1. (4) Mrs. Field's doctor is concerned that she may suffer from gestational diabetes (high blood glucose levels during pregnancy). There is variation in both a person's actual glucose level and in the blood test that measures the level. A patient is classified as having gestational diabetes if the glucose level is above 140 milligrams per deciliter one hour after a measured amount of a sugary drink is ingested.

Suppose that Mrs. Field's measured glucose level one hour after ingesting the sugary drink varies according to the Normal distribution with \( \mu = 125 \) mg/dl and \( \sigma = 10 \) mg/dl. If measurements are made on 4 separate days and the mean result is compared with the criterion 140 mg/dl, what is the probability that Mrs. Field is diagnosed as having gestational diabetes? (numeric answer; show your work)

2. The Federal Trade Commission annually rates varieties of domestic cigarettes according to their tar, nicotine, and carbon monoxide content. This problem is based on a subset of such FTC data. The SAS output below concerns carbon monoxide (cot) levels in milligrams in a random sample of 25 brands of cigarette.

<table>
<thead>
<tr>
<th>Analysis Variable</th>
<th>CDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>25</td>
</tr>
<tr>
<td>Mean</td>
<td>12.528000</td>
</tr>
<tr>
<td>Std Dev</td>
<td>4.7396835</td>
</tr>
<tr>
<td>Lower 90.0% CLM</td>
<td>10.9661921</td>
</tr>
<tr>
<td>Upper 90.0% CLM</td>
<td>14.1498079</td>
</tr>
</tbody>
</table>

(a) (3) To correct interpretation of the 90% confidence interval for cot is
i. We are 90% confident that the mean cot in the population of all types of cigarettes is between 10.966 and 14.149 mg.
ii. We are 90% confident that the sample mean of cot is between 14.960 and 14.150 mg.
iii. There is a 90% probability that the mean cot in the population of all types of cigarettes is between 14.960 and 14.150 mg.
iv. There is a 90% probability that the sample mean of cot is between 10.966 and 14.150 mg.

(b) (3) If we computed a 95% confidence interval based on the same dataset (circle one)

i. it would be narrower than the 90% confidence interval
ii. it would be wider than the 90% confidence interval
iii. insufficient information is given in the problem to tell whether the 95% interval would be wider or narrower than the 90% interval

(c) (2) SAS computed a t confidence interval. This type of confidence interval is appropriate when (circle one):

i. the sample is drawn from a population that is assumed to be approximately normal with \( \mu \) known
ii. the sample is drawn from a population that is assumed to be approximately normal with \( \sigma \) NOT known
iii. the sample is drawn from a population that is assumed to be approximately normal with \( \mu \) known
iv. the margin of error is very small

(d) (3) In computing the \( t \) confidence interval from these data, SAS used a \( t \) distribution with how many degrees of freedom?

3. This data contains the labor force participation rate (LFPR) of women in 19 cities in the United States in each of two years (1968 and 1972). The data help to measure the presence of women in the labor force over this period. The variables are:

1. city: City in the United States
2. lfp1972: Labor Force Participation rate of women in 1972

Labor force participation is a quantitative variable.

Suppose that these 19 cities could be considered a simple random sample of cities in the U.S. We wish to use these data to draw conclusions as to whether the mean labor force participation of women in all U.S. cities stayed the same or changed between 1968 and 1972. We do not know in advance in which direction a change might have gone.

(a) (1) The populations of interest are (circle one):

i. all women in the labor force
ii. all cities in the U.S. in 1968 and all cities in the U.S. in 1972
iii. the 19 cities included in the study
iv. labor force participation of women
v. the average labor force participation of women in all U.S. cities in 1968 and the average labor force participation of women in all U.S. cities in 1972
vi. the average labor force participation of women in the 19 cities in 1968 and the average labor force participation of women in the 19 cities in 1972
vii. none of the above

(b) (1) The parameters of interest are (circle one):
   i. all women in the labor force
   ii. all cities in the U.S., in 1968 and all cities in the U.S., in 1972
   iii. the 19 cities included in the study
   iv. labor force participation of women
   v. the average labor force participation of women in all U.S. cities in 1968 and the average labor force participation of women in all U.S. cities in 1972
   vi. the average labor force participation of women in the 19 cities in 1968 and the average labor force participation of women in the 19 cities in 1972
   vii. none of the above

(c) (1) The variable of interest is (circle one):
   i. all women in the labor force
   ii. all cities in the U.S., in 1968 and all cities in the U.S., in 1972
   iii. the 19 cities included in the study
   iv. labor force participation of women
   v. the average labor force participation of women in all U.S. cities in 1968 and the average labor force participation of women in all U.S. cities in 1972
   vi. the average labor force participation of women in the 19 cities in 1968 and the average labor force participation of women in the 19 cities in 1972
   vii. none of the above

(d) (1) The study has been set up as a
   i. single sample problem
   ii. paired sample problem
   iii. two independent sample problem
   iv. none of the above

(e) (2) Write appropriate null and alternative hypotheses for a test that would address the nutritionists’ research question. Use conventional symbols.

