CollegeBoard

AP® Statistics Practice Exam

From the 2016 Administration

NOTE: This is a modified version of the 2016 AP Statistics Exam.

Contents

Note: This publication shows the page numbers that appeared in the *2015–16 AP Exam Instructions* book and in the actual exam. This publication was not repaginated to begin with page 1.

Exam Instructions

The following contains instructions taken from the *2015–16 AP Exam Instructions* book.

AP® Statistics Exam

Regularly Scheduled Exam Date: Thursday afternoon, May 12, 2016

Late-Testing Exam Date: Wednesday morning, May 18, 2016

Section I Total Time: 1 hr. 30 min. Section II Total Time: 1 hr. 30 min.

Section I

Total Time: 1 hour 30 minutes
Graphing calculator expected
Number of Questions: 40*
Percent of Total Score: 50%
Writing Instrument: Pencil required

**The number of questions may vary slightly depending on the form of the exam.*

Section II

Total Time: 1 hour 30 minutes
Graphing calculator expected
Number of Questions: 6
Percent of Total Score: 50%
Writing Instrument: Either pencil or pen with black or dark blue ink

What Proctors Need to Bring to This Exam

- Exam packets
- Answer sheets
- AP Student Packs
- *2015-16 AP Coordinator's Manual*
- This book — *AP Exam Instructions*
- AP Exam Seating Chart template(s)
- School Code and Home-School/Self-Study Codes
- Extra graphing calculators
- Pencil sharpener
- Container for students' electronic devices (if needed)
- Extra No. 2 pencils with erasers
- Extra pens with black or dark blue ink
- Extra paper
- Stapler
- Watch
- Signs for the door to the testing room
  - "Exam in Progress"
  - "Cell phones are prohibited in the testing room"

Students are expected to bring graphing calculators with statistical capabilities to the AP Statistics Exam. Nongraphing scientific calculators are permitted as long as they have the required computational capabilities. Before starting the exam administration, make sure each student has a graphing calculator from the approved list on page 47 of the *2015-16 AP Coordinator's Manual* or a scientific calculator. It is up to the student to determine if a nongraphing scientific calculator has the required computational capabilities. If a student does not have a graphing calculator from the approved list or an appropriate scientific calculator, you may provide one from your supply. See pages 44–47 of the *AP Coordinator's Manual* for more information. If the student does not want to use the calculator you provide, or does not want to use a calculator at all, he or she must hand copy, date, and sign the release statement on page 45 of the *AP Coordinator's Manual*.

Students may have no more than two calculators on their desks. Calculators may not be shared. Calculator memories do not need to be cleared before or after the exam. Students with Hewlett-Packard 48–50 Series and Casio FX-9860 graphing calculators may use cards designed for use with these calculators. Proctors should make sure infrared ports (Hewlett-Packard) are not facing each other. **Since graphing calculators can be used to store data, including text, proctors should monitor that students are using their calculators appropriately. Attempts by students**

to use the calculator to remove exam questions and/or answers from the room may result in the cancellation of AP Exam scores.

SECTION I: Multiple Choice

Do not begin the exam instructions below until you have completed the appropriate General Instructions for your group.

Make sure you begin the exam at the designated time. Remember, you must complete a seating chart for this exam. See pages 305–306 for a seating chart template and instructions. See the *2015-16 AP Coordinator's Manual* for exam seating requirements (pages 49–52).

If you are giving the regularly scheduled exam, say:

It is Thursday afternoon, May 12, and you will be taking the AP Statistics Exam.

If you are giving the alternate exam for late testing, say:

It is Wednesday morning, May 18, and you will be taking the AP Statistics Exam.

In a moment, you will open the packet that contains your exam materials. By opening this packet, you agree to all of the AP Program's policies and procedures outlined in the *2015-16 Bulletin for AP Students and Parents*. You may now remove the shrinkwrap from your exam packet and take out the Section I booklet, but do not open the booklet or the shrinkwrapped Section II materials. Put the white seals aside. . . .

Carefully remove the AP Exam label found near the top left of your exam booklet cover. Now place it on page 1 of your answer sheet on the light blue box near the top right-hand corner that reads "AP Exam Label."

If students accidentally place the exam label in the space for the number label or vice versa, advise them to leave the labels in place. They should not try to remove the label; their exam can still be processed correctly.

Read the statements on the front cover of Section I and look up when you have finished. . . .

Sign your name, and write today's date. Look up when you have finished. . . .

Now print your full legal name where indicated. Are there any questions? . . .

Turn to the back cover of your exam booklet and read it completely. Look up when you have finished. . . .

Are there any questions? . . .

You will now take the multiple-choice portion of the exam. You should have in front of you the multiple-choice booklet and your answer sheet. Open your answer sheet to page 2. You may never discuss these specific multiple-choice questions at any time in any form with anyone, including your teacher and other students. If you disclose these questions through any means, your AP Exam score will be canceled.

You must complete the answer sheet using a No. 2 pencil only. Mark all of your responses beginning on page 2 of your answer sheet, one response per question. Completely fill in the circles. If you need to erase, do so carefully and completely. No credit will be given for anything written in the exam booklet. Scratch paper is not allowed, but you may use the margins or any blank space in the exam booklet for scratch work. Calculators may be used for both sections of this exam. You may place your calculators on your desk. Are there any questions? . . .

You have 1 hour and 30 minutes for this section. Open your Section I booklet and begin.

Note Start Time here _____. Note Stop Time here _____. Check that students are marking their answers in pencil on their answer sheets and that they are not looking at their shrinkwrapped Section II booklets. Proctors should walk around and make sure Hewlett-Packard calculators' infrared ports are not facing each other and that students are not sharing calculators. After 1 hour and 20 minutes, say:

There are 10 minutes remaining.

After 10 minutes, say:

Stop working. Close your booklet and put your answer sheet on your desk, face up. Make sure you have your AP number label and an AP Exam label on page 1 of your answer sheet. Sit quietly while I collect your answer sheets.

Collect an answer sheet from each student. Check that each answer sheet has an AP number label and an AP Exam label. After all answer sheets have been collected, say:

Now you must seal your exam booklet using the white seals you set aside earlier. Remove the white seals from the backing and press one on each area of your exam booklet cover marked "PLACE SEAL HERE." Fold each seal over the back cover. When you have finished, place the booklet on your desk, face up. I will now collect your Section I booklet. . . .

Collect a Section I booklet from each student. Check that each student has signed the front cover of the sealed Section I booklet.

There is a 10-minute break between Sections I and II. When all Section I materials have been collected and accounted for and you are ready for the break, say:

Please listen carefully to these instructions before we take a 10-minute break. All items you placed under your chair at the beginning of this exam must stay there, and you are not permitted to open or access them in any way. Leave your shrinkwrapped Section II packet on your desk during the break. You are not allowed to consult teachers, other students, notes, or textbooks during the break. You may not make phone calls, send text messages, use your calculators, check email, use a social networking site, or access any electronic or communication device. Remember, you may never discuss the multiple-choice questions at any time in any form with anyone, including your teacher and other students. If you disclose these questions through any means, your AP Exam score will be canceled. Are there any questions? . . .

You may begin your break. Testing will resume at _____.

SECTION II: Free Response

After the break, say:

May I have everyone's attention? Place your Student Pack on your desk. . . .

You may now remove the shrinkwrap from the Section II packet, but do not open the exam booklet until you are told to do so. . . .

Read the bulleted statements on the front cover of the exam booklet. Look up when you have finished. . . .

Now take an AP number label from your Student Pack and place it on the shaded box. If you don't have any AP number labels, write your AP number in the box. Look up when you have finished. . . .

Read the last statement. . . .

Using a pen with black or dark blue ink, print the first, middle, and last initials of your legal name in the boxes and print today's date where indicated. This constitutes your signature and your agreement to the statements on the front cover. . . .

Turn to the back cover and, using your pen, complete Item 1 under "Important Identification Information." Print the first two letters of your last name and the first letter of your first name in the boxes. Look up when you have finished. . . .

In Item 2, print your date of birth in the boxes. . . .

In Item 3, write the school code you printed on the front of your Student Pack in the boxes. . . .

Read Item 4. . . .

Are there any questions? . . .

I need to collect the Student Pack from anyone who will be taking another AP Exam. You may keep it only if you are not taking any other AP Exams this year. If you have no other AP Exams to take, place your Student Pack under your chair now. . . .

Read the information on the back cover of the exam booklet. Do not open the booklet until you are told to do so. Look up when you have finished. . . .

Collect the Student Packs. Then say:

Are there any questions? . . .

Section II has two parts. You have 1 hour and 30 minutes to complete all of Section II. You are responsible for pacing yourself and may proceed freely from one part to the next. You must write your answers in the exam booklet using a pen with black or dark blue ink or a No. 2 pencil. If you use a pencil, be sure that your writing is dark enough to be easily read. If you need more paper during the exam, raise your hand. At the top of each extra sheet of paper you use, be sure to write only your AP number and the question number you are working on. Do not write your name. Are there any questions? . . .

You may begin Part A.

Note Start Time here _____. Note Stop Time here _____. You should also make sure that Hewlett-Packard calculators' infrared ports are not facing each other and that students are not sharing calculators. After 1 hour and 5 minutes, say:

There are 25 minutes remaining and you may want to move on to Part B, if you have not already started answering that question.

After 15 minutes, say:

There are 10 minutes remaining.

After 10 minutes, say:

Stop working and close your exam booklet. Place it on your desk, face up. . . .

If any students used extra paper for a question in the free-response section, have those students staple the extra sheet(s) to the first page corresponding to that question in their exam booklets. Complete an Incident Report. A single Incident Report may be completed for multiple students per exam subject per administration (regular or late testing) as long as all of the required information is provided. Include all exam booklets with extra sheets of paper in an Incident Report return envelope (see page 60 of the *2015-16 AP Coordinator's Manual* for complete details). Then say:

Remain in your seat, without talking, while the exam materials are collected. . . .

Collect a Section II booklet from each student. Check for the following:

- Exam booklet front cover: The student placed an AP number label on the shaded box and printed his or her initials and today's date.
- Exam booklet back cover: The student completed the "Important Identification Information" area.

When all exam materials have been collected and accounted for, return to students any electronic devices you may have collected before the start of the exam.

If you are giving the regularly scheduled exam, say:

You may not discuss or share these specific free-response questions with anyone unless they are released on the College Board website in about two days. Your AP Exam score results will be available online in July.

If you are giving the alternate exam for late testing, say:

None of the questions in this exam may ever be discussed or shared in any way at any time. Your AP Exam score results will be available online in July.

If any students completed the AP number card at the beginning of this exam, say:

Please remember to take your AP number card with you. You will need the information on this card to view your scores and order AP score reporting services online.

Then say:

You are now dismissed.

All exam materials must be placed in secure storage until they are returned to the AP Program after your school's last administration. Before storing materials, check the "School Use Only" section on page 1 of the answer sheet and:

- Fill in the appropriate section number circle in order to access a separate AP Instructional Planning Report (for regularly scheduled exams only) or subject score roster at the class section or teacher level. See "Post-Exam Activities" in the *2015-16 AP Coordinator's Manual*.
- Check your list of students who are eligible for fee reductions and fill in the appropriate circle on their registration answer sheets.

Be sure to give the completed seating chart to the AP Coordinator. Schools must retain seating charts for at least six months (unless the state or district requires that they be retained for a longer period of time). Schools should not return any seating charts in their exam shipments unless they are required as part of an Incident Report.

Student Answer Sheet for the Multiple-Choice Section

Use this section to capture student responses. (Note that the following answer sheet is a sample, and may differ from one used in an actual exam.)

AP Answer Sheet 2016

AP Number Label
(from Student Pack)

B123456789T

AP Exam Label
(from Section I Booklet)

CollegeBoard

COMPLETE THIS AREA AT EVERY EXAM. **USE NO. 2 PENCIL ONLY**

To maintain the security of the exam and the validity of my AP score, I will allow no one else to see the multiple-choice questions. I will seal the multiple-choice booklet when asked to do so, and I will not discuss these questions with anyone at any time after completing the section. I am aware of and agree to the AP Program's policies and procedures as outlined in the *2015-16 Bulletin for AP Students and Parents*, including using testing accommodations (e.g., extended time, computer, etc.) only if I have been preapproved by College Board Services for Students with Disabilities.

A. SIGNATURE Sign your legal name as it will appear on your college applications. Date

B. LEGAL NAME Omit apostrophes, Jr., II.

Legal Last Name — First 15 Letters | Legal First Name — First 12 Letters | MI

C. YOUR AP NUMBER

D. EXAM DATE Month | Day

E. EXAM START TIME AM | PM

F. MULTIPLE-CHOICE BOOKLET SERIAL NUMBER S

G. ONLINE PROVIDER CODE

H. AP EXAM I AM TAKING USING THIS ANSWER SHEET Exam Name: | Form: | Form Code:

COMPLETE THIS AREA ONLY ONCE.

I. AREA CODE AND PHONE NUMBER

INTERNATIONAL PHONE

J. SCHOOL YOU ATTEND

SCHOOL CODE | School Name | City | State | Country

K. DATE OF BIRTH

Month (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec) | Day | Year

L. SOCIAL SECURITY NUMBER (Optional)

M. COLLEGE TO RECEIVE YOUR AP SCORE REPORT

COLLEGE CODE

Using the college code listed in the AP Student Pack, indicate the **ONE** college that you want to receive your AP score report.

College Name | City | State | Country

N. CURRENT GRADE LEVEL

- Not yet in 9th grade
- 9th
- 10th
- 11th
- 12th
- No longer in high school

O. STUDENT SEARCH SERVICE®

Colleges and scholarship programs may request your information to inform you of educational opportunities and financial aid.

Would you like us to supply your information?

- Yes
- No

If you don't answer and previously chose to participate in this service, we will continue providing your information.

785208

SCHOOL USE ONLY

Section Number: 1 2 3 4 5 6 7 8 9

Fee Reduction Granted: 1 Option 1 2 Option 2

108603-00657 • TF216E47575 • Printed in U.S.A.
Q4068/1-4

COMPLETE THIS AREA AT EACH EXAM (IF APPLICABLE).

P. SURVEY QUESTIONS — Answer the survey questions in the AP Student Pack. Do not put responses to exam questions in this section.

1	A B C D E F G H I	4	A B C D E F G H I	7	A B C D E F G H I
2	A B C D E F G H I	5	A B C D E F G H I	8	A B C D E F G H I
3	A B C D E F G H I	6	A B C D E F G H I	9	A B C D E F G H I

Q. LANGUAGE — Do not complete this section unless instructed to do so.

If this answer sheet is for the French Language and Culture, German Language and Culture, Italian Language and Culture, Spanish Language and Culture, or Spanish Literature and Culture Exam, please answer the following questions. Your responses will not affect your score.

