FAMAT Statistics Test Guidelines & Common Disputes by MRG, January 2022

NOTE: Statistics tests are to be written at the Alpha / Precalculus level or below. At no time should any questions or concepts ever require any knowledge of Calculus.

January Tests:

I. Descriptive Statistics & Exploratory Data Analysis
   A. Types of data & levels of measurement
      1. Qualitative / Categorical: nominal vs. ordinal
      2. Quantitative: interval vs. ratio
   B. Interpreting graphs of distributions of categorical variables & data
      1. Bar graphs, side-by-side bar graphs, and stacked or segmented bar graphs, pie charts
      2. Two-way tables: marginal, conditional, and joint distributions
   C. Interpreting graphs of distributions of a sample of data from one quantitative variable (stemplots, histograms, boxplots, dotplots, ogives, etc.)
      1. Shape (symmetric vs. skewed), center, spread, outliers, and unusual features
      2. Summarizing distributions of one variable data
      3. Measures of center and location in a distribution: mean, median, quartiles, and mode(s)
      4. Measures of variation: range, interquartile range, variance and standard deviation
      5. The 1.5(IQR) Rule for outliers
   D. Quartiles, percentiles, z-scores - NOTE: The concept of percentiles or percentile ranks should only be used in a context involving a continuous distribution, such as a Normal distribution. Please see the common disputes section below for details.
   E. The effect of linear and non-linear transformations on the summary statistics of a data set
   F. Comparing distributions of one variable data via back-to-back stemplots, parallel boxplots, etc.
   G. Compare measures of center, spread, clusters and gaps
   H. Compare shape (skewness vs. symmetry), outliers, and any other unusual features

II. Exploring two-variable data sets via least squares linear correlation and regression
   A. Create and interpret scatterplots and residual plots and identify outliers and influential points
   B. Calculate and interpret the Pearson Product Moment Correlation Coefficient
   C. Assess linear relationships and the least squares regression line (slope and y-intercept)
   D. Calculate and interpret the coefficient of determination, residuals, and standard deviation (standard error) of the residuals
   E. Logarithmic or other non-linear transformations and linearity

III. Planning a Study and well-designed versus poorly designed studies
   A. Overview of methods of data collection and planning and conducting surveys and studies
   B. Census vs. sample survey and experiment vs. observational study
   C. Sampling methodology, sampling error / variability, and sources of bias
      1. Simple random, stratified, cluster, and systematic sampling
      2. Convenience samples, voluntary response samples, and voluntary response bias
      3. Undercoverage bias, nonresponse bias, response bias
   D. Planning and conducting studies and experiments
      1. Explanatory variables / factors, levels, treatments, and response variables
      2. Control group, placebo, placebo effect, blinding, and replication
      3. Experimental units / subjects, randomization and random assignment processes
      4. Completely randomized design, randomized block design, and matched pairs design

E. Scope of inference and Statistical Significance as a concept
    1. Inference about a population from a sample
    2. Inference about cause and effect from an experiment

IV. Fundamentals of Probability and Simulation
  A. The definition of a probability and the Law of Large Numbers concept
  B. Probability models, definitions, and calculations
    1. Sample, space, events and outcomes
    2. Types of events: independent, mutually exclusive / disjoint, complementary, exhaustive
  C. Probability rules and axioms
    1. Addition and multiplication rules and notation for event unions and intersections
    2. Conditional probability rules and notation
    3. Venn diagrams, two-way tables, and tree diagrams
    4. The 10% condition when sampling without replacement - see common dispute section
  D. Combinatorial (hypergeometric) probability calculations

V. Discrete and continuous random variables
  A. Binomial and geometric distributions and random variables
  B. Mean / expected value, variance, and standard deviation of a discrete random variable
  C. Continuous, uniform and non-uniform random variables and density curves and functions
  D. Z-scores, Normal distributions and their properties
    1. Empirical Rule (68, 95, 99.7 Rule)
    2. Forward and backward Normal calculations using charts and / or calculators
  E. Normal approximation to the binomial distribution
  F. Linear and non-linear combinations and transformations of independent random variables
    1. Notion of independent vs. non-independent random variables
    2. Mean, standard deviation, and probability calculations

