_{i=1}^{\infty} p(x_i,y)$.

Solution

As a warm-up computation, $\Pr(X=2,Y=3) = p(2,3) = \frac{e^{-4}3^2}{2!3!} = 0.013737$.

The probabilities need to be positive (they are) and sum to one. Recall that $e^{\alpha} = \sum_{n=0}^{\infty}\frac{\alpha^n}{n!} = 1+\alpha+\frac{\alpha^2}{2!}+\frac{\alpha^3}{3!}+\cdots$.

$$\begin{aligned}\sum_{x,y} p(x,y) &= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\frac{e^{-4}3^x}{x!\cdot y!} = \sum_{x=0}^{\infty}\left(\frac{e^{-4}3^x}{x!}\sum_{y=0}^{\infty}\frac{1}{y!}\right)\\ &= \sum_{x=0}^{\infty}\left(\frac{e^{-4}3^x}{x!}\cdot e^1\right)\\ &= \sum_{x=0}^{\infty}\left(\frac{e^{-3}3^x}{x!}\right) = e^{-3}\cdot\sum_{x=0}^{\infty}\frac{3^x}{x!} = e^{-3}\cdot e^3 = 1.\end{aligned}$$

Similarly,

$$p_Y(y) = \sum_{x=0}^{\infty} p(x,y) = \sum_{x=0}^{\infty}\frac{e^{-4}3^x}{x!\cdot y!} = \frac{e^{-4}}{y!}\sum_{x=0}^{\infty}\frac{3^x}{x!} = \frac{e^{-4}}{y!}\cdot e^3 = \frac{e^{-1}}{y!}$$

for all $y=0,1,2,\dots$. ❒

Exercise 7-1 Consider random variables X and Y with joint probability distribution, $p(x,y)=\frac{e^{-4}3^x}{x!\cdot y!}$ for $x=0,1,2,\dots$ and $y=0,1,2,\dots$.

(a) Find the marginal probabilities for X, $p_X(x)=\sum_{j=1}^{\infty} p(x,y_j)$.

(b) Compute the mean of the random variable X.

Example 7.1-5 Joint Distributions

Given that X and Y are discrete random variables with joint probability distribution:

		Y		
		0	**1**	**4**
	1	.10	.05	.15
X	**3**	.05	.20	.25
	5	.15	.00	.05

(a) Calculate the marginal probability functions for X and Y.
(b) Calculate the expected value of X from both the joint distribution and the marginal distribution for X.

Solution

(a) The joint table with marginal distributions displayed is:

		Y			
		0	**1**	**4**	$p_X(x)$
	1	.10	.05	.15	**.30**
X	**3**	.05	.20	.25	**.50**
	5	.15	.00	.05	**.20**
$p_Y(y)$		.30	.25	.45	1.00

Therefore the marginal distributions are:

Y	$p_Y(y)$
0	.30
1	.25
4	.45
	1.00

X	$p_X(x)$
1	**.30**
3	**.50**
5	**.20**
	1.00

(b) Using the joint probability function, we have

$$
\begin{aligned}
E[X] &= \sum\sum x \cdot p(x,y) \\
&= 1\cdot p(1,0)+1\cdot p(1,1)+1\cdot p(1,4) \\
&\qquad +3\cdot p(3,0)+3\cdot p(3,1)+\cdots+5\cdot p(5,4) \\
&= 1\cdot(.10)+1\cdot(.05)+1\cdot(.15)+3\cdot(.05) \\
&\qquad +3\cdot(.20)+3\cdot(.25)+5\cdot(.15)+5\cdot(.00)+5\cdot(.05) \\
&= \mathbf{1\cdot(.30)+3\cdot(.50)+5\cdot(.20)} \\
&= 2.80.
\end{aligned}
$$

Using our marginal distribution for X,

$$
\begin{aligned}
E[X] &= \sum x \cdot p_X(x) \\
&= 1\cdot p_X(1)+3\cdot p_X(3)+5\cdot p_X(5) = \mathbf{1\cdot(.30)+3\cdot(.50)+5\cdot(.20)} = 2.80.
\end{aligned}
$$

❒

Exercise 7-2 Regarding the random variables in Example 7.1-5, compute $E[Y^2]$ using both the joint probability distribution and the marginal probability distribution for *Y*.

Exercise 7-3 Suppose *X* and *Y* are discrete random variables with $p(x,y) = k\cdot(x+2y)$ for $x = 0,1,2,3$ and $y = 1,2$.

(a) Determine the constant *k*.

(b) Find $E\left[\sqrt{Y}\right]$.

