AP Statistics Exam Common Errors

Many students spend lots of time constructing graphs only to forget the labels. This will always cause a deduction in the students score. It is very important for students to make sure their graph displays the data with the proper labels and scaling.

When students are asked to describe the distribution on the AP exam, many students only comment on the shape. To get full credit, they need to address shape, center, spread, and any outliers. Emphasizing the SOCS and acronym will help.

When students are asked to compare distributions, there are two very common errors that students make. First, they forget to address all four characteristics of the distribution (think SOCS). Second, they don't explicitly compare each characteristic. Discussing the SOCS for each distribution separately will not receive credit. Students must use phrases like "about the same as" or "is much greater than" to receive full credit.

When making a stem plot, a common error is to forget the key or labels. Make sure you don't forget these features or you will lose points on the AP exam.

When making a stem plot, common error is to forget the key or labels resulting in a deduction of points on the AP exam.

When making graphs, students often forget to label their axes. Students will lose points on the AP exam.

Many students refer to the box in a boxplot as the IQR. For example, they will say things like "the median is in the middle of the IQR". This is incorrect since the IQR is a number (the length of the box), not the box itself.

When describing the shape of a box plot, many students will describe a symmetric box plot as approximately normal. This is incorrect since a boxplot does not reveal where the modes of a distribution are.

Students sometimes try to apply the $68 – 95 – 99.7$ Rule to non-normal distributions. This rule works reasonably well only for distributions that are approximately normal

When doing normal distribution calculations, students often report the area to the left of a boundary when they are asked about the area to the right of the boundary it. Insisting that your students do a quick sketch will prevent many of these mistakes.

Students often do not get full credit on the AP exam because they use "calculator-speak" to show their work on normal calculation questions. This is not considered clear communication because many AP readers do not use the same type of calculators and may not understand the syntax for each calculator to get full credit. Students should include a sketch of a normal curve with the mean and standard deviation clearly identified or, at a minimum carefully label each of the inputs in the calculator command.

Many students lose points because they forget to label their graphs. Including the proper labels is more important than graphing each point in precisely the right place.

On the AP exam, if students are asked to describe a scatterplot, they are expected to discuss the direction, form, and strength of the association in context as well as identify any possible outliers. To help students remember this, consider and acronym such as the DOFS (direction, outliers, form, strength).

When you are asked to interpret correlation, be careful when addressing the form. While a strong positive linear relationship will have a correlation near 1, a correlation near 1 doesn't mean that the relationship is linear. It is easy to create curved data with the correlation close to 1. For example, graphing the function $y = x^2$ for the integers 1 to 10 gives a correlation equal to 0.97, even though the data are clearly not linear.

When writing the equation of a regression line, the students often forget to include the hat on the y. Make sure the students understand the difference between the actual value of y (denoted y) and the predicted value of y (denoted ŷ). This also applies to defining the variables used in a regression equation. In this case, y = the actual number and ŷ (y-hat) equals the predicted value.

Many students lose credit on the AP exam for not stating that the slope is the predicted (or estimated or expected or average) change in the y - variable for increased for each increase of one unit in the x-variable. Students will have to understand the difference between the actual data and the equation we are using to model the data. For example, a response that says, "The price will go down by $0.1629 for each additional mile driven" mistakenly implies a deterministic relationship where all the data values will be on the regression line.

When asked if a linear model is appropriate, students will sometimes incorrectly use the correlation to justify linearity. However, a strong correlation doesn't mean the association is linear. Associations can be clearly nonlinear and still have a correlation close to plus or minus 1. Only a residual plot can adequately address whether a line is an appropriate model for the data by showing the pattern of deviations from the line.

When working with a table of random digits, it is very important that each label have the same number of digits. For example, if you need 50 labels, use $01 - 50$ rather than $1 - 50$. The single-digit "1" has a 10% chance of occurring but the two digit combination "01" has a 1% chance of occurring.

When describing how to select a sample using a random integer generator, many students forget to address what to do with repeated integers. Students must explicitly state that repeated integers should be ignored or say that they will generate random integers until they get "n" numbers in the specified range.

