The Type-I and Type-II Errors in Business Statistics

<table>
<thead>
<tr>
<th>Your Decision Based On a Random Sample</th>
<th>Given the Null Hypothesis Is</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td>Type I Error</td>
</tr>
<tr>
<td>Correct Decision</td>
<td>False</td>
</tr>
<tr>
<td></td>
<td>Correct Decision</td>
</tr>
<tr>
<td>Do Not Reject</td>
<td>Type II Error</td>
</tr>
</tbody>
</table>

Two Types of Errors in Decision Making

As indicated in the above matrix a Type-I error occurs when, based on your data, you reject the null hypothesis when in fact it is true. The probability of a type-I error is the level of significance of the test of hypothesis and is denoted by $\alpha$.

Type-I error is often called the producer’s risk that consumers reject a good product/service indicated by the null hypothesis. That is, a producer introduces a good product, in doing so, he/she take a risk that consumer will reject it.

A type II error occurs when you do not reject the null hypothesis when it is in fact false. The probability of a type-II error is denoted by $\beta$.

Type-II error is often called the consumer’s risk for not rejecting possibly a worthless product or service indicated by the null hypothesis.

The foundation and logic of Statistics (i.e. Inferential Statistics):
Consider the test of hypothesis with Null Ho, verses a two-side alternative Ha, since the sample is random (i.e., unbiased) to get such a large (absolute) computed statistics under the null hypothesis is very rare (say $\alpha= 5\%$), however we got such a large statistics
surprisingly; the question is what is wrong here? Well, the only possibility is that your null hypothesis is wrong. That is why we reject the null hypothesis.

Since there is a duality between the test of hypothesis and estimation with confidence, the above logic is applicable to the estimation and construction of confidence interval.

- Confusion! Which one, Non-Rejection Region OR Non-Rejection Interval?

- Give two important applications of descriptive statistics, such as Histogram?

Three ways of doing statistical hypotheses:
1. Based on significance level (say $\alpha = 5\%$),
2. Based on P-value,
3. The hybrid of both

- What is the use of Standardize $Z$?

Among other useful questions, for example you may ask why we are interested in estimating the population's expected value $\mu$ and its Standard Deviation $\sigma$? Here are some applicable reasons. Business Statistics must provide justifiable answers to the following concerns for every consumer and producer:

1. What is your (or your customers) Expectation of the product/service you buy (or that your sell)? That is, what is a good estimate for $\mu$?
2. Given the information about your (or your customers) expectation, what is the Quality of the product/service you buy (or that you sell)? That is, what is a good estimate for $\sigma$?
3. Given the information about what you buy (or your sell) expectation, and the quality of the product/service, how does the product/service compare with other existing similar types? That is, comparing several $\mu$s’ and several $\sigma$s’.
4. Finding any (linear) relationship for prediction purposes. For example, sales ($S$) as function of advertising rate ($A$) for a specific budget and duration of campaign ($T$). That is, estimation of $S = mA + b$. 