AP STATISTICS

AP EXAM STUDY GUIDE

Table of Contents

You are responsible for…

- Completing this study guide (5 points per topic)
- Completing the Practice Problems (5 points per topic)
- *Studying hard and doing your best!*

Topic 1: Sampling Techniques and Sources of Bias (2.2)

1. Know and understand the difference between a *population* and *sample*
   - How is each one measured (what do we use to measure them)?

     taking a measurement from every subject/object creates a population parameter, taking a measurement from a subset creates a sample statistic
   - Why do we often measure samples instead of populations?

     populations take too much time, may be impossible

2. Know the different types of *bias* and how to spot them in different situations
   - *Bias* is anything that causes a sample to be **not representative of the population of interest**
     - You must be able to articulate *what* the bias is, *why* it should be considered bias, and *how* it distorts the results from what they otherwise might be.
   - What is the difference between *sampling error* and *sampling bias*?

     Sampling error should be random, due to differences in subjects not design.
     sampling bias is a problem with the design.
   - How can a small sample size affect the validity of the sample? (*this is related to sampling error rather than bias)*

     Increasing sample size will not reduce bias it can reduce sampling error

Define the types of **sampling bias** (a bias in *who* was in the sample)	Define the types of **response bias** (a bias in *what* the sample is saying)
Undercoverage Not everyone who should be in the study is included in the sample	Loaded Questions Cause the response to differ from the subjects true response.
Nonresponse bias Not everyone included in the study has a recorded response	False answers Measurement devices that are miscalibrated or broken leads to measurements that are different from true.
Voluntary response bias People in the study have a reason to be included, volunteers are not a part of a random process	

3. Know the different types of sampling techniques and how to identify which one is being used (as well as the *advantages* and *disadvantages* of each)

Simple Random Sample (SRS)	Stratified Random Sample
A sample taken in such a way that all samples of size n have an equal chance of being selected	When pop is easily divided on a variable that may affect the response, then take a SRS from each subset **Stratifying will reduce variability of possible sample results!*
Systematic Random Sample	**Cluster Sample**
Pick a random # from 1 to n, then start there and every nth person after that is included in the study	When pop is divided into ~~homogeneous~~ heterogeneous sub groups randomly select n sub groups and include everyone in the cluster
Multistage Sample	**Convenience Sample**
X	Sample taken with out any random process

4. Know how to design a random sampling procedure

- **Random number generator** will be your friend!
- "Describe a method…" (NOTE: blanks will be filled in with the context of the problem!)
  - START WITH: *Assign each __________(unit, subject, etc.) a different number between _____ and _____*
  - *Describe how you will implement the sampling method you want to use*
  - *Randomly select _________ numbers, ignoring repeats, and include the __________(unit, subject, etc.) that corresponds with those numbers in your sample.*

Example: Mr. Frederick wants to create an advisory committee of 20 randomly-selected students out of the 1,950 students at Grant High School. Describe how he could do so using a…

Simple random sample	Systematic Random Sample
# the students from 1 to 1950 use a RNG to generate 20 unique #'s from 1 to 1950, The students whose name corresponds to the #, will be included in the study	# the students from 1 to 1950, using a RNG select 1 number btw 1 & 97, start with that student to be in the study then every 97th student after that will be included in the study
Stratified Random Sample	**Cluster Sample**
# the students fr 1 to 1950 Divide the H.S. into 4 groups by grade in school, using a RNG select 5 students from each grade, those 20 students will be included in the study	# the 97 homerooms from 1 to 97 use a RNG to get a number from 1 to 97. All students in that room are to be included in the study
Multistage Sample	**Convenience Sample**
X	Survey the first 20 students that enter the main doors

Topic 2: Experimental Design (Notes: 2.1, 2.2, 2.3, 2.4)

1. Know the vocabulary of experiments and experimental design

- What is the difference between an Experiment and an Observational Study? Which one lets us establish cause-and-effect relationships? ***HINT:*** *There is one "dead giveaway" keyword when identifying an experiment. It starts with the letter A.*

  Treatments are assigned to subjects / objects

- Define *Treatment* –

  ~~Combination of experimental~~
  Explanatory variable manipulated by the researcher

- Define *Confounding* –

  ~~two or more~~ variables that are not of interest to the study that may effect the response variable

- Define *Experimental Units* (*Subjects* when human) –

  one member of a set of objects/subjects that are initially equivalent / smallest unit to which a treatment can be applied

2. Know the four principles of a good experiment

- Direct Control
- Blocking
- Randomization
- Replication

3. Know methods for **controlling** an experiment to prevent bias

- Control group (what is it, and what does it allow us to do?)
  (***NOTE:*** *A control group is NOT mandatory; it is just one way to get comparison, which IS mandatory*)

  allows the researcher to assess how the response behaves when the treatment is not used

- Placebo effect –

  ~~to~~ a treatment with no active ingredients, used to compare if the process of the treatment has an effect on the response

- Blind study – The subject does not know which treatment was recieved

- Double-blind study – The subject and the person measuring the response do not know the treatment recieved

4. Know the different types of experimental design and how to identify which one is being used (as well as the *advantages* and *disadvantages* of each)

- Completely Randomized Design a design that uses randomization of factors to control the effect of extraneous variables

- Randomized Block Design ("Blocking") treatments are assigned within blocks of ~~[illegible]~~ subjects, ~~each treat~~ all treatments are used in each block used to control an extraneous variable that randomization alone may not

- Matched Pairs Design Subjects are grouped into pairs based on an extraneous variable ~~then each~~ then within each pair subjects are randomly assigned a treatment

5. Be able to discuss *generalizability* – the extent to which the results of a sample (or experimental group) can be applied to a certain population

- You can generalize to the population *from which the sample or experimental group was taken*
- **BIAS** can hurt (or even eliminate) generalizability. You need **RANDOMNESS** to avoid this!
    - For example, a study that consists of **volunteers** should only be generalized to those volunteers! You *might* be able to generalize to "people who are similar to the volunteers," but absolutely no further, because they weren't *randomly selected*!
    - ***NOTE:** Even a relatively small sample size (not ridiculously small, but somewhat small) can be valid as long as it's random!*

Example:

A researcher studied a random sample of 100 teens in Oklahoma. To which populations will the results of this researcher's findings be generalizable? (Circle ALL that apply)

(A.) The 100 Oklahoma teens in the study

(B.) All teens in Oklahoma

C. All teens no not all teens were in the sampling frame

D. All Oklahomans no not all Okl. were in sampling frame

Topic 3: Analyzing Data (Notes: 1.4, 3.2, 3.3, 4.1, 4.2, 4.3)

1. The 5 things you should discuss when analyzing a **distribution** of data:

shape, center, spread, and any other interesting features

***NOTE:** If asked to compare data sets, make sure you explicitly compare them (For example, "The first distribution has a greater mean than the second distribution, while the second distribution has a greater spread than the first")*