1. Have you lived or studied for one month or more in a country where the language of the exam you are now taking is spoken?

○ Yes ○ No

2. Do you regularly speak or hear the language at home?

○ Yes ○ No

QUESTIONS 1–75

Indicate your answers to the exam questions in this section (pages 2 and 3). Mark only one response per question for Questions 1 through 120. If a question has only four answer options, do not mark option E. Answers written in the multiple-choice booklet will not be scored.

COMPLETE MARK ●

EXAMPLES OF INCOMPLETE MARKS 

You must use a No. 2 pencil and marks must be complete. Do not use a mechanical pencil. It is very important that you fill in the entire circle darkly and completely. If you change your response, erase as completely as possible. Incomplete marks or erasures may affect your score.

1	A B C D E	26	A B C D E	51	A B C D E
2	A B C D E	27	A B C D E	52	A B C D E
3	A B C D E	28	A B C D E	53	A B C D E
4	A B C D E	29	A B C D E	54	A B C D E
5	A B C D E	30	A B C D E	55	A B C D E
6	A B C D E	31	A B C D E	56	A B C D E
7	A B C D E	32	A B C D E	57	A B C D E
8	A B C D E	33	A B C D E	58	A B C D E
9	A B C D E	34	A B C D E	59	A B C D E
10	A B C D E	35	A B C D E	60	A B C D E
11	A B C D E	36	A B C D E	61	A B C D E
12	A B C D E	37	A B C D E	62	A B C D E
13	A B C D E	38	A B C D E	63	A B C D E
14	A B C D E	39	A B C D E	64	A B C D E
15	A B C D E	40	A B C D E	65	A B C D E
16	A B C D E	41	A B C D E	66	A B C D E
17	A B C D E	42	A B C D E	67	A B C D E
18	A B C D E	43	A B C D E	68	A B C D E
19	A B C D E	44	A B C D E	69	A B C D E
20	A B C D E	45	A B C D E	70	A B C D E
21	A B C D E	46	A B C D E	71	A B C D E
22	A B C D E	47	A B C D E	72	A B C D E
23	A B C D E	48	A B C D E	73	A B C D E
24	A B C D E	49	A B C D E	74	A B C D E
25	A B C D E	50	A B C D E	75	A B C D E

ETS USE ONLY

Exam		0 1 2 3 4 5 6 7 8 9
		0 1 2 3 4 5 6 7 8 9
Exam		0 1 2 3 4 5 6 7 8 9
		0 1 2 3 4 5 6 7 8 9

SELECTED MEDIA EXAMS	R	W	O	OTHER EXAMS	R	W	O
PT02				TOTAL			
PT03				Subscore (if applicable)			
PT04				Subscore (if applicable)			

DO NOT WRITE IN THIS AREA

QUESTIONS 76–120

Be sure each mark is dark and completely fills the circle. If a question has only four answer options, do not mark option E.

76	A B C D E	91	A B C D E	106	A B C D E
77	A B C D E	92	A B C D E	107	A B C D E
78	A B C D E	93	A B C D E	108	A B C D E
79	A B C D E	94	A B C D E	109	A B C D E
80	A B C D E	95	A B C D E	110	A B C D E
81	A B C D E	96	A B C D E	111	A B C D E
82	A B C D E	97	A B C D E	112	A B C D E
83	A B C D E	98	A B C D E	113	A B C D E
84	A B C D E	99	A B C D E	114	A B C D E
85	A B C D E	100	A B C D E	115	A B C D E
86	A B C D E	101	A B C D E	116	A B C D E
87	A B C D E	102	A B C D E	117	A B C D E
88	A B C D E	103	A B C D E	118	A B C D E
89	A B C D E	104	A B C D E	119	A B C D E
90	A B C D E	105	A B C D E	120	A B C D E

QUESTIONS 121–126

For Students Taking AP Biology

Write your answer in the boxes at the top of the griddable area and fill in the corresponding circles. Mark only one circle in any column. You will receive credit only if the circles are filled in correctly.

121 122 123 124 125 126

QUESTIONS 131–142

For Students Taking AP Physics 1 or AP Physics 2

Mark two responses per question. You will receive credit only if both correct responses are selected.

131	A B C D	135	A B C D	139	A B C D
132	A B C D	136	A B C D	140	A B C D
133	A B C D	137	A B C D	141	A B C D
134	A B C D	138	A B C D	142	A B C D

DO NOT WRITE IN THIS AREA

COMPLETE THIS AREA ONLY ONCE.

R. YOUR MAILING ADDRESS Use the address abbreviations from your AP Student Pack. Fill in only one circle per column. Indicate a space in your address by leaving a blank box; do not grid that column.

STREET ADDRESS (include street number, street name, apartment number, etc.)

CITY

ZIP OR POSTAL CODE

COUNTRY CODE

STATE

AK	HI	MI	NY	VT	
AL	IA	MN	OH	WA	
AR	ID	MS	OR	WV	
AZ	IL	MT	PA	WY	
CA	IN	NC	RI	Puerto Rico	
CO	KS	ND	SC		
CT	KY	NE	SD	AA	
DC	LA	NH	TN	AE	
DE	MA	NJ	TX	AP	
FL	MD	NM	UT	Other	
GA	ME	NV	VA		

V. SEX

- Female
- Male

W. WHICH LANGUAGE DO YOU KNOW BEST?

- English
- English and another language about the same
- Another language

X. RACIAL/ETHNIC GROUP

Please answer both questions about Hispanic origin and about race. For the following questions about your identity, Hispanic origins are not races.

(You may mark all that apply.)

a. Are you of Hispanic, Latino, or Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Cuban
- Yes, Mexican
- Yes, Puerto Rican
- Yes, another Hispanic, Latino, or Spanish origin

b. What is your race?

- American Indian or Alaska Native
- Asian (including Indian subcontinent and Philippines origin)
- Black or African American (including Africa and Afro-Caribbean origin)
- Native Hawaiian or other Pacific Islander
- White (including Middle Eastern origin)

Y. PARENTAL EDUCATION LEVEL

In the first column, indicate the highest level of education of one parent/guardian, and indicate whether this is your mother/female guardian or father/male guardian. Then, if applicable, indicate the highest level of education of your other parent/guardian in the second column, and indicate whether this is your mother/female guardian or father/male guardian.

Mother or female guardian	
Father or male guardian	
Grade school	
Some high school	
High school diploma or equivalent	
Vocational or trade school	
Some college	
Associate or two-year degree	
Bachelor's or four-year degree	
Some graduate or professional school	
Graduate or professional degree	

S. FOR STUDENTS OUTSIDE THE UNITED STATES ONLY If your address does not fit in the spaces provided in Item R, fill in as many circles as you can, then fill in the circle in Item S and print the remainder of your address in the spaces provided.

T. STUDENT IDENTIFIER (Student ID Number)

Address	City	State or Province	Country	ZIP or Postal Code

U. EMAIL ADDRESS By providing your email address, you are granting the College Board permission to use your email in accordance with the policies in the *2015-16 Bulletin for AP Students and Parents*.

Section I: Multiple-Choice Questions

This is the multiple-choice section of the 2016 AP exam.
It includes cover material and other administrative instructions
to help familiarize students with the mechanics of the exam.
(Note that future exams may differ in look from the following content.)

For purposes of test security and/or statistical analysis, some questions
have been removed from the version of the exam that was administered
in 2016. Therefore, the timing indicated here may not be appropriate
for a practice exam.

AP® Statistics Exam

SECTION I: Multiple Choice

2016

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.

At a Glance

Total Time
1 hour, 30 minutes
Number of Questions
40
Percent of Total Score
50%
Writing Instrument
Pencil required
Electronic Device
Graphing calculator expected

Instructions

Section I of this exam contains 40 multiple-choice questions. Fill in only the circles for numbers 1 through 40 on your answer sheet.

Indicate all of your answers to the multiple-choice questions on the answer sheet. No credit will be given for anything written in this exam booklet, but you may use the booklet for notes or scratch work. After you have decided which of the suggested answers is best, completely fill in the corresponding circle on the answer sheet. Give only one answer to each question. If you change an answer, be sure that the previous mark is erased completely. Here is a sample question and answer.

Sample Question	Sample Answer
Chicago is a (A) state (B) city (C) country (D) continent (E) village	Ⓐ ● Ⓒ Ⓓ Ⓔ

Use your time effectively, working as quickly as you can without losing accuracy. Do not spend too much time on any one question. Go on to other questions and come back to the ones you have not answered if you have time. It is not expected that everyone will know the answers to all of the multiple-choice questions.

Your total score on the multiple-choice section is based only on the number of questions answered correctly. Points are not deducted for incorrect answers or unanswered questions.

Form I
Form Code 4MBP4-S

90

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1}\sum\left(x_i - \bar{x}\right)^2}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\sum\left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\sum\left(x_i - \bar{x}\right)^2}}$$

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

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

$$E(X) = \mu_X = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_X = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If $\overline{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\overline{x}} = \mu$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\bullet$ (standard deviation of statistic)

Single-Sample

Statistic	Standard Deviation of Statistic
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

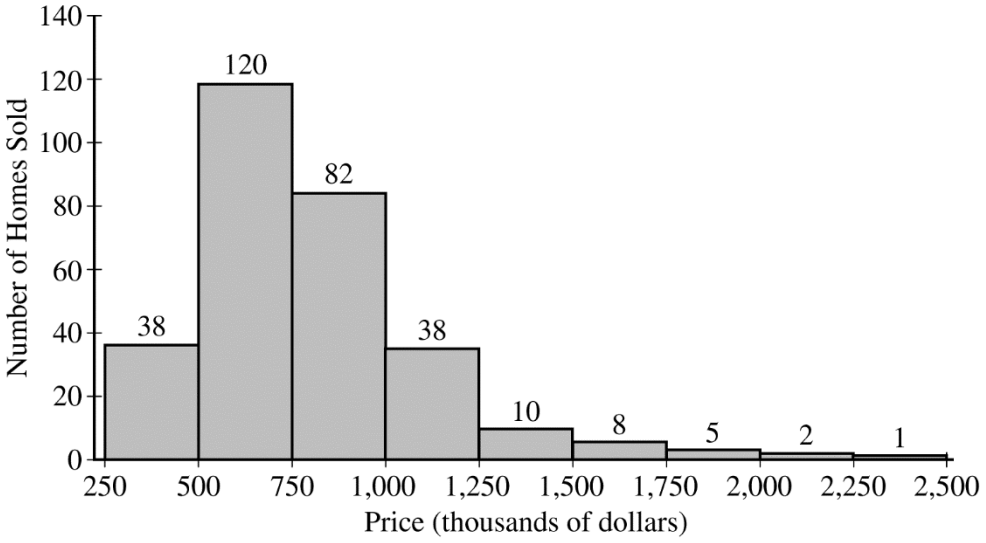
Statistic	Standard Deviation of Statistic
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

STATISTICS
SECTION I
Time—1 hour and 30 minutes
Number of questions—40
Percent of total score—50

Directions: Solve each of the following problems, using the available space for scratch work. Decide which is the best of the choices given and fill in the corresponding circle on the answer sheet. No credit will be given for anything written in the test book. Do not spend too much time on any one problem.

1. The prices, in thousands of dollars, of 304 homes recently sold in a city are summarized in the histogram below.



Based on the histogram, which of the following statements must be true?

(A) The minimum price is $250,000.

(B) The maximum price is $2,500,000.

(C) The median price is not greater than $750,000.

(D) The mean price is between $500,000 and $750,000.

(E) The upper quartile of the prices is greater than $1,500,000.

GO ON TO THE NEXT PAGE.

2. As part of a study on the relationship between the use of tanning booths and the occurrence of skin cancer, researchers reviewed the medical records of 1,436 people. The table below summarizes tanning booth use for people in the study who did and did not have skin cancer.

	Used a Tanning Booth	Did Not Use a Tanning Booth	Total
Skin cancer	190	706	896
No skin cancer	75	465	540
Total	265	1,171	1,436

Of the people in the study who had skin cancer, what fraction used a tanning booth?

(A) $\frac{190}{265}$

(B) $\frac{190}{896}$

(C) $\frac{190}{1,436}$

(D) $\frac{265}{1,436}$

(E) $\frac{896}{1,436}$

GO ON TO THE NEXT PAGE.

3. A researcher is conducting a study of charitable donations by surveying a simple random sample of households in a certain city. The researcher wants to determine whether there is convincing statistical evidence that more than 50 percent of households in the city gave a charitable donation in the past year. Let p represent the proportion of all households in the city that gave a charitable donation in the past year. Which of the following are appropriate hypotheses for the researcher?

(A) $H_0 : p = 0.5$ and $H_a : p > 0.5$

(B) $H_0 : p = 0.5$ and $H_a : p \neq 0.5$

(C) $H_0 : p = 0.5$ and $H_a : p < 0.5$

(D) $H_0 : p > 0.5$ and $H_a : p \neq 0.5$

(E) $H_0 : p > 0.5$ and $H_a : p = 0.5$

---

4. A company determines the mean and standard deviation of the number of sick days taken by its employees in one year. Which of the following is the best description of the standard deviation?

(A) Approximately the mean distance between the number of sick days taken by individual employees and the mean number of sick days taken by all employees

(B) Approximately the median distance between the number of sick days taken by individual employees and the median number of sick days taken by all employees

(C) The distance between the greatest number of sick days taken by an employee and the mean number of sick days taken by all employees

(D) The number of days separating the fewest sick days taken and the most sick days taken when considering all employees

(E) The number of days separating the fewest sick days taken and the most sick days taken when considering the middle 50 percent of the distribution

GO ON TO THE NEXT PAGE.

5. In one region of the country, the mean length of stay in hospitals is 5.5 days with standard deviation 2.6 days. Because many patients stay in the hospital for considerably more days, the distribution of length of stay is strongly skewed to the right. Consider random samples of size 100 taken from the distribution with the mean length of stay, $\bar{x}$, recorded for each sample. Which of the following is the best description of the sampling distribution of $\bar{x}$?

   (A) Strongly skewed to the right with mean 5.5 days and standard deviation 2.6 days

   (B) Strongly skewed to the right with mean 5.5 days and standard deviation 0.26 day

   (C) Strongly skewed to the right with mean 5.5 days and standard deviation 0.026 day

   (D) Approximately normal with mean 5.5 days and standard deviation 2.6 days

   (E) Approximately normal with mean 5.5 days and standard deviation 0.26 day

GO ON TO THE NEXT PAGE.

