

# AP Statistics

## Module 8 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:

### BASICS FOR $\chi^2$ TESTS

Actual AP Exam Expectations	Notes
1. Verify all conditions	<p>The sample is a random sample.            *State expected counts*            All expected counts are at least 1.            No more than 20% of expected counts are less than 5.</p>
2. Use the correct distribution	<p><b>Goodness of Fit:</b>            Compares the observed distribution of a single categorical variable to an expected distribution for one sample. No two way table is given.</p> <p><b>Independence:</b>            Examines counts from a single sample for evidence of an association between two categorical variables. Uses a two way table.</p> <p><b>Homogeneity:</b>            Compares the distribution of two samples for one categorical variable. Uses a two way table.</p>

### TESTS OF SIGNIFICANCE

1. SHOW all steps of the test	<p>1) Parameter: <b>Chi Squared Goodness of Fit (GOF)</b>            "We want to test if <u>context of problem</u> are all equal in the population."  <b>Chi Squared test of Homogeneity</b> "We want to test if the <u>variable in context</u> are equal for <u>the two populations in context.</u>"  <b>Chi Squared test of Independence</b> "We want to test if there is evidence of association between <u>variable 1 and variable 2 in context of problem</u>"</p> <p>2) Conditions  <b>(see above BASICS #1 for specifics on each of these)</b>            *SRS            *Expected Counts</p> <p>Name the test: Chi Squared GOF, Homogeneity, or Independence and state the degrees of freedom</p>
-------------------------------	--

# AP Statistics

## Module 8 Free Response and Essay Tips

<p>SHOW all steps of the test</p>	<p>State the null (<math>H_0</math>) and alternative hypotheses (<math>H_a</math>) in context.  <b>GOF</b> <math>H_0</math>: The <u>variable proportions</u> are equal.  <math>H_a</math>: The <u>variable proportions</u> are not equal.</p> <p><b>Homogeneity</b> <math>H_0</math>: The proportion of <u>variable in context</u> is equal for every <u>context populations</u>.  <math>H_a</math>: For at least one of the <u>context populations</u>, the <u>variable in context</u> is different.</p> <p><b>Independence</b> <math>H_0</math>: The <u>variable 1 in context</u> is independent of <u>variable 2 in context</u>.  <math>H_a</math>: The <u>variable 1 in context</u> is not independent of the <u>variable 2 in context</u>.</p> <p>3) Show all work:          Write down ALL input information, and ALL output information</p> <p>4) Interpret the results in the context of the problem and make a connection to the given information. <b>Remember the 3 C's</b></p>
<p>2. GOF equation and degrees of freedom</p>	$\chi^2 = \sum \frac{(O - E)^2}{E}$ $df = n - 1$ <p>* n represents the number of categories</p>
<p>3. Independence and Homogeneity equation and degrees of freedom</p>	$\chi^2 = \sum \frac{(O - E)^2}{E}$ $df = (r - 1)(c - 1)$ <p>*r represents number of rows and c represents number of columns</p>
<p>4. Interpret the results of the test</p>	<p>Conclusions should be given in terms of the context of the question.</p> <p><b>GOF:</b> Reject or Fail to reject the null hypothesis that the <u>context of the problem</u> are not equal because the p-value is &lt; &gt; level of significance. There is/is not sufficient evidence to suggest that <u>restate the null hypothesis</u>.</p> <p><b>Homogeneity:</b> Reject or Fail to reject the null hypothesis that the <u>variable 1 in context</u> is independent of <u>variable 2 in context</u> because the p value is &gt; &lt; level of significance. There is/is not sufficient evidence to suggest that <u>restate the null hypothesis</u>.</p>

# AP Statistics

## Module 8 Free Response and Essay Tips

Interpret the results of the test	<p><b>Independence:</b> Reject or Fail to reject the null hypothesis that the proportion of <u>variable in context</u> is equal for every <u>context populations</u> because the p value is <math>&gt;</math> <math>&lt;</math> level of significance. There is/is not sufficient evidence to suggest that <u>_____ restate the null hypothesis _____</u>.</p>
-----------------------------------	--

- **P-value:** The probability of seeing a result from a random sample that is as extreme as or more extreme than the result you got from your random sample, if the null hypothesis is true.
- You can also find the p value, once you have the chi square value, by using  $2^{\text{nd}}$  -> Distr -> xcdf(lower, upper, df)
- The chi square distribution gives the probability to the right of the test statistic.
- As the degrees of freedom increase, the distribution becomes less skewed and more like a normal distribution.
- Sketch the curve and label the  $x^2$  value and shade the probability to the right