Exercise 7-4 SOA EXAM P Sample Exam Questions #100

A car dealership sells 0, 1, or 2 luxury cars on any day. When selling a car, the dealer also tries to persuade the customer to buy an extended warranty for the car. Let *X* denote the number of luxury cars sold in a given day, and let *Y* denote the number of extended warranties sold.

$$P(X=0, Y=0) = \tfrac{1}{6} \qquad P(X=1, Y=0) = \tfrac{1}{12}$$
$$P(X=1, Y=1) = \tfrac{1}{6} \qquad P(X=2, Y=0) = \tfrac{1}{12}$$
$$P(X=2, Y=1) = \tfrac{1}{3} \qquad P(X=2, Y=2) = \tfrac{1}{6}$$

What is the variance of *X*? Author's note: Solve this problem two different ways; (1) using marginal probabilities and (2) directly.

(A) 0.47 (B) 0.58 (C) 0.83 (D) 1.42 (E) 2.58

Exercise 7-5 A cooler has 6 Gatorades®, 2 colas, and 4 waters. You select three beverages from the cooler at random. Let *B* denote the number of Gatorades® selected and let *C* denote the number of colas selected. For example, if you grabbed a cola and two waters, then $C = 1$ and $B = 0$.

(a) Construct a joint probability distribution for *B* and *C*.

(b) Find the marginal distribution $p_B(b)$.

(c) Compute $E[C]$.

(d) Compute $E[3B - C^2]$.

11.7 CHAPTER 11 SAMPLE EXAMINATION

1. **CAS Exam 3 Fall 2005 #1**

The following sample was taken from a distribution with probability density function $f(x) = \theta x^{\theta-1}$, where $0 < x < 1$ and $\theta > 0$.

0.21 0.43 0.56 0.67 0.72

Let R and S be the estimators of θ using the maximum likelihood method and method of moments, respectively. Calculate the value of $R - S$.

(A) Less than 0.3
(B) At least 0.3, but less than 0.4
(C) At least 0.4, but less than 0.5
(D) At least 0.5, but less than 0.6
(E) At least 0.6

2. **CAS Exam 3 Fall 2006 #6**

You are testing the hypothesis H_0 that the random variable X has a uniform distribution on $[0,10]$ against the alternative hypothesis H_1 that X has a uniform distribution on [5,10]. Using a single observation and a significance level 5%, calculate the probability of a Type II error.

(A) Less than 0.2
(B) At least 0.2, but less than 0.4
(C) At least 0.4, but less than 0.6
(D) At least 0.6, but less than 0.8
(E) At least 0.8

3. **CAS Exam 3 Spring 2005 #21**

An actuary obtains two independent, unbiased estimates, Y_1 and Y_2, for a certain parameter. The variance of Y_1 is four times that of Y_2. A new unbiased estimator of the form $k_1 * Y_1 + k_2 * Y_2$ is to be constructed. What is the value of k_1 that minimizes the variance of the new estimate?

(A) Less than 0.148
(B) At least 0.18, but less than 0.23
(C) At least 0.23, but less than 0.28
(D) At least 0.28, but less than 0.33
(E) 0.33 or more

15. **CAS Exam ST Spring 2014 #23**

You are given the following information:

- $X_1, \ldots, X_5$ are a random sample from a Poisson distribution with parameter λ, where λ follows a Gamma distribution with $\alpha = 2$ and β.
- The mean of this Poisson-Gamma conjugate pair can be represented as a weighted average of the maximum likelihood estimator for the mean and the mean of the prior distribution.
- Let W_{MLE} be the weight assigned to the maximum likelihood estimator.
- The maximum likelihood estimate for the mean is 1.2.
- The variance of the prior Gamma is 8.

Calculate W_{MLE}.

16. **CAS Exam ST Fall 2014 #5**

Calculate the Fisher information $I(q)$ based on a single observation from a Bernoulli trial random variable with probability of success q.

17. **CAS Exam ST Fall 2014 #7**

You are given the following information:

- $X_1, \ldots, X_7$ is a random sample of size 7 from an exponential distribution with mean θ.
- The sample mean is 400.

Calculate the MVUE (minimum variance unbiased estimator) of θ^2.

18. **CAS Exam ST Fall 2014 #22**

You are given:

i. For a general liability policy, the log of paid claims conditionally follows the normal distribution with mean μ, which varies by policyholder, and variance 1.

ii. The posterior distribution of μ follows the normal distribution with mean $\dfrac{n\bar{X}+2}{n+1}$ and variance $\dfrac{1}{n+1}$ where $\bar{X}$ denotes the sample mean and n denotes the sample size.

iii. The following sample of observed log of paid claims:

3.22	4.34	5.98	7.32	2.78

Calculate the upper bound of the symmetric 95% Bayesian confidence interval for μ.