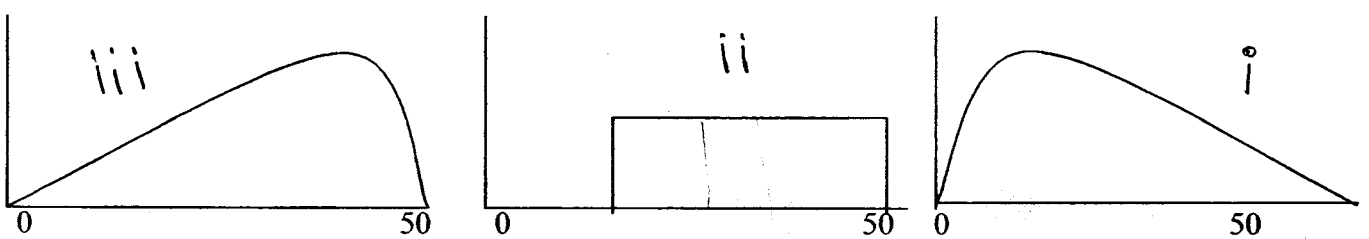
1. Please indicate whether each statement is true or false **(one point each)**

	True	False	Statement
A		X	In a randomized controlled experiment utilizing a placebo, if the control group is comparable to the treatment group, then the difference in the responses of the two groups is likely to be a result of the placebo effect.
B	X		Double blind experiments are better at preventing bias than blind experiments
C	X		A histogram is a graphical summary which represents percentages as areas
D	X		Probability methods are used in sampling because they are impartial and they can minimize bias
E	X		The area under the histogram between two values is equal to the percentage of cases in a class interval defined by those values
F	X		Incorporating a density scale in a histogram means that the height of each block times the width of the class interval equals the percentage of cases in that class interval
G	X		Confounding is a major source of bias
H		X	Larger samples are better than smaller samples because they are better at preventing bias
I	X		If a large number of those selected for a sample do not respond, problems of response bias are likely
J	X		The total area of a histogram is always 100% when areas are expressed as percentages

2. Here are three histograms, assume they have been correctly drawn:



Match **each** histogram above to the best choice listed below: **(2 points each, 6 points total)**

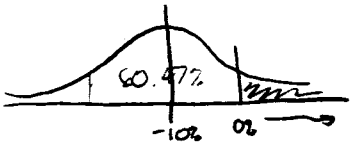
i. The average is larger than the median.
ii. The average is equal to the median
iii. The average is smaller than the median
iv. Cannot determine the median for this graphic
v. Cannot determine the average for this graphic

3. In an observational study (choose one) **(2 points)**
   a. Investigators do not assign subjects to treatment or to control groups
   b. There is a control group
   c. Investigators can establish association but not causation
   d. Confounding factors can occasionally be controlled for by comparing homogeneous subgroups
   e. All of the above are true
   f. Only A and C are true
   g. Only D is false

4. The next 3 questions refer to this statement, but each question is separate (i.e. you can get the first one wrong and it won't affect the others): Los Angeles has been suffering economically as a result of this year's stock market decline and loss of tax revenue. The 20,000 restaurants and eateries in Los Angeles reflect this economic downtown in their percentage change in sales. Suppose the sales for all of these restaurants are normally distributed with an average percentage change in sales of -10% over the last year (they loss 10% of their sales on average) and a standard deviation of 12% SHOW YOUR WORK FOR FULL CREDIT.

