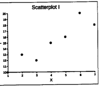
Question 1 For each scatterplot in (a), (b) and (c) below, select the correct form (linear or non-linear), strength of the association (weak or strong), direction (positive, negative or neither), and the best approximate value for the correlation r.

(a)[3 points]



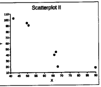
form: ☑ linear ☐ non-linear

strength: ☐ weak ☑ strong

direction: ☑ positive ☐ negative ☐ neither

value of r: ☐ −1.0 ☐ −0.7 ☐ −0.1 ☐ 0 ☑ 0.7 ☐ 1.0

(b)[3 points]



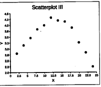
form: ☑ linear ☐ non-linear

strength: ☐ weak ☑ strong

direction: ☐ positive ☑ negative ☐ neither

value of r: ☐ −1.0 ☑ −0.7 ☐ −0.1 ☐ 0 ☐ 0.7 ☐ 1.0

(c)[3 points]



form: ☐ linear ☑ non-linear

strength: ☐ weak ☑ strong

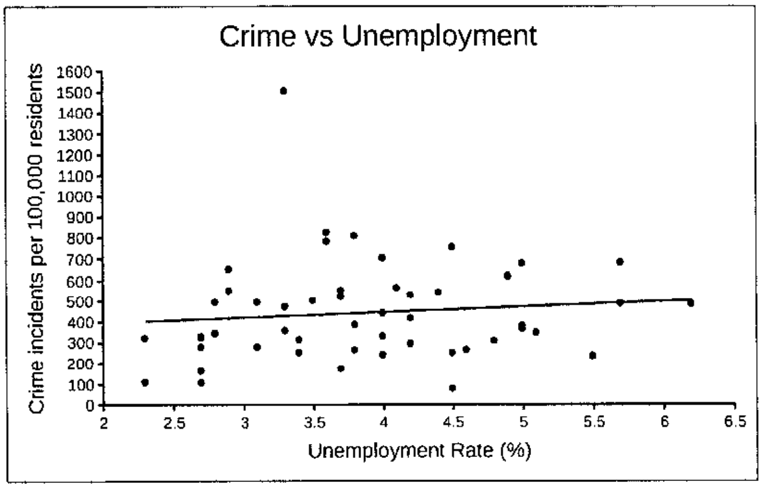
direction: ☐ positive ☐ negative ☑ neither

value of r: ☐ −1.0 ☐ −0.7 ☐ −0.5 ☑ 0 ☐ 0.7 ☐ 1.0

(d)[1 point] A scatterplot contains only two data points: $(1, 10)$ and $(10, 1)$. What must be the value of r?

$r = -1$. Both data points lie on a line, and y decreases with x.

Question 2 The following scatterplot shows crime and unemployment rates for fifty US states:



Letting x denote the unemployment rate and y the crime incidents per 100,000 residents, the least squares regression line shown has equation

$$y = 25.6x + 344.0$$

(a)[3 points] Based on the least squares model, what is the approximate change in crime rate corresponding to a two percent increase in unemployment rate?

As x increases by 2, y increases by $(25.6)(2) = \boxed{51.2}$
(incidents per 100,000 residents.)

(b)[3 points] Based on the least squares model, what unemployment rate corresponds to a crime rate of 450 incidents per 100,000 residents?

When $y = 450$ what is x?

$$450 = 25.6x + 344.0$$
$$25.6x = 450 - 344$$
$$x = \frac{450-344}{25.6} \doteq \boxed{4.1\%}$$

(c)[2 points] The square of the correlation in this case is $r^2 = 0.01$. A national newspaper claims that, based on the least squares model, crime increases with unemployment. Is this a fair statement?

No. $r^2 = 0.01$ says that only 1% of the variation in y is explained by the linear relationship between y and x.

(d)[2 points] If the data point corresponding to 1500 on the y-axis is removed from the scatterplot would you expect r to increase or decrease? Explain briefly.

Increase, since the data would then follow a stronger linear relationship.

