Name_____ AP Statistics Date_____

Notes – Simpson's Paradox

Simpson's Paradox refers to the reversal of the direction of a comparison or an association when data from several groups are combined to form a single group.

Example 1

Simpson's paradox is well known in statistics. It occurs when a small population having a high proportion of significant members is agglomerated with a large population that has a low proportion of significant members.

It is illustrated below by a constructed example that *prima facie* shows sexual discrimination in a college selection test.

College X offered 1000 places in two courses in the faculty of Economics. In year 2000AD there were 1000 female applicants, and 1000 male applicants. 181 females were accepted, and 819 males were accepted.

So although 50% of applicants were male and 50% were female, only 18.1% of females were accepted, but 81.9% of male applicants were accepted.

Was there a clear case of sexual discrimination in the selection process?

Not so!

The faculty of Economics at college X offered two courses. One was *Accountancy* and 900 places were offered. The other was *Ecology* and because the course was new, only 100 places were offered.

There were 950 male applicants for the Accountancy qualification, and 100 females. In Ecology there were 50 male applicants and 900 female applicants.

Acceptances	Places	Male Applicants	Male Acceptances	% Males Accepted	Female Applicants	Female Acceptances	% Females Accepted
Accountancy	900	950	814	86%	100	86	86%
Ecology	100	50	5	10%	900	95	11%
Total Faculty	1000	1000	819	81.9%	1000	181	18.1%

Despite the fact (row 3) that 81.9% of male applicants to Accountancy and Ecology were accepted,
and that only 18.1% of female applicants to Accountancy and Ecology were accepted,

there was no discrimination.

There was no discrimination in Accountancy. (row 1) because 814 of 950 males and 86 of 100 females obtained places. So 86% of male applicants students and 86% of females applicant students gained places in Accountancy.

There was no discrimination in Ecology,.(row 2) because 5 of 50 males and 95 of 900 females obtained places, so 10% of male applicants and 11% of females applicants for places in Ecology gained places.

Example 2

A curious thing happened to two baseball players this year during the first two weeks of the season. Here are the data:

	Player 1			Player 2		
Week	At Bats	Hits	BA	At Bats	Hits	BA
1	5	2		25	9	
2	20	5		5	1	

a. Show that for each week, Player 1 had a higher batting average (BA = Hits/At Bats) than does player 2.

b. Show that at the end of the two weeks, the cumulative results for Player 2 were better than the cumulative results for Player 1.

c. Explain the phenomenon. What is the lurking variable?