CollegeBoard AP

STATISTICS

Course Description

Effective Fall 2010

AP Course Descriptions are updated regularly. Please visit AP Central® (apcentral.collegeboard.com) to determine whether a more recent Course Description PDF is available.

The College Board

The College Board is a not-for-profit membership association whose mission is to connect students to college success and opportunity. Founded in 1900, the College Board is composed of more than 5,700 schools, colleges, universities and other educational organizations. Each year, the College Board serves seven million students and their parents, 23,000 high schools, and 3,800 colleges through major programs and services in college readiness, college admission, guidance, assessment, financial aid, and enrollment. Among its widely recognized programs are the SAT®, the PSAT/NMSQT®, the Advanced Placement Program® (AP®), SpringBoard® and ACCUPLACER®. 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.

For further information, visit www.collegeboard.com.

The College Board and the Advanced Placement Program encourage teachers, AP Coordinators and school administrators to make equitable access a guiding principle for their AP programs. The College Board is committed to the principle that all students deserve an opportunity to participate in rigorous and academically challenging courses and programs. All students who are willing to accept the challenge of a rigorous academic curriculum should be considered for admission to AP courses. The Board encourages the elimination of barriers that restrict access to AP courses for students from ethnic, racial and socioeconomic groups that have been traditionally underrepresented in the AP Program. Schools should make every effort to ensure that their AP classes reflect the diversity of their student population.

Contents

Welcome to the AP® Program

AP® is a rigorous academic program built on the commitment, passion and hard work of students and educators from both secondary schools and higher education. With more than 30 courses in a wide variety of subject areas, AP provides willing and academically prepared high school students with the opportunity to study and learn at the college level.

Through AP courses, talented and dedicated AP teachers help students develop and apply the skills, abilities and content knowledge they will need later in college. Each AP course is modeled upon a comparable college course, and college and university faculty play a vital role in ensuring that AP courses align with college-level standards. For example, through the AP Course Audit, AP teachers submit their syllabi for review and approval by college faculty. Only courses using syllabi that meet or exceed the college-level curricular and resource requirements for each AP course are authorized to carry the "AP" label.

AP courses culminate in a suite of college-level assessments developed and scored by college and university faculty members as well as experienced AP teachers. AP Exams are an essential part of the AP experience, enabling students to demonstrate their mastery of college-level course work. Strong performance on AP Exams is rewarded by colleges and universities worldwide. More than 90 percent of four-year colleges and universities in the United States grant students credit, placement or both on the basis of successful AP Exam scores. But performing well on an AP Exam means more than just the successful completion of a course; it is the gateway to success in college. Research consistently shows that students who score a 3 or higher typically experience greater academic success in college and improved graduation rates than their non-AP student peers.

AP Course Audit

The intent of the AP Course Audit is to provide secondary and higher education constituents with the assurance that an "AP" designation on a student's transcript is credible, meaning the AP Program has authorized a course that has met or exceeded the curricular requirements and classroom resources that demonstrate the academic rigor of a comparable college course. To receive authorization from the College Board to label a course "AP," teachers must participate in the AP Course Audit. Courses authorized to use the "AP" designation are listed in the AP Course Ledger made available to colleges and universities each fall. It is the school's responsibility to ensure that its AP Course Ledger entry accurately reflects the AP courses offered within each academic year.

The AP Program unequivocally supports the principle that each individual school must develop its own curriculum for courses labeled "AP." Rather than mandating any one curriculum for AP courses, the AP Course Audit instead provides each AP teacher with a set of expectations that college and secondary school faculty nationwide have established for college-level courses. AP teachers are encouraged to develop or maintain their own curriculum that either includes or exceeds each of these expectations; such courses will be authorized to use the "AP" designation. Credit for the success of AP courses belongs to the individual schools and teachers that create powerful, locally designed AP curricula.

Complete information about the AP Course Audit is available at www.collegeboard.com/apcourseaudit.

AP Development Committees

An AP Development Committee is a group of nationally renowned subject-matter experts in a particular discipline that includes professionals in secondary and postsecondary education as well as from professional organizations. These experts ensure that AP courses and exams reflect the most up-to-date information available, as befitting a college-level course, and that student proficiency is assessed properly. To find a list of current AP Development Committee members, please visit: apcentral.collegeboard.com/developmentcommittees.

AP Reading

AP Exams — with the exception of AP Studio Art, which is a portfolio assessment — consist of dozens of multiple-choice questions scored by machine, and free-response questions scored at the annual AP Reading by thousands of college faculty and expert AP teachers. AP Readers use scoring standards developed by college and university faculty who teach the corresponding college course. The AP Reading offers educators both significant professional development and the opportunity to network with colleagues. For more information about the AP Reading, or to apply to serve as a Reader, visit apcentral.collegeboard.com/readers.

AP Exam Scores

The Readers' scores on the free-response questions are combined with the results of the computer-scored multiple-choice questions; the weighted raw scores are summed to give a composite score. The composite score is then converted to a score on AP's 5-point scale. While colleges and universities are responsible for setting their own credit and placement policies, AP scores signify how qualified students are to receive college credit or placement:

AP SCORE	QUALIFICATION
5	Extremely well qualified
4	Well qualified
3	Qualified
2	Possibly qualified
1	No recommendation

AP Exam scores of 5 are equivalent to A grades in the corresponding college course. AP Exam scores of 4 are equivalent to grades of A–, B+ and B in college. AP Exam scores of 3 are equivalent to grades of B–, C+ and C in college.

Credit and Placement for AP Scores

Thousands of two- and four-year colleges and universities grant credit, placement or both for qualifying AP Exam scores because these scores represent a level of

achievement equivalent to that of students who have taken the comparable college course. This college-level equivalency is ensured through several AP Program processes:

- College faculty are involved in course and exam development and other AP activities. Currently, college faculty:
  - Serve as chairs and members of the committees that develop the Course Descriptions and exams for each AP course.
  - Are responsible for standard setting and are involved in the evaluation of student responses at the annual AP Reading. The Chief Reader for each AP exam is a college faculty member.
  - Lead professional development seminars for new and experienced AP teachers.
  - Serve as the senior reviewers in the annual AP Course Audit, ensuring AP teachers' syllabi meet the curriculum guidelines for college-level courses.
- AP courses and exams are reviewed and updated regularly based on the results of curriculum surveys at up to 200 colleges and universities, collaborations among the College Board and key educational and disciplinary organizations, and the interactions of committee members with professional organizations in their discipline.
- Periodic college comparability studies are undertaken in which the performance of college students on a selection of AP Exam questions is compared with that of AP students to ensure that grades earned by college students are aligned with scores AP students earn on the exam.

For more information about the role of colleges and universities in the AP Program, visit the Value of AP to Colleges and Universities section of the College Board website at http://professionals.collegeboard.com/higher-ed/placement/ap.

Setting Credit and Placement Policies for AP Scores

The College Board website for education professionals has a section specifically for colleges and universities that provides guidance in setting AP credit and placement policies. Visit http://professionals.collegeboard.com/higher-ed/placement/ap/policy.

Additional resources, including links to AP research studies, released exam questions and sample student responses at varying levels of achievement for each AP Exam are also available. To view student samples and scoring guidelines, visit http://apcentral.collegeboard.com/apc/public/exam/exam_questions/index.html.

To review recent validity research studies, visit http://professionals.collegeboard.com/data-reports-research/cb/ap.