VI. Sampling Distributions, their properties, and conditions
  A. Parameters vs. Statistics, sampling variability and unbiased vs. biased estimators of parameters
  B. The Sampling Distribution of the Sample Proportion and its properties and notation
    1. Mean and standard deviation
    2. Probability calculations and interpretations
  C. Sampling Distribution of the Sample Mean and its properties and notation
    1. Central Limit Theorem
    2. Mean, standard deviation, and probability calculations and interpretations
  D. Linear combinations of independent sampling distributions

February Tests - all of the above plus the following:

VII. One-sample Statistical Inference
  A. One-sample Confidence Intervals for an unknown proportion or mean
    1. Point estimators and point estimates
    2. Student's t-distribution definition and properties
    3. Critical values and standard errors vs. standard deviations of sampling distributions
    4. Definition and interpretation of the confidence level and margin of error
    5. Calculation and interpretation of confidence intervals
    6. Conditions for inference and the effects of input changes on the interval
    7. Minimum sample size for a confidence interval with a desired margin of error
    8. Matched-pairs confidence interval for a mean difference
    9. Know how and when to use a Z-interval vs. a t-interval for a mean

B. One-sample hypothesis testing of a population proportion or mean
   1. Definition and properties of hypothesis tests
   2. State hypotheses and check assumptions and conditions for inference
   3. Interpret p-value, significance level, and power
   4. Calculate and interpret Type-I and Type-II error probabilities and power
   5. Matched-pairs hypothesis test for a mean difference
   6. Statistical significance vs. practical significance and scope of inference
   7. Know how and when to use a Z-test vs. a t-test for a mean

VIII. Two-sample Statistical Inference
   A. Two-sample Confidence Intervals for a difference between unknown proportions or means
      1. Check assumptions and conditions for inference and interpret confidence level
      2. Calculate and interpret margin of error and confidence intervals
      3. Determine the effects of input changes on the interval and its interpretation
   B. Two-sample Hypothesis Tests for a difference between unknown proportions or means
      1. Definition and properties of two-sample hypothesis tests
      2. State hypotheses and check assumptions and conditions for inference
      3. Calculate and interpret test statistics and p-values
      4. Interpret significance level, Type-I and Type-II error probabilities and power
      5. Interpret results in context

March Tests - all of the above plus the following:

IX. Chi-Square Distributions (and their properties) and hypothesis tests
   A. Mean, variance, standard deviation, and shape of chi-square distributions
   B. Testing for Goodness of Fit, Independence, and Homogeneity
      1. State hypotheses and check assumptions and conditions for inference
      2. Calculate and interpret observed counts, expected counts, and components of the test
      3. Calculate and interpret test statistics and p-values
      4. Interpret results of the test and follow-up analysis in context

X. Inference for slope of least squares line and inference on transformed data to achieve linearity
   A. The Sampling Distribution of the Sample Slope: shape, mean, and standard deviation
   B. State hypotheses and check assumptions and conditions for regression inference
   C. Calculate and interpret standard error of the residuals and the slope of the least squares line
   D. Construct and interpret a confidence interval for the population regression slope
   E. Perform a significance test for the population regression line slope and interpret results
   F. Perform all of the above for transformed data (logarithmically or otherwise)

Common and / or Potential Disputes:

1. Unless explicitly stated or sufficient context implies otherwise, any set of "data" or other numbers is treated as a sample and not a population for calculation of the variance and standard deviation.

2. The quartiles of a data set are computed with the exclusion of the median when determining the "lower half" and "upper half" of the data set. In the case of repeated values at the median for a data set with an odd number of values, one of those values is omitted when determining the quartiles.

3. The notion of a "percentile" or "percentile rank" of a single data value within a small, finite, and discrete data set is not well defined. It is possible to legitimately interpret it as either the proportion of data values "strictly below" or "below or equal to" the individual data value in question. Therefore, rather than use the terms "percentile" or "percentile rank" in the context of such a data set, questions should ask for the more explicit "percent of data below" or "percent of data below or equal to" the data value in question so as to avoid this ambiguity.