On the AP exam, many students lose credit when describing what can go wrong in sample surveys because they use the wrong terminology. While it is important for students to understand and use the vocabulary of statistics correctly, they are rarely required to use specific vocabulary in their answers on the AP exam. To be safe, do not worry about naming the error. Instead

clearly described the error in the context of the question. The explanation is especially important since just correctly naming the error will also result in no credit as well.

On the AP exam, many students lose credit when describing what can go wrong in sample surveys because they use incorrect terminology. While it is important for students to understand and use the vocabulary of statistics correctly, they are rarely required to use specific vocabulary in their responses on the AP exam. To be safe, do not worry about naming a specific problem or source of bias. Instead, clearly describe the problem and its consequences in the context of the question.

Many students lose credit on the AP exam for using statistical vocabulary incorrectly. Instead of relying on vocabulary, which by itself usually won't earn credit, students should explain the concept in the context of the problem. For example, just saying that, "wealth is a confounding variable" will probably not earn credit without additional explanation. Students need to explain not only that wealthier people tend to go to the doctor more often and to be healthier overall, but that wealthier people are also more likely to get hormone replacement. Thus, it could actually be the overall healthiness of the wealthy women cause and the reduction in heart attacks, not the hormones they were taking. In fact, any variable that is more common in the group of women taking replacement hormones could be the cause of the reduction in heart attacks.

Many students lose credit on the AP exam for failing to adequately describe how they assign the treatments to experimental units in an experiment. Most importantly, the method the students use must be *random*. In addition, the method must be described in *sufficient detail* so that two knowledgeable users of statistics could follow the students description and carry out the method in exactly the same way. For example, saying "Assign students to the two groups using random digits" isn't sufficiently detailed since there are many ways to use random digits. If a student chooses to use random digits, he or she must use labels of the same length (for example $01 - 30$, not $1 - 30$) and ignore repeated numbers, since each subject can be assigned to a group only once.

A very good randomization method is the "hat" method. It is easy to describe and helps students avoid making the mistakes that are possible when using random digits or coin flips. Just make sure that the slips of paper are the same size and that they are well mixed.

Many students confuse blocks and treatment groups. Blocks are not formed at random. Instead, they are formed by grouping experimental units that are similar in some way that is expected to systematically affect the response to the treatments. This means that each block should be very different from other blocks. However, a treatment group is formed at random, with the goal that the treatment groups be as similar as possible to each other.

Many students believe that an experiment cannot be double-blind because someone has to know who gets which treatment. There is always someone who knows which subjects are receiving which treatment. As long as this person doesn't interact with the subjects or measure the response, the experiment can still be double-blind.

Confusing stratified random sampling with a randomized block is another way that students mix up the language of sample surveys and experiments. These two ideas are similar in some ways, because they both involve forming groups of similar subjects before random selection or random assignment. Both methods also help to account for the variability created by other variables. However, stratified random sampling is done only when taking a sample from a population. Likewise, blocking happens only when assigning units to treatments in an experiment.

Students often confuse the two types of inferences we can make: inferences about a population and inferences about cause and effect. Over the years there have been many questions on the AP exam that Ask students to discuss what types of inferences (conclusions) are appropriate based on the design of the study and the problem. Make sure that you understand how the role of randomization in a study helps to decide which type of inference, if any, is appropriate.

Students sometimes report probabilities is a number greater than one. In many cases, this is because their calculator is expressing a very small number in scientific notation, such as 2.383223456E -8. This number is actually a very small decimal where you move the decimal eight places to the left, filling in with zeros. Remember that probabilities must be between zero and one.

When working with a random digit table, students often make mistakes or do not clearly communicate method they are using. every label needs to be the same length. For example, if you are labeling 1500 items, the labels should be $0001 - 1500$, not $1 - 1500$. Also, if students are not using all the labels of a certain length, they should explain that the remaining labels will be ignored. Finally, if the process involves sampling without replacement, then students should clearly acknowledge they will ignore repeated labels or look for a certain number of different labels.