6. A local television news station includes a viewer survey question about a current issue at the beginning of every evening news broadcast. Viewers are invited to use social media to respond to the question. The results of the survey are shared with the audience at the end of each broadcast. In relation to the opinions of the population of the region, which of the following is a possible reason why the results of such surveys could be biased?

   I. Viewers with strong opinions about the current issue are more likely to respond than are viewers without strong opinions.
   II. The opinions of viewers of one television station are not necessarily representative of the population of a region.
   III. Viewers with access to social media are not necessarily representative of the population of a region.

   (A) I only

   (B) II only

   (C) III only

   (D) II and III only

   (E) I, II, and III

---

7. A graduate student conducted a study of field mice in rural Kansas. The student obtained a sample of 100 field mice and recorded the weight, in grams, of each mouse. After the measurements were taken, it was discovered that the scale was not calibrated correctly. The student adjusted the 100 recorded measurements by subtracting 3 grams from each measurement. Which of the following statistics for the weight, in grams, of the field mice has the same value before and after the adjustment?

   (A) The median

   (B) The mean

   (C) The first quartile

   (D) The third quartile

   (E) The interquartile range

GO ON TO THE NEXT PAGE.

8. A statistician proposed a new method for constructing a 90 percent confidence interval to estimate the median of assessed home values for homes in a large community. To test the method, the statistician will conduct a simulation by selecting 10,000 random samples of the same size from the population. For each sample, a confidence interval will be constructed using the new method. If the confidence level associated with the new method is actually 90 percent, which of the following will be captured by approximately 9,000 of the confidence intervals constructed from the simulation?

(A) The sample mean

(B) The sample median

(C) The sample standard deviation

(D) The population mean

(E) The population median

---

9. The distribution of monthly rent for one-bedroom apartments in a city is approximately normal with mean $936 and standard deviation $61. A graduate student is looking for a one-bedroom apartment and wants to pay no more than $800 in monthly rent. Of the following, which is the best estimate of the percent of one-bedroom apartments in the city with a monthly rent of at most $800 ?

(A) 1.3%

(B) 2.5%

(C) 50%

(D) 95%

(E) 97.5%

GO ON TO THE NEXT PAGE.

10. A news article reported that college students who have part-time jobs work an average of 15 hours per week. The staff of a college newspaper thought that the average might be different from 15 hours per week for their college. Data were collected on the number of hours worked per week for a random sample of students at the college who have part-time jobs. The data were used to test the hypotheses

$$H_0 : \mu = 15$$
$$H_a : \mu \neq 15,$$

where μ is the mean number of hours worked per week for all students at the college with part-time jobs. The results of the test are shown in the table below.

Sample Mean	Std Error	df	*t*-stat	*p*-value
13.755	0.707	25	−1.761	0.090

Assuming all conditions for inference were met, which of the following represents a 95 percent confidence interval for μ ?

(A) 13.755 ± 0.244

(B) 13.755 ± 0.286

(C) 13.755 ± 0.707

(D) 13.755 ± 1.245

(E) 13.755 ± 1.456

11. A team of psychologists studied the effect of multitasking on the completion of cognitive tasks. A group of 40 women participated in the study. Each woman owned a smartphone equipped with the same type of keyboard. The women typed a text passage on the phone twice, one time while sitting in a quiet room (a single task) and the other time while walking (a multitask). The order of the single task and the multitask was randomly determined for each woman. The psychologists recorded the time it took each woman to type the text for both tasks. If the conditions of inference are met, which of the following tests is most appropriate to analyze the data?

(A) A two-sample t-test for a difference between means

(B) A matched-pairs t-test for a mean difference

(C) A one-sample z-test for a proportion

(D) A two-sample z-test for a difference between proportions

(E) A chi-square test of independence

---

12. The random variable X is normally distributed with mean 5 and standard deviation 25. The random variable Y is defined by $Y = 2 + 4X$. What are the mean and the standard deviation of Y?

(A) The mean is 20 and the standard deviation is 102.

(B) The mean is 20 and the standard deviation is 50.

(C) The mean is 22 and the standard deviation is 102.

(D) The mean is 22 and the standard deviation is 100.

(E) The mean is 22 and the standard deviation is 50.

GO ON TO THE NEXT PAGE.

13. In northwest Pennsylvania, a zoologist recorded the ages, in months, of 55 bears and whether each bear was male or female. The data are shown in the back-to-back stemplot below.

Female		Male
4	15	
	14	
2 0	13	
	12	
	11	
	10	
4	9	
2	8	
1	7	0
1 0	6	6
8 3 2	5	3 5
6 5 4 2 1	4	6
8 8 5 4 3 2	3	2 4 7 8
9 7 6 2 1 0 0	2	0 0 1 2 2 3 4 4 4 5 6 7 7 9
	1	5 7 8

7|0 represents 70 months

Based on the stemplot, which of the following statements is true?

(A) The median age and the range of ages are both greater for female bears than for male bears.

(B) The median age and the range of ages are both less for female bears than for male bears.

(C) The median age is the same for female bears and male bears, and the range of ages is the same for female bears and male bears.

(D) The median age is less for female bears than for male bears, and the range of ages is greater for female bears than for male bears.

(E) The median age is greater for female bears than for male bears, and the range of ages is less for female bears than for male bears.

GO ON TO THE NEXT PAGE.

14. A produce supplier ships boxes of produce to individual customers. The distribution of weights of shipped boxes is approximately normal with mean 36 pounds and standard deviation 4 pounds. Which expression represents the weight, in pounds, at the 75th percentile of the distribution?

(A) $-1.96(4) + 36$

(B) $-0.25(4) + 36$

(C) $0.25(4) + 36$

(D) $0.67(4) + 36$

(E) $0.75(4) + 36$

15. A polling agency conducted a survey by selecting 100 random samples, each consisting of 1,200 United States citizens. The citizens in each sample were asked whether they were optimistic about the economy. For each sample, the polling agency created a 95 percent confidence interval for the proportion of all United States citizens who were optimistic about the economy. Which of the following statements is the best interpretation of the 95 percent confidence level?

(A) With 100 confidence intervals, we can be 95% confident that the sample proportion of citizens of the United States who are optimistic about the economy is correct.

(B) We would expect about 95 of the 100 confidence intervals to contain the proportion of all citizens of the United States who are optimistic about the economy.

(C) We would expect about 5 of the 100 confidence intervals to not contain the sample proportion of citizens of the United States who are optimistic about the economy.

(D) Of the 100 confidence intervals, 95 of the intervals will be identical because they were constructed from samples of the same size of 1,200.

(E) The probability is 0.95 that 100 confidence intervals will yield the same information about the sample proportion of citizens of the United States who are optimistic about the economy.

GO ON TO THE NEXT PAGE.

16. As part of a national sleep study, a random sample of adults was selected and surveyed about their physical activity and the number of hours they sleep each night. Of the 183 adults who exercised regularly (exercisers), 59 percent reported sleeping at least seven hours at night. Of the 88 adults who did not exercise regularly (nonexercisers), 52 percent reported sleeping at least seven hours at night. Which of the following is the most appropriate standard error for a confidence interval for the difference in proportions of adults who sleep at least seven hours at night among exercisers and nonexercisers?

(A) $\sqrt{\frac{(0.59)(0.41)}{183} + \frac{(0.52)(0.48)}{88}}$

(B) $\sqrt{\frac{(0.59)(0.41) + (0.52)(0.48)}{183 + 88}}$

(C) $\sqrt{(0.57)(0.43)\left(\frac{1}{183} + \frac{1}{88}\right)}$

(D) $\sqrt{(0.5)(0.5)\left(\frac{1}{183} + \frac{1}{88}\right)}$

(E) $\sqrt{(0.5)(0.5)\left(\frac{1+1}{183 + 88}\right)}$

17. A representative of a car manufacturer in the United States made the following claim in a news report.

> Ten years ago, only 53 percent of Americans owned American-made cars, but that figure is significantly higher today.

A research group conducted a study to investigate whether the claim was true. The group found that 56 percent of a randomly selected sample of car owners in the United States owned American-made cars. A test of the appropriate hypotheses resulted in a p-value of 0.283. Assuming the conditions for inference were met, is there sufficient evidence to conclude, at the significance level of $\alpha = 0.05$, that the proportion of all car owners in the United States who own American-made cars has increased from what it was ten years ago?

(A) Yes, because $0.56 > 0.53$.

(B) Yes, because a reasonable interval for the proportion is 0.56 ± 0.283.

(C) Yes, because $0.56 - 0.53 = 0.03$ and $0.03 < 0.05$.

(D) No, because $0.283 < 0.53$.

(E) No, because $0.283 > 0.05$.

GO ON TO THE NEXT PAGE.

18. Researchers wanted to investigate whether babies have a sense of right and wrong. They showed each of 60 babies a puppet show in which a red puppet was trying to open a heavy box lid. A second puppet, called the helper, would try to help the red puppet open the box, while a third puppet, called the hinderer, would try to slam the box lid down. After watching the show, each baby was presented with a tray containing the helper puppet and the hinderer puppet, and the researchers recorded which puppet the baby reached for. The researchers wanted to determine whether the babies would reach for the helper puppet more than for the hinderer puppet.

    As part of the show, a green puppet and a yellow puppet served as the helper and hinderer. For each baby, a coin was tossed to determine which color would serve which role. Which of the following is the most important reason for the random assignment of color to role in the study?

    (A) Slamming the lid might cause wear on the hinderer puppet after 60 shows. The random assignment of color to role permits more even wear between the helper and the hinderer.

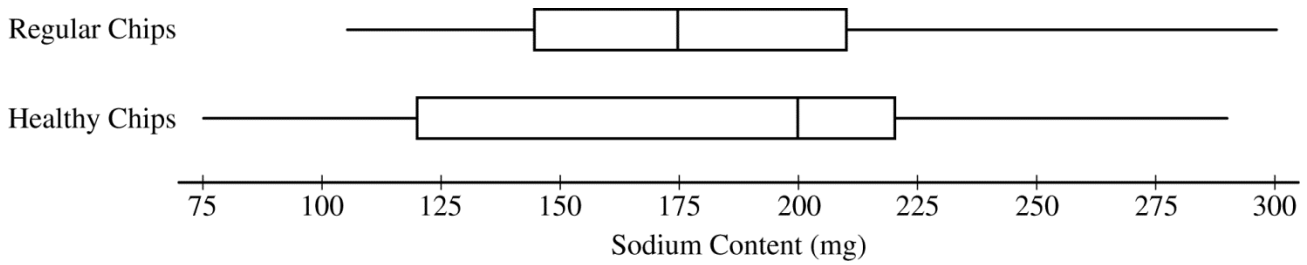
    (B) The puppeteer might grow tired of doing the same show with the same puppet colors. The random assignment of color to role keeps the show fresh for each performance.

    (C) If the same role is played by the same color puppet, the babies might show a preference for the color instead of a preference for the role.

    (D) The random assignment of color to role allows the researchers to determine if one color is better than another in teaching babies right from wrong.

    (E) Boys and girls might prefer different colors. The random assignment of color to role ensures that the show is equally accessible to boys and girls.

19. Nutritionists examined the sodium content of different brands of potato chips. Each brand was classified as either healthy or regular based on how the chips were marketed to the public. The sodium contents, in milligrams (mg) per serving, of the chips are summarized in the boxplots below.



Based on the boxplots, which statement gives a correct comparison between the two classifications of the sodium content of the chips?

(A) The number of brands classified as healthy is greater than the number of brands classified as regular.

(B) The interquartile range (IQR) of the brands classified as healthy is greater than the IQR of the brands classified as regular.

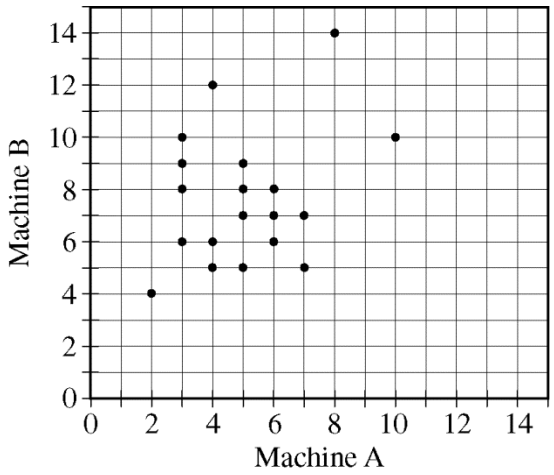
(C) The range of the brands classified as healthy is less than the range of the brands classified as regular.

(D) The median of the brands classified as healthy is more than twice the median of the brands classified as regular.

(E) The brand with the least sodium content and the brand with the greatest sodium content are both classified as healthy.

GO ON TO THE NEXT PAGE.

20. A factory has two machines, A and B, making the same part for refrigerators. The number of defective parts produced by each machine during the first hour of operation was recorded on 19 randomly selected days. The scatterplot below shows the number of defective parts produced by each machine on the selected days.



Which statement gives the best comparison between the number of defective parts produced by the machines during the first hour of operation on the 19 days?

(A) Machine A always produced the same number of defective parts as machine B.

(B) Machine A always produced fewer defective parts than machine B.

(C) Machine A always produced more defective parts than machine B.

(D) Machine A usually, but not always, produced fewer defective parts than machine B.

(E) Machine A usually, but not always, produced more defective parts than machine B.

Item 21 was not scored.

GO ON TO THE NEXT PAGE.

22. The faces of a four-sided fair die are numbered 1 through 4, respectively. For a certain game, the die is tossed and the number that lands facedown is recorded. The table below summarizes the points a player earns for the number that lands facedown.

Number landing facedown	1	2	3	4
Points	0	1	0	2

Consider two independent tosses of the die. Let the random variable S represent the sum of the points earned from the two tosses. Which table represents the probability distribution of S ?

(A)

S	0	1	2
Probability	0.5	0.25	0.25

(B)

S	0	2	4
Probability	0.5	0.25	0.25

(C)

S	0	2	4
Probability	0.25	0.5	0.25

(D)

S	0	1	2	3	4
Probability	0.25	0.25	0.3125	0.125	0.0625

(E)

S	0	1	2	3	4
Probability	0.0625	0.25	0.375	0.25	0.0625

GO ON TO THE NEXT PAGE.

23. A botanist collected one leaf at random from each of 10 randomly selected mature maple trees of the same species. The mean and the standard deviation of the surface areas for the 10 leaves in the sample were computed. Assume the distribution of surface areas of maple leaves is normal. What is the appropriate method for constructing a one-sample confidence interval to estimate the population mean surface area of the species of maple leaves, and why is the method appropriate?

