READ SECTION 12.1 (pages 675 - 688)

- TAKE YOUR OWN NOTES OR ANNOTATE HANDOUT WITH YOUR NOTES (NO NOTES WILL BE GIVEN)
- FOLLOW INSTRUCTION IN NOTES
- THEN DO PROBLEMS BELOW.

SECTION 12.1A Exercises

1. **Aw, nuts!** A company claims that each batch of ~~its deluxe mixed nuts contains 52% cashews~~, 27% almonds, 13% macadamia nuts, and 8% brazil nuts. To test this claim, a quality control inspector takes a random sample of 150 nuts from the latest batch. The one-way table below displays the sample data.

SEE... pg 679

Nut:	Cashew	Almond	Macadamia	Brazil
Count:	83	29	20	18

(a) State appropriate hypotheses for performing a test of the company's claim.

(b) Calculate the expected counts for each type of nut. Show your work.

3. **Aw, nuts!** Calculate the chi-square statistic for the data in Exercise 1. Show your work.

SEE pg 680

5. **Aw, nuts!** Refer to Exercises 1 and 3.

(a) Confirm that the expected counts are large enough to use a chi-square distribution. Which distribution (specify the degrees of freedom) should you use?

(b) Sketch a graph like Figure 11.4 (page 683) that shows the P-value.

(c) SEE pg 683 find the P-value. Then use your calculator's χ^2cdf command.

12.1 A HW

(A)

(1) PARAMETER: P_{nuts} = true population proportion of nuts

HYPOTHESIS: $H_0: P_{cashew} = .52$ $\quad P_{almond} = .27$

$P_{macadamia} = .13$ $\quad P_{brazil} = .08$

H_A: at least one of the p_i's is incorrect

(B)

NUT	# OBSERVED	% EXPECTED	# EXPECTED	$\frac{(O-E)^2}{E}$
Cashew	83	.52	78.0	0.3205
ALmond	29	.27	40.5	3.2654
Macadamia	20	.13	19.5	0.01282
Brazil	18	.08	12.0	3.0
Total	150	1.00	150	$\boxed{\Sigma x = 6.599}$
Take advantage of calculator	L1	L2	L3 = 150(L2)	$L4 = \frac{(L-L3)^2}{L3}$

1VARSTAT L4

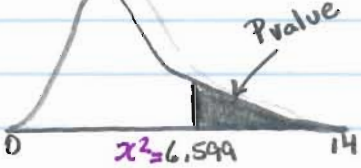
(3) $\chi^2 = .3205 + 3.2654 + .01282 + 3.0 = 6.599$

DO BY HAND - TIP USE LISTS IN CALC

$$\chi^2 = \frac{(83-78)^2}{78} + \frac{(29-40.5)^2}{40.5} + \frac{(20-19.5)^2}{19.5} + \frac{(18-12)^2}{12} = 6.599$$

(5) (a) The expected counts are all at least 5.
There are 4 categories - df = 3 for χ^2 distribution.

(b)



Chi-SQUARE distribution with 3 df

(c) $P(\chi^2 \geq 6.599) = .0858 \quad > .05$

$\chi^2cdf(6.599, E99, 3)$

Since th pvalue > .05, we fail to reject Ho. We do not have enough evidence to say the Companies claim about the distribution of nuts is wrong.

HW 12.1B Chi-Square Goodness-of-Fit Tests

7. **Birds in the trees** Researchers studied the behavior of birds that were searching for seeds and insects in an Oregon forest. In this forest, 54% of the trees were Douglas firs, 40% were ponderosa pines, and 6% were other types of trees. At a randomly selected time during the day, the researchers observed 156 red-breasted nuthatches: 70 were seen in Douglas firs, 79 in ponderosa pines, and 7 in other types of trees.[2] Do these data suggest that nuthatches prefer particular types of trees when they're searching for seeds and insects? Carry out a chi-square goodness-of-fit test to help answer this question.

9. **No chi-square** A school's principal wants to know if students spend about the same amount of time on homework each night of the week. She asks a random sample of 50 students to keep track of their homework time for a week. The following table displays the average amount of time (in minutes) students reported per night:

Night:	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Average time:	130	108	115	104	99	37	62

Explain carefully why it would not be appropriate to perform a chi-square goodness-of-fit test using these data.

χ^2 is not appropriate because the data collected is NOT counts but average amount of time spent on HW.

17. **Mendel and the peas** Gregor Mendel (1822–1884), an Austrian monk, is considered the father of genetics. Mendel studied the inheritance of various traits in pea plants. One such trait is whether the pea is smooth or wrinkled. Mendel predicted a ratio of 3 smooth peas for every 1 wrinkled pea. In one experiment, he observed 423 smooth and 133 wrinkled peas. The data were produced in such a way that the Random and Independent conditions are met. Carry out a chi-square goodness-of-fit test based on Mendel's prediction. What do you conclude?

pg 688

11. **Benford's law** Faked numbers in tax returns, invoices, or expense account claims often display patterns that aren't present in legitimate records. Some patterns are obvious and easily avoided by a clever crook. Others are more subtle. It is a striking fact that the first digits of numbers in legitimate records often follow a model known as Benford's law.[3] Call the first digit of a randomly chosen record X for short. Benford's law gives this probability model for X (note that a first digit can't be 0):

USE CALC

EXPECTED %

First digit (X):	1	2	3	4	5	6	7	8	9
Probability:	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046 — $\boxed{L1}$

A forensic accountant who is familiar with Benford's law inspects a random sample of 250 invoices from a company that is accused of committing fraud. The table below displays the sample data.

