AP STATISTICS EXAM
(Topics and Format)

I. Topics

A. Exploratory Analysis (20% - 30%)

In examining distributions of data, you should be able to detect important characteristics such as shape, location, variability and unusual values. From careful observations of patterns in data, you should 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 accompanies this conceptual development throughout.

B. Planning A Study (10% - 15%)

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 the 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.

C. Probability (20% - 30%)

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 oriented toward using probability distributions to describe data.

D. Statistical Inference (30% - 40%)

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.

II. Format

A. 40-Multiple Choice Questions (50%)

1. 90-minutes
2. ~~Guessing Penalty~~

B. 5-Free Response Questions (37.5%)

1. 60-minutes
2. Scored on a scale of 0 (no understanding) to 4 (complete understanding)

C. 1-Free Response Question (12.5%)

1. 30 minutes
2. Scored on a scale of 0 to 4
3. Designed to assess ability to integrate statistical ideas and apply them in a new context

III. Scoring Guidelines (Free Reponses)

Take a look at the sample scoring rubrics & responses on the AP Central website - note that these are generally 3 & 4 point responses (so they are essentially complete - commentary indicates any problems). What is needed is:

Identification of the test/procedure "by name or formula" (2-sample t, etc.)

Clear statement of parameter and/or hypotheses - for two-sample tests (and estimates especially) the symbols need to be associated with the populations (which is 1? which is 2?). It is usually important that the statement be in terms of the parameter, not the statistic.

Check (not just statement) of **conditions** [a list of the conditions with check marks is *not* considered evidence of a "check of conditions" -reasons why the conditions are satisfied is needed - or statement that a certain condition is assumed because it can't be verified (ideally with an argument as to why the assumption is plausible - approximate normality *of the population* and large enough population are the most common examples - arguments based on distribution of the sample (not too weird, no outliers...) or on reasonable expectation that the population is large are most common examples)]

Indication of the relevant statistics (sample means, standard deviations, z- or t-value, p-value) - labeled so they can be attached to the meanings (p1 x-bar1 etc. are OK if 1 & 2 are distinguished before this – otherwise the distinction needs to be made).

For tests, a **statement of the decision criterion** is needed. This can be based on a critical value or not - with your students it's likely to be "the p-value is small enough that the observed difference is not likely explainable as chance" [or the opposite] or "this p-value is less than (whatever alpha-value is set)" [or the reverse]. Could also be of the form "The sample z is greater than (or less than) the critical z-value ..." [or is not beyond ...]. This can be included in the statement of the conclusion ("since the probability of a difference this extreme occurring by chance is ....") but has to appear somewhere.

A statement of the **conclusion (in context)**.

Actually, the statistics (other than the z, t, chi-square) are not always critical as long as the two populations (or population and sample) are clearly identified/distinguished. They only become critical if the student makes a button-pushing error. Then they can make the difference between "essentially correct" and "partially correct". In at least one of the rubrics you will find (I saw it as I was checking them twenty minutes ago, but have now forgotten where) the statement "an identifiable arithmetic (or was it calculation?) error will in an otherwise correct solution should be scored essentially correct" - "Identifiable" is an important word there - if the reader can see what the student did, and it's clear there's just a number error - *even if it leads to an incorrect conclusion (but correct for the student's work)* that work can be scored as essentially correct. The student must give enough information that the source of the error is clear for this happy event to occur. That usually means identifying all the sample statistics (input and output to the calculator). Writing the appropriate formula *with the problem statistics entered* is another way to show understanding in the face of incorrect results. [If the student is working from raw data only the summary statistics would be expected. If summary statistics were given, showing the substitution in the formula would be most useful].

Shorter answer - the formula and the statistics are not critical for a student who never makes mistakes. For a human student, they provide insurance against silly arithmetic. The calculator should be used for calculation where practical (including 2-sample t's, etc.).

Dr. Charlie Peltier
Math Professor/AP Reader
Saint Mary's College