Histidine is an amino acid that is a component of urine. The parallel boxplots show total histidine excretion (mg) from 24-hour urine samples for five men and ten women.

Summary statistics are as follows.

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>( \bar{y} )</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>men</td>
<td>5</td>
<td>300.8</td>
<td>123.7</td>
</tr>
<tr>
<td>women</td>
<td>10</td>
<td>153.2</td>
<td>49.8</td>
</tr>
</tbody>
</table>

(a) Test the hypothesis that the two population mean histidine excretion totals are equal versus the two-sided alternative. State hypotheses, calculate a test statistic, and report a p-value (with a range). There are 4.66 degrees of freedom. Summarize your conclusions in the context of the problem.

For each problem, circle True or False. Briefly justify your response.

(b) **True** or **False**:

Based on the graph alone, we can be essentially certain that the mean histidine excretion levels in men is greater than that in women in this population.

(c) **True** or **False**:

We could compute using R the exact p-value for the hypothesis test in (a) by finding the area to the right of the \( t \) test statistic (men minus women) under a \( t \) distribution with 4.66 degrees of freedom.

(d) **True** or **False**:

The one-sided test with alternative hypothesis that the mean for men is greater than the mean for women is significant at the 5% level.

(e) **True** or **False**:

A p-value of 0.06 means that the null hypothesis is true.

Solution:

(a) The hypotheses are \( H_0 : \mu_1 = \mu_2 \) versus \( H_A : \mu_1 \neq \mu_2 \) where \( \mu_1 \) and \( \mu_2 \) are the population mean 24-hour histidine excretion levels (mg) for men and women, respectively. The test statistic is

\[
t = \frac{300.8 - 153.2}{\sqrt{\frac{123.7^2}{5} + \frac{49.8^2}{10}}} \approx 2.57
\]

Rounding 4.66 degrees of freedom down to 4 to use the table, we find a one-sided p-value between 0.03 and 0.04 corresponding to two-sided p-values between 0.06 and 0.08. In the context of the problem,
There is weak evidence of a difference in mean 24-hour histidine excretion levels between men and women ($p < 0.08$, two-sided independent sample $t$-test with unequal variances).

(Note that with R we can use 4.66 degrees of freedom and find $p = 0.053$.)

(b) FALSE. We cannot be essentially certain about a hypothesis test from graphs alone, especially when the sample sizes are so small. That is why we have quantitative inference procedures.

(c) FALSE. The right $p$-value would be two-sided. We would need to double the calculated area under the curve.

(d) TRUE. The one-sided $p$-value was between 0.03 and 0.04, thus less than 0.05.

(e) FALSE. $P$-values do not measure the probabilities of hypotheses. A $p$-value of 0.06 means that if the null hypothesis is true, then the probability of observing a new test-statistic at least as extreme as that actually observed is 6 percent.