The "AP Credit Policy Info" online search tool provides links to credit and placement policies at more than 1,000 colleges and universities. This tool helps students find the credit hours and/or advanced placement they may receive for qualifying exam scores within each AP subject at a specified institution. AP Credit Policy Info is available at www.collegeboard.com/ap/creditpolicy. If the information for your institution is not listed or is incorrect, please contact aphighered@collegeboard.org.

AP Statistics

INTRODUCTION

The Advanced Placement Program offers a course description and exam in statistics to secondary school students who wish to complete studies equivalent to a one-semester, introductory, non-calculus-based, college course in statistics.

Statistics and mathematics educators who serve as members of the AP Statistics Development Committee have prepared the Course Description and exam to reflect the content of a typical introductory college course in statistics. The exam is representative of such a course and therefore is considered appropriate for the measurement of skills and knowledge in the field of introductory statistics.

In colleges and universities, the number of students who take a statistics course is almost as large as the number of students who take a calculus course. A July 2002 article in the *Chronicle of Higher Education* reports that the enrollment in statistics courses from 1990 to 2000 increased by 45 percent — one testament to the growth of statistics in those institutions. An introductory statistics course, similar to the AP Statistics course, is typically required for majors such as social sciences, health sciences and business. Every semester about 236,000 college and university students enroll in an introductory statistics course offered by a mathematics or statistics department. In addition, a large number of students enroll in an introductory statistics course offered by other departments. Science, engineering and mathematics majors usually take an upper-level calculus-based course in statistics, for which the AP Statistics course is effective preparation.

THE COURSE

The purpose of the AP course in statistics is to introduce students to the major concepts and tools for collecting, analyzing and drawing conclusions from data. Students are exposed to four broad conceptual themes:

1. Exploring Data: Describing patterns and departures from patterns
2. Sampling and Experimentation: Planning and conducting a study
3. Anticipating Patterns: Exploring random phenomena using probability and simulation
4. Statistical Inference: Estimating population parameters and testing hypotheses

Students who successfully complete the course and exam may receive credit, advanced placement or both for a one-semester introductory college statistics course. This does not necessarily imply that the high school course should be one semester long. Each high school needs to determine the length of its AP Statistics course to best serve the needs of its students. Statistics, like some other AP courses, could be effectively studied in a one-semester, a two-trimester or a one-year course. Most schools, however, offer it as a one-year course.

Student Selection

The College Board and the Advanced Placement Program encourage teachers, AP Coordinators and school administrators to make equitable access a guiding principle for their AP programs. The College Board is committed to the principle that all students deserve an opportunity to participate in rigorous and academically challenging courses and programs. All students who are willing to accept the challenge of a rigorous academic curriculum should be considered for admission to AP courses. The Board encourages the elimination of barriers that restrict access to AP courses for students from ethnic, racial and socioeconomic groups that have been traditionally underrepresented in the AP Program. Schools should make every effort to ensure that their AP classes reflect the diversity of their student population.

The AP Statistics course is an excellent option for any secondary school student who has successfully completed a second-year course in algebra and who possesses sufficient mathematical maturity and quantitative reasoning ability.

Because second-year algebra is the prerequisite course, AP Statistics usually will be taken in either the junior or senior year. The decisions about whether to take AP Statistics and when to take it depend on a student's plans:

- Students planning to take a science course in their senior year will benefit greatly from taking AP Statistics in their junior year.
- For students who would otherwise take no mathematics in their senior year, AP Statistics allows them to continue to develop their quantitative skills.
- Students who wish to leave open the option of taking calculus in college should include precalculus in their high school program and perhaps take AP Statistics concurrently with precalculus.

Students with the appropriate mathematical background are encouraged to take both AP Statistics and AP Calculus in high school.

Students who take the AP Statistics course are strongly encouraged to take the exam.

Teaching the Course

The AP Statistics course lends itself naturally to a mode of teaching that engages students in constructing their own knowledge. For example, students working individually or in small groups can plan and perform data collection and analyses where the teacher serves in the role of a consultant, rather than a director. This approach gives students ample opportunity to think through problems, make decisions and share questions and conclusions with other students as well as with the teacher.

Important components of the course should include the use of technology, projects and laboratories, cooperative group problem-solving, and writing, as a part of concept-oriented instruction and assessment. This approach to teaching AP Statistics will allow students to build interdisciplinary connections with other subjects and with their world outside school.

The AP Statistics course depends heavily on the availability of technology suitable for the interactive, investigative aspects of data analysis. Therefore, schools should make every effort to provide students and teachers easy access to computers to facilitate the teaching and learning of statistics.

Providing instructional information and educational opportunities for teachers is an important component of the AP Program. The College Board offers workshops, summer courses and institutes for teachers in all AP courses. Further information about these and other training opportunities may be obtained at AP Central (apcentral.collegeboard.com) and from your College Board regional office (contact information is on the inside back cover). The Teachers' Resources section of AP Central offers reviews of textbooks, articles, websites and other teaching resources. The electronic discussion groups (EDGs) accessible through AP Central also provide a moderated forum for exchanging ideas, insights and practices among members of the AP professional community.

Additionally, the following publications provide some insight into the philosophy of the AP Statistics course.

Cobb, George, "Teaching Statistics: More Data, Less Lecturing." 1992. In *Heeding the Call for Change: Suggestions for Curricular Action,* ed. Lynn Arthur. Washington, D.C.: Mathematical Association of America. To order MAA publications call 800-331-1622 or go to www.maa.org.

Gordon, Florence and Sheldon, eds. 1992. *Statistics for the Twenty-First Century,* MAA Notes No. 26. Washington, D.C.: Mathematical Association of America.

Moore, Thomas, ed. 2000. *Teaching Statistics: Resources for Undergraduate Instructors,* MAA Notes Vol. 52. Washington, D.C.: Mathematical Association of America.

National Council of Teachers of Mathematics. 2003. *Principles and Standards for School Mathematics,* 3rd ed. Reston, Va.: National Council of Teachers of Mathematics. www.nctm.org.

Course Content Overview

The topics for AP Statistics are divided into four major themes: exploratory analysis (20–30 percent of the exam), planning and conducting a study (10–15 percent of the exam), probability (20–30 percent of the exam), and statistical inference (30–40 percent of the exam).

I. *Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns.* In examining distributions of data, students should be able to detect important characteristics, such as shape, location, variability and unusual values. From careful observations of patterns in data, students can generate conjectures about relationships among variables. The notion of how one variable may be associated with another permeates almost all of statistics, from simple comparisons of proportions through linear regression. The difference between association and causation must accompany this conceptual development throughout.

II. *Data must be collected according to a well-developed plan if valid information is to be obtained.* If data are to be collected to provide an answer to a question of interest, a careful plan must be developed. Both the type of analysis that is appropriate and the nature of conclusions that can be drawn from that analysis depend in a critical way on how the data was collected. Collecting data in a reasonable way, through either sampling or experimentation, is an essential step in the data analysis process.

III. *Probability is the tool used for anticipating what the distribution of data should look like under a given model.* Random phenomena are not haphazard: they display an order that emerges only in the long run and is described by a distribution. The mathematical description of variation is central to statistics. The probability required for statistical inference is not primarily axiomatic or combinatorial but is oriented toward using probability distributions to describe data.

