a/s/m

Actuarial Study Materials

Study Manual for
Exam P/Exam 1
Probability

17-th Edition (for 2016)

by

Dr. Krzysztof Ostaszewski
FSA, CERA, FSAS, CFA, MAAA

Note:
NO RETURN IF OPENED

TO OUR READERS:

Please check A.S.M.'s web site at www.studymanuals.com for errata and updates. If you have any comments or reports of errata, please e-mail us at mail@studymanuals.com or e-mail Dr. Krzysztof Ostaszewski, at krzysio@krzysio.net. His web site is www.krzysio.net. All errata of current Dr. Ostaszewski's books are posted at http://smartURL.it/errata.

The author would like to thank the Society of Actuaries and the Casualty Actuarial Society for their kind permission to publish questions from past actuarial examinations.

INTRODUCTION

Before you start studying for actuarial examinations you need to familiarize yourself with the following:

Fundamental Rule for Passing Actuarial Examinations
You should greet every problem you see when you are taking the exam with these words: "Been there, done that."

I will refer to this Fundamental Rule as the *BTDT Rule*. If you do not follow the BTDT Rule, neither this manual, nor any book, nor any tutorial, will be of much use to you. And I do want to help you, so I must beg you to follow the BTDT Rule. Allow me now to explain its meaning.

If you are surprised by any problem on the exam, you are likely to miss that problem. Yet the difference between a 5 and a 6 is one problem. This "surprise" problem has great marginal value. There is simply not enough time to think on the exam. *Thinking is always the last resort on an actuarial exam*. You may not have seen this very problem before, but you must have seen a problem like it before. If you have not, you are not prepared.

If you have not thoroughly studied all topics covered on the exam you are taking, you must have subconsciously wished to spend more time studying … a half a year's, or even a year's worth, more. But, clearly, the biggest reward for passing an actuarial examination is: *Not having to take it again!* By spending extra hours, days, or even weeks, studying and memorizing *all* topics covered on the exam, you are saving yourself possibly as much as a year's worth of your life. Getting a return of one year on an investment of one day is better than anything you will ever make on Wall Street, or even lotteries (unearned wealth is destructive, thus objects called "lottery winnings" are smaller than they appear). Please study thoroughly, without skipping any topic or any kind of a problem. To paraphrase my favorite quote from Ayn Rand: *for zat, you will be very grateful to yourself*.

I had once seen Harrison Ford being asked what he answers to people who tell him: "*May the Force be with you!"*? He said: "*Force Yourself!"* That's what you need to do. You have much more will than you assume, so go make yourself study.

Good luck!

Krzys' Ostaszewski
Bloomington, Illinois, September 2004

P.S. I want to thank my wife, Patricia, for her help and encouragement in writing of this manual. I also want to thank Hal Cherry for his help and encouragement. Any errors in this work are mine and mine only. If you find any, please be kind to let me know about them, because I definitely want to correct them.

TABLE OF CONTENTS

INTRODUCTION	**- 1 -**
SECTION 1: GENERAL PROBABILITY	**- 4 -**
SECTION 2: RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS	**- 22 -**
SECTION 3: MULTIVARIATE DISTRIBUTIONS	**- 57 -**
SECTION 4: RISK MANAGEMENT AND INSURANCE	**- 92 -**
INDEX	**- 105 -**
PRACTICE EXAMINATIONS: AN INTRODUCTION	**- 109 -**
PRACTICE EXAMINATION NUMBER 1	**- 110 -**
PRACTICE EXAMINATION NUMBER 1: SOLUTIONS	**- 118 -**
PRACTICE EXAMINATION NUMBER 2	**- 141 -**
PRACTICE EXAMINATION NUMBER 2: SOLUTIONS	**- 149 -**
PRACTICE EXAMINATION NUMBER 3	**- 172 -**
PRACTICE EXAMINATION NUMBER 3: SOLUTIONS	**- 179 -**
PRACTICE EXAMINATION NUMBER 4	**- 199 -**
PRACTICE EXAMINATION NUMBER 4: SOLUTIONS	**- 207 -**
PRACTICE EXAMINATION NUMBER 5	**- 228 -**
PRACTICE EXAMINATION NUMBER 5: SOLUTIONS	**- 236 -**
PRACTICE EXAMINATION NUMBER 6:	**- 256 -**
PRACTICE EXAMINATION NUMBER 6: SOLUTIONS	**- 263 -**
PRACTICE EXAMINATION NUMBER 7:	**- 289 -**
PRACTICE EXAMINATION NUMBER 7: SOLUTIONS	**- 296 -**
PRACTICE EXAMINATION NUMBER 8:	**- 315 -**
PRACTICE EXAMINATION NUMBER 8: SOLUTIONS	**- 322 -**
PRACTICE EXAMINATION NUMBER 9:	**- 343 -**
PRACTICE EXAMINATION NUMBER 9: SOLUTIONS	**- 351 -**

PRACTICE EXAMINATION NUMBER 10:	**- 374 -**
PRACTICE EXAMINATION NUMBER 10: SOLUTIONS	**- 381 -**
PRACTICE EXAMINATION NUMBER 11:	**- 400 -**
PRACTICE EXAMINATION NUMBER 11: SOLUTIONS	**- 408 -**
PRACTICE EXAMINATION NUMBER 12:	**- 433 -**
PRACTICE EXAMINATION NUMBER 12: SOLUTIONS	**- 440 -**
PRACTICE EXAMINATION NUMBER 13:	**- 463 -**
PRACTICE EXAMINATION NUMBER 13: SOLUTIONS	**- 470 -**
PRACTICE EXAMINATION NUMBER 14:	**- 493 -**
PRACTICE EXAMINATION NUMBER 14: SOLUTIONS	**- 500 -**
PRACTICE EXAMINATION NUMBER 15:	**- 524 -**
PRACTICE EXAMINATION NUMBER 15: SOLUTIONS	**- 531 -**
PRACTICE EXAMINATION NUMBER 16:	**- 551 -**
PRACTICE EXAMINATION NUMBER 16: SOLUTIONS	**- 557 -**
PRACTICE EXAMINATION NUMBER 17:	**- 578 -**
PRACTICE EXAMINATION NUMBER 17: SOLUTIONS	**- 585 -**
PRACTICE EXAMINATION NUMBER 18:	**- 606 -**
PRACTICE EXAMINATION NUMBER 18: SOLUTIONS	**- 614 -**
PRACTICE EXAMINATION NUMBER 19:	**- 635 -**
PRACTICE EXAMINATION NUMBER 19: SOLUTIONS	**- 643 -**
PRACTICE EXAMINATION NUMBER 20:	**- 666 -**
PRACTICE EXAMINATION NUMBER 20: SOLUTIONS	**- 673 -**

SECTION 1: GENERAL PROBABILITY

Basic probability concepts

Probability concepts are defined for elements and subsets of a certain set, a universe under consideration. That universe is called the *probability space,* or *sample space*. It can be a finite set, a countable set (a set whose elements can be put in a sequence, in fact a finite set is also countable), or an infinite uncountable set (e.g., the set of all real numbers, or any interval on the real line). Subsets of a given probability space for which probability can be calculated are called *events*. An event represents something that can possibly happen. In the most general probability theory, not every set can be an event. But this technical issue does not come up on any lower level actuarial examinations. It should be noted that while not all subsets of a probability space must be events, it is always the case that the empty set is an event, the entire space S is an event, a complement of an event is an event, and a set-theoretic union of any sequence of events is an event. The entire probability space will be usually denoted by S (always in this text) or Ω (in more theoretical probability books). It encompasses everything that can possibly happen.

