AP Statistics / Mr. Hansen Name: ___________________________
5/18/1999 (rev. 10/10/1999, 2/25/2000, April 2000, 2/26/02, 12/11/02, 3/11/03, 3/3/04, 3/5/05, 3/15/06, 4/15/06, 4/6/07, 4/9/08, 3/2/09, 3/20/010, 2/16/011, 4/27/22, 4/4/23)

TI-83/84 STAT TESTS Summary

#	Common name	Assumptions	Comments
~~1~~	~~1-sample *z* test for mean (or for mean diff. of **matched pairs**)~~	~~1. SRS~~ ~~2. normal pop.~~ ~~3. known sigma~~	• Rarely used, since a 1-sample *t* test is more realistic and accurate. (We never know σ in the real world.) • CLT lets us relax normality assumption for larger samples (see p.251 of Barron's, p.606 of Yates textbook, or bottom of tan boxes on p.553 and p.595 of Peck/Olsen/Devore textbook).
2	1-sample *t* test for mean (or for mean diff. of **matched pairs**)	1. SRS 2. normal pop. 3. unknown σ	• CLT lets us relax normality assumption for larger samples (see p.251 of Barron's, p.606 of Yates textbook, or bottom of tan boxes on p.553 and p.595 of Peck/Olsen/Devore textbook).
~~3~~	~~2-sample *z* test for difference of means~~	~~1. two indep. SRS's~~ ~~2. two normal pops.~~ ~~3. two known sigmas~~	[rare; use 2-sample *t* procedures instead]
4	2-sample *t* test for difference of means	1. two indep. SRS's 2. two normal pops. 3. unknown σ_1, σ_2	• CLT lets us relax normality assumptions when n_1 and n_2 are both ≥ 30, if no extreme skewness or extreme outliers are evident. See p.251 of Barron's, p.606 of Yates textbook, or bottom of tan boxes on p.553 and p.595 of Peck/Olsen/Devore textbook for details. Some textbooks say you can add sample sizes together to reach your target, but play it safe and check for both sample sizes at least 30. • Test is especially robust when $n_1 \approx n_2$. • **When prompted by calculator for "Pooled," always answer "No."** The pooled ("equal variances") method was formerly listed on the AP formula sheet, but since the College Board has dropped that formula, you should never use it. The pooled method unrealistically assumes $\sigma_1 = \sigma_2$, and the pooled method was already obsolete when the AP Statistics course was created in 1997. Since then, better calculators and software have made the pooled method completely pointless, and the College Board dropped the "pooled"/"equal variances" formula in 2019.
5	1-prop. *z* test (used for testing a single proportion against a benchmark)	1. SRS 2. $N \geq 10n$	• Need large pop. so that SRS resembles indep. trials (sampling w/ replacement) to justify binomial model.

		3. $np \geq 10$ 4. $nq \geq 10$	• Need *np* and *nq* (expected # of successes and failures) to be at least 10 as a rule of thumb so that the *z* approx. to binomial is reasonable. • We use binomial model as an approx., then *z* model as an approx. of that. Thus there are 2 approximations of the true sampling distribution. • Use hypothesized proportions (p_0 and q_0) to estimate *p* and *q* when checking assumptions 3-4 and when computing s.e. by the AP formula.
6	2-prop. *z* test (test for difference between 2 props.)	1. two <u>indep.</u> SRS's 2. $N_1 \geq 10n_1$ 3. $N_2 \geq 10n_2$ 4. $n_1 p_1 \geq 10$ 5. $n_1 q_1 \geq 10$ 6. $n_2 p_2 \geq 10$ 7. $n_2 q_2 \geq 10$	• Similar comments as for #5. Some textbooks give 5 as the number of expected successes and failures in each sample, but 10 is what our textbook uses. These are merely rules of thumb to ensure that the *z* approx. is reasonable. **Important:** We nearly always have H_0: $p_1 = p_2$. In that case, do not use $\hat{p}_1,\ \hat{p}_2,\ \hat{q}_1,\ \hat{q}_2$ when checking assumptions 4-7. This is a new rule that the AP graders started cracking down on in 2021. Instead, compute $\hat{p}_c = \dfrac{X_1 + X_2}{n_1 + n_2}$ and $\hat{q}_c = 1 - \hat{p}_c$, and use those two values in place of all the *p*'s and *q*'s when checking your assumptions. Think of $\hat{p}_c$ as the "combined estimate of sample proportion." In plain English: Add up all the successes in both samples, and divide by the sum of both sample sizes. You must also use that combined estimate of sample proportion with the AP formula sheet's "when $p_1 = p_2$ is assumed" formula for the standard error: $s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\hat{p}_c \hat{q}_c \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$. **Rare exception:** It is possible, but extremely unlikely, that H_0 could be different from H_0: $p_1 = p_2$. For example, you could have H_0: $p_1 = p_2 + .03$. Your TI calculator cannot handle that case automatically. You would have to compute the s.e. manually by using the other s.e. formula on the AP formula sheet, which coincidentally is what you would *always* use when computing a 2-prop. *z* confidence interval: $s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$.
~~7~~	~~1-sample *z* C.I. for mean (or for mean diff. of **matched pairs**)~~	~~Same as #1.~~	[rare; use 1-sample *t* procedures instead]
8	1-sample *t* C.I. for mean (or for mean diff. of **matched pairs**)	Same as #2.	