Question 3 This question deals with the roulette wheel discussed in class and also in your homework. A roulette wheel has 38 slots: 18 red, 18 black, and 2 green slots numbered 0 and 00. On any spin of the roulette wheel the marble is equally likely to land in any of the 38 slots. Round all answers below to two decimals.

(a)[3 points] On two successive spins of the wheel, what is the probability of first landing on red, and on the second spin landing on black or green?

P(first red, second black or green)

= P(first red) P(second black or green) by independence.

$= \left(\frac{18}{38}\right)\left(\frac{20}{38}\right)$

$\doteq \boxed{0.25}$

(b)[3 points] On two successive spins of the wheel, what is the probability of landing on green at least once?

P(green at least once)

= 1 − P(not green both times)

$= 1 - \left(\frac{36}{38}\right)\left(\frac{36}{38}\right)$ by independence

$\doteq \boxed{0.10}$

(c)[4 points] Suppose landing on red or black pays you \$2, landing on the green 0 pays you \$12, and landing on the green 00 pays you \$20. What is the expected payout for this game (ignore what it costs you to play the game– assume it costs you nothing).

Let X = payout

$E(X) = (2)\,P(\text{red or black}) + (12)\,P(\text{green } 0) + (20)\,P(\text{green } 00)$

$= (2)\left(\frac{36}{38}\right) + (12)\left(\frac{1}{38}\right) + (20)\left(\frac{1}{38}\right)$

$= \boxed{\$2.74}$

Question 4 If a coin is flipped 10,000 times, the number of times heads comes up is approximately normally distributed with mean $\mu = 5000$ and standard deviation $\sigma = 50$.

(a)[2 points] If a coin is flipped 10,000 times what is the probability of getting fewer than 5075 heads?

Let X = # heads.

z-score for $X = 5075$: $z = \frac{5075-5000}{50} = 1.5$

$\therefore P(X < 5075) = P(Z < 1.5) = \frac{93.32}{100} = \boxed{0.9332}$

(b)[4 points] Suppose two people each flip a coin 10,000 times. Find the probability that at least one of the sequences of 10,000 coin flips results in 5075 or more heads.

P(at least one results in 5075 or more heads)

= 1 − P(both result in fewer than 5075 heads)

= 1 − (0.9332)(0.9332) by independence, using (a)

= $\boxed{0.1291}$

(c)[4 points] If a coin is flipped 10,000 times, the probability of getting fewer than n heads is 0.92. What is n?

$P(X \leq n) = 0.92$

z-score for $X = n$ is $\frac{n-5000}{50}$.

z-score corresponding to 92nd percentile is 1.4

$\therefore \frac{n-5000}{50} = 1.4$,

$n - 5000 = 70$, so $n = 5000 + 70 = \boxed{5070}$

Question 5

(a)[**5 points**] Consider the following game: draw a card from a standard 52 card deck. If the card is a king, you receive \$10, if it is a red queen you pay \$12. If any other card comes up you don't pay or receive anything. What amount of money do you expect to win or lose in this game?

Let X = amount won.

$$E(X) = (10)P(\text{King}) + (-12)\,P(\text{red queen}) + \cancel{(0)\,P(\text{not King or red queen})} \to 0$$

$$= (10)\left(\frac{4}{52}\right) + (-12)\left(\frac{2}{52}\right)$$

$$= \boxed{\$0.31}$$

(b)[**5 points**] We wish to modify the game in part (a) so that the expected amount won or lost is zero. If you still must pay \$12 when you draw a red queen, what must be the amount received if you draw a king?

Let C = amount received if King is drawn.

Then we want $E(X) = 0$, so

$$0 = (c)\left(\frac{4}{52}\right) + (-12)\left(\frac{2}{52}\right)$$

$$0 = \frac{4c}{52} - \frac{24}{52}$$

$$0 = \frac{4c-24}{52}$$

$$\therefore\ 4c - 24 = 0$$

$$4c = 24$$

$$\boxed{C = \$6}$$