The simplest possible event is a one-element set. If we perform an experiment, and observe what happens as its outcome, such a single observation is called an *elementary event*. Another name commonly used for such an event is: a *sample point*.

A *union* of two events is an event, which combines all of their elements, regardless of whether they are common to those events, or not. For two events A and B their union is denoted by $A \cup B.$ For a more general finite collection of events $A_1, A_2, \ldots, A_n,$ their union consists of all elementary events that belong to any one of them, and is denoted by $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n.$ An infinite union of a sequence of sets $\bigcup_{n=1}^{\infty} A_n$ is also defined as the set that consists of all elementary events that belong to any one of them.

An *intersection* of two events is an event, whose elements belong to both of the events. For two events A and B their intersection is denoted by $A \cap B.$ For a more general finite collection of events $A_1, A_2, \ldots, A_n,$ their intersection consists of all elementary events that belong to all of them, and is denoted by $\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n.$ An infinite intersection of a sequence of sets $\bigcap_{n=1}^{\infty} A_n$ is also defined as the set that consists of all elementary events that belong to all of them.

Two events are called *mutually exclusive* if they cannot happen at the same time. In the language of set theory this simply means that they are disjoint sets, i.e., sets that do not have any elements in common, or $A \cap B = \varnothing.$ A finite collection of events $A_1, A_2, \ldots, A_n$ is said to be *mutually exclusive* if any two events from the collection are mutually exclusive. The concept is defined the same way for infinite collections of events.

We say that a collection of events forms *exhaustive outcomes,* or that this collection forms a *partition* of the probability space, if their union is the entire probability space, and they are mutually exclusive.

An event A is a *subevent* (although most commonly we use the set-theoretic concept of a *subset)* of an event B, denoted by $A \subset B$, if every elementary event (sample point) in A is also contained in B. This relationship is a mathematical expression of a situation when occurrence of A automatically implies that B also occurs. Note that if $A \subset B$ then $A \cup B = B$ and $A \cap B = A$.

For an event E, its *complement*, denoted by E^C, consists of all elementary events (i. e., elements of the sample space S) that do not belong to E. In other words, $E^C = S - E$, where S is the entire probability space. Note that $E \cup E^C = S$ and $E \cap E^C = \varnothing$. Recall that $A - B = A \cap B^C$ is the set difference operation.

An important rule concerning complements of sets is expressed by *DeMorgan's Laws*:

$$(A \cup B)^C = A^C \cap B^C, \qquad (A \cap B)^C = A^C \cup B^C,$$

$$\left(\bigcup_{i=1}^{n} A_i\right)^C = \left(A_1 \cup A_2 \cup \ldots \cup A_n\right)^C = A_1^C \cap A_2^C \cap \ldots \cap A_n^C = \bigcap_{i=1}^{n} A_i^C,$$

$$\left(\bigcap_{i=1}^{n} A_i\right)^C = \left(A_1 \cap A_2 \cap \ldots \cap A_n\right)^C = A_1^C \cup A_2^C \cup \ldots \cup A_n^C = \bigcup_{i=1}^{n} A_i^C,$$

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^C = \bigcap_{n=1}^{\infty} A_n^C, \qquad \left(\bigcap_{n=1}^{\infty} A_n\right)^C = \bigcup_{n=1}^{\infty} A_n^C.$$

The *indicator function* for an event E is a function $I_E : S \to \mathbb{R}$, where S is the entire probability space, and $\mathbb{R}$ is the set of real numbers, defined as $I_E(x) = 1$ if $x \in E$, and $I_E(x) = 0$ if $x \notin E$.

The simplest, and commonly used, example of a probability space is a set consisting of two elements $S = \{0,1\}$, with 1 corresponding to "success" and 0 representing "failure". An experiment in which only such two outcomes are possible is called a *Bernoulli Trial*. You can view taking an actuarial examination as an example of a Bernoulli Trial. Tossing a coin is also an example of a Bernoulli trial, with two possible outcomes: heads or tails (you get to decide which of these two would be termed success, and which one is a failure).

Another commonly used finite probability space consists of all outcomes of tossing a fair six-faced die. The sample space is $S = \{1,2,3,4,5,6\}$, each number being a sample point representing the number of spots that can turn up when the die is tossed. The outcomes 1 and 6, for example, are mutually exclusive (more formally, these outcomes are events {1} and {6}). The outcomes $\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}$, usually written just as 1, 2, 3, 4, 5, 6, are exhaustive for this probability space and form a partition of it. The set {1, 3, 5} represents the event of obtaining an odd number when tossing a die. Consider the following events in this die-tossing sample space:

$A = \{1, 2, 3, 4\}$ = "a number less than 5 is tossed",
$B = \{2, 4, 6\}$ = "an even number is tossed",
$C = \{1\}$ = "a 1 is tossed",
$D = \{5\}$ = "a 5 is tossed".

Then we have:

$A \cup B = \{1,2,3,4,6\} = D^C$,
$A \cap B = \{2,4\}$,
$A \cap D = \varnothing$, i.e., events A and D are mutually exclusive,
$(B \cup C)^C = \{1,2,4,6\}^C = \{3,5\} = \{1,3,5\} \cap \{2,3,4,5,6\} = B^C \cap C^C$.

Another rule concerning operations on events:

$$A \cap (E_1 \cup E_2 \cup \ldots \cup E_n) = (A \cap E_1) \cup (A \cap E_2) \cup \ldots \cup (A \cap E_n),$$
$$A \cup (E_1 \cap E_2 \cap \ldots \cap E_n) = (A \cup E_1) \cap (A \cup E_2) \cap \ldots \cap (A \cup E_n),$$

so that the distributive property holds the same way for unions as for intersections. In particular, if $E_1, E_2, \ldots, E_n$ form a partition of *S,* then for any event *A*

$$A = A \cap S = A \cap (E_1 \cup E_2 \cup \ldots \cup E_n) = (A \cap E_1) \cup (A \cap E_2) \cup \ldots \cup (A \cap E_n).$$

As the events $A \cap E_1, A \cap E_2, \ldots, A \cap E_n$ are also mutually exclusive, they form a partition of the event *A*. In the special case when $n = 2$, $E_1 = B$, $E_2 = B^C$, we see that $A \cap B$ and $A \cap B^C$ form a partition of *A*.

Probability (we will denote it by Pr) is a function that assigns a number between 0 and 1 to each event, with the following defining properties:

- $\Pr(\varnothing) = 0$,
- $\Pr(S) = 1$,

and

- If $\{E_n\}_{n=1}^{+\infty}$ is a sequence of mutually exclusive events, then $\Pr\left(\bigcup_{n=1}^{+\infty} E_n\right) = \sum_{n=1}^{+\infty} \Pr(E_n)$.