~~9~~	~~2-sample z C.I. for diff. of means~~	~~Same as #3.~~	[rare; use 2-sample t procedures instead]
0	2-sample t C.I. for diff. of means $(\mu_1 - \mu_2)$	Same as #4.	
A	1-prop. z C.I. for p	Same as #5, except as noted at right.	• This time, do not use the hypothesized values of p and q when checking assumptions and calculating s.e. Instead, use the observed values of $\hat{p}$ and $\hat{q}$.
B	2-prop. z C.I. for $p_1 - p_2$	Same as #6, except as noted at right.	• This time, always use the first s.e. formula on AP sheet, namely $s_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$. Do not use the second formula, and do not compute the "combined estimate for proportion," $\hat{p}_c$, since we are not hypothesizing that $p_1 = p_2$. Instead, use $\hat{p}_1$, $\hat{p}_2$, $\hat{q}_1$, and $\hat{q}_2$ to estimate p_1, p_2, q_1, and q_2 when checking assumptions and when computing the s.e.
C	χ^2 test for **independence** or for **homogeneity of proportions** (both tests are sometimes called "χ^2 matrix tests")	1. SRS or census (for independence test); multiple SRS's (one per column) for homogeneity of proportions test 2. all exp. counts $\geq$ 5	• Note that all cells must be **<u>counts</u>** even though H_0 is stated in terms of probabilities. • Mechanics are exactly the same for both matrix tests. Only the hypotheses change. <u>χ^2 test for independence:</u> H_0: hair color and personality type are indep. H_a: some assoc. exists betw. hair color & pers. type <u>χ^2 test for homogeneity of proportions:</u> H_0: all 4 schools have same distrib. of hair colors H_a: not all 4 schools have same distrib. of hr. colors • TI-83/84 computes each expected count cell by this formula: *cell* = *rowtotal* · *coltotal* / *grandtotal*. • Use df = (rows – 1)(cols – 1). • Expected counts are usually not integers. • Some textbooks have more complicated assumptions regarding expected counts. • Look at the cells whose contributions to χ^2 are the biggest when making inference about a rejected H_0.
D	χ^2 goodness-of-fit test (now available as part of the built-in TI-84 commands; formerly supported by CHISQGOF or CSDELUXE)	1. SRS 2. all exp. counts $\geq$ 5	• Note that all cells must be **<u>counts</u>** even though H_0 is stated in terms of probabilities. • Use df = (# of bins – 1). • Expected counts are usually not integers. • Some textbooks have more complicated assumptions regarding expected counts. • Look at the cells whose contributions to χ^2 are the biggest when making inference about a rejected H_0.
~~E~~	~~F test~~		[beyond the scope of our course]

F,G	linear regression *t* test, confidence interval for LSRL slope	Memory aid: **LINER.** **L**inear is the true fit (scatterplot looks linear, r^2 is close to 1, and resid. plot is random). **I**ndependent observations (independently generated, or independently chosen SRS with $N \geq 10n$). Note that most time-series data will *not* satisfy this requirement. **N**ormally distributed residuals for each *x*. **E**qual variance of the residuals: Variance of the residuals must be uniform for all *x*. **R**andom selection of points, or the points come from an experiment that involved random assignment of treatments.	• Use df = $n - 2$, where n = number of data points. • If the r^2 value is reasonably close to 1, you could theoretically satisfy both L (linear) and E (equal variance/homoscedasticity) assumptions by stating the r^2 value, commenting that it is good since it is close to 1, and sketching the scatterplot and resid. plot. A good resid. plot will show that the LSRL fit is valid and that the fit does not get noticeably worse as *x* changes. (Be sure to *write* these observations so that the AP graders know that you know what you are talking about.) However, beginning approximately 2021, the AP graders want to see the E (equal variance/homoscedasticity) assumption split out separately. • The N (normality) assumption cannot be verified, since you are never going to have an acceptable number of residuals *for each x*. Instead, we have to settle for looking at the residuals as a group to see if they seem normally distributed. On the AP exam, all you have to do to satisfy the N (normality) assumption is to show a histogram or stemplot of the residuals and assert that no gross departure from normality is visible. Or, you can sketch the NPP of the residuals (6th pictograph on 2nd STATPLOT Plot 1 menu). If the NPP shows a relatively straight line, you know that the residuals are approximately normal. A bend to the right (as you move your finger from left to right) means right skewness, and a bend to the left means left skewness. Sketch your NPP, and write a description of what the NPP is telling you. • The E (equal variances) assumption should be listed separately, beginning approximately 2021. This is a College Board thing, since if you were paying attention, satisfying the L (linearity) assumption and showing the resid. plot already covers the E step. The only good news is that in a 4/27/2022 Zoom call involving College Board–experienced administrators and AP teachers from around the country, the consensus was that any question on LSRL *t*-test or C.I. for LSRL slope will probably state, “Assume that the assumptions for inference have been met.” • Note the R (randomness) assumption. As always, some degree of random selection is required for valid inference. If the data points came from an experiment, you would not choose an SRS; you would use all of your data and say that the randomness assumption was satisfied by the design of the experiment, since if the experiment was any good, it surely used random assignment of treatments.
~~H~~	~~ANOVA (analysis of variance)~~		• This is an optional topic to cover after the AP exam. • ANOVA provides a way to check for a statistically significant difference among several means.