(A) The t-interval is appropriate, because the population standard deviation is not known.

(B) The t-interval is appropriate, because the t-interval is narrower than the z-interval.

(C) The z-interval is appropriate, because the z-interval is narrower than the t-interval.

(D) The z-interval is appropriate, because the central limit theorem applies.

(E) The z-interval is appropriate, because the sample standard deviation is known.

24. A state educational agency was concerned that the salaries of public school teachers in one region of the state, region A, were higher than the salaries in another region of the state, region B. The agency took two independent random samples of salaries of public school teachers, one from region A and one from region B. The data are summarized in the table below.

	Region A	Region B
Mean salary	\$62,583	\$60,117
Standard deviation	\$6,274	\$9,319
Number of salaries	117	78

Assuming all conditions for inference are met, do the data provide convincing statistical evidence that the salaries of public school teachers in region A are, on average, greater than the salaries of public school teachers in region B?

(A) Yes, there is evidence at the significance level of $\alpha = 0.001$.

(B) Yes, there is evidence at the significance level of $\alpha = 0.01$ but not at $\alpha = 0.001$.

(C) Yes, there is evidence at the significance level of $\alpha = 0.05$ but not at $\alpha = 0.01$.

(D) Yes, there is evidence at the significance level of $\alpha = 0.10$ but not at $\alpha = 0.05$.

(E) No, there is no evidence at the significance level of $\alpha = 0.10$.

GO ON TO THE NEXT PAGE.

25. A florist wanted to investigate whether a new powder added to the water of cut flowers helps to keep the flowers fresh longer than just water alone. For a shipment of roses that was delivered to the store, the florist flipped a coin before placing each rose in its own individual container with water. If the coin landed heads up, the rose was placed in water with the new powder; otherwise, the rose was placed in water alone. Which of the following is the best description of the method used by the florist?

(A) A census, because all roses are assigned to a container
(B) An experiment with a completely randomized design
(C) An experiment with a blocked design, with blocking by type of water
(D) An experiment with a matched-pairs design
(E) An observational study

GO ON TO THE NEXT PAGE.

26. A commercial for a breakfast cereal is shown during a certain television program. The manufacturer of the cereal wants to estimate the percent of television viewers who watch the program. The manufacturer wants the estimate to have a margin of error of at most 0.02 at a level of 95 percent confidence. Of the following, which is the smallest sample size that will satisfy the manufacturer's requirements?

(A) 40

(B) 50

(C) 100

(D) 1,700

(E) 2,500

---

27. A contestant's final winnings on a game show are determined by a random selection of a base amount and a possible multiplier. For the base amount, the contestant randomly selects one of four cards, where two cards are marked $1,000, one card is marked $2,000, and one card is marked $5,000. After choosing the card, the contestant randomly selects one of five chips, where three chips are red and two chips are white. If the selected chip is red, the contestant's final winnings are twice the value of the base amount. If the selected chip is white, the contestant's final winnings are the value of the base amount. What is the probability that a contestant's final winnings are exactly $2,000 ?

(A) 0.100
(B) 0.200
(C) 0.250
(D) 0.325
(E) 0.400

GO ON TO THE NEXT PAGE.

28. Meteorologists are interested in the relationship between minimum pressure and maximum wind speed of hurricanes. The minimum pressure, in millibars, and maximum wind speed, in knots, were collected for a random sample of 100 hurricanes from the year 1995 to the year 2012. A regression analysis of maximum wind speed on minimum pressure produced a 95 percent confidence interval of $(-1.42, -1.20)$ for the slope of the least-squares regression line. Which statement is a correct interpretation of the interval?

(A) The probability is 0.95 that wind speed will decrease, on average, between 1.20 knots and 1.42 knots for each millibar increase in minimum pressure.

(B) The probability is 0.95 that a different sample of 100 hurricanes will result in an increase, on average, of wind speed between 1.20 knots and 1.42 knots for each millibar increase in minimum pressure.

(C) We can be 95% confident that wind speed decreases, on average, between 1.20 knots and 1.42 knots for each millibar increase in minimum pressure.

(D) We can be 95% confident that wind speed increases, on average, between 1.20 knots and 1.42 knots for each millibar increase in minimum pressure.

(E) We can be 95% confident that, for any sample of hurricanes, the wind speed will decrease, on average, between 1.20 knots and 1.42 knots for each millibar increase in minimum pressure.

GO ON TO THE NEXT PAGE.

29. Some contact lens wearers report problems with dryness in their eyes. A study was conducted to evaluate the effectiveness of a new eye-drop solution to relieve dryness for contact lens wearers. Twenty-five volunteers who wore contact lenses agreed to use the new solution for one month. At the end of the month, 36 percent of the volunteers reported that the new solution was effective in relieving dryness. The company that produced the new eye-drop solution concluded that using the new solution is more effective in relieving dryness than using no solution. Which of the following best explains why the study does not support such a conclusion?

(A) The sample size was too small.

(B) The study had no control group.

(C) The participants were volunteers.

(D) The participants self-reported the frequency with which they used the new solution.

(E) The participants self-reported the effectiveness of the new solution.

GO ON TO THE NEXT PAGE.

30. The management team of a company with 10,000 employees is considering installing charging stations for electric cars in the company parking lots. In a random sample of 500 employees, 15 reported owning an electric car. Which of the following is a 99 percent confidence interval for the proportion of all employees at the company who own an electric car?

(A) $0.03 \pm 2.326\sqrt{\frac{(0.03)(0.97)}{500}}$

(B) $0.15 \pm 2.326\sqrt{\frac{(0.15)(0.85)}{500}}$

(C) $0.03 \pm 2.576\sqrt{\frac{(0.03)(0.97)}{500}}$

(D) $0.15 \pm 2.576\sqrt{\frac{(0.15)(0.85)}{500}}$

(E) $0.03 \pm 2.576\sqrt{\frac{(0.03)(0.97)}{500} + \frac{(0.03)(0.97)}{10,000}}$

GO ON TO THE NEXT PAGE.

31. A test of the hypotheses $H_0 : \mu = 0$ versus $H_a : \mu > 0$ was conducted using a sample of size 7. The test statistic was $t = 1.935$. Which of the following is closest to the *p*-value of the test?

(A) 0.0125

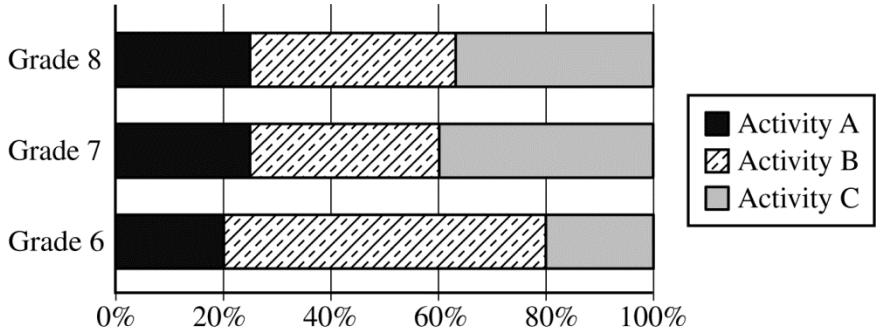
(B) 0.0265

(C) 0.0471

(D) 0.0506

(E) 0.1012

32. As part of a community service program, students in three middle school grades (grade 6, grade 7, grade 8) each chose to participate in one of three school-sponsored volunteer activities. The graph below shows the distribution for each class for the three activities.



Based on the graph, which statement must be true?

(A) Of all the students who chose activity B, the greatest number of students were in grade 6.

(B) Grade 7 and grade 8 had the same number of students who did not choose activity A.

(C) The grade with the greatest percentage of students who chose activity C was grade 8.

(D) For students in grade 7, the number who chose activity C was greater than the number who chose activity B.

(E) For students in grade 8, the number who chose activity A was greater than the number who chose activity B.

GO ON TO THE NEXT PAGE.

33. At a large airport, data were recorded for one month on how many baggage items were unloaded from each flight upon arrival as well as the time required to deliver all the baggage items on the flight to the baggage claim area. A scatterplot of the two variables indicated a strong, positive linear association between the variables. Which of the following statements is a correct interpretation of the word "strong" in the description of the association?

(A) A least-squares model predicts that the more baggage items that are unloaded from a flight, the greater the time required to deliver the items to the baggage claim area.

(B) The actual time required to deliver all the items to the baggage claim area based on the number of items unloaded will be very close to the time predicted by a least-squares model.

(C) The time required to deliver an item to the baggage claim area is relatively constant, regardless of the number of baggage items unloaded from a flight.

(D) The variability in the time required to deliver all items to the baggage claim area is about the same for all flights, regardless of the number of items unloaded from a flight.

(E) The time required to unload baggage items from a flight is related to the time required to deliver the items to the baggage claim area.

GO ON TO THE NEXT PAGE.

34. A group of men and women were surveyed to investigate the association between gender and the number of friends the person has on a social media Web site. Results are shown in the table below.

	Number of Friends				
	0 to 50	51 to 100	101 to 150	151 to 200	201 or more
Men	17	82	73	56	42
Women	45	73	98	87	75

Which of the following procedures is the most appropriate for investigating whether an association exists between gender and the number of friends a person has on a social media Web site?

(A) A matched-pairs *t*-test for a mean difference

(B) A two-sample *t*-test for the difference between means

(C) A *t*-test for the slope of the regression line

(D) A chi-square goodness-of-fit test

(E) A chi-square test of independence

GO ON TO THE NEXT PAGE.

35. Carly commutes to work, and her commute time is dependent on the weather. When the weather is good, the distribution of her commute times is approximately normal with mean 20 minutes and standard deviation 2 minutes. When the weather is not good, the distribution of her commute times is approximately normal with mean 30 minutes and standard deviation 4 minutes. Suppose the probability that the weather will be good tomorrow is 0.9. Which of the following is closest to the probability that Carly's commute time tomorrow will be greater than 25 minutes?

(A) 0.0056

(B) 0.0894

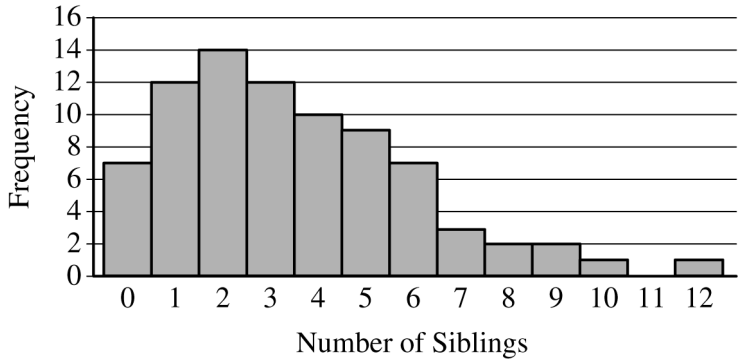
(C) 0.0950

(D) 0.8055

(E) 0.9006

36. The number of siblings was recorded for each student of a group of 80 students. Some summary statistics and a histogram displaying the results are shown below.

Mean	Standard Deviation	Q1	Q3
3.5	2.535	2	5



An outlier is often defined as a number that is more than 1.5 times the interquartile range below the first quartile or above the third quartile. Using the definition of an outlier and the given information, which of the following can be concluded?

(A) The median is greater than the mean, and the distribution has no outliers.

(B) The median is greater than the mean, and the distribution has only one outlier.

(C) The median is greater than the mean, and the distribution has two outliers.

(D) The median is less than the mean, and the distribution has only one outlier.

(E) The median is less than the mean, and the distribution has two outliers.

GO ON TO THE NEXT PAGE.

37. In the states of Florida and Colorado, veterinarians investigating obesity in dogs obtained random samples of pet medical records and recorded the weights of the dogs in the samples. A test was conducted of $H_0 : p_1 = p_2$ versus $H_a : p_1 \neq p_2$, where p_1 represents the proportion of all overweight dogs in Florida and p_2 represents the proportion of all overweight dogs in Colorado. The resulting test statistic for a two-sample z-test for a difference between proportions was 1.85. At the significance level $\alpha = 0.05$, which of the following is a correct conclusion?

(A) There is not sufficient statistical evidence to conclude that the proportion of all overweight dogs in Florida is different from the proportion of all overweight dogs in Colorado because the p-value is greater than 0.05.

(B) There is not sufficient statistical evidence to conclude that the proportion of all overweight dogs in Florida is different from the proportion of all overweight dogs in Colorado because the z-test statistic is greater than 0.05.

(C) There is sufficient statistical evidence to conclude that the proportion of all overweight dogs in Florida is different from the proportion of all overweight dogs in Colorado because the p-value is greater than 0.05.

(D) There is sufficient statistical evidence to conclude that the proportion of all overweight dogs in Florida is different from the proportion of all overweight dogs in Colorado because the p-value is less than 0.05.

(E) There is sufficient statistical evidence to conclude that the proportion of all overweight dogs in Florida is greater than the proportion of all overweight dogs in Colorado because the z-test statistic is positive.

GO ON TO THE NEXT PAGE.

38. A newspaper editor wants to investigate whether residents of the city support a proposal to build a new high school football stadium. The editor hires a polling firm to conduct a survey and requests that a sample of 500 residents be selected using a stratified sampling design based on voting districts within the city. Which of the following methods will achieve the desired sampling design?

(A) Send a survey to all city residents and use the first 500 returned surveys for the sample.

(B) Select a random sample from each voting district based on the proportion of city residents in the district so that a total of 500 is obtained.

(C) Select one voting district at random, and then select a random sample of 500 from the selected voting district.

(D) Alphabetize a list of all city residents, and then select the first 500 residents on the list, classifying those selected by voting district.

(E) Select the first 500 city residents who attend the next high school football game.

GO ON TO THE NEXT PAGE.

39. A simulation was conducted using 10 fair six-sided dice, where the faces were numbered 1 through 6, respectively. All 10 dice were rolled, and the average of the 10 numbers appearing faceup was recorded. The process was repeated 20 times. Which of the following best describes the distribution being simulated?

(A) A sampling distribution of a sample mean with $n = 10$, $\mu_{\bar{x}} = 3.5$, and $\sigma_{\bar{x}} \approx 0.54$

(B) A sampling distribution of a sample mean with $n = 10$, $\mu_{\bar{x}} = 3.5$, and $\sigma_{\bar{x}} \approx 1.71$

(C) A sampling distribution of a sample mean with $n = 20$, $\mu_{\bar{x}} = 3.5$, and $\sigma_{\bar{x}} \approx 0.38$

(D) A sampling distribution of a sample proportion with $n = 10$, $\mu_{\hat{p}} = \frac{1}{6}$, and $\sigma_{\hat{p}} \approx 0.118$

(E) A sampling distribution of a sample proportion with $n = 20$, $\mu_{\hat{p}} = \frac{1}{6}$, and $\sigma_{\hat{p}} \approx 0.083$

GO ON TO THE NEXT PAGE.