While the last condition is stated for infinite unions and a sum of a series, it applies equally to finite unions of mutually exclusive events, and the finite sum of their probabilities.

A *discrete probability space,* or *discrete sample space* is a probability space with a countable (finite or infinite) number of sample points. The assignment of probability to each elementary event in a discrete probability space is called the *probability function,* sometimes also called *probability mass function*.

For the simplest such space described as Bernoulli Trial, probability is defined by giving the probability of success, usually denoted by p. The probability of failure, denoted by q is then equal to $q = 1 - p$. Tossing a fair coin is a Bernoulli Trial with $p = 0.5$. Taking an actuarial examination is a Bernoulli Trial and we are trying to get your p to be as close to 1 as possible.

When a Bernoulli Trial is performed until a success occurs, and we count the total number of attempts, the resulting probability space is discrete, but infinite. Suppose, for example, that a fair coin is tossed until the first head appears. The toss number of the first head can be any positive integer and thus the probability space is infinite. Repeatedly taking an actuarial examination creates the same kind of discrete, yet infinite, probability space. Of course, we hope that after reading this manual, your probability space will not only be finite, but a *degenerate* one, consisting of one element only.

Tossing an ordinary die is an experiment with a finite probability space $\{1,2,3,4,5,6\}$. If we assign to each outcome the same probability of $\frac{1}{6}$, we obtain an example of what is called a *uniform probability function*. In general, for a finite discrete probability space, a uniform probability function assigns the same probability to each sample point. Since probabilities have to add up to 1, if there are n points in a finite probability space, uniform probability function assigns to each point equal probability of $\frac{1}{n}$.

In a finite probability space, for an event E, we always have $\Pr(E)=\sum_{e\in E}\Pr(\{e\})$. In other words, you can calculate probability of an event by adding up probabilities of all elementary events contained in it. For example, when rolling a fair die, the probability of getting an even number is

$$\Pr(\{2\})+\Pr(\{4\})+\Pr(\{6\})=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{1}{2}.$$

Discrete probability space is only one type of a probability space. Other types analyzed involve a continuum of possible outcomes, and are called *continuous probability spaces*. For example we assume usually that an automobile physical damage claim is a number between 0 and the car's value (possibly after deductible is subtracted), i.e., it can be any real number between those two boundary values.

Some rules concerning probability:

- If $A\subset B$ then $\Pr(A)\le\Pr(B)$. Also, $A\subset B$ implies that $\Pr(B-A)=\Pr(B)-\Pr(A)$.
- $\Pr(A\cup B)=\Pr(A)+\Pr(B)-\Pr(A\cap B)$,
- $\Pr(A\cup B\cup C)=$
  $=\Pr(A)+\Pr(B)+\Pr(C)-\Pr(A\cap B)-\Pr(A\cap C)-\Pr(B\cap C)+\Pr(A\cap B\cap C)$.
- $\Pr(E^C)=1-\Pr(E)$.
- $\Pr(A)=\Pr(A\cap E)+\Pr(A\cap E^C)$.
- If events $E_1,E_2,\ldots,E_n$ form a partition of the probability space S then:

$$\Pr(A)=\Pr(A\cap E_1)+\Pr(A\cap E_2)+\ldots+\Pr(A\cap E_n).$$

The last statement is commonly referred to as *The Law of Total Probability*.

In some of the actuarial examinations you will see the words "percentage" and "probability" used interchangeably. If you are told, for example, that 75% of a certain population watch the television show *White Collar,* you should interpret this 75% = 0.75 as the probability that a person randomly chosen from this population watches this television show.

Exercise 1.1. May 2003 Course 1 Examination, Problem No. 1, also P Sample Exam Questions, Problem No. 1, and Dr. Ostaszewski's online exercise posted April 18, 2009

A survey of a group's viewing habits over the last year revealed the following information:
(i) 28% watched gymnastics,
(ii) 29% watched baseball,
(iii) 19% watched soccer,
(iv) 14% watched gymnastics and baseball,
(v) 12% watched baseball and soccer,
(vi) 10% watched gymnastics and soccer,
(vii) 8% watched all three sports.
Calculate the percentage of the group that watched none of the three sports during the last year.

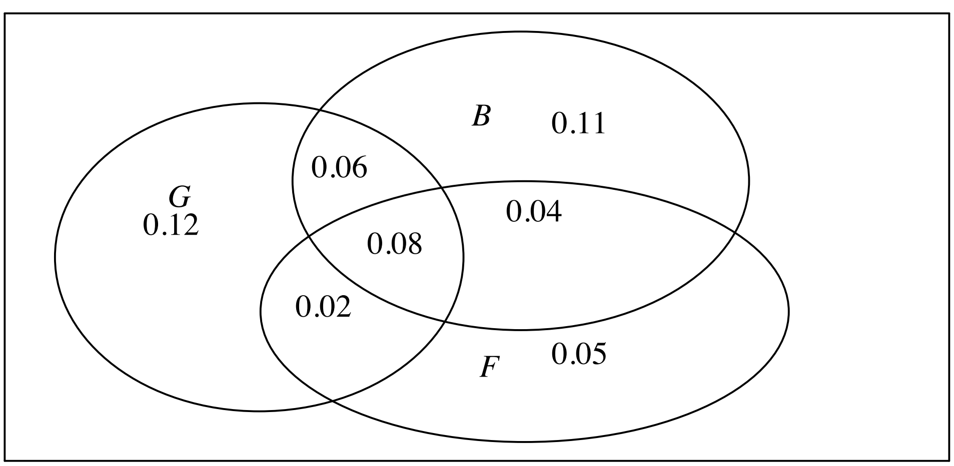
A. 24 B. 36 C. 41 D. 52 E. 60

Solution.

Treat the groups' preferences as events: G – watching gymnastics, B – watching baseball, F – watching soccer (F for "football", as the rest of the world calls it, because we use S for the probability space). We have the following probabilities:

$$\Pr(G)=0.28, \quad \Pr(B)=0.29, \quad \Pr(F)=0.19,$$
$$\Pr(G\cap B)=0.14, \quad \Pr(B\cap F)=0.12, \quad \Pr(G\cap F)=0.10,$$
$$\Pr(G\cap B\cap F)=0.08.$$

Therefore, the non-overlapping pieces of these sets have the probabilities shown in the figure below.



We are interested in the "area" (probability) outside of the ovals, i.e.,

$$1-0.12-0.02-0.06-0.08-0.05-0.04-0.11=0.52.$$

We can also formally calculate it as:

$$\Pr\left((G\cup B\cup F)^{C}\right)=1-\Pr(G\cup B\cup F)=1-\Pr(G)-\Pr(B)-\Pr(F)+$$
$$+\Pr(G\cap B)+\Pr(G\cap F)+\Pr(B\cap F)-\Pr(G\cap B\cap F)=$$
$$=1-0.28-0.29-0.19+0.14+0.12+0.10-0.08=0.52.$$

Answer D.

