Appendix A
About the AP® Exam

Chapter 1

• If you learn to distinguish categorical from quantitative variables now, it will pay big rewards later. You will be expected to analyze categorical and quantitative variables correctly on the AP® exam.

• When comparing distributions of quantitative data, it's not enough just to list values for the center and spread of each distribution. You have to explicitly *compare* these values, using words like "greater than," "less than," or "about the same as."

• If you're asked to make a graph on a free-response question, be sure to label and scale your axes. Unless your calculator shows labels and scaling, don't just transfer a calculator screen shot to your paper.

• You may be asked to determine whether a quantitative data set has any outliers. Be prepared to state and use the rule for identifying outliers.

• Use statistical terms carefully and correctly on the AP® exam. Don't say "mean" if you really mean "median." Range is a single number; so are Q_1, Q_3, and IQR. Avoid colloquial use of language such as "the outlier *skews* the mean." Skewed is a shape. If you misuse a term, expect to lose some credit.

Chapter 2

• Normal probability plots are not included on the AP® Statistics topic outline. However, these graphs are very useful for assessing Normality. You may use them on the AP® exam if you wish—just be sure that you know what you're looking for (a linear pattern).

Chapter 3

• If you are asked to make a scatterplot for a free-response question, be sure to label and scale both axes. Don't just copy an unlabeled calculator graph directly onto your paper.

• If you're asked to interpret a correlation, start by looking at a scatterplot of the data. Then be sure to address direction, form, strength, and outliers (sound familiar?) and put your answer in context.

• When displaying the equation of a least-squares regression line, the calculator will report the slope and intercept with much more precision than is needed. However, there is no firm rule for how many decimal places to show for answers on the AP® exam. Our advice: Decide how much to round based on the context of the problem you are working on.

• Students often have a hard time interpreting the value of r^2 on AP® exam questions. They frequently leave out key words in the definition. Our advice: Treat this as a fill-in-the-blank exercise. Write "________% of the variation in [response variable name] is accounted for by the linear model relating [response variable name] to [explanatory variable name]."

• The formula sheet for the AP® exam uses different notation for these equations: $b_1 = r\frac{s_y}{s_x}$ and $b_0 = \bar{y} - b_1\bar{x}$. That's because the least-squares line is written as $\hat{y} = b_0 + b_1x$. We prefer our simpler versions without the subscripts!

Chapter 4

• If you're asked to describe how the design of a study leads to bias, you're expected to do two things: (1) identify a problem with the design, and (2) explain how this problem would lead to an underestimate or overestimate. Suppose you were asked, "Explain how using your statistics class as a sample to estimate the proportion of all high school students who own a graphing calculator could result in bias." You might respond, "This is a convenience sample. It would probably include a much higher proportion of students with a graphing calculator than would the population at large because a graphing calculator is required for the statistics class. So this method would probably lead to an overestimate of the actual population proportion."

• If you are asked to identify a possible confounding variable in a given setting, you are expected to explain how the variable you choose (1) is associated with the explanatory variable and (2) affects the response variable.

• If you are asked to describe the design of an experiment on the AP® exam, you won't get full credit for a diagram like Figure 4.5 (page 246). You are expected to describe how the treatments are assigned to the experimental units and to clearly state what will be measured or compared. Some students prefer to start with a diagram and then add a few sentences. Others choose to skip the diagram and put their entire response in narrative form.

• Don't mix the language of experiments and the language of sample surveys or other observational studies. You will lose credit for saying things like "Use a randomized block design to select the sample for this survey" or "This experiment suffers from nonresponse since some subjects dropped out during the study."

Chapter 5

• On the AP® exam, you may be asked to describe how you will perform a simulation using rows of random digits. If so, provide a clear enough description of your simulation process for the reader to get the same results you did from *only* your written explanation.

• Many probability problems involve simple computations that you can do on your calculator. It may be tempting to write down just your final answer without showing the supporting work. Don't do it! A "naked answer," even if it's correct, will usually earn you no credit on a free-response question.

• You can write statements like $P(B \mid A)$ if events A and B are defined clearly, or you can use a verbal equivalent, such as P(reads *New York Times* | reads *USA Today*). Use the approach that makes the most sense to you.