### TEST OF SIGNIFICANCE -- Linear Regression

Actual AP Exam Expectations	Notes
1. SHOW all steps of the test	<p>1. <b>Parameter:</b> Identify the x variable that we are analyzing as the sample statistic.</p> <p>2. <b>Conditions:</b></p> <ul style="list-style-type: none"> <li>*Normality – The values of the response variable y are normally distributed.</li> <li>*Independence – The values of the response variable y are independent.</li> <li>*Linear – There is a linear relationship between x and y represented by <math>\mu_y = \alpha + \beta x</math></li> </ul> <p>Name the test: Linear Regression T test</p>

SHOW all steps of the test	<p>State the null (<math>H_0</math>) and alternative hypotheses (<math>H_a</math>) in context.  <b>Note:</b> <math>H_0</math>: Let <math>\beta</math> = the true slope of the regression line for prediction <u>y variable</u> on <u>x variable</u> .  <math>\beta = 0</math> The slope of the true regression line is equal to zero.  <math>H_a</math>: <math>\beta &gt; &lt; \neq 0</math> The slope of the true regression line is ____ to zero.</p> <p>3. <b>Show all work:</b> ALL inputted information, and ALL output information</p> <p>4. <b>Conclusion:</b> Interpret the results in the context of the problem and make a connection to the given information.  *Reject or Fail to reject the null hypothesis that the slope of the true regression line for <u>y variable on x variable</u> is = 0 because <math>p &lt; &gt;</math> level of significance. There is/is not evidence that the true slope of the regression line is ____.</p>
2. Regression t equation	$t = \frac{b_1 - \beta}{s_{b_1}}$ <p>* <math>b_1</math> and <math>s_{b_1}</math> come from the sample</p>
3. Standard error of the slope	$s_{b_1} = \frac{s}{t}$ <p>* <math>b_1</math> is from the LSRL and t is from the calculator</p>
4. Degrees of freedom	$n - 2$ <p>*n is the number of points in the sample</p>

- Most linear regression inference questions will include a computer printout for you to read in order to get your statistics. Here is an example of what information you would need:

Successfully Completed Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	23.514	1.684	13.95	0.000
GPA	-2.7555	0.4668	-5.90	0.000

S = 0.5658      R-Sq = 76.0%

**CONFIDENCE INTERVAL -- Linear Regression**

Actual AP Exam Expectations	Notes
<p>1. SHOW all steps of the interval</p>	<p>1. <b>Parameter:</b> Identify the x variable that we are analyzing as the sample statistic.</p> <p>2. <b>Conditions: Same as Significance Test</b>            *Normality            *Independence            *Linear</p> <p>Name the interval: Linear Regression T Interval</p> <p>3. <b>Show all work:</b> ALL inputted information, and ALL output information</p> <p>4. <b>Conclusion:</b> Interpret the results in the context of the problem and make a connection to the given information.            *We are _____% confident that the slope of the true linear relationship between <u>y variable</u> and <u>x variable</u> is between <u>lower value</u> and <u>upper value</u></p>
<p>2. Regression t interval</p>	<p><math>b_1 \pm t^* s_{b_1}</math>    * <math>b_1</math> and <math>s_{b_1}</math> come from the sample</p>
<p>3. Find the <math>t^*</math> for the confidence interval</p>	<p>Using the calculator: Subtract the level of significance from 1. So <math>1-\alpha</math> and then divide that value by 2. On the calculator DISTR-&gt; INV T -&gt; AREA: <math>((1-\alpha)/2)</math> df: <math>n - 2</math> -&gt; ENTER This will give you the critical value <math>t^*</math></p>

Table entry for  $p$  is the point ( $\chi^2$ ) with probability  $p$  lying above it.

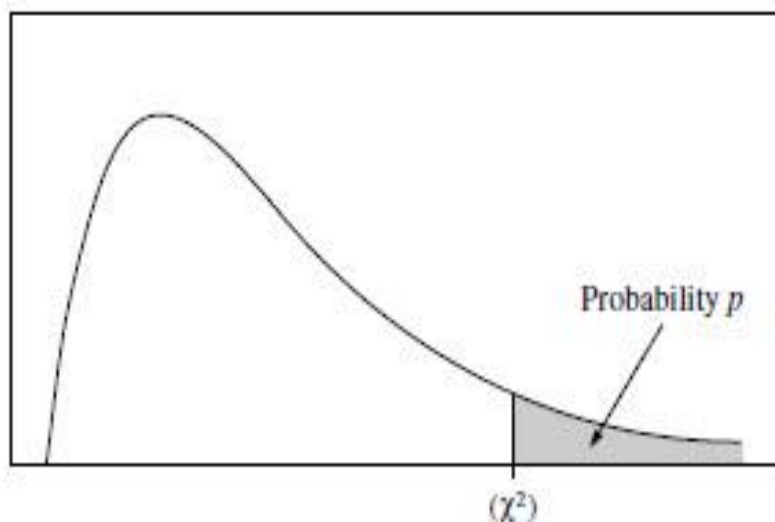


Table C  $\chi^2$  critical values

df	Tail probability $p$											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2