40. The SC Electric Company has bid on two electrical wiring jobs. The owner of the company believes that

- the probability of being awarded the first job (event A) is 0.75;
- the probability of being awarded the second job (event B) is 0.5; and
- the probability of being awarded both jobs (event (A and B)) is 0.375.

If the owner's beliefs are correct, which of the following statements must be true concerning event A and event B ?

(A) Event A and event B are mutually exclusive and are independent.

(B) Event A and event B are mutually exclusive and are not independent.

(C) Event A and event B are not mutually exclusive and are independent.

(D) Event A and event B are not mutually exclusive and are not independent.

(E) Event A and event B are not mutually exclusive, and independence cannot be determined with the information given.

END OF SECTION I

IF YOU FINISH BEFORE TIME IS CALLED, YOU MAY CHECK YOUR WORK ON THIS SECTION.

DO NOT GO ON TO SECTION II UNTIL YOU ARE TOLD TO DO SO.

MAKE SURE YOU HAVE DONE THE FOLLOWING.

- **PLACED YOUR AP NUMBER LABEL ON YOUR ANSWER SHEET**
- **WRITTEN AND GRIDDED YOUR AP NUMBER CORRECTLY ON YOUR ANSWER SHEET**
- **TAKEN THE AP EXAM LABEL FROM THE FRONT OF THIS BOOKLET AND PLACED IT ON YOUR ANSWER SHEET**

Table entry for z is the probability lying below z.

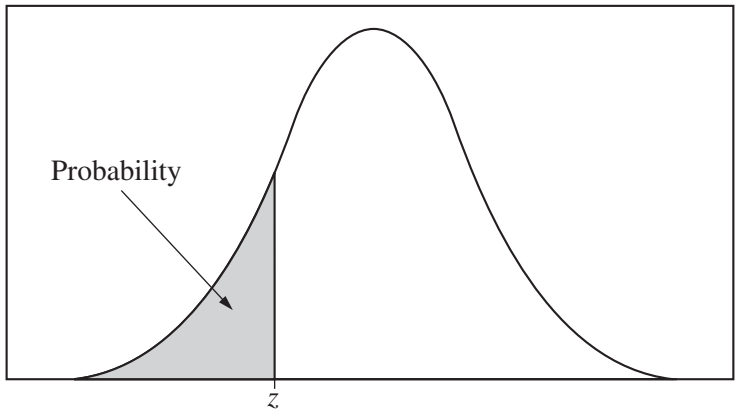


Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Table entry for z is the probability lying below z.

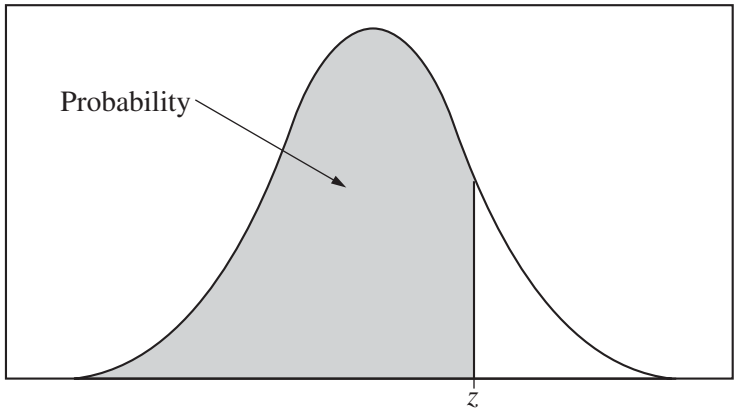


Table A (*Continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

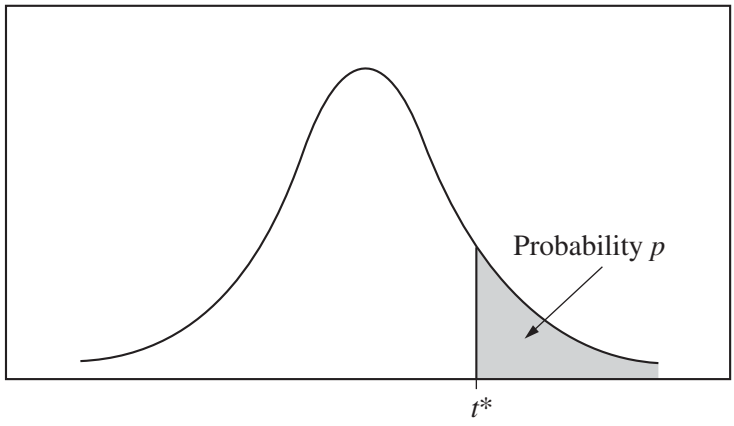


Table B ***t* distribution critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Table entry for p is the point (χ^2) with probability p lying above it.

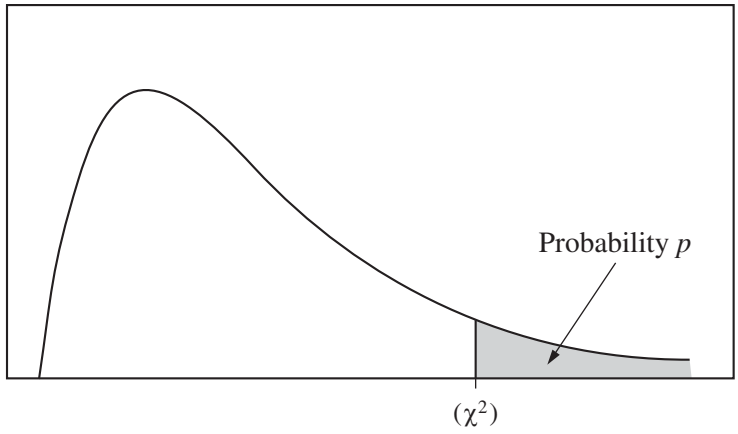


Table C χ^2 critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

Section II: Free-Response Questions

This is the free-response section of the 2016 AP exam.
It includes cover material and other administrative instructions
to help familiarize students with the mechanics of the exam.
(Note that future exams may differ in look from the following content.)

AP® Statistics Exam

SECTION II: Free Response

2016

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO.

At a Glance

Total Time
1 hour, 30 minutes
Number of Questions
6
Percent of Total Score
50%
Writing Instrument
Either pencil or pen with black or dark blue ink
Electronic Device
Graphing calculator expected

Part A

Number of Questions
5
Suggested Time
1 hour, 5 minutes
Percent of Section II Score
75%

Part B

Number of Questions
1
Suggested Time
25 minutes
Percent of Section II Score
25%

IMPORTANT Identification Information

PLEASE PRINT WITH PEN:

1. First two letters of your last name ☐☐
   First letter of your first name ☐
2. Date of birth
   ☐☐ Month ☐☐ Day ☐☐☐☐ Year
3. Six-digit school code
   ☐☐☐☐☐☐
4. Unless I check the box below, I grant the College Board the unlimited right to use, reproduce, and publish my free-response materials, both written and oral, for educational research and instructional purposes. My name and the name of my school will not be used in any way in connection with my free-response materials. I understand that I am free to mark "No" with no effect on my score or its reporting.
   No, I do not grant the College Board these rights. ☐

Instructions

The questions for both Part A and Part B are printed in this booklet. You may use any blank space in the booklet to organize your answers and for scratch work, but you must write your answers in the spaces provided for each answer. Pages containing statistical tables and useful formulas are printed in this booklet.

You may wish to look over the questions before starting to work on them. It is not expected that everyone will be able to complete all parts of all questions. Show all your work. Indicate clearly the methods you use because you will be scored on the correctness of your methods as well as the accuracy and completeness of your results and explanations. Correct answers without supporting work may not receive credit. Write your solution to each part of each question in the space provided for that part. Write clearly and legibly. Cross out any errors you make; erased or crossed-out work will not be scored.

Manage your time carefully. The proctor will announce the suggested time for Part A and Part B, but you may proceed freely from one question to the next. You may review your responses if you finish before the end of the exam is announced.

Form I
Form Code 4MBP4-S

90

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\sum x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1}\sum\left(x_i - \bar{x}\right)^2}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum\left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$r = \frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\sum\left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\sum\left(x_i - \bar{x}\right)^2}}$$

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

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

$$E(X) = \mu_X = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_X = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If $\overline{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\overline{x}} = \mu$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\bullet$ (standard deviation of statistic)

Single-Sample

Statistic	Standard Deviation of Statistic
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

Statistic	Standard Deviation of Statistic
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

STATISTICS

SECTION II

Part A

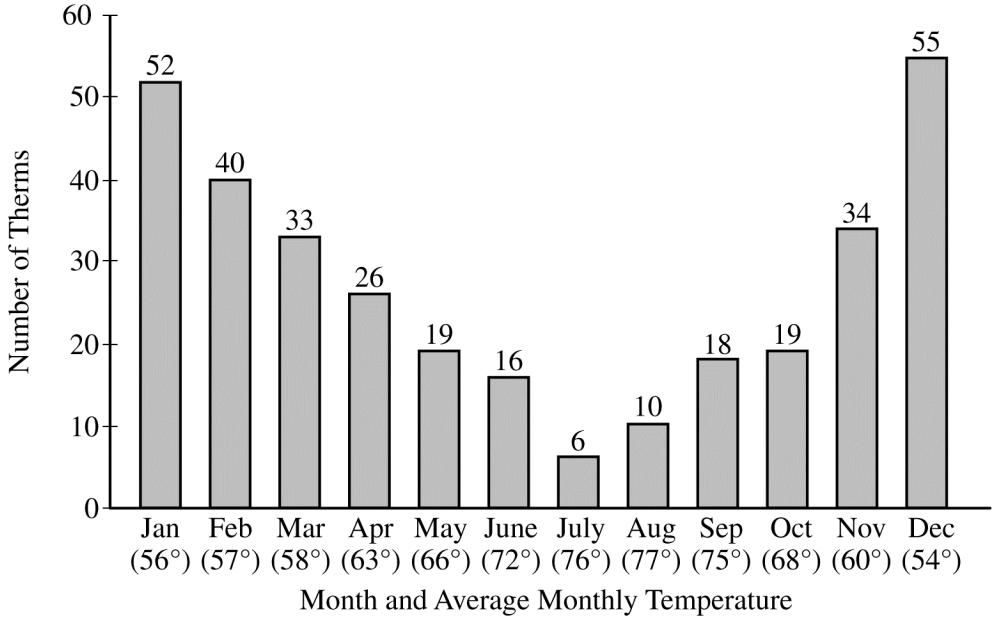
Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II score—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

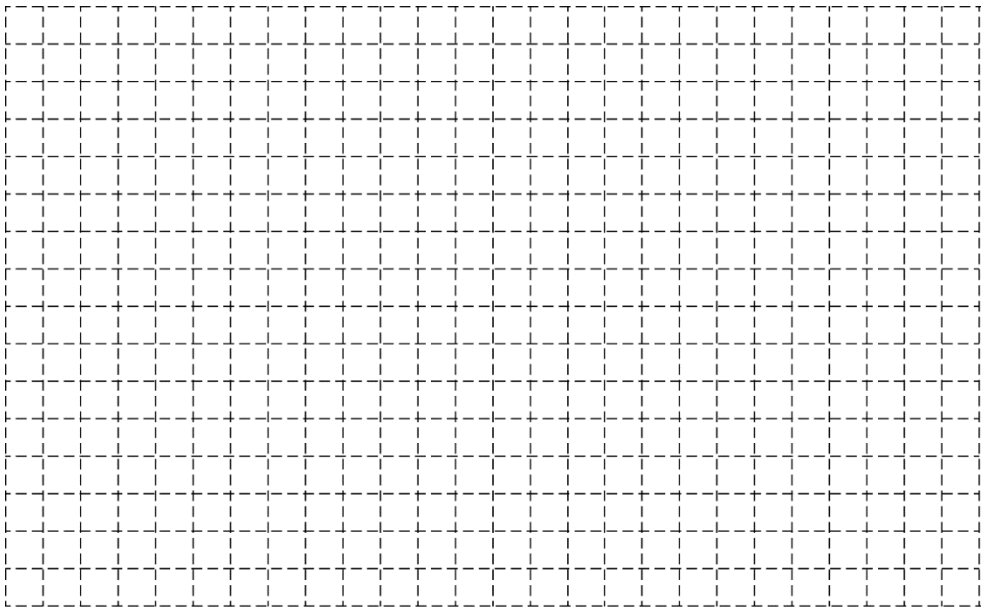
1. Natural gas is used in some households to heat the home, to heat the water, and to cook. A utility company sent the following bar chart to a household to show the amount of natural gas, measured in therms (a unit of heat energy), that the household used last year. The chart shows the number of therms and the average monthly temperature, in degrees Fahrenheit, for each month of the year.



(a) Describe how the number of therms used each month changed over the year.

GO ON TO THE NEXT PAGE.

(b) Construct an appropriate graph that shows the relationship between the number of therms used and the average monthly temperature.



(c) Describe what your graph in part (b) reveals about the relationship between the number of therms used and the average monthly temperature that is not revealed on the bar chart sent by the utility company.

GO ON TO THE NEXT PAGE.

2. Swedish researchers investigated the relationship between chocolate consumption and stroke. The researchers gave a questionnaire about eating habits to a randomly selected sample of Swedish men. Based on the responses to the questionnaire, the men were classified into two groups. Group A consisted of the 9,250 men who ate the most chocolate per week, and group B consisted of the 9,250 men who ate the least chocolate per week. The researchers tracked the men's health for ten years. During that time, there were 458 cases of stroke among the men in group A and 543 cases of stroke among the men in group B.

   (a) Do the data provide convincing statistical evidence that Swedish men who would be classified into group A have a lower probability of stroke than Swedish men who would be classified into group B?

If you need more room for your work in part (a), use the space below.

(b) A report in a newspaper concluded that Swedish men can reduce their probability of stroke by eating more chocolate. Based on the description of the investigation, was the conclusion appropriate? Justify your answer.

3. A large retail company has 500 stores in the United States and 300 stores in Europe. The average number of employees per store is 200, for a total of 100,000 employees in the United States and 60,000 employees in Europe. The company is considering offering employees one of two new benefits—one additional day of paid vacation per year or a small increase in pay. A survey will be given to a sample of employees to investigate which benefit is preferred and whether there is a difference in preference between employees in the United States and employees in Europe.

   Two sampling methods have been proposed.