2. **Center**

Measure	*How to find it*	*Resistant to the effects of outliers?*
Mean *Population:* μ *Sample:* $\bar{x}$	$\frac{\Sigma x}{n}$	no
Median	$\left(\frac{n}{2} + .5\right)^{th}$ value in an ordered data set	Yes

- The best one to use is usually mean, unless the data is skewed, at which point median should be used

3. **Shape**

Shape	**Normal**	**Skewed Left**	**Skewed Right**	**Uniform**	**Bimodal**
Sketch					
Which is greater, mean or median? (or are they =)	=	$\bar{x} < q_2$	$\bar{x} > q_2$	$\bar{x} = q_2$	$\bar{x} = q_2$

4. **Spread**

Measure	*Paired with... (mean or median)*	*How to find it*	*Resistant to the effects of outliers?*
Standard Deviation *Population:* σ *Sample:* s	mean	$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{n}}$ $s = \sqrt{\frac{\Sigma(x-\mu)^2}{n-1}}$ ***Or use 1-Var Stats!***	no
Variance *Population:* σ^2 *Sample:* s^2	mean	$\sigma^2 = \frac{\Sigma(x-\mu)^2}{n}$ $s^2 = \frac{\Sigma(x-\mu)^2}{n-1}$ ***Or use 1-Var Stats!***	no
Lower Quartile (Q1)	median	Midpoint of Minimum and Median ***Or use 1-Var Stats!***	yes
Upper Quartile (Q3)	median	Midpoint of Median and Maximum ***Or use 1-Var Stats!***	yes

Range	either	max - min	no
Interquartile Range (IQR)	med.	$Q_3 - Q_1$	yes

5. **Outliers** (You may ALSO want to point out gaps, clusters, and any other "interesting" features a data set may have)

- *What is an outlier?*
  any value more than 1.5 IQR away from a quartile
- **NOTE:** An outlier **CAN** change the value of the Median, Q1, Q3, etc. if the addition of an outlier causes the *position* of numbers to change. However, this change will ***usually*** be slight
- *How to identify outliers:* **IQR TEST** (remember, this is a general guideline, not a strict rule!)

How it works:

lower < Q_1 − 1.5 IQR

upper > Q_3 + 1.5 IQR

Example: Min = 11, Q1 = 32, Med = 36, Q3 = 44, Max = 51

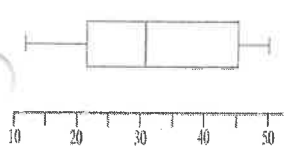
lower < 32 − 1.5(12)
lower < 14

upper > 44 + 1.5(IQR)
upper > 62

Any point *below* 14 or *above* 62 can be considered an outlier. **Outliers in this data set:** at least one lower

6. **Graphs**

Boxplot



Notes:
- Min, Q1, Med, Q3, Max
- **Cannot show shape** (but *can* show skews)
- **Outliers** should be marked with a *

Stemplot

stem	leaf
6	9
7	
8	7 8 8 9
9	0 6 7 7
10	0

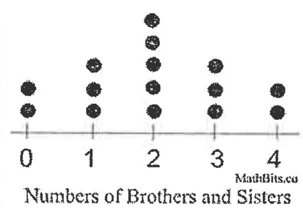
Key: 6|8 means 68

Notes:
- Remember to give a *key* to show what the numbers mean
- **Do not skip stems**
- If given a **back to back** stemplot, *always* read stem first, then leaf

Boys		Girls
7	0	
1	1	1
146	2	268
458	3	3446689
122289	4	436
3479	5	4

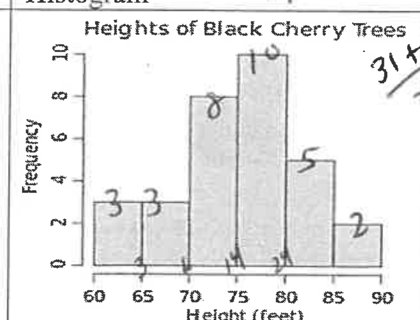
This data point is 24, **NOT** 42

Dotplot



So easy a caveman could do it!

Histogram n = 31



$\frac{31+1}{2} = 16$

Notes:
- X-axis shows *intervals*, y-axis shows the *frequency* (number of data points that belong in that interval)
- **Finding the median:** Figure out how many data points there are, use $\frac{n+1}{2}$ to find the *position* of the median, then figure out which interval contains that position!

EXAMPLE:

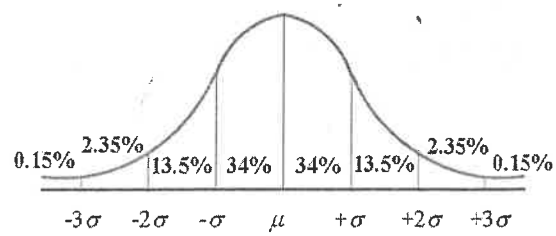
Number of data points: 31

Position of median: 16th

Interval containing median: 75 to 80

Topic 4: Normal Distributions and Z-Scores (Notes: 4.4 and 7.6)

1. Know how to analyze a normal distribution
   - *THEORETICAL* distribution (in reality, we consider data to be approximately normal)
   - It's like a **histogram** in which the center is the mean and the intervals are each one std. deviation



2. Know how to use the **Empirical Rule**
   - About 68 % of data is within 1 Standard Deviation of the mean
   - About 95 % of data is within 2 Standard Deviations of the mean
   - About 99.7 % of data is within 3 Standard Deviations of the mean

3. Know how to calculate and interpret **z-score**
   - A data point's z-score is the # of standard deviations away from the mean
   - **Formula** (NOT in AP exam): $\mathbf{z} = \frac{x-\mu}{\sigma}$
   - Z-scores can help us compare two **unalike** measurements
     Example: *Suppose the weights of apples are normally distributed with a mean of 85 grams and a standard deviation of 8 grams. The weights of oranges are also normally distributed with a mean of 131 grams and a standard deviation of 20 grams. Amy has an apple that weighs 90 grams and an orange that weighs 155 grams.*

     1. Calculate **and interpret** the z-score of Amy's apple

        $Z_a = \frac{90-85}{8} = .625$

     2. Which is *relatively* larger, Amy's apple or her orange? **Explain.**

        $Z_o = \frac{155-131}{20} = \frac{24}{20} = 1.2$

        The orange, it is more std dev above the mean makes it larger

     3. How large would Amy's apple have to be in order to be comparable to her orange?

        $1.2 = \frac{x-85}{8}$

        $9.6 = x - 85$ $\quad$ $x = 94.6$ g

4. Know how to use Z-scores to calculate the percentage of data points above, below, or between certain boundaries
   This works **ONLY for normally-distributed data!! **DO NOT** do these procedures if you do not **know** that your data is normally distributed!*

With Z-table	With Calculator
• Z-table gives the percentage of values **below** a given z-score • You can use the z-table **backwards** – if you know the percentage, find it on the z-table, and see what z-score it equates to!	• NormalCDF (if *looking for* percentage/probability) • InvNorm (if *given* percentage or probability) • To adequately *show work*, you must write…

Topic 5: Probability Rules (Notes: Chapter 6)

1. Understand what probability *is*

- How do you calculate the probability of an outcome?

$$P(s) = \frac{\#\text{ of } S's}{\#\text{ of trials}}$$

- What is the Law of Large Numbers?