When students make conclusions, they often lose credit for suggesting that a claim is definitely true or that the evidence proves that a claim is incorrect. A better response would be to say that there is sufficient evidence (or there isn't sufficient evidence) to support a particular claim.

Although a statistically significant association data from a well-designed experiment does imply causation, many students incorrectly respond that a statistically significant association in data from an observational study implies causation as well. This isn't true. The design of the study is what determines if the cause and effect conclusion is possible.

Many students confuse blocks and treatment groups. Blocks are not formed at random. Instead, they are formed by grouping experimental units that are similar in some way that is expected to systematically affect the response to the treatments. This means that each block should be very different from other blocks. However, a treatment group is formed at random, with the goal that the treatment groups be as similar as possible to each other, other than the treatment.

When students make conclusions, they often lose credit for suggesting that a claim is definitely true or saying that the evidence proves that a claim is incorrect. If the conclusion is based on a probability, students should always knowledge that they can't be certain that their collusion is

correct. A better response would be to say that there is convincing evidence (or there isn't convincing evidence) that seem to support a particular claim.

When using the complement rule, some students use incorrect notation such as $0.25^c = 0.75 \; or \; P(A)^c = 0.75$. Events have complements, not probabilities. Here is a correct statement: $P(A^c) = 1 - 0.25 = 0.75$. Notice that the "C" appears in the superscript of event A, not its probability.

Many students lose credit on probability questions because they do not show their work when using the conditional probability formula. Students should show the numerator and the denominator of the probability, not just decimal answer.

Many students incorrectly believe that the expected value of a random variable must be equal to one of possible values of the variable. This is not the case.

When students are asked to calculate the mean (expected value) or standard deviation of a random variable, they often lose credit for not showing adequate work, or for showing no work at all. Because not all AP readers are familiar with every brand of calculator, students need to go beyond just reporting the commands they use on the calculator. In this case, filling the first couple of terms with an ellipsis, which is… For the calculation of the mean or standard deviation is sufficient.

When students are asked about the effect a linear transformation on summary statistics, some students forget that adding or subtracting a constant to or from every value in the distribution has no effect on measures of spread, including the standard deviation, variance, range, and interquartile range.

On the AP exam, many students lose credit because they add standard deviations instead of adding variances when combining random variables.

On many questions involving binomial settings, students do not recognize that using binomial distribution is appropriate. In fact, free response questions about the binomial distribution are among the lowest scoring questions on the exam. Spend plenty of time learning how to identify a binomial distribution. Check to see it is a binomial setting whenever you are not sure how to answer a probability question.

Many students lose credit are not clearly identified a distribution and the values and interests. To avoid this mistake, follow the three steps listed below:

> State the distribution and the values of interests.
> Perform calculations and shall your work.
> Answer the question.

On the AP exam, many students lose credit because they fail to make the distinction between parameters and statistics. For example, in a significance test, stating hypotheses in terms of statistics instead of parameters will always lose credit. Likewise, when defining a parameter, it is

important that the student use phrases like "population mean" or "true mean" rather than simply saying "mean", which can refer to the mean of a population or the mean of a sample.

On the AP exam, a common error is to write an ambiguous statement such as "the variability decreases when the sample size increases". This statement may be true for the distribution of a sample statistic, but it almost certainly isn't true for the distribution of the sample. Whenever students are describing a distribution, they should always say "the distribution of ____" to avoid ambiguity.

Many students lose credit for using inappropriate symbols or formulas on the AP exam. The best way to avoid losing credit is should know exactly what each of the symbols mean. When in doubt, however, leave them out. Students will still get credit for showing work if they start with numbers substituted into a formula and don't risk using an incorrect symbol.

Students often forget to divide by the square root of the sample size when finding probabilities involving the sample mean.

Students often refer to the central limit theorem when describing how the variability of the sampling distribution decreases as a sample size increases. While this is true, it isn't what the central limit means. The central limit theorem is only about the shape of the sampling distribution of the sample mean, not the spread.

Many students confuse interpreting confidence levels with interpreting confidence intervals. Every time students construct their confidence interval, they are expected to interpret the interval they calculated. They are expected to interpret the confidence level only when they are specifically asked to do so. If students interpret the confidence level when they are not asked, they won't gain any credit and will lose credit if the interpretation is incorrect.