   Sampling method 1: The company will randomly select 8 stores from its 800 stores. All employees at the 8 selected stores will be asked which benefit they prefer.

   Sampling method 2: The company will randomly select 1,000 employees from a list of all employees at the United States stores and 600 employees from a list of all employees at the European stores. All 1,600 selected employees will be asked which benefit they prefer.

   (a) One of the two methods results in a stratified sample of employees and the other results in a cluster sample of employees.

      (i) Identify the sampling method that results in a stratified sample of employees, and identify the strata.

      Sampling method number: Strata:

      (ii) Identify the sampling method that results in a cluster sample of employees, and identify the clusters.

      Sampling method number: Clusters:

   (b) Give one statistical advantage and one statistical disadvantage of using sampling method 1.

      Advantage:

      Disadvantage:

GO ON TO THE NEXT PAGE.

(c) Explain a statistical advantage of using sampling method 2 rather than using a simple random sample.

GO ON TO THE NEXT PAGE.

4. In ecology, a measure of diversity of a species in a certain population is called the diversity index. The diversity index D is defined as the probability that any two organisms selected at random from the population will be from different species.

   (a) If every organism in a population is from the same species, what is the value of the diversity index D ? Explain.

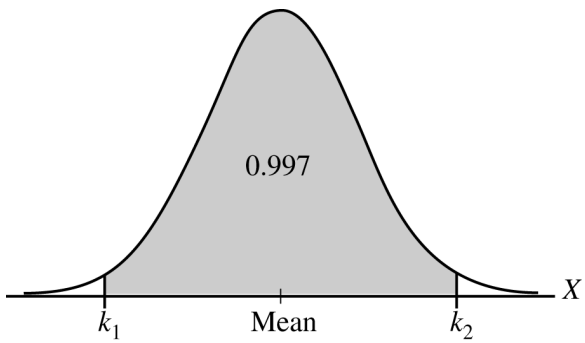
   (b) Consider a large population of turtles in which 30 percent of the population are snapping turtles and 70 percent of the population are box turtles. What is the value of the diversity index D for the population?

   (c) Suppose a large population of turtles consists only of two species, snapping turtle and box turtle, but the percentages of the species in the population are not known. What is the maximum possible value of the diversity index D for the population? Justify your answer.

5. A city council is considering funding a proposal to create a new city park. The council members will fund the proposal if they conclude that more than 60 percent of the city residents support the proposal. A survey of 2,000 randomly selected city residents will be conducted to investigate the level of support for the proposal. Let X represent the number of city residents in the sample who support the proposal. Assume that X is a binomial random variable.

   (a) Determine the mean and the standard deviation of the random variable X, assuming that 60 percent of city residents support the proposal.

The figure below shows a normal distribution that can be used to approximate the binomial probability distribution of X. The shaded region is bounded by k_1 and k_2 on the figure and represents the middle 0.997 of the area under the curve.



   (b) Assume that 60 percent of the city residents support the proposal. Use a normal approximation and the mean and standard deviation from part (a) to determine the values of k_1 and k_2.

   $k_1 =$

   $k_2 =$

GO ON TO THE NEXT PAGE.

(c) The survey was conducted, and 1,293 of the 2,000 city residents surveyed supported the proposal. Do your answers in part (b) and the survey results support the funding of the proposal? Justify your answer.

GO ON TO THE NEXT PAGE.

STATISTICS

SECTION II

Part B

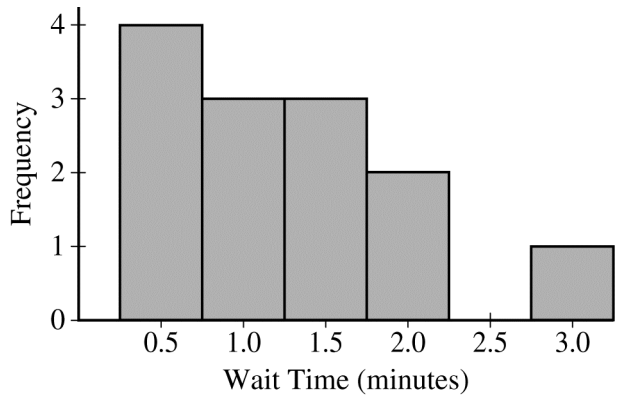
Question 6

Spend about 25 minutes on this part of the exam.

Percent of Section II score—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Phone callers to a bank's customer service center must wait until a service representative is available to answer the phone call. The bank manager is interested in estimating the mean customer wait time. Thirteen calls were selected at random. A histogram of the 13 wait times, in minutes, is shown below.



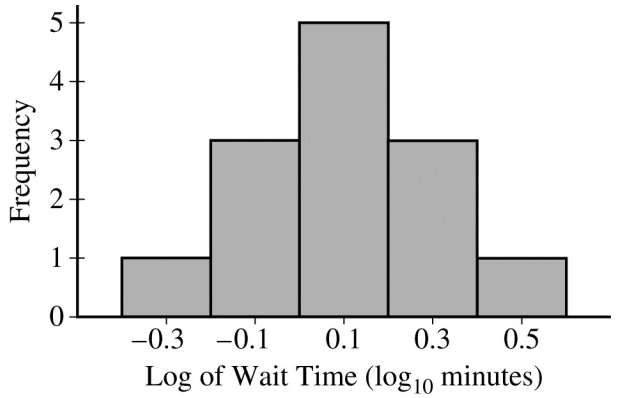
(a) Based on the histogram, explain why it might not be appropriate to use a one-sample *t*-interval to estimate the mean wait time for all customers.

GO ON TO THE NEXT PAGE.

A logarithmic transformation is often used to transform data such as wait times. Let x represent a customer's wait time. The log transformation of the customer's wait time is given by $\log_{10} x$. The table below shows the original 13 wait times, the respective log-transformed wait times, and the corresponding means, medians, and standard deviations.

	Wait Time, x	$\log_{10} x$
	0.40	−0.3979
	0.66	−0.1805
	0.71	−0.1487
	0.71	−0.1487
	1.10	0.0414
	1.16	0.0645
	1.20	0.0792
	1.29	0.1106
	1.29	0.1106
	1.70	0.2304
	1.90	0.2788
	2.15	0.3324
	2.82	0.4502
Mean	1.31	0.0632
Median	1.20	0.0792
Standard deviation	0.679	0.235

A histogram of the 13 log-transformed data values is shown below.



GO ON TO THE NEXT PAGE.

The summary statistics for x and $\log_{10} x$ are repeated below.

	Wait Time, x	$\log_{10} x$
Mean	1.31	0.0632
Median	1.20	0.0792
Standard deviation	0.679	0.235

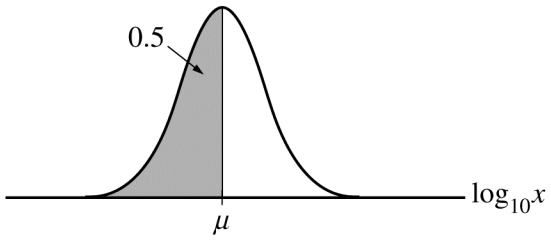
(b) Based on the histogram, the conditions for inference have been met for the log-transformed data. Construct and interpret a 95 percent confidence interval for the population mean μ of the log of the wait times.

(c) The mean of the log-transformed data is 0.0632 $\log_{10}$ minutes, which can be converted back to 1.157 minutes by calculating $10^{0.0632}$. Convert the endpoints of your interval in part (b) back to minutes and write the resulting interval.

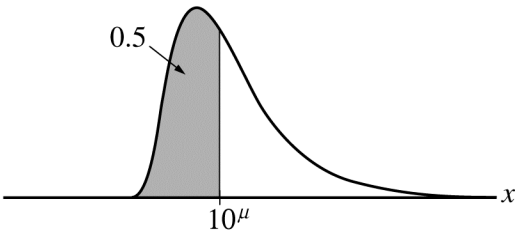
GO ON TO THE NEXT PAGE.

Graph 1 below shows a population distribution of the log of wait times, in $\log_{10}$ minutes, which is normal with mean μ. Graph 2 shows the result of converting the population distribution in Graph 1 back to the population distribution of wait times, in minutes. The lower 50 percent of the distribution is shaded in each graph.

Graph 1
Population Distribution of Log of Wait Time



Graph 2
Population Distribution of Wait Time



(d) Consider the parameter 10^{μ} in Graph 2.

(i) How does the parameter 10^{μ} compare with the median of the population distribution of wait times?

(ii) How does the parameter 10^{μ} compare with the mean of the population distribution of wait times?

(e) Write an interpretation of the interval you constructed in part (c).

GO ON TO THE NEXT PAGE.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX ON THE COVER.**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

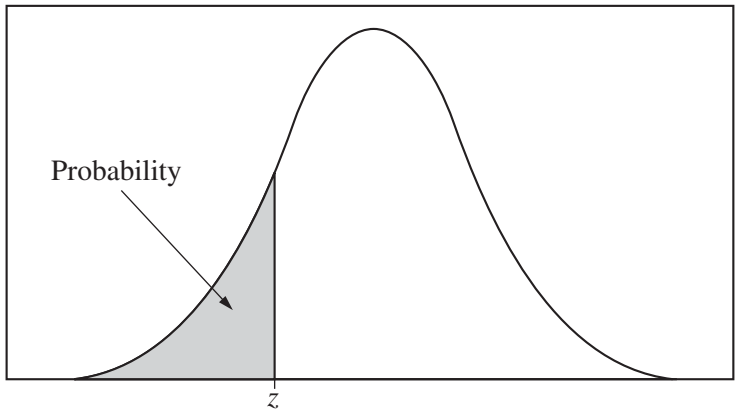


Table entry for z is the probability lying below z.

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

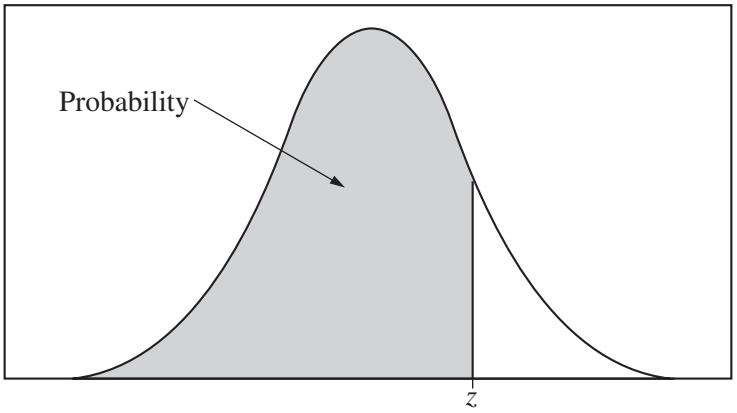


Table entry for z is the probability lying below z.

Table A (*Continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

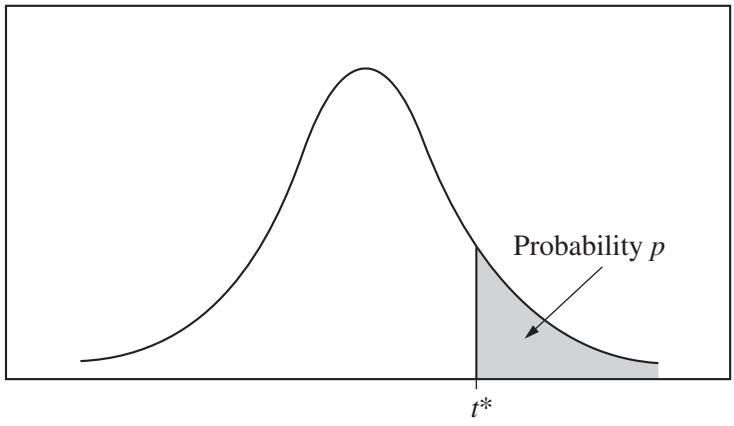


Table B ***t* distribution critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Table entry for p is the point (χ^2) with probability p lying above it.

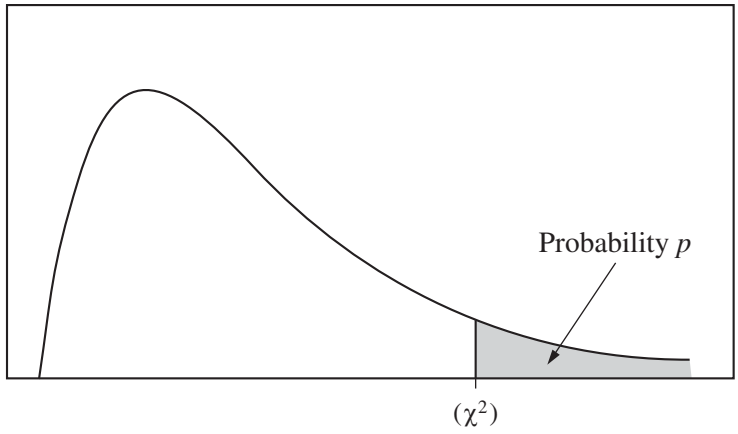


Table C χ^2 critical values

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

Multiple-Choice Answer Key

The following contains the answers to
the multiple-choice questions in this exam.

Answer Key for AP Statistics Practice Exam, Section I

Question 1: C

Question 2: B

Question 3: A

Question 4: A

Question 5: E

Question 6: E

Question 7: E

Question 8: E

Question 9: A

Question 10: E

Question 11: B

Question 12: D

Question 13: A

Question 14: D

Question 15: B

Question 16: A

Question 17: E

Question 18: C

Question 19: B

Question 20: D

Question 21: *

Question 22: D

Question 23: A

Question 24: C

Question 25: B

Question 26: E

Question 27: E

Question 28: C

Question 29: B

Question 30: C

Question 31: D

Question 32: D

Question 33: B

Question 34: E

Question 35: C

Question 36: E

Question 37: A

Question 38: B

Question 39: A

Question 40: C

*Item 21 was not used in scoring.

Free-Response Scoring Guidelines

The following contains the scoring guidelines
for the free-response questions in this exam.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 1

Intent of Question

The primary goals of this question are to assess a student's ability to (1) interpret a graph showing trends across time; (2) create an appropriate plot to show the relationship between two quantitative variables; and (3) determine what information can be obtained from a scatterplot that is not obtainable from a bar graph.

