AP Statistics Semester Final Exam Study Guide – Chapter 1 to 6

**79 M/C Questions (110 minutes) – worth 200 test points**

You bring:

1. Your graphing calculator (or your ID to check one out)
2. 2 blank pieces of scratch paper
3. Pencil and eraser

I will provide you with:

1. AP formula sheet

**Suggestions for studying:**

1. Read the Chapter Summary and Section Summary for each chapter and take notes on the concepts, vocabulary, and formulas. Pay close attention to formulas you need that are NOT on the AP formula sheet that will be provided on the exam.
2. Do all of the review assignments and ask lots of questions
3. Review your hand-written notes for each chapter
4. Review the extra examples we have done together in class
5. Review calculator functions - see notes or “Technology Corners Reference” at back of book
6. Go through the PowerPoint slides for each chapter
7. Take the multiple choice tests on the book publisher’s website
8. As needed, do the “Check for Understanding” problems – especially Chapters 2 and 3
9. Review the guided reading notes
10. Do steps #1 – 9 ahead of time so you can relax when it is time for the final exam

Note: You do not need to worry about the 10% condition (sample size no more than 10% of the population size) for independence or the approximation to normal condition (np and n(1 – p) >= 10). We will save those two concepts for Chapter 7 next semester.

Chapter 1: Graphing with One Variable and Interpreting Graphs **(14 questions)**

* Analyzing and displaying categorical and quantitative data
* Categorical data (doesn’t make sense to find the mean)
  + See graphically with a pie chart or a bar graph
* Quantitative data
  + Graphically: (Describe using your SOCS, and remember outlier formula – 1.5 x IQR)
    - Dot-plot
    - Stem-plot
    - Histogram
  + Numerically:
    - Mean and standard deviation
    - Median and IQR

Chapter 2: Normal Distributions **(13 questions)**

* Describing location in **any** distribution
  + The percentile tells you the percent of values that are less than or fall below the given value.
  + The z-score tells you the number of standard deviations above or below the mean your value lies.
* And **if you have a normal** distribution, you can then you can use the z-score and Table A ( which describes the areas under a Standard Normal Curve given the z-score to compute probabilities. [ or use normalcdf (p. 123) – you don’t need z-score at all ]
  + You can go the other direction. Use invNorm (percentile, mean, std dev) on your calculator to obtain the value, given the percentile. (p. 124)
* With a Normal Distribution – the mean and standard deviation completely specify the distribution.
* Normal distributions obey the 68-95-99.7 rule which describes what percent of observations lie within one, two or three standard deviations of the mean.
* Transforming Data – The shape never changes
  + Adding a constant (adds to center and location, but does not change spread
  + Multiplying by a constant (multiplies center and location, and spread)

Chapter 3: Describing Relationships between two Variables **(13 questions)**

* Scatterplots help see graphically the relationship between two variables.
* Explanatory and Response variables – which is which?
* Correlation, r – measures the strength and direction of the linear association between two quantitative variables x and y
* Coefficient of determination, – the percent of variation in [response variable] accounted by the regression line
* LSRL - Least-squares regression line – Be SURE you can find it on your calculator and also the

strength of the correlation, r, by making sure “Diagnostics on”. Also make a residual plot.

Try to re-do the Valentine’s Day activity from scratch.

* This regression line “predicts”, “estimates”, “approximates” the value of y given x.
* Be careful not to make the mistake of extrapolation – using values outside of a reasonable range for context of problem.
* Residual plot – is a scatterplot of the residuals against the eXplanatory variables. If a linear model is appropriate in the original data, the residual plot should show no obvious patterns and the residuals should be relatively small. For a “perfect fit”, the residual plot would be a horizontal line.

Chapter 4: Designing Experiments **(13 questions)**

* Don’t confuse the language for doing a sample survey from that for experiments.

Sample Survey language:

Biased: convenience sample, voluntary response sample

Unbiased: SRS, stratified, cluster, multi-stage (meaning you combine one or more

of the previous 3 methods)

Sources of error: undercoverage, non-response, wording of questions

Experimental Design language:

Completely randomized, randomized block, matched pair

Use blocking to control effects of lurking variables (reduce confounding)

Control what you can, block what you can’t, then randomize to create comparable groups.

* Association does \*not\* mean causation. Well-designed experiments are the only way we can

establish causation.

Chapter 5: Probability **(14 questions)**

* “or”
  + Two events are mutually exclusive if they have no outcomes in common so they can never both occur.
  + General addition rule for two events
    - P (A or B) = P(A) + P(B) – P (A and B) , however P (A and B) is zero if events are mutually exclusive
* “and”
  + Two events are independent if P(B | A) = P (B)
    - So use P (A and B) = P(A) \* P (B)
  + Otherwise, you need to use one of 3 methods for Conditional Probability
    - 1) Two-way table
    - 2) Tree diagram
    - 3) Formula -
* Note – a tool you can use for ANY of the above is a Venn Diagram!

Chapter 6: Random Variables **(12 questions)**

* For a discrete random variable with a certain given probability distribution, be able to find:
  + Expected value
  + Variance
  + Standard deviation =
* For a continuous random variable – the probability falls in a range of values so you find the area under the curve. You learned about this in Chapter 2 when you have a Normal Distribution….nuff said….we’ll talk about other distributions later.
* Transformations of ONE Random Variable (meaning you add or multiply by a constant)
  + If (meaning that Y is the result of the transformation on X)
    - Then new probability distribution has the same shape
    - But the mean of Y is:
    - The standard deviation of Y is:
    - The variance of Y is:
* Sum and Difference of TWO Random Variables
  + The mean is (The mean of the sum or difference of two random variables is the sum or difference of their means.)
  + And as long as the variables are independent:
    - The variance is (The variance of the sum or difference of two random variables is the sum of their variances.)
    - The standard deviation is

\*\* Note: you can NOT directly add standard deviations!! Instead you add the variances and take the square root of the sum.

* In a Binomial setting, meanings BINS is true:
  + You can use the binomial probability formula where n = fixed number of trials and p = probability of success OR use binompdf (n, p, k) or binomcdf (n, p, k) (see p. 389)
  + The mean of X is:
  + The standard deviation of X is:
* In a Geometric setting, meaning BITS is true:
  + You can use the geometric probability formula, (p)
  + The mean of X is:
  + The standard deviation of X is: Don’t worry about it.