Exercise 1.2. November 2001 Course 1 Examination, Problem No. 9, also P Sample Exam Questions, Problem No. 8, and Dr. Ostaszewski's online exercise posted August 18, 2007
Among a large group of patients recovering from shoulder injuries, it is found that 22% visit both a physical therapist and a chiropractor, whereas 12% visit neither of these. The probability that a patient visits a chiropractor exceeds by 0.14 the probability that a patient visits a physical therapist. Determine the probability that a randomly chosen member of this group visits a physical therapist.

A. 0.26 B. 0.38 C. 0.40 D. 0.48 E. 0.62

Solution.
Let *C* be the event that a patient visits a chiropractor, and *T* be the event that a patient visits a physical therapist. We are given that $\Pr(C\cap T)=0.22$, $\Pr\left(C^{C}\cap T^{C}\right)=\Pr\left((C\cup T)^{C}\right)=0.12$, and $\Pr(C)=\Pr(T)+0.14$. Therefore,

$$1-0.12=0.88=\Pr(C\cup T)=\Pr(C)+\Pr(T)-\Pr(C\cap T)=$$
$$=\Pr(T)+0.14+\Pr(T)-0.22=2\Pr(T)-0.08.$$

This implies that

$$\Pr(T)=\frac{0.88+0.08}{2}=0.48.$$

Answer D.

Conditional probability
The concept of conditional probability is designed to capture the relationship between probabilities of two or more events happening. The simplest version of this relationship is in the question: given that an event *B* happened, how does this affect the probability that *A* happens? In order to tackle this question, we define the *conditional probability of A given B* as:

$$\Pr(A|B)=\frac{\Pr(A\cap B)}{\Pr(B)}.$$

In order for this definition to make sense we must assume that $\Pr(B)>0$. This concept of conditional probability basically makes *B* into the new probability space, and then takes the probability of *A* to be that of only the part of *A* inside of *B*, scaled by the probability of *B* in relation to the probability of the entire *S* (which is, of course, 1). Note that the definition implies that

$$\Pr(A \cap B) = \Pr(A|B) \cdot \Pr(B),$$

of course as long as $\Pr(B) > 0$.

Note also the following properties of conditional probability:

• For any event E with positive probability, the function $A \mapsto \Pr(A|E)$, assigning to any event A its conditional probability $\Pr(A|E)$ meets the conditions of the definition of probability, thus it is itself a probability, and has all properties of probability that we listed in the previous section. For example,

$$\Pr(A^C|E) = 1 - \Pr(A|E),$$

or, for any two events A and B,

$$\Pr(A \cup B|E) = \Pr(A|E) + \Pr(B|E) - \Pr(A \cap B|E).$$

We also have:

• If $A \subset B$ then $\Pr(A|B) = \dfrac{\Pr(A \cap B)}{\Pr(B)} = \dfrac{\Pr(A)}{\Pr(B)}$, and $\Pr(B|A) = \dfrac{\Pr(A \cap B)}{\Pr(A)} = \dfrac{\Pr(A)}{\Pr(A)} = 1.$

• If $\Pr(A_1 \cap A_2 \cap \ldots A_n) > 0$ then

$$\Pr(A_1 \cap A_2 \cap \ldots A_n) = \Pr(A_1) \cdot \Pr(A_2|A_1) \cdot \Pr(A_3|A_1 \cap A_2) \cdot \ldots \cdot \Pr(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

Exercise 1.3. May 2003 Course 1 Examination, Problem No. 18, also P Sample Exam Questions, Problem No. 7, and Dr. Ostaszewski's online exercise posted August 11, 2007

An insurance company estimates that 40% of policyholders who have only an auto policy will renew next year and 60% of policyholders who have only a homeowners policy will renew next year. The company estimates that 80% of policyholders who have both an auto and a homeowners policy will renew at least one of those policies next year. Company records show that 65% of policyholders have an auto policy, 50% of policyholders have a homeowners policy, and 15% of policyholders have both an auto and a homeowners policy. Using the company's estimates, calculate the percentage of policyholders that will renew at least one policy next year.

A. 20 B. 29 C. 41 D. 53 E. 70

Solution.

Let A be the event that a policyholder has an auto policy, and H be the event that a policyholder has a homeowners policy, and R be the event that a policyholder renews a policy. We are given: $\Pr(A) = 0.65$, $\Pr(H) = 0.50$, $\Pr(A \cap H) = 0.15$, $\Pr(R|A \cap H^C) = 0.40$, $\Pr(R|H \cap A^C) = 0.60$, and $\Pr(R|A \cap H) = 0.80$. We are looking for $\Pr(R)$. Note that

$$\Pr(A \cap H^C) = \Pr(A - H) = \Pr(A - (A \cap H)) = \Pr(A) - \Pr(A \cap H) = 0.65 - 0.15 = 0.50,$$

and

$$\Pr(A^C \cap H) = \Pr(H - A) = \Pr(H - (H \cap A)) = \Pr(H) - \Pr(A \cap H) = 0.50 - 0.15 = 0.35.$$

Also note that the events $A \cap H^C$, $A^C \cap H$, and $A \cap H$ form a partition of the probability space considered. Therefore,

$$\Pr(R) = \Pr(R \cap (A \cap H^C)) + \Pr(R \cap (A^C \cap H)) + \Pr(R \cap (A \cap H)) =$$

$$= \Pr(R | A \cap H^C) \cdot \Pr(A \cap H^C) + \Pr(R | A^C \cap H) \cdot \Pr(A^C \cap H) +$$

$$+ \Pr(R | A \cap H) \cdot \Pr(A \cap H) = 0.4 \cdot 0.5 + 0.6 \cdot 0.35 + 0.8 \cdot 0.15 = 0.53.$$

Answer D.

Exercise 1.4. May 2003 Course 1 Examination, Problem No. 5, also P Sample Exam Questions, Problem No. 9, and Dr. Ostaszewski's online exercise posted August 25, 2007

An insurance company examines its pool of auto insurance customers and gathers the following information:

(i) All customers insure at least one car.
(ii) 70% of the customers insure more than one car.
(iii) 20% of the customers insure a sports car.
(iv) Of those customers who insure more than one car, 15% insure a sports car.

Calculate the probability that a randomly selected customer insures exactly one car and that car is not a sports car.

A. 0.13 B. 0.21 C. 0.24 D. 0.25 E. 0.30

Solution.

Always start by labeling the events. Let *C* (stands for Corvette) be the event of insuring a sports car (not *S*, because we reserve this for the entire probability space), and *M* be the event of insuring multiple cars. Note that M^C is the event of insuring exactly one car, as all customers insure at least one car. We are given that $\Pr(M) = 0.70$, $\Pr(C) = 0.20$, and $\Pr(C | M) = 0.15$. We need to find $\Pr(M^C \cap C^C)$. You need to recall De Morgan's Laws and then we see that:

$$\Pr(M^C \cap C^C) = \Pr((M \cup C)^C) = 1 - \Pr(M \cup C) = 1 - \Pr(M) - \Pr(C) + \Pr(M \cap C) =$$

$$= 1 - \Pr(M) - \Pr(C) + \Pr(C | M)\Pr(M) = 1 - 0.70 - 0.20 + 0.15 \cdot 0.70 = 0.205.$$

Answer B.