OBSERVED

First digit:	1	2	3	4	5	6	7	8	9
Count:	61	50	43	34	25	16	7	8	6 — $\boxed{L2}$

(a) Are these data inconsistent with Benford's law? Carry out an appropriate test at the $\alpha = 0.05$ level to support your answer. If you find a significant result, perform a follow-up analysis.

(b) Describe a Type I error and a Type II error in this setting, and give a possible consequence of each. Which do you think is more serious?

$L3 = L1 * 250$

12.1B

7

TREES IN FOREST	%	BIRDS OBSERVED	EXPECTED #	$\frac{(O-E)^2}{E}$
DOUGLAS FIRS	.54	70	84.24	2.4071
PINES	.40	79	62.40	4.416
OTHER TYPES	.06	7	9.36	0.595
	1.00	156	156	Σ 7.418 $= \chi^2$

USE L1, L2, L3 or do by hand

TEST: χ^2 GOODNESS OF FIT TEST FOR $\alpha = .05$

Hypothesis P_t = true proportion of trees in forest

H_0: $P_{Firs} = .54$ $P_{pines} = .40$ $P_{other} = .06$

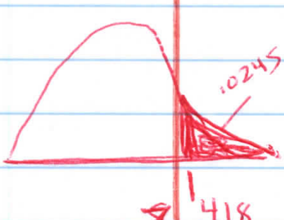
H_A: At least one of the P_t's is incorrect

CONDITIONS

Random - a random sample was used

Independent - reasonable 156(10) = 1,560 red breasted nut hatches

Large sample size - The expected counts in each category was greater than 5 (84.24, 62.4, 9.36)



MECHANICS $\chi^2 = 7.418$ df = 2

pvalue $\rightarrow P(\chi^2 \geq 7.418) = \chi^2cdf(7.418, E99, 2) = .0245$

Conclude: Since the pvalue (.0245) < .05, We Reject H_0, and conclude these birds prefer particular types of trees when they are searching for food.

12.1B HW

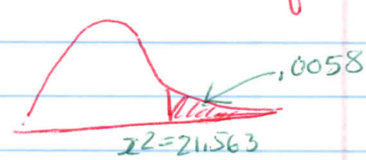
(9) See handout

(11) p_{digit} = true proportion of Benford's law digit

H_0: $p_1 = .301$ $p_3 = .125$ $p_5 = .079$ $p_7 = .058$ $p_9 = .046$
$p_2 = .176$ $p_4 = .097$ $p_6 = .067$ $p_8 = .051$

H_A: at least one of the p_{digits} is incorrect

STATE TEST: CHI SQUARE (χ^2) Goodness of fit test
$\alpha = .05$



CONDITIONS

Random - random sample of 250 invoices

Independent - reasonable their are 10(250) = 2500 invoices at the company

Large sample size - The expected counts are at least 5:

Must Give all expected Counts and round 2 decimals

75.25, 44, 31.25, 24.25, 19.75, 16.75, 14.50, 12.75, 11.5

MECHANICS:

$$\chi^2 = \sum \frac{(\text{observed} - \text{Expected})^2}{\text{Expected}} = \underbrace{\frac{(61-75.25)^2}{75.25} + \dots + \frac{(6-11.5)^2}{11.5}}_{\text{Can show 1st + last}}$$

$\chi^2 = 21.563$ df = 8

pvalue = $P(\chi^2 \geq 21.563) = \chi^2cdf(21.563, E99, 8) = .0058$

CONCLUDE: Since the pvalue is less than .05, we reject H_0 and conclude that the invoices are inconsistent with Benford's Law

#11 cont

(A) FOLLOW UP ANALYSIS:

DIGIT	OBSERVED		EXPECTED	x^2
1	61		75.25	2.7
2	50		44.00	0.8
3	43	>	31.25	4.4 *
4	34	>	24.25	3.9 *
5	25		19.75	1.4
6	16		16.75	0.03
7	7	<	14.50	3.9 *
8	8		12.75	1.8
9	6		11.5	2.6

Reviewing x^2 contribution- 3,4,7 have the largest contribution. Digits 3+4 have too many and Digit 7 has not enough.

11B TYPE I ERROR: SAYS THAT THE COMPANY'S INVOICES DID NOT FOLLOW BENFORD'S LAW (SUGGESTING FRAUD) WHEN IN FACT THEY WERE CONSISTENT WITH BENFORD'S LAW.

TYPE II ERROR: SAYS THAT THE INVOICES WERE CONSISTENTS WITH BENFORD'S LAW (SUGGESTING FRAUD) WHEN IN FACT THEY WERE NOT.

A TYPE I ERROR WOULD BE MORE SERIOUS HERE. ALLEGING THAT THE COMPANY HAD COMMITTED FRAUD WHEN IT HAD NOT

12.1 B HW

17

TEST: χ^2 goodness-of-fit test $\alpha = .05$

H_0: $P_{SMOOTH} = .75$ $P_{WRINKLED} = .25$

H_A: AT LEAST ONE OF THE P_i's is incorrect.

Conditions

Random and Independent Conditions were given

Large enough sample size -

The expected counts 417 and 139 are both greater than 5.

PEAS	%	OBS	EXPECTED	$\frac{(O-E)^2}{E}$
SMOOTH	.75	423	417	.0863
WRINKLED	.25	133	139	.2589
	1.00	n=556	556	.3452

Mechanics

$\chi^2 = .352$ $df = 1$

PVALUE $= P(\chi^2 \geq .352) = \chi cdf(.3452, E99, 1) = .5568$

Conclude:

Since the pvalue is very large and greater than .05, we fail to reject Ho, we do not have enough evidence to dispute Mendel's belief.