IV. *Statistical inference guides the selection of appropriate models.* Models and data interact in statistical work: models are used to draw conclusions from data, while the data are allowed to criticize and even falsify the model through inferential and diagnostic methods. Inference from data can be thought of as the process of selecting a reasonable model, including a statement in probability language, of how confident one can be about the selection.

TOPIC OUTLINE

Following is an outline of the major topics covered by the AP Statistics Exam. The ordering here is intended to define the scope of the course but not necessarily the sequence. The percentages in parentheses for each content area indicate the coverage for that content area in the exam.

I. Exploring Data: Describing patterns and departures from patterns (20%–30%)

Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns. Emphasis should be placed on interpreting information from graphical and numerical displays and summaries.

A. Constructing and interpreting graphical displays of distributions of univariate data (dotplot, stemplot, histogram, cumulative frequency plot)
   1. Center and spread
   2. Clusters and gaps
   3. Outliers and other unusual features
   4. Shape

B. Summarizing distributions of univariate data
   1. Measuring center: median, mean
   2. Measuring spread: range, interquartile range, standard deviation
   3. Measuring position: quartiles, percentiles, standardized scores (z-scores)
   4. Using boxplots
   5. The effect of changing units on summary measures

C. Comparing distributions of univariate data (dotplots, back-to-back stemplots, parallel boxplots)
   1. Comparing center and spread: within group, between group variation
   2. Comparing clusters and gaps
   3. Comparing outliers and other unusual features
   4. Comparing shapes

D. Exploring bivariate data
   1. Analyzing patterns in scatterplots
   2. Correlation and linearity
   3. Least-squares regression line
   4. Residual plots, outliers and influential points
   5. Transformations to achieve linearity: logarithmic and power transformations

E. Exploring categorical data
   1. Frequency tables and bar charts
   2. Marginal and joint frequencies for two-way tables
   3. Conditional relative frequencies and association
   4. Comparing distributions using bar charts

II. Sampling and Experimentation: Planning and conducting a study (10%–15%)

Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.

A. Overview of methods of data collection
   1. Census
   2. Sample survey
   3. Experiment
   4. Observational study

B. Planning and conducting surveys
   1. Characteristics of a well-designed and well-conducted survey
   2. Populations, samples and random selection
   3. Sources of bias in sampling and surveys
   4. Sampling methods, including simple random sampling, stratified random sampling and cluster sampling

C. Planning and conducting experiments
   1. Characteristics of a well-designed and well-conducted experiment
   2. Treatments, control groups, experimental units, random assignments and replication
   3. Sources of bias and confounding, including placebo effect and blinding
   4. Completely randomized design
   5. Randomized block design, including matched pairs design

D. Generalizability of results and types of conclusions that can be drawn from observational studies, experiments and surveys

III. Anticipating Patterns: Exploring random phenomena using probability and simulation (20%–30%)

Probability is the tool used for anticipating what the distribution of data should look like under a given model.

A. Probability
   1. Interpreting probability, including long-run relative frequency interpretation
   2. "Law of Large Numbers" concept
   3. Addition rule, multiplication rule, conditional probability and independence
   4. Discrete random variables and their probability distributions, including binomial and geometric
   5. Simulation of random behavior and probability distributions
   6. Mean (expected value) and standard deviation of a random variable, and linear transformation of a random variable

B. Combining independent random variables
   1. Notion of independence versus dependence
   2. Mean and standard deviation for sums and differences of independent random variables

C. The normal distribution
   1. Properties of the normal distribution
   2. Using tables of the normal distribution
   3. The normal distribution as a model for measurements

D. Sampling distributions
   1. Sampling distribution of a sample proportion
   2. Sampling distribution of a sample mean
   3. Central Limit Theorem
   4. Sampling distribution of a difference between two independent sample proportions
   5. Sampling distribution of a difference between two independent sample means
   6. Simulation of sampling distributions
   7. t-distribution
   8. Chi-square distribution

IV. Statistical Inference: Estimating population parameters and testing hypotheses (30%–40%)

Statistical inference guides the selection of appropriate models.

A. Estimation (point estimators and confidence intervals)
   1. Estimating population parameters and margins of error
   2. Properties of point estimators, including unbiasedness and variability
   3. Logic of confidence intervals, meaning of confidence level and confidence intervals, and properties of confidence intervals
   4. Large sample confidence interval for a proportion
   5. Large sample confidence interval for a difference between two proportions

6. Confidence interval for a mean
7. Confidence interval for a difference between two means (unpaired and paired)
8. Confidence interval for the slope of a least-squares regression line

B. Tests of significance
1. Logic of significance testing, null and alternative hypotheses; p-values; one- and two-sided tests; concepts of Type I and Type II errors; concept of power
2. Large sample test for a proportion
3. Large sample test for a difference between two proportions
4. Test for a mean
5. Test for a difference between two means (unpaired and paired)
6. Chi-square test for goodness of fit, homogeneity of proportions, and independence (one- and two-way tables)
7. Test for the slope of a least-squares regression line

The Use of Technology

The AP Statistics course adheres to the philosophy and methods of modern data analysis. Although the distinction between graphing calculators and computers is becoming blurred as technology advances, at present the fundamental tool of data analysis is the computer. The computer does more than eliminate the drudgery of hand computation and graphing — it is an essential tool for structured inquiry.

Data analysis is a journey of discovery. It is an iterative process that involves a dialogue between the data and a mathematical model. As more is learned about the data, the model is refined and new questions are formed. The computer aids in this journey in some essential ways. First, it produces graphs that are specifically designed for data analysis. These graphical displays make it easier to observe patterns in data, to identify important subgroups of the data and to locate any unusual data points. Second, the computer allows the student to fit complex mathematical models to the data and to assess how well the model fits the data by examining the residuals. Finally, the computer is helpful in identifying an observation that has an undue influence on the analysis and in isolating its effects.

In addition to its use in data analysis, the computer facilitates the simulation approach to probability that is emphasized in the AP Statistics course. Probabilities of random events, probability distributions of random variables and sampling distributions of statistics can be studied conceptually, using simulation. This frees the student and teacher from a narrow approach that depends on a few simple probabilistic models.

Because the computer is central to what statisticians do, it is considered essential for teaching the AP Statistics course. However, it is not yet possible for students to have access to a computer during the AP Statistics Exam. Without a computer and under the conditions of a timed exam, students cannot be asked to perform the amount of computation that is needed for many statistical investigations. Consequently, standard computer output will be provided as necessary and students will be expected to interpret it. (See two examples of computer output in the Multiple-Choice Questions section on pages 21 and 23.)

A graphing calculator is a useful computational aid, particularly in analyzing small data sets, but should not be considered equivalent to a computer in the teaching of statistics. If a graphing calculator is used in the course, its computational capabilities should include standard statistical univariate and bivariate summaries through linear regression. Its graphical capabilities should include common univariate and bivariate displays such as histograms, boxplots and scatterplots. Students find calculators where data are entered into a spreadsheet format particularly easy to use. Ideally, students should have access to both computers and calculators for work in and outside the classroom.

Currently, the graphing calculator is the only computational aid that is available to students for use as a tool for data analysis on the AP Exam. Students who utilize graphing calculators on the exam should be aware of the following policy.