As the number of chance experiments increase, the difference btw the true value and the relative frequency of success approaches zero

- What are *mutually exclusive* outcomes?

Two events that cannot occur ~~simultaneously~~ simultaneously or two events with no common outcomes

- What are *independent* events?

Two events in which the occurance of one event does not effect the occurance of the 2nd event.

- Why can two events that are mutually exclusive *never* be independent?

If one of the mutually exclusive events occur then we know the other will not occur so they can not be independent (the occurance of one changed likely hood of other)

2. Know the basic rules of probability

- When calculating the probability of getting more than one outcome for a given event, what formula should you use? ***HINT:*** *Always account for any overlap between outcomes!*

$$P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$$

- When analyzing events with multiple outcomes, what visual aide will be the most beneficial?

Venn diagram, tree diagram, two way table

- When calculating the probability of *multiple* events, what rule or formula should you use?

- When, and *how*, do you use the *combinations* (nCr) function in your calculator?

Use when you need a number of arrangements

- When analyzing a series of multiple events, each with multiple possible outcomes, what visual aide will be helpful?

tree or two way table

- When calculating the probability of multiple independent events, what three things should you account for? ***HINT:*** *The formula on the formula sheet may help you!*

~~p # of success~~ n # of trials
p robability of success
x # of success

- How does the above procedure change when the events are dependent?

- What is *conditional* probability, and how do you calculate the conditional probability of a given event?

A probability that is dependent on another event occuring

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Situation	**Rule**	**Formula**
"At least one"	Opposite of "none"	1 – P(0)
Multiple outcomes – mutually exclusive	Add probabilities	P(A∪B) = P(A) + P(B) ***NOTE:*** *P(A∩B) = 0 (no overlap for mutually exclusive events)*
Multiple outcomes – NOT mutually exclusive	Add probabilities but ***subtract the overlap*** *If using a Venn Diagram, just add up the 3 sections in the diagram	P(A∪B) = P(A) + P(B) – P(A∩B)
Multiple events – Independent	Multiply probabilities, and account for COMBINATIONS in which these events can occur (nCr)	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ nCr • $(p_{success})^{\text{\# of successes}}$ • $(p_{fail})^{\text{\# of fails}}$
Multiple events – Dependent	Multiply probabilities *Account for the **change** in probability with each trial *Account for **combinations** (nCr)	nCr • $p_{event\ 1}$ • $p_{event\ 2}$ • $p_{event\ 3}$... ***NOTE:*** *Remember these probabilities CHANGE!!*
Conditional Probability (A *given* B)	$\frac{\text{Probability of both events}}{\text{Probability of first event}}$	$P(A\|B) = \frac{P(A \cap B)}{P(B)}$

Topic 6: Probability Distributions (Notes: Chapter 7)

1. Know the different types of random variables and how their distributions work

- What is the difference between a discrete and a continuous random variable?

A discrete random variable has values that are isolated points on a number line.

A continuous random variable has values that could fill an interval on a number line

- For continuous random variables, what is the probability of getting *exactly* one given outcome? 0

- How do you calculate the **expected value** of a discrete random variable?

$$\mu_X = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

- What is the **definition** of expected value? (It mean something *very specific*)

The average outcome in the long run

- What formula can you use to calculate the spread (st. dev.) of a discrete random variable *by hand*?

$$\sigma_X^2 = \sum (x - \mu_X)^2 \cdot P(x)$$

- How are variance and standard deviation related?

$$\sigma_X = \sqrt{\sigma_X^2}$$

2. Know how transforming and combining a random variable changes that variable's distribution

Action	Effect on **Center** (mean)	Effect on **Spread** (standard deviation)
Adding/Subtracting a CONSTANT (number)	If $Y = X + c$ $\mu_Y = \mu_X + c$	$\sigma_Y = \sigma_X$
Multiplying/Dividing by a CONSTANT (number)	If $Y = aX$ $\mu_Y = a\mu_X$	$\sigma_Y = \lvert a \rvert \sigma_X$ or $\sigma_Y^2 = a^2\sigma_X^2$
Combining (adding or subtracting two random variables to each other)	If $Z = X + Y$ $\mu_Z = \mu_X + \mu_Y$	If $Z = X + Y$ $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$ $\sigma_Z = \sqrt{\sigma_X^2 + \sigma_Y^2}$

HINTS:

- If *X* and *Y* are normally distributed, so are *X* + *Y* and *X* – *Y*. This means **use normalCDF!**
- *X* > *Y* is the same as *X* – *Y* > *0* (likewise, *X* < *Y* is the same as *X* – *Y* < *0*)

Topic 7: Binomial and Geometric Distributions (Notes: 7.5)

1. Know and understand how to use a Binomial Distribution

- **Using the Binomial Distributions**

  – Only works in *binomial* settings, which occurs when the following conditions are met (**"BINS"**)

    - B: Only 2 mutually exclusive outcomes
    - I: Prob. of success same on each trial
    - N: Each trial is independent
    - S: Set # of trials

  – BinomPDF: finds $P(X = x)$

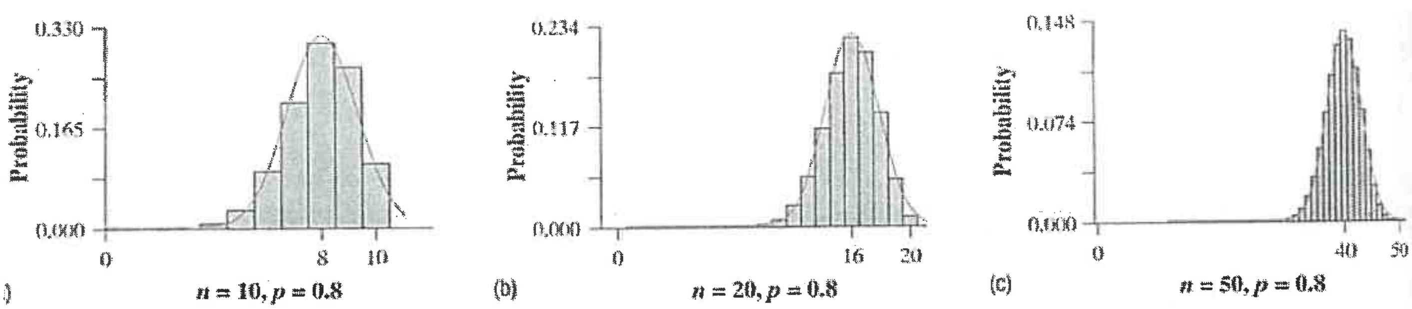
  – BinomCDF: finds $P(X \leq x)$

- **Binomial Curve**

  – CENTER: $n\pi$ (number of trials • probability of success = expected # of successes)