Independence of events

We say that events *A* and *B* are *independent* if

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B).$$

Note that any event of probability 0 is independent of any other events. For events with positive probabilities, independence of *A* and *B* is equivalent to $\Pr(A | B) = \Pr(A)$ or $\Pr(B | A) = \Pr(B)$. This means that two events are independent if occurrence of one of them has no effect on the probability of the other one happening. You must remember that the concept of independence should never be confused with the idea of two events being mutually exclusive. If two events have positive probabilities and they are mutually exclusive, then they must be dependent, as if one of them happens, the other one cannot happen, and the conditional probability of the second event given the first one must be zero, while the unconditional (regular) probability is not zero. It can be shown, and should be memorized by you that if A and B are independent, then so are A^C and B, as well as A and B^C, and so are A^C and B^C.

For three events A, B, and C, we say that they are *independent*, if $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$, $\Pr(A \cap C) = \Pr(A) \cdot \Pr(C)$, $\Pr(B \cap C) = \Pr(B) \cdot \Pr(C)$, and

$$\Pr(A \cap B \cap C) = \Pr(A) \cdot \Pr(B) \cdot \Pr(C).$$

In general, we say that events $A_1, A_2, \ldots, A_n$ are *independent* if, for any finite collection of them $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$, where $1 \le i_1 < i_2 < \ldots < i_k \le n$, we have

$$\Pr\left(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}\right) = \Pr\left(A_{i_1}\right) \cdot \Pr\left(A_{i_2}\right) \cdot \ldots \cdot \Pr\left(A_{i_k}\right).$$

Exercise 1.5. May 2003 Course 1 Examination, Problem No. 37, also P Sample Exam Questions, Problem No. 17, and Dr. Ostaszewski's online exercise posted October 13, 2007
An insurance company pays hospital claims. The number of claims that include emergency room or operating room charges is 85% of the total number of claims. The number of claims that do not include emergency room charges is 25% of the total number of claims. The occurrence of emergency room charges is independent of the occurrence of operating room charges on hospital claims. Calculate the probability that a claim submitted to the insurance company includes operating room charges.

A. 0.10 B. 0.20 C. 0.25 D. 0.40 E. 0.80

Solution.
As always, start by labeling the events. Let O be the event of incurring operating room charges, and E be the event of emergency room charges. Then, because of independence of these two events,

$$0.85 = \Pr(O \cup E) = \Pr(O) + \Pr(E) - \Pr(O \cap E) = \Pr(O) + \Pr(E) - \Pr(O) \cdot \Pr(E).$$

Since $\Pr\left(E^C\right) = 0.25 = 1 - \Pr(E)$, it follows that $\Pr(E) = 0.75$. Therefore

$$0.85 = \Pr(O) + 0.75 - 0.75 \cdot \Pr(O),$$

and $\Pr(O) = 0.40$.
Answer D.

The Bayes Theorem
Note that as long as $0 < \Pr(E) < 1$:

$$\Pr(A) = \Pr(A \cap E) + \Pr\left(A \cap E^C\right) = \Pr(A|E) \cdot \Pr(E) + \Pr\left(A|E^C\right) \cdot \Pr\left(E^C\right).$$

Therefore, if also $\Pr(A) > 0$,

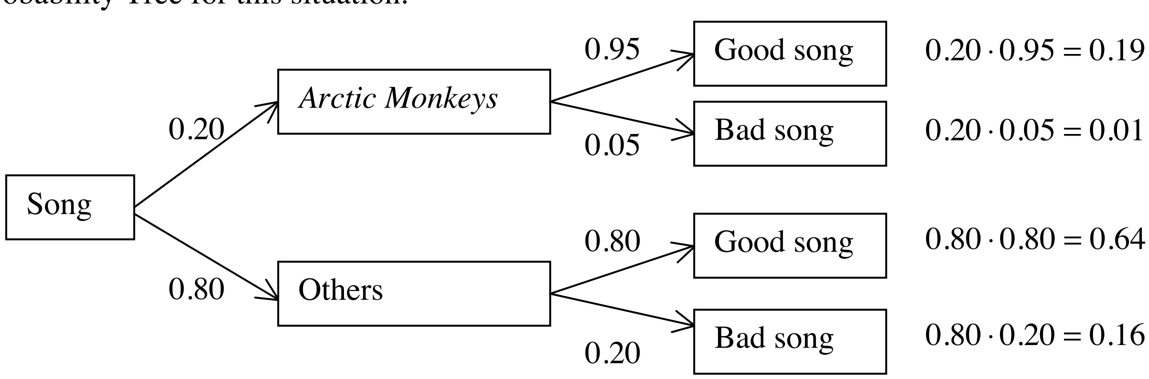
$$\Pr(E|A) = \frac{\Pr(E \cap A)}{\Pr(A)} = \frac{\Pr(A|E) \cdot \Pr(E)}{\Pr(A|E) \cdot \Pr(E) + \Pr(A|E^C) \cdot \Pr(E^C)}.$$

The more general version of the above statement is this very important *Bayes Theorem:*
If events (of positive probability) $E_1, E_2, \ldots, E_n$ *form a partition of the probability space under consideration, and A is an arbitrary event with positive probability, then for any* $i = 1,2,\ldots,n$

$$\Pr(E_i|A) = \frac{\Pr(A|E_i) \cdot \Pr(E_i)}{\sum_{j=1}^{n} \Pr(A|E_j) \cdot \Pr(E_j)} = \frac{\Pr(A|E_i) \cdot \Pr(E_i)}{\Pr(A|E_1) \cdot \Pr(E_1) + \ldots + \Pr(A|E_n) \cdot \Pr(E_n)}.$$

In solving problems involving the Bayes Theorem (also known as the *Bayes Rule)* the key step is to identify and label events and conditional events in an efficient way. In fact, labeling events named in the problem should always be your starting point in any basic probability problem. The typical pattern you should notice in all Bayes Theorem is the “flip-flop”: reversal of the roles of events *A* and E_i in the conditional probability. The values of $\Pr(E_i)$ are called *prior probabilities,* and the values of $\Pr(E_i|A)$ are called *posterior probabilities*. The applications of Bayes Theorem occur in situations in which all $\Pr(E_i)$ and $\Pr(A|E_i)$ probabilities are known, and we are asked to find $\Pr(E_i|A)$ for one of the *i's*.