> It is not only inappropriate, but unethical, for students who are taking the AP Statistics Exam to have access to any information in their graphing calculators or elsewhere that is not directly related to upgrading the statistical functionality of older graphing calculators to make them comparable to statistical features found on newer models.
>
> During the exam, students are not permitted to have access to any information in their graphing calculators or elsewhere that is not directly related to upgrading the statistical functionality of older graphing calculators to make them comparable to statistical features found on newer models.
>
> Acceptable upgrades include improving the calculator's computational functionalities and/or graphical functionalities for data that students key into the calculator while taking the exam.
>
> Unacceptable enhancements include, but are not limited to, keying or scanning text or response templates into the calculator. Students attempting to augment the capabilities of their graphing calculators in any way other than for the purpose of upgrading features, as described above, will be considered to be cheating on the exam.

Formulas and Tables

Students enrolled in the AP Statistics course should concentrate their time and effort on developing a thorough understanding of the fundamental concepts of statistics. They do not need to memorize formulas.

The following list of formulas and tables will be furnished to students taking the AP Statistics Exam. Teachers are encouraged to familiarize their students with the form and notation of these formulas by making them accessible at the appropriate times during the course.

I. Descriptive Statistics

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

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

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

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

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

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

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

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

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

II. Probability

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

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

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

$$Var(X) = \sigma_x^2 = \sum (x_i - \mu_x)^2 p_i$$

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

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

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

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	$\frac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\frac{p(1-p)}{n}}$

Two-Sample

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

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

Table entry for z is the probability lying below z.

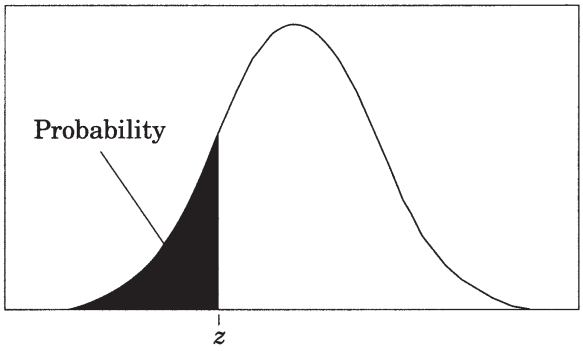


Table A Standard normal probabilities

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.

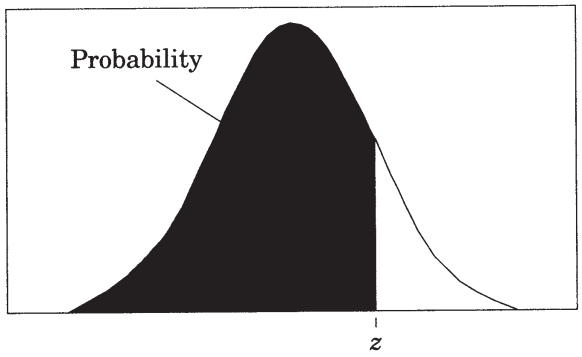


Table A *(Continued)* Standard normal probabilities

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^*.

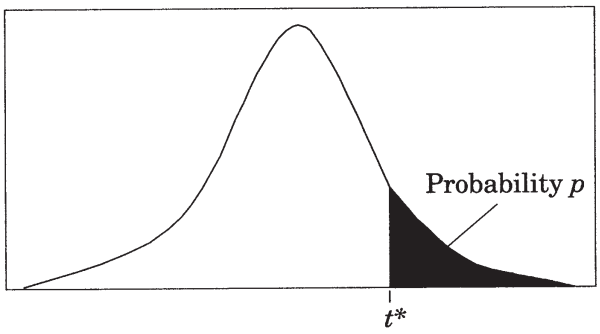


Table B

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.

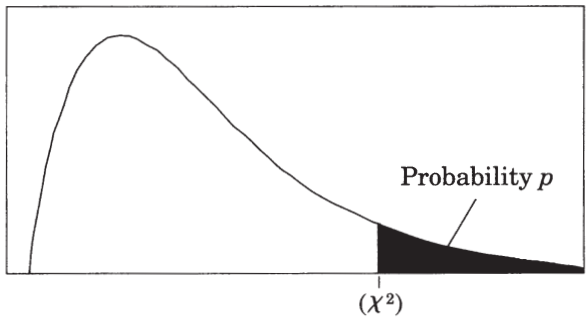


Table C

χ^2 critical values

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

THE EXAM

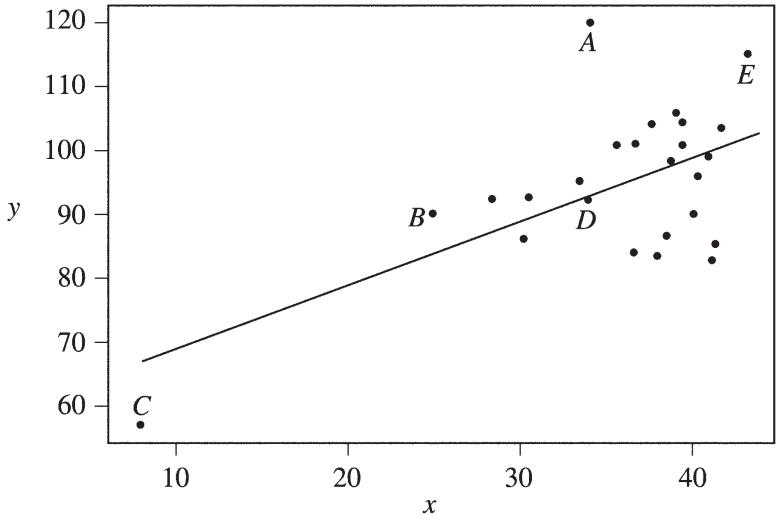
The AP Statistics Exam is 3 hours long and seeks to determine how well a student has mastered the concepts and techniques of the subject matter of the course. This paper-and-pencil exam consists of (1) a 90-minute multiple-choice section testing proficiency in a wide variety of topics, and (2) a 90-minute free-response section requiring the student to answer open-ended questions and to complete an investigative task involving more extended reasoning. In the determination of the score for the exam, the two sections will be given equal weight.

Each student will be expected to bring a graphing calculator with statistical capabilities to the exam. The expected computational and graphic features for these calculators are described in an earlier section. Minicomputers, pocket organizers, electronic writing pads (e.g., Newton) and calculators with QWERTY (i.e., typewriter) keyboards will not be allowed. Calculator memories will not be cleared. However, calculator memories may be used only for storing programs, not for storing notes. During the exam, students are not permitted to have access to any information in their graphing calculators or elsewhere that is not directly related to upgrading the statistical functionality of older graphing calculators to make them comparable to statistical features found on newer models. Acceptable upgrades include improving the calculator's computational functionalities and/or graphical functionalities for data that students key into the calculator while taking the exam. Unacceptable enhancements include, but are not limited to, keying or scanning text or response templates into the calculator. Students attempting to augment the capabilities of their graphing calculators in any way other than for the purpose of upgrading features as described above will be considered to be cheating on the exam. A student may bring up to two calculators to the exam.

Multiple-Choice Questions

The following are examples of the kinds of multiple-choice questions found on the AP Statistics Exam; the answers to these questions follow question 18. The distribution of topics and the levels of difficulty are illustrative of the composition of the exam; however, this group of questions does not constitute a complete exam, nor does it show the complete range of questions that appear in an exam.

Multiple-choice scores are based on the number of questions answered correctly. Points are not deducted for incorrect answers, and no points are awarded for unanswered questions. Because no points are deducted for incorrect answers, students are encouraged to answer all multiple-choice questions. On difficult questions, students should eliminate as many incorrect answer choices as they can, and then make an educated guess among the remaining choices.