Chapter 6

• If the mean of a random variable has a noninteger value but you report it as an integer, your answer will not get full credit.

• When showing your work on a free-response question, you must include more than a calculator command. Writing `normalcdf(68,70,64,2.7)` will *not* earn you full credit for a Normal calculation. At a minimum, you must indicate what each of those calculator inputs represents. Better yet, sketch and label a Normal curve to show what you're finding.

• Don't rely on "calculator speak" when showing your work on free-response questions. Writing `binompdf(5,0.25,3)` = 0.08789 will not earn you full credit for a binomial probability calculation. At the very least, you must indicate what each calculator input represents. For example, "I used `binompdf(trials 5,p:0.25, x value:3)`."

Chapter 7

• Terminology matters. Don't say "sample distribution" when you mean *sampling distribution*. You will lose credit on free-response questions for misusing statistical terms.

• Notation matters. The symbols $\hat{p}$, $\bar{x}$, p, μ, σ, $\mu_{\hat{p}}$, $\sigma_{\hat{p}}$, $\mu_{\bar{x}}$, and $\sigma_{\bar{x}}$ all have specific and different meanings. Either use notation correctly—or don't use it at all. You can expect to lose credit if you use incorrect notation.

Chapter 8

• On a given problem, you may be asked to interpret the confidence interval, the confidence level, or both. Be sure you understand the difference: the confidence interval gives a set of plausible values for the parameter and the confidence level describes the long-run capture rate of the method.

• If a free-response question asks you to construct and interpret a confidence interval, you are expected to do the entire four-step process. That includes clearly defining the parameter, identifying the procedure, and checking conditions.

• You may use your calculator to compute a confidence interval on the AP® exam. But there's a risk involved. If you give just the calculator answer with no work, you'll get either full credit for the "Do" step (if the interval is correct) or no credit (if it's wrong). We recommend showing the calculation with the appropriate formula and then checking with your calculator. If you opt for the calculator-only method, be sure to name the procedure (e.g., one-proportion z interval) and to give the interval (e.g., 0.514 to 0.607).

• If a question of the AP® exam asks you to calculate a confidence interval, all the conditions should be met. However, you are still required to state the conditions and show evidence that they are met.

• It is not enough just to make a graph of the data on your calculator when assessing Normality. You must *sketch* the graph on your paper to receive credit. You don't have to draw multiple graphs—any appropriate graph will do.

Chapter 9

• The conclusion to a significance test should always include three components: (1) an explicit comparison of the P-value to a stated significance level, (2) a decision about the null hypothesis: reject or fail to reject H_0, and (3) a statement in the context of the problem about whether or not there is convincing evidence for H_a.

• When a significance test leads to a fail to reject H_0 decision, be sure to interpret the results as "We don't have enough evidence to conclude H_a." Saying anything that sounds like you believe H_0 is (or might be) true will lead to a loss of credit. And don't write text-message-type responses, like "FTR the H_0."

• You can use your calculator to carry out the mechanics of a significance test on the AP® exam. But there's a risk involved. If you give just the calculator answer with no work, and one or more of your values are incorrect, you will probably get no credit for the "Do" step. We recommend doing the calculation with the appropriate formula and then checking with your calculator. If you opt for the calculator-only method, be sure to name the procedure (one-proportion z test) and to report the test statistic ($z = 1.15$) and P-value (0.1243).

• It is not enough just to make a graph of the data on your calculator when assessing Normality. You must *sketch* the graph on your paper to receive credit. You don't have to draw multiple graphs—any appropriate graph will do.

• Remember: If you give just calculator results with no work and one or more values are wrong, you probably won't get any credit for the "Do" step. If you opt for the calculator-only method, name the procedure (t test) and report the test statistic ($t = -0.94$), degrees of freedom (df = 14), and P-value (0.1809).

Chapter 10

• The formula for the two-sample z interval for $p_1 - p_2$ often leads to calculation errors by students. As a result, we recommend using the calculator's `2-PropZInt` feature to compute the confidence interval on the AP® exam. Be sure to name the procedure (two-proportion z interval) and to give the interval (0.076, 0.143) as part of the "Do" step.