Many students lose credit for referring to the sample when interpreting a confidence interval. For example, the following interpretation is incorrect because it makes a conclusion about the members of the sample rather than the members of the population: We are 95% confident that the interval from 0.167 to 0.213 captures the true proportion of US adults who said that they use Twitter." Replacing the word "said" with "would say" makes the conclusion about the larger population and not just the people who were in the sample.

Does increasing the sample size increase the capture rate? Increasing the sample size should not affect the confidence level. The intervals will be shorter or more precise, but the capture rate should still be the same. It would be incorrect to say "Increasing the sample size makes us more confident." A correct statement would be that, "Increasing the sample size makes our estimate more precise."

Many students use the 1-PropZInt feature will to correctly calculate the confidence interval and then try to "show their work" with an incorrect formula. This will result in a loss of credit because the two attempts are considered as "parallel solutions" and students are scored on the "worst response". If students want to include a formula in their response, they should make sure it produces the same results as a calculator. If the results aren't the same, students should choose

one of the answers and cross out the other. Also, if students choose to include a formula, they should skip the symbolic formula and start with numbers substituted in. In some cases, students choose the right formula and use the correct values, but include an incorrect symbol, for example p instead of $\hat{p}$, and then they lose credit.

Many students incorrectly state the Normal/Large Sample condition by saying that the sample must have a normal distribution. This is not true or even possible. Instead, we use the sample data to make an inference about the shape of a population from which the sample came.

On the AP exam, many students incorrectly use the data to form the hypotheses rather than the wording of the question. In real life, the hypotheses is formed before the data is collected.

On the AP exam, many students incorrectly use sample statistics such as $\bar{x}$ or $\hat{p}$ when stating their hypotheses. We already know the value of the sample statistic. It is the population parameter that we are trying to make an inference about.

Instead of failing to reject the null hypotheses, many students accept the null hypotheses or use equivalent language such as, "We can conclude that teens are getting enough calcium." Accepting the null hypotheses will always lose credit on the AP exam. When we do a significance test, we have some evidence that the null hypotheses is true. If the P-value is large, it simply means that the evidence we have is not convincing enough to rule out change. It does not mean that there is convincing evidence that the null hypotheses is true.

On the AP exam, students often reverse Type I and Type II errors. To avoid this mistake, remember that type II error occurs when we fail to "(II)" reject the null hypothesis when we really should.

Many students lose credit on the AP exam for making statements such as, "We conclude that the true proportion of potatoes with blemishes is 0.08." or "We conclude that the shipment of potatoes is OK.". Both of these statements suggest that there is evidence the null hypotheses is true. Instead, the results of the test only say that there isn't convincing evidence the proportion of blemished potatoes is greater than 0.08.

Students often use a one-sided H_a when H_a should be two-sided. Use the wording of the question to determine the alternative hypotheses, not the data.

Students often mix up the role of random selection and the role of random assignment in a statistical study. Random selection from a population allows us to make an inference about the population based on sample data. Random assignment of subjects to treatments allows us to make cause-and-effect conclusions.

Many students lose credit for not including the word "mean" or "average" in their interpretations of a confidence interval or difference in means or the conclusion to a significance test. The interval is about the difference in means and not the difference in variables.

Make sure students report the degrees of freedom that he used to calculate the P-value. This will help the greater understand which method, conservative or technology, the student is using.

Many students have trouble deciding when it is appropriate to use a two-sample t-test and when it is appropriate to use paired t-test. The key distinction is how the data were produced. In an experiment, groups were formed using a completely randomized design, then the two sample t-test is the right choice. However if subjects were paired and then split at random into the two treatment groups, or each subject received the treatments, the paired t-test is appropriate. Paired data can also arise when the data are produced by random sampling. For example, if wells are selected at random and the quality of water is measured at the top and the bottom of each well we should treat these measurements as paired data.