1. In the scatterplot of y versus x shown above, the least squares regression line is superimposed on the plot. Which of the following points has the largest residual?

   (A) A
   (B) B
   (C) C
   (D) D
   (E) E

2. Under which of the following conditions is it preferable to use stratified random sampling rather than simple random sampling?

   (A) The population can be divided into a large number of strata so that each stratum contains only a few individuals.
   (B) The population can be divided into a small number of strata so that each stratum contains a large number of individuals.
   (C) The population can be divided into strata so that the individuals in each stratum are as much alike as possible.
   (D) The population can be divided into strata so that the individuals in each stratum are as different as possible.
   (E) The population can be divided into strata of equal sizes so that each individual in the population still has the same chance of being selected.

3. All bags entering a research facility are screened. Ninety-seven percent of the bags that contain forbidden material trigger an alarm. Fifteen percent of the bags that do not contain forbidden material also trigger the alarm. If 1 out of every 1,000 bags entering the building contains forbidden material, what is the probability that a bag that triggers the alarm will actually contain forbidden material?

   (A) 0.00097
   (B) 0.00640
   (C) 0.03000
   (D) 0.14550
   (E) 0.97000

4. A candy company claims that 10 percent of its candies are blue. A random sample of 200 of these candies is taken, and 16 are found to be blue. Which of the following tests would be most appropriate for establishing whether the candy company needs to change its claim?

   (A) Matched pairs *t*-test
   (B) One-sample proportion *z*-test
   (C) Two-sample *t*-test
   (D) Two-sample proportion *z*-test
   (E) Chi-square test of association

DESCRIPTIVE STATISTICS

Variable	N	Mean	Median	TrMean	StDev	SE Mean
score	50	1045.7	1024.7	1041.9	221.9	31.4

Variable	Minimum	Maximum	Q1	Q3
score	628.9	1577.1	877.7	1219.5

5. Some descriptive statistics for a set of test scores are shown above. For this test, a certain student has a standardized score of $z = -1.2$. What score did this student receive on the test?

   (A) 266.28
   (B) 779.42
   (C) 1008.02
   (D) 1083.38
   (E) 1311.98

6. In a test of H_0: $\mu = 8$ *versus* H_a: $\mu \neq 8$, a sample of size 220 leads to a *p*-value of 0.034. Which of the following must be true?

   (A) A 95% confidence interval for μ calculated from these data will not include $\mu = 8$.
   (B) At the 5% level if H_0 is rejected, the probability of a Type II error is 0.034.
   (C) The 95% confidence interval for μ calculated from these data will be centered at $\mu = 8$.
   (D) The null hypothesis should not be rejected at the 5% level.
   (E) The sample size is insufficient to draw a conclusion with 95% confidence.

7. A summer resort rents rowboats to customers but does not allow more than four people to a boat. Each boat is designed to hold no more than 800 pounds. Suppose the distribution of adult males who rent boats, including their clothes and gear, is normal with a mean of 190 pounds and standard deviation of 10 pounds. If the weights of individual passengers are independent, what is the probability that a group of four adult male passengers will exceed the acceptable weight limit of 800 pounds?

   (A) 0.023
   (B) 0.046
   (C) 0.159
   (D) 0.317
   (E) 0.977

8. Consider a data set of positive values, at least two of which are not equal. Which of the following sample statistics will be changed when each value in this data set is multiplied by a constant whose absolute value is greater than 1?

   I. The mean
   II. The median
   III. The standard deviation

   (A) I only
   (B) II only
   (C) III only
   (D) I and II only
   (E) I, II, and III

9. Each person in a simple random sample of 2,000 received a survey, and 317 people returned their survey. How could nonresponse cause the results of the survey to be biased?

   (A) Those who did not respond reduced the sample size, and small samples have more bias than large samples.
   (B) Those who did not respond caused a violation of the assumption of independence.
   (C) Those who did not respond were indistinguishable from those who did not receive the survey.
   (D) Those who did not respond represent a stratum, changing the simple random sample into a stratified random sample.
   (E) Those who did respond may differ in some important way from those who did not respond.

10. In a certain game, a fair die is rolled and a player gains 20 points if the die shows a "6." If the die does not show a "6," the player loses 3 points. If the die were to be rolled 100 times, what would be the expected total gain or loss for the player?

    (A) A gain of about 1,700 points
    (B) A gain of about 583 points
    (C) A gain of about 83 points
    (D) A loss of about 250 points
    (E) A loss of about 300 points

11. The Attila Barbell Company makes bars for weight lifting. The weights of the bars are independent and are normally distributed with a mean of 720 ounces (45 pounds) and a standard deviation of 4 ounces. The bars are shipped 10 in a box to the retailers. The weights of the empty boxes are normally distributed with a mean of 320 ounces and a standard deviation of 8 ounces. The weights of the boxes filled with 10 bars are expected to be normally distributed with a mean of 7,520 ounces and a standard deviation of

    (A) $\sqrt{12}$ ounces
    (B) $\sqrt{80}$ ounces
    (C) $\sqrt{224}$ ounces
    (D) 48 ounces
    (E) $\sqrt{1,664}$ ounces

12. Exercise physiologists are investigating the relationship between lean body mass (in kilograms) and the resting metabolic rate (in calories per day) in sedentary males.

Predictor	Coef	StDev	T	P
Constant	264.0	276.9	0.95	0.363
Mass	22.563	6.360	3.55	0.005

S = 144.9 R-Sq = 55.7% R-Sq(adj) = 51.3%

    Based on the computer output above, which of the following is the best interpretation of the value of the slope of the regression line?

    (A) For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 22.563 calories per day.
    (B) For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 264.0 calories per day.
    (C) For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 144.9 calories per day.
    (D) For each additional calorie per day for the resting metabolic rate, the lean body mass increases on average by 22.563 kilograms.
    (E) For each additional calorie per day for the resting metabolic rate, the lean body mass increases on average by 264.0 kilograms.

13. An investigator was studying a territorial species of Central American termites, *Nasutitermes corniger*. Forty-nine termite pairs were randomly selected; both members of each of these pairs were from the same colony. Fifty-five additional termite pairs were randomly selected; the two members in each of these pairs were from different colonies. The pairs were placed in petri dishes and observed to see whether they exhibited aggressive behavior. The results are shown in the table below.

	Aggressive	Nonaggressive	Total
Same colony	40 (33.5)	9 (15.5)	49
Different colonies	31 (37.5)	24 (17.5)	55
Total	71	33	104

A Chi-square test for homogeneity was conducted, resulting in $\chi^2 = 7.638$. The expected counts are shown in parentheses in the table. Which of the following sets of statements follows from these results?

(A) χ^2 is not significant at the 0.05 level.
(B) χ^2 is significant, $0.01 < p < 0.05$; the counts in the table suggest that termite pairs from the same colony are less likely to be aggressive than termite pairs from different colonies.
(C) χ^2 is significant, $0.01 < p < 0.05$; the counts in the table suggest that termite pairs from different colonies are less likely to be aggressive than termite pairs from the same colony.
(D) χ^2 is significant, $p < 0.01$; the counts in the table suggest that termite pairs from the same colony are less likely to be aggressive than termite pairs from different colonies.
(E) χ^2 is significant, $p < 0.01$; the counts in the table suggest that termite pairs from different colonies are less likely to be aggressive than termite pairs from the same colony.