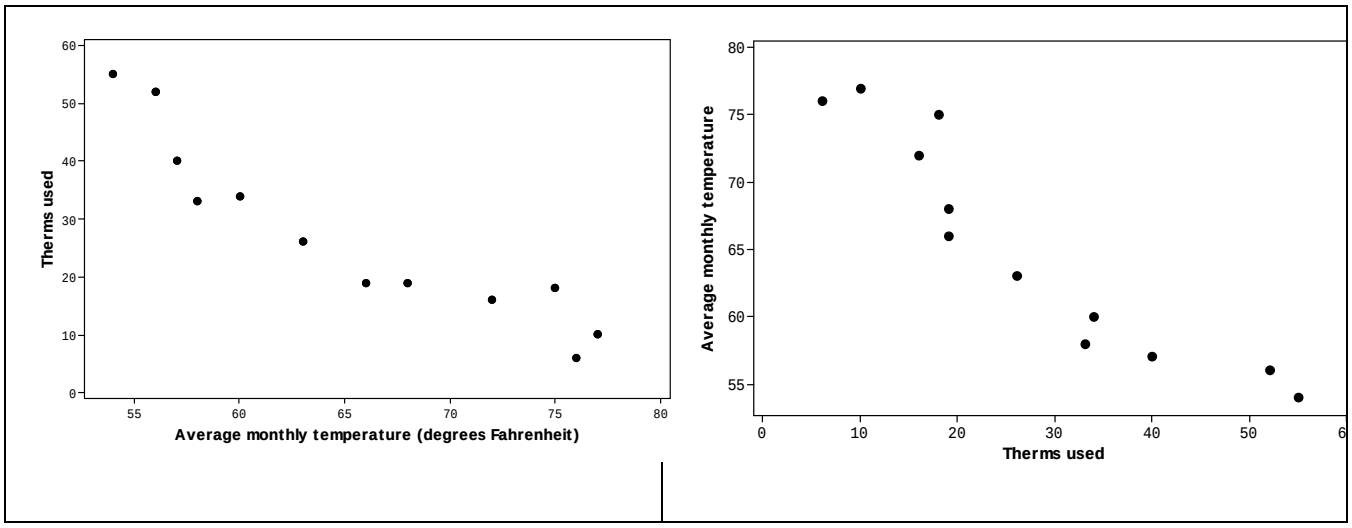
Solution

Part (a):

The number of therms used decreased each month from January to July, and then increased each month from July to December. In general, the lower the average monthly temperature, the more therms were used.

Part (b):

The appropriate graph is a scatter plot. Both options for the order of the axes are shown below.



Part (c):

There is a strong, negative, and somewhat curved relationship between therms used and average monthly temperature.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 1 (continued)

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response includes the following three components:
(1) the response correctly notes the decreasing and then increasing trend for therms used,
(2) notes that it was lowest in July, and
(3) describes the trend in context.

Partially correct (P) if the response includes components (1) and (2) but not (3);

OR

if the response has component (3) but describes the trend in terms that are too general, such as "therms used has a U-shaped distribution across the months" or "therms used were higher in winter than in summer" OR describes the trend over only a portion of the year, such as "therms decrease from January to July, and increase from July to September."

Incorrect (I) if the response does not meet the criteria for E or P.

Part (b) is scored as follows:

Essentially correct (E) if the response includes the following three components:
(1) A scatter plot of average monthly temperature and therms (with the axes in either order),
(2) All points in approximately the correct locations, and
(3) Scales and labels correctly placed on both axes.

Note: If one point is misplaced (or missing) or there is one minor error on one of the axes but everything else is correct, part (b) should be scored as essentially correct.

Partially correct (P) if the response has component (1) but is missing either component (2) or component (3).

Incorrect (I) if the response does not meet the criteria for E or P.

Part (c) is scored as follows:

Essentially correct (E) if the response notes that the relationship has the following four components:
(1) (moderately) strong,
(2) negative (or states that number of therms decreases as temperature increases),
(3) curved or (approximately) linear, and
(4) uses the context of the relationship between therms and temperature

Note: The word "correlation" is not synonymous with linear.

Partially correct (P) if the response includes two or three of the four components necessary for an E.

Incorrect (I) if the response does not meet the criteria for E or P.

Question 1 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

AP® STATISTICS
2016 SCORING GUIDELINES

Question 2

Intent of Question

The primary goals of this question are to assess a student's ability to perform an appropriate hypothesis test to address a particular question. More specific goals are to assess students' ability to (1) state appropriate hypotheses; (2) identify the name of an appropriate statistical test and check appropriate assumptions/conditions; (3) calculate the appropriate test statistic and *p*-value; (4) draw an appropriate conclusion, with justification, in the context of the study; and (5) determine whether a cause and effect conclusion is appropriate for a particular study.

Solution

Part (a):

Step 1: States a correct pair of hypotheses.

Let p_A and p_B represent the probability of a stroke during the 10 years of the study for the populations of Swedish men who would be classified into group A and group B, respectively.

The hypotheses to be tested are $H_0 : p_A = p_B$ versus $H_a : p_A < p_B$, or equivalently, $H_0 : p_A - p_B = 0$ versus $H_a : p_A - p_B < 0$.

Step 2: Identifies a correct test procedure (by name or by formula) and checks appropriate conditions.

The appropriate procedure is a two-sample *z*-test for comparing proportions.

The men in the study were randomly sampled from the population of Swedish men.

The second condition is that the sample sizes are large relative to the proportions involved. This condition is satisfied because all sample counts are larger than standard thresholds such as 5 and 10. There were 458 cases with strokes and 8,792 cases without strokes in group A, and 543 cases with strokes and 8,707 cases without strokes in group B.

An additional condition may be checked: The population sizes are much larger than 10 (or 20) times the sample sizes.

Step 3: Correct mechanics, including the value of the test statistic and *p*-value.

The sample proportions who had strokes are $\hat{p}_A = \frac{458}{9250} = 0.0495$ and $\hat{p}_B = \frac{543}{9250} = 0.0587$.

The combined proportion who had a stroke is $\hat{p}_{combined} = \frac{458 + 543}{9250 + 9250} = 0.0541$.

The test statistic is $z = \frac{0.0495 - 0.0587}{\sqrt{0.0541(1 - 0.0541)\left(\frac{1}{9250} + \frac{1}{9250}\right)}} \approx -2.77$.

The *p*-value is $P(Z \leq -2.77) = 0.003$, where Z has a standard normal distribution.

Question 2 (continued)

Step 4: State a correct conclusion in the context of the study, using the result of the statistical test.

Because the *p*-value is very small (for instance, less than 0.05), we reject the null hypothesis and conclude that the data provide convincing statistical evidence that the Swedish men who would be classified into group A have a lower probability of stroke than the Swedish men who would be classified into group B.

Part (b):

The conclusion in the report was not appropriate. The report implied that eating more chocolate would *cause* the probability of having a stroke to go down. But the results were based on an observational study, not a randomized experiment. So a cause and effect conclusion is not justified

Scoring

The scoring has 4 sections. Section 1 consists of part (a) steps 1 and 3, section 2 consists of part (a) step 2, section 3 consists of part (a) step 4, and section 4 consists of part (b). Sections 1, 2, 3, and 4 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response includes the following components:
1. Identifies correct parameters;
2. Includes both hypotheses with labels and stating the correct relationships between the parameters;
3. Correctly calculates or states the value of the test statistic, and;
4. Provides a correct *p*-value.

Partially correct (P) if the response includes only two or three of the components necessary for E.

Incorrect (I) if the response includes at most one of the components necessary for E.

Notes:
- For step 1, either defining the parameter symbols in context, or simply using common parameter notation, such as p_A and p_B, is sufficient. However, if parameter symbols are defined they must be done correctly, and must implicitly or explicitly refer to the population and not the samples.
- For step 3, the *p*-value is considered correct if it is consistent with the alternative hypothesis stated in the response and the calculated test statistic, even if those are incorrect.

Section 2 is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:
1. Identifies the correct test procedure (by name or by formula);
2. Notes the use of a random sample, and;
3. Checks for approximate normality by citing that all four counts are larger than some standard criterion such as 5 or 10, or by stating that all four counts are clearly greater than standard criteria.

Question 2 (continued)

Partially correct (P) if the response correctly includes only two of the three components required for E.

Incorrect (I) if the response correctly includes one or none of the three components required for E.

Note: Correctly using the combined proportion is sufficient for checking approximate normality.

Section 3 is scored as follows:

Essentially correct (E) if the response correctly provides a conclusion in context and in terms of the alternative hypothesis AND correctly provides a justification based on linkage between the *p*-value and the conclusion. (No value of alpha was provided, so it is not necessary to state a specific alpha for the linkage.)

Partially correct (P) if the response provides a correct decision to reject H_0 but with an incorrect conclusion, AND with either correct linkage, or context, or both;
OR
if the response provides a correct conclusion in terms of the alternative hypothesis but with only correct linkage or only context, but not both.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes:

- If the conclusion is consistent with an incorrect *p*-value from step 3, and also in context with justification based on linkage to the *p*-value, then section 3 is scored as E, unless the conclusion is stated as *accepting* the null hypothesis.
- If an appropriate value for α and the *p*-value are stated together, linkage is implied.
- If an interpretation of the *p*-value is included, it must be correct in order to receive credit for the conclusion.
- *For sections 1 to 3:* If the response includes a correct chi-squared test for homogeneity of proportions, then the response should be scored no higher than 3. If the chi-squared test *p*-value is halved to reflect the one-sided alternative hypothesis, the response should still be scored no higher than 3.

Section 4 is scored as follows:

Essentially correct (E) if the response states that the newspaper conclusion was not appropriate, and provides justification based on a specific characteristic of the study, such as being an observational study, not using random assignment, not being an experiment, or not *controlling* for confounding variables.

Partially correct (P) if the response states that the newspaper conclusion was not appropriate, and gives a reasonable but generic justification such as association doesn't imply causation or a vague reference to the existence of confounding variables.

Incorrect (I) if the response does not meet the criteria for E or P.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 2 (continued)

Each essentially correct (E) section counts as 1 point, and a partially correct (P) section counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and communication.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 3

Intent of Question

The primary goals of this question are to assess a student's ability to (1) identify which of two sampling plans is stratified sampling and which is cluster sampling; (2) provide an advantage and a disadvantage of a particular sampling plan; (3) provide an advantage of a particular type of sampling as compared to simple random sampling.

Solution

Part (a):

(i) Sampling method: 2 Strata: The two strata are US employees and European employees

(ii) Sampling method: 1 Clusters: Each store is a cluster

Part (b):

Advantage: Sampling method 1 might be easier to implement than other possible methods because the employees selected for the survey all work at one of the 8 stores in the sample. Therefore, it might be easier to go to those stores and survey the employees than to try to survey employees at hundreds of locations. Another advantage is that this method is likely to result in a higher response rate, and thus non-response bias is less likely to be a problem.

Disadvantage: The employees at the 8 selected stores may not be representative of all employees that work for the company. It is even possible that all 8 stores would be in the United States, or all in Europe, and there may be differences in how the employees in the United States and Europe would respond to the survey. Another disadvantage might be that responses of employees who work in the same store might not be independent of each other.

Part (c):

An advantage of sampling method 2 compared to a simple random sample is that the method guarantees that the proportions of US and European employees in the sample are the same as the proportions in the population. This proportional allocation results in unbiased estimates and decreased sampling variability. With a simple random sample it is possible to get too many employees from one region and not enough from the other. Also, with a simple random sample there may not be enough employees from one of the two regions to make a comparison.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

AP® STATISTICS
2016 SCORING GUIDELINES

Question 3 (continued)

Part (a) is scored as follows:

Essentially correct (E) if the response correctly identifies stratified sampling as sampling method 2 and cluster sampling as sampling method 1, *AND* identifies the regions as strata and the stores as clusters. (Strata may be identified as the regions or employees within each of the regions. Clusters may be identified as either the 8 stores in the sample or all possible stores. Clusters may also be identified as the employees at stores instead of the stores themselves.)

Partially correct (P) if the response switches the two methods, but correctly identifies that the regions are strata and the stores are clusters OR correctly identifies sampling method 2 as stratified sampling and sampling method 1 as cluster sampling, but does not correctly explain what constitutes one or both of "strata" and "clusters."

Incorrect (I) if the response says that stratified sampling is sampling method 1 with stores as the strata and cluster sampling is sampling method 2 with regions as clusters OR otherwise does not meet the criteria for E or P.

Part (b) is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:
1. A plausible statistical advantage of cluster sampling (sampling method 1) based on employees being located in only 8 stores
2. A plausible statistical disadvantage of the method based either on the fact that the employees in the 8 stores selected may not be representative of all employees, or on the fact that responses within stores may not be independent

Partially correct (P) if the response correctly includes only one of the two components required for E.

Incorrect (I) if the response correctly includes neither of the components required for E.

Part (c) is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:
1. Indicates that sampling method 2 results in proportions equal to those in the population; or results in decreased sampling variability; or assures representation from both regions.
2. Realizes that a simple random sample might result in proportions from the two regions that are not consistent with the population proportions or that there might not be enough responses from one of the regions to get accurate estimates.

Partially correct (P) if the response correctly includes only one of the two components required for E.

Incorrect (I) if the response correctly includes neither of the components required for E.

Notes:
- Statements that may be true but that do not truly represent advantages or disadvantages should not be scored as correct. An example is stating only "It is completely random." as an advantage for sampling method 1 in part (b).

- It is acceptable if the response notes as a statistical advantage that stratified random sampling (sampling method 2) results in equal proportions of employees drawn from both regions.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 3 (continued)

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

AP® STATISTICS
2016 SCORING GUIDELINES

Question 4

Intent of Question

The primary goals of this question are to assess a student's ability to (1) determine the consequence of a newly defined term involving probability; (2) perform a probability calculation involving the multiplication rule for independent events and the addition rule for disjoint events; and (3) apply understanding of variability to determine how to maximize a probability function.

Solution

Part (a):

If every organism in the population is from the same species, then the probability that two organisms chosen at random will be from different species is 0. Therefore, the value of the diversity index is 0.

Part (b):

The diversity index is the probability that when two turtles are chosen at random from the population they are different species, which is the probability that the first turtle is a snapping turtle and the second turtle is a box turtle, or vice versa. This probability is $(0.3)(0.7) + (0.7)(0.3) = 0.21 + 0.21 = 0.42$. Independence of the species for the two turtles is justified by random selection from a large population.

Part (c):

The maximum possible value of the diversity index is achieved when the population is as diverse as possible, meaning that each species is equally represented. In this case, that means the population would have 50% snapping turtles and 50% box turtles. In this situation the diversity index is the probability that two randomly chosen turtles are from different species, so it is:

$P(\text{snapper and box}) + P(\text{box and snapper}) = (0.5)(0.5) + (0.5)(0.5) = 0.25 + 0.25 = 0.5.$

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response has the following three components:
(1) provides the correct answer (0),
(2) provides a reasonable explanation based on the idea of probability (or impossibility), and
(3) is stated in context.