Some students mistakenly rounded the expected values, believing that they must be integers. Students will lose points for this on the AP exam because it shows a misunderstanding of what expected counts represent… the average number of observations in a given category in many, many random samples

When stating conclusions for chi-square tests, students often accept the null hypotheses. Never accept the null hypotheses. Points will always be deducted for this error. And we never conclude that the hypothesis is correct or true, incorrect or false.

When checking conditions, many students examine the observed counts rather than the expected counts. Also, many students forget to list the expected counts, especially when doing the calculations on their calculators. Simply stating that the expected counts are at least five is not sufficient. Students must list the expected counts to prove that they have actually check them.

On the AP exam, students often use the t-test statistic from the computer output as a critical value. The value of t in the computer output measures how far the sample slope is from zero in terms of standard errors. However the t^* critical value measures how many standard errors we need to go from the mean to have C% of the area under a t-curve with specified degrees of freedom. To help avoid this error, we use an asterisk to denote the critical value t^*.

When students interpret the slope, they often forget to include the word "predicted" in their response.

AP Statistics Exam Tips

If you learn to distinguish categorical from quantitative variables now, it will pay big rewards later. The type of data determines what kind of graphs and which numerical summaries are appropriate. You will be expected to analyze categorical and quantitative data effectively on the AP exam.

When comparing distributions of qualitative data, it's not enough just to list values were the center and spread of each distribution. You have to us 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.

On the AP exam, students will be allowed to use any reasonable definition of quartiles on the free response section. On the multiple-choice section the answer choices will be designed so that any reasonable method of finding quartiles will give the correct answer.

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, like "the outlier skews the mean". Skewed is a shape. If you misuse of term, expect to lose some credit.

On the AP exam, students will be allowed to use any reasonable definition of percentile on the free-response section. On the multiple-choice section, the answer choices will be designed so that any reasonable method of finding percentiles will give the correct answer. When working with large data sets, the different definitions will give very similar values for the percentiles of a distribution.

Don't use "calculator-speak" when showing your work on free-response questions. Writing normalcdf (305, 325, 304, 8) will not earn you credit for a normal calculation. At the very least you must indicate what each of these calculator inputs represents. For example "I use normalcdf on my calculator with lower bound 305, upper bound 325, mean 304, and standard deviation eight." Better yet, sketch and label a normal curve to show what you're finding.

Scatterplots are the only choice for displaying the relationship between two quantitative variables. For a single quantitative variable, there are many choices, including. dotplots, histograms, box plots, and stem plots.

Pick nice values to mark on each axis when doing scatter plots. Make sure to cover the range of each variable. The axis usually does not start at zero and are often not on the same scale. However the scale on each axis must be consistent. Clearly label the variable name on each axis.

When trying to identify outliers in a scatterplot, look for points that fall outside the overall pattern of the association. Unlike with distributions of a single variable, there isn't a specific rule for identifying outliers on a scatter plot.

When discussing the form of the scatterplot relationship, do not let one or two unusual values influence your description. For example, if covering up one value makes the form go from nonlinear to linear, you should call it a linear association with an outlier. This is especially true for small data sets.

Use the acronym, **S**tatistics **P**roblems **D**emand **C**onsistency when using the four- step process **(State, Plan, Do, Conclude).**

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

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

The formula for correlation is on the formula sheet provided on both sections of the AP exam. There is a very small chance that you will actually use this formula on the AP exam.

On the formula sheet provided on both sections of the AP exam, the least-squares regression line is written $\hat{y} = b_0 + b_1 x$. Likewise, formulas for the slope (b_1) and y-intercept (b_0) of the least - squares regression line also use this notation. The coefficient of x is always the slope, no matter what notation. Make sure that you are aware of the difference in notation.

There's no for rule for how many decimal places to show for answers on the AP exam. Our advice: give your answer correct to two or three nonzero decimal places. Exception: if you're using one of the tables in the back of the book, give the value shown in the table.

When you calculate the residuals, interpret what the residual means. Make sure you address both the distance from what we predicted and the direction of the error.

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

Both s and r^2 are calculated using the sum of squared residuals, and both measure how well the **line** fits the data. However, s is expressed in the units of the y variable, whereas r^2 is expressed as a percent between 0% and 100%.