14. Consider n pairs of numbers (x_1, y_1), (x_2, y_2), . . . , and (x_n, y_n). The mean and standard deviation of the x-values are $\bar{x} = 5$ and $s_x = 4$, respectively. The mean and standard deviation of the y-values are $\bar{y} = 10$ and $s_y = 10$, respectively. Of the following, which could be the least squares regression line?

(A) $\hat{y} = -5.0 + 3.0x$
(B) $\hat{y} = 3.0x$
(C) $\hat{y} = 5.0 + 2.5x$
(D) $\hat{y} = 8.5 + 0.3x$
(E) $\hat{y} = 10.0 + 0.4x$

15. The mayor of a large city will run for governor if he believes that more than 30 percent of the voters in the state already support him. He will have a survey firm ask a random sample of n voters whether or not they support him. He will use a large sample test for proportions to test the null hypothesis that the proportion of all voters who support him is 30 percent or less against the alternative that the percentage is higher than 30 percent. Suppose that 35 percent of all voters in the state actually support him. In which of the following situations would the power for this test be highest?

    (A) The mayor uses a significance level of 0.01 and n = 250 voters.
    (B) The mayor uses a significance level of 0.01 and n = 500 voters.
    (C) The mayor uses a significance level of 0.01 and n = 1,000 voters.
    (D) The mayor uses a significance level of 0.05 and n = 500 voters.
    (E) The mayor uses a significance level of 0.05 and n = 1,000 voters.

16. George and Michelle each claimed to have the better recipe for chocolate chip cookies. They decided to conduct a study to determine whose cookies were really better. They each baked a batch of cookies using their own recipe. George asked a random sample of his friends to taste his cookies and to complete a questionnaire on their quality. Michelle asked a random sample of her friends to complete the same questionnaire for her cookies. They then compared the results. Which of the following statements about this study is <u>false</u>?

    (A) Because George and Michelle have a different population of friends, their sampling procedure makes it difficult to compare the recipes.
    (B) Because George and Michelle each used only their own respective recipes, their cooking ability is confounded with the recipe quality.
    (C) Because George and Michelle each used only the ovens in their houses, the recipe quality is confounded with the characteristics of the oven.
    (D) Because George and Michelle used the same questionnaire, their results will generalize to the combined population of their friends.
    (E) Because George and Michelle each baked one batch, there is no replication of the cookie recipes.

17. A large company is considering opening a franchise in St. Louis and wants to estimate the mean household income for the area using a simple random sample of households. Based on information from a pilot study, the company assumes that the standard deviation of household incomes is σ = \$7,200. Of the following, which is the least number of households that should be surveyed to obtain an estimate that is within \$200 of the true mean household income with 95 percent confidence?

    (A) 75
    (B) 1,300
    (C) 5,200
    (D) 5,500
    (E) 7,700

18. Courtney has constructed a cricket out of paper and rubber bands. According to the instructions for making the cricket, when it jumps it will land on its feet half of the time and on its back the other half of the time. In the first 50 jumps, Courtney's cricket landed on its feet 35 times. In the next 10 jumps, it landed on its feet only twice. Based on this experience, Courtney can conclude that

    (A) the cricket was due to land on its feet less than half the time during the final 10 jumps, since it had landed too often on its feet during the first 50 jumps
    (B) a confidence interval for estimating the cricket's true probability of landing on its feet is wider after the final 10 jumps than it was before the final 10 jumps
    (C) a confidence interval for estimating the cricket's true probability of landing on its feet after the final 10 jumps is exactly the same as it was before the final 10 jumps
    (D) a confidence interval for estimating the cricket's true probability of landing on its feet is more narrow after the final 10 jumps than it was before the final 10 jumps
    (E) a confidence interval for estimating the cricket's true probability of landing on its feet based on the initial 50 jumps does not include 0.2, so there must be a defect in the cricket's construction to account for the poor showing in the final 10 jumps

Answers to Multiple-Choice Questions

1 – A	4 – B	7 – A	10 – C	13 – E	16 – D
2 – C	5 – B	8 – E	11 – C	14 – D	17 – C
3 – B	6 – A	9 – E	12 – A	15 – E	18 – D

Free-Response Questions

In the free-response section of the AP Statistics Exam, students are asked to answer five questions and complete an investigative task. Each question is designed to be answered in approximately 12 minutes. The longer investigative task is designed to be answered in approximately 30 minutes.

Statistics is a discipline in which clear and complete communication is an essential skill. The free-response questions on the AP Statistics Exam require students to use their analytical, organizational and communication skills to formulate cogent answers and provide students with an opportunity to:

- Relate two or more different content areas (i.e., exploratory data analysis, experimental design and sampling, probability, and statistical inference) as they formulate a complete response or solution to a statistics or probability problem.
- Demonstrate their mastery of statistics in a response format that permits the students to determine *how* they will organize and present each response.

The purpose of the investigative task is not only to evaluate the student's understanding in several content areas but also to assess his or her ability to integrate statistical ideas and apply them in a new context or in a nonroutine way.

Scoring of Free-Response Questions

The evaluation of student responses on the free-response section of the AP Statistics Exam reflects the dual importance of statistical knowledge and good communication. The free-response questions and the investigative task are scored "holistically"; that is, each question's response is evaluated as "a complete package." With holistic scoring, after reading through the details of a student's response, the scorer makes a judgment about the *overall quality* of the response. This is different from "analytic" scoring, where the individual components to be evaluated in a student's response are specified in advance, and each component is given a value counting toward the overall score.

Holistic scoring is well suited for questions where the student is required to synthesize information and respond at least partially in written paragraphs, and for questions that could potentially generate multiple and diverse, but equally correct, responses. For example, an open-ended question may present data from a real-life study and ask the student not only to analyze the data but also to comment on how the study's protocol might be improved. Comments on improving the protocol might focus on improving the sampling method, controlling confounding variables, or seeking more power by increasing the sample size. In this context, holistic scoring represents a recognition not only of the existence of multiple reasonable approaches to a statistical analysis, but a realization of the existence of a statistical synergy — i.e., that a quality student response is more than just the sum of its parts.

The AP Statistics scoring guideline (rubric) for each free-response question has five categories, numerically scored on a 0 to 4 scale. Each of these categories represents a level of quality in the student response. These levels of quality are defined on two dimensions: statistical knowledge and communication. The specific rubrics for each question are tied to a general template, which represents the descriptions of the quality levels as envisioned by the Development Committee. This general template is given in the following table, "A Guide to Scoring Free-Response Statistics Questions."