4. Skewness vs. symmetry vis-a-vie the relationship between the mean and median in a data set:
   a. If the distribution of a data set is distinctly skewed to the left or the right, then the mean is almost always pulled away from the median and in the direction of the skew. The notable exception is the presence of an extreme value (usually an outlier) in the opposite direction.
   b. If a data set is distinctly and monotonically skewed to the right (or left), then the mean is greater (or less) than the median. However, the converse of this statement is not necessarily true. Namely, just because the mean is greater (or less) than the median, one cannot definitively conclude that the distribution of the data set is skewed to the right (or left) solely based on this information.
   c. If the distribution of a data is perfectly symmetric, then the mean and the median are exactly equal. However, the converse of this statement is not necessarily true. Likewise, if the distribution of a data set is approximately symmetric, then the mean and median are likely to be approximately (but not necessarily exactly) equal.
   d. The above facts therefore lead us to the following: merely knowing the relationship between the mean and the median (i.e. which one is greater or that they are either equal or approximately equal) is insufficient information to definitely conclude that the distribution of a data set is distinctly skewed in either direction or that it is symmetric or even approximately symmetric. This is especially true of small data sets where even a single outlier is highly influential.

5. The term influential is extremely vague and nonspecific. It is preferable that questions that assess this concept involve some form of calculation or other criteria for determining the "influentiality" of a value.
6. Given a deck of cards:
   a. The deck is always assumed to be a single, ordinary, fair, well-shuffled deck with no additional cards or missing cards and with each card having equal probability of being drawn unless specifically stipulated otherwise.
   b. Cards are always drawn without replacement unless specifically stipulated otherwise.
   c. Face cards include only Jacks, Queens, and Kings. Aces are not included in the set of face cards (since, strictly speaking, they have no "face" on them) nor do they have any integral value (such as 1) since this is only dependent upon a specific game being played, such as Blackjack.

7. Given a coin (or coins) or a die (or dice):
    a. Coins are always assumed to be fair and two-sided unless specifically stipulated otherwise.
    b. Dice are always assumed to be ordinary, fair, and six-sided unless specifically stipulated otherwise. If a 10-sided die is used, the numbering on the sides must be specified as being either 0 through 9, inclusive, or 1 through 10, inclusive. Otherwise, a 10-sided can legitimately be interpreted as either one, which is ambiguous.
    c. If more than one die is rolled, the exact number of dice <u>must</u> be specified; for example: "two dice" or a "pair of dice," etc. Simply writing "dice" is <u>not</u> sufficiently specific and open to dispute. What is being done with the results of each die must also be specified, such as taking their sum or product or looking at each die result independently of the others. If it is not specified, the question is open to interpretation and disputable and any valid interpretation should be granted.

8. When using the Normal distribution to approximate the binomial or in any other discrete situation:
    a. Whenever the conditions for using the Normal approximation to the binomial are met (namely, independence of observations / trials or the 10% condition is met and $np \geq 10$ and $nq \geq 10$), there may be two potentially correct answers depending on whether or not the questions explicitly instructs students to use the Normal approximation to the binomial distribution. This is because the binomcdf and normalcdf functions in the calculator may return slightly different results depending on sample size, rounding, and the use of any <u>correction for continuity</u>.
    b. Students should <u>never</u> use any <u>correction for continuity</u> unless specifically told to do so in the question. This also applies to any other situation where a probability in a discrete setting is approximated by a continuous random variable, such as a Normal distribution.

9. Unless the context of a problem is such that successive trials, outcomes, events, selection of individuals from a population, or groups being compared are obviously or reasonable assumed to be independent (such as successive die rolls or coin flips) the question must specifically stipulate independence between such trials, outcomes, events, selections, or groups or that "all assumptions and conditions are met," as in the case of inference questions. Otherwise, the question is open to dispute. For example: someone "practicing" an activity reasonably implies that subsequent trials are not necessarily independent. Therefore, independence of trials must be specified to avoid a valid dispute on these grounds. In the case of inference questions, simply stating that one may assume all inference assumptions and conditions are met will suffice to avoid this potential dispute.

