HW 5

26] A χ^2 variate with degrees of freedom equal to df has representation: $Z_1^2 + \ldots + Z_{df}^2$, where $Z_1, \ldots, Z_{df}$ are independent standard normal variates.

a) If Z has a standard normal distribution, what distribution does Z^2 have?

Z^2 has a χ^2 distribution with df degrees of freedom, where df is the number of variates/terms represented by Z

b) Show that, if Y_1 and Y_2 are independent χ^2 variates with degrees of freedom df_1 and df_2, then $Y_1 + Y_2$ has a χ^2 distribution with $df = df_1 + df_2$.

Y_1 is represented by: $Y_{11}^2 + Y_{12}^2 \ldots + Y_{1df_1}^2$

Y_2 is represented by: $Y_{21}^2 + Y_{22}^2 \ldots + Y_{2df_2}^2$

then

$Y_1 + Y_2$ is represented by: $Y_{11}^2 + Y_{21}^2 + Y_{12}^2 + Y_{22}^2 \ldots + Y_{1df_1}^2 + Y_{2df_2}^2$

therefore

if df_1 and df_2 represent the counts of Y_1 terms and Y_2 terms respectively, then the df for the χ^2 distribution of $Y_1 + Y_2$ will be equal to the total number of terms in the χ^2 distribution, i.e. $df = df_1 + df_2$

33] In murder trials in 20 Florida Counties during 1976 and 1977, the death penalty was given in 19 out of 151 cases in which a white killed a white, in 0 out of 9 cases in which a white killed a black, in 11 out of 63 cases when a black killed a white, and in 6 out of 103 cases when a black killed a black.

a) Exhibit the data in a 3-way contingency table

Murderer	Victim	Decision (DP Y/N) Y	N	Total	% Received DP
White	White	19	132	151	12.583%
	Black	0	9	9	0%
Black	White	11	52	63	17.46%
	Black	6	97	103	5.825%
Total	White	30	184	214	14.019%
	Black	6	106	112	5.357%

b) Construct the partial tables necessary to study the conditional association between defendant's race and death penalty verdict. Find and interpret the sample conditional odds ratios, adding 0.5 to each cell to reduce the impact of the 0 cell count.

X = Murderer's Race Y = Death Penalty Decision Z = Victim's Race

Z = White Victim

Murderer	Decision Y	N
White	19.5	132.5
Black	11.5	52.5

$$\hat{\theta}_{Z=\text{White}} = \frac{19.5 \cdot 52.5}{132.5 \cdot 11.5} = 0.67186$$

Z = Black Victim

Murderer	Decision Y	N
White	0.5	9.5
Black	6.5	97.5

$$\hat{\theta}_{Z=\text{Black}} = \frac{0.5 \cdot 97.5}{6.5 \cdot 9.5} = 0.78947$$

In both cases, the odds ratios show that the odds of a white murderer receiving the death penalty are lower than those of a black murderer receiving the death penalty, leading us to beleive that black murderers receive the death penalty more often.

c) Compute and interpret the sample marginal odds ratio between defendant's race and death penalty verdict. Do these exhibit Simpson's Paradox? Explain.

		Decision: Y	Decision: N
murderer	White	19	141
	Black	17	149

$\hat{\theta} = \frac{19 \cdot 149}{141 \cdot 17} = 1.18106$

This odds ratio indicates that the odds that a white murderer receives the death penalty are higher than those of a black murderer. These results do exhibit Simpson's Paradox: inference made using conditional data will likely contradict inference made using marginal data.

X vs. Z

Victim	White	Black
White	151	9
Black	63	103

murderer

$\hat{\theta} = \frac{151 \cdot 103}{63 \cdot 9} = 27.4$

Z vs. Y

Decision	Y	N
White	30	184
Black	6	106

victim

$\hat{\theta} = \frac{30 \cdot 106}{6 \cdot 184} = 2.88$

Overall: By examining other partial tables, we can see why Simpson's paradox arose: because white victim crimes are more likely to receive the death penalty, and whites are <u>much</u> more likely to kill whites, making it appear that whites receive the death penalty more, as opposed to the more refined conclusion that black murderers are more likely to receive the death penalty.

35] At each age level, the death rate is higher in South Carolina than in Maine, but overall the death rate is higher in Maine. Explain how this could be possible.

It is typical for older age groups to have higher death rates than younger ones, so it is possible that Maine has an "older" population than South Caroling, creating a confound and causing Maine to appear to be a more deadly place to live than South Caroling.

37] Based on murder rates in the U.S., the Associated Press reported that the probability of a newborn child has of eventually being a murder victim is 0.0263 for non-white males, 0.0049 for white males, 0.0072 for non-white females, and 0.0023 for white females.

a) Find the conditional odds ratios between race and victim status, given gender. Interpret.

X = race Y/N Y = victim status Y/N Z = gender M/F

Z = Male

X - race white	Y - victim: Y	Y - victim: N
Y	.0049	.9951
N	.0263	.9737

$$\hat{\theta}_{Z=m} = \frac{.0049 \cdot .9737}{.9951 \cdot .0263} = .18231$$

Z = Female

X - race white?	Y - victim: Y	Y - victim: N
Y	.0023	.9977
N	.0072	.9928

$$\hat{\theta}_{Z=F} = \frac{.0023 \cdot .9928}{.9977 \cdot .0072} = 0.31788$$

The odds ratios indicate that whites are much less likely to be murder victims than non-whites. Also, the difference is less drastic for females.

b) If half the newborns are of each gender, for each "race," find the marginal odds ratio between race and whether a murder victim.

Y - victim status

X - race white?	Y	N
Y	$\frac{.0049}{2} + \frac{.0023}{2} = .0036$	$\frac{.9951}{2} + \frac{.9977}{2} = .9964$
N	$\frac{.0263}{2} + \frac{.0072}{2} = .01675$	$\frac{.9737}{2} + \frac{.9928}{2} = .98325$

$$\hat{\theta} = \frac{.0036 \cdot .98325}{.9964 \cdot .01675} = .21209$$

In this case, Simpson's paradox fails to emerge. Gender is not a confound in this example because there was no disparity between number of male cases and number of female cases in general or for either race.