AP Statistics

Module 1 Free Response and Essay Tips

Below you will find a breakdown of different AP topics for this module. The sections include expectations for answering questions over each topic and examples of how these areas should be handled:

DATA & GRAPHS

	Notes
Qualitative Data (Categorical)	*Data consists of categories* **Ordinal:** Has no natural order [i.e. hair color: {brown, black, blonde}] **Nominal:** Has a natural order to it [i.e. level in high school: {freshman, sophomore, junior, senior}]
	Qualitative Data could be graphed using a bar chart or a pie chart
Quantitative Data (Numerical)	*Data consists of counts or measurements* **Discrete:** Data is only integers **Continuous:** Data could be decimals
	***Quantitative Data could be graphed using a histogram, dotplot, stem and leaf plot, cumulative frequency plot ***

DESCRIBING DISTRIBUTION – "SOCS"

Shape	1) **Unimodal/Symmetric:** Describes a distribution that has one mode, generally mound-shape (or bell curve) 2) **Bimodal:** more than one peak 3) **Uniform:** roughly flat, all bars approximately the same height 4) **Skewed Left:** tail to the left 5) **Skewed Right:** tail to the right
Outliers	Extreme values that don't appear to belong with the rest of the data Tukey's Rule: observations below Q1 – (1.5*IQR) and values above Q3 +(1.5*IQR) will be outliers ***You <u>must</u> show all work in how you calculate that there are/ aren't outliers for a set of data.***
Center	What number "typifies" the data? a. Use median for skewed data or data with outliers (Q2 or M) since it is resistant to skewness and outliers. b. Use mean (x bar)for normal, unimodal data If distribution is perfectly symmetrical, mean and median will be the same. Explain why you are using the mean/median as the measure of center. Mean < median when it is skewed left Mean > median when it is skewed right.

AP Statistics
Module 1 Free Response and Essay Tips

Spread	Give description/range "about # - #" Quartiles and Interquartile Range Standard Deviation and Variance 5 Number Summary Choose most appropriate measure of spread for the context ***When data is symmetric or uniform, you should be giving the standard deviation as the measure of spread. When the data is skewed or has outliers, you should be giving the IQR as the measure of spread.***
Comparing Distributions of two sets of data	a. Describe relevant differences in their distribution b. Substantiate those differences with numerical/graphical evidence c. Make appropriate conclusions in the context of the problem based on your observations ***Make sure you always compare the SOCS of both distributions***

Normal Distributions

Standardized Values	$z=\frac{x-\mu}{\sigma}$ ***We standardize values to compare different values even if they have different units and are on different scales.*** ***Z-score tells us how many standard deviations a value is from the mean***
Density Curves	Area under the curve is 1, represents 100% of the observations
Empirical Rule	68-95-99.7 Rule 68% of observations fall within 1 σ of the mean μ 95% of observations fall within 2 σ of the mean μ 99.7% of observations fall within 3 σ of the mean μ
Normalcdf	Normalcdf(lower limit, upper limit, mean, standard deviation) ***Use this when we have the z value and we are looking for a percentage/probability***
InvNorm	InvNorm(percentile, mean, standard deviation) ***Use this when we have the percentage/probability and we are looking for the z value***

AP Statistics

Module 1 Free Response and Essay Tips

Cumulative frequency plots & Relative frequency plots	*For cumulative frequency plots, your y-axis should be the number of your sample size. You should be graphing the cumulative amounts for each bar (the amount in that bar and the total amount for the bars below it). The last bar in the histogram should be the total number in your sample size. *For relative frequency plots, you should have the percentile on the y-axis (up to 1 or 100%). Each point in the plot should be (value, percentile). To get the percentile, take the number of sample units that falls in that bin and divide by the total sample size.
Normal probability plots	*Used to assess the normality of a data set. To see if the data is approximately normal, the plot should appear linear. A curved plot shows skewness.

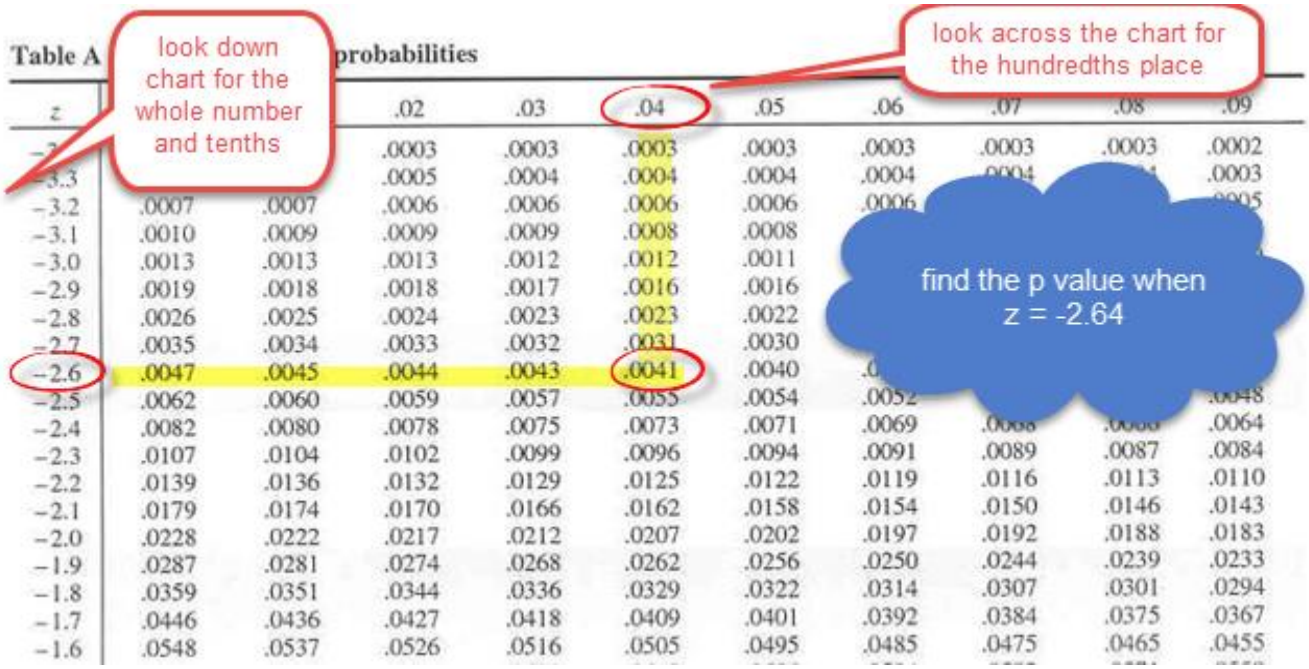

Table A probabilities

z	.02	.03	.04	.05	.06	.07	.08	.09
		.0003	.0003	.0003	.0003	.0003	.0003	.0002
		.0005	.0004	.0004	.0004	.0004		.0003
−3.2	.0007	.0006	.0006	.0006	.0006	.0006		
−3.1	.0009	.0009	.0009	.0008	.0008			
−3.0	.0013	.0013	.0012	.0012	.0011			
−2.9	.0018	.0018	.0017	.0016	.0016			
−2.8	.0025	.0024	.0023	.0023	.0022			
−2.7	.0034	.0033	.0032	.0031	.0030			
−2.6	.0045	.0044	.0043	.0041	.0040			
−2.5	.0060	.0059	.0057	.0055	.0054	.0052		.0048
−2.4	.0080	.0078	.0075	.0073	.0071	.0069		.0064
−2.3	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455