  – SPREAD: Standard Deviation, $\sigma = \sqrt{n\pi(1-\pi)}$

  – SHAPE: Approaches **normality** if you can *expect* at least ______ successes and ______ failures



Example:

Genetics says that children receive genes from each of their parents independently. Each child of a particular set of parents has probability a probability of 0.25 of having Type O blood. Suppose these parents have 6 children. Let X = the number of children with Type O blood.

a. Calculate the mean and standard deviation of the number of children who will have Type O blood

$\mu_X = n\pi = 6(.25) = 1.5$ $\sigma_X = \sqrt{n\pi(1-\pi)}$

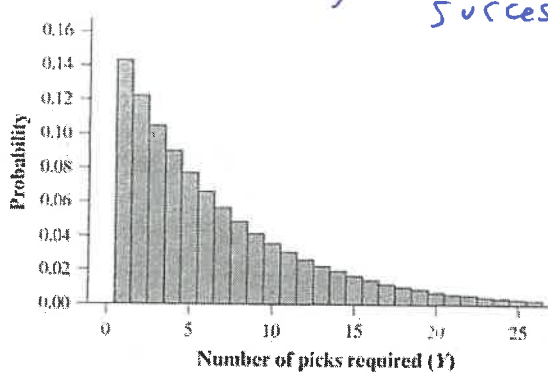
$\sigma_X = \sqrt{6(.25)(.75)} = 1.061$

b. Find the probability of each of the following

P(X = 4); exactly 4 children will have Type O blood	P(X ≤3); 3 or fewer children have Type O blood	P(X > 1); More than 1 child will have Type O blood	P(X ≥3); 3 or more children will have Type O blood.
$\binom{6}{4}(.25)^4(.75)^2$ bpdf(6,.25,4)	bcdf(6,.25,3)	$P(X>1) = 1 - P(X \leq 1)$ 1 − bcdf(6,.25,1)	$P(X \geq 3) = 1 - P(X \leq 2)$ 1 − bcdf(6,.25,2)

2. Know and understand how to use a Geometric distribution

- Geometric Distribution – a density curve that allows us to determine how many trials it will take to get one success (also think of it as wait time)
  - Events need to be Independent (of course)
- How to calculate it
  - **Calculator**
    - GeometPDF is used for ~~geometric~~ equal ~~distribution~~, the probability that the first success will happen on the Kth trial
    - GeometCDF is used for inequalities, the probability that the first success will happen on or before the Kth trial
    - Type in probability and the trial #
- EXPECTED VALUE (mean) of a Geometric Random Variable is $\frac{1}{p}$ *(If $n = \frac{1}{p}$, then $np = 1$)*
- Shape is always right skew
  - As you continue, the probability of having your first success gets lower



Examples:

1. A slot machine has a win rate of 8%. A gambler wants to play at this slot machine until they win – then, they will leave.

a. What is the expected number of games the gambler will have to play in order to win? $\frac{1}{.08} = 12.5$

b. Find the probability that it will take the gambler…

7 spins to win	10 or fewer spins	More than 20 spins
$(.92)^6(.08)^1$ $.0485$	$P(X \le 10) = gcdf(.08, 10)$	$P(X > 20) = 1 - P(X \le 20)$

Topic 8: Sampling Distributions (Notes: Chapter 8)

1. Know the basics of *sampling distributions*

- What is the difference between a *parameter* and a *statistic*?
  A parameter is a measurement taken from a population
  A statistic is a measurement taken from a sample
- What is the difference between a *proportion* and a *mean*?
  Proportion is the ratio of the # of times a categorical variable appears divided by ~~the~~ sample size.
  mean is the sum of ~~all~~ the numerical variable for every object in the sample divided by the sample size
- What is a *sampling* distribution?
  A collection of all possible sample statistic taken from all the possible samples in a population
- Know the difference between a sampl<u>e</u> distribution and a sampl<u>ing</u> distribution
  - Sampl*e* distribution – a graph of data taken from one sample
  - Sampl*ing* distribution – a graph of statistics taken from multiple samples

2. Know the importance of the **Central Limit Theorem** (define it below)

When n is sufficiently large the sampling dist of $\bar{x}$ is well approximated by a normal curve, even when the population is not itself normal

3. Know how to analyze a **normal distribution**, and use it to find the probability of a sample statistic occurring, *given* an assumed population mean and standard deviation

- *What function in the calculator should we used to do this?* normalcdf

From the AP Formula Sheet:

If X has a binomial distribution with parameters n and p, then…	If $\bar{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then…
$\mu_{\hat{p}} = p$	$\mu_{\bar{x}} = \mu$
$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

- **REMEMBER:** These formulas are for **CONVERSION** from the population standard deviation! If you're already given the standard deviation of the sampling distribution, just use it!

4. Know the **CONDITIONS** that must be met for the Central Limit Theorem to apply, and thus for **inference** to occur

Condition	*How to meet the condition*	*Ensures* shape *of the* ***sampling distribution*** *is appropriate for inference (center, shape, or spread)*
1. Large Enough (Normal)	**For Proportions:** $n\pi \geq 10$ and $n(1-\pi) \geq 10$ **For Means:** Parent Pop is normal or $n \geq 30$ or graphical display shows no outliers + if symm.	***NOTE:*** *If the* *population* *has an approximately normal distribution, this condition can be considered "met"* ***regardless*** *of sample size!*
2. SRS	State that the sample was randomly selected or that treatments were randomly assigned →	don't just say "✓"
3. Independence for proportions $20n < N$		
σ is ?? for means	σ known use z σ unknown use t	spread

What must we do if the conditions are not met? Stop or tell reader you are Continuing but results are only valid if conditions are met

PROCEDURES FOR CONFIDENCE INTERVALS AND SIGNIFICANCE TESTS (Chapter 9-11)

1. State what you're doing

Confidence Intervals	Significance Tests
• Procedure you're using • The *parameter* (population) *of interest!* • Confidence level **"We will use a __________ Interval to estimate, with _____% confidence, the *true* (mean/proportion) of ____*(context)*____"**	• Procedure you're using • The *parameter* (population) *of interest!* • Hypotheses, H_0 and H_a • Significance Level, α (If none is given, use .05) **"We will use a __________ Test to test the following hypotheses at the α = _____ level"**

Additional Notes:

- Remember, H_0 implies "no change" or "no difference"
- If you are doing a 2-Sample or 2-Proportion test, state **both** populations – indicate which one is which!
- For a **Paired** t-test, find the *difference* between the matched pairs, and use these *differences* as your one sample! **H_0:** $\mu_{Difference} = 0$, **H_a:** $\mu_{Difference}$ is >, <, or $\neq 0$

