Investigators take a sample of DINKS (dual-income families, where husband and wife both work and have no kids). The investigators have data on the husband’s income and the wife’s income. By definition,

\[
\text{family income} = \text{husband’s income} + \text{wife’s income}.
\]

The average family income was around $50,000, and 10% of the couples had family income in the range $45,000–$55,000. Fill in the blanks, using the options given below, and explain briefly:

1. (10 Points) The correlation between wife’s income and family income is [\_\_\_\_] somewhat positive.

   *Textbook: “Although wife’s income must be less than family income, the two are positively associated.”*

   Options: (a) -1   (b) nearly -1   (c) somewhat negative   (d) 0
   (e) somewhat positive   (f) nearly 1   (g) 1   (h) -1.1   (i) 1.1

   “lightly” wrong answer: 1: (d), (f)
   “slightly” wrong answer: 2: (a), (c)

2. (10 Points) Among couples whose family income is in the range $45,000–$55,000, the correlation between wife’s income and husband’s income is [\_\_\_\_] nearly -1.

   *Textbook: “As family income is practically constant, the more the wife makes, the less the husband can make.”*

   “slightly” wrong answer: 2: (a), (c)
   “slightly” wrong answer: 5: (b), (d)
   “slightly” wrong answer: 6: (b), (d)
   “slightly” wrong answer: 8: (f), (g)
Question 2: Histograms (35 Points)

When the Tribbles invaded the spaceship Enterprise, suppose that crew member Spock decided to take the logical step of seeing what the crew was up against, and he wanted to graphically represent the size of the Tribbles. Suppose that the table below summarizes the heights of the 50 Tribbles he found on the bridge. (Class intervals include the left but not the right endpoints.) [If you don't know what Tribbles are, take a look at http://www.startrek.com/startrek/view/series/TOS/episode/68744.html – there are several photos available at this Web site.]

<table>
<thead>
<tr>
<th>Tribble Height (inches)</th>
<th>Number of Tribbles</th>
<th>Percentage</th>
<th>Width</th>
<th>Height = Percentage/Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5</td>
<td>19</td>
<td>38%</td>
<td>2</td>
<td>38%/2 = 19%</td>
</tr>
<tr>
<td>5-7</td>
<td>10</td>
<td>20%</td>
<td>2</td>
<td>20%/2 = 10%</td>
</tr>
<tr>
<td>7-9</td>
<td>12</td>
<td>24%</td>
<td>2</td>
<td>24%/2 = 12%</td>
</tr>
<tr>
<td>9-11</td>
<td>7</td>
<td>14%</td>
<td>2</td>
<td>14%/2 = 7%</td>
</tr>
<tr>
<td>11-13</td>
<td>2</td>
<td>4%</td>
<td>2</td>
<td>4%/2 = 2%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. (25 Points) Draw a histogram of these height data, with the vertical axis on the usual density scale, both axes labeled, and heights of bars clearly indicated.

2. (10 Points) If a Tribble is in the 82nd percentile for height, about how tall is it? (Note: Use the histogram, NOT the normal curve here.)

Its height is about: 9 inches
Question 3: Normal Curve (45 Points)

Spock did some additional height measurements of all Tribbles aboard Enterprise (and not only of those found on the bridge) and determined that their overall size closely follows the normal curve, with an average of 6.5 inches and an SD of 2.5 inches.

Fill the blanks in the statements below and show all the work needed to obtain the answer:

1. (15 Points) The percentage of Tribbles that are between 8.0 and 10.0 inches tall is \[ \frac{8.0 - 6.5}{2.5} = 0.6 \] \[ \frac{10.0 - 6.5}{2.5} = 1.4 \] area between 0.6 and 0.6: \[ 45.15\% \] area between -1.4 and 1.4: \[ 83.85\% \] area between 0.6 and 1.4: \[ \frac{83.85\% - 45.15\%}{2} = 19.35\% \]

2. (15 Points) The percentage of Tribbles that are less than 4.5 inches tall is \[ \frac{4.5 - 6.5}{2.5} = -0.8 \] area between -0.8 and 0.8: \[ 57.63\% \] area below -0.8: \[ \frac{100\% - 57.63\%}{2} = 21.19\% \]

3. (15 Points) When using the normal curve, a Tribble that is in the 82nd percentile for height is about \[ \frac{8.75 \text{ inches}}{2} \] inches tall.
Question 4: Controlled Experiment/Observational Study (60 Points)

A recent study in Europe looked at a large group of women of childbearing age. The researchers asked each woman how much alcohol they had consumed over the past 12 months. The researchers found that women who drank moderate amount of alcohol were somewhat less likely to have infertility problems than women who did not drink alcohol at all (November, 2001). The study said it “controlled for age, income, and religion”.

1. (15 Points) Based on the information above, was this a controlled experiment or an observational study? Circle your answer and explain briefly.

   - No intervention was used—nobody was told to drink or not to drink.
     - Correct explanation
     - Incorrect explanation

2. (15 Points) Why did they “control for” age, income, and religion?

   - There may be confounding factors
   - -5 for missing keyword
   - -5 for otherwise correct explanation

3. (15 Points) Is this convincing evidence that infertility would decrease if women with infertility problems started to drink moderate amounts of alcohol? (Note: we are only asking about infertility. There may be other problems introduced by such behavior, but ignore them for answering this question).

   - No! We only know that there is an association between drinking and fertility; drinking does not cause infertility
   - Correct explanation
   - Incorrect explanation

4. (15 Points) Suggest a possible confounding factor (other than age, income, or religion) and clearly explain why you think it might be a confounding factor.

   - general health (condition): someone who has some other medical problem may not drink and also be less fertile
   - Correct confounding factor
   - Correct explanation
   - Incorrect explanation
Question 5: Regression (40 Points)

In one study, the correlation between the educational level of husbands and wives in a certain town was about 0.50; both averaged 12 years of schooling completed, with an SD of 3 years.

1) husband, 2) wife: \( \bar{\text{avg}}_x = 12 \) \( \sigma_x = 3 \) \( r = 0.50 \)

Show your work.

1. (15 Points) Predict the educational level of a woman whose husband has completed 18 years of schooling.
   
   The answer is: \( \underline{15} \) years

   \[
   s_{\bar{y}} = \frac{x - \text{avg}_x}{\sigma_x} = \frac{18 - 12}{3} = 2 \quad (5) \\
   s_{\bar{y}} = s_{\bar{x}} \cdot r = 2 \cdot 0.50 = 1 \quad (5) \\
   \gamma = s_{\bar{y}} \cdot \sigma_y + \text{avg}_y = 1 \cdot 3 + 12 = 15 \quad (5)
   \]

2. (15 Points) Predict the educational level of a man whose wife has completed 15 years of schooling.
   
   The answer is: \( \underline{13.5} \) years

   \[
   s_{\bar{x}} = \frac{x - \text{avg}_x}{\sigma_x} = \frac{15 - 12}{3} = 1 \quad (5) \\
   s_{\bar{x}} = 1 \cdot 0.50 = 0.50 \quad (5) \\
   \gamma = 0.50 \cdot 3 + 12 = 13.5 \quad (5)
   \]

3. (10 Points) Apparently, well-educated men marry women who are less well-educated than themselves. But the women marry men with even less education. How is this possible?

   Nothing unexpected—this is just the regression effect!

   The explanation given is an example of the regression fallacy.

   \( \underline{10} \) for correct keyword
   \( \underline{5} \) for reasonable explanation
   \( \underline{1} \) for some explanation