Problems involving the Bayes Theorem can also be conveniently handled by using a *probability tree* expressing all probabilities involved. Let us illustrate this with an example. Suppose that you have a collection of songs on your iPod and you play them at a party. 20% of your songs are by *Arctic Monkeys,* and 80% by other performers. If you pick a song randomly from *Arctic Monkeys* songs, the probability that it is a good song is 0.95. For other performers, this probability is 0.80. You pick a song randomly from your iPod, play it, and the song turns out to be bad. What is the probability that you picked a song by *Arctic Monkeys*? Consider the figure below, i.e., the Probability Tree for this situation:



The probabilities listed on the right are

$\Pr($A random song is by *Arctic Monkeys* and is good$) = 0.19$,

$\Pr($A random song is by *Arctic Monkeys* and is bad$) = 0.01$,

$$\Pr(\text{A random song is not by } \textit{Arctic Monkeys} \text{ and is good}) = 0.64,$$

$$\Pr(\text{A random song is not by } \textit{Arctic Monkeys} \text{ and is bad}) = 0.16.$$

Therefore,

$$\Pr(\text{A song is by } \textit{Arctic Monkeys} | \text{That song is bad}) = \frac{0.01}{0.01+0.16} = \frac{1}{17}.$$

Notice that we simply take the probability of both things happening (a song by *Arctic Monkeys* and bad) from the right column and put it in the numerator, and the sum of all ways that the song is bad from the right column and put that in the denominator. You can also get the same result by direct application of the Bayes Theorem:

$$\Pr(\text{A song is by } \textit{Arctic Monkeys} | \text{That song is bad}) =$$

$$= \frac{\Pr(\text{Bad song} | \textit{Arctic Monkeys}) \cdot \Pr(\textit{Arctic Monkeys})}{\Pr(\text{Bad song} | \textit{Arctic Monkeys}) \cdot \Pr(\textit{Arctic Monkeys}) + \Pr(\text{Bad song} | \text{Other}) \cdot \Pr(\text{Other})} =$$

$$= \frac{0.05 \cdot 0.20}{0.05 \cdot 0.20 + 0.20 \cdot 0.80} = \frac{0.01}{0.01+0.16} = \frac{1}{17}.$$

Exercise 1.6. May 2003 Course 1 Examination, Problem No. 31, also P Sample Exam Questions, Problem No. 22, also Dr. Ostaszewski's online exercise posted March 1, 2008

A health study tracked a group of persons for five years. At the beginning of the study, 20% were classified as heavy smokers, 30% as light smokers, and 50% as nonsmokers. Results of the study showed that light smokers were twice as likely as nonsmokers to die during the five-year study, but only half as likely as heavy smokers. A randomly selected participant from the study died over the five-year period. Calculate the probability that the participant was a heavy smoker.

A. 0.20 B. 0.25 C. 0.35 D. 0.42 E. 0.57

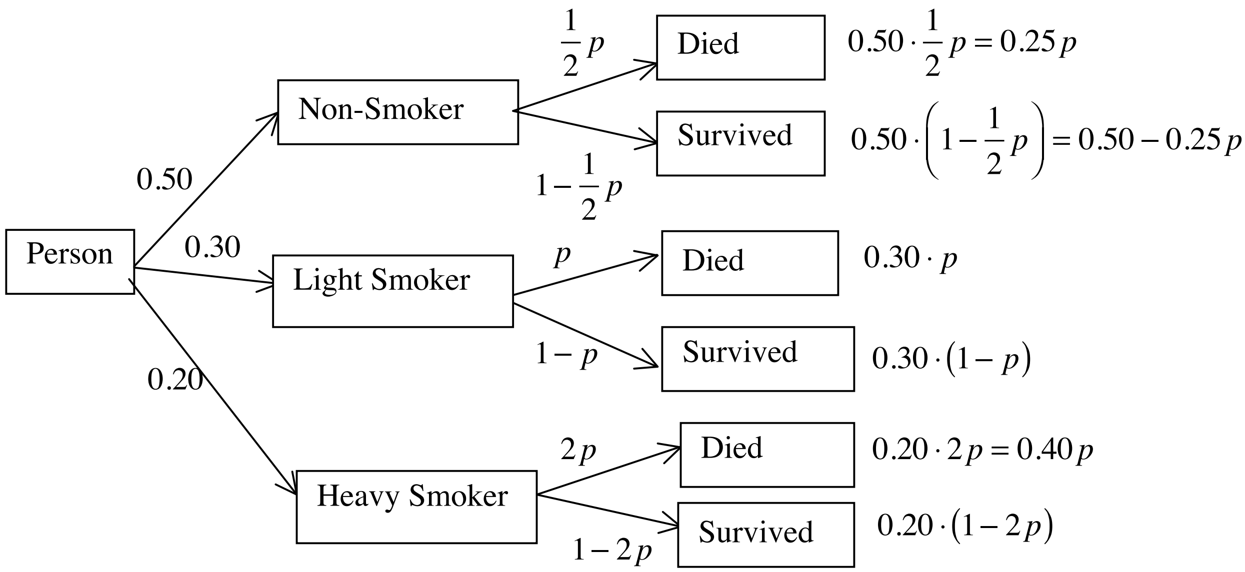
Solution.

Let H be the event of studying a heavy smoker, L be the event of studying a light smoker, and N be the event of studying a non-smoker. We are given that $\Pr(H) = 0.20$, $\Pr(L) = 0.30$, and $\Pr(N) = 0.50$. Additionally, let D be the event of a death within five-year period. We know that $\Pr(D|L) = 2\Pr(D|N)$ and $\Pr(D|L) = \frac{1}{2}\Pr(D|H)$. We are looking for $\Pr(H|D)$. Using the Bayes Theorem, we conclude:

$$\Pr(H|D) = \frac{\Pr(D|H) \cdot \Pr(H)}{\Pr(D|N) \cdot \Pr(N) + \Pr(D|L) \cdot \Pr(L) + \Pr(D|H) \cdot \Pr(H)} =$$

$$= \frac{2\Pr(D|L) \cdot 0.2}{\frac{1}{2}\Pr(D|L) \cdot 0.5 + \Pr(D|L) \cdot 0.3 + 2\Pr(D|L) \cdot 0.2} = \frac{0.4}{0.25+0.3+0.4} \approx 0.4211.$$

Answer D. Alternatively, we can draw a Probability Tree, as shown in the figure below:



Therefore,

$$\Pr(H|D) = \frac{0.40p}{0.25p+0.30p+0.40p} = \frac{0.4}{0.25+0.3+0.4} \approx 0.4211.$$

Answer D, again.

Exercise 1.7. November 2001 Course 1 Examination, Problem No. 4, also P Sample Exam Questions, Problem No. 21, and Dr. Ostaszewski's online exercise posted February 23, 2008
Upon arrival at a hospital's emergency room, patients are categorized according to their condition as critical, serious, or stable. In the past year:
(i) 10% of the emergency room patients were critical;
(ii) 30% of the emergency room patients were serious;
(iii) The rest of the emergency room patients were stable;
(iv) 40% of the critical patients died;
(v) 10% of the serious patients died; and
(vi) 1% of the stable patients died.
Given that a patient survived, what is the probability that the patient was categorized as serious upon arrival?