(f) (2) Use the SAS output to give the test statistic and p-value for the test (numeric answers)

(g) (2) Should we reject the null hypothesis at the .05 significance level? (yes/no)
   Briefly explain.

(h) (1) The appropriate interpretation of the result of the hypothesis test alone is:
   i. the data provided sufficient evidence to reject the null hypothesis that the average of women’s labor force participation in U.S. cities did not change between 1968 and 1972
   ii. the data did not provide sufficient evidence to reject the null hypothesis that the average of women’s labor force participation in U.S. cities did not change between 1968 and 1972
   iii. the data prove that the average of women’s labor force participation in U.S. cities did not change between 1968 and 1972
   iv. the data prove that the average of women’s labor force participation in U.S. cities changed between 1968 and 1972
   v. the data prove that the average of women’s labor force participation in U.S. cities increased between 1968 and 1972
   vi. none of the above

(i) (2) Does any part of the SAS output give clear evidence of a change in one particular direction between 1968 and 1972? (yes/no) If so, cite the specific output and state the direction of change.

4. (4) A cigarette manufacturer claims that the mean nicotine content of its lower brand is 1.5 mg. We might investigate this claim by testing

\[ H_0 : \mu = 1.5 \]
\[ H_A : \mu > 1.5 \]

where \( \mu \) is the true population mean nicotine content in all cigarettes of this manufacturer’s lower brand. We will select a random sample of \( n = 36 \) cigarettes. Imagine that we know that the standard deviation \( \sigma \) of nicotine content is .20 mg. We choose a significance level \( \alpha \) of .01. This results in a critical value of \( z \) of 1.645 mg. That is, we must reject \( H_0 \) if \( z \) is greater than 1.645 mg, and otherwise fail to reject.

Suppose it would be important to us if the true \( \mu \) were as far away from 1.5 as 1.6 mg. What is the power of our test against this alternative? Give a numeric answer; show your work.
<table>
<thead>
<tr>
<th>Obs</th>
<th>city</th>
<th>lbfp1972</th>
<th>lbfp1968</th>
<th>diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N.Y.</td>
<td>0.45</td>
<td>0.42</td>
<td>0.03</td>
</tr>
<tr>
<td>2</td>
<td>L.A.</td>
<td>0.50</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>Chicago</td>
<td>0.52</td>
<td>0.52</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>Philadelphia</td>
<td>0.45</td>
<td>0.45</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>Detroit</td>
<td>0.46</td>
<td>0.43</td>
<td>0.03</td>
</tr>
<tr>
<td>6</td>
<td>San Francisco</td>
<td>0.15</td>
<td>0.56</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>Boston</td>
<td>0.60</td>
<td>0.46</td>
<td>0.15</td>
</tr>
<tr>
<td>8</td>
<td>Pitt.</td>
<td>0.49</td>
<td>0.34</td>
<td>0.15</td>
</tr>
<tr>
<td>9</td>
<td>St. Louis</td>
<td>0.35</td>
<td>0.45</td>
<td>0.10</td>
</tr>
<tr>
<td>10</td>
<td>Connecticut</td>
<td>0.55</td>
<td>0.54</td>
<td>0.01</td>
</tr>
<tr>
<td>11</td>
<td>Wash., D.C.</td>
<td>0.52</td>
<td>0.42</td>
<td>0.10</td>
</tr>
<tr>
<td>12</td>
<td>Cinn.</td>
<td>0.53</td>
<td>0.51</td>
<td>0.02</td>
</tr>
<tr>
<td>13</td>
<td>Baltimore</td>
<td>0.57</td>
<td>0.49</td>
<td>0.08</td>
</tr>
<tr>
<td>14</td>
<td>Newark</td>
<td>0.53</td>
<td>0.54</td>
<td>0.01</td>
</tr>
<tr>
<td>15</td>
<td>Minn./St. Paul</td>
<td>0.59</td>
<td>0.50</td>
<td>0.09</td>
</tr>
<tr>
<td>16</td>
<td>Buffalo</td>
<td>0.64</td>
<td>0.58</td>
<td>0.06</td>
</tr>
<tr>
<td>17</td>
<td>Houston</td>
<td>0.60</td>
<td>0.49</td>
<td>0.11</td>
</tr>
<tr>
<td>18</td>
<td>Patterson</td>
<td>0.57</td>
<td>0.56</td>
<td>0.01</td>
</tr>
<tr>
<td>19</td>
<td>Dallas</td>
<td>0.64</td>
<td>0.63</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The MEANS Procedure

Variable N Mean Std Dev CL for Mean CL for Mean
-----------------------------------------------
lbfp1968 19 0.493179 0.0079912 0.4891672 0.4971926
lbfp1972 19 0.5268421 0.0707933 0.4827208 0.5609634
diff 19 0.0336642 0.0597412 0.0048899 0.0624785

The UNIVARIATE Procedure

Variable: diff (lbfp1972 - lbfp1968)

Tests for Location: \( \mu = 0 \)

Test Statistic Value

Student's t 2.457704 Pr > |t| 0.0244
Sign M 0.7 Pr >= |M| 0.0074
Signed Rank S 46 Pr >= |S| 0.0062