A GUIDE TO SCORING FREE-RESPONSE STATISTICS QUESTIONS: THE CATEGORY DESCRIPTORS

Score Descriptors	Statistical Knowledge	Communication
	Identification of the important components of the problem Demonstration of the statistical concepts and techniques that result in a correct solution of the problem	Explanation of what was done and why, along with a statement of conclusions drawn in context
4 Complete	• shows complete understanding of the problem's statistical components • synthesizes a correct relationship among these components, perhaps with novelty and creativity • uses appropriate and correctly executed statistical techniques • may have minor arithmetic errors but answers are still reasonable	• provides a clear, organized, and complete explanation, using correct terminology, of what was done and why • states appropriate assumptions and caveats • uses diagrams or plots when appropriate to aid in describing the solution • states an appropriate and complete conclusion in context

3 Substantial	• shows substantial understanding of the problem's statistical components • synthesizes a relationship among these components, perhaps with minor gaps • uses appropriate statistical techniques • may have arithmetic errors but answers are still reasonable	• provides a clear but not perfectly organized explanation, using correct terminology, of what was done and why, but explanation may be slightly incomplete • may miss necessary assumptions or caveats • uses diagrams or plots when appropriate to aid in describing the solution • states a conclusion that follows from the analysis but may be somewhat incomplete
2 Developing	• shows some understanding of the problem's statistical components • shows little in the way of a relationship among these components • uses some appropriate statistical techniques but misses or misuses others • may have arithmetic errors that result in unreasonable answers	• provides some explanation of what was done, but explanation may be vague and difficult to interpret and terminology may be somewhat inappropriate • uses diagrams in an incomplete or ineffective way, or diagrams may be missing • states a conclusion that is incomplete
1 Minimal	• shows limited understanding of the problem's statistical components by failing to identify important components • shows little ability to organize a solution and may use irrelevant information • misuses or fails to use appropriate statistical techniques • has arithmetic errors that result in unreasonable answers	• provides minimal or unclear explanation of what was done or why it was done, and explanation may not match the presented solution • fails to use diagrams or plots, or uses them incorrectly • states an incorrect conclusion or fails to state a conclusion
0	• shows little to no understanding of statistical components	• provides no explanation of a legitimate strategy

Some important points that students should remember when answering free-response questions on the AP Statistics Exam are given below.

1. Read the questions carefully and answer them in context; for example, the results of a hypothesis test should always be followed by a conclusion in context, and a confidence interval should always be followed by an interpretation of the interval in context. Explanations and conclusions in context are always required for a complete answer.
2. Know the vocabulary of statistics, and use that vocabulary correctly in all written responses.
3. Remember to define all symbols. Specifically, remember to distinguish between population parameters and sample statistics.
4. Remember to state and check all necessary assumptions when performing hypothesis tests and constructing interval estimates.
5. Be able to interpret data displayed in a variety of ways, including graphs and computer outputs. Be able to represent data in a variety of forms and base sound statistical arguments on these representations.

AP Central contains free-response questions, scoring guidelines and selected student responses from past AP Statistics Exams. This is an excellent place to become more familiar with the content of past free-response questions and how they were scored.

The following questions are examples of free-response questions. These questions were administered as part of a previous year's exam.

1. The summary statistics for the number of inches of rainfall in Los Angeles for 117 years, beginning in 1877, are shown below.

N	MEAN	MEDIAN	TRMEAN	STDEV	SE MEAN
117	14.941	13.070	14.416	6.747	0.624

MIN	MAX	Q1	Q3
4.850	38.180	9.680	19.250

   (a) Describe a procedure that uses these summary statistics to determine whether there are outliers.
   (b) Are there outliers in these data?
   Justify your answer based on the procedure that you described in part (a).
   (c) The news media reported that in a particular year, there were only 10 inches of rainfall. Use the information provided to comment on this reported statement.

2. A department supervisor is considering purchasing one of two comparable photocopy machines, *A* or *B*. Machine *A* costs \$10,000 and machine *B* costs \$10,500. This department replaces photocopy machines every three years. The repair contract for machine *A* costs \$50 per month and covers an unlimited number of repairs. The repair contract for machine *B* costs \$200 per repair. Based on past performance, the distribution of the number of repairs needed over any one-year period for machine *B* is shown below.

Number of Repairs	0	1	2	3
Probability	0.50	0.25	0.15	0.10

You are asked to give a recommendation based on overall cost as to which machine, *A* or *B*, along with its repair contract, should be purchased. What would your recommendation be? Give a statistical justification to support your recommendation.

3. Every Monday a local radio station gives coupons away to 50 people who correctly answer a question about a news fact from the previous day's newspaper. The coupons given away are numbered from 1 to 50, with the first person receiving coupon 1, the second person receiving coupon 2, and so on, until all 50 coupons are given away. On the following Saturday, the radio station randomly draws numbers from 1 to 50 and awards cash prizes to the holders of the coupons with these numbers. Numbers continue to be drawn without replacement until the total amount awarded first equals or exceeds \$300. If selected, coupons 1 through 5 each have a cash value of \$200, coupons 6 through 20 each have a cash value of \$100, and coupons 21 through 50 each have a cash value of \$50.

   (a) Explain how you would conduct a simulation using the random number table provided below to estimate the distribution of the number of prize winners each week.

   (b) Perform your simulation three times. (That is, run three trials of your simulation.) Start at the leftmost digit in the first row of the table and move across. Make your procedure clear so that someone can follow what you did. You must do this by marking directly on or above the table. Report the number of winners in each of your three trials.

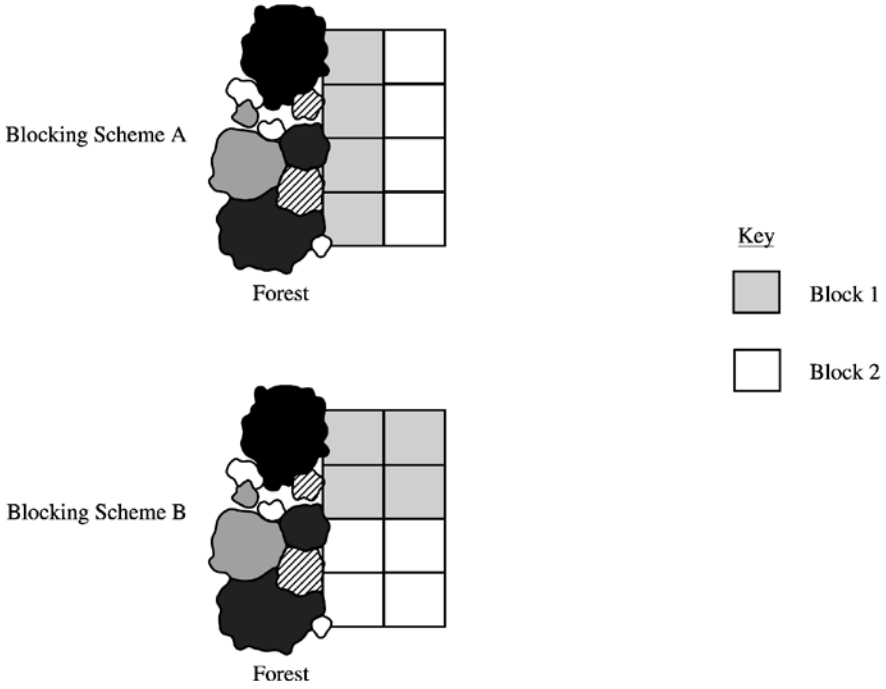
72749 13347 65030 26128 49067 02904 49953 74674 94617 13317

81638 36566 42709 33717 59943 12027 46547 61303 46699 76423

38449 46438 91579 01907 72146 05764 22400 94490 49833 09258

4. Students are designing an experiment to compare the productivity of two varieties of dwarf fruit trees. The site for the experiment is a field that is bordered by a densely forested area on the west (left) side. The field has been divided into eight plots of approximately the same area. The students have decided that the test plots should be blocked. Four trees, two of each of the two varieties, will be assigned at random to the four plots within each block, with one tree planted in each plot.

   The two blocking schemes shown below are under consideration.
   For each scheme, one block is indicated by the white region and the other block is indicated by the gray region in the figures.