The population is a group of individuals we want to know about, not necessarily the population we can make inferences about.

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

Suppose that the parent of a football player is concerned about proposed budget cuts to the athletic program at school, so she asked spectators at a game if they agree with the proposed cuts. This design is biased since it systematically favors a "No" outcome. That is, if the parent were to conduct a similar survey over and over again, each time the result will probably overestimate the proportion of people who would respond "No".

Another problem with voluntary response samples is that people can respond more than once. Television shows such as *American Idol* and *Dancing with the Stars* are examples of polls that use voluntary response samples.

The word "error" in the phrase "margin of error" does not mean a mistake is made. The margin of error compensates for the variability that results from taking a random sample from a population. It does not account for any mistake made during the sampling or data collection process.

Be sure that you know how to use the table of random digits. When working with a table of random digits, it is very important that each label have the same number of digits. For example, if you need 50 labels, use 01 – 50. The single-digit 1 has a 10% chance of occurring but the two digit combination 01 has the proper 2% chance of occurring. Students often lose credit for failing to address what they will do with repeated label.

Voluntary response is a sampling error since it has to do with choosing the sample. Nonresponse is a non-sampling error since it occurs after the sample has been chosen.

Lurking variables create problems in two different ways. They can create extra variability in the response variable and also have the potential to become confounded with the explanatory variable if not controlled.

Each lurking variable is a potential confounding variable. However, by following the principles of experimental design, you can usually prevent lurking variables from becoming confounding variables.

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.

A good experiment needs comparison to prevent some lurking variables from becoming confounded with the explanatory variable.

An experiment without random assignment is never a good idea. Not only does random assignment create roughly equivalent treatment groups, but it is the basis for the inference methods we use.

If you're asked to describe the design of an experiment on the AP exam, you will not get full credit for just the diagram. You're expected to describe how the treatments are assigned in the experimental units and then 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 the entire response in narrative form.

An inactive treatment is often called a placebo.

Sometimes students get lost in the vocabulary of experiments and lose sight of the big picture. The main goal in experimental design is to create treatment groups with no systematic differences between them, other than the treatment.

Many students think that "control" refers only to a control group. While good experiments will always include comparison, an experimenter should do whatever is possible to control lurking variables by making them the same for all treatment groups.

A good illustration of the placebo effect is when a parent kisses a child's "boo-boo" when the child gets injured. Even though the kiss has no "active treatment", it makes the child feel better.

Be sure that you are able to explain why you chose a particular blocking variable. The best explanation is usually that the blocking variable has a strong association with the response variable. Of course you need to give your explanation in the context of the question.

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 the survey" or "this experiment suffers from nonresponse since some subject struck out during the study."

A matched-pairs design is a special case of a randomized block design that uses blocks of size 2.

When you write the conclusion to a simulation question, emphasize that the answer is always an approximation. For example, instead of saying that a probability is 18/100, it is better to say that the probability is approximately 18/100.

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.

If students express a probability as a fraction, it's okay if they don't reduce it. In fact, it's often easier for students to keep the denominator the same for the probability of each outcome in a particular problem so that they can easily compare the probabilities and verify that they add up to one.

Many probability problems involved simple computations that you can do on your calculator. It may be tempting to just write down 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.

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 credit for 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.

You may use your calculator to compute a confidence interval on the AP exam. But there is a risk involved. If you just give 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. If you opt for the calculator only method, be sure to name the procedure (for example, one proportion Z-interval) and to give the interval (for example, 0.514 to 0.607).

If a free-response question asks you to construct and interpret a confidence interval, you're expected to do the entire four-step process. That includes clearly defining the parameter, identifying the procedure, and checking conditions. The four step process is STATE, PLAN, DO, and CONCLUDE.

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.

When a significance test leads to a fail to reject the H_0 decision, be sure to interpret the results as "We don't have convincing evidence to conclude H_0." 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".

The formula for two-sample z-interval for $p_1 - p_2$ often leads to calculation errors by students. As a result, we recommend using the 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 as part of the "DO" step.

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

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 and to report the test statistic, P value, and degrees of freedom on the part of the "DO" step.