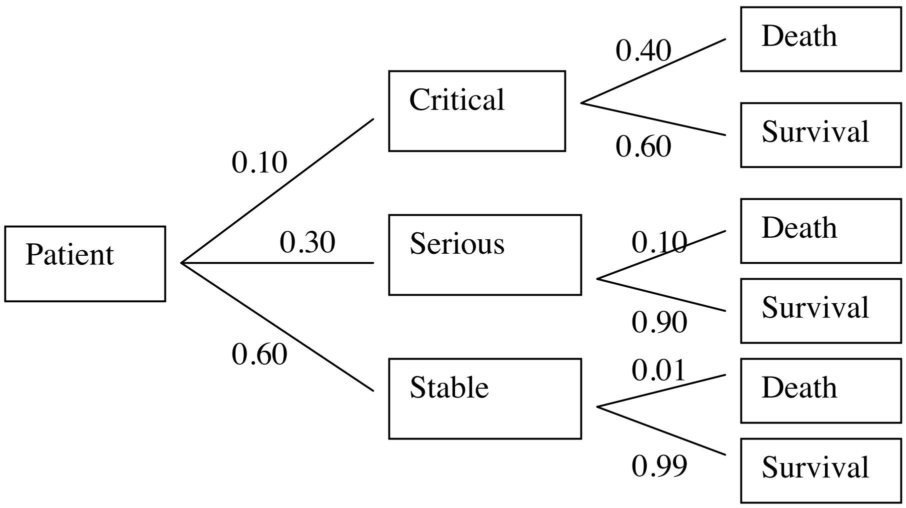
A. 0.06 B. 0.29 C. 0.30 D. 0.39 E. 0.64

Solution.
Let U be the event that a patient survived, and E_S be the event that a patient was classified as serious upon arrival, E_C -- the event that a patient was critical, and E_T -- the event that the patient was stable. We apply the Bayes Theorem:

$$\Pr(E_S|U) = \frac{\Pr(U|E_S)\cdot\Pr(E_S)}{\Pr(U|E_C)\cdot\Pr(E_C)+\Pr(U|E_S)\cdot\Pr(E_S)+\Pr(U|E_T)\cdot\Pr(E_T)} =$$

$$= \frac{0.9\cdot 0.3}{0.6\cdot 0.1+0.9\cdot 0.3+0.99\cdot 0.6} \approx 0.2922.$$

Answer B. This answer could have been also derived using the probability tree shown below:



Exercise 1.8. May 2003 Course 1 Examination, Problem No. 8, also P Sample Exam Questions, Problem No. 19, and Dr. Ostaszewski's online exercise posted January 12, 2008
An auto insurance company insures drivers of all ages. An actuary compiled the following statistics on the company's insured drivers:

Age of Driver	Probability of Accident	Portion of Company's Insured Drivers
16-20	0.06	0.08
21-30	0.03	0.15
31-65	0.02	0.49
66-99	0.04	0.28

A randomly selected driver that the company insures has an accident. Calculate the probability that the driver was age 16-20.

A. 0.13 B. 0.16 C. 0.19 D. 0.23 E. 0.40

Solution.
This is a standard application of the Bayes Theorem. Let A be the event of an insured driver having an accident, and let

B_1 = Event: driver's age is in the range 16-20,
B_2 = Event: driver's age is in the range 21-30,
B_3 = Event: driver's age is in the range 30-65,
B_4 = Event: driver's age is in the range 66-99.

Then

$$\Pr(B_1|A) = \frac{\Pr(A|B_1)\Pr(B_1)}{\Pr(A|B_1)\Pr(B_1) + \Pr(A|B_2)\Pr(B_2) + \Pr(A|B_3)\Pr(B_3) + \Pr(A|B_4)\Pr(B_4)} =$$

$$= \frac{0.06 \cdot 0.08}{0.06 \cdot 0.08 + 0.03 \cdot 0.15 + 0.02 \cdot 0.49 + 0.04 \cdot 0.28} = 0.1584.$$

Answer B.

Combinatorial probability

If we have a set with *n* elements, there are *n* ways to pick one of them to be the first one, then $n-1$ ways to pick another one of them to be number two, etc., until there is only one left. This implies that there are $n \cdot (n-1) \cdot \ldots \cdot 1$ ways to put all the elements of this set in order. The expression $n \cdot (n-1) \cdot \ldots \cdot 1$ is denoted by $n!$ and termed *n-factorial*. The orderings of a given set of *n* elements are called *permutations*. In general, a *permutation* is an ordered sample from a given set, not necessarily containing all elements from it. If we have a set with *n* elements, and we want to pick an ordered sample of size *k,* then we have *n* – *k* elements remaining, and since all of their orderings do not matter, the total number of such ordered *k*-samples (i.e., permutations) possible is $\frac{n!}{(n-k)!}$.

A *combination* is an unordered sample (without replacement) from a given finite set, i.e., its subset. Given a set with *n* elements, there are $\frac{n!}{(n-k)!}$ ordered *k*-samples that can be picked from it. But a combination does not care what the order is, and since *k* elements can be ordered in $k!$ ways, the number of combinations (i.e., subsets) of size *k* that can be picked is reduced by the factor of $k!$ and is therefore equal to: $\frac{n!}{k!(n-k)!}$. This expression $\frac{n!}{k!(n-k)!}$ is written usually as $\binom{n}{k}$ and read as *n choose k*.

In addition to unordered samples without replacement, i.e., combinations, we often consider *samples with replacement,* which are, as the names indicates, samples obtained by taking elements of a finite set, with elements returned to the set after they have been picked. Those problems are best handled by common sense and some practice. Every time an element is picked this way, its chances of being picked are simply the ratio of the number of elements of its type to the total number of elements in the set.

There are also some other types of combinatorial principles that might be useful in probability calculations that we will list now.

Given *n* objects, of which n_1 are of type 1, n_2 are of type 2, …, and n_m are of type *m,* with $n_1+n_2+\ldots+n_m=n,$ the number of ways to order all *n* objects, with objects of each type indistinguishable from other objects of the same type, is $\frac{n!}{n_1!n_2!\ldots n_m!}$, sometimes denoted by $\begin{pmatrix} & n & & \\ n_1 & n_2 & \ldots & n_m \end{pmatrix}$.

Given *n* objects, of which n_1 are of type 1, n_2 are of type 2, …, and n_m are of type *m,* with $n_1+n_2+\ldots+n_m=n$, the number of ways to choose a subset of size $k\le n$ (without replacement), with k_1 objects of type 1, k_2 objects of type 2, …, and k_m are of type *m,* with $k_1+k_2+\ldots+k_m=k,$ is $\binom{n_1}{k_1}\cdot\binom{n_2}{k_2}\cdot\ldots\cdot\binom{n_m}{k_m}$.

The $\binom{n}{k}$ expression plays a special role in the *Binomial Theorem,* which states that

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}.$$

We also have this general application of $\binom{n}{k}$: In the power series expansion of $(1+t)^n$ the coefficient of t^k is $\binom{n}{k}$. There is also the following multivariate version: In the power series expansion of $(t_1+t_2+\ldots+t_m)^n$ the coefficient of $t_1^{n_1}\cdot t_2^{n_2}\cdot\ldots\cdot t_m^{n_m}$, where $n_1+n_2+\ldots+n_m=n,$ is $\begin{pmatrix} & n & & \\ n_1 & n_2 & \ldots & n_m \end{pmatrix}$.