• The formula for the two-sample z statistic for a test about $p_1 - p_2$ often leads to calculation errors by students. As a result, we recommend using the calculator's `2-PropZTest` feature to perform calculations on the AP® exam. Be sure to name the procedure (two-proportion z test) and to report the test statistic ($z = 1.17$) and P-value (0.2427) as part of the "Do" step.

• The formula for the two-sample t interval for $\mu_1 - \mu_2$ often leads to calculation errors by students. As a result, we recommend using the calculator's `2-SampTInt` feature to compute the confidence interval on the AP® exam. Be sure to name the procedure (two-sample t interval) and to give the interval (3.9362, 17.724) and df (55.728) as part of the "Do" step.

• When checking the Normal condition on an AP® exam question involving inference about means, be sure to include graphs. Don't expect to receive credit for describing a graph that you made on your calculator but didn't put on paper.

• The formula for the two-sample t statistic for $\mu_1 - \mu_2$ often leads to calculation errors by students. As a result, we recommend using the calculator's `2-SampTTest` feature to perform calculations on the AP® exam. Be sure to name the procedure (two-sample t test) and to report the test statistic ($t = 1.60$), P-value (0.0644), and df (15.59) as part of the "Do" step.

Chapter 11

• You can use your calculator to carry out the mechanics of a significance test on the AP® exam. But there's a risk involved. If you give just the calculator answer with no work, and one or more of your values is incorrect, you will probably get no credit for the "Do" step. We recommend writing out the first few terms of the chi-square calculation followed by ". . .". This approach might help you earn partial credit if you enter a number incorrectly. Be sure to name the procedure (χ^2 GOF-Test) and to report the test statistic ($\chi^2 = 11.2$), degrees of freedom (df = 3), and P-value (0.011).

• In the "Do" step, you aren't required to show every term in the chi-square statistic. Writing the first few terms of the sum followed by ". . ." is considered as "showing work." We suggest that you do this and then let your calculator tackle the computations.

• You can use your calculator to carry out the mechanics of a significance test on the AP® exam. But there's a risk involved. If you give just the calculator answer with no work and one or more of your values is incorrect, you will probably get no credit for the "Do" step. We recommend writing out the first few terms of the chi-square calculation followed by ". . .". This approach might help you earn partial credit if you enter a number incorrectly. Be sure to name the procedure (χ^2-Test for homogeneity) and to report the test statistic ($\chi^2 = 18.279$), degrees of freedom (df = 4), and P-value (0.0011).

• If you have trouble distinguishing the two types of chi-square tests for two-way tables, you're better off just saying "chi-square test" than choosing the wrong type. Better yet, learn to tell the difference!

Chapter 12

• The AP® exam formula sheet gives $\hat{y} = b_0 + b_1x$ for the equation of the sample (estimated) regression line. We will stick with our simpler notation, $\hat{y} = a + bx$, which is also used by TI calculators. Just remember: The coefficient of x is always the slope, no matter what symbol is used.

• The AP® exam formula sheet gives the formula for the standard error of the slope as

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

The numerator is just a fancy way of writing the standard deviation of the residuals s. Can you show that the denominator of this formula is the same as ours?

• The formula for the t interval for the slope of a population (true) regression line often leads to calculation errors by students. As a result, we recommend using the calculator's `LinRegTInt` feature to compute the confidence interval on the AP® exam. Be sure to name the procedure (t interval for slope) and to give the interval (−0.217, −0.108) and df (14) as part of the "Do" step.

• When you see a list of data values on an exam question, don't just start typing the data into your calculator. Read the question first. Often, additional information is provided that makes it unnecessary for you to enter the data at all. This can save you valuable time on the AP® exam.

• The formula for the test statistic in a t test for the slope of a population (true) regression line often leads to calculation errors by students. As a result, we recommend using the calculator's `LinRegTTest` feature to perform calculations on the AP® exam. Be sure to name the procedure (t test for slope) and to report the test statistic ($t = 3.065$), P-value (0.002), and df (36) as part of the "Do" step.