Question 4 (continued)

Partially correct (P) if the response has components (2) and (3) but, for component (1), answers "undefined" rather than 0;
OR
if the response has components (1) and (3) but has a weak explanation for component (2) not based on the idea of probability or impossibility (for example, stating "because there's no diversity" or "because all organisms are the same");
OR
if the response has components (1) and (2) but not component (3).

Incorrect (I) if the response otherwise does not meet the criteria for E or P.

Part (b) is scored as follows:

Essentially correct (E) if the response provides the correct answer and shows work.

Note: The response does not need to justify the use of the independence assumption.

Partially correct (P) if the response doesn't demonstrate that there are two different orderings of snapping and box turtles or completes only half of the calculation, such as $(0.3)(0.7) = 0.21$;
OR
if the response demonstrates correct use of the multiplication rule using the correct values for the population proportions but makes an error in performing the calculation.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: A response that in part (a) states that $D = 1$ and in part (b) states $D = 0.58$ should receive a score of E in part (b). However, a response that in part (b) states $D = 0.42$ should receive a score of P in part (b). A response with any other value of D in part (b) should receive a score of I for part (b).

Part (c) is scored as follows:

Essentially correct (E) if the response has the following three components:
(1) provides the correct answer that $D = 0.5$,
(2) shows computation of the value of D,
(3) justifies $D = 0.5$ in one of the following ways:
   i. with a reasonable explanation based on the idea that the maximum diversity index occurs when 50% of each species is present; *OR*
   ii. computes $2p(1 - p)$ for at least one other population proportion that differs from 0.5; *OR*
   iii. recognizes that the answer is whatever maximizes $2p(1 - p)$ AND notes that the answer would be the same value as when doing sample size or margin of error computations, which is known to be 0.5; *OR*
   iv. uses calculus to find the maximum or algebra to find the vertex of the parabola $2p(1 - p)$.

Question 4 (continued)

Notes:

- If the response in part (b) completes only half the calculation, and answers 0.21, then in part (c), component (1) should indicate that the maximum value of D is 0.25, and computations and/or formulas in components (2) and (3) should read $p(1-p)$.
- If the response in part (a) states that $D = 1$, then a response in part (c) of $D = 1$, with explanation that the answer in part (a) was $D = 1$, should receive a score of E. Otherwise, the response in part (c) receives a score of I.

Partially correct (P) if the response provides component (1) and either component (2) or component (3).

Incorrect (I) if the response does not meet the criteria for E or P.

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct

OR

No parts essentially correct and two parts partially correct

AP® STATISTICS
2016 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this question are to assess a student's ability to (1) calculate the mean and standard deviation of a binomial random variable; (2) use the normal distribution to determine an interval of values that have very high probability to occur, based on a binomial distribution; and (3) make and explain an inferential statement about a parameter based on an interval of likely values for a binomial random variable.

Solution

Part (a):

X has a binomial distribution with parameters $n = 2{,}000$ and $p = 0.6$. Therefore, $E(X) = np = (2{,}000)(0.6) = 1{,}200$ people. $\text{Var}(X) = np(1 - p) = (2{,}000)(0.6)(0.4) = 480$, and $\text{SD}(X) = \sqrt{480} \approx 21.91$ people.

Part (b):

According to the empirical rule, the probability is about 0.997 that the value of X will fall within 3 standard deviations of its mean. Based on the answer to part (a), the interval is $1{,}200 \pm 3(21.91) = 1{,}200 \pm 65.73$, which is the interval from 1,134.27 through 1,265.73. So $k_1 = 1{,}134.27$ and $k_2 = 1{,}265.73$. Using the calculator normal probability function, the values fall 2.97 standard deviations above and below the mean, so the endpoints are $1{,}200 \pm 65.07$, or 1,134.93 and 1,265.07.

Part (c):

Because 1,293 is greater than the entire interval determined in part (b), we know that a sample result of 1,293 supporting the proposal would be very unlikely to occur if in fact 60% of the population supports the proposal for a new park. The survey result is consistent with a higher mean than 1,200 for the distribution shown in part (b). Therefore, there is convincing evidence that more than 60% of the city residents support the new park, and that the proposal should therefore be funded.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:

1. The response provides the correct mean.
2. The response provides the correct standard deviation.
3. The response correctly shows how the standard deviation was calculated. (The response is not required to show how the mean was calculated.)

Question 5 (continued)

Partially correct (P) if the response includes only two of the three components
OR
if the response provides the correct mean and correct variance and shows how the variance was calculated but does not provide the correct standard deviation;
OR
if the response provides the correct values for μ, n and p and the correct formula for calculating the standard deviation, but makes a minor error in carrying out the calculations, such as writing $(0.6)(1 - 0.4)$ instead of $(0.6)(1 - 0.6)$ when calculating the standard deviation.

Incorrect (I) if the response otherwise does not meet the criteria for E or P.

Note: If the mean and standard deviation are labelled $\bar{x}$ and s respectively, a response that otherwise would earn an "E" should be given a "P;" a response that otherwise would earn a "P" should be given an "I."

Part (b) is scored as follows:

Essentially correct (E) if the response gives the correct values for the endpoints and justifies them either by showing work that includes the correct use of z-values of ±3 or ±2.97 or by demonstrating sufficient work that these values must have been used correctly to obtain the endpoints. Due to rounding at various stages of the calculation and using the table or the calculator, correct values may be anything in the range of 1,134 to 1,135 for k_1 and 1,265 to 1,266 for k_2.

Note: If the mean and standard deviation given in part (a) are incorrect, Part (b) should be scored according to the values given in part (a).

Partially correct (P) if the response gives the correct values for the endpoints and justifies them by showing reasonable work that is not sufficient to demonstrate the correct use of z-values of ±3 or ±2.97.
OR
if the response indicates that the interval includes 3 (or 2.97) standard deviations on either side of the mean, but calculates values corresponding to 3 (or 2.97) standard deviations not using the mean and standard deviations found in part (a);
OR
if the response uses a probability of 0.9997 rather than 0.997 and correctly uses any z-values between ±(3.39 and 3.48) to obtain anything in the range of 1,123.7 to 1,125.7 for k_1 and 1,274.3 to 1,276.3 for k_2;
OR
if the response uses a mean and standard deviation for the proportion instead of the counts, and therefore defines k_1 and k_2 as 0.57 and 0.63, unless the mean and standard deviation for the proportion were given as the response in part (a).

Incorrect (I) if the response does not meet the criteria for E or P.

Note: If the response does not provide values for the mean and/or standard deviation in part (a), part (b) may be scored as essentially correct if the formulas are given for the endpoints using the symbols or words for mean and standard deviation.

Question 5 (continued)

Part (c) is scored as follows:

Essentially correct (E) if the response correctly includes the following two components:
1. Gives the correct conclusion that there is convincing evidence that more than 60% of the city residents support the new park OR that the survey results support the funding of the proposal.
2. Bases the conclusion on the fact that the sample count (1,293) is above the entire interval from part (b).

Partially correct (P) if the response gives the correct conclusion that there is convincing evidence that more than 60% of the city residents support the new park OR that the survey results support the funding of the proposal, but only notes that 1,293 is not in the interval.

Incorrect (I) if the response gives a conclusion that is not consistent with the interval determined in part (b);
OR
if the response otherwise does not meet the criteria for E or P.

Notes:
- If the interval determined in part (b) does include the value 1,293, then the conclusion in part (c) must be consistent with that interval to earn E.
- If the response includes a hypothesis test for the null hypothesis that $p = 0.6$ or a confidence interval for p based on the data and ignores the result in part (b) it is scored as I.

4 Complete Response

Three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and no parts partially correct
OR
One part essentially correct and one or two parts partially correct
OR
Three parts partially correct

1 Minimal Response

One part essentially correct
OR
No parts essentially correct and two parts partially correct

AP® STATISTICS
2016 SCORING GUIDELINES

Question 6

Intent of Question

The primary goals of this question are to assess a student's ability to (1) check conditions for applying a one-sample *t*-interval procedure; (2) construct and interpret a one-sample *t*-interval; (3) convert log data back to original units; and (4) understand how a log transformation affects the mean and median of a distribution.

Solution

Part (a):

Using a one-sample *t*-procedure to produce a confidence interval for the population mean would not be appropriate because the distribution of wait times is clearly skewed to the right, and the sample size ($n = 13$) is not large.

Part (b):

The interval is found as $\bar{x} \pm t_{(12,\ 0.975)} \frac{s}{\sqrt{n}}$ or $0.063 \pm 2.179\left(\frac{0.235}{\sqrt{13}}\right) = 0.063 \pm 0.142$, giving the interval $(-0.079, 0.205)$. We can be 95% confident that the mean of the population of log wait times is between -0.079 log minutes and 0.205 log minutes.

Part (c):

The endpoints are $10^{-0.079}$ and $10^{0.205}$, so the interval is 0.834 minutes to 1.604 minutes.

Part (d):

(i) The parameter 10^{μ} is equal to the median of the population of wait times.

(ii) The parameter 10^{μ} is less than the mean of the population of wait times.

Part (e):

The interval in part (c) is a 95% confidence interval for the population *median* of the wait times. Therefore, we are 95% confident that the median of the population of wait times is between 0.834 minutes and 1.604 minutes.

AP® STATISTICS
2016 SCORING GUIDELINES

Question 6 (continued)

Scoring

This question is scored in 4 sections. Section 1 consists of part (a); section 2 consists of parts (b) and (c); section 3 consists of part (d); and section 4 consists of part (e). Sections 1, 2, 3, and 4 are scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response:
1. Comments that the distribution of the sample data is skewed (not symmetric).
2. Comments that the sample size is small.

Partially correct (P) if the response correctly includes only one of the two components.

Incorrect (I) if the response correctly includes neither of the two components.

Section 2 is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:
1. In part (b), reports the interval correctly.
2. In part (b), interprets the interval correctly including that it is for the mean of the log of the population of wait times.
3. In part (c), correctly computes both endpoints.

Partially correct (P) if the response correctly includes only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P.

Section 3 is scored as follows:

Essentially correct (E) if the response:
1. States that the parameter 10^{μ} is equal to the median of the population of wait times.
2. States that the parameter 10^{μ} is less than the mean of the population of wait times.

Partially correct (P) if the response correctly includes only one of the two components.

Incorrect (I) if the response correctly includes neither of the two components.

Section 4 is scored as follows:

Essentially correct (E) if the response correctly states that the interval is a confidence interval for the population median wait time AND gives a correct interpretation of a 95 percent confidence interval;
OR
if the response states that the interval is a confidence interval for 10^{μ} *AND* gives a correct interpretation of a 95% confidence interval.

Question 6 (continued)

Partially correct (P) if the response correctly states that the interval is a confidence interval for the population median of the wait times *OR* for 10^{μ}, but does not give a correct confidence interval interpretation;

Incorrect (I) if the response states that the confidence interval is for some other parameter, such as the mean of the population of wait times, or the median of the log wait time, and then gives a correct confidence interval interpretation based on that parameter;
OR
if the response otherwise does not meet the criteria for E or P.

Each essentially correct (E) section counts as 1 point, and a partially correct (P) section counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to decide whether to score up or down, depending on the strength of the response and communication.

Scoring Worksheet

The following provides a scoring worksheet and conversion table used for calculating a composite score of the exam.

2016 AP Statistics Scoring Worksheet

Section I: Multiple Choice

_____ × 1.2820 = _____

Number Correct (out of 39*) — Weighted Section I Score (Do not round)

Section II: Free Response

Question 1 _____ (out of 4) × 1.8750 = _____ (Do not round)

Question 2 _____ (out of 4) × 1.8750 = _____ (Do not round)

Question 3 _____ (out of 4) × 1.8750 = _____ (Do not round)

Question 4 _____ (out of 4) × 1.8750 = _____ (Do not round)

Question 5 _____ (out of 4) × 1.8750 = _____ (Do not round)

Question 6 _____ (out of 4) × 3.1250 = _____ (Do not round)

Sum = _____

Weighted Section II Score (Do not round)

Composite Score

_____ + _____ = _____

Weighted Section I Score — Weighted Section II Score — Composite Score (Round to nearest whole number)

AP Score Conversion Chart
Statistics

Composite Score Range	AP Score
70-100	5
58-69	4
44-57	3
35-43	2
0-34	1

*Although 40 multiple-choice items were administered in Section I, item 21 was not used in scoring.

Question Descriptors and Performance Data

The following contains tables showing the content assessed,
the correct answer, and how AP students performed on each question.

2016 AP Statistics

Question Descriptors and Performance Data

Multiple-Choice Questions

Question	Topic	Key	% Correct
1	Exploring Data	C	65
2	Exploring Data	B	90
3	Statistical Inference	A	89
4	Exploring Data	A	68
5	Probability and Simulation	E	31
6	Sampling and Experimentation	E	90
7	Exploring Data	E	84
8	Statistical Inference	E	28
9	Probability and Simulation	A	75
10	Statistical Inference	E	27
11	Statistical Inference	B	66
12	Probability and Simulation	D	50
13	Exploring Data	A	78
14	Probability and Simulation	D	56
15	Statistical Inference	B	55
16	Statistical Inference	A	88
17	Statistical Inference	E	80
18	Sampling and Experimentation	C	86
19	Exploring Data	B	91
20	Exploring Data	D	85
21	Probability and Simulation	*	*
22	Probability and Simulation	D	58
23	Statistical Inference	A	62
24	Statistical Inference	C	74
25	Sampling and Experimentation	B	71
26	Statistical Inference	E	39
27	Probability and Simulation	E	65
28	Statistical Inference	C	62
29	Sampling and Experimentation	B	56
30	Statistical Inference	C	72
31	Statistical Inference	D	54
32	Exploring Data	D	74
33	Exploring Data	B	41
34	Statistical Inference	E	63
35	Probability and Simulation	C	41
36	Exploring Data	E	59

* Item not included in scoring

2016 AP Statistics

Question Descriptors and Performance Data

Question	Topic	Key	% Correct
37	Statistical Inference	A	42
38	Sampling and Experimentation	B	80
39	Probability and Simulation	A	20
40	Probability and Simulation	C	50

AP Statistics

The College Board

The College Board is a mission-driven not-for-profit organization that connects students to college success and opportunity. Founded in 1900, the College Board was created to expand access to higher education. Today, the membership association is made up of over 6,000 of the world's leading educational institutions and is dedicated to promoting excellence and equity in education. Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services in college readiness and college success — including the SAT® and the Advanced Placement Program®. The organization also serves the education community through research and advocacy on behalf of students, educators, and schools. The College Board is committed to the principles of excellence and equity, and that commitment is embodied in all of its programs, services, activities, and concerns.