2. Check your conditions

NOTE: *If a problem says* ***"assume conditions are met"****, you* do not *have to go through this process!!*

- Sample Size (also known as "Large Counts")
  - **If met, the SHAPE of the sampling distribution is Normal (or χ^2 distribution for χ^2 tests)**
  - *Means (μ):*
    - 30 or more, OR
    - Graph of the sample shows no obvious skews or outliers (**t-test only**), OR
    - Population is *known* to be normal
  - *Proportions (p):*
    - At least 10 expected successes and 10 expected failures (find *expected* value of each)

- Randomness
  - **Ensures that the CENTER (the <u>sample statistic</u>) is legitimate**
  - *Samples and Observational Studies:* Randomly selected from the population
  - *Experiments:* Randomly assigned into treatment or control group(s)
  - ***Note:*** *If you are running a 2-sample interval or test, you must <u>check</u> and <u>STATE</u> that <u>both</u> samples are random!*
- Independence
  - **Ensures that the SPREAD (the <u>standard deviation</u>) formulas that you're given are reliable**
  - *Samples and Observational Studies:* sample must be *less* than 10% of the population
  - *Experiments:* Groups should be independent of each other (i.e. not matched pairs)
    - If there ARE matched pairs, do a PAIRED t-test; find the *difference* between each pair and use *those* numbers in a 1-sample t-test!

3. Do the calculation (create the interval or run the test)

Confidence Intervals	*Significance Tests*
• Re-state *type* and *confidence level* (just to be safe) • Give interval: (lower, upper)	• Test Statistic (z, t, or χ^2) • Degrees of Freedom (t and χ^2 ONLY) • *p-value*

4. State your conclusion

Confidence Intervals	*Significance Tests*
• Give the % confidence • Give the interval *in context* (including **PROPER UNITS**) **"I am _______% confident that the *true* mean (or *true* proportion) of ___*(context)*____ is between _____ and _____."** The true mean value is between (low) and (up) and _% of all intervals created the same way will contain the true mean CONTEXT!!!	• State whether $p < \alpha$ (reject) or $p > \alpha$ (fail to reject) • Give the consequences *in context* • **Chi-Squared:** You may be asked to perform a follow-up analysis to see where the biggest gaps between observed and expected values are. REJECT: **"Because $p < \alpha$, we can reject H_0. There is statistically significant evidence to suggest _____*(whatever H_a was)*_____** FAIL TO REJECT: **"Because $p > \alpha$, we fail to reject H_0. There is NO statistically significant evidence to suggest _____*(whatever H_a was)*_____**

IMPORTANT: The p-value is **ALWAYS** between __0__ and __1__. If you calculator gives something *other* than this, I <u>*guarantee*</u> there will be an E at the end. This represents *scientific notation* (# • 10^x). This means your p-value is **very small** (in fact, many statisticians just write "p < .001" and call it a day). **As far as we're concerned, p-values this low will <u>always</u> be significant!**

ALSO IMPORTANT: Know the difference between "interpret the p-value" and "draw conclusions"

- **Interpretation:** IF H_0 is true, the probability that we would get a test statistic as or more extreme as the one we got in our sample (by random chance) is __*(p-value)*__
  - **NOTE:** If there is a *direction* involved (< or >), <u>*state*</u> that direction ("as high or higher" or "as low or lower")

- **Draw conclusions:** Rejecting or Failing to Reject H_0 (and associated context)

Topic 9: Confidence Intervals (Notes: 9.1,9.2,9.3,11.1-11.3)

. Understand the purpose of confidence intervals and how they work

- What does a confidence interval allow us to do?

Estimate a population parameter with a certain degree of confidence

- How do we *interpret* a confidence interval? (For instance, to interpret 95% confidence level, what *words* would you say?)

The true parameter is between __ and __, and __% of all intervals created the same way will contain the true parameter

- How do we interpret a confidence *level*? (For instance, in a 95% confidence interval, what does the 95% tell us? What does it ***mean*** to be "95% confident"?)

It tells us the percent of time, in the long run, that the true parameter falls inside the interval created

- Know how to use the FORMULA for confidence interval:
  - Statistic ± Critical Value • Standard Deviation of Statistic
  - **Critical Values** can be found in the *t* table (for *z* distributions, use the ________ row)
  - **Standard Deviation:** Use the formula sheet (they are *very* clearly laid out!)
    - *In this context,* St. Dev. of the Sampling Distribution is also called **Standard Error**
- What is the margin of error, and how do we calculate it?

The difference btw. the sample statistic and the lower or upper bound of the interval.

ME = ~~[illegible]~~ (critical value)(std error)

2. Know what type of confidence interval to calculate, and *when* to calculate it

When estimating a **population proportion** $p \pm z^* \sqrt{\frac{p(1-p)}{n}}$	When estimating the ***difference*** **between two population proportions** $p_1 - p_2 \pm z^* \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
When estimating a **population mean** and the population standard deviation is *known* (**RARE**) $\bar{x} \pm z^* \frac{\sigma_x}{\sqrt{n}}$	When estimating the ***difference*** **between two population means** and the population standard deviations are *known* (**RARE**) $\bar{x}_1 - \bar{x}_2 \pm z^* \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Note: *The true name of this procedure and the calculator name are slightly different. Know* ***both!***	
When estimating a **population mean** and the population standard deviation is **NOT** known $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$ ***Note:*** *The true name of this procedure and the calculator name are slightly different. Know* ***both!***	When estimating the ***difference*** **between two population means** and the population standard deviations are **NOT** known $\bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

3. Know the essentials of the *t*-distribution

- When do we use it?

When pop. std. dev is unknown

- How do we calculate the *degrees of freedom* of a t-distribution?

$df = n-1$

4. Know the **general process** of statistical inference (in this case, creating a confidence interval)

1. check conditions
2. calculate critical value
3. construct interval
4. conclusion

5. Know how to *check conditions*

- What conditions must you check, and where in the study guide can you look to find them?

SRS, Large enough, Ind / σ

- If dealing with a *t*-distribution and your sample size is not 30 or more, what *other* methods can you use to check for normality? **Be specific!**

create a boxplot and look for symmetry and no outliers

6. Know how to *manipulate* confidence intervals

- Be able to solve for n or z^* *(or t^*)* (**NOTE:** In multiple choice, you can always *plug in* the choices!)
  - If a sample proportion is not given in this case, assume $p =$ __.5__ (this gives us the greatest margin of error to work with)
- Remember that the sample statistic ("point estimate") is in the __center__ of the confidence interval (and that the distance between the sample statistic and the ends of the confidence interval is the __margin of error__
- Know what happens to the margin of error (and thus *width* of the confidence interval) if we...
  - Increase sample size:
  - Decrease sample size:
  - Increase confidence level:
  - Decrease confidence level:
- If you adjust sample size, confidence interval changes by the **square root** of that amount (since n is inside the square root in all standard deviation formulas)
  - ***Example:*** *What will happen to the confidence interval if you multiply the sample size by 4?*

The interval width will be $\frac{1}{2}$ as wide

Topic 10: Significance Tests (Notes: Chapters 10 & 11)

1. Understand what significance tests are for and allow us to do

- What are the two types of hypotheses used in significance tests, and what *symbols* do we use to represent them?