10. Current AP Statistics curriculum states that <u>all</u> expected counts are at least 5 in order to perform chi-square tests. However, some older texts use less stringent criteria. Please use the current curriculum standards.

11. Logarithmic or any other non-linear transformations of data do not necessarily <u>guarantee</u> a linear relationship after the transformation. Likewise, if a linear relationship between transformed data sets exists, then it does not necessarily <u>guarantee</u> that the relationship was non-linear prior to the transformation.

12. Regarding rounding instructions and answer formats:
    a. All individual test questions that require rounded final answers (such as in the case of a repeating decimal or an irrational number computed in the calculator) must explicitly state the decimal places to which the final answer is rounded, for example, to the thousandths place. Mere rounded answer choices <u>do not</u> sufficiently imply to what decimal place the final answer is

rounded. For example, if the correct final answer is $\sqrt{2}$, and 1.414 is in an answer choice, it is not an acceptable answer unless specifically told to round to the thousandths place.

b. The above rule applies to such things as Normal distribution calculations in the calculator as well: P(Z < 1) = 0.8413 is not correct and acceptable (even if it is amongst the choices listed) unless explicitly told to round to the ten-thousandths place (or to four decimal places).

c. An otherwise unflawed question with rounded answer choices but without explicit rounding instructions given in the question should switch to E since a technically correct answer exists but it is not among the choices listed. Additionally, if the rounded answer choices do not match the given rounding instructions, the question should switch to E.

d. If the correct final answer (and any distractor answer) is a simple terminating decimal, the answer choices may be expressed in either simplified fraction form or as any correct equivalent terminating decimal but preferably without any trailing zeros. For example ½ = 0.5 but preferably not 0.50 or 0.5000, although these answers are considered correct and acceptable. In this case, rounding instructions are not necessary, but preferred. However, tests should be written with consistently rounded answer choices and answer choices should (preferably) not have trailing zeros unless rounding instructions deem them necessary.

e. If a question is phrased as "Which of the following is closest to the [correct answer]," then more explicit rounding instructions are not necessary (yet are preferred), so long as one and only one of the listed answer choices is indeed closest to the exact correct final answer.

f. In general, if a question is phrased as "which of the following is…" then students must choose from the listed choices in A, B, C, and D even if there is a technically "better" answer than those listed and should only choose E if the question is truly flawed or it is indeed the correct answer since none of the other choices are correct. For example, if a question asks: "Which of the following values is the largest ____?" then students must choose from the listed options even if a larger valid value exists than the largest value listed in the choices.

g. All team question free response answers are to be exact unless specific rounding or format instructions are given in the question. For example, if a simplified fraction is requested, then the decimal equivalent is not acceptable. Thus, in this case, 0.5 is not acceptable in place of ½. However, if no particular answer format (or rounding) is specified, then any correct equivalent form is acceptable: 0.5 = ½ = 0.50 = 50% etc.

h. In the case of a final team answer with trailing zeros when asked to round to a specific decimal place, such as 0.5000, the trailing zeros must be present to receive credit for a correct answer. However, in the interest of fairness to all students, team questions should be written in such a way that the final answer does not require any trailing zeros when rounded and asked to round to a specific decimal place. For example, if the final answer is 1.959963986… but rounded to a specific decimal place, the final answer is best expressed rounded to the hundredths place (1.96) or the hundred-thousandths place (or five decimal places: 1.95996) rather than to the thousandths place (1.960) or to the ten-thousandths place (1.9600). This will avoid an influx of disputes over this rounding issue. Leading zeros in decimal valves less than 1 are never required. For example, .5 is always acceptable in place of 0.5.

13. In general, if a slightly flawed question has a correct answer but it is not among the choices listed, it should switch to E. However, if a question is sufficiently flawed, ambiguous, vague, or open to multiple potentially valid interpretations or it is not explicitly clear what it is stating or what is being requested, it should get thrown out.