The combinatorial principles are very useful in calculating probabilities. The probability of the outcome desired is calculated as the ratio of the number of favorable outcomes to the total number of outcomes. We will illustrate this in some exercises below.

Exercise 1.9. November 2001 Course 1 Examination, Problem No. 1, also P Sample Exam Questions, Problem No. 4, and Dr. Ostaszewski's online exercise posted June 30, 2007
An urn contains 10 balls: 4 red and 6 blue. A second urn contains 16 red balls and an unknown number of blue balls. A single ball is drawn from each urn. The probability that both balls are the same color is 0.44. Calculate the number of blue balls in the second urn.

A. 4 B. 20 C. 24 D. 44 E. 64

Solution.

For $i = 1,2$ let R_i be the event that a red ball is drawn from urn i, and B_i be the event that a blue ball is drawn from urn i. Then, if x is the number of blue balls in urn 2, we can assume that drawings from two different urns are independent and obtain

$$0.44 = \Pr\left(\left(R_1 \cap R_2\right) \cup \left(B_1 \cap B_2\right)\right) = \Pr\left(R_1 \cap R_2\right) + \Pr\left(B_1 \cap B_2\right) = \Pr\left(R_1\right) \cdot \Pr\left(R_2\right) +$$

$$+ \Pr\left(B_1\right) \cdot \Pr\left(B_2\right) = \frac{4}{10} \cdot \frac{16}{x+16} + \frac{6}{10} \cdot \frac{x}{x+16} = \frac{1}{5} \cdot \left(\frac{32}{x+16} + \frac{3x}{x+16}\right).$$

Therefore, by multiplying both sides by 5 we get

$$2.2 = \frac{32}{x+16} + \frac{3x}{x+16} = \frac{3x+32}{x+16}.$$

This can be immediately turned into an easy linear equation, whose solution is $x = 4$.

Answer A.

Exercise 1.10. February 1996 Course 110 Examination, Problem No. 1, also Dr. Ostaszewski's online exercise posted June 19, 2010

A box contains 10 balls, of which 3 are red, 2 are yellow, and 5 are blue. Five balls are randomly selected with replacement. Calculate the probability that fewer than 2 of the selected balls are red.

A. 0.3601 B. 0.5000 C. 0.5282 D. 0.8369 E. 0.9167

Solution.

This problem uses sampling with replacement. The general principle of combinatorial probability is that in order to find probability sought, we need to take the ratio of the number of outcomes giving our desired result to the total number of outcomes. The total number of outcomes is always calculated more easily, and in this case, we have 10 ways to choose the first ball, again 10 ways to choose the second one (as the choice is made with replacement), etc. Thus the total number of outcomes is 10^5. We are interested in outcomes that give us only one red ball, or no red balls. The case of no red balls is easy: in such a situation we simply choose from the seven non-red balls five times, and the total number of such outcomes is 7^5. Now let us look at the case of exactly one red ball. Suppose that the only red ball chosen is the very first one. Then we have three choices in the first selection, and 7^4 choices in the remaining selections, for a total of $3 \cdot 7^4$, as the consecutive selections are independent. But the red ball could be in any of the five spots, not just the first one. This raises the total number of outcomes with only one red ball to $5 \cdot 3 \cdot 7^4$. Therefore, the total number of outcomes giving the desired result (no red balls or one red ball) is

$$5 \cdot 3 \cdot 7^4 + 7^5 = 15 \cdot 7^4 + 7 \cdot 7^4 = 22 \cdot 7^4.$$

The desired probability is

$$\frac{22 \cdot 7^4}{10^5} = 2.2 \cdot 0.7^4 \approx 0.52822.$$

Answer C. Note that this problem can be also solved using the methodologies of a later section in this manual, by treating each ball drawing is a Bernoulli Trial with probability of success being 0.3 (3 red balls out of 10), and the number of successes X in five drawings as having the binomial distribution with $n = 5, p = 0.3$. Then

$$\Pr(X<2)=\Pr(X=0)+\Pr(X=1)=\binom{5}{0}\cdot 0.3^0\cdot 0.7^5+\binom{5}{1}\cdot 0.3^1\cdot 0.7^4=$$

$$=\frac{7^5}{10^5}+\frac{5\cdot 3\cdot 7^4}{10^5}=\frac{22\cdot 7^4}{10^5}=2.2\cdot 0.7^4\approx 0.52822.$$

Answer C, again.

Exercise 1.11. February 1996 Course 110 Examination, Problem No. 7, also Dr. Ostaszewski's online exercise posted June 26, 2010

A class contains 8 boys and 7 girls. The teacher selects 3 of the children at random and without replacement. Calculate the probability that the number of boys selected exceeds the number of girls selected.

A. $\frac{512}{3375}$ B. $\frac{28}{65}$ C. $\frac{8}{15}$ D. $\frac{1856}{3375}$ E. $\frac{36}{65}$

Solution.

The number of boys selected exceeds the number of girls selected if there are two or three boys in the group selected. First, the total number of outcomes is the total number of ways to choose 3 children out of 15, without consideration for order, and that is $\binom{15}{3}$. If we choose two boys and one girl, there are $\binom{8}{2}$ ways to choose the boys and $\binom{7}{1}$ ways to choose the girl, for a total of $\binom{8}{2}\cdot\binom{7}{1}$. If we choose three boys, there are $\binom{8}{3}$ ways to pick them, and $\binom{7}{0}$ ways to choose no girls. Thus the desired probability is

$$\frac{\binom{8}{2}\cdot\binom{7}{1}+\binom{8}{3}\cdot\binom{7}{0}}{\binom{15}{3}}=\frac{28\cdot 7+56\cdot 1}{455}=\frac{196+56}{455}=\frac{252}{455}=\frac{36}{65}.$$

Answer E.

Exercise 1.12. May 1983 Course 110 Examination, Problem No. 39, also Dr. Ostaszewski's online exercise posted July 3, 2010

A box contains 10 white marbles and 15 black marbles. If 10 marbles are selected at random and without replacement, what is the probability that x of the 10 marbles are white for $x = 0, 1, \ldots,$

10?

A. $\frac{x}{10}$ B. $\binom{10}{x}\left(\frac{2}{5}\right)^{x}\left(\frac{3}{5}\right)^{10-x}$ C. $\frac{\binom{10}{x}\binom{15}{10-x}}{\binom{25}{10}}$ D. $\frac{\binom{10}{x}}{\binom{25}{10}}$ E. $\frac{\binom{10}{x}}{\binom{25}{x}}$

Solution.

The total number of ways to pick 10 marbles out of 25 is $\binom{25}{10}$. This is the total number of possible outcomes. How many favorable outcomes are there? There are $\binom{10}{x}$ ways to choose x white marbles from 10, and $\binom{15}{10-x}$ ways to choose $10-x$ black marbles out of 15. This gives a total number of favorable outcomes as $\binom{10}{x}\cdot\binom{15}{10-x}$, and the desired probability as

$$\frac{\binom{10}{x}\cdot\binom{15}{10-x}}{\binom{25}{10}}.$$

Answer C.