Null – H_0

Alternative – H_a

- What is a *null hypothesis*, and what does the null hypothesis *always* assume to be true?
NOTE: The answer is *slightly* different for 1-sample and 2-sample tests – know **both!**

The null is what we assume to be true. We assume that the parameter of interest is equal to the hypothesized value

- What is an *alternative hypothesis?* What are the 3 types of alternative hypotheses you could have?
NOTE: The answer is *slightly* different for 1-sample and 2-sample tests – know **both!**

A competing claim. <, >, or ≠

- Significance levels (alpha-levels) determine the p-value below which a test's results should be considered significant. If no alpha level is given, it is a good *general* rule to use 0.05

2. Know what type of significance test to run, and *when* to run it

When testing a claim about a **population proportion** Large sample z test of proportion	When testing a claim about the ***difference*** **between two population proportions** 2 independent sample z test, difference of proportion
When testing a claim about a **population mean** and the population standard deviation is *known* **(RARE)** Large sample z test of mean	When testing a claim about the ***difference*** **between two population means** and the population standard deviations are *known* **(RARE)** 2 independent sample z test, difference of means
When testing a claim about a **population mean** and the population standard deviation is **NOT** known Large sample t test of mean	When testing a claim about the ***difference*** **between two population means** and the population standard deviations are **NOT** known 2 independent sample t test, difference of means
When testing a claim about a study or experiment that utilizes *matched pairs* 2 dependent sample t test difference of means	*In the calculator, which type of test would you select?* 1 sample t test, using the list of differences

3. Know how to *interpret* the results of a significance test

- What two (for t-tests, three) things should you report after running a significance test in your calculator?

  test stat, p-value, df

- How do you *interpret* a p-value? What does that number *mean*?

  If p-value is < α then the Ho is rejected
  p-value is the probability of getting a sample this rare or more rare assuming the null is true

- How do you analyze (interpret the results of) a test for which the p-value is *less* than alpha (for instance, p < .05). *What would you write?*

  Reject the Ho, There is sufficient evidence to support the ~~claim that~~ Alternative Hyp.

- How do you analyze (interpret the results of) a test for which the p-value is *greater* than alpha (for instance, p > .05). *What would you write?*

  Fail to Reject the Ho, There is not sufficient evidence to support the alternative hyp.

Topic 11: Chi-Squared Tests and Types of Error (Notes: Chapter 12)

1. Know the similarities and differences between Chi-squared and the other types of significance tests

- When do we use Chi-squared tests? In other words, what do Chi-squared tests allow us to measure?

When analyzing the dist. of a categorical variable.

- What are the three types of chi-squared tests, and when do we use each?

Type	Purpose/When to use	Name in Calculator
χ^2 GOF test	One sample, comparing the dist. of a categorical variable to an expected dist.	χ^2-GOF
χ^2 test of Homogeniety	two samples or more, comparing the distribution of a categorical variable to two or more populations	χ^2-Test
χ^2 test of Independence	One sample, two categorical variables, checking to see if relationship or association b/tw the two variables	χ^2-Test

NOTE: *The biggest difference between the second and third type is* *context.* *Other than that, they are essentially the same.*

- What are the null and alternative hypotheses of a Chi-squared test?

H_0: The distribution is as expected
H_a: The distribution is not as expected

2. Know the *conditions* of a Chi-Squared test

- Same conditions as other significance tests
- How is the *sample size* condition different for Chi-Squared tests, and how do we check it?

SRS &
Large Enough: all expected cells ≥ 5
Independence of samples and subjects

3. Know how to calculate and interpret the Chi-squared statistic

- How can we find *expected counts?*
  - Goodness-of-fit: **READ THE PROBLEM!**
    - Sometimes, you may *expect* certain proportions out of a total (like we did with M&Ms).
    - Sometimes, you may *expect* that the data is *equally distributed* among the categories (in this case, just use simple division!)
  - Homogeneity and Independence: *What formula do we use to calculate each expected value?*

$$\text{Exp} = \frac{\text{row total} \cdot \text{col total}}{\text{grand total}}$$

- How do we calculate *degrees of freedom* for a chi-squared test?
  - Goodness-of-Fit: # cat − 1
  - Homogeneity and Independence: (row − 1)(col − 1)

- When running a Chi-Squared test, what three things must you report? ***NOTE:*** *The interpretation and analysis/drawing conclusions aspects of these are the same as the other significance tests.*

χ^2, p-value, df

---

4. Know what Type I and Type II error are; be able to spot them in context, *and* discuss what the *consequences* of these types of error would be if they happened in a real-life situation (including possibly evaluating which one would be worse in that situation) (Notes for 10.2)

HINT: *The chart on your 5.4 notes may be a handy tool to help you understand and remember which is which!*

- What is a Type I error?

Rejecting the Null when the Null is true

- What is a Type II error?

Failing to reject the null when the null was false

- What variables are used to represent the probability that Type I error and Type II error, respectively, will happen?

$P(\text{type I}) = \alpha$

$P(\text{type II}) = \beta$

5. Know what *power* is, why it's important, and how it can be influenced. (Activity for 10.5)

- What is the definition of *power*?

  The probability of rejecting a false null

- How is power calculated?

  Power = $1 - \beta$

- How can power be *increased*? List 3 ways.

  ① Increase sample size ← ↙ researcher control
  ② Increase α (significance level))
  ③ Increase distance from hyp. value to actual value ← can't be controlled

6. Understand the relationship between Power, Type I Error, and Type II error

Power	Type I Error (α)	Type II Error (β)
Increases ⬆	↑	↓
Decreases ⇩	↓	↑

Fill in each of the following blanks with either "same" or "opposite"

Type I and Type II error always go the opposite direction

Power and Type I error always go the same direction

Power and Type II error always go the opposite direction

Suppose you want to avoid a *Type I* error at all costs. Should you use a significance level of .10, .05, or .01? Explain.

.01

Topic 11: Bivariate Data (Notes: Chapter 5)

1. Know how to analyze a correlation between two variables

- Explanatory and Response variables (which one is x and which one is y?)

  x is the explanatory
  y is the response

- 5 things we should look for in bivariate data:

Characteristic	*Possibilities*	*What the **r-value** tells us*
Shape	are the ordered pairs have a linear shape	R-value assumes that shape is… linear
Strength	are the points close to the LSRL	r is close to 1 or −1
Direction	Are the points inc from left to right or are they dec. ??	+ if inc − if dec
Outliers (especially if they substantially alters the equation of the *regression line,* or line of best fit)		
Context (as always) – what two variables are we examining?		