   (a) Which of the blocking schemes, A or B, is better for this experiment? Explain your answer.
   (b) Even though the students have decided to block, they must randomly assign the varieties of trees to the plots within each block. What is the purpose of this randomization in the context of this experiment?

5. A growing number of employers are trying to hold down the costs that they pay for medical insurance for their employees. As part of this effort, many medical insurance companies are now requiring clients to use generic brand medicines when filling prescriptions. An independent consumer advocacy group wanted to determine if there was a difference, in milligrams, in the amount of active ingredient between a certain "name" brand drug and its generic counterpart. Pharmacies may store drugs under different conditions. Therefore, the consumer group randomly selected ten different pharmacies in a large city and filled two prescriptions at each of these pharmacies, one for the "name" brand and the other for the generic brand of the drug. The consumer group's laboratory then tested a randomly selected pill from each prescription to determine the amount of active ingredient in the pill. The results are given in the following table.

ACTIVE INGREDIENT
(in milligrams)

Pharmacy	1	2	3	4	5	6	7	8	9	10
Name brand	245	244	240	250	243	246	246	246	247	250
Generic brand	246	240	235	237	243	239	241	238	238	234

Based on these results, what should the consumer group's laboratory report about the difference in the active ingredient in the two brands of pills? Give appropriate statistical evidence to support your response.

6. The statistics department at a large university is trying to determine if it is possible to predict whether an applicant will successfully complete the Ph.D. program or will leave before completing the program. The department is considering whether GPA (grade point average) in undergraduate statistics and mathematics courses (a measure of performance) and mean number of credit hours per semester (a measure of workload) would be helpful measures. To gather data, a random sample of 20 entering students from the past 5 years is taken. The data are given below.

Successfully Completed Ph.D. Program

Student	A	B	C	D	E	F	G	H	I	J	K	L	M
GPA	3.8	3.5	4.0	3.9	2.9	3.5	3.5	4.0	3.9	3.0	3.4	3.7	3.6
Credit hours	12.7	13.1	12.5	13.0	15.0	14.7	14.5	12.0	13.1	15.3	14.6	12.5	14.0

Did Not Complete Ph.D. Program

Student	N	O	P	Q	R	S	T
GPA	3.6	2.9	3.1	3.5	3.9	3.6	3.3
Credit hours	11.1	14.5	14.0	10.9	11.5	12.1	12.0

The regression output below resulted from fitting a line to the data in each group. The residual plots (not shown) indicated no unusual patterns, and the assumptions necessary for inference were judged to be reasonable.

Successfully Completed Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	23.514	1.684	13.95	0.000
GPA	–2.7555	0.4668	–5.90	0.000
S = 0.5658	R-Sq = 76.0%			

Did Not Complete Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	24.200	3.474	6.97	0.001
GPA	–3.485	1.013	–3.44	0.018
S = 0.8408	R-Sq = 70.3%			

(a) Use an appropriate graphical display to compare the GPAs for the two groups. Write a few sentences commenting on your display.
(b) For the students who successfully completed the Ph.D. program, is there a significant relationship between GPA and mean number of credit hours per semester? Give a statistical justification to support your response.
(c) If a new applicant has a GPA of 3.5 and a mean number of credit hours per semester of 14.0, do you think this applicant will successfully complete the Ph.D. program? Give a statistical justification to support your response.

Teacher Support

AP Central® (apcentral.collegeboard.com)

You can find the following Web resources at AP Central:

- AP Course Descriptions, information about the AP Course Audit, AP Exam questions and scoring guidelines, sample syllabi and feature articles.
- A searchable Institutes and Workshops database, providing information about professional development events.
- The Course Home Pages (apcentral.collegeboard.com/coursehomepages), which contain articles, teaching tips, activities, lab ideas and other course-specific content contributed by colleagues in the AP community.
- Moderated electronic discussion groups (EDGs) for each AP course, provided to facilitate the exchange of ideas and practices.

Additional Resources

Teacher's Guides and **Course Descriptions** may be downloaded free of charge from AP Central; printed copies may be purchased through the College Board Store (store.collegeboard.com).

Course Audit Resources. For those looking for information on developing syllabi, the AP Course Audit website offers a host of valuable resources. Each subject has a syllabus development guide that includes the guidelines reviewers use to evaluate syllabi as well as multiple samples of evidence for each requirement. Four sample syllabi written by AP teachers and college faculty who teach the equivalent course at colleges and universities are also available. Along with a syllabus self-evaluation checklist and an example textbook list, a set of curricular/resource requirements is provided for each course that outlines the expectations that college faculty nationwide have established for college-level courses. Visit www.collegeboard.com/apcourseaudit for more information and to download these free resources.

Released Exams. Periodically the AP Program releases a complete copy of each exam. In addition to providing the multiple-choice questions and answers, the publication describes the process of scoring the free-response questions and includes examples of students' actual responses, the scoring standards and commentary that explains why the responses received the scores they did. Released Exams are available at the College Board Store (store.collegeboard.com).

Additional, **free AP resources** are available to help students, parents, AP Coordinators and high school and college faculty learn more about the AP Program and its courses and exams. Visit www.collegeboard.com/apfreepubs for details.

Contact Us

AP Services
P.O. Box 6671
Princeton, NJ 08541-6671
609-771-7300
888-225-5427 (toll free in the U.S. and Canada)
610-290-8979 (Fax)
E-mail: apexams@info.collegeboard.org

National Office
45 Columbus Avenue
New York, NY 10023-6992
212-713-8000

AP Canada Office
2950 Douglas Street, Suite 550
Victoria, BC, Canada V8T 4N4
250-472-8561
800-667-4548 (toll free in Canada only)
E-mail: gewonus@ap.ca

International Services
Serving all countries outside the U.S. and Canada

45 Columbus Avenue
New York, NY 10023-6992
212-373-8738
E-mail: international@collegeboard.org

Middle States Regional Office
Serving Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania, Puerto Rico and the U.S. Virgin Islands

Three Bala Plaza East
Suite 501
Bala Cynwyd, PA 19004-1501
610-227-2550
866-392-3019
610-227-2580 (Fax)
E-mail: msro@info.collegeboard.org

Midwestern Regional Office
Serving Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, West Virginia and Wisconsin

6111 N. River Road, Suite 550
Rosemont, IL 60018-5158
866-392-4086
847-653-4528 (Fax)
E-mail: mro@info.collegeboard.org

New England Regional Office
Serving Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont

1601 Trapelo Road, Suite 12
Waltham, MA 02451-1982
866-392-4089
781-663-2743 (Fax)
E-mail: nero@info.collegeboard.org

Southern Regional Office
Serving Alabama, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee and Virginia

3700 Crestwood Parkway NW, Suite 700
Duluth, GA 30096-7155
866-392-4088
770-225-4062 (Fax)
E-mail: sro@info.collegeboard.org

Southwestern Regional Office
Serving Arkansas, New Mexico, Oklahoma and Texas

4330 Gaines Ranch Loop, Suite 200
Austin, TX 78735-6735
866-392-3017
512-721-1841 (Fax)
E-mail: swro@info.collegeboard.org

Western Regional Office
Serving Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, Oregon, Utah, Washington and Wyoming

2099 Gateway Place, Suite 550
San Jose, CA 95110-1051
866-392-4078
408-367-1459 (Fax)
E-mail: wro@info.collegeboard.org

apcentral.collegeboard.com

I.N. 090083307