AP Stats Ch 26- Extra Practice **Name:** Solutions **Per:** ___

17. Montana. A 1992 poll conducted by the University of Montana classified respondents by gender and political party, as shown in the table. We wonder if there is evidence of an association between gender and party affiliation.

	Democrat	Republican	Independent
Male	36	45	24
Female	48	33	16

(are sample)

a) Is this a test of homogeneity or independence?
b) Write an appropriate hypothesis.
c) Are the conditions for inference satisfied?
d) Find the P-value for your test.
e) State a complete conclusion.

b) H_0: There is no association between party and gender.
H_A: There is an association between party and gender.

c) conditions
✓ — counts?
✓ — SRS? not stated but likely to be representative
✓ — expected counts [check [B] after) lowest = 19.2

d) Perform χ^2 test in TI-84 using obs data in [A]
$\chi^2 = 4.85$, p-value = .088, $df = 2$

e) with $\alpha = .05$, $p = .088$ is high, so we fail to reject H_0.
We do not have sufficient statistical evidence to conclude there is an association between party and gender

19. Montana revisited. The poll described in Exercise 17 also investigated the respondents' party affiliations based on what area of the state they lived in. Test an appropriate hypothesis about this table, and state your conclusions.

	Democrat	Republican	Independent
West	39	17	12
Northeast	15	30	12
Southeast	30	31	16

H_0: party is independent of region. (still a χ^2 test of independence)
H_A: party depends upon region.

conditions
✓ — counts?
✓ — SRS? (representative)
✓ — expected counts ≥ 5?
(check [B] after
lowest = 11.28

Perform a χ^2 test in TI-84 with obs counts in [A]
$\chi^2 = 13.849$
$p = .0078$
$df = 4$

with $\alpha = .05$, $p = .0078$ is low, so we reject H_0.
We do have sufficient statistical evidence to conclude that party depends upon region.

29. Race and education. Data from the U.S. Census Bureau show levels of education attained by age 30 for a sample of U.S. residents.

	Not HS Grad	HS Diploma	College Grad	Adv. Degree
White	810	6429	4725	1127
Black	263	1598	549	117
Hispanic	1031	1269	412	99
Other	66	341	305	197

Do these data highlight significant differences in education levels attained by these groups?

Problem states "a sample" so should be independence.

conditions: counts? ✓ SRS? assumed exp counts ≥ 5? yes ✓

$\chi^2 = 2815$

p-value = 0

H_0: level of education is independent of race.

H_A: level of education depends upon race.

With $\alpha = .05$, p-value is low so we reject H_0. There is strong evidence that education level depends upon race.

exp:
$$\begin{bmatrix} 1469 & 6524 & 4056 & 1042 \\ 283 & 1259 & 783 & 201 \\ 815 & 1401 & 879 & 224 \\ 102 & 453 & 282 & 72 \end{bmatrix}$$

res:
$$\begin{bmatrix} -14 & -1 & 11 & 3 \\ -1 & 10 & -8 & -6 \\ 40 & -4 & -16 & -8 \\ -4 & -5 & 1 & 15 \end{bmatrix}$$

Hispanics more likely not to graduate from HS, whites less likely not to graduate. 'Other' races have proportionally more adv. degrees.

10. Pi. Many people know the mathematical constant π is approximately 3.14. But that's not exact. To be more precise, here are 20 decimal places: 3.14159265358979323846. Still not exact, though. In fact, the actual value is irrational, a decimal that goes on forever without any repeating pattern. But notice that there are no 0's and only one 7 in the 20 decimal places above. Does that pattern persist, or do all the digits show up with equal frequency? The table shows the number of times each digit appears in the first million digits. Test the hypothesis that the digits 0 through 9 are uniformly distributed in the decimal representation of π.

The first million digits of π

Digit	Count
0	99959
1	99758
2	100026
3	100229
4	100230
5	100359
6	99548
7	99800
8	99985
9	100106

expected = $\frac{1{,}000{,}000}{10} = 100{,}000$

counts → L1 df = 9

100,000 → L2

$\chi^2 = 5.5$

$p = .7878$

H_0: digits equally likely

H_A: digits not equally likely

with $\alpha = .05$, $p = .7878$ is high so we fail to reject H_0.

there is no evidence ~~that digits~~ are not equally likely.

16. Cars. A random survey of autos parked in the student lot and the staff lot at a large university classified the brands by country of origin, as seen in the table. Are there differences in the national origins of cars driven by students and staff?

Origin	Driver: Student	Driver: Staff
American	107	105
European	33	12
Asian	55	47

E

115 96

24 21

55 46

residuals

-1.7 1.9

1.8 -2

0 0.1

a) Is this a test of independence or homogeneity?

b) Write appropriate hypotheses.

c) Check the necessary assumptions and conditions.

d) Find the P-value of your test.

e) State your conclusion and analysis.

homogeneity (2 populations, 1 variable)

H_0: The distribution of car origin is the same for students and staff.

H_A: The distribution of car origin is different for students and staff.

(c) counts? ✓ SRS/indep? ✓ random survey stated — expected > 5? yes (6 matrix after analysis)

d) data → [A] χ^2-Test

$\chi^2 = 7.828$

$p = .0199$

$df = 2$

e) For $\alpha = .05$, p-value of .02 is low, so we reject H_0.

This is evidence the distribution of car origin is not the same for students and staff.

Residuals show more European cars for students and less European cars for staff.