- X and Y are correlated. Does this mean that X *causes* Y?

  No correlation does not imply causation

2. Know how to analyze the least-squares regression line (line of best fit): $\hat{y} = mx + b$

- $\hat{y}$ is the predicted value value of y for a given value of x
- *Interpretation* of Slope:

  The amount we would expect, on average, change in y for every 1 unit increase in x → remember to add contex + units

- *Interpretation* of Y-intercept:

- r^2 value **("coefficient of determination")**

  The percent of variation in y that is due to the linear relationship btw x and y. → remember context

- Extrapolation

  only should predict values for x's that are with the domain used to create the LSRL

3. Know how to analyze *residuals* and *residual* plot

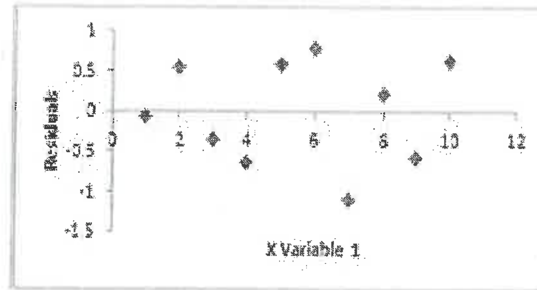
- What *is* a residual?

The diff. btw the actual and predicted values
Resid = $y - \hat{y}$

- How do you calculate a residual?

Resid = $y - \hat{y}$

Residual Plot



- What information does a residual plot give you?

If there is a relationship btw x's and residuals

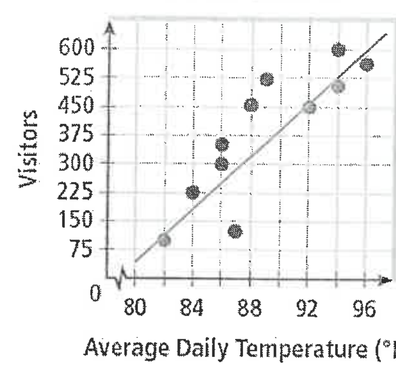
4. Know how to handle *curved* data (linear transformations)

- *Be aware that one* **or both** *variables may be transformed* (square root, log, natural log (ln), etc.) *in order to* *linearize curved data*
- Make sure that all **interpretations** (see above) take all transformations into account!

Example

Analyze the correlation shown r = .85

There is a strong positive linear relationship btw temp and # of beach visitors



b. Give **and interpret** the value of the slope of the regression line

b = 32.23

We would expect, on average, a 32.23 person increase in # of visitors to beach for every 1 degree increase in temp.

Predictor	Coef	SE Coef	t	P
Constant	-2486.13	96.84	-2.11	.03
Temp	32.23	15.3	4.76	.000
r = .85		$r^2 = .72$		

c. Give **and interpret** the value of the y-intercept of the regression line

d. Give **and interpret** the r^2 value of the regression line

72% of the variation in the number of visitors is due to the linear relationship btw # of visitors and temperature

e. If tomorrow's temperature is going to be 90°, predict how many visitors the beach will have tomorrow. **Show work!**

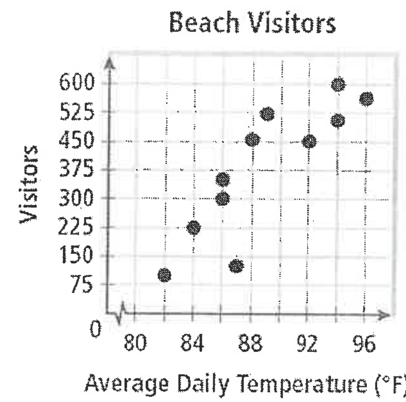
$\hat{y} = -2486.13 + 32.23x$

$\hat{y} = -2486.13 + 32.23(90) = 414.57$ visitors

Topic 13: Confidence Intervals and Significance Tests with Bivariate Data (Notes: Chapter 13)

- A regression line is created using a bivariate set of data. Confidence intervals and significance tests allow us to predict and test the amount slope of the relationship between the explanatory and response variables (x and y)
  - You can also do this for y-intercept, but this is not something to worry about for the exam
- The AP exam will most likely ask you to use a Computer out put to make inference
  - Remember, everything dealing with slope is in the row with the variable name ("constant" refers to the **y-intercept**)
  - If you need to do them in the calculator…
    - 1. Put all Xs in one list and Ys in another list
    - 2. Go to **LinRegInterval** or **LinRegT-Test,** type in the inputs, and get your results!
- Confidence Interval
  - Confidence interval = Statistic ± Critical Value • Standard Deviation of Statistic
  - For a linear regression, this becomes $b \pm t^* S_b$
    - "SE Coef" can be found in the Computer output col. 3
    - t^* can be found in your calculator
      - For **degrees of freedom (DF),** use $n-2$
  - Interpretation (assuming 95% confidence)
    - **I am 95% confident that the slope of the *true* regression line of the relationship between __x__ and __y__ is between ______ and ______.**
- Significance Test
  - Ho: Assume that there is no relationship between the variables (this means **slope (β)** = 0)
  - Ha can be <, >, or ≠ (just like before)
  - t and p can *both* be found in the row of slope. Interpret as usual!
    - The **formula** for the test statistic is $t = \frac{b-\beta}{S_b}$

- Conditions!! (Use the acronym **LINEaR**)
  - L: The dist. of e at any x is zero. That is $\mu_e = 0$
  - I: The std. dev of e at any x is the same (or use ______)
  - N: The dist of e at any x is approx. normal
  - E: The e's associated with diff observations are independent (can think of this as ______)
  - and
  - R: andom sample



Example

The Florida Tourism Department is studying the habits of beachgoers across the state. They observe a certain beach on 11 randomly-selected days during the peak season (May thru August) and record the Average Daily Temperature and the number of visitors who come to the beach that day. A scatterplot of the data is shown, as is a computer output of the data. Assume that temperatures and number of visitors are both normally distributed.

Predictor	Coef	SE Coef	*T*	*P*
Constant	-2486.13	96.84	-2.11	.03
Temp	32.23	6.76	4.76	.000
r = .85		r^2 = .72		

a. CHECK conditions for inference:

A linear model seems appropriate so all 4 basic assumptions are met
The sample was randomly selected

b. Construct and interpret a 95% confidence interval of the slope of this regression line

$t^* = \text{invt}(.025, 9) = 2.262$

$b \pm t^* s_b$

$32.23 \pm 2.262(6.76)$

$(16.94, 47.52)$

The true slope of the relationship btw # of visitors and temp is btw 16.94 ppl/deg and 47.52 ppl/deg. And 95% of all intervals created the same way will contain the true slope.

c. Is there significant evidence at the α = .05 level to suggest that there is a relationship between average daily temperature and number of visitors?

