Chapter 2

Q 1 (3 marks)

There is some evidence that drinking moderate amounts of wine helps prevent heart attacks. The table below gives data on yearly wine consumption and yearly deaths from heart disease.

Wine	Deaths	Country	Wine	Deaths
2.5	211	Netherlands	1.8	167
3.9	167	New Zealand	1.9	266
2.9	131	Norway	0.8	227
2.4	191	Spain	6.5	86
2.9	220	Sweden	1.6	207
0.8	297	Switzerland	5.8	115
9.1	71	United Kingdo	m 1.3	285
0.8	211	United States	1.2	199
0.7	300	West Germany	/ 2.7	172
7.9	107			
	Wine 2.5 3.9 2.9 2.4 2.9 0.8 9.1 0.8 0.7 7.9	WineDeaths2.52113.91672.91312.41912.92200.82979.1710.82110.73007.9107	WineDeathsCountry2.5211Netherlands3.9167New Zealand2.9131Norway2.4191Spain2.9220Sweden0.8297Switzerland9.171United Kingdo0.8211United States0.7300West Germany7.9107	WineDeathsCountryWine2.5211Netherlands1.83.9167New Zealand1.92.9131Norway0.82.4191Spain6.52.9220Sweden1.60.8297Switzerland5.89.171United Kingdom 1.30.8211United States1.20.7300West Germany2.77.9107107

- a) Make a scatterplot that shows how national wine consumption helps explain heart disease death rates.
- b) Describe the form of the relationship. Is there a linear pattern? How strong is the relationship?
- c) Is the direction of the association positive or negative? Explain in simple language what this says about wine and heart disease. Do you think these data give good evidence that drinking wine causes a reduction in heart disease deaths?

Answers



- b) This is a fairly strong relationship.
- c) There is a negative association between wine consumption and number of heart disease. Countries with high wine consumption have fewer heart disease deaths, while low wine consumption tends to go with more deaths from heart disease. This does not prove causation; there may be some other reasons (lurking variables) for the link.

Q 2. (4 marks)

Answers

b)

After birth, a male white rat gains 40 grams per week.

- a) If the rat weighed 100 g at birth, give an equation for his weight after x weeks. What is the slope of the line?
- b) Draw a graph of this like between birth and 10 weeks of age.
- c) Would you be willing to use this line to predict a rats weight at age 2 years? Do the prediction and think about the reasonableness of the result. (There are 454 g in a pound. Note that a cat is about 10 pounds).



Weeks

c) 2 years = 104 weeks

In two years the rat will weigh 4260 g. or 9.4 pounds. Rats do not grow at a constant rate through out their years.

Q. 3 (8 marks)

As running speed increases the stride rate should also increase for runners. In a study of 21 female runners, researchers measured the stride rate for different speeds. The following table gives the speeds in feet per second, and the mean stride rate for these runners.

Speed	15.86	16.88	17.50	18.62	19.97	21.06	22.11
Stride	3.05	3.12	3.17	3.25	3.36	3.46	3.55

a) Plot the data with speed on the x axis and stride on the y axis. Does a straight like adequately describe these data?

- b) Find the equation of the regression line of stride rate on the y axis. Draw this like on your plot.
- c) For each of the speeds given, obtain the predicted value and the residual. Verify that the residuals add to 0.

Answers

- a) The relationship seems linear
- b) R = 0.999, a = 1.77, b = 0.0803
 Regression line = 1.77 + 0.0803 X



speeds (feet/sc)

Regression

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999 ^a	.998	.998	9.068E-03

a. Predictors: (Constant), speeds (feet/sc)

b. Dependent Variable: stride rate

Coefficients^a

		Unstandardized Coefficients		Standardi zed Coefficien ts		
Model		В	Std. Error	Beta	t	Sig.
1	(Constant)	1.766	.031		57.566	.000
	speeds (feet/sc)	8.028E-02	.002	.999	49.658	.000

a. Dependent Variable: stride rate

c) Residuals are 0.01062, -0.0127, -0.00104, -0.01096, -0.00934, 0.00315, 0.00885 and they add up to 0.0001

Residuals are positive for the low and high speeds and negative for moderate speeds. This suggests that a curve maybe a better fit.

We cannot plot residuals vs time of observation because we do not have enough information

Graph of Residuals vs. Speed



speeds (feet/sc)

Q 2.96 (5 marks)

4. The following three way table classifies 326 cases in which the defendant was convicted of murder. The variables are defendant's race, vistim's race, and whether defendant was sentenced to death.

White defendant			Black defendant		
	Death Penalty		Death Penalty		
	Yes	No	Yes	No	
White victim	19	132	11	52	
Black victim	0	9	6	97	

- a) From these data make a two way table of the defendant's race by death penalty.
- b) Show that Simpson's paradox holds: a higher percent of white defendants are sentenced to death overall, but for both black and white victims a higher percent of black defendants are sentenced to death.
- c) Basing your reasoning on the data, explain why the paradox holds in language that a judge could understand.

Answers:

a)

	Death Penalty			
	Yes	No	Total	
White Def.	19	141	160	
Black Def.	17	149	166	

b)

Overall white defendants are sentenced to death more often than black defendants (11.9 % versus 10.2%).

But,

Brake down the death penalties by race of victim and race of defendant.

- White victim white defendant = 12.6% (19/151)
- White victim black defendant = 17.5% (11/63)

- Black victim white defendant = 0% (0/9)
- Black Victim Black defendant = 5.8% (6/103)

The Simpson's Paradox holds true because when you consider the race of the victim (lurking variable) then black defendants are sentenced to death more often than white defendants.

c) In cases involving white victims, 14% (19+11/214) of defendants got the death penalty.

White defendants killed whites 94.3% of the time (151/160) but are less likely to get the death penalty than blacks who killed whites.