A. What percentage of restaurants experienced no decrease or an increase in sales over the last year? **(5 points)**

we want return = 0% $\bar{x} = -10\%$ $SD = 12\%$



$$z_0 = \frac{0-(-10)}{12}$$

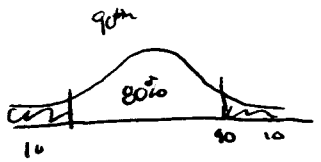
$$z_0 = \frac{10}{12} = .833 \approx .85$$

between $-z$ and z
$(-.85)$ and $(.85) = 60.47\%$

$100 - 60.47 = 39.53 - 2 = 19.765\%$

$\rightarrow 19.8\% \approx$ **20%**

B. The percentage change in sales for your favorite restaurant was at the 90th percentile. What exactly was the value of its percentage change? **(5 points)**



find z with 80%

$$1.30 = \frac{x-(-10)}{12}$$

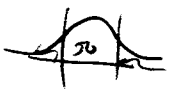
$$15.6 = x + 10$$
$$-10 \quad -10$$
$$5.6 = x$$

$\rightarrow$ **5.6% increase**

C. Using only the information given at the beginning (i.e. ignore any additional information from parts A and B), is it possible to calculate the inter-quartile range for this distribution of percentage change in sales? If yes, please calculate the IQR for this variable. If no, please write "not calculable" in the space below and explain verbally why it is not calculable. **(7 points)**

Yes. IQR = 75th percentile - 25th percentile (value)

75th percentile: 50% in middle



.65 standard deviations from 0

75th percentile, use +.65

25th percentile, use -.65

$$.65 = \frac{Q_3-(-10)}{12} = 7.8 = Q_3 + 10$$
$$-10 \quad -10$$
$$-2.2 = Q_3$$

$$-.65 = \frac{Q_1-(-10)}{12} = -7.8 = Q_1 + 10$$
$$-17.8 = Q_1$$

$$Q_3 - Q_1 = -2.2 - (-17.8)$$
$$= \boxed{15.6}$$

The Dull Computer Company manufacturers its own computers and delivers them directly to customers who order them via the Internet. Dull's market dominance has arisen from its quick delivery and competitive pricing. The CEO (Chief Executive Officer) of Dull has stated publicly that if customers make unassisted online purchases of their computers, the computers will have a mean delivery time of 45 hours from the time of purchase (with a standard deviation of 11 hours) and have a mean cost of $1,503 with a standard deviation of $678. Please assume that the delivery time and the cost of all the computers is normally distributed.

A consumer research organization decided to test the CEO's mean delivery time claim by purchasing 100 computers from Dull at randomly selected times and days. The 100 purchases were randomly divided into two groups: 51 were purchased by telephone and involved talking to a live salesperson, the remaining 49 were unassisted online purchases. The delivery time of the 49 had a mean of 52 hours with a standard deviation of 16 hours and they also had a mean cost of $1,588 with a standard deviation of $401. 11 of the 49 computers were delivered in less than 45 hours.

5A. (2 points) The population of interest to the consumer research organization is

(a) all Dull Computers
(b) all Dull Computers purchased online and unassisted
(c) 100 computers purchased from Dull by the consumer research organization
(d) 51 computers purchased by telephone
(e) 49 computers purchased online and unassisted

5B. (2 points) The sample of interest to the consumer research organization is

(a) all Dull Computers
(b) all Dull Computers purchased online and unassisted
(c) 100 computers purchased from Dull by the consumer research organization
(d) 51 computers purchased by telephone
(e) 49 computers purchased online and unassisted

5C. (2 points) The statistic of greatest interest to the consumer research organization is

(a) 52 hours
(b) 16 hours
(c) 45 hours
(d) 11 hours
(e) $1,503
(f) $1,588
(g) $401
(h) $678

5D. (2 points) The parameter of greatest interest to the consumer research organization is

(a) 52 hours
(b) 16 hours
(c) 45 hours
(d) 11 hours
(e) $1,503
(f) $1,588
(g) $401
(h) $678

5E. (2 points) The median cost of all Dull computers purchased online and unassisted is

(a) $1, 503
(b) $1, 588
(c) Greater than $1,503 and also greater than $1,588
(d) Less than $1,503 and also less than $1,588
(e) Greater than $1,503 but less than $1,588
(f) Not possible to calculate the median from the information provided