Yes, p-value for slope is approx. zero, therefore we would reject the null in favor of there being a relationship between # of visitors and temperature

Additional Topics

1. Know the **SYMBOLS** for parameters and statistics. **Mis-using a symbol WILL cause you to get docked on the exam!!**

Measure	*Parameter Symbol* (Population)	*Statistic Symbol* (Sample)
Mean	μ	$\bar{x}$
Standard Deviation *(also applies to Variance)*	σ	s
Proportion	p or π	$\hat{p}$ or p
Sample size	n	

2. Know how to work with *percentiles* ("relative frequency")

- A data point's percentile tells the percentage of the data that is less than or equal to that data point
  - **Example:** If you're in the 85th percentile, 85% of the population is at or below your level
  - This means **Q1** is the 25th percentile, **Median** is the 50th, and **Q3** is the 75th
- The numbers in the z-table can be considered *percentiles* (for instance, the z-score 0.45 corresponds with .6736 in the z-table, which is the 67th percentile)

AP EXAM ADVICE

General advice for ALL your exams:

- **Be prepared**
  - Have your pencils and materials ready to go
  - Get a good night's sleep! (This will feel strange to some of you)
  - **Be on time.** You WILL NOT be admitted to the testing room if you are late.
  - *Leave the personal drama at the door.* Do not let it bring you down on an exam this important!

- *Don't try and do too much!* I have seen many students write great answers, only to get docked because they added an incorrect piece of information or tried to make a claim that wasn't there. *Answer the question as **fully yet concisely** as possible, and then get out!*

- Read each question **VERY** carefully! AP loves to throw curveballs and you need to be sure of what the question is asking you to do!

- TIME IS OF THE ESSENCE. If you are stuck on a question, **OR** you know that question may take a while to figure out, *come back to it.* Knock out the easier ones first.

- **Two minute warning** is the best time to start guessing (*especially* on Multiple Choice).
  - The WORST answer you can possibly have is a blank!

Specific advice for THIS exam:

- TIMING:
  - 2 minutes and 15 seconds for each multiple choice
  - 13 minutes for Free Response #1 – 5
  - 25 minutes for Free Response #6
  - *Some questions will take more or less than this. That's fine. Just pace yourself!*

- **Calculator Check!**
  - Is it charged and/or have working batteries?
    - If your TI-84 is okay at the start of the test but then says "low battery" in the middle of the test, it will last through the duration of the test. DO NOT WORRY!

- **Show work!** You HAVE to show enough to prove to the AP Readers that you understand the *process* behind your answers (you WILL get docked for not showing enough work.)
  - It doesn't matter *how* simple the calculation is. If it's 1+1 = 2, **write that down.**

- **Formula sheet** is your friend! ***Especially*** the 2nd and 3rd pages (as well as the **t-table** because it gives you all the *critical values* you could ever want!). Sometimes the formula sheet gives away an otherwise tricky answer.
  - But be careful: do not, and I repeat, **DO NOT TEAR OUT THE FORMULA SHEET FROM THE TEST BOOKLET. *THIS WILL INVALIDATE YOUR EXAM.*** This happened to someone I knew on the AP Chem Exam; her score was invalidated and she had to take the test again next year.

- If you need to make a graph, **LABEL YOUR AXES!!**
  - If you're doing it to check the Normality (Sample Size) condition for inference, make sure you **write** whether you see any skews or outliers. **Just showing the graph is not enough** (but don't *forget* to put the graph, either! You need BOTH the graph AND the analysis of skew/outliers)
    - Remember that *boxplots* are the most efficient (but not the *only*) way of checking for this!

- **Watch your language!** Words like *average, range, skew*, and *significant* have very specific meanings in statistics, so DO NOT use these words unless you are using them in the correct *statistical* context (otherwise, **find synonyms)**
  - Average → Typical
  - "Ranges from" → "Goes from"
  - Skews → Distorts
  - Significant → Substantial
  - ***NOTE:*** *It is okay to use these words for their statistical definitions. Just use synonyms if you're going to venture outside of that.*
  - **If you aren't sure what a word means, avoid using it!!**

- *Stick to the script!* Know how to phrase your analyses of the following (*they are in your study guide*). **These phrasings help ensure you have covered all important aspects of the analysis in a clear and concise manner!**
  - Confidence intervals
  - Confidence *levels*
  - *Interpreting* p-values
  - *Analyzing* or *drawing conclusions* about p-values
  - Slope of a regression line
  - Interpreting r^2

- **Randomization** and a **large sample size** can solve most of life's problems – they make for better, more accurate, and more reliable (unbiased) results

- DO NOT mix up the language of *sampling* and the language of *experiments.*
  - For example, subjects of experiments are usually not randomly selected (often times that's *highly* unethical). They *are,* however, randomly *assigned* to groups (at least they *should* be)

- If you use symbols, **DEFINE** what that symbol means. OR you can weave the context *into* your symbol
  - *Both ways are acceptable* (although one is definitely **quicker!**)

Symbols with definitions	**Symbols with context *interwoven***
P(A ∩ B), where *A* represents being a girl and *B* represents being a senior	P(Girl ∩ Senior)
$\mu = 23$, where μ represents the mean weight of the *population* of piglets (or *true* mean weight of piglets)	$\mu_{\text{piglets}} = 23$
$p_1 > p_2$, where p_1 represents the *true* proportion of adults who like snacks, and p_2 represents the *true* proportion of children who like snacks	$p_{\text{adults}} > p_{\text{children}}$

- For **sampling distributions,** make sure you use $\mu_{\bar{x}}$ (or $\mu_{\hat{p}}$) for mean and $\sigma_{\bar{x}}$ (or $\sigma_{\hat{p}}$) for standard deviation
  - **<u>IF YOU DON'T KNOW WHAT SYMBOL TO USE, DON'T USE A SYMBOL AT ALL!!</u>** There's nothing wrong with writing out an answer in words. An incorrect symbol **WILL** get you docked.

- For inference problems (confidence intervals and significance test), **LOOK** for the statement "assume all conditions are met". **If it is not there, you had <u>better</u> check those conditions!**
  - Also be on the eye out for *randomness* – is it stated? And for 2-sample problems, is it stated for *both* samples?

- If you're doing an interval or test, always provide the **name** of the procedure when you do it!

- Remember, **NEVER** claim H_0 or H_a are "true" or "false". We "reject" or "fail to reject" based on the *probability* of getting a certain result by chance (that's what significance tests are all about!) and we *know* that probability is NEVER a guarantee!

- **BREATHE!!** We've been working for this all year. *You've got this!* One wrong answer won't kill you. Heck, just getting half of the questions right is *almost guaranteed* to be a 3! Don't overthink – just do your best.